Lending Club Case Study

Group members:
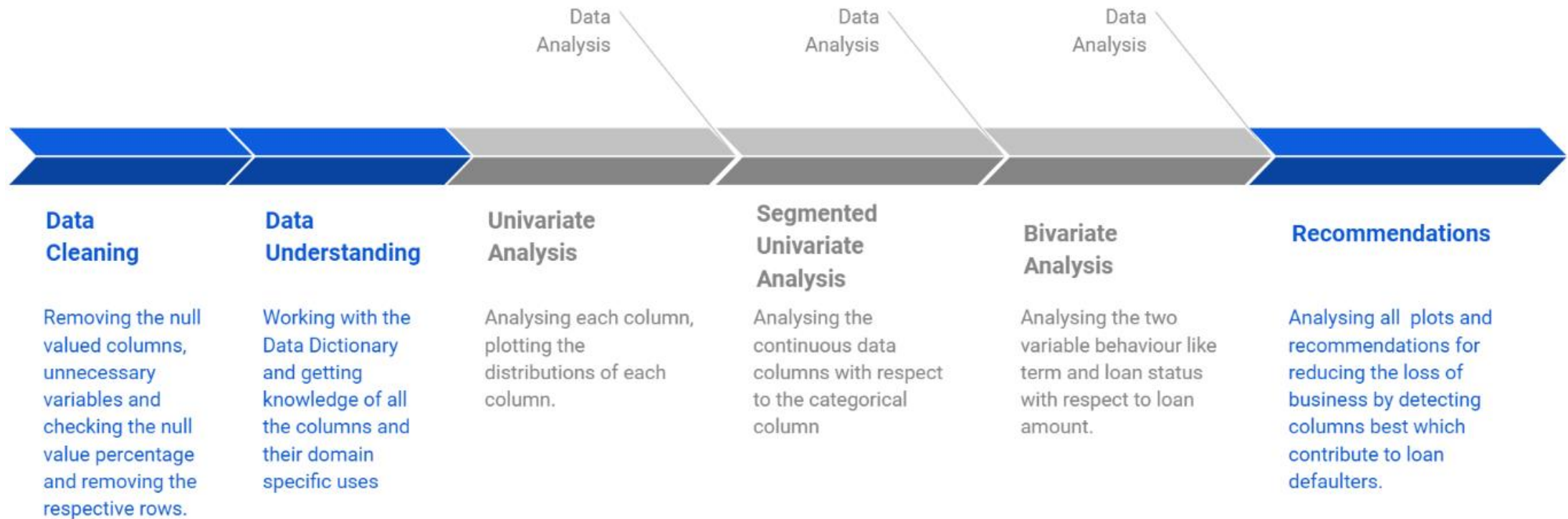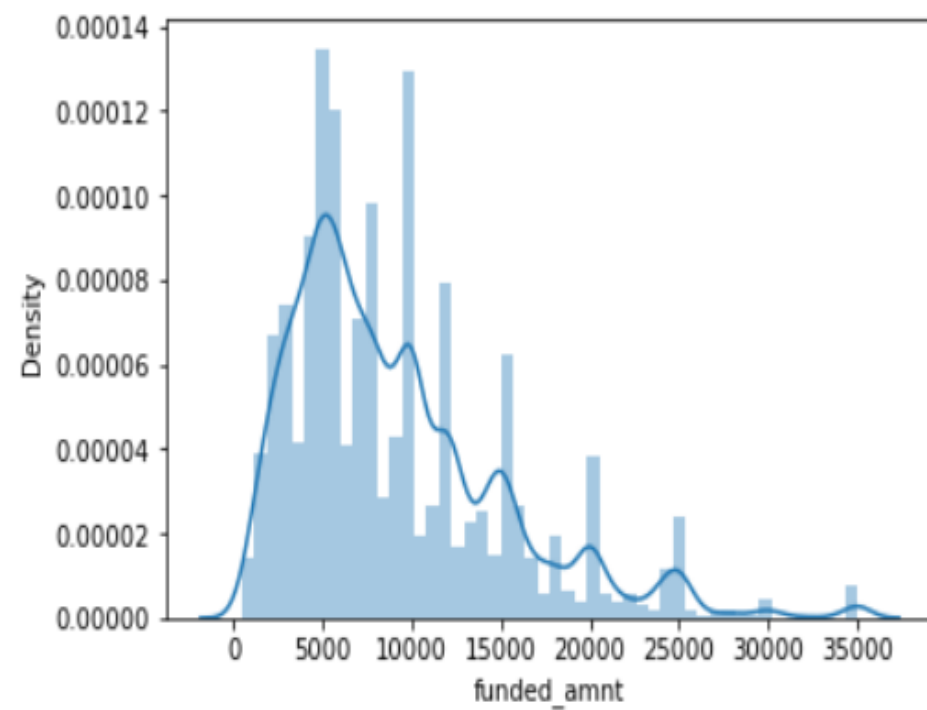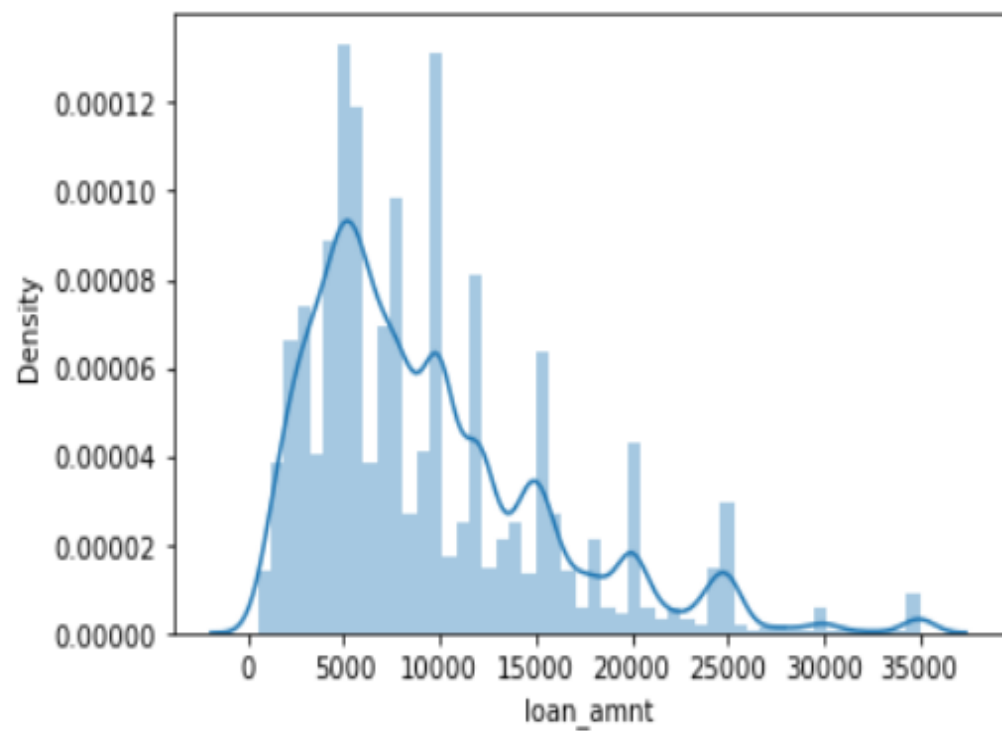
Rakesh Panigrahy

Chinmay Nayak

Abstract

- Lending club is the largest online loan market place, facilitating personal loans, business loan and financing of medical procedures

- Borrowers can easily access lower interest rate loans through a fast online interface

- The objective of the analysis is to use the information of the past loan applicants and find whether they 'defaulted or not'
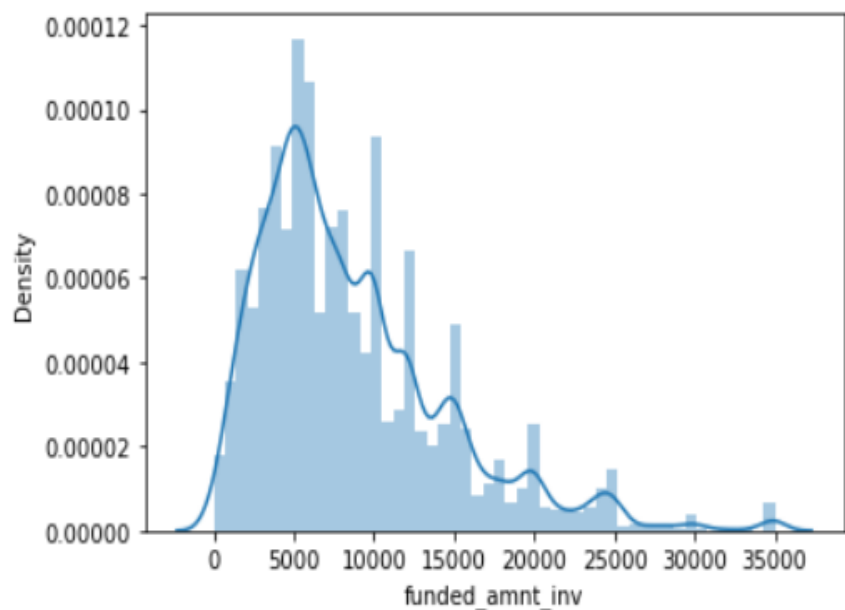
Problem solving methodology

Data Analysis

Data Analysis

Data Analysis

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 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.
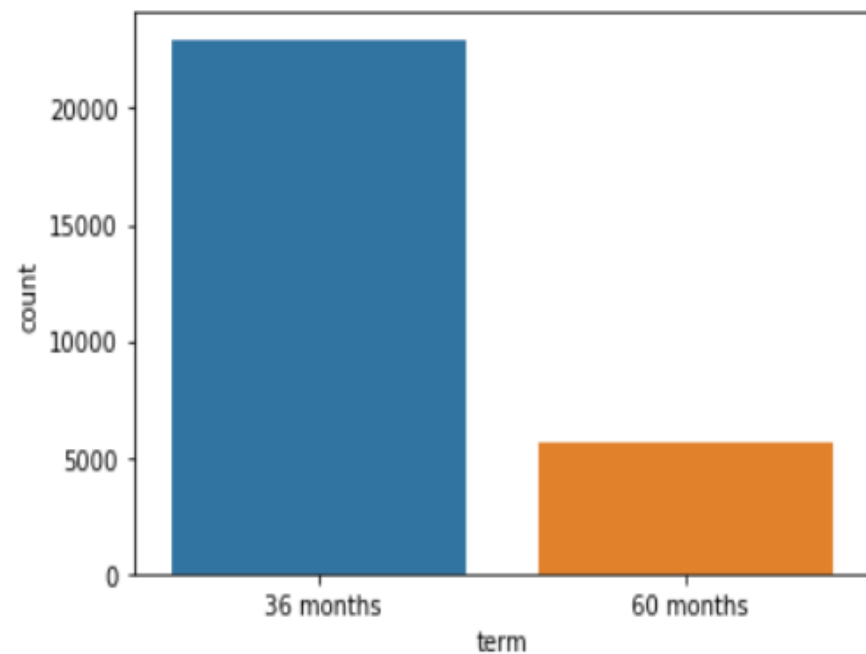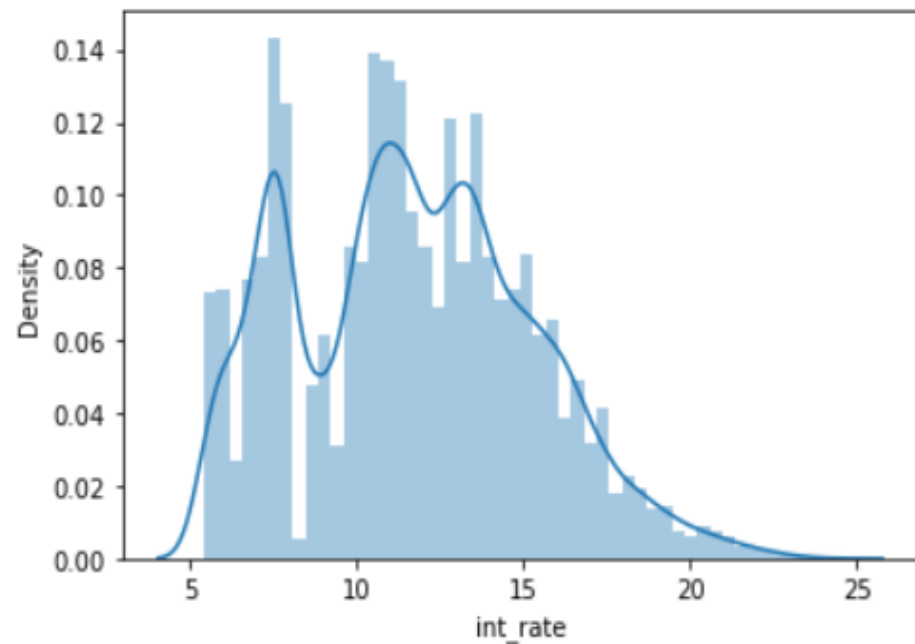
Analysis

Analysis

Mostly people apply for loan for 4000 to 7000 and we can see spikes at 5000, 10000, 15000, 20000 and so on. So mostly people prefer to go for such kind of amounts.
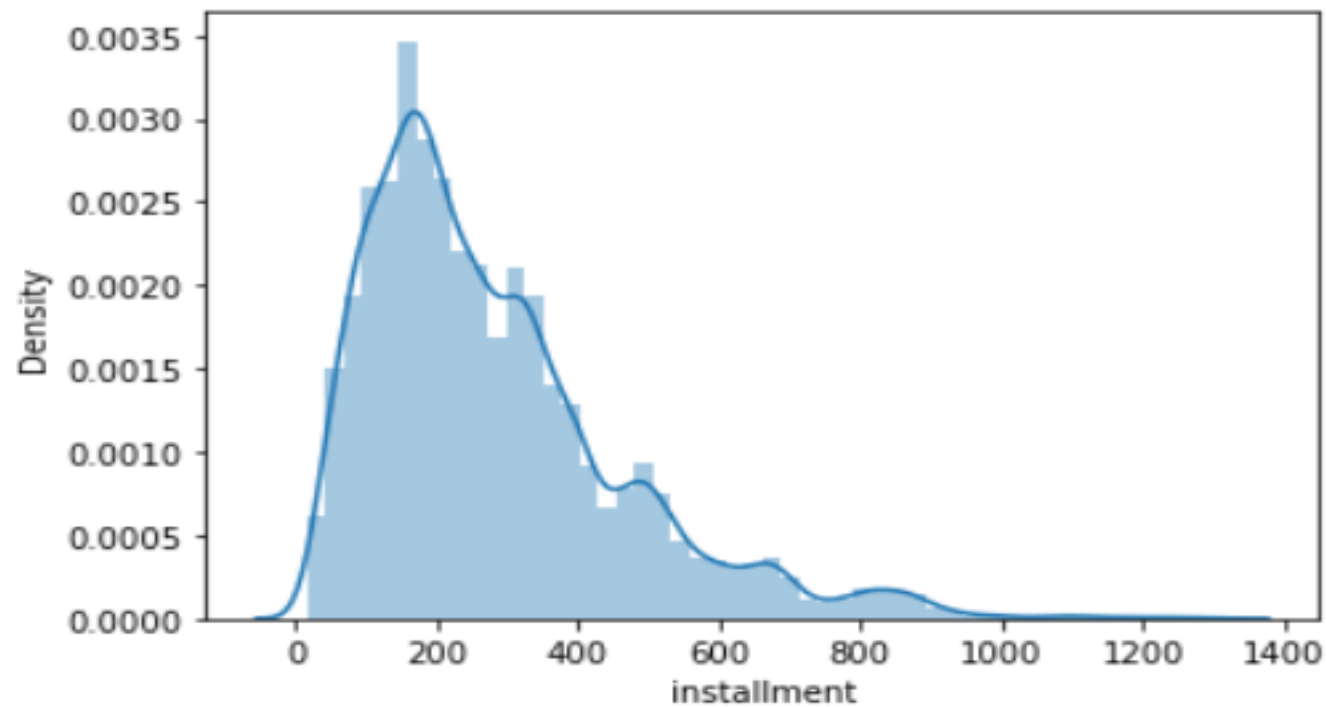
Analysis

Mostly people go for 36 months plan than 60 months plan. People are interested in short term loan than long term loan
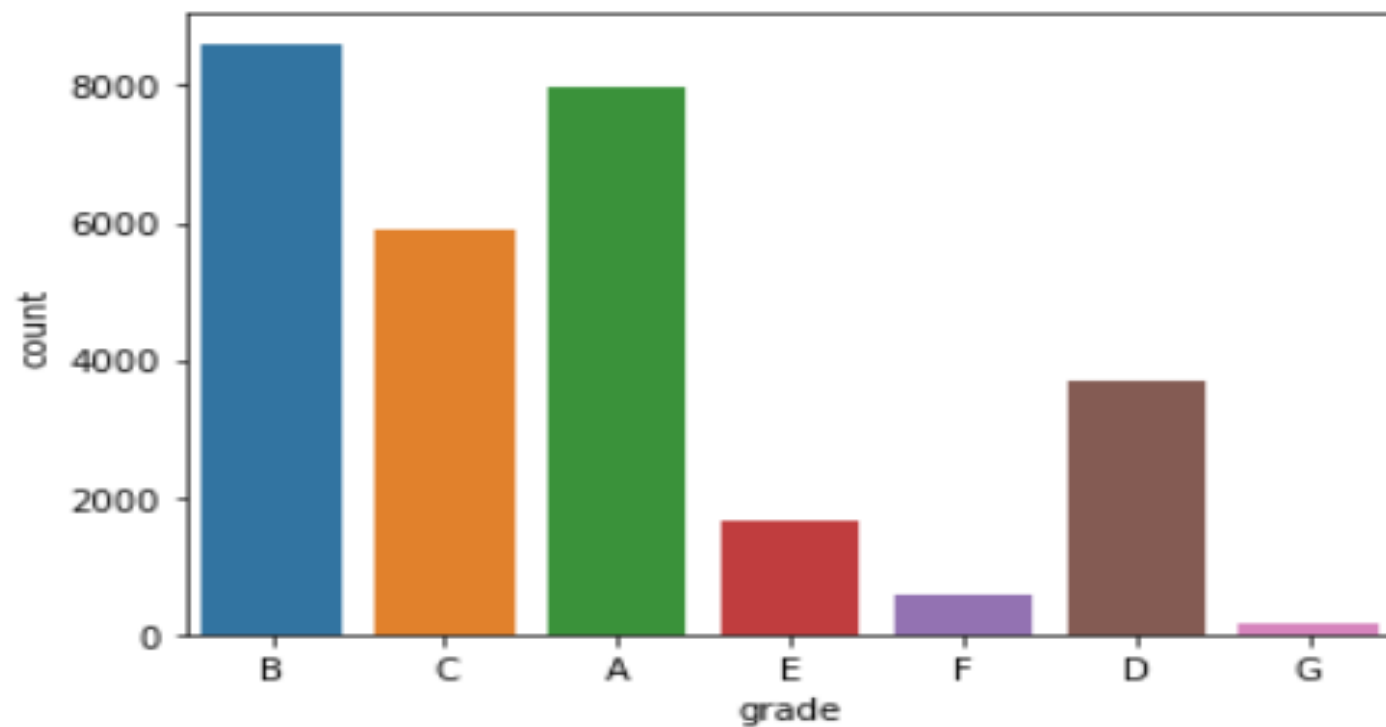
Analysis

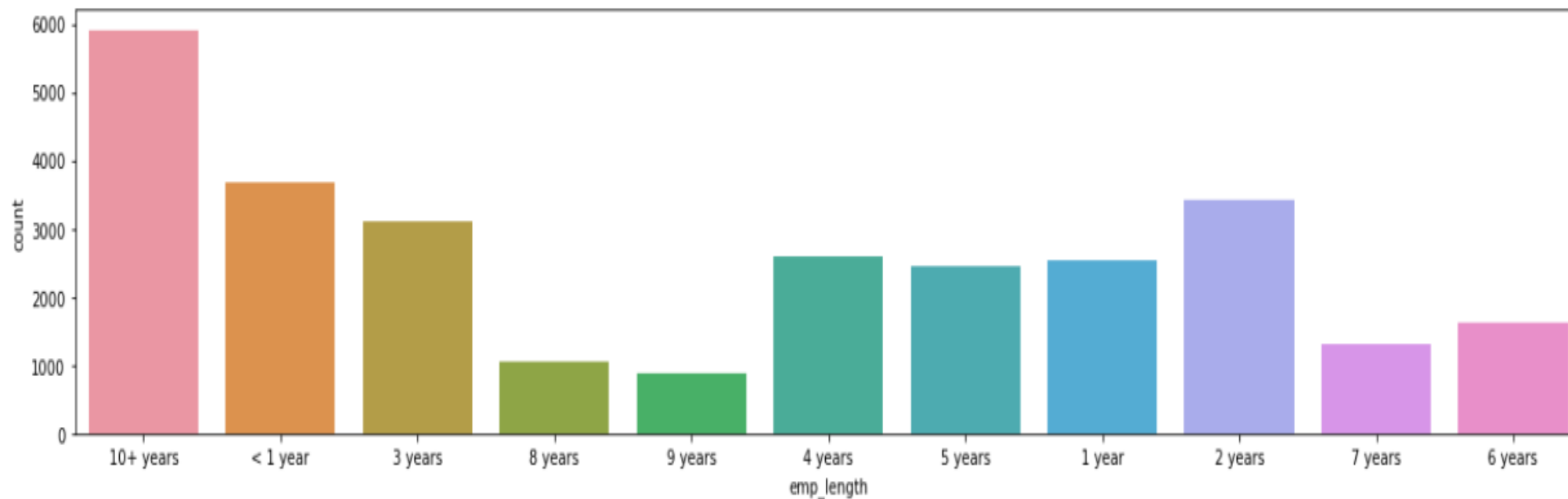Mostly interest rate is between 10% to 15%. But there is also spike between 6% to 8% interest rate
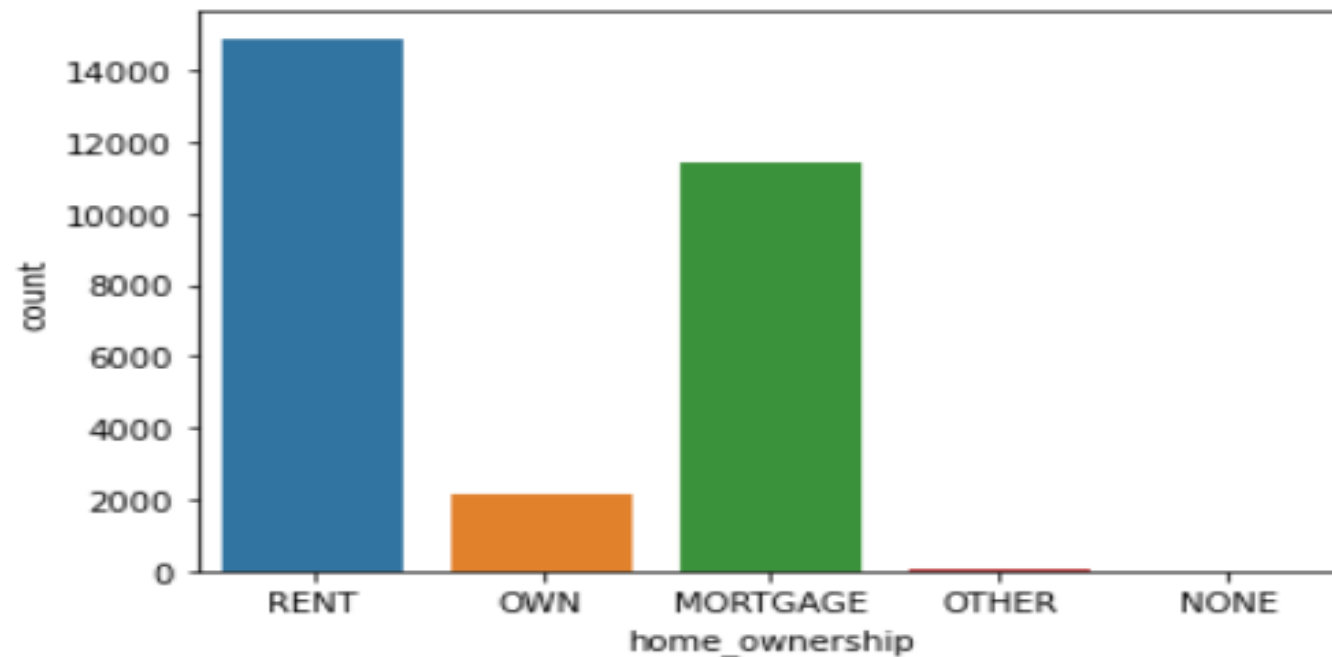
Analysis

Most installments are between 100 and 300

Analysis

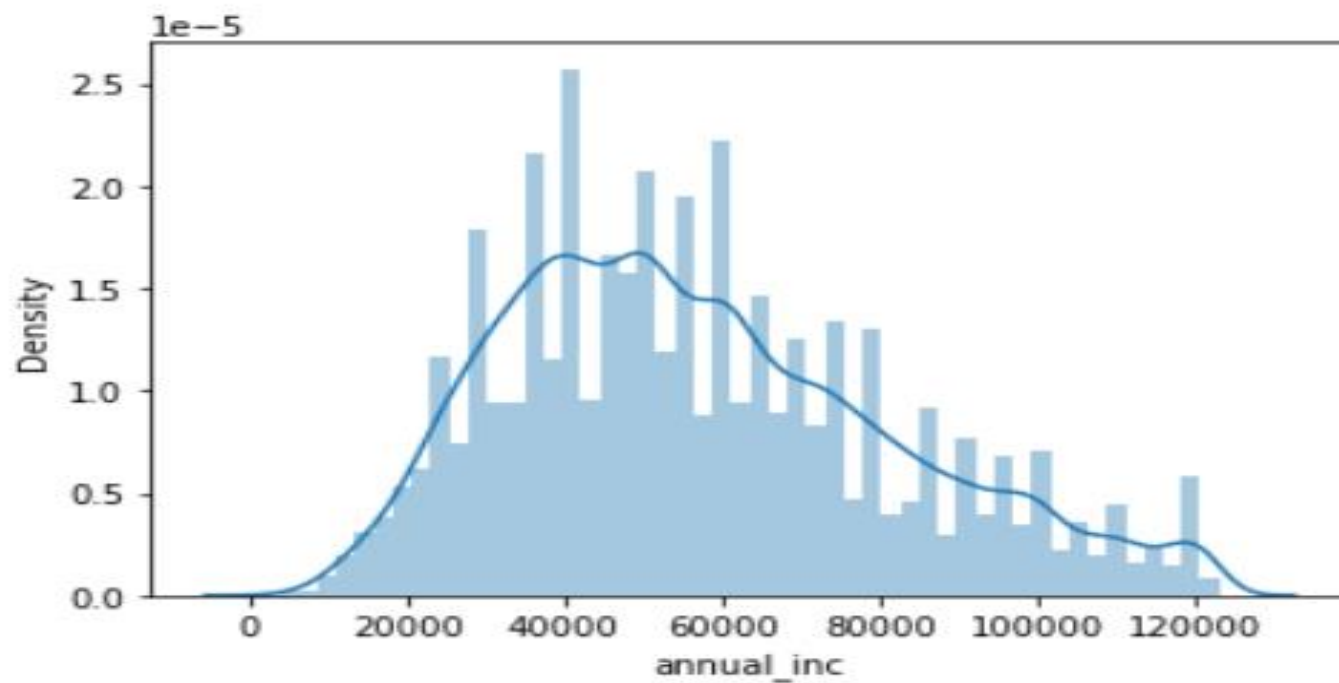Most of the borrower have grade B and A

Analysis

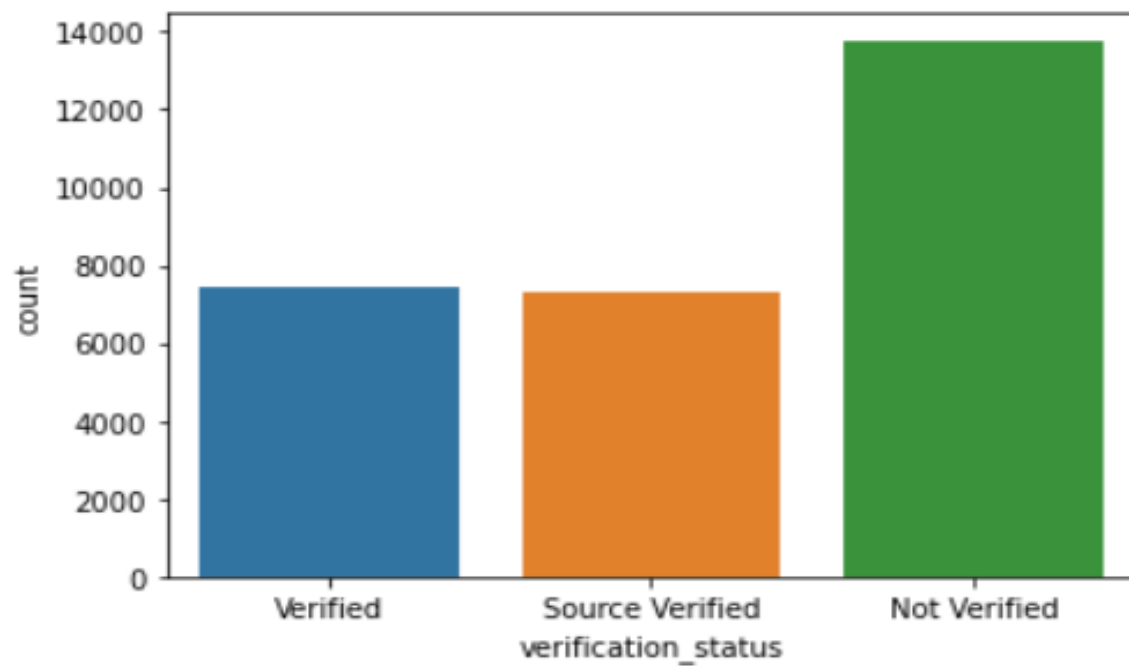Most of the borrowe have more than 10 years of experience

Analysis

Mostly the people who are in rent or mortage require loan than others
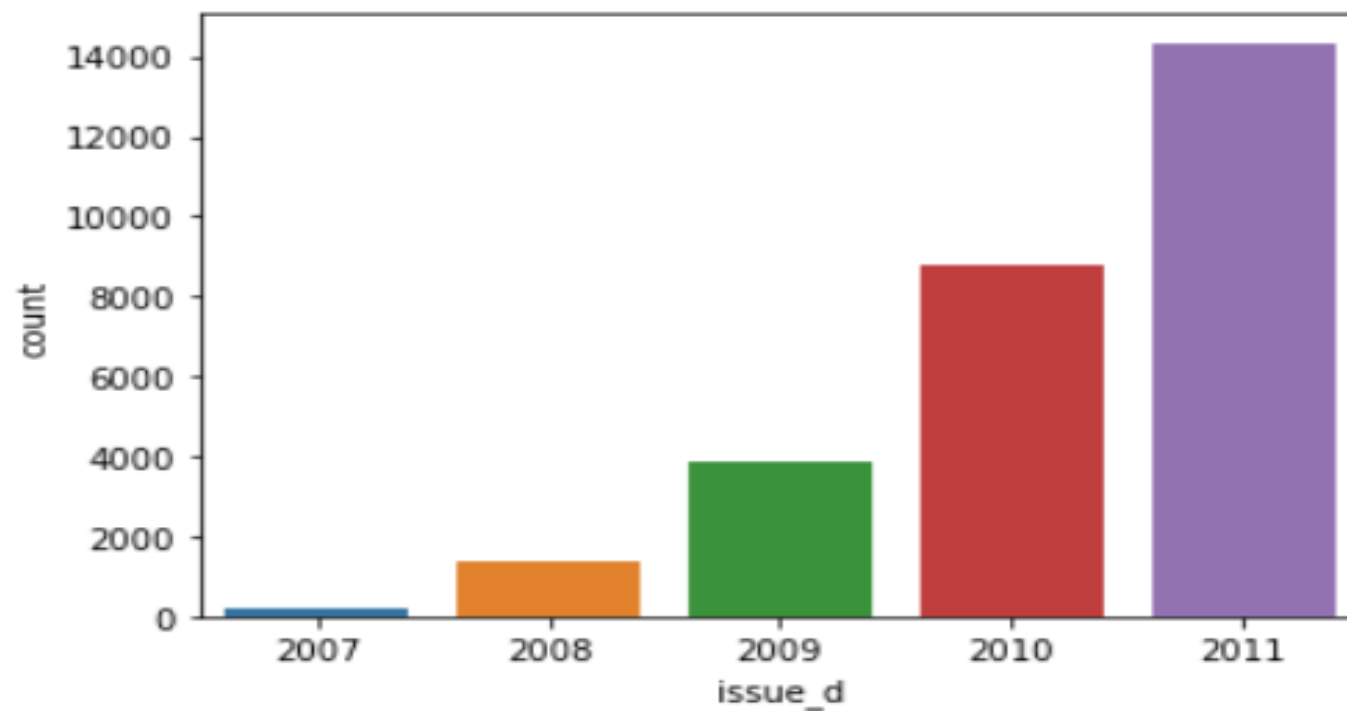
Analysis

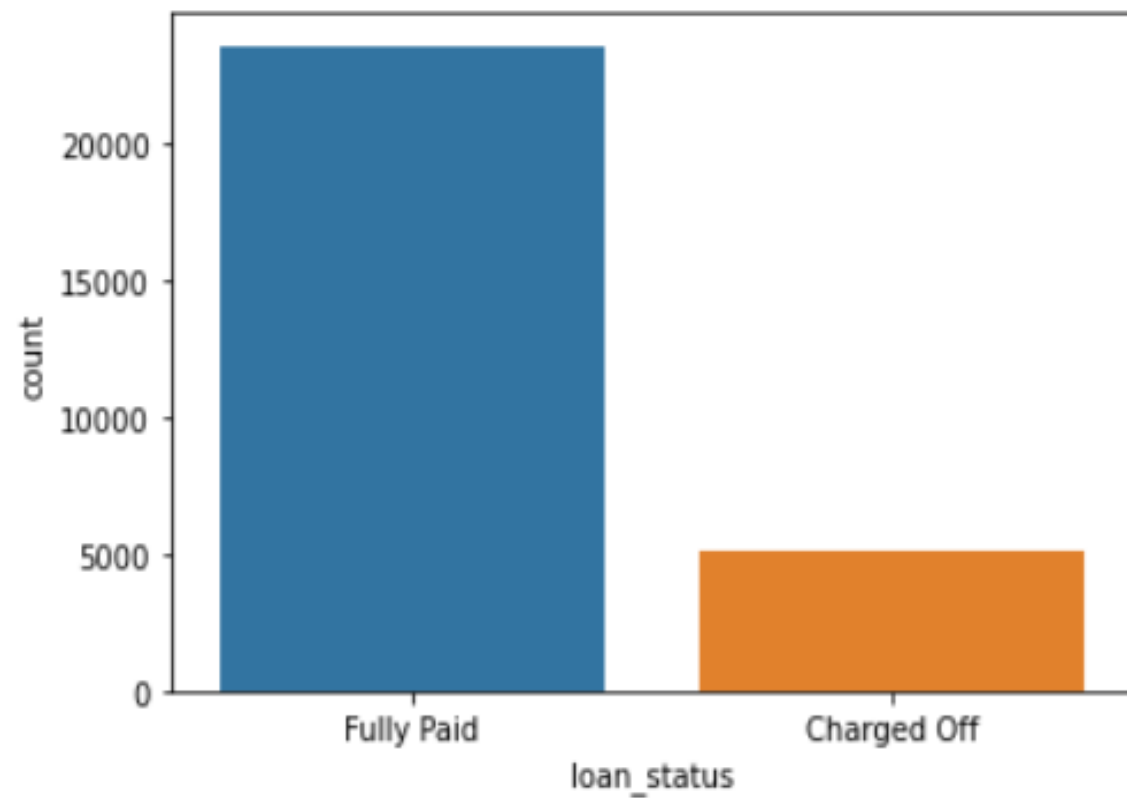People whose salary between 30000 and 60000 take loans

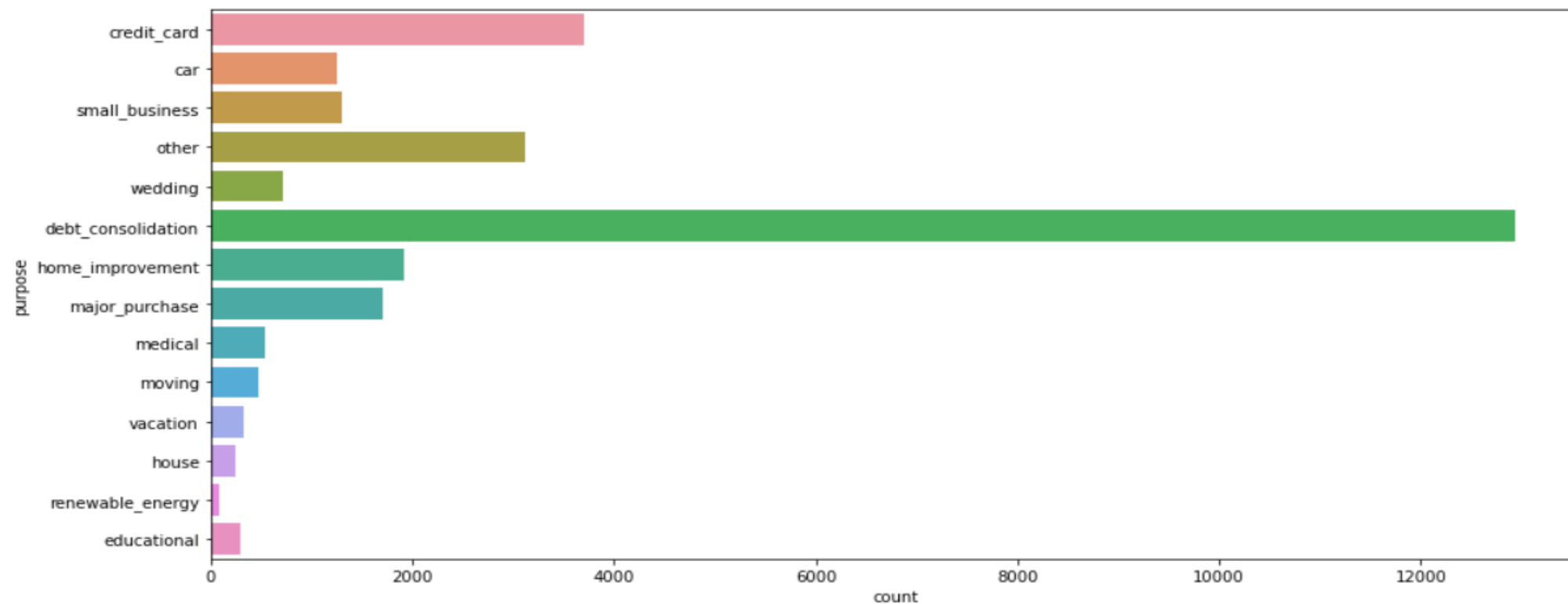# Analysis

Most of the borrowers are not verified

Analysis

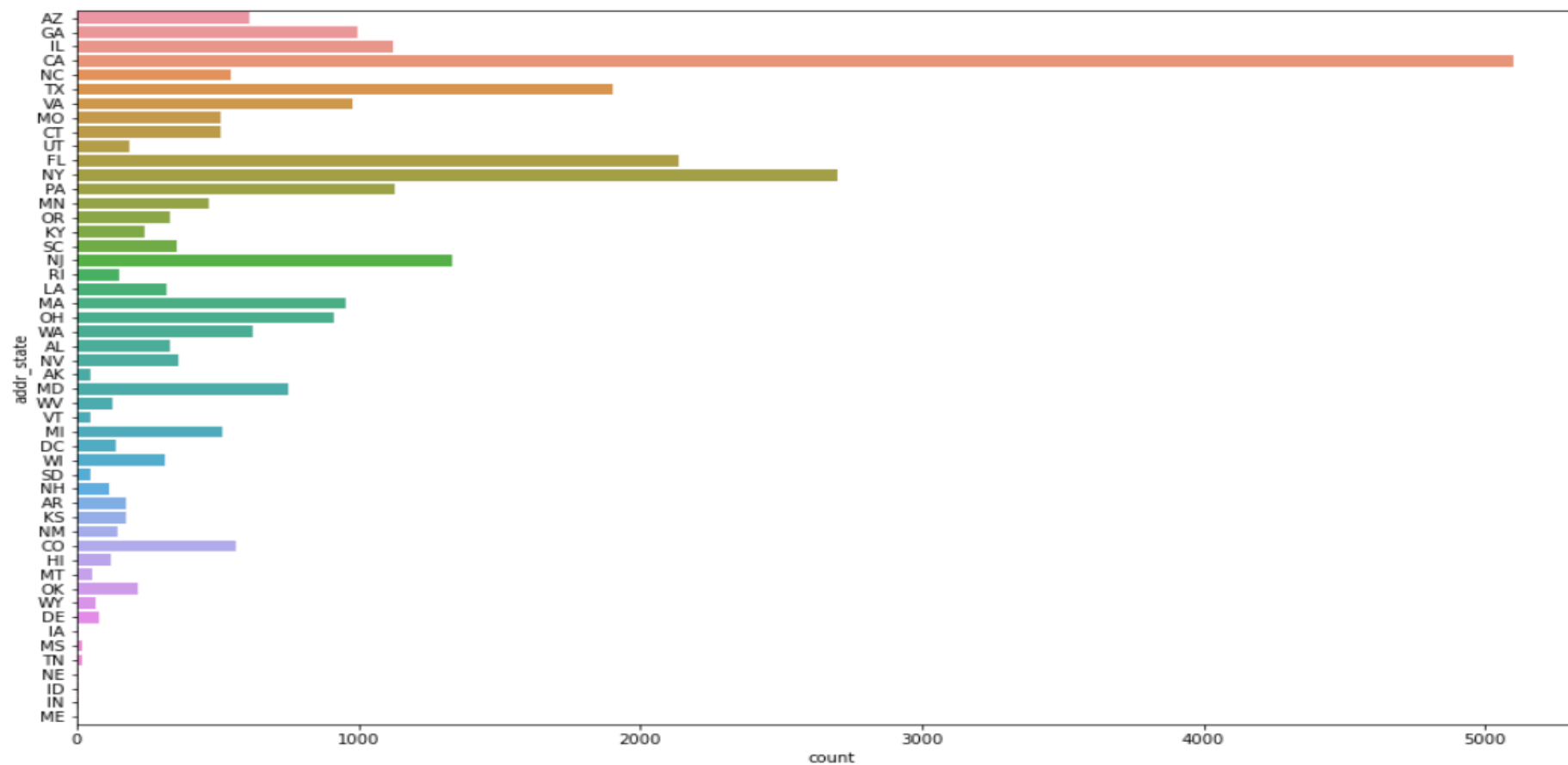Issuing loans are increased by 2x every year
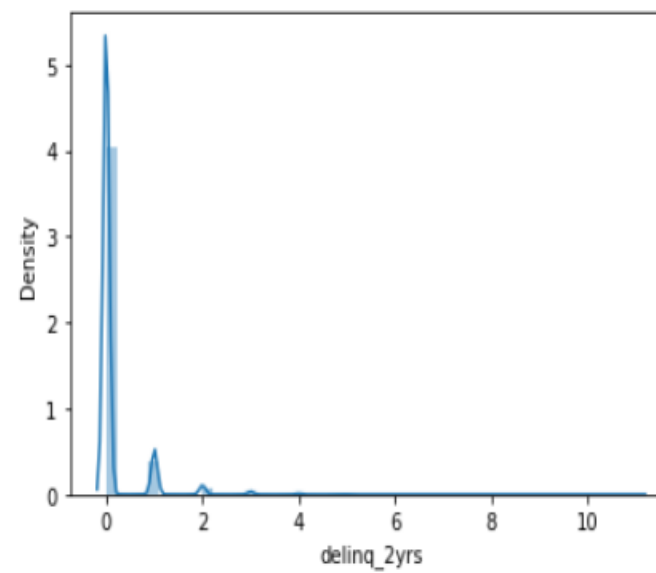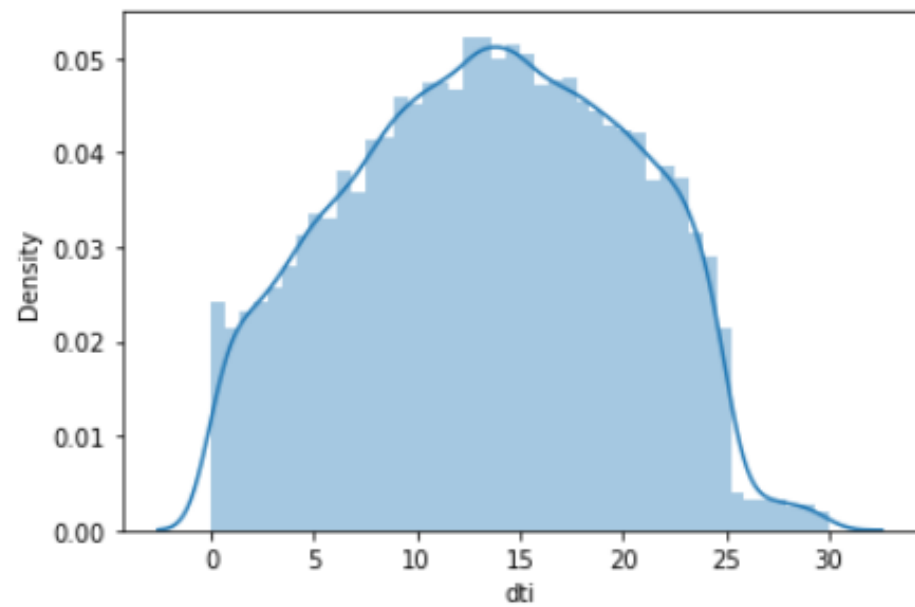
Analysis

Analysis

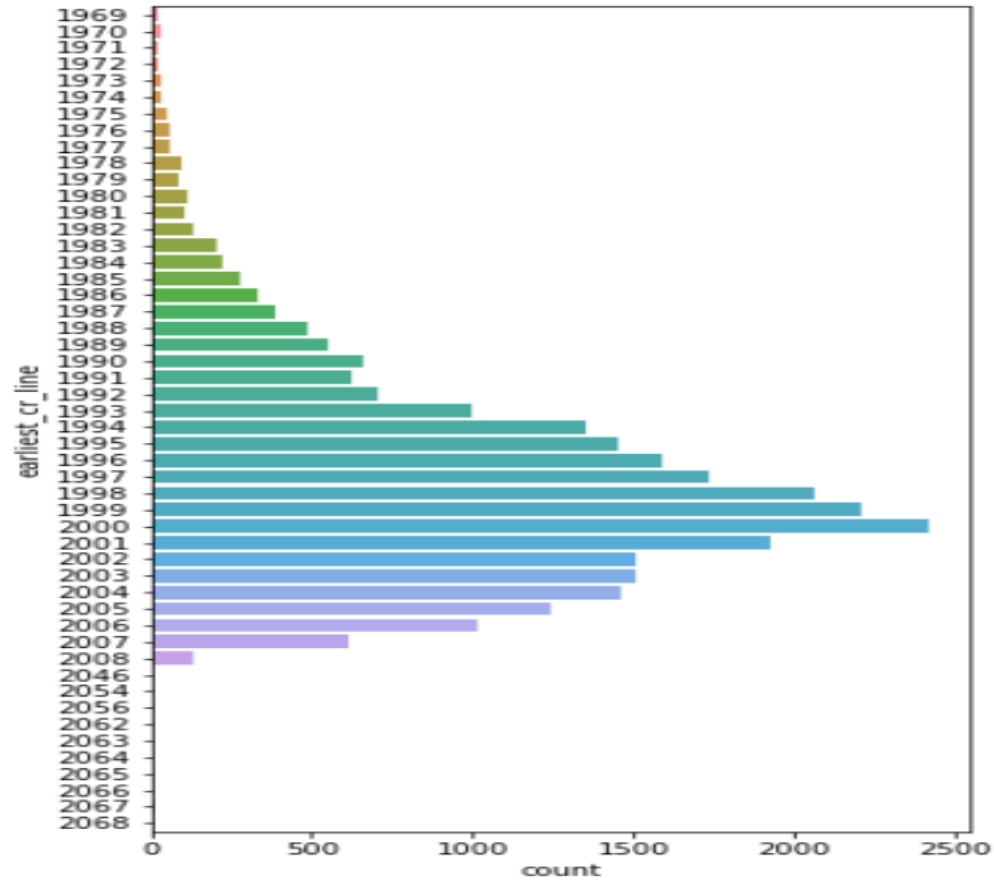Mostly people take loan for debt consolidation, credit card or for other usage

Analysis

Most of the borrowers are from Callifornia, New York, Florida and Texas
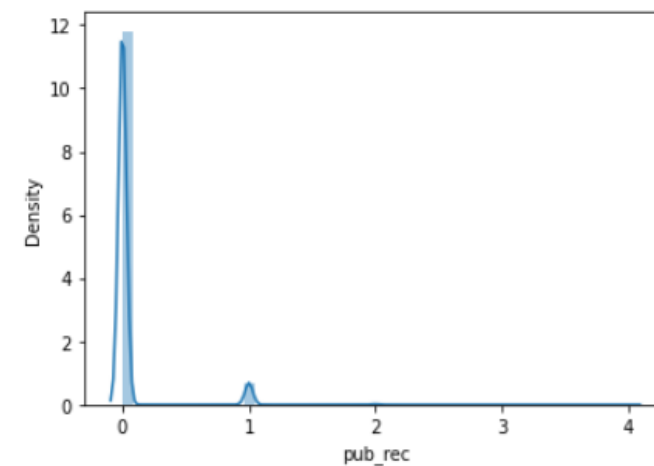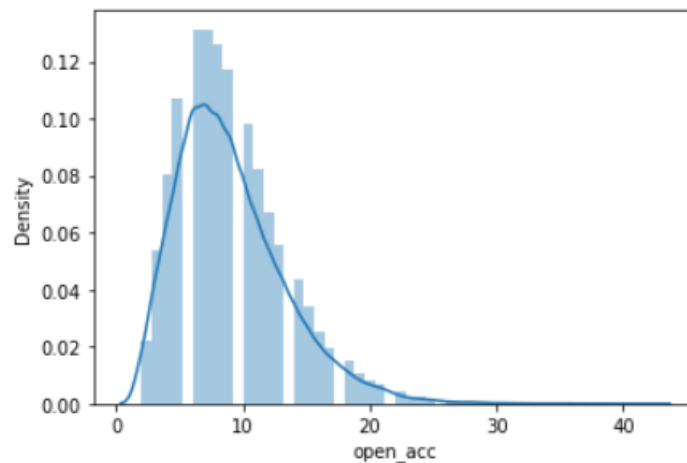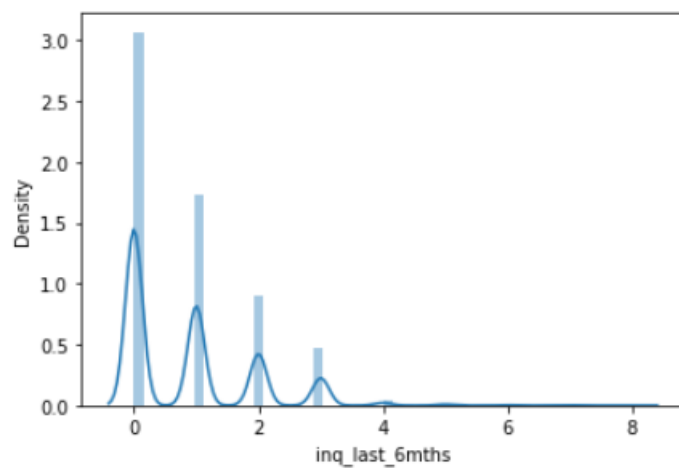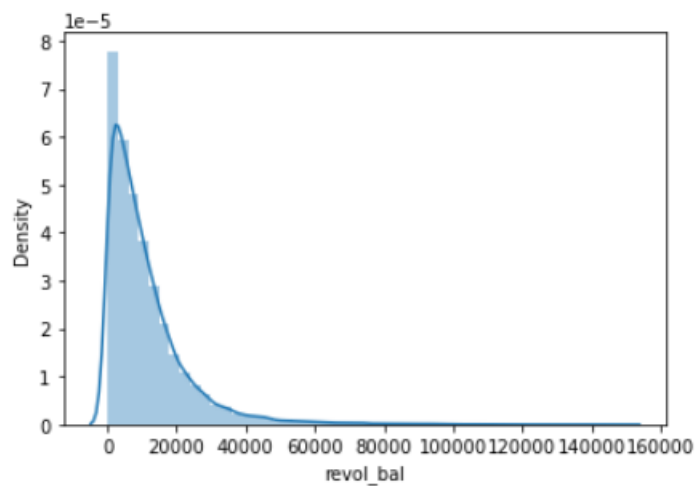
Analysis

Analysis

The opening of credit line increased till the year of 2000 and then it started to decrease
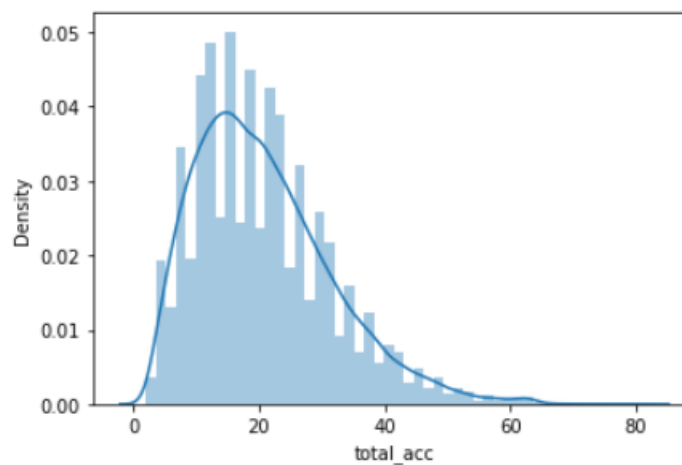
Analysis

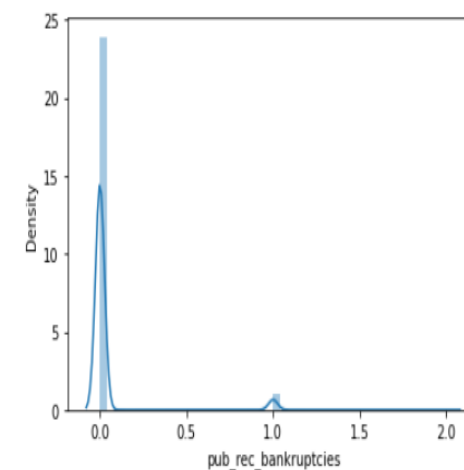Mostly borrowers have 5 to 10 open credit lines

Analysis

Mostly revolving balances are below 20000

Most of the customers have 10 to 25 credit lines

Analysis
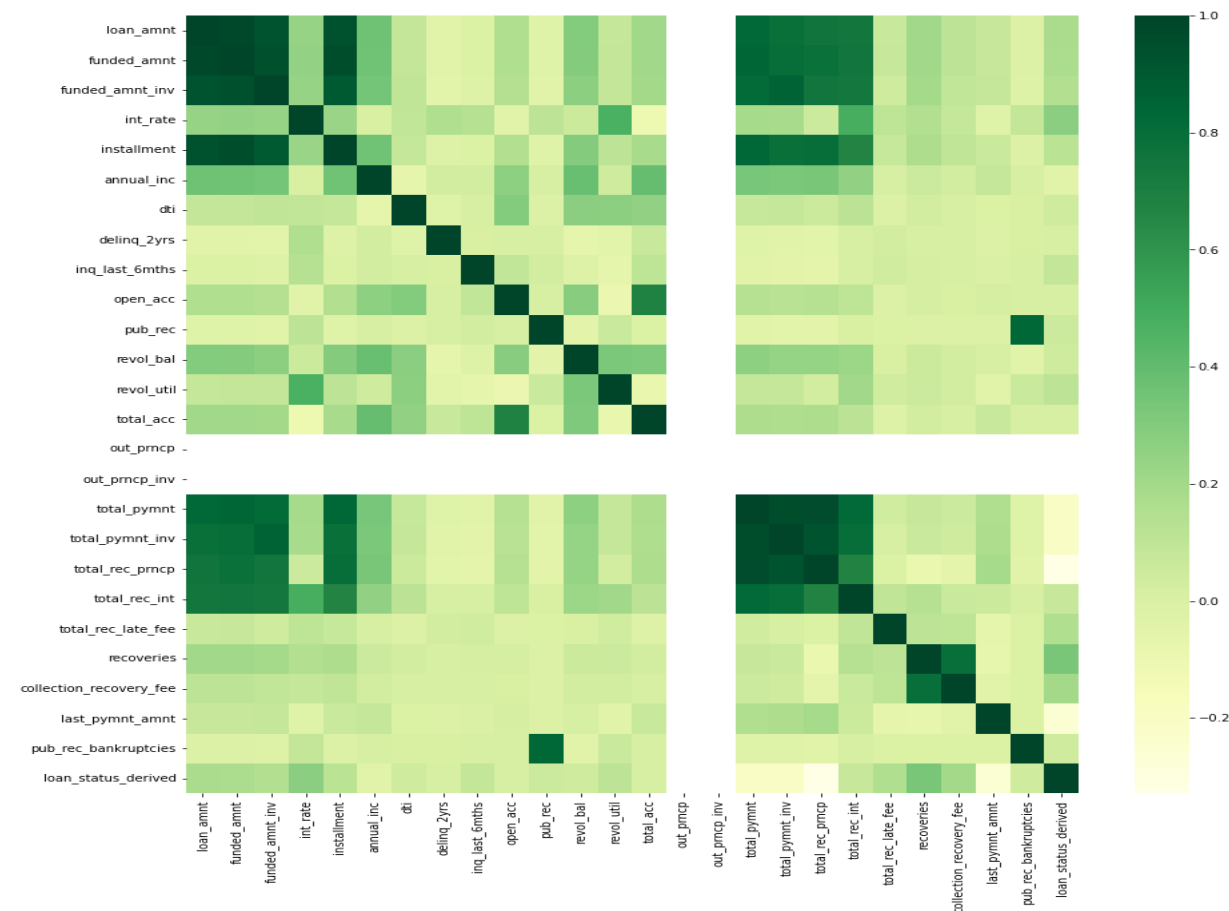
Analysis
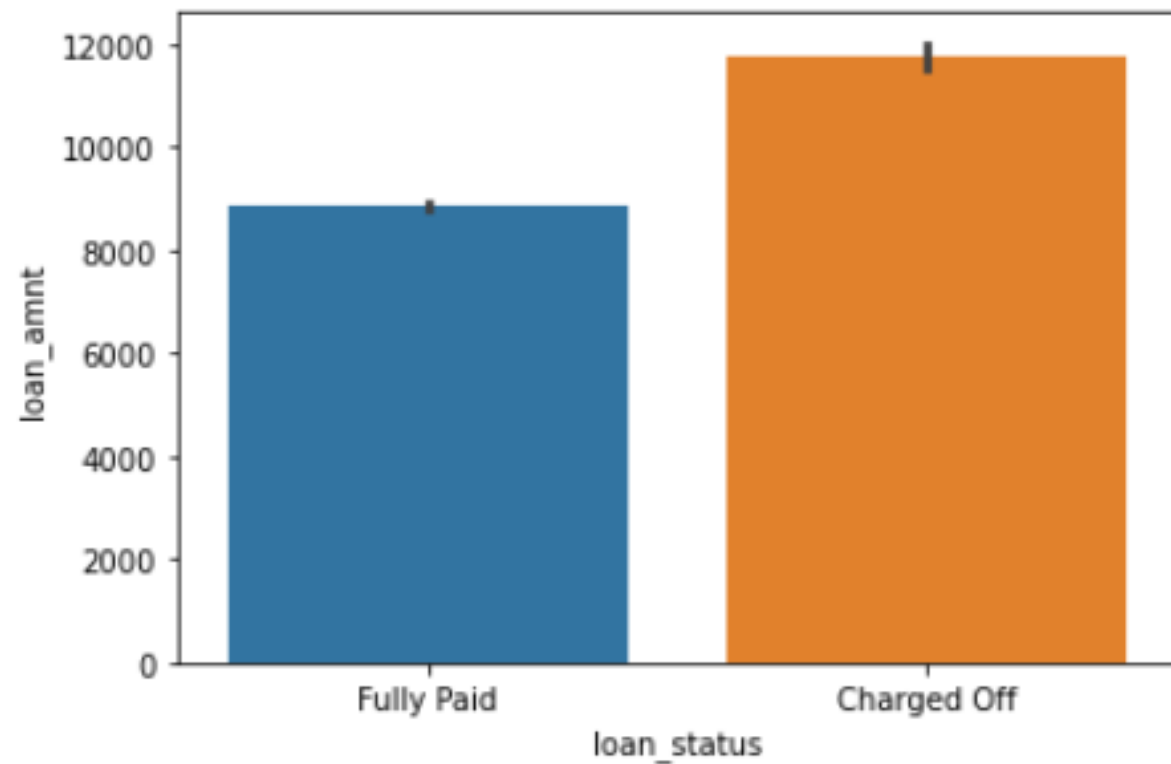
We can see non default customers have comparatively low loan amount than defaulters

Analysis
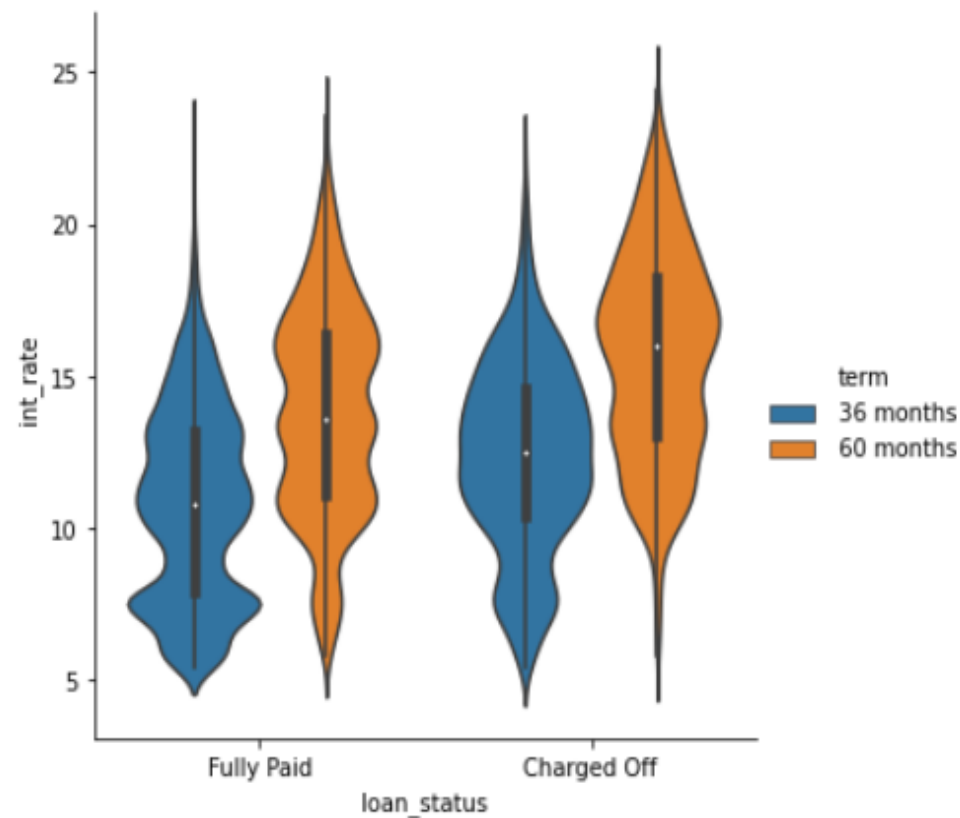
It seems who have taken loan for 36 months with 13% interest (apprx.) and 60 months above 15% interest have higher chance of being default

Analysis
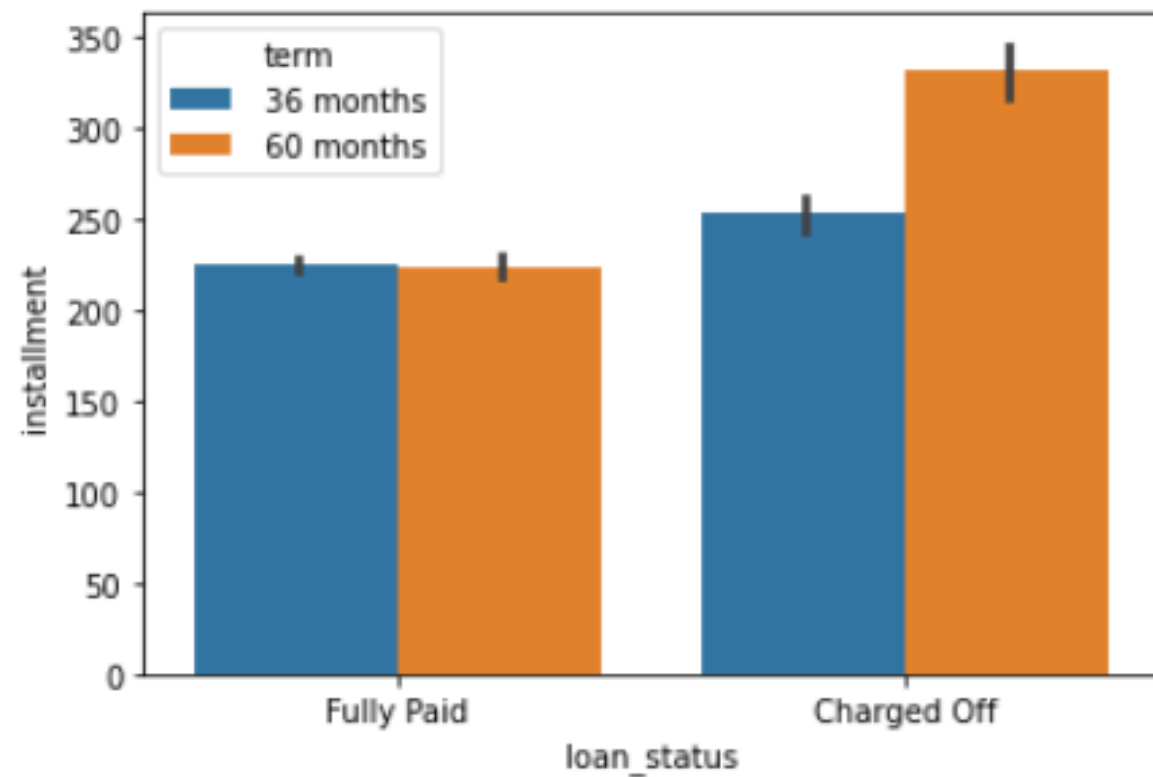
Defaulters have more than avg 200 dollar monthly installments

Analysis
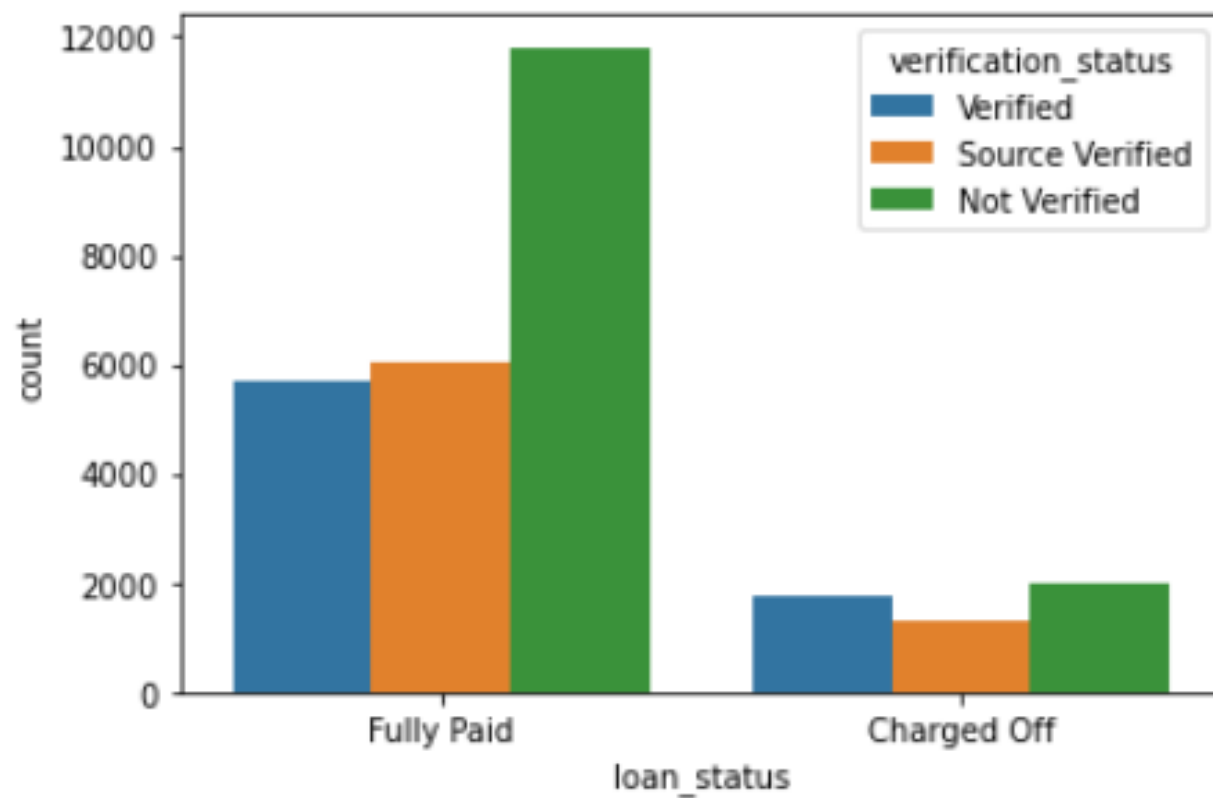
Most of defaulters are not verified

Analysis

Increase in no of enquiries can lead to default

Analysis
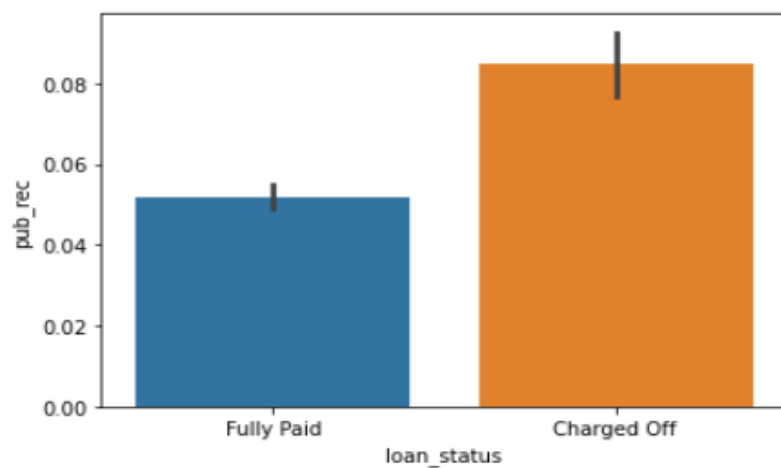
Dafaulters have higher number of derogatory public records

It's obvious that dafaulters will have higher revolving balance

Analysis
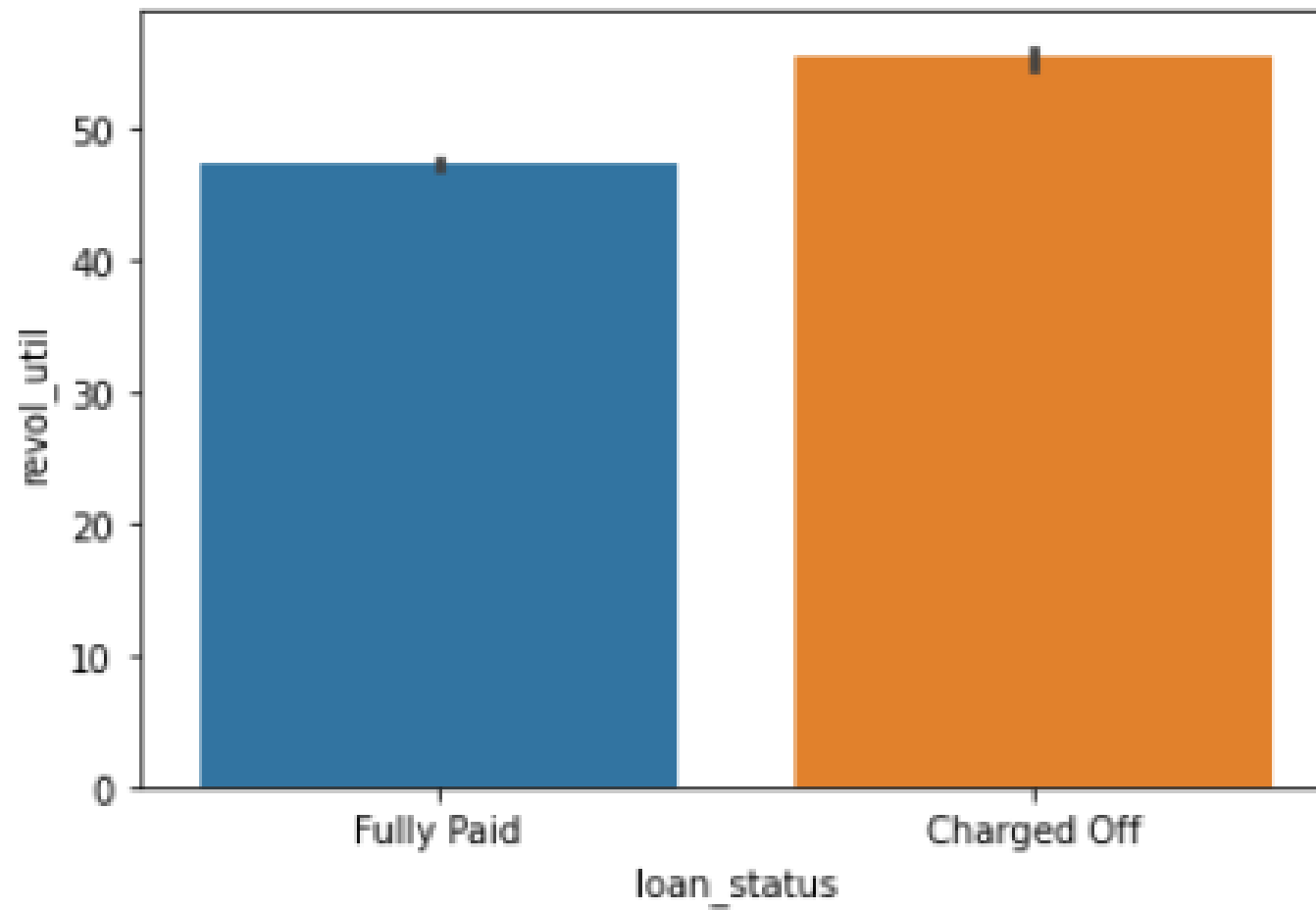
Analysis
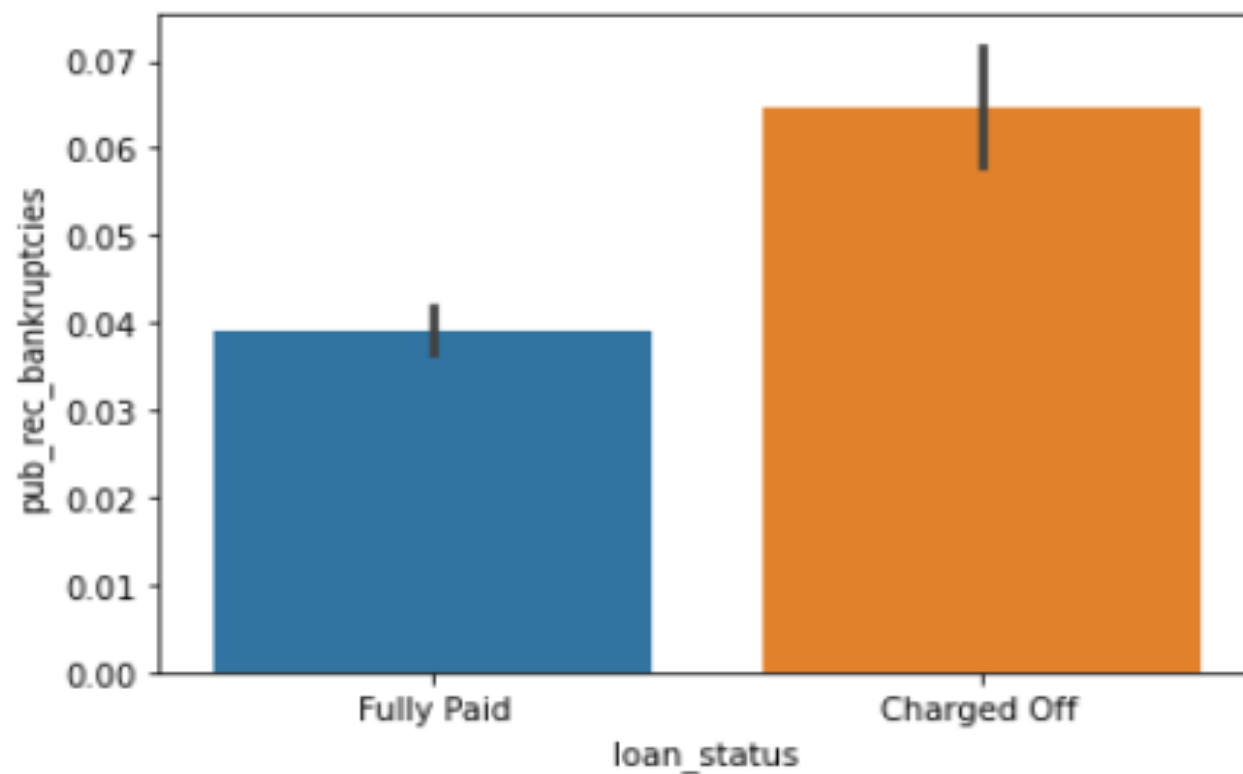
Dafaulters have the most bank rupts files

Recommendations

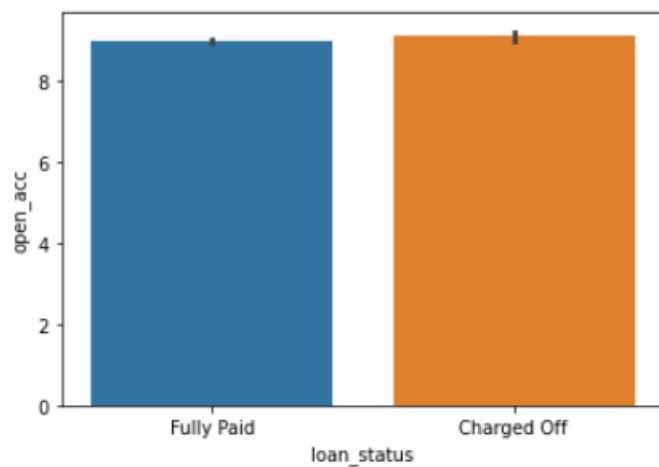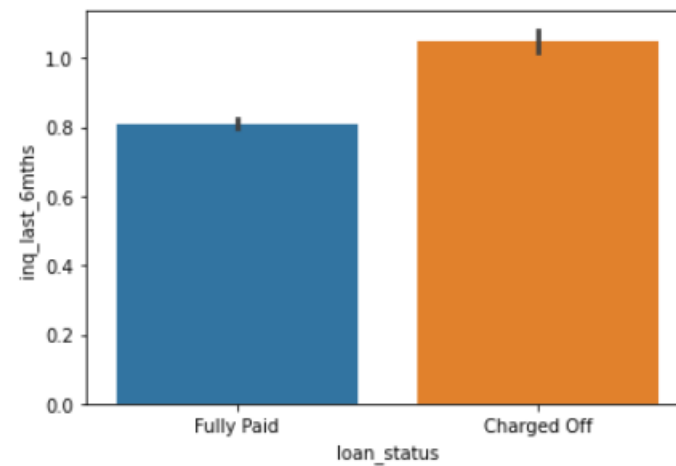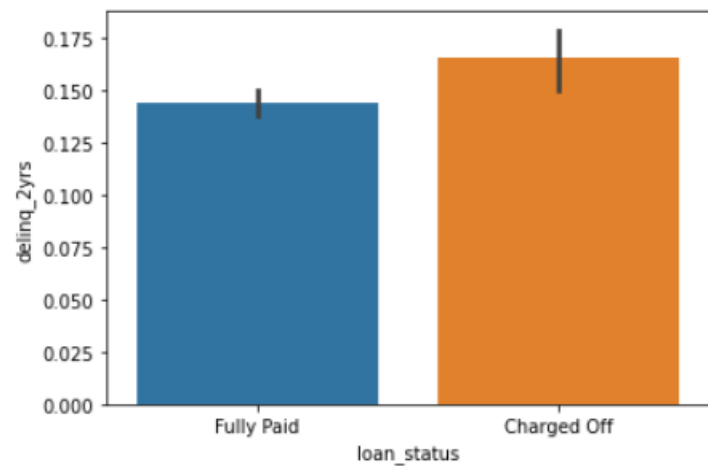Hereby we come to an end of the EDA of the loan data set and finding some of the drivers for loan default. Apart from the ones highlighted below, I am sure there will be multiple others too; however, according to me, these are the most impactful ones.

Minor Impact
- Higher loan amount (above 16K)
- Higher installment amount (above 327)
- Lower annual income (below 37K)
- Higher debt to income ratio (above 15%)
- Applicant's address state (NV, SD, AK, FL, etc.)
- Loan issue month (Dec, May, Sep)

Heavy impact

- Higher interest rate (above 13%)
- Higher revolving line utilization rate (above 58%)
- Repayment term (5 years)
- Loan grade & sub-grade (D to G)
- Missing employment record
- Loan purpose (small business, renewable energy, educational)
- Derogatory public records (1 or 2)
- Public bankruptcy records (1 or 2)

Combined impact

- High loan amount & interest rate for lower income group
- High installment and longer repayment term
- Home ownership (other) and loan purpose (car, moving or small business)
- Residential state and loan purpose
- Income group and loan purpose

Conclusion

- Lending Club should reduce the high interest loans for 60 months tenure, they are prone to loan default
- Grades are good metric for defecting faulters. 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 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 no public derogatory records for borrowers