Identifying Loan Default Risk Using Exploratory Data Analysis (EDA)

Members:
Maruthi Komera
Shriya Magadal

8/28/2024

Problem Statement:

As an analyst at Lending Club, explore the dataset of past loan applicants through Exploratory Data Analysis (EDA) to gain insights into how applicants and loan attributes impact the possibilities of default.

Business Cases:

Lending companies encounter two main types of risks:

- Loan Accepted: There are three possible outcomes—fully paid, currently active, or charged-off (defaulted).

- Loan Rejected: The dataset lacks transactional history for applicants whose loans were rejected.

Business Objective:

The goal is to reduce credit losses by identifying high-risk applicants and potential defaulters through EDA. The emphasis is on understanding the key factors that lead to loan defaults and mitigating financial losses.

Loan Approval Scenarios

When a loan is approved, there are three possible outcomes:

- **Fully Paid**: The loan has been completely repaid.

- **Current**: The loan is still in the process of being repaid.

- **Charged-Off**: The loan has defaulted.

# Data Understanding

Dataset Overview:

- **Key Variables**: Includes int_rate, dti, loan_amnt, grade, annual_inc.

- **Target Variable**: loan_status, which indicates whether a loan is fully paid or charged-off.

Problem Approach

Through EDA, we will analyze the dataset to gain insights into:

- Consumer attributes.

- Loan characteristics that impact the risk of default.
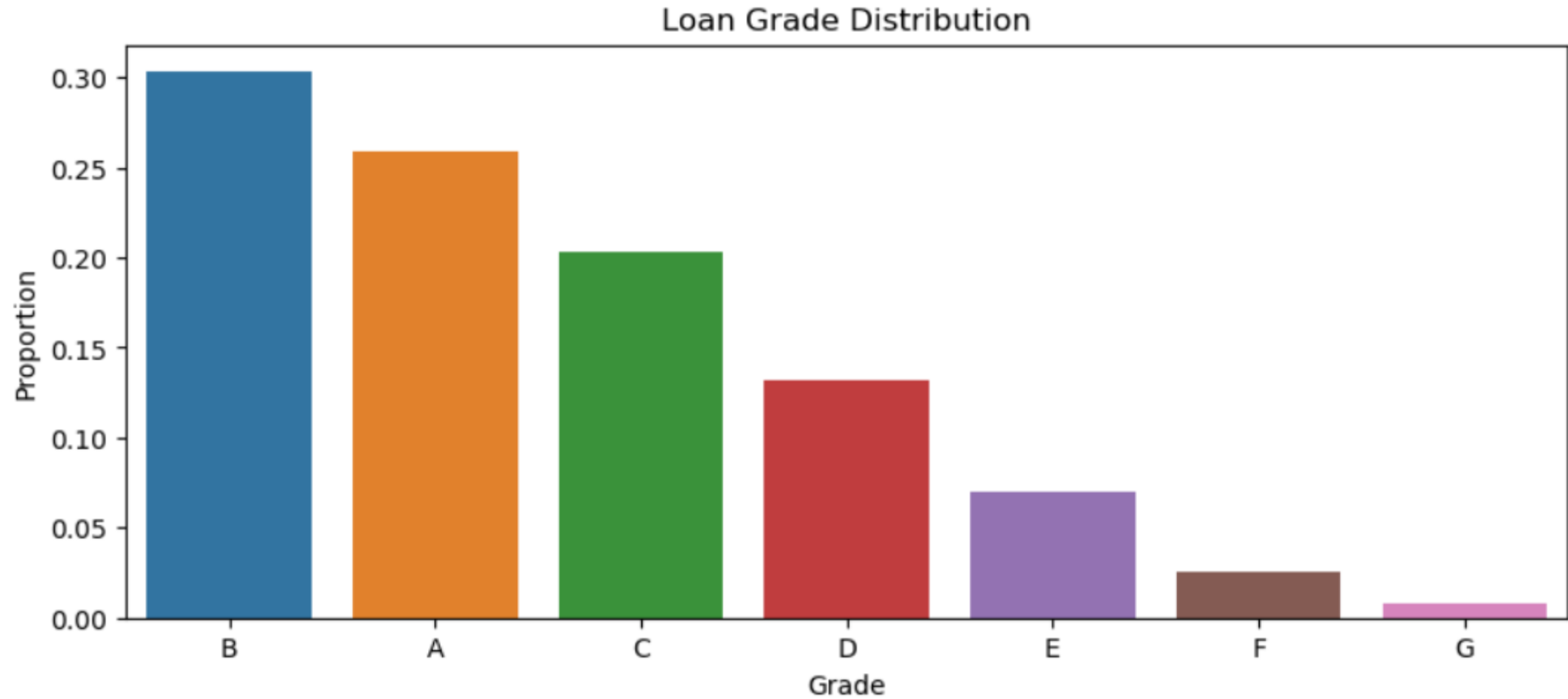
Loan Amount Distribution
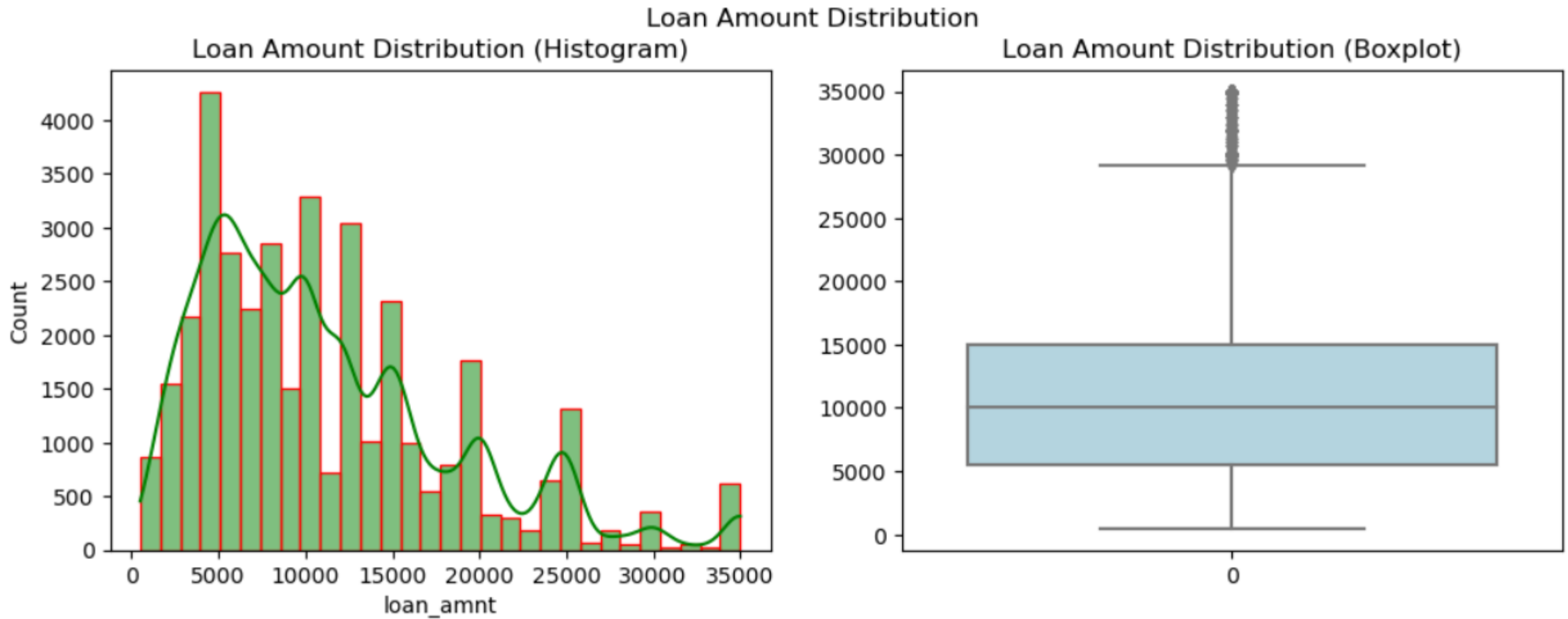
Distribution of the loan amount using Histogram and Boxplot

Analysis Of Issued Year, Month and Week Distribution

Lending Club (LC) has been increasing its loan issuance annually. Moreover, there is a surge in loan issuances during the final quarter of the year, specifically in October, November, and December.
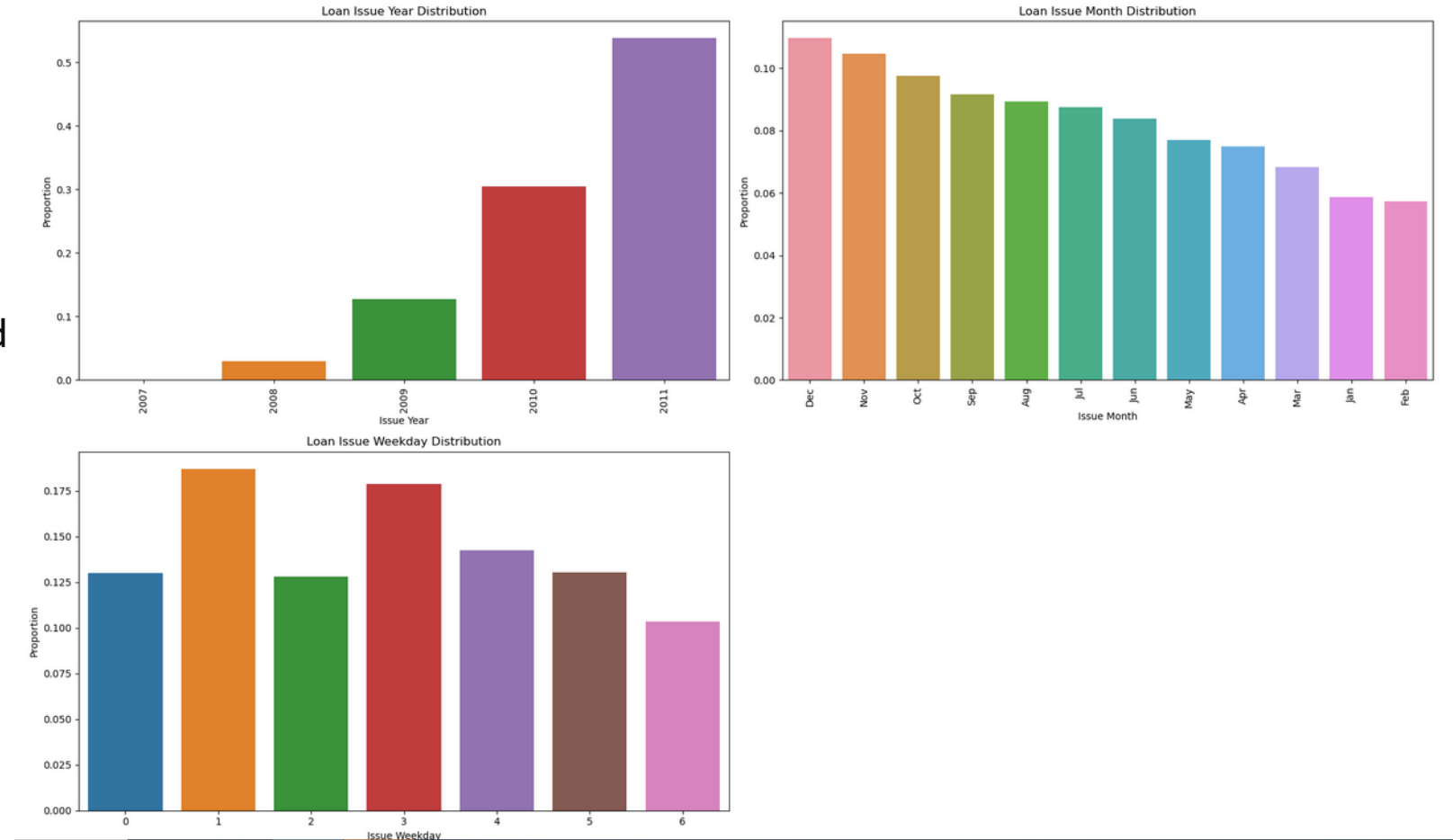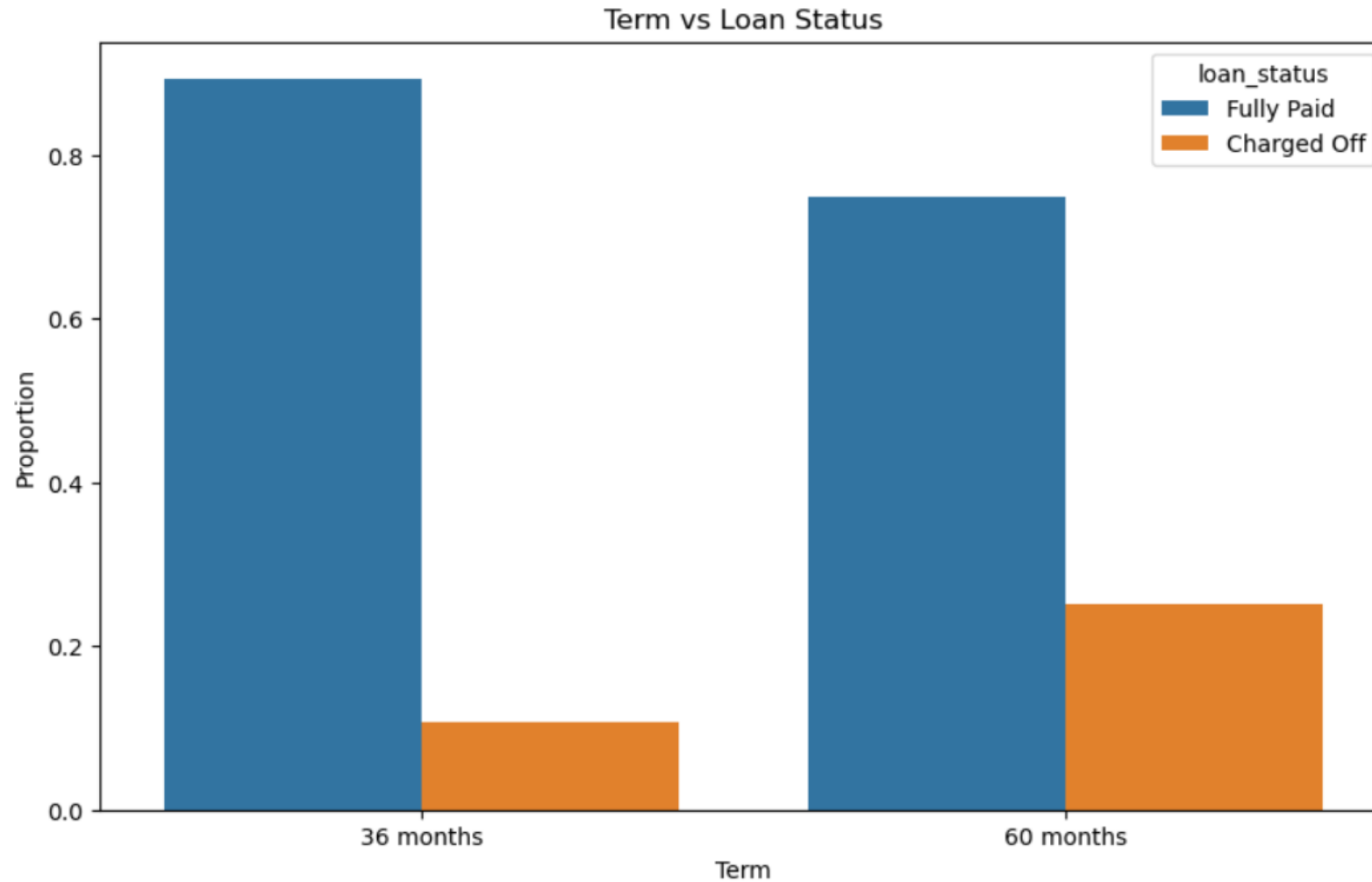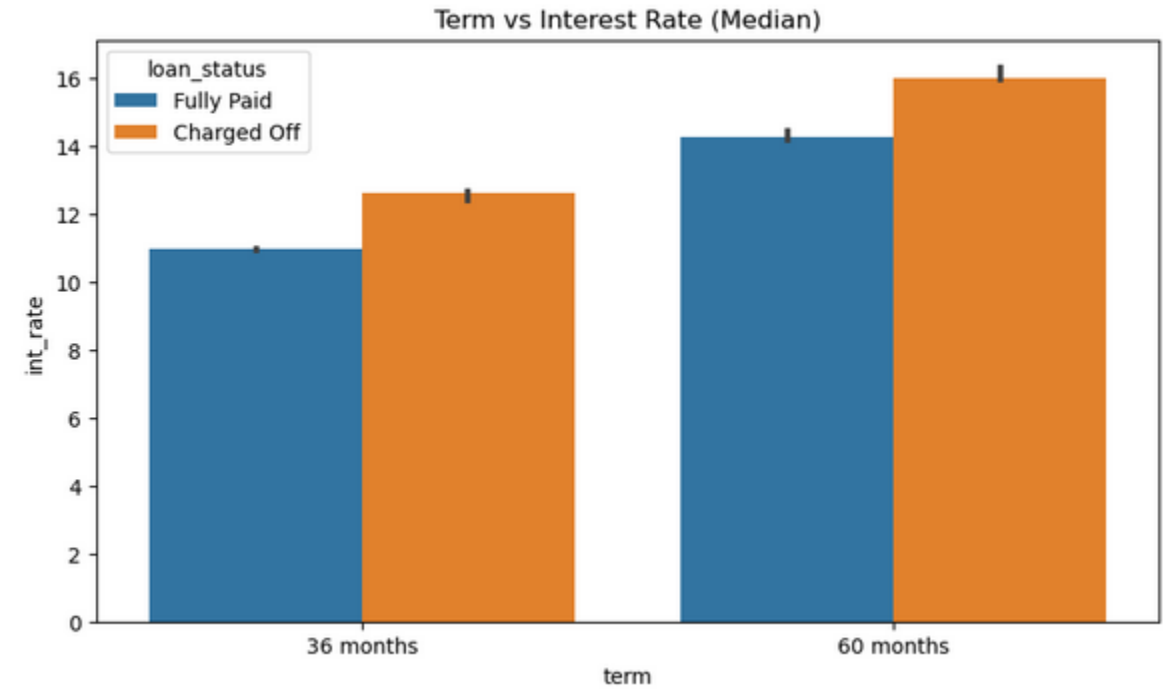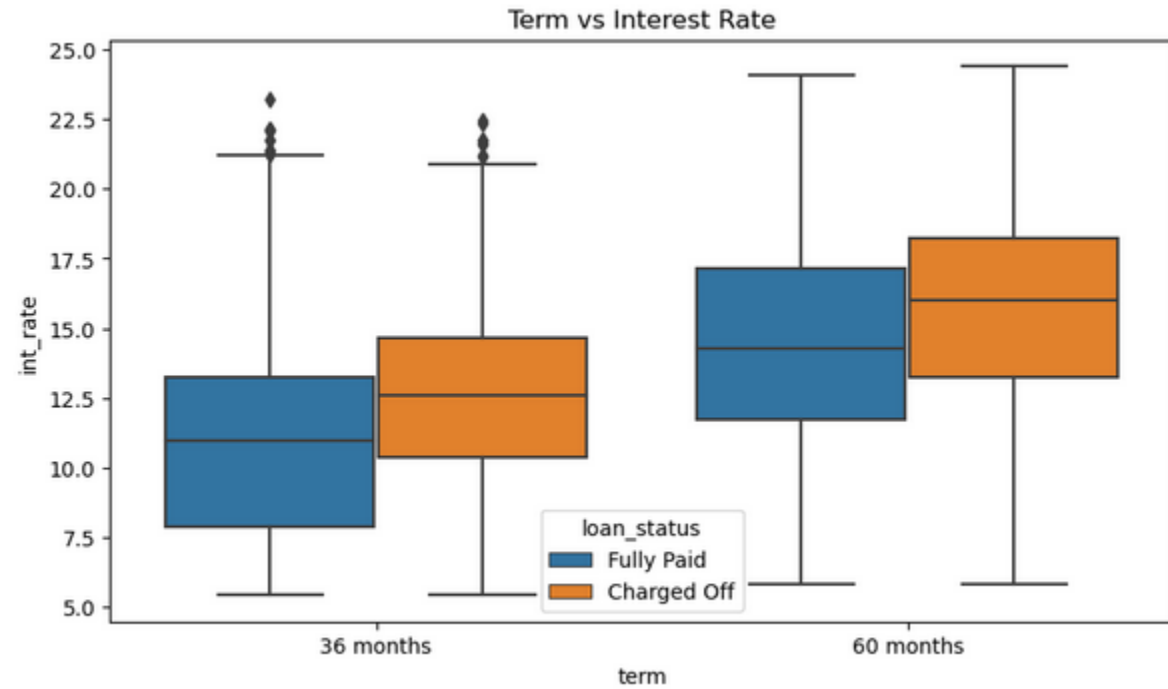
Loan Grade Distribution

Most of the loans are classified as Grade B and A.

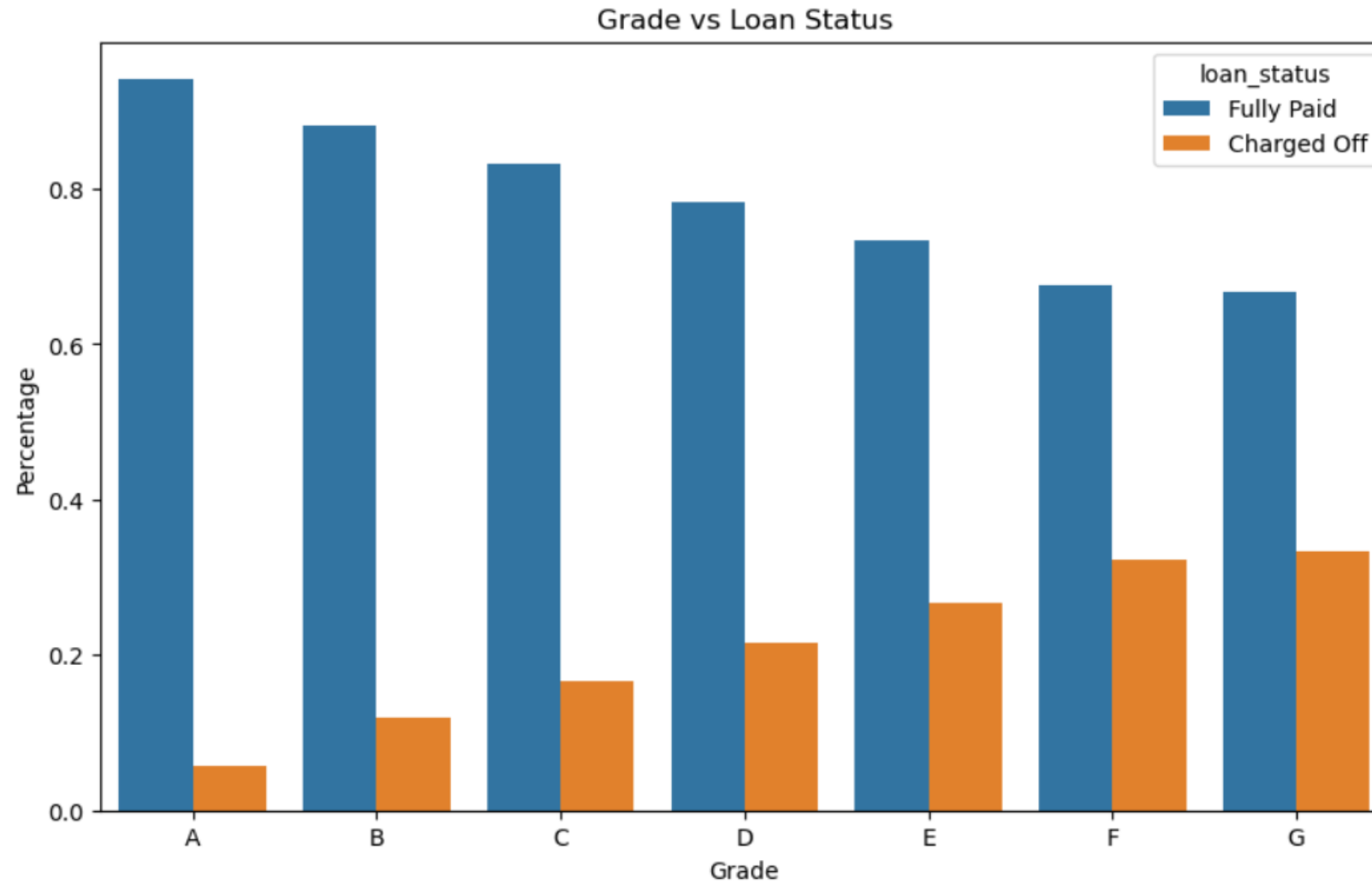# Analysis of Term Vs Loan Status

Approximately 75% of borrowers opted for loans with a 36-month term.

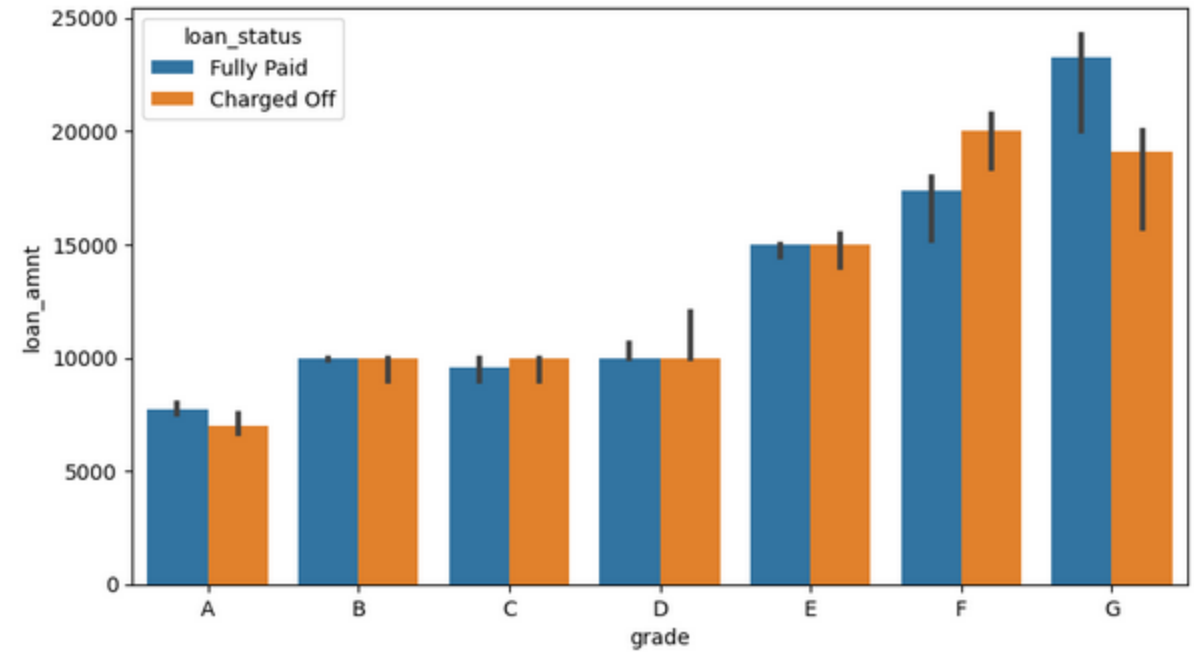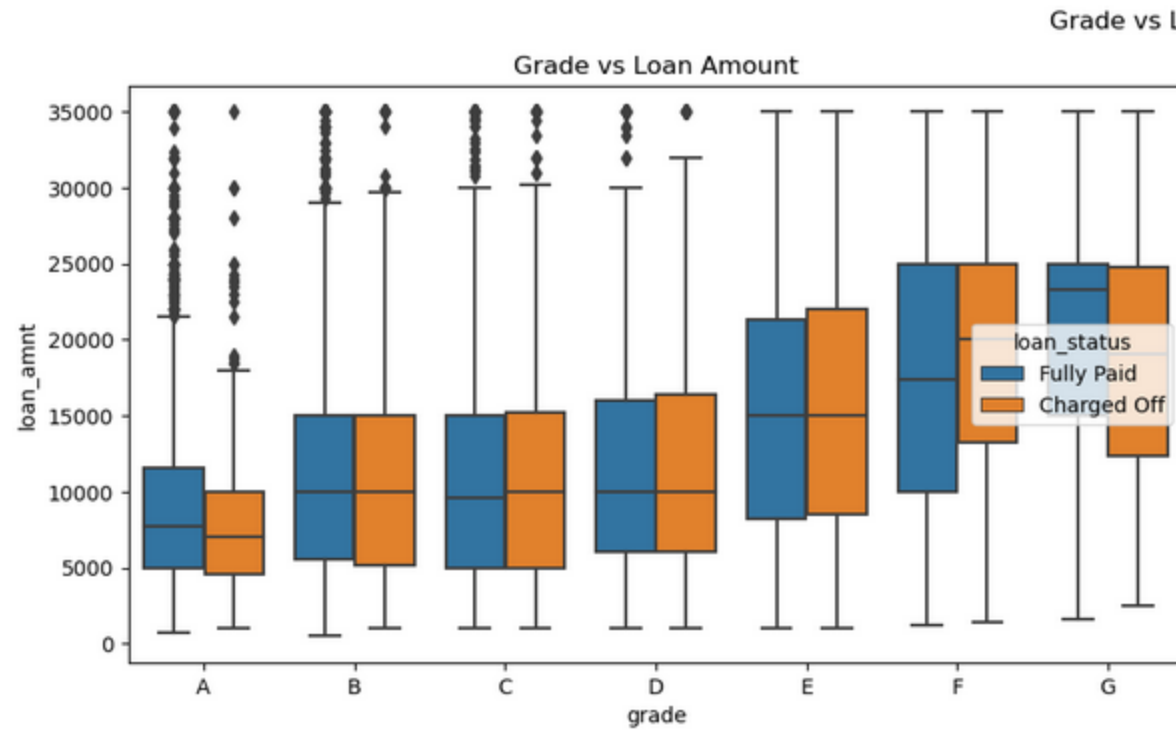# Analysis Of Loan Status Vs Interest rate

The interest rate rises as the loan term increases.

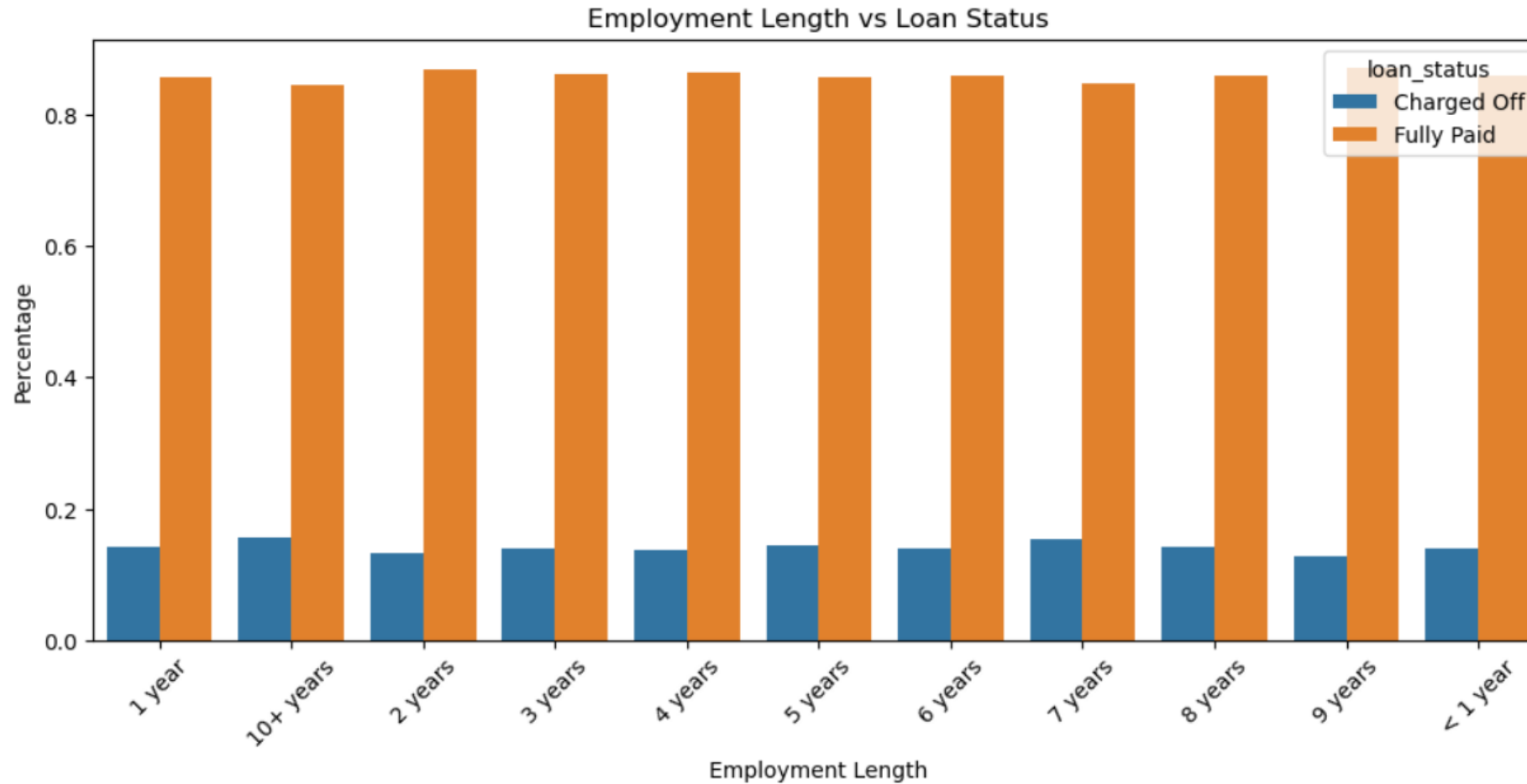## Analysis Of Loan Status Vs Grades

Number of charged-off loans increases as credit grades decrease.

Analysis Of Grade Vs Loan Amount

Loans are more often charged off in lower credit grades, where borrowers usually take bigger loans and are more likely to default.

Analysis Of Employment Years Vs Loan Status

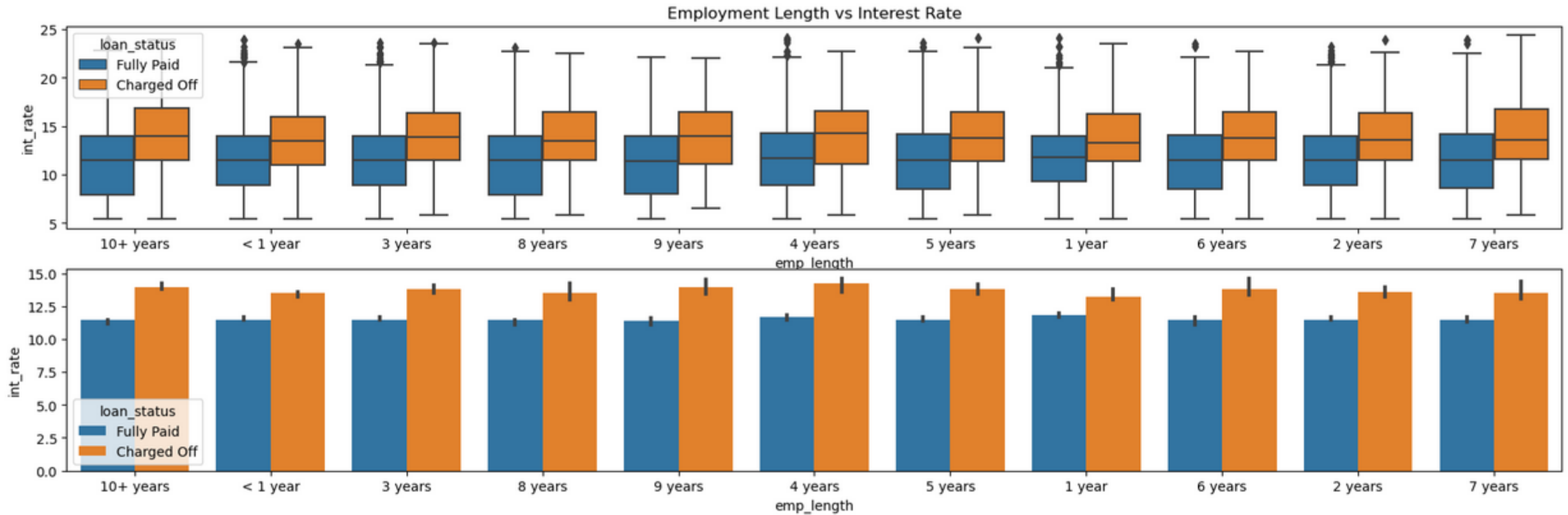No significant patterns in default rates across different employment lengths.

Analysis Of Employment Years Vs Loan Amount

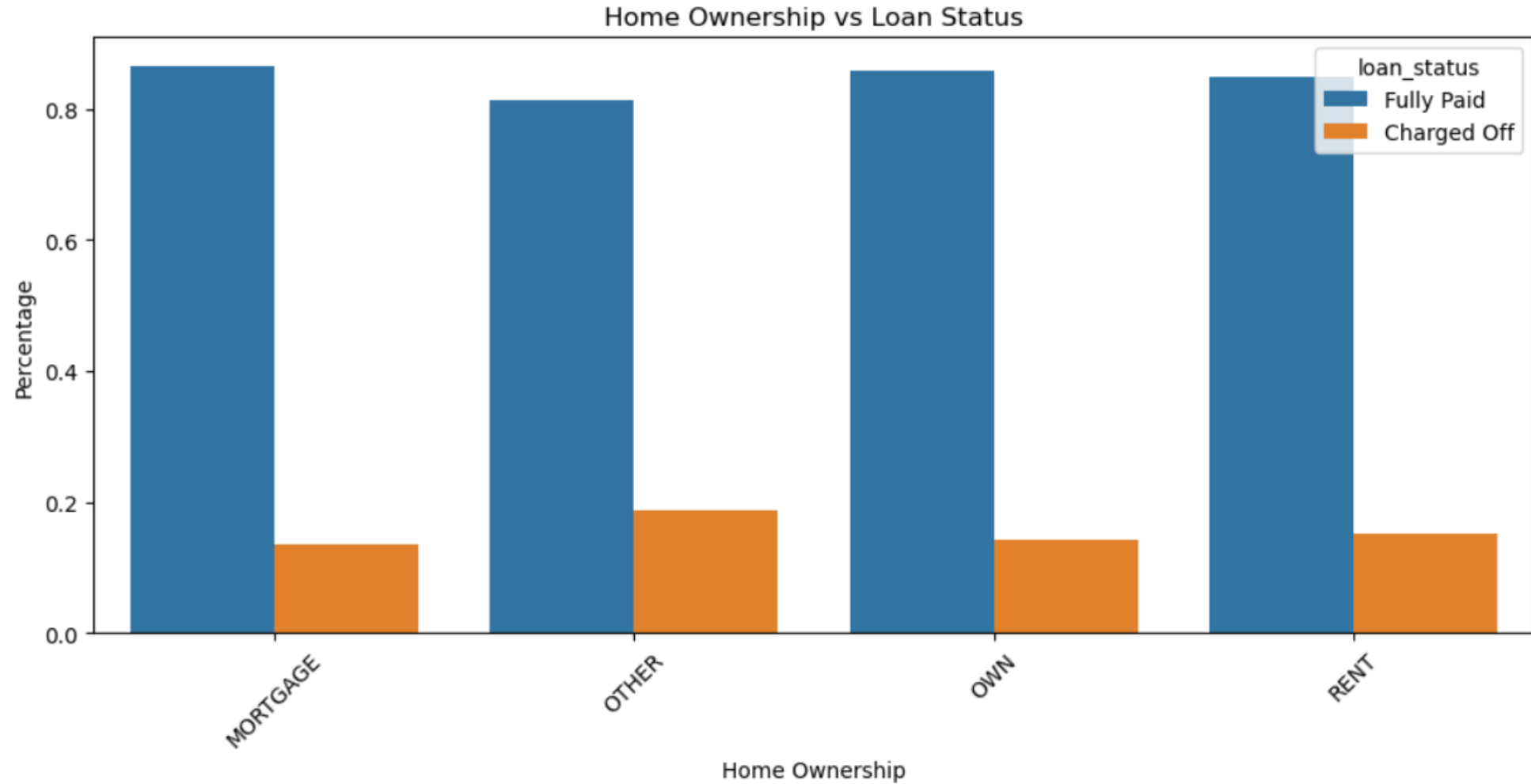Employment Length vs Loan Amount
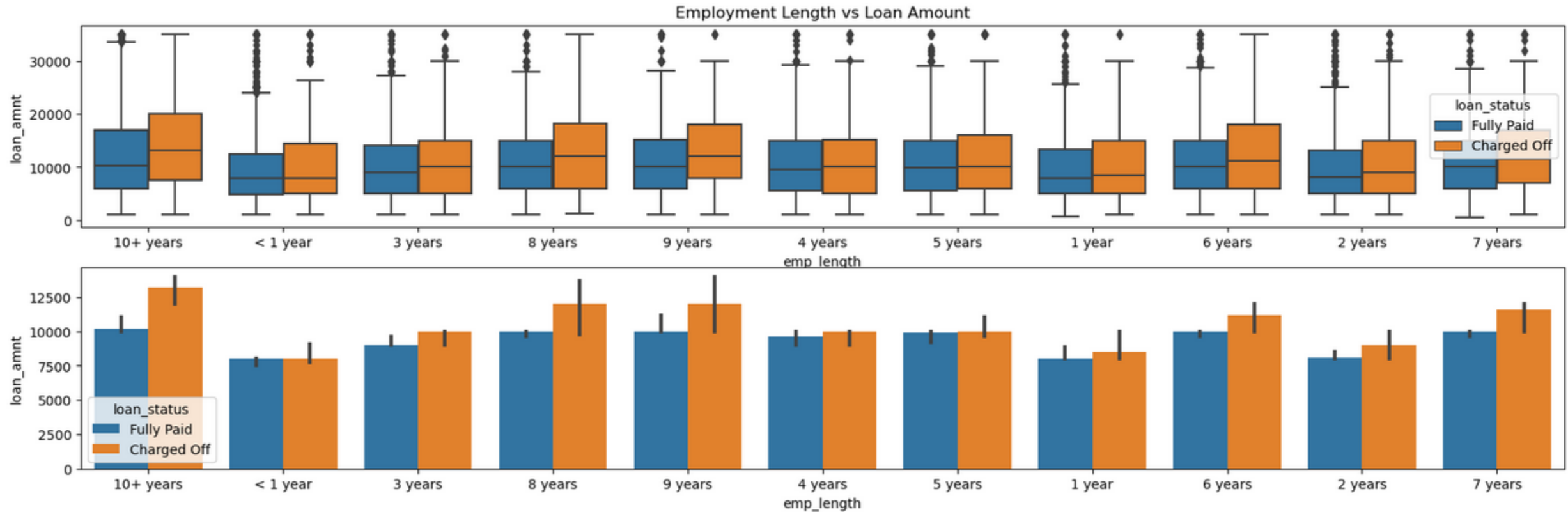
Borrowers with more job years and larger loan amounts typically experience higher default rates

Analysis Of Employment Years Vs Interest Rate

Regardless of employment length, loans with higher interest rates tend to have higher default rates.

Analysis Of Home Ownership Vs Loan Status

Slightly higher percentage of defaults in the "Other" home ownership category.

EDA Insights:

Interest Rates and Defaults:

- Higher interest rates are strongly linked to increased default rates, regardless of loan duration or verification status.

- Borrowers with high revolving credit utilization face higher interest rates and tend to default more often.

Loan Amount and Defaults:

- Larger loan amounts, especially in small business and home loan categories, are associated with higher default rates.

- High loan amounts in small business loans have a strong correlation with defaults.

Credit Grades and Defaults:

- Lower credit grades (F and G) are associated with higher default rates and larger loan amounts, with borrowers in these grades having more public derogatory records.

- A-grade borrowers have fewer derogatory records and a lower default rate.

Home Ownership and DTI:

- Borrowers with mortgages have higher incomes but also larger loan amounts, which correlate with higher default rates.

- Other home ownership categories have a lower DTI ratio but a similar likelihood of default compared to other categories.

Recommendations:

Refine Home Ownership Criteria:

- Modify lending criteria for various home ownership categories based on observed default rates and DTI ratios.

- Consider creating customized financial products to match the risk profiles of borrowers in different home ownership categories.

Improve Verification Processes:

- Enhance verification procedures to ensure loans are issued to verified borrowers

- Explore ways to improve verification for categories with higher default rates, like the "Not Verified" group.

Monitor and Analyze Trends:

- Continuously track default rates and the effects of interest rates across different loan types, credit grades, and home ownership categories.