

Lending Club Case Study

Gopal Pawar

Haribabu Aryasomayajula



Business Understanding

The objective of the case study is to analyze a real world scenario (Approval of Loan) using EDA techniques and provide recommendations to business stakeholders to minimize the risk of default.

Benefits

- ❖ Use EDA techniques to come up with data analysis and preparation
- ❖ Understand Risk Analytics in Loan approval process
- ❖ Provide a visual representation of data to arrive at useful insights
- ❖ Minimize the risk of default / losses by identifying key impacting variables

Solution Approach

- ❖ Data Pre Processing
- ❖ Univariate Analysis
- ❖ Bivariate Analysis
- ❖ Multivariate Analysis
- ❖ Outcomes / Recommendations

Data Pre Processing

❖ Focused on column

- Finding columns that value percentage of missing values and drop the column that has missing value more than 30%
- Column that has all unique like Id etc. values are not so important for analysis so dropping it

❖ Focused on rows

- Removing rows that has NA value

❖ Focusing on Data Conversion

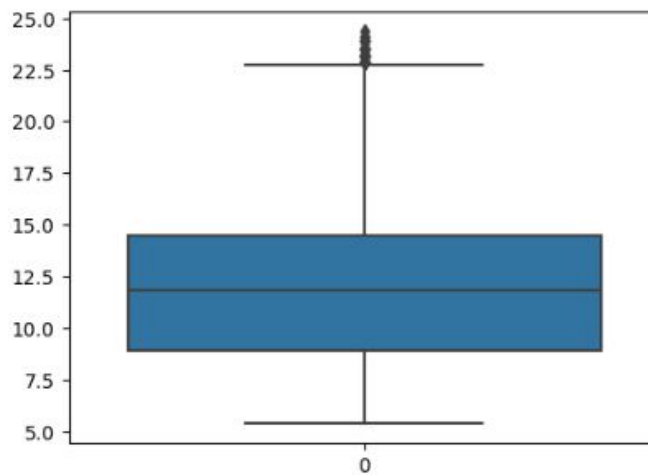
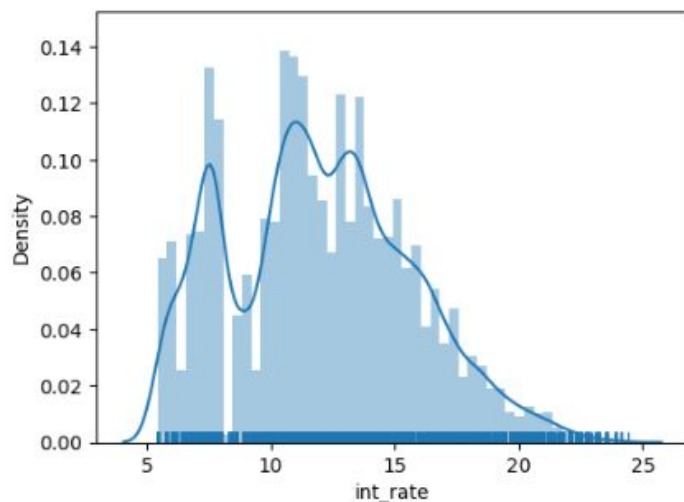
- Converting the column with data like 10 % or 36 months to 10 and 36
- Creating derived variables such as month, date and year from issue_d column

Univariate Analysis

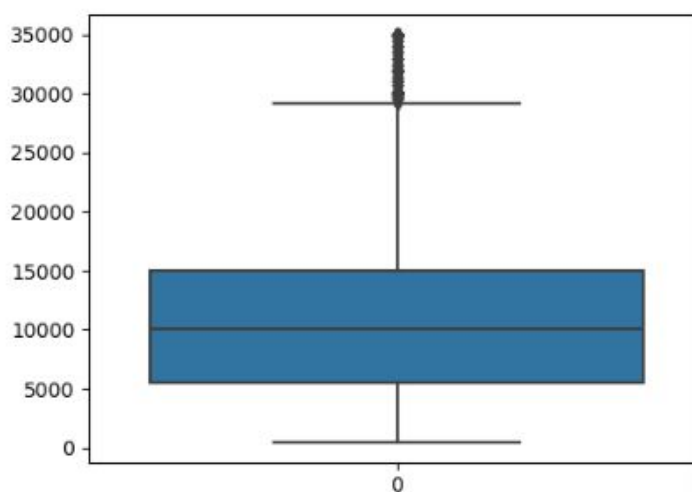
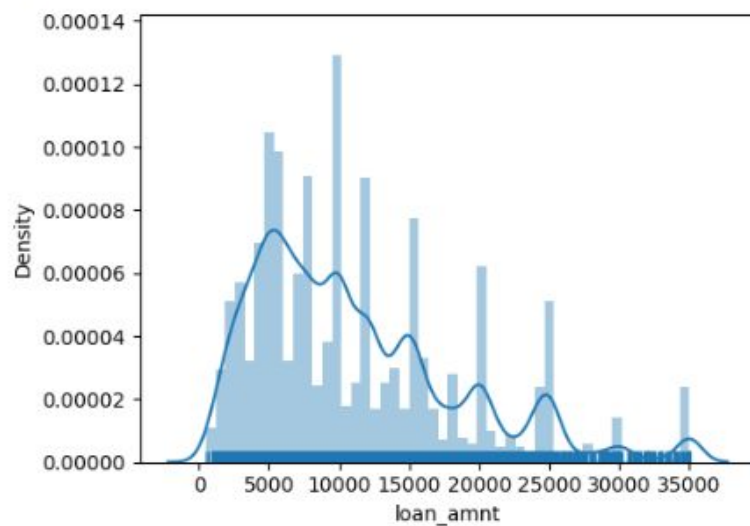
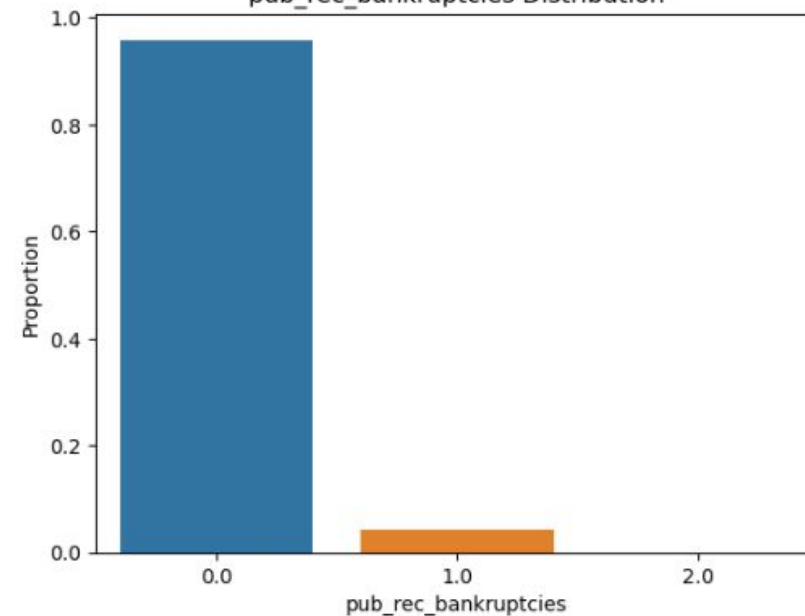
- ❖ By doing some univariate analysis on Loan amount, interest rate, funded amount
 - From the above plots, more number of people took loan amount of 10000, and also median of distribution is 10000. And very few people took more than 30000 loan amount.
 - Funded amount investment data behaves similar to loan Amount, Which tells us that Lending club approved atmost of Applied loan amount.
 - 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%**
 - Around 361 people are having more the 235000 annual income. and 99% of people are having incomes below 235000.
 - Around 90% borrower's are having no public derogatory records.
 - 99% people have not went bankrupt.

Univariate Analysis

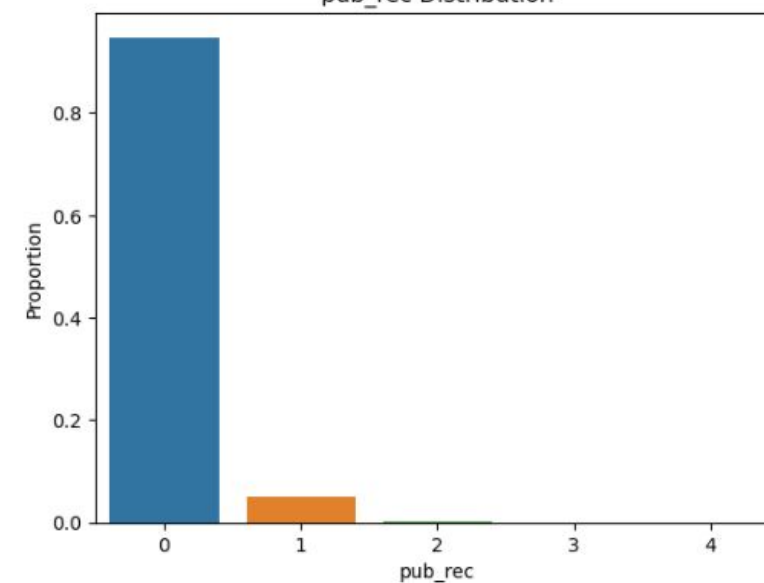
interest rate distribution



pub_rec_bankruptcies Distribution

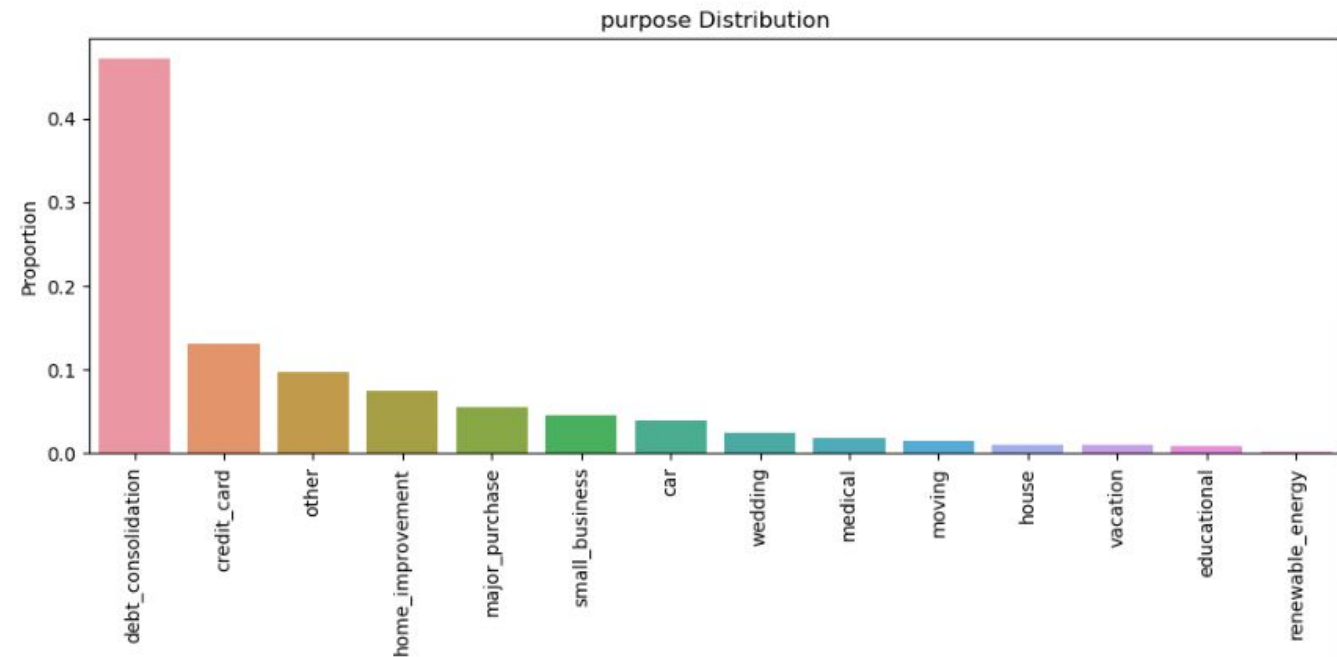
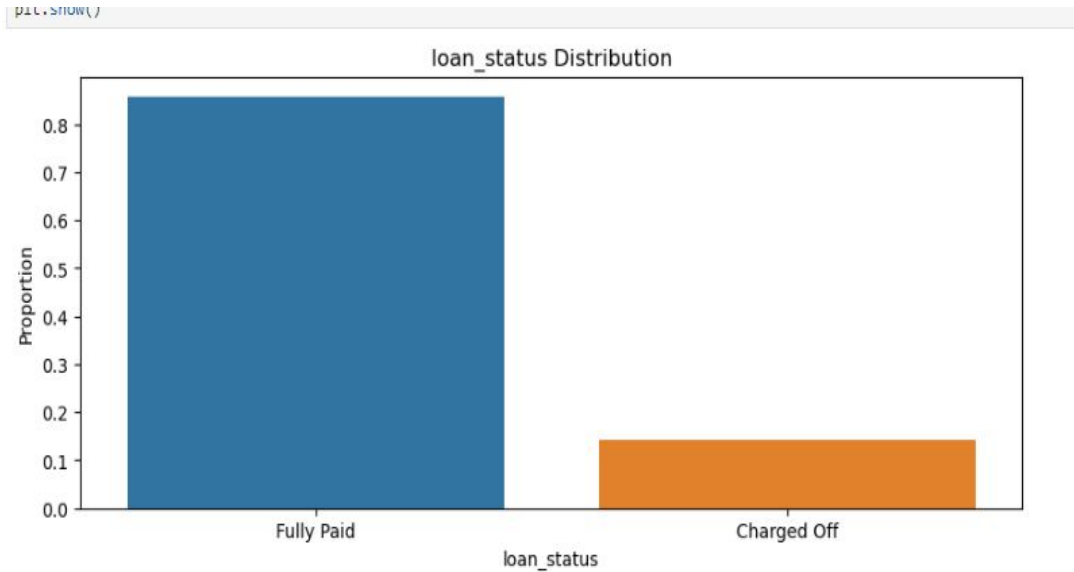


pub_rec Distribution



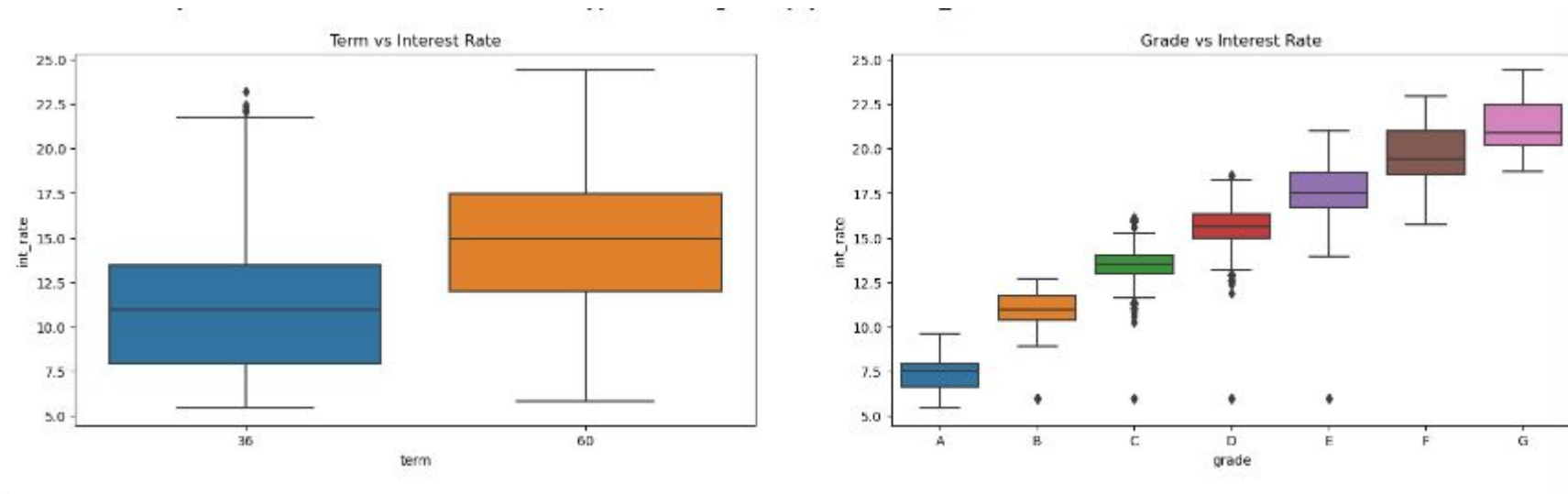
Univariate Analysis

- ❖ 85% of borrowers has paid the loan fully whereas 14% are defaulted the loan.
- ❖ Looks like more number of people took loan for *debt consolidation* and a very few people took for *renewable energy*



Segmented Univariate Analysis

- The interest rates are higher for Higher tenure loans.
- And Also Interest Rates are Higher as Grades are Lowering (A to G).



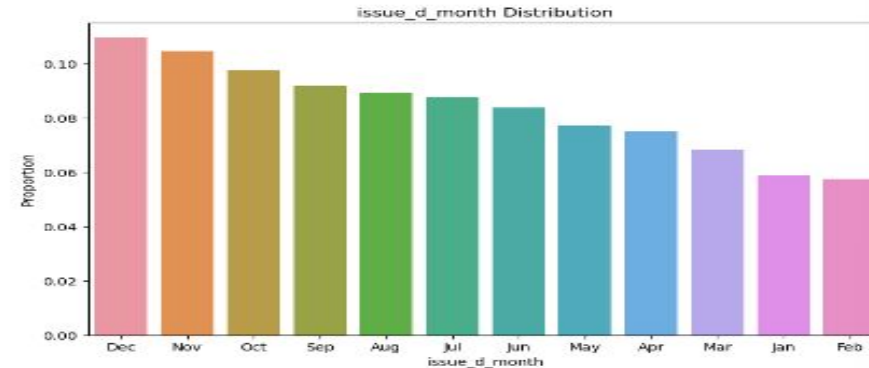
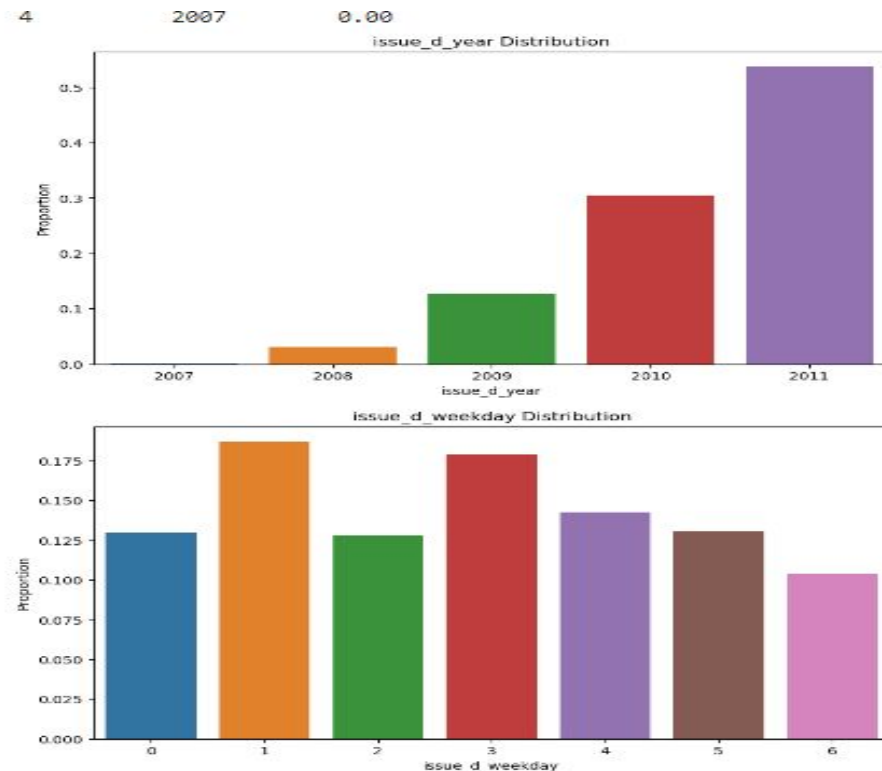
Few Observations with Bar Plot

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

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

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

Borrowers have taken more loans with 36 months tenure as compared to 60 months.



Bivariate Analysis

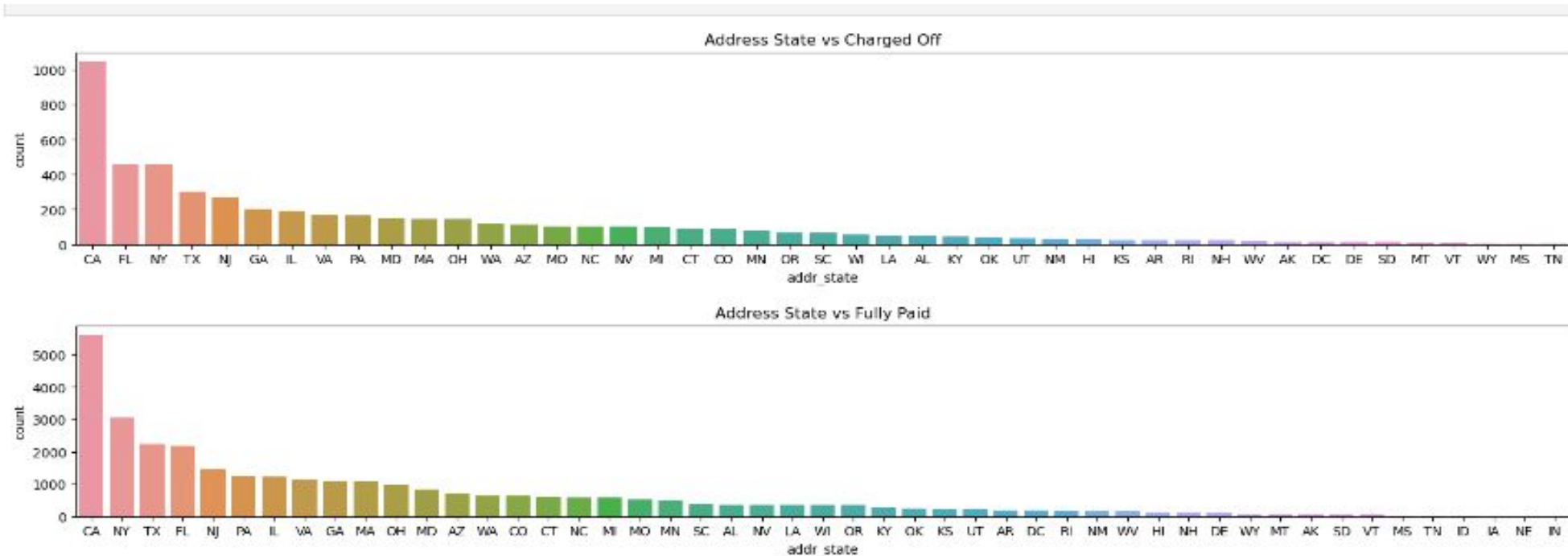
Looks like there are more proportion of borrowers defaulted loan in 60 months term then 36 months. Also the Fully Paid rate is higher in 36 months tenure.

Loan amount is not a decider for defaults in both 36 and 60 months.

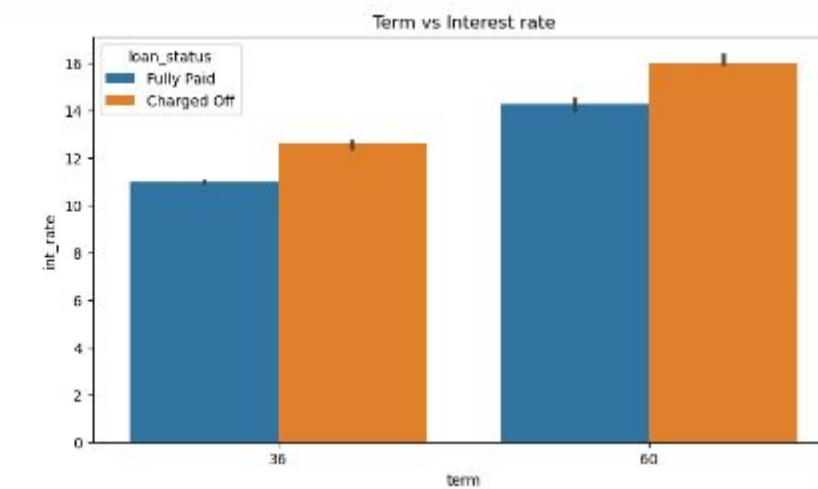
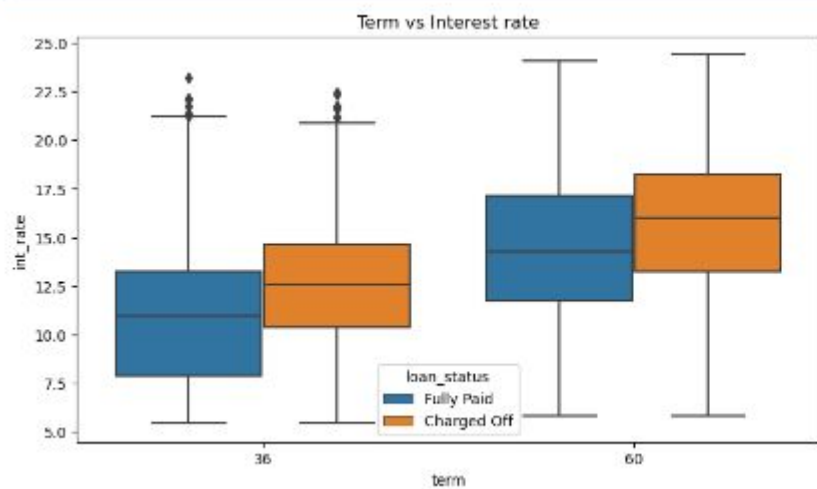
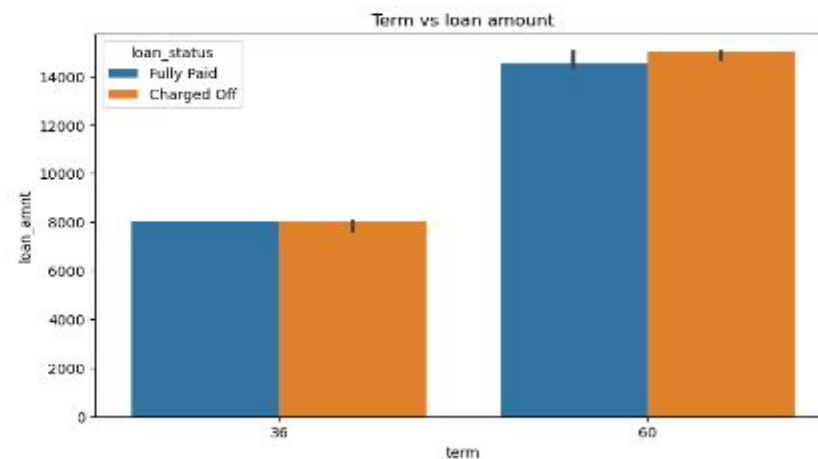
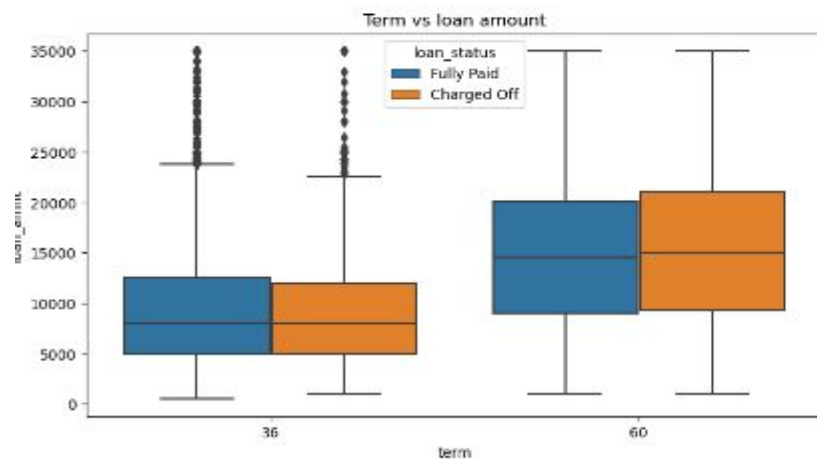
Borrowers have equal distribution is both default and non default for 36 and 60 months tenures.

For higher interest rates the default rate is higher in both 36 and 60 months tenure.

More number of borrowers defaulted in *CA* , *FL* and *NY* states.

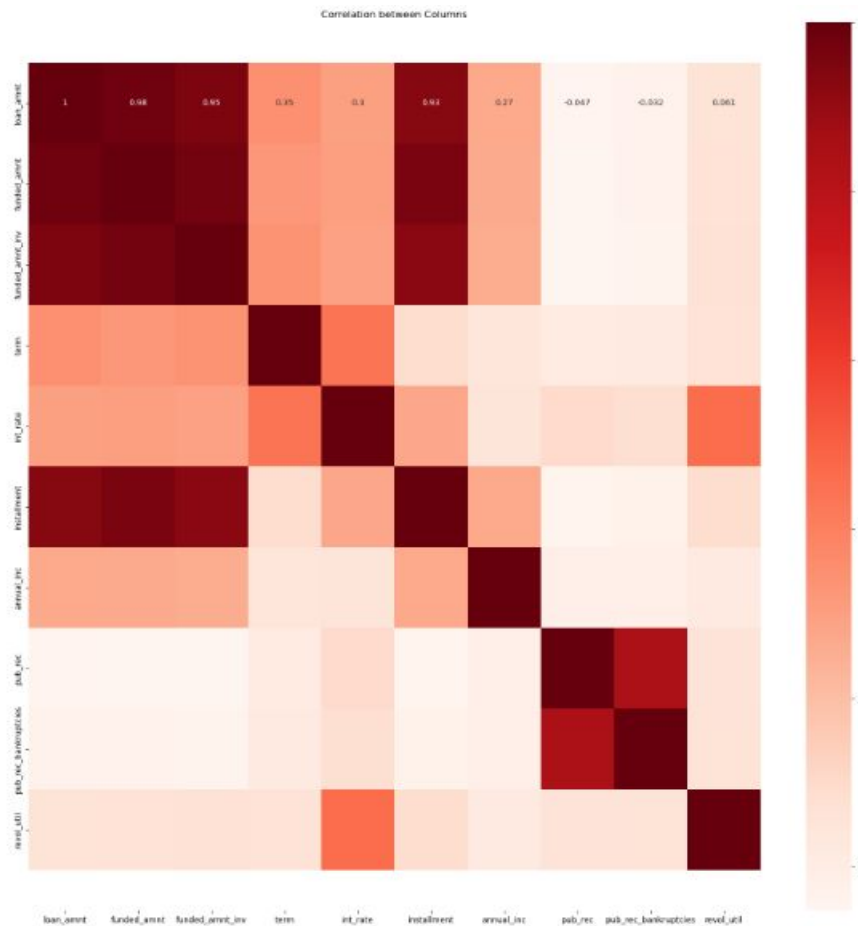


Bivariate Analysis



Multivariate Analysis

- * The public derogatory records column is highly (+ve) correlated with public bankruptcies records.
- * Interest rates are high for people with high revol utilisation.



Conclusion

- Key variables to consider while approving the loan are Loan purpose, Grade, Loan tenure, Employment tenure, Funded Amount, Interest Rates, Loan Issue Month, Subgrade and Home Ownership
- Minimize the loans for setting up small business for the purpose of debt consolidations and CC settlements
- Loans with very high interest rates and high funded amounts are likely to get charged off. Reducing the interest rate or not approving the loans for such scenarios is recommended.
- Scrutinize the loan applicants who have Other/Rented houses
- Number of loan approvals is increasing each year indication a healthy growth rate for the lending club