Lending Club Case Study

PRESENTED BY :

K DILEEP KUMAR REDDY

PRIYANKA PARAB

Agenda

I. Problem Statement
II. Methodology
III. Univariate Analysis
IV. Segmented Univariate Analysis
V. Bivariate Analysis
VI. Conclusions

Problem Statement

A company 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. Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'. If one can identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicant's using EDA is the aim of this case study. In other words, the company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

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

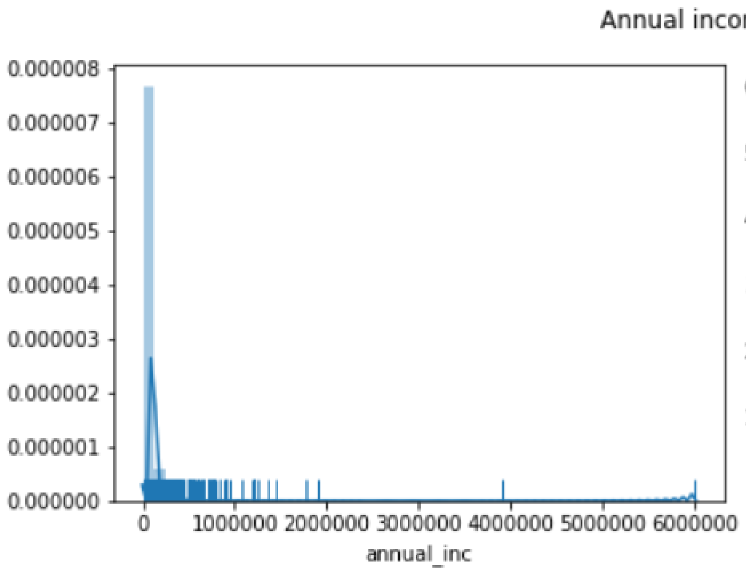
Bivariate Analysis

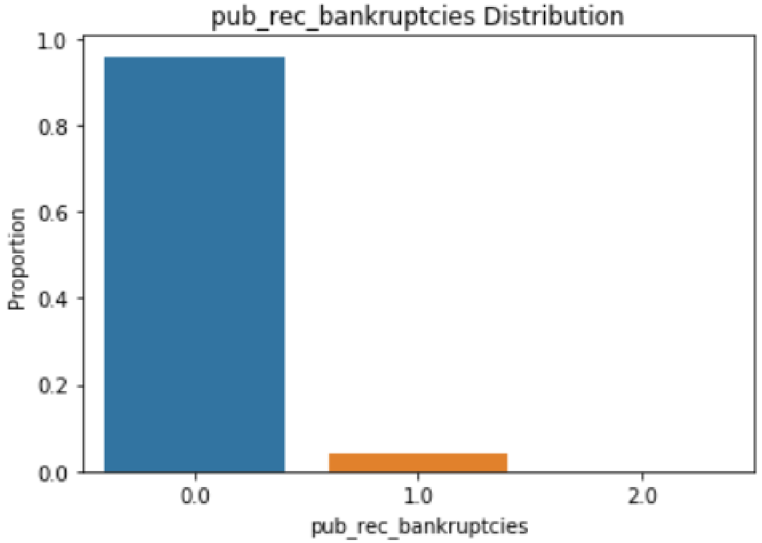
Analysing the two variable behaviour like term and loan status with respect to loan amount.

Recommendations

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

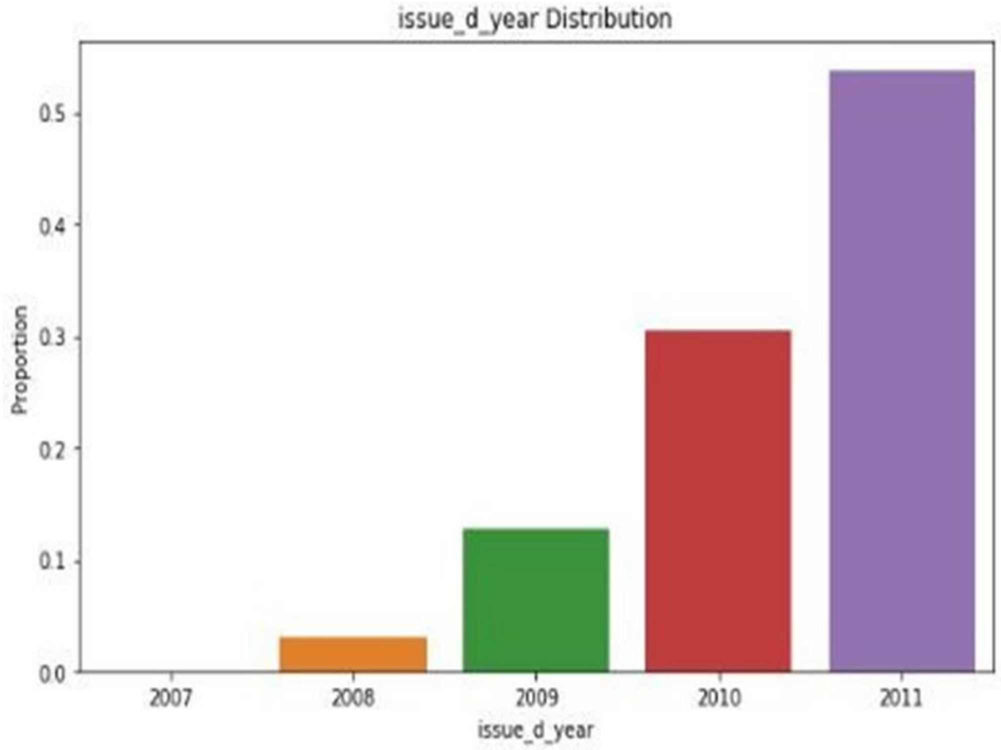
Univariate Analysis

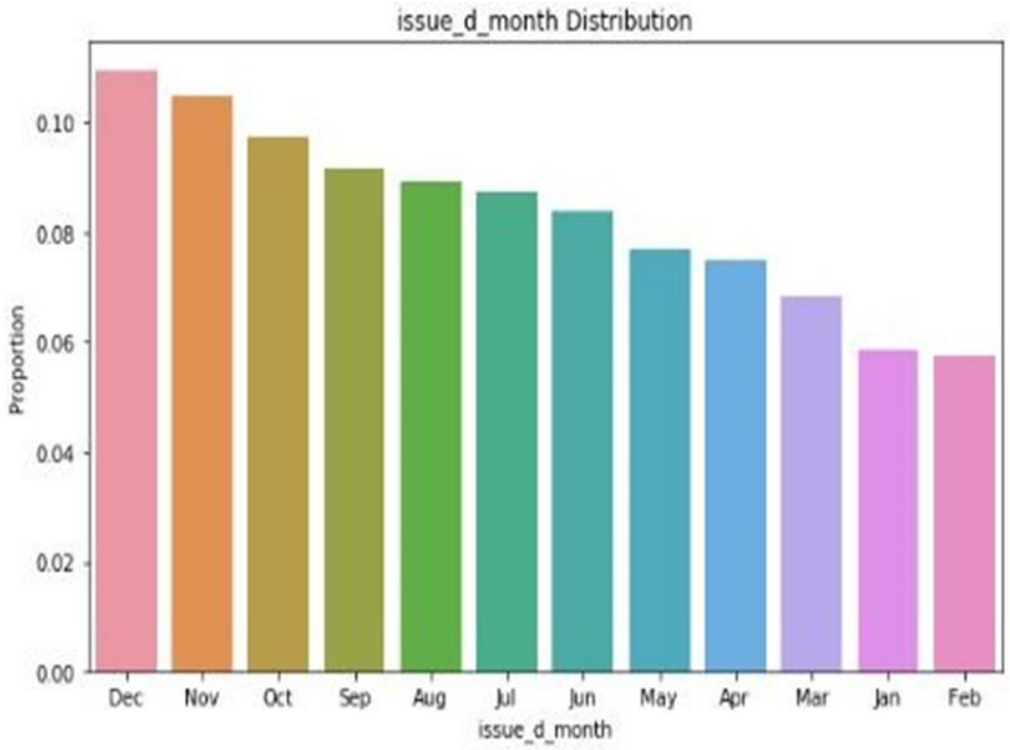




1. Around 350 people are having more the 235000 and 99% of people are having incomes below 235000.
2. 99 Percentage have not went to bankrupt.

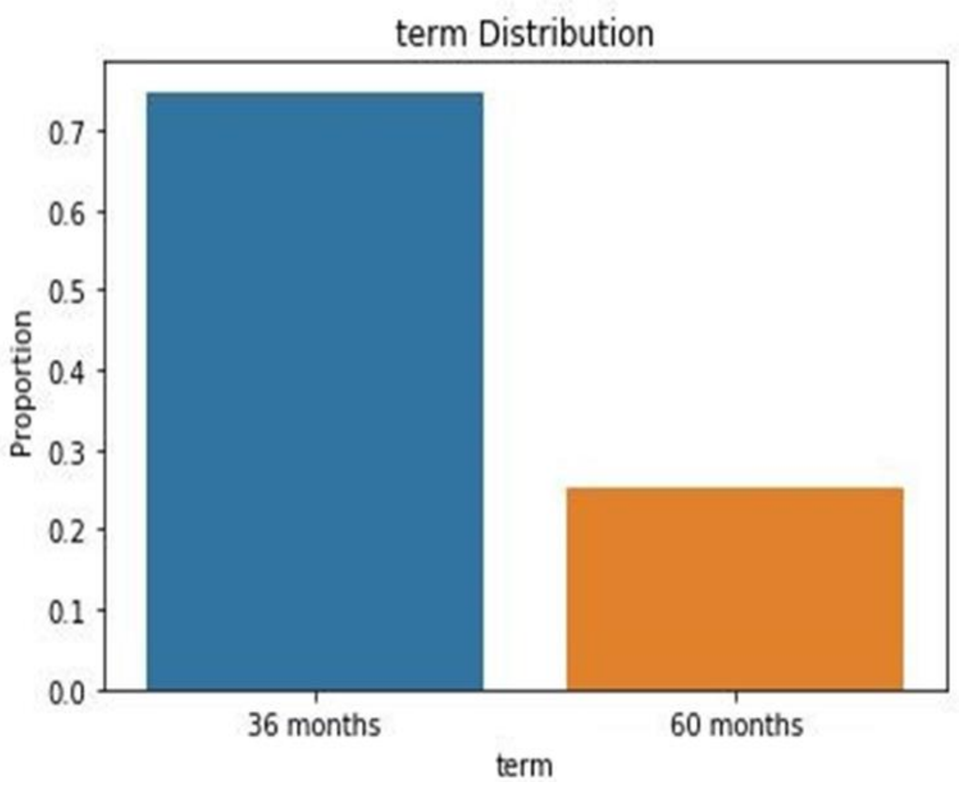
Univariate Analysis

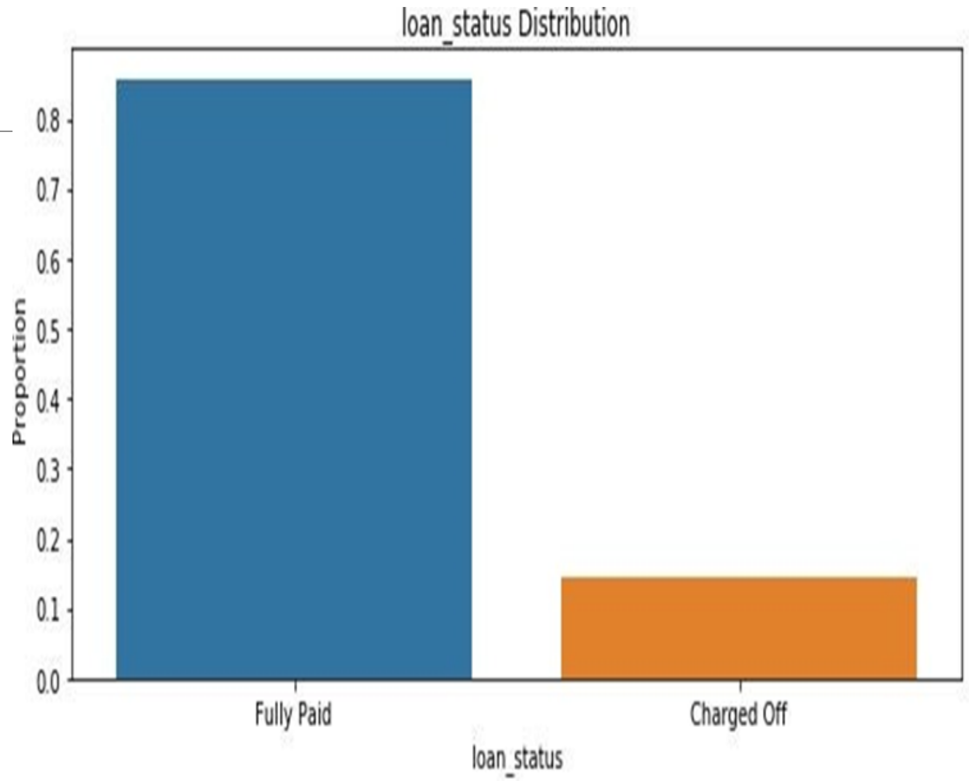




1. Lending club has really expanded year by year, the number of loan issued are doubled every year.
2. Also the issued month of loans is also increasing from January to December. In the final quarter of year there are more loans.

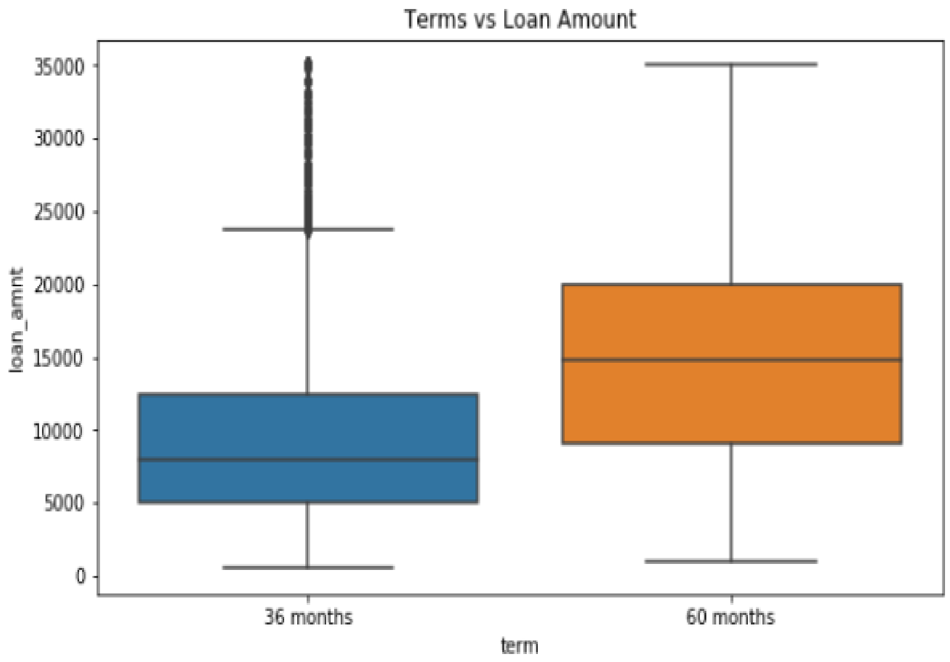
Univariate Analysis

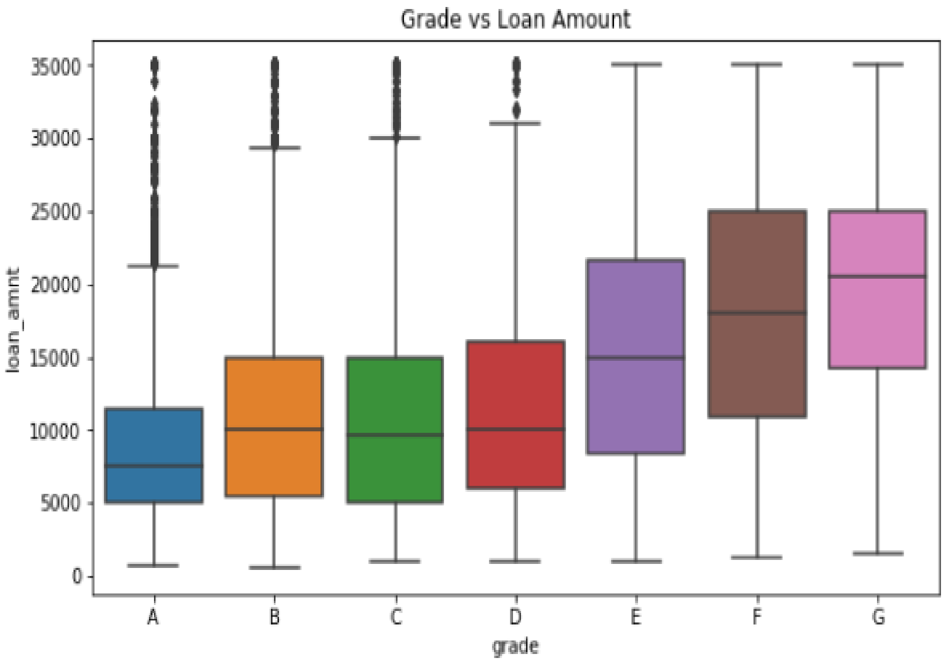




1. Around 75% borrowers took loans with 36 months term.
2. The charged off borrowers are around 15% and fully paid is around 85% in the given data set.

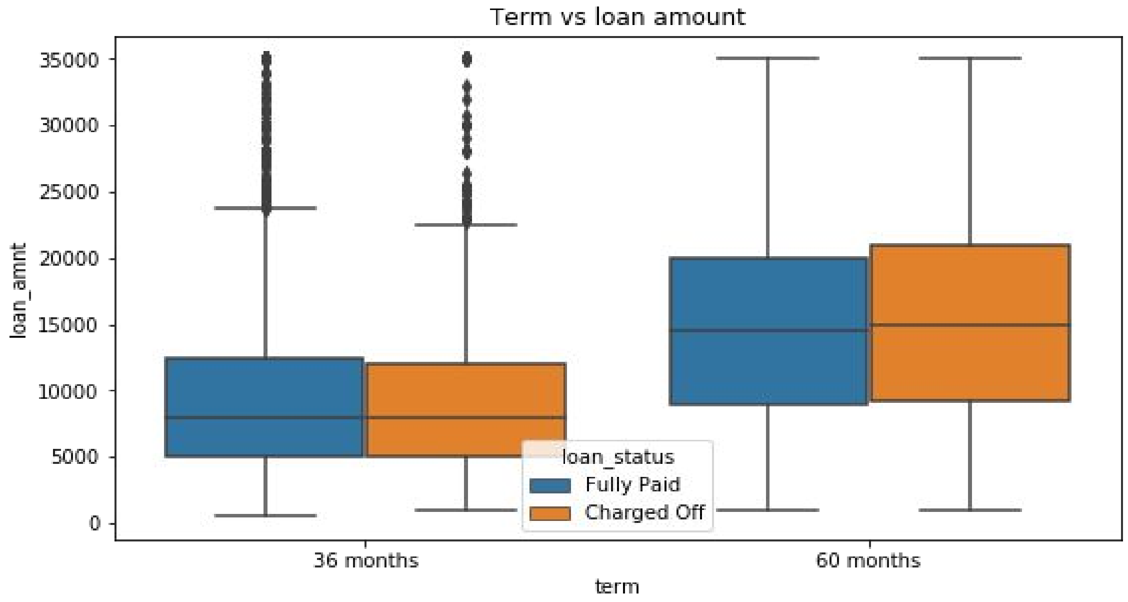
Segmented Univariate Analysis

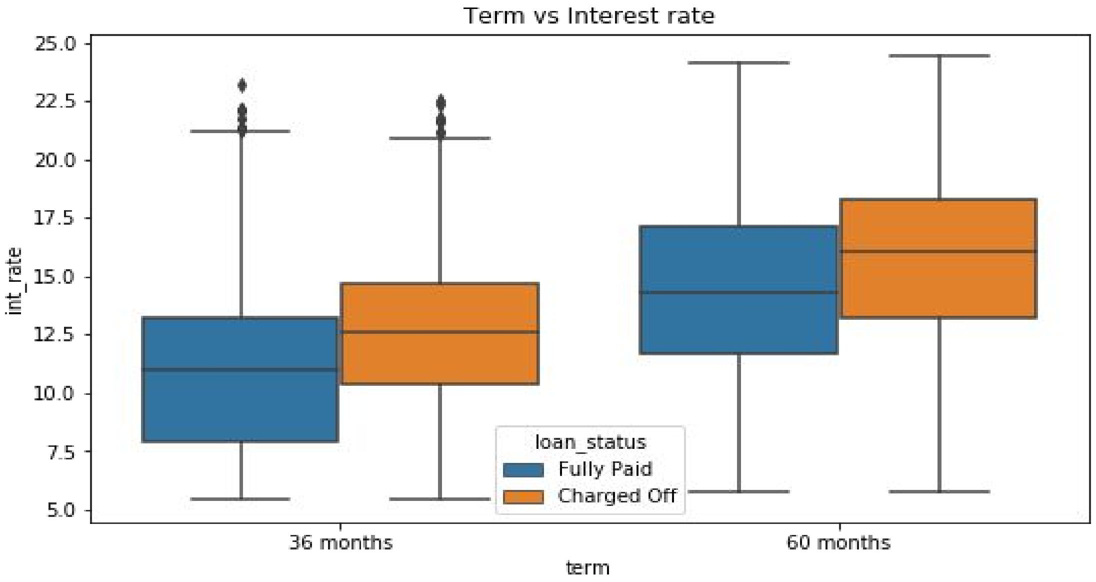


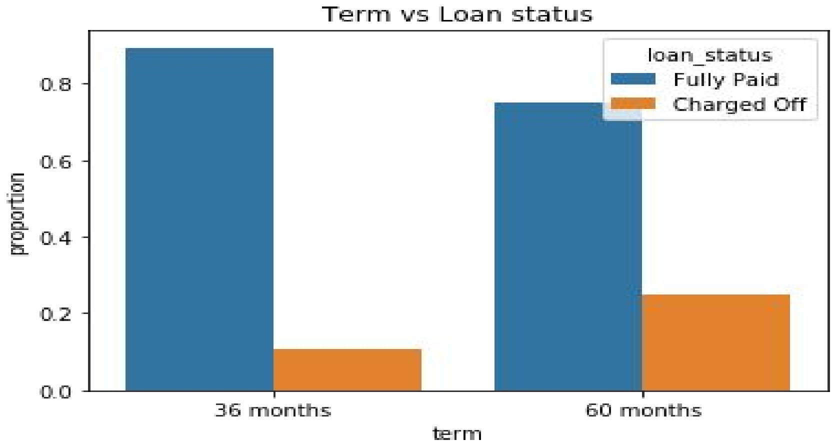


Bigger amount loans have high tenure of 60 months. Grade G and F have taken maximum loan amount

Segmented Univariate Analysis

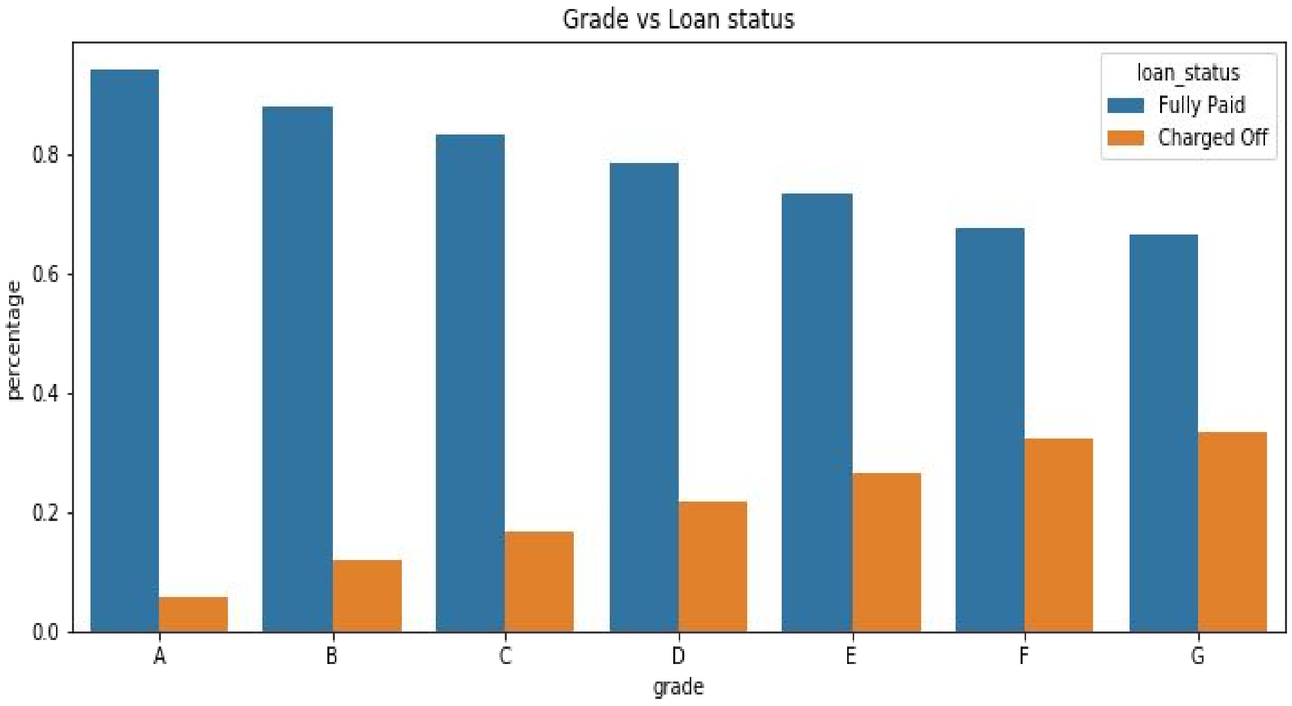


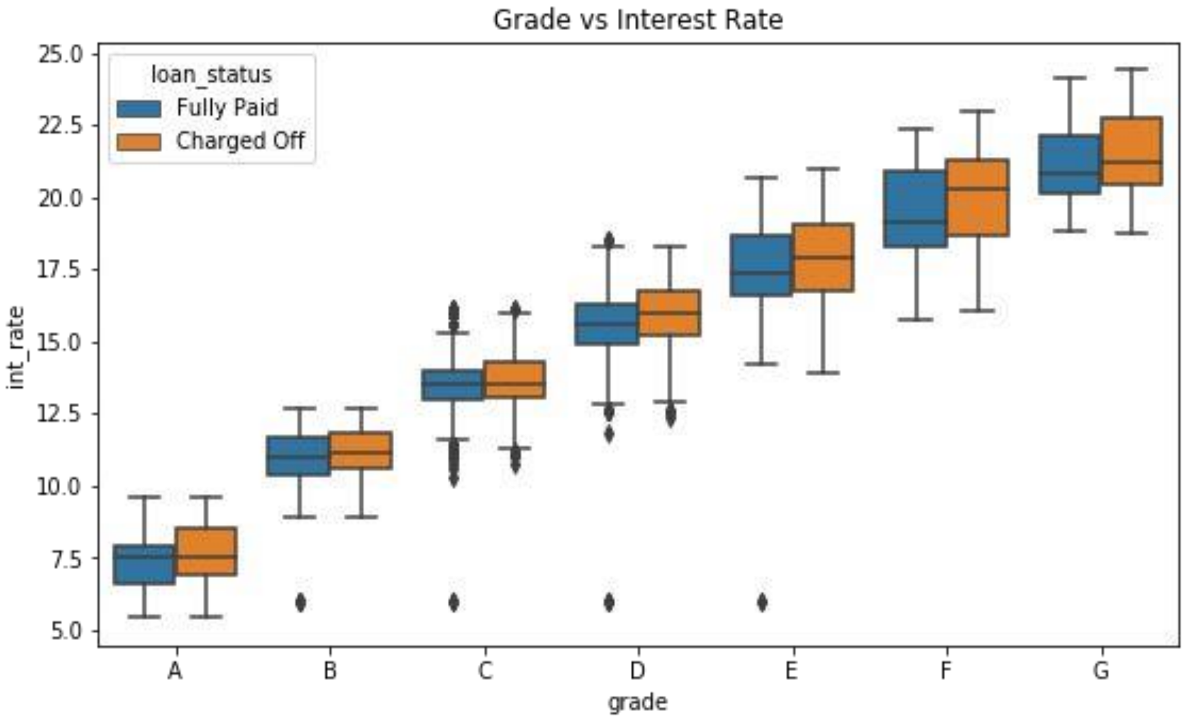




The default rate is high in 60 months tenure because most people took high loan amount with high interest rate in it and they faced difficulties in returning the sum to bank.

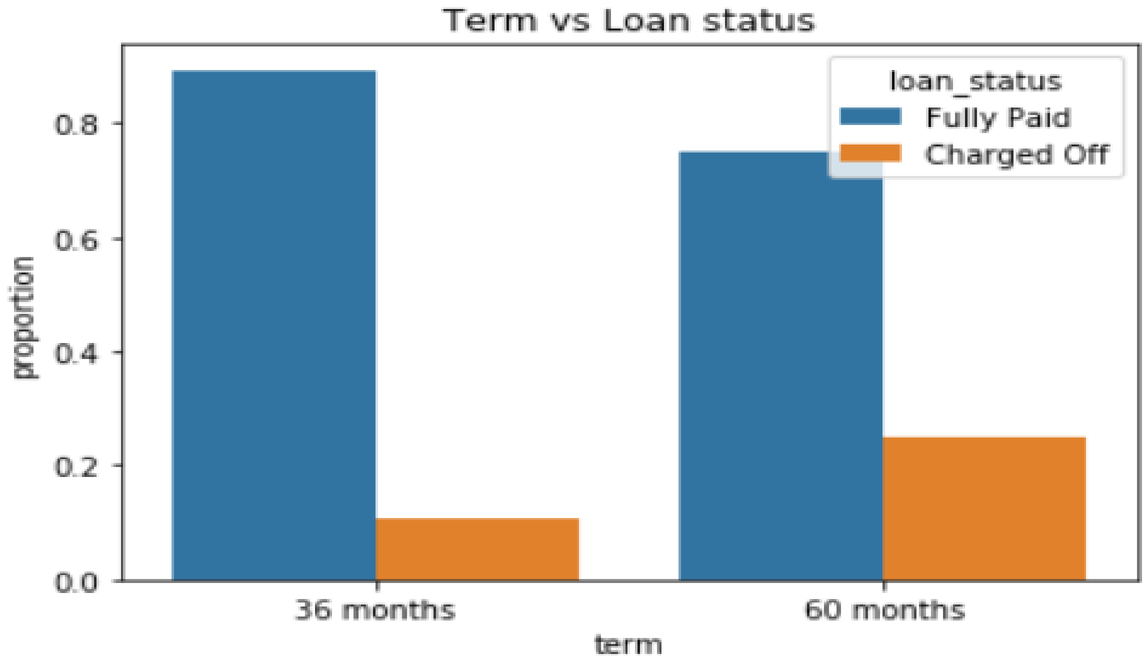
Segmented Univariate Analysis





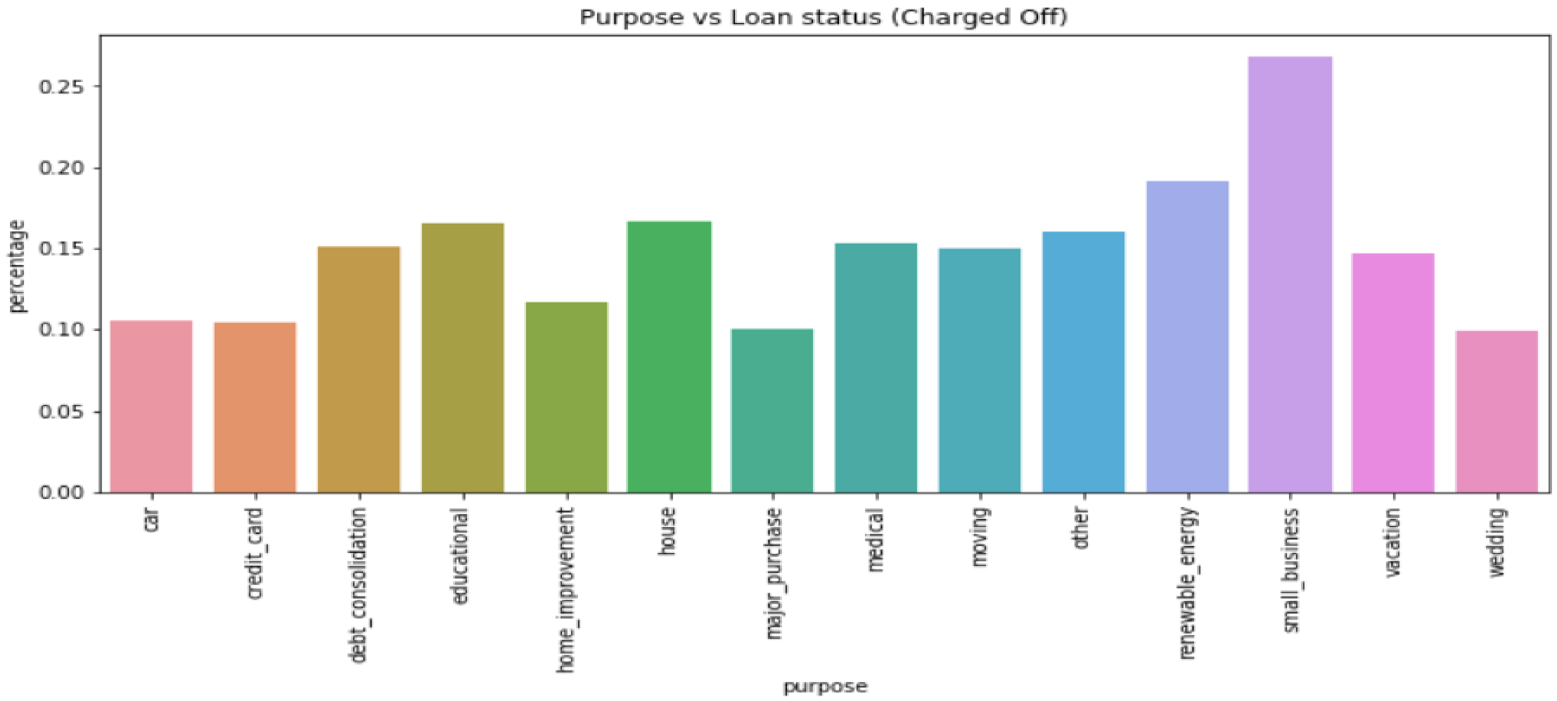
1. Grades are very good category to tell the borrower probability of defaulting the loan.
2. The Lower grades(E,F,G) have higher chances of defaulting the loan than Higher ones(A,B)
3. Also the Lower grades are getting loans for higher interest rates which might be the cause for loan default.

Bivariate Analysis



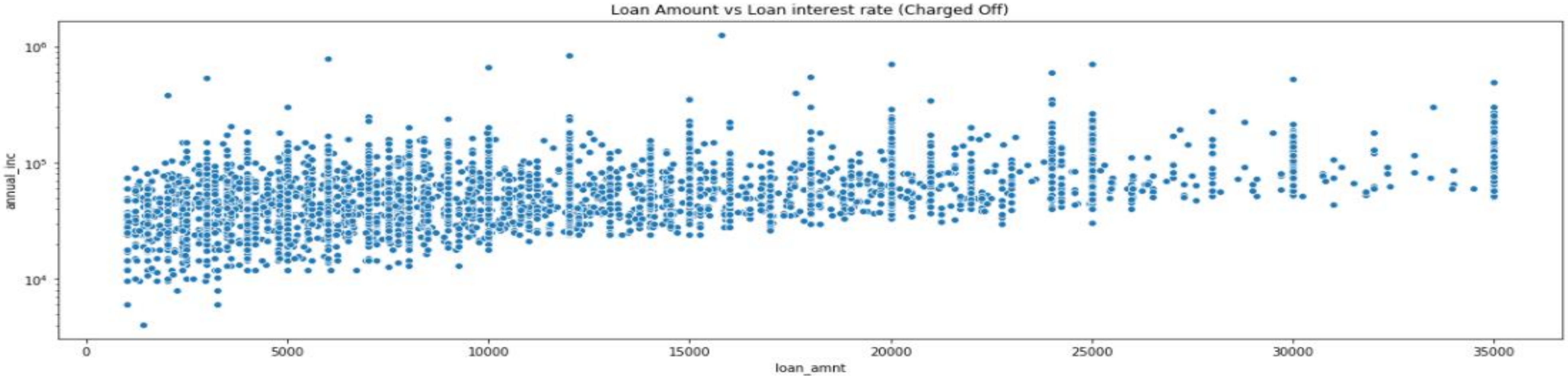
1. 36month borrower is paying better than the 60month borrowers
2. Longer tenure borrowers are delaying the payment most of the time

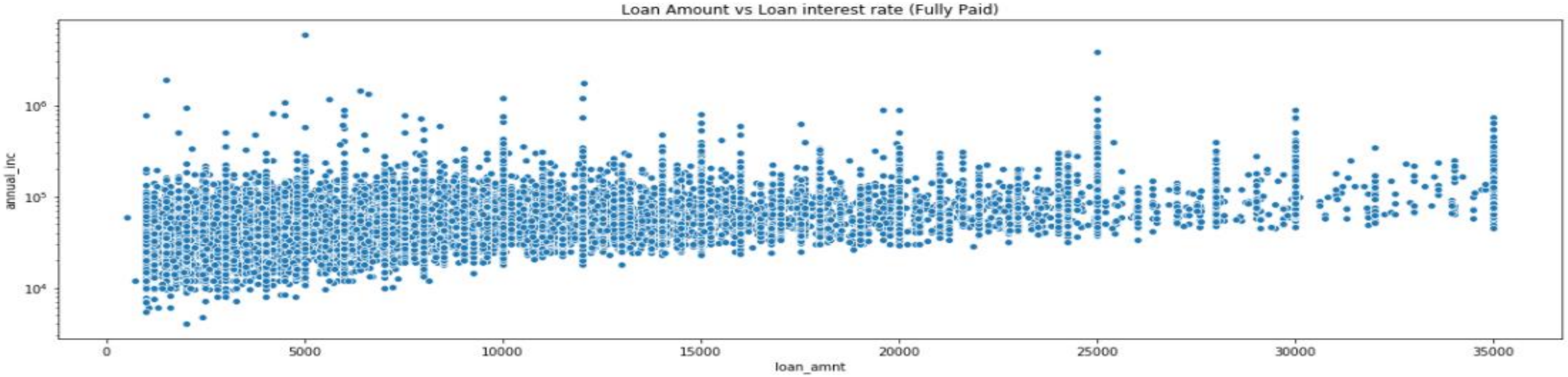
Bivariate Analysis



Charged-off is higher for small business and car loan borrowers are paying in time.

Bivariate Analysis





Both Fully paid and Charged Off loans are having similar pattern versus Annual income. We can fit a linear patter with a line which has very much less slope

Conclusion

- Lending club should reduce the Loans whose interest falls between 15-20% which are more likely defaulted.
- Small business loans are defaulted more. Lending club should stop/reduce issuing the loans to them.
- Lending club should reduce the loans from AK,UT,WY to reduce the defaulters and make profits
- 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.
- Debt consolidation loans must be reduced as there are a greater number of defaulters.

Thankyou.