Loans Credit Risk Analysis

Rafael Cabral dos Santos

1 Summary

Credit risk management is a cornerstone of financial stability, playing a vital role in mitigating losses from borrowers' inability to fulfill their obligations. The 2008 housing crisis highlighted the catastrophic impact of inadequate risk assessment, where the widespread issuance of sub- prime loans without proper evaluation of borrowers' creditworthiness led to a global economic downturn. Understanding and assessing key metrics like Probability of Default (PD), Expo- sure at Default (EAD), and Loss Given Default (LGD) is essential for financial institutions to quantify potential risks, optimize capital allocation, and comply with regulatory standards. PD estimates the likelihood of a borrower defaulting, EAD measures the exposure at the time of default, and LGD reflects the percentage of loss after recovery efforts. Together, these met- rics form the foundation of credit risk modeling, enabling lenders to make informed decisions, maintain profitability, and safeguard the broader financial system.

2 Introduction

2.1 Data Overview

The dataset spans loan issuance and performance between 2007 and 2014, focusing on borrower attributes and loan details like requested amounts, repayment terms, interest rates, and monthly installments. It reflects the broader effects of the mortgage crisis on non-mortgage lending, offering insights into how the financial turmoil influenced consumer borrowing trends.

In this analysis, the term *"Prime loans"* refers to loans that are either currently being paid or have been fully paid off. On the other hand, *"Risky"* loans represent those that have defaulted or are significantly overdue. These definitions are used to categorize loans based on their payment status and risk level in the context of this analysis.

2.2 Economic Context

In the years preceding the Great Financial Crisis, financial institutions issued a significant volume of subprime mortgages, targeting borrowers with limited creditworthiness. These loans were securitized into Mortgage-Backed Securities (MBS) and sold to investors under the guise of high-yield, low-risk assets, thanks to the flawed risk assessments by credit rating agencies. This practice concealed the underlying systemic risk and significantly contributed to the housing bubble. Key contributing factors included inadequate down payments, high loan-to-value (LTV) and debt-to-income (DTI) ratios, and the non-recourse nature of the loans. Moreover, the non-recourse structure allowed borrowers to default with limited repercussions, as the property served as the sole collateral. When housing prices dropped below the mortgage balance—commonly referred to as "negative equity"—borrowers had an economic incentive to default and abandon their properties.

3 Historical Analysis

3.1 Home Ownership Distribution

The chart below illustrates the home ownership distribution among loan borrowers, highlighting the relationship between home ownership status and loan acquisition.

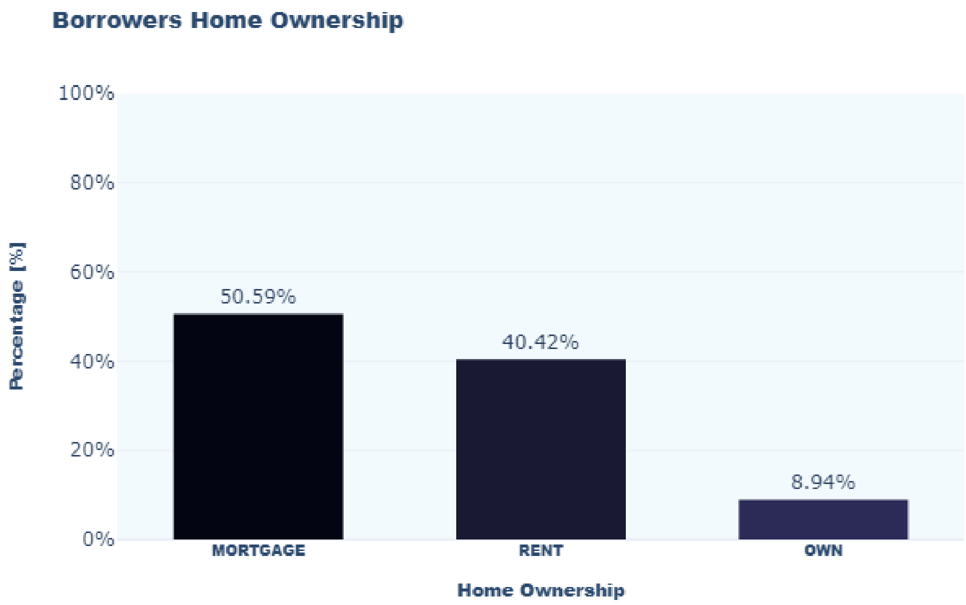

Figure 1: Home ownership distribution among loan borrowers

- **Mortgage Holders**: Representing 50% of borrowers, this group constitutes the largest segment. These individuals already have significant ongoing financing commitments through their mortgages. This increases their debt-equity ratio, making them more leveraged and elevating their default risk. Such dynamics were a critical factor during the Great Financial Crisis, as highly leveraged borrowers faced significant challenges in meeting financial obligations during economic downturns.

3.2 Loan Grade

The chart below presents the distribution of loan grades, providing insight into the creditworthiness of borrowers and the associated risk profile.

- **Subprime Loans (C, D, E, F, G)**: Comprising over 50% of all loans, this category highlights the prevalence of high-risk borrowers. Subprime loans are typically issued to individuals with lower credit scores or limited credit history, resulting in higher interest rates and default risks.
- **Prime Loans (A)**: Accounting for only 16% of loans, this group reflects the most creditworthy borrowers with strong financial stability and lower default risk. Their limited representation indicates a skewed risk profile in the loan portfolio.

The relationship between loan grade and default rates, with higher-grade loans exhibiting significantly lower default rates compared to lower-grade loans can be seen below.

3.3 Risk Over Time

The chart bellow illustrates the proportion of **Prime Loans** (green) and **Risky Loans** (red) from 2007 to 2016. The shaded red region highlights the housing crisis period (2007–2009), marking a significant turning point for borrower classifications.

- The impact of the Great Recession is evident in the 20% increase in risky loans during the crisis period, which led to significant default challenges in the years that followed.

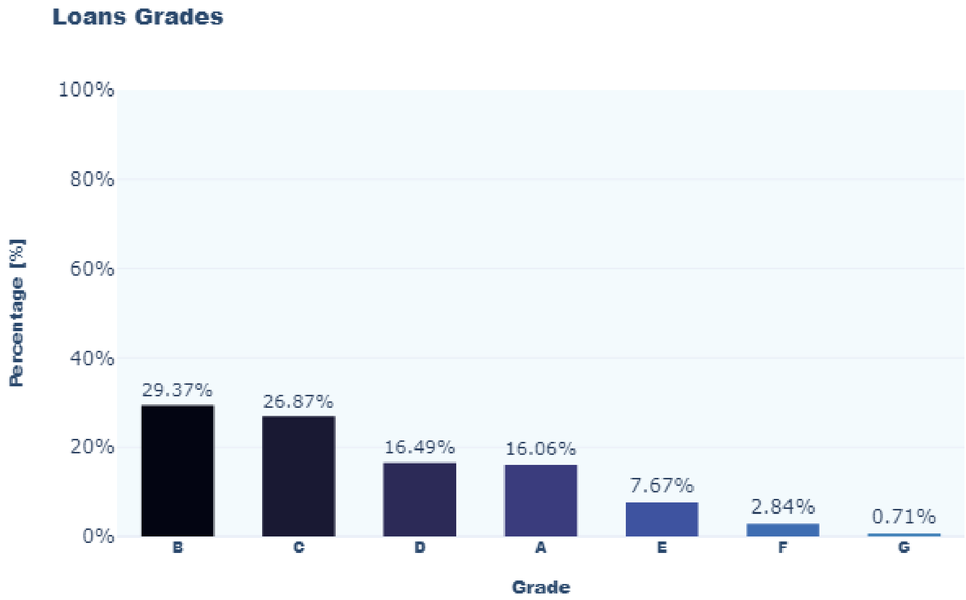

Figure 2: Loan grade distribution among borrowers

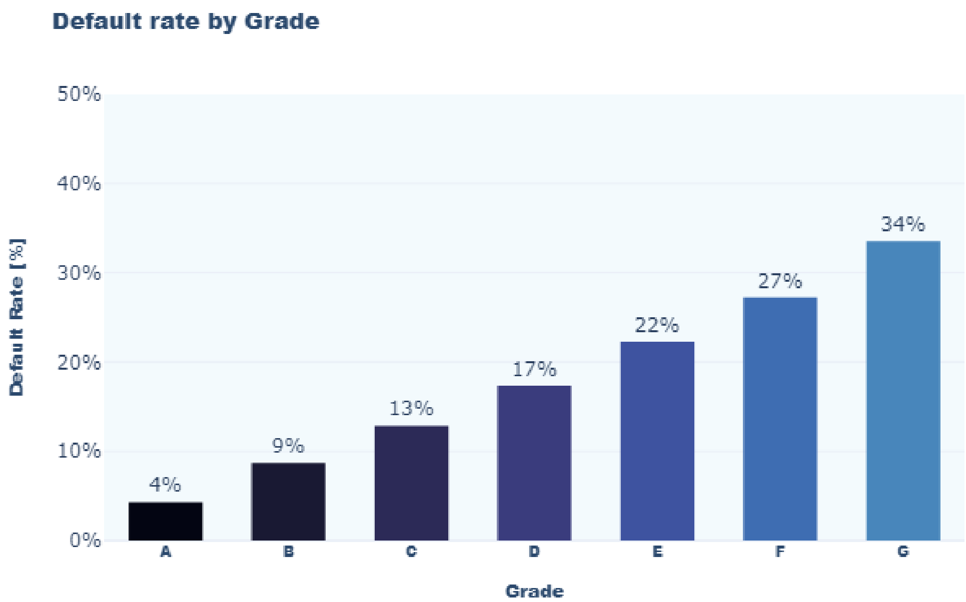

Figure 3: Percentage of Loans Defaulted by Grade

- The passage of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 introduced stringent regulations for loans and mortgages, contributing to a decline in the prevalence of risky loans. These developments highlight a notable shift in lending standards over the decade, with a clear focus on higher-quality borrowers in the aftermath of the financial crisis.
- By 2016, 91% of the company's loan portfolio was classified as Prime, indicating an overwhelming concentration of loans with minimal default risk.

3.4 Interest Rate Drift

- **2007-2009 (Crisis Period):** Both the prime and risky loan curves exhibit low standard deviations, suggesting that prior to the crisis, interest rate calculations may have been less sophisticated, possibly based on predefined assumptions rather than detailed risk assessments.
- **Post-Crisis Period:** Following the crisis, the dispersion of interest rates increased, likely due to the introduction

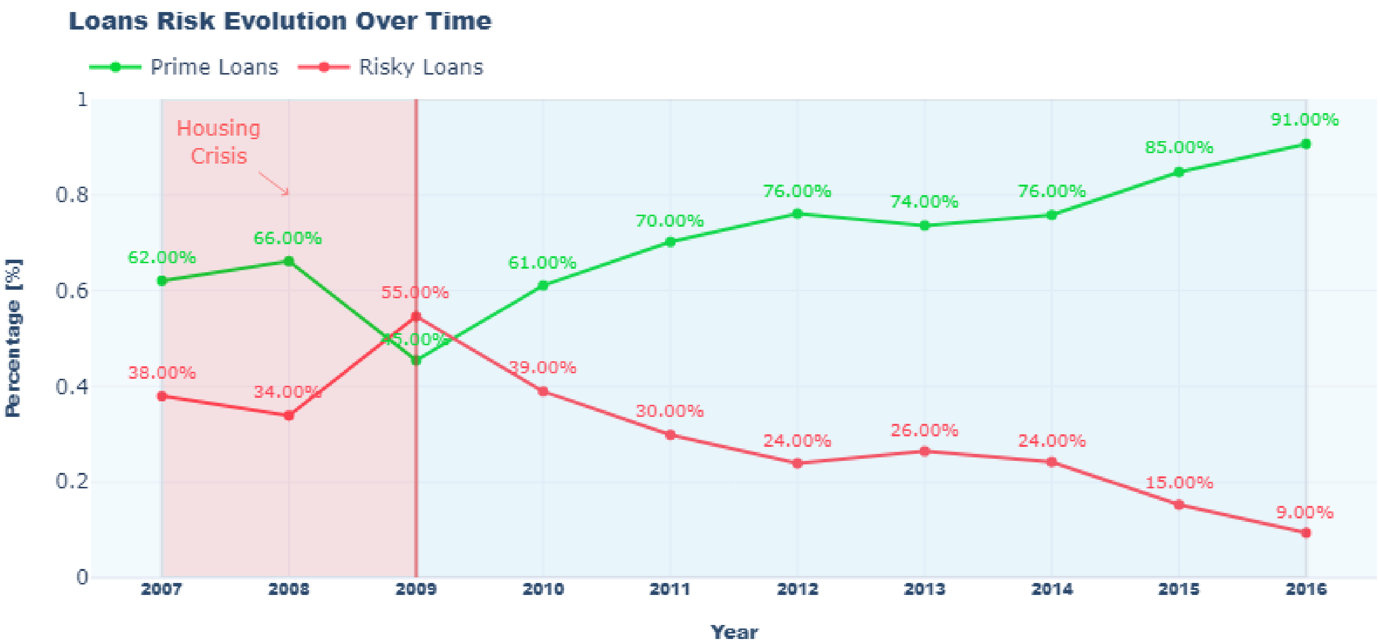

Figure 4: Loans Risk Evolution from 2007 to 2016

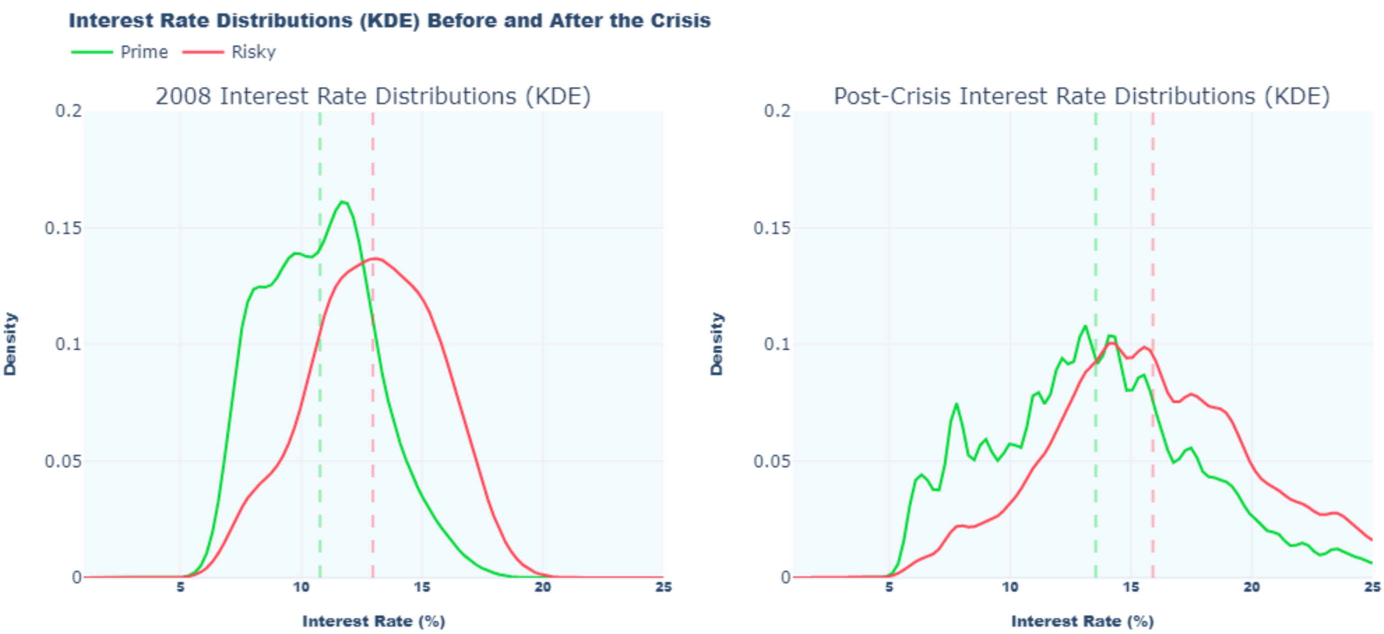

Figure 5: Interest Rate Distributions (KDE) Before and After the Crisis

of more refined and specific risk evaluation practices.

- The average interest rates rose by approximately 5% for both prime and risky loans in the post-crisis period. This increase can be attributed to factors such as inflation, leading to higher treasury interest rates, as well as improvements in risk assessment methodologies.

As illustrated by the trend in the chart below, higher interest rates are associated with riskier loans, which in turn result in higher default rates.

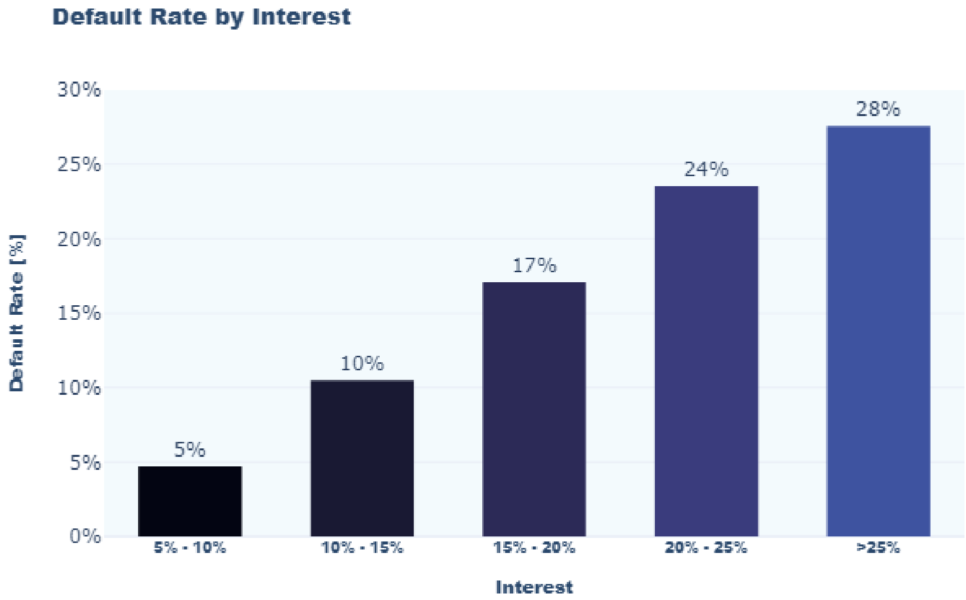

Figure 6: Percentage of Loans Defaulted by Interest Rate

4 Default Sensitivity

This analysis examines the sensitivity of default rates to key parameters within the loan dataset. By exploring how loan and borrower characteristics influence default behavior, we aim to identify critical drivers of loan performance and assess the potential risks associated with different lending conditions.

4.1 DTI

The Debt-to-Income (DTI) ratio is a key financial metric that compares an individual's monthly debt payments to their monthly income. It is commonly used by lenders to assess a borrower's ability to repay a loan. A higher DTI ratio indicates that a larger portion of income is allocated to debt obligations, which can signal financial strain.

The chart below demonstrates a direct correlation between higher DTI ratios and increasing default rates. This suggests that individuals with larger amounts of debt relative to their income are more likely to face difficulties in repaying their loans, thereby significantly increasing the risk of loan defaults.

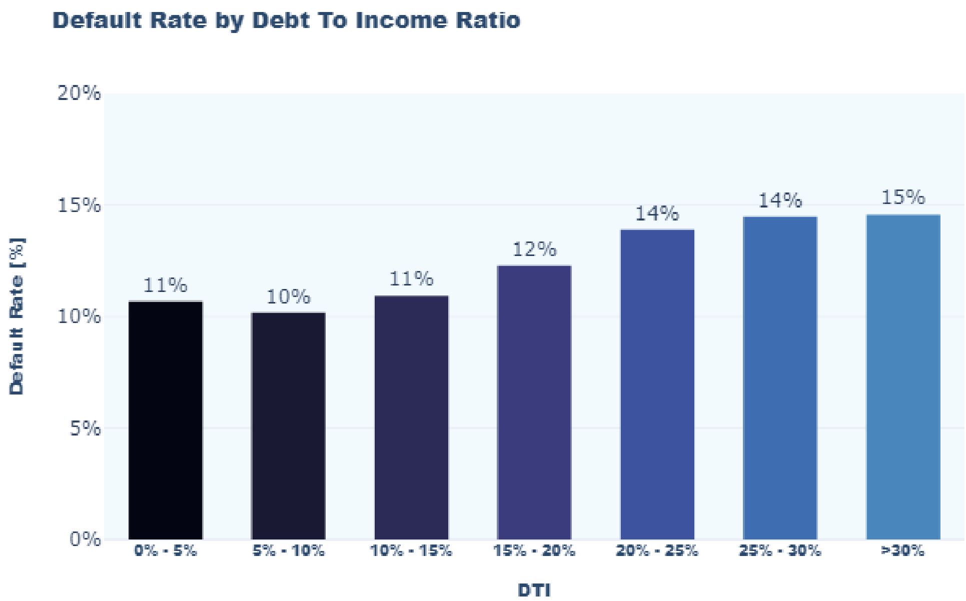

Figure 7: Percentage of Loans Defaulted by DTI Ratio

4.2 Annual Income

The chart below illustrates that as income increases, default rates tend to decrease slightly. This suggests that individuals with higher income are more likely to meet their loan obligations, as they are less financially strained compared to those with lower income.

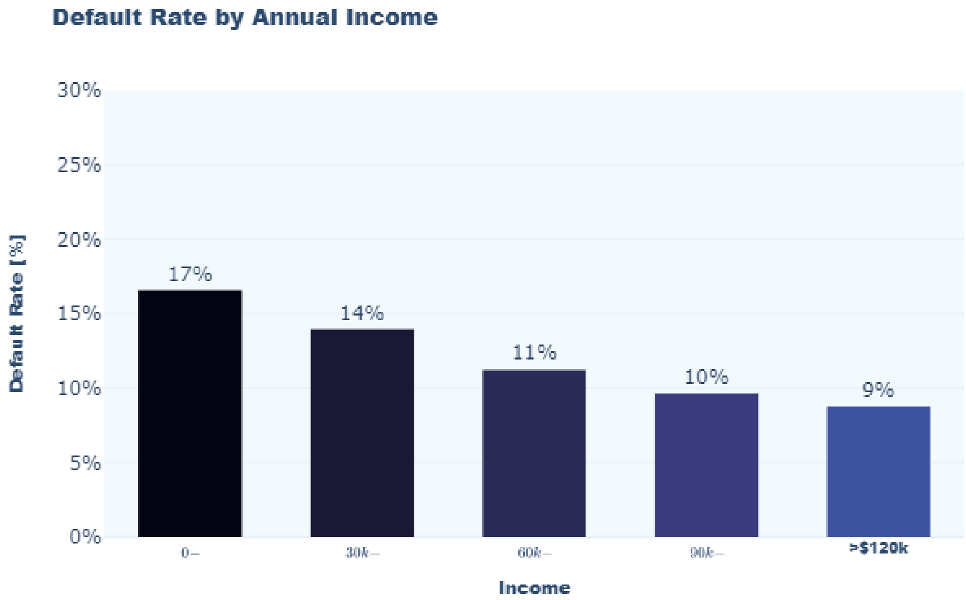

Figure 8: Percentage of Loans Defaulted by Income

5 Current Portfolio Analysis

This section provides a brief analysis of the current loan portfolio, excluding loans that have already defaulted or been fully paid. The metrics are summarized in the following table.

Table 1: Current Portfolio Metrics Summary

Metric	Value
Exposure at Default (EAD)	$204.76M
Probability of Default (PD)	21.64%
Recovery Rate (RR)	50.98%
Loss Given Default (LGD, $)	$100.37M
Loss Given Default (LGD, %)	49.02%
Expected Loss (EL)	$21.72M
Weighted Average Interest Rate	14.18%
Portfolio Yield	20.27%
Median Debt-to-Income Ratio (DTI)	17.7

1. **High Probability of Default (PD):** The portfolio's Probability of Default (**21.64%**) indicates that a significant portion of the portfolio is at risk. This highlights the need for active risk management and mitigation strategies.

2. **Moderate Recovery Rate (RR):** With a Recovery Rate of **50.98%**, the portfolio is recovering approximately half of the outstanding balances from defaulted loans.

3. **Significant Loss Given Default (LGD):** The Loss Given Default (**49.02%**) represents substantial potential losses after recovery from defaults, amounting to **$100.37MM**. This emphasizes the importance of reducing exposure to high-risk segments and improving recovery processes.

4. **Expected Loss (EL):** The Expected Loss is calculated at **$21.72MM**, which represents a significant proportion of the portfolio's total exposure. This highlights the material risk associated with the portfolio, necessitating strategies such as diversification or increased provisioning to manage these potential losses.

5.1 Recommendations

- Reduce the Probability of Default by tightening credit policies or improving borrower assessment criteria.
- Enhance the Recovery Rate through more effective collection mechanisms and improved collateral valuation processes.
- Diversify the portfolio to lower the concentration of risk in high-PD segments.

6 Problem Addressing in Future Work

As part of future work in my portfolio, I plan to expand this analysis by developing machine learning models for the key components of credit risk: Probability of Default (PD), Exposure at Default (EAD), and Loss Given Default (LGD). These models will leverage advanced techniques to predict credit risk more accurately, offering deeper insights into borrower behavior and potential losses. The full study and future models will be accessible through my GitHub portfolio at `https://github.com/rafaelCabralDS/quant-portfolio`, where I will continue to refine and expand upon the models to better address the evolving challenges in credit risk management.