Loan Default Analysis for Risk Assessment

Understanding Key Indicators of Loan Default

Prepared by Donthula Vikas and Gunupudi Lakshman

Problem Statement

Goal: To identify the primary factors contributing to loan default to strengthen portfolio risk assessment.

Business Need: The company requires insights into variables that strongly indicate potential loan default, enabling more effective risk management.

Data Analysis Approach

1. **Data Overview:** Summary of the dataset and key columns (e.g., loan amount, interest rate, employment length, loan status).

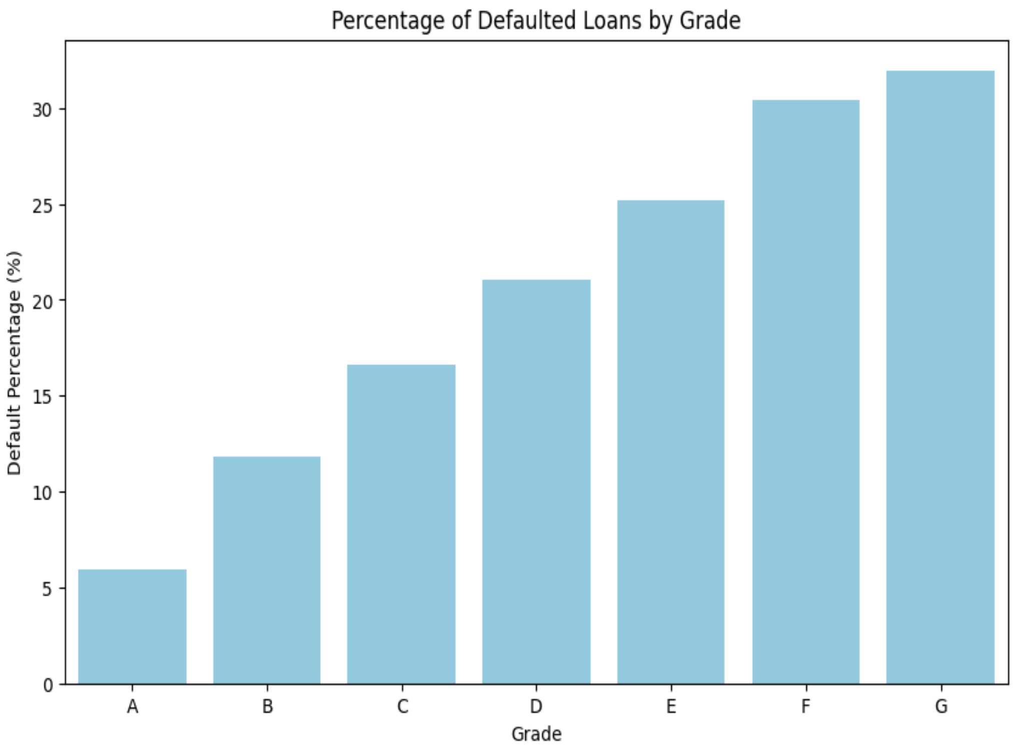
2. **Analytical Steps**:

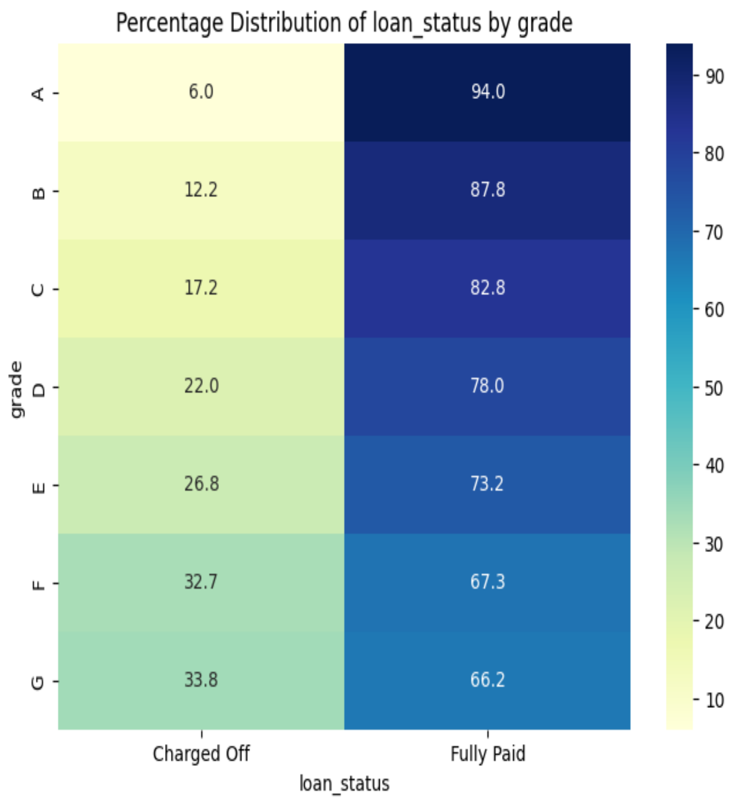
    Univariate Analysis: Examined distribution and outliers, particularly focusing on skewed distributions (e.g., income,loan amount etc..)

    Bivariate Analysis: Explored relationships between loan default and variables like loan grade, income, and Interest Rate

    Multivariate Analysis: Identified interdependencies using correlation techniques.

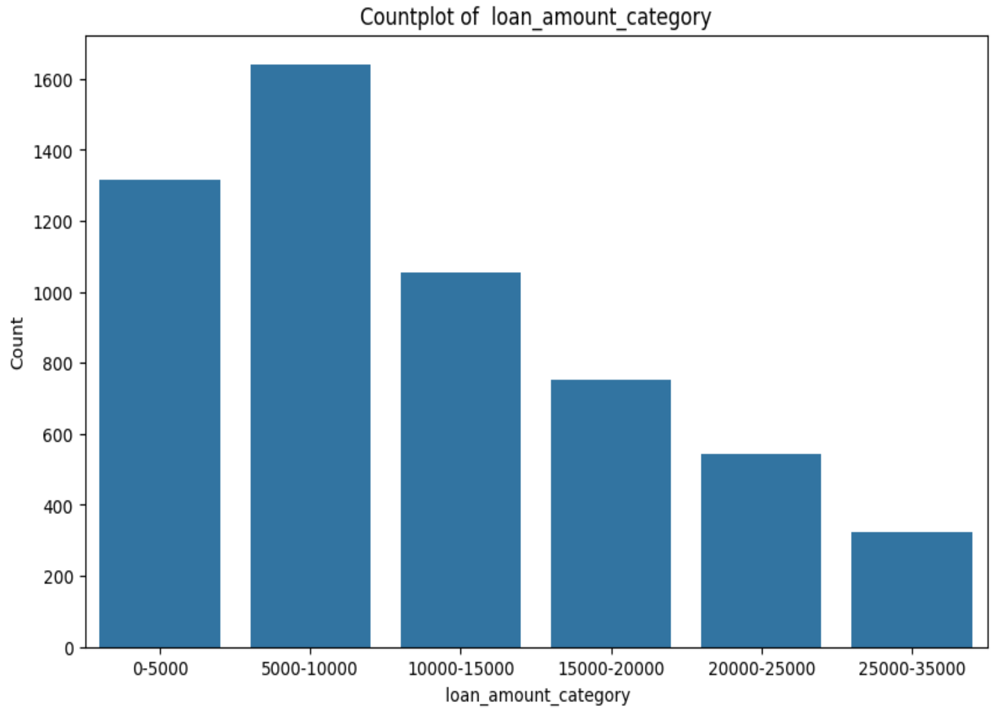
Analysis: Loan Grade and Default Rate

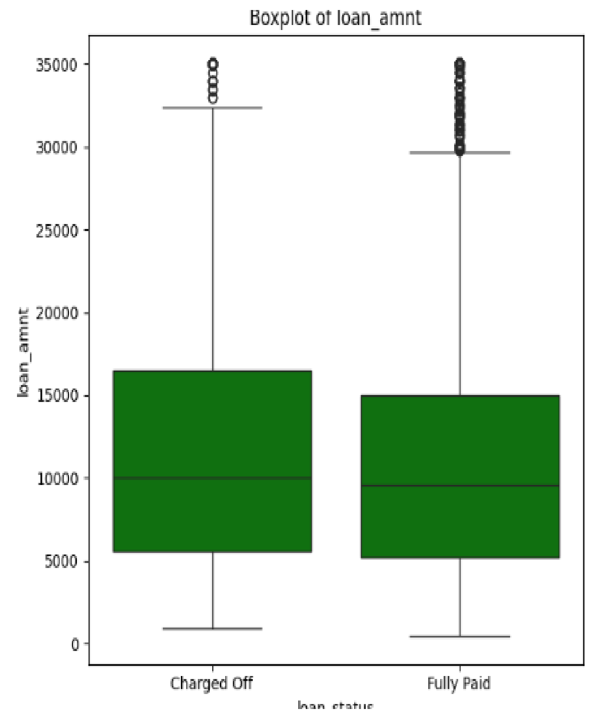




- Finding: Default rates vary significantly by loan grade.
- Insight: Lower grades (e.g., D,E,F,G) show higher default percentages, indicating correlation between grade riskiness and default likelihood.

Loan Amount and Default Rate





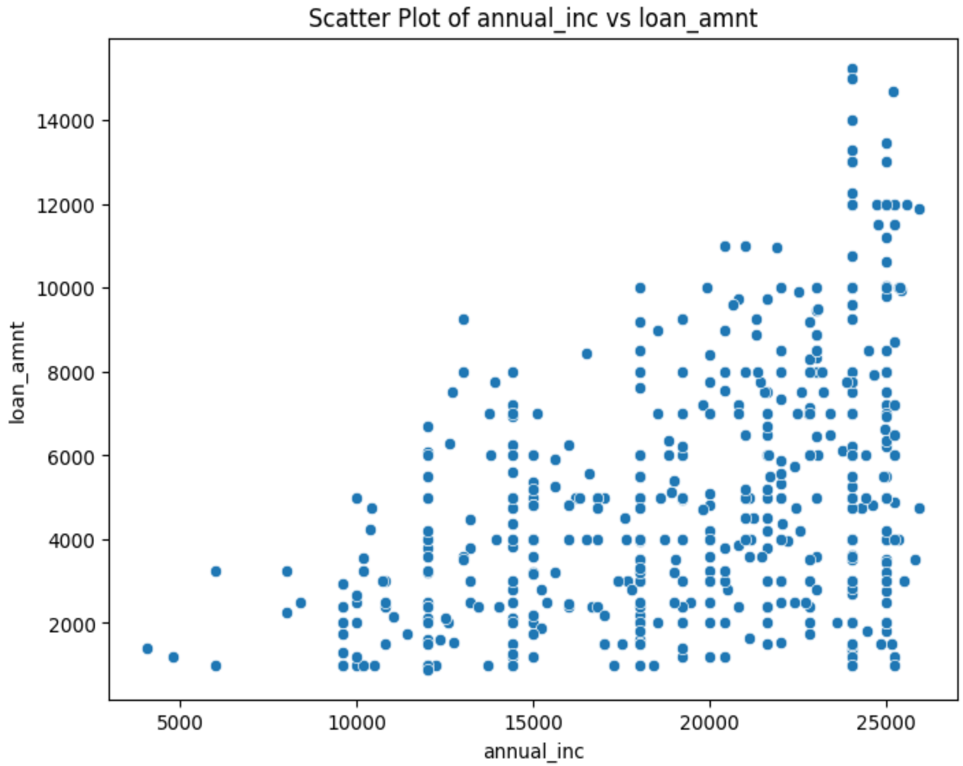
Loan Amounts Concentrated Around Average (~$12,000)

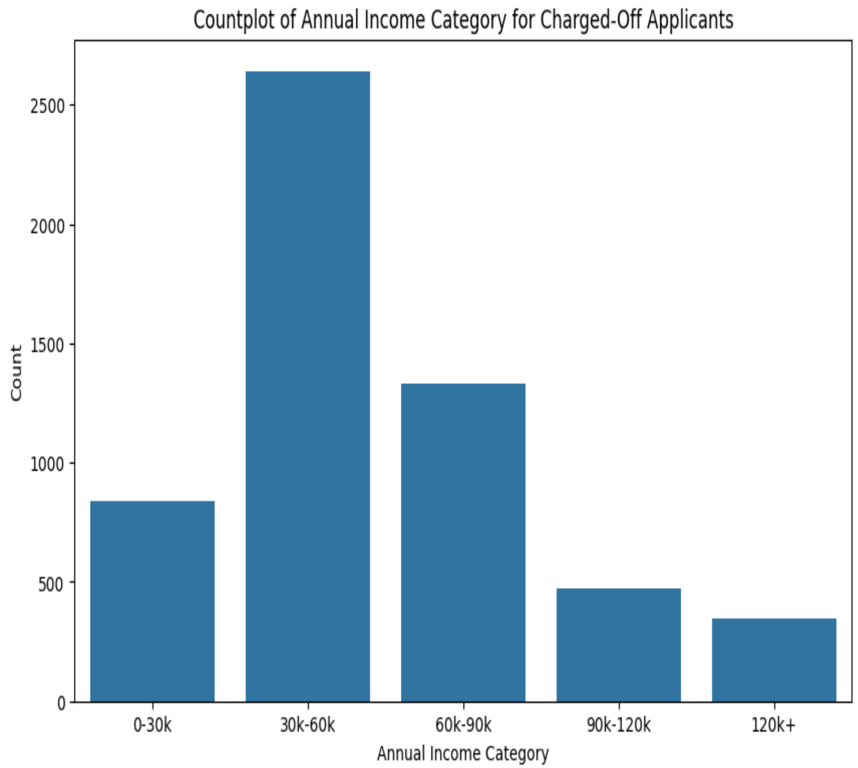
Insight: As the loan amount increases, the number of defaults generally decreases. This suggests that borrowers with larger loan amounts may have better creditworthiness or are seeking loans for more substantial needs, making them less likely to default.

Annual Income vs Loan Default rate

- **Insight**: A clear trend shows that as annual income rises, the number of defaults tends to decrease. Most defaults occur among individuals in the lower income bracket, suggesting a negative correlation. This indicates that roughly 50% of people earn under $55,000.

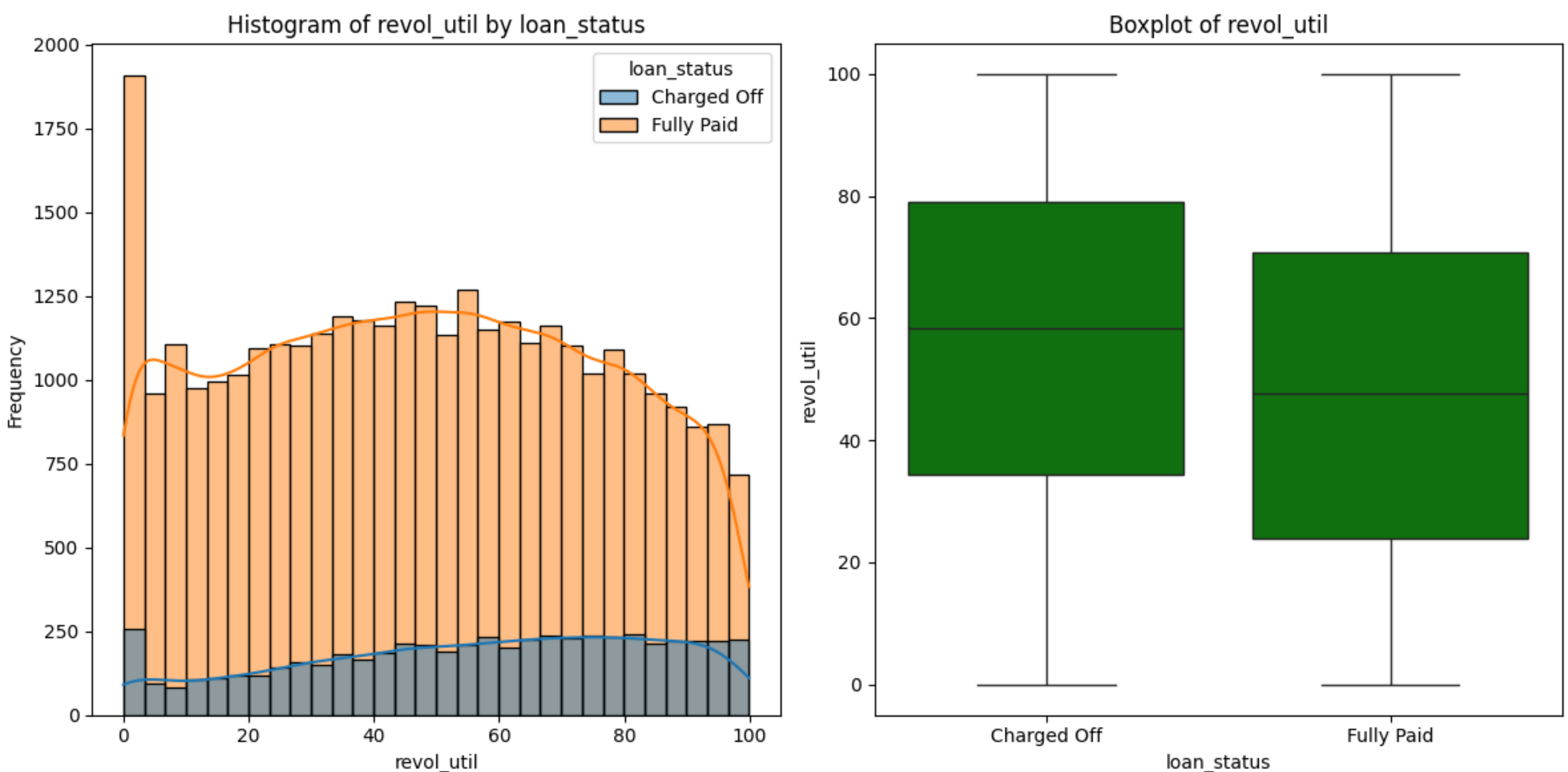
Additionally, there is a Good correlation between annual income and loan amount.



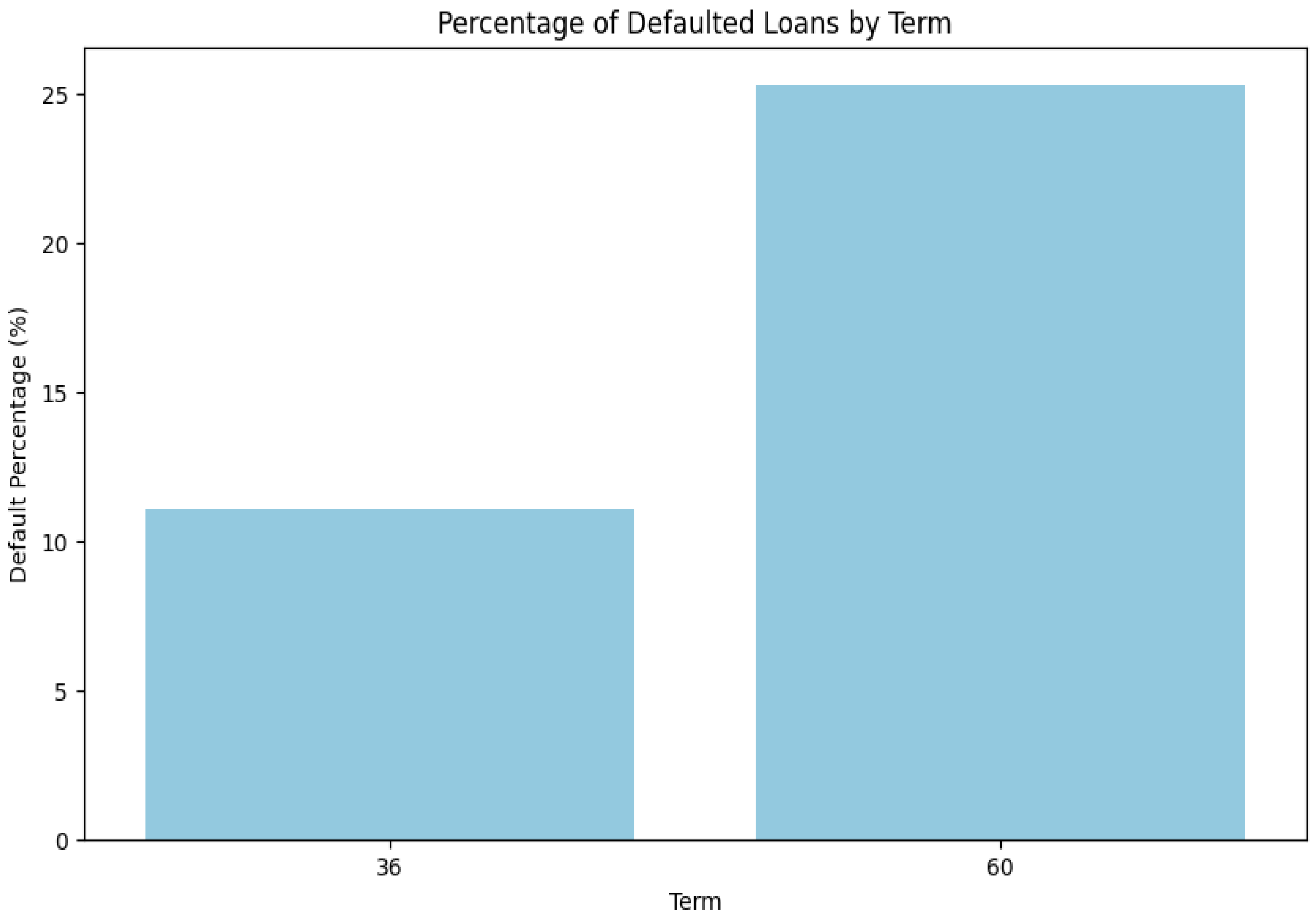


Revolving Balance and Default Rate

The data suggests that borrowers with higher revolving utilization ratios are more likely to default on their loans. This aligns with the intuition that individuals with high credit card debt and low available credit are at greater risk of financial distress.



No of Months(term) Opted for Loans



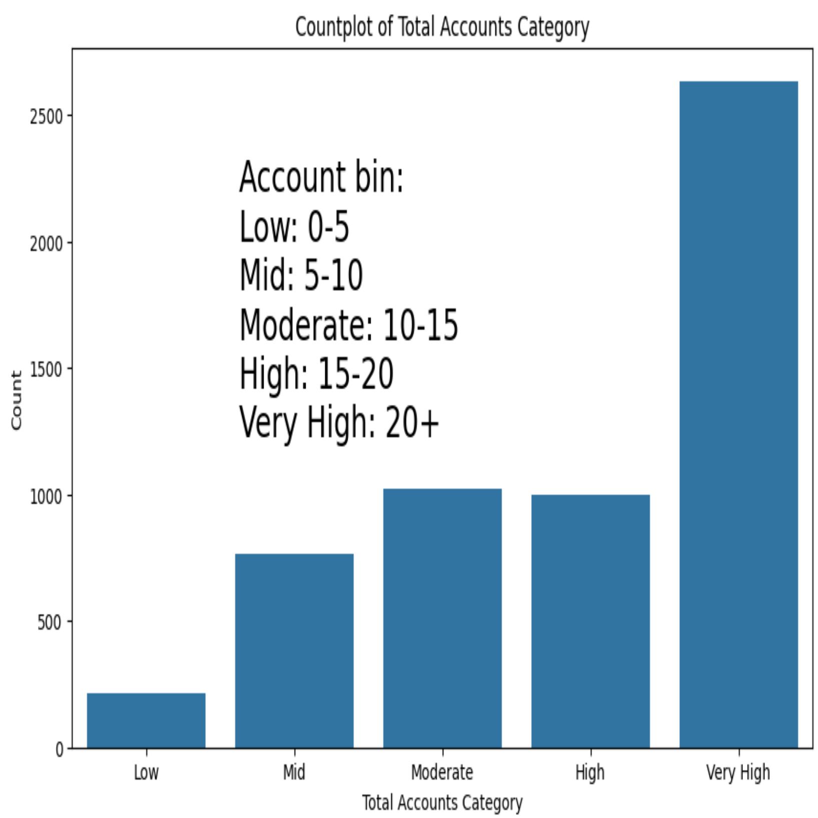
There is a noticeable trend showing that the loan default rate is elevated among those who have opted for 60 months term loan.

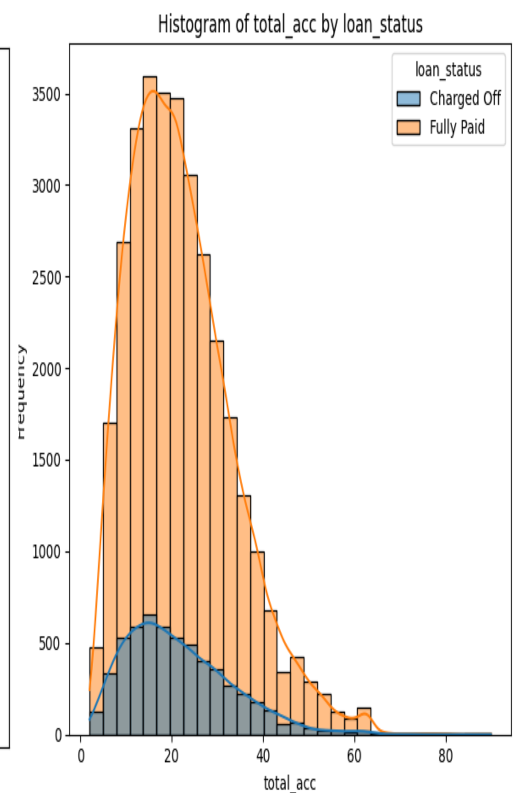
Insights : Since there is high default rate for 60 month the lender can decrease the interest rate for them to decrease the default likeliness.

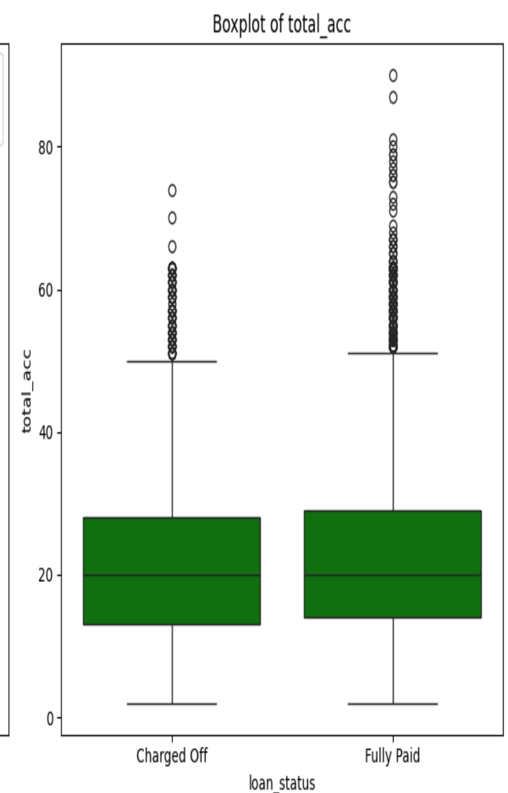
Analysis: Total accounts and Default Rate

Our analysis reveals a correlation between the number of accounts a customer holds (Category class) and their likelihood of default. As the number of accounts increases, the default rate also tends to rise.

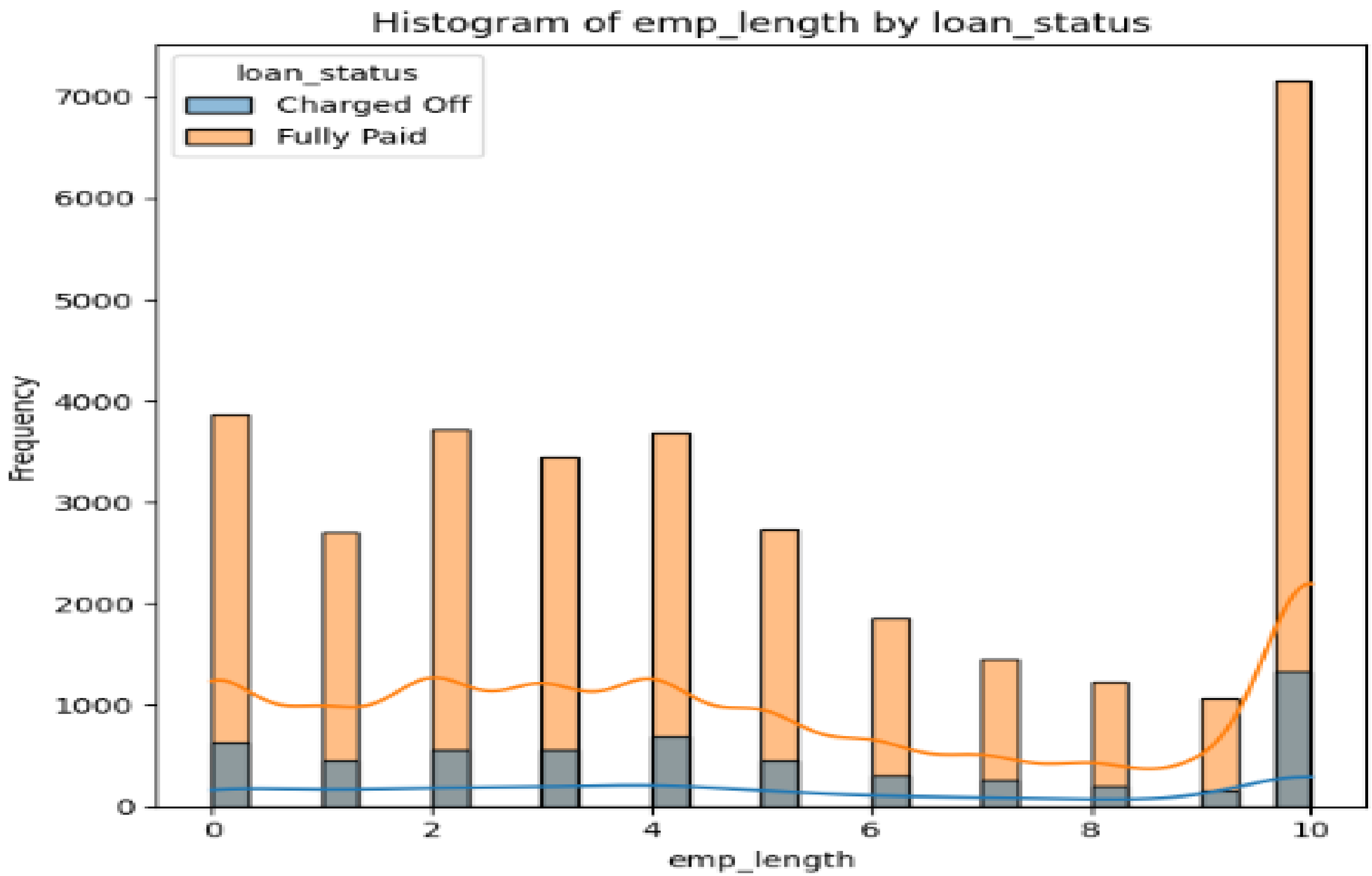
Insights :It Demonstrates the value of the binning strategy we employed, as it enables us to segment customers based on their account activity and assess their associated risk profiles more effectively.





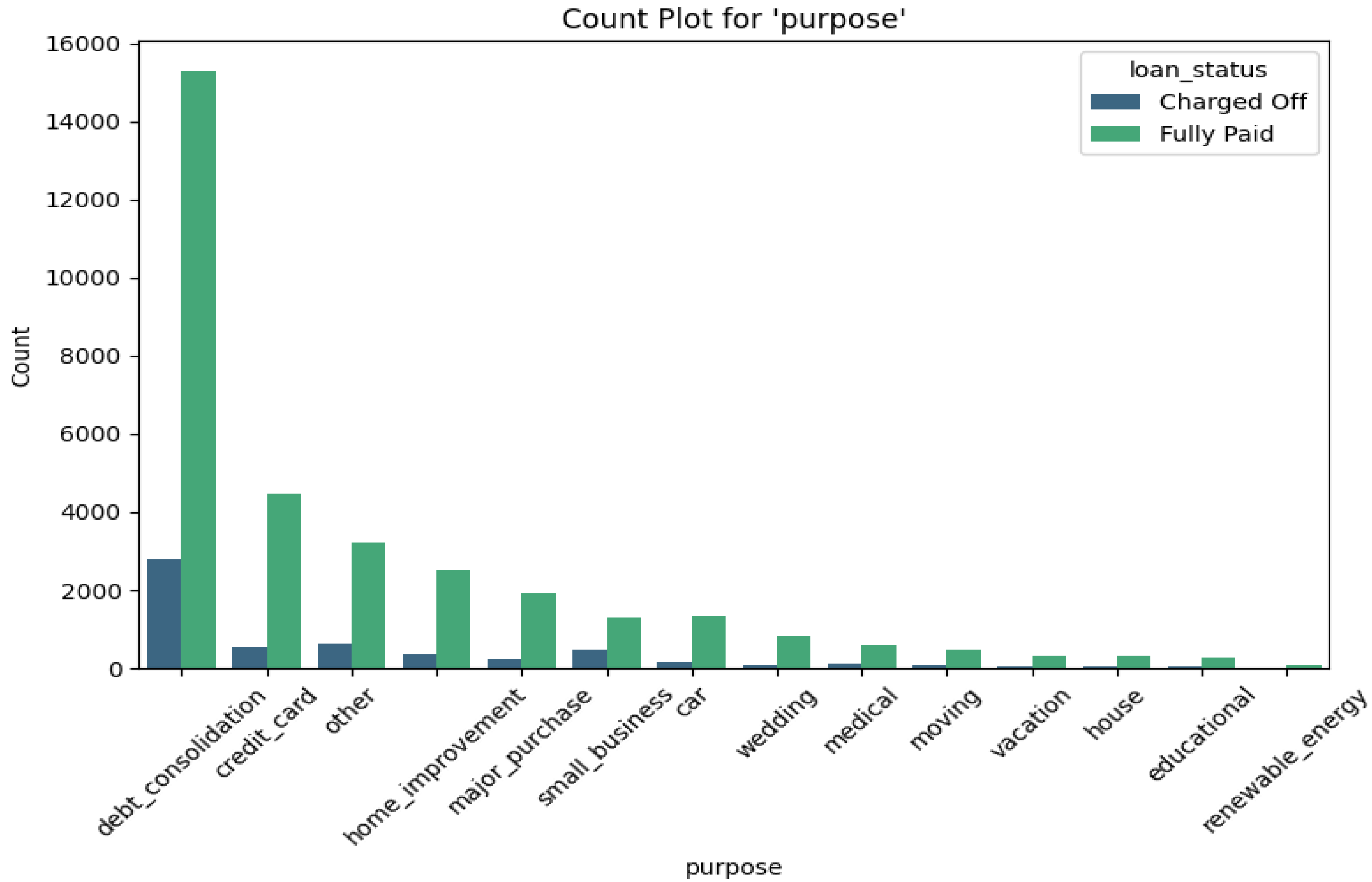


Analysis: Employment Length and Default Rate



- Finding: Shorter employment lengths correlate with higher default rates.
- Insight: Borrowers with longer employment histories may have more stable income, reducing default risk.

Analysis: Purpose mentioned by borrower



Finding : Debt Consolidation Purpose stands out in Default rate of loan

Insight : A high proportion of Charged Off loans in debt consolidation and credit card suggests that these loans might be taken to manage existing financial burdens, potentially leading to default.

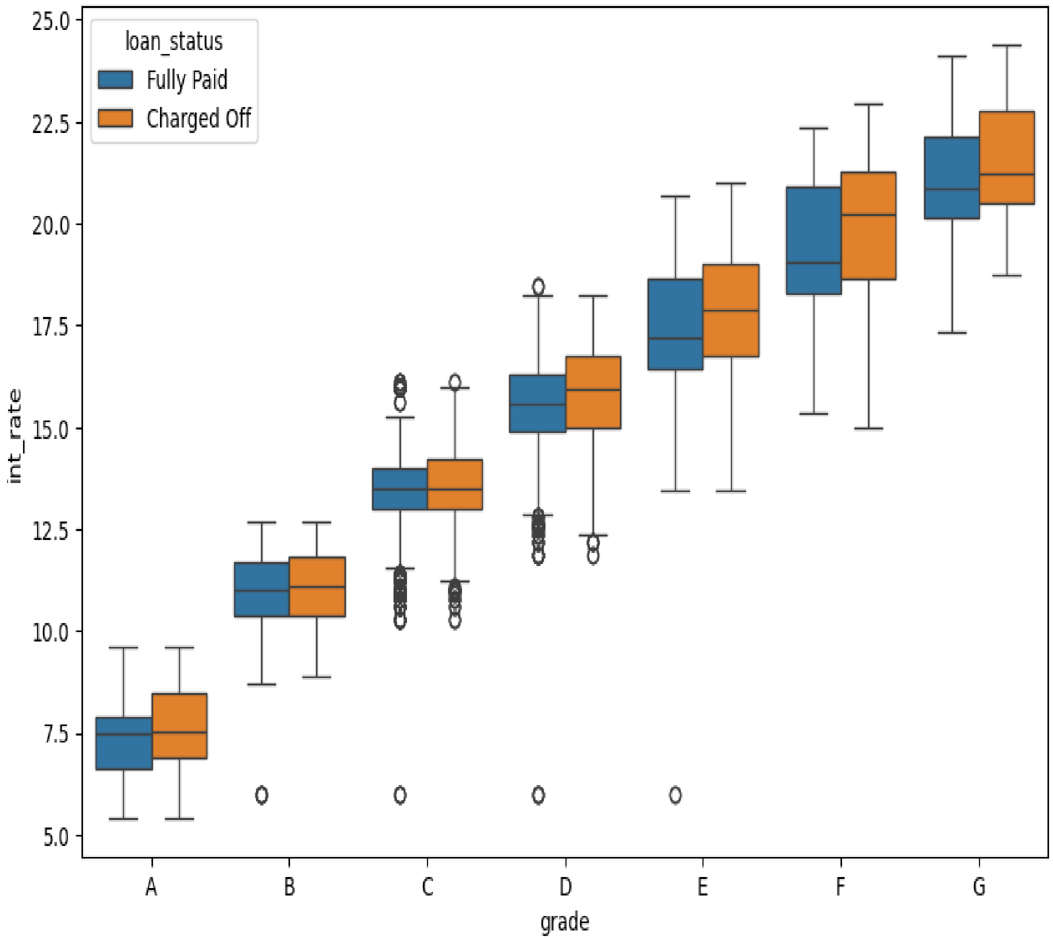
Relationship between Loan Grade, Interest Rate, and Loan Status

Interest Rate Increases with Grade: There is a clear upward trend in interest rates as loan grades decline (A to G), indicating a higher perceived risk for lower grades.

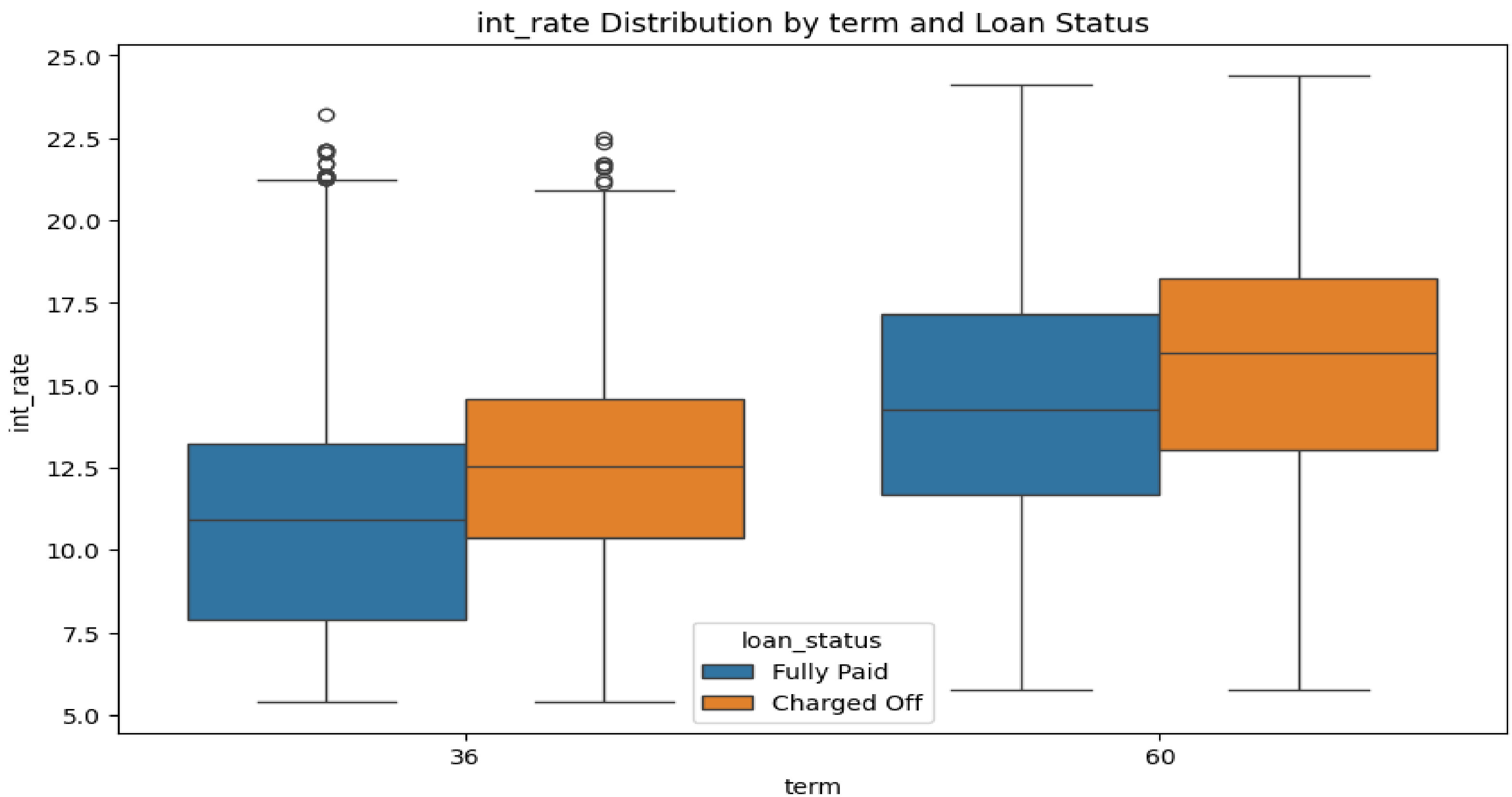
Insight: The higher-risk borrowers are typically assigned lower grades and charged higher interest rates.

Implications for Business

- **Risk Mitigation**: Consider stricter screening for loans with high interest rates, especially in the D-G grade range, as they are more prone to defaults.
- **Interest Rate Strategy**: Evaluate the interest rate strategy for each grade, as higher rates may increase borrower risk of default.

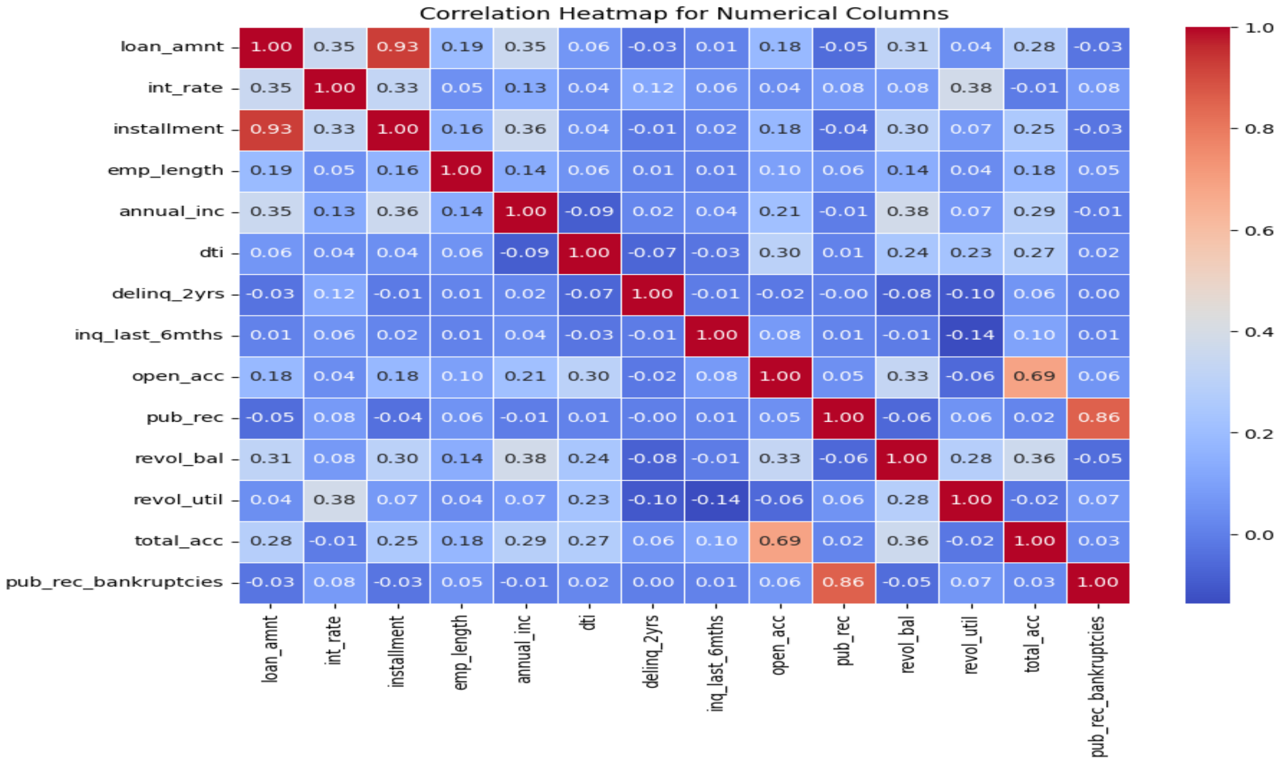


Relationship between int_rate Distribution by term and Loan Status



- Individuals are facing elevated installment payments for their respective term loans because the charges are higher than those for fully paid loans. This situation might also suggest that there are other factors contributing to lower interest rates.
- Additionally, this could indicate that investors are willing to take on more risk by offering higher interest rates.

Multivariate Analysis: Correlation Heatmap



Moderate Correlations:

•**loan_amnt and installment**: A notable positive correlation shows that larger loan amounts typically result in higher monthly payments.

•**loan_amnt and annual_inc**: A moderate positive correlation indicates that borrowers with higher incomes are often approved for larger loan amounts.

•**int_rate and revol_util**: A moderate positive correlation suggests that individuals with greater revolving utilization rates (credit card debt) may face higher interest rates.

•**revol_bal and total_acc**: A moderate positive correlation implies that borrowers who have more accounts usually carry larger revolving balances.

Key Insights on Default Risk Indicators and Prevention

Key High-Risk Indicators:

Factors such as loan grades, loan amounts, low annual income, loan term (in months), and revolving balance utilization play crucial roles in determining the likelihood of default.

Insights for Prevention:

Impact of Loan Grade: There is a noticeable trend where lower loan grades (from A to G) correlate with higher default rates. Adjusting the terms for these lower grades may help reduce associated risks.

Revolving Utilization Rate: Higher revolving utilization rates are commonly found in lower grades, such as F and G. This highlights the need for more thorough credit evaluations for individuals with significant credit card debt ratios.

Correlation Between Loan Amount and Income: While higher income often permits larger loans, defaults are frequently observed among those with lower income levels. Confirming income can improve the precision of loan approvals and assessments of repayment ability.

Bankruptcy History: A previous bankruptcy is significantly associated with increased default risks, making it vital to track bankruptcy records for effective risk assessment.

Additionally, there is a relationship between low annual income and elevated default rates, underscoring the importance of setting income thresholds during loan evaluations.

Lastly, an increase in the total number of accounts is correlated with a higher likelihood of default, stressing the necessity for tailored risk assessments.