

Lending Club Case Study

By :-

Bhagvan Bux Verma

What is Lending Club Case Study

Lending Club is a marketplace for personal loans that matches borrowers who are seeking a loan with investors looking to lend money and make a return.

Abstract

- Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.
- Borrowers can easily access lower interest rate loans through a fast online interface.
- The objective of analysis is to use the information about past loan applicants and find whether they ‘defaulted’ or not.

Problem solving methodology

Data Cleaning

- Removing the 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

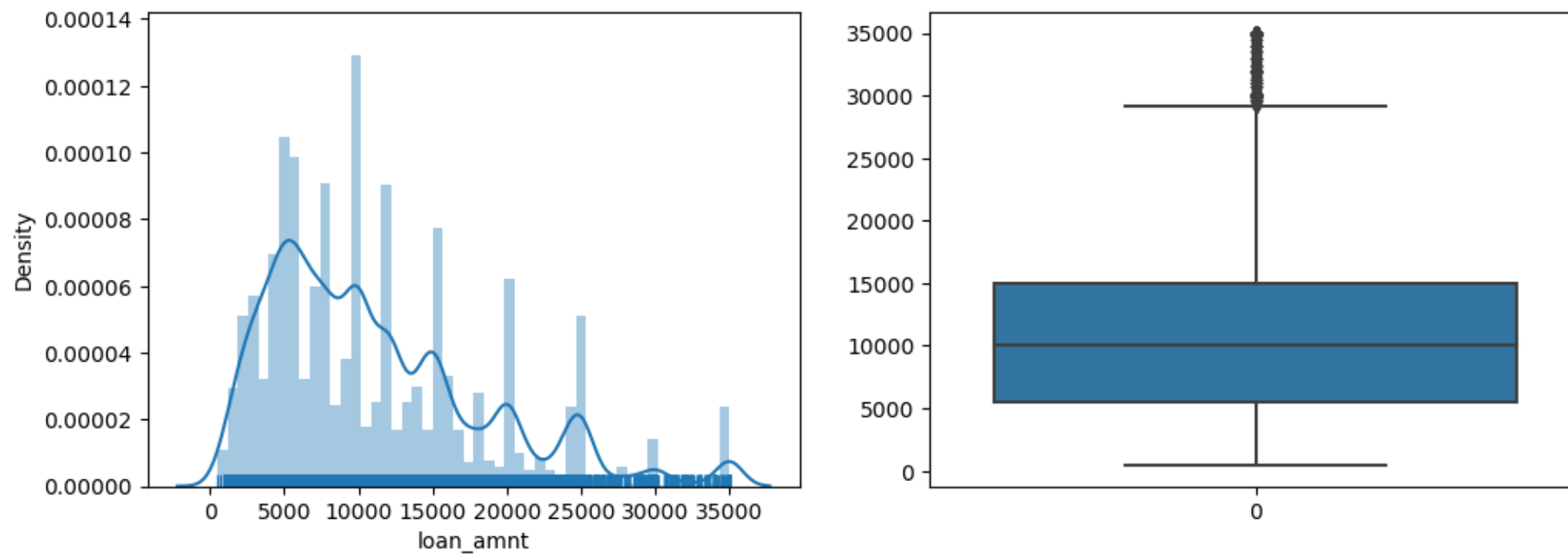
- Analysing the two variables behaviour like term and loan status with respect to loan amount..

Recommendations

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

Analysis

Loan Amount distribution

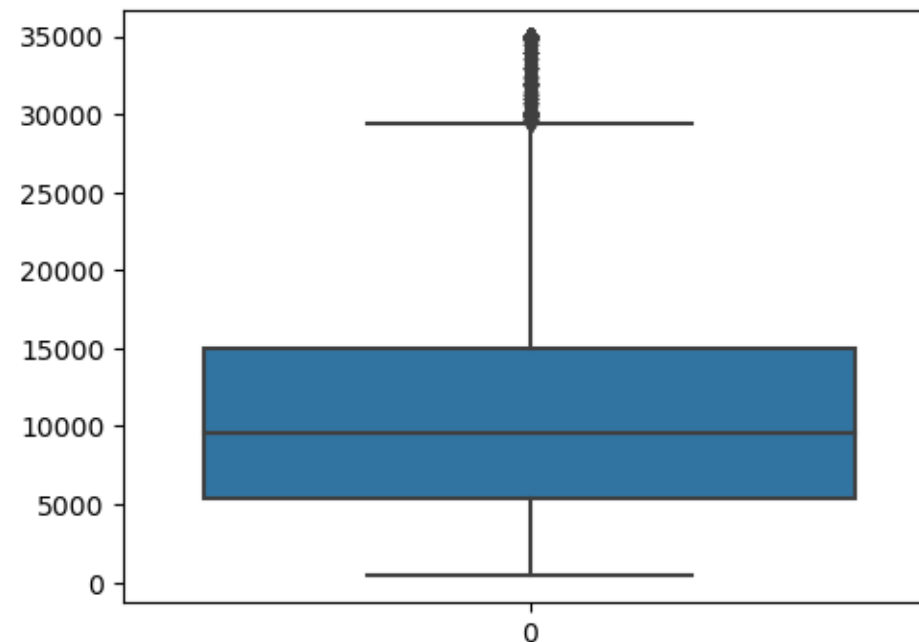
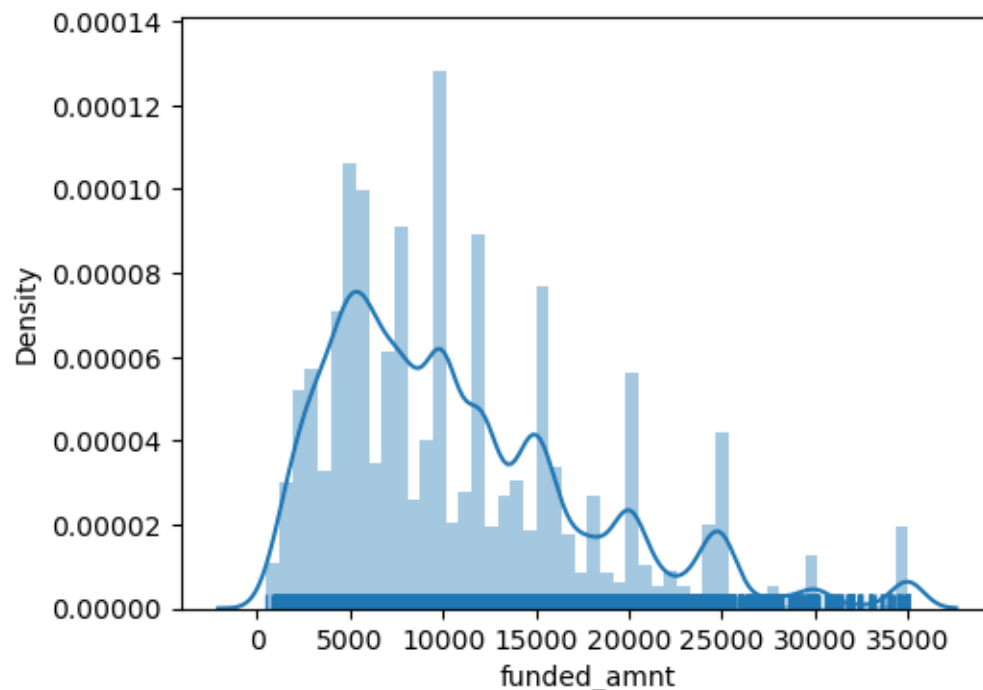


Observations:

From the above plots, more number of people took loan amount of 10000, and also meadian of distribution is 10000. And very few people took more than 30000 loan amount.

Analysis

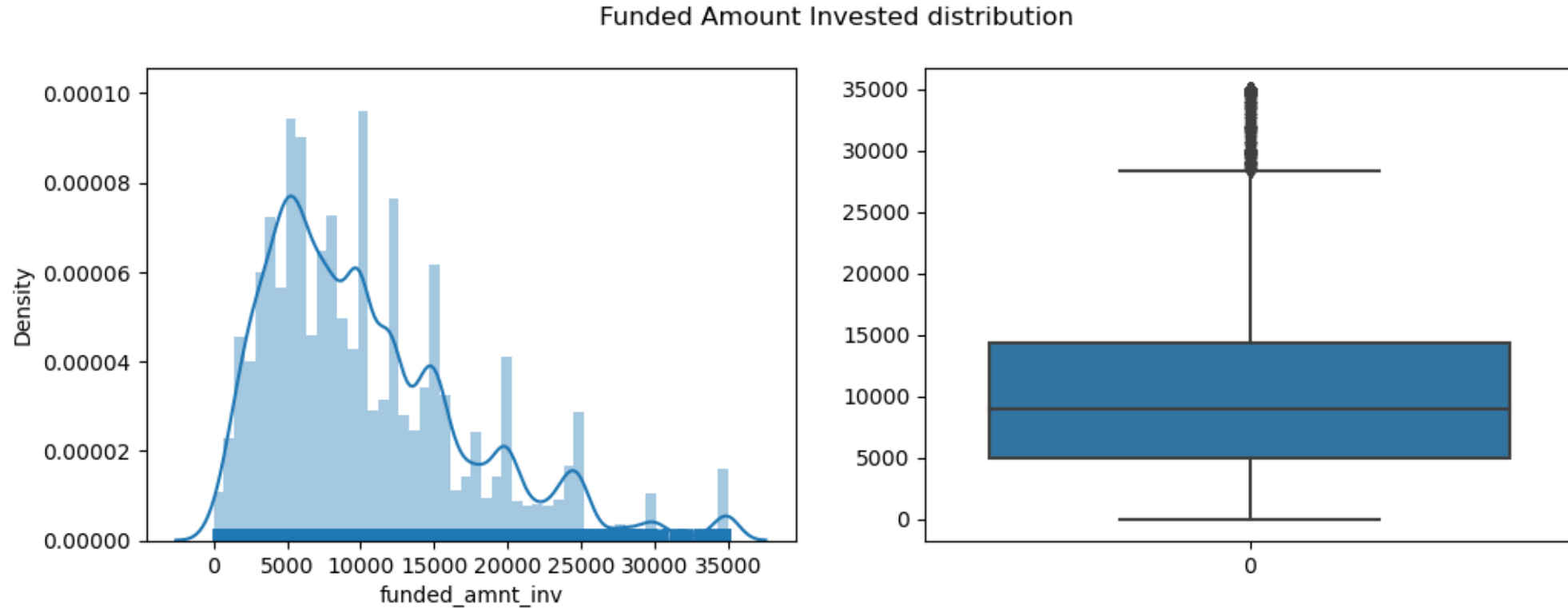
Funded Amount distribution



Observations:

Funded amount data behaves similar to loan Amount, Which tells us that Lending club approved most of Applied loan amount.

Analysis

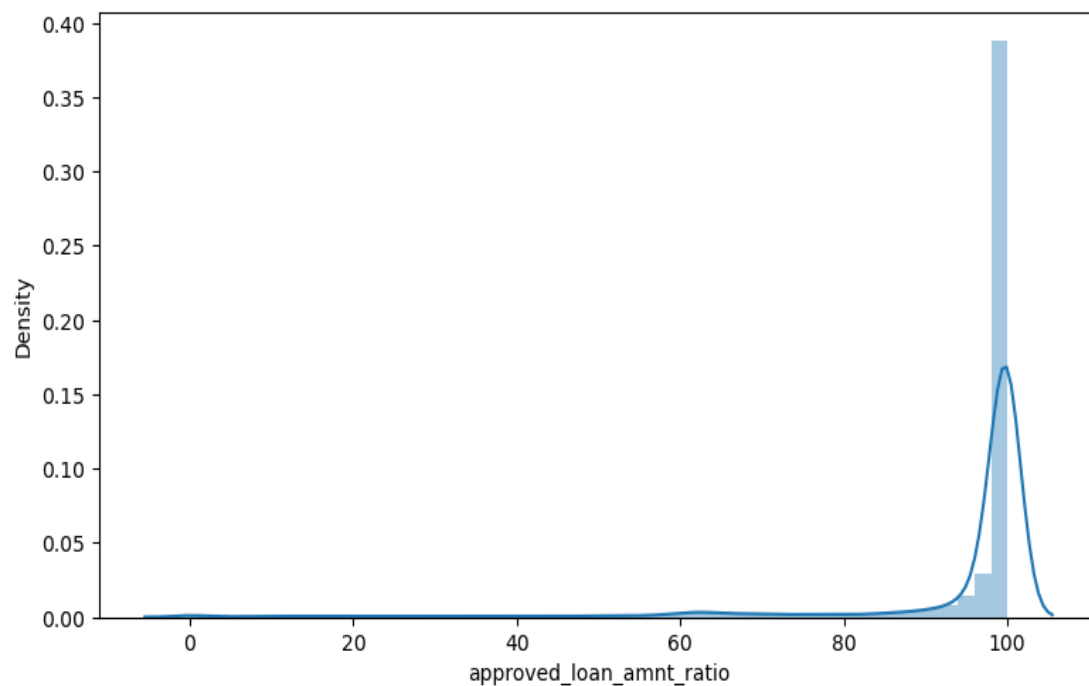


Observations:

Funded amount investment data behaves similar to loan Amount, Which tells us that Lending club approved atmost of Applied loan amount.

Analysis

Approved Loan Amount Ratio distribution

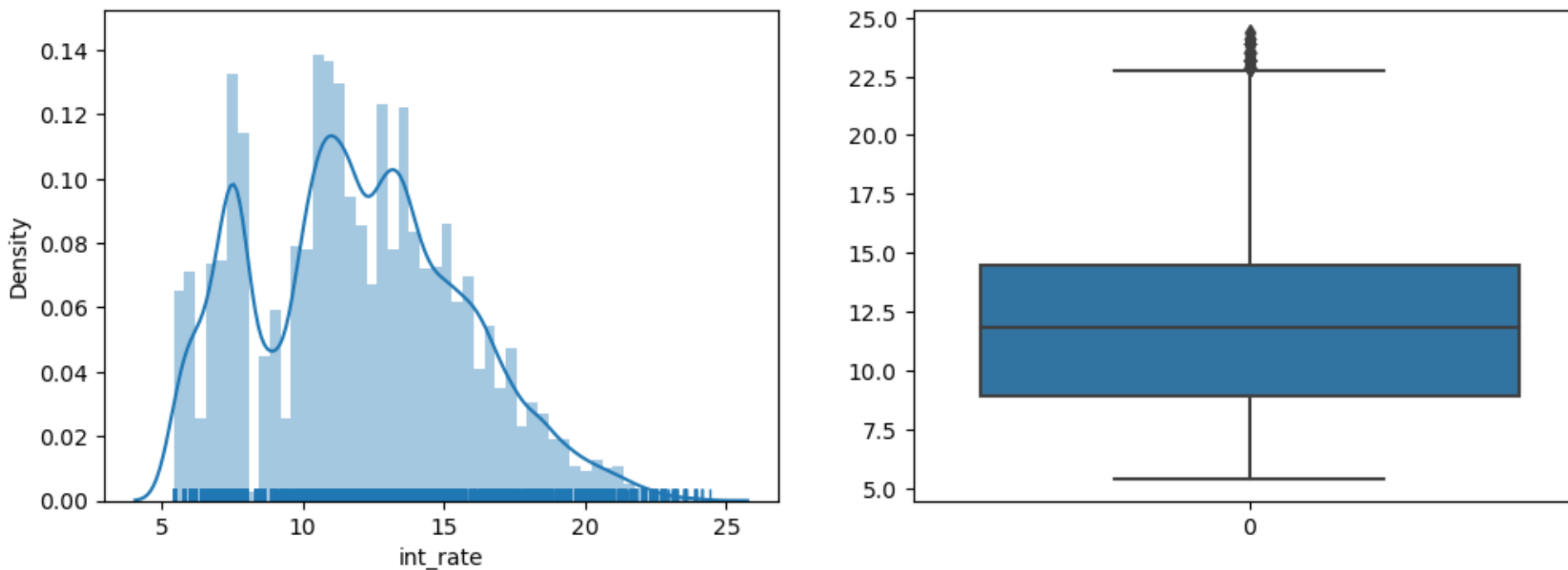


Observations:

70% of Borrowers got 100% loan amount from investors.

Analysis

interest rate distribution

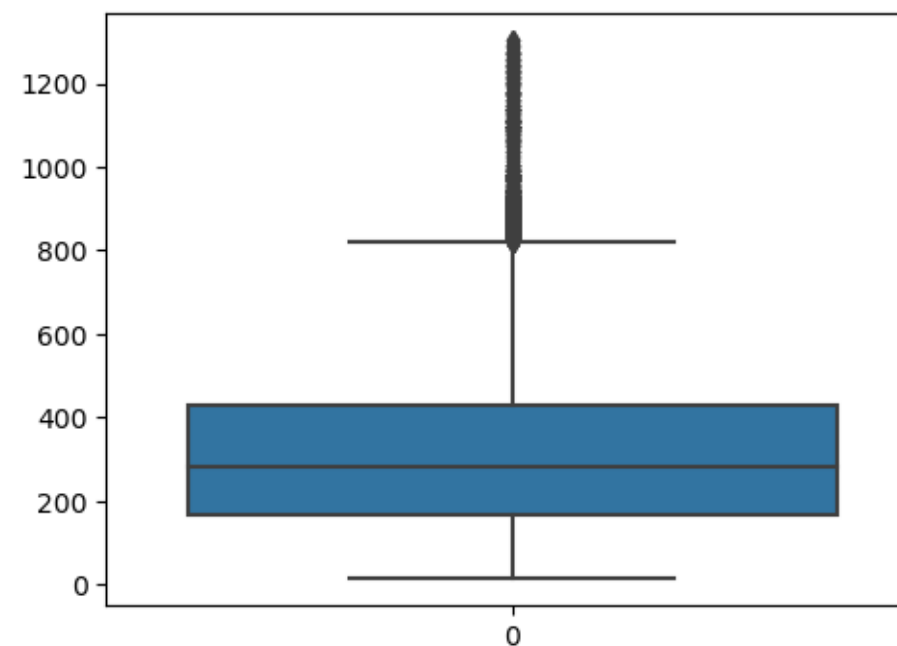
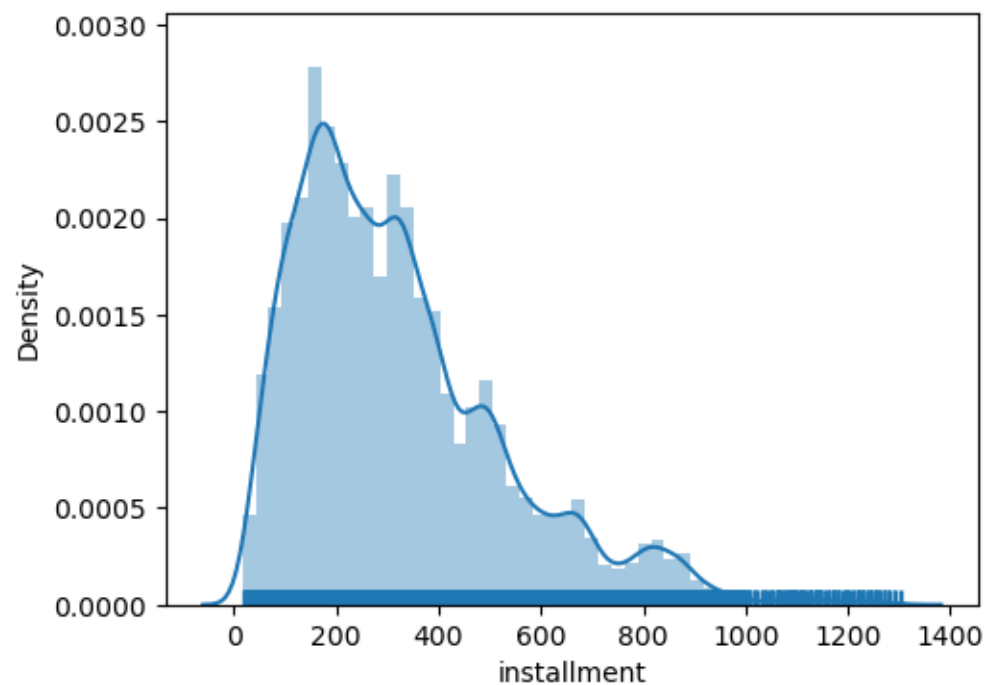


Observations:

From the above 2 plots and statistics of interest rates we can conclude that most of the interest rates lies between **9% to 14.5%**. Some people took loan at higher rates of interest i.e., **22.5%**

Analysis

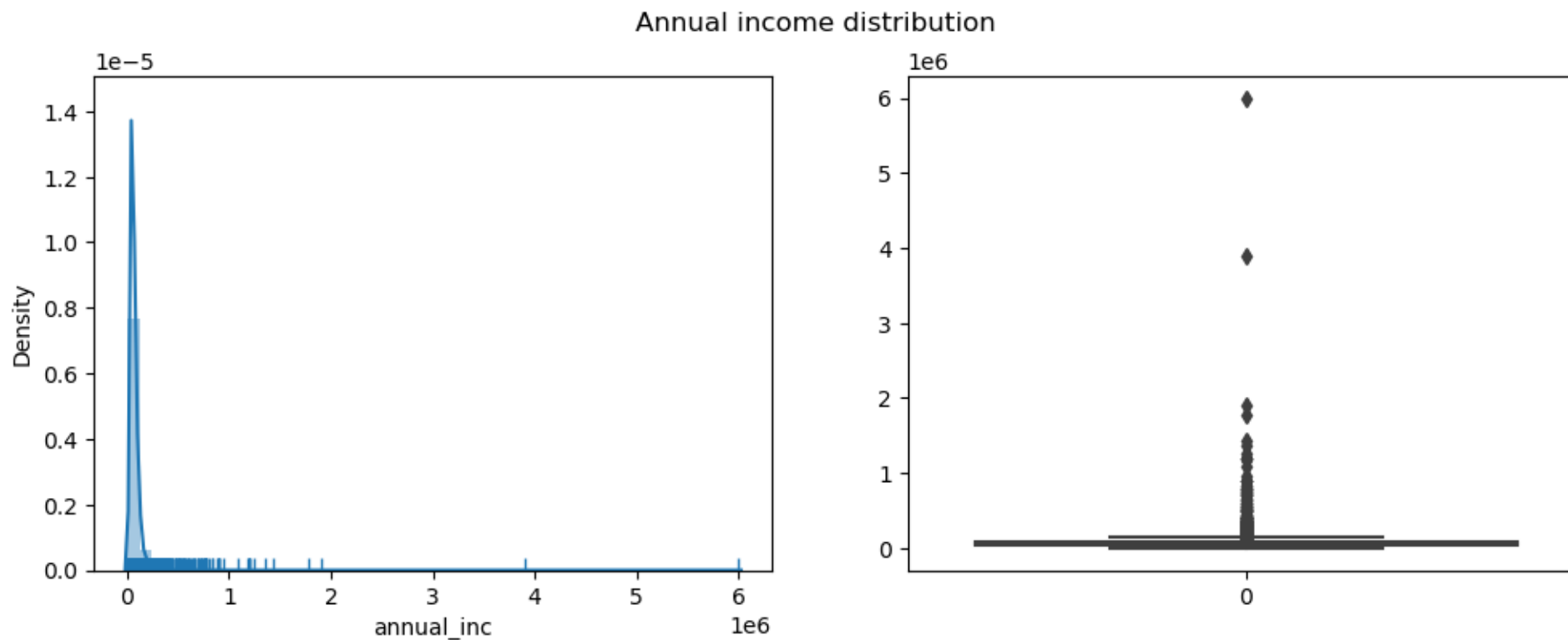
Installment Amount distribution



Observations:

The Most representative value of Installment amount is around 280.

Analysis

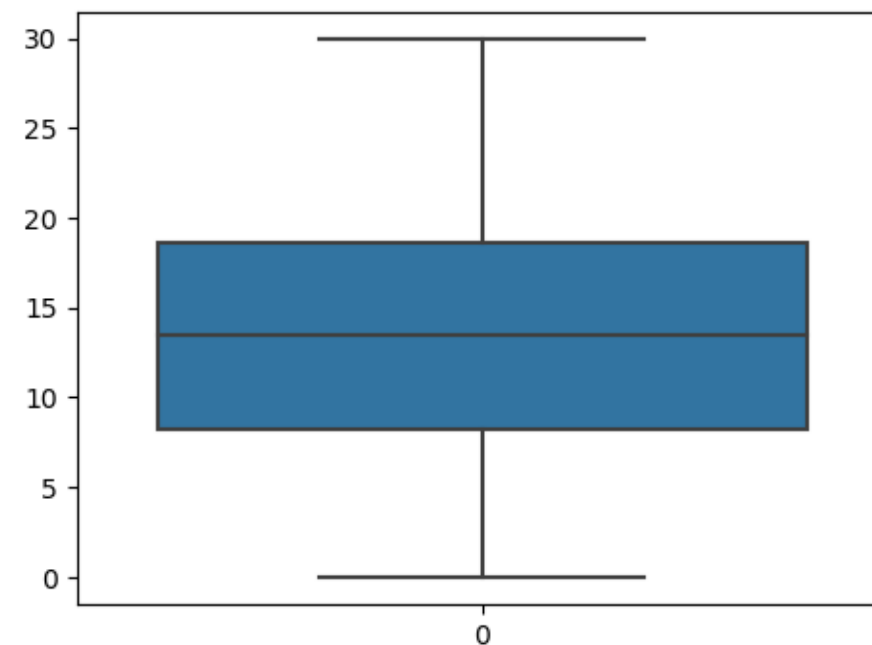
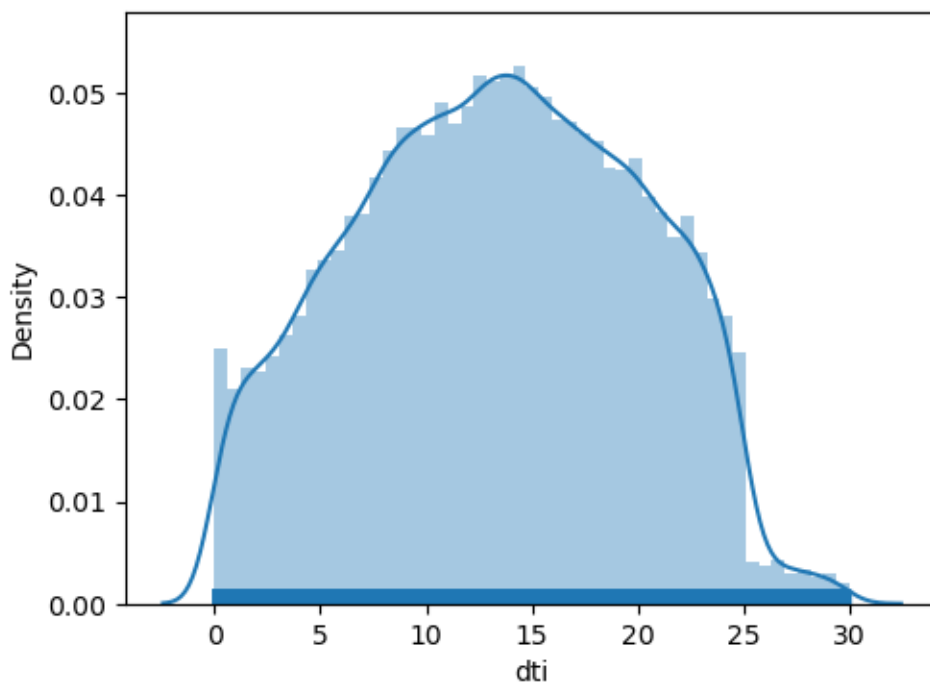


Observations:

Around 361 people are having more the 235000 annual income. and 99% of people are having incomes below 235000.

Analysis

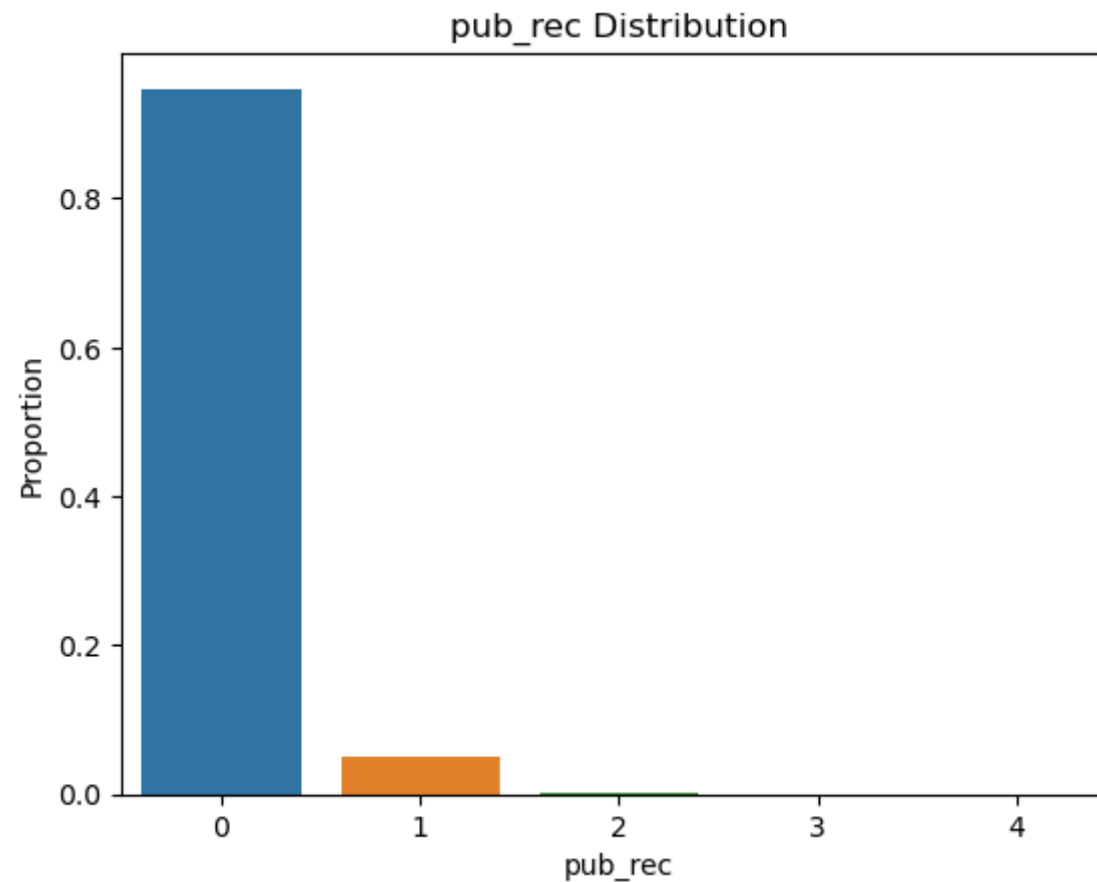
Debt to Income distribution



Observations:

Looks like there are no outliers and the distribution is very much similar to normal distribution. This is good sign that all the loans are given to barrower's who have Debt to Income ration less than 30.

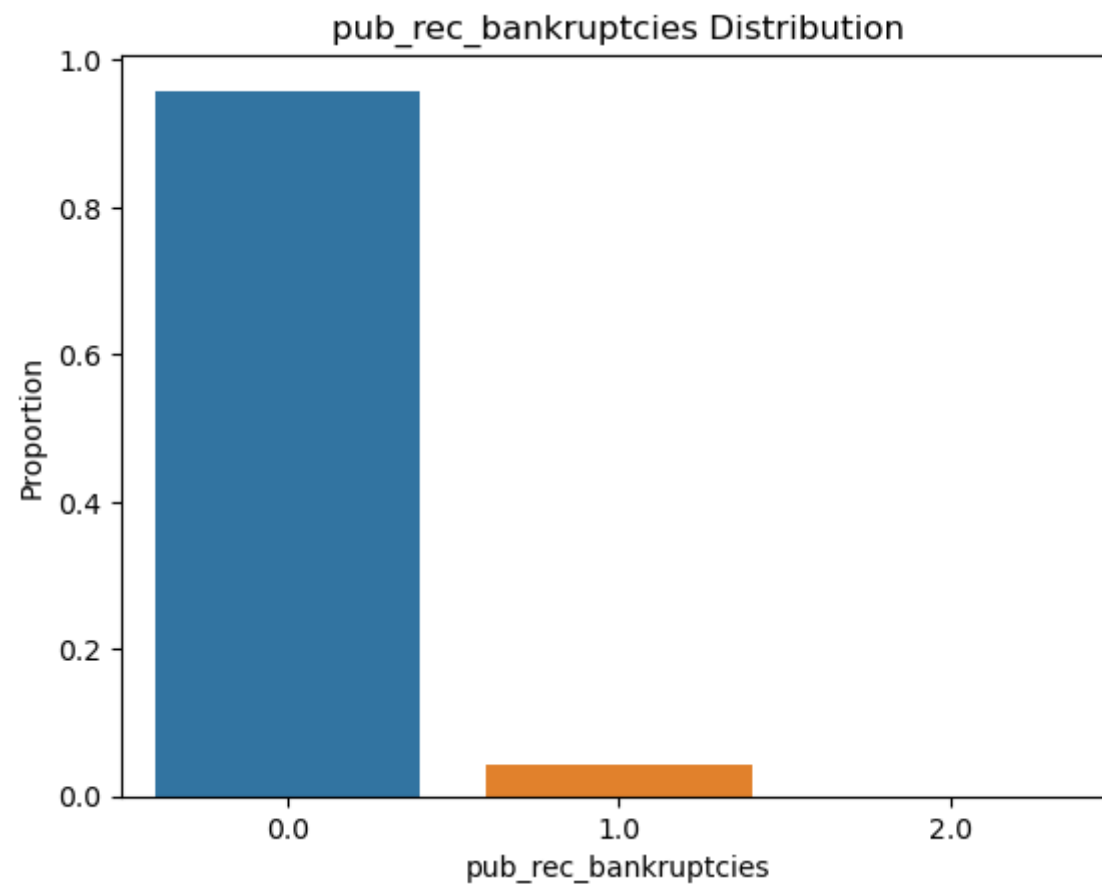
Analysis



Observations:

Around 90% borrower's are having no public derogatory records.

Analysis

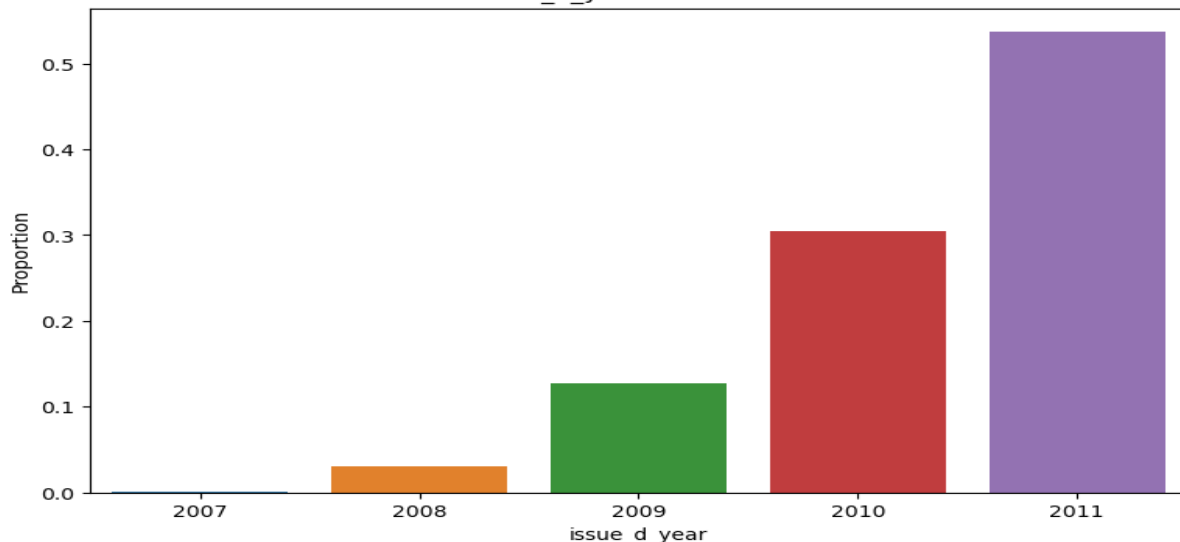


Observations:

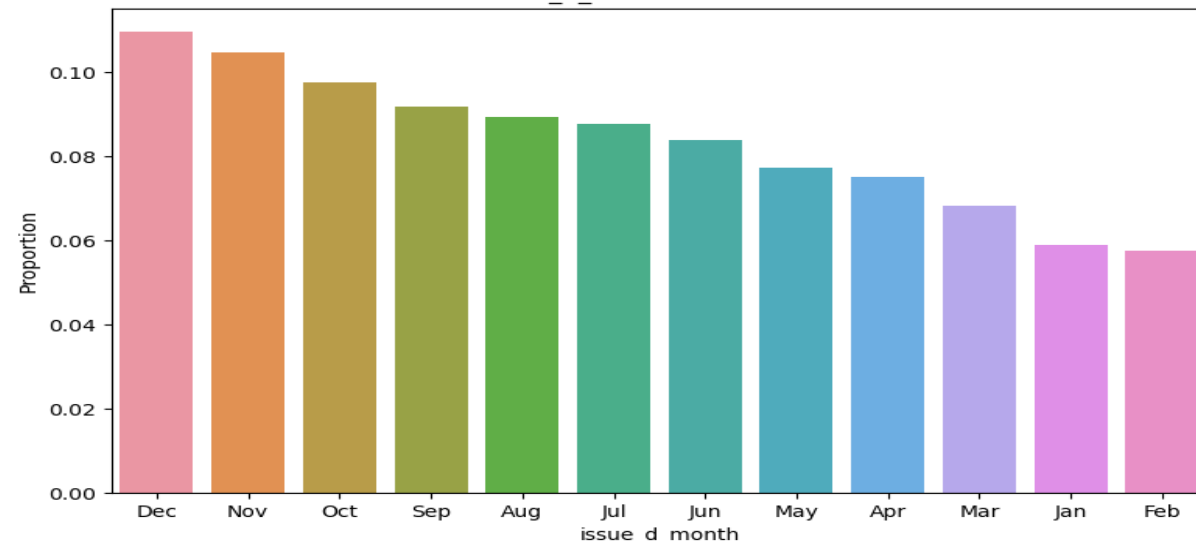
Around 99% people have not went bankrupt..

Analysis

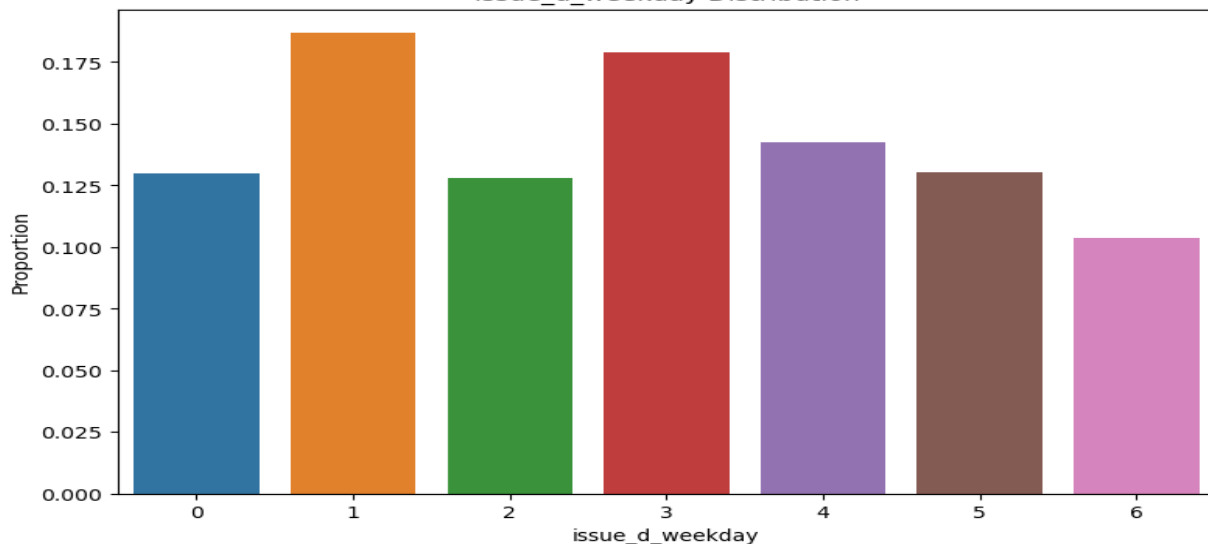
issue_d_year Distribution



issue_d_month Distribution



issue_d_weekday Distribution



Observations:

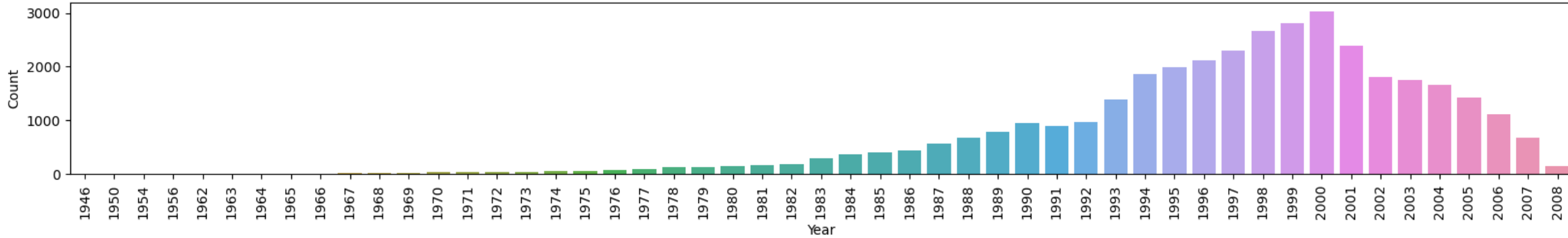
The lending club(LC) has doubling loan issues every year.

There are more issues of loan in last 3 months every end of the ear i.e., Oct, Nov and Dec.

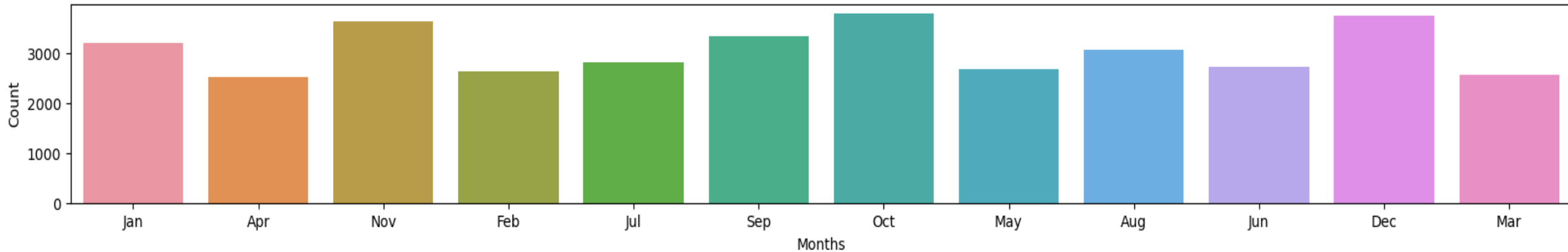
LC has issued more loans on Tuesday and Wednesday than other week-days.

Analysis

Count vs Earliest cr line year



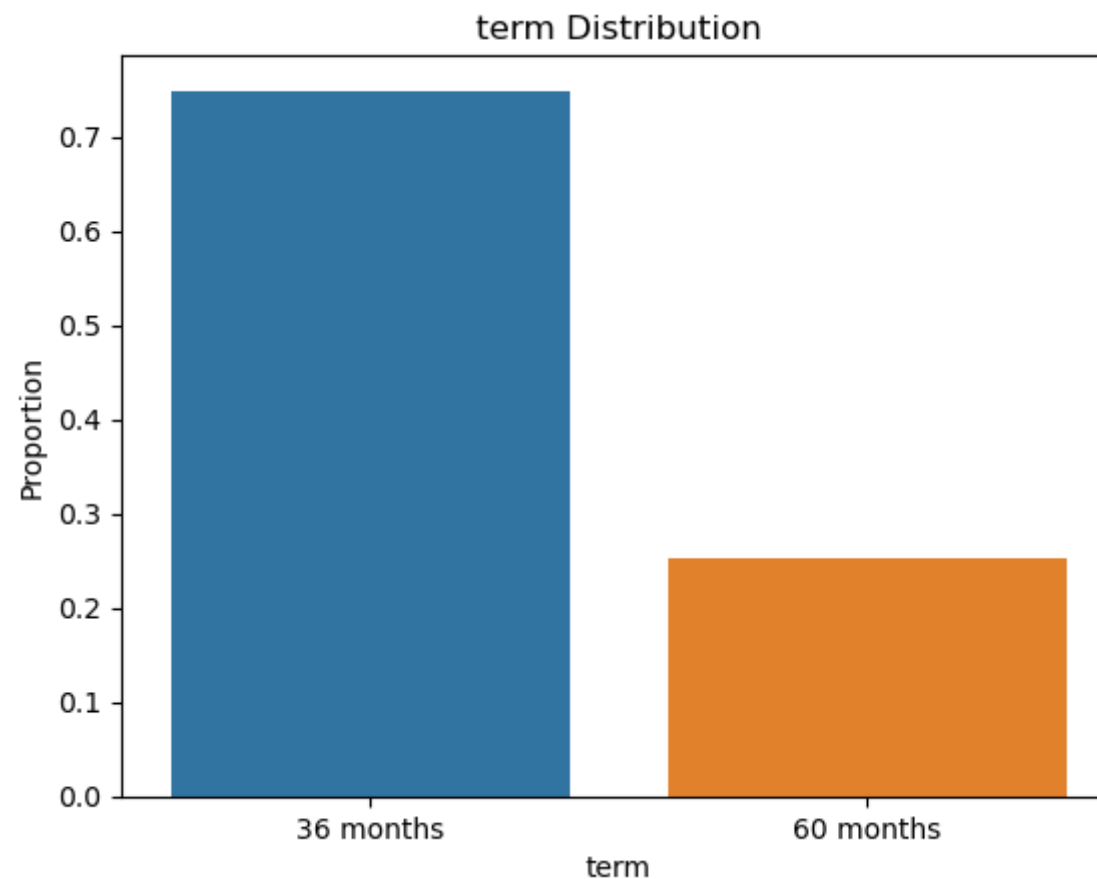
Count vs Earliest cr line month



Observations:

Many of Loan borrowers of LC have got earlier credit line in 2000 year, and also most have got earlier credit line on end of the year i.e., Oct, Nov, Dec

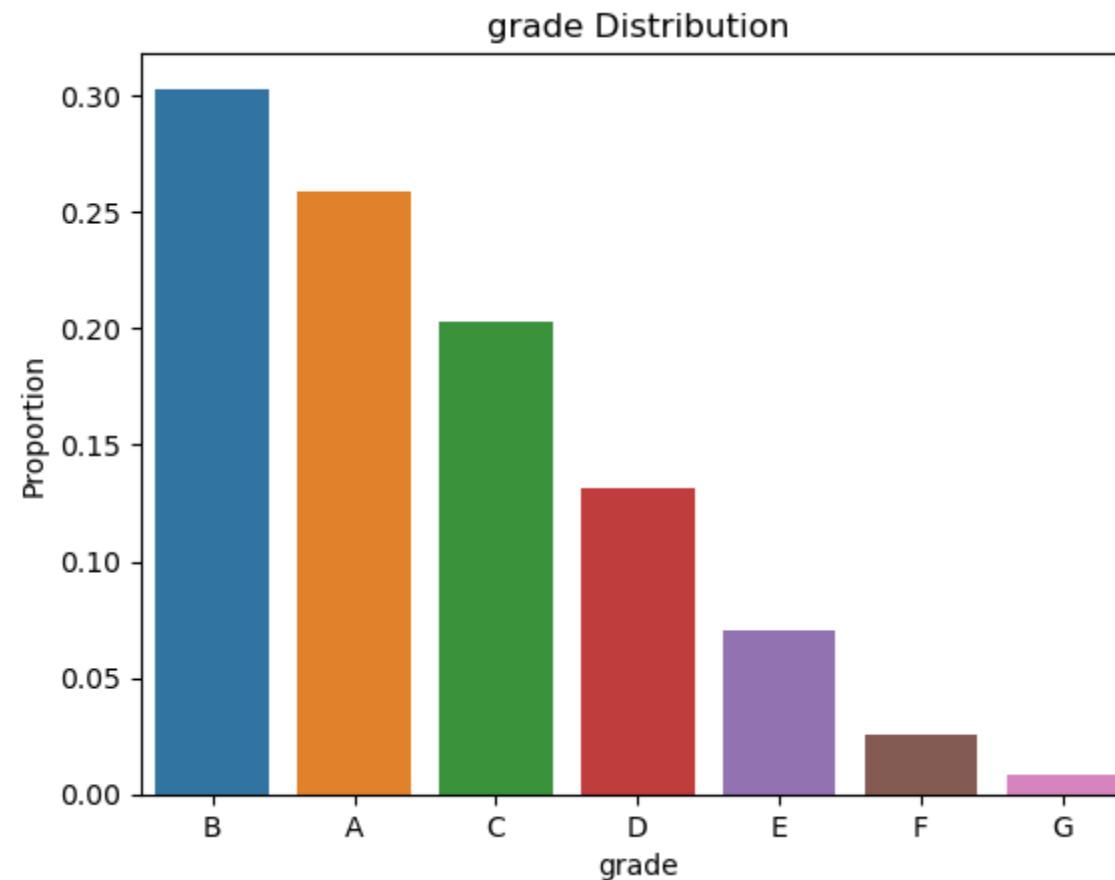
Analysis



Observations:

Barrows have taken 36 months tenure more than 60 months.

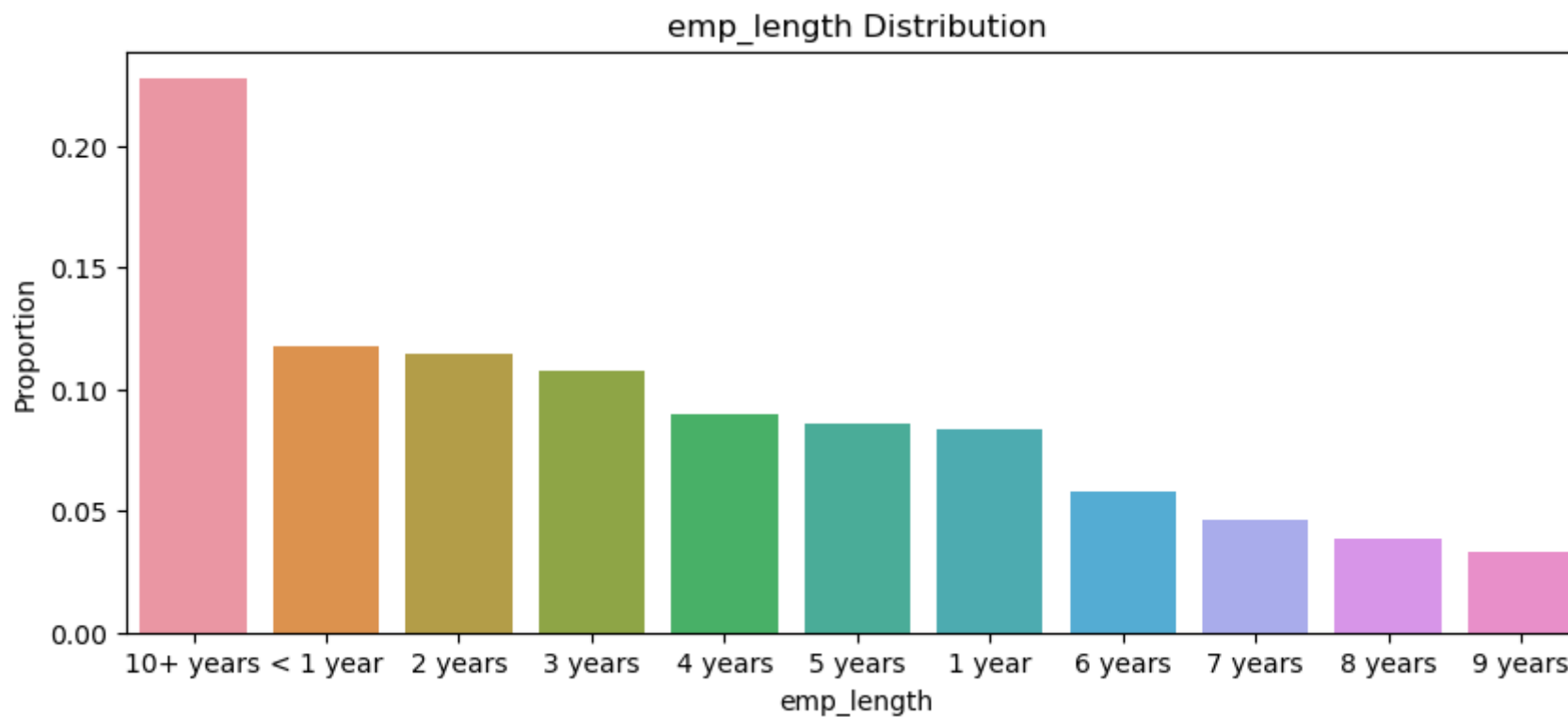
Analysis



Observations:

Most borrowers fall under A and B grades then other grades

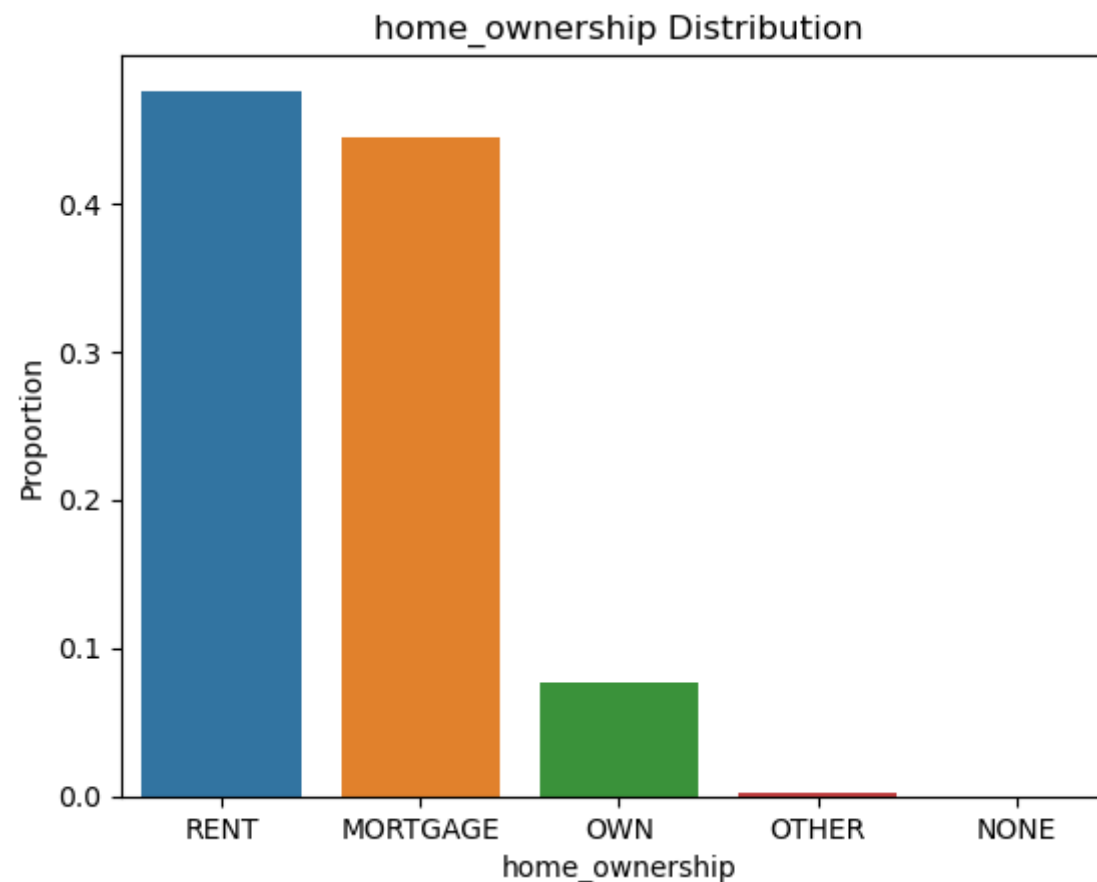
Analysis



Observations:

Borrowers are mostly 10+ years employment length.

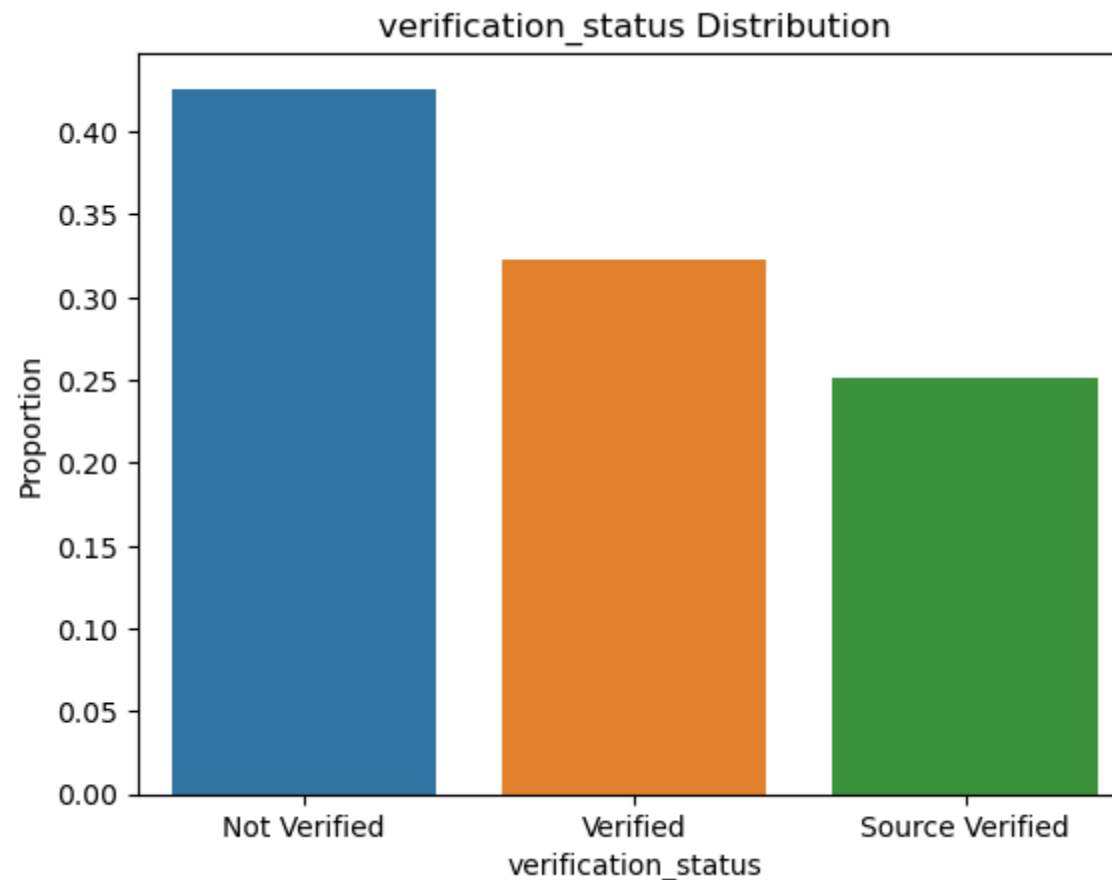
Analysis



Observations:

The loan borrowers are mostly having rented and mortgage houses.

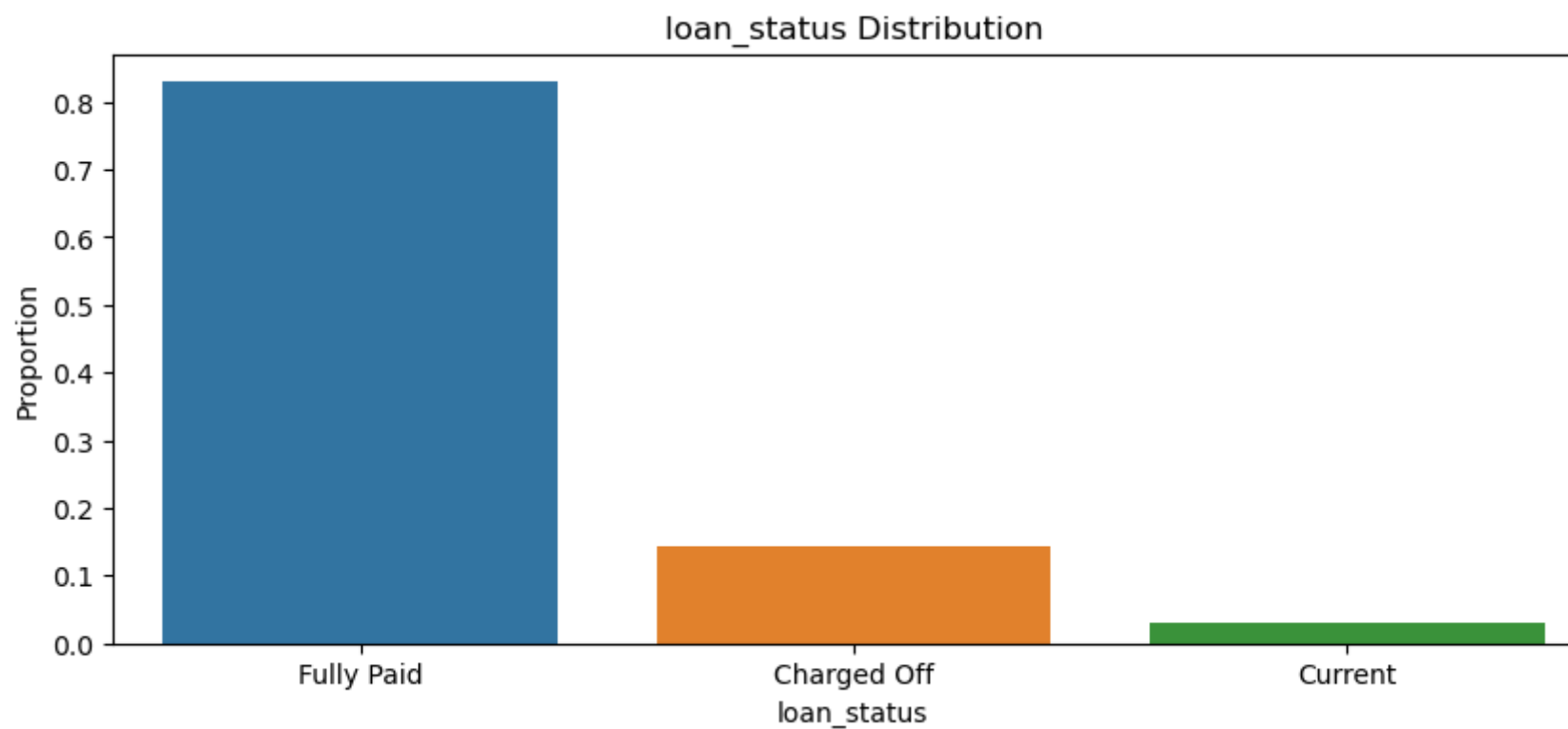
Analysis



Observations:

Majority of loans were given without verification of applicants income.

Analysis

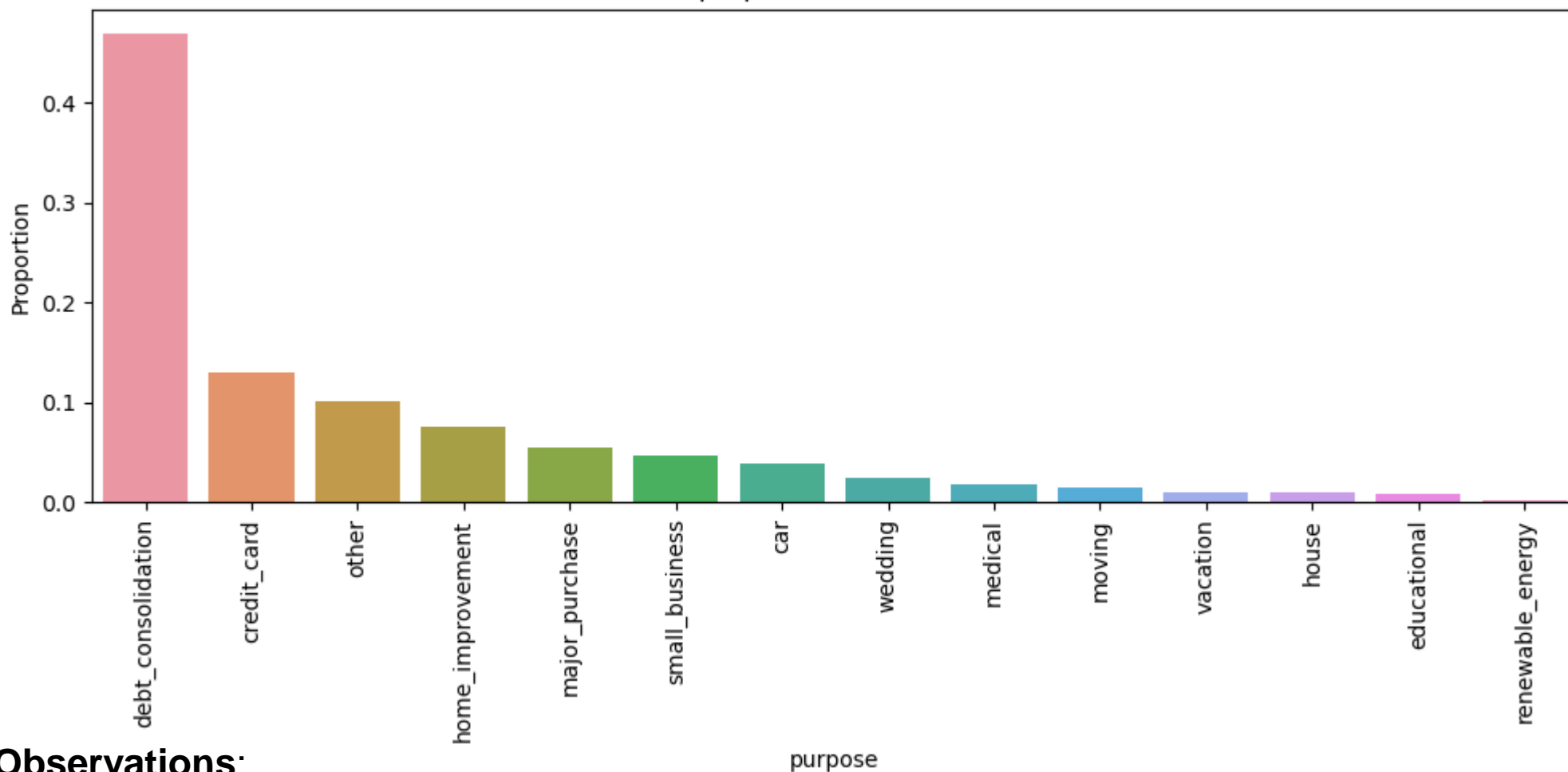


Observations:

85% of borrowers has paid the loan fully. where are 14% are defaulted the loan.

Analysis

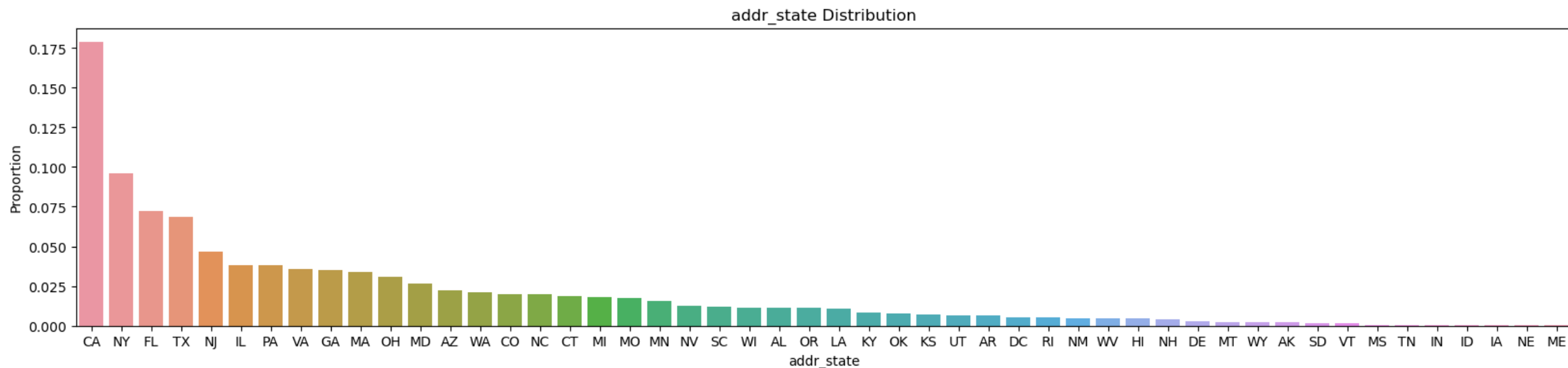
purpose Distribution



Observations:

Looks like more number of people took loan for *debt consolidation* and a very few people took for *renewable energy*

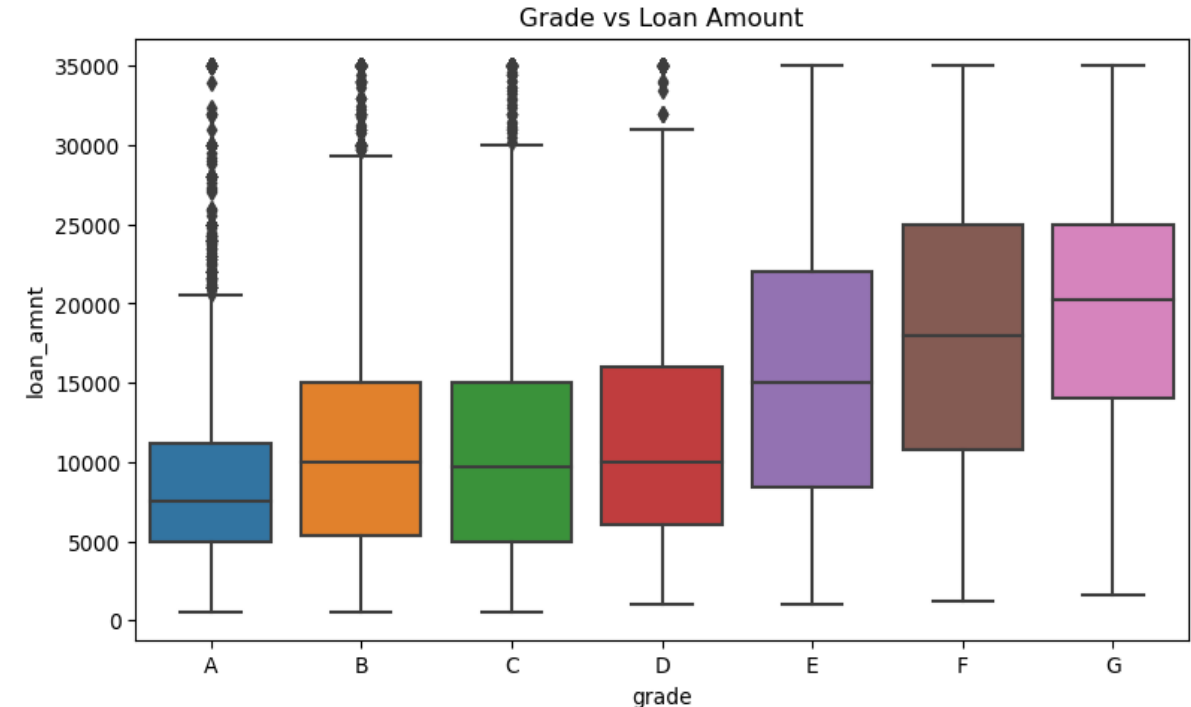
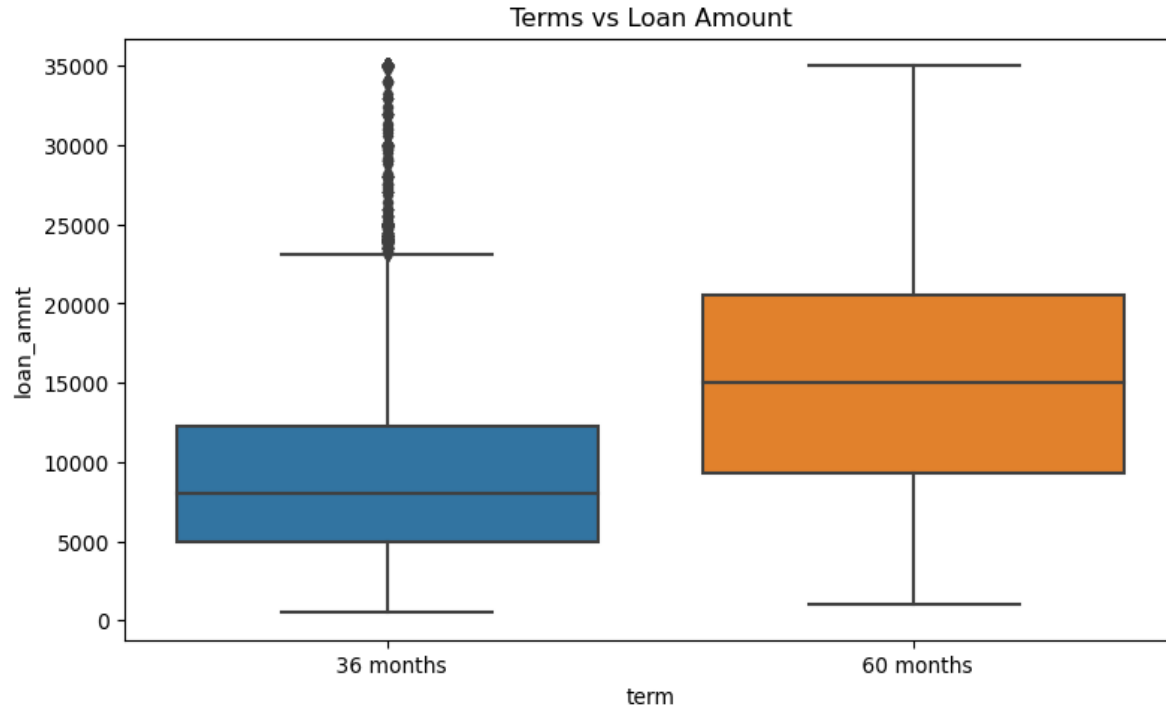
Analysis



Observations:

Most of the borrowers are from *CA* and *NY*

Analysis

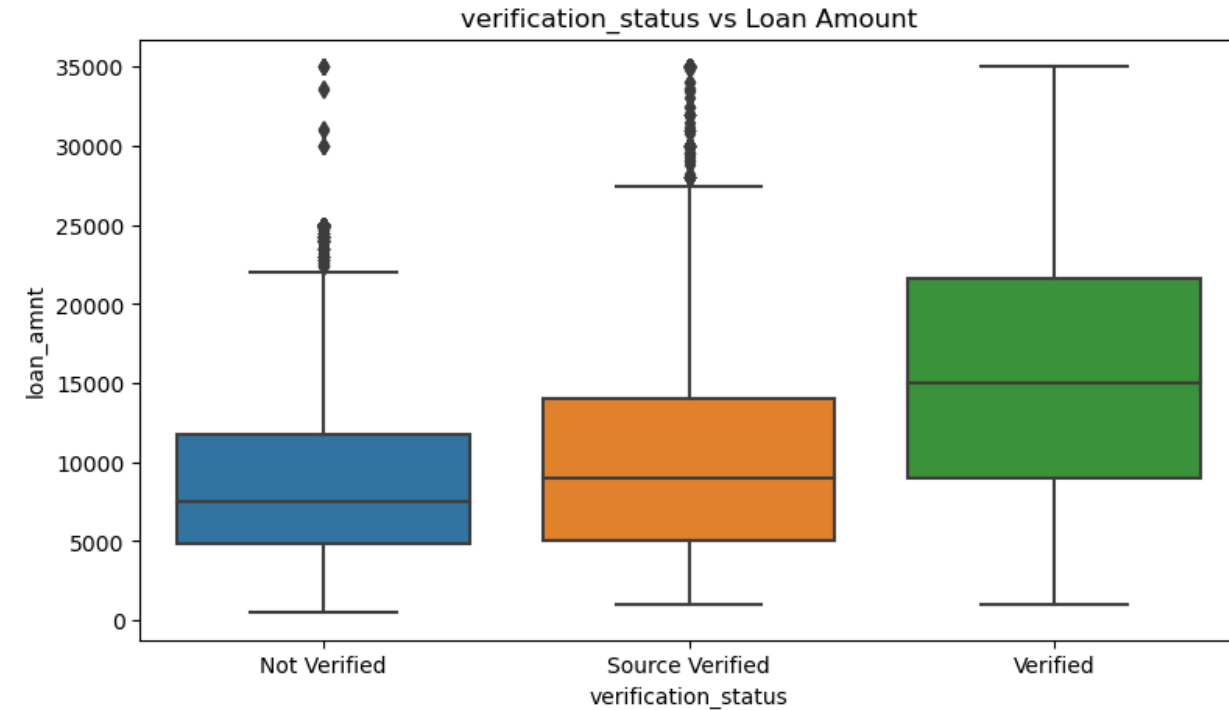
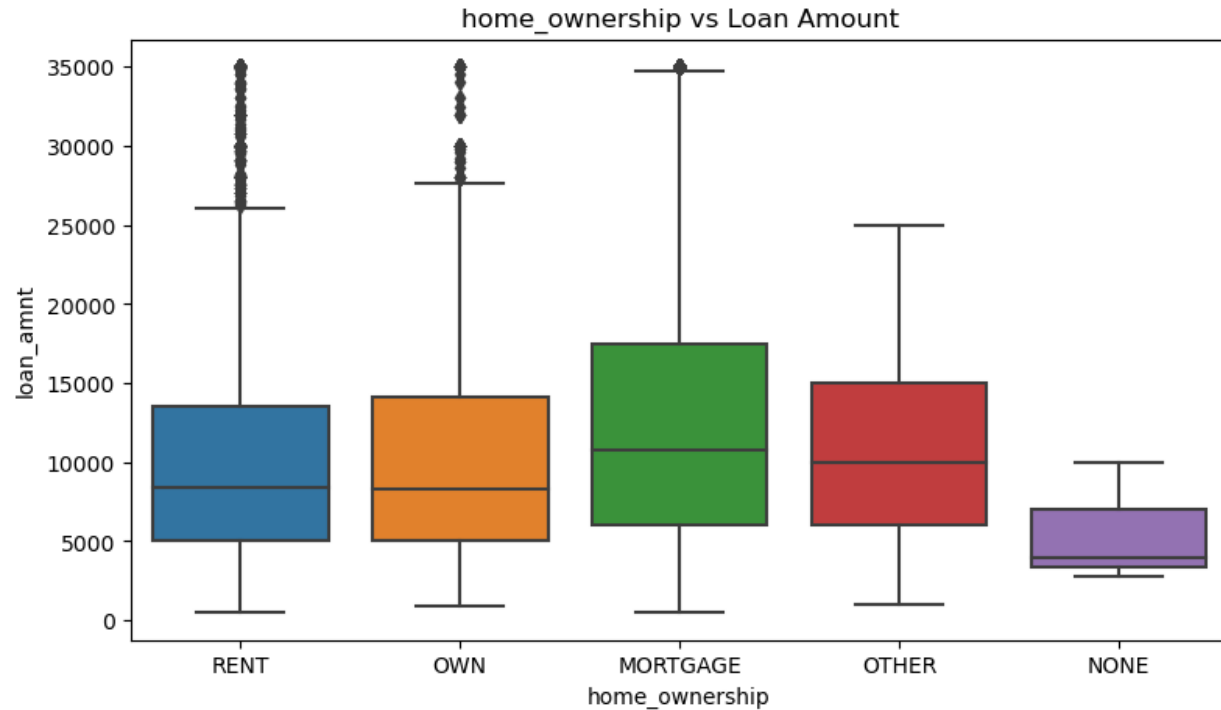


Observations:

Higher amount loans have high tenure i.e, 60 months.

Grade 'G' and 'H' have taken max loan amount. As Grades are decreasing the loan amount is increasing.

Analysis

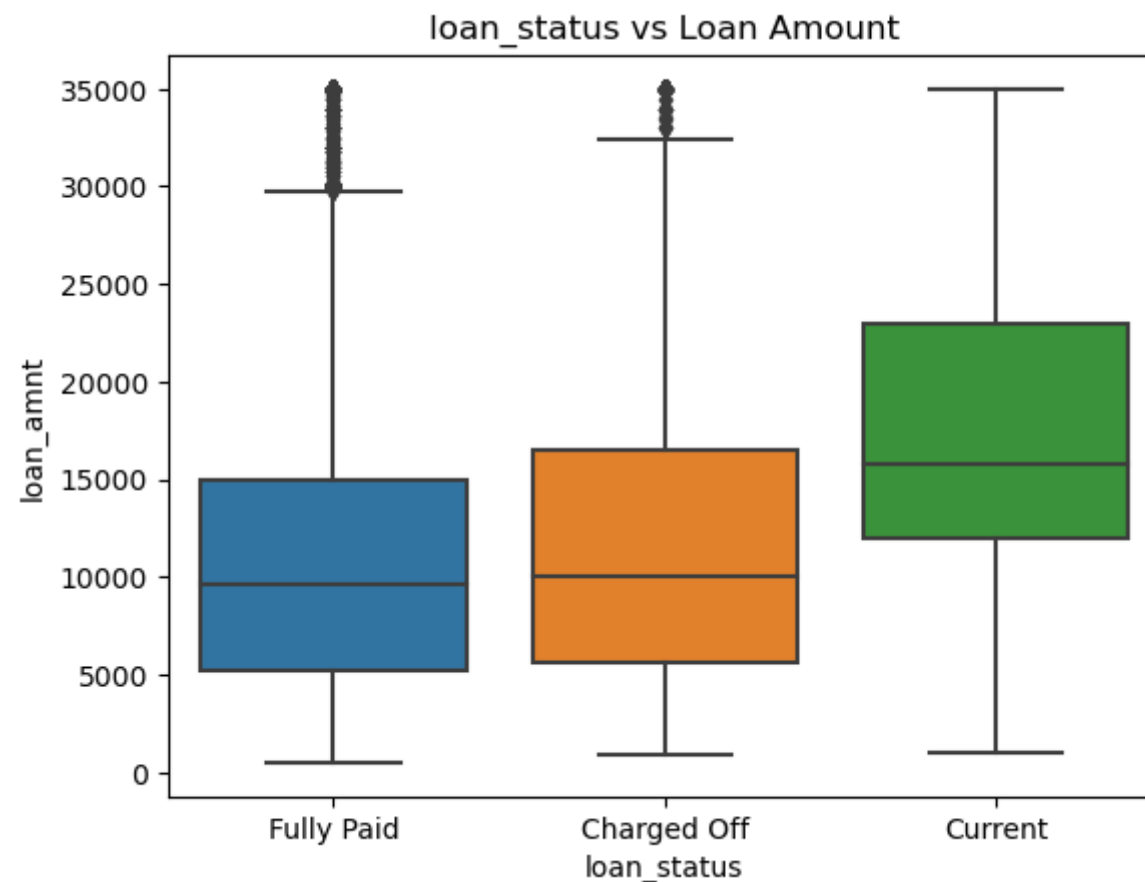


Observations:

more borrowers are from MORTGAGE and also the median loan amount also high for MORTGAGE owned borrowers.

And most of borrowers are verified for borrowing loan >9k

Analysis

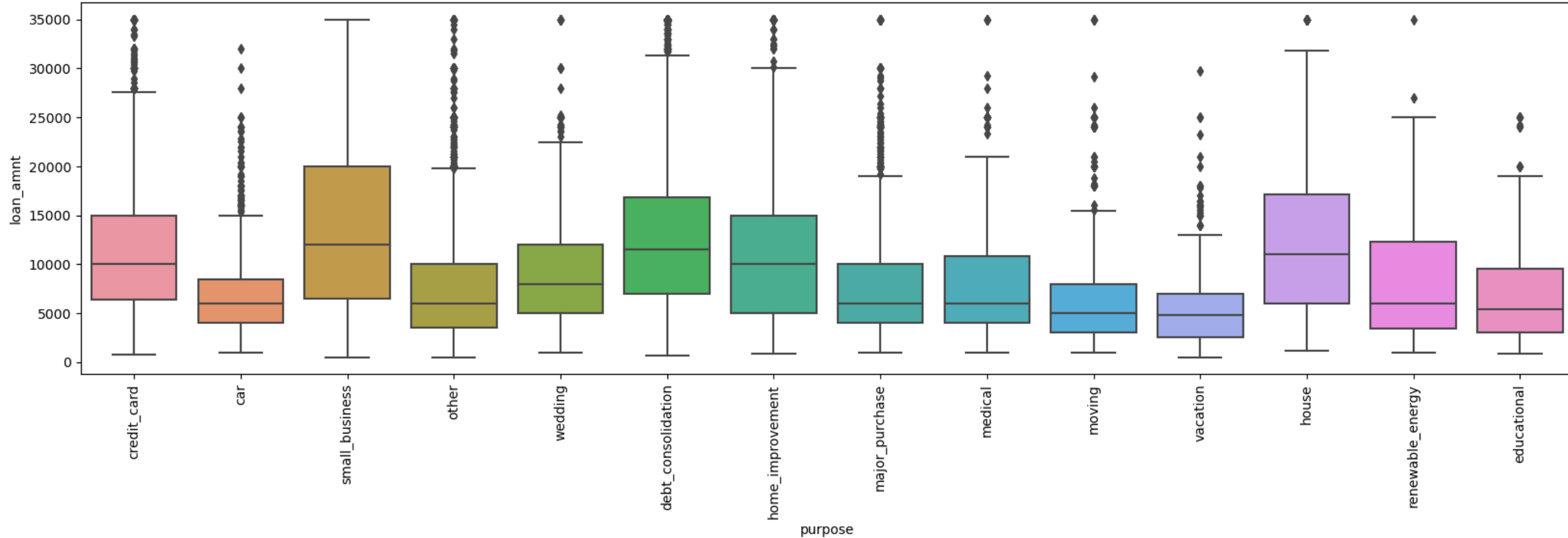


Observations:

Charged Off loans have higher amounts than Fully Paid ones.

Analysis

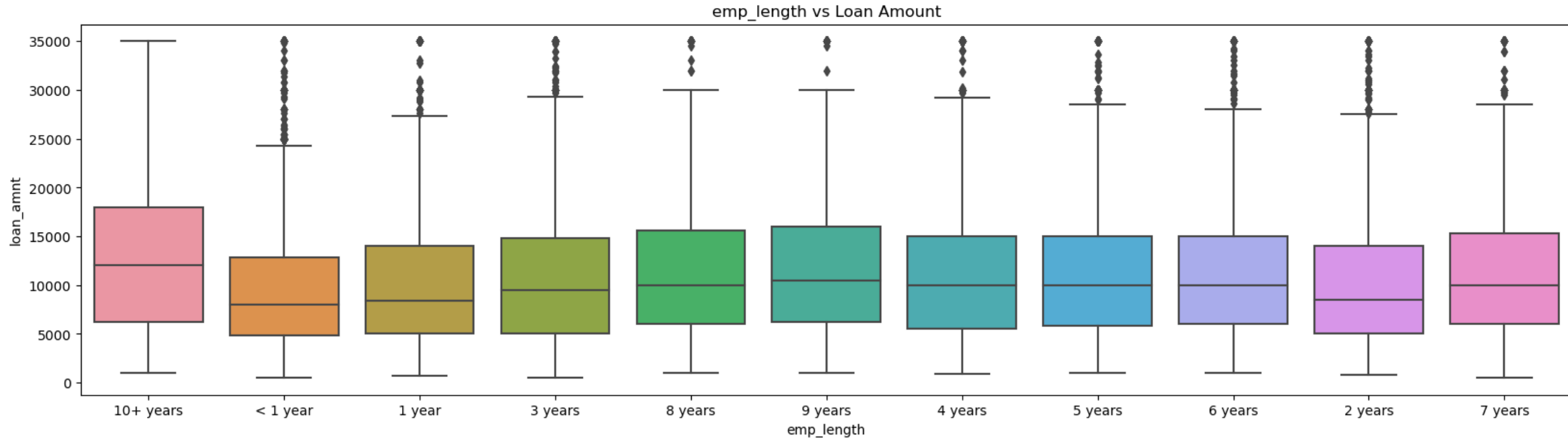
purpose vs Loan Amount



Observations:

More loan amount is from Small bussiness followed by debt consolidation.

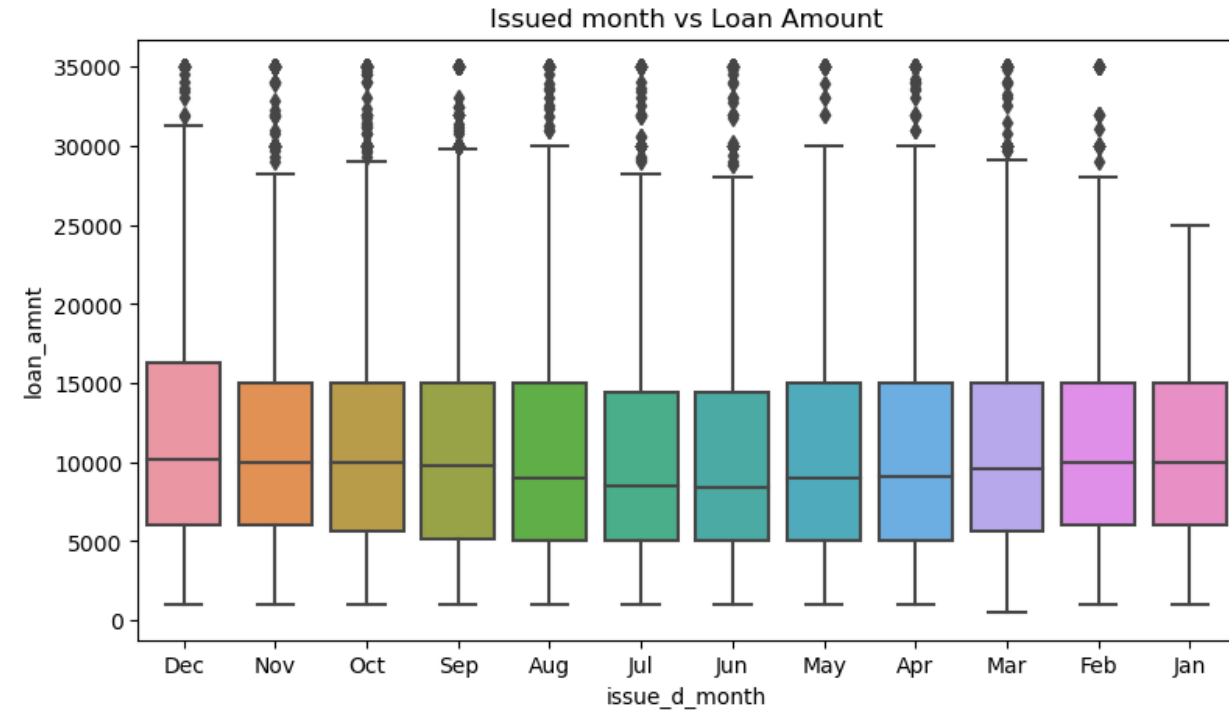
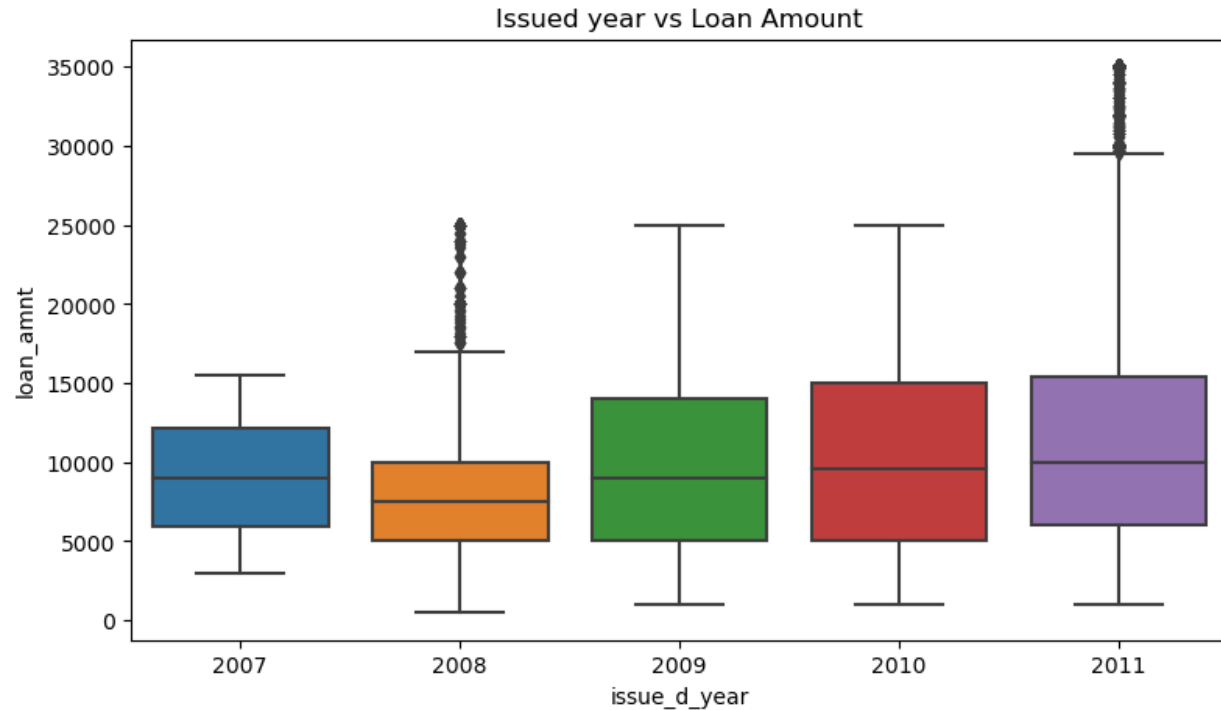
Analysis



Observations:

more borrowers are from 10+ years and least is <1 year.

Analysis

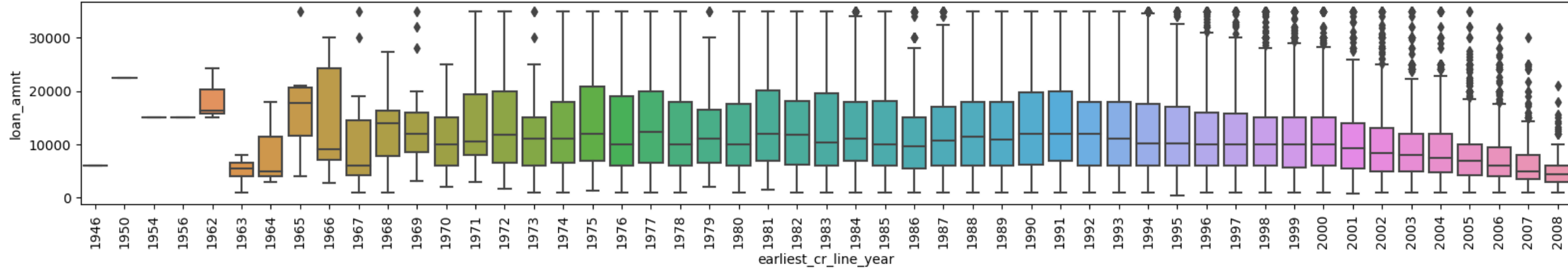


Observations:

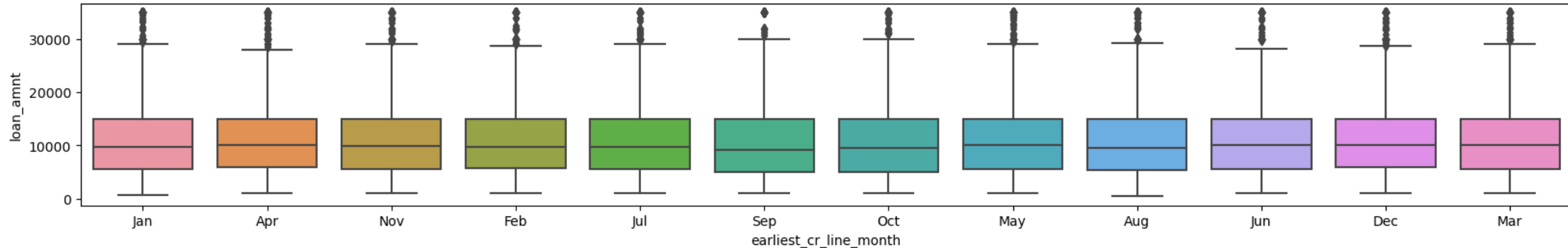
- The median loan amount in each year did not change much but the distribution is more spread as the years increase, which means people have taken different loan amounts in each year.
- Few Loan borrowers took higher loan amount in 2008 and 2011 which are plotted as outliers.
- Loan borrowers took almost similar amounts in all the months except in December, people have taken higher amounts as distribution is high above median.

Analysis

Earliest Credit Line year vs Loan Amount



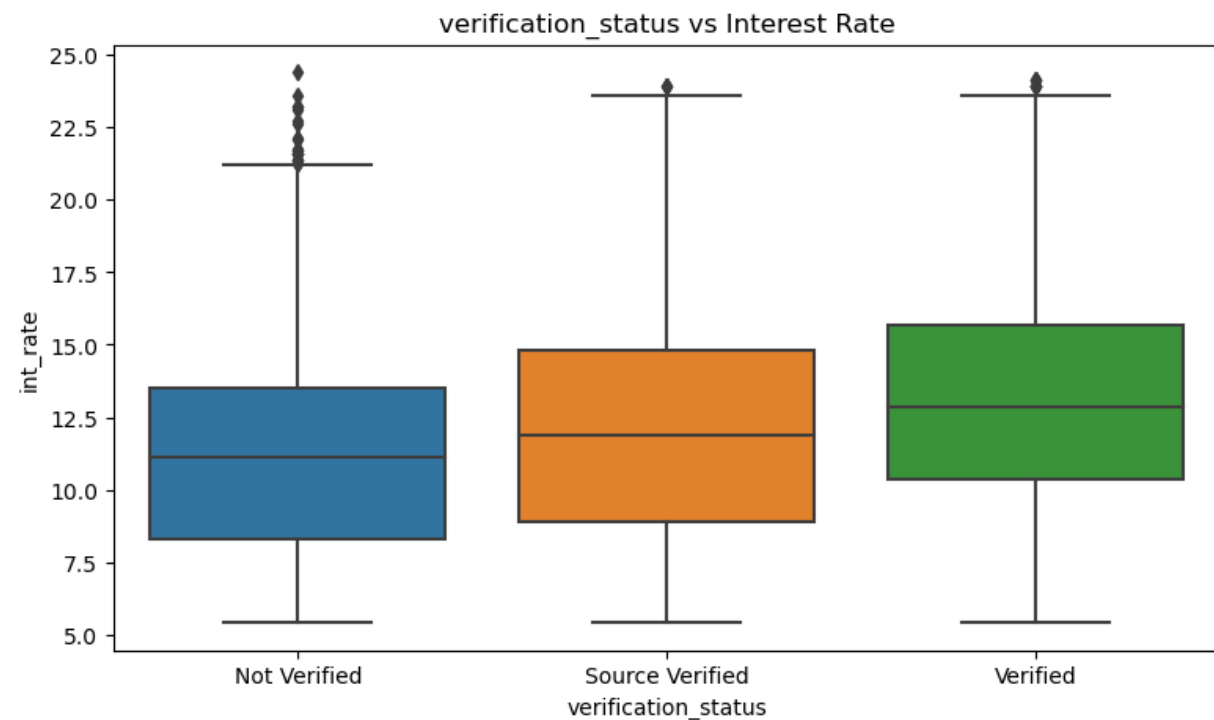
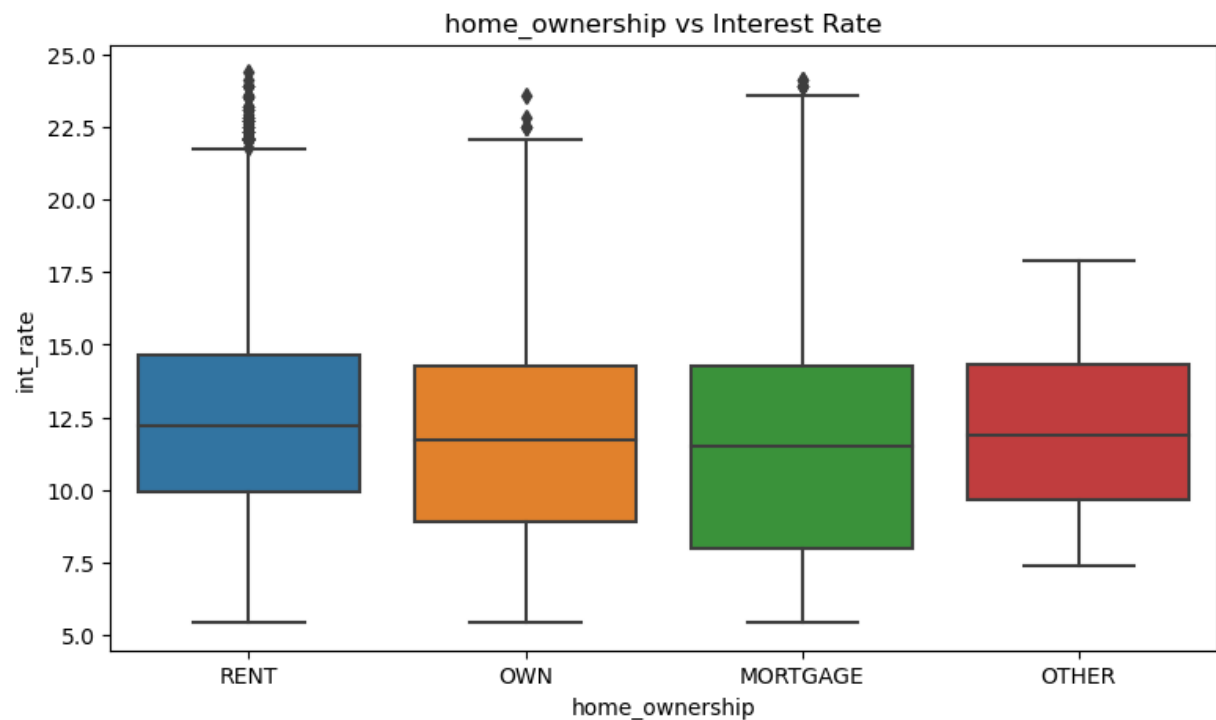
Earliest Credit Line month vs Loan Amount



Observations:

Borrowers who go earliest credit line in 1966 got wide spreaded amount of loans than others.

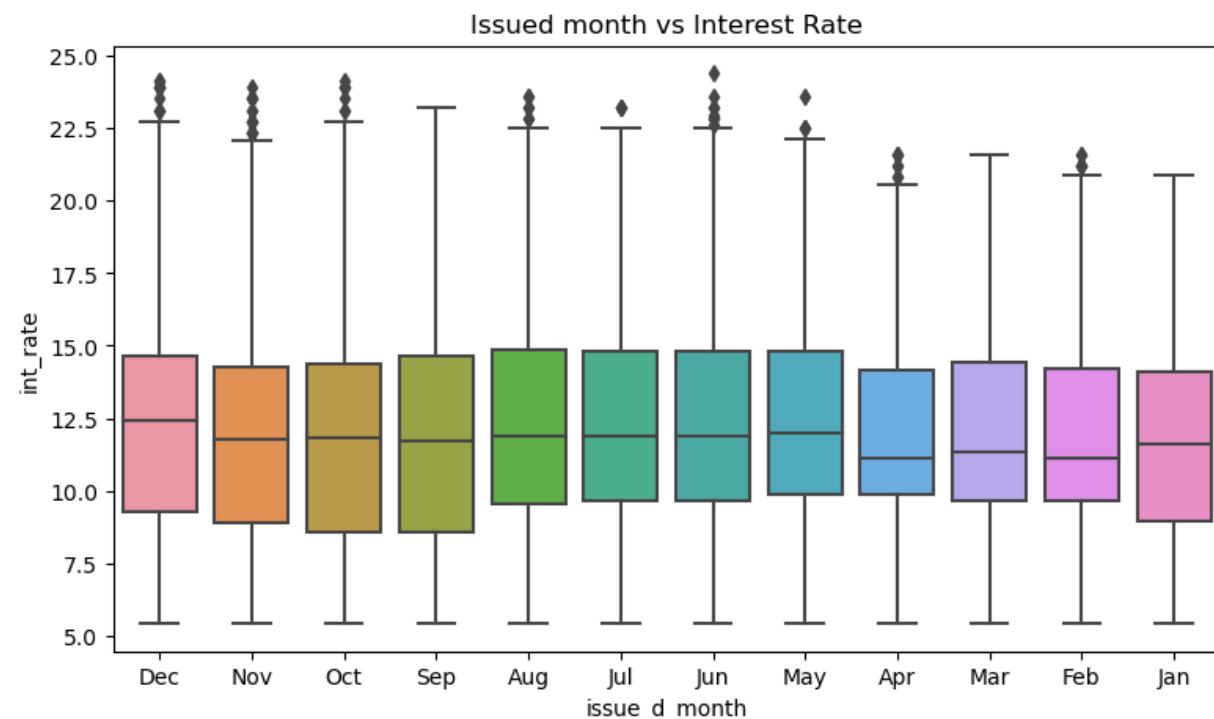
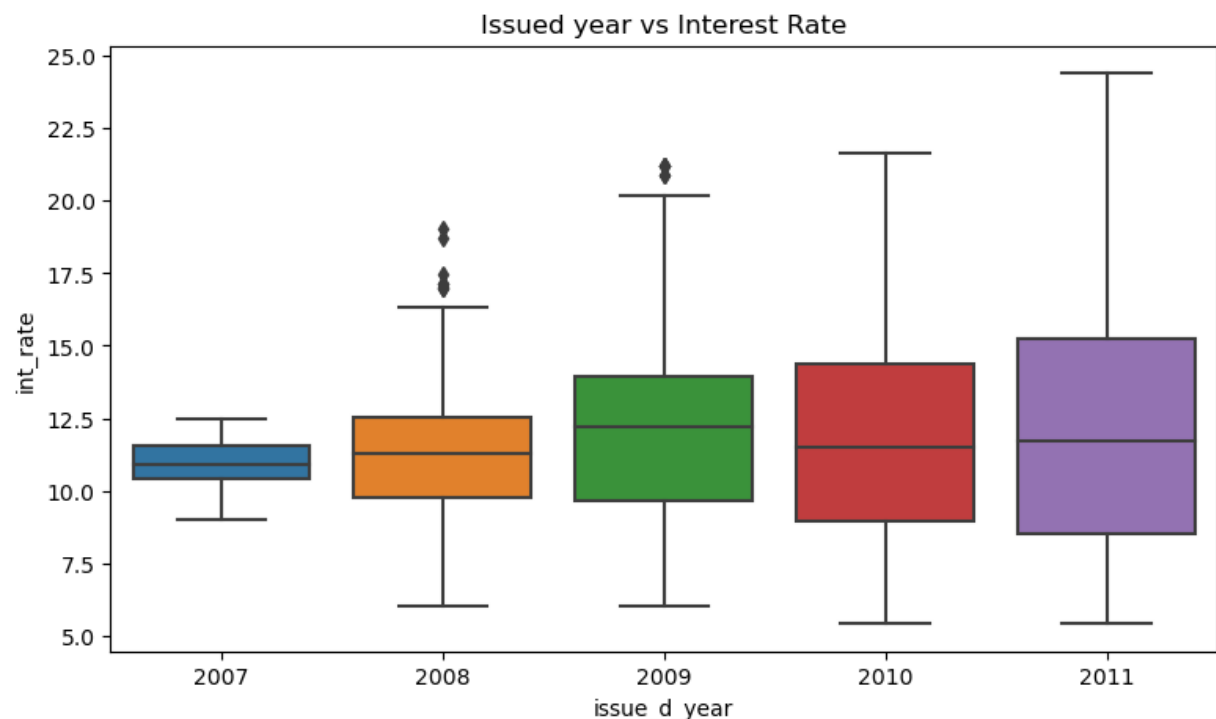
Analysis



Observations:

- Barrowers with Own and Mortgage got loans with less interest rates as the security will be there.
- And The Non-Verified Barrowers got less interest rates compared to Verified and Source Verified barrowers.

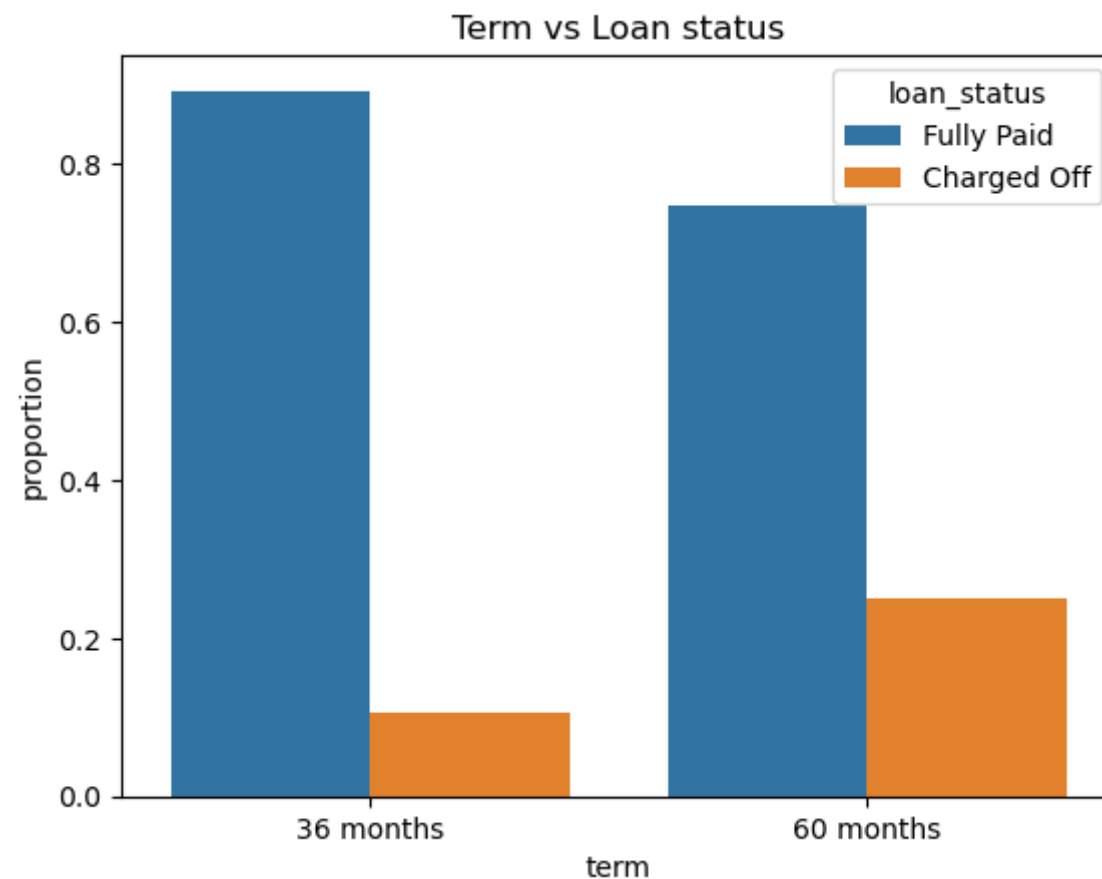
Analysis



Observations:

As the years of business increase the interest rates are getting more diversified but median of of interest rate is quite same in all the years.

Analysis



Observations:

Looks like there are more proportion of borrowers defaulted loan in 60 months term then 36 months. Also the Fully Paid rate is higher in 36 months tenure.

Conclusions

- ☐ Lending club should reduce the high interest loans for 60 months tenure, they are prone to loan default.
- ☐ Grades are good metric for detecting 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 issues 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 stop giving loans to this category when loan amount requested is more than 12000.

- ☐ People with more number of public derogatory records are having more chance of filing a bankruptcy.

Lending club should make sure there are no public derogatory records for borrower.