

Lending Club Case Study

Sonipriya Singh

Anuja Kudumula

Problem Statement

- Our objective is to assist a finance company in making smarter lending decisions by predicting loan defaults from past data.

Business Objective

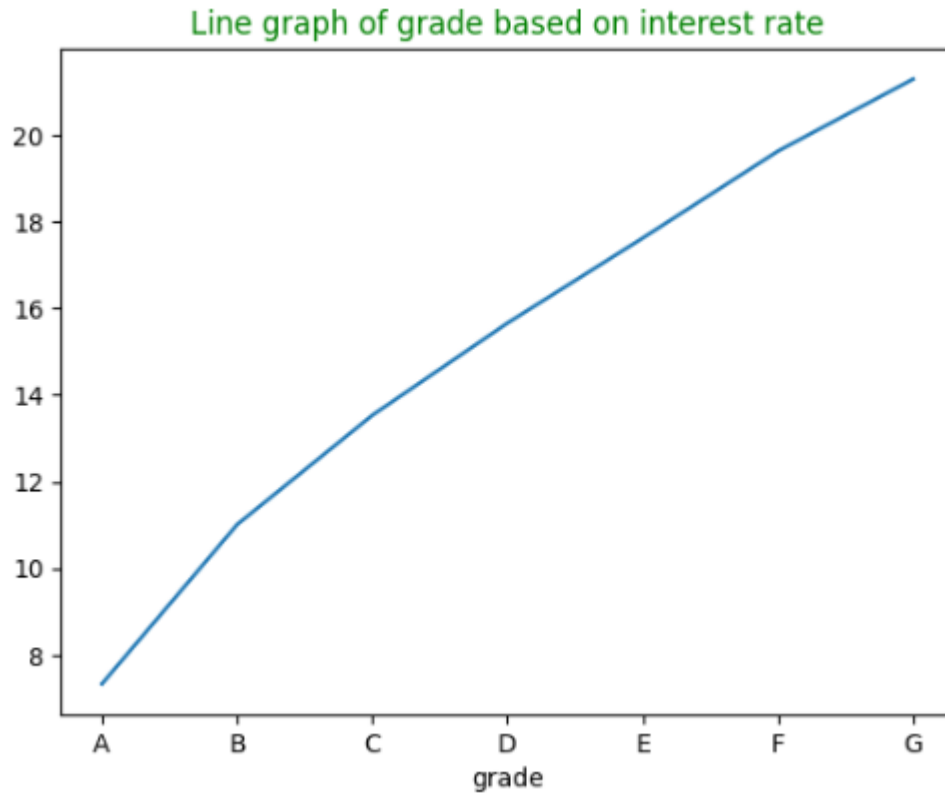
- Identify key factors driving loan defaults to minimize credit losses and improve risk assessment for the finance company.

Approach

- Data cleaning and manipulation
- Handling missing values & outliers
- Univariate analysis
- Segmented univariate analysis
- Bivariate analysis
- Multivariate analysis

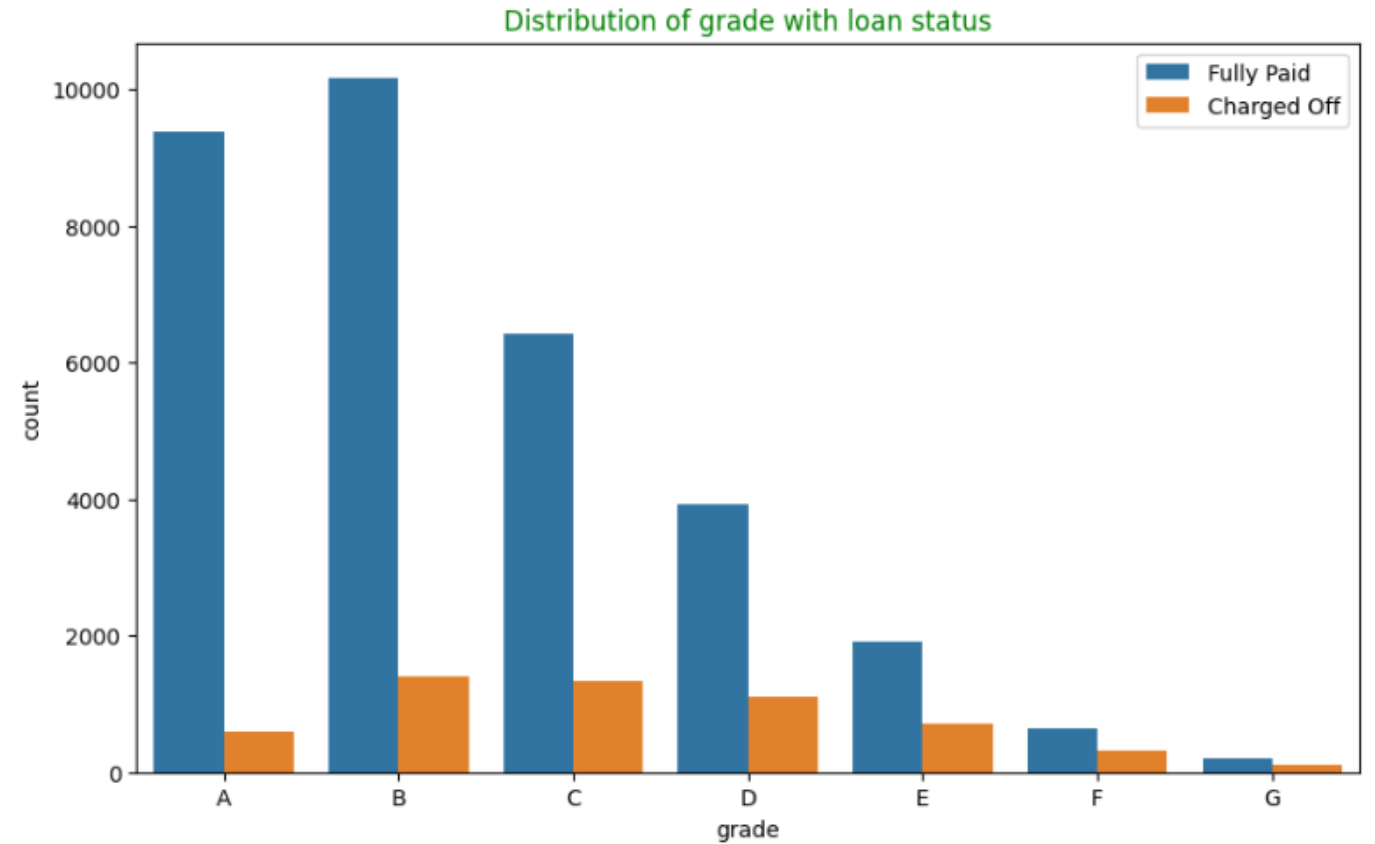
Analysis

- Grade and Interest rate



As grade increases(meaning from A to G), interest rate has also increased

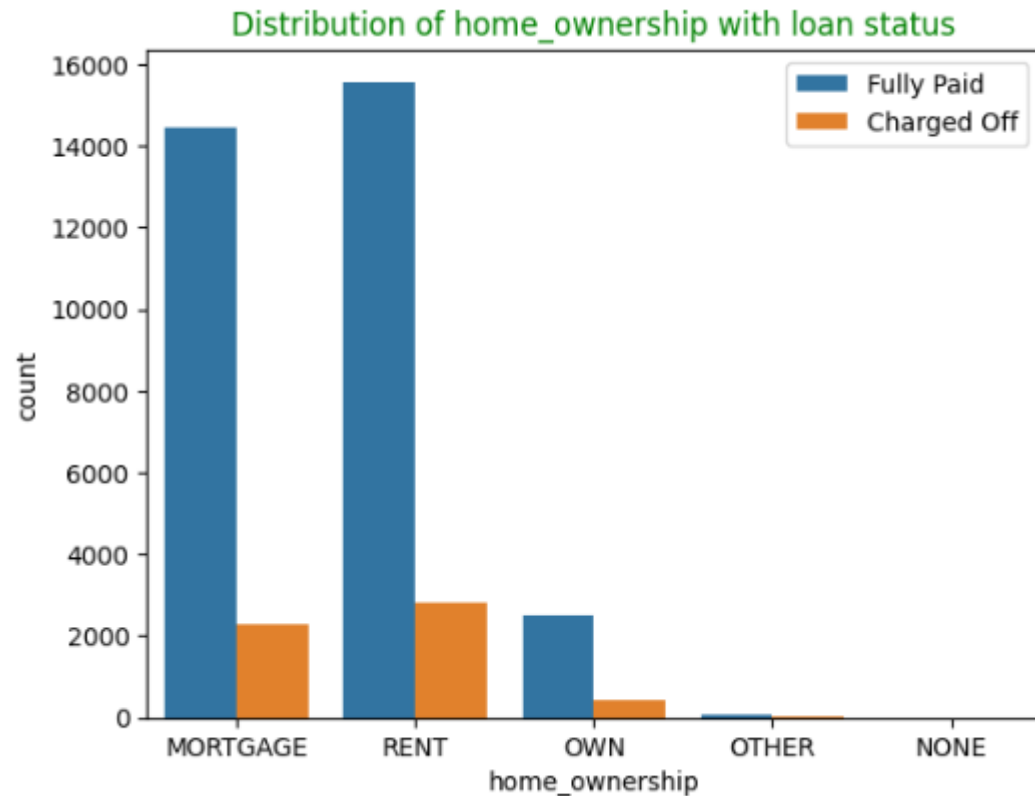
- Grade and loan status



Borrowers with higher Grade E,F,G have higher chance of defaulting the loan.

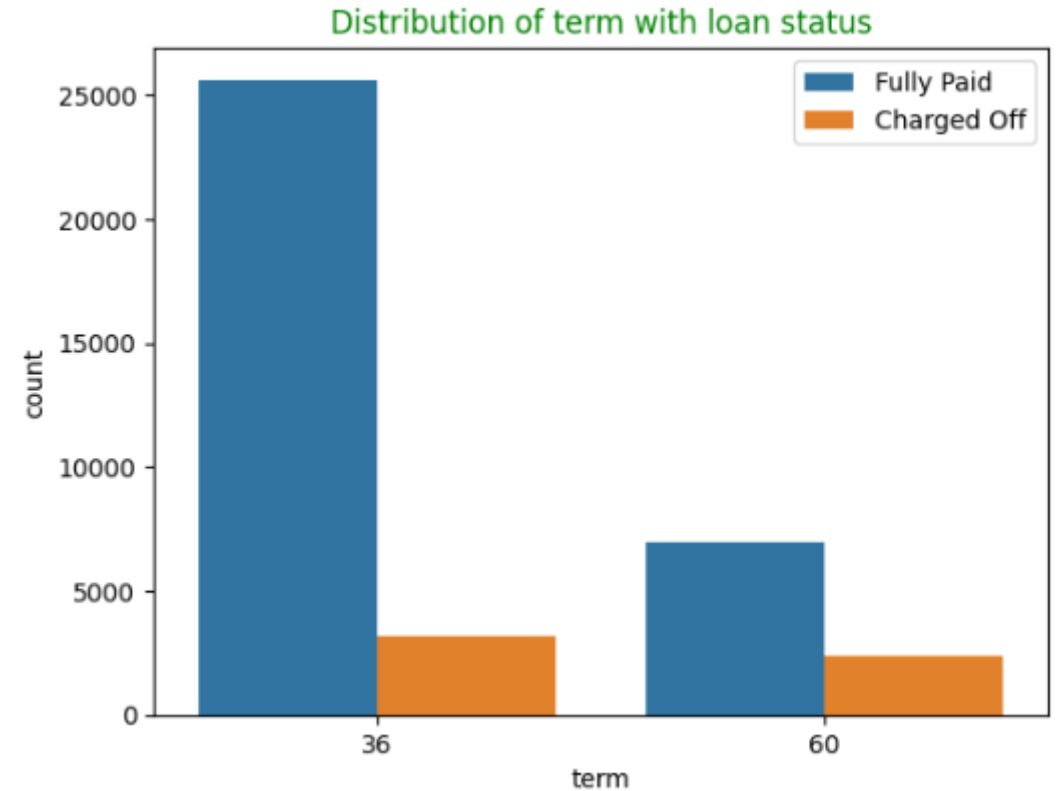
Analysis

- Home Ownership



Applicants who don't own a property and on are rent/mortgage are more likely to default the loan.

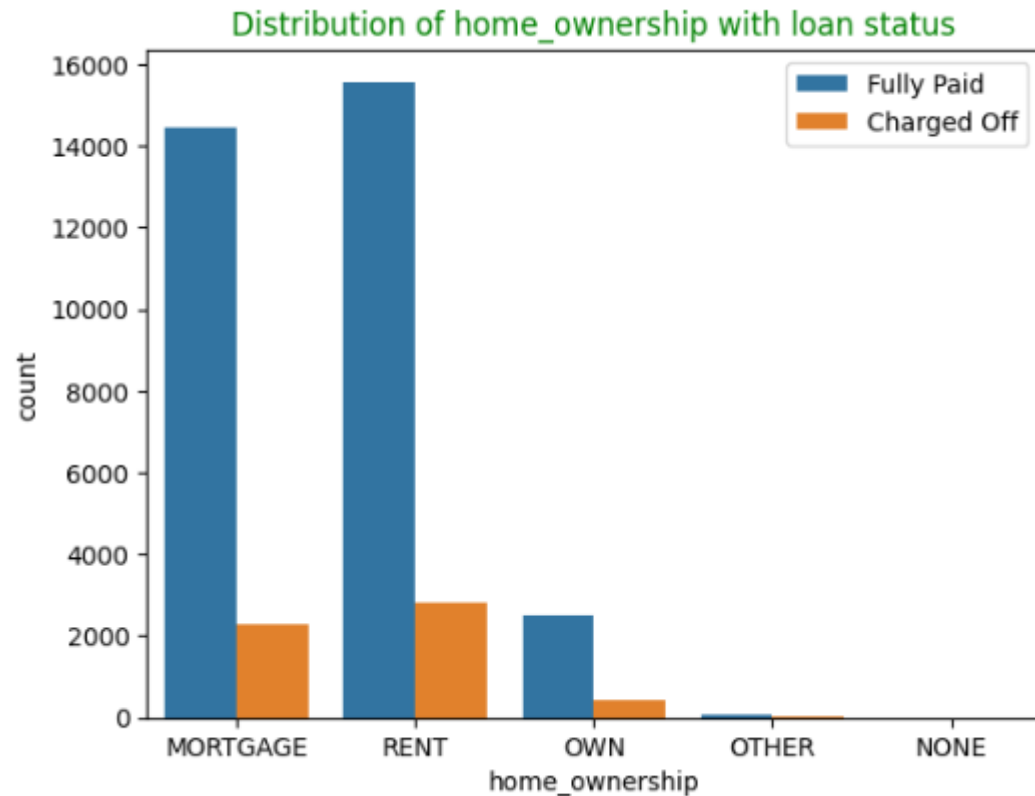
- Term



Majority of loan given to 36 months tenure and are less likely to default

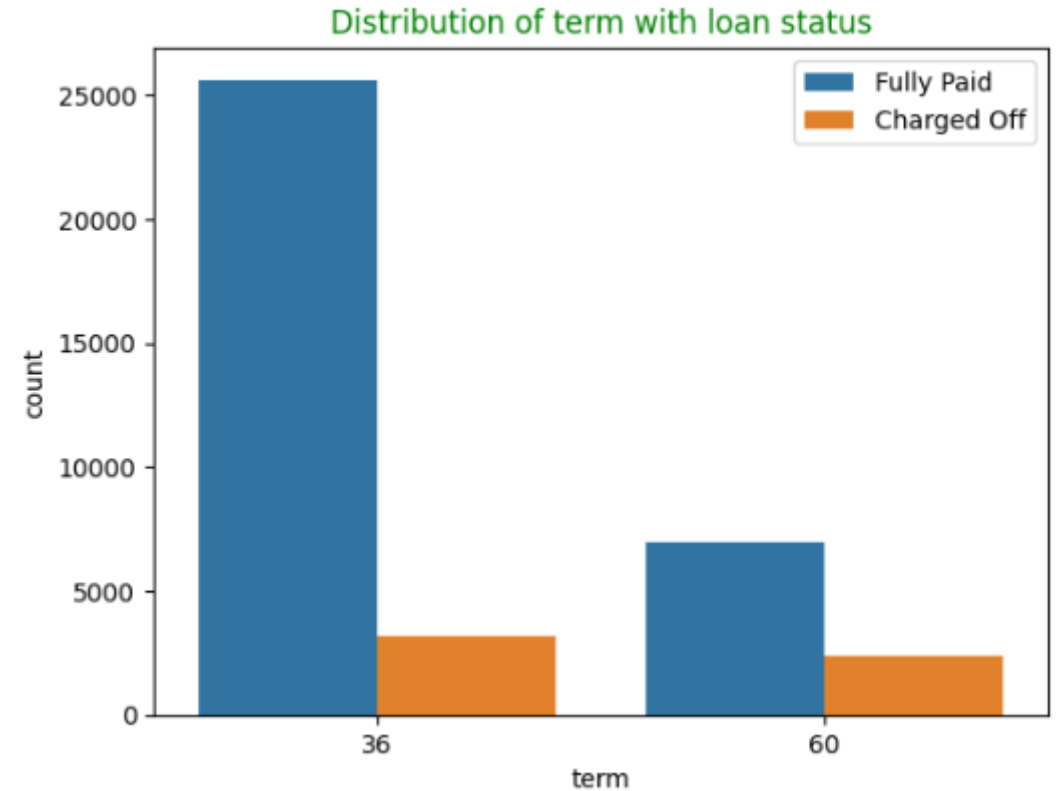
Analysis

- Home Ownership



Applicants who don't own a property and on are rent/mortgage are more likely to default the loan.

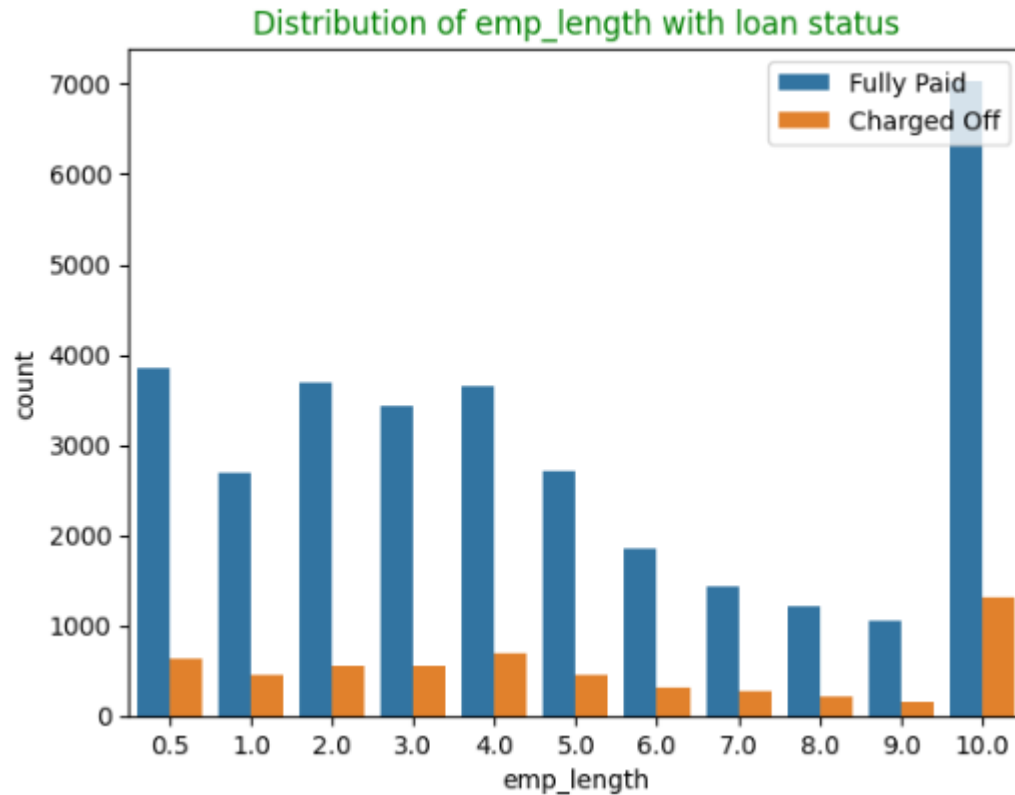
- Term



Majority of loan given to 36 months tenure and are less likely to default

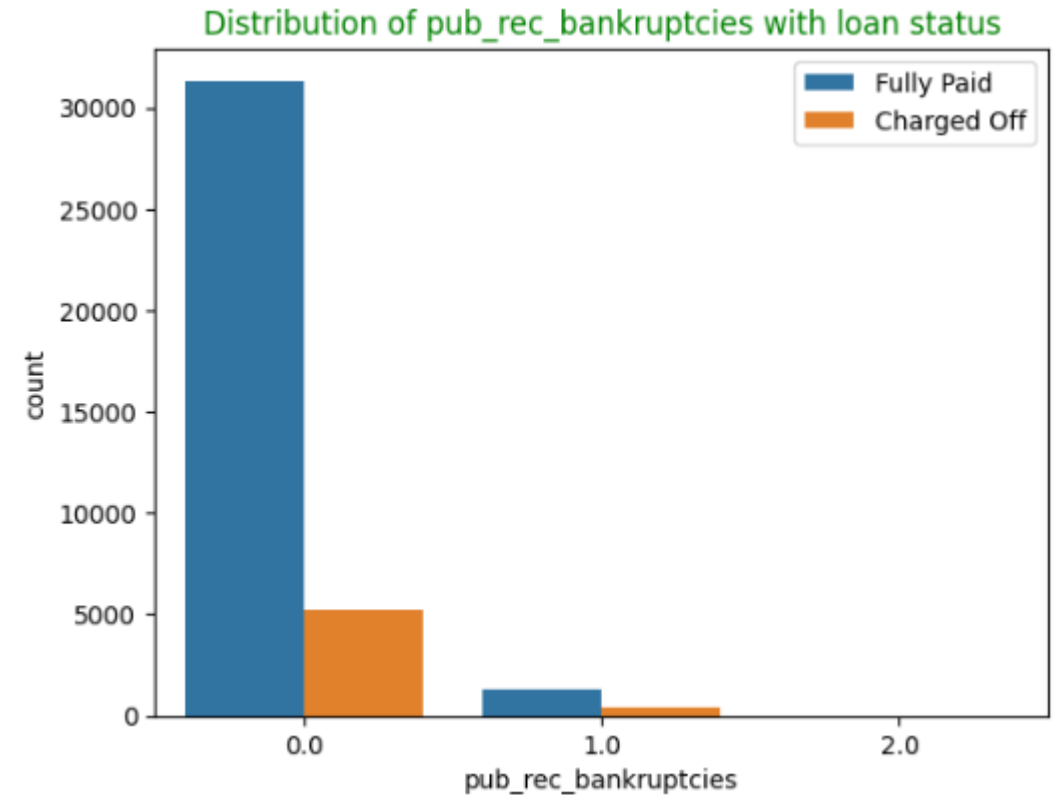
Analysis

- **Employment Length**



Majority of the borrowers have 10+ years of experience and have highest number of defaulters.

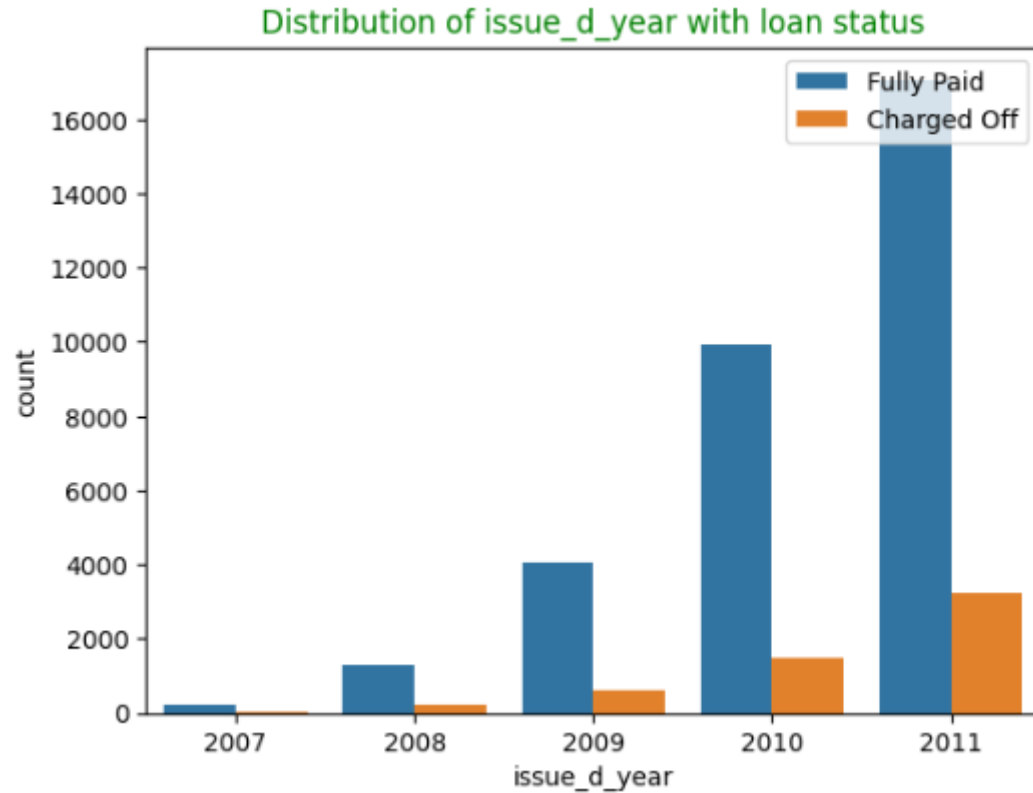
- **Public record bankruptcies**



Majority of the borrowers have no history of bankruptcy. Borrowers with large number of public records bankruptcies are more likely to default.

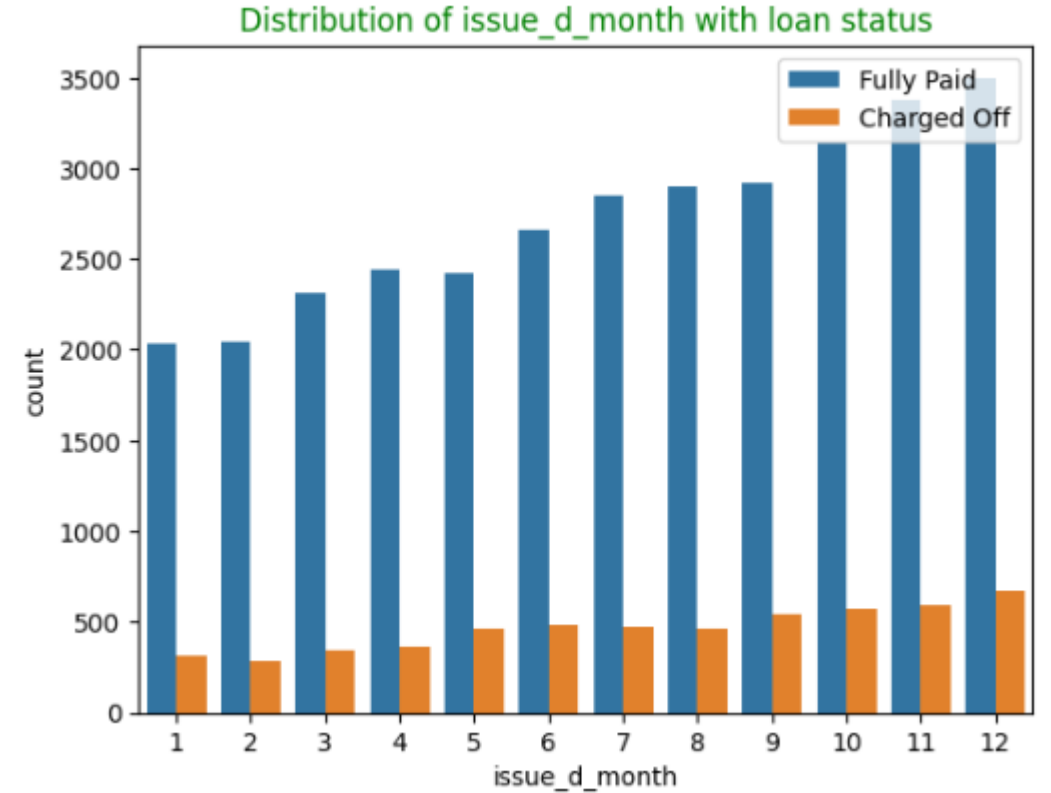
Analysis

- **Loan trend over years**



Number of loans issued have increased from 2007 to 2011.

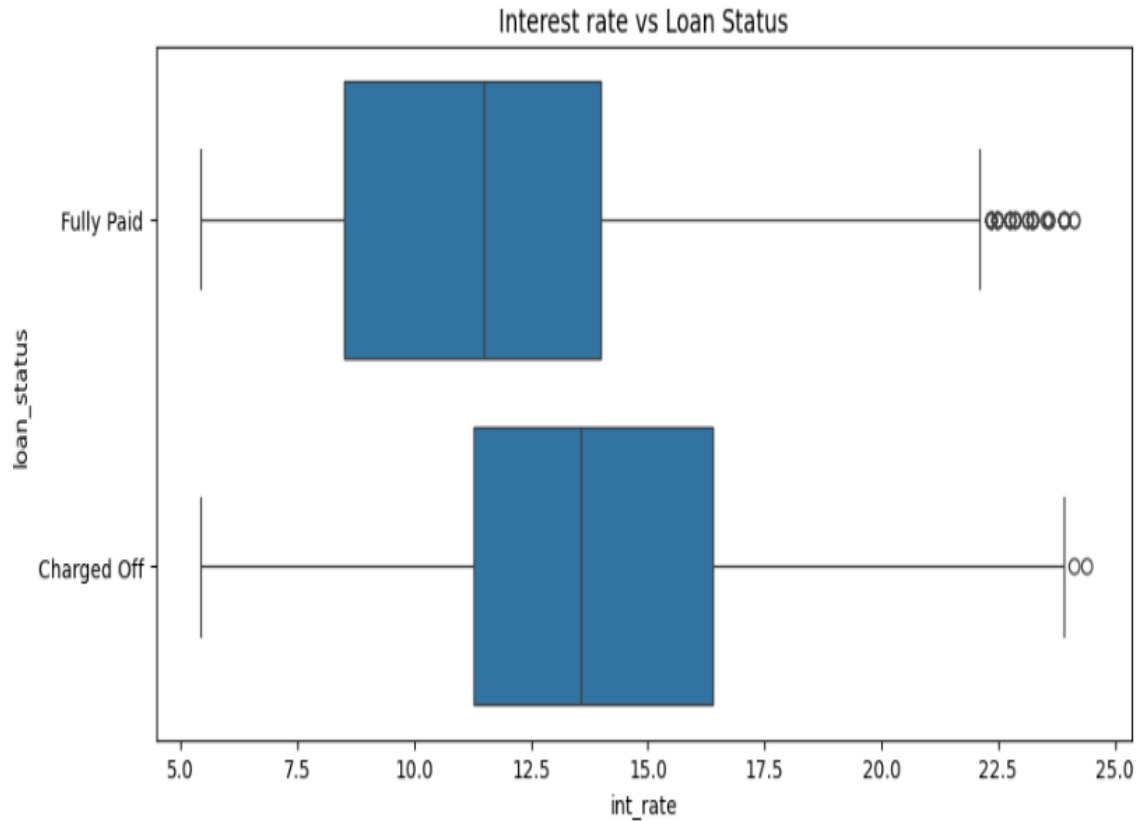
- **Loan trend over months**



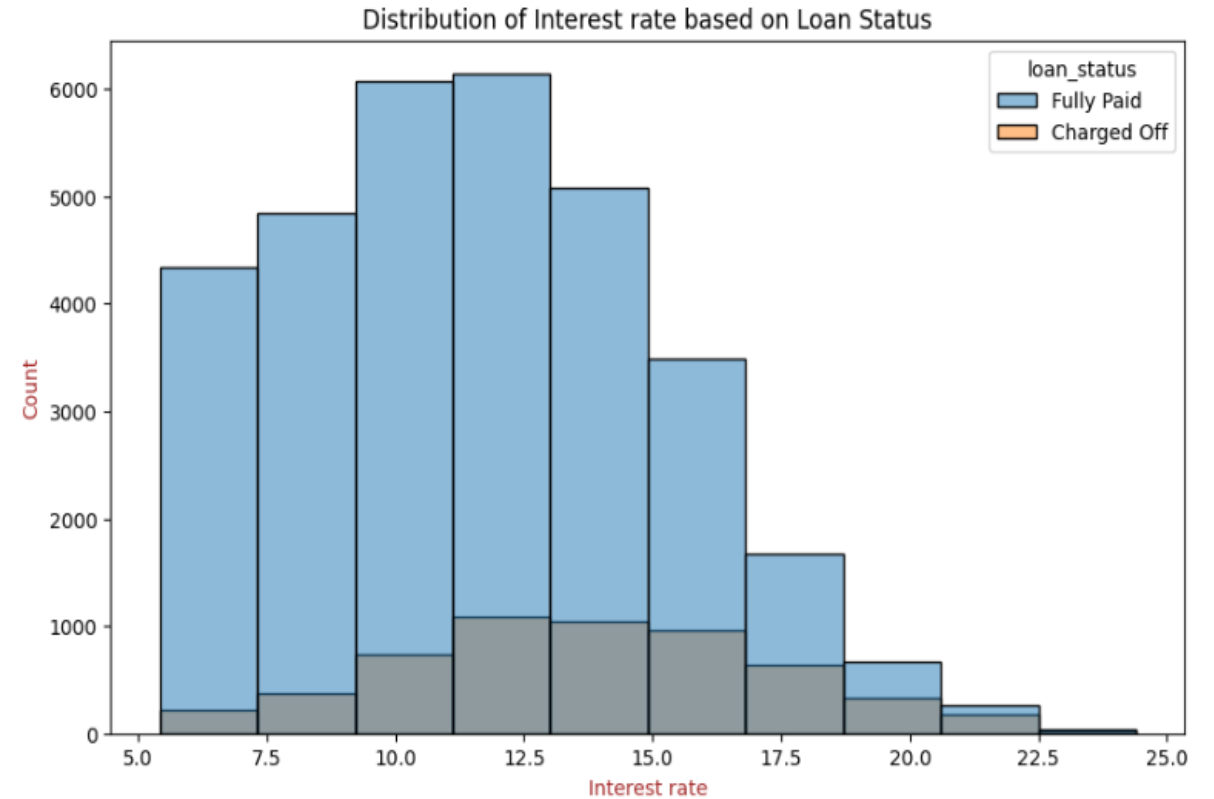
Number of the loans issued have increased in the last quarter.

Analysis

- Interest rate over loan status



- Interest rate over loan status

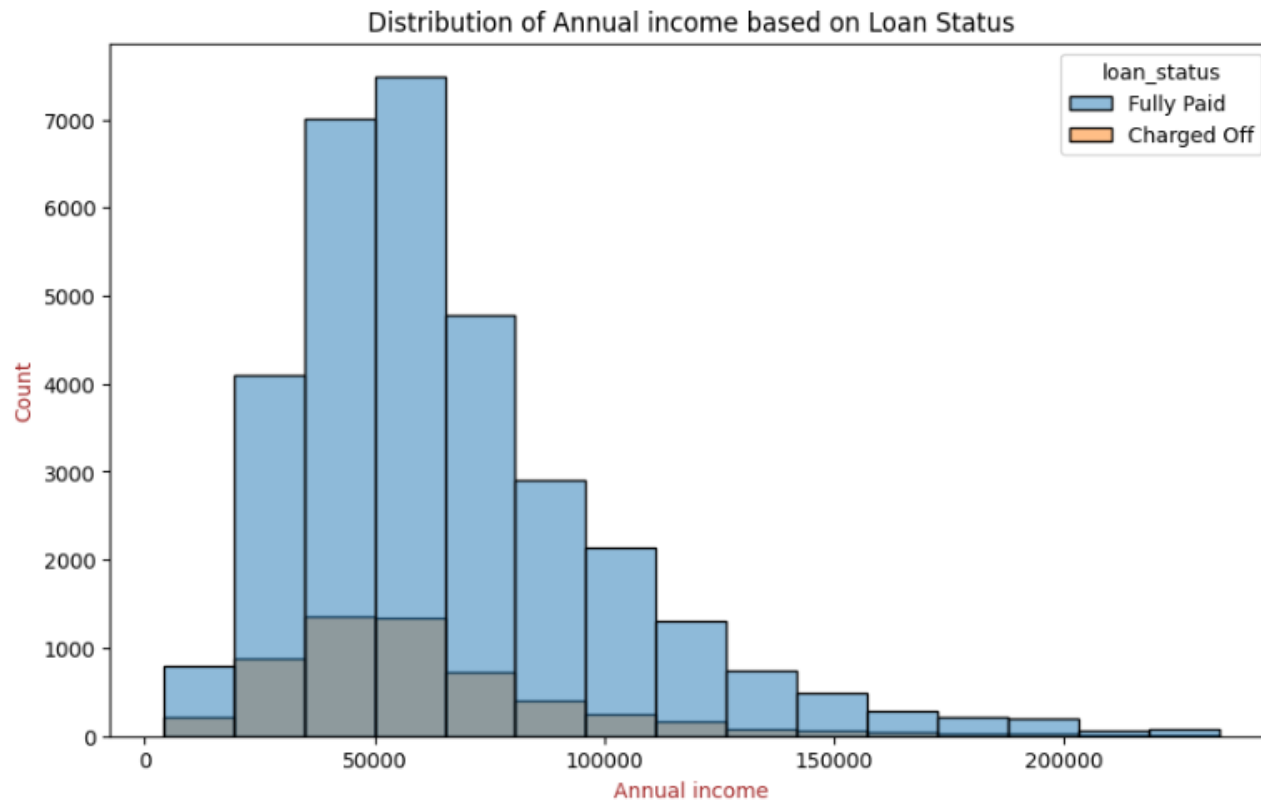


Majority of borrowers are in the interest range 10-15.

High interest rate borrowers are more likely to default.

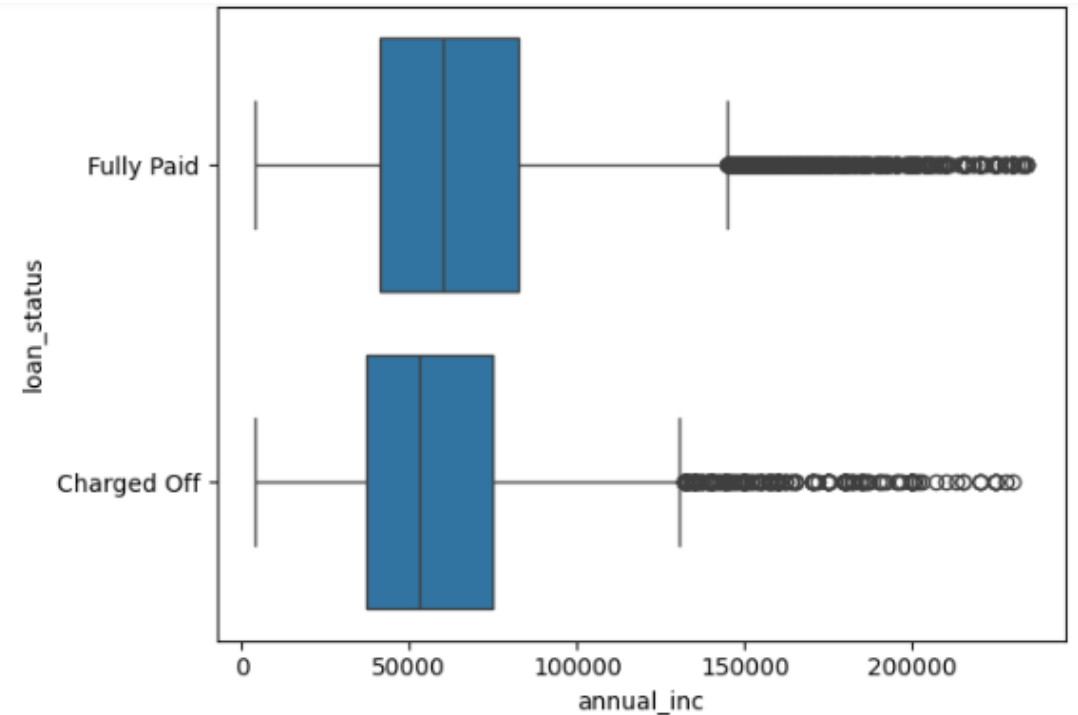
Analysis

- Annual income



Majority of Borrowers are with annual income around 50000.

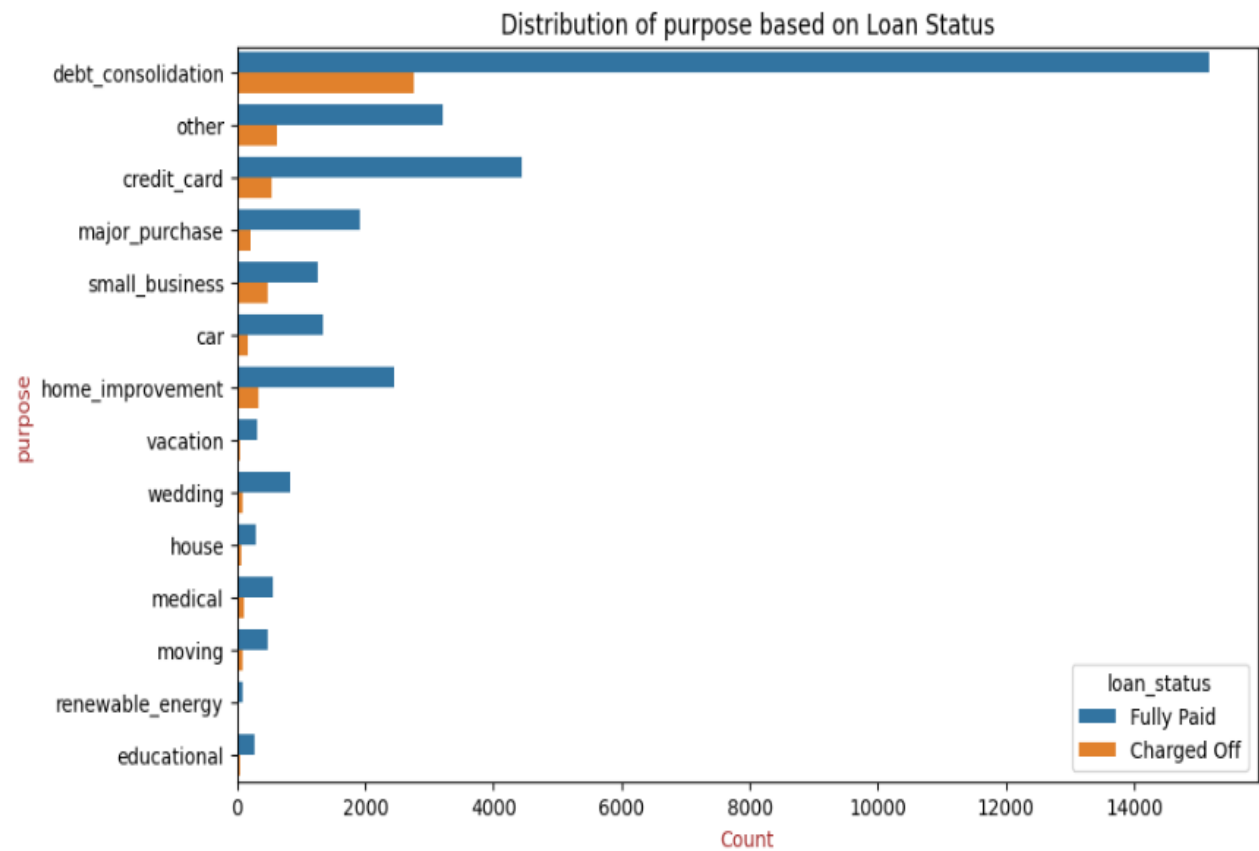
- Annual income based on loan status



People with low annual income are more likely to default.

Analysis

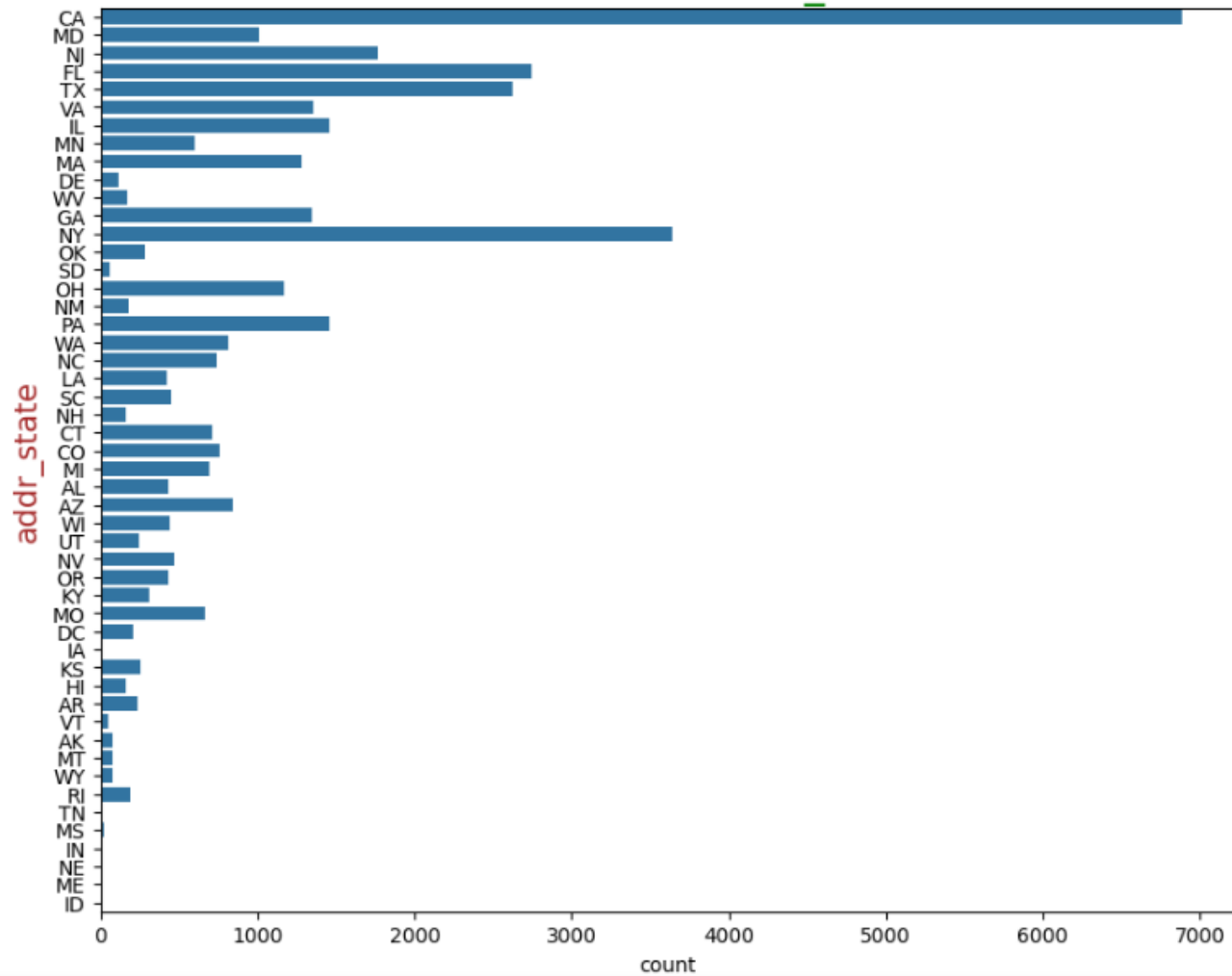
- **Loan Purpose**



Majority of loans were taken for the purpose of debt_consolidation followed by credit_card.

- **State of the borrower**

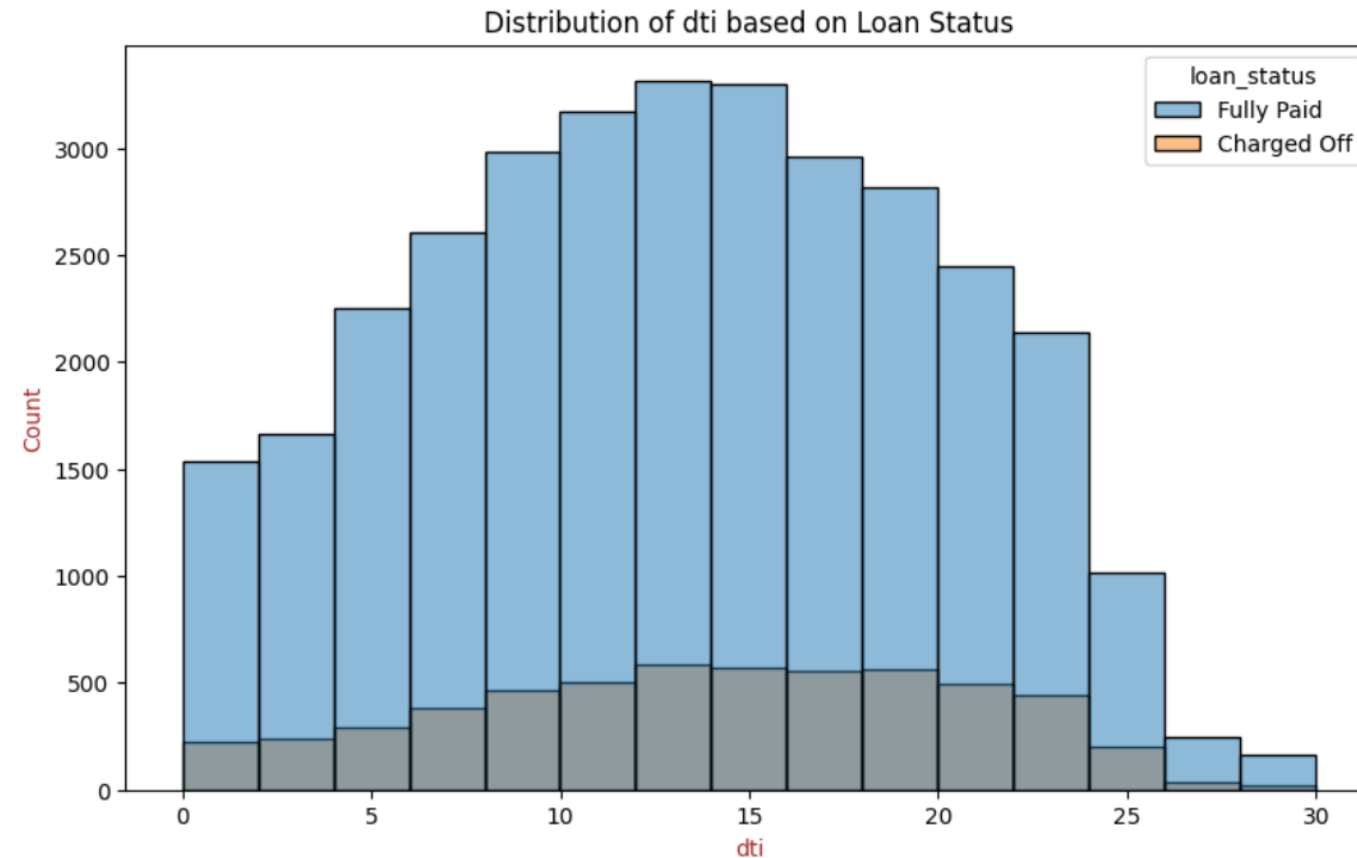
Distribution of addr_state



Majority of borrowers are from CA followed by NY, FL, TX.

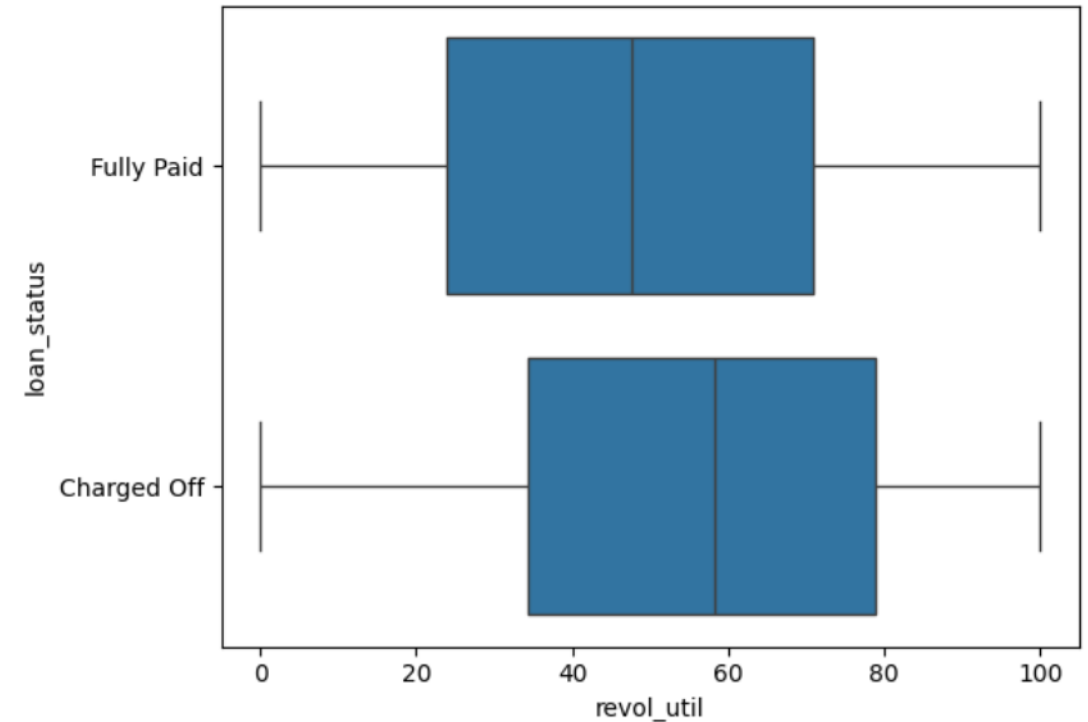
Analysis

- **dti**



Majority of the borrowers are with dti range 10 -15, the defaulters are also more.

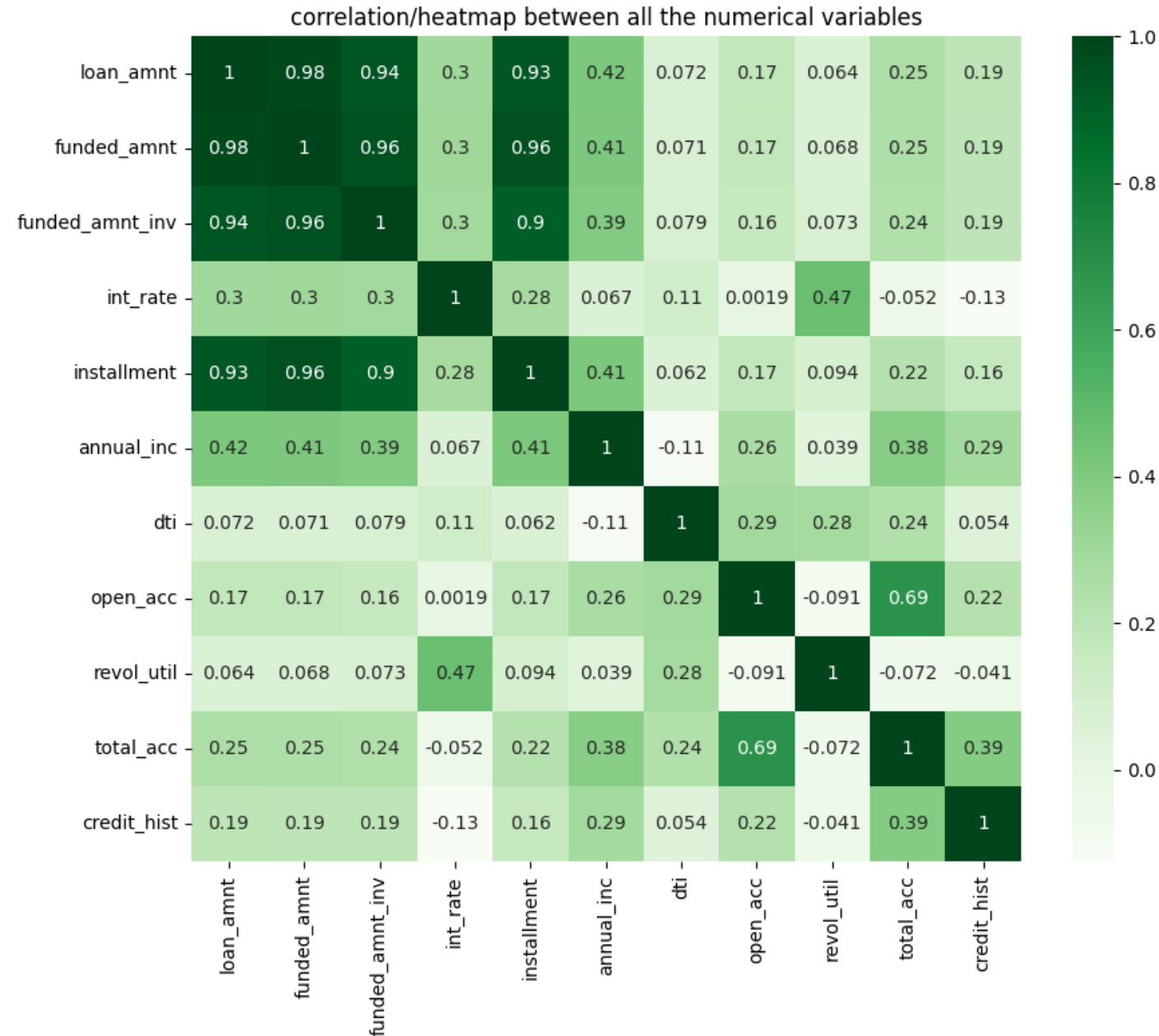
- **Revolving line utilization**



Borrowers having high revolving utilization are more likely to default.

Analysis

- Correlation matrix



- There is strong correlation between installment and loan_amnt, funded_amnt and funded_amnt_inv. These parameters are directly proportional to one another.

Conclusion

Below are the Major Driving factors which can be used to predict the loan defaulters.

- home_ownership
- Grade
- revolving_util
- int_rate
- Pub_rec
- annual_income