Loan Default Analysis for Risk Assessment

Understanding Key Indicators of Loan Default

Prepared by vikas and lakshman

upGrad

iiit-b

ज्ञानमुत्तमम

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.

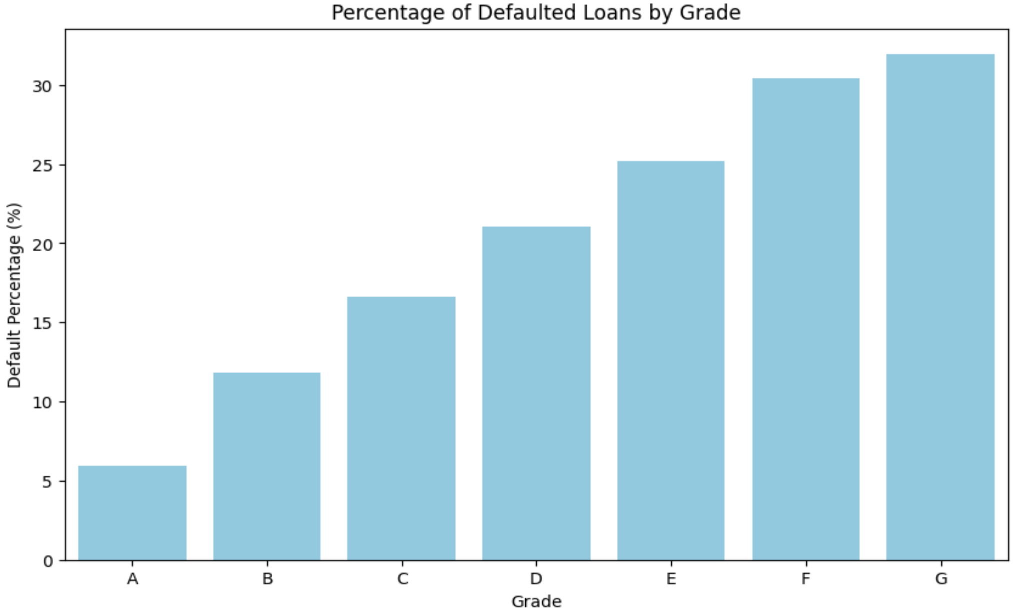
Data Analysis Approach

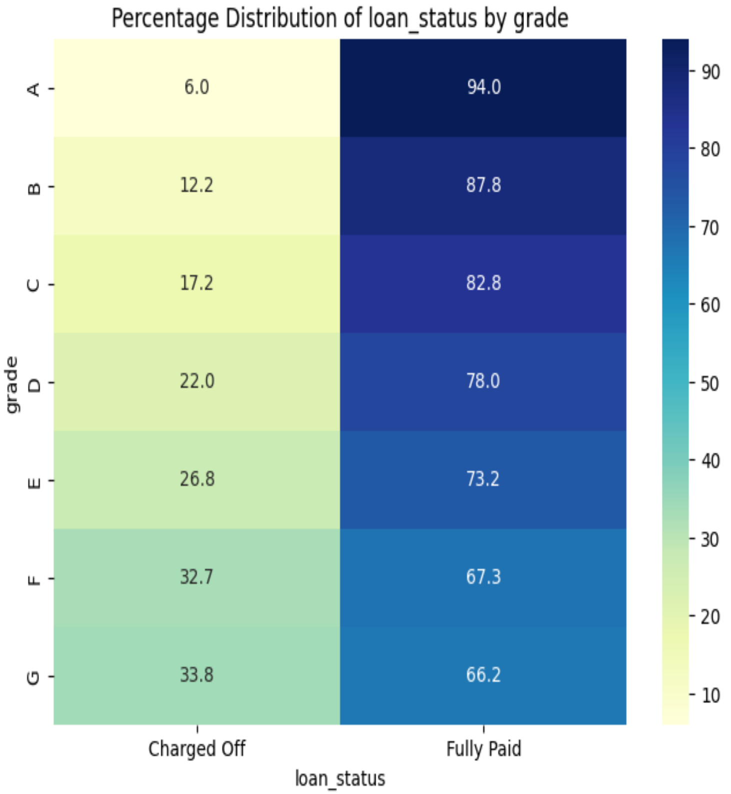
1. Data Overview: Summary of the dataset and key columns (e.g., loan amount, interest rate, employment length, loan status).

2. Analytical Steps:

- - Univariate Analysis: Examined distribution and outliers, particularly focusing on skewed distributions (e.g., income,loan amount).
- - Bivariate Analysis: Explored relationships between loan default and variables like loan grade, income, and credit history.
- - Multivariate Analysis: Identified interdependencies using correlation techniques.

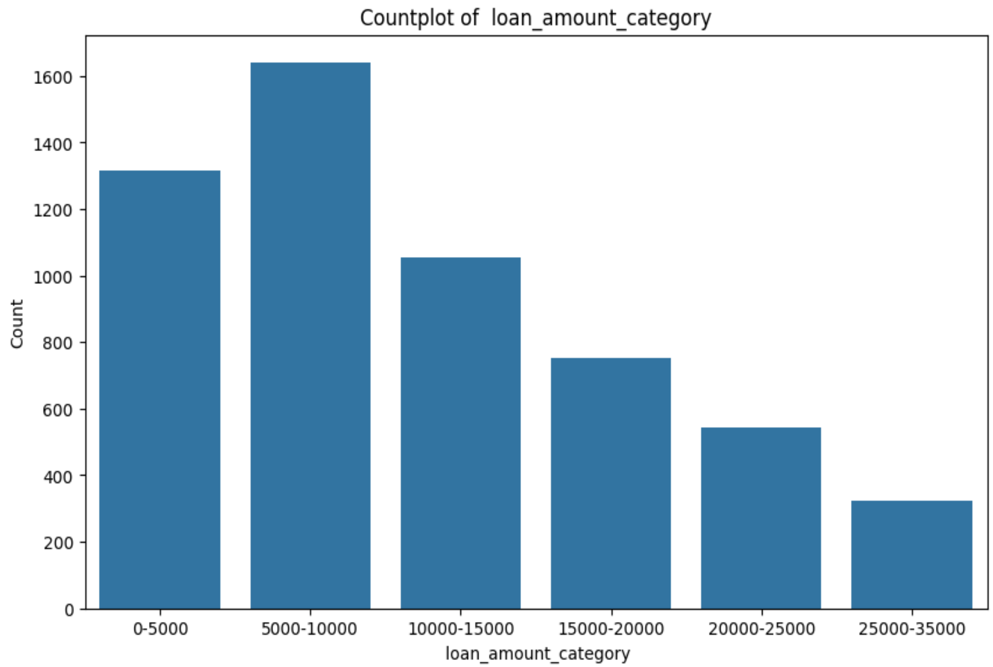
Analysis: Loan Grade and Default Rate

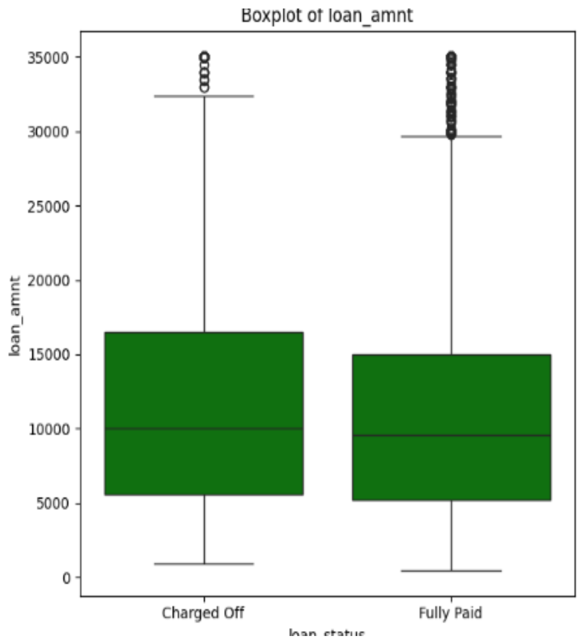




- Finding: Default rates vary significantly by loan grade.
- Insight: Lower grades (e.g., D,E,F,G) show higher default percentages, indicating correlation between grade riskiness and default likelihood.

Loan Amount and Default Rate

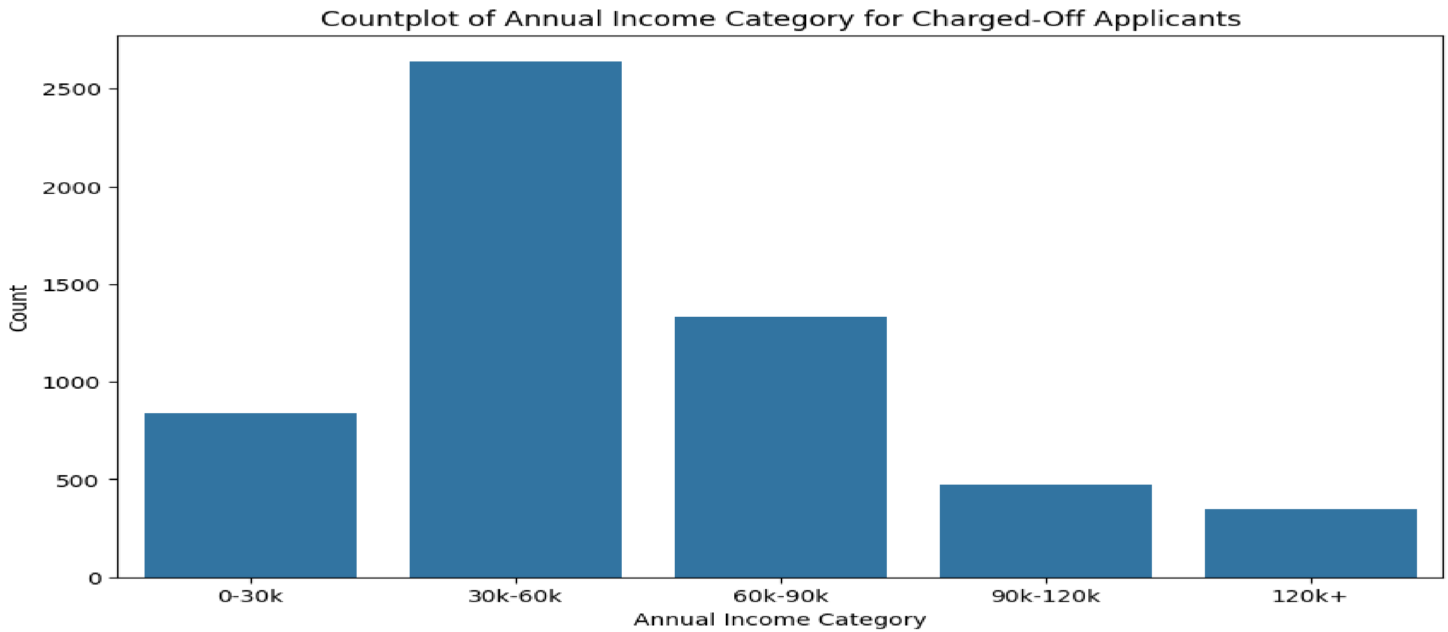




Loan Amounts Concentrated Around Average (~$12,000)

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

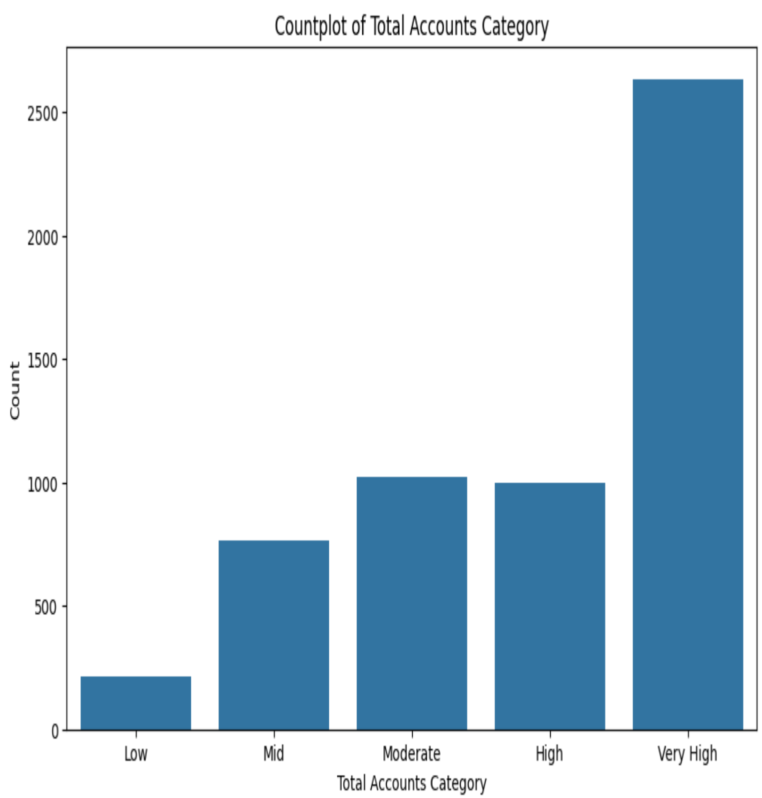


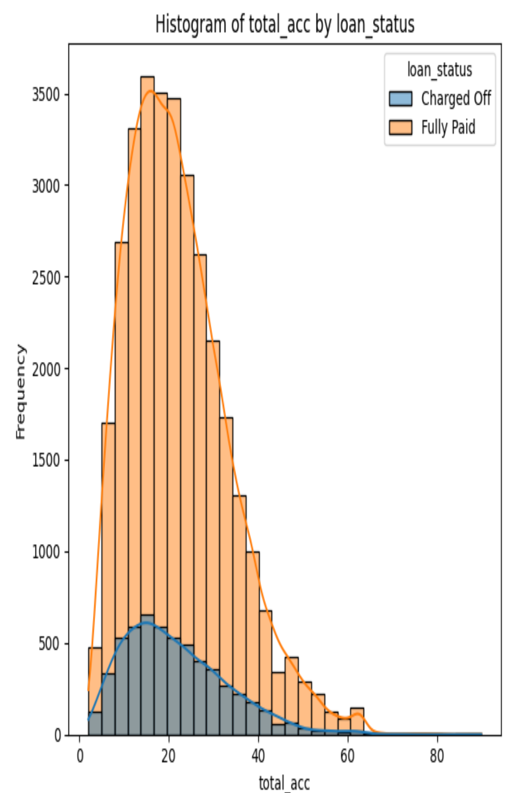
Insight: There is a noticeable trend where an increase in annual income correlates with a decrease in the number of defaults. The majority of those who defaulted fall within the lower income bracket, indicating a negative correlation. This serves as a strong indicator that approximately 50% of individuals earn less than $55,000.

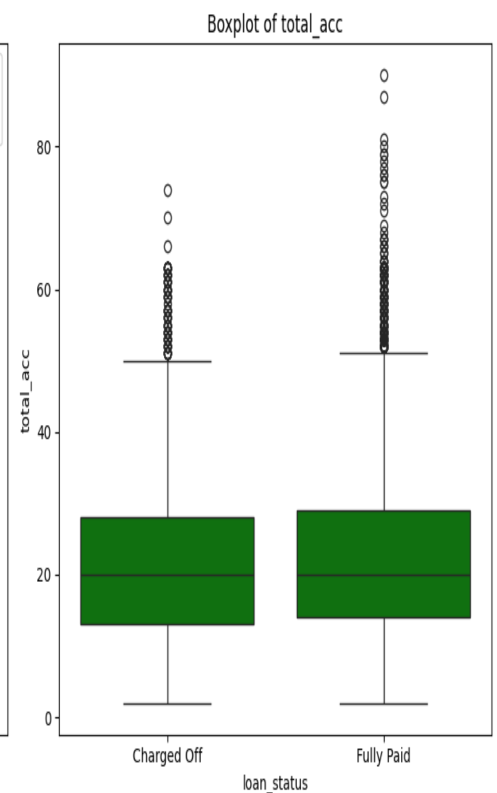
Analysis: Total accounts and Default Rate

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.

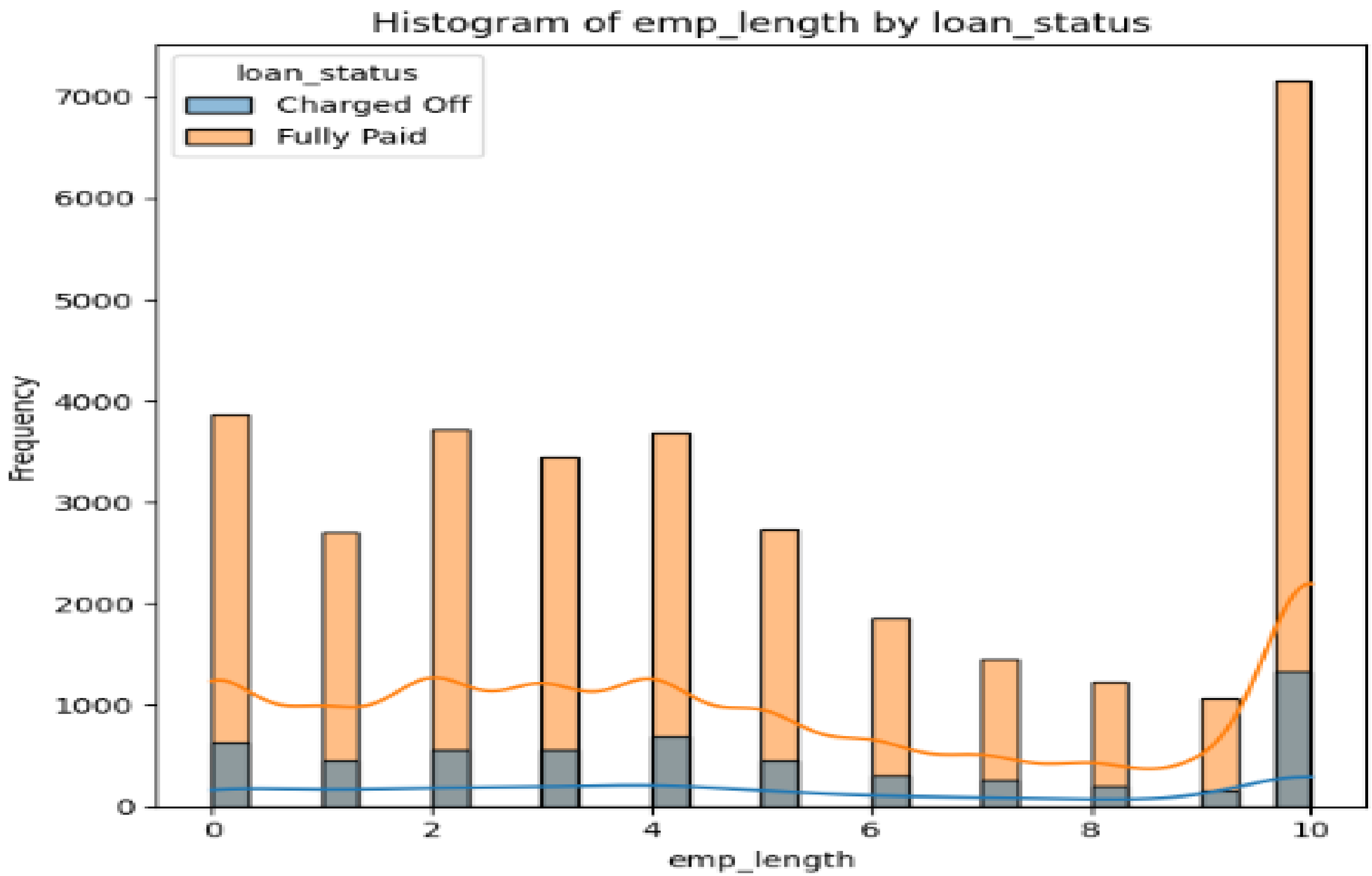
Insights :it Demonstrates the value of the binning strategy we employed, as it enables us to segment customers based on their account activity and assess their associated risk profiles more effectively.





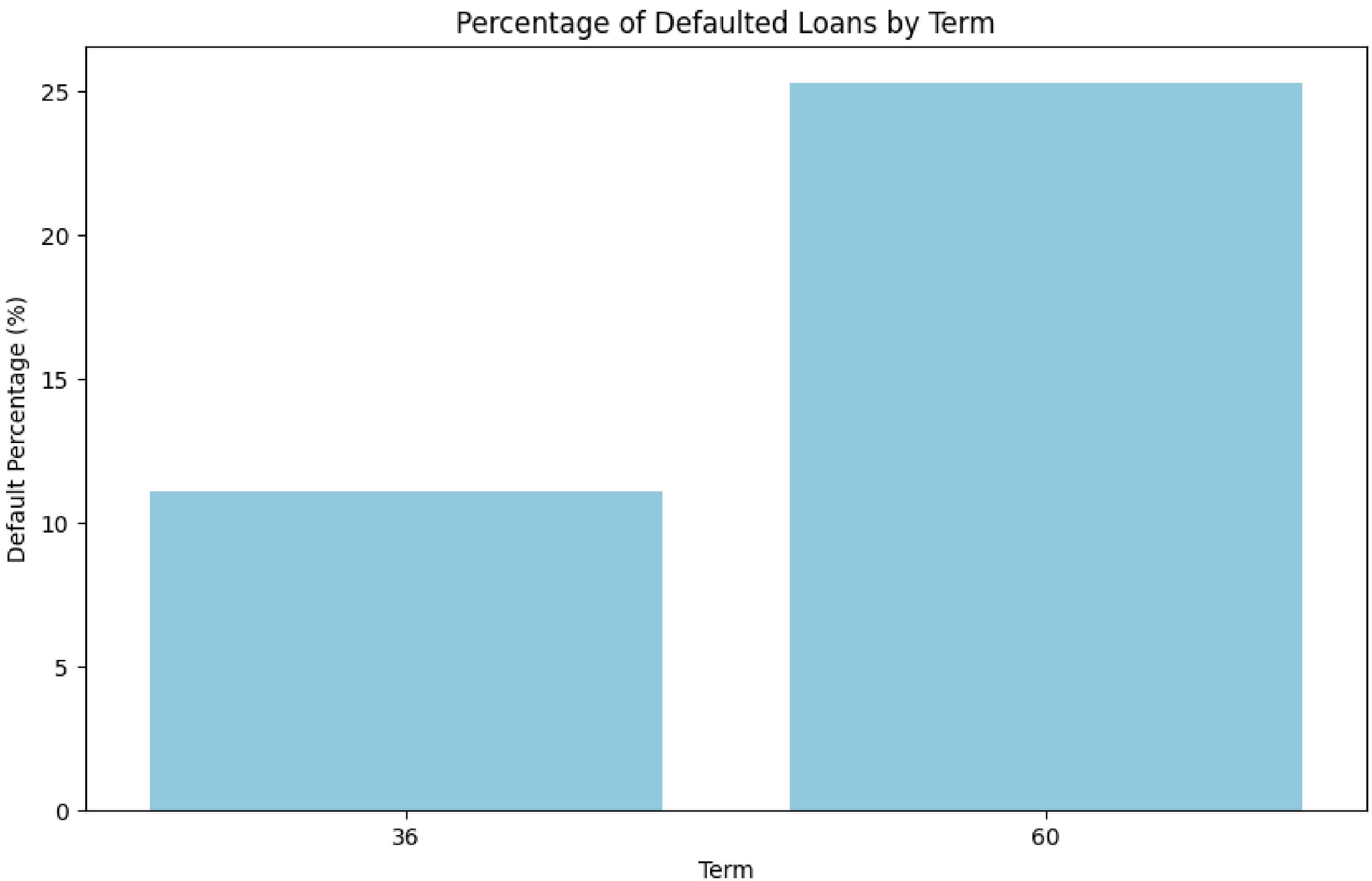


Analysis: Employment Length and Default Rate



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

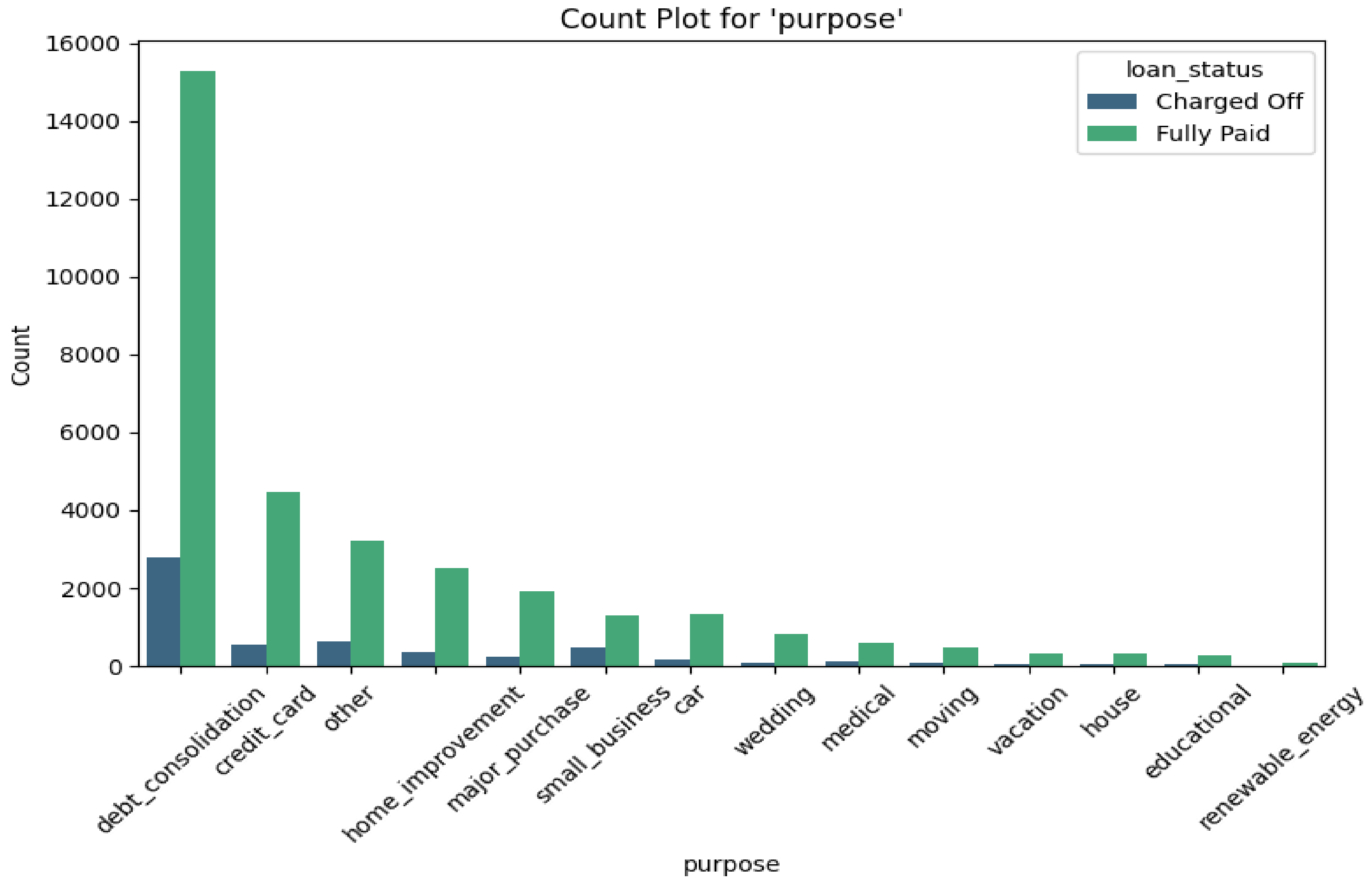
No of Months(term) Opted for Loans



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: 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.

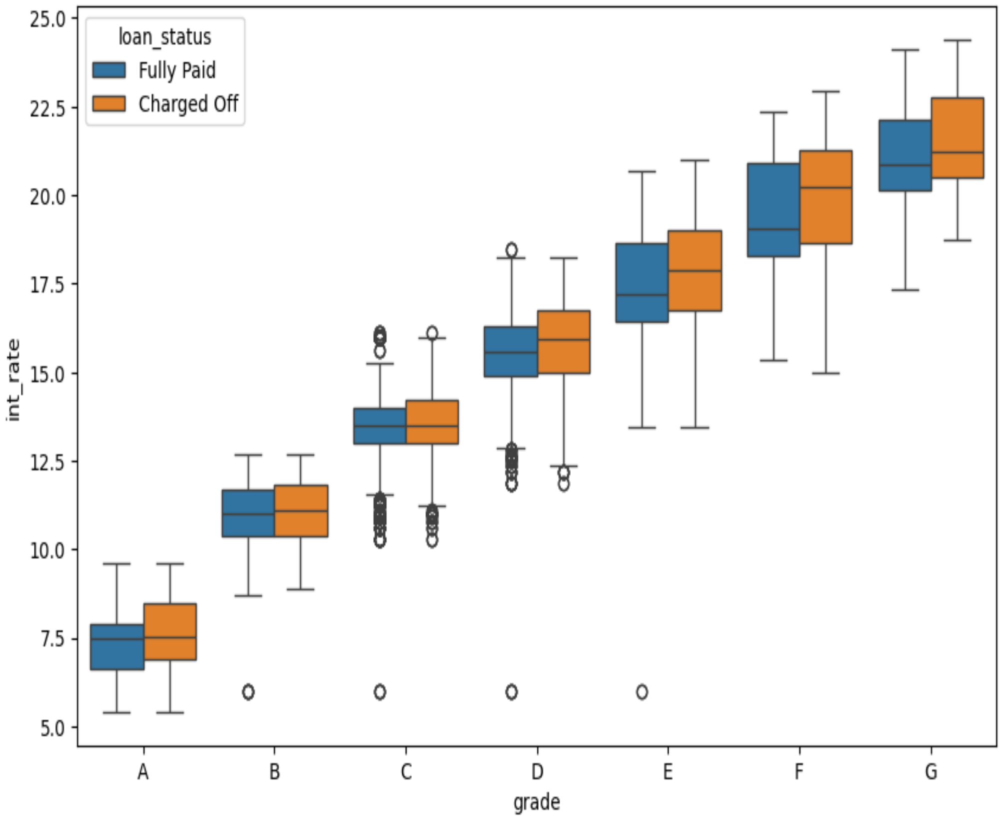
Relationship between Loan Grade, Interest Rate, and Loan Status

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.

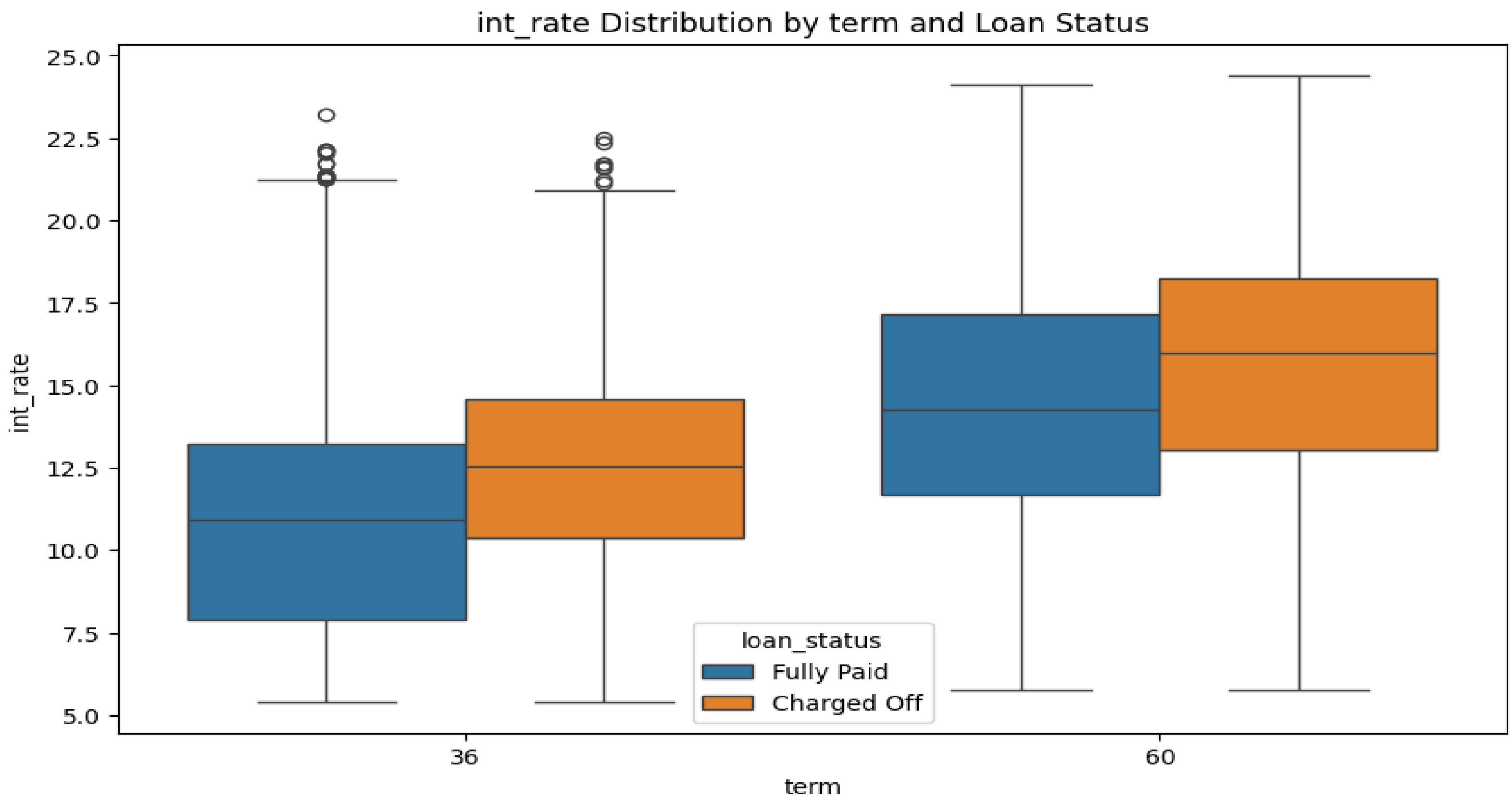
Insight: The higher-risk borrowers are typically assigned lower grades and charged higher interest rates.

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.

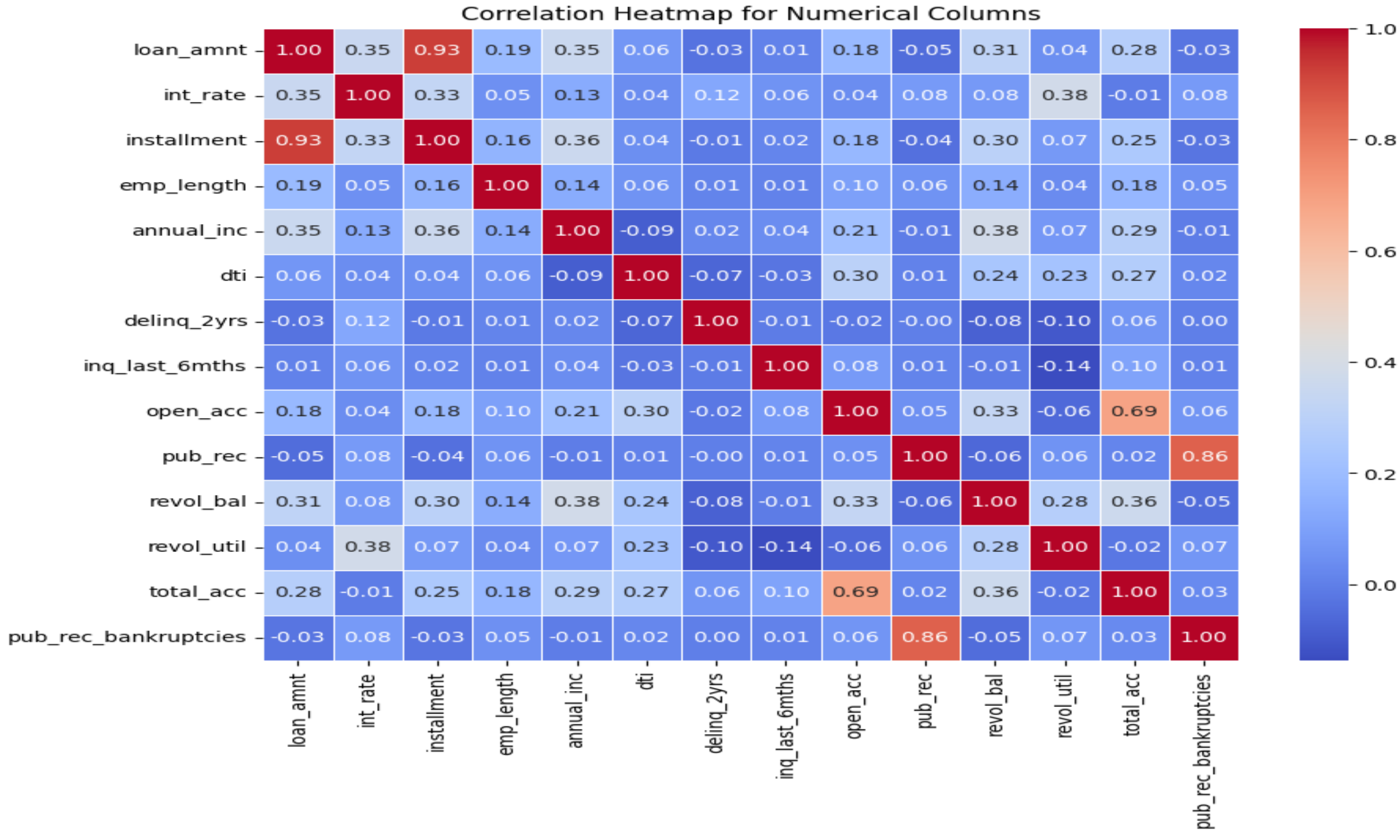


Relationship between int_rate Distribution by term and Loan Status



- 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



- Finding: Variables like loan grade, income, and credit history indicators exhibit moderate correlations.

- Insight: These factors jointly impact the likelihood of loan default, supporting their role in risk assessment.

Key Insights on Default Risk Indicators and Prevention

Key High-Risk Indicators:

Factors such as loan grades, loan amount, low Annual Income amounts ,term and revolving Balance Utilisation are decisive factor for default Likeliness.

Insights for Prevention:

Bankruptcy History: A history of bankruptcy is closely linked to increased default risks, making it essential to monitor bankruptcies for effective risk evaluation.

Impact of Loan Grade: There is a clear trend of higher default rates associated with lower loan grades (ranging from A to G). Modifying terms for lower grades could potentially mitigate risk.

Revolving Utilization Rate: Elevated revolving utilization rates are often seen in lower grades (like F and G), indicating the necessity for more rigorous credit assessments for individuals with high credit card debt ratios.

Correlation Between Loan Amount and Income: Although higher incomes generally allow for larger loans, defaults frequently occur within lower range . Verifying income can enhance the accuracy of loan approvals and assessments of repayment capabilities.

Risk Factors in Loan Default Analysis

- Many applicants either having low employment experience or possess more than a decade of work history.
- There is a connection between low annual income and higher default rates, highlighting the need for income thresholds during loan assessments.
- An increase in the number of total accounts is linked to a greater probability of default, emphasizing the importance of customized risk evaluations.