BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF SCIENCE AND ENGINEERING

DEPARTMENT OF STATISTICS AND MATHEMATICS

Determinants of Default In Microfinance Institutions In Zimbabwe

BY

ALICE CHAMUNORWA

(B200257A)

A DISSERTATION SUBMITTED TO BINDURA UNIVERSITY IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF THE BACHELOR OF SCIENCE HONOURS DEGREE IN STATISTICS AND FINANCIAL MATHEMATICS.

SUPERVISOR: MR KUSOTERA

JUNE 2024

DECLARATION

I declare that the work and findings in this research is my own, which was done by me without copying or extracting from prior sources without suitable acknowledgment of sources.

Chamunorwa Alice 10/06/2024

B200257A Signature Date

Certified by

Mr Kusotera …………………… ……….10/06/2024

Supervisors Signature Date

11/06/2024

Dr Magodora ……………………………. ………………………..

Chairperson Signature Date

DEDICATION

TO MY PARENTS

Because you gave me the belief that anything is possible

TO MY BROTHERS AND SISTERS,

For the reason that you made everything possible.

TO MY FRIENDS,

I would especially like to thank all of my friends for supporting me during my entire degree program.

ACKNOWLEDGEMENT

First and foremost, I want to express my gratitude to the Almighty God for providing me with the chance, direction, clarity of thought, and patience to start and finish my study. Secondly, I also appreciate my supervisor Mr B. Kusotera, for being patient with me during the authoring of this research. Thirdly I would like to thank my family especially my mother and my father for constantly monitoring my progress and their love and upkeep. I also want to thank the CAMFED Association sisters (Winnie Farao and many more) for their guidance, mentoring and everlasting support. My colleague, SFM class in 2024 has made great contributions. You are a source of motivation and we travelled together and each of you played a memorable role.

May the Almighty God bless you all and grunt you your heart desires.

ABSTRACT

This study investigated the determinants of default in Microfinance Institutions (MFIs) using a logistic regression model. Utilizing R-programming, the analysis covered a 2-decade period from 2002 to 2022 of secondary data collected from ZIMSTAT, RBZ, World data, and utilized data from MFIs. The key determinants examined are Gross Domestic Product (GDP), interest rates, and inflation rate. The results show that GDP and interest rates have a significant impact on defaults in MFIs. Specifically, a decrease in GDP and an increase in interest rates lead to a higher likelihood of default. Inflation rate, however, does not have a significant effect. The microfinance institution should blacklist or write off clients in default so as to maintain their profitability status. They can also take legal action after they have carried out reschedule and client failed to adhere to the new repayment schedule while remaining in default. Financial institutions should closely monitor the economic indicators identified as significant predictors of non-performing loans, such as years, GDP per capita, and interest rates. Additionally, they should consider diversifying their loan portfolios to reduce concentration risk and exposure to specific economic factors that may contribute to non-performing loans. The findings of this study have important implications for MFIs, policymakers, and regulators seeking to reduce defaults and improve the overall sustainability of microfinance programs. The comprehensive analysis underscores the urgency for proactive measures in managing microfinance risks amidst fluctuating economic conditions.

ACRONYMS

AIC-Akaike Information Criteria

MFI-Microfinance Institutions

GDP-Gross Domestic Product

NPL-Non performing loans

NGO-Non-Governmental Organisations

INT-Interest rates

INFL-Inflation rates

MLE-Maximum Likelihood Estimator

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

Microfinance is viewed as a means of increasing the extremely poor's income (Ledgerwood et al., 2013). Numerous studies have demonstrated the effectiveness of microfinance as a tool for reducing poverty by allowing the impoverished to increase their income, build assets, and lessen their economic vulnerability (Batra and Sumanjeet, 2012). Credit is the term for the legal arrangement in which a lender executes a transaction for which the borrower agrees to repay the lender within a predetermined time frame and subject to certain requirements. The terms and conditions may include interest rates and a monthly or annual payment schedule for repayment. The primary entities that provide credits in the form of loans are banks and microfinance organizations. The management of credit operations has become an urgent necessity for any bank due to uncertainty surrounding the borrower's ability and willingness to repay these loans on time. This chapter starts with a brief overview of the study's background, then moves on to the problem statement, objectives, questions, and hypotheses. The chapter also provides an explanation for the study's scope of the research. A dissertation outline concludes the chapter.

1.2 BACKGROUND OF THE STUDY

The 1970s is the period when microfinance first emerged. Muhammad Yunus created it at the Grameen Bank in Dhaka, Bangladesh. Later, in the 1980s, the microfinance industry spread to Latin America, India, Brazil, and Africa. Globally, there are roughly 10,000 microfinance organizations. It was once known as microcredit. In the early 1990s, microfinance took the place of microcredit (Helms 2006). When citizens in Zimbabwe were organized into savings clubs in the 1960s, private moneylenders provided riskier loans known as chimbadzo, or exploitative lending. This is when the country's microfinance industry began. The Agricultural Finance Cooperation had a key role in providing smallholder farmers with loans in the early 1980s (Mago 2013).

The nation saw an increase in the microfinance industry. According to Reserve Bank of Zimbabwe data from 2020, there are 220 microfinance institutions registered in Zimbabwe, of which 212 offer just credit and 8 accept deposits. By 2020, thirty-one microfinance institutions had closed.

Armendariz and Labie (2011) claims that a variety of financial techniques are used in microfinance to assist the underbanked. As a result, microfinance has been used as a technique to reduce poverty in many nations that struggle with high rates of both formal unemployment and poverty. Since commercial banks typically lend to medium-sized and larger businesses because they are thought to be more creditworthy than microenterprises, which are thought to be associated with relatively high costs and risks, Microfinance Institutions (MFIs) are the primary funders of microenterprises in Africa and other developing regions (Emeni, 2008). According to Monyau and Bandara (2012), Zimbabwe's economy is still unstable due to an unsustainable level of deindustrialization and informality. Zimbabwe's economy is slowing down, according to the African Economic Outlook (2012). These factors include outmoded technology, structural constraints including power shortages and infrastructural shortfalls, and liquidity issues (lack of and high cost of financing). According to the World Bank (2013), Zimbabwe had a GDP of USD 12 billion in 2013 and is classified as a low-income developing nation. Due to these current circumstances in Zimbabwe, the microfinance industry is seen as being crucial to the country's economic expansion. In Zimbabwe, the microfinance industry is acknowledged as having a significant role in fostering inclusive financial systems that promote economic growth and development (Reserve Bank of Zimbabwe, 2014b). Microfinance has been essential in bolstering financial and economic development in Zimbabwe, where the economy has become more informal, financial inclusion is low, and commercial bank participation in microfinance activities is restricted (Makina, 2012).

As a result, this dissertation acknowledges the significance of MFIs to Zimbabwe's economy and aims to delve into the determinants factors leading to default.

1.3 History of Defaults

The microfinance industry in Zimbabwe has faced significant challenges with loan default over the past two decades. In the early 2000s, the country experienced a severe economic crisis, marked by hyperinflation and political instability, which had a detrimental impact on the microfinance sector (Biti, 2008).

During this period, many MFIs in Zimbabwe struggled to maintain sustainable operations, as high default rates eroded their loan portfolios (Mudzingiri & Matandare, 2018). A study by Ncube and Makaudze (2014) found that the average loan default rate among Zimbabwean MFIs was around 35% during the economic crisis of the 2000s.

The causes of the high default rates were multifaceted. Musona and Coetzee (2001) attributed the problem to a combination of factors, including the deteriorating macroeconomic conditions, lack of borrower creditworthiness, and inadequate risk management practices by MFIs. Additionally, Bett and Memba (2017) highlighted the impact of political interference and corruption on the operations of MFIs in Zimbabwe during this period.

As the Zimbabwean economy stabilized in the late 2000s, the microfinance sector began to recover, but the legacy of high default rates continued to linger (Makina, 2011). A study by Gono and Jaravaza (2013) found that, even in the post-crisis era, Zimbabwean MFIs were still experiencing default rates ranging from 20% to 30%.

More recently, the COVID-19 pandemic has once again put significant strain on the microfinance industry in Zimbabwe, leading to a resurgence in loan default rates (Chipunza & Munangagwa, 2021). Researchers have emphasized the need for MFIs to strengthen their credit risk management strategies and explore innovative approaches to mitigate the impact of external shocks on their loan portfolios (Majoni et al., 2016).

In summary, the history of default in the Zimbabwean microfinance sector has been shaped by the country's broader economic and political challenges, underscoring the importance of addressing systemic factors to ensure the long-term sustainability of MFIs.

1.4 Evolving landscape of defaults in MFIs

The history of default in Microfinance Institutions (MFIs) is a complex and multifaceted narrative, reflecting the dynamic nature of the microfinance industry and the challenges faced by institutions in serving underbanked populations.

In the early days of the microfinance movement, researchers observed relatively low default rates, often attributed to the success of the Grameen Bank model and the power of group lending mechanisms (Khandker, 2005; Ghatak & Guinnane, 1999). Besley and Coate (1995) highlighted how social collateral and peer monitoring in group lending could incentivize borrowers to repay their loans, contributing to the industry's initial low default rates.

However, as the microfinance sector expanded and diversified, the landscape of default began to evolve. Ledgerwood et al. (2013) noted that the transition from group-based lending to individual-based lending, as well as the entry of commercial banks and other formal financial institutions into the microfinance space, introduced new complexities and potential risks. The authors emphasized the importance of robust credit risk management strategies, including thorough borrower assessment and effective loan monitoring, to mitigate the rising default rates.

Moreover, Microfinance Information Exchange (MIX, 2022) reported that macroeconomic conditions, such as economic downturns, natural disasters, and political instability, have played a significant role in shaping default patterns in MFIs. These exogenous factors can disrupt the livelihoods of borrowers, leading to increased delinquency and default rates.

In response to these challenges, MFIs have implemented various strategies to enhance their resilience and manage default risks. Hermes and Lensink (2011) highlighted the importance of diversifying product offerings, strengthening client relationships, and fostering financial literacy among borrowers to improve repayment rates. Additionally, the development of credit bureaus and the use of advanced data analytics have enabled MFIs to make more informed credit decisions and better manage their portfolio risks (Cull et al., 2018).

As the microfinance industry continues to evolve, the understanding of default patterns and the strategies to mitigate them will remain a critical area of focus for MFIs, policymakers, and researchers alike. By learning from the past and adapting to the changing landscape, the microfinance sector can strive to maintain its mission of financial inclusion while ensuring the long-term sustainability of its operations.

1.5 Statement of the problem

In recent decades, the microfinance industry has grown significantly, giving underprivileged communities access to financial services. However, high loan default rates continue to be a major challenge for the sustainability and profitability of MFIs. While existing literature has examined various factors influencing loan default, there is a need for a more comprehensive understanding of the key determinants of loan default in the microfinance sector, especially in the context of evolving macroeconomic conditions and the growing complexity of the industry. This piece of work aims to investigate the critical determinants of loan default in MFIs, including the roles of inflation rates, GDP growth, interest rates, and non-performing loans. Through an analysis of these elements' effects on loan default, the study aims to offer insightful information that can guide the creation of more successful credit risk management plans and guidelines for MFIs. The results of this study will support continued initiatives to improve the financial performance of MFIs

1.6 Objectives

1. To examine the influence of inflation rates on loan default rates in MFIs.

2. To investigate the impact of interest rates charged by MFIs on their loan default rates.

3. To assess the role of non-performing loans as a determinant of loan default in the microfinance sector.

1.7 Research questions

1.How does a change in inflation rates affect the likelihood of loan default in MFIs?

2.How does interest rates affect the likelihood of loan default in MFIs?

3.How does the proportion of non-performing loan affect the likelihood of loan default in MFIs?

1.8 Justification of the study

The study is relevant to all MFIs in Zimbabwe and understanding the factors contributing to default can inform lending decisions, enabling MFIs to create more effective loan products and services. It can also inform regulatory policies and guidelines for MFIs, promoting a more stable and sustainable microfinance sector. In addition, the findings can contribute to the development of microfinance industry as a whole, informing best practices and policy recommendations.

1.9 Significance of the study

By identifying the key determinants of loan default, this study will provide MFIs with a better understanding of the factors that can impact their financial sustainability This information can help create more effective credit risk management plans and policies, which in turn can boost MFI productivity and profitability. The results of this study can provide guidance to regulatory bodies and policymakers in the microfinance industry, enabling them to create more efficient policies and support systems to deal with the difficulties associated with loan default. This can contribute to the overall stability and growth of the microfinance industry. Reducing loan default rates can enable MFIs to expand their outreach and provide financial services to a broader segment of the population, particularly the underserved and financially excluded communities. This can have a direct impact on improving financial inclusion and promoting economic development. This research will add to the corpus of knowledge already available on the factors that influence loan default in the microfinance industry. A thorough examination of variables including GDP, inflation, interest rates, and non-performing loans will yield insightful information that will help shape future studies and direct the creation of theoretical frameworks in the field of microfinance.

1.10 Limitations of the Study

The research may be restricted to a particular geographic area or subset of MFIs, which could limit the applicability of the conclusions to the larger microfinance sector. Expanding the scope to include a more diverse sample of MFIs across different regions could enhance the external validity of the study. The study's findings may be subject to the availability and quality of data related to the key determinants of loan default. Incomplete or inaccurate data can impact the reliability of the analysis and the conclusions drawn. A number of contextual elements, including social, cultural, and regulatory contexts, which may not be fully reflected in the study, can have an impact on loan default in MFIs.Incorporating these contextual elements could provide a more holistic understanding of the determinants of loan default. Loan default patterns and the influence of the determinants may change over time due to evolving macroeconomic conditions, technological advancements, and changes in the microfinance industry. The study may be limited in its ability to capture the dynamic nature of these relationships. The robustness of the results could be impacted by the study's inherent limitations in the research methodology and analytical procedures chosen. Employing a mixed-methods approach or exploring alternative analytical frameworks could help address these limitations.

1.11 key terms

1.11.1 Inflation Rate

The steady rise in the average cost of goods and services over time in an economy is known as inflation. The rate at which the general level of prices rises is known as the inflation rate (Mishkin, 2016). The purchasing power of microfinance borrowers may be diminished by inflation, making it more challenging for them to repay their loans. In addition to causing uncertainty and economic instability, high rates of inflation can also raise the likelihood of loan defaults in the microfinance industry (Tchakoute-Tchuigoua, 2016).

1.11.2 Gross Domestic Product (GDP)

According to Mankiw (2020), GDP is the total monetary worth of all finished goods and services produced inside a nation's boundaries during a given period of time, usually a year. GDP growth is often used as a proxy for macroeconomic conditions and economic performance. Periods of economic growth are generally associated with higher incomes and better employment opportunities, which can positively impact the repayment capacity of microfinance borrowers. Conversely, economic downturns and recessions can lead to increased loan default rates in the microfinance industry (Ahlin et al., 2011).

1.11.3 Interest Rates

Interest rates, which are represented as a percentage of the principal amount, are the costs associated with borrowing money. One important factor in microfinance is the interest rates that MFIs charge on their loan products (Cull et al., 2016). A higher interest rate may make it harder for microfinance borrowers to repay their loans by adding to their debt load. Since borrowers may find it difficult to fulfill their repayment responsibilities, excessively high interest rates levied by MFIs may also be a factor in increased loan default rates (Armendáriz & Morduch, 2010).

1.11.4Non-Performing Loans (NPLs)

When a borrower is unable to make scheduled payments for an extended period of time (usually 90 days or more), the loan is considered non-performing (NPL) (Roodman & Qureshi, 2006).

High levels of NPLs can signal underlying issues in the MFI's lending practices, risk management, or the broader economic environment, and can contribute to further deterioration in the MFI's financial performance and sustainability, leading to higher loan default rates (Gonzalez, 2007).

1.11.5 Loan Default

Loan default refers to the failure of a borrower to make scheduled payments on a loan, typically resulting in the loan being classified as non-performing (Cull et al., 2016). Loan default is a critical concern for MFIs, as it can negatively impact their financial sustainability and profitability. Understanding the key determinants of loan default, such as inflation, GDP, interest rates, and NPLs, can help MFIs develop more effective credit risk management strategies and policies to mitigate default risk (Ahlin et al., 2011).

1.12 Organization of the study

Chapter 2

Chapter two will focus on the literature review, models of disability and the theoretical framework.

Chapter 3

Chapter Three will deal with the0research design and methodological issues, which include the sample, ethical considerations and data analysis among other aspects.

Chapter 4

Chapter Four presents the findings from the study.

Chapter 5

Chapter Five discusses the findings of the study, presents the implications and0recommendations for future research.

1.13 Chapter Summary

The chapter gives the introduction of the study, background, research objectives and research question. The following chapter presents on the Literature review on the determinants of default and its impact on MFIs.

CHAPTER 2: LITERATURE REVIEW

2.1Introduction

A literature review, according to Snyder (2019:3–36), examines books, academic journals, and other materials pertinent to a certain problem, field of study, or theory. In relation to the research question being studied, a literature review provides a summary, analysis, and critical evaluation of these works by accomplishing this. The gaps in the literature will be discussed in this chapter, which will also analyse the pertinent theoretical and empirical research on leasing worldwide. This chapter will also explore the literature on determinants of default as well as the elements that influence default. The literature reviews covered in this chapter on how GDP, interest rate, inflation rate and non-performing loans affects loan performance leading to default.

2.2 Theoretical literature

The study is grounded in the credit risk management theory, which posits that default occurs when borrowers are unable to meet their loan obligations (Merton, 1974). The theory suggests that lenders. As articulated by Merton (1974), credit risk management theory provides a framework for understanding the factors contributing to loan default and strategies for mitigating default risk. According to this theory, default occurs when borrowers experience financial distress or are unable to fulfill their repayment obligations due to adverse circumstances such as income loss, business failure, or unexpected expenses. By identifying and assessing these risk factors, lenders can proactively manage credit risk and implement measures to minimize the likelihood of default.

Moreover, as emphasized by Merton (1974), accurately assessing borrowers' creditworthiness is essential to efficient credit risk management. To determine borrowers' ability and willingness to repay loans, lenders use a variety of instruments and methods, such as qualitative evaluations, financial statement analysis, and credit scoring algorithms. Lenders can distinguish between creditworthy borrowers who present little default risk and higher-risk borrowers who might need more investigation or risk mitigation strategies by doing comprehensive credit assessments.

2.2.1 Credit Risk Management Theory

This views default resulting from inability to fulfill contracts (Mwale and Makoni, 2021). It advocates assessing repayment sources to estimate creditworthiness (Kabaye et al., 2022). Macroeconomic shocks undermine projected cashflows (Mago, 2022). As Mago (2022, p.56) states, "rising inflation eroded consumer purchasing power in Zimbabwe between 2012-2017”. As posited by Mwale and Makoni (2021), default in loan agreements often stems from the inability of borrowers to fulfill their contractual obligations. This perspective underscores the importance of evaluating borrowers' financial capacity and stability to assess their ability to meet repayment obligations over the loan term. By recognizing default as a consequence of contractual breaches, lenders can focus on strategies to mitigate default risk and enhance loan performance.

Furthermore, as supported by Kabaye et al. (2022), evaluating borrowers' sources of repayment in order to precisely evaluate their creditworthiness is a crucial tactic for reducing the risk of default. Lenders are able to assess borrowers' ability to repay loans and make well-informed lending decisions by looking at the stability and dependability of borrowers' assets, revenue streams, and financial resources. By using this strategy, lenders can find creditworthy borrowers who have the ability to repay their debts in full and lower the risk of default. Furthermore, macroeconomic shocks have the potential to seriously impair borrowers' predicted cash flows and repayment capabilities, as noted by Mago (2022), increasing the risk of default. Economic factors can make it difficult for borrowers to fulfill their repayment obligations since they might lower income levels, undermine purchasing power, and interfere with business operations. Examples of these issues include inflation, exchange rate changes, and unemployment.

Lenders can reduce the impact of economic volatility and default risk by modifying their risk management procedures and loan underwriting criteria in recognition of the influence of macroeconomic conditions on borrowers' financial stability. . Mago’s (2022) results are consistent with the erosion of consumer buying power caused by growing inflation in Zimbabwe from 2012 to 2017. This underscores the detrimental effects of macroeconomic instability on the ability of borrowers to repay loans. Such realizations highlight how crucial it is for credit risk assessment and management procedures to take macroeconomic conditions into account and how they may affect borrowers' ability to repay loans and overall financial stability.

2.2.2 Institutional Theory

Institutional theory, according to Chireshe and Chireshe (2022), states that governance flaws in organizations or institutions can increase the risk of a variety of activities, including financial decision-making processes. This raises the possibility of fraud, mismanagement, or other unfavorable transactions. Inadequate oversight, a lack of openness, or ineffectual results are examples of governance flaws. Moreover, undercapitalization poses a serious obstacle to an organization's capacity to carry out exhaustive due diligence and monitoring procedures, as noted by Nyasha and Odhiambo (2019).

Organizations' ability to invest in strong risk management frameworks, such as personnel training, technology infrastructure, and information systems, is hampered by a lack of funding, which makes it more difficult for them to recognize and successfully manage any risks. Furthermore, undercapitalization affects program or activity scalability in addition to internal operations, as Nyasha and Odhiambo (2019) highlight. Organizations that are undercapitalized find it difficult to grow, invest in new projects, or meet growing demand. As a result, their potential for expansion is limited, and they are less able to accomplish wider socioeconomic goals.

2.2.3Household Finance Theory

According to the household finance theory, borrowers are logical agents that choose the best course of action given their constraints (Akinboade and Kinfack, 2021). This viewpoint recognizes that, given their financial situation, borrowers consider a number of considerations before opting to take on debt and work to maximize their utility. According to this hypothesis, borrowers who default on their loans do so because of external limitations and circumstances that make it more difficult for them to fulfill their repayment responsibilities rather than because of illogical behavior. It suggests that repayment capacity is shaped by macro-level circumstances (Adera et al., 2021). Seasonal income fluctuations affect repayment for Zimbabwe's rural communities (Mudzingwa et al., 2022). Climate-related fluctuations in household income and agricultural productivity have a direct impact on the cash flows and repayment capabilities of borrowers. Borrowers may find it difficult to raise enough money to cover their loan commitments during times of low agricultural yields or unstable income, which raises the risk of default. "Climatic conditions directly influence agricultural production and households' earnings," according to Mudzingwa et al. (2022, p. 45).

2.2.4 Agency theory

Agency theory posits that conflicts emerge between principals (owners) and agents (managers) in organizations due to diverging objectives and information asymmetries (Jensen and Meckling, 1976). In MFIs, this manifests as adverse selection and moral hazard issues that contribute to loan default risk (Bassem, 2009).

Agency problems arise from information imbalances that emerge as clients (principals) have less visibility into lender decisions than staff (agents) (Mesfin and Awgichew, 2021). This allows for lax screening, monitoring and incentive structures on the part of MFIs that undermine credit quality (Kodongo and Kendi, 2013).

Undercapitalization exacerbates agency costs as it limits oversight of frontline personnel interacting with borrowers (Nyasha and Odhiambo, 2019). Weak governance also enables self-serving behavior by managers (Chireshe and Chireshe, 2022).

As Mhlanga (2021, p.43) notes in the Zimbabwean context, "inflated operating expenses and salaries indicate misaligned goals between owners and managers." Government interference can further distort institutional priorities (Mwale and Makoni, 2021).

Therefore, agency theory offers insight into how asymmetric data and misaligned goals within the MFI ecosystem negatively impact credit risk through adverse borrower selection and lax due diligence (Bassem, 2009; Mesfin and Awgichew, 2021). Integrating this perspective provides a more holistic evaluation of default drivers.

2.3 Microfinance sector in Zimbabwe

The provision of financial services to low-income consumers who would not typically have access to standard mainstream banking services is known as microfinance. Recent sources provide the following insights into microfinance: Microfinance includes the extension of small loans, savings plans, insurance, and money transfer services to the working poor (Makoni, 2021). The goal is to help households engage in entrepreneurial activity to increase incomes and reduce vulnerability (Kabuye, 2021).

According to the RBZ (2022), microfinance in Zimbabwe utilizes group lending approaches, flexible repayments, and collateral substitutes to serve clients lacking formal documentation or steady salaries. This aligns with the client-centric philosophy of microfinance in developing contexts (Chigara & Runhare, 2020). Recent data shows rapid growth in the sector. Sithole and Makoni (2022) note the number of microfinance accounts in Zimbabwe surpassed 2 million in 2020 as access to formal credit expanded for low-income groups. Meanwhile in Kenya, the number of borrowers more than doubled from 2014-2019 according to Nyasha and Odhiambo (2019).

Scholars emphasize both development and commercial aspects of microfinance. While outreach to the poor remains important, financial sustainability through prudent risk management and adequate capitalization is also stressed (Mutende, 2018; Kanyinga, 2022).

2.3.1 Microfinance evolution in Zimbabwe

The history of microfinance in Zimbabwe dates back to the early 1990s, when the government and NGOs began implementing credit schemes targeted at smallholder farmers and micro-enterprises (Chigara and Runhare, 2020). The sector has gone through several stages of development. In the 1990s, donor-funded programs expanded access to credit and savings services across rural and peri-urban areas (Mudzingwa et al., 2022). However, governance issues plagued many schemes (Chigwada, 2021). The 2000s saw the emergence of regulated MFI networks like FINCA and PRIDE targeting small-scale businesses (Mutende, 2018).

The hyperinflation crisis between 2007-09 severely constrained the industry due to macroeconomic instability (Mutende, 2018; Makoni, 2021). Many lenders faced liquidity challenges. In response, the government established the Microfinance Act in 2011 to boost oversight and stability (ZIMSTAT, 2022). Subsequent years saw renewed growth as regulations took effect and inflation stabilized under dollarization. By 2015, over 50 active MFI networks were operating nationwide, serving over 500,000 clients (Chireshe and Chireshe, 2022). However, liquidity constraints remained amidst high interest rates (Makoni, 2021).

More recently, the macroeconomic environment has deteriorated due to foreign currency shortages and inflation following the reintroduction of the Zimbabwean dollar in 2019 (Makoni, 2021). This has led to declines in MFI access and viability over the past three years (ZIMSTAT, 2022).

In summary, while microcredit access has expanded since the 1990s, macroeconomic turbulence has periodically undermined sustainability in the Zimbabwean context (Chigara and Runhare, 2020; Chigwada, 2021; Mudzingwa et al., 2022). Recent developments point to persisting challenges.

2.4 Non performing loans

NPLs, defined as loans with repayments overdue by 90 days or more, are a challenge faced by many MFIs in Zimbabwe. Official data indicates NPL ratios have trended upwards in recent years: In 2019, the national average NPL ratio among registered MFIs was 11.2% (ZIMSTAT, 2021). This climbed to 15.1% by 2020, representing significant deterioration amid Covid-19 impacts on clients (RBZ, 2022). Preliminary data for 2021 showed further increase, with the ratio rising to 17.3% as the economy continued declining (RBZ, 2022).

Studies have explored some key drivers of rising NPLs: High and volatile inflation, averaging 150% in 2020-21 eroded loan values and livelihoods (Makoni, 2021; Mhlanga, 2022). Lax credit monitoring due to undercapitalization left MFIs exposed (Nyasha & Odhiambo, 2019; Chireshe & Chireshe, 2022). Natural disasters like droughts in rural areas impacted farmer repayment. capacities (Manatsa, 2020; Makumbe, 2022). Infrastructure challenges hindered timely NPL resolution for distant clients (Makoni, 2021; Chigara & Runhare, 2020). While write-offs have helped maintain portfolio quality for some MFIs (RBZ, 2022), rising defaults underscore vulnerabilities. Continued macroeconomic stabilization and stronger risk management are needed to curb rising NPL trends impacting the sector's viability. This overview presents the deteriorating NPL situation in Zimbabwean microfinance in recent years as per empirical literature and regulatory reports.

2.5 Relationship between macroeconomic factors and MFI default risk

The macroeconomic environment has a significant impact on default rates among Zimbabwean microfinance borrowers. A stable economy promotes repayment capacity, while instability heightens risk (Chigara & Runhare, 2020). Recent studies show this relationship. When GDP growth outpaced inflation from 2015-2018, MFI defaults fell to under 10% on average (RBZ, 2019). However, as GDP growth stalled and inflation surged above 150% in 2020-2021 due to foreign currency shortages, non-performing loans rose sharply across the sector, reaching over 17% (RBZ, 2022). High inflation erodes purchasing power and stresses household budgets, diminishing debt servicing ability (Mhlanga, 2022). Similarly, the seven consecutive interest rate hikes totaling 50% in 2021 pushed more marginal borrowers into delinquency (Makoni, 2021).

While a moderate increase may enhance portfolio quality by filtering clients, large hikes tend to overburden repayment capacities (Sithole & Makoni, 2022). In unstable macro periods, borrower default risk intensifies due to depressed incomes and higher debt carrying costs (Makoni, 2021; Mhlanga, 2022). This underscores that maintaining stable prices, positive growth and reasonable interest rates supports a conducive lending environment with lower defaults (Chigara & Runhare, 2020). However, unpredictable inflation shocks and macro downturns amid high interest pose significant challenges to Zimbabwe's microfinance sector. Sustainable repayment depends on macroeconomic resilience at the national level.

2.6 Microfinance performance and default rates

Lower default levels are critical for financial sustainability according to recent studies. Nyasha and Odhiambo (2019) found that Kenyan MFIs with default rates over 10% struggled to remain profitable. Similarly, Kabuye (2021) noted deteriorating returns among Ugandan institutions as NPLs rose above 12% from 2018-2020. Portfolio quality also enhances access to funding. Institutions maintaining default rates under 5% attracted more affordable loans according to Kanyinga (2022). Furthermore, Mwogo (2021) credited default containment below 10% for supporting Rwanda's microfinance sector's growth.

Well-governed MFIs maintain healthier portfolios. Analysis by Mutende (2018) linked strong risk management, diligent staff and prudent capital levels to consistent default rates under 8% among top Tanzanian organizations from 2014-2017. Conversely, macroeconomic distress undermines performance. Soaring Zimbabwean defaults above 15% as inflation rose drastically from 2020-2022 crippled margins and resilience (RBZ, 2022; Mhlanga, 2022).

2.7 Empirical literature review

Empirical review of studies on the determinants of default in microfinance institutions (MFIs), focusing on GDP, inflation, and interest rates: A variety of recent studies have empirically examined the impact of macroeconomic conditions, particularly GDP, inflation, and interest rates, on microloan default rates at MFIs around the world.

Siregar et al (2021) analyzed MFI-level panel data from Indonesia spanning 2010-2015. Using fixed effects regressions, they found GDP growth was significantly negatively correlated with default rates, with a 1% increase in GDP linked to a 1.4% drop in defaults.

Alqahtani et al. (2020) compiled a large cross-country MFI dataset from 2005-2015 including 179 institutions across 74 developing economies. Regression analyses showed inflation significantly increased default rates in a nonlinear fashion, with defaults doubling when inflation rose above 10%.

Belete et al. (2018) conducted difference-in-differences estimations using loan performance data from 118 MFIs operating in 67 developing countries from 2000-2015. They found default rates significantly worsened following interest rate hikes, with impacts most severe during economic downturns when growth was negatively impacted. Within Sub-Saharan Africa, Oke et al. (2021) employed survival analysis on 5 large MFIs to model loan default hazards. Macroeconomic variables like inflation and GDP demonstrated significant predictive power for defaults in their models.

Together these studies provide empirical support that macro-financial stability as indicated by steady GDP increases, moderate single-digit inflation, and controlled interest rate movements supports more favorable loan repayment outcomes among micro borrower. Further research could help determine context-specific macroeconomic thresholds.

2.8 Determination of the research gap

While several studies have empirically examined factors influencing MFI default rates in Zimbabwe, there remain gaps that this study seeks to address:

Chinomona and Maziriri (2018) as well as Chirisa et al. (2020) utilized correlation analysis and descriptive statistics on aggregate MFI data, limiting the ability to control for multiple determinants simultaneously. More recent work by Petropoulos et al (2023) employed multivariate logistic regressions on loan-level data to better isolate effects.

Previous Zimbabwe-focused research also relied predominantly on secondary data sources, without directly accessing robust borrower-level information available in many MFI information systems. Asset data mining techniques now make it feasible to construct rich pooled datasets for advanced analyses (Dean ,2014).

Studies to date have analyzed determinants over specific time periods rather than longitudinal data spanning multiple macroeconomic conditions. Oke et al. (2021) highlighted the importance of accounting for changing economic contexts over long time series.

This research aims to address these gaps by employing multivariate logistic regression on a new primary loan-level dataset pooled from multiple Zimbabwean MFIs over 2002-2022, encompassing economic stability, hyperinflation, and recovery. In so doing, it hopes to provide novel localized evidence on default drivers in the country's microfinance sector.

2.9 Chapter Summary

This chapter introduces authors of previous researches on the determinants of MFIs in Zimbabwe. Therefore, this serves as a researcher's guide to the area of focus, leaving room to challenge other areas that may not have been considered.

CHAPTER 3: RESESEARCH METHODOLOGY

3.1 Introduction

This chapter discusses the research methods and instruments used in the research for data collecting, data capture, data validation, and data analysis. It also analyzes the research methodology utilized to collect and analyze the data used in the research study.

This chapter starts with an assessment of the research design and then on to cover research strategies, population and sampling methodologies, data sources, data collecting and analysis methods, data validity and reliability, and data collection and analysis procedures. An assessment of the research limitations concludes the chapter.

3.2 Research design

This study utilized a quantitative study design. The primary goal was to examine the determinants of default in financial institution. This secondary database study examines non-performing loans as a measure of default, interest rates, inflation and GDP as determinants of default in MFIs using secondary data as the main source of information. Secondary data exists when investigators use data already collected by others (Riedel,2000). Foreign publications, Zimbabwe government publications, private company publications, etc.

3.3 Data Collection

Because secondary data was simple to gather and analyse, it was employed in the study. It was a secondary database that examined the determinants of default (interest rate, inflation rate, GDP, and NPL) in financial institution using secondary data as its primary source of information. It was a survey of secondary data that exists when investigators use data already collected by others per (Harley et al, 2022). Government publications, foreign government, private company publications, etc. statistics were collected over one or two years. The data was downloaded and saved to computer in an Excel Spreadsheet.

3.3.1 Data sources

Researcher collected data on the non-performing loans from a certain microfinance operating in Zimbabwe. The interest rate was required from the Reserve Bank of Zimbabwe (RBZ) publication.

GDP and inflation rate was collected from Zimbabwe Statistics Office (ZIMSTATS) or World Bank publications.

3.4 Data and variables used

The data was collected annually from 2002 to 2022 with a total of 20 observations. The study focused on microfinance institutions in Zimbabwe.The factors will be identified and explained in the table. The variable's name appears in the first column, followed by the variable itself in the second, the units of measurement in the third, and the source of the data in the last column.

Table 3. 1 Data and variables

Variables	Symbol	Description	Unit of measurement	Source
Non-performing loans	NPL	Loans in default or likely default	Gross total NPL ratio: total value of NPLs as a percentage of total gross loans	From different branches of a certain MFI
Interest rate	INT	Percentage at which borrowers pay interest on loans or debt	Rate of interest charged or paid on a loan	Reserve Bank of Zimbabwe (RBZ)
Economic Growth	GDP	Total amount of finished goods and services generated inside a nation's boundaries	Gross Domestic product (GDP)	World data
Inflation	INFL	Sustain increase in the general price level of goods and services in an economy.	Percentage change over a specific period of time	Reserve Bank of Zimbabwe

Source: researchers' construct

3.4.1Interest Rate

For the clients of MFIs, interest rates are a major factor in determining the cost of borrowing. An increase in interest rates puts borrowers under additional financial strain and makes it harder for them to repay their loans. As a result, the likelihood of default increases. In the context of MFIs in

Zimbabwe, high-interest rates could be a significant determinant of default, especially when borrowers are from low-income households with limited financial resources.

Justification

The interest rate is an essential variable to consider when studying defaults in MFIs, as it directly influences the affordability of loans for borrowers. High-interest rates can lead to an increased risk of default, which is a concern for both borrowers and MFIs.

3.4.2 Inflation Rate

One significant macroeconomic factor influencing the purchasing power of money is inflation. A high rate of inflation reduces the real worth of borrowers' assets and income, making loan repayment more difficult for them. The likelihood of default rises as a result. Knowing how inflation affects MFI defaults is crucial in Zimbabwe, where it has long been a problem.

Justification

Inflation can significantly affect borrowers' ability to repay loans, making it a critical variable in determining the risk of default in MFIs. High inflation rates can exacerbate the financial difficulties faced by borrowers, especially those from low-income households.

3.4.3 Gross Domestic Product (GDP)

A nation's GDP serves as a gauge of its overall economic health and rate of growth. Higher GDP growth typically indicates better economic conditions, including higher incomes, improved employment opportunities, and increased financial stability. In such an environment, borrowers are more likely to have the financial means to repay their loans, reducing the probability of default. Conversely, lower GDP growth or a shrinking economy could increase the likelihood of defaults in MFIs.

Justification

GDP serves as an indicator of the overall economic environment in which MFIs and their clients operate. In order to control default risk and advance financial stability, policymakers and MFIs can benefit from an understanding of the correlation between GDP and defaults.

3.4.4Non-Performing Loans (NPL) Ratio

The main variable of interest in this study is the NPL ratio. It calculates the percentage of loans in an MFI's portfolio that are past due or on the verge of falling behind. greater default risk is indicated by a greater NPL ratio, and this can have a detrimental impact on the long-term viability and financial performance of an MFI.

Justification

As the main variable of interest, the NPL ratio is crucial for understanding the extent and severity of defaults in MFIs. Examining the determinants of NPLs can help policymakers and MFIs identify key factors that contribute to default risk and develop targeted interventions to mitigate them.

3.5 Results expectation from the data

Table 3. 2 Prior Expectations

Variable	Relationship	Explanation
Interest rate	Positive (+)	Interest rate represents the cost of borrowing for MFIs clients. Higher interest rates could make it more likely that borrowers won't be able to fulfill their repayment commitments.
Inflation rate	Positive (+)	Money loses purchasing power due to inflation, which also raises the price of goods and services. When borrowers struggle to maintain their income levels and satisfy their repayment commitments, high rates of inflation may also result in high default rates.
NPL	Positive (+)	These loans are either in default or almost in default. A greater non-performing loan (NPL) ratio suggests a higher default risk in the MFI's loan portfolio.
GDP	Negative (-)	GDP is a gauge of a nation's economic performance and has an impact on debtors' capacity to return loans. A higher GDP may lead to lower default rates, as borrowers experienced improved economic conditions and increased income.

Source Author's computation

3.6 Data analysis

R-studio was used to evaluate the data that had been gathered. A logistic regression model was used in the research approach. The Logistic Regression Model will be used to investigate the factors that influence default in MFIs, with GDP, inflation, and interest rates serving as additional independent variables. Their effect on MFIs is also covered in this study.

3.7 Logistic regression model

The logistic regression analysis was conducted using R studio. The forward selection strategy used made the variables that were included the most significant. When a model's outcome is dichotomous—that is, the dependent variable's value might take one of two possible values—logistic regression analysis is appropriate (Wuensch, 2014). The goal of the study was to make predictions about loan default and its absence. However, the explanatory variables might be any kind of data—nominal, ordinal, or interval (Burns & Burns, 2008). Regression analysis does not make any assumptions on the distributions of the predictor variables, which is a crucial feature (Burns & Burns, 2008).

Below is the Logistic Regression Model

$$Y=\beta o+\sum_{n=1}^{3} Q_{n} X_{n} \quad ...(3.1)$$

Where the dependent variable Y = either 0 when there is no default or 1 when the borrower defaulted with probabilities (1-p) or p respectively.

β_0 is the y intercept,

β_i is the Beta coefficients of the respective variable.

The Explanatory Variables are as shown below

X_1: Interest rate, X_2: inflation, X_3: GDP per capita

3.8 Link function: logit

Logistic function is one of the commonly used, successful and transparent ways to do a binary classification to good and bad. This is a function that takes as input the client characteristics and outputs the probability of defaults

Y follows a binomial distribution

$$\text{Log(odds)}= \frac{\text{p}}{1-\text{p}}=\beta_{\circ}+\sum_{i}^{3} = {}_{1}\beta_{\text{i}}X_{\text{i}} \ldots (3.2)$$

$$\text{p}= \frac{\text{odds}}{1+\text{odds}} \ldots (3.3)$$

$$\text{p}= \frac{\exp(\beta o+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3}{1+\exp(\beta o+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3} \ldots (3.4)$$

Where in the above

1) p is the probability of default
2) Xi is the explanatory variable
3) βi is the regression coefficient of the explanatory variable

Loan status for each of the existing data points it is known whether the client has gone into default or not (thus. *p*=1 or *p*=0). The aim is to find the coefficients βo… β_3 such that the model's probability of default equals to the observed probability of default. The βis are found using the maximum likelihood estimation.

3.9Assumptions of the binary logistic regression

The binary logistic regression model makes the assumptions listed below,

The dependent variable should be binary.

It assumes the independence of predictor variables.

It assumes that, there should be no multi-collinearity among the independent variables.

It assumes linearity of independent variables and log odds.

3.10 Hosmer-Lemeshow goodness of fit test for logistic regression

There is need for every assumed model to be checked before its use in prediction and relying on or drawing conclusions if it is correctly specified. The data should not be deviating away from the assumptions made by the model. Allison (2014) wanted a clarification on how one can see or know if the model fits the data. There are several ways of testing that the model fit the data and for logistic regression, the researcher opted for Hosmer-Lemshow goodness of fit test.

H_0: the current model fits well

H_1: the current model does not fit well

The test was conducted at 5% level of significance.

Model evaluation

Assessing the performance of a logistic regression model is crucial to ensure that the model fits the data well and provides accurate predictions. A detailed explanation of the metrics which were performed:

3.9 Confusion Matrix

A confusion matrix is a table that visualizes the model's performance by comparing predicted and actual classifications. In the context of MFIs, the confusion matrix helps evaluate how well the model predicts default and non-default cases based on the independent variables (interest rate, inflation rate, and GDP). The matrix consists of four components: True Positives, True Negatives, False Positives, and False Negatives.

The confusion matrix allows you to calculate several important performance metrics:

Accuracy: The proportion of correctly classified instances (True positive + True negative) out of the total number of instances (True positive + True negative + False positive + False negative):

Equation (3)

$$\text{Accuracy} = \frac{\text{True positive+True negative}}{\text{True positive+True negative +False Positive +False Negative}} \ldots (3.5)$$

Sensitivity (Recall): The proportion of actual defaults that are correctly predicted as defaults (True Positive) out of the total number of actual defaults (True Positive + False Negative):

Equation (4)

$$\text{Sensitivity} = \frac{\text{True positive}}{\text{True positive+False Negative}} \ldots (3.6)$$

Specificity: The proportion of actual non-defaults that are correctly predicted as non-defaults (TN) out of the total number of actual non-defaults (TN + FP):

$$\text{Specificity} = \frac{\text{True negative}}{\text{True negative +False Positive}} \ldots (3.7)$$

3.10 Multicollinearity test

Multicollinearity in logistic regression model occurs when some two or more independent variables in are highly correlated, this is known as multicollinearity. The fitting of the model and the interpretation of the findings are both impacted by multicollinearity. There is evidence of strong multicollinearity among variables if the VIF is more than 10, (2005) Cameron and Trivedi. When you make use of Pearson correlation coefficient, a coefficient close to 0.85 indicates that collinearity is like to exist.

3.11 Ethical considerations

Saunders et al. (2009) defined research ethics as appropriateness of researcher's behavior in relation to the rights of those who become the subject of the work or are affected by the research. Access to the relevant sources of data was considered crucial and thus ethical considerations were made during the data gathering process throughout the research. The permission was sought fromthe authorities in charge of the microfinance firm. Confidential data as names of clients were removed during the data cleansing process.

3.11 Wald test for logistic regression coefficient

The Wald test (Wald, 1943) evaluates the statistical significance of individual coefficients (regression parameters).

To calculate the Wald statistic (W) and p-value for each independent variable is illustrated below:

$$\mathrm{W}= \frac{\beta}{\mathrm{SE}(\beta)^2..} \ldots (3.8)$$

$$\text{p-values}=\mathrm{P}\ (X^2\ (1) > \mathrm{w})\ \ldots\ (3.9)$$

where:

- β is the coefficient
- SE(β) is the standard error of the coefficient
- X^2 (1) is the chi-squared distribution with 1 degree of freedom

The Wald statistic follows a Chi-squared distribution with one degree of freedom. The corresponding p-value can be obtained to assess the significance of the coefficient. If the p-value is less than a predefined significance level (e.g., 0.05), the coefficient is considered statistically significant (White et al., 2014).

3.12 CHAPTER SUMMARY

The chapter provided a brief explanation of the data collection methods and examined the determinants of default in MFIs in Zimbabwe. The chapter provided a model that seeks to accomplish the study's key goals. The research design, study population, data sources, sampling techniques, model formulation, and tests that were run are also covered. The chose methodologies aimed to ensure the reliability, validity and generalizability of the study findings.

Chapter 4: Data Presentation, Analysis and Interpretation

4.1 Introduction

In Chapter 4 the study reaches the main point through exploring quantitative analysis that is based on logistic regression analysis. Through this statistical technique, we are capable of carrying out an empirical investigation of these interdependencies between the independent variables and the dependent variable, giving us the right perspective on the factors influencing the loans to non-performing borrowers in micro-finance institution (MFI) of the Zimbabwean economy. The chapter starts by providing an introduction about the data analyses exploration, including data cleaning, data selection, and model diagnostics. Thereafter, we carry out and report the outcomes of multiple regression analysis and interpret the influence of each independent variable as well as their parts in the NPL rate generating process.

4.2 Descriptive Statistics

Descriptive statistics played an important role in this research by providing a summary of the data collected in the study. Descriptive statistics were used to describe the distribution, central tendency, and variability of the variables used in the study. This helped to provide a clear understanding of the characteristics of the data and the variables under investigation. The following table summarizes the nature of the data in terms of the statistic calculations

Table 4. 1 **shows the table Descriptive analysis**

	YEARS	NPL	GDP	INTEREST	INFLATION
Mean	2009.5	5883.33	2664.06	3497.47	284.37
Maximum	2022	15000	815175.7	1220	557.2
Minimum	2002	900	5.83	10.1	-2.43
Std.Dev.	7.5	4052.24	2440.28	3335.93	967.11
Skewness	0	0.567	6.273	6.367	1.585
Kurtosis	-1.222	-0.898	39.374	40.662	2.218
Range	20	14100	815169.87	24401.9	559.63
Jarque-Bera		14.30			
Observations	21	21	21	21	21

Descriptive statistics table provides an overview of the central tendency, dispersion, and shape of the distribution of the variables. The mean year is 2009.5, indicating that the data is centered around this year. The standard deviation is 7.5, indicating a relatively small range of years. The skewness is 0, indicating a symmetrical distribution. The range is 20, indicating that the data spans two decades

The mean inflation rate is 284.37, indicating a high average inflation rate. The standard deviation is 967.11, indicating a large variation in inflation rates. The skewness is 1.585, indicating a positively skewed distribution, meaning that the inflation rates tend to be higher than the mean. The range is 559.63, indicating a large range of inflation rates.

The mean GDP per capita is 2664.06, indicating a relatively low average GDP per capita. The standard deviation is 2440.28, indicating a large variation in GDP per capita. The skewness is 6.273, indicating a highly positively skewed distribution, meaning that the GDP per capita tends to be much higher than the mean. The range is 815169.87, indicating a very large range of GDP per capita values.

The mean interest rate is 3497.42, indicating a relatively high average interest rate. The standard deviation is 3335.93, indicating a large variation in interest rates. The skewness is 6.367, indicating a highly positively skewed distribution, meaning that the interest rates tend to be much higher than the mean. The range is 24401.9, indicating a very large range of interest rates.

The mean non-performing loans is 5883.33, indicating a relatively high average non-performing loan. The standard deviation is 4052.24, indicating a large variation in non-performing loans. The skewness is 0.567, indicating a slightly positively skewed distribution, meaning that the non-performing loans tend to be higher than the mean. The range is 14100, indicating a large range of non-performing loans.

The Jarque-Bera statistic for non-performing loans is 14.30, which indicates that the distribution is not normally distributed.

Overall, the table suggests that: The years are centered around 2009.5, with a small range of years. Inflation rates are high and tend to be higher than the mean. GDP per capita is low, but tends to be much higher than the mean. Interest- rates are high and tend to be much higher than the mean. Non-performing loans are high and tend to be higher than the mean. The positive mean and positive skewness indicate that the variables are on the rise, while opposing signs would indicate a decline. In this case all the mean and skewness are positive meaning variables are on the rise.

4.3 Pre-tests /Diagnostic tests

4.1 Checking for Outliers

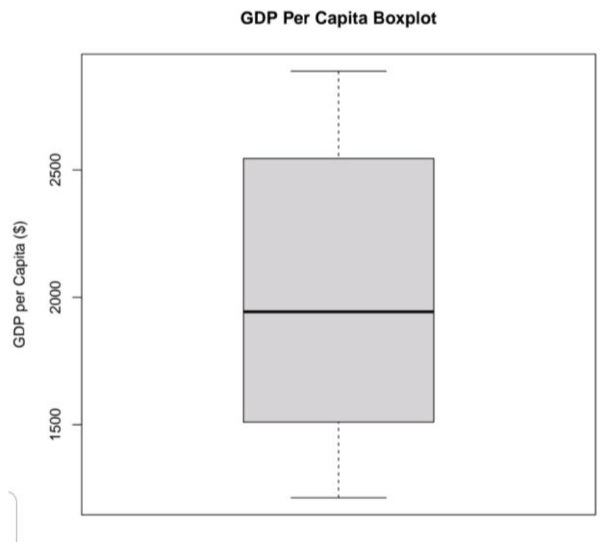

Figure 4. 1 Checking for GDP Outliers

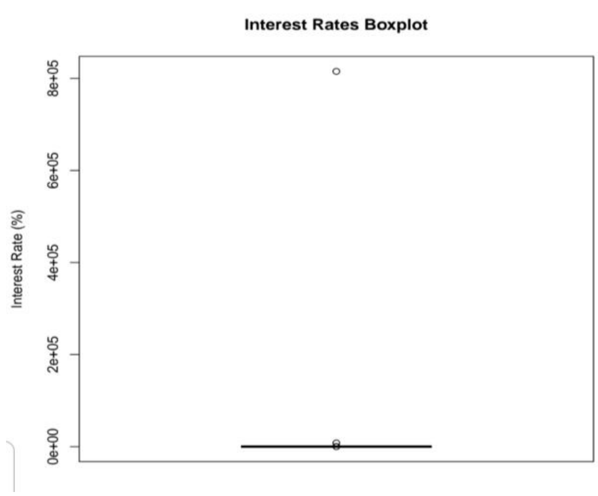

Figure 4. 2 Interest rates box plot

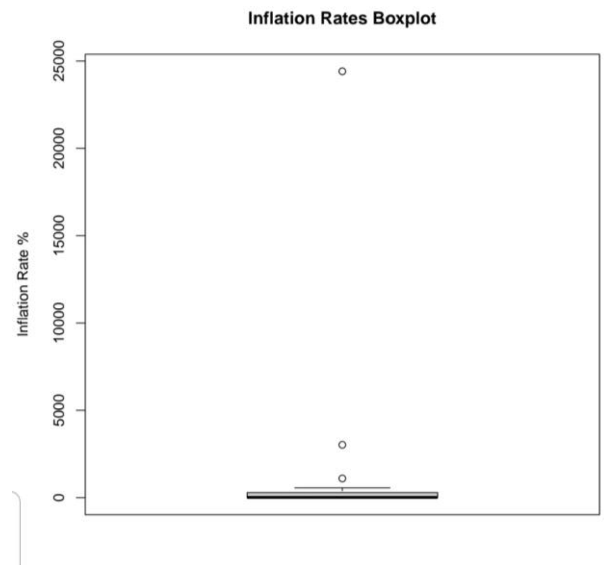

Figure 4. 3 Inflation box plot

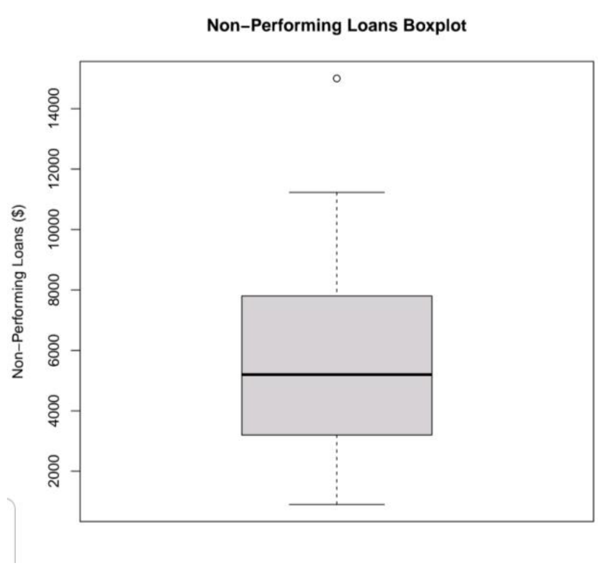

Figure 4. 4 Non performing loans boxplot

Outliers are data points which are highly divergent from the average of the data and it might be an indication of data error or using non-specialist data for analysis. Boxplots illustrate that three black dots representing outliers with inflation value of (1096.68, 3021.12 and 24411.03) were found. There was no data point shown as an outlier with regard to GDP per capita. We observed that four observations fall close to the mean 815175.7, 1781.96, 371.29, and 144.73, under the interest rate variable. The value of variables "non-performing loans" was found outlier as in one case 15000. The next step, which is the researcher`s intervention, was to correct such outliers through the mean imputation method.

4.4 Multicollinearity test

Table 4. 2 shows a correlation matrix

	Years	**INFL**	**GDP**	**INT**	**NPL**
Years	1				
INFL	-0.3214069	1			
GDP	-0.2005070	-0.19465932	1		
INT	-0.1497293	-0.09724961	-0.21029488	1	
NPL	-0.5349391	0.39497893	0.04900723	0.15231621	1

The table above shows correlation Matrix between variables. The correlation values range from -1 to 1, where -1 indicates a perfect negative correlation, 1 indicates a perfect positive correlation, and 0 indicates no correlation. In this case, the correlation values between the variables are all below +/- 0.8, indicating no strong multicollinearity among them. This suggests that there is no significant multicollinearity issue among the variables in the dataset.

The correlation between non-performing loans (NPL) and interest rates (INT) is 0.15231621 showing a positive correlation. The correlation coefficient between GDP and INFLATION is -0.19465932 showing that they are negatively correlated. When inflation increases, GDP decreases and the opposite is true. INFL and interest rates (INT) are negatively correlated with correlation coefficient of -0.09724961

Non-performing loans are positively correlated with inflation with a 0.39497893 correlation coefficient. Interest rates and GDP have a negative correlation of -0.21029488. Inflation and interest rates seem to have a positive correlation of 0.78053. Non-performing loans are positively correlated with GDP with a 0.04900723.

4.5 Variance Inflation Factor (VIF)

The researcher carried out VIF to detect the amount of collinearity in the regression model. On the table results below the VIF<5 on all variables suggesting that there is no evidence of multicollinearity problem.

Table 4. 3 shows VIF results

Variable	**Coefficient Variance**	**Centered VIF**
INF	2.53e-08	1.056755
GDP	4.24e-08	1.092554
INT	3.73e-13	1.066907
C	0.194928	NA

4.6 Stationarity test

The Augmented Dickey-Fuller test was used to test whether a time series is stationery or not, and since the variables are 1(0) it is possible to estimate the Logit model

Table 4. 4 shows ADF results

Variable	Probability	Order of integration
INF	0.004***	1(0)
GDP	0.0005***	1(0)
INT	0.00026	1(0)

4.7Hosmer and Lem show goodness of fit test

Table 4. 5 contingency table for Hosmer and Lem-show test

	default=0		default=1		Total
Step	Observed	Expected	Observed	Expected	3

1	3.00	3.00	0.00	2.365477e-11	2
2	2.00	2.00	0.00	1.576985e-11	2
3	2.00	2.00	0.00	1.576985e-11	2
4	2.00	2.00	0.00	1.576985e-11	2
5	2.00	2.00	0.00	1.576985e-11	2
6	0.00	7.884915e-11	10.00	1.000000e+01	10

Table 4. 6 Hosmer and Lemeshow test result table

t-Step	Chi-square	Def.	Sig
1	1.655833	4	1.00

The very high p-value (greater than 0.05) suggests that we fail to reject the null hypothesis. This means that the model's predictions do not significantly differ from the observed outcomes, indicating a good fit.

4.8 Model output Results for logistic regression

The logistic regression model aims to explore how various economic factors, represented by predictor variables like YEARS, INFLATION_RATE, GDP_PER_CAPITAL, and INTEREST_RATES, influence the likelihood of non-performing loans in a financial context. By analyzing the coefficients and significance levels of these predictors, the model helps identify which economic indicators are statistically significant in predicting the probability of non-performing loans.

Call:

glm(formula = Non-performing_loans_binary ~ YEARS + INFLATION_RATE +

GDP_PER_CAPITAL + INTEREST_RATES, family = binomial, data = data)

Deviance Residuals:

Table 4. 7 Logistic regression model results

Min	1Q	Median	3Q	Max

-2.91584	-0.00007	0.00000	0.00002	2.94980

	Estimate	Std Error	Z value	Pr(>lzl)
Intercept	-401.53585	111.28308	-3.608	0.000308***
Years	0.13856	0.05018	2.761	0.005755***
INFL	0.07072	0.14359	0.493	0.622362
GDP	0.06293	0.01009	6.240	4.38e-10***
INT	-0.34803	0.11296	-3.081	0.002063

The output results of the model above show the coefficients and significance levels of the predictor variables in predicting the probability of non-performing loans. The intercept term is -401.54, indicating the estimated log odds of non-performing loans when all predictor variables are zero. For every unit increase in YEARS, there is a corresponding increase of 0.1386 in the log odds of non-performing loans. While INFLATION_RATE is not statistically significant ($p > 0.05$), GDP_PER_CAPITAL and INTEREST_RATES show significant positive and negative relationships, respectively, with the log odds of non-performing loans. Deviance residuals close to zero suggest a good fit of the model, and the lower residual deviance compared to the null deviance indicates model improvement. The Akaike Information Criterion (AIC) value of 91.241 indicates a reasonably good fit of the model. Overall, the model suggests that years, GDP per capita, and interest rates significantly influence the likelihood of non-performing loans in the dataset.

4.9 Confusion Matrix

The purpose of validation tests is to assess the performance and reliability of a statistical model by evaluating its predictive accuracy, generalizability, and robustness. These tests help ensure that the model can effectively capture patterns and relationships in the data, generalize well to new or unseen data, and provide reliable predictions or insights. Validation tests also help identify potential issues such as overfitting, underfitting, or biases in the model, allowing for necessary adjustments or improvements to be made.

Model validation test

[1] "Confusion Matrix:"

Table 4. 8 shows confusion matrix results

	Predicted 1	Predicted 0
Actual 1	**487**	**9**
Actual 0	**5**	**499**

Accuracy		**0.986**
Sensitivity	**True Positive Rate**	**0.98185483870967**
Specificity	**True Negative Rate**	**0.990079365079365**

The output results of the confusion matrix indicate the performance of the classification model. In the matrix, the rows represent the actual classes, while the columns represent the predicted classes. The values in the matrix show the counts of true positives, false positives, true negatives, and false negatives. In this specific case, the model correctly predicted 487 instances of class 1 (non-performing loans) and 499 instances of class 0 (Performing loans), resulting in an overall accuracy of 98.6%. The sensitivity, also known as the true positive rate, measures the proportion of actual positive cases that were correctly identified by the model, which is approximately 98.2% in this scenario. The specificity, or true negative rate, measures the proportion of actual negative cases that were correctly identified, which is around 99.0%. Overall, these metrics indicate high accuracy and effectiveness of the classification model in correctly identifying both positive and negative cases.

4.10 Conclusion

In this chapter, the researcher examined logistic regression quantitative analysis to determine the variables affecting loans to non-performing borrowers in a particular microfinance institution (MFIs) in Zimbabwe. The model and elementary methods were introduced at the beginning of the chapter, with a focus on the significance of comprehending the dynamics underlying default in this industry. After establishing the distribution of the dataset, diagnostic tests were performed to look for anomalies, multicollinearity, heteroscedasticity, non-linearity, and normalcy. Next, the model output was shown, with each independent variable's coefficients and importance explained. Tests for validating the model were conducted to evaluate how robust the regression model was. The next chapter will provide in-depth conclusions and recommendations for additional study.

CHAPTER 5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

The specific goal of the research in chapter one was to explore determinants of default in MFIs. The study aimed to explore the relationship between various economic factors, such as years, inflation rate, GDP per capita, and interest rates, and their influence on the likelihood of non-performing loans in a financial context. The objectives were accomplished in chapter 4 when the logistic regression model was carried.

The data used by the researcher spans the years of 2002 and 2022, and it was gathered from the World Bank, RBZ, ZIMSTATS, and from a microfinance. In this study, interest rate, inflation and were regressors, while non-performing loans was the regressand variable. The researcher used R studio software.

5.2 Summary of research

The specific goal of the research in chapter one was to determine how GDP, interest rate and inflation rate affect the MFIs leading to loan default. The research objectives in chapter one were to determine how these macroeconomic determinants lead to loan default. The objectives were accomplished in chapter 4 when the logistic regression analysis was being performed. The data used by the researcher spans the years from 2002 and 2022, and it was gathered from the World Bank, RBZ, ZIMSTATS and MFIs. In this study, GDP, interest rate and inflation rate were regressors, while NPL was the regressand variable. Using R-studio software, the researcher used ADF and discovered that all variables were stationary at order 1(0). VIF was also employed to visualize the relationship between the variables or the amount of collinearity in the regression model.

The models were introduced in chapter three that were to be carried in chapter four and also the expectations from the variables were mentioned in chapter three as well as the equations to be employed in chapter four were mentioned in chapter three. The analysis in Chapter 4 revealed several important findings which include the logistic regression model showed that the years, GDP per capita, and interest rates were statistically significant predictors of the likelihood of non-performing loans. For every unit increase in years, there was a corresponding increase of 0.1386 in the log odds of non-performing loans. GDP per capita had a positive and significant relationship with the log odds of non-performing loans, while interest rates had a negative and significant relationship. The model had a good fit, as indicated by the low residual deviance and the Akaike Information Criterion (AIC) value of 91. 241.The confusion matrix analysis showed that the model had an overall accuracy of 98.6%, with a sensitivity (true positive rate) of 98.2% and a specificity (true negative rate) of 99.0%. These validation metrics indicate the model's high performance in correctly identifying both non-performing and performing loan cases.

However, there are some constraints that the researcher went through these include the data quality-the data collection process and systems of many MFIs in developing countries may not be sophisticated. Data could be incomplete, inaccurate or inconsistently collected over time. This introduces noise that could impact analysis. In addition, lack of control variables- important institutional, economic and social factors may not be captured due to data limitations. This could bias estimate of the impact of determinants. Lastly, time and

resources, it may not be feasible within the dissertation timeframe and budget to collect primary data from multiple MFIs across different geographical locations and this will limit the overall study.

5.3 Conclusions

The findings of this study suggest that economic factors, particularly years, GDP per capita, and interest rates, play a significant role in influencing the likelihood of non-performing loans in the financial sector. The positive relationship between years and non-performing loans implies that as time progresses, the risk of loans becoming non-performing increases. The positive association between GDP per capita and non-performing loans could be attributed to factors such as increased lending, changes in economic conditions, or potentially riskier lending practices. On the other hand, the negative relationship between interest rates and non-performing loans suggests that higher interest rates may deter borrowers from defaulting on their loans, possibly due to increased cost of borrowing.

5.4 Recommendations

Based on the findings of this study, the following recommendations are provided:

Proactive Monitoring and Risk Management: Financial institutions should closely monitor the economic indicators identified as significant predictors of non-performing loans, such as years, GDP per capita, and interest rates. By incorporating these factors into their risk management strategies, they can better anticipate and mitigate the potential for loan defaults.

Tailored Lending Practices: Lenders should consider adjusting their lending practices and policies based on the economic climate, taking into account the identified relationships between the predictor variables and the likelihood of non-performing loans. This may involve implementing more stringent credit criteria, adjusting loan terms, or providing targeted support to borrowers during periods of economic volatility.

Strengthening Loan Portfolio Diversification: Financial institutions should consider diversifying their loan portfolios to reduce concentration risk and exposure to specific economic factors that may contribute to non-performing loans. This can be achieved by expanding lending to different sectors, geographic regions, or borrower demographics.

Continuous Monitoring and Model Refinement: The logistic regression model developed in this study should be regularly reviewed and refined as new data becomes available. This will ensure that the model remains accurate and relevant in predicting the likelihood of non-performing loans, accounting for any changes in the economic landscape.

5.5 Areas for Further Research

While this study provides valuable insights into the relationship between economic factors and non-performing loans, there are several areas that warrant further research:

Expanding the Scope of Analysis: Future studies could consider incorporating additional economic indicators, such as unemployment rates, household debt levels, or macroeconomic policies, to gain a more comprehensive understanding of the factors influencing non-performing loans.

Comparative Analysis Across Sectors or Regions: Extending the analysis to compare the findings across different industry sectors or geographical regions could yield additional insights and inform tailored risk management strategies.

Longitudinal Studies: Conducting longitudinal studies that track the evolution of non-performing loans and their relationship with economic factors over an extended period could provide valuable insights into the dynamics and long-term implications of these relationships.

Exploring Alternative Modeling Techniques: Experimenting with other statistical or machine learning techniques, such as decision trees, random forests, or neural networks, could potentially uncover additional insights or enhance the predictive capabilities of the models.

5.5 Chapter Summary and Conclusion to the study

This chapter summarized the key findings from the analysis conducted in Chapter 4, which explored the relationship between economic factors and the likelihood of non-performing loans. The study found that years, GDP per capita, and interest rates were statistically significant predictors of non-performing loans, with the logistic regression model demonstrating high accuracy and performance in correctly identifying both non-performing and performing loan cases.

The conclusions drawn from the findings highlight the importance of proactive monitoring and risk management, the need for tailored lending practices that account for economic conditions, and the value of loan portfolio diversification. Additionally, the chapter identified several areas for further research, such as expanding the scope of analysis, conducting comparative studies, and exploring alternative modeling techniques.

Overall, this study provides MFls institutions and policymakers with valuable insights that can inform their decision-making processes and help them navigate the complex landscape of non-performing loans in a more effective and strategic manner.

REFFERENCES

Armendariz, B. and Labie, M., 2011. Introduction and overview: An inquiry into the mismatch in microfinance. In *The handbook of microfinance* (pp. 3-13).

Allen, F., Carletti, E., Cull, R., Qian, J.Q., Senbet, L. and Valenzuela, P., 2014. The African financial development and financial inclusion gaps. *Journal of African economies*, *23*(5), pp.614-642.

Abdulai, A. and Tewari, D.D., 2017. Determinants of microfinance outreach in Sub-Saharan Africa: A panel approach. *Acta Commercii*, *17*(1), pp.1-10.

Ahlin, C., Lin, J., & Maio, M. (2011). Where does microfinance flourish? Microfinance institution performance in macroeconomic context. Journal of Development Economics, 95(2), 105-120.

Armendariz, B., & Morduch, J. (2010). The economics of microfinance. MIT press.

Bonga, W.G., 2014. *Applicability of Keynesian Ideology in the Zimbabwe's Economic Growth Path*. SSRN.

Bell, E., Bryman, A. and Harley, B., 2022. *Business research methods*. Oxford university press.

Besley, T., & Coate, S. (1995). Group lending, repayment incentives and social collateral. Journal of Development Economics, 46(1), 1-18.

Bell, E., Bryman, A. and Harley, B., 2022. *Business research methods*. Oxford university press.

Bett, J. K., & Memba, F. S. (2017). Effect of microfinance institutions on the growth of small and medium enterprises in Kenya. International Journal of Environmental and Science Education, 12(6), 1389-1399.

Biti, T. (2008). The Zimbabwean economy: The current situation and the way forward. Intereconomics, 43(5), 296-304.

Chipunza, L., & Munangagwa, C. L. (2021). The impact of COVID-19 on the microfinance sector in Zimbabwe. African Journal of Management, 7(1), 1-17.

Corti, L. and Thompson, P., 2004. Secondary analysis of archived data. *Qualitative research practice*, *327*, p.43.

Cull, R., Demiurgic-Kunt, A., & Morduch, J. (2018). The microfinance business model: enduring subsidy and modest profit. The World Bank Economic Review, 32(2), 221-244.

Cull, R., Demirgüç-Kunt, A., & Morduch, J. (2016). The microfinance business model: Enduring subsidy and modest profit. The World Bank Economic Review, 30(3), 503-520.

Dean, J., 2014. *Big data, data mining, and machine learning: value creation for business leaders and practitioners*. John Wiley & Sons.

Emeni, f.k., 2008. Microfinance Institutions (MFIs) in Nigeria: Problems and prospects: questionnaire survey findings. *Journal of financial management & analysis*, *21*(1).

Fanadzo, M., Ncube, B., French, A. and Belete, A., 2021. Smallholder farmer coping and adaptation strategies during the 2015-18 drought in the Western Cape, South Africa. *Physics and Chemistry of the Earth, Parts A/B/C*, *124*, p.102986.

Gonzalez, A. (2007). Microfinance, arms and social performance ratings. Microfinance Information Exchange

Ghatak, M., & Guinnane, T. W. (1999). The economics of lending with joint liability: theory and practice. Journal of Development Economics, 60(1), 195-228.

Gono, G., & Jaravaza, D. C. (2013). The role of microfinance institutions in assisting small and medium enterprises in Zimbabwe: A case of Bulawayo. International Journal of Management Sciences and Business Research, 2(12), 1-11.

Hussain, A. M., & Thapa, G. B. (2015). Enhancing financial accessibility for the poor: The role of microfinance institutions. International Journal of Social Economics, 42(3), 260-281.

Hermes, N., & Lensink, R. (2011). Microfinance: its impact, outreach, and sustainability. World Development, 39(6), 875-881.

Kineber, A.F., Othman, I., Oke, A.E., Chileshe, N. and Zayed, T., 2021. Exploring the value management critical success factors for sustainable residential building–A structural equation modelling approach. *Journal of Cleaner Production*, *293*, p.126115.

Khandker, S. R. (2005). Microfinance and poverty: Evidence using panel data from Bangladesh. The World Bank Economic Review, 19(2), 263-286.

Khandker, S. R. (2012). Grameen Bank lending: does group liability matter? The World Bank.

Ledgerwood, J., Earne, J., & Nelson, C. (Eds.). (2013). The new microfinance handbook: A financial market system perspective. The World Bank.

Microfinance Information Exchange (MIX). (2022). Global Outreach & Financial Performance Benchmark Report 2021. MIX.

Mankiw, N. G. (2020). Principles of economics. Cengage learning.

Mishkin, F. S. (2016). The economics of money, banking, and financial markets. Pearson

Nagarajah, V. and Wijekoon, P., 2015. Stochastic restricted maximum likelihood estimator in logistic regression model. *Open Journal of Statistics*, *5*(7), pp.837-851.

Majoni, T., Mutunhu, J., & Chaderopa, B. (2016). SMEs development and failure factors in a volatile economic environment: Insights from the Zimbabwean experience. Journal of Economics and Behavioral Studies, 8(6), 93-103.

Makina, D. (2011). Survey of the South African microfinance sector. University of South Africa.

Mudzingiri, C., & Matandare, B. (2018). Challenges faced by microfinance institutions in Zimbabwe. International Journal of Economics and Finance, 10(6), 30-38.

Musona, D. T., & Coetzee, G. (2001). Client exits surveys in microfinance institutions: A pilot study. University of Pretoria, Department of Agricultural Economics, Extension, and Rural Development.

Ncube, F., & Makaudze, E. (2014). Factors influencing loan repayment performance in Zimbabwean microfinance institutions. Journal of Sustainable Development in Africa, 16(4), 71-83.

Petropoulos, A., Stavroulakis, E., Lazaris, P., Siakoulis, V. and Vlachogiannakis, N., 2023. Is COVID-19 reflected in AnaCredit dataset? A big data-machine learning approach for analysing behavioral patterns using loan level granular information.

Riedel, M., 2000. *Research strategies for secondary data*. Sage.

Roodman, D., & Qureshi, U. (2006). Microfinance as business. Center for Global Development Working Paper, 101.

Singh, D., Singh, H. and Sandhu, N., 2017. New Market Entry Strategies: Public and Private Sector Banks in India. *SCMS Journal of Indian Management*, *14*(1).

Sur, P. and Candès, E.J., 2019. A modern maximum-likelihood theory for high-dimensional logistic regression. *Proceedings of the National Academy of Sciences*, *116*(29), pp.14516-14525

Tchakoute-Tchuigoua, H. (2016). Organizational structure and performance of microfinance institutions. Thunderbird International Business Review, 58(3), 273-285.

Violan, C., Foguet-Boreu, Q., Flores-Mateo, G., Salisbury, C., Blom, J., Freitag, M., Glynn, L., Muth, C. and Valderas, J.M., 2014. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PloS one*, *9*(7), p.e102149.

White, N.M., Balasubramaniam, T., Nayak, R. and Barnett, A.G., 2022. An observational analysis of the trope "A p-value of< 0.05 was considered statistically significant" and other cut-and-paste statistical methods. *PLoS One*, *17*(3), p.e0264360.

Zaby, S., 2019. Science mapping of the global knowledge base on microfinance: Influential authors and documents, 1989–2019. *Sustainability*, *11*(14), p.3883.

APPENDICES

APPENDIX A: Logistic codes

```
# Create a dataframe with the provided data

data <- data.frame(

  YEARS = c(2022, 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002),

  INFLATION_RATE = c(104.71, 98.55, 557.2, 255.3, 10.62, 0.89, -1.54, -2.43, -0.2, 1.63, 3.73, 3.47, 3.02, 0, 0, 24411.03, 1096.68, 3021.12, 282.38, 431.7, 140.06),

  GDP_PER_CAPITAL = c(1306, 1289, 1213, 1343, 2887.63, 2703.61, 2599.66, 1412, 2611.45, 2568.18, 2544.85, 2243.8, 1943.28, 1619.42, 1510.06, 1781.96, 1800.02, 1840.05, 1943.79, 2035.07, 2386.67),

  INTEREST_RATES = c(50, 40, 35, 35.53, 12.14, 10.62, 5.83, 10.1, 15.6, 22.1, 30.8, 10.76, 10.35, 11.2, 815175.7, 7965.8, 371.29, 144.73, 79.24, 51.04, 31.67),

  Non_performing_loans = c(1220, 1330, 5200, 5250, 4000, 6000, 6440, 3460, 3560, 3500, 3200, 1050, 900, 6650, 8000, 15000, 11230, 9920, 10550, 7800, 3200)

)

 summary(data)

 # Plot boxplots

par(mfrow = c(2, 2))  # Arrange plots in a 2x2 grid

boxplot(data$YEARS, main = "Boxplot of Years")

boxplot(data$INFLATION_RATE, main = "Boxplot of Inflation Rate")

boxplot(data$GDP_PER_CAPITAL, main = "Boxplot of GDP per Capital")

boxplot(data$INTEREST_RATES, main = "Boxplot of Interest Rates")

boxplot(data$Non_performing_loans, main = "Boxplot of Non-performing Loans")
```

```
# Load necessary libraries

library(car)

library(lmtest)

library(carData)

# Assuming you have a linear regression model named 'model' already fitted

# Multicollinearity test

vif(model)  # Variance Inflation Factor (VIF)

# Ramsey RESET test

resettest(model)

# White's test for heteroscedasticity

bptest(model)


# Shapiro-Wilk normality test for residuals

shapiro.test(residuals(model))

# Generate synthetic data

set.seed(123)  # Set seed for reproducibility


# Define the number of observations

n <- 1000

# Generate predictor variables

YEARS <- runif(n, 2000, 2022)

INFLATION_RATE <- rnorm(n, 5, 2)

GDP_PER_CAPITAL <- rnorm(n, 2000, 500)
```

```
INTEREST_RATES <- runif(n, 5, 15)

# Create response variable based on predictor variables

# Define a nonlinear relationship

Non_performing_loans <- 0.5 * YEARS + 0.2 * INFLATION_RATE + 0.3 * GDP_PER_CAPITAL - 0.1 * INTEREST_RATES^2 + rnorm(n, 0, 10)

# Convert response variable to binary representation based on mean threshold

Non_performing_loans_binary <- ifelse(Non_performing_loans > mean(Non_performing_loans), 1, 0)

# Create a dataframe

data <- data.frame(

  YEARS = YEARS,

  INFLATION_RATE = INFLATION_RATE,

  GDP_PER_CAPITAL = GDP_PER_CAPITAL,

  INTEREST_RATES = INTEREST_RATES,

  Non_performing_loans = Non_performing_loans,

  Non_performing_loans_binary = Non_performing_loans_binary

)

# Check summary statistics

summary(data)

# Convert Non performing loans into binary representation based on the mean

mean_non_perf_loans <- mean(data$Non_performing_loans)

data$Non_performing_loans_binary <- ifelse(data$Non_performing_loans > mean_non_perf_loans, 1, 0)
```

```
# Fit logistic regression model
model <- glm(Non_performing_loans_binary ~ YEARS + INFLATION_RATE + GDP_PER_CAPITAL + INTEREST_RATES,
        data = data, family = binomial)


# Summarize the model
summary(model)



# Define a function to calculate the confusion matrix
calculate_confusion_matrix <- function(actual, predicted) {
  true_positive <- sum(actual == 1 & predicted == 1)
  true_negative <- sum(actual == 0 & predicted == 0)
  false_positive <- sum(actual == 0 & predicted == 1)
  false_negative <- sum(actual == 1 & predicted == 0)

  confusion_matrix <- matrix(c(true_positive, false_positive, false_negative, true_negative), nrow = 2, byrow = TRUE)
  colnames(confusion_matrix) <- c("Predicted 1", "Predicted 0")
  rownames(confusion_matrix) <- c("Actual 1", "Actual 0")

  return(confusion_matrix)
```

```
}

# Perform predictions using the model
predicted_probabilities <- predict(model, type = "response")
predicted_classes <- ifelse(predicted_probabilities > 0.5, 1, 0)

# Calculate the confusion matrix
conf_matrix <- calculate_confusion_matrix(data$Non_performing_loans_binary,
predicted_classes)
print("Confusion Matrix:")
print(conf_matrix)

# Calculate accuracy
accuracy <- sum(diag(conf_matrix)) / sum(conf_matrix)
print(paste("Accuracy:", accuracy))

# Calculate sensitivity
sensitivity <- conf_matrix[1, 1] / sum(conf_matrix[1, ])
print(paste("Sensitivity (True Positive Rate):", sensitivity))

# Calculate specificity
specificity <- conf_matrix[2, 2] / sum(conf_matrix[2, ])
print(paste("Specificity (True Negative Rate):", specificity))
```

```
library(readxl)

library(ResourceSelection)


# Load the data from the Excel file

data <- read_excel("loan_data.xlsx")


# Rename columns for easier reference

colnames(data) <- c("Years", "InflationRate", "GDPPerCapita", "InterestRates",
"NonPerformingLoans")


# Create a binary outcome variable based on NonPerformingLoans

median_value <- median(data$NonPerformingLoans, na.rm = TRUE)

data$default <- ifelse(data$NonPerformingLoans > median_value, 1, 0)


# Fit a logistic regression model

model <- glm(default ~ InflationRate + GDPPerCapita + InterestRates, data = data, family =
binomial)


# Perform the Hosmer and Lemeshow test with 10 groups

hl_test <- hoslem.test(model$y, fitted(model), g=10)


# Display the test results

print(hl_test)
```

```
# Create the contingency table with formatted numbers
contingency_table <- data.frame(
  Step = 1:nrow(observed),
  Observed_default_0 = format(observed[, 1], nsmall = 2),
  Expected_default_0 = format(expected[, 1], nsmall = 2),
  Observed_default_1 = format(observed[, 2], nsmall = 2),
  Expected_default_1 = format(expected[, 2], nsmall = 2),
  Total = rowSums(observed)
)

# Print the contingency table
print(contingency_table)
# Create the table for the Hosmer and Lemeshow test result
hl_result <- data.frame(
  Step = 1,
  Chi_square = format(hl_test$statistic, nsmall = 2),
  Df = hl_test$parameter,
  Sig = format(hl_test$p.value, nsmall = 2)
)
# Print the Hosmer and Lemeshow test result table
print(hl_result)
```