CREDIT SCORE IN DEFAULT PREDICTION FOR P2P LENDING

SIM HUI XIAN

BACHELOR OF INTERNATIONAL BUSINESS (HONOURS)

UNIVERSITI TUNKU ABDUL RAHMAN

FACULTY OF ACCOUNTANCY AND MANAGEMENT DEPARTMENT OF INTERNATIONAL BUSINESS

APRIL 2024

CREDIT SCORE IN DEFAULT PREDICTION FOR P2P LENDING

BY

SIM HUI XIAN

A final year project submitted in partial fulfilment of the requirement for the degree of

BACHELOR OF INTERNATIONAL BUSINESS (HONOURS)

UNIVERSITI TUNKU ABDUL RAHMAN

FACULTY OF ACCOUNTANCY AND MANAGEMENT DEPARTMENT OF INTERNATIONAL BUSINESS

APRIL 2024

DECLARATION

I hereby declare that:

(1) This undergraduate FYP is the end result of my own work and that due acknowledgement has been given in the references to ALL sources of information be they printed, electronic, or personal.

(2) No portion of this FYP has been submitted in support of any application for any other degree or qualification of this or any other university, or other institutes of learning.

(3) Sole contribution has been made by me in completing the FYP.

(4) The word count of this research report is ______________________.

Name of student: ____________________

Student ID: ________________

Signature: _________________

Date: _________________

ACKNOWLEDGMENT

It is no doubt that a research project cannot be done alone, as so in this research. There are some remarkable individuals who assist me in during a year period in completing this research. Therefore, I would like to take this opportunity to give a sincere appreciation sentence to all the remarkable individuals.

I would like to express the sincere gratitude and appreciation to my research thesis supervisor, Ms. Loh Yin Xia who guide, motivated and support me during the process of this thesis. Without her patient, support, efforts, and supervision. I could not complete this thesis in the right way and on time. Her strong knowledge in producing a research paper is valuable and precise. Her feedback and comments on my thesis increase my knowledge of producing a thesis paper.

Hereby, I would like to express my sincere thanks to Ms. Yeong Wai Mun, my research coordinator who provide useful guidance and instructions on the first step of the research. Her comments and suggestions on guiding me to have a great topic was valuable.

Furthermore, I would like to take this opportunity to thanks to my second examiners Dr Farah Waheeda binti Jalaudin in providing valuable recommendation and feedback on my research presentation in order to enhance the thesis.

I would like to acknowledge the help of Universiti Tunku Abdul Rahman for providing the experience of research as well as providing us the resources for our thesis.

Lastly, I appreciate to all the remarkable individuals who involve in this research both directly and indirectly.

DEDICATION

This dissertation is dedicated to: My supervisor, Ms Loh Yin Xia for guiding and helping me throughout the whole research study.

Tertiary educational institution, Universiti Tunku Abdul Rahman (UTAR) for giving me the opportunity to conduct this research study.

Families and friends, for giving me their supports, encouragement, and help throughout the whole process of research study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

P2P	Peer to Peer
Fintech	Financial Technology
PD	Probability of Default
DTI	Debt to Income Ratio
FICO	Fair Isaac Corporation
DV	Dependent Variable
IV	Independent Variable
RUR	Revolving Utilization Rate
LA	Loan Amount
IR	Interest Rate
TOTRB	Total open to buy on revolving bankcard
BU	Bankcard Utilization Rate
NORA	Number of open revolving accounts
SPSS	Statistical Package for Social Science
NPL	Non Performing Loan

LIST OF APPENDIXES

PERFACE

In recent years, the rapid advancement of technology has disrupted traditional financial services, prompting banks to innovate to stay competitive. The rise of digitalization in finance and the advent of Peer to Peer (P2P) lending offers an alternative to conventional bank loans, particularly for individuals underserved by traditional banking, such as the unbanked or those denied loans. Moreover, there's a growing recognition of social comparison as a factor influencing borrowing habits and debt levels. This trend suggests that individuals may seek to enhance their social status through increased borrowing, potentially leading to higher levels of indebtedness. These developments present both opportunities and challenges for financial institutions. While P2P lending platforms provide new avenues for lending, they also expose lenders to risks such as defaults and loan write-offs. Consequently, lenders commonly rely on the FICO credit score as a tool to assess an individual's creditworthiness for various types of credit, including credit cards, mortgages, or loans. This credit scoring model considers platforms as defaulters if they fail to meet repayment deadlines to investors or experience service disruptions.

ABSTRACT

This research aims to investigate the factors and criteria influencing default in Peer to Peer (P2P) lending, with a focus on providing valuable insights for the future of Fintech and contributing to industry growth and sustainability. The study examines the intention to adopt P2P lending and its implications on financial decision-making. Utilizing quantitative methods, the analysis incorporates variables including loan amount, interest rate, total open-to-buy on revolving bank cards, bank card utilization rate, number of open revolving accounts, debt to income ratio (DTI), and revolving utilization rate. Data from a Kaggle dataset for the year 2018 comprising 445 samples with charge-offs, late payments of 16-30 days, and late payments of 31-120 days is analyzed. Results indicate a highly positive relationship with revolving utilization rate and negative relationships with the number of open revolving accounts and total open-to-buy on revolving bank cards. The implications suggest enhancing credit monitoring within credit assessment processes and implementing alternative data for more accurate evaluations. By accessing FICO scores to assess creditworthiness based on consumer payment behaviour is recommended to improve loan approval processes.

CHAPTER 1: RESEARCH OVERVIEW

1.0 Introduction

This chapter introduce of the background of research, problem statement, research questions, research objectives and significance of study.

1 .1 Research Background

The emergence of Financial Technology (FinTech) companies has revolutionized the financial services industry in innovative business models and cutting-edge technology to offer customers more convenient, accessible, and customer-centric products and services (Windasari et al., 2022).

Over the past decade, social lending platforms, particularly Peer-to-Peer (P2P) lending, have disrupted traditional credit risk assessment services by facilitating direct interaction between lenders and borrowers, bypassing the need for traditional financial institutions (Dharmastuti & Laurentxius, 2021). (Refer Figure 1.1)

From consumer perspective, this evolution has significant in the potential to provide enhanced services (Fast-Moving FinTech Poses Challenge for Regulators, 2022). The digitalization of the financial system and the emergence of P2P lending is an alternative to bank lending, particularly for customers situated beyond the margins of traditional banking services (Octavini & Puspitasari, 2023).

Fintech not only reshape the financial system but also influence how individuals approach their financial decision-making processes (Leckė et al., 2022). There is a growing acceptance of carrying significant levels of debt among the younger generation, who view possessions such as higher education or houses as status symbols within society (Mahdzan et al., 2022). This trend of borrowing to fulfill desired purchases contributes to increased indebtedness, driven by social comparison and the desire to enhance one's power and prestige (Mahdzan et al., 2022). For example, among college students have play a significant role in shaping consumption choices funded through borrowing, the utilization of credit cards, often used as a means of self-definition through consumption (Mahdzan et al., 2022).

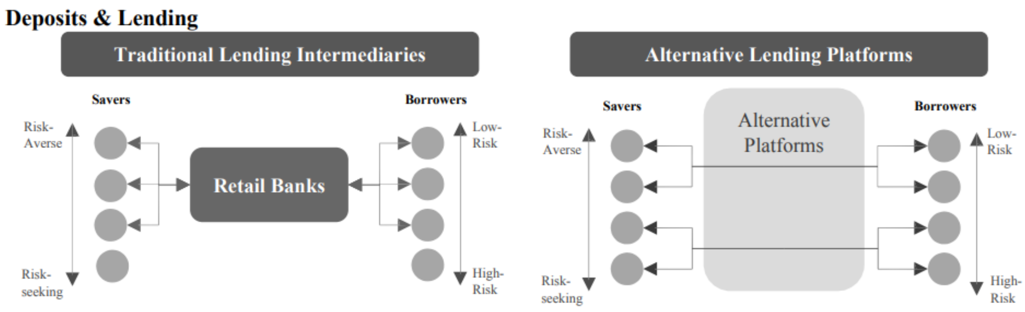
P2P lending platforms offer an efficient and accessible alternative to traditional banks, especially during economic downturns when traditional banks tighten lending requirements and increase interest rates (Croux et al., 2020). However, the adaptability of P2P lending also introduces default risks (W. Gao et al., 2023), posing challenges for regulators in terms of risk management and ensuring the resilience of financial institutions (Fast-Moving FinTech Poses Challenge for Regulators, 2022). To mitigate these risks, P2P lending platforms need to accurately predict the probability of default (PD) of individual loans to optimize profitability and sustainability (Fast-Moving FinTech Poses Challenge for Regulators, 2022).

These circumstances have created the opportunities for financial to provide the platform of P2P lending. However, it also adding default risk of P2P lending (W. Gao et al., 2023). These factors also create a challenge for regulators which risk management system and the resilience of financial institutions, more exposed to risks themselves on default loan or charged off. (Fast-Moving FinTech Poses Challenge for Regulators, 2022).

In order to predict the default loan or charged off loan. P2P lending must determine the probability of default (PD) of individual loans to remain and optimize their profitability and sustainability of businesses.

As a result, this study will examine the factors and characteristics that the default of P2P lending.

Figure 1.1: Traditional Lending Intermediaries and Alternative Lending Platforms



Adapted from: Croux, C., Jagtiani, J., Korivi, T., & Vulanović, M. (2020). Important factors determining Fintech loan default: Evidence from a lendingclub consumer platform. Journal of Economic behaviour and Organization

1.1 Problem Statement

Firstly, P2P lending recently is the new business model that have define in Fintech, their target audience have sharing demographic background (A, 2023). P2P lending directly match lending needs of borrowers with willing lenders, fostering contractual obligations without traditional financial intermediaries growing popularity and utilization of these P2P lending platforms, for instance, like Sofi and Lending Club (Dharmastuti & Laurentxius, 2021). This growth has gained immense popularity due to its cost-effectiveness and convenience, facilitated by technological advancements (Sun, 2020).

However, the challenge of credit risk remains a pressing issue that cannot be overlooked. The factors such as borrowers' domicile, interest rates, inflation, and loan characteristics significantly influence the probability of default (PD) (Siering, 2023). In 2019, Chinese P2P lending platforms faced an exceptionally high default rate of 87.2%, highlighting the severity of the challenges within this industry. These services are highly susceptible to significant losses due to this trend (Gao et al., 2021). Hence, there is a crucial provide insights for enhancing financial literacy, promoting P2P lending, and assessing default risk within the P2P lending market.

To investigate the likelihood of loan repayment and predict potential defaults, the focus revolves around two primary risks in credit analysis. The loss of potential revenue by either rejecting suitable candidates or approving too few, leading to missed business opportunities. The second risk involves financial loss incurred by approving applicants who ultimately fail to repay the loan (Mariiagusarova, 2022).

Credit analysis is integral to borrowing money, with the credit scoring model considering platforms as defaulters if they miss repayment deadlines to investors or encounter service unavailability (Gao et al., 2021). Additionally, there exists a correlation between defaulting on P2P loans and borrowers' creditworthiness. The

PD rises in tandem with borrowers' credit risk, with key factors for loan defaults including borrowers' credit grade, debt to income ratio (DTI), and FICO score, enabling the transfer of borrowers' creditworthiness to lenders (Gao et al., 2021).

Hence, two critical issues have emerged for lending services including the high lending risk for default loan and adverse impact on sustaining profitability within P2P lending services.

In order to conduct this study to identify the risks facing and sustainability of the P2P lending platform. This study conduct based on the historical customer segment of P2P lending to analyze their customer segment in order to provide the default prediction analysis help to reduce lending risk. To optimize the trade-off between revenue and default loss to yield maximum benefit to the business (Mariiagusarova, 2022).

1.2 Research Question

The core of this research is to scrutinize the relationship between the factors in the default of P2P lending. The research questions that align in this business research are listed as below:

1. What are the factors that influence the default of P2P lending?

2. What are the criteria that influence the default of P2P lending?

1.4 Research Objective

The main objective is to examine the factors and characteristics that the default of P2P lending.

1. To examine the factors that influence the default of P2P lending.

2. To exact the criteria that influence the default of P2P lending.

1.5 Research Significance

This research serves a dual purpose which predicting customer default based on metrics and evaluating profitability and sustainability while managing credit risk for lenders within the context of P2P lending. By conducting credit risk assessment, the study facilitates an understanding of spending behaviour and individual financial obligations, thereby assisting P2P lending platforms in evaluating the lending risk linked to spending patterns and informing their lending decisions.

1.6 Definition of Terms

1.6.1 Fintech

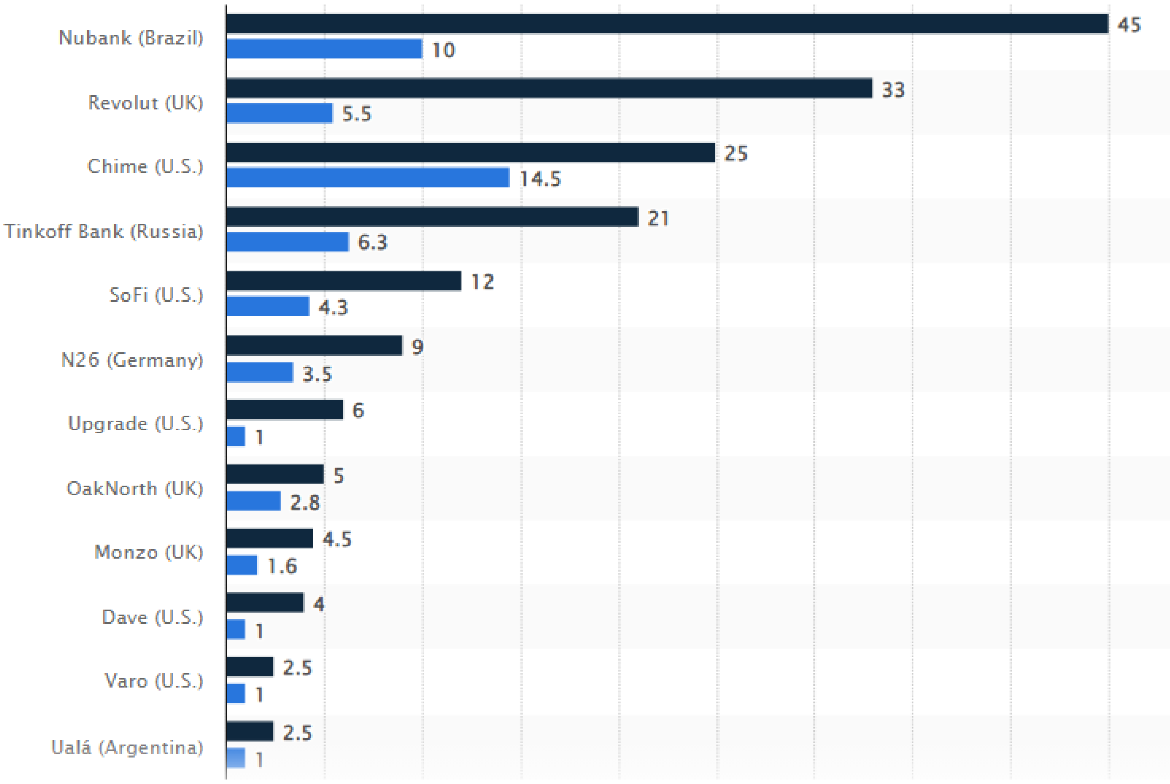
Fintech as known as Financial Technology that represents companies leveraging innovative technology to challenge traditional financial approaches in providing financial services (A, 2023). Examples of their products incorporating technology include credit cards, ATMs, and bank mainframe computers. Within the FinTech, consist of eight categories which payment, insurance, planning, lending and crowdfunding, blockchain, trading and investment, data and analytics, and security (Giglio, 2021).

1.6.2 Neobanks

Neobanks, also known as “challenger banks,” are fintech companies providing applications, software, and other technological solutions for streamlined mobile and online banking, operating without physical branches (Reepu, 2023). These fintech firms typically focus on specific financial products like wealth management, P2P lending, crowdfunding, capital markets, and insurance. They achieve this specialization by minimizing operational expenses and aiming their services toward more specialized or niche market segments (Monis & Pai, 2023). Despite often partnering with larger financial institutions to back their products, neobanks are generally known for their agility and transparency compared to traditional megabanks (Walden, 2021). According to Statista (2023), the example of most valued

Neobank in worldwide include Nubank, Revolut, Chime and Tinkoff Bank. (Refer Figure 1.2)

Figure 1.2: Most valued independent neobanks worldwide



Adapted from: Most valued independent neobanks worldwide (Statista, 2023).

1.6.3 Lending Club

Lending Club is a financial services company, founded since 2007. It was the pioneer in P2P lending that offer loan trading on a secondary market. Over 4.7 million members have joined Lending Club to achieve their financial objectives (Wang et al., 2019). Serving as the singular digital marketplace bank on a large scale, the platform leverages technology to offer members access to an extensive range of financial products and services. By 2015, it had established itself as the largest P2P lending platform globally. Lending Club's primary goal was to assist users in reducing borrowing costs and increasing savings earnings (Chang et al., 2022).

1.6.4 Peer to Peer (P2P) Lending

Peer to Peer (P2P) lending is an online marketplace where individual lenders extend unsecured microloans to individual borrowers in an anonymous setting, facilitating direct lending between individuals and bypassing traditional financial intermediaries. Presently, this innovative form of lending has proliferated globally, with numerous platforms available (Dharmastuti & Laurentxius, 2021). P2P lending platforms act as connectors between borrowers and investors, setting interest rates and terms, thereby facilitating direct lending transactions between people without the involvement of banks. Borrowers in P2P lending typically seek alternatives to conventional banks or aim for lower interest rates (Wang et al., 2019).

1.6.5 Default Prediction

By studying customer segments, default prediction can help financial institutions decide whether to accept or reject their loan applications. Predicting loan defaults is an essential process for financial lenders, aiding in determining the likelihood of a loan facing default (Nagashree, 2023). Successful prediction of loan defaults empowers financial institutions to mitigate the occurrence of problematic loans, ultimately leading to increased profitability. In order to identify characteristics and behaviours that correlate with higher or lower default rates, helping them allocate resources more efficiently and reduce credit losses (Uwais, 2022).

1.6.6 FICO (Fair Isaac Corporation)

FICO is a leading analytics software company, it stands for Fair Isaac Corporation and renowned for creating the credit scoring model bearing the same name. Lenders widely utilize the FICO credit score as a tool to assess an individual's eligibility for various forms of credit, such as credit cards, mortgages, or loans (Hayes, 2023). To determine a borrower's credit worthiness. The scoring system, encompassing payment history, credit utilization, length of credit history, new credit, and credit mix, spans from 300 to 850. Financial institutions widely utilize this system to assess the viability of lending money or granting credit. Typically, a FICO score falling between 670 to 739 signifies a "good" credit history, often viewed favourably by most lenders. Conversely, individuals within the 580 to 669 score range may encounter challenges securing financing at competitive interest rates (Doroghazi, 2020).

1.7 Summary

In summary, this chapter review into research background and research objective, addresses the problem statement concerning P2P lending, with a focus on Lending Club and Neobanks. It emphasizes the significance of the study and its contribution to the field of P2P lending. The next chapter will explore the suggested conceptual framework and variables in more detail.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

This chapter discussed the factors that influencing the profitability and sustainability of the Peer to Peer (P2P) lending business model. A detailed review based on previous literature; the development of hypothesis will be included to be discussed further.

2.1 Underlying Theories

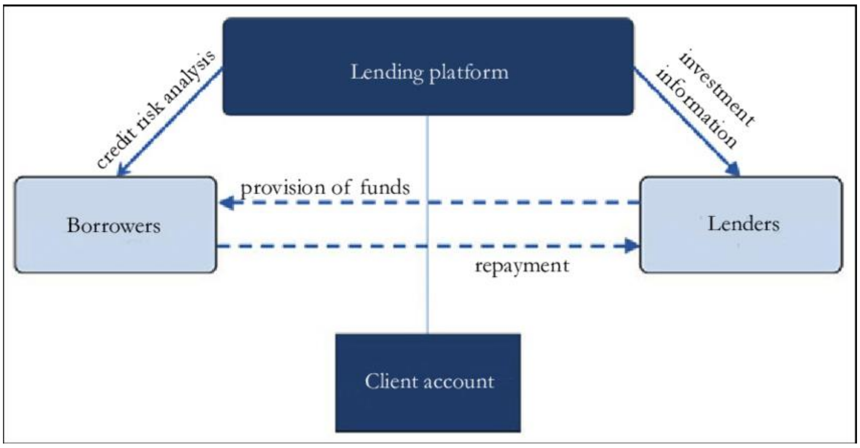
2.1.1 Peer to Peer (P2P) Lending Models

The first online P2P lending platform emerged in 2005 with public launch in UK. The business model of P2P lending that presents the internet technology merges with financial services, enabling individuals and businesses to both lend and borrow money among themselves in small amounts of funds from various sources (Wang et al., 2019). The P2P lending model acts as a catalyst for the loan application process, enabling the entire loan procedure, from fund collection to contract signing and information exchange, to take place entirely over the internet (Suryono et al., 2019).

The overall model conducted over Internet, allowing for the entire loan process included fund, contracts and information procedures (Suryono et al., 2019).

P2P lending model facilities the lending and borrowing money directly between individuals and businesses without involvement in traditional financial intermediaries such as banks. In the academic field, analyzing user behavioural patterns and developing credit or trust models within the context of P2P lending scenarios (Zhang et al., 2023).

Figure 2.1: P2P Lending Model



Adapted from: Figure 1. Stylized Traditional P2P Lending Model., n.d.

2.1.2 Credit Score Model

Credit score model was developed in 1960 where through percentage risk to calculate the lending amount and evaluating the probability of customers repaying borrowed funds within a particular period (Baldo et al., 2023). Credit risk assessment plays a key role to correct and supporting financial institutes in defining their policies and commercial strategies (Doroghazi, 2020).

In the credit loan practices of lending platforms, the credit scoring model is used for accurate analysis the historical credit data in order to distinguish between defaulters and valid customers (Lyócsa et al,. 2022). Credit scoring model concerns with assist financial institutions to aid in their decision making processes particularly in assessing the risk linked to a credit applicant that categorized in non-trustworthy “or non-creditworthy” group (Trivedi, 2020).

While a borrower's FICO score plays a significant role in determining creditworthiness, lenders also consider additional factors like income, employment duration, and the specific type of credit requested (Hayes, 2023). In the credit loan practices of lending platforms, the fundraising process, enabling the participation of any number and size of lenders (Ismawati & Faturohman, 2023).

2.1.3 Risk Management Model

Effective credit risk management is crucial for financial institutions to mitigate the possibility of significant economic losses due to loan defaults (Lee et al., 2021). This involves assessing various data sources when borrowers apply for loans, aiding in the creation of credit risk scoring models to evaluate and manage potential risks associated with lending (Siering, 2023).

There is a high correlation between the possibility of a loan default and the credit risk of the borrower. When borrowers default, it not only impacts the profitability of P2P companies but also jeopardizes the investments of the lenders (Gu et al., 2022). Therefore, the higher interest rates charged to high-risk borrowers are not sufficient to cover the possibility of default on their loans. In other words, borrowers with higher credit scores are often perceived as more trustworthy, displaying a lower PD.

2.1.4 Impulsive Buying Behaviours

In the current era, there's a growing acceptance of the mindset that being in debt is not a significant concern. This shift in attitude has led to an increased willingness among individuals to spend more and rely on borrowing and credit to fulfil their desired purchases. This consumer behaviour aligns with a rising interest in consumption and materialism. This trend towards impulse buying, characterized by spontaneous and emotionally driven purchase behaviours (Lim et al., 2020).

Moreover, impulse buying is defined as among individuals experiencing lower subjective social status, leading to feelings of hopelessness or a lack of control in their lives. Coping with these emotions, some individuals resort to impulsive buying (Abdelsalam et al., 2020). Particularly among individuals perceiving themselves as having lower social status. It can trigger feelings of inferiority and shame, ultimately contributing to negative emotions and impulsive spending habits (Gantulga & Dashrentsen, 2023).

2.2 Review of Variables

The data was collected from Kaggle and contains 7 variables: Loan Amount, Interest Rate, Total open to buy on Revolving Account, Bankcard Utilization Rate, Number of Revolving Account, Debt to Income ratio (DTI) and Revolving Utilization Rate. The variables and their descriptions are shown in Table 2.2.

Table 2.2: Dependent Variables and Independent Variables

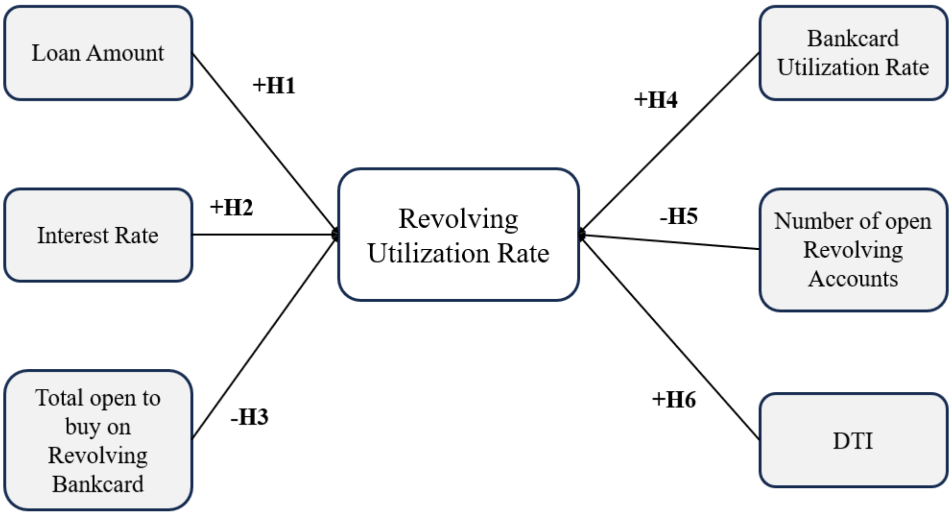
	Terms	**Definition**
Dependent Variables (DV)	**Revolving Utilization Rate**	The proportion of available revolving credit that a borrower is currently utilizing.
	Loan Amount	The specific loan amount applied in loan application.
	Interest Rate	Rate of interest applied to the loan.
	Total open to buy on revolving bankcard	The aggregate credit limit available across multiple credit cards that allows account

Independent Variables (IV)		holders to borrow money repeatedly up to a predetermined limit while repaying the borrowed amount in instalments.
	Bankcard Utilization Rate	The proportion of the total current balance across all bankcard accounts to the high credit limit or maximum credit limit.
	Number of open revolving accounts	The currently active and continuing credit or loan accounts that offer borrowers a maximum limit and flexibility in credit availability.
	Debt to Income (DTI) Ratio	A ratio calculated monthly debt payments excluding mortgage obligations divided by monthly income. Representing a measure of their financial obligations relative to their income.

<u>Source:</u> Kaggle (Mariiagusarova, 2022)

2.3 Conceptual Framework

Figure 2.3: Variables Framework



Adapted from: Developed for the research

2.4 Hypothesis Development

2.4.1 Loan Amount

When borrowers apply for loans through P2P lending platforms, their creditworthiness and the risk of default are assessed to determine if they can be approved (Croux et al., 2020).

From the lending platform's perspective, the loan amount is a crucial detail that defines the terms and conditions of the loan, impacting both the borrower's financial obligations and the lender's risk exposure (Hughes et al., 2022). Including credit history and financial indicators, are considered to evaluate loan applicants, who are then assigned a subgrade or credit rating (Croux et al., 2020).

The loan amount reflects borrowers' behaviour in meeting repayment obligations, revealing their preferences and risk recognition in lending activities (C. Wang et al., 2019). These factors can influence and constrain a borrower’s ability to repay loans (Tang et al., 2023).

Typically, borrowers with high credit limits demonstrate greater reliability in obtaining loans, as they have the capacity to cover instalment payments or the total loan amount (Croux et al., 2020). A higher available revolving credit limit indicates greater credibility in an individual, reflecting a more secure and dependable borrower.

H1: The relationship between Loan Amount and Revolving Utilization Rate

2.4.2 Interest Rate

Borrowers' default probability heavily correlates with the interest rate set by Lending Club. Typically, lenders use borrower characteristics to assess credit risk and determine loan eligibility, interest rates, and loan terms (Lee et al., 2021).

In traditional financial markets, there are the correlation between interest rates and the likelihood of default. In high-interest rate environments, lenders often impose higher charges on loans, increasing the financial burden on borrowers and making it harder for them to meet their obligations (Nagashree, 2023). This situation mirrors P2P lending, where higher real interest rates result in increased costs associated with repaying debts, limiting borrowers' ability to fulfill their obligations and raising the risk of default (Nigmonov et al., 2022).

According to Lending Club reports that interest rates range from 9.57% to 35.99% Annual Percentage Rate. Higher loan risk corresponds to wider interest rate spreads. Consequently, consumers with good credit can access lower interest rates through lending. As high-interest rates of borrowing, it often involves more risk on the part of the investors as these investments or lending prospects are not volatile, riskier and more vulnerable to failure.

H2: The relationship between Interest Rate and Revolving Utilization Rate

2.4.3 Total open to buy Revolving Account

Total open-to-buy revolving account refers to the available credit limit that a purchaser has for making open account purchases, accessible as long as the account remains in good standing. This metric offers valuable insights to lenders, assisting them in assessing a borrower's ability to manage additional debt, analyzing recent credit-seeking behaviour, and evaluating the responsible utilization of existing credit lines when making lending decisions (Nigmonov et al., 2022).

According to Abbassi & F, (2021), it reflects the borrower behaviour, risk assessment during lending, the impact of borrower information on loan repayment processes are crucial aspects of P2P lending. Lenders can utilize this metric to evaluate the risk associated with a loan application and distinguish between high-risk and low-risk applicants. By Examining various dimensions of a borrower aids in constructing a precise borrower profile, allowing them to evaluate credit risk more comprehensively beyond solely relying on credit history (Tang et al., 2023).

H3: The relationship between Total open to buy Revolving Account and Revolving Utilization Rate

2.4.4 Bankcard Utilization Rate

Credit card information is essential for assessing an individual's creditworthiness and financial behaviour. Lenders use these credit card-related metrics to evaluate a borrower's credit risk, financial stability, and capacity to handle additional credit card debt when making lending decisions (Croux et al., 2020).

Bankcard Utilization Rate is an essential factor in P2P lending that lenders consider when assessing the creditworthiness of borrowers and making lending decisions. It serves as a valuable indicator of financial responsibility and influences both the funding success of loan listings and the probability of loan repayment, ultimately impacting the default risk associated with P2P loans (Croux et al., 2020).

Bankcards represent one of the most common and adaptable forms of credit available. Bankcards issuance by banks entails more stringent requirements compared to those established by Lending Club (Croux et al., 2020).

The Bankcard Utilization Rate reflects the proportion of available credit currently being used, with a lower ratio usually indicating responsible credit management and borrower solvency (Mariiagusarova, 2022). For example, possessing bankcards, especially with higher limits, can contribute to borrower solvency. However, increasing financial stress is indicated by a rise in bankcard utilization rates, suggesting a decline in consumer spending. In addition, the fluctuations in the utilization rate are often driven by changes in credit card spending levels (Zhu et al,.2023).

H4: The relationship between Bankcard Utilization Rate and Revolving Utilization Rate

2.4.5 Number of open Revolving Account

The "Number of open revolving accounts" refers to the count of active credit card accounts held by an individual (Mariiagusarova, 2022). This metric is significant because it reflects the diversity of an individual's credit history and contributes to assessing lending risk. In the context of P2P lending platforms, where various types of loans are offered, including personal loans, small business loans, and debt consolidation loans, the number of revolving accounts plays a crucial role (Mariiagusarova, 2022).

Having a higher number of revolving accounts can positively influence the likelihood of successful funding on P2P lending platforms (Xu et al., 2021). It demonstrates a borrower's creditworthiness and responsible credit management, which can reduce default rates and potentially lead to lower interest rates for borrowers. Lenders are more likely to trust borrowers who have demonstrated a history of managing multiple credit accounts responsibly (C. Wang et al., 2019).

Number of open revolving accounts serves as a predictive factor for both successful funding in loan listings and the probability of loan repayment. This means that borrowers with a higher number of active revolving accounts are perceived as lower lending risks and are more likely to receive funding for their loan requests (Uwais, 2022).

Furthermore, the impact of the number of revolving accounts on lending and repayment behaviour directly influences an individual's ability to settle debts due to its financial implications. Borrowers with a higher number of open revolving accounts are typically better equipped to manage their finances and honour their loan obligations, reducing the likelihood of default (Uwais, 2022).

H5: The relationship between Number of open Revolving Account and Revolving Utilization Rate

2.4.6 Debt to Income Ratio (DTI)

DTI is a fundamental financial metric used by lenders to assess an individual's or entity's financial health and creditworthiness. It measures the proportion of one's monthly debt obligations to their gross monthly income. In simpler terms, it shows how much of a person's income is being used to repay debts each month (Nagashree, 2023).

According to Nier (2019), when total borrower monthly debt service obligations exceed 43 percent of monthly gross income, it indicates a higher risk of default. This threshold serves as a guideline for lenders to evaluate the borrower's ability to manage additional debt responsibly. Borrowers with lower DTI ratios are generally considered to have a lower risk of default because they have more disposable income after covering their debt payments (Nigmonov et al., 2022).

High DTI ratios are often associated with a greater likelihood of mortgage default, as borrowers with a significant portion of their income allocated to debt payments may struggle to keep up with mortgage payments, especially if their income decreases or if they encounter unexpected financial challenges (Nigmonov et al., 2022). Moreover, high DTI ratios may reflect poor financial management strategies, potentially leading to difficulties in meeting future financial obligations.

In summary, the DTI ratio is a critical indicator of financial stability and risk for lenders, helping them assess an individual's capacity to take on additional debt responsibly and manage their financial obligations effectively.

H6: The relationship between DTI and Revolving Utilization Rate

Below is the result summary of the relationship between dependent variable and independent variable.

Table 2.4: Summary of Hypothesis Development

L**oan Amount**	H1: The relationship between Loan Amount and Revolving Utilization Rate
Interest Rate	H2: The relationship between Interest Rate and Revolving Utilization Rate
Total open to buy on Revolving Account	H3: The relationship between Total open to buy on Revolving Account and Revolving Utilization Rate
Bankcard Utilization Rate	H4: The relationship between Bankcard Utilization Rate and Revolving Utilization Rate

Number of open Revolving Account	H5: The relationship between Number of Revolving Account and Revolving Utilization Rate
Debt To Income (DTI)	H6: The relationship between DTI and Revolving Utilization Rate

Source: Developed for the research

2.5 Summary

A key challenge faced by individuals investors in P2P lending should allocate effectively their funds across various loans, requiring accurately assessing each loan's credit risk. However, these risks can be can be minimized by capitalizing on the benefits derived from investments.

As a result, this chapter have focussed the investigation of P2P lending in the factors and criteria that influencing the profitability and sustainability of P2P lending. Through the examination the relationship between IVs to in depth understanding of the probability of default loan.

CHP 3: METHODLOGY

3.0 Introduction

This chapter will discuss the research methods used during the research process. In this chapter included research design, sampling design, data sources and data analysis tool.

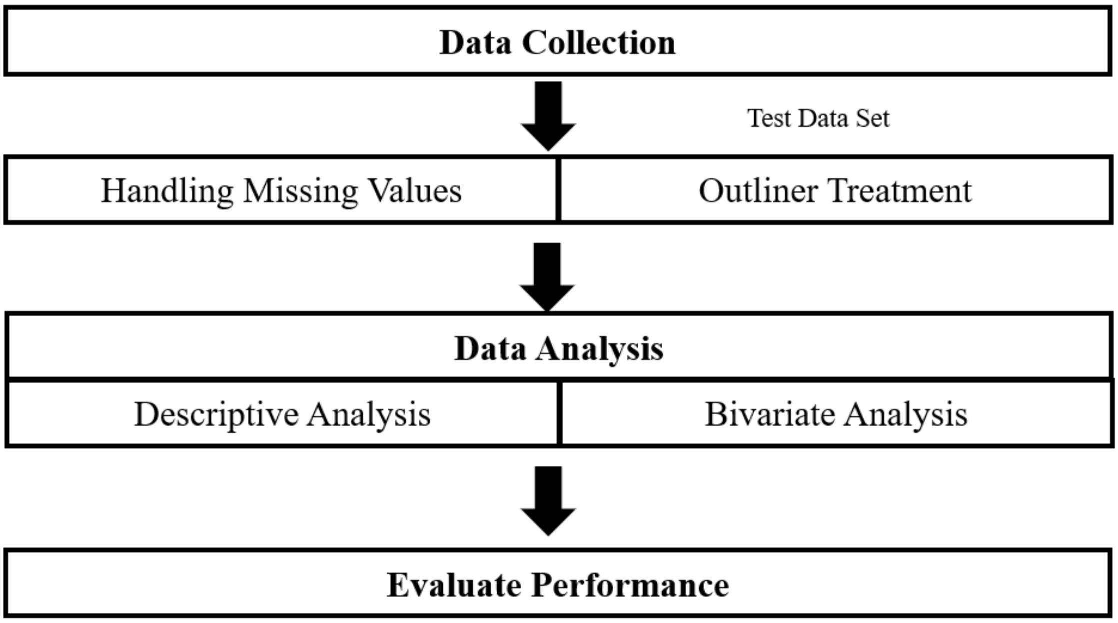
3.1 Research Design

Research designs are plans and the procedures for research that span the decisions from broad assumptions to detailed methods of data collection and analysis (Creswell & Creswell, 2022). It provides an appropriate framework for formulate their problem and objective and present the findings from the data obtained during the study period.

In this study, quantitative approaches are appropriate for large sample sizes of target respondents and contains multiple types of analysis and measurement. Since the data obtained is numerical and evaluated using mathematical and statistical approaches (Creswell & Creswell, 2022).

Moreover, descriptive approach is used in this research. It can analyze one or more variables using a range of research methodologies as well as determine trends, characteristics, categories, and frequencies (Creswell & Creswell, 2022).

Figure 3.1: Research Design



Adapted from: (Creswell & Creswell, 2022)

3.2 Sampling Design

3.2.1 Target Population

Population is referred to the parent group from which a sample is to be drawn to understand the complete mass of observations (Pandey, P., & Pandey, M. M., 2021). It is important for researchers to ensure the eligibility of respondents to obtain the accurate data and to generate a valid and reliable insight. The objective of this study is to understand the solvency of the borrowers from P2P lending. Therefore, the target population of this study are those people have the intention to apply P2P loan.

P2P lending targets the unbanked section of the population or people with limited institutional credit exposure (Musari, 2022). P2P lending platforms also serve as a means for recovery of loans and can be used by borrowers who are unable to obtain loans from traditional sources ("Money Market Executive's Perception Towards Peer-to-peer (P2p) Lending," 2021). It is also a financing option for individuals and small businesses with little or poor credit history (Gupta, 2019).

3.2.2 Sampling Frame

A sampling frame is used in this study as a tool to define the population of specific interest (Berndt, A. E., 2020). The sampling frame is narrowed to the population that is active in P2P lending market.

The sampling frame is narrowed to the population who applied for P2P loan and the target customer of P2P business model. Thus, in this research the sampling frame focus on the customer who have go through the application process and applied successfully on the P2P loan. P2P lending offers a means of household finance, particularly in economically deprived areas affected by austerity measures, by providing consumer loans to individuals facing banking and digital exclusion.

3.2.3 Sampling Size

Sample size refers to the number of individuals or observations selected from a larger population to be included in a statistical sample (Lakens, D., 2022).

The target population in this study are those unserved banked and seeking for loan approval. However, the size of the target population in this study is hard to define.

This study uses data from the Lending Club platform. The data set contains information about 2260701 entries and 150 features between January 2015 and December 2018 that have established in Kaggle which includes 672377 positive instances fully paid.

Over the Kaggle study period of Lending Club lent into borrowers. The loans in the data set have 7 different statuses which are Fully Paid, Charged Off, Current, Late in 31 to 120days and Late in 16 to 30 days, In Grace Period, Default. However, to determine between "good" and "bad" loans and estimate the probability of loan default, certain categories such as Current, Default, Fully Paid, and In Grade Period were excluded. This exclusion was necessary to focus specifically on the outcome of default. Additionally, the default category was converted into the charge-off category, as charge-off pertains to the borrower's side, while default indicates the status from the lender's perspective.

In short, the data used 2018 to single out loan applications filed by year and selected in 2018 of 1500 respondents.

3.2.4 Sampling Technique

Purposeful sampling refers to a group of non-probability sampling techniques in which units are selected because they have characteristics that are needed sample (Thomas, F. B.,2022). Unlike other methods of random sampling, a purposive sampling strategy ensures that the exact types of instances that can be included are included in the final sample of the study. Hence, the final sample consist of funded loans in "Charged Off", "Late in 31 to 120days" and "Late in 16 to 30 days" in 2018 of 1500 respondents.

3.3 Data Sources

3.3.1 Secondary Data

Secondary data is information that researchers can easily use in their research that has been gathered from primary sources (Renbarger et al., 2019). When conducting secondary data analysis, researchers conduct secondary data analysis by repurposing information collected by others to address their own unique research questions (Renbarger et al., 2019). Secondary data sources include books, personal records, newspapers, journals, websites, government reports, and more. Acquiring secondary data is easier and requires less effort compared to primary data collection. With the advent of electronic media and the Internet, accessing secondary data sources has become even more convenient (Renbarger et al., 2019).

Following research, it was observed that many articles utilize the Lending Club dataset as their primary case study. This dataset is publicly available and can also be found on platforms like Kaggle.

3.4 Data Analytical Software

Initially, to determine the sample size for this study, G*Power was employed, a tool enabling researchers to ascertain sample sizes based on parameters like effect size, significance level, and desired power (Kang, 2021). After inputting the parameters, the minimum sample size required was found to be 200.

Subsequently, Statistical Package for Social Science (SPSS) Statistics, a widely used statistical data analysis software, was employed (Watkins, 2021). Following data processing, the data file was inserted into SPSS, and numerical data were labelled for analysis purposes. SPSS facilitated various analyses such as reliability testing, descriptive analysis, and Pearson correlation analysis.

3.5 Data Analysis

Data analysis is the application of reasoning to understand the data that have been gathered (Zikmund et al., 2010).

Lending Club data set contains a comprehensive list of features that enable to employ to train our model for prediction and forecasting. The data set includes detailed information for every loan issued by Lending Club from 2007 to 2015, including a borrower's annual incomes, subgrade, revolving balances, purpose for borrowing and so on. Through the dataset enable to optimize the trade-off between revenue and default loss to yield maximum benefit to the business (Mariiagusarova, 2022).

However, in this study that the data was collected from Kaggle and contains 6 individual variables and 1 dependent variable which loan amount, interest rate, total open to buy on revolving bankcard, bankcard utilization rate, number of revolving accounts, Debt to Income ratio (DTI) and revolving utilization rate.

3.5.1 Bivariate Analysis

Bivariate analysis refers to the statistical analysis of data collected on 7 variables, which may be correlated with each other. It is commonly used when examining phenomena that involve multiple factors or variables. In this analysis, gaining business-level insights on what factors are causing the label outcome, in this case, how each feature impacts loan payment outcome.

The following categorical variables which ' loan amount ', ' interest rate ' , ' total open to buy on revolving bankcard ', 'bankcard utilization rate ' , ' number of revolving account ' and 'DTI' will be compared to the dependent variable 'Revolving Utilization Rate' (Mariiagusarova, 2022).

3.5.2 Descriptive Analysis

Descriptive Analysis refers to data that consists solely of observations on a single attribute. The basic goal of descriptive analysis is helping to explore each feature in a dataset which don’t look into the relationship between features (Siedlecki, 2020). The data is visually shown using graphing. During this study, visual representation of data in graph. The primary purpose of graphs is to convey data, summarize data, enhance verbal descriptions, describe and explore data, facilitate comparisons, avoid distortions and stimulate thinking about the data. The categorical and ordinal features explored 'term of years', 'subgrade', 'home ownership', 'loan status' , 'employee length', 'months since rent inquiries' and 'annual income' (Mariiagusarova, 2022).

3.6 Summary

In conclusion, the research methodology to conduct the study is through deductive approach and quantitative research is applied. The sample size, sampling design, sampling techniques are all applied in secondary data that through Python to understand the default in P2P lending. These enables for the further data analysis which would be discussed in next chapter.

CHP 4: DATA ANALYSIS

4.0 Introduction

This chapter briefly discusses data analysis, providing a comprehensive overview of descriptive and bivariate analysis. It includes sections on data preparation, normality testing, Confirmatory Factor Analysis, descriptive analysis, Pearson correlation, hypothesis testing, and a summary of the chapter's summary.

4.1 Data Preparation

Data preprocessing is crucial for preparing and transforming data to ensure it fits into data mining algorithms and produces accurate results (Garcia et al., 2016). This process involves cleaning, reducing, transforming, and integrating data (Garcia et al., 2016).

In the case of the Lending Club borrowing data from 2005 to 2018 obtained from Kaggle, measures were taken to eliminate noise from the dataset in a Comma-separated values (CSV) file (Çetin & YILDIZ, 2022). Additionally, normalization processing was applied using the logarithm function ("Log") in Excel to ensure a more uniform distribution across selected features.

Furthermore, a thorough filtering process was conducted to select relevant independent variables as features for analysis and to convert categorical string features while encoding each detailed value according to the range of the P2P

Lending Club dataset (Howitt, 2020)..

Removing outliers from the dataset was also performed to enhance the accuracy of subsequent analyses and models (Smiti, 2020). In this scenario, outliers were identified in several key variables: Revolving Utilization Rate, Bankcard Utilization Rate, Months Since Recent Inquiries, Number of Revolving Accounts, and Bankcard Open to Buy were eliminated to improve overall coherence and logical consistency.

Finally, attributes that were replicated or did not contain predictive variables, such as address status, were removed from the dataset and resulting in 445 rows and 16 columns. The remaining columns include Debt to Income (DTI), Revolving Utilization Rate, Loan Amount, Bankcard Utilization, Number Open Revolving Account, Interest Rate, Total Open to buy revolving card, subgrade, home ownership, annual income, loan status, employee length, term of years, and months since recent inquiry.

4.2 Data Coding

Since using SPSS as the data technique, imported the data file that into SPSS and labelling all the data can be categorized into several distinct phases, each indicating the state.

Table 4.1: Summary of Data Coding

Subgrade	1 = A1 – A5 2 = B1 – B5 3 = C1 – C5 4 = D1 – D5 5 = E1 – E5

	6 = F1 – F2 7 = G1 – G5
Homeownership	1 = Own 2 = Rent 3 = Mortgage
Loan Status	1 = Current 2 = In grace period 3 = Late (16-30 days) 4 = Late (31-120 days) 5 = Charged off
Employee Length	1 = ≥ 10 years 2 = 6 – 9 years 3 = ≤ 5 years

Source: Developed for the research.

After encoding, a normality test was conducted to determine whether the sample data was drawn from a normally distributed population. The results of the normality test indicated that all IVs were drawn from a normally distributed population and identified outliers. The outliers were then assessed to determine if they were valid values or not, and subsequently either removed or normalized.

In the normality test, if the significance value exceeds 0.05, the data is considered to be normally distributed. Conversely, if the significance value is below 0.05, it indicates a significant deviation from a normal distribution (Park, 2015). Based on the results, the significance values for Loan Amount, Interest Rate, Number of Open Revolving Accounts, Revolving Utilization Rate, Total open to buy on revolving bankcard, Bankcard Utilization Rate, and DTI demonstrated normal distribution.

Table 4.2: Normality Test

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Loan Amount	.119	445	**.000**	.938	445	**.000**
Interest Rate	.058	445	**.001**	.983	445	**.000**
Number of Revolving Account	.134	445	**.000**	.919	445	**.000**
Revolving Utilization Rate	.123	445	**.000**	.822	445	**.000**
Total open to buy on Revolving Bankcard	.198	445	**.000**	.754	445	**.000**
Bankcard Utilization Rate	.202	445	**.000**	.683	445	**.000**
DTI	.209	445	**.000**	.811	445	**.000**

a. Lilliefors Significance Correction

Source: Develop from SPSS

4.3 Data Validation

The aim of conducting Confirmatory Factor Analysis (CFA) is used to test how well the measured variables represent the number of constructs (Sureshchandar, 2021)**.** A total of 445 samples were selected, and 14 items were chosen for analysis. The results of the CFA, presented in Table 4.17, indicate that factor extraction ranged from 0.40 to 0.85. The result demonstrated the factor extraction lies between 0.40 to 0.85. The result implied that Loan Amount, Interest Rate, Total open to buy on Revolving Bankard, Bankcard Utilization Rate, Number of Revolving Account, DTI and Revolving Utilization Rate more than 0.5 that indicated significant. Conversely, Months since Recent Inquiries had a loading below 0.5, indicating insignificance.

Furthermore, given that Annual Income exhibited similar characteristics to DTI but with a lower loading, it was decided to exclude Annual Income from further analysis.

Table 4.3： Confirmatory Factor Analysis

Communalities

	Initial	Extraction
Loan Amount	1.000	.685
Interest Rate	1.000	.600
Revolving Utilization Rate	1.000	.829
Total open to buy on Revolving Bankcard	1.000	.691
Bankcard Utilization Rate	1.000	.771
Number of Revolving Account	1.000	.564
DTI	1.000	.827
Months since recent inquiries	1.000	**.489**
Annual Income	1.000	.723

Extraction Method: Principal Component Analysis.

Sources: Source: Develop from SPSS

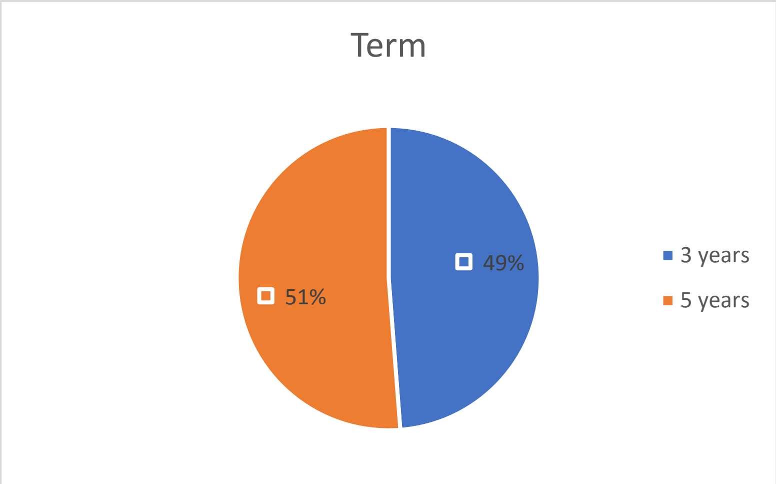
4.4 Data Analysis

4.4.1 Descriptive Analysis

All the data have present in graph in below. In this study, the data in Dataset can be downloaded from the Kaggle. Below showed that the data set is largely representative of 445 sample in terms of terms of years, subgrade, employee length, home ownership, loan status, months since recent inquiry and annual income.

4.4.1.1 Term of year

Figure 4.4: Term of Year

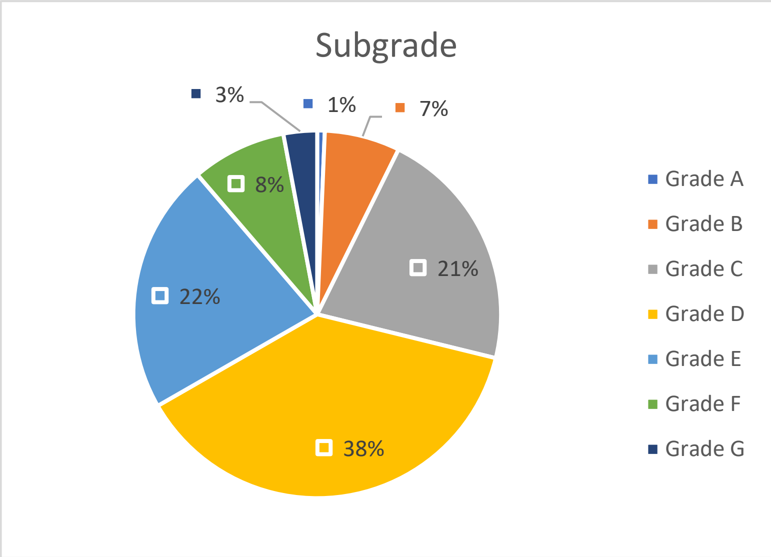


Adapted from: Developed for the research

Figure 4.4 show that for the 'Term' variable, 3 represents 3 years and 5 represents 5 years. 3 years are around 49%, and 5 years are around 51 %, which is roughly the same.

4.4.1.2 Subgrade

Figure 4.5: Subgrade

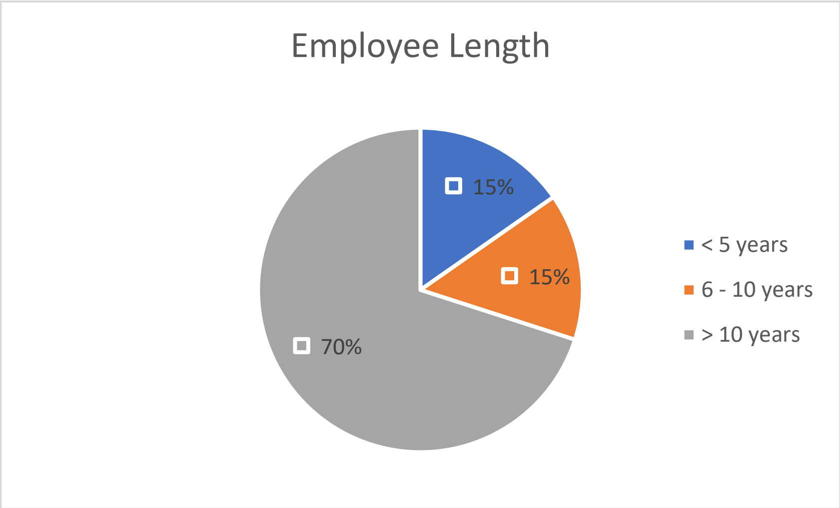


Adapted from: Developed for the research

The descriptive statistics show that for the 'Subgrade' variable, 1 represents A1 to A5 (1%), 2 represents B1 to B5 (7%), 3 represents C1 to C5 (21%), 4 represent D1 to D5 (38%), 5 represents E1 to E5 (22%), 6 represents F1 to F2 (8%) and 7 represents G1 to G5 (3%). From the statistics shown that majority of borrowers are from Grade D in their credit score.

4.4.1.3 Employee Length

Figure 4.6: Employee Length

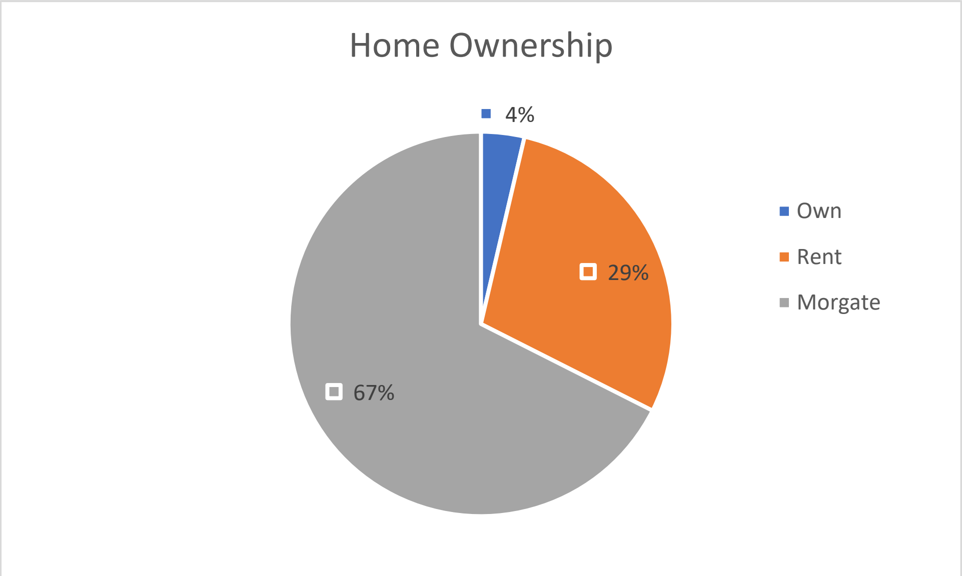


Adapted from: Developed for the research

The sample of Employee Length in year consisted of more than 10 years (15%), 6 years to 9 years (15%) and less than 5 years (15%) in the dataset (N=445).

4.4.1.4 Home Ownership

Figure 4.7: Home Ownership

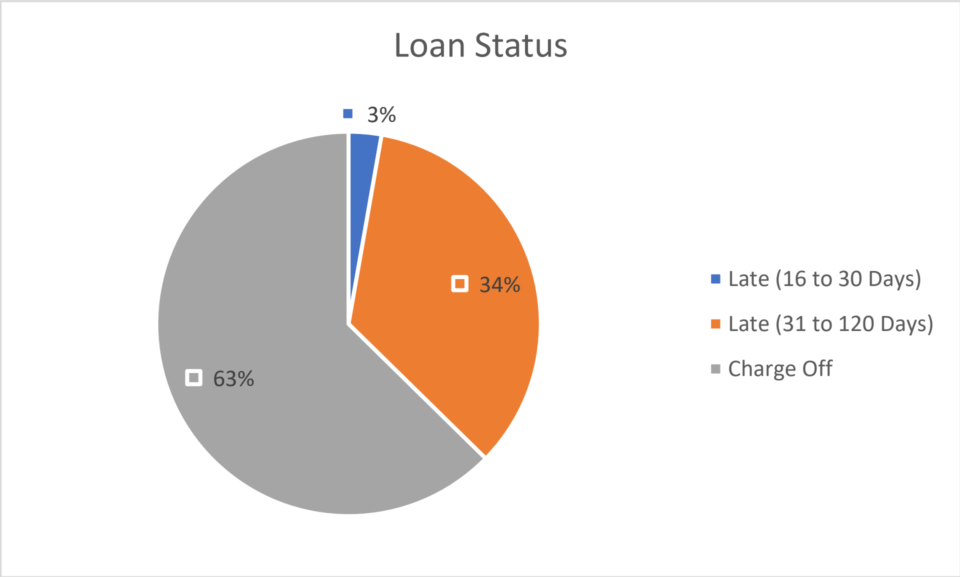


Adapted from: Developed for the research

There are only 3 numerical attributes chosen in this variable, which include own (4%), rent (29%) and mortgage (67%).

4.4.1.5 Loan status

Figure 4.8: Loan Status



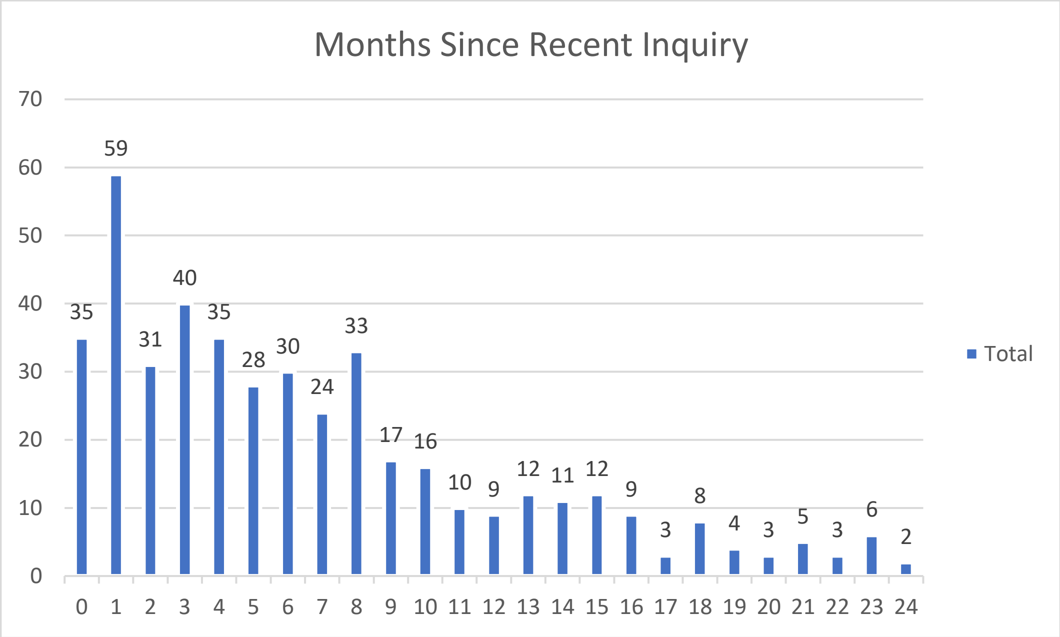
Adapted from: Developed for the research

Figure 4.8 show that the ‘Loan Status variables are represented by status of borrowers, 1 for Late (16 to 30 Days), 2 for Late (31 to 120 Days) and 3 for Charge Off.

According to figure 4.8 Late (16 to 30 Days) have the lowest percentage in 3%, whileCharge Off have the highest percentage in 63%. As a result, the category for this attribute with the status have recorded accordingly.

4.4.1.6 Months since recent Inquiry

Figure 4.9: Months Since Recent Inquiry

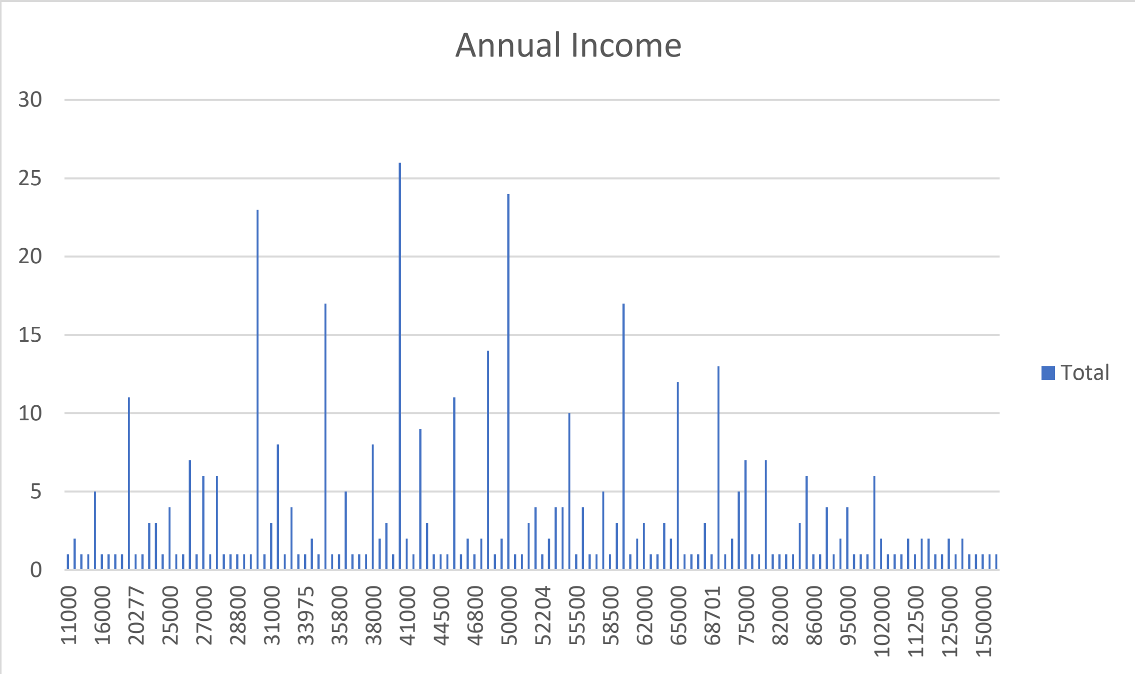


Adapted from: Developed for the research

The descriptive statistics show that for the 'Months Since Recent Inquiry' variable in past 24 months. According to figure 4.9, the lowest in the duration in months since the latest inquiries occurred shown in after 24 months and the higher toward less than a month to 8 months.

4.4.1.7 Annual Income

Figure 4.10: Annual Income



Adapted from: Developed for the research

The descriptive statistics show that for the ‘Annual Income’ variable of borrowers. According to figure 4.10, it indicated the annual income of borrowers majority are around 31000 USD to 50000 USD.

4.4.2 Bivariate Analysis

Table 4.11: Result of Bivariate Analysis

Hypothesis	Path	P- value	Pearson Correlation	Result
H1	LA → IR	0.217	0.730	**Positive**
H2	LA → RUR	0.029	0.103*	**Positive**
H3	LA → TOTRB	0.277	0.520	**Positive**
H4	LA → BU	0.790	0.083	**Positive**
H5	LA → NORA	0.052	0.092	**Positive**
H6	LA → DTI	0.000	0.179**	**Positive**
H7	IR → RUR	0.033	0.101*	**Positive**
H8	IR → TOTRB	0.000	0.214	Negative
H9	IR → BU	0.052	0.092	**Positive**
H10	IR → NORA	0.026	0.105*	Negative
H11	IR → DTI	0.171	0.65	Negative
H12	RUR → TOTBR	0.000	0.4**	Negative
H13	RUR → BU	0.000	0.75**	**Positive**
H14	RUR → NORA	0.000	0.182**	Negative
H15	RUR → DTI	0.346	0.045	**Positive**
H16	TOTBR → BU	0.000	0.312**	Negative
H17	TOTBR → NORA	0.000	0.441**	**Positive**
H18	TOTBR → DTI	0.011	0.121*	Negative
H19	BU → NORA	0.352	0.044	Negative
H20	BU → DTI	0.479	0.034	**Positive**
H21	NORA → DTI	0.755	0.015	Negative

Source: Developed for the research

Based on the table 4.18, all the relationships between IV and DV from H1 to H7, H9, H13, H15, H17, and H20 are statistically significant as their p-values are below 0.05. The correlation coefficient for H3 falls between 0.50 and 0.70, indicating a moderate positive correlation, while H17 shows a correlation between 0.30 and 0.50, suggesting a low positive correlation.

Relationships with correlation coefficients between 0.00 and 0.30, such as H2, H4, H5, H6, H7, H9, and H15, exhibit negligible positive correlations.

Moreover, the relationship between Interest Rate and DTI (H11) demonstrates a moderate negative correlation within the range of 0.50 to 0.70. Similarly, H16 and H12 both indicate low negative correlations, falling between 0.30 and 0.50. Conversely, H10, H14, H18, H19, and H21 exhibit negligible negative correlations.

4.5 Summary

Based on the decision rule for assessing the Pearson Correlation if p-value less than 0.05, test is significant and not significant in more than 0.05. (Howitt, 2020). There are 12 significant relationships between the IVs in 21 hypotheses testing.

CHP 5: DISCUSSION, RECOMMENDATION AND LIMITATION

5.0 Introduction

Chapter 5 provides an overview of the statistical analysis conducted in the study, encompassing bivariate analysis and descriptive analysis. The chapter also includes discussions on the main findings, theoretical implications, practical implications, limitations of the study, recommendations, and conclusions.

5.1 Summary of Statistical Analysis

In this study have utilized data from Lending Club, the largest P2P platform in the United States, to verify actual default and charge-off results. Table 5.1 provides detailed characteristics of the data.

Chapter 4 results revealed 11 significant relationships among the IVs in 21 hypotheses tested. A majority of loans, around 51%, had a term of 5 years, with the majority of subgrades falling in grade D and around 70% having an employee length of more than 10 years. Additionally, 67% of loans were for mortgage purposes, and approximately 63% ended up in charge-offs. The median annual income was $50,000.

Borrowers were classified into seven credit score categories from A to G, with lower scores indicating higher risk for investors. Notably, borrowers with a "D" score were more prone to default on their loans. Predicting which loans are likely to turn bad is crucial for Lending Club's future growth, especially considering that many of these loans were applied for mortgage purposes and eventually charged off.

Logically, lower credit scores entail higher risk, leading to higher interest rates paid by clients to investors. Loans with longer terms, specifically 5 years, are more likely to turn bad. A high number of bad loans increases lending risk for Lending Club, potentially deterring investors and disrupting the capital chain, possibly leading to bankruptcy. Thus, these loans tend to pose a higher risk of turning bad (Ma &, W., 2021).

5.1.1 Discussion of Major Findings

Table 5.1: Summary of Major Findings

Hypothesis	**Relationship**	**Result**
H1:	Revolving Utilization Rate has significance influence on Loan Amount during P2P Lending in Lending Club	**Accepted**
H2:	Revolving Utilization Rate has significance influence on Interest Rate during P2P Lending in Lending Club	**Accepted**
H3:	Revolving Utilization Rate has significance influence on Total open to buy on Revolving Bankcard during P2P Lending in Lending Club	**Accepted**
H4:	Revolving Utilization Rate has significance influence on Bankcard Utilization Rate during P2P Lending in Lending Club	**Accepted**
H5:	Revolving Utilization Rate has significance influence on Number of open Revolving Accounts during P2P Lending in Lending Club	**Accepted**
H6:	Revolving Utilization Rate has significance influence on DTI during P2P Lending in Lending Club	Rejected

Source: Developed for the research

5.1.1.1 Relationship between Revolving Utilization Rate has significant influence on Loan Amount during P2P Lending in Lending Club

A significant influence of loan amount on revolving utilization rate suggests that as the loan amount increases or decreases, there is a consequential effect on how much of the available revolving credit a borrower utilizes. For instance, as the loan amount increases, borrowers may tend to utilize a higher proportion of their available revolving credit, potentially because they have more financial obligations to meet (Leckė et al., 2022).

On the other hand, individuals may also rely more heavily on their available revolving credit to cover additional expenses, leading to higher utilization rates, which can decrease the likelihood of loan approval (Madeira, 2023). Conversely, as individuals repay loans or reduce their debt burden, they may utilize less of their available revolving credit, resulting in lower utilization rates (Field, 2024).

Moreover, borrowers who take out larger loans may need to manage their revolving credit more cautiously to avoid exceeding their credit limits or negatively impacting their credit scores. Therefore, higher loan amounts carried by borrowers often increase the potential risk of default (Leckė et al., 2022).

5.1.1.2 Relationship between Revolving Utilization Rate has significant influence on Interest Rate during P2P Lending in Lending Club

A low positive relationship between interest rate and revolving utilization rate indicates that as the revolving utilization rate increases, there is a slight tendency for the interest rate to also increase.

Borrowers with higher loan amounts might have higher credit utilization ratios, potentially impacting their available credit limits and affect the interest rate also tends to increase significantly (Zhu et al,.2023). Typically, higher loan amounts might signal higher risk to lenders, as larger loans may pose a greater risk of default. Consequently, lenders may increase interest rates to compensate for this higher risk. especially in developing countries where consumer credit is widely used, often at high interest rates and high default risk (Madeira 2018).

Moreover, riskier borrowers may be willing to accept higher interest rates to compensate lenders for the increased risk of default (Hughes et al., 2022). This can lead to adverse selection, where lenders attract a disproportionate number of high-risk borrowers willing to pay higher rates. Therefore, investors charge a higher interest rate on borrowing, given the increased risk of lending or defaulting on the part of borrowers. which is transferred to investors in the P2P market.

5.1.1.3 Relationship between Revolving Utilization Rate has significant influence on Total open to buy on Revolving Bankcard during P2P Lending in Lending Club

A negative significant relationship could indicate that individuals with low revolving utilization rate tend to have more total open to buy on revolving bankcard.

As remain the revolving utilization rate on 30% below as a good revolving utilization rate which have the low total debt holding and large portion of credit available (Gupta, 2019). If a borrower's revolving utilization ’ate Is high, indicating that they are already using a significant portion of their available credit. As such, borrowers with high revolving utilization rates may exhibit riskier financial behaviour, which could influence lenders’ decisions regarding total open to buy on revolving bankcards. As the revolving utilization rate increases, the total open-to-buy decreases, and vice versa.

Moreover, borrowers with lower credit scores are less likely to be funded and more likely to default and end up with higher interest rates. However, while low and medium levels of bankcard use signal the creditworthiness of borrowers, very high levels of bankcard use lead to lower funding probabilities and higher interest rates due to the risk of high leverage and vulnerability to shocks (Serrano-Cinca & Nieto, 2016).

5.1.1.4 Relationship between Revolving Utilization Rate has high significant influence Bankcard Utilization Rate during P2P Lending in Lending Club

Bankcard Utilization Rate is focuses on the portion of available credit that a person is using on their credit cards which looking only at the credit card portion of revolving credit (Kim, 2021). The more accounts with a past due balance, and larger balances past due.

The bankcard utilization rate has reflected the financial behaviour of borrowers that could indicate a certain financial behaviour or spending pattern. Individuals who may take out larger loans may also tend to use a higher proportion of their available credit on their bankcards (Lim et al., 2020). Thus, the bankcard has higher debt amount might lead individuals to utilize more of their available credit on bankcards to manage their finances, possibly due to increased expenses or a need for liquidity and as a sign of financial strain or increased credit risk. Thus Lenders may consider borrowers with a high bankcard utilization rate as having a higher capacity to take on additional debt, which could affect their financial stability.

Moreover, the bankcard utilization rate and the revolving utilization rate have affected by the credit score calculations. Higher utilization rates can negatively impact credit scores, leading to higher interest rates or rejection from lenders.

Furthermore, borrowers with a high total open to buy may be perceived as riskier if they have a history of maxing out their credit limits or carrying high balances on their credit cards. Thus Lenders may be more conservative in extending additional credit to borrowers who are already heavily indebted. Hence, the portion of credit available will decrease in revolving utilization rate.

5.1.1.5 Relationship between Revolving Utilization Rate has significant influence on Number of open Revolving Account during P2P Lending in Lending Club

A negative significant relationship could indicate that individuals with low revolving utilization rate tend to have more open revolving accounts. This could suggest that individuals with larger loans have established credit histories or higher creditworthiness, allowing them to access more credit accounts. Lenders might also consider the number of open revolving accounts when determining loan amounts. For instance, individuals with a higher number of open revolving accounts might be seen as less risky borrowers. Therefore, lenders may be more cautious about approving additional accounts for borrowers with high utilization rates to mitigate their exposure to risk.

However, a high revolving utilization rate also may indicate that a borrower is heavily reliant on credit and may have difficulty managing additional debt. In such cases, the number of open revolving accounts may be limited as lenders may perceive borrowers with high utilization rates as riskier. High revolving utilization rates may indicate that borrowers are utilizing a significant portion of their available credit, potentially signalling financial strain or instability. Lenders may consider this factor when evaluating the number of open revolving accounts, they are willing to extend to borrowers.

Borrowers with high utilization rates may be perceived as less financially stable, leading lenders to limit the number of accounts they can open (Ma &, W., 2021).

5.1.1.6 Relationship between Revolving Utilization Rate has significant influence on DTI during P2P Lending in Lending Club

An insignificant relationship between revolving utilization rate and DTI ratio. Both revolving utilization rate and DTI ratio are key indicators of an applicant's financial health and ability to manage debt.

Debt-to-Income (DTI) ratio, on the other hand, measures the percentage of a person's monthly income that goes toward debt payments. This ratio is used by lenders to assess an individual's ability to manage monthly payments and is often considered in loan approval process (Ismawati & Faturohman, 2023). A high DTI ratio may indicate that a person has a significant amount of debt relative to their income, which could impact their ability to take on additional debt, leading to rejection or less favourable loan terms.

While there can be some overlap between the factors influencing revolving utilization rate and DTI ratio, they are not directly correlated (Ismawati & Faturohman, 2023). Someone could have a low revolving utilization rate but a high DTI ratio if they have significant non-revolving debt like mortgages or car loans. Similarly, someone could have a high revolving utilization rate but a low DTI ratio if they have high credit card balances, but relatively low overall debt compared to their income.

In addition, although the borrower’s DTI doesn’t affect the decision to prepay, but borrowers with large revolving balance are less likely to prepay, which implies that borrowers use P2P loans to manage their revolving debt and high probability of default.

5.2 Practical Implications

5.2.1 Regular Credit Monitoring

Understanding borrower behaviour can help identify early warning signs of financial distress and potential defaults. P2P platforms can employ behavioural analysis techniques to monitor borrower activities, transaction patterns, and communication interactions. By detecting changes in borrower behaviour indicative of financial stress or default intentions, platforms can take proactive measures, such as offering financial counselling, restructuring loans, or implementing risk mitigation strategies to minimize losses (Zhang et al., 2023).

Furthermore, the adoption of the non-performing loan (NPL) ratio enhances credit assessment accuracy and efficiency. A high NPL ratio suggests that the lender is extending riskier loans prone to default, signalling a lack of proficiency in credit assessment and loan management (Guo et al., 2020).

For example, based on the risk-based pricing to adjust the interest rates and terms based on the borrower’s credit risk. Therefore, riskier borrowers are charged higher interest rates to compensate for the increased risk.

5.2.2 Supplement with Alternative Data

According to Lending Club website (n.d), Lending Club have conducted either income verification on 69.6% of issued loans since 2008. However, supplemented with alternative data sources to provide a more comprehensive assessment of a borrower's creditworthiness. Zopa, the UK's largest P2P lending service, tracks the applicants from another willing lender (Kim, 2021).

Due to different customer information and are subject to different regulations, which affects their loan terms and the ability to target borrowers. Hence transformed the analysis from passive information retrieval into proactive big data analytics. In other words, in traditional credit evaluation, lenders passively depend on the borrowers providing information about themselves; while in the big data era, lenders can proactively search the 360-degree online footprint of potential borrowers and assess it from multiple perspectives to gain new insights. For instance, social media, surfing, ecommerce purchases, financial transactions.

Integrating such data can enrich the predictive power of default models and help identify creditworthy borrowers who may have limited or no traditional credit history.

5.2.3 Prediction of future change borrowers

Borrower behaviour in P2P lending is offers insights into how borrowers make decisions under uncertainty, risk preferences, time preferences, and social influences. Prospect Theory, for example, suggests that borrowers may exhibit loss aversion and risk-seeking behaviour when faced with potential losses from default (Khan el al,2021).

The current credit scoring systems are divided into the forward-looking mechanism, which assesses the creditworthiness of borrowers based on their uploaded information, and the backward-looking mechanism, which assesses the creditworthiness of borrowers based on their historical repayment performance (Gao, & Zhou, 2017). These two systems are used to identify default risk. Understanding these behavioural aspects can help lenders design borrower incentives, communication strategies, and default prevention mechanisms that align with borrowers' preferences and motivations, ultimately improving credit risk management in P2P lending (Rosdini et al., 2022).

If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company. If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company (Y. Guo et al., 2021).

5.3 Theoretical Implications

This study addressed theoretical gaps in existing research and identified crucial variables for determining credit scoring. While it provided comprehensive data, there remains a lack of a solid foundation for prediction, mainly due to the continuous expansion of the P2P lending sector and the diverse characteristics of borrowers across different lenders. Although various research conducted from all this research have examined the factor that different perspective influenced credit scoring in default prediction.

Furthermore, examining the spending pattern and financial behaviour of borrowers has an interrelationship with credit scores, affecting access to credit and reducing default rates. By assessing FICO scores to determine customers' creditworthiness for loan approval, drawing from credit history data such as credit card debt, outstanding loans, payment history, and the number of revolving accounts.

The findings of this study suggest a notable positive relationship between bankcard utilization rates and revolving utilization rate, suggesting a potential avenue for further research. This contributes to the existing literature and lays a foundation for future investigations and explore study context. It underscores the importance of revolving utilization rates in credit scoring, particularly in the dynamic landscape of P2P lending, where such rates can significantly influence credit scores and subsequent lending decisions. Additionally, the study explores how borrowers' spending patterns and financial behaviours impact their credit scores, enhancing the understanding of credit scoring mechanisms.

5.4 Limitations

Generally, the larger the dataset the greater statistical power for pattern recognition enable very large-scale data analysis. However, according to this research sample size have only selected 445 sample, it still has limited sample size for further research.

Secondly, the research did not target any specific industry which in order to not bias in certain variable like specially in certain factor but also to included and with wide range for the better understanding of the target segment and the probability of default (PD). For instance, included employee title, the time of loan application, last payment date and so on to have further research on specially included the variable. This can lead to inaccurate assessments of credit risk, potentially resulting in loans being given to borrowers who are not creditworthy or denying credit to borrowers who are.

Thirdly, Kaggle was posted using the machine learning in Python that as the programming to run through the data analysis and all the process of the dataset. However, in the context of this study, it is using SPSS for statistics research due to Machine learning techniques may require more advanced programming skills and a deeper understanding of algorithms and model evaluation techniques, especially for complex analyses or large-scale datasets.

5.5 Recommendations

For future analysis, the study should not carry out on a specific variable to narrow down the area of study in order to provide more accurate result in this study. It is able to with wide range for the better understanding of the target segment and the probability of default (PD). In term of wider range of sample would have more interesting result in the diversify and without bias.

Nowadays, machine learning methods have been applied to predict important factors affecting loan repayment, such as verification of identity, assets, and education level (Xu et al., 2021). In the further study, it is recommended to use Python instead of using SPSS for statistics research as SPSS can only support simple statistics and inflexibility and adaptability to Changing Data. Machine learning algorithms are adept at capturing complex, nonlinear relationships within data. They can identify patterns and interactions among variables that might be missed by traditional statistical techniques (Xia et al., 2022).

Moreover, judged more comprehensively because certain other variables have also important effects on repayment performance outcomes such as inquiries for new credit, employee title, the time of loan application, last payment date and types of credit accounts. Hence analysis with wide range of variables for better prediction of probability of default, better distinguish between low-risk and high-risk borrowers.

5.6 Conclusion

In conclusion, this study examined credit score in default prediction for P2P lending. It can be concluded that the bankcard utilization will be affected revolving utilization rate while accessing the credit assessment and studying the borrowers' spending behaviour to predict the financial decision making of borrowers. And to ensure long-term viability of P2P lending. On the other hand, DTI will not be affected in credit assessment. Future studies should focus on bankcard utilization rates in P2P lending, leveraging machine learning and various software tools. This study serves as a reference for future researchers to validate their own findings.

REFERENCES

A, T. V. (2023). FinTech is enabler or disruptive to the Banking Industry: An analytical study. *World Journal of Advanced Research and Reviews*, *17*(1), 067–072. https://doi.org/10.30574/wjarr.2023.17.1.1472

Abbassi, P., Bräuning, F., & Schulze, N. (2021). Bargaining power and outside options in the interbank lending market. *Financial Management*, *50*(2), 553-586.

Abdelsalam, S., Salim, N., Alias, R. A., & Husain, O. (2020). Understanding Online impulse buying Behavior in Social Commerce: A Systematic Literature review. *IEEE Access*, *8*, 89041–89058. https://doi.org/10.1109/access.2020.2993671

About income verification | LendingClub. (n.d.). https://www.lendingclub.com/public/income-verification.action

Agnihotri, A., & Bhattacharya, S. (2023). Chime Neobank: a new business model in the banking sector. In SAGE Publications: SAGE Business Cases Originals eBooks. https://doi.org/10.4135/9781529610482

Baldo, D. R., Regio, M. S., & Manssour, I. H. (2023). Visual analytics for monitoring credit scoring models. *Information Visualization*, *22*(4), 340–357. https://doi.org/10.1177/14738716231180803

Berndt, A. E. (2020). Sampling methods. Journal of Human Lactation, 36(2), 224-226.

Çetin, V., & YILDIZ, O. (2022). A comprehensive review on data preprocessing techniques in data analysis. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, *28*(2), 299-312.

Chang, A., Yang, L., Tsaih, R., & Lin, S. (2022). Machine learning and artificial neural networks to construct P2P lending credit-scoring model: A case using

Lending Club data. *Quantitative Finance and Economics*, *6*(2), 303–325. https://doi.org/10.3934/qfe.2022013

Creswell, J. W., & Creswell, J. D. (2022). Research design: Qualitative, Quantitative, and Mixed Methods Approaches. SAGE Publications.

Croux, C., Jagtiani, J., Korivi, T., & Vulanović, M. (2020). Important factors determining Fintech loan default: Evidence from a lendingclub consumer platform. Journal of Economic Behavior and Organization, 173, 270–296. https://doi.org/10.1016/j.jebo.2020.03.016

Dharmastuti, C. F., & Laurentxius, J. (2021). Factors and Benefits that Affect Lender's Interest in Giving Loans in Peer to Peer (P2P) Lending Platform. *Binus Business Review : Management, Accounting and Hospitality Management*, *12*(2), 121–130. https://doi.org/10.21512/bbr.v12i2.6359

Doroghazi, R. M. (2020). FICO scores. American Journal of Cardiology, 130, 157–158. https://doi.org/10.1016/j.amjcard.2020.06.001

Fast-Moving FinTech poses challenge for regulators. (2022, April 13). IMF. https://www.imf.org/en/Blogs/Articles/2022/04/13/blog041322-sm2022-gfsr-ch3

Field, Andy. *Discovering statistics using IBM SPSS Statistics*. SAGE Publications Limited, 2024.

Gantulga, U., & Dashrentsen, D. (2023). FACTORS INFLUENCE IMPULSIVE BUYING BEHAVIOR. *Zeszyty Naukowe Wyższej Szkoły Humanitas*, *24*(1), 9–25. https://doi.org/10.5604/01.3001.0053.4041

Gao, B. M., Yen, J., & Liu, M. T. (2021). Determinants of defaults on P2P lending platforms

Gao, W., Ju, M., & Yang, T. (2023). Severe Weather and Peer-to-Peer Farmers' Loan Default Predictions: Evidence from Machine Learning Analysis. Finance Research Letters, 58, 104287. https://doi.org/10.1016/j.frl.2023.104287

Gao, Y., Sun, J., & Zhou, Q. (2017). Forward looking vs backward looking. *China Finance Review International (Print)*, *7*(2), 228–248. https://doi.org/10.1108/cfri-07-2016-0089

Giglio, F. (2021). Fintech: A literature review. *European Research Studies Journal*, *24*(2B), 600-627.

Gu, Y., Guo, L., Ma, C., Wang, H., & Wei, X. (2022, November). Peer to Peer Lending Risk Analysis: Predictions from Lending Club. In *2022 3rd International Conference on E-commerce and Internet Technology (ECIT 2022)* (pp. 750-759). Atlantis Press.

Guo, H., Peng, K., Xu, X., Tao, S., & Wu, Z. (2020). The Prediction Analysis of Peer-to-Peer lending platforms default risk based on comparative models. *Scientific Programming*, *2020*, 1–10. https://doi.org/10.1155/2020/8816419

Guo, Y., Jiang, S., Zhou, W., Luo, C., & Xiong, H. (2021). A predictive indicator using lender composition for loan evaluation in P2P lending. *Financial Innovation (Heidelberg)*, *7*(1). https://doi.org/10.1186/s40854-021-00261-1

Gupta, S. (2019). P2P Lending: bridging India's credit gap. *IIBM's Journal of Management Research*, 173–180. https://doi.org/10.33771/iibm.v4i1-2.1079

Hayes, A. (2023, February 18). FICO Score. Investopedia. https://www.investopedia.com/terms/f/ficoscore.asp

Howitt, Dennis, and Duncan Cramer. *Understanding statistics in psychology with SPSS*. Pearson, 2020.

Hughes, J. P., Jagtiani, J., & Moon, C. (2022). Consumer lending efficiency: commercial banks versus a fintech lender. *Financial Innovation (Heidelberg)*, *8*(1). https://doi.org/10.1186/s40854-021-00326-1

in China. International Review of Economics & Finance, 72, 334–348. https://doi.org/10.1016/j.iref.2020.11.012

Ismawati, I. Y., & Faturohman, T. (2023). Credit risk scoring model for consumer financing: Logistic Regression Method. In *International symposia in economic theory and econometrics* (pp. 167–189). https://doi.org/10.1108/s1571-038620230000031023

Kaggle. https://www.kaggle.com/code/mariiagusarova/complete-data-science-fintech-project-part-i#%F0%9F%93%8C-Problem-Statement

Kang, H. (2021). Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*, *18*, 17. https://doi.org/10.3352/jeehp.2021.18.17

Khan, M. T. I., Yee, G. H., & Goh, G. G. G. (2021). Antecedents of intention to use online Peer-to-Peer platform in Malaysia. *Vision*, *27*(5), 680–694. https://doi.org/10.1177/09722629211039051

Kim, D. (2021). Empirical evidence of faulty credit scoring and business failure in P2P lending. *Global Business and Finance Review*, *26*(2), 67–82. https://doi.org/10.17549/gbfr.2021.26.2.67

Lakens, D. (2022). Sample size justification. Collabra: Psychology, 8(1), 33267.

Leckė, G., Milašiūtė, G., & Šeškevičiūtė, A. (2022). Assessment of the factors affecting the intention to invest in a P2P lending platform. Socialiniai Tyrimai, 45(2), 8–25. https://doi.org/10.15388/soctyr.2022.45.2.1

Lee, J. W., Lee, W. K., & Sohn, S. Y. (2021). Graph convolutional network-based credit default prediction utilizing three types of virtual distances among borrowers. *Expert Systems With Applications*, *168*, 114411. https://doi.org/10.1016/j.eswa.2020.114411

Lim, W. M., Phang, C. S. C., & Lim, A. L. (2020). The effects of possession- and social inclusion-defined materialism on consumer behavior toward economical versus luxury product categories, goods versus services product types, and individual versus group marketplace scenarios. *Journal of Retailing and Consumer Services*, *56*, 102158. https://doi.org/10.1016/j.jretconser.2020.102158

Lopez, I., Mahdzan, N. S., & Rahman, M. (2022). Identifying salient beliefs underlying intention to save regularly: an elicitation study of Malaysian Gen Y. *Qualitative research in financial markets*, *15*(1), 98-118.

Lyócsa, Š., Vašaničová, P., Misheva, B. H., & Vateha, M. D. (2022). Default or profit scoring credit systems? Evidence from European and US peer-to-peer lending markets. *Financial Innovation*, *8*(1). https://doi.org/10.1186/s40854-022-00338-5

Ma, Z., Hou, W., & Zhang, D. (2021). A credit risk assessment model of borrowers in P2P lending based on BP neural network. *PloS One*, *16*(8), e0255216. https://doi.org/10.1371/journal.pone.0255216

Madeira, C. (2023). Adverse selection, loan access and default behavior in the Chilean consumer debt market. *Financial Innovation*, *9*(1). https://doi.org/10.1186/s40854-023-00458-6

Mahdzan, N. S., Zainudin, R., & Shaari, M. S. (2022). The influence of religious belief and psychological factors on borrowing behaviour among Malaysian public sector employees. Asia-pacific Journal of Business Administration, 15(3), 361–385. https://doi.org/10.1108/apjba-10-2020-0362

Mariiagusarova. (2022, September 1). Complete Data Science Fintech Project: Part I.

Money market executive's perception towards peer-to-peer (p2p) lending. (2021). *The Journal of Contemporary Issues in Business and Government*, *26*(02). https://doi.org/10.47750/cibg.2020.26.02.078

Monis, E., & Pai, R. (2023). Neo Banks: A Paradigm shift in banking. *Zenodo (CERN European Organization for Nuclear Research)*. https://doi.org/10.5281/zenodo.8011125

Musari, K. (2022). P2P lending and philanthropy platform. *Routledge eBooks*, 185. https://doi.org/10.4324/9781003224532-11

Nagashree, H. G. (2023). Loan default prediction using machine learning techniques. *Indian Scientific Journal of Research in Engineering and Management*, *07*(07). https://doi.org/10.55041/ijsrem24519

Nigmonov, A., Shams, S., & Alam, K. (2022). Macroeconomic determinants of loan defaults: Evidence from the U.S. peer-to-peer lending market. *Research in International Business and Finance*, *59*, 101516. https://doi.org/10.1016/j.ribaf.2021.101516

Octavini, N. A., & Puspitasari, A. D. (2023). Fintech dan Bank : P2P Lending Vs Bank Lending. *Ekombis Sains: Jurnal Ekonomi, Keuangan Dan Bisnis*, *8*(1), 49–58. https://doi.org/10.24967/ekombis.v8i1.2029

Pandey, P., & Pandey, M. M. (2021). Research methodology tools and techniques. Bridge Center.

Park, H. M. (2015). Univariate analysis and normality test using SAS, Stata, and SPSS.

Reepu, R. (2023). Banking of the upcoming age. In *Advances in marketing, customer relationship management, and e-services book series* (pp. 38–50). https://doi.org/10.4018/978-1-6684-8312-1.ch004

Renbarger, R., Sulak, T., & Kaul, C. R. (2019). Finding, accessing, and using secondary data for research on gifted education and advanced academics. Journal of Advanced Academics. https://doi.org/10.1177/1932202x19864117

Rong, Y., Liu, S., Yan, S., Huang, W., & Chen, Y. (2023). Proposing a new loan recommendation framework for loan allocation strategies in online P2P lending. *Industrial Management and Data Systems*, *123*(3), 910–930. https://doi.org/10.1108/imds-07-2022-0399

Rosdini, D., Wahyuni, E. T., & Sari, P. Y. (2022). The bane of P2P lending: credit scoring governance on the ASEAN fintech triumvirate. *Journal of Science & Technology Policy Management*. https://doi.org/10.1108/jstpm-02-2022-0043

Rosenberg, E. (2023, November 19). LendingClub Personal Loan review (2023): pros, cons, approval odds · TIME Stamped: Personal Finance made easy. *TIME Stamped: Personal Finance Made Easy*. https://time.com/personal-finance/article/lendingclub-personal-loans-review/

Serrano-Cinca, C., & Nieto, B. G. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. *Decision Support Systems*, *89*, 113–122. https://doi.org/10.1016/j.dss.2016.06.014

Siedlecki, S. L. (2020). Understanding descriptive research designs and methods. *Clinical Nurse Specialist*, *34*(1), 8-12.

Siering, M. (2023). Peer-to-Peer (P2P) lending Risk Management: Assessing credit risk on social lending platforms using textual factors. *ACM Transactions on*

Management Information Systems, *14*(3), 1–19. https://doi.org/10.1145/3589003

Sileyew, K. J. (2019). Research design and methodology. Cyberspace, 1-12.

Smiti, A. (2020). A critical overview of outlier detection methods. *Computer Science Review*, *38*, 100306.

Statista. (2023). *Most valued independent neobanks worldwide 2021*. https://www.statista.com/statistics/1239713/most-valued-independent-neobanks-worldwide/#:~:text=The%20most%20valued%20independent%20neobank,with%2010%20billion%20U.S.%20dollars

Sun, X. (2020). Prediction of the Borrowers' Payback to the Loan with Lending Club Data. 2020 International Conference on Modern Education and Information Management (ICMEIM). https://doi.org/10.1109/icmeim51375.2020.00092

Sureshchandar, G. (2021). Quality 4.0 – a measurement model using the confirmatory factor analysis (CFA) approach. *International Journal of Quality and Reliability Management/International Journal of Quality & Reliability Management*, *40*(1), 280–303. https://doi.org/10.1108/ijqrm-06-2021-0172

Suryono, R. R., Purwandari, B., & Budi, I. (2019). Peer to Peer (P2P) lending Problems and Potential Solutions: A Systematic Literature review. *Procedia Computer Science*, *161*, 204–214. https://doi.org/10.1016/j.procs.2019.11.116

Tang, X., Zhu, J., He, M., & Feng, C. (2023). How can we learn from a borrower's online behaviors? The signal effect of a borrower's platform involvement on its credit risk. *Electronic Commerce Research and Applications*, *59*, 101272. https://doi.org/10.1016/j.elerap.2023.101272

Thomas, F. B. (2022). The role of purposive sampling technique as a tool for informal choices in a social Sciences in research methods. *Just Agriculture*, *2*(5), 1-8.

Trivedi, S. K. (2020). A study on credit scoring modeling with different feature selection and machine learning approaches. *Technology in Society*, *63*, 101413. https://doi.org/10.1016/j.techsoc.2020.101413

Uwais, A. M. (2022, March 10). Loan default prediction using spark machine learning algorithms. University of Portsmouth. https://researchportal.port.ac.uk/en/publications/loan-default-prediction-using-spark-machine-learning-algorithms

Walden, S. (2021, June 24). What Is A Neobank? Forbes Advisor. https://www.forbes.com/advisor/banking/what-is-a-neobank/

Wang, C., Zhang, W., Zhao, X., & Wang, J. (2019). Soft information in online peer-to-peer lending: Evidence from a leading platform in China. Electronic Commerce Research and Applications, 36, 100873. https://doi.org/10.1016/j.elerap.2019.100873

Watkins, M. W. (2021). A Step-by-Step Guide to Exploratory Factor Analysis with SPSS. In *Routledge eBooks*. https://doi.org/10.4324/9781003149347

Windasari, N. A., Kusumawati, N., Larasati, N., & Amelia, R. P. (2022). Digital-only banking experience: Insights from gen Y and gen Z. *Journal of Innovation & Knowledge*, *7*(2), 100170.

Xia, H., Wang, P., Wang, T., Zhang, Z., Weng, J., & Jasimuddin, S. M. (2022). Peer-to-peer lending platform risk analysis: an early warning model based on multi-dimensional information. ˜ *the œ Journal of Risk Finance*, *23*(3), 303–323. https://doi.org/10.1108/jrf-06-2021-0102

Xu, J., Lu, Z., & Xie, Y. (2021). Loan default prediction of Chinese P2P market: a machine learning methodology. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-98361-6

Zhang, Z. X., Gu, Y., Jiang, L., Yu, W., & Dai, J. (2023). Internet of Things and Blockchain-Based Smart Contracts: Enabling continuous risk monitoring and assessment in Peer-to-Peer lending. *Journal of Emerging Technologies in Accounting*, *20*(2), 181–194. https://doi.org/10.2308/jeta-2022-003

Zhu, X., Chu, Q., Song, X., Hu, P., & Peng, L. (2023). Explainable prediction of loan default based on machine learning models. *Data Science and Management*, *6*(3), 123–133. https://doi.org/10.1016/j.dsm.2023.04.003

Zikmund, W. G., Babin, B. J., Griffin, M., & Carr, J. C. (2010). Business research methods. Cengage Learning.

Appendix

Appendix 4.1: Summary of Hypothesis Development

Hypothesis	Outcome	Determination
H1: Loan Amount has not significance influence on Interest Rate during P2P Lending in Lending Club	Significant Value: 0.217, $p< 0.05$	Rejected
H2: Loan Amount has significance influence on Revolving Utilization Rate during P2P Lending in Lending Club	CFA Extraction Value: 0.685, >0.5 Significant Value: 0.029, $p< 0.05$	**Accepted**
H3: Loan Amount has not significance influence on Total open to buy on Revolving Bankcard during P2P Lending in Lending Club	Significant Value: 0.277, $p< 0.05$	Rejected
H4: Loan Amount has not significance influence on Bankcard Utilization Rate during P2P Lending in Lending Club	Significant Value: 0.79, $p> 0.05$	Rejected
H5: Loan Amount has not significance influence on Number of open Revolving Accounts during P2P Lending in Lending Club	Significant Value: 0.052, $p< 0.05$	Rejected
H6: Loan Amount has significance influence on DTI during P2P Lending in Lending Club	CFA Extraction Value: 0.685, >0.5 Significant Value: 0, $p< 0.05$	**Accepted**

H7: Interest Rate has significance influence on Revolving Utilization Rate during P2P Lending in Lending Club	CFA Extraction Value: 0.600, >0.5 Significant Value: 0.033, p< 0.05	**Accepted**
H8: Interest Rate has significance influence on Total open to buy on Revolving Bankcard during P2P Lending in Lending Club	CFA Extraction Value: 0.600, >0.5 Significant Value: 0, p< 0.05	**Accepted**
H9: Interest Rate has not significance influence on Bankcard Utilization Rate during P2P Lending in Lending Club	Significant Value: 0.052, p< 0.05	Rejected
H10: Interest Rate has significance influence on Number of open Revolving Accounts during P2P Lending in Lending Club	CFA Extraction Value: 0.600, >0.5 Significant Value: 0.026, p< 0.05	**Accepted**
H11: Interest Rate has not significance influence on DTI during P2P Lending in Lending Club	Significant Value: 0.171, p< 0.05	Rejected
H12: Revolving Utilization Rate has significance influence on Total open to buy on Revolving Bankcard during P2P Lending in Lending Club	CFA Extraction Value: 0.829, >0.5 Significant Value: 0, p< 0.05	**Accepted**
H13: Revolving Utilization Rate has significance influence on Bankcard Utilization Rate during P2P Lending in Lending Club	CFA Extraction Value: 0.829, >0.5 Significant Value: 0, p< 0.05	**Accepted**
H14: Revolving Utilization Rate has significance influence on Number of open Revolving Accounts during P2P Lending in Lending Club	CFA Extraction Value: 0.829, >0.5 Significant Value: 0 p< 0.05	**Accepted**

H15: Revolving Utilization Rate has not significance influence on DTI during P2P Lending in Lending Club	CFA Significant Value: 0.346, p< 0.05	Rejected
H16: Total open to buy on Revolving Bankcard has significance influence on Bankcard Utilization Rate during P2P Lending in Lending Club	CFA Extraction Value: 0.691, >0.5 Significant Value: 0, p< 0.05	**Accepted**
H17: Total open to buy on Revolving Bankcard has significance influence on Number of open Revolving Accounts during P2P Lending in Lending Club	CFA Extraction Value: 0.691, >0.5 Significant Value: 0, p< 0.05	**Accepted**
H18: Total open to buy on Revolving Bankcard has significance influence on DTI during P2P Lending in Lending Club	CFA Extraction Value: 0.691, >0.5 Significant Value: 0.011, p< 0.05	**Accepted**
H19: Bankcard Utilization Rate has not significance influence on Number of open Revolving Accounts during P2P Lending in Lending Club	CFA Significant Value: 0.352, p< 0.05	Rejected
H20: Bankcard Utilization Rate has not significance influence on DTI during P2P Lending in Lending Club	CFA Significant Value: 0.479, p< 0.05	Rejected
H21: Number of open Revolving Accounts has not significance influence on DTI during P2P Lending in Lending Club	CFA Significant Value: 0.755, p< 0.05	Rejected

Source: Developed for the research