

Lending Club Case Study and Report

Presented by

Gowtham Neelapareddy
Chandana

Date Presented

20 November 2024

Agenda

Overview

Data Description

Analysis Stages

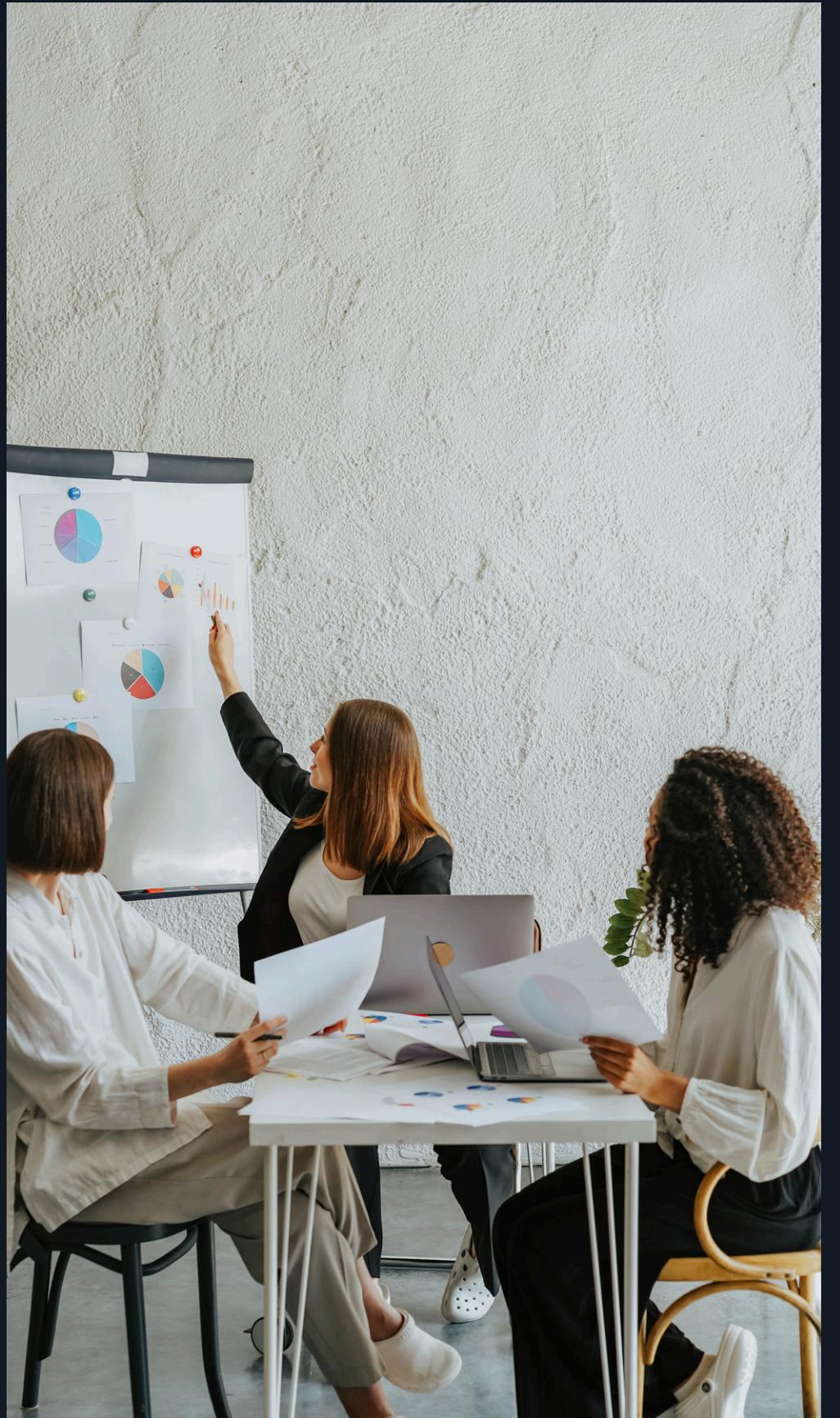
Univariate Analysis

Segmented Univariate Analysis

Bivariate Analysis

Default Rate Analysis

Summary

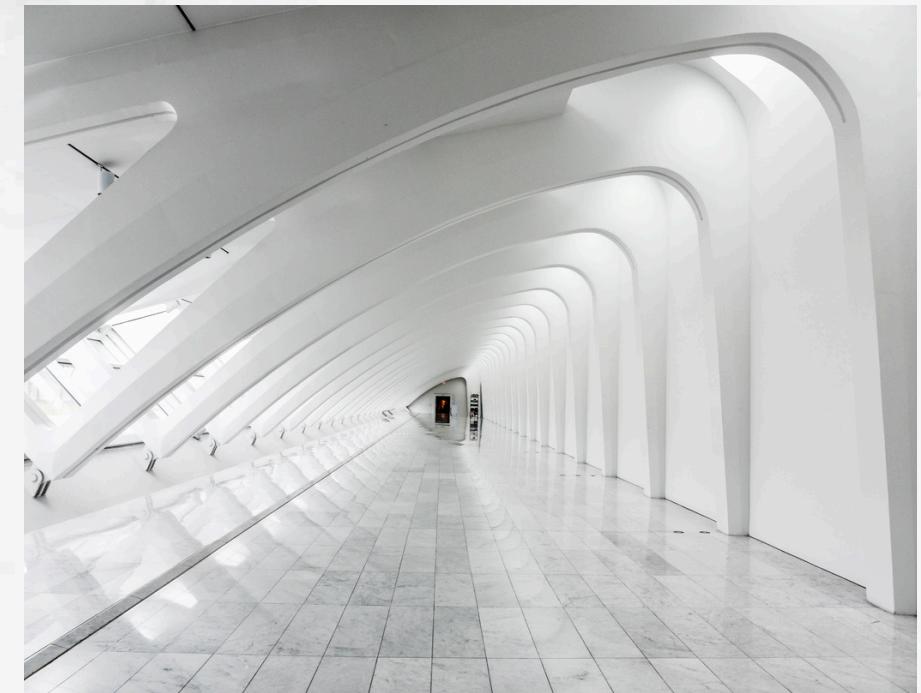
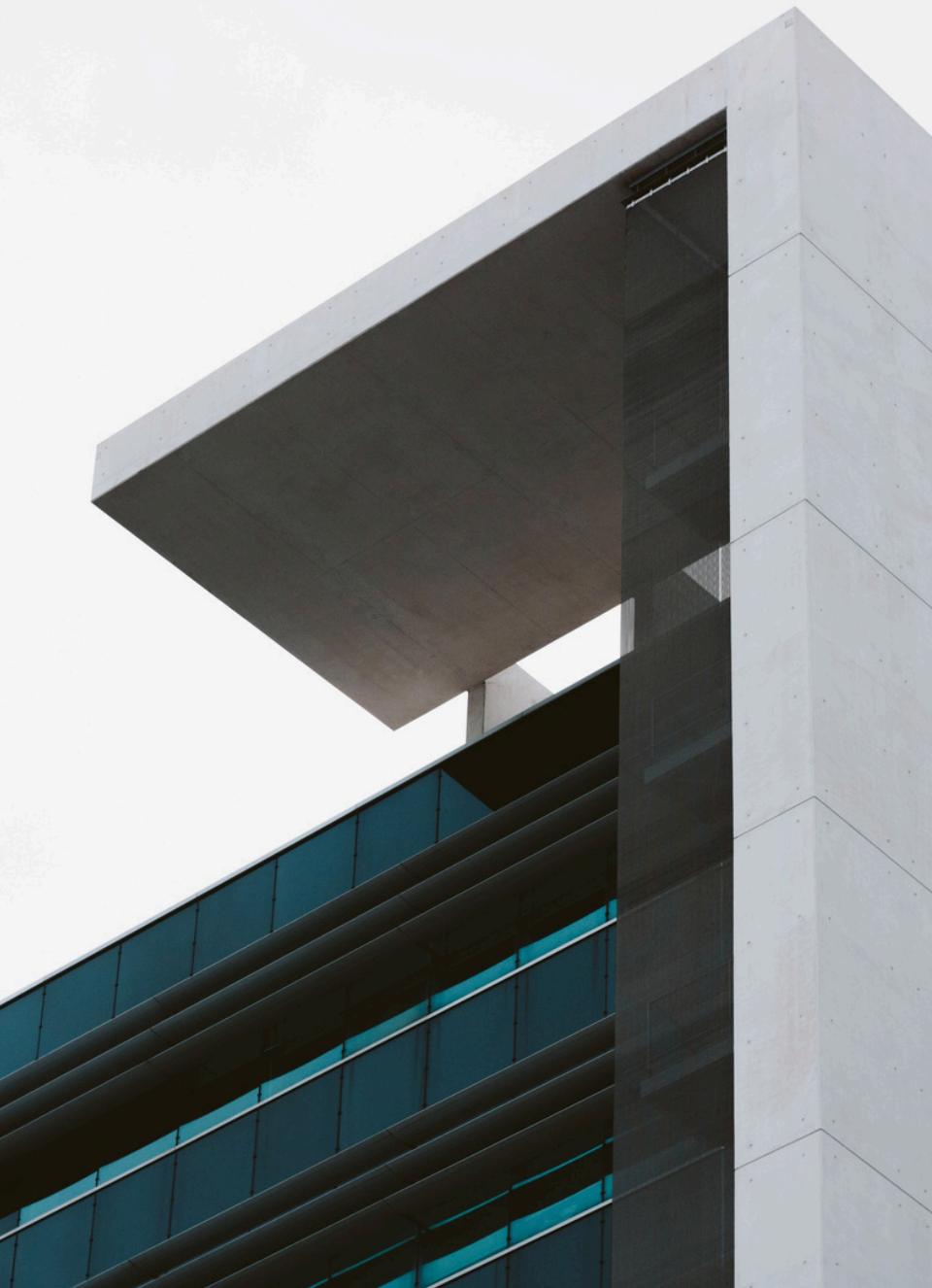




[Back to Agenda](#)

Overview

- The project aims to understand the dynamics of loan defaults and assess risk factors associated with borrowers.
- We explore various aspects of the loan data, including income levels, employment history, loan amounts, interest rates, and more, to identify patterns and trends that predict loan performance.





[Back to Agenda](#)

Data Description

- The dataset consists of loan data from a leading financial institution, including information on borrower demographics, financial behaviour, and loan characteristics.
- Key attributes include loan amount, interest rate, employment length, home ownership status, annual income, and loan outcome.



[Back to Agenda](#)

Analysis Stages

Data Understanding

The process of comprehending the data's structure, attributes, and sources to ensure alignment with the analysis objectives.

Data Cleaning

The process of refining the dataset by fixing errors, removing duplicates, and dropping irrelevant or sparse columns to enhance its quality and relevance.

Data Analysis

The systematic application of statistical and exploratory methods to interpret and extract meaningful information from the data.

Data Correlation

The study of relationships between variables to understand their interactions and influence on each other.

Business Insights

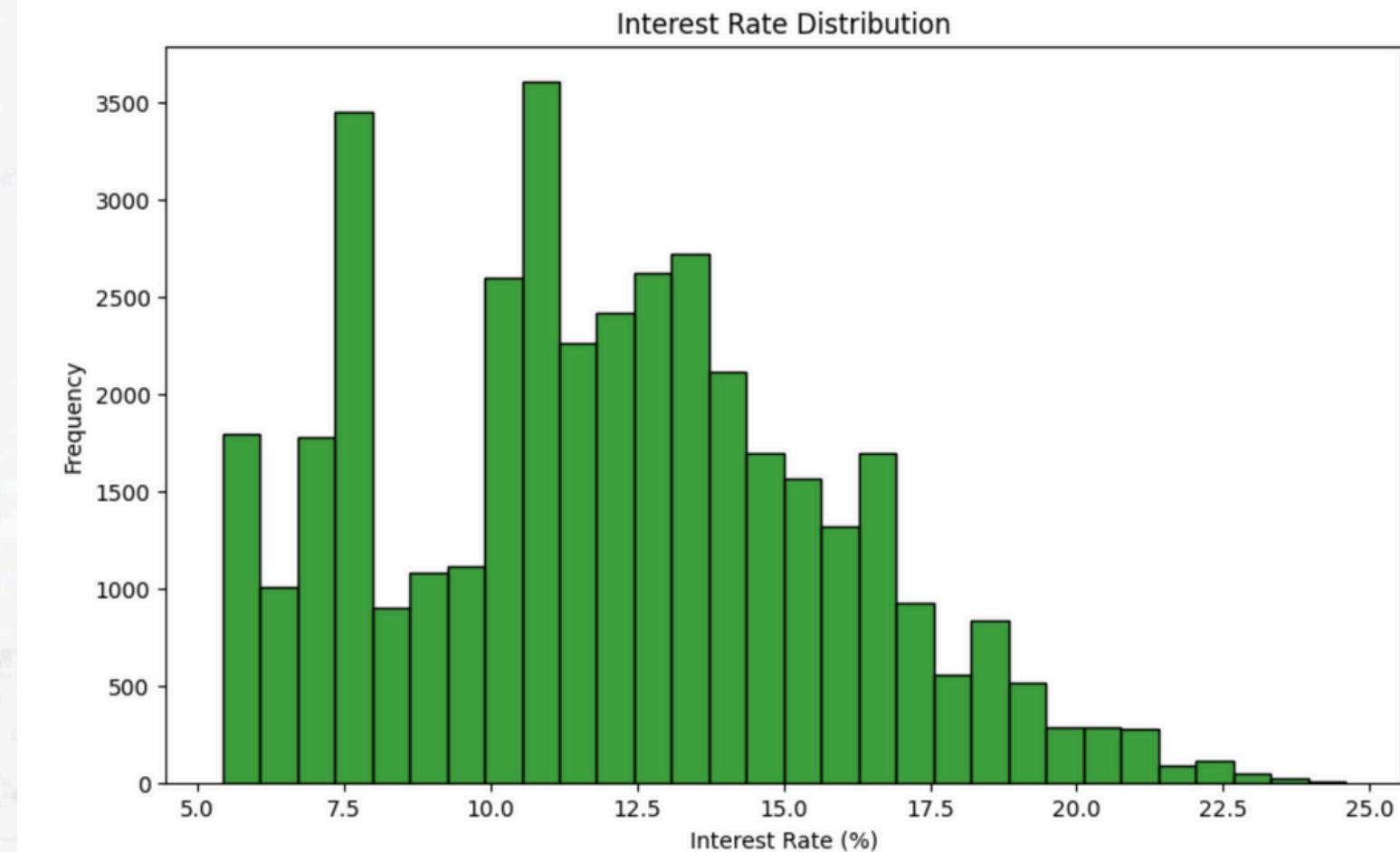
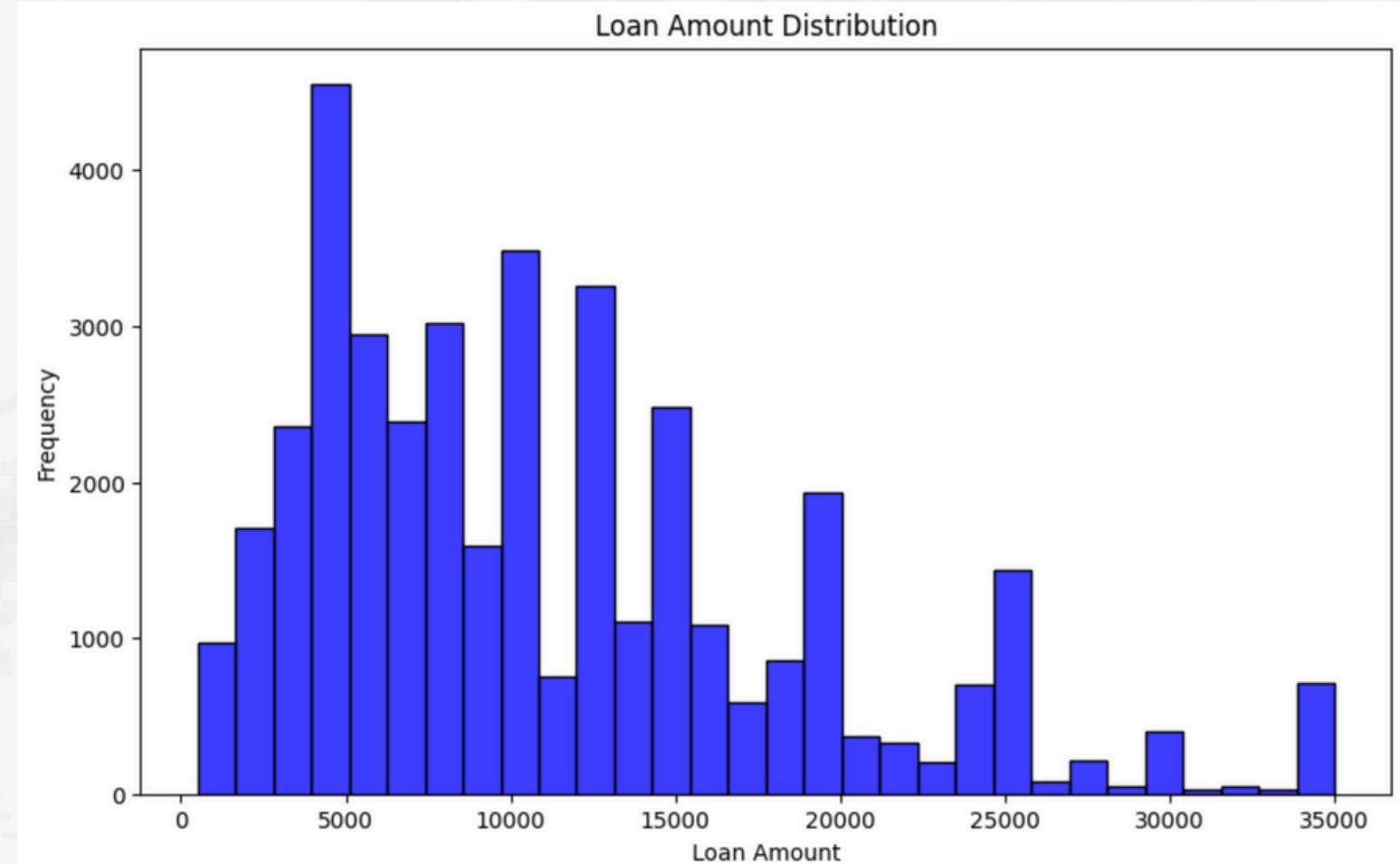
The actionable conclusions drawn from data analysis to guide strategic decision-making and optimize business outcomes.



[Back to Agenda](#)

Univariate Analysis

The analysis identified that the distribution of loan amounts showed multiple peaks, typically around 5000, 10,000, 13,000, and 15,000, indicating popular loan amounts. Interest rates mostly clustered around 10% to 15%, highlighting common lending rates.

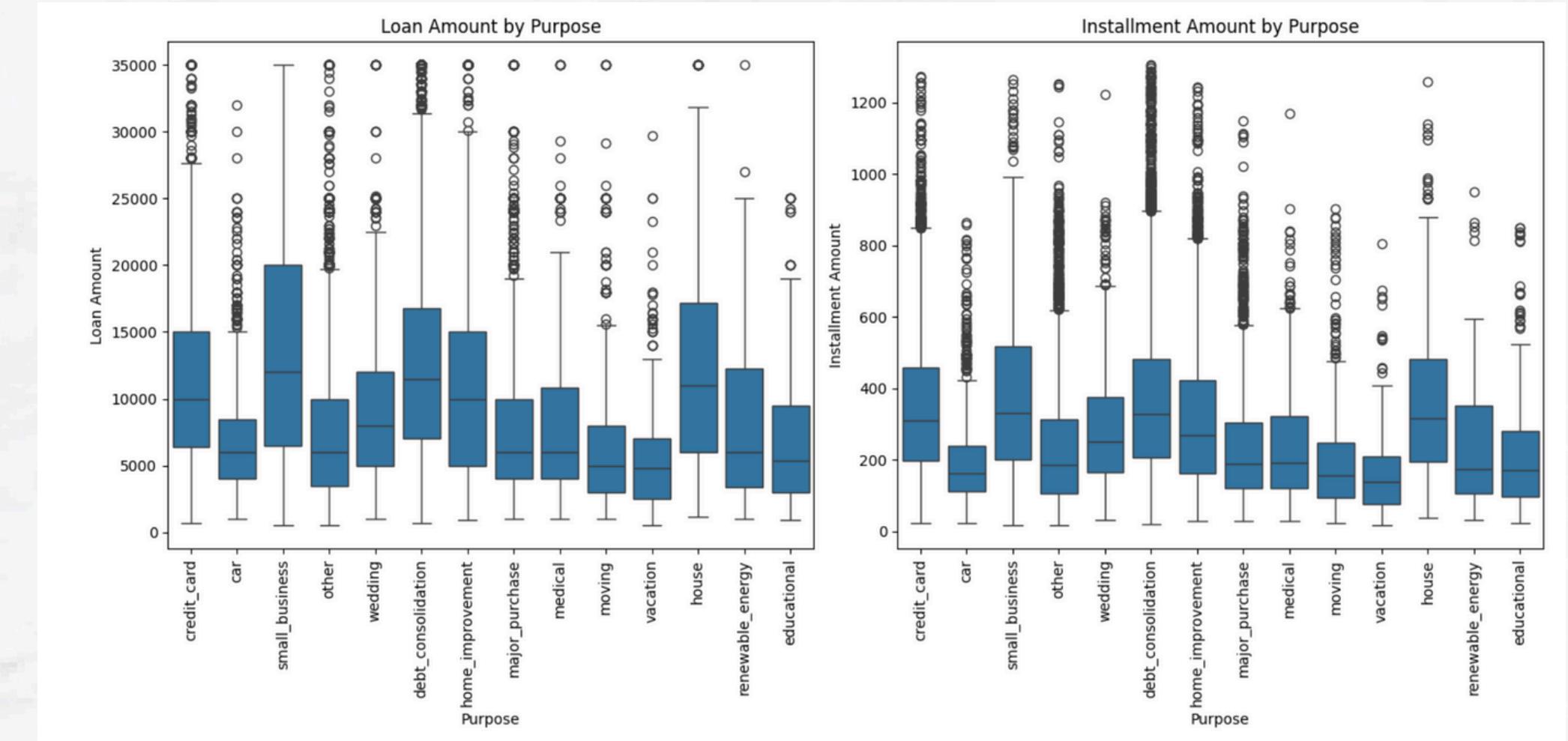




[Back to Agenda](#)

Segmented Univariate Analysis

Segmented analysis by loan purpose showed that loans for small business, debt consolidation and house had higher average amounts and rates, identifying them as high-risk areas. This segmentation helps in tailoring risk management strategies for different loan purposes.

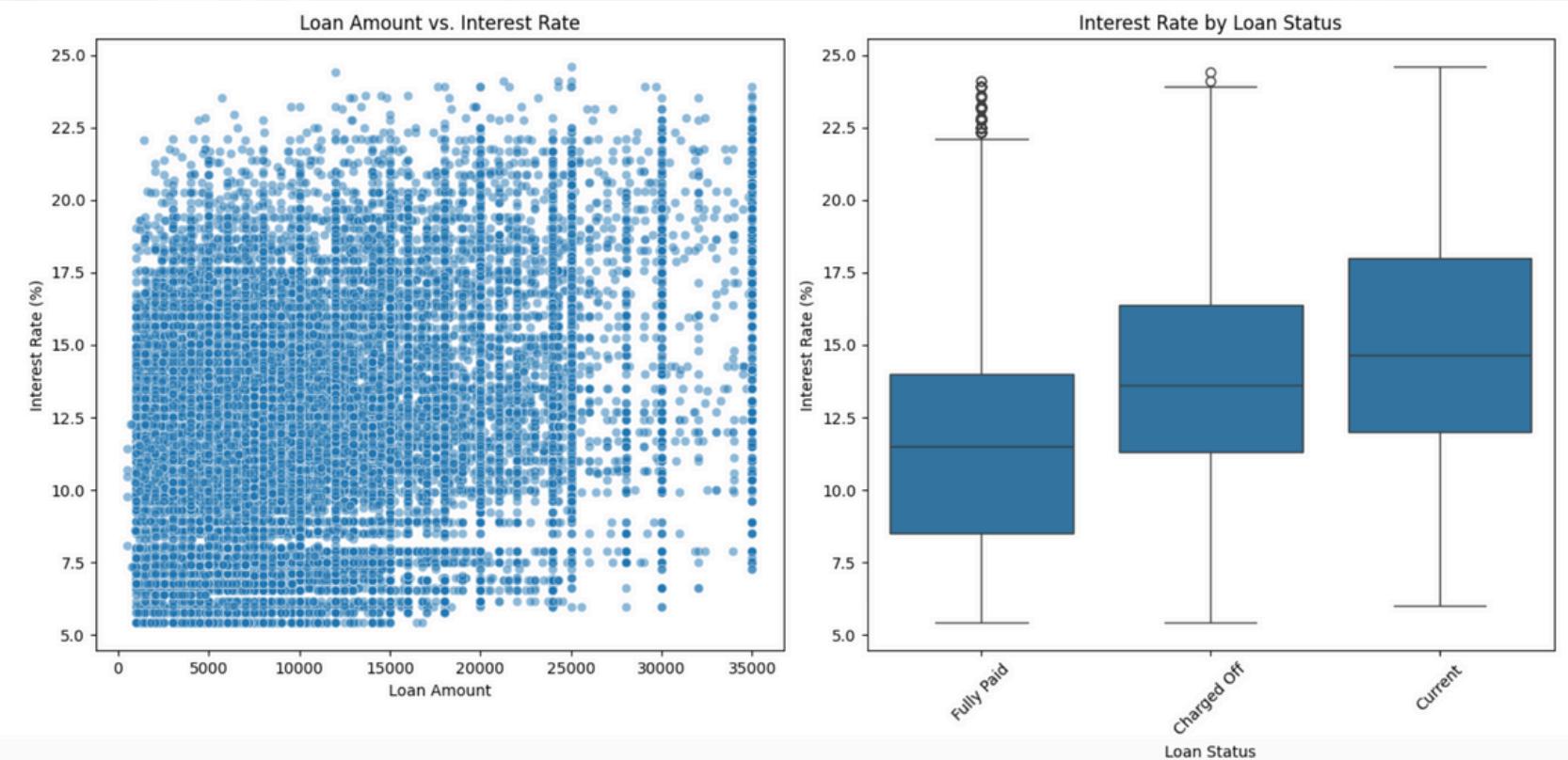
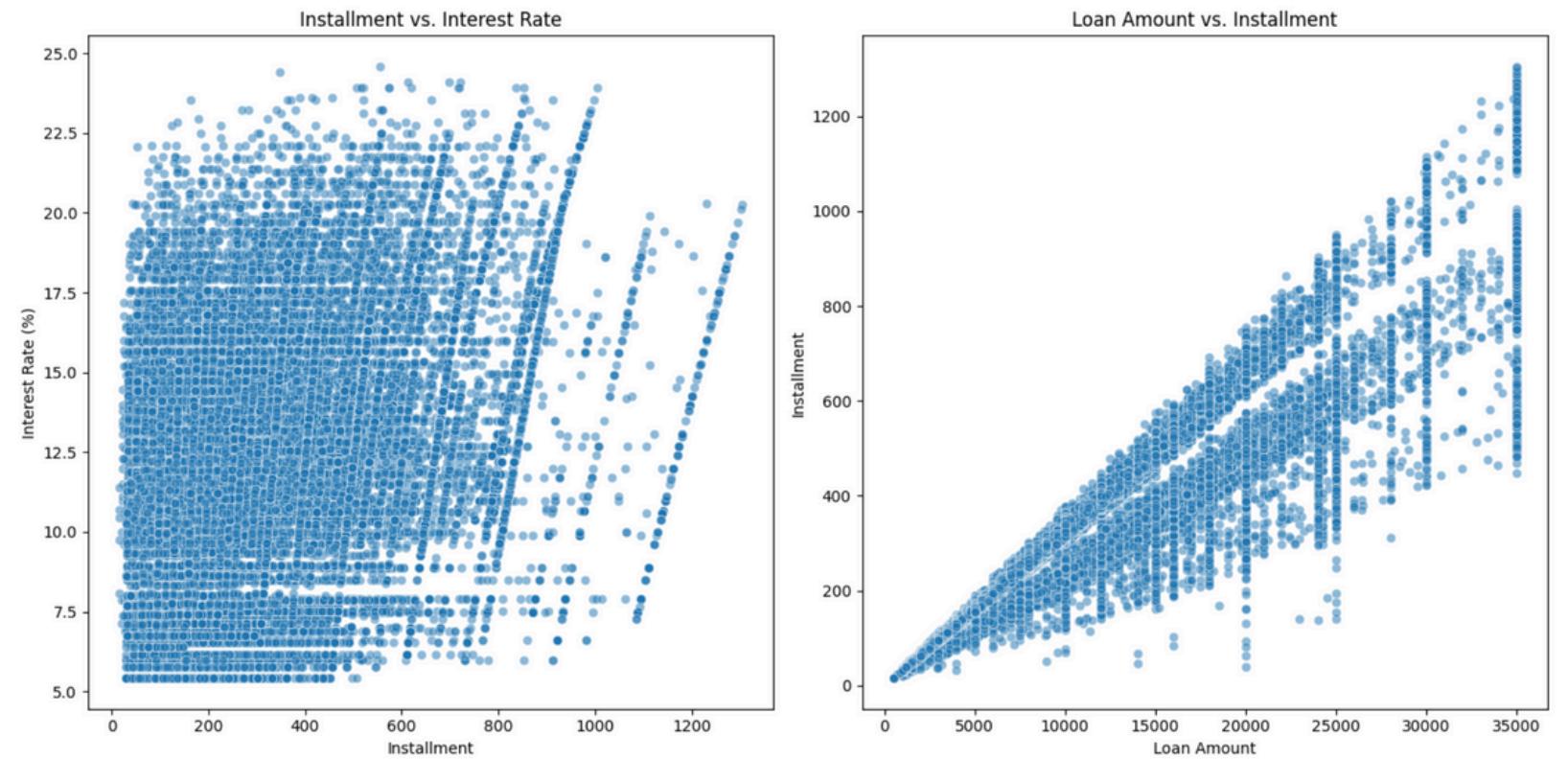




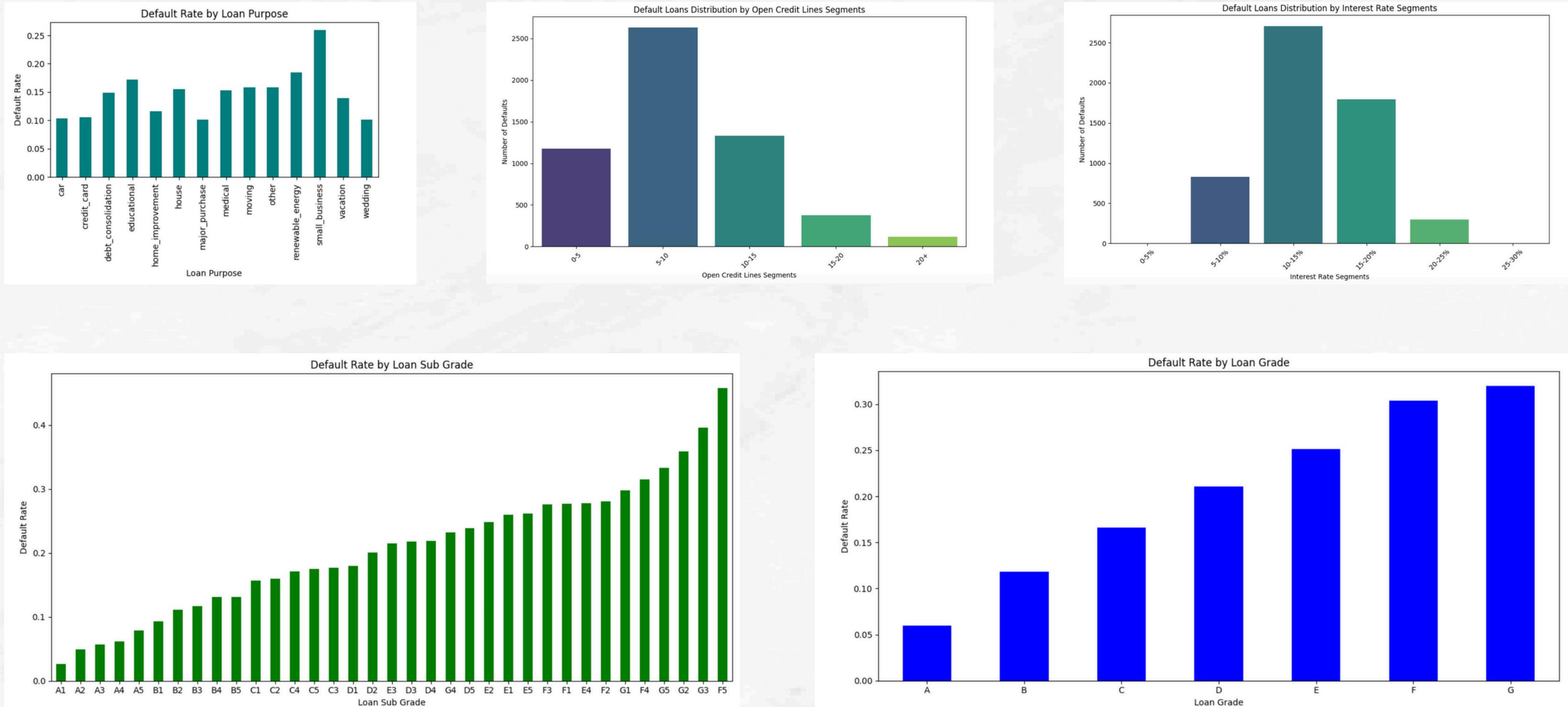
[Back to Agenda](#)

Bivariate Analysis

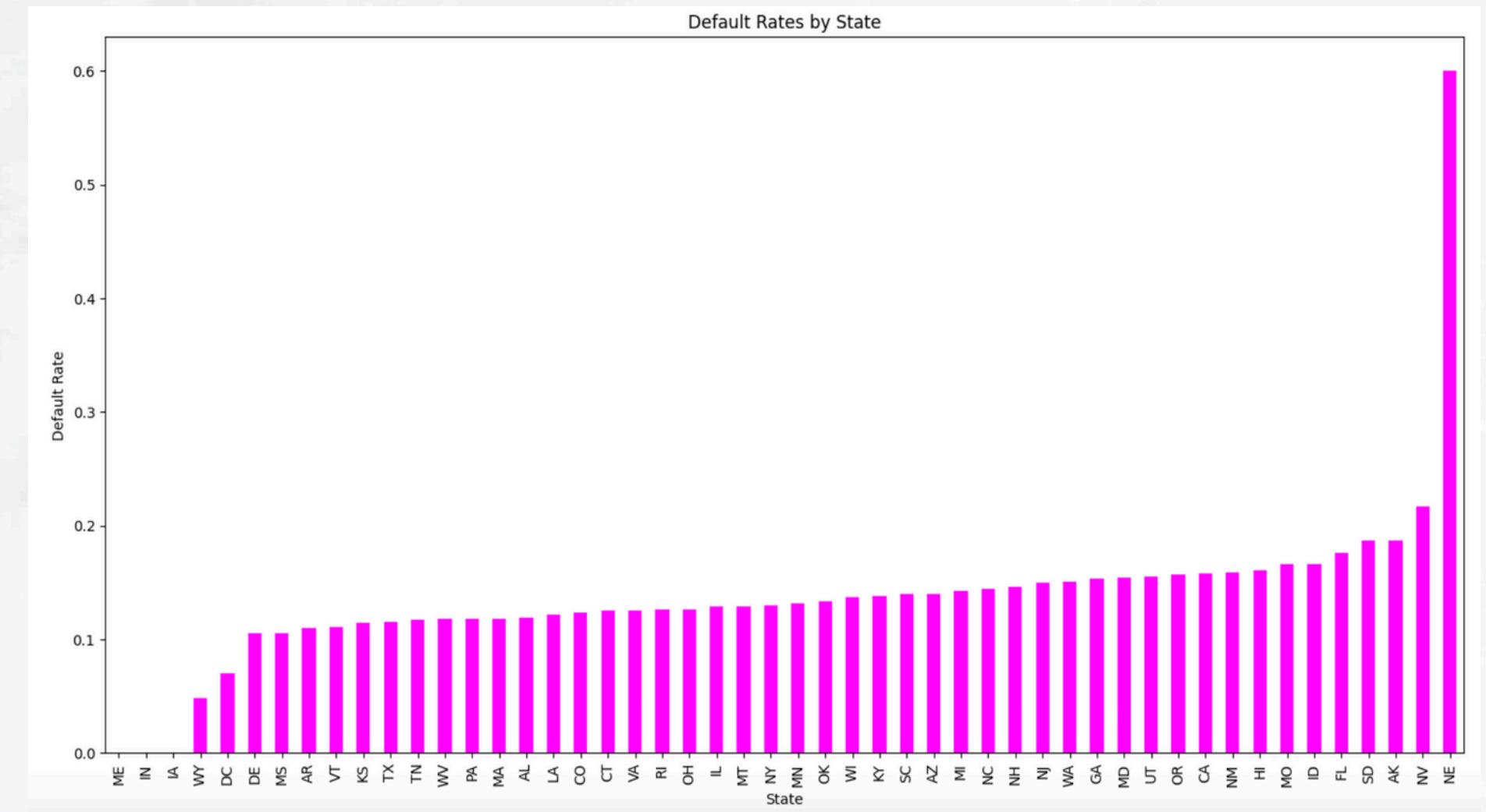
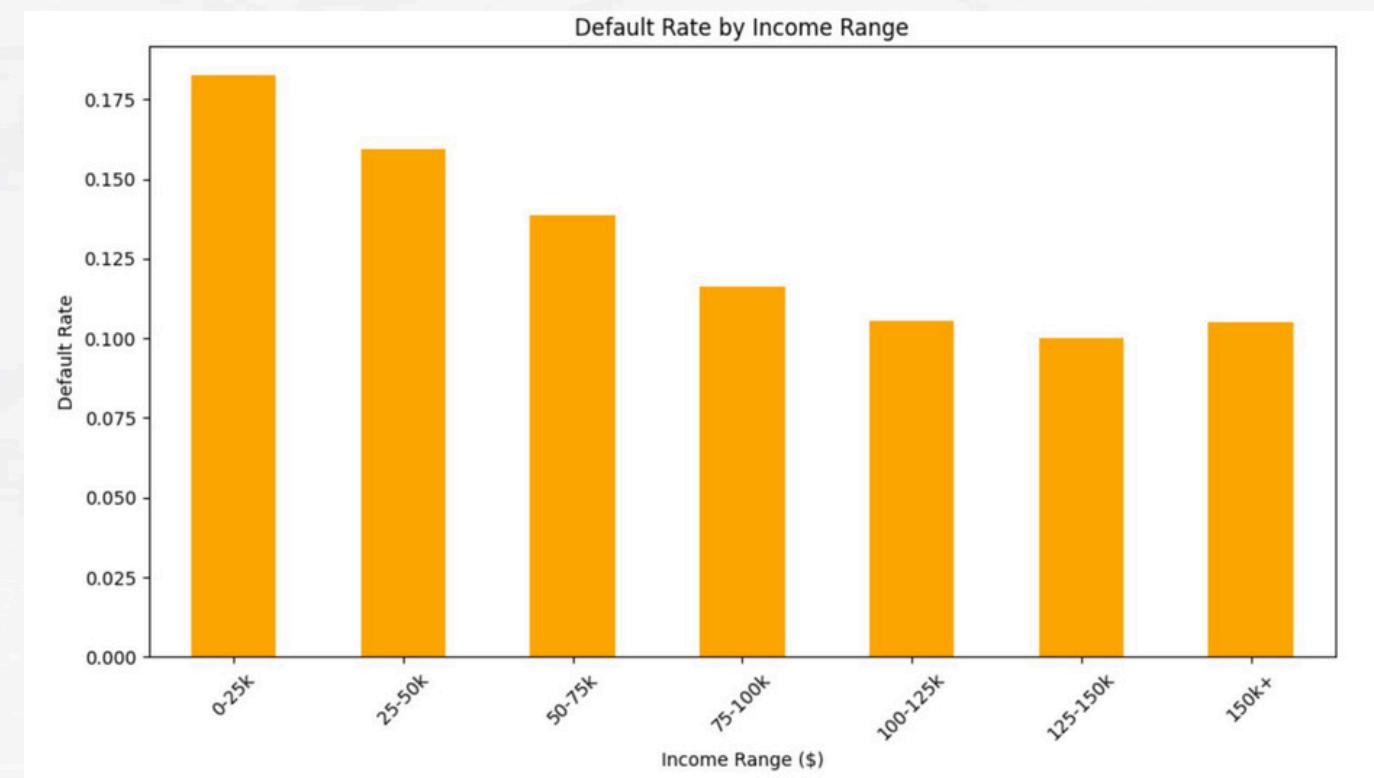
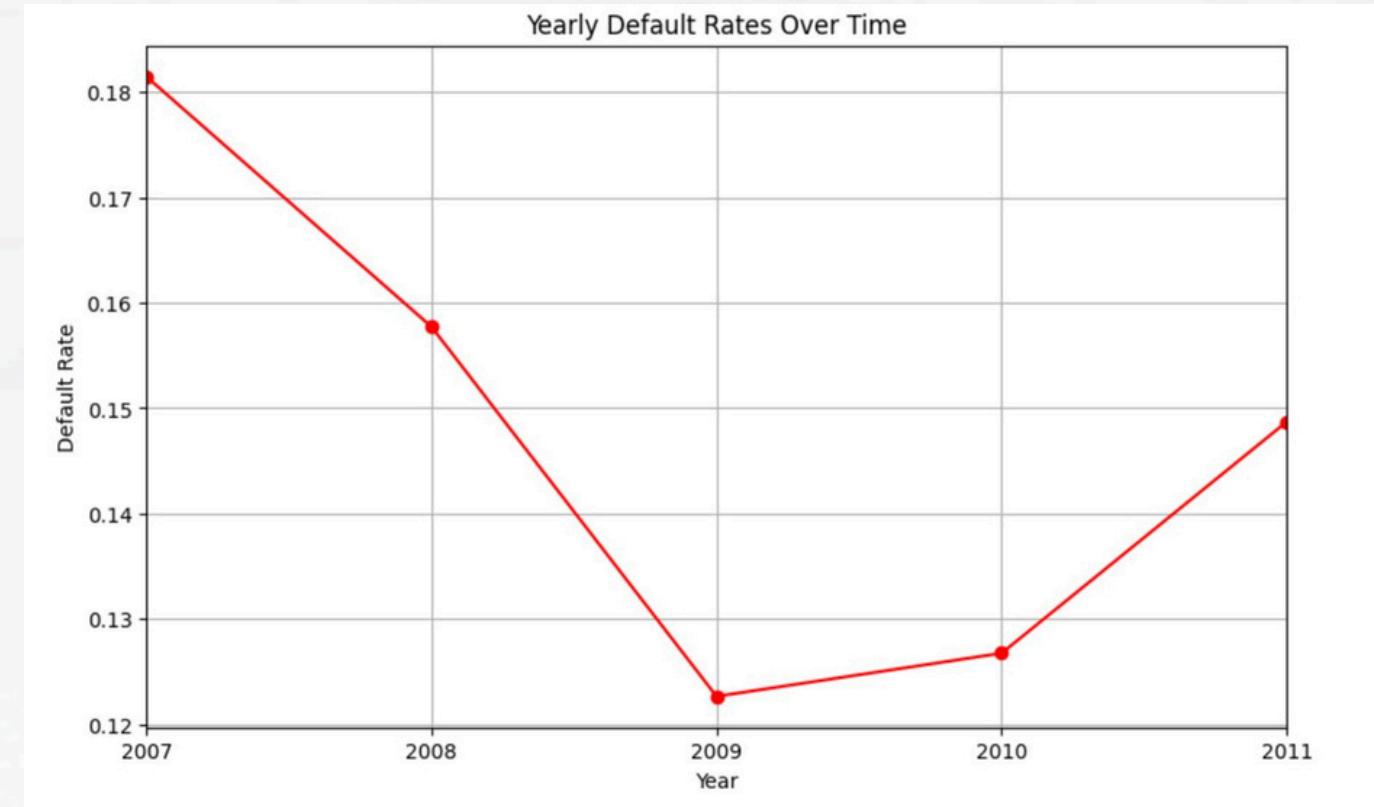
The relationship between loan amount and interest rate revealed a moderate positive correlation, suggesting that higher loan amounts generally come with higher interest rates due to increased risk. This insight is vital for risk-based pricing strategies.



Defaults Analysis



Defaults Analysis



Summary

- Income Distribution Among Borrowers: Shows most borrowers earn under \$100,000, highlighting a middle-income customer base.
- Default Loans Distribution by Interest Rate Segments: Higher default rates in higher interest brackets, especially above 10%.
- Default Loans Distribution by Open Credit Lines Segments: Applicants with number of open credit lines between 5 and 10 tend to show an increasing trend in default rates.
- Average Loan Amount by Income Bracket and State: Heatmap showing higher income brackets typically receive larger loans across different states.
- Employment Length Distribution by State: Heatmap showing diverse employment stability across states and its impact on loan performance.
- Default Rate by Issue Year: Line chart tracking fluctuations in default rates over years, reflecting economic conditions or lending changes.
- Default Rate by Loan Grade: Bar chart showing higher default rates in higher grades, affirming higher risk.
- Default Rate by Loan Sub Grade: Bar chart shows higher default rates in the sub grades of F and G, affirming higher risk.
- Loan Amount vs. Annual Income: Scatter plot showing positive correlation between income and loan amounts.
- Average Loan Amount & Interest Rate Distribution by Year: Line chart indicating a general increase in loan amounts over time.
- Default Rate by Employment Length: Shows that longer employment does not necessarily correlate with lower default rates.
- Default Rates by State: Bar chart identifying states with higher financial risk based on default rates.
- Financial Maturity and Default Trends: Bar chart showing lower default rates among borrowers with longer credit histories.
- Default Rate by Financial Maturity: Bar chart indicating that longer financial maturity correlates with lower default rates.
- Income vs. Loan Amount: Scatter plot supporting that higher incomes correlate with larger loan amounts.