Lending Club Case Study

Presented by

Tejashwini

Guru Vamsi Krishna Uppalapati

Problem Overview

The data Provided contains information about past loan applicants and whether they 'defaulted' or not. The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.
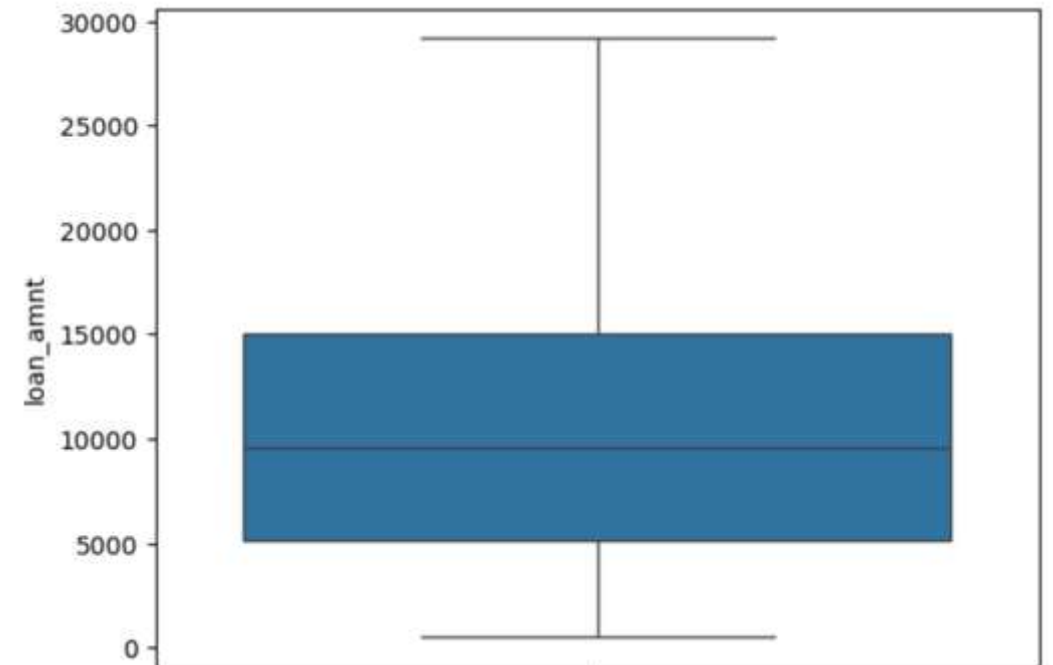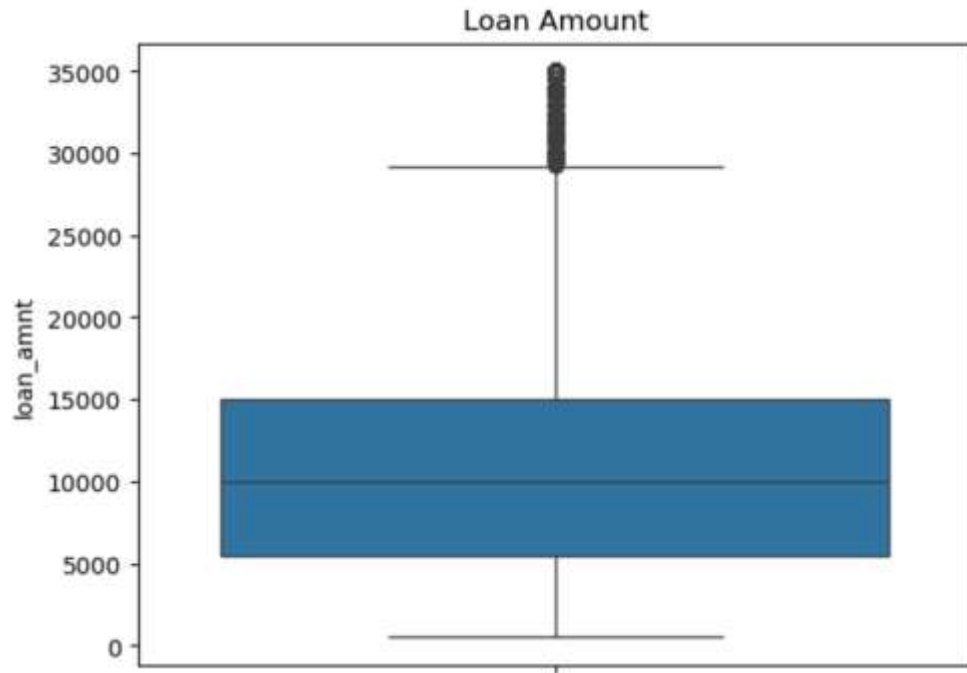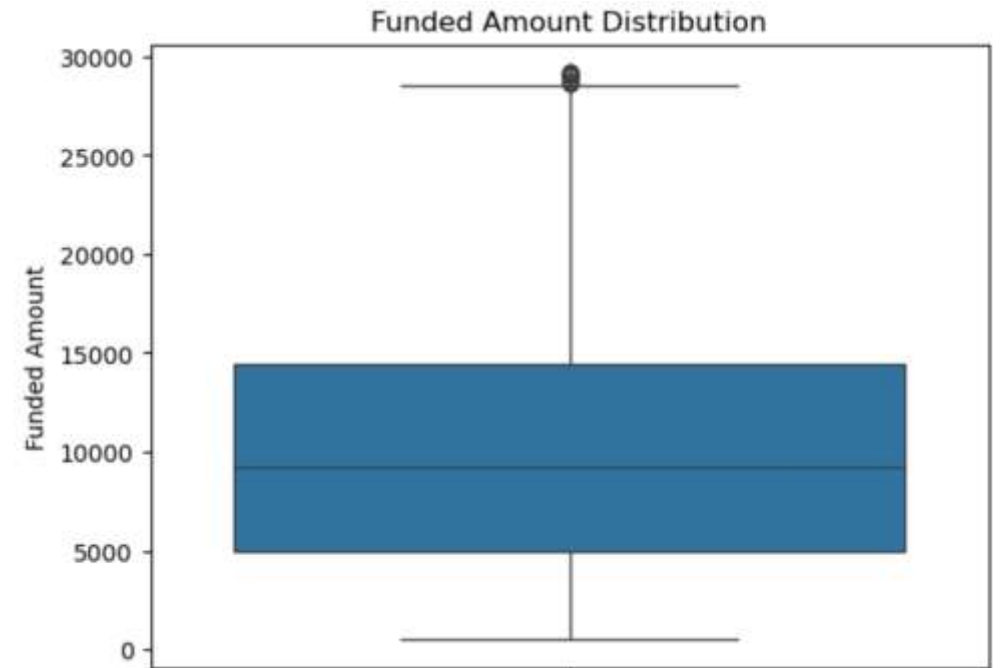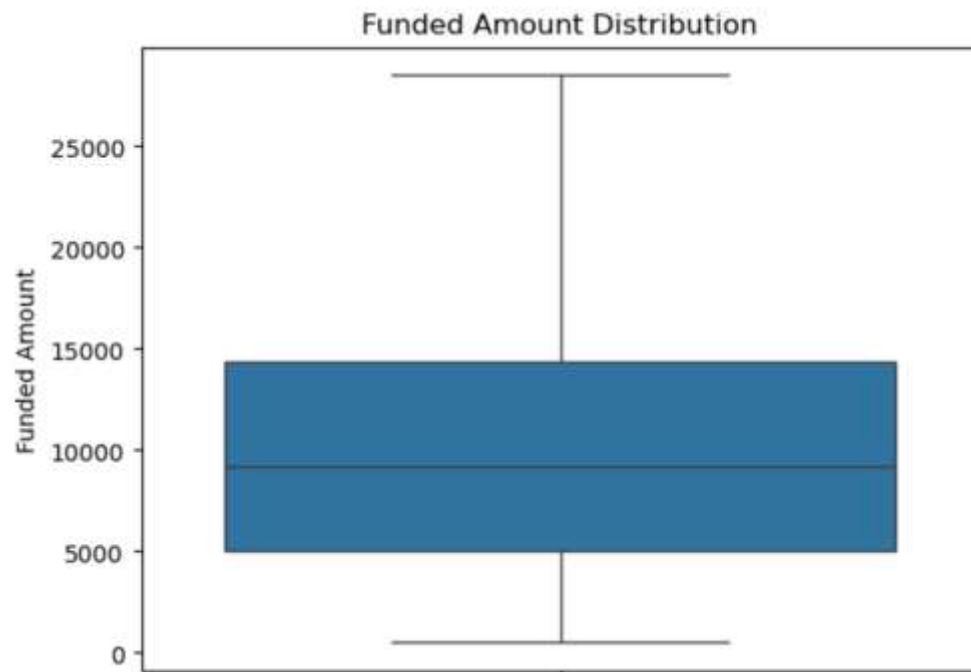
Approach

- Data Cleaning.

- Outlier Treatment.

- Univariate Analysis.

- Bi-Variate Analysis.

- Identifying driving factors behind loan default, i.e. the variables which are strong indicators of default with correlation matrix.
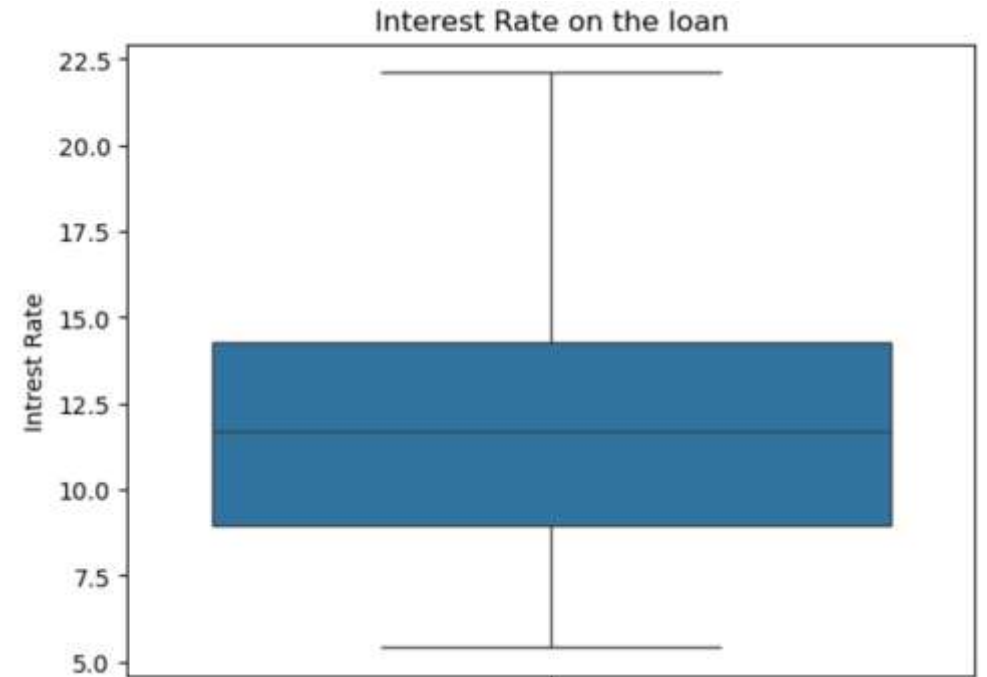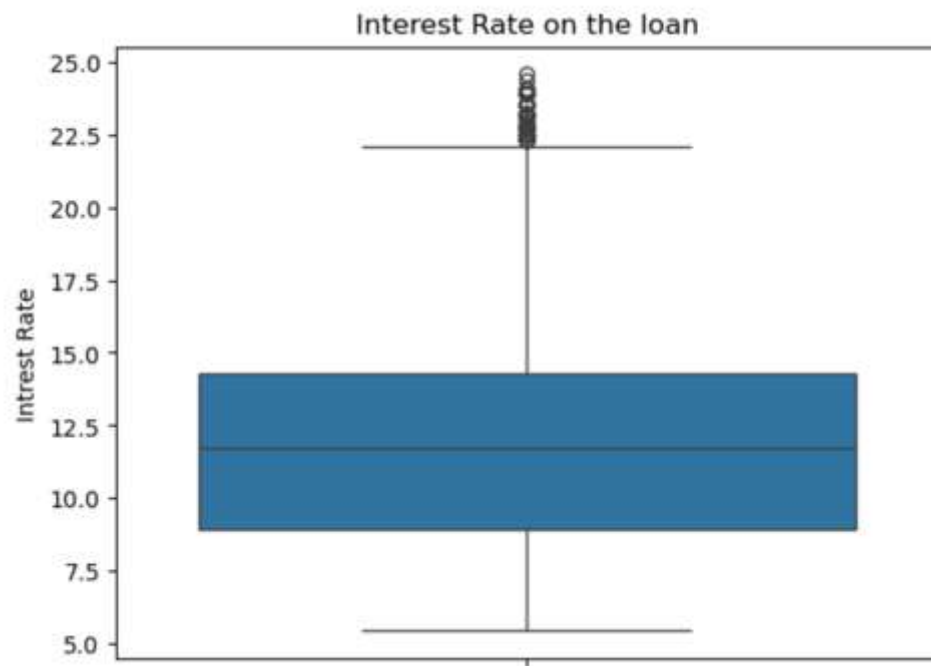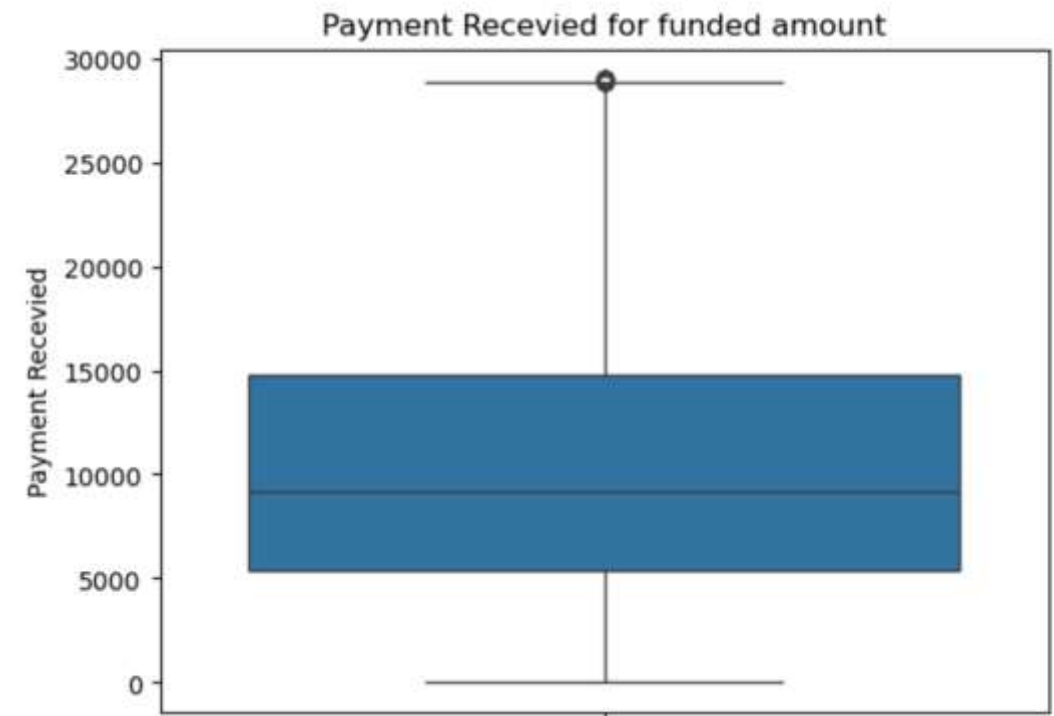
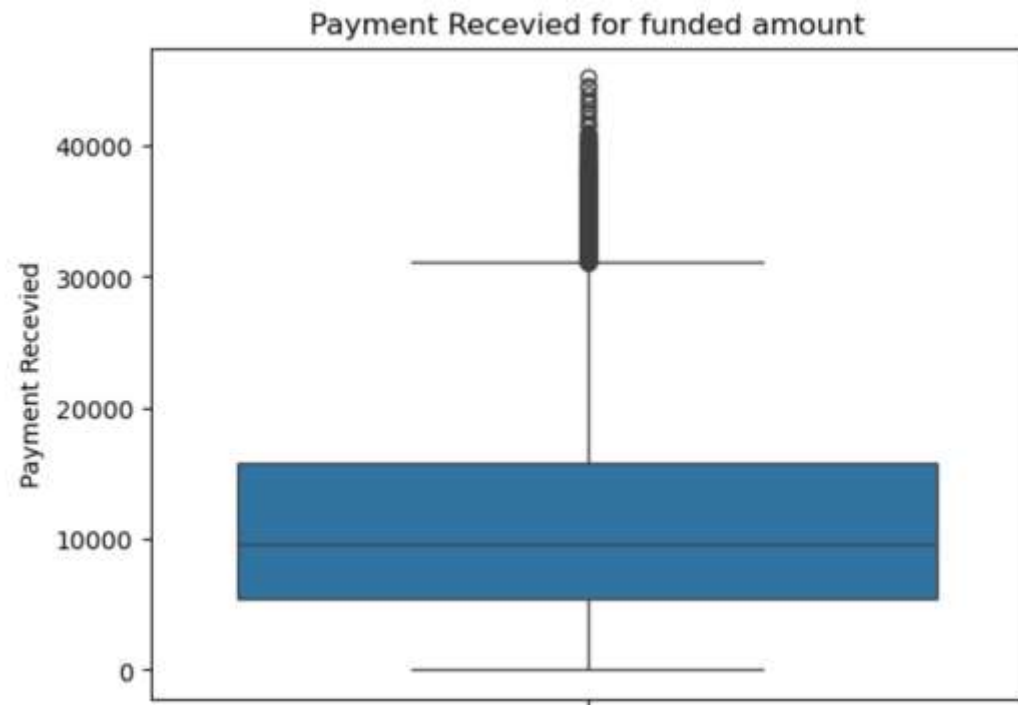# Outlier Treatment – Loan Amount

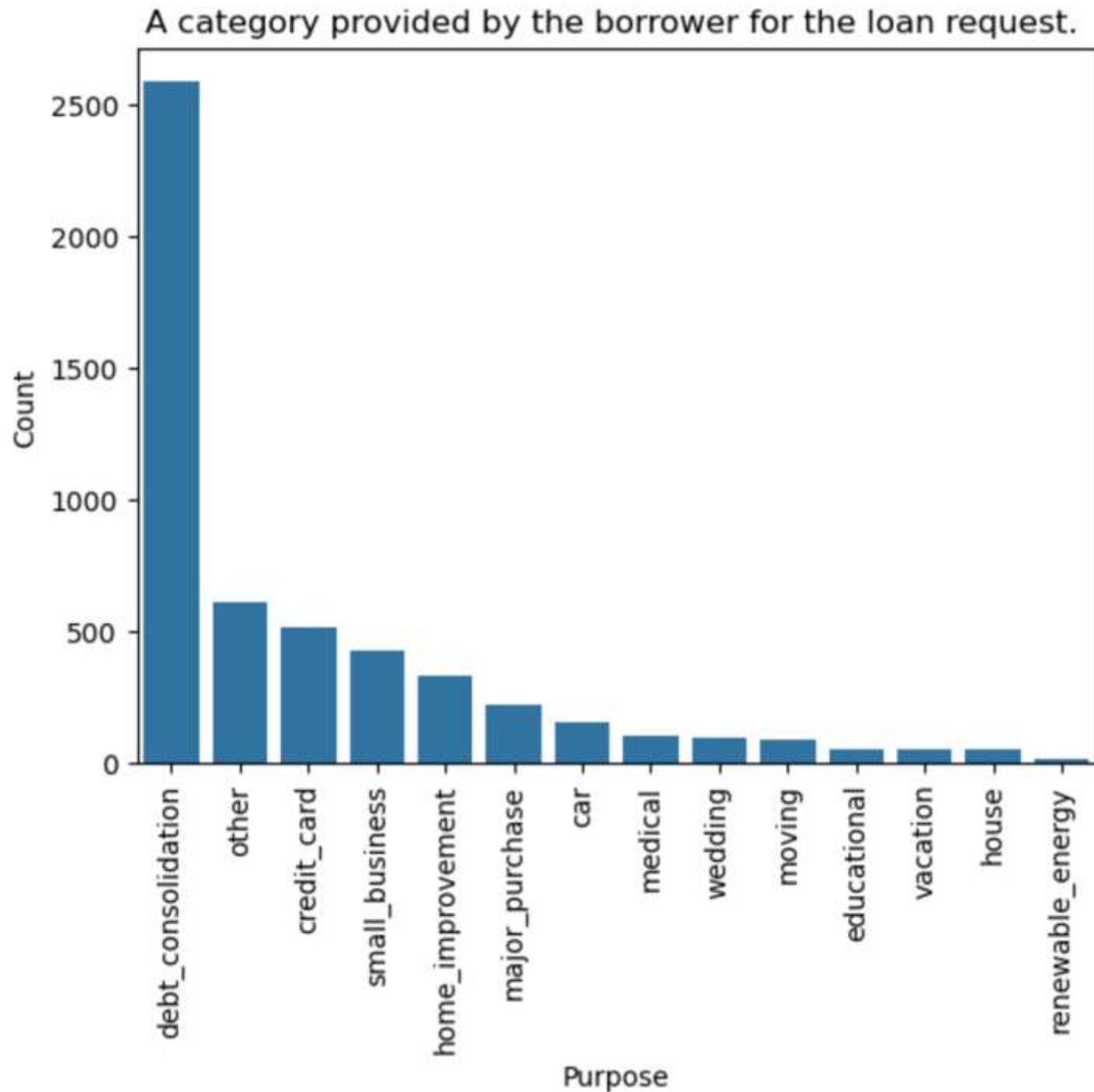# Funded Amount

Interest Rate

Total Payment : Payments received to date for total amount funded

Univariate Analysis

- Category provided by the borrower for the loan request.
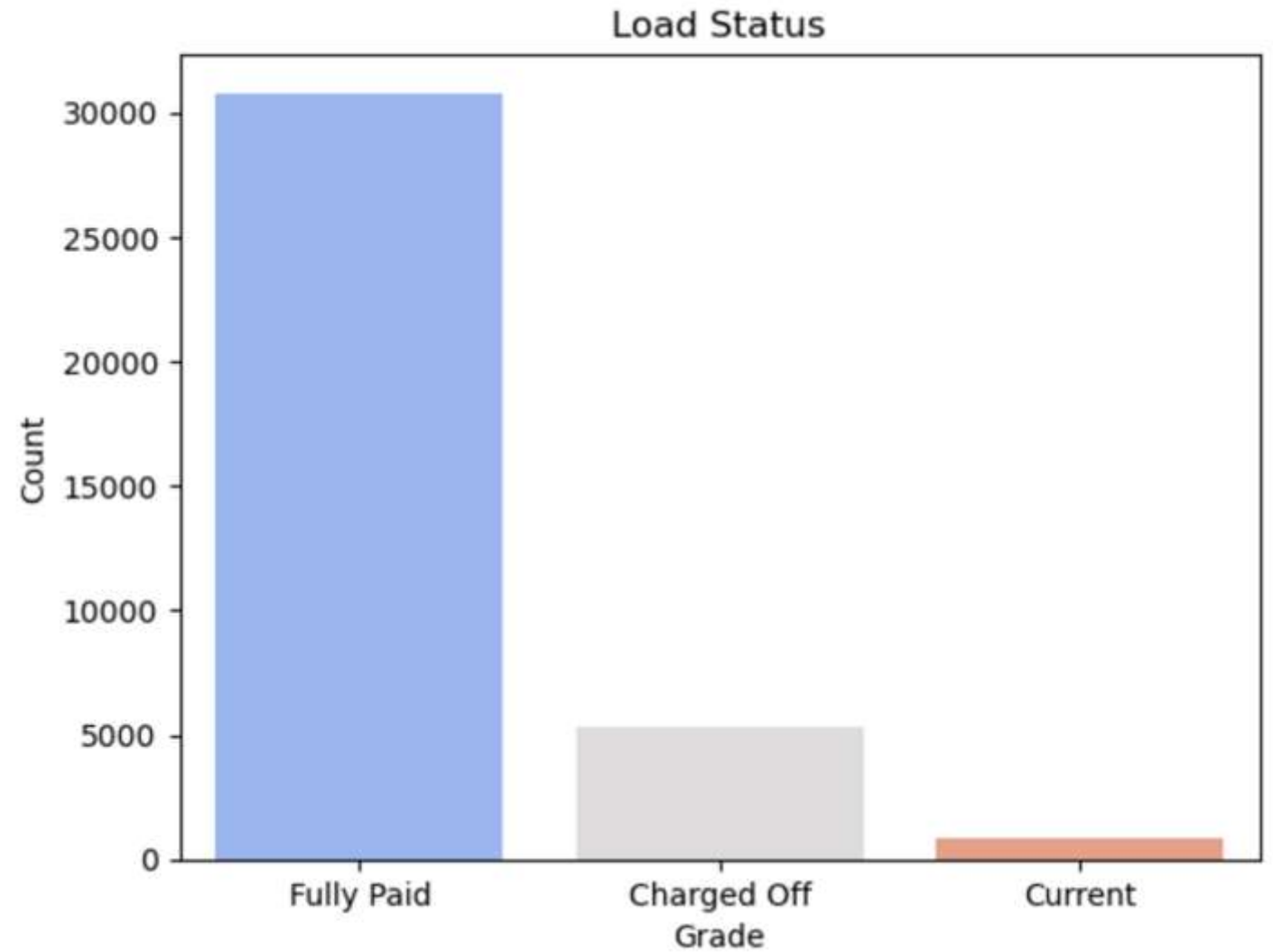

A category provided by the borrower for the loan request.

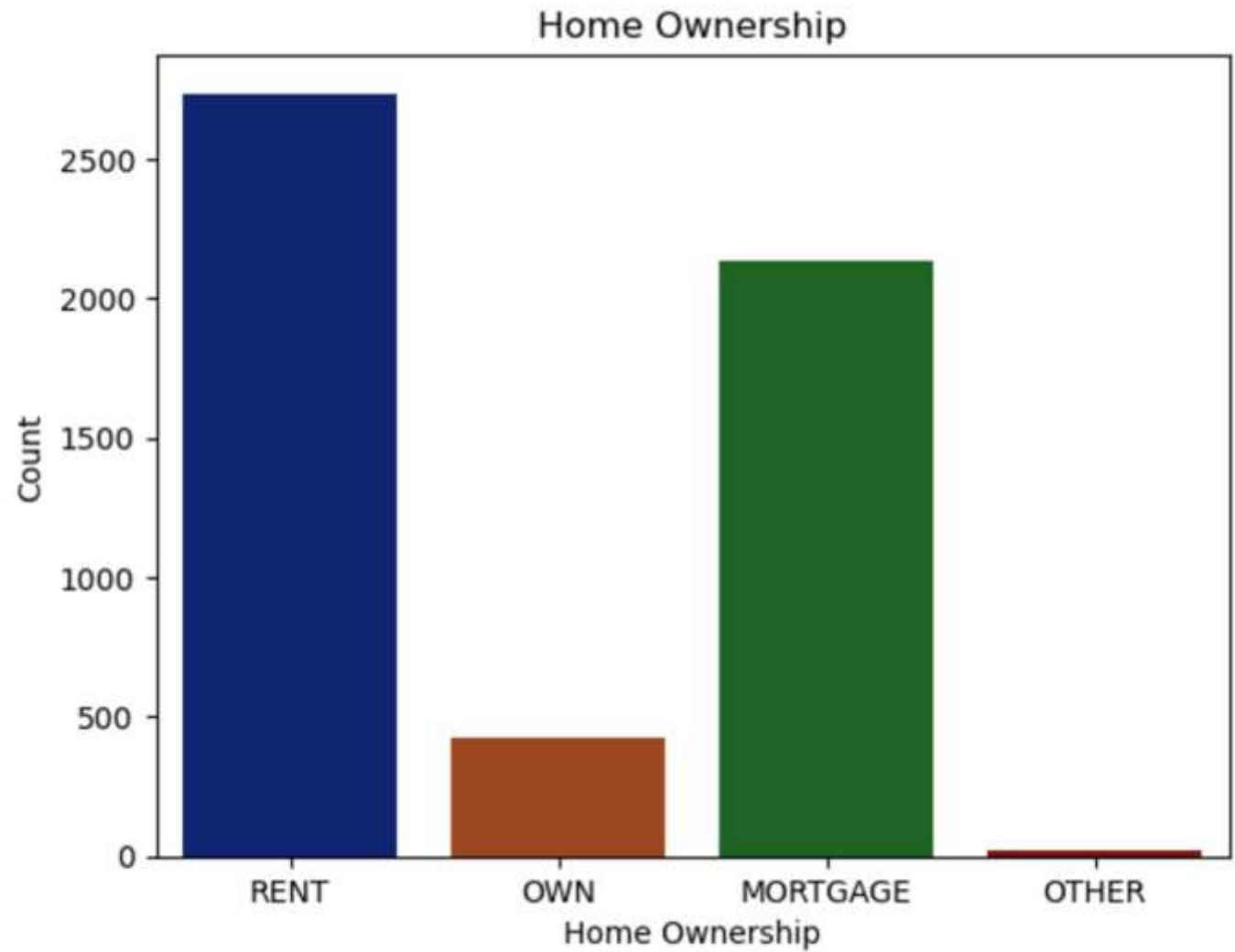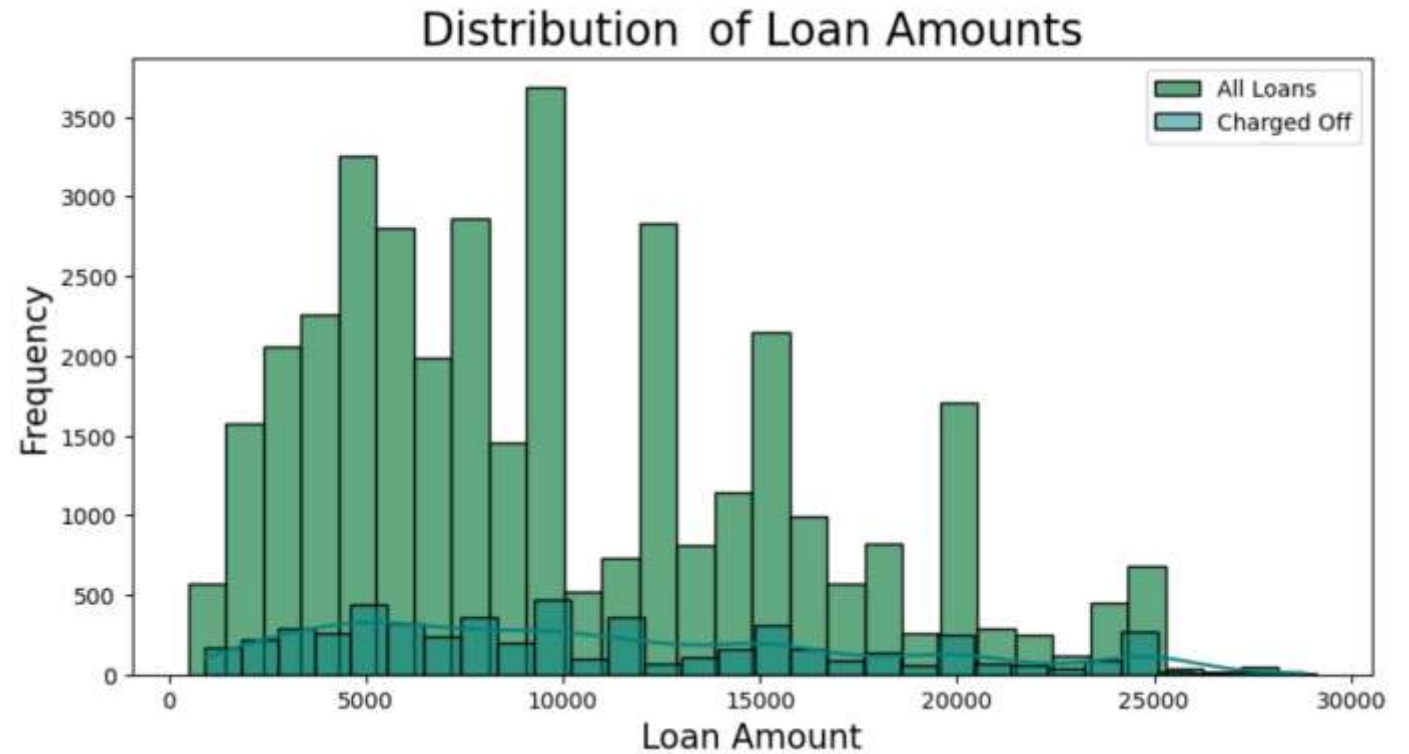# Distribution of Loan Status

Distribution of Home Ownership
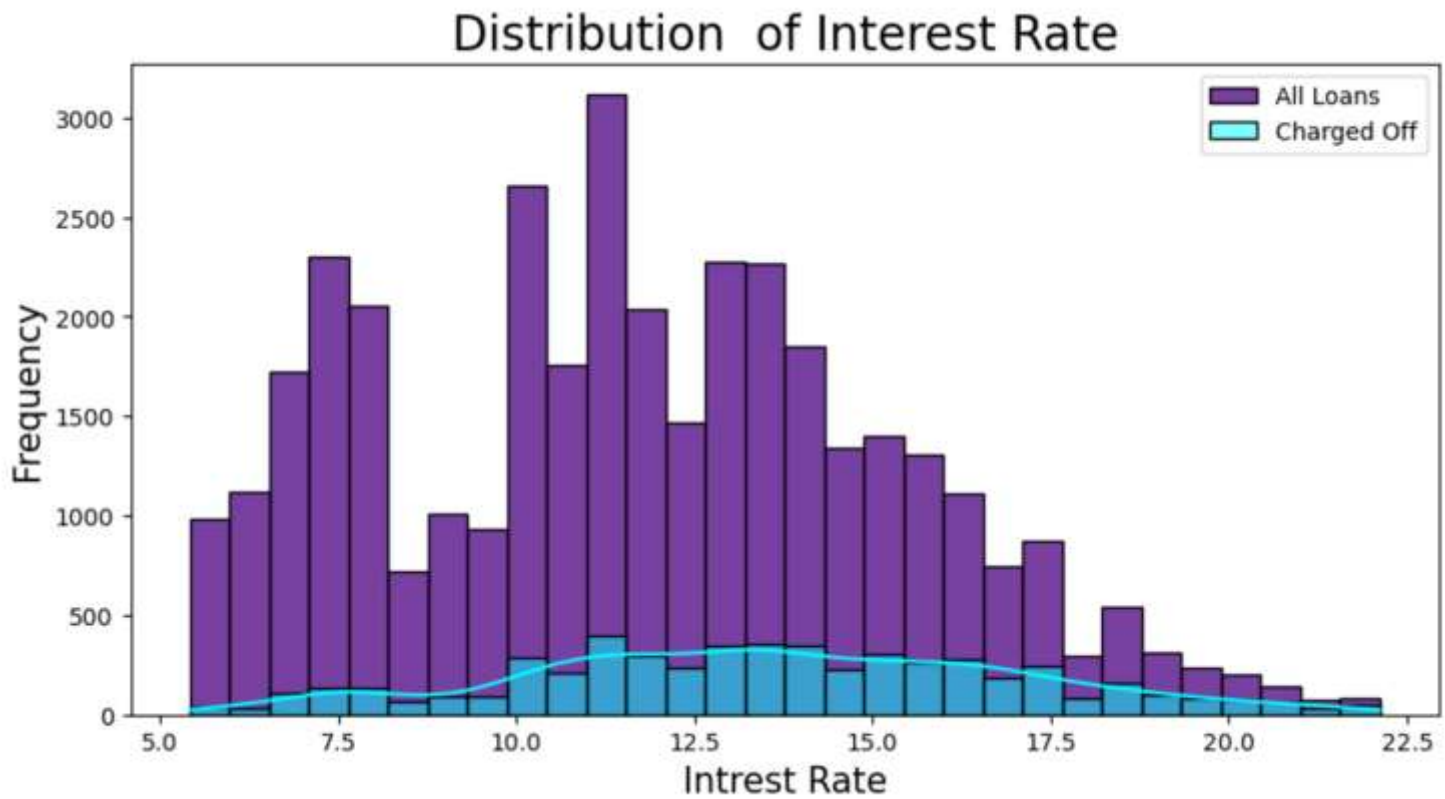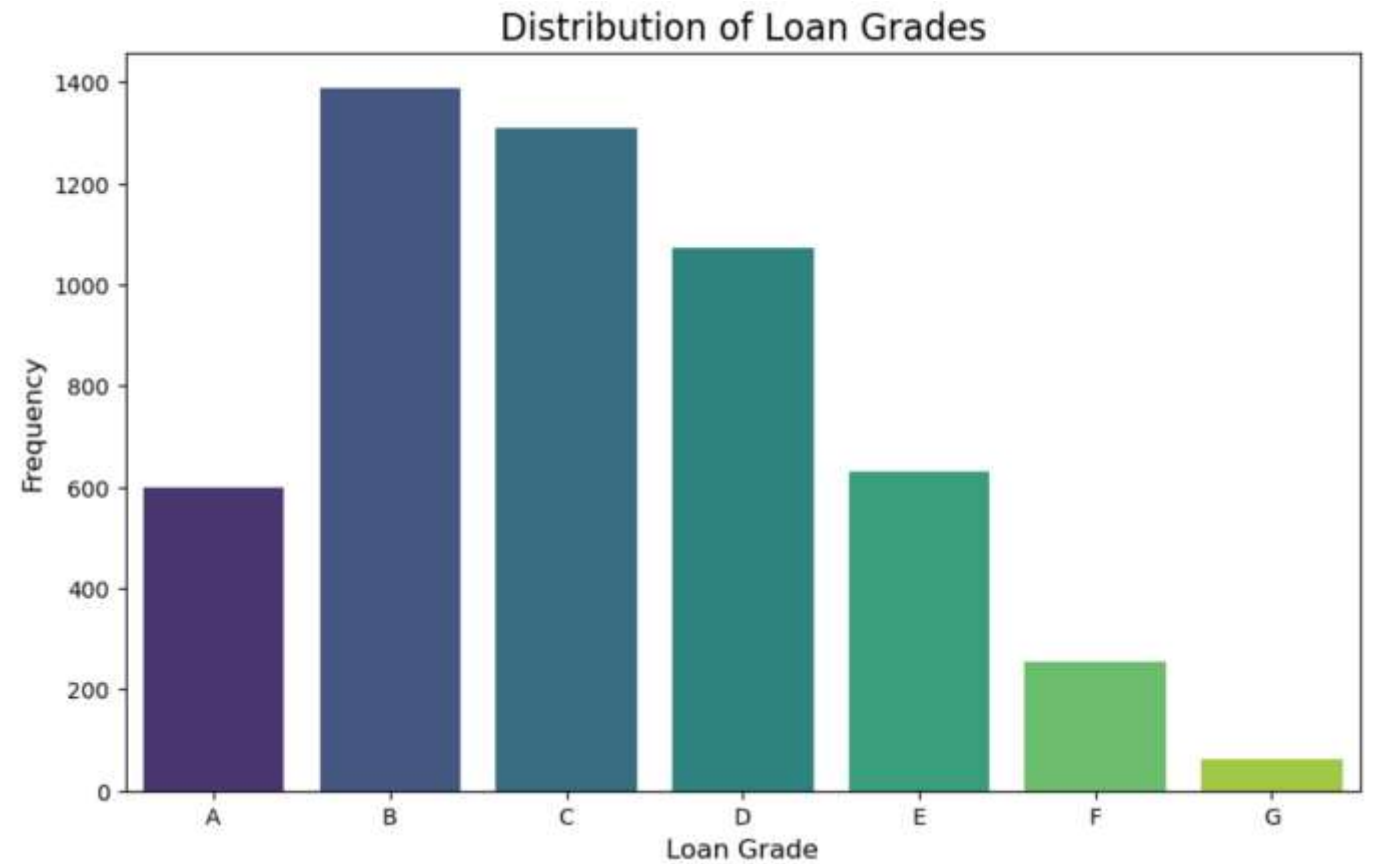
Distribution of Loan Amount

Distribution of Loan Amounts

Distribution of Interest Rate

Distribution of Loan Grades

Univariate Analysis Insights

Purpose : dept_consolidation is the maximum case when the customer has charged off

Home ownership : Most of the charged off customers are in 'RENT' and 'MORTGAGE'

Loan marked as 'Charged Off' are in lower to mid range [5000 to 10000] of loan amounts compared to over all distribution
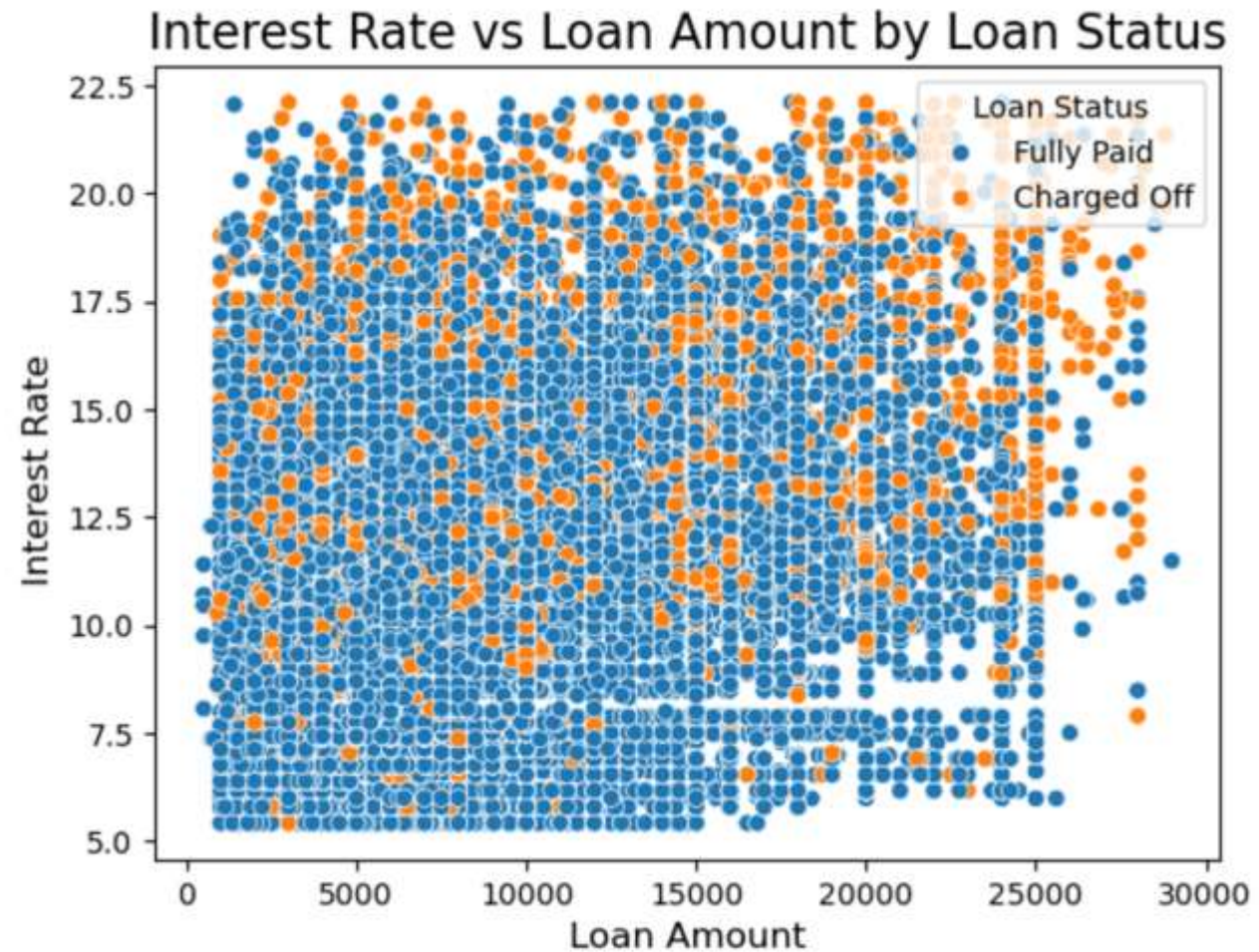
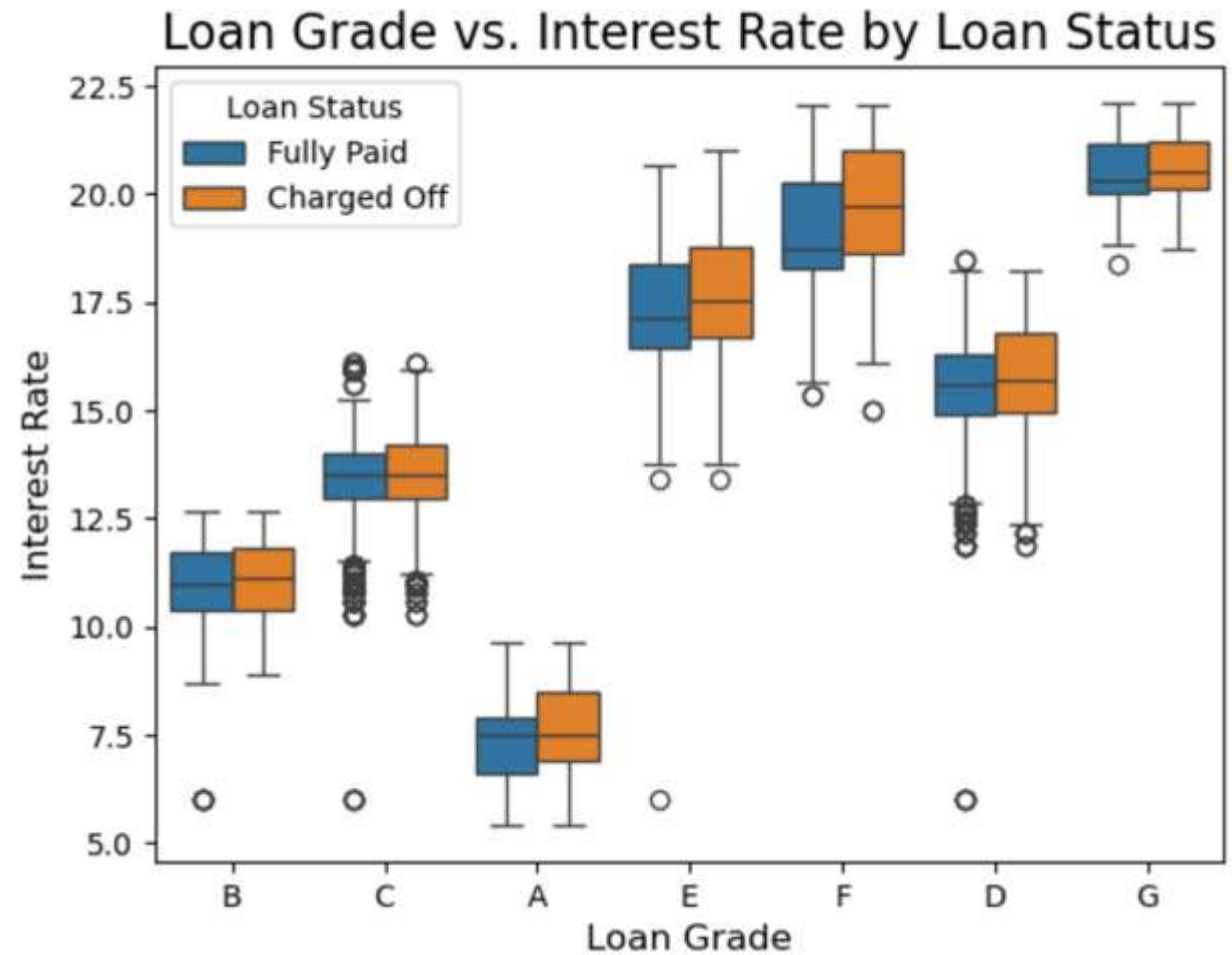Interest Rate: Defaults are more among the loans with higher interest rates.

Loan grades : Grades B,C, D are common among defaults, Lower grade higher risk of defaults.

Bivariate Analysis

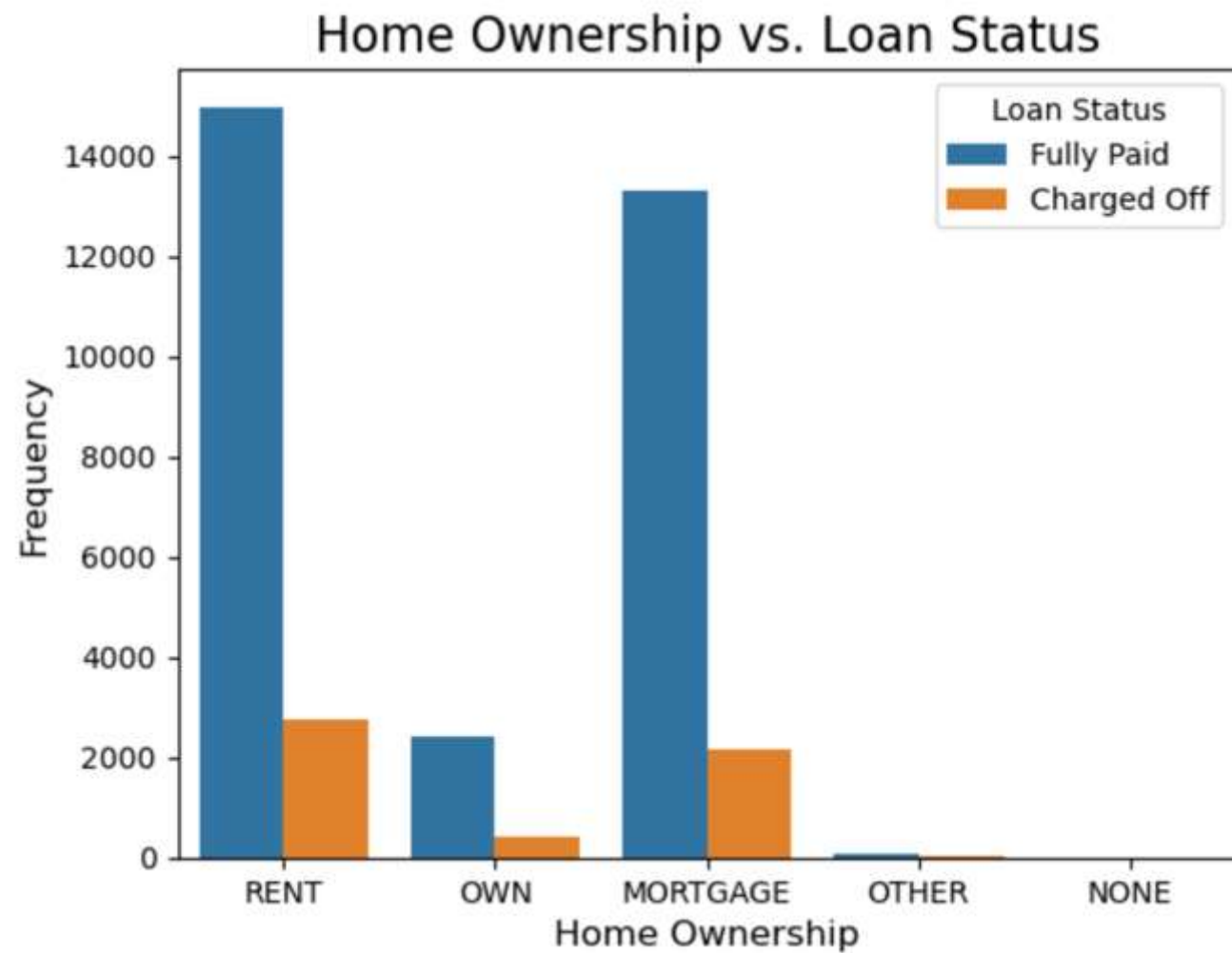Interest Rate vs Grade

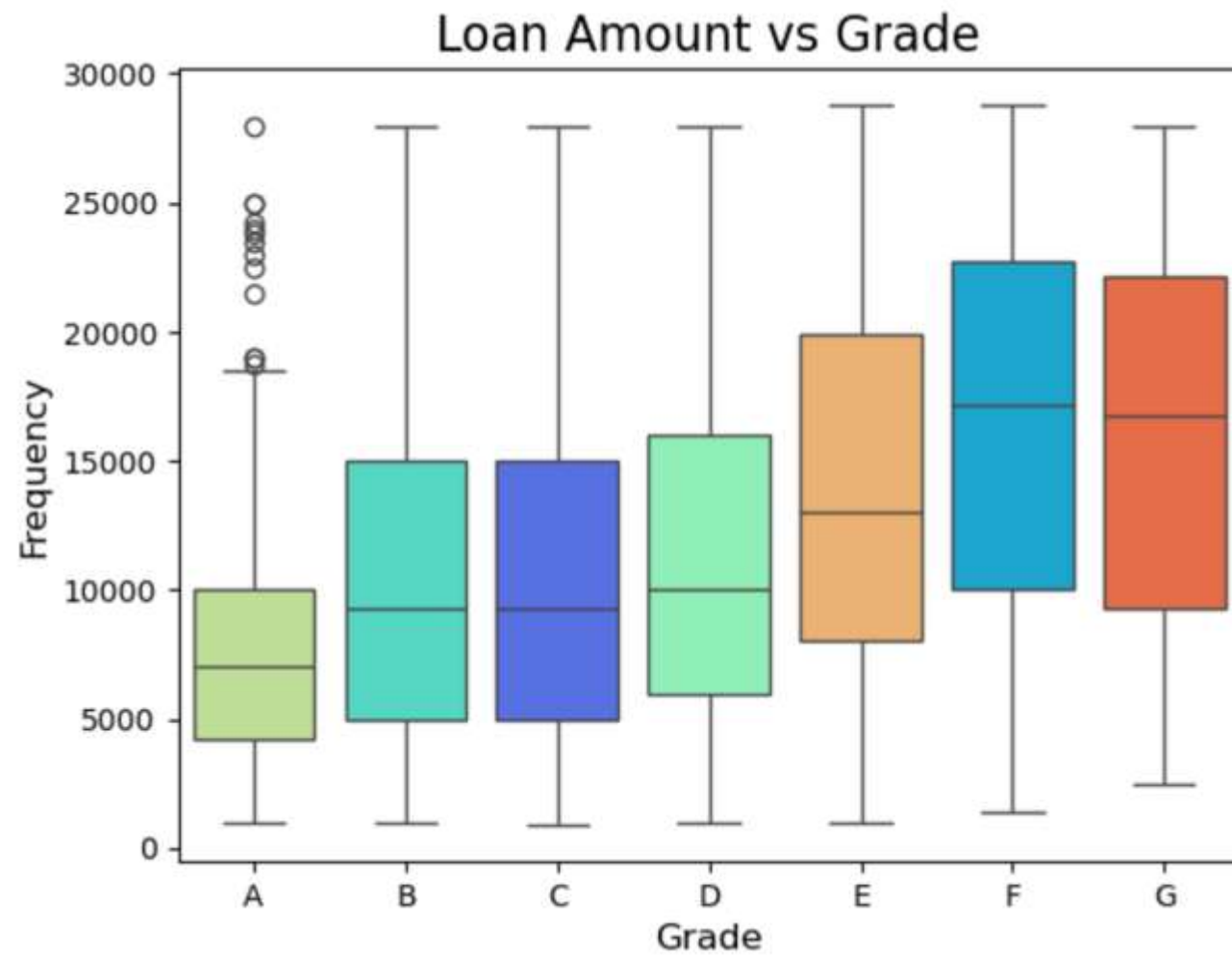Loan Grade vs. Interest Rate by Loan Status

Home Ownership vs Loan status

Loan amount vs Grade

Loan Amount vs Grade

The December 2011 where the loan was funded that were charged off

The month and year which the loan was funded
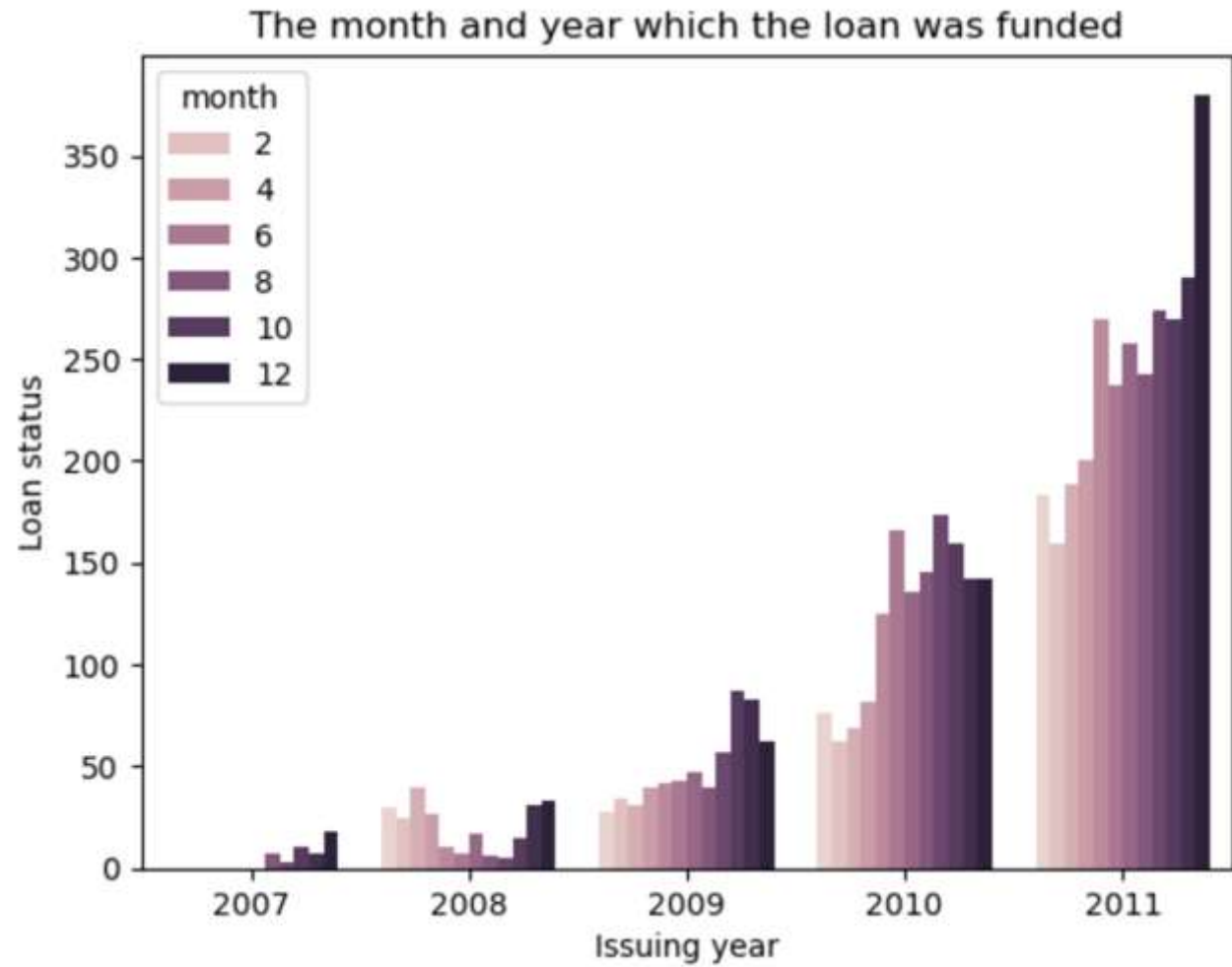
Correlation Matrix of key indicator

- Loan Amount

- Interest Rate

- Funded Amount

- Issued ID (Year and Month)

Correlation Matrix of Key Indicators
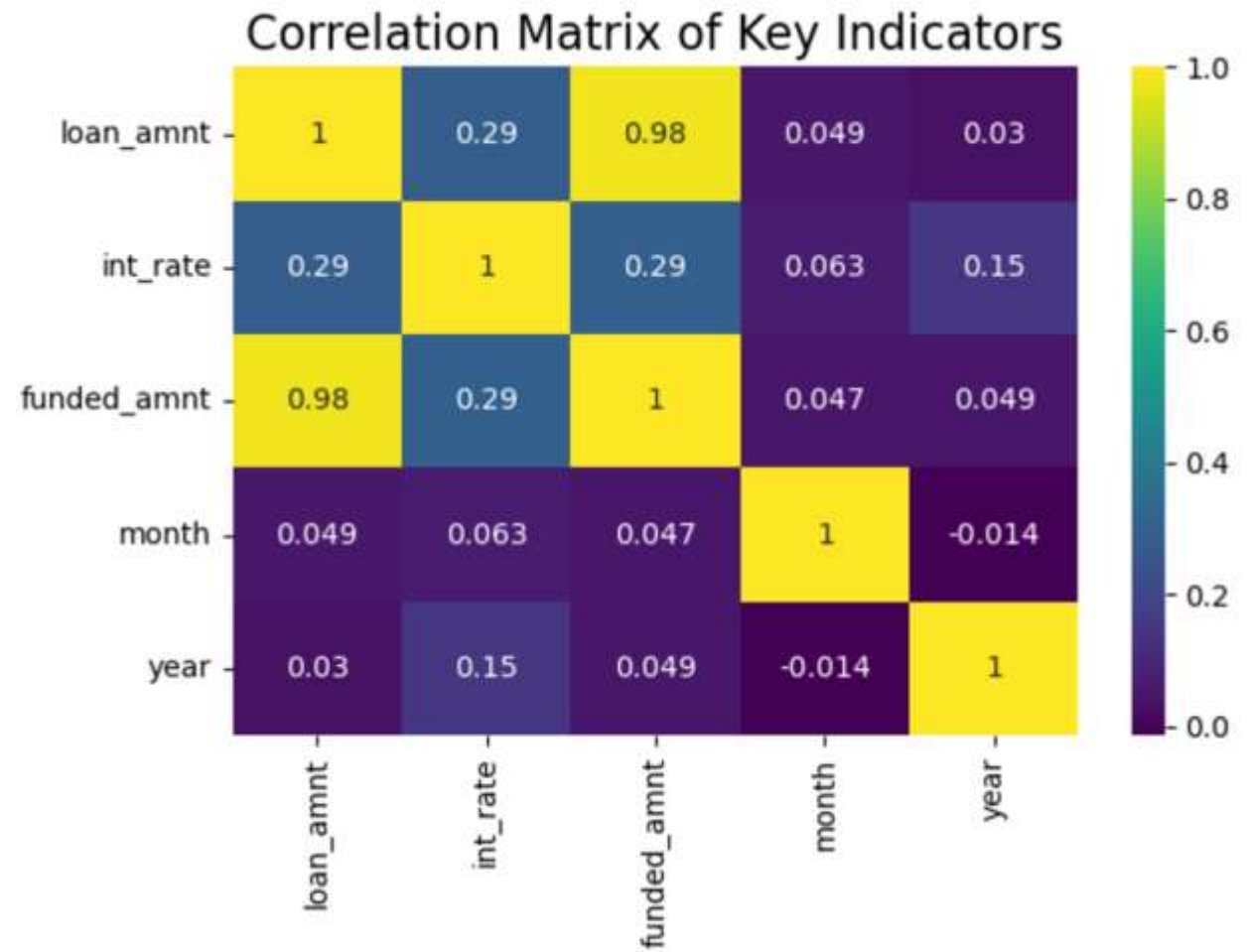
Bivariate Analysis insights

- Interest Rate vs Loan status : Loan with higher interest rate > 12 % are more likely to Charge off.

- Interest Rate vs Grade : Loan with grade D, E , F with higher interest rate are more likely to charge off.

- Home Ownership vs Loan status : Applicants with home rent and mortgage are charged off

- Loan amount vs Grade: Grade 'E,F,G' corresponds to larger loan amount are charged off

- Issue ID : The December 2011 where the loan was funded that were charged off