Lending Club Case Study

Abstract

- Lending club is the largest loan marketplace for personal loans, business loans and other type of loans.
- Borrowers can easily access to lower interest rates loans through fast online interface.
- The objective of this analysis is to use the past information about loan applicants and find whether they have 'defaulted' or not.

Problem solving methodology

Data Analysis | Data Analysis | Data Analysis

Data Cleaning

Removing all null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.

Data Understanding

Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses

Univariate Analysis

Analysing each column, plotting the distributions of each column.

Segmented Univariate Analysis

Analysing the continuous data columns with respect to the categorical column

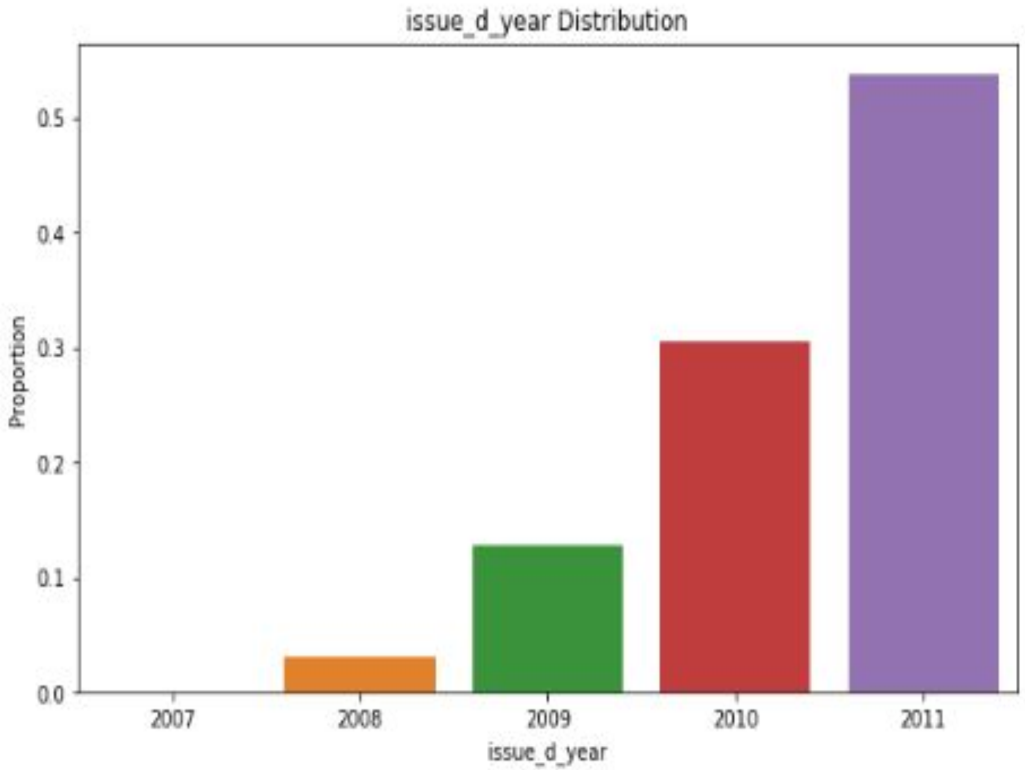
Bivariate Analysis

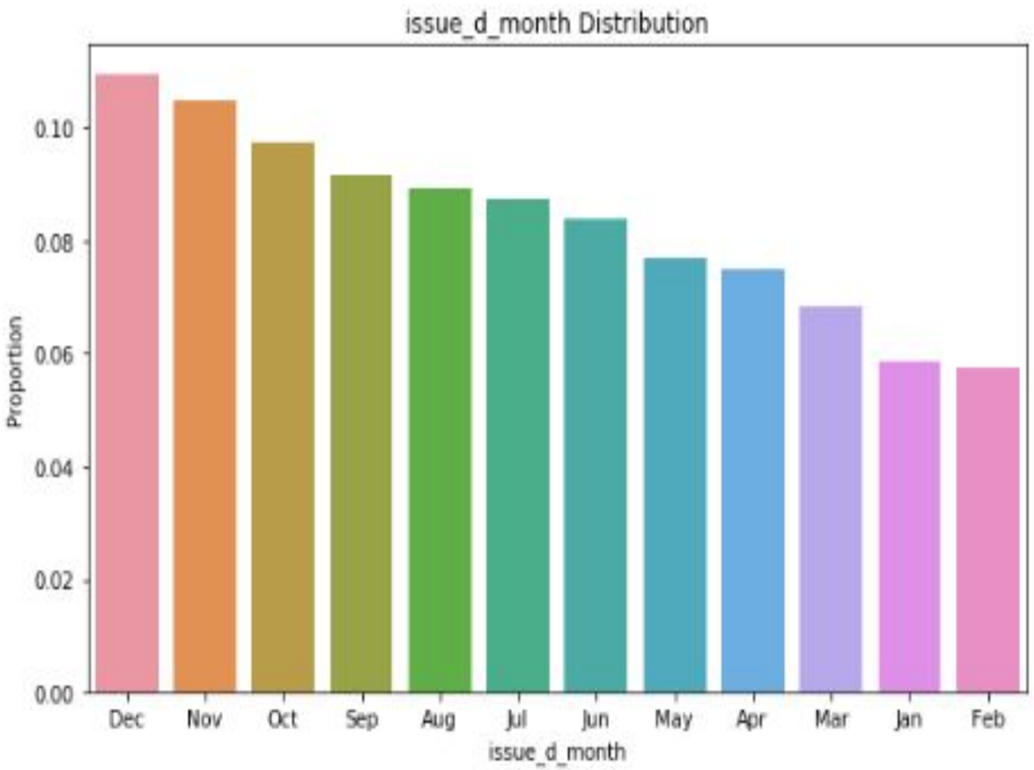
Analysing the two variable behaviour like term and loan status with respect to loan amount.

Recommendations

Analysing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

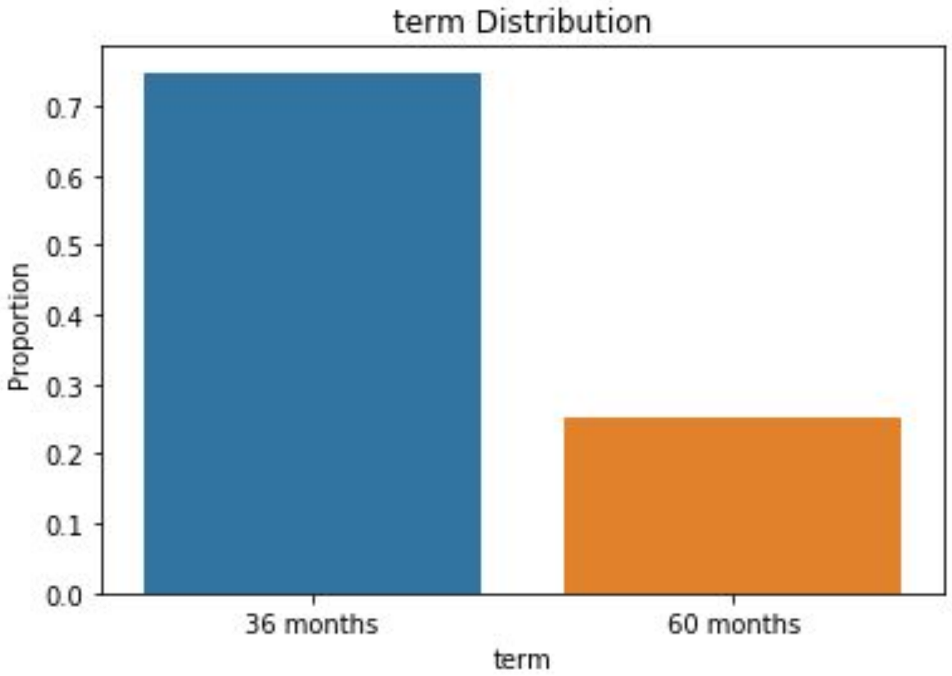
Analysis

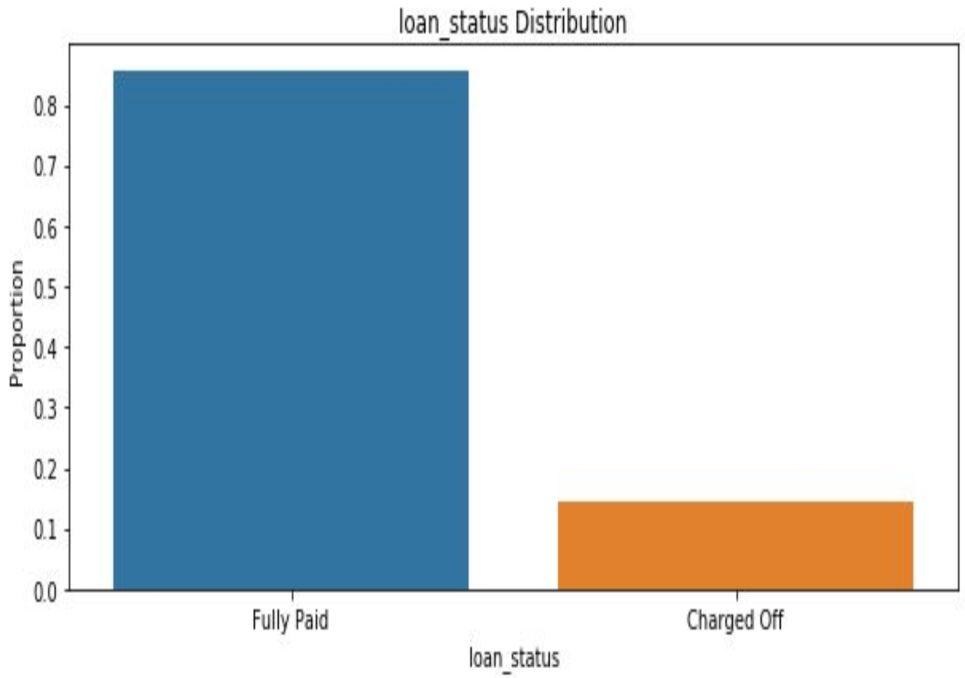




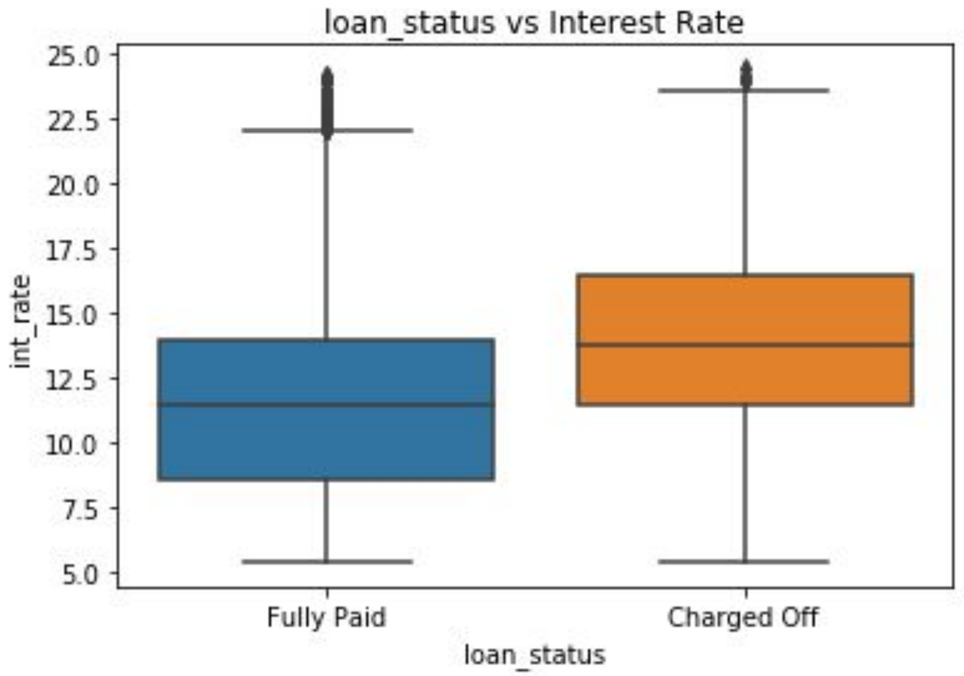
- Lending club has really expanding every year, the number of loans issued are doubled year by year.
- The issued month of loans is also increasing from January to December. In the final quarter of year there are more loans issued this could be because of vacation and christmas.

Analysis



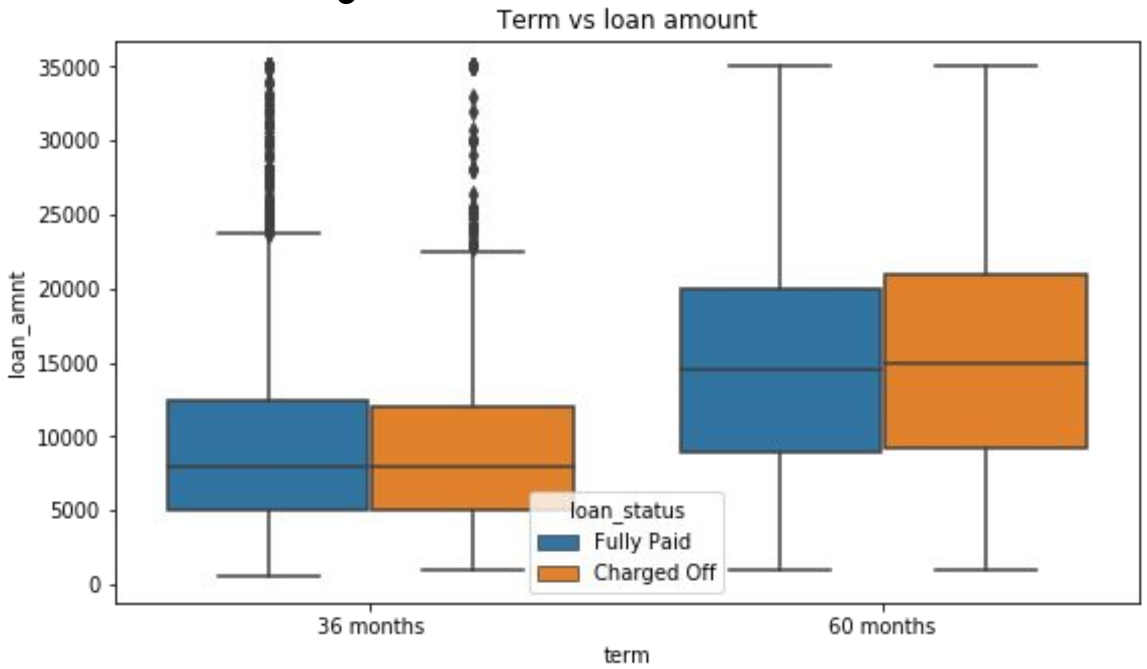


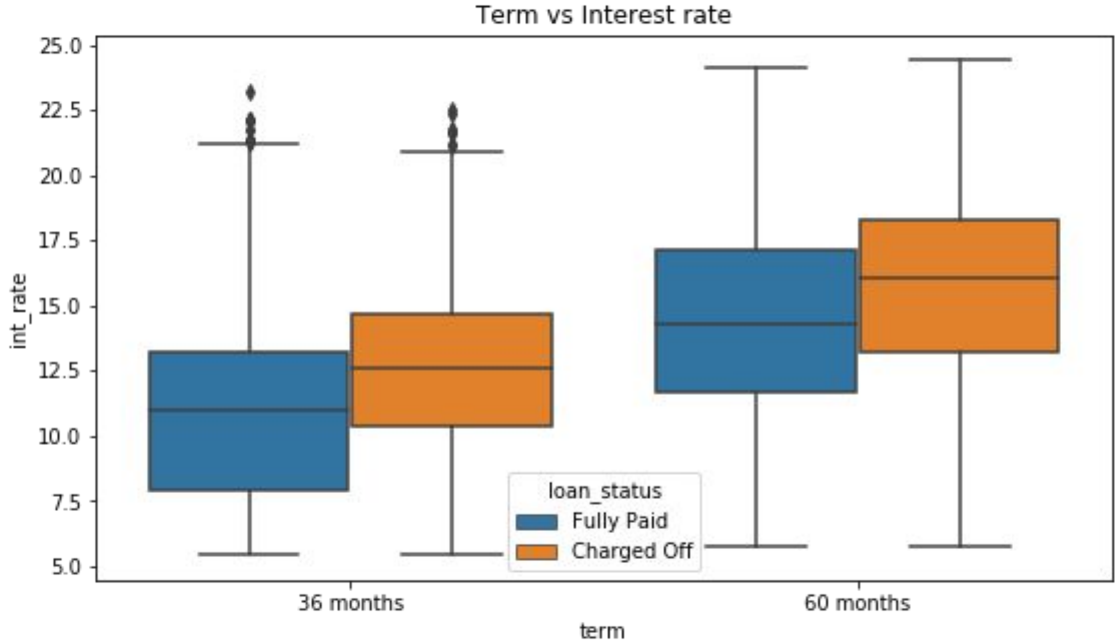
- There are only two loan terms 36 and 60 months. Around 75% borrowers took loans with 36 months term.
- The charged off borrowers are around 15% and fully paid is around 85% in the given data set.

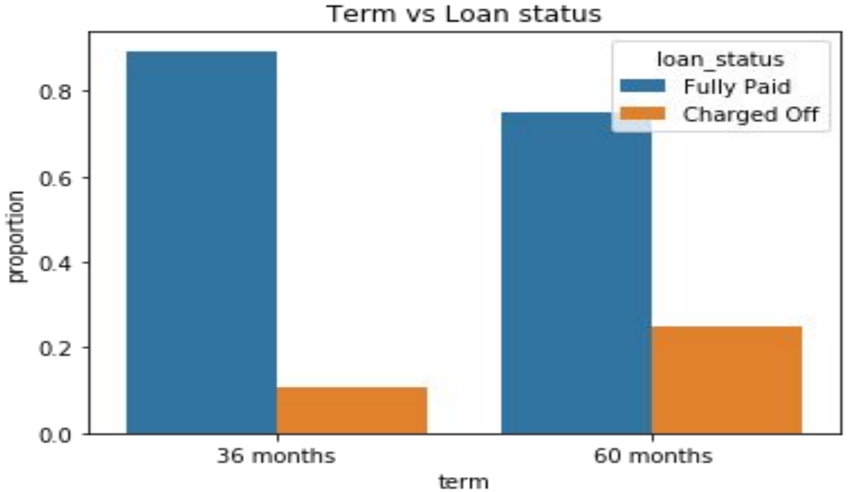


When the interest rate of loan is high there is a high chance of loan getting defaulted.

Analysis

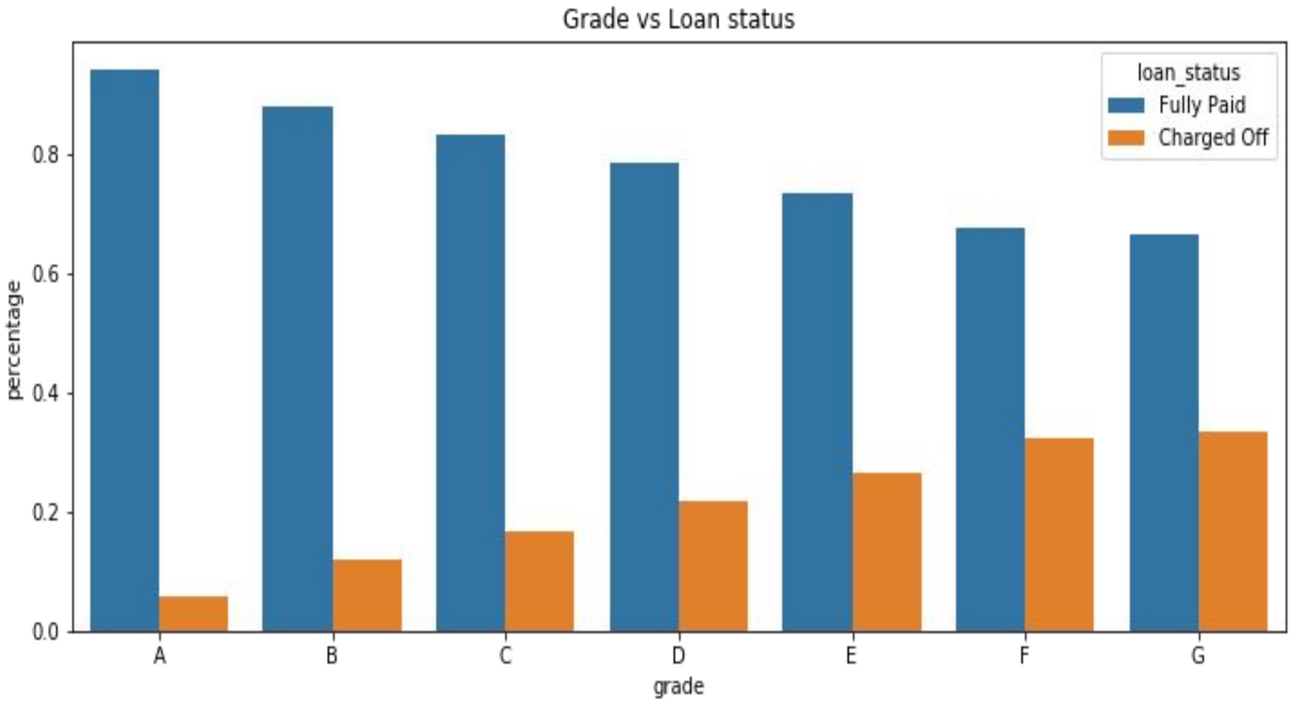


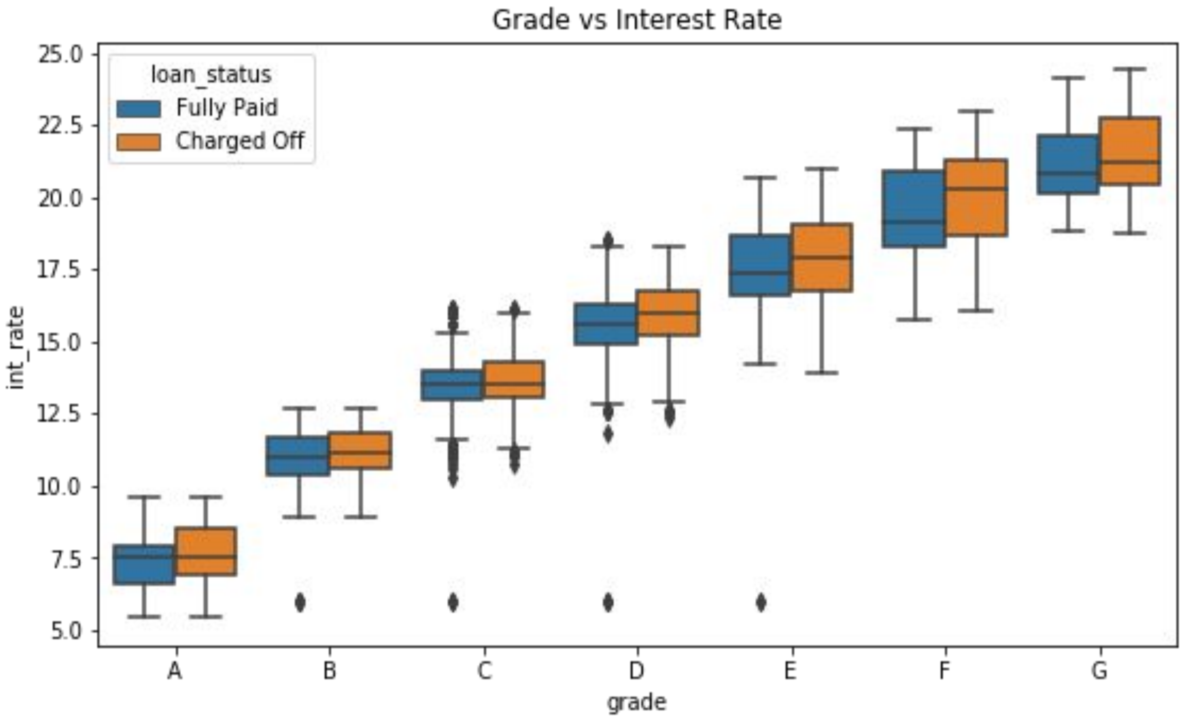




The default rate is high in 60 months tenure because most people took high loan amount with high interest rate in it and find difficult to pay back them to bank.

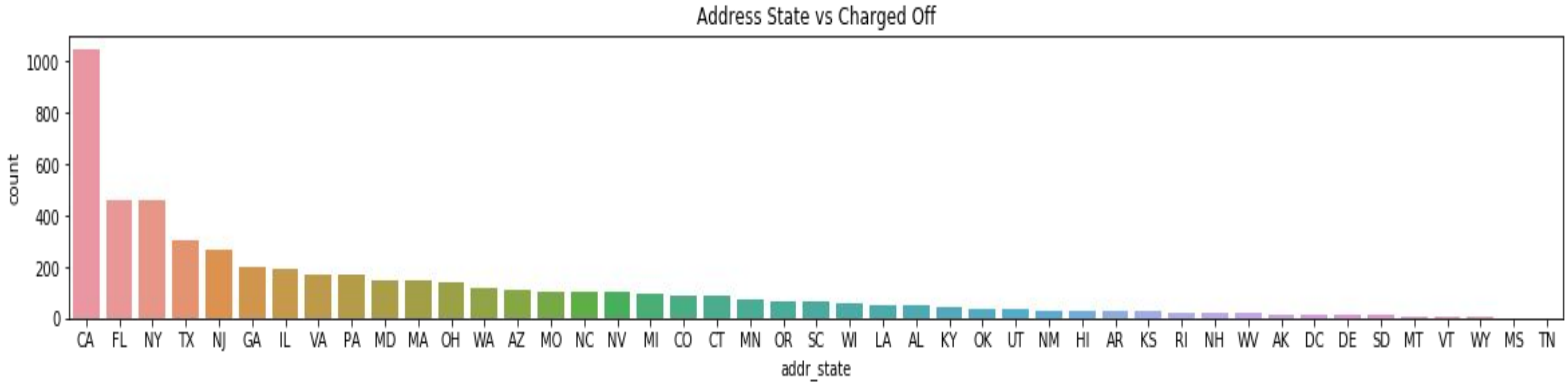
Analysis





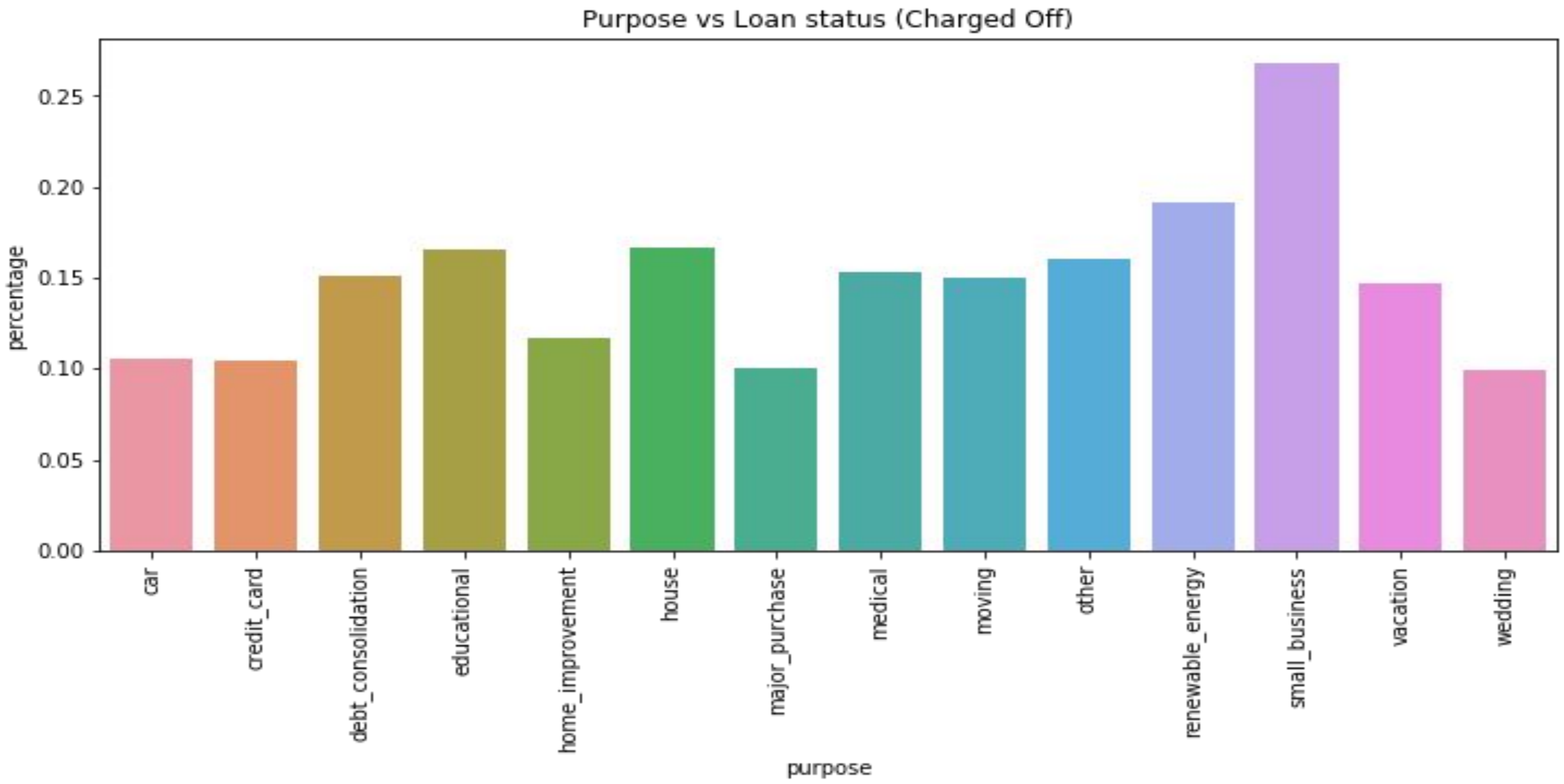
- Grades given by banks are very good category to tell the borrower probability of defaulting the loan.
- The lower grades(E,F,G) have higher chances of defaulting the loan than higher ones(A,B)
- Also the lower grades are getting loans for high rate of intrest which might be the cause for loan default.

Analysis



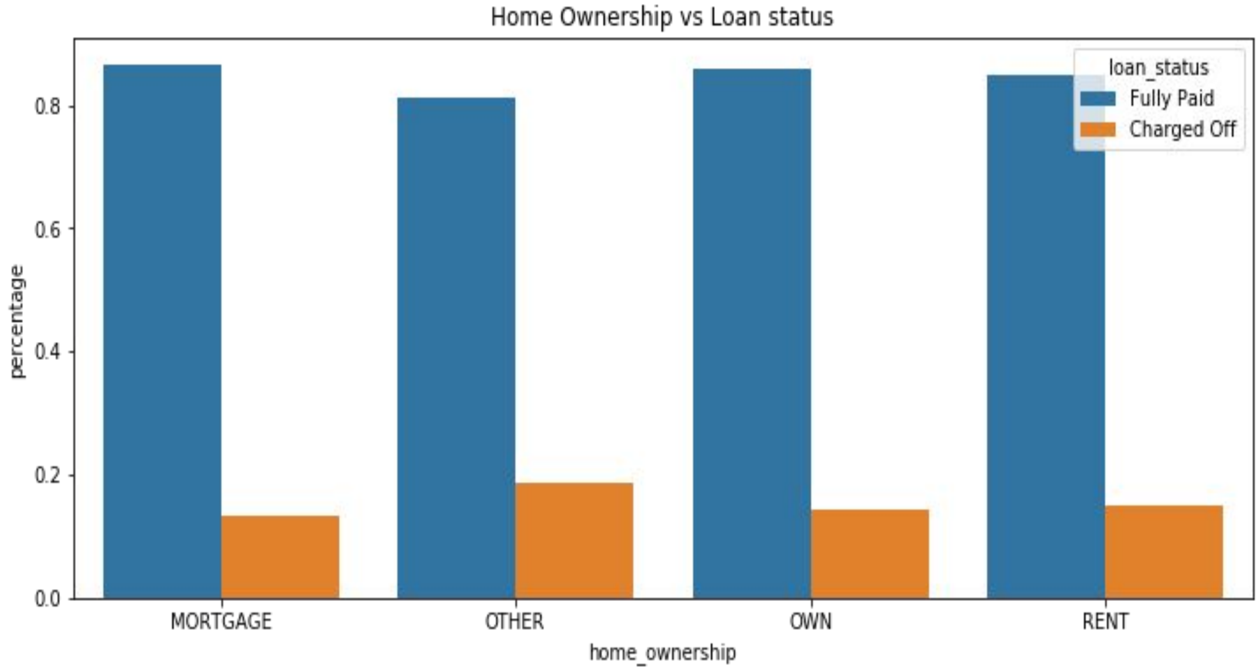
The loans which are given to CA, FL and TX state borrower’s have defaulted more than other states.

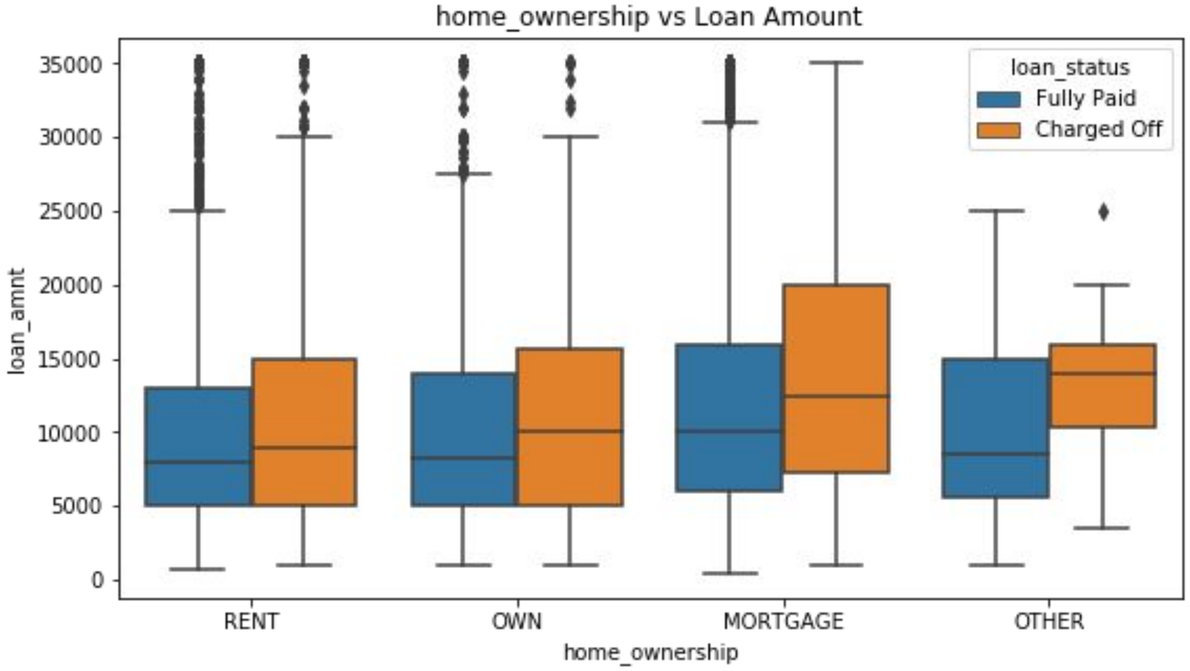
Analysis



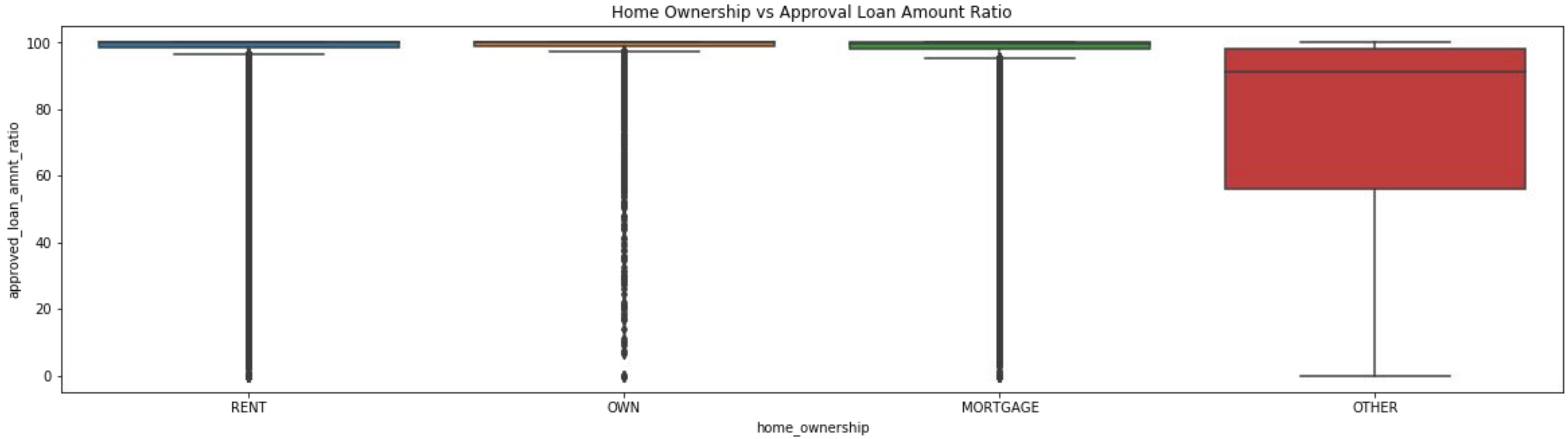
Borrower’s who have taken loans for small business purpose have defaulted more.

Analysis





- There is around 20% chance of default in loan in each home ownership category.
- From the 2nd plot we can see the people with higher loan amounts in **mortgage** home ownership has high default rate than others.



Approved loan (Funded Amount by investor) is less than the requested loan amount by borrowers for Other Home ownership category.

Conclusions

- Lending club should reduce the rate of interest for loans with 60 months tenure, as they are prone to loan default.
- Grades are good metric for detecting loan defaulters. Lending club should examine more information from borrowers before issuing loans to Low grade (G to A).
- Lending Club should control their number of loan issued to borrowers who are from CA, FL and NY to make profits.
- Small business loans are defaulted more. Lending club should stop/reduce issuing the loans to them.
- Borrowers with mortgage home ownership are taking higher loans and defaulting the approved loans. Lending club should reduce the loan amount giving to this category more than 12000.