Lending Club Case Study

Presented by - Bhargav Meka & Ishani Garg

Introduction

- Lending Club Case is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.
- Like most other lending companies, lending loans to ‘risky’ applicants leads to credit loss.
- This case study aims to identify the risky loan applicants using the EDA approach so that such loans can be reduced thereby cutting down the amount of credit loss.

RoadMap

01

02

03

Step 1

Data Understanding & Cleaning

Step 2

Univariate and Segmented Analysis

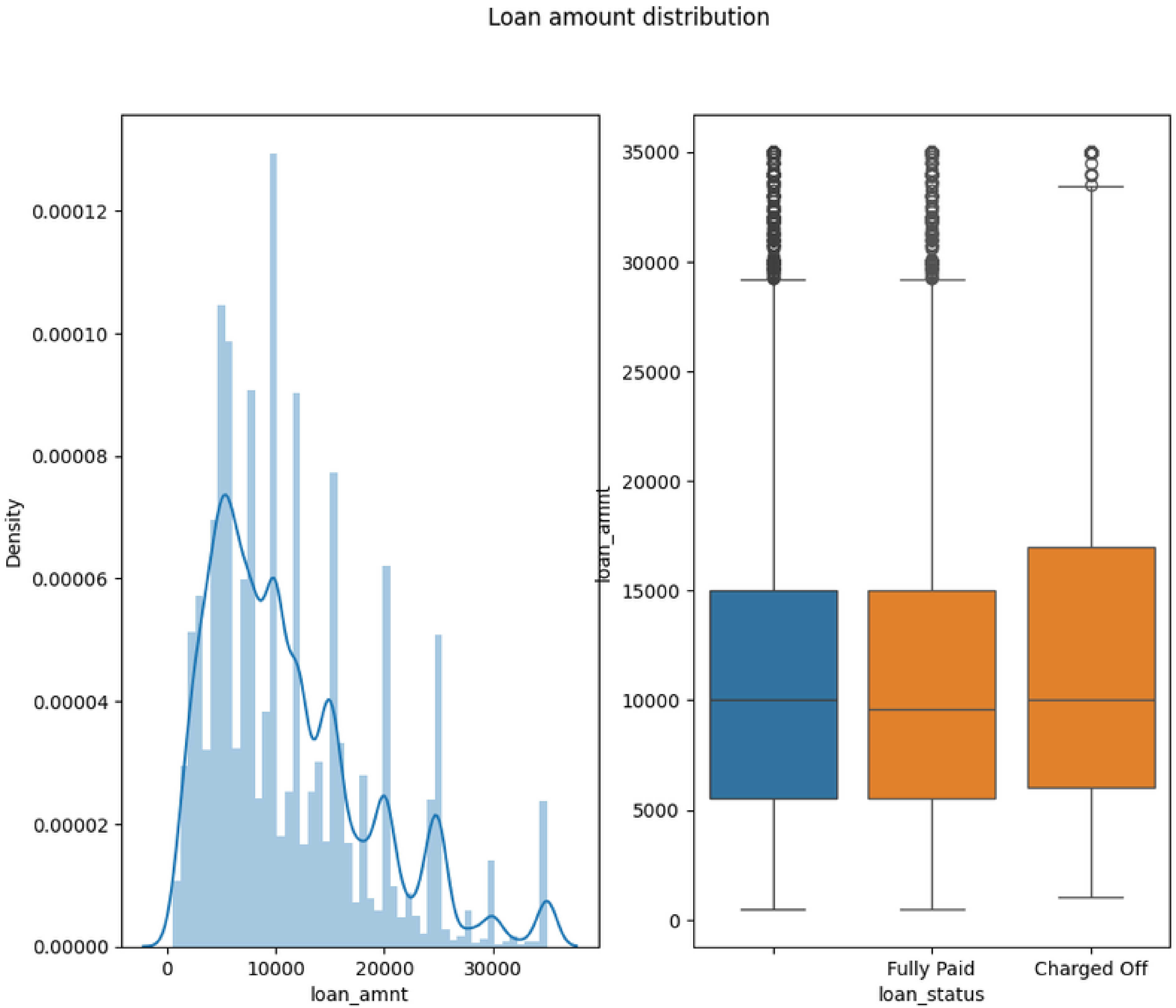
Step 3

Biraviate Analysis and Derive Conclusions

Data Understanding & Cleaning

- Removed all the columns with Null/NA values
- Removed all the rows with majority Null values
- Removed columns having only one unique value
- Removed Columns that are not relevant to the analysis
- Removed leading and trailing spaces from the columns
- Converted appropriate datatypes e.g. Numeric columns to float, columns having date to data type datetime.

Univariate Analysis

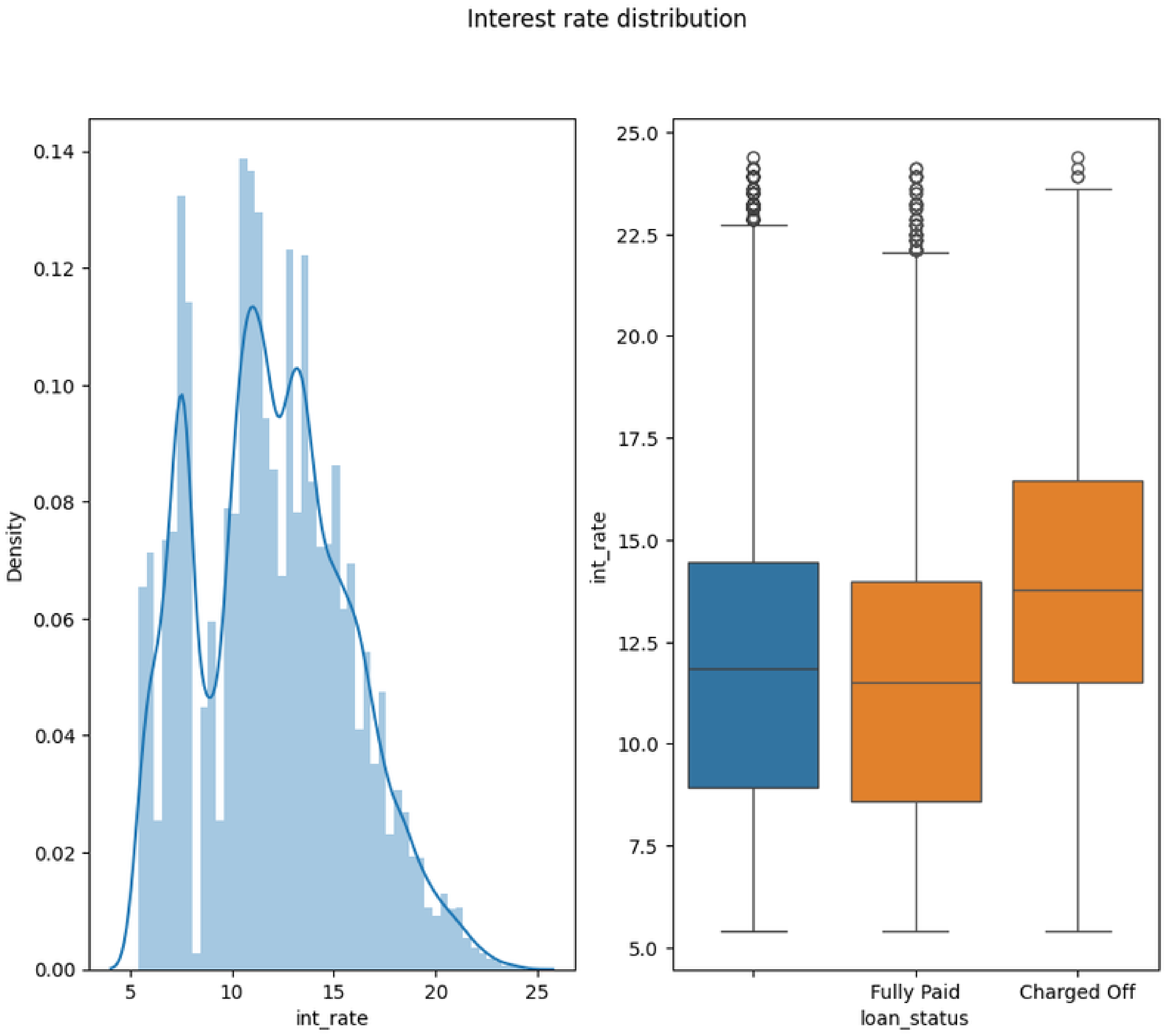


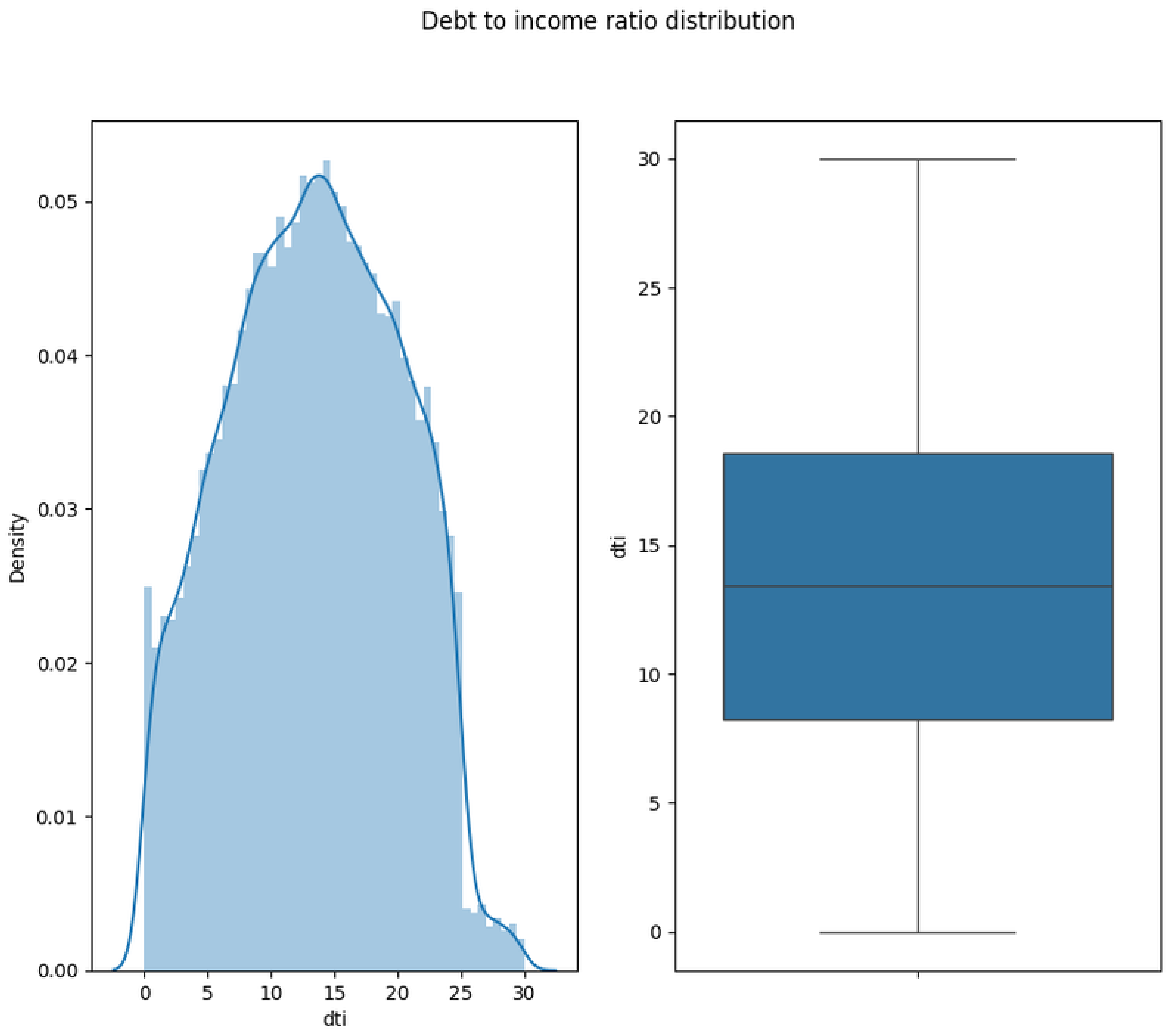
Loan Amount Distribution

- The distribution of loan_amnt is right-skewed with mean > median
- The median loan amount is 10000 and most of the loans are distributed
- Very few loans are distributed for amounts > 20000
- The median loan amount for charged-off loans is higher than fully paid loans

Interest Rate Distribution

- The interest rate is between 5.42% to 24.4% with a median of 13.49%
- The median interest rate for charged-off loans is 15.61% and for fully paid loans is 13.11%
- The interest rate for charged-off loans is higher than fully paid loans
- Most of the loans have an interest rate between 10% to 15%



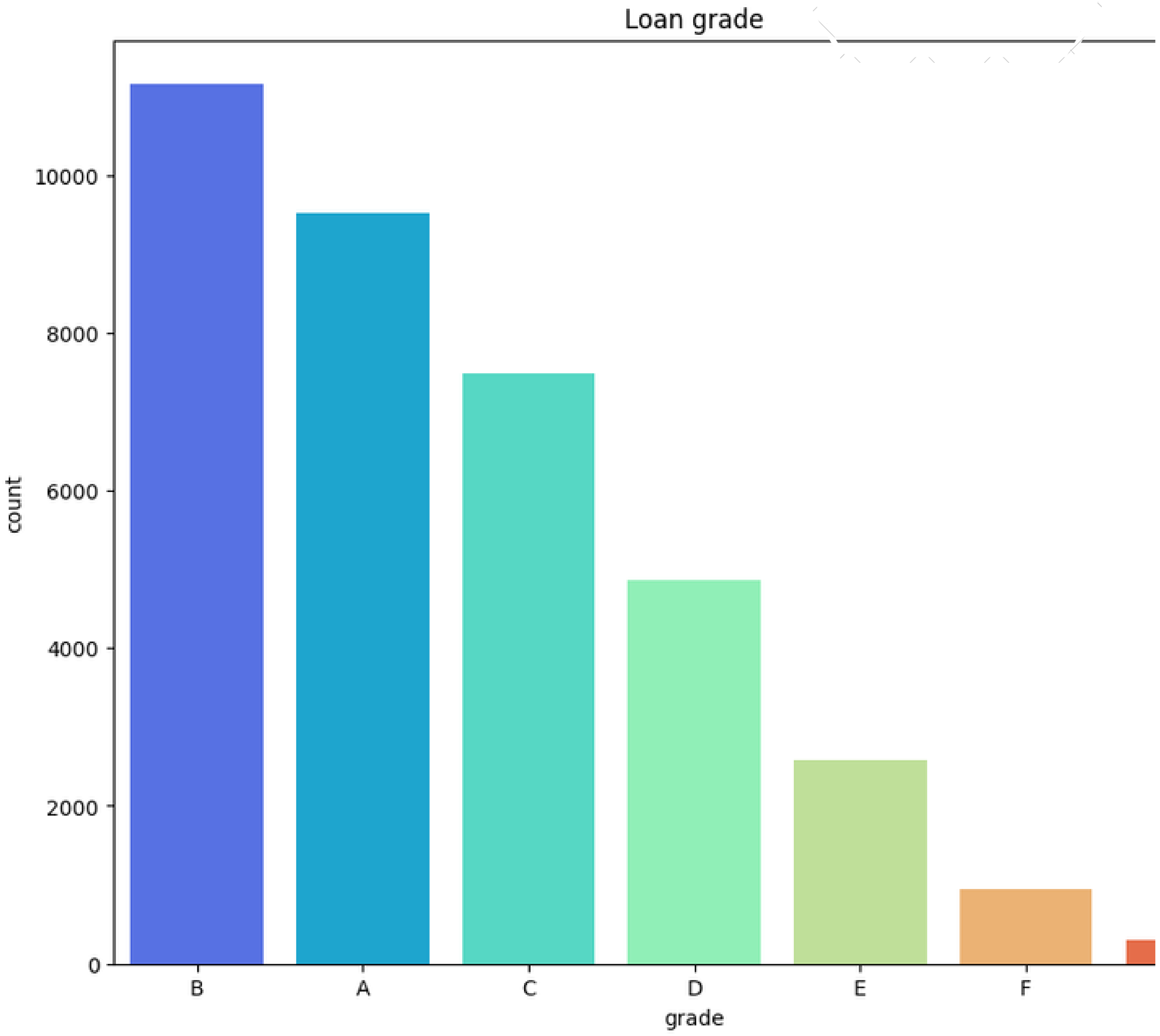


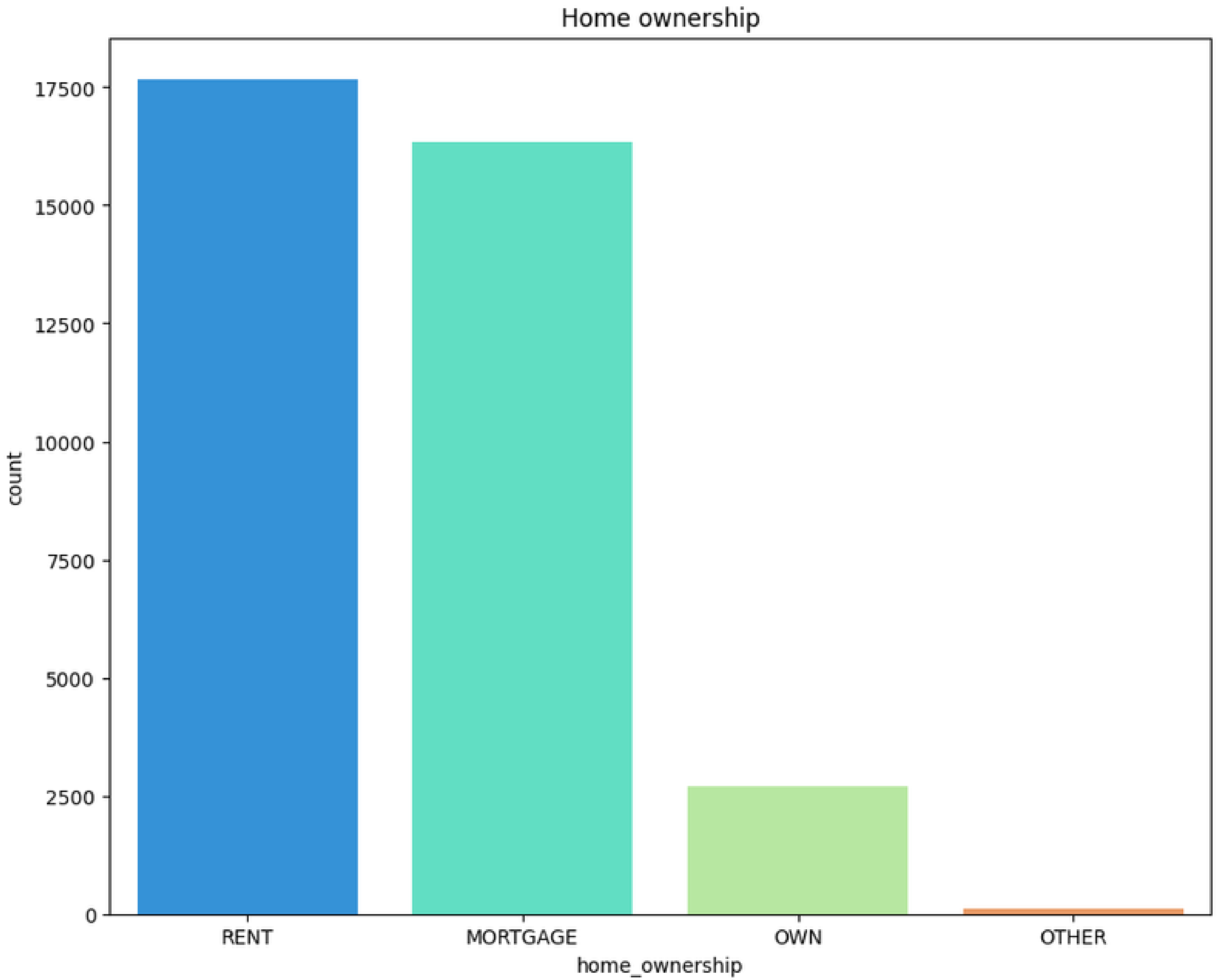
Debt To Income Ratio

Most of the loans are issued to applicants with debt to income ratio less than 20

Loan Grade

- Most of the loans are of grade A, B and C



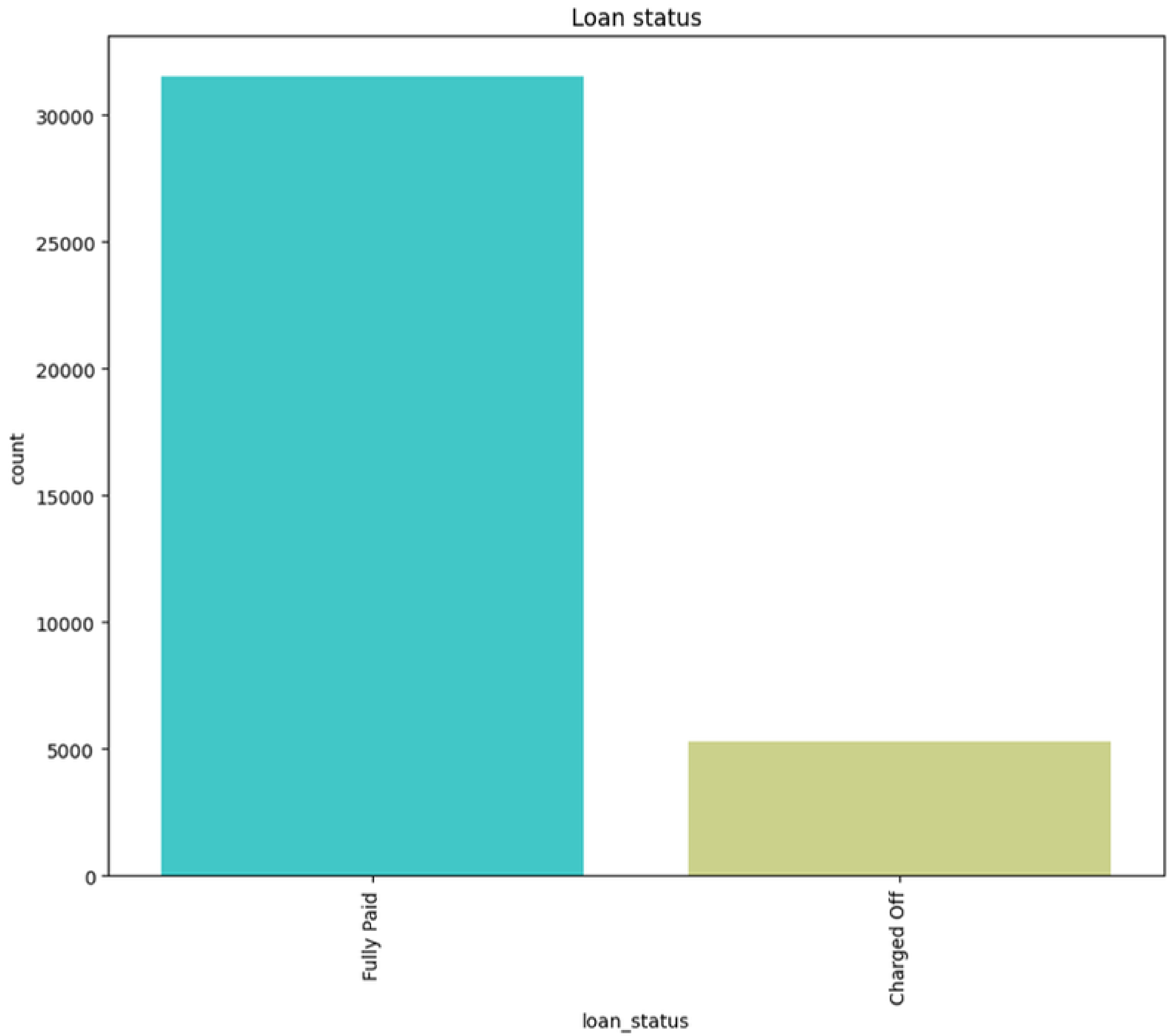


Home Ownership

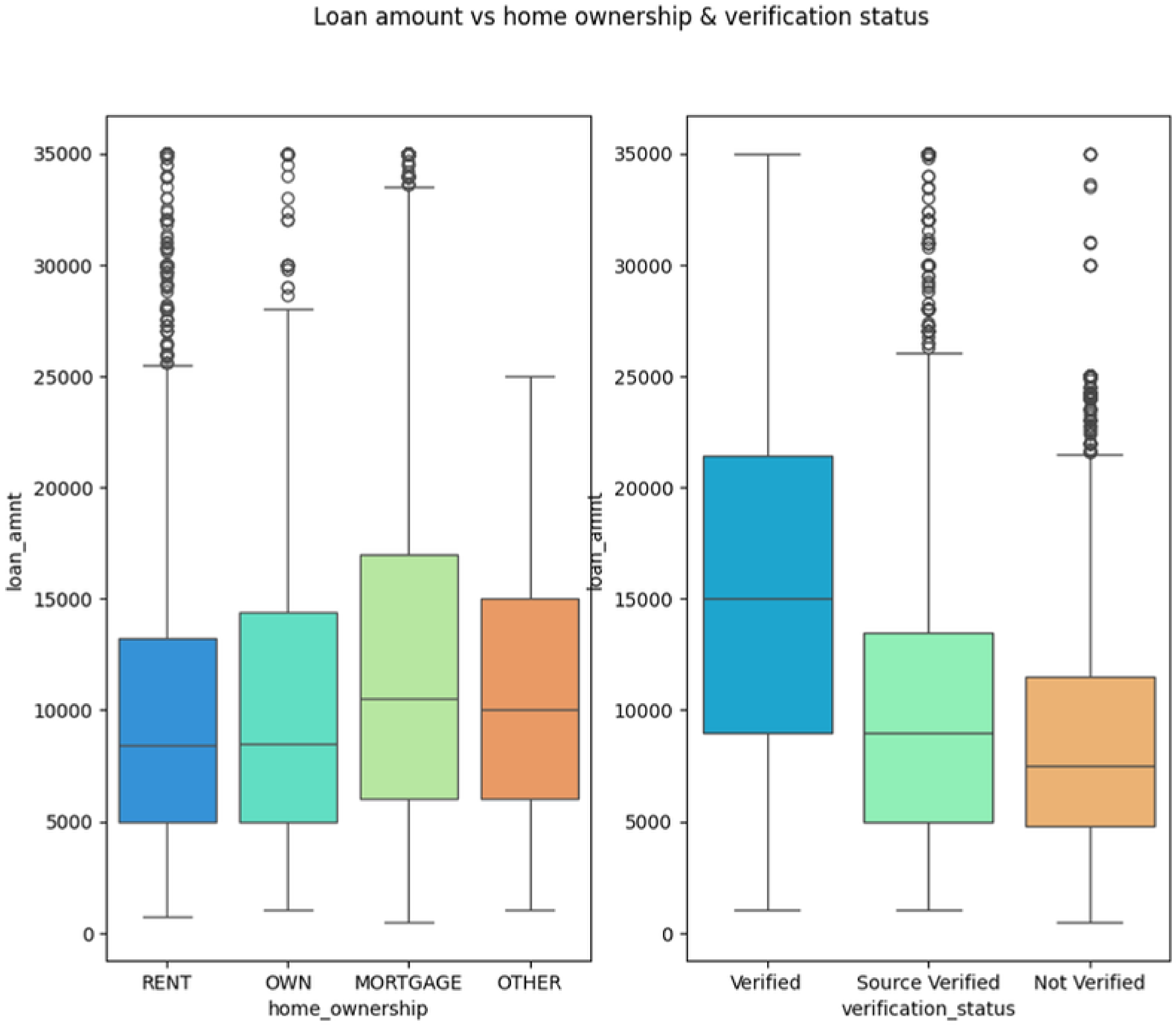
Majority of the applicants are either on rent or mortgage holders

Fully Paid vs Charged Off

15% of the loans are charged off (defaulted) and 85% are fully paid

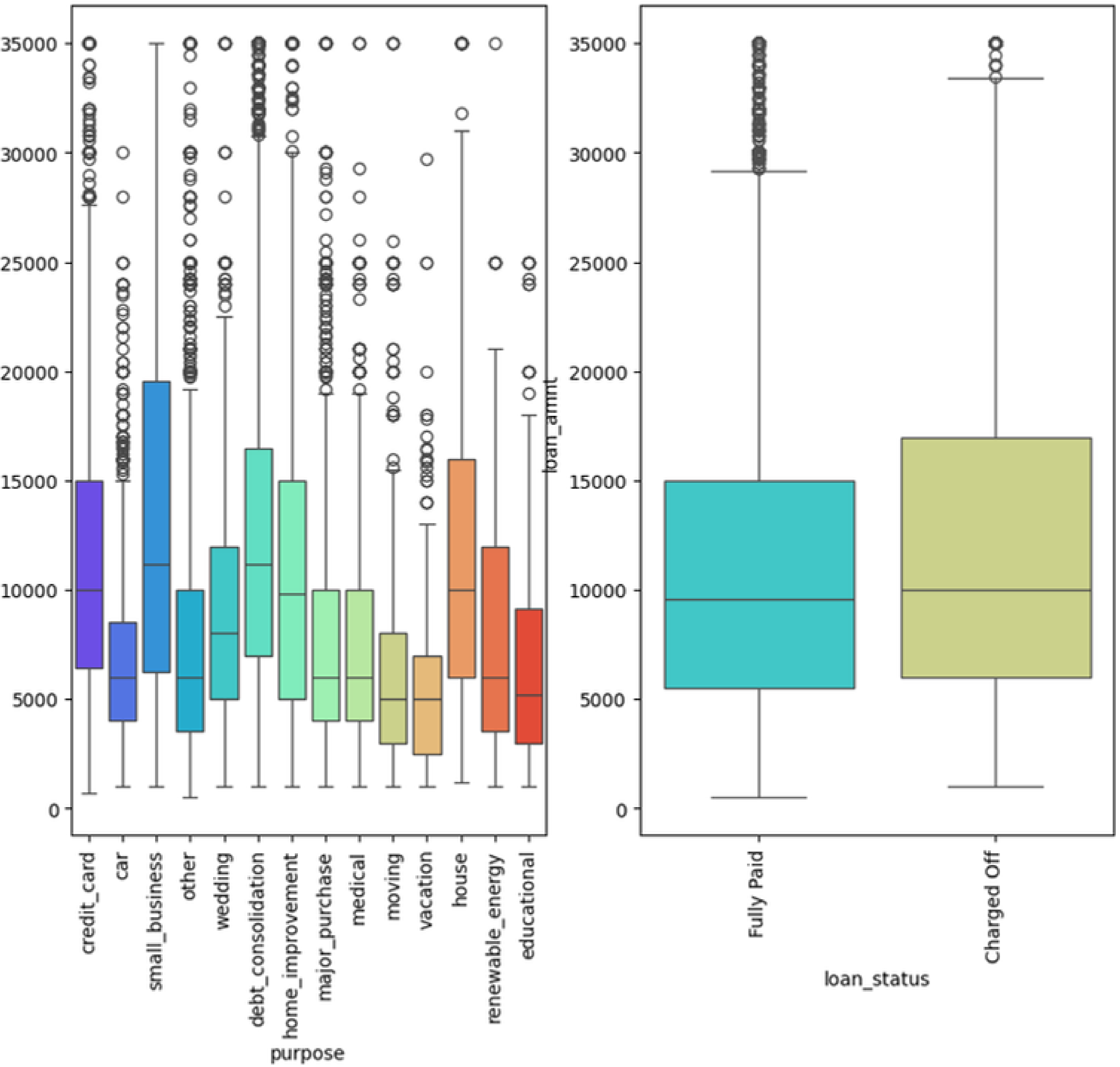


Segmented Univariate Analysis

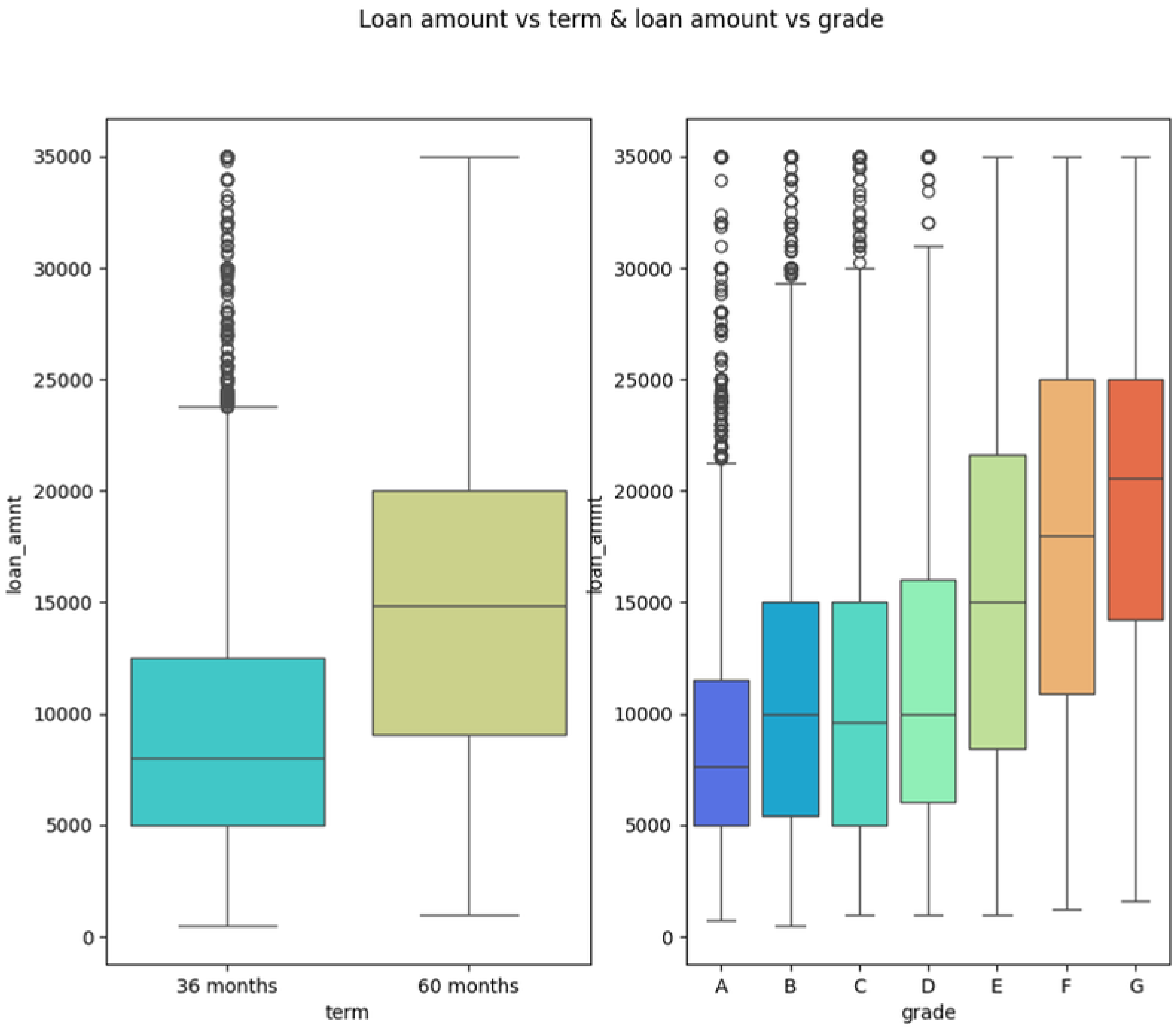


- Mortgage has the highest loan amount
- Verified borrowers have the highest loan amount
- Not verified borrowers have the lowest loan amount
-

Loan amount vs purpose & loan amount vs loan status



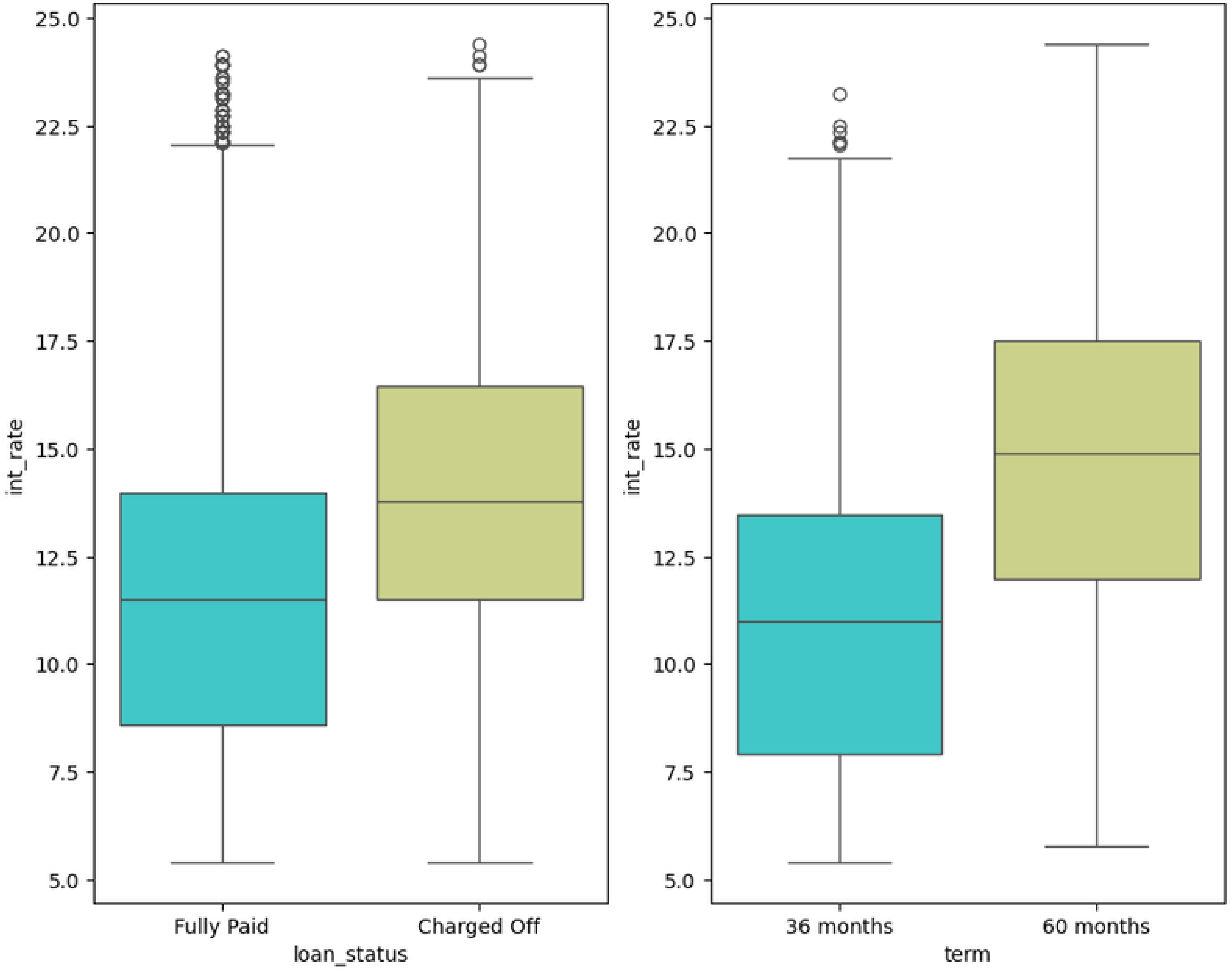
- **Charged-off loans have higher loan amounts than fully paid loans**
- **Loans for small businesses have higher loan amounts than other purposes**



Higher loan amounts have higher loan term

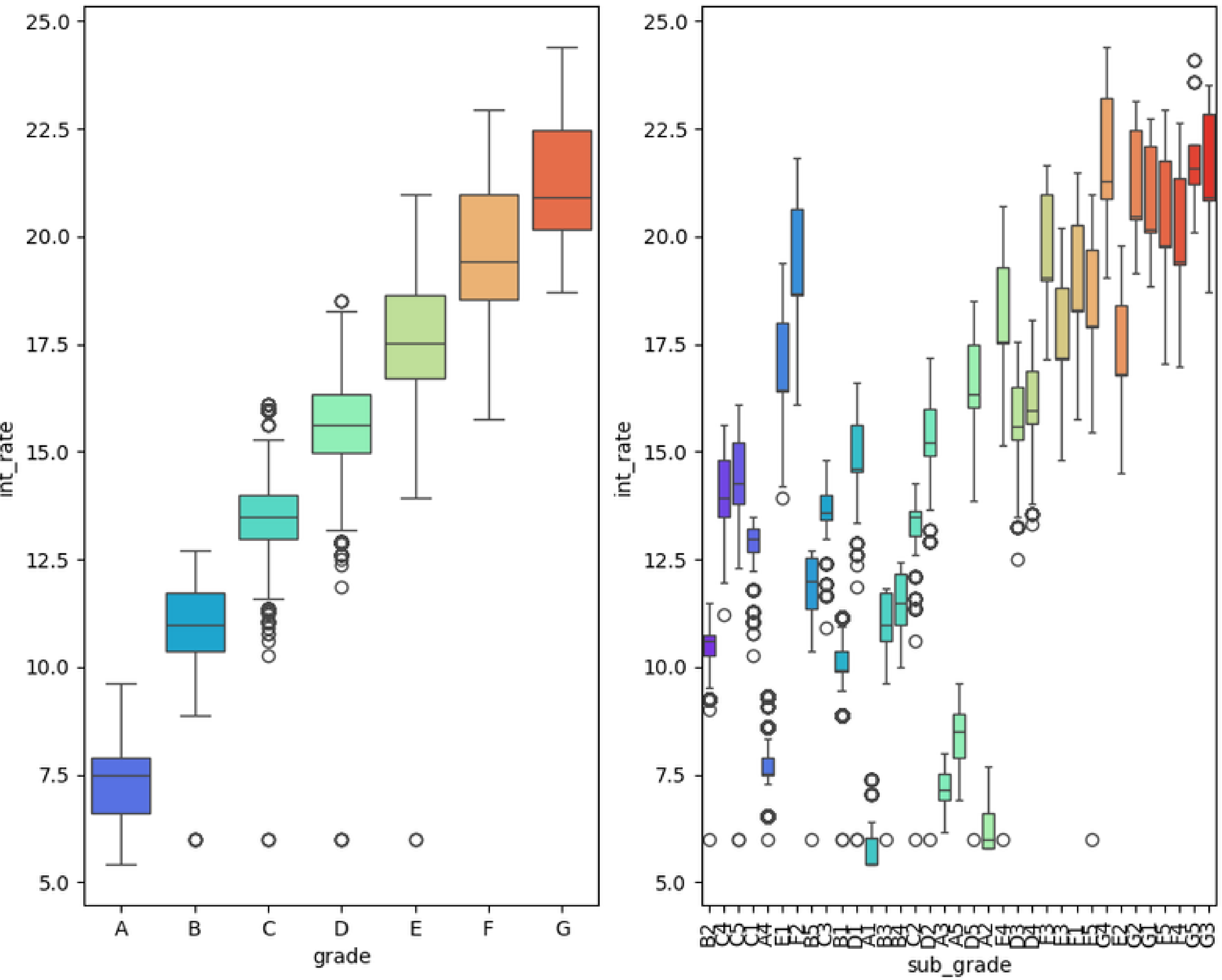
Higher loan amounts are in grades E, F, and G

Interest rate vs loan status and loan status vs term



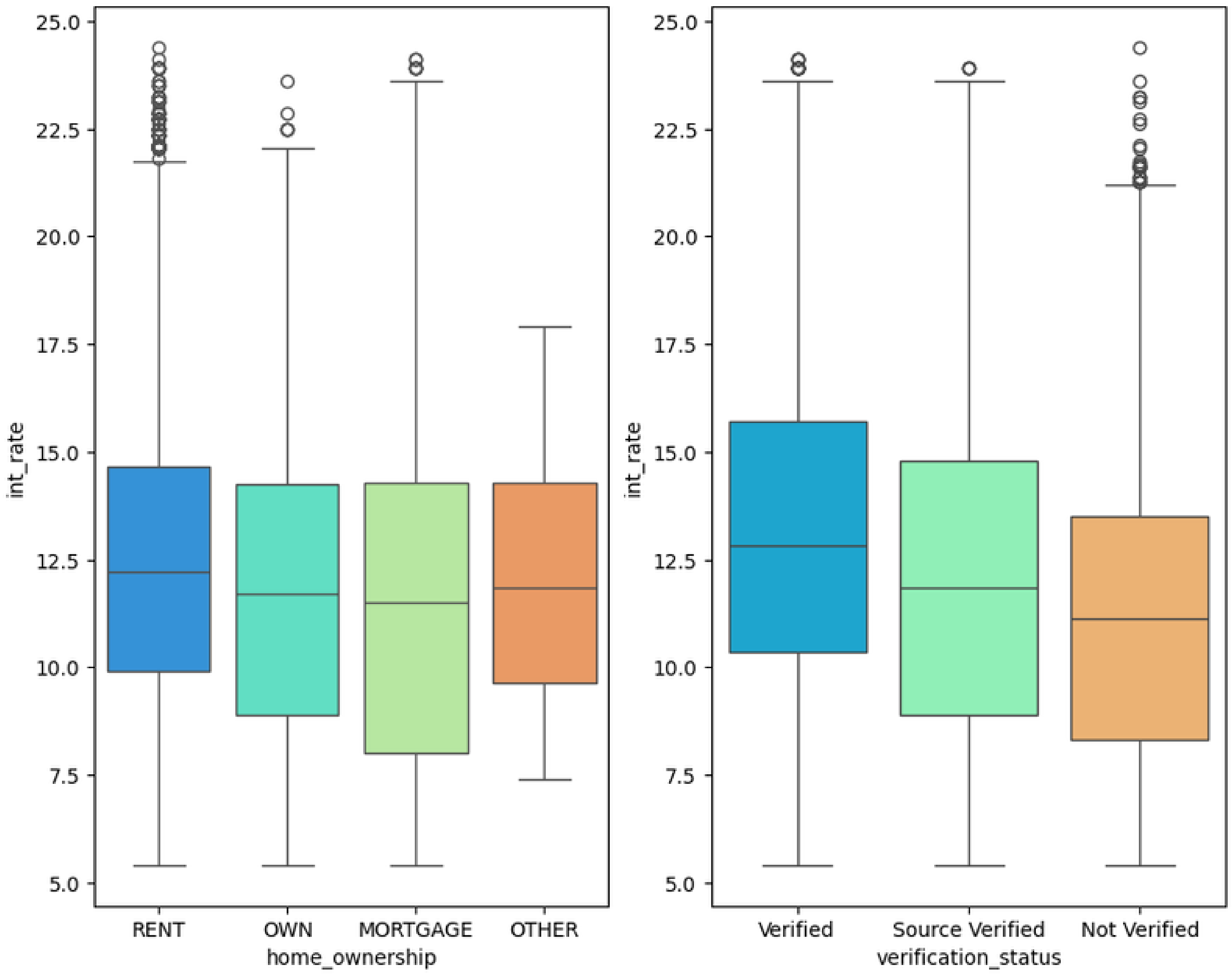
- **Interest rate is higher for charged-off loans**
- **Interest rate is higher for 60 month term loans**

Interest rate vs grade and interest rate vs sub grade



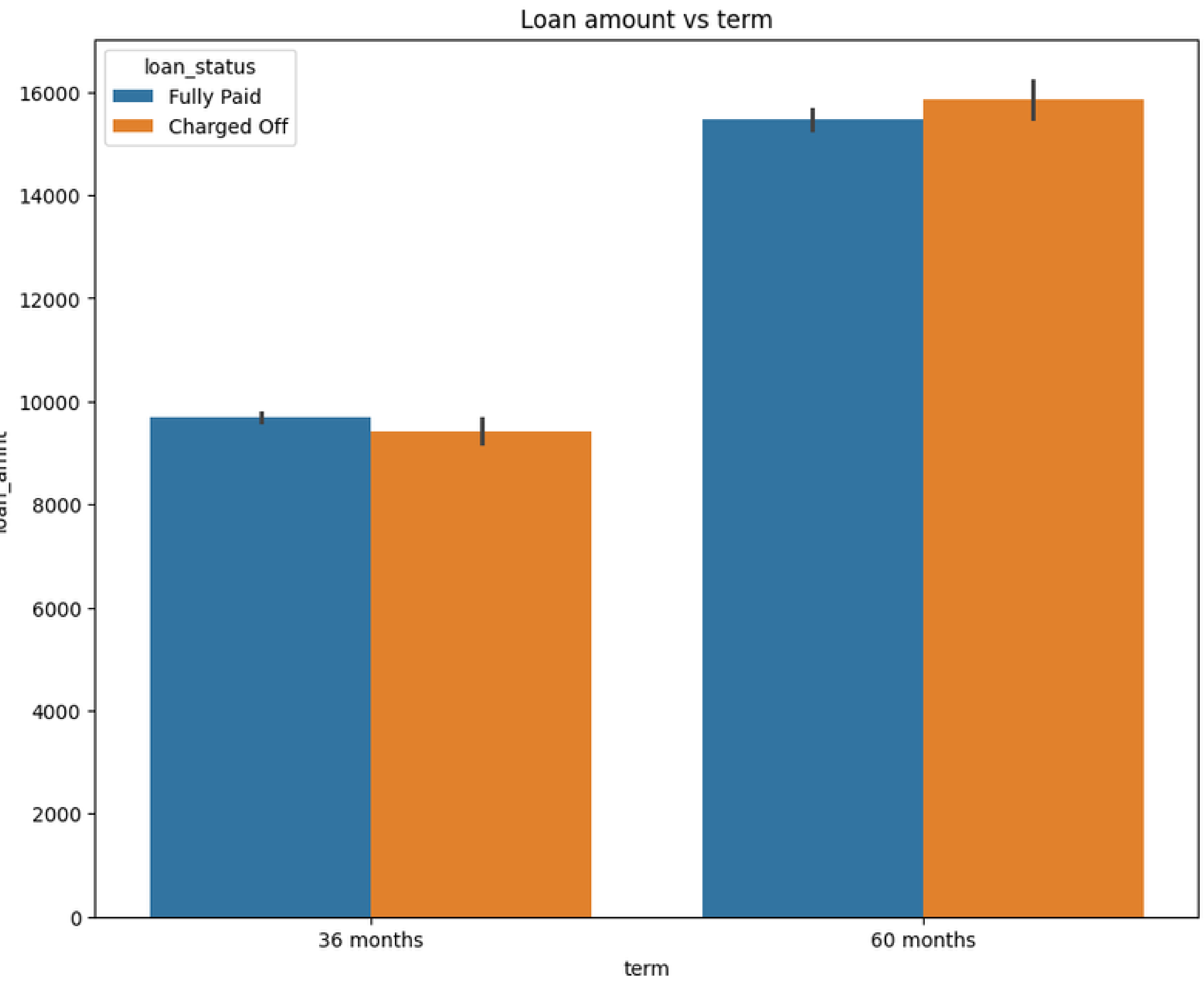
- Interest rates are higher as the grade and sub grade decreases

Interest rate vs home ownership and interest rate vs verification status

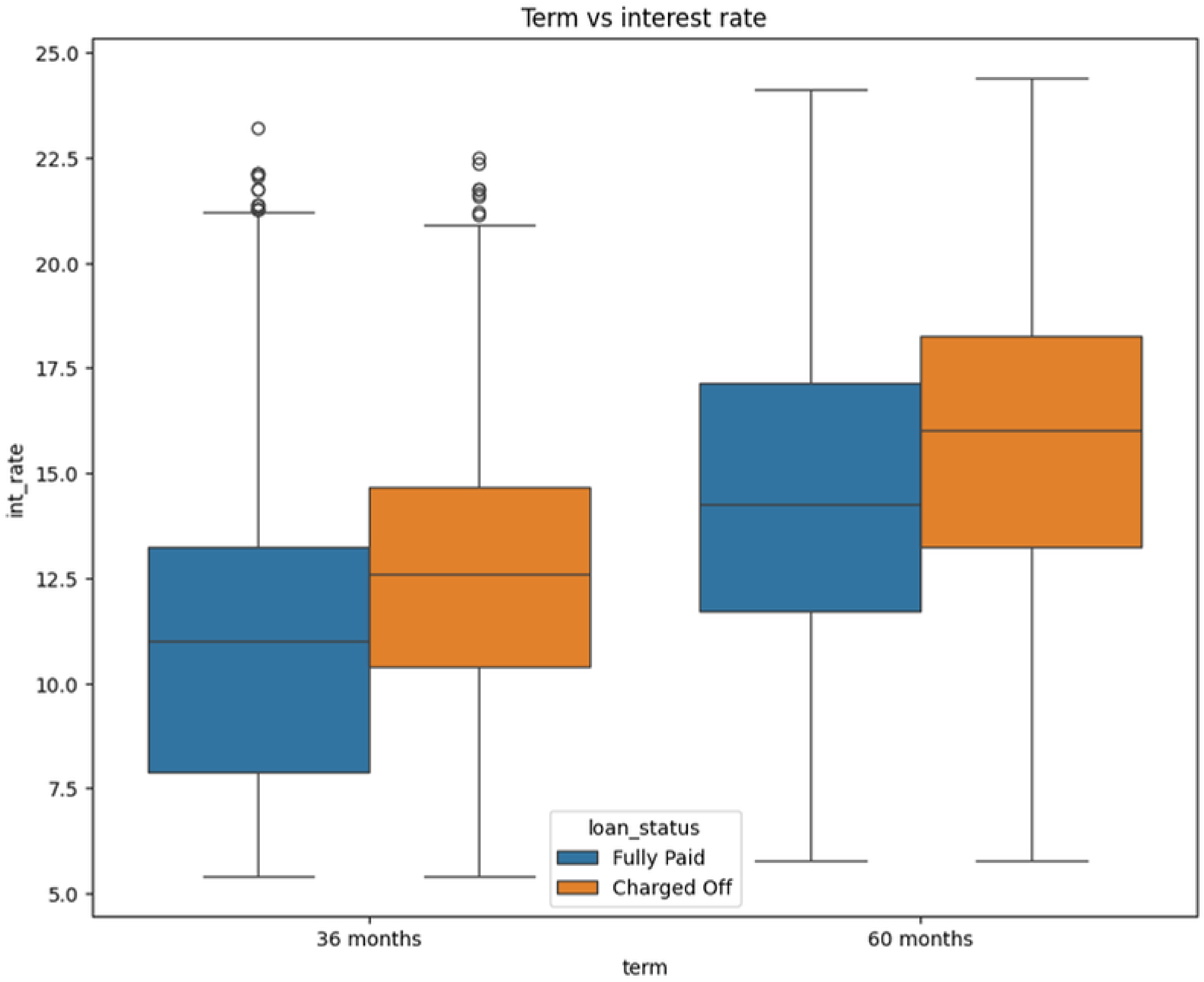


- **Borrowers with home ownership as Own and Mortgage have lower interest rates**
- **Not verified borrowers have lower interest rates**

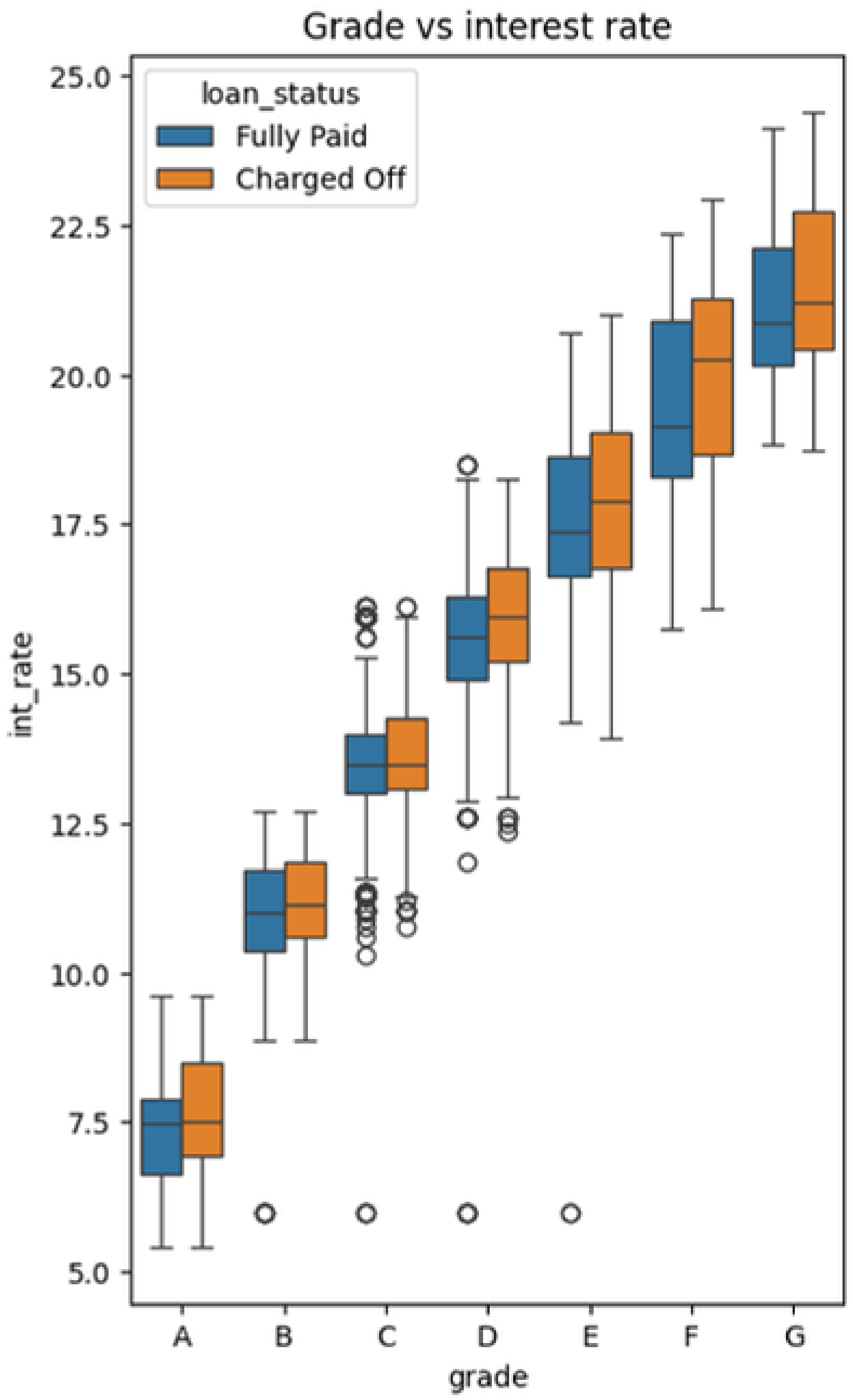
Bivariate Analysis

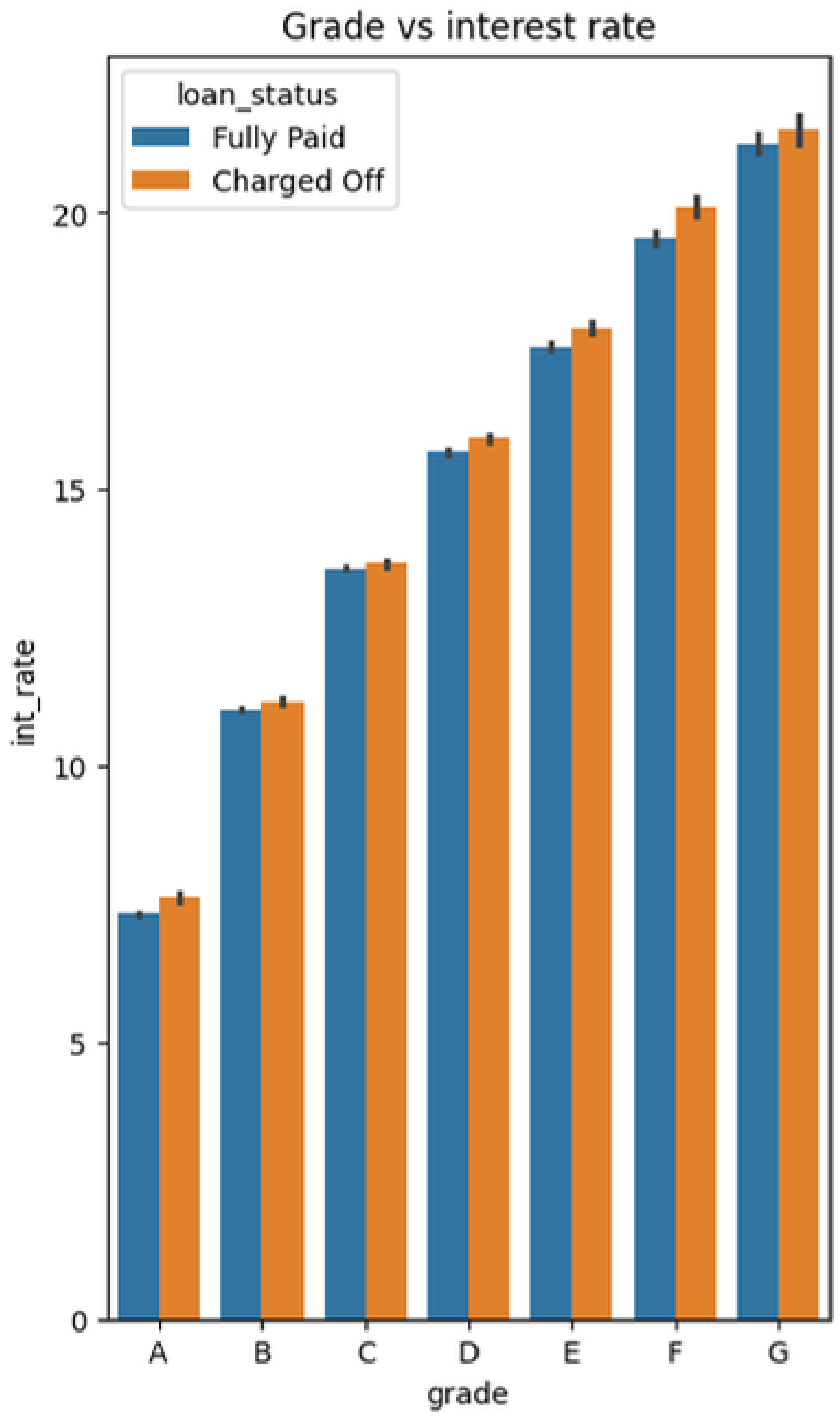


- The loan amount is higher for 60 months term than 36 months term

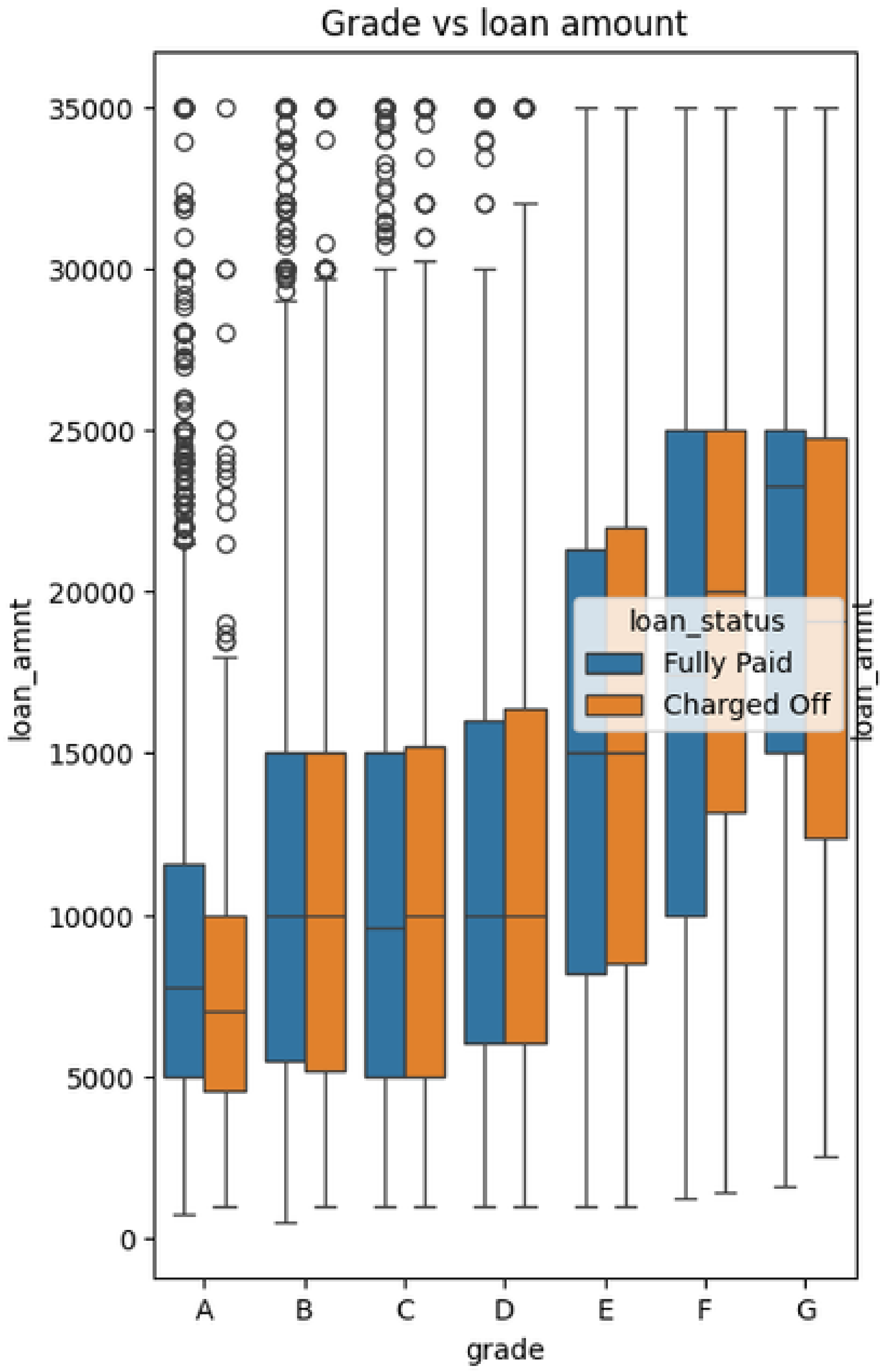


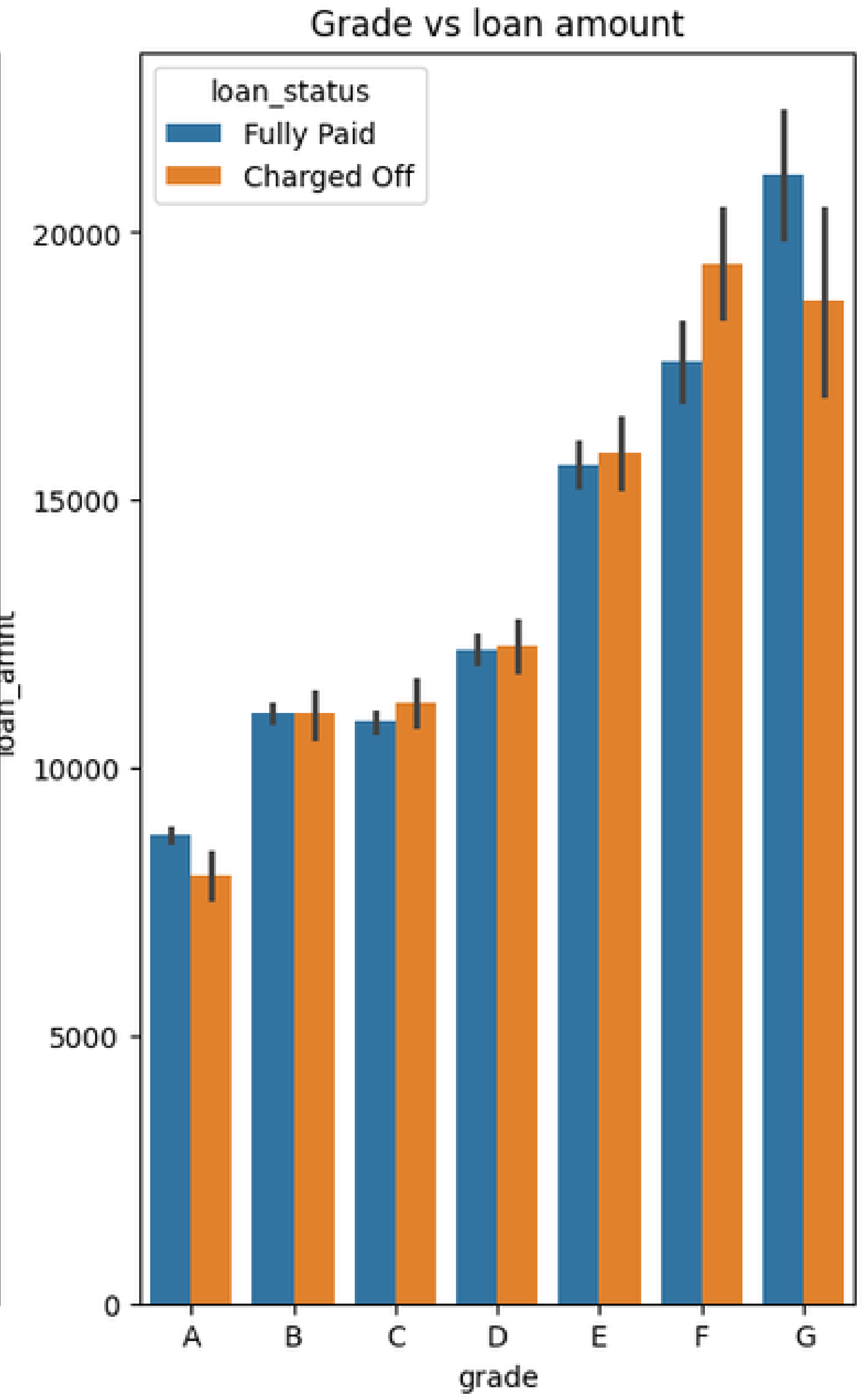
- **Higher interest rates have higher default rates in both 36 months and 60 months term loans**



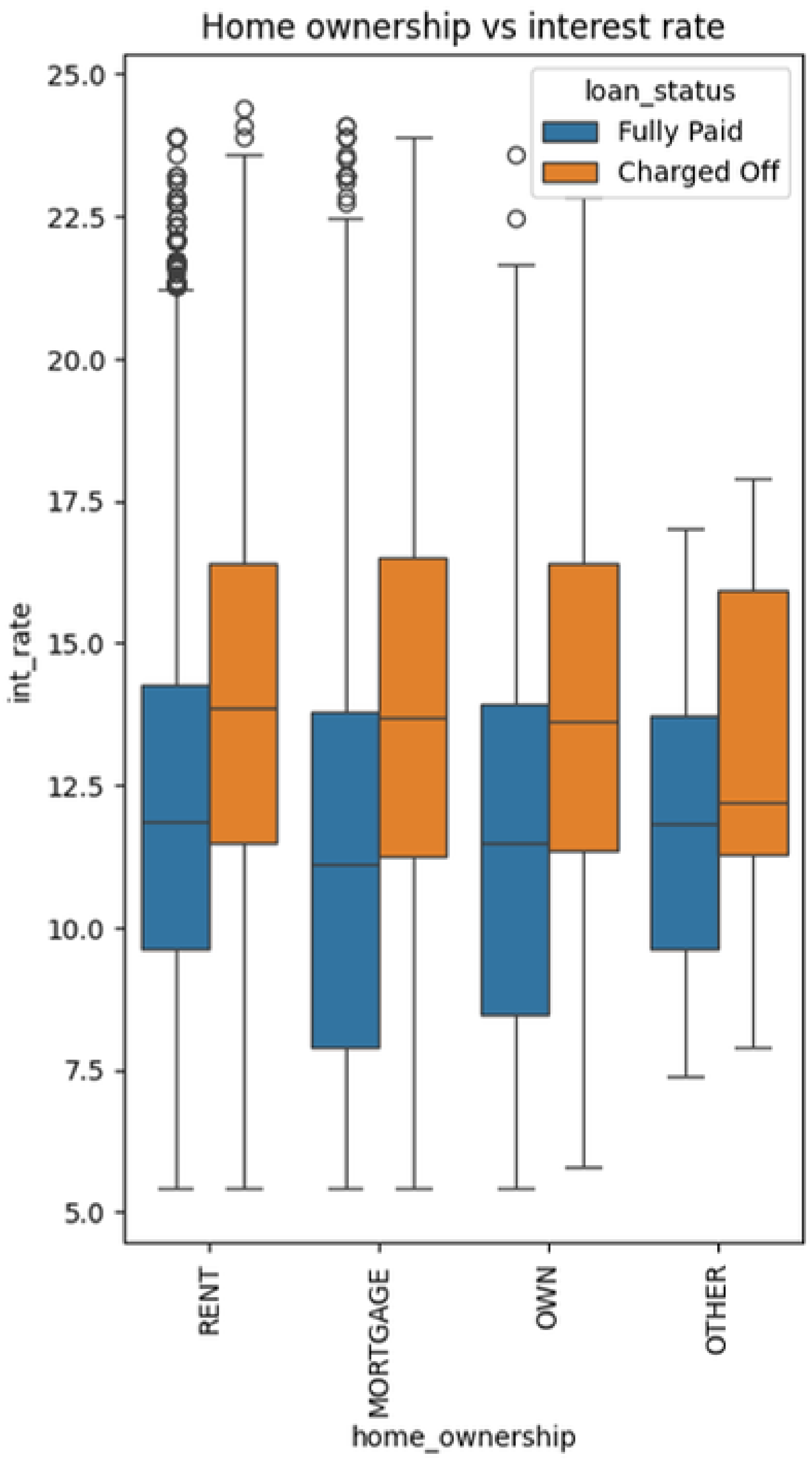


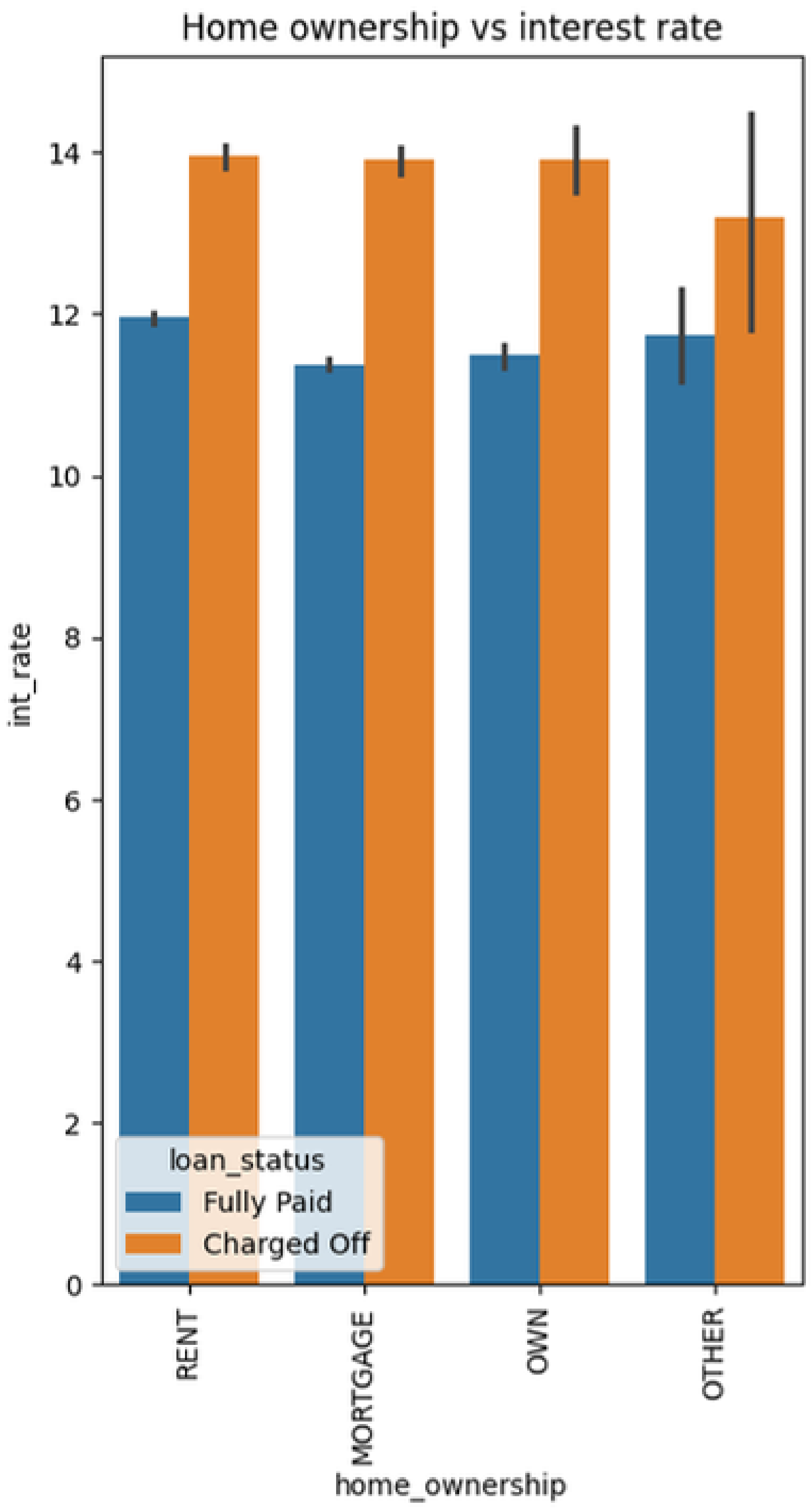
- Lower the grade, higher the interest rate



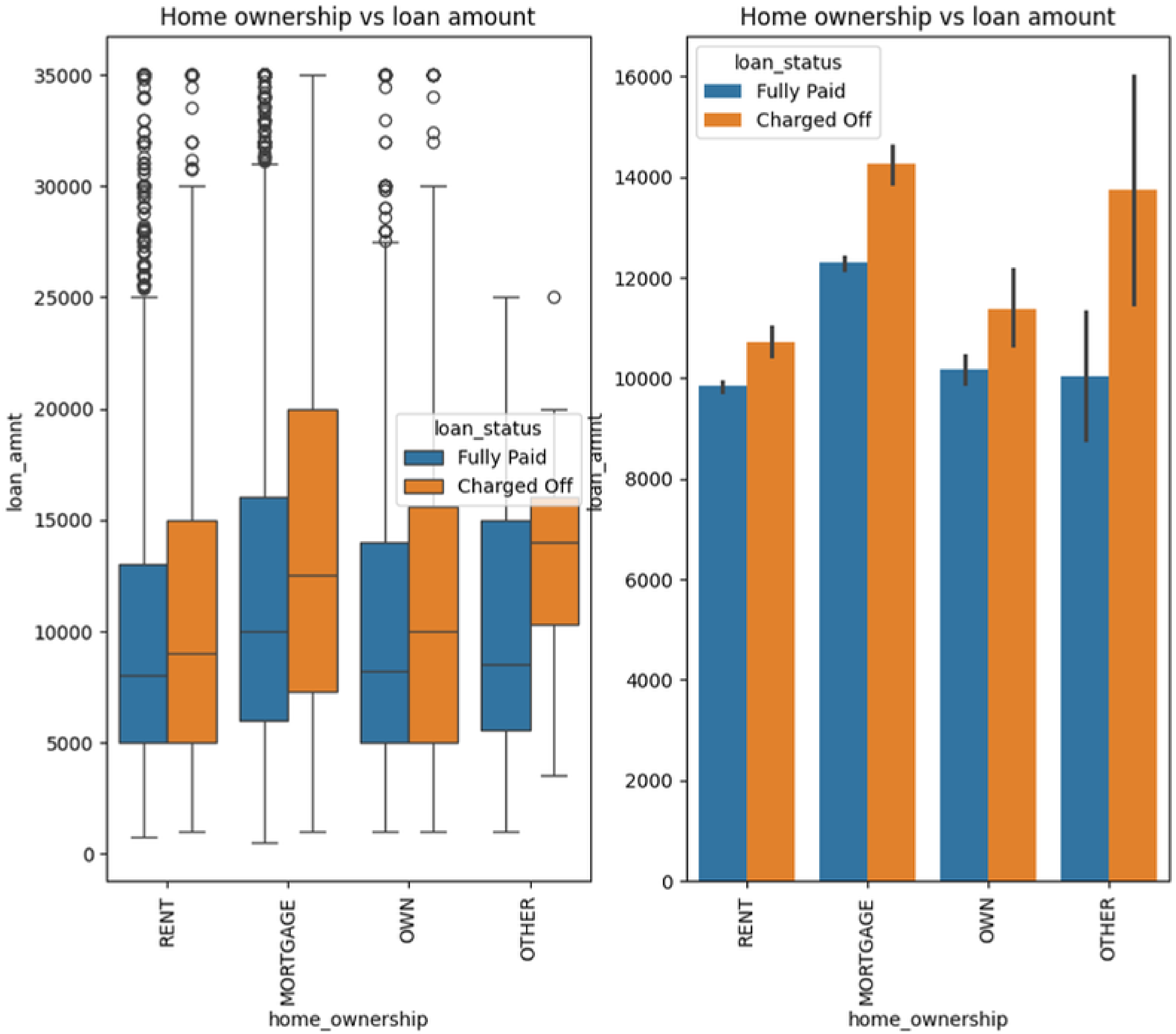


- **Loan amount is higher for lower grades like E,F,G**

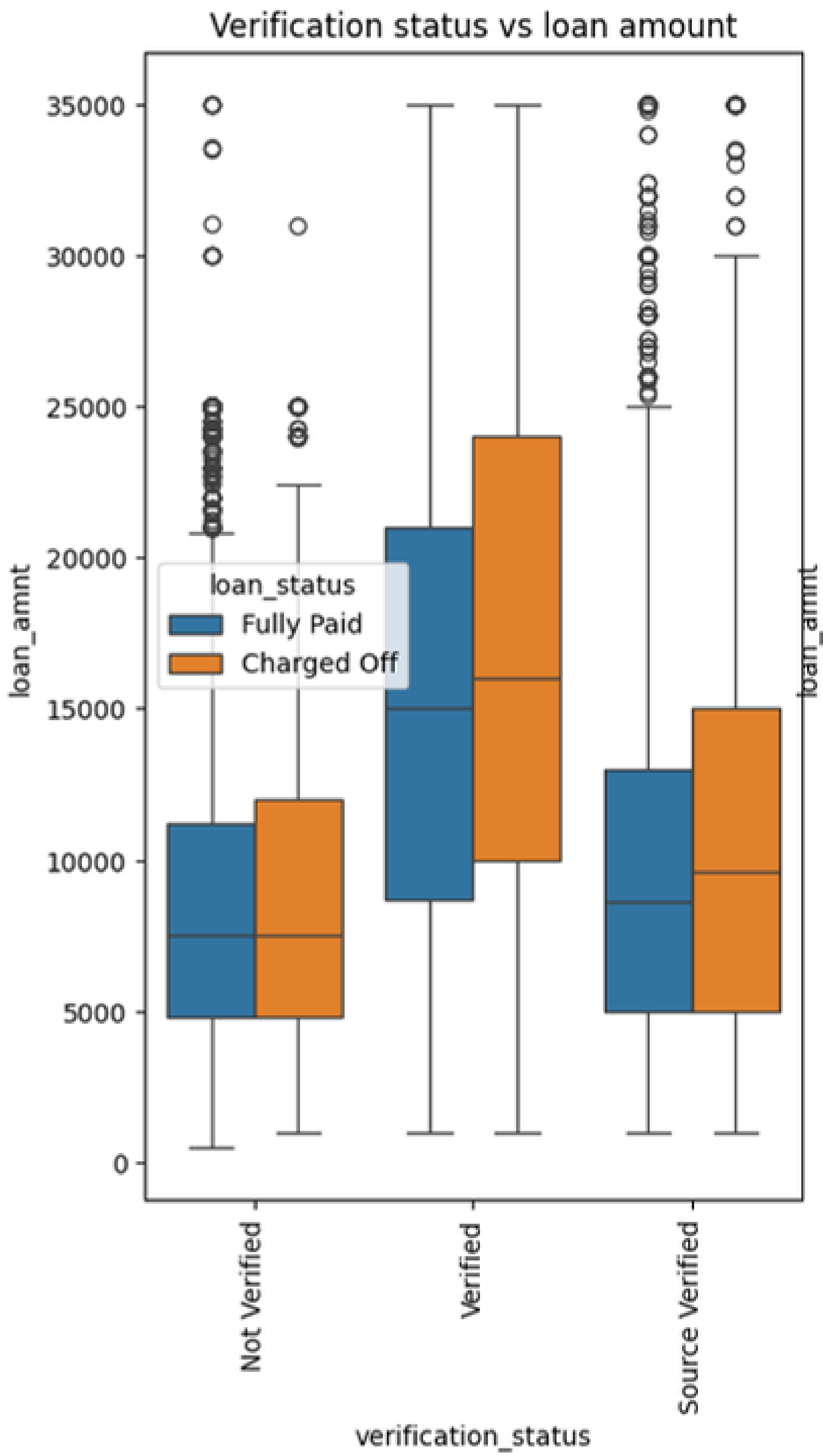


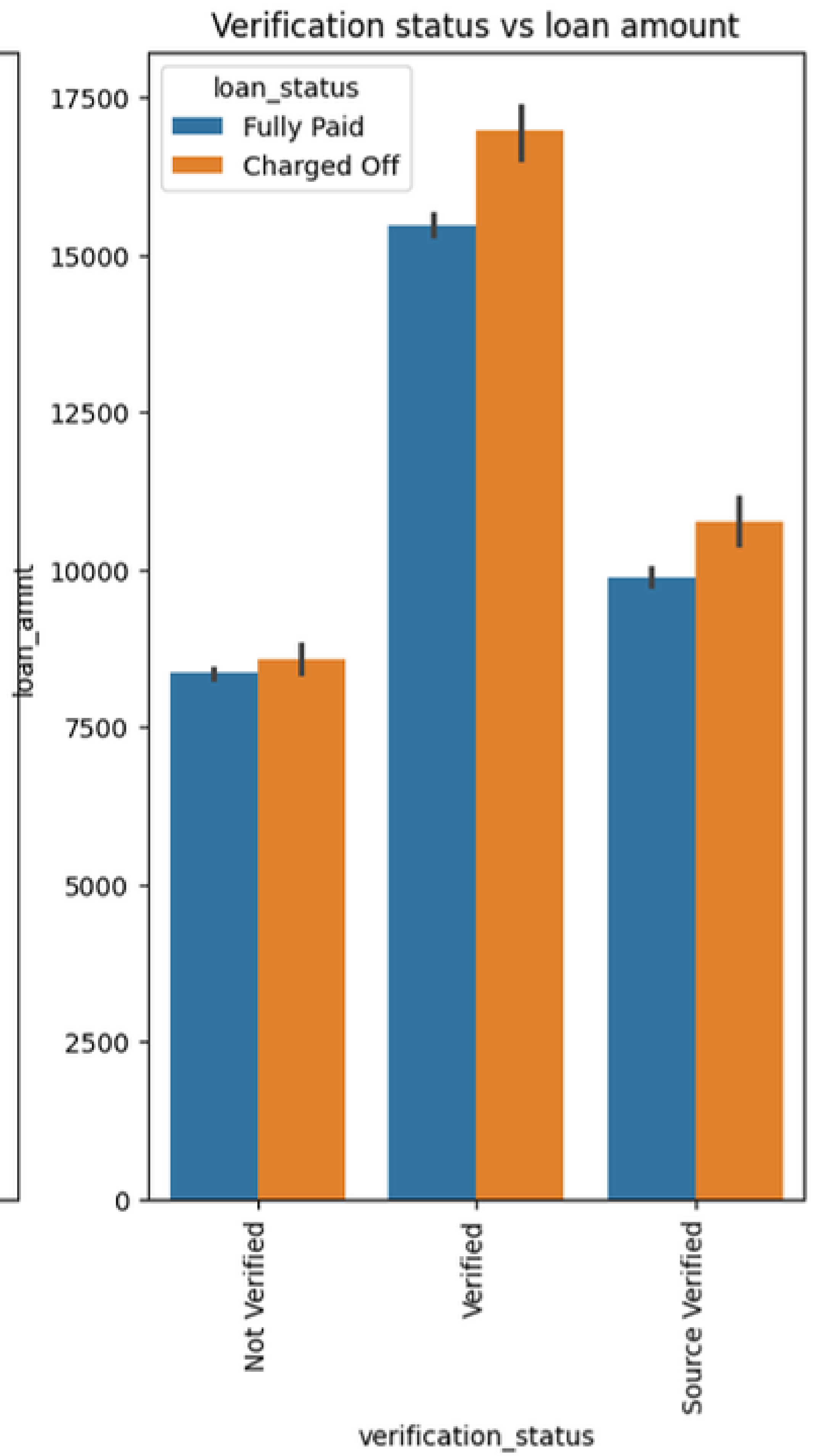


- Higher the interest rate, higher the risk of default

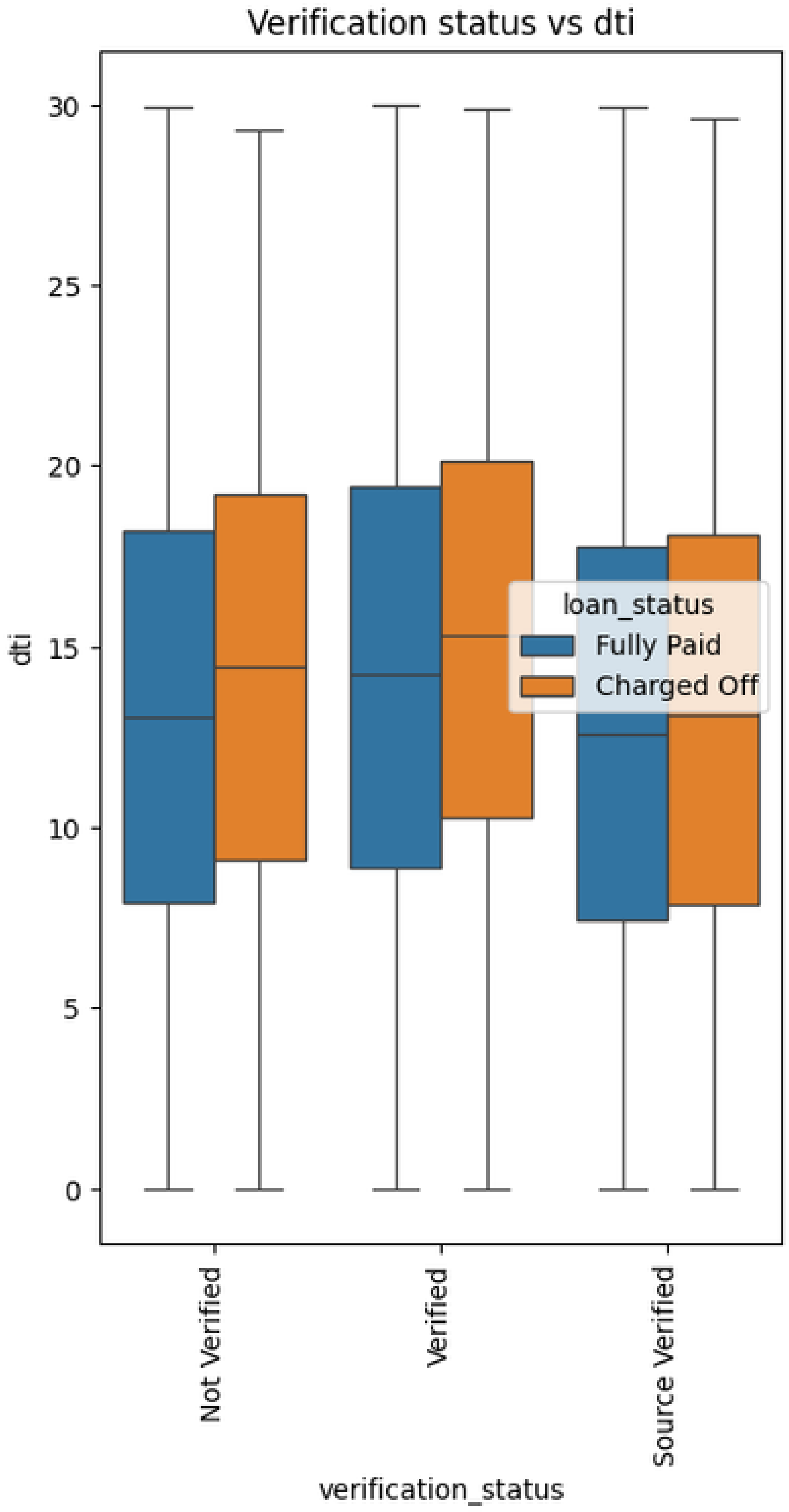


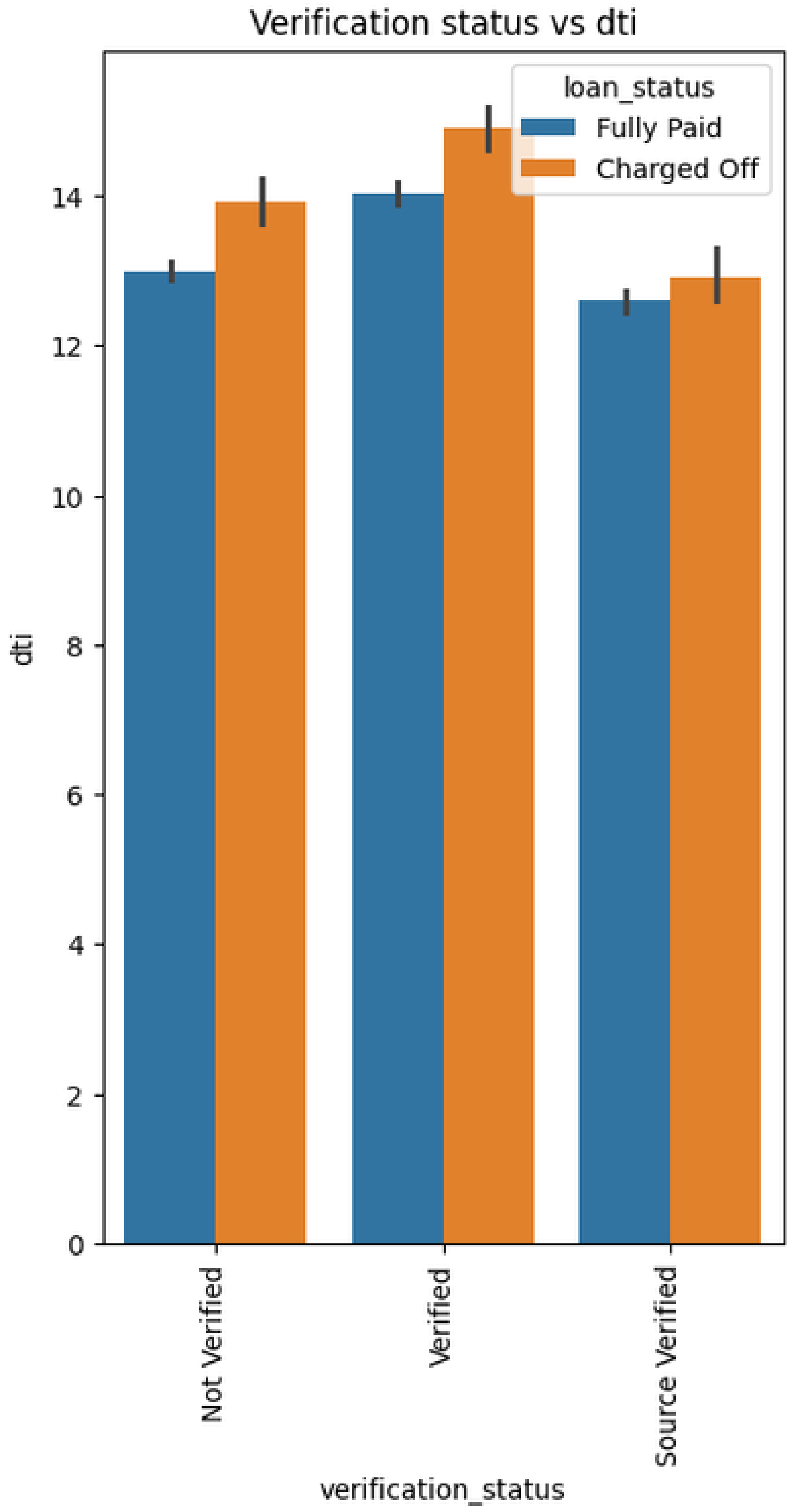
- Mortgage has the highest loan amount and highest default rate



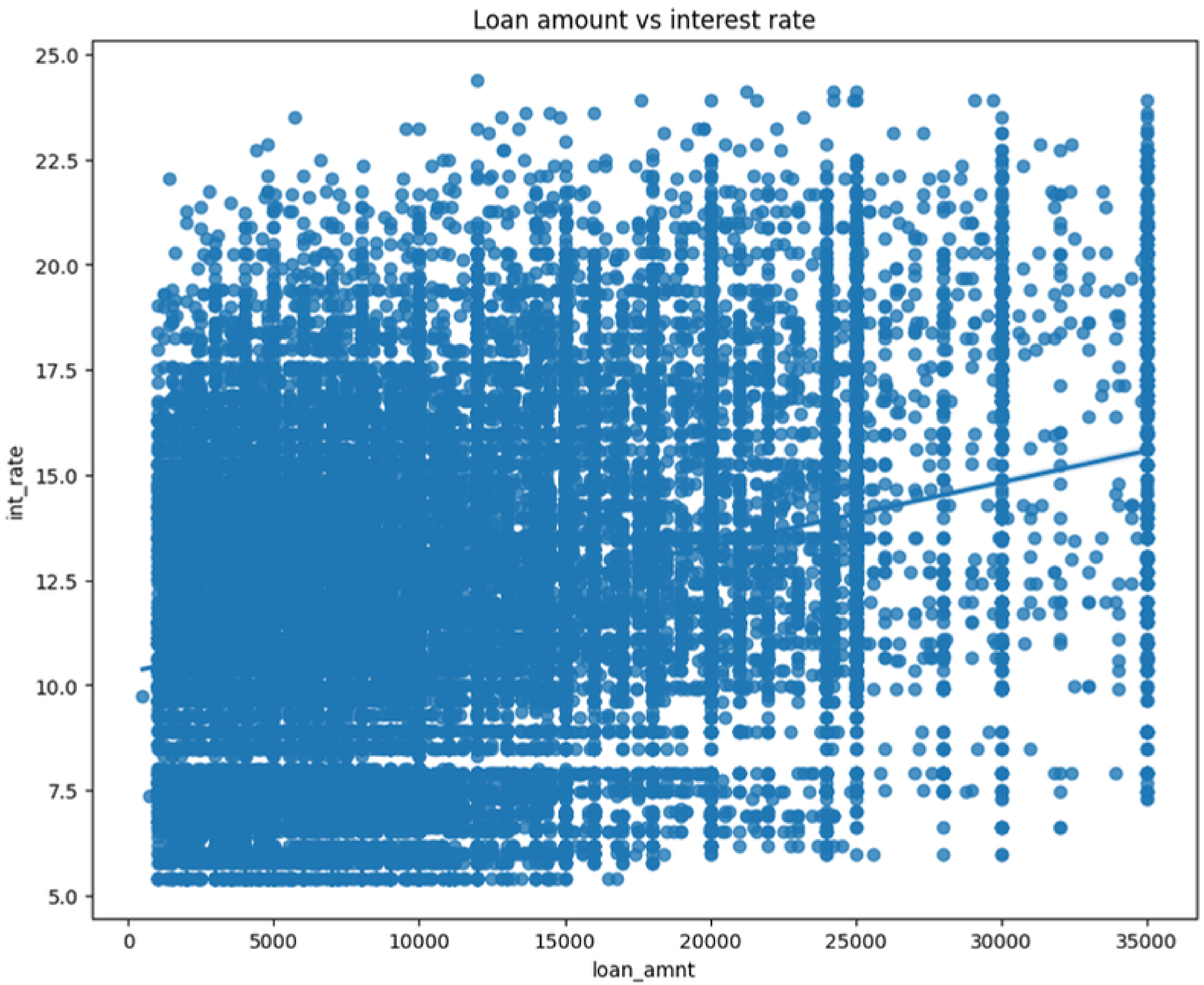


- Verified loans have higher loan amounts than not-verified loans
- More defaulters in verified loans than not verified loans

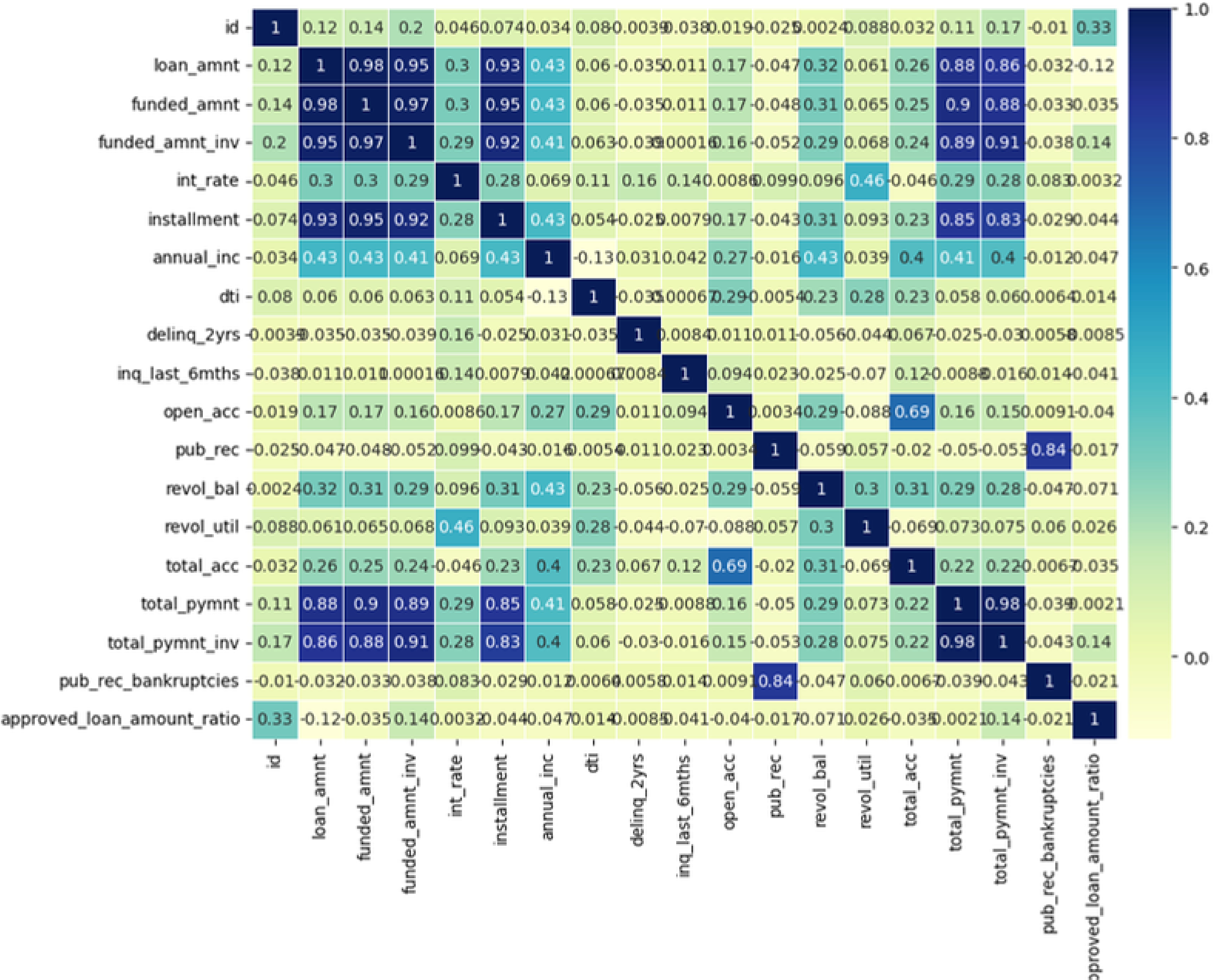




- **Defaulters have high DTI**



- As the loan amount increases, the interest rate also increases. This is because the risk of default increases with the loan amount



- The loan amount and funded amount are highly correlated
- Loan amount and installment are highly correlated
- Funded amount and installment are highly correlated
- Interest rate and grade are highly correlated
- Loan amount and approved loan amount ratio are highly correlated

Observations

- Applicants taking loans for 'home improvement' and have an income of 60k -70k
- Applicants whose home ownership is 'MORTGAGE and have an income of 60-70k
- Applicants who receive interest at the rate of 21-24% and have an income of 70k-80k
- Applicants who have taken a loan in the range 30k - 35k and are charged interest rate of 15-17.5 %
- Applicants who have taken a loan for a small business and the loan amount is greater than 14k
- Applicants whose home ownership is 'MORTGAGE and have a loan of 14-16k
- When a grade is F and the loan amount is between 15k-20k
- When employment length is 10yrs and loan amount is 12k-14k
- When the loan is verified and the loan amount is above 16k
- For grade G and interest rate above 20%

Observations

Strong correlation

- The loan amount and funded amount are highly correlatedLoan amount and installment are highly correlated
- Funded amount and installment are highly correlated
- Interest rate and grade are highly correlated
- The loan amount and approved loan amount ratio are highly correlated

Moderate correlation

- 1. Loan amount and int_rate are moderately correlated
- 2. Loan amount and DTI are moderately correlated
- 3. Funded amount and int_rate are moderately correlated
- 4. Funded amount and DTI are moderately correlated
- 5. Int_rate and DTI are moderately correlated

Observations

Low correlation

- 1. Loan amount and annual income are low correlated
- 2. Funded amount and annual income are low correlated
- 3. Int_rate and annual income are low correlated
- 4. Dti and annual income are low correlated
- 5. Approved loan amount ratio and annual income are lowly correlated

Income and DTI

Lower annual income is associated with higher DTI, indicating a financial strain.

Income and Employment

There is a positive correlation between annual income and employment years, suggesting that income tends to increase with work experience.

Observations

Loan Status
Approximately 14% of loans were charged off.

Loan Purposes
Debt consolidation and paying credit card bills are the most common loan purposes.
These purposes also have a relatively high number of charged-off loans.

Home Ownership
Most borrowers either rent or have mortgaged their homes.
These categories also have a relatively high number of charged-off loans.

Observations

Loan Term

Loans with a 60-month term have a higher percentage of applicants charged off compared to 36-month loans.

Income and Default Risk

Higher-income ranges, especially $80,000 and above, have lower chances of loan default.

Loan Purpose and Default Risk

Small business applicants have a higher likelihood of loan default.
Renewable energy loans have a comparatively lower default rate.

Grade and Default Risk

Loan grades "A" have lower chances of default, while grades "F" and "G" have higher chances.

Conclusions

- Lending clubs should reduce the high-interest loans for 60 months tenure, as they are prone to loan default.
- Grades are a good metric for detecting defaulters. The lending club should examine more information from borrowers before issuing loans to Low grades (G to A).
- Lending Club should control the number of loans issued to borrowers who are from CA, FL, and NY to make profits.
- Small business loans are defaulted more. The lending club should stop/reduce issuing loans to them.
- Borrowers with mortgage homeownership are taking higher loans and defaulting on the approved loans. The lending club should stop giving loans to this category when the loan amount requested is more than 12000.
- People with more public derogatory records have having higher chance of filing a bankruptcy. The lending club should make sure there are no public derogatory records for the borrower.

Conclusions

Correlations and Relationships:

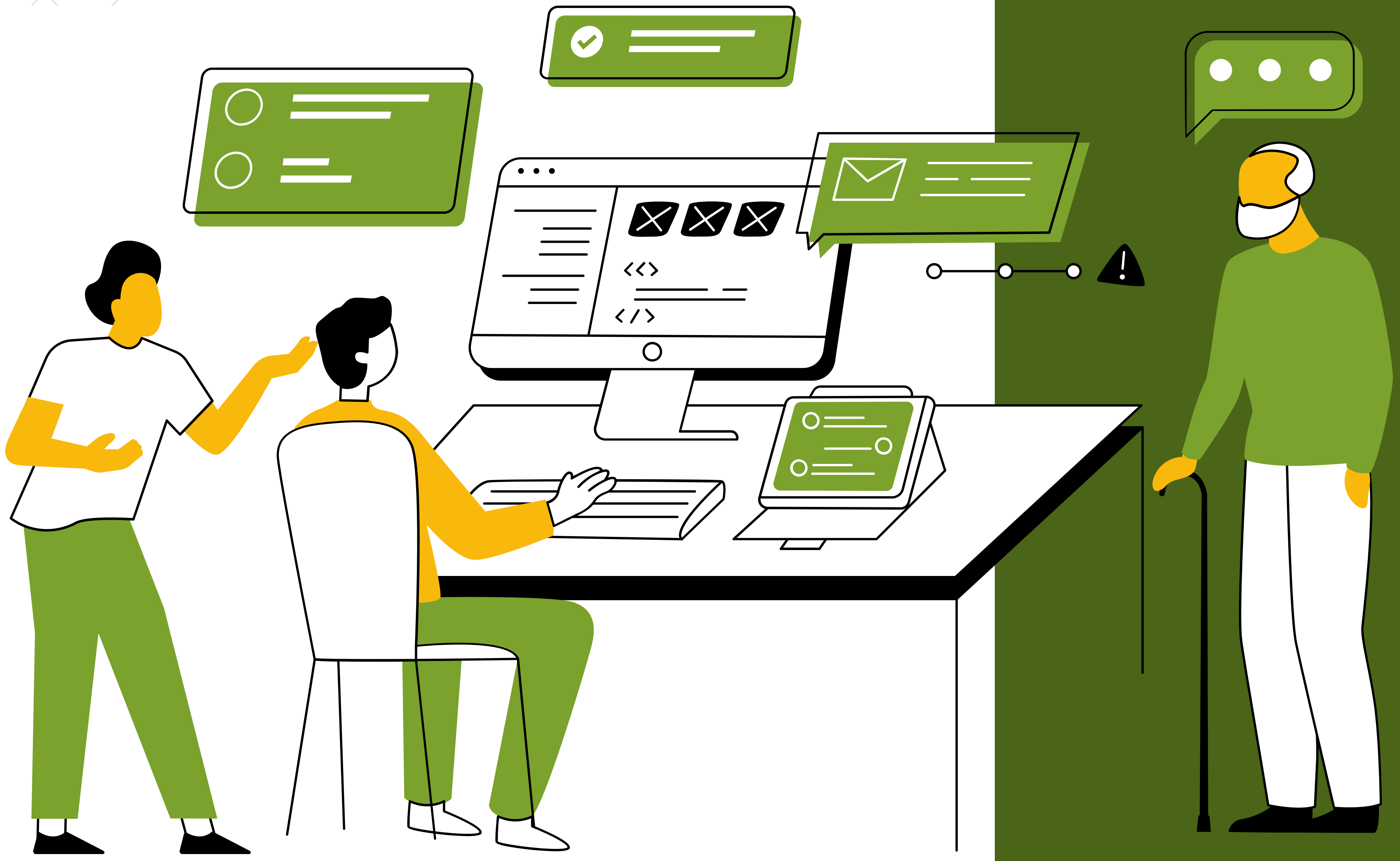
- Strong correlations exist between loan amount, investor amount, and funding amount.
- Annual income is inversely related to the debt-to-income ratio (DTI).

Income and Employment:

- Annual income tends to increase with years of employment, suggesting a positive correlation between income and experience

Risk Factors:

- Loan term impacts default rates, with longer-term loans having higher default rates.
- Higher-income ranges are associated with lower default risk.
- Loan purpose, grade, interest rate, and employment history significantly affect default risk.



THANK YOU