Lending Club Case Study

Lipika Panda

Overview

Context - Lending Club is a consumer finance company which specialises in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant’s profile.

Problem Statement - Lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

Solution - To identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study i.e. understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.

Analysis Steps

01

Data Understanding

Understand the business problem and dataset. Use data dictionary.

02

Data Cleaning

. Dropping irrelevant columns and records.
. Missing value treatment
. Standardising columns
. Getting derived columns
. Outlier treatment

03

Univariate & Segmented Univariate Analysis

Perform univariate and segmented univariate analysis on both categorical and numerical variables.

04

Bivariate & Multivariate Analysis

Perform bivariate and multivariate analysis to understand the relation between variables.

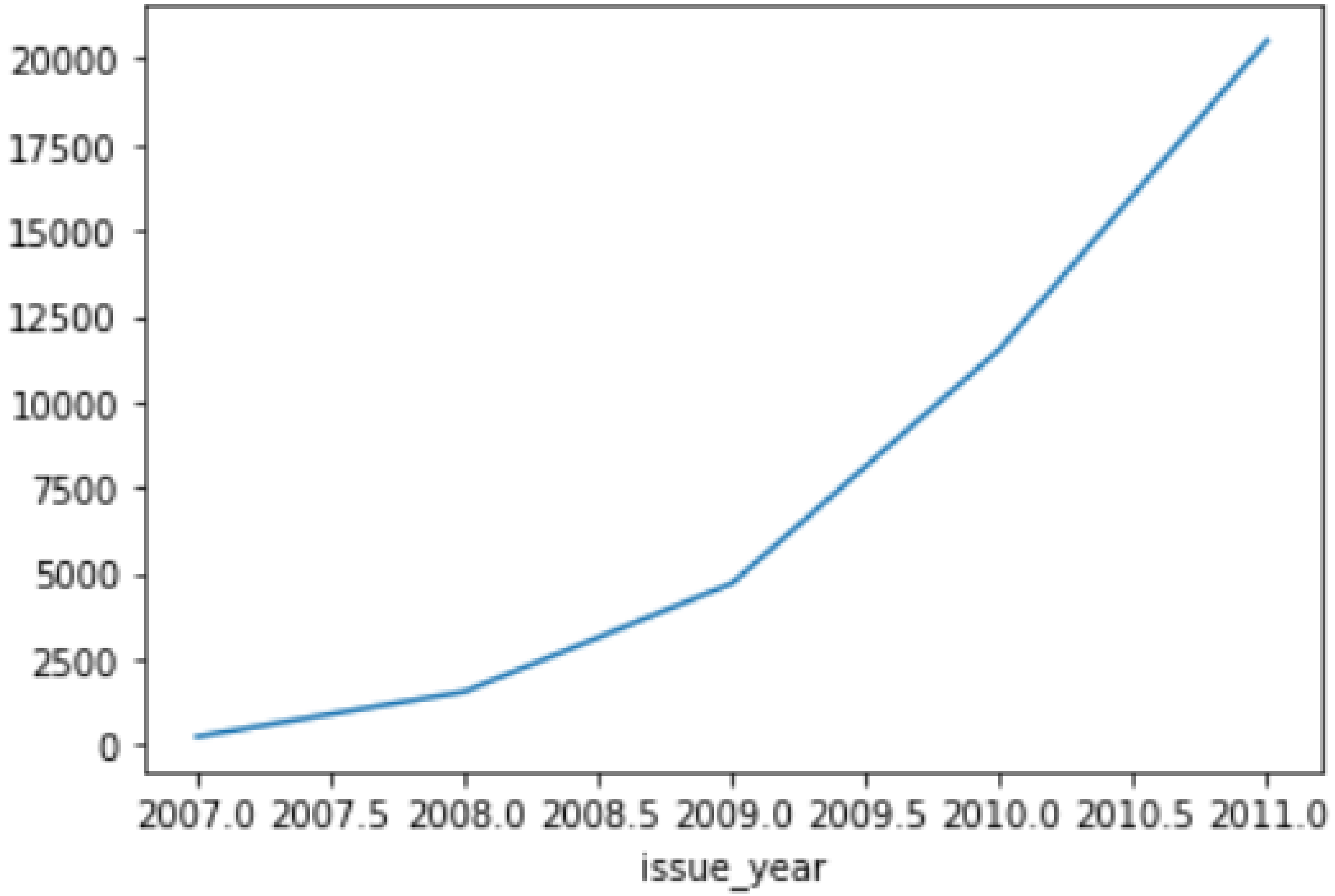
05

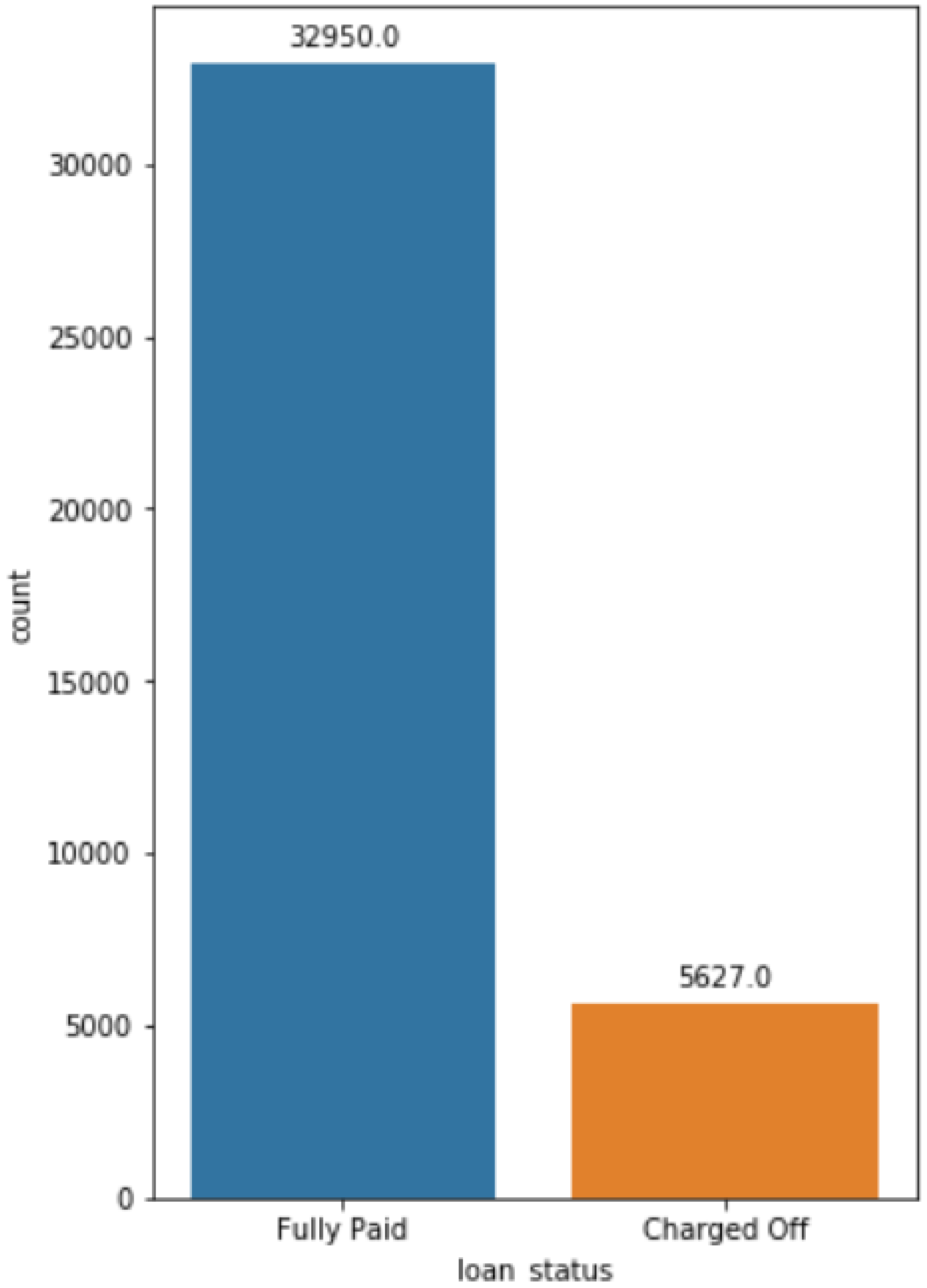
Recommendations

Inference from the analysis and visual plots.

Loan Status

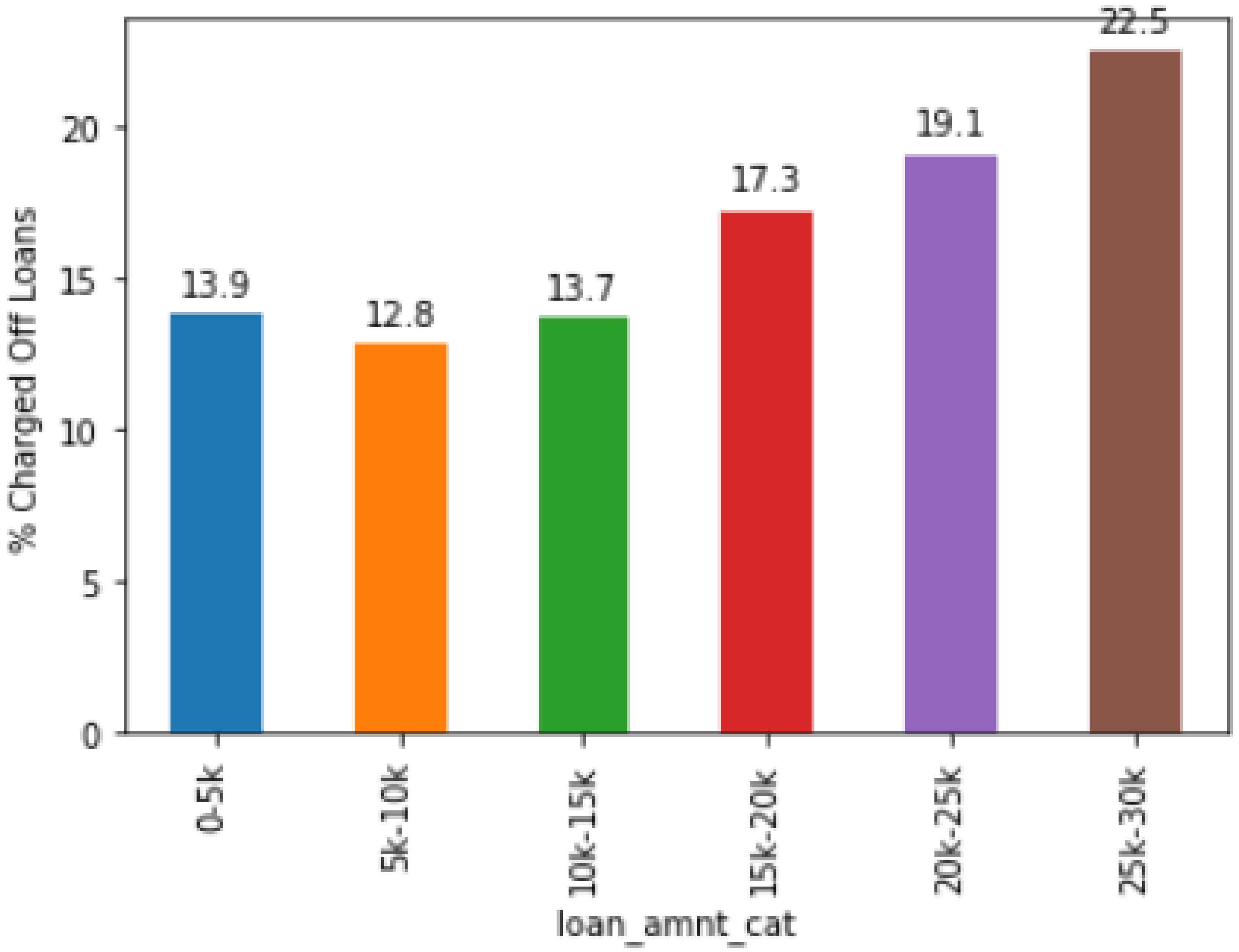
- Most of the loans are Fully Paid.
- About 15%(5627 out of 38577) of loan are having status as defaulters.
- Loans issued is showing a raise over the years.

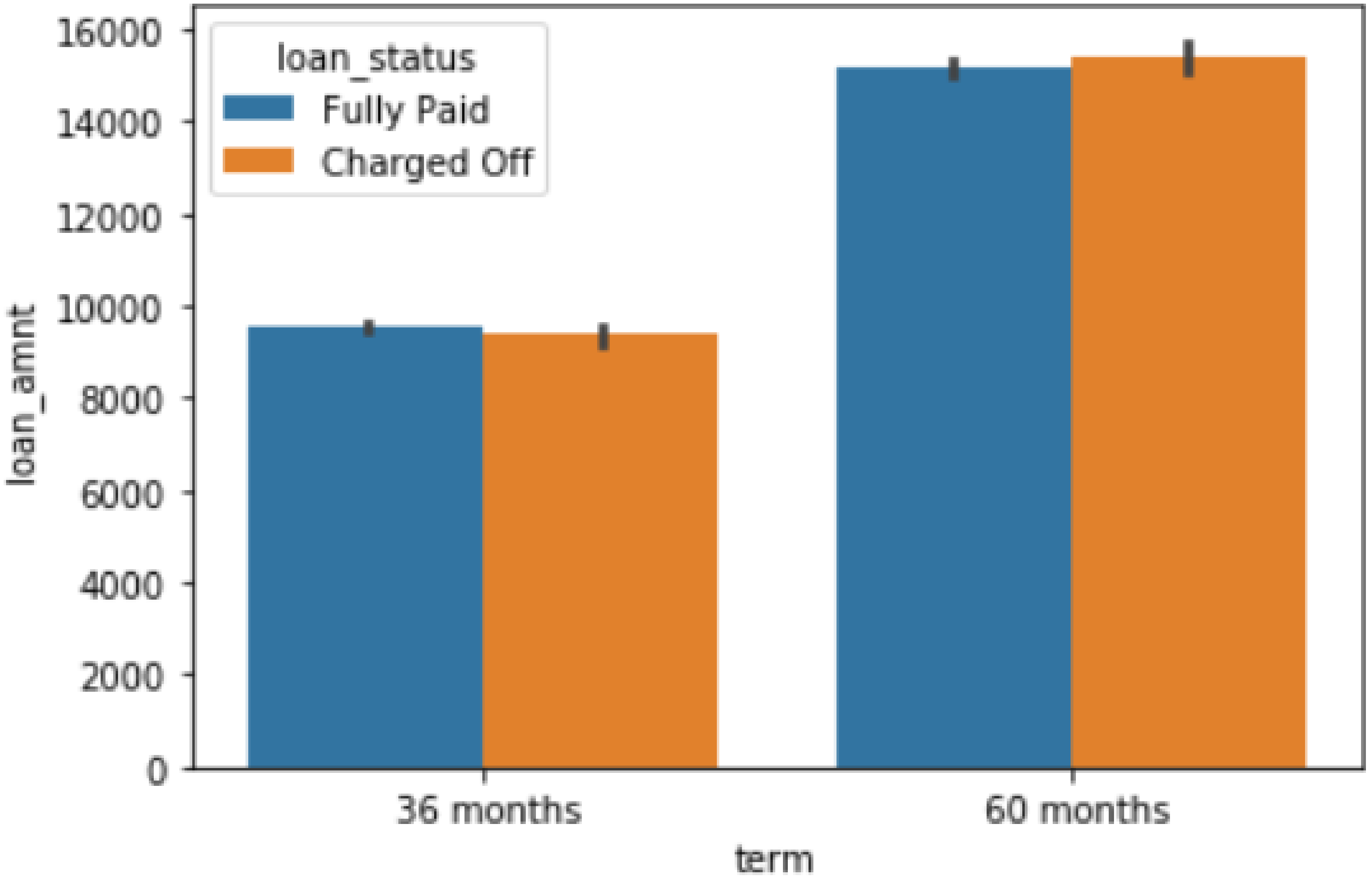




Loan Amount

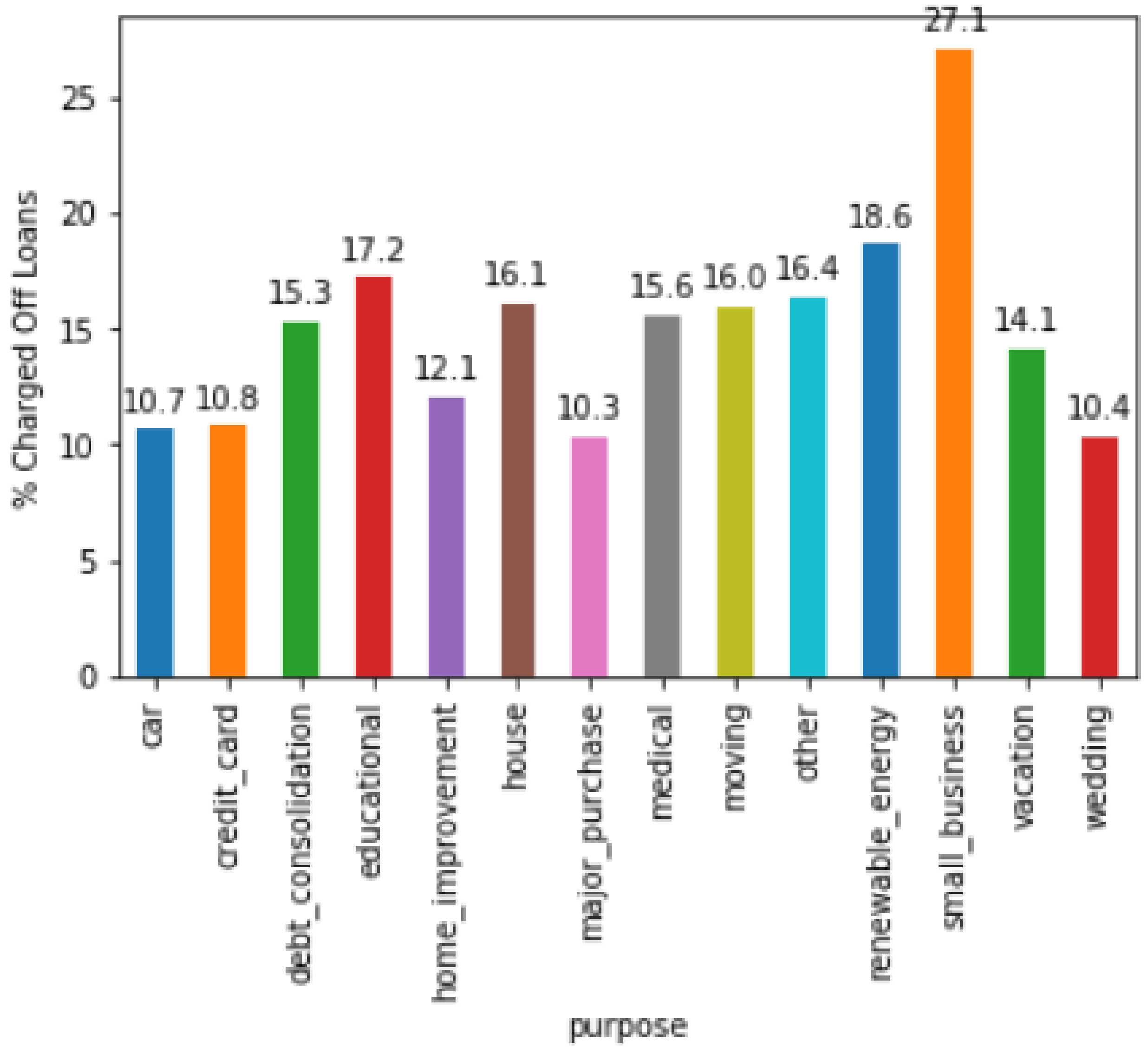
- Higher loan amounts have higher percentages of defaults.

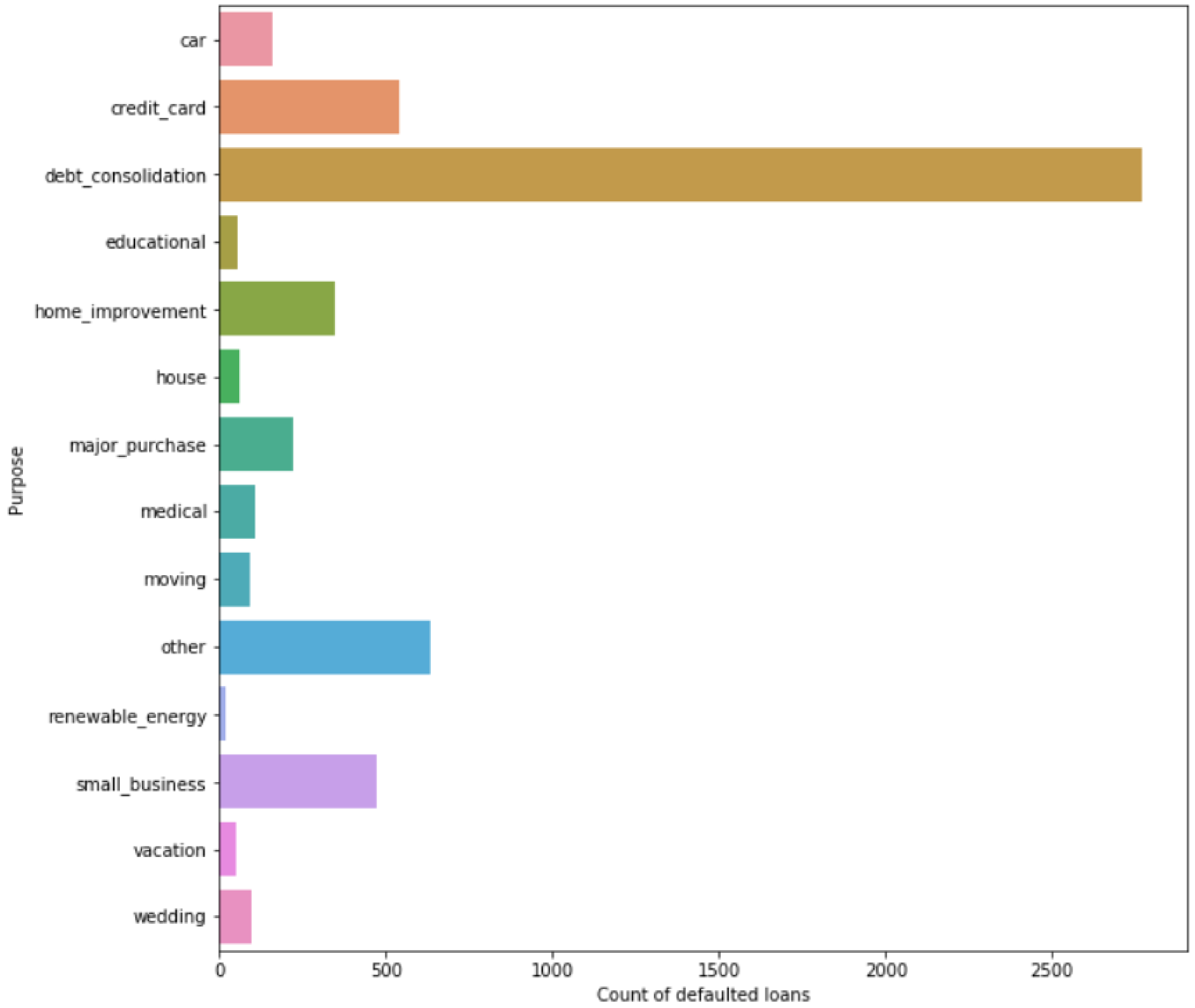




- Higher loan amounts also have higher terms and have higher defaults.

Loan Purpose

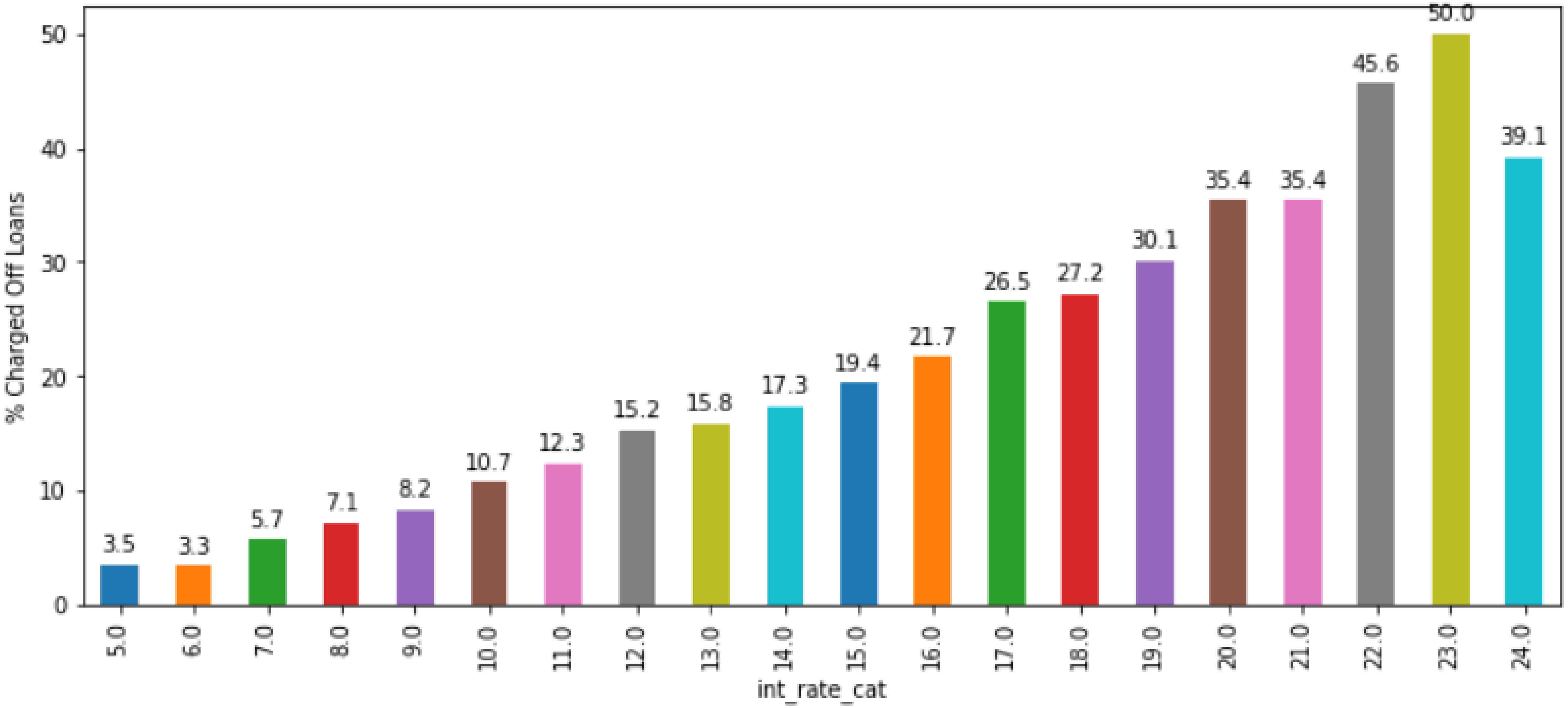




- There are high number of defaulted loans taken on the purpose of Debt Consolidation.
- The percentage of defaulters is however high incase of small business.

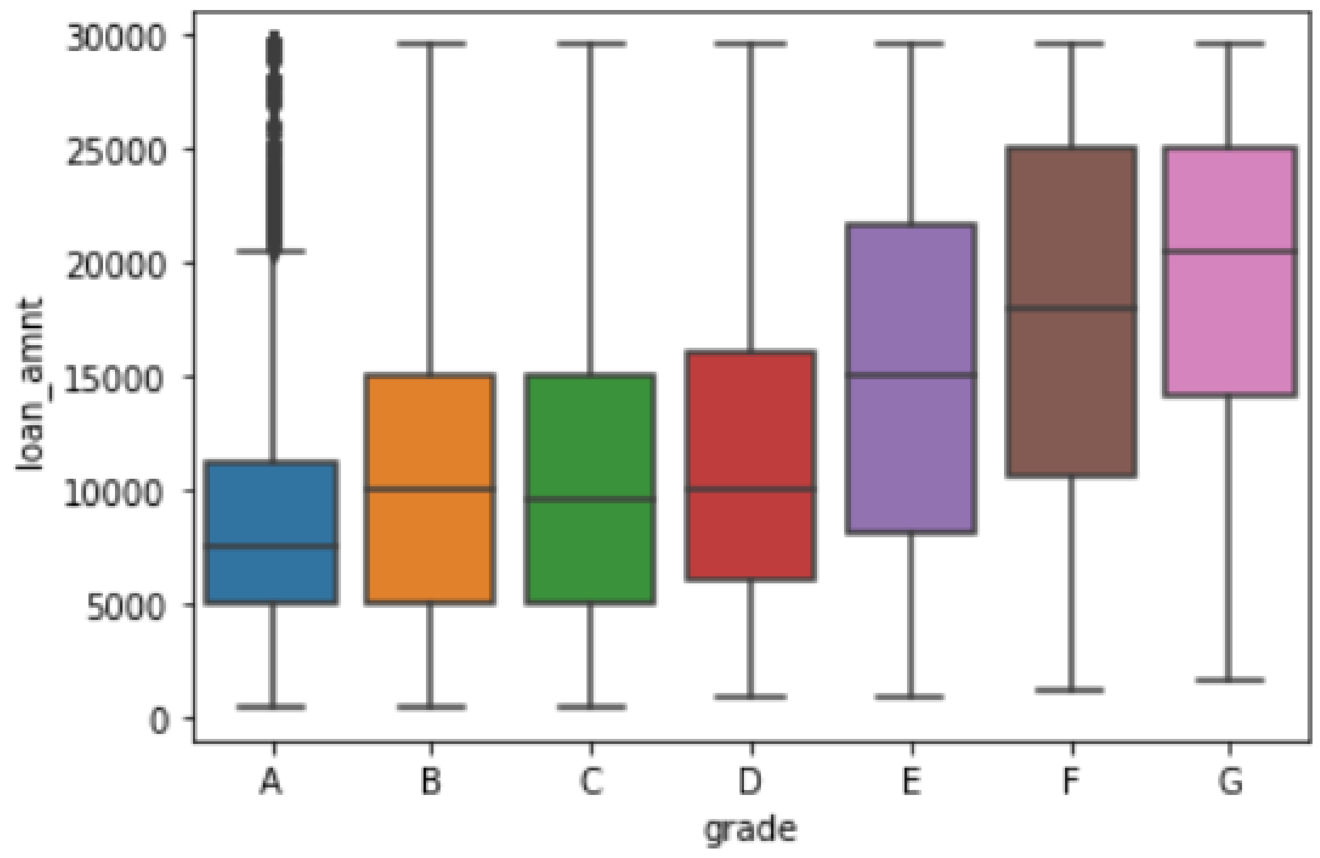
Intrest Rate

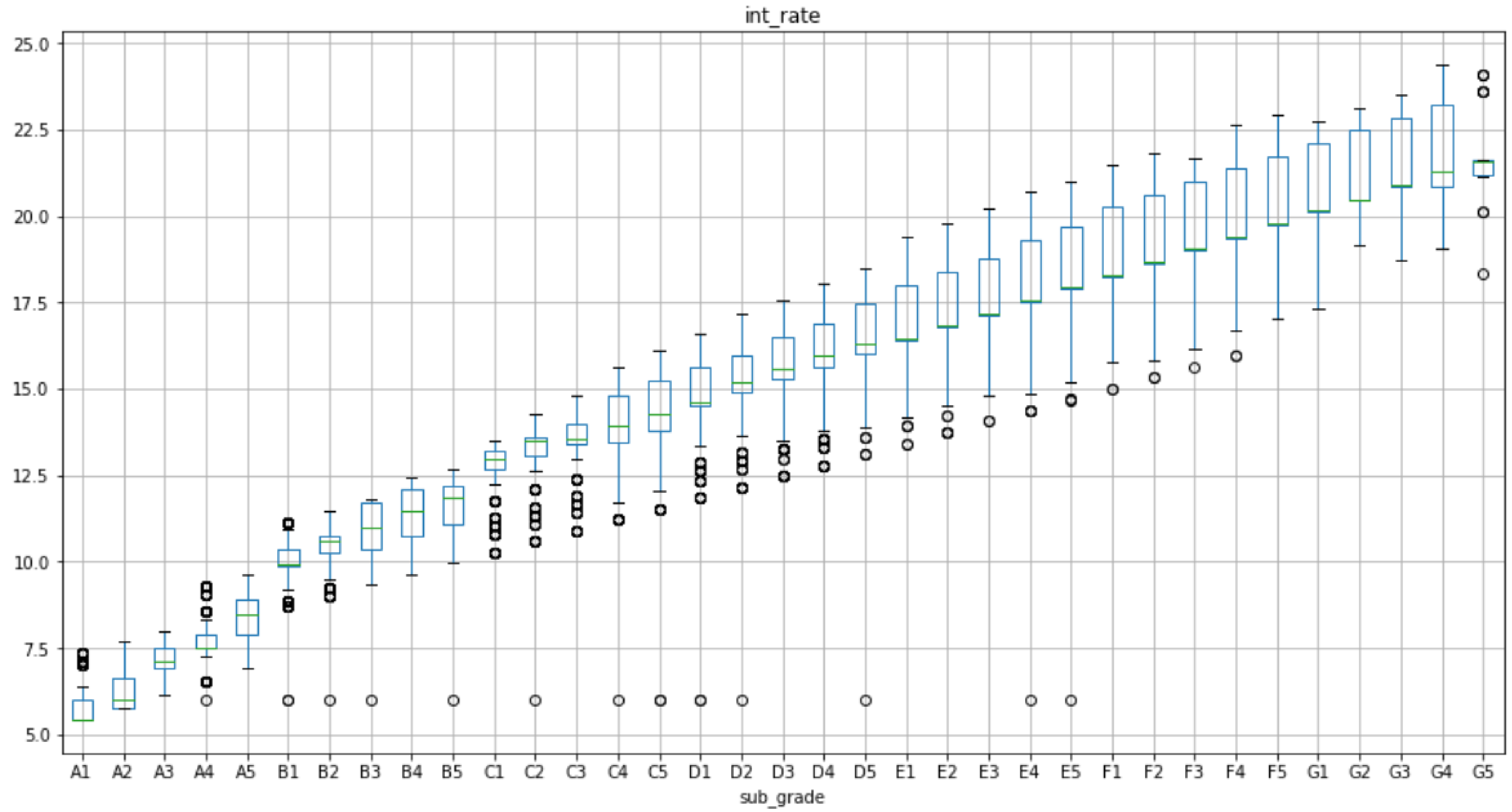
- With increase in intrest rate the defaulters also increase.

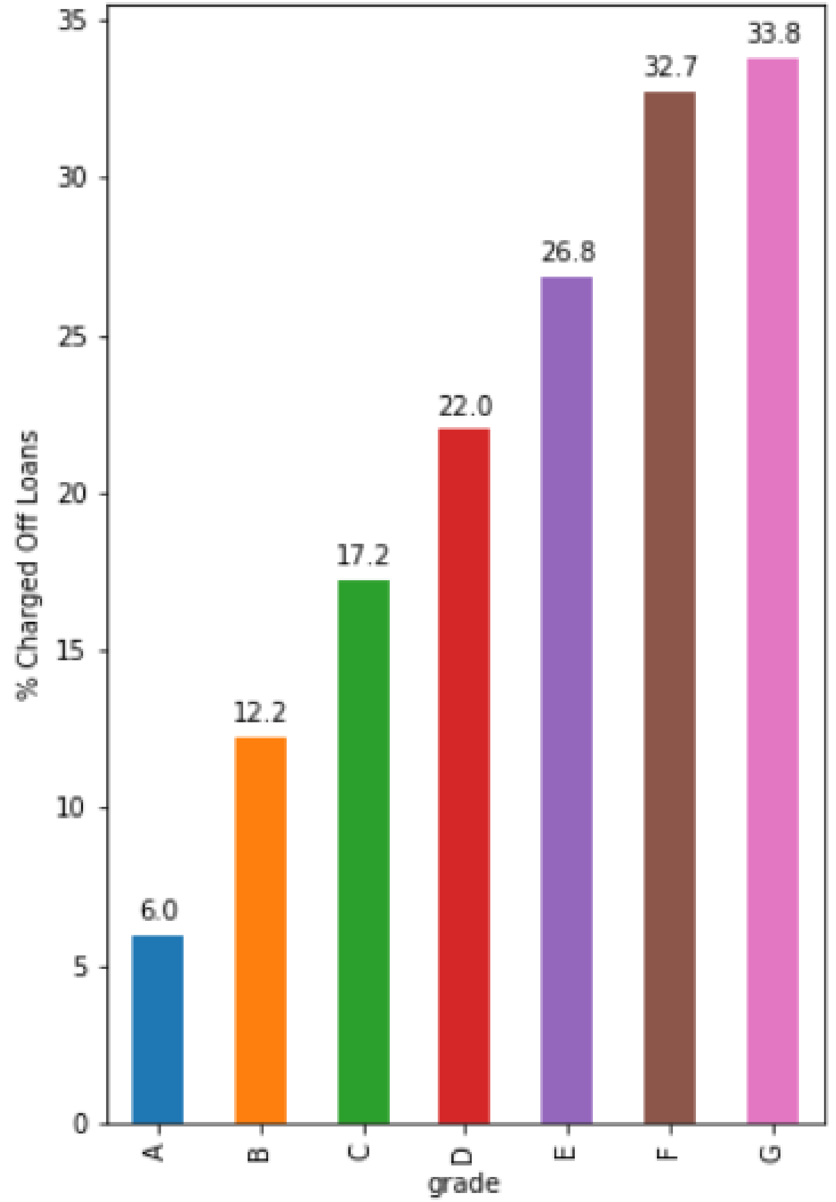


Grades

- Lower grades have higher percentage of defaulters.
- We observe a continious increase in intrest rates for lower grades.
- Also lower grades apply for high loan amount which in turn defaults.

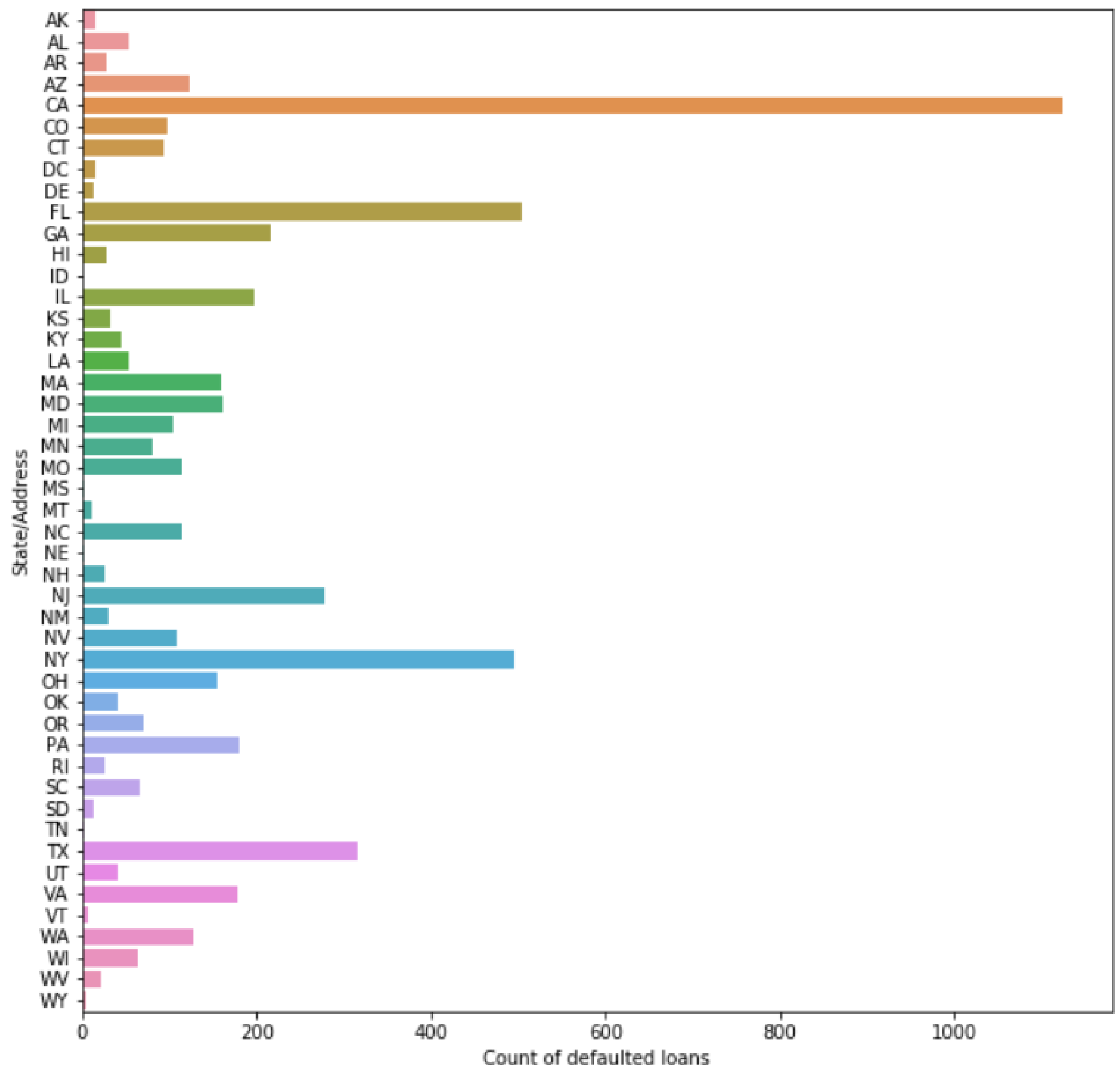






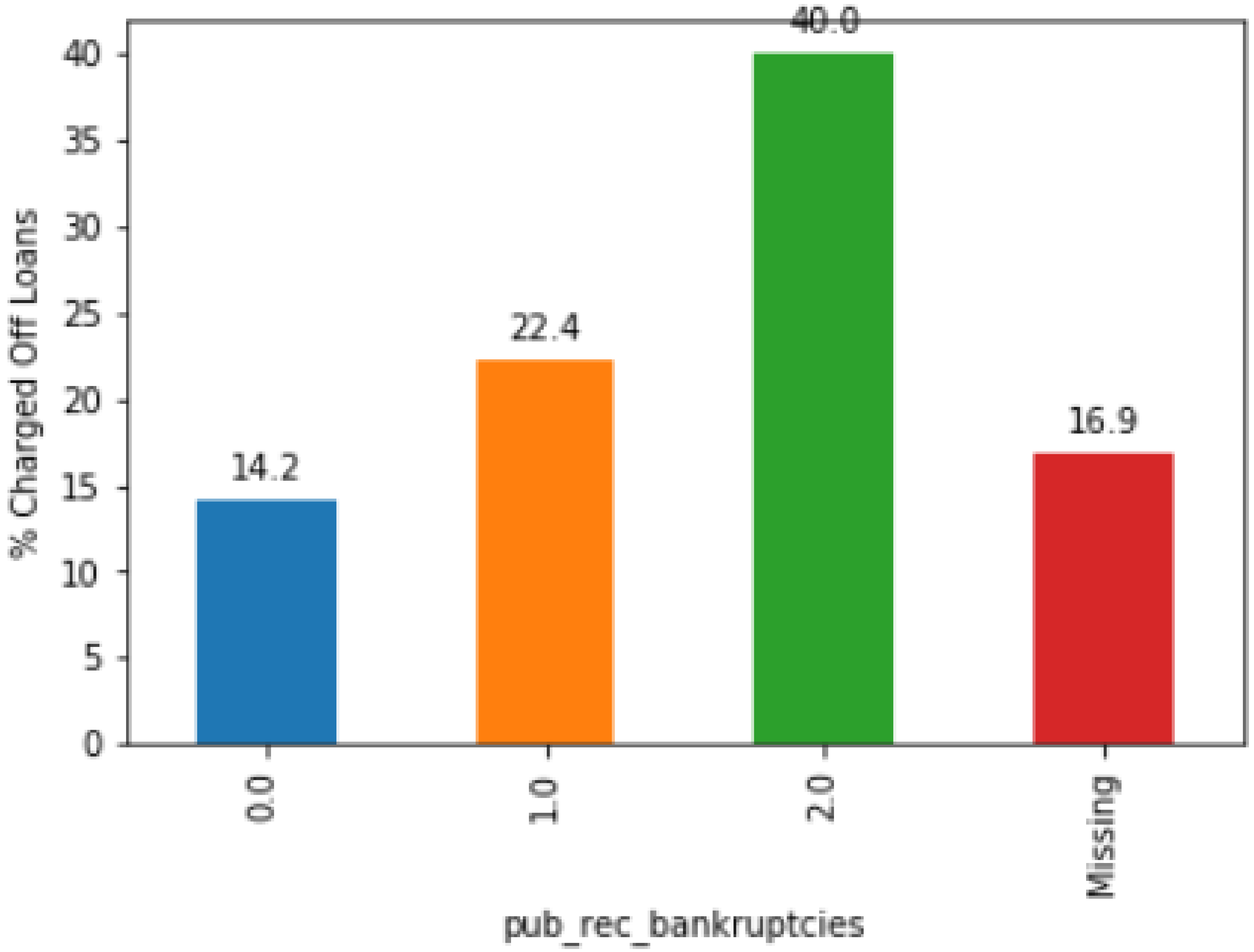
State

- There is a high number of defaulted loans from California state followed by Florida and New York

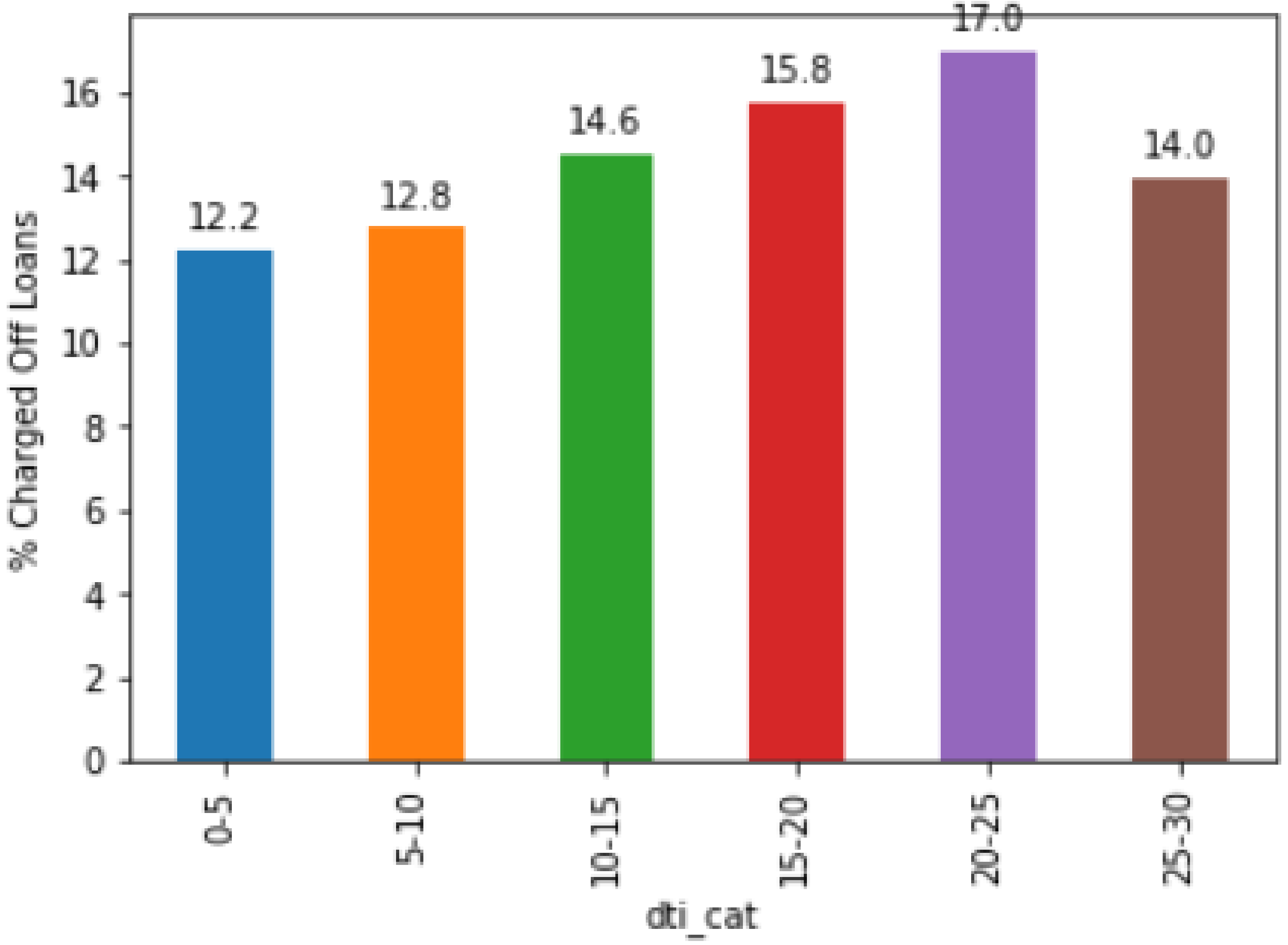


Public Record Bankrupties

- Candidates with prior public records of bankrupties have higher default rates.



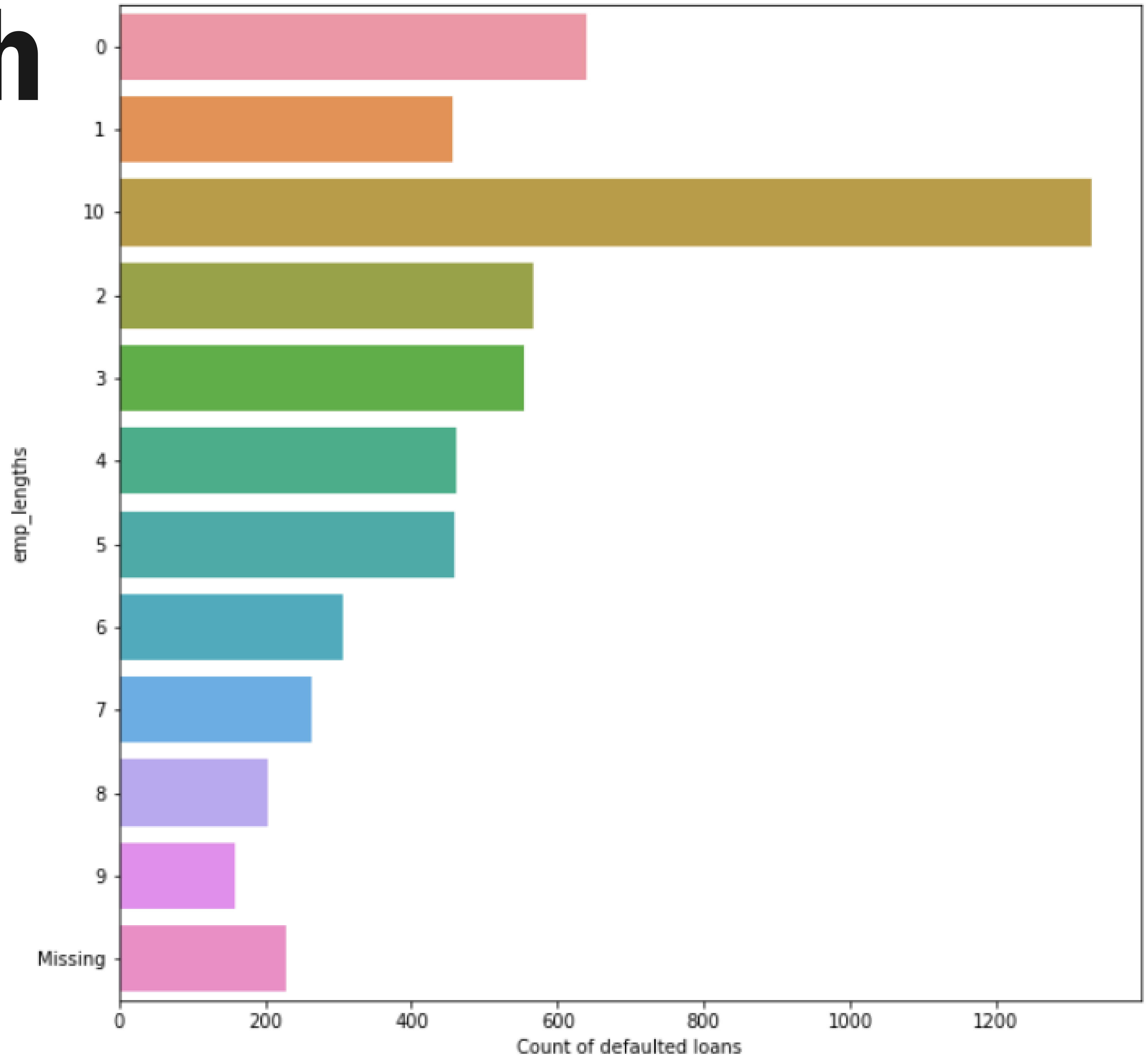
Debt To Income Ratio



- With higher debt to income ratio, the risk of defaulting increases. We observe the highest percentage of defaulters in the bucket 20-25.

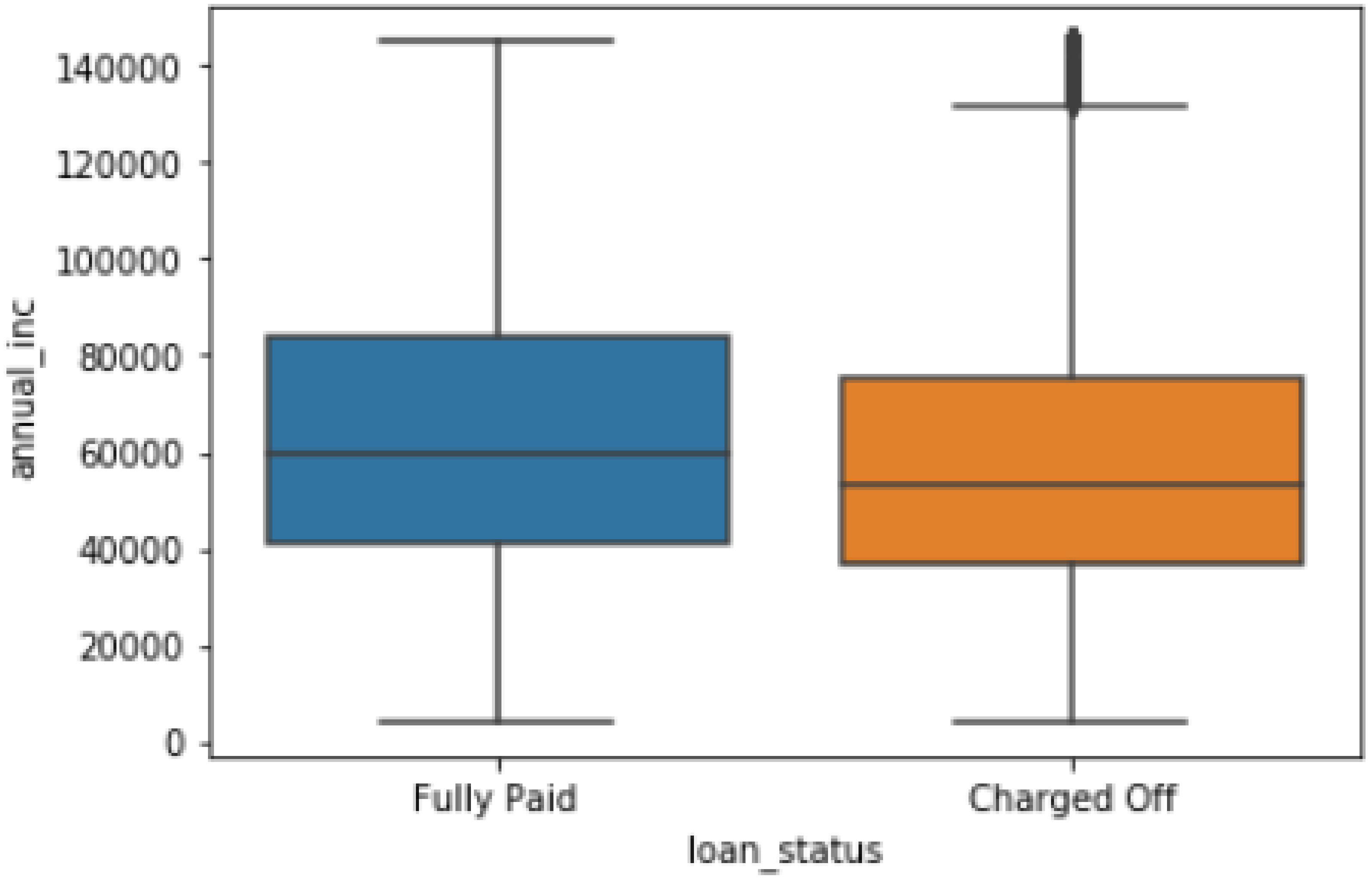
Employee Length

- Maximum of loan applicants are from 10+ years group.
- 10+ years group loan applicants default the most.



Annual Income

- Applicants with higher annual income tend to payback the loan.



Recommendations

01 Applicants from lower grades or with public record bankrupties should not be approved.

02 Loans for the purpose of debt consolidation and small business cause defaulters.

03 Applicants from Calfornia state have high count of defaulters and hence should be checked on.

04 Applicants with higher rate of intrests tend to default more especially as the grade lowers.

05 Applicants with higher annual income should be preferred as they tend to payback. However, those with higher debt to income ratio should be reduced.

Thank you!