DATA MINING AND PERFORMANCE OF MICROFINANCE INSTITUTIONS IN KENYA

KOKONYA N. LORINE

A MANAGEMENT RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF BUSINESS ADMINISTRATION, SCHOOL OF BUSINESS, UNIVERSITY OF NAIROBI

OCTOBER 2014
DECLARATION

I declare that this research project is my original work and has not been submitted for any academic qualification in this or any other University for examination.

Signed: _______________________________  Date: __________________

Kokonya N. Lorine

D61 / 61709 / 2010

This research project has been submitted with my approval as the project supervisor

Signed: _______________________________  Date: __________________

Dr. Kate Litondo

Lecturer, Department of Management Science
DEDICATION

This project is dedicated to my loving parents for instilling in me the virtue of hard work from a very early age. God bless you.
ACKNOWLEDGEMENT

I am deeply indebted to all those who in their own way contributed to successful completion of this study. First and foremost I thank the almighty God, to whom all knowledge, wisdom and power belong for sustaining me in good health, sound judgment and strength to move on and complete my master’s studies. Special appreciation goes to my supervisor for her dedication, guidance, valuable suggestion and ideas throughout the course of this project. Without her enormous support this study would not have been successful. Thanks to my family and friends who always inspired me in every step to accomplish this study. I am eternally grateful for your love, encouragement and support in all my endeavors.
ABSTRACT

Mining customer data to measure productivity and enhance performance management is not only feasible in this information era but also in line with the transformation of a microfinance institution into a "customer driven organization". In this paper, we look at application of data mining techniques to performance management in the microfinance industry. Data Mining or knowledge discovery in databases can be defined as an activity that extracts some new nontrivial information contained in large databases. The goal is to discover hidden patterns, unexpected trends or other subtle relationships in the data using a combination of techniques from machine learning, statistics and database technologies. This new discipline today finds application in a wide and diverse range of business, scientific and engineering scenarios. Given the above arguments, this study aimed to answer the following research questions: Are microfinance institutions in Kenya using data mining techniques? What challenges do microfinance institutions face when using these data techniques? What is the relationship between data mining techniques and the success of microfinance institutions? Population of this study comprised of 56 Microfinance Institutions with the population of interest being 56 respondents who are IT officers or database administrators of the 56 microfinance institutions in Kenya. This study was limited to the institutions that are registered and regulated by the AMFI. The findings revealed that data mining has strong and positive correlation with the performance of microfinance institutions in Kenya. Data mining makes it possible to analyze routine business transactions and glean a significant amount of information about individual’s such as buying habits and preferences, banking information and customer details. Businesses collect information about their customers in many ways to understand their purchasing behaviors trends and with the adoption of data mining, microfinance institutions are able to obtain data with ease. Data mining and data warehousing tend to be self-reinforcing. The more powerful the data mining queries, the greater the utility of the information being gleaned from the data, and the greater the pressure to increase the amount of data being collected and maintained, which increases the pressure for faster, more powerful data mining queries. This increases pressure for larger, faster systems, which are more expensive and enable microfinance to increase in performance efficiency. Based on the study findings the study recommends that microfinance institutions should adopt data mining to enhance their
performance. The organizations need to make sure that there is enough data to analyze as well as assure quality of data. Organizations should ensure that the analysts are trained well and deduct the correct information which serves the purposes of the problem in the first place. The process of using data mining should be a learning experience.
# TABLE OF CONTENTS

DECLARATION

1.1 Background of the study ............................................................................................................. 2
  1.1.1 Data Mining ...................................................................................................................... 2
  1.1.2 Performance of Microfinance Institutions .......................................................................... 4
  1.1.3 Microfinance Industry in Kenya .......................................................................................... 6
1.2 Statement of the problem ........................................................................................................... 7
1.3 Research Objectives ................................................................................................................... 8
1.4 Importance of the study ............................................................................................................. 9

CHAPTER TWO: LITERATURE REVIEW ................................................................................... 10

  2.1 Introduction ............................................................................................................................ 10
  2.2 The concept of Data Mining .................................................................................................. 10
    2.2.1 History of Data Mining ................................................................................................. 11
    2.2.2 Data Mining Techniques ............................................................................................... 12
    2.2.3 Benefits of Data Mining ............................................................................................... 14
    2.2.4 Challenges of Data Mining ........................................................................................... 17
  2.3 Performance of Microfinance Institutions ............................................................................ 19
  2.4 Data Mining and Performance of organizations .................................................................... 21
  2.5 Theoretical Perspective ........................................................................................................ 23
  2.6 Summary and Conceptual Framework .................................................................................. 24
    2.6.1 Conceptual Framework .................................................................................................. 24
    2.6.2 Summary of Literature Review .................................................................................... 25

CHAPTER THREE: RESEARCH METHODOLOGY ..................................................................... 27

  3.1 Introduction ............................................................................................................................ 27
  3.2 Research Design ..................................................................................................................... 27
  3.3 Population of the Study ......................................................................................................... 27
  3.4 Sample Population ................................................................................................................ 27
  3.5 Data collection ....................................................................................................................... 28
  3.6 Data Analysis ........................................................................................................................ 28

CHAPTER FOUR: DATA ANALYSIS, RESULTS AND DISCUSSION ..................................... 29

  4.1 Introduction ............................................................................................................................. 29
  4.2 Response Rate ........................................................................................................................ 29
4.3 Descriptive Statistics .................................................................................................................. 30
  4.3.1 Classify Ownership of Institution ....................................................................................... 30
  4.3.2 Number of Local branches ................................................................................................. 31
  4.3.3 Number of years in operation .............................................................................................. 31
  4.3.4 Designation Level of Respondents ...................................................................................... 32
4.4 Data Mining Techniques .............................................................................................................. 33
4.5 Challenges of Data Mining Techniques ...................................................................................... 34
4.6 Benefits of Data Mining Techniques .......................................................................................... 35
4.7 Critical performance Measures .................................................................................................. 36
4.8 Correlation Co-efficient ............................................................................................................ 37
4.9 Regression Analysis .................................................................................................................. 38
  4.9.1 Goodness of Fit Statistics .................................................................................................. 39
  4.9.2 Analysis of Variance (ANOVA) ......................................................................................... 40
  4.9.3 Regression Coefficients .................................................................................................... 40
4.10 Summary of the findings and interpretations ............................................................................ 45

CHAPTER FIVE: SUMMARY OF FINDINGS, CONCLUSION AND RECOMMENDATIONS .......................................................................................................................... 46
  5.1 Introduction .............................................................................................................................. 46
  5.2 Summary of findings ................................................................................................................ 46
  5.3 Conclusion ............................................................................................................................... 47
  5.4 Recommendation .................................................................................................................... 48
  5.5 Limitation of the study ............................................................................................................ 49
  5.6 Suggestions for further Research ............................................................................................. 49

References ......................................................................................................................................... 50

Appendix I: - LIST OF MFIs ........................................................................................................... 53

Appendix II: - Sample Questionnaire .............................................................................................. 54
LIST OF FIGURES

Figure 4.1: The three domains of the real. Source: (Mingers, 2004, p. 98) ..................................23

Figure 4.2: Conceptual Framework Source: researcher (2014) ..................................................25
LIST OF TABLES

Table 4.1: Effectiveness of Micro Finance Institutions (Lindsay (2010)) ........................... 21
Table 4.2: Response Rate ........................................................................................................ 30
Table 4.3: Challenges hindering adoption of data mining techniques .................................... 34
Table 4.4: Benefits of Data Mining Techniques ...................................................................... 35
Table 4.5: Average Representation of critical performance measures ................................. 37
Table 4.6: Pearson Correlation Matrix ...................................................................................... 38
Table 4.7: Goodness of Fit Statistics .......................................................................................... 39
Table 4.8: Analysis of Variance (ANOVA) ............................................................................ 40
Table 4.9: Regression Results: Number of people served ...................................................... 41
Table 4.10: Regression Results on Moderating Variables: No. of people served ............... 42
Table 4.11: Regression Results: Number loans funded per year ........................................... 42
Table 4.12: Regression Results on Moderating Variables: No. loans funded per year ....... 43
Table 4.13: Regression Results: Total amount loaned per year ............................................. 44
Table 4.14: Regression Results on Moderating Variables: Total amount loaned per year .... 44
CHAPTER ONE: INTRODUCTION

1.1 Background of the study

Performance management (PM) has become one of the most important initiatives in the microfinance industry today. One urgent problem facing the Microfinance institutions is how to link the organization’s performance to growth and profitability so that resources can be optimally allocated and fully utilized to meet competition and support increasing demand of quality products/services from customers. Mining customer data to measure productivity and enhance performance management is not only feasible in this information era but also in line with the transformation of a microfinance institution into a "customer driven organization". In this paper, we look at application of data mining techniques to performance management in the microfinance industry.

1.1.1 Data Mining

Data Mining or knowledge discovery in databases can be defined as an activity that extracts some new nontrivial information contained in large databases. The goal is to discover hidden patterns, unexpected trends or other subtle relationships in the data using a combination of techniques from machine learning, statistics and database technologies. This new discipline today finds application in a wide and diverse range of business, scientific and engineering scenarios.

The overall knowledge discovery process is outlined as an interactive and iterative process involving more or less the following steps: understanding the application domain, selecting the data, data cleaning and preprocessing, data integration, data reduction and transformation,
selecting data mining algorithms, data mining, and interpretation of the results and using the
discovered knowledge Fayyad et al (1996)

According to Silltow (2006), in its simplest form, data mining automates the detection of
relevant patterns in a database, using defined approaches and algorithms to look into current
and historical data that can then be analyzed to predict future trends. Because data mining
tools predict future trends and behaviors by reading through databases for hidden patterns,
they allow organizations to make proactive, knowledge-driven decisions and answer
questions that were previously too time-consuming to resolve. Silltow (2006) mentions the
core data mining techniques as: association, classification, link analysis, sequence analysis,
clustering, prediction, sequential patterns, decision trees, combinations, long-term (memory)
processing.

Data mining has got its own challenges and Han (2006) mentions these challenges to include:
lack individual privacy; issue of data integrity whereby he states that data analysis can only
be as good as the data that is being analyzed; the issue of cost; most databases are dynamic as
databases usually change continually therefore getting representation data may be a
challenge; databases may be huge and there may be difficulty in accessing data.

Newman (2013) also argues that a lot of good can come from data mining. He highlights the
top benefits of data mining as more money as a result of profits and investments; tapping into
new markets; the principle of share and share alike which involves sharing of information
that may be useful to similar organizations; learning from the past; help in competitor
analysis where data mining helps companies to get information that they can use effectively
to stand out from competition
1.1.2 Performance of Microfinance Institutions

Microfinance, according to Otero (1999) is the provision of financial services to low-income poor and very poor self-employed people. These financial services according to Ledgerwood (1999) generally include savings and credit but can also include other financial services such as insurance and payment services. Schreiner et al (2001) define microfinance as the attempt to improve access to small deposits and small loans for poor households neglected by banks. Therefore, microfinance involves the provision of financial services such as savings, loans and insurance to poor people living in both urban and rural settings who are unable to obtain such services from the formal financial sector.

Microfinance institutions (MFIs) play a vital role in the economic development of many developing countries. They offer loans and/or technical assistance in business development to low-income communities in developing countries Hartungi (2007). For some microfinance is a world where as many poor and near-poor households as possible have permanent access to an appropriate range of high quality financial services including not just credit but also savings, insurance and fund transfers.

Micro finance developed from banking systems dating back to the early 1700s, but it wasn’t until the 1970s that it began to branch into the three forms of micro finance institutions used today namely; commercial, quasi-commercial and nonprofit micro finance institutions. Each type’s characteristics equip it to serve its members differently Lindsay (2010). As micro finance institutions are developing, they are becoming overwhelmingly commercialized. Though commercial MFIs have the most financial support, their desire to profit prohibits them from being as effective as nonprofit organizations that only seek to help the poor. On the other hand, nonprofit micro finance institutions only seek social return, but do not have as
much capital to execute their goals as commercialized MFIs Lindsay (2010). Lindsay (2010) identified the critical performance measures in microfinance institutions as: number of people served / number of loans funded per year; total amount loaned per year, loan repayment rates per year; average loan size and average interest rate.

According to Sinha (1998), Microfinance provides financial services to low income or poor credit record clients which have become an alternative and easy way of borrowing money for the poor. However, extremely high interest rates from loans sometimes put low income people into worse poverty which in return affects the performance of these institutions. Due to the varying characteristics of this marginal group of clients and in order to ensure success of these institutions, Microfinance Institutions need to develop their own loan risk assessment and performance evaluation systems. Although data mining methods have the potential for developing such a system, the relative performance of the different data mining methods on such data is not known.

Although research by FSD (2008) and IJRET (2014) show that most Microfinance institutions apply the Credit scoring model for data mining, where the classification technique to analyze credit risk is applied on the dataset of the previous customers available at the financial institutions to distinguish them as good or delinquent customers, to find if there exists a relationship between the characteristics and reasons for delinquency of loans and accordingly choose an accurate classifier to implement on the new applicants. This paper focuses on comparing different data mining methods when applied to such data as loan repayment rates and number of people served for Microfinance institutions and their influence on the success of these institutions.
1.1.3 Microfinance Industry in Kenya

According to Johnston (2006), Kenya’s microfinance industry has come a long way since the 1980s, and particularly since the landmark Microfinance Intermediaries Act of 2006. The country now has over 20 deposit-taking microfinance intermediaries (MFIs) operating under a regulatory framework assessed by the Economist Intelligence Unit (EIU) as the best in Africa (EIU 2010). Overall, the EIU rates Kenya as having the second best business environment for MFIs in all of Africa (and one of the top ten in the world). Kenya has the second largest borrower base in the continent (MIX and CGAP 2010), and its largest savings and credit cooperatives (SACCO) movement Johnston (2006). As of June 2003, there were an estimated 3,460 legally constituted microfinance service providers in Kenya, including 3,397 savings and credit cooperatives and cooperative-like community-based intermediaries, 56 microfinance institutions (MFIs), four commercial banks, two building societies, and the Kenya Post Office Savings Bank. Excluded from this list were 17,305 rotating savings and credit associations (ROSCAs), 115,884 registered women groups, and 1,342 primary agricultural producer and marketing cooperative societies, also involved in providing credit countrywide. Research reveals that approximately 4 million Kenyans depend entirely on Microfinance Institutions Mutua (2011).

Information is a crucial input to all financial intermediation; yet the management information systems of many intermediaries in Kenya’s microfinance sector are inadequate or rudimentary Ndulu (2010). This does not bode well for the industry with almost no ability to use collateral, it needs to be able to manage information asymmetries and spot early signs of problematic borrowing. Indeed, in countries where the microfinance sector has recently been hit hardest by rising delinquencies, a key problem was multiple borrowing, which solid management information systems (MIS) could have spotted early on Chen et al (2010).
Hence this study aimed to look at importance of data mining techniques to the success of this industry. This study however focused on the 56 microfinance institutions (MFIs) that are registered and regulated by the AMFI as listed in Appendix I. This choice is by the fact that they have led in the implementation of ICT as compared to other financial institutions.

1.2 Statement of the problem

In the recent years microfinance industry has grown steadily and has been largely identified as a key instrument through which poverty can be alleviated. However, the rapid growth has dramatically changed the way in which microfinance organizations operate. This has created a major shift of participants from government and non-government organizations which provide credit to the very poor with the objective of alleviating poverty. The success of these institutions in achieving both very high repayment rates and, often, profits has attracted new entrants to the microcredit market Rosenberg (2007).

The performance of microfinance institutions in terms of institutional sustainability seems not encouraging despite the fact that international and national development programs have been giving high priority on sustainable microfinance to the poor for many years Yunus (1996). Locally, a few studies have been done on the issue of sustainability which included; sustainability of pilot purpose community telecentres in Kenya and Uganda Munyua (2003) which looked on the factors of sustainability of microfinance institutions. To ensure continuous measurement of performance, Microfinance Institutions need to develop their own performance evaluation systems.

There have been different arguments from past researchers on the effectiveness and benefits of data mining techniques. Sherwood et al (2010) argue that all that data mining faces is
challenges related to data integrity. The outcome of data mining can only be as good as the underlying data. Duplicate records, incomplete records, timeliness of updates, and human error all create data integrity problems. For instance names change is recorded differently to addresses and other identifiers change, or data are entered incorrectly. Original authors such as Coppock (2003) bring their own perspectives and biases to the interpretation of past events and these biases are more difficult to ascertain in historical resources. Due to the lack of control over external variables, it is rare that the entirety of historical data needed to fully address a specific problem is available for interpretation.

On the other hand Hermiz (1999), Pechenizkiy et al (2008) found four critical success factors for data mining projects as follows: organizations have a sound business problem and data mining is the right tool to for it; when pursuing the problem, the organization need to ensure the right type of data and sufficient quality and quantity of data is delivered to data mining; data mining is a complex process and needs to be managed accordingly and the project of using data mining as a tool to control the business problem should be a process of learning, regardless of the outcome, as well as there is no guarantee that the process was successful.

Given the above arguments, this study aimed to answer the following research questions: Are microfinance institutions in Kenya using data mining techniques? What challenges do microfinance institutions face when using these data techniques? What is the relationship between data mining techniques and the success of microfinance institutions?

1.3 Research Objectives

The objectives of this study were:

i. To establish the extent to which microfinance Institutions are using data mining
techniques.

ii. To establish the challenges of data mining in microfinance institutions.

iii. To establish the benefits of data mining in the microfinance Industry.

iv. To determine the effect of data mining on the success of microfinance institutions.

1.4 Importance of the study

The beneficiaries of this study are as follows:

Kenyan MFIs: - This study provides a theoretical reference to the management of Kenyan MFIs to analyze their data using data mining techniques to identify the strengths and weaknesses of this industry since its introduction into the country. Policy Makers in the Microfinance Industry in Kenya: - The study contributes towards understanding the state of Microfinance industry in Kenya and Information derived from this study will help in making future decisions and policies on how these important microfinance lenders can be well positioned in the country to grow and reach the millions of potential clients who do not currently have access to mainstream financial services as well as create wealth in in Kenya today and in the future. Academicians: - The study provides a foundation for knowledge that makes it possible for improvement as it will form a base for further research and development. This study also adds to the stock of global knowledge on data mining and organizational performance and provides the source of new ideas, methods, techniques and innovation across a whole range of disciplinary and multi-disciplinary areas.
CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

This chapter presents a review of literature on data mining and performance of microfinance institutions. The literature explores the concept of data mining by looking at different definitions of data mining by different scholars and the different data mining techniques. This chapter also looks at benefits and challenges of data mining, performance of Microfinance institutions as well as data mining and performance of organizations. This chapter will also discuss contribution of this study to scientific theories and finally we’ll discuss the conceptual framework of this study.

2.2 The concept of Data Mining

Data Mining is the process of extracting knowledge hidden from large volumes of raw data. The knowledge must be new, not obvious, and one must be able to use it Hulten (2001). Pregibon (1997) describes Data Mining as an analytic process that is designed to explore large amounts of data mostly business or market in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. Data Mining is often considered to be "a blend of statistics, AI (artificial intelligence), and data base research with the main objective being is prediction Pregibon (1997).

Hulten (2001) describes that the process of data mining consists of three stages: The initial exploration which usually involves data preparation that is cleaning data, data transformations and selecting subsets of records in order to identify the most relevant variables and determine the complexity and/or the general nature of models that can be taken
into account in the next stage. Model building or pattern identification with validation/verification which involves considering various models and choosing the best one based on their predictive performance. Deployment, which is the application of the model to new data in order to generate predictions. According to Clifton (2010), fundamentally, data mining is about processing data and identifying patterns and trends in that information so that you can decide or judge.

Data mining principles have been around for many years, but with the advent of more complex models and techniques it is even more prevalent. Data mining has caused an explosion in the use of more extensive data mining techniques, partially because the size of the information is much larger and because the information tends to be more varied and extensive in its very nature and content. Business-driven needs have changed simple data retrieval and statistics into more complex data mining. The business problem drives an examination of the data that helps to build a model to describe the information that ultimately leads to the creation of the resulting report.

2.2.1 History of Data Mining

The term "Data mining" was introduced in the 1990s, but data mining is the evolution of a field with a long history Lyman et al (2002). According to Ragman (2007) Data mining roots are traced back along three family lines: classical statistics, artificial intelligence and machine learning. Statistics are the foundation of most technologies on which data mining is built, for example regression analysis, standard distribution, standard deviation, standard variance, discriminant analysis, cluster analysis, and confidence intervals. All of these are used to study data and data relationships. Artificial intelligence or AI, which is built upon heuristics as opposed to statistics, attempts to apply human-thought-like processing to statistical problems.
Certain AI concepts which were adopted by some high-end commercial products, such as query optimization modules for Relational Database Management Systems (RDBMS). Machine learning is the union of statistics and AI. It could be considered an evolution of AI, because it blends AI heuristics with advanced statistical analysis. Machine learning attempts to let computer programs learn about the data they study, such that programs make different decisions based on the qualities of the studied data, using statistics for fundamental concepts, and adding more advanced AI heuristics and algorithms to achieve its goals.

Witten et al (2011) describes data mining, in many ways, as fundamentally the adaptation of machine learning techniques to business applications. Data mining is best described as the union of historical and recent developments in statistics, AI, and machine learning. These techniques are then used together to study data and find previously-hidden trends or patterns within.

### 2.2.2 Data Mining Techniques

According to Silltow (2006), Data mining can be performed with comparatively modest database systems and simple tools, including using off the shelf software packages. Complex data mining benefits from the past experience and algorithms defined with existing software and packages, with certain tools gaining a greater affinity or reputation with different techniques. For example a data warehouse provides data sourcing, preprocessing, mining, and analysis information in a single package, which allows one to take information from the source database straight to the final report output.

In the recent past introduction of very large data sets and large-scale data processing have been able to allow data mining to collate and report on groups and correlations of data that are more complicated. An entirely new range of tools and systems have been introduced
including combined data storage and processing systems. For example, one can now mine data with various different data sets, including, traditional SQL databases, raw text data, key/value stores, and document databases.

Silltow (2006) mentions the following core techniques that are used in data mining based on the type of mining and data recovery operation. Unfortunately, the different companies and solutions do not always share terms, which can add to the confusion and apparent complexity. These techniques include:

- **Association**: where a simple correlation is made between two or more items, often of the same type to identify patterns. **Classification**: describing multiple attributes to identify a particular class. For example, you can easily classify cars into different types. Classification maps a data item into one of several pre-defined categories. These algorithms normally output "classifiers", for example, in the form of decision trees or rules.

- **Link analysis**: determines relations between fields in the database. Finding out the correlations in given sets of data will provide insight for selecting the right set of system features for analysis.

- **Sequence analysis**: models sequential patterns. These algorithms can help us understand what (time-based) sequence of events is frequently encountered together. These frequent event patterns are important elements of the behavior profile of a user or program.

- **Clustering**: examining one or more attributes or classes and grouping individual pieces of data together to form a structure opinion. That is using one or more attributes as your basis for identifying a cluster of correlating results. **Prediction**: involves analyzing trends, classification, pattern matching, and relation. By analyzing past events or instances, you can make a prediction about an event. **Sequential patterns**: Often used for identifying trends, or regular occurrences of similar events. **Decision trees**: can be used either as a part of the
selection criteria, or to support the use and selection of specific data within the overall structure. They are often used with classification systems to attribute type information and with predictive systems where different predictions might be based on past historical experience that helps drive the structure of the decision tree and the output. Combinations: - Involves combining more than one technique for example Classification, clustering and decision trees to identify, refine, build and finally identify sequences and patterns on the classifications. Long-term (memory) processing: - as new information, events, and data points are identified, it is sometimes necessary to build more branches, or even entirely new trees, to cope with the additional information. This technique is mostly used during this time.

This study however focused on association, classification, clustering and prediction as the proposed data mining techniques due to their simplicity and are straightforward to implement.

2.2.3 Benefits of Data Mining

Newman (2013) argues that a lot of good can come from data mining and identifies the following as the top benefits of Data mining:

More Money: Money is always a good thing in business. When data is mined that unearths the kinds of information that can contribute to growth and more profits, it can result in serious cash. Once a business knows what their customers want, they can customize approaches and outreach. Computer algorithms can slice and dice everything from a customer's age and gender to credit scores and buying history. By carefully mining this information, analytics software can help identify patterns in customer behavior that can increase sales and reduce customer turnover. Huge retailers like Target (TGT) and Wal-Mart (WMT) have long taken advantage of data mining with costly servers and high-priced data
scientists. But as cost-effective, web-based alternatives emerge, small businesses are also putting these resources to use.

*Tap into New Markets:* Business owners can use the mined data from various databases to find out more information about potential consumers and identify any holes in the current tactics. This information can be useful in making decisions on new potential markets. The same technology that has made the customer faceless also offers us more information about them than we’ve ever had access to. Although often a user name can give us little clue to customers’ gender, age or location, computer cookies can tell us the important things about our customers what interests them, disinterests them, whether it is price, size or color that stops them from making that purchase. Such information allows a more level playing field for all those in the market, big and small. Where customer insight used to be the privilege of those with large marketing budgets, customer data mining is just as available to businesses and startups now as their larger competitors. This means the technical savvy enterprise was able to tap into different markets, tailor content and present wares to markets home and away.

*Share and Share Alike:* Sharing of information that may be useful to similar organizations. For example, some coalitions may share information on consumers in order to provide better services. This can be dangerous grounds, but if it’s legally acceptable, some business owners can access the data of other partner organizations too. This largely expands the availability of information and can provide more data and likely in turn more accurate data--to improve the bottom line, services and research. Advancements in technology have made relationship marketing a reality in the recent years. Technologies such as data mining have made customer relationship management a new area where firms can gain a competitive advantage. Particularly through data mining, the extraction of hidden predictive information from large
databases, organizations can identify valuable customers, predict future behaviors and enable firms to make proactive, knowledge-driven decisions.

*Learn from the Past:* Data mining past information and comparing it to the current situation can reveal a lot. Graphs can easily show any troubling sales years, spikes or other trends that should be taken into consideration. Seeing the flow of a business via data can provide insight that otherwise might be overlooked. For example, a business that knows there's a history of high sales in July can work on maximizing that month, while giving extra attention to periods where sales slack. To play it safe, business intelligence services might be in order no matter the business. Experts can help businesses determine what's legal and appropriate, as well as provide structure and security. If data mining is new to a business, it can be difficult to figure out where to start. Just figuring out how to weed out the right information can be time-consuming however, with a little research and maybe the help of some experts, it can open up a world of possibilities.

*Competitor analysis:* Data mining helps companies to get information that they can use effectively to stand out from competition. Extracted information helps to make important business decisions that would affect the whole business in a better way. It’s based on the assumption that you can predict future customer behavior by past performance, including purchases and preferences. This looks at when customers bought, and tries to predict when they will buy again. This type of analysis can be used to determine a strategy of planned obsolescence or figure out complimentary products to sell. Of course for a marketer to get any value out of a database, it must continue to grow and evolve. Database information from sales, surveys, subscriptions and questionnaires is fed and then customers are targeted based upon this intelligence. In a world where price wars occur, customers jump ship every time a
competitor offers lower prices. Data mining can help to minimize this churn, especially with social media. These results to quick and correct access to useful information which makes companies to concentrate more on decision making and other important processes and would make them gain competitive advantage over others.

2.2.4 Challenges of Data Mining

According to Han (2006) the challenges of Data Mining include:-

*Lack Individual privacy:* Data mining makes it possible to analyze routine business transactions and glean a significant amount of information about individual’s such as buying habits and preferences, banking information, customer details which interfere with privacy. The concerns about the personal privacy have been increasing enormously recently especially when internet is booming with social networks, e-commerce, forums, blogs. Because of privacy issues, people are afraid of their personal information being collected and used in unethical way that is potentially causing them a lot of troubles. Businesses collect information about their customers in many ways for understanding their purchasing behaviors trends. However businesses don’t last forever, some days they may be acquired by other or gone. At this time the personal information they own probably is sold to other or leak.

*Issue of data integrity:* data analysis can only be as good as the data that is being analyzed. A key implementation challenge is integrating conflicting or redundant data from different sources. This poses a great challenge and may impact on the accuracy of information derived or extracted. Information is collected through data mining intended for the ethical purposes can be misused. This information may be exploited by unethical people or businesses to take benefits of vulnerable people or discriminate against a group of people. In addition, data
mining technique is not perfectly accurate. Therefore if inaccurate information is used for decision-making, it will cause serious consequence.

*The issue of cost:* Data mining and data warehousing tend to be self-reinforcing. The more powerful the data mining queries, the greater the utility of the information being gleaned from the data, and the greater the pressure to increase the amount of data being collected and maintained, which increases the pressure for faster, more powerful data mining queries. This increases pressure for larger, faster systems, which are more expensive.

*Most Databases Dynamic:* Databases usually change continually therefore getting representation data may be a challenge. Many existing data mining systems require that all the representative data are given at once. If something is changed at a later time, the whole research process may have to be conducted again. An important challenge for data mining systems is to avoid this, and instead change its current rules according to updates performed which is quite expensive. The data are growing with unpredictable rate. Discovering knowledge in these data is a very expensive operation. Running data mining algorithms each time when there is a change in data is a challenging problem. Therefore updating knowledge dynamically will solve these problems. Due to the continuous, unbounded, and high speed characteristics of dynamic data, there is a huge amount of data in both offline and online data streams, and thus there is not enough time to rescan the whole database or perform a rescan as in traditional data mining algorithms whenever an update occurs. Furthermore, there is not enough space to store all the stream data for online processing. Therefore, a process model has been designed to guide the user through a sequence of steps that will lead to good results.

*Databases may be Huge:* The size of databases seems to be ever increasing. Having very much data is advantageous since they probably will show relations really existing, but the
number of possible descriptions of such a dataset is enormous. Some possible ways of coping with this problem, are to design algorithms with lower complexity and to use heuristics to find the best classification rules. Simply using a faster computer is a good solution. An important issue in data mining is how the various techniques for exploratory, visual, and particularly predictive data mining perform when applied to extremely large data sets. It is not uncommon in many domains of application to deal with data sets in the multiple gigabyte range, with tens of millions of observations. Analyzing data sets of this size will require some planning to avoid unnecessary performance bottlenecks, and inappropriate analytic choices. For example, using advanced neural network techniques to analyze all of 20 million observations is simply inappropriate, because in some cases it could take as long as several days to complete even using a designated supercomputer-class mainframe and the same information can be quickly extracted from the data by applying an appropriate sub-sampling method first, and then analyzing a reasonable subset of the input data.

**Difficulty in Accessing Data:** This may be as a result of scattered data throughout an organization or more commonly accessing data because it does not exist. This may also be due to the lack of a plan or strategy for what data is needed, how it can be obtained, how quality can be assured or improved and how it can be maintained.

### 2.3 Performance of Microfinance Institutions

Key performance measures and Data Mining management systems, methodologies and tools are closely related. An understanding of these relationships can help organizations accelerate productivity gains in their operations and thereby improve profitability Angoss (2011). Paris (1981) defines Measures of performance as the criteria that we believe show the impact of our work. The measures may be quantifiable or qualitative, but they are observable in some
way. Lindsay (2010) explains that performance effectiveness of micro finance institutions is measured on 6 criteria namely: - number of people served per week; - number of loans funded per year; total amount loaned per year; loan repayment rates per year; average Loan size; average interest rate.

Lindsay (2010) came up with the table below where a comparison between a non-profit MFI (Jamii Bora) and a commercial MFI (Blue Orchard) have been analyzed because they are leaders in their institution distinctions type. In terms of total amount loaned and amount loaned per year, Commercial micro finance institutions clearly have the advantage. Blue Orchard has loaned $714 million to date, an average of $131 million a year. Nonprofit MFIs loan the least amount per year; around $3.5 million which also puts them last for total amount loaned. This distinct advantage of commercial institutions to loan out the most amount of money is explained by their ability to mobilize deposits into loans. They also have higher capital than the other institutions because they have more financial support from investors. Jamii Bora has funded 145,000. Blue Orchard, the commercial bank has only given 800 loans to date. Aforementioned, commercial banks generally give loans out less frequently, but do so in larger amounts. The average loan size for Blue Orchard is more than $1,600 while the average loan size for Jamii Bora is $95. These definitions also express that commercial banks generally gave out few loans but in large secured amounts. There is no doubt that commercial banks would have the best repayment rate of the two Micro finance institutions. They select loan candidates based on the belief that they can pay back their loan which is backed by collateral. There is a direct relationship between loan repayment rates and loan interest rates. It would be presumed that an interest rate of .5% per week (Jamii
Bora) would have a much higher repayment rate than an institution that charged an average of 31% interest.

<table>
<thead>
<tr>
<th>Type of MFI</th>
<th>Non Profit</th>
<th>Commercial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of MFI</td>
<td>Jamii Bora</td>
<td>Blue Orchard</td>
</tr>
<tr>
<td>Total Amount Loaned</td>
<td>$21 Million</td>
<td>$714 Million</td>
</tr>
<tr>
<td>Amount Loaned Per Year ($)</td>
<td>$3.5 Million</td>
<td>$131 Million</td>
</tr>
<tr>
<td>Countries Represented</td>
<td>1</td>
<td>37</td>
</tr>
<tr>
<td>Number of Loans Funded</td>
<td>144,760</td>
<td>800</td>
</tr>
<tr>
<td>Average Loan Size</td>
<td>$95</td>
<td>$1,653</td>
</tr>
<tr>
<td>Repayment Rate</td>
<td>82.29%</td>
<td>97%</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>.5% per week</td>
<td>31% average</td>
</tr>
</tbody>
</table>

Table 4.1: Effectiveness of Micro Finance Institutions (Lindsay (2010))

2.4 Data Mining and Performance of organizations

According to Hermiz (1999), Data mining is an evolving technology whose implementation in the future has become mandatory in order for business organizations to remain competitive. In his article, he mentions four main critical success factors which organizations need to consider in a given situation for a successful data mining initiative: The organization needs to have a sound business problem which it acquires a solution for, and data mining technology is the right choice for solving this problem. Example in MFI’s: Who are my best customers? How do my customers behave in transactions? How many customers are in debt
and how many are using credit facilities less frequently? The organizations need to make sure that there is enough data to analyze as well as assure quality of data. The organizations have to know which methods are the best to use in analyzing algorithms.

Example: historical data on the consumer, such as the history of the transactions. The data mining process needs to be managed as exploratory data analysis. Organizations must ensure that the analysts are trained well and deduct the correct information which serves the purposes of the problem in the first place. Example: If the analysis is about searching for customer attrition, the organization needs a time frame with checkpoints on which data transformation can be difficult to attain as data changes and fluctuates, thus proceeding with prediction. The process of using data mining should be a learning experience. There are no guarantees that the projects which involve the usage of data mining will yield successful outcomes. Knowledge and experience are key skills which support people who seek to extract information which delivers valuable opportunities for organizations.

Sambuli & Crandall (2013) of iHub research carried out a study on viability of passive crowdsourcing as a data mining tool during elections in Kenya and their findings were that in the case of the Kenyan 2013 election, social media provided more information than other sources, although the level of severity of the incidents was different compared to traditional media; social media data contained smaller, real-time interest stories. Regardless of whether they were also picked up by traditional media or captured by Uchaguzi, such stories from social media were worth verifying and following up on as they were of interest to citizens, particularly if they lived in areas where incidents were taking place.
2.5 Theoretical Perspective

The theoretical perspective adopted in this paper is Critical Realism also known as Bashkar’s philosophy because it was Roy Bhaskar that first developed the concept of Critical Realism in 1975 from Realism Mingers (2000). Critical Realism is relatively new in Information Systems research. Critical Realism is a realist philosophy which states that there is a world outside and independent of our perception and we as humans are only capable of viewing some aspects of this world by using our own senses Easton (2009).

Since events are independent of our sensing, the causes that generate those events are also not perceived Easton (2009). There are three levels of reality, ontological domains Mingers (2004). Although they are separated domains i.e. fields, the empirical field is a subset of the field actual, and the latest is subset to the field real (see Figure 1).

---

**Figure 4.1: The three domains of the real. Source: (Mingers, 2004, p. 98)**

*Experiences (Empirical):* Represent our experience about the world and what we see in it. This view can be subjective and limited to the perspective view. This is a level considered to
be nominalist and our experiences as individuals are not mirrors of reality. We use our own perspective to view things that we experience Fisher, et al (2007). Events (Actual): Events occur in the world whether we perceive them or not. They are independent of our view on reality and thus objective. When they are perceived, they are observed within experiences. They belong to the second level of reality Fisher, et al. (2007) Easton (2009).

Mechanisms (Real): Mechanisms are the causes or reasons of events within the deepest and third level of reality.

However, at this level Critical Realism sustains that there are multiple mechanisms which may occur and they cannot be directly experienced, instead they need to be logically inferred from events, respectively, the events that we experience Fisher et al. (2007). Critical Realism differs from other approaches because it allows investigators to understand why events occur the way they do. The reason for choosing Critical Realism approach for this study is that only from this perspective was the research able to identify the real causes behind the events and understand why things are the way they are, i.e., what causes depletion of nonprofit Microfinance Institutions? What are the necessary and contingent relations for these factors to have an outcome? Which structures of Data Mining Techniques support the success of Microfinance Industry in Kenya? This theory contributed to the analysis of the influence of data mining on the success of microfinance institutions in Kenya.

2.6 Summary and Conceptual Framework

2.6.1 Conceptual Framework

In a conceptual framework there are three types of variables: dependent, independent and intervening variable. In this study, as conceptualized by the author (2014), the independent variables are the data mining techniques. These independent variables influence the
performance measures of MFIs which in this study are the dependent variables. These variables are the key performance measures or indicators that are used in performance management in microfinance institutions. The intervening variables are the moderating factors which can also influence performance of MFIs.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Dependent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Association</td>
<td>• Number of people served per week</td>
</tr>
<tr>
<td>• Classification</td>
<td>• Number of Loans funded per year</td>
</tr>
<tr>
<td>• Clustering</td>
<td>• Total amount Loaned per year</td>
</tr>
<tr>
<td>• Prediction</td>
<td>• Loan repayment rates per year</td>
</tr>
</tbody>
</table>

Moderating factors

- Number of employees
- Ownership

Figure 4.2: Conceptual Framework Source: researcher (2014)

2.6.2 Summary of Literature Review

Microfinance Industry has now evolved into a fertile land for data mining. In the last two decades, Microfinance Industry has evolved to a stage that gigantic amounts of data are constantly being generated and collected, and data mining and knowledge discovery becomes the essential scientific discovery process. We have proceeded to the era of data science and data engineering. Data Mining then becomes a subset of corporate organizational growth that focuses on whether IT sustains and extends the organization’s strategies and objectives. With the found four critical success factors for Data Mining as mentioned by Hermiz (1999),
Pechenizkiy et al (2008) namely: organizations have a sound business problem and Data Mining is the right tool to for it, when pursuing the problem, the organization need to ensure the right type of data and sufficient quality and quantity of data is delivered to Data Mining. Data Mining is a complex process and needs to be managed accordingly, and the project of using Data Mining as a tool to control the business problem should be a process of learning, regardless of the outcome, as well as there is no guarantee that the process was successful in mind. Microfinance institutions can successfully implement any technique for data mining successfully.

Despite the over-reliance on IT for their operations, the concept of data mining as an umbrella framework encompassing a wide spectrum of arrangements, including the measurement of benefits is yet to emerge. Most MFIs in Kenya are knowingly or unknowingly practicing data mining mainly driven by regulations. The current literature focuses more on the application of ICT by MFIs but less on whether this leads to value delivery, IT strategic alignment and managed IT risks. Despite the increasing need for Data mining in Kenyan MFIs, the current practice of IT in Kenyan MFIs (e.g. do they follow a structured framework? How is performance/value of the outcome measured?) remains not known and this presented a research gap that the researcher seeked to fill.
CHAPTER THREE: RESEARCH METHODOLOGY

3.1 Introduction

The chapter looks at the methods that were used in the study of data mining and performance of Microfinance Institutions in Kenya. This chapter is structured into research design, population of study, data collection and data analysis.

3.2 Research Design

A descriptive study design was used in this study. Descriptive study was preferred for it is used to obtain information concerning the current status of a phenomena and the purpose of this method is to describe “what exists” with respect to situational variables i.e. this helps to explain the relationship between and among variables. This method has been used successfully by various researchers and so the method is perceived to be the best in obtaining in-depth data. This kind of design allowed data to be collected within the setting of the respondents and data analysis inductively building from particular to general themes with the researcher interpreting the meanings of the data.

3.3 Population of the Study

Population of this study comprised of 56 Microfinance Institutions. The low number and accessibility of MFIs through a head office in Nairobi made a census the most ideal.

3.4 Sample Population

The population of interest in this study consisted of 56 respondents who are either IT officers or database administrators of the 56 microfinance institutions in Kenya. This study was limited to the institutions that are registered and regulated by the AMFI.
3.5 Data collection

This study was facilitated by the use of primary data. A semi-structured questionnaire was used for data collection in this study. The questionnaire addressed the objectives of the study. A questionnaire was administered in each institution. The questionnaires had five sections; Section A was used to collect profile data of the respondents, section B focused on Data mining techniques used in their organization, section C collected data on challenges of data mining techniques, Section D collected data on benefits of data mining techniques and Section E focused on the critical performance measures used in their institutions.

3.6 Data Analysis

Collected data from the questionnaires was coded and tabulated. This was to ensure comprehensive analysis. Objectives (i),(ii) and (iii) were measured using descriptive statistics whereby the findings was presented using tables, frequencies and percentages, graphical representations and quantitative measures of dependence. Objective (iv) was measured using the following regression model:

\[ Y = a_0 + a_1 x_1 + a_2 x_2 + e \]

Where by:

\( Y \) = Success Measures  
\( X_1 \) = Data mining techniques  
\( X_2 \) = Intervening variables  
\( e \) = error term  
\( a_0, a_1 \) and \( a_2 \) were the parameters to be estimated.
CHAPTER FOUR: DATA ANALYSIS, RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents the findings of the study established from the sample questionnaire in seeking to determine the data mining techniques and performance of Microfinance institutions in Kenya. The findings of the study were guided by this objective in determining the impact of the data mining techniques on performance of Microfinance institutions in Kenya. This chapter also explains the findings in comparison with relevant literature as established by other authors in the same field of study.

4.2 Response Rate

The targeted sample size was 56 Microfinance Institutions. Those filled and returned questionnaires were 46 Microfinance Institutions making a response rate of 82%. According to Mugenda and Mugenda (1999), a response rate of 50% is adequate for analysis and reporting; a rate of 60% is good and a response rate of 70% and over is excellent. This means that the response rate for this study which was 82% was excellent and therefore enough for data analysis and interpretation.

This high response rate can be attributed to the data collection procedures, where the researcher pre-notified the potential participants and applied the drop and pick method where the questionnaires were picked at a later date to allow the respondents ample time to fill the questionnaires.
Table 4.2: Response Rate

<table>
<thead>
<tr>
<th>Questionnaires administered</th>
<th>Questionnaires filled &amp; returned</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>56</td>
<td>46</td>
<td>82%</td>
</tr>
</tbody>
</table>

Source: Research Data

4.3 Descriptive Statistics

The study sought to determine the classification on the level of ownership of the institutions where they work; local branches available; numbers of years the institutions have been in operation and the designation level of respondents.

4.3.1 Classification of Ownership of Institution

The study sought to establish the ownership of the institutions where the respondents work. The findings were presented in the figure below.

Source: Research Data

From the findings in the figure above, the majority of the respondents at 45.7% represented those who indicated that the ownership of the Microfinance institution as local ownership.
36.9% accounted for respondents who indicated that Microfinance institutions have both local and foreign ownership while 17.4% accounted for respondents who indicated that microfinance institutions have foreign ownership.

4.3.2 Number of Local branches

The study sought to determine the number of local branches the Microfinance Institutions have. The findings were presented in the figure below.

![Number of Local branches](image)

**Source:** Research Data

It is evident from the findings that the majority of the respondents at 39.1% have 6 to 10 local branches; 32.6% indicated that the Microfinance Institutions have 1 to 5 local branches; 17.4% indicated that the Microfinance Institutions have 11 to 25 branches while 10.9% indicated that the Microfinance Institutions have 26 to 50 branches locally.

4.3.3 Number of years in operation

The study sought to determine the number of years the Microfinance institutions have been in operation. The findings were presented in the figure below.
From the findings in the figure below, the majority of the Microfinance institutions have been in operation for 11 to 25 years accounting for 32%. The findings established further that the Microfinance institutions that have been in operation for 6 to 10 years accounted for 26%. 20% of the institutions have been in operation for 1 to 5 years.

4.3.4 Designation Level of Respondents
The study sought to determine the designation level of the respondents and the findings were presented in the figure below.
The findings in the figure above showed that the majority of the respondents at 34% were business unit managers/supervisors in the institution. 23% of the respondents indicated that they had a designated position of a Chief Information Officer. 21% of the respondents comprised of the board of directors while 14% comprised of the executive managers. 8% of the respondents had other designated positions relevant to the Microfinance Institution.

4.4 Data Mining Techniques

The study sought to determine the data mining techniques the organization uses. The findings were presented in the figure below.

Source: Research Data

From the findings in the figure above, it is evident that the data mining technique mostly used in the organizations is association which accounted for 36%. 26% accounted for Microfinance institutions that use classification as a data mining technique where by a model or classifier is constructed to predict categorical labels. 16% represented the Microfinance institutions that used other means of data mining techniques in the organization. These other means of data mining techniques include link analysis, sequence analysis, sequential patterns,
decision trees, combinations and long-term (memory) processing. 12% accounted for microfinance institutions that use clustering as a data mining and 10% represented microfinance institutions that use prediction as a data mining where by the model is constructed to predicts a continuous-valued function, or ordered value, or trend.

4.5 Challenges of Data Mining Techniques

The study sought to determine the challenges that may have hindered successful adoption of appropriate data mining techniques in the organizations. The respondents were required to rate the extent to which they agreed or disagreed on the challenges hindering the successful adoption of appropriate data mining techniques. The findings are presented in the table below.

Table 4.3: Challenges hindering adoption of data mining techniques

Responses to the extent to which the respondents agreed to the challenges hindering successful adoption of appropriate data mining techniques in the organization

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs:- Too expensive</td>
<td>3.87</td>
<td>.1543</td>
<td>agree</td>
</tr>
<tr>
<td>Lack information privacy</td>
<td>3.62</td>
<td>.2151</td>
<td>agree</td>
</tr>
<tr>
<td>Issue of data integrity</td>
<td>3.46</td>
<td>.5436</td>
<td>neutral</td>
</tr>
<tr>
<td>Issue of dynamic databases where data grows at a very high rate</td>
<td>3.58</td>
<td>.1546</td>
<td>agree</td>
</tr>
<tr>
<td>Issue of accessibility where data is not stored in one place</td>
<td>3.51</td>
<td>.7112</td>
<td>agree</td>
</tr>
<tr>
<td>Lack of enough training on data mining</td>
<td>3.74</td>
<td>.4541</td>
<td>agree</td>
</tr>
<tr>
<td>Lack of clear understanding of data mining techniques.</td>
<td>3.82</td>
<td>.4568</td>
<td>agree</td>
</tr>
<tr>
<td>Lack of skilled personnel to impact the knowledge</td>
<td>3.42</td>
<td>.3351</td>
<td>neutral</td>
</tr>
<tr>
<td>Lack of clear communication strategies</td>
<td>3.81</td>
<td>.2547</td>
<td>agree</td>
</tr>
<tr>
<td>Absence of documentation on data mining</td>
<td>3.63</td>
<td>.3141</td>
<td>agree</td>
</tr>
<tr>
<td>Lack of working concept and knowledge by the board and executive management on issues related to Data mining</td>
<td>3.61</td>
<td>.1521</td>
<td>agree</td>
</tr>
</tbody>
</table>
In Table 4.3 above, the respondents’ decisions show some level of agreement to the fact that these challenges have actually hindered the effective adoption of data mining techniques in Microfinance institutions. Majority of the respondents’ agreed to the challenges. If these challenges of were not addressed in the organizations, a significant hindrance in adopting data mining techniques was reflected. The analysis of the respondents confirms that every microfinance institution faces its challenges some of which are similar and other specific to individual microfinance institution.

4.6 Benefits of Data Mining Techniques

The study sought to determine the benefits that the microfinance institutions enjoy from using mining techniques. The respondents were required to rate the extent to which they agreed or disagreed on the benefits of using data mining techniques in their organization. The findings were presented in the table below.

**Table 4.4: Benefits of Data Mining Techniques**

The table below shows responses to the extent to which the respondents agreed to the benefits of data mining techniques in the organization.
<table>
<thead>
<tr>
<th>Benefits</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribute to growth and more profits to the organization</td>
<td>3.86</td>
<td>.2154</td>
<td>Agree</td>
</tr>
<tr>
<td>Use of the mined data out more information about potential consumers and identify any holes in the current tactics.</td>
<td>3.67</td>
<td>.2851</td>
<td>Agree</td>
</tr>
<tr>
<td>Sharing of information that may be useful to similar organizations</td>
<td>3.94</td>
<td>.3541</td>
<td>Agree</td>
</tr>
<tr>
<td>Data mining pass information and comparing it to the current situation can reveal a lot i.e. learning from the past</td>
<td>3.54</td>
<td>.3874</td>
<td>Agree</td>
</tr>
<tr>
<td>Data mining helps to get information that can be used effectively to stand out from competition</td>
<td>3.62</td>
<td>.1541</td>
<td>Agree</td>
</tr>
<tr>
<td>Mined data helps to improve customer experience and increase profits</td>
<td>3.53</td>
<td>.2871</td>
<td>Agree</td>
</tr>
<tr>
<td>Data mining improves planning and decision making</td>
<td>3.91</td>
<td>.1644</td>
<td>Agree</td>
</tr>
<tr>
<td>Information about loans and credit reporting can be easily accessed</td>
<td>3.86</td>
<td>.1384</td>
<td>Agree</td>
</tr>
</tbody>
</table>

Source: Research Data

\[ X \geq 3.50 = \text{Agree} \]

The findings revealed that majority of the respondents agreed that data mining techniques greatly benefits their various microfinance institutions. These benefits were as presented in table 4.4 above. The benefits of data mining techniques have overall been identified as a contributor to growth and more profits to the organization. Data mining has been used to identify loopholes in the organization and also to give out more information about potential consumers. In so doing sharing of useful information to similar organization is made easier.

Data mining can pass information that if compared to the current situation can reveal a lot i.e. learning from the past according to Newman (2013) who identified learning form the past as one of the benefits of using data mining techniques.

### 4.7 Critical performance Measures

The study sought to determine the levels of critical performance of measures used the organizations in measuring performance within a specific period of time.
Table 4.5: Average representation of critical performance measures

<table>
<thead>
<tr>
<th>Critical performance Measures</th>
<th>Per week:</th>
<th>Per year:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of people served</td>
<td>100,000 clients</td>
<td>-</td>
</tr>
<tr>
<td>Total amount Loaned</td>
<td>-</td>
<td>9,000,000</td>
</tr>
<tr>
<td>Loan repayment rates</td>
<td>-</td>
<td>70%</td>
</tr>
<tr>
<td>Average Loan size</td>
<td>-</td>
<td>200,000</td>
</tr>
<tr>
<td>Average interest rate</td>
<td>-</td>
<td>8%</td>
</tr>
</tbody>
</table>

Source: Research Data

From the findings in the table above, the results revealed that on average, the number of people served with loans is approximately 100,000 clients in a week. The study further established the total amount loaned to the clients per annum is approximately Ksh. 9,000,000 on average. The loan repayment rates per annum are approximately 70%. The average loan size given to the clients is approximately Ksh. 200,000 per annum. The average interest rate per annum is approximately 8%.

4.8 Correlation Co-efficient

Correlation analysis was used to determine both the significance and degree of association of the variables. The correlation technique is used to analyze the degree of relationship between two variables. It varies between -1 and + 1 with both ends of the continuum indicating perfect negative and perfect positive relationship between any two variables respectively. The results of the correlation analysis are summarized in table 4.6
Table 4.6: Pearson Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>DM</th>
<th>Ass</th>
<th>Class</th>
<th>Clust</th>
<th>Pred</th>
<th>ROA</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of People Served</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tot. Amt Loaned</td>
<td>-.476</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loan Repayment rates</td>
<td>.027</td>
<td>-.172</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Loan Size</td>
<td>.018</td>
<td>.315</td>
<td>-.838</td>
<td></td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Average Interest Rates</td>
<td>-.423</td>
<td>-.039</td>
<td>-.194</td>
<td>.440</td>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td>ROA</td>
<td>.657</td>
<td>.623</td>
<td>.439</td>
<td>-.184</td>
<td>-.540</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Source: Research Data

From the findings we see that Classification as a data mining technique has a negative correlation on Average Loan size. This implies that Using Classification in data mining reduces the average loan size significantly as compared to the other techniques. Clustering on the other hand has a strong positive correlation on Average Interest rates implying that using clustering as a data mining technique to mine information on Interest rates increases or produces higher interest rates.

4.9 Regression Analysis

The study conducted regression analysis to establish the effect data mining on performance of microfinance institutions in Kenya using the following equation:

\[ Y = a_0 + a_1 x_1 + a_2 x_2 + e \]

Where by:

\( Y = \) Success Measures

\( x_1 = \) Data mining techniques

\( x_2 = \) Intervening variables
Regression analysis was done on 3 dependent variables namely: number of people served per week, number of Loans funded per year, total amount Loaned per year

The table below shows the results for the goodness of fit statistics.

4.9.1 Goodness of Fit Statistics

The study also established the goodness of fit for the model. This was to check on the significance of explanatory variable in explaining the variation in loan repayment.

Table 4.7 below gives a summary of the result.

### Table 4.7: Goodness of Fit Statistics

<table>
<thead>
<tr>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
<th>Durbin-Watson</th>
</tr>
</thead>
<tbody>
<tr>
<td>.812*</td>
<td>.754</td>
<td>.795</td>
<td>.787</td>
<td>1.817</td>
</tr>
</tbody>
</table>

Source: Research Data

Determination coefficients ($R^2$) were also carried out to determine the strength of the relationship between independent and dependent variables. The study established an adjusted $R^2$ of 0.795. $R^2$ of 79.5% indicates that 79.5% of the variation in the performance of microfinance institutions in Kenya is attributed to data mining and the control variables. Durbin Watson test was also run to establish if the model would be affected by autocorrelation. Since the DW value of 1.817 was close to 2, therefore it can be concluded that there was no autocorrelation among the model residual.
4.9.2 Analysis of Variance (ANOVA)

Table 4.8: Analysis of Variance (ANOVA)

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>110.450</td>
<td>5</td>
<td>30.075</td>
<td>35.037</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>22.659</td>
<td>1</td>
<td>.573</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>133.109</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Research Data  
Dependent Variable: Loan Repayment rates

The study used ANOVA statistics to establish the significance of the relationship between performance and data mining and the control variables. The ANOVA statistics presented in Table 4.8 was used to present the significance of regression model. An F-significance value of p < 0.000 was established which is less than p= 0.05 at 5% level of significance hence the model is significant for regression analysis.

4.9.3 Regression Coefficients

Multiple regression analysis was used to determine the significance of the relationship between the dependent variable and all the independent variables pooled together and also the effect on the dependent variables when the moderating variables are included. This analysis was used to answer the questions; how do the independent variables influence the dependent variable collectively; to what extent does each independent variable affect the dependent variable in such a collective set-up, what are the effects on the dependent variables when the moderating variables are included and which are the more significant factors? The results are given in the model summary in Tables 4.9, 4.10, 4.11, 4.12, 4.13 and 4.15.
Table 4.9: Regression Results: Number of people served per week

<table>
<thead>
<tr>
<th>Dependent No. of people served per week</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Explanatory variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association</td>
<td>6.4562</td>
<td>0.890008</td>
<td>0.092</td>
</tr>
<tr>
<td>Classification</td>
<td>2.2464</td>
<td>1.806656</td>
<td>0.021</td>
</tr>
<tr>
<td>Clustering</td>
<td>0.1351</td>
<td>2.392872</td>
<td>0.000</td>
</tr>
<tr>
<td>Prediction</td>
<td>0.2419</td>
<td>2.670336</td>
<td>0.012</td>
</tr>
<tr>
<td>Constant</td>
<td>0.09124</td>
<td>4.517277</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Source:** Research Data

\[ Y = 0.09124 + 6.4562 \text{ASS} + 2.2464 \text{CLS} + 0.1351 \text{CLUS} + 0.2419 \text{PRE} \]

From the output in table 4.9, we can see that the predictor variables of Clustering and Prediction have a significant effect on Number of people served because both of their p-values are 0.000 and 0.012 respectively. Classification also has a significant effect with a P-Value of 0.021. However, the p-value for Association (0.092) is greater than the common alpha level of 0.05, which indicates that it is not statistically significant hence, does not have a significant effect on the Number of people served.
Table 4.10: Regression Results on moderating variables: No. of people served per week

<table>
<thead>
<tr>
<th>Dependent</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of people served per week</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explanatory variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association</td>
<td>6.4562</td>
<td>0.890008</td>
<td>0.092</td>
</tr>
<tr>
<td>Classification</td>
<td>2.2464</td>
<td>1.806656</td>
<td>0.021</td>
</tr>
<tr>
<td>Clustering</td>
<td>0.1351</td>
<td>2.392872</td>
<td>0.000</td>
</tr>
<tr>
<td>Prediction</td>
<td>0.2419</td>
<td>2.670336</td>
<td>0.012</td>
</tr>
<tr>
<td>Moderating Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Employees</td>
<td>0.06402</td>
<td>0.048790</td>
<td>0.962</td>
</tr>
<tr>
<td>Ownership</td>
<td>0.03333</td>
<td>0.436022</td>
<td>0.666</td>
</tr>
<tr>
<td>Constant</td>
<td>0.09124</td>
<td>4.517277</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Source: Research Data

From the output in table 4.10, we see that the moderating variables; Number of employees and Ownership have p-values of 0.962 and 0.666 respectively which are much greater than the common alpha level of 0.05. This indicates that they are not statistically significant hence, do not have any effect on the Number of people served.

Table 4.11: Regression Results: Number of Loans funded per year

<table>
<thead>
<tr>
<th>Dependent</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Loans funded per year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explanatory variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association</td>
<td>5.3451</td>
<td>0.800905</td>
<td>0.126</td>
</tr>
<tr>
<td>Classification</td>
<td>3.1352</td>
<td>1.914462</td>
<td>0.000</td>
</tr>
<tr>
<td>Clustering</td>
<td>0.1950</td>
<td>2.429872</td>
<td>0.016</td>
</tr>
<tr>
<td>Prediction</td>
<td>0.2308</td>
<td>2.543015</td>
<td>0.032</td>
</tr>
<tr>
<td>Constant</td>
<td>0.08054</td>
<td>4.517677</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Source: Research Data
From the regression analysis, the following model was established:

\[ Y = 0.08054 + 5.3451\text{ASS} + 3.1352\text{CLS} + 0.1950\text{CLUS} + 0.2308\text{PRE} \]

From the output in table 4.11, we can see that the predictor variables of Classification, Clustering and Prediction have a significant effect on Number of Loans funded per year as they all have p-values that are less than 0.05 i.e. 0.000, 0.016 and 0.032 respectively. Association on the other hand has a p-value of 0.126 which indicates that it is not statistically significant hence, does not have a significant effect on the Number of Loans Funded per year.

### Table 4.12: Regression Results on moderating variables: No. of Loans funded per year

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Loans funded per year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Explanatory variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association</td>
<td>5.3451</td>
<td>0.800905</td>
<td>0.126</td>
</tr>
<tr>
<td>Classification</td>
<td>3.1352</td>
<td>1.914462</td>
<td>0.000</td>
</tr>
<tr>
<td>Clustering</td>
<td>0.1950</td>
<td>2.429872</td>
<td>0.016</td>
</tr>
<tr>
<td>Prediction</td>
<td>0.2308</td>
<td>2.543015</td>
<td>0.032</td>
</tr>
<tr>
<td><strong>Moderating Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Employees</td>
<td>0.09014</td>
<td>0.184508</td>
<td>0.601</td>
</tr>
<tr>
<td>Ownership</td>
<td>0.05301</td>
<td>0.326602</td>
<td>0.943</td>
</tr>
<tr>
<td>Constant</td>
<td>0.08054</td>
<td>4.517677</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Source: Research Data**

From the table above, the moderating variables; Number of employees and Ownership have p-values of 0.601 and 0.943 respectively. This implies that they are not statistically significant hence; do not have any effect on the Number of loans served per year.
Table 4.13: Regression Results: Total amount Loaned per year

<table>
<thead>
<tr>
<th>Dependent Total amount Loaned per year</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association</td>
<td>5.1264</td>
<td>0.890102</td>
<td>0.222</td>
</tr>
<tr>
<td>Classification</td>
<td>2.2655</td>
<td>2.837094</td>
<td>0.003</td>
</tr>
<tr>
<td>Clustering</td>
<td>0.2351</td>
<td>2.190642</td>
<td>0.000</td>
</tr>
<tr>
<td>Prediction</td>
<td>0.1416</td>
<td>1.963138</td>
<td>0.001</td>
</tr>
<tr>
<td>Constant</td>
<td>0.06024</td>
<td>3.200341</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Source: Research Data

From the regression analysis, the following model was established:

\[ Y = 0.06024 + 5.1264 \text{ASS} + 2.2655 \text{CLS} + 0.2351 \text{CLUS} + 0.1416 \text{PRE} \]

From the output in table 4.13, we see that Clustering, Classification and Prediction have a significant effect on Total amount loaned per year with the \( p \)-values of 0.000, 0.003 and 0.001 respectively. Association on the other hand does not have a significant effect on Total amount loaned per year with a \( p \)-value of 0.222.

Table 4.14: Regression Results on moderating variables: Total amount Loaned per year

<table>
<thead>
<tr>
<th>Dependent Total amount Loaned per year</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association</td>
<td>5.1264</td>
<td>0.890102</td>
<td>0.222</td>
</tr>
<tr>
<td>Classification</td>
<td>2.2655</td>
<td>2.837094</td>
<td>0.003</td>
</tr>
<tr>
<td>Clustering</td>
<td>0.2351</td>
<td>2.190642</td>
<td>0.000</td>
</tr>
<tr>
<td>Prediction</td>
<td>0.1416</td>
<td>1.963138</td>
<td>0.001</td>
</tr>
<tr>
<td>Moderating Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Employees</td>
<td>3.21004</td>
<td>0.650801</td>
<td>0.971</td>
</tr>
<tr>
<td>Ownership</td>
<td>1.25001</td>
<td>0.821067</td>
<td>0.416</td>
</tr>
<tr>
<td>Constant</td>
<td>0.06024</td>
<td>3.200341</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Source: Research Data
The table 4.14 above shows that the moderating variables; Number of employees and Ownership are not statistically significant hence; do not have any effect on the Total amount Loaned per year have with p-values of 0.971 and 0.416 respectively.

4.10 Summary of the findings and interpretations

The study had a response rate of 82% which was quite impressive. This high response rate was attributed to the data collection procedures, where the researcher pre-notified the potential participants and applied the drop and pick method where the questionnaires were picked at a later date to allow the respondents ample time to fill the questionnaires. The study also shows that most Microfinance Institutions are owned locally with most institutions having between 6-10 local branches. The study reveals that majority of the institutions use Association as a data mining technique. If all the explanatory variables are held constant, the study shows that Classification, Clustering and Prediction have a significant effect on Number of people served, Number of Loans served per year and Total amount loaned per year as their p-values are less than the common alpha level of 0.05 making them statistically significant. While Association has a greater P-value making it not statistically significant hence has very low effect on Number of people served, Number of Loans served per year and Total amount loaned per year as dependent variables.
CHAPTER FIVE: SUMMARY OF FINDINGS, CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

In this chapter, we discuss the summary of the main study findings. The chapter also covers conclusion, recommendations limitations and suggestions of the study.

5.2 Summary of findings

The purpose of the study was to establish the effect of data mining techniques on the performance of the microfinance institutions in Kenya. The study shows that most Microfinance Institutions are owned locally with most institutions having between 6-10 local branches with majority of the institutions using Association as a data mining technique. The findings indicated that majority of the respondents agreed to the fact that every microfinance institution faces its challenges; some of which are similar and others specific to individual microfinance institution with the main challenges of data mining as discussed being: lack individual privacy; issue of data integrity; issue of cost; the fact that most databases dynamic; databases may be huge and difficulty in accessing data. However, the study further indicated that use of data mining techniques also comes with some benefits. Such benefits as discussed in this study included more money resulting to company profits; ability to tap into new markets; sharing of information among organizations; learning from the past and competitor analysis. The benefits of data mining techniques have overall been identified as a contributor to growth and more profits to the organization by identifying loopholes in the organization and also giving out more information about potential consumers.

The study also established that using Classification as a data mining technique introduces negative correlation on Average Loan size. This means that Classification in data mining
reduces the average loan size significantly as compared to the other techniques. Clustering on the other hand showed a strong positive correlation on Average Interest rates implying that using clustering as a data mining technique to mine information on Interest rates increases or produces higher interest rates. Based on the regression analysis the study shows that Classification, Clustering and Prediction as data mining techniques tend to have a significant effect on Number of people served, Number of Loans served per year and Total amount loaned per year as their p-values were than the common alpha level of 0.05 as displayed in figures; 4.9, 4.11 and 4.13 respectively. This indicates than they are statistically significant. Association on the other hand had a p-value that was greater than the common alpha level of 0.05, meaning it is not statistically significant hence has very low effect on Number of people served, Number of Loans served per year and Total amount loaned per year as dependent variables. The study also established that the moderating variables used; Number of employees and Ownership were not statistically significant with their p-values greater than 0.05. This therefore means that they do not have any effect on Number of people served, Number of Loans served per year and Total amount loaned per year as dependent variables. Results from the study therefore indicate that many of the respondents felt that data mining techniques positively impact on the performance on Microfinance institutions.

5.3 Conclusion

The factors affecting successful adoption of data mining techniques have been tested and proven from the research data. These challenges in the process need to be reviewed. Microfinance institutions should act as models to the other business fraternity in implementation and adoption of data mining techniques for performance management. It’s evident that there is need to look into Association as a data mining technique to resolve the
factors affecting it to ensure that it has a significant impact like the other data mining techniques.

5.4 Recommendation

Effective adoption of data mining techniques to assist in performance management is a function of many facets as evidenced by the study. This is strengthened by the responses in the research. The microfinance institutions have to have the best systems to ensure that transparency and data integrity are practiced. This in return ensures efficiency and validity of data that will aid in decision making.

Every officer with a responsibility on mining company data should be well endowed with necessary theory and skills necessary to ensure conformance and the championing of policies and procedures on information and data management structures. Organizational performance management and risk management are central to the realization of value, strategic alignment and management of IT related risks.

The organizations need to make sure that there is enough data to analyze as well as assure quality of data. Organizations should ensure that the process of using data mining should be a learning experience. Organizations should also ensure that challenges affecting data mining are discovered and tackled in time to ensure continuous data integrity.

The results of the study reinforce the possibility of application of data mining techniques to the other areas of microfinance industry, particularly in customer segmentation and prediction as well as financial management.
5.5 Limitation of the study

The study relied on both secondary data and questionnaires for the analysis. Some microfinance institutions were not willing to provide full information on their organizations financial statements. Therefore the result of the findings based on the data might not give true result.

5.6 Suggestions for further Research

The study suggests that similar studies should be done on deposit taking microfinance institutions that have been in operation for a long time to establish if data mining contributes negatively or positively on their growth. It would be helpful to replicate the study in another setting particularly in other sectors of the economy where data mining is practiced or is in the process of adoption. The factors affecting successful adoption of data mining in institutions in Kenya are varied and may even be better brought out if the study is extended to other financial institutions and even conventional business organizations.
REFERENCES


IJRET (2014).Improving the Credit Scoring Model of Microfinance Institutions by Support Vector Machine, Volume: 03, Special Issue: 07


Newman D., (2013).*Introduction to Data Mining*. California: Irvine

Ndulu K J. (2010) Factors affecting Institutional Transformation a case for a microfinance regulatory framework in Kenya, Masters in Public Finance project University of Stellenbosch

Oscar, M., et al (2009); A Data Mining & Knowledge Discovery Process Model. *In Data Mining and Knowledge Discovery in Real Life Applications*.


Rosenberg (2007). Banker to the Poor: Micro Lending and the Battle Against World Poverty, New York: Public Affairs


Silltow J., (2006).*Data mining tools and techniques*. United Kingdom


APPENDIX I: - LIST OF MFIs

1. AAR Credit Services
2. Aga Khan Foundation
3. Africa Credit Ltd
4. Africashare Partnership
5. American Advance
6. Amsoft
7. Bimas Ltd
8. Bidii Development Programme
9. Biashara Factors Limited
10. Blue Limited
11. BCF Kenya Limited
12. Canyon Rural Credit Limited
13. Elite Microfinance
14. Faulu Kenya DTM Limited
15. Fusion Capital Ltd
16. Jamii Bora
17. Jitegemea Credit Scheme
18. Juhudi Kilimo Company Limited
19. K-rep Development Agency
20. KADET
22. Kenya Entrepreneur Empowerment Foundation
23. Kenya Women Finance Trust
24. MIC Microcredit limited
25. Micro Africa
26. cro-Finance Institutions
27. FSI Capital Ltd
28. Molyn Credit Limited
29. Oiko Credit
30. Opportunity International
31. Pamoja Women Development Programme
32. Renewable Energy Technology Assistance Programme
33. SISDO
34. SMEP
35. Swiss Contact
36. Taifa Option Microfinance
37. U & I Microfinance Limited
38. Uwezo Microfinance Bank
39. WEEC
40. Yehu Enterprises Support Services
41. Capital Sacco Society Ltd
42. Ace Capital & Credit Ltd
43. Rafiki
44. Business Capital Access Ltd
45. Fast Link Capital Limited
46. Faidi Ltd
47. Equinox Solutions Ltd
48. Emara Credit Ltd
49. Elite Group
50. Creative Leadership & Microenterprise Project
51. Changamka Kenya
52. Century Ltd
53. Jambo Capital Micro-finance
54. FSD Kenya
55. Jisaidie Trade Credit Ltd
56. Jitegemee Trust Ltd