

Lending Club Case Study

Group Member : Arijit Paul

Problem statement

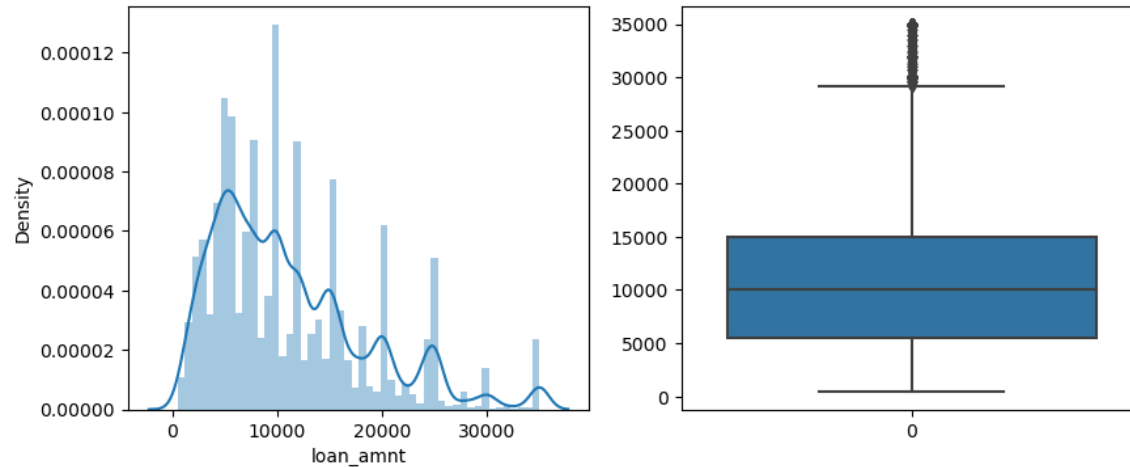
- Lending Club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.
- Like most other lending companies, lending loans to 'risky' applicants is the largest source of financial loss (called credit loss). 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 is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants 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 utilise this knowledge for its portfolio and risk assessment.

Approach taken for the analysis

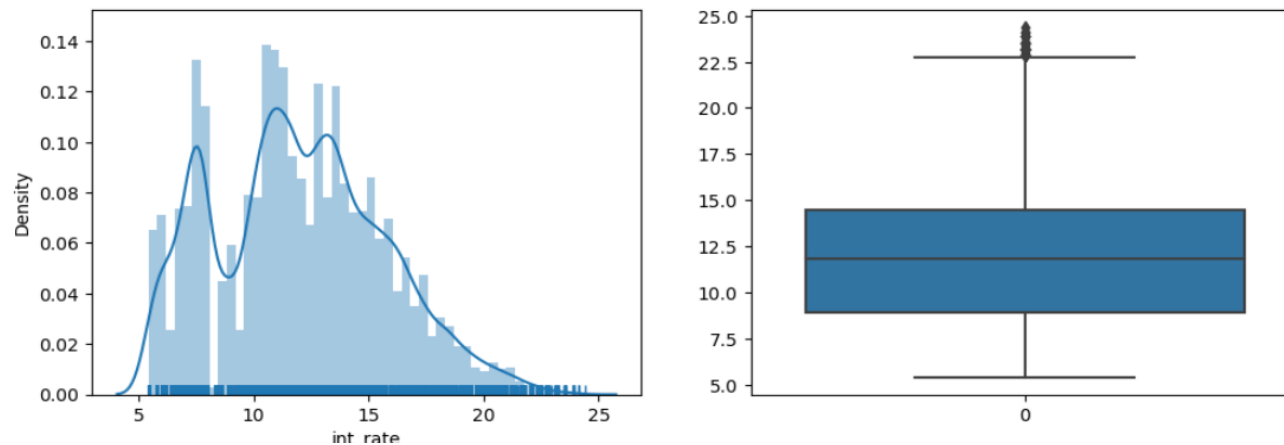
- Cleaning of the data
 - Null value columns were removed
 - Single value columns were dropped since they don't add any value
 - Irrelevant columns such as those which pertain to post loan approval were dropped
 - Null value rows were dropped since their percentage was relatively low
 - Formatting of the data was done
- Clarity of the data
 - Data dictionary was studied to get clarity of different columns and its purpose
- Univariate Analysis
 - Plots generated for single variable which provides insightful data
 - Distribution of the data for each variable
- Segmented Univariate Analysis
 - Plots generated for continuous data columns with respect to categorical data columns
 - Understand the relation of these columns
- Bivariate Analysis
 - Plots generated for two variables with respect to a categorical data
 - Understand the patterns
- Observations and Recommendation
 - Derive insights from the generated plots and identify the key indicators of default

Univariate Analysis

Distribution of Loan Amount

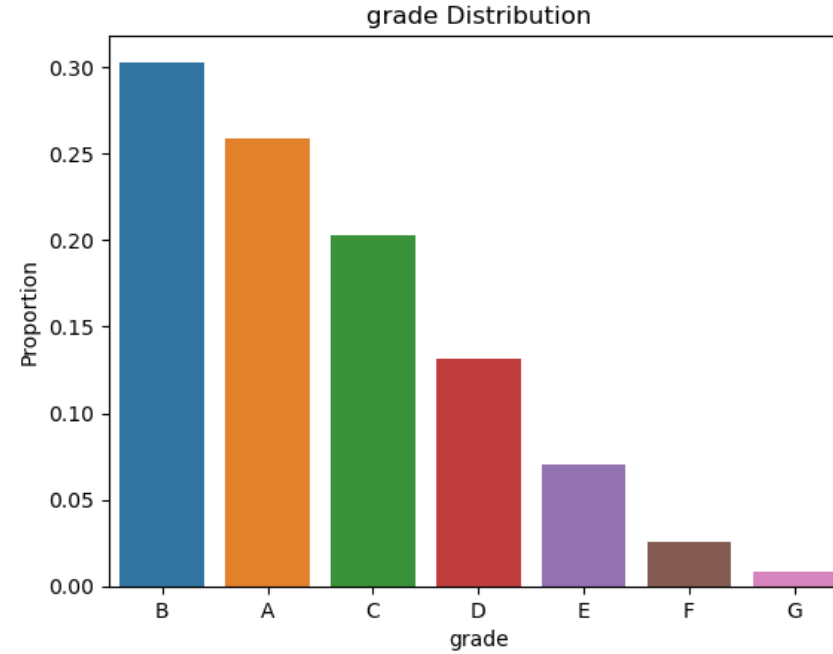
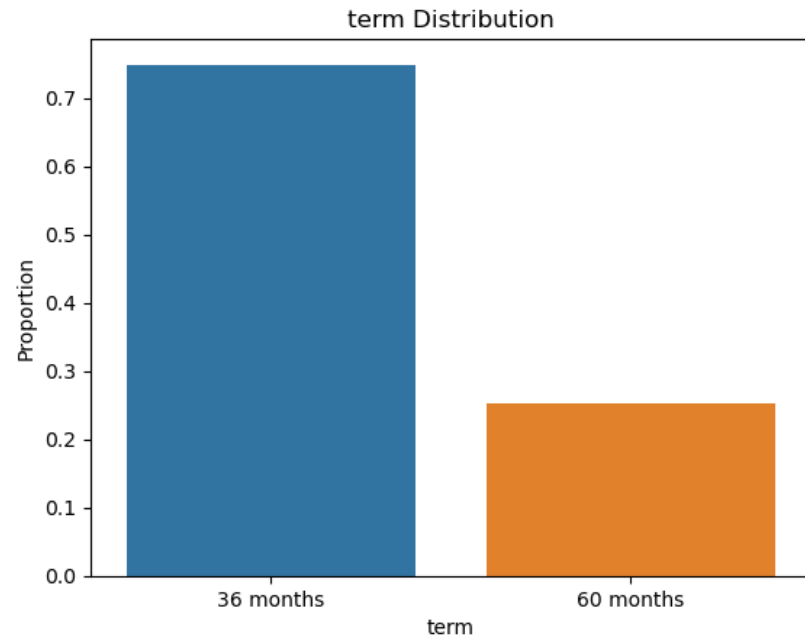


Distribution of interest rate



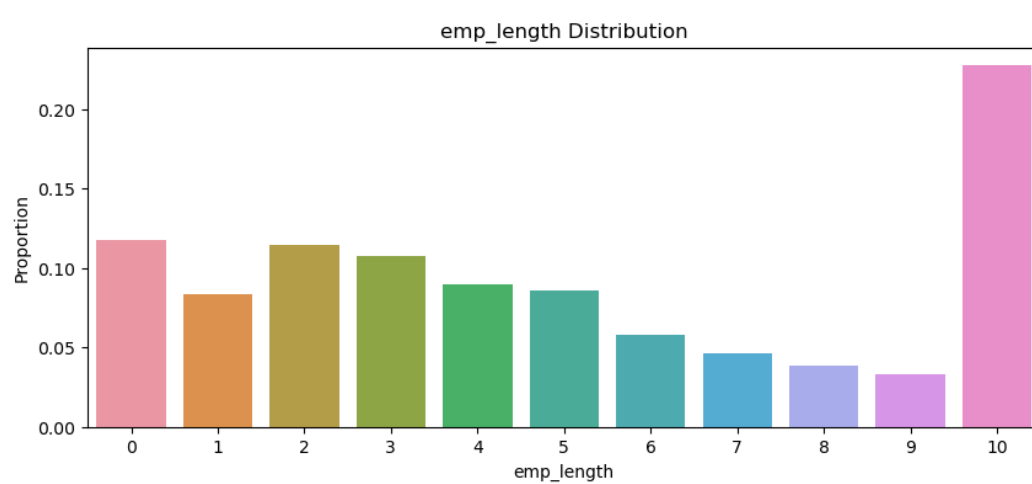
- From the plot we can see that most people took a loan amount of 10000 and only few took loan > 30000
- From the plot we can see that most of the interest rates lies between 9% to 14.5%. Few people took loan at higher rates of interest i.e., 22.5%

Univariate Analysis

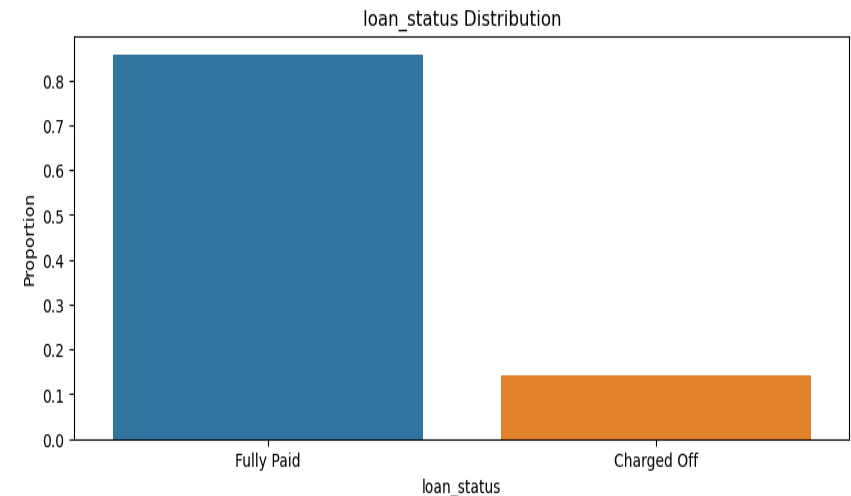
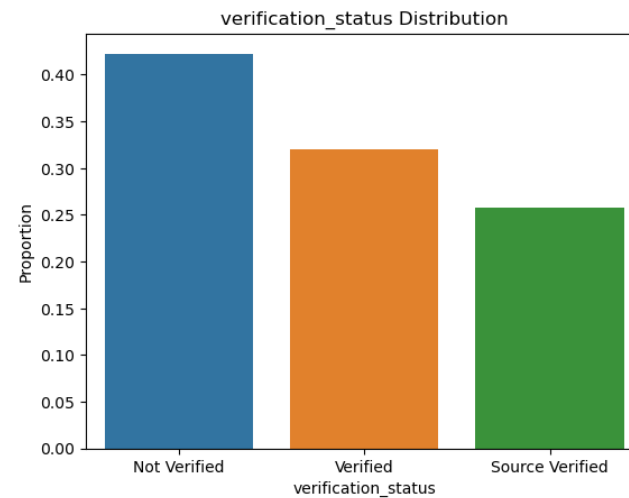
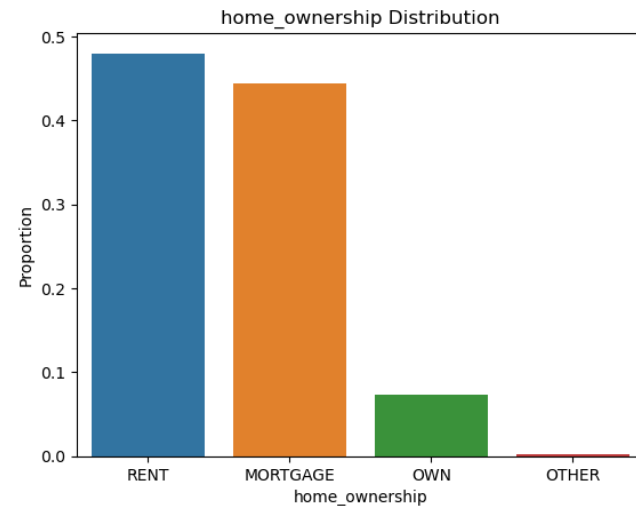


- Borrowers have taken 36 months tenure more than 60 months.
- Most borrowers fall under A and B grades then other grades

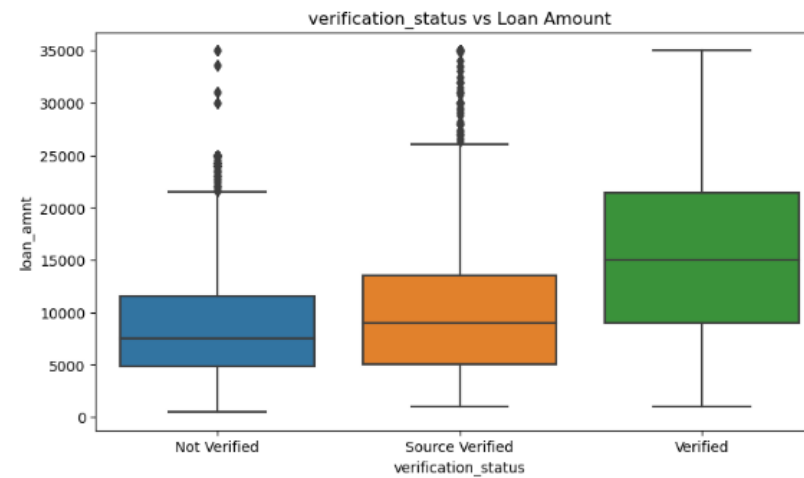
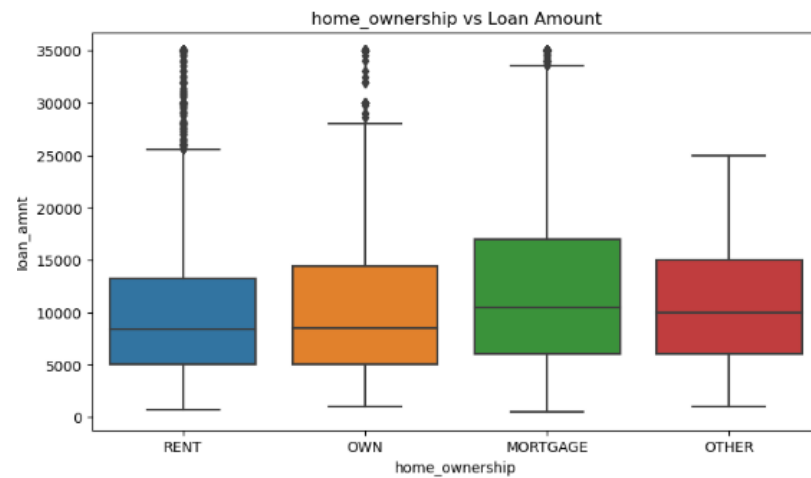
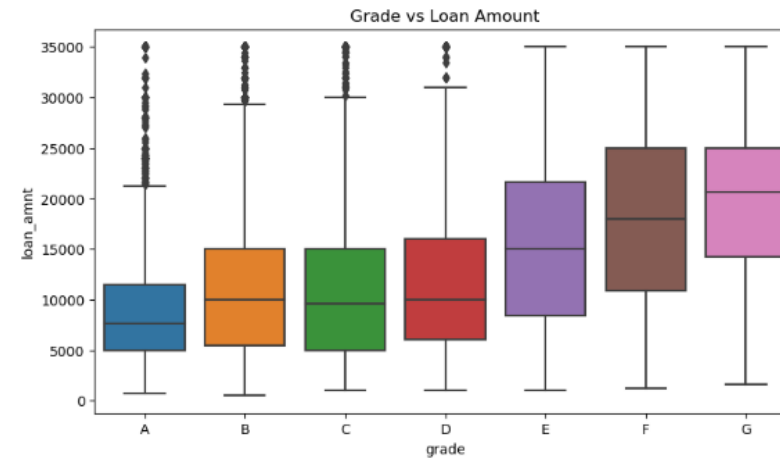
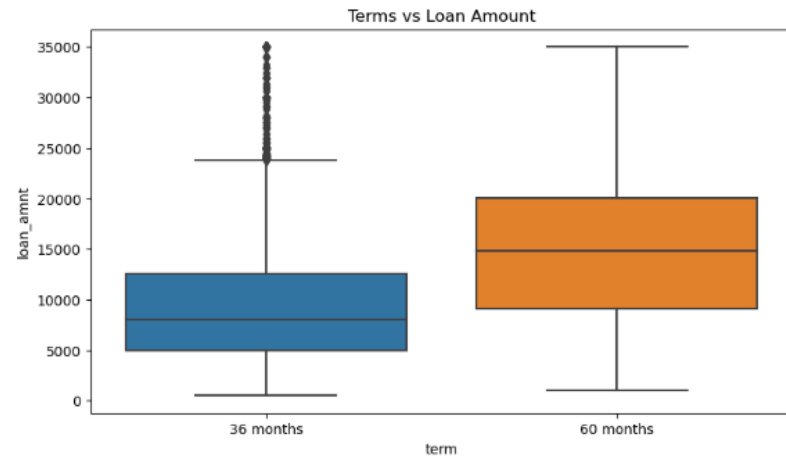
Univariate Analysis



- Borrowers are mostly 10+ years employment length.
- The loan borrowers are mostly having rented and mortgage houses.
- Majority of loans were given without verification of applicant's income.
- 85% of borrowers has paid the loan fully, whereas 15% has defaulted the loan.

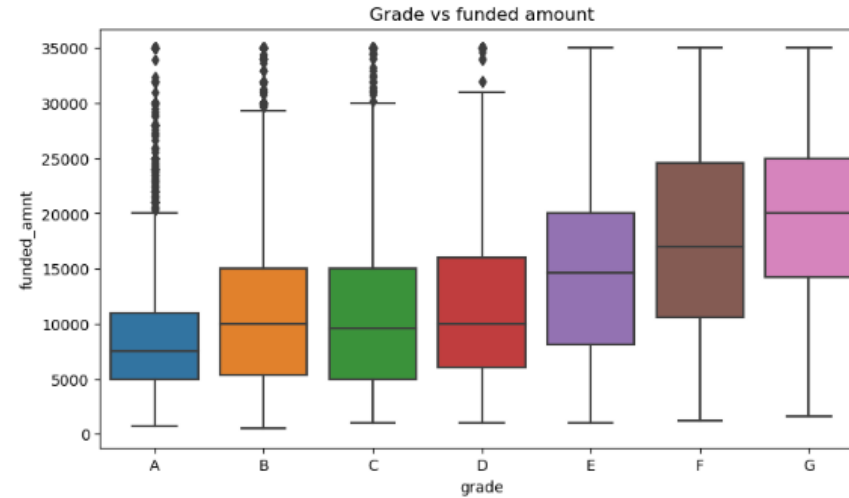
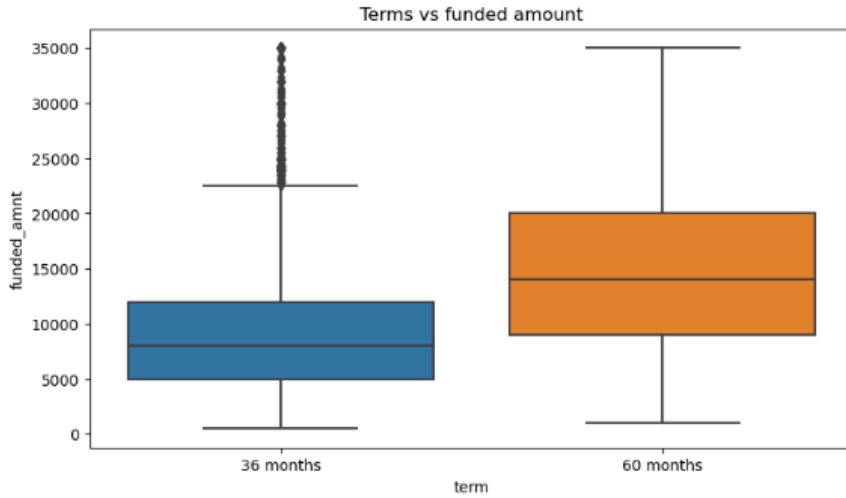


Segmented Univariate Analysis

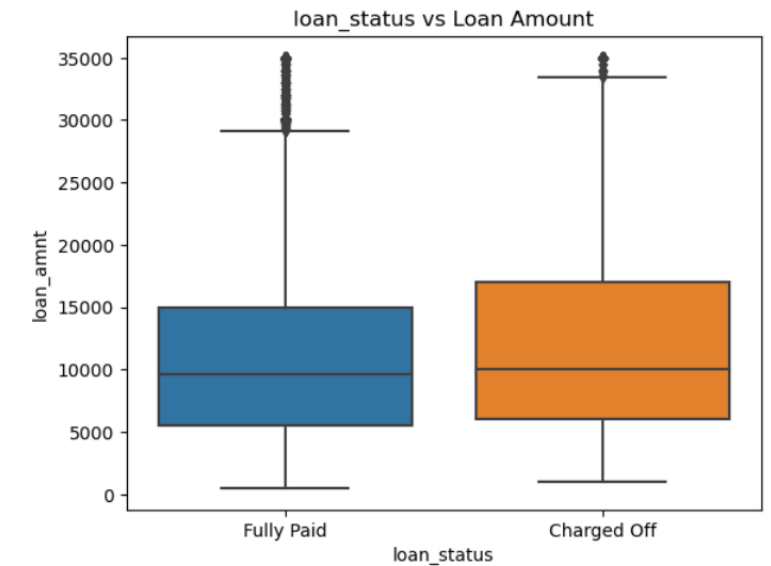
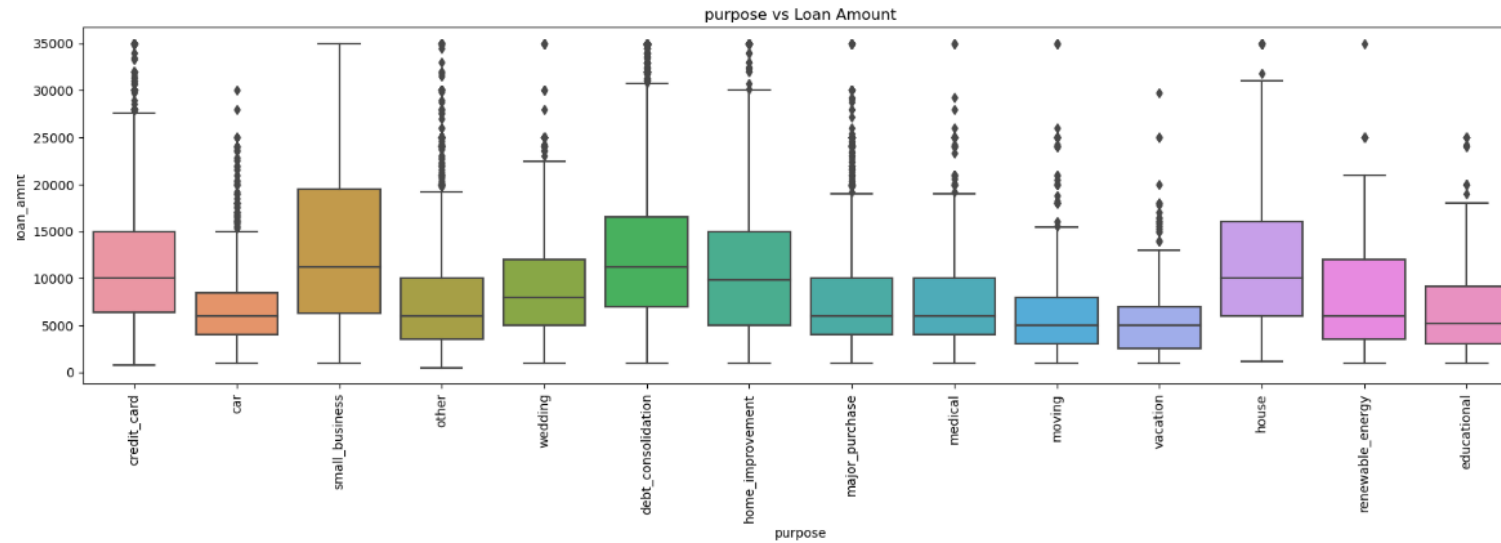


- Amount of loan increases with tenure, in this case loan amount is higher for 60 months tenure.
- Grade 'F' and 'G' have taken max loan amount. As Grades are decreasing the loan amount is increasing.
- More borrowers are from MORTGAGE and also the median loan amount also high for MORTGAGE owned borrowers.
- And most of borrowers are verified for borrowing loan >9k

Segmented Univariate Analysis

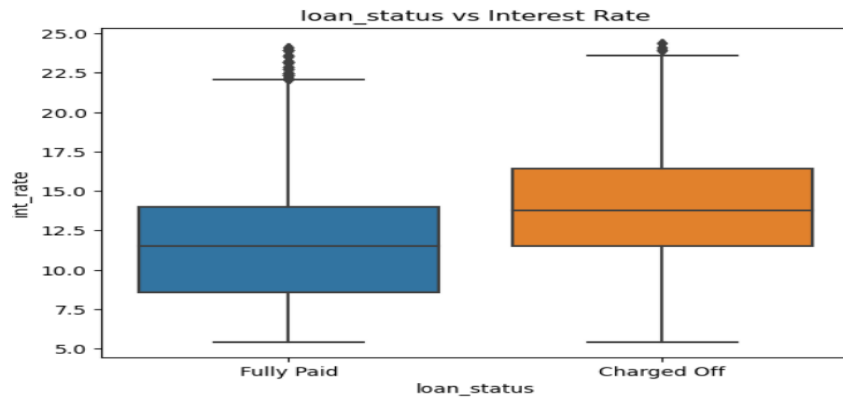
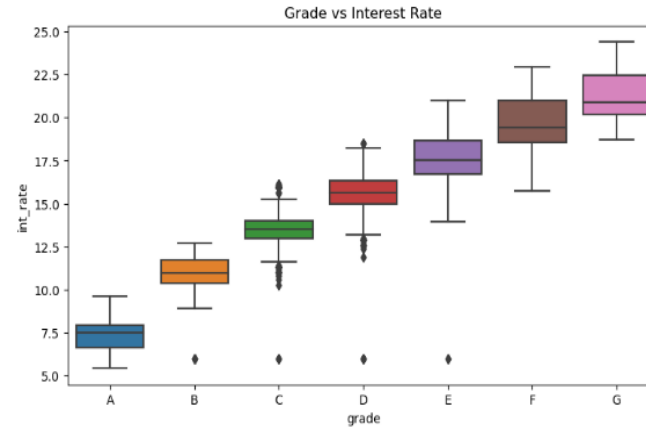
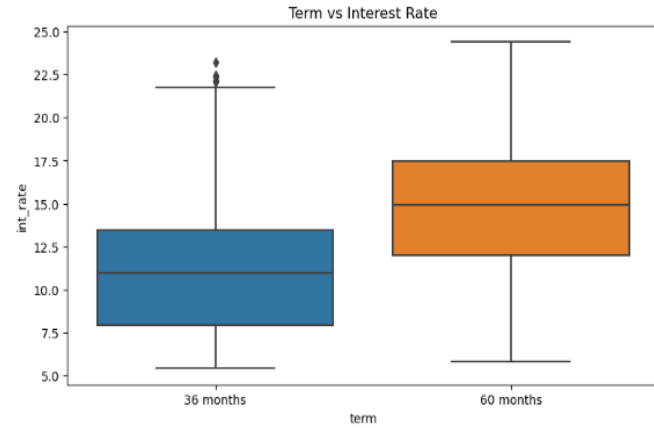


- Higher the tenure, higher is the Funded amount. With lesser grades (F&G) funded amount increases.
- More loan amount is from Small business followed by debt consolidation.
- Charged Off loans have higher amounts than Fully Paid ones.

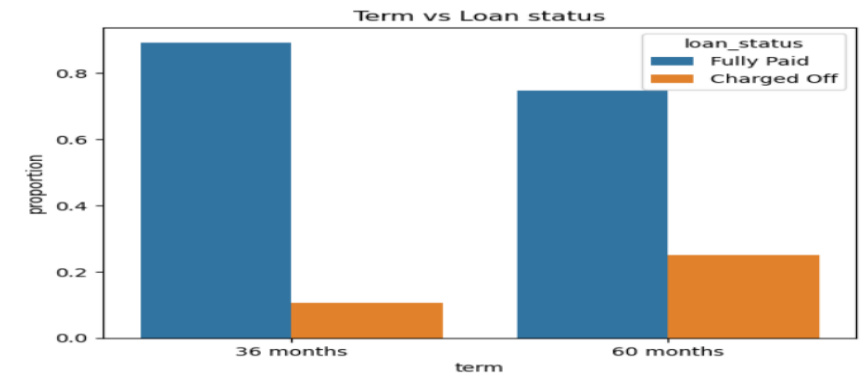
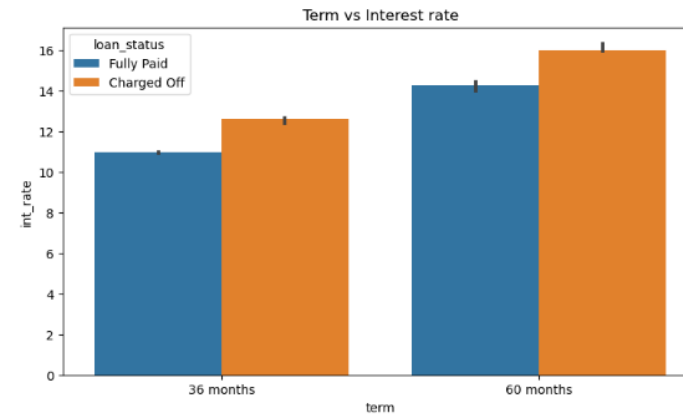
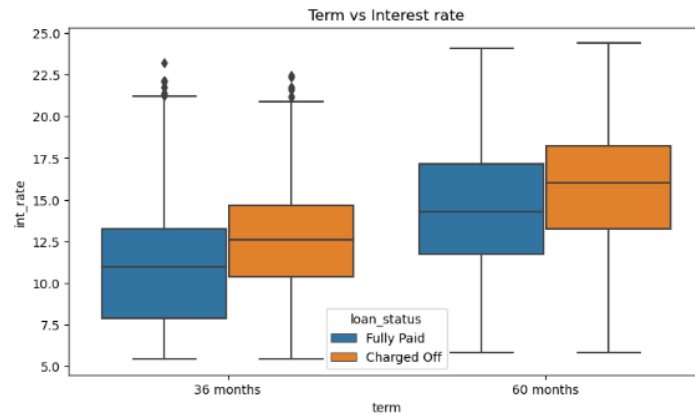
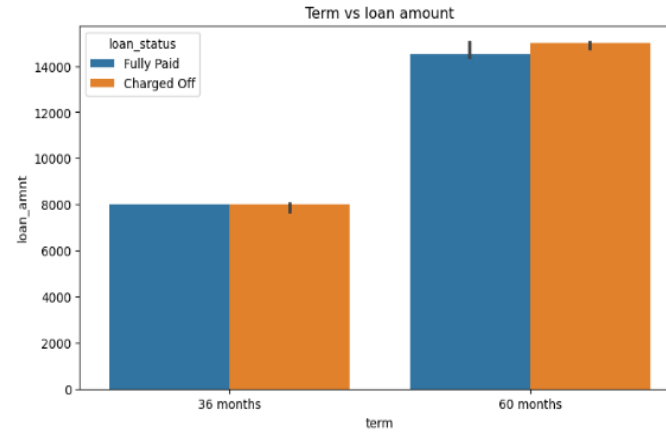
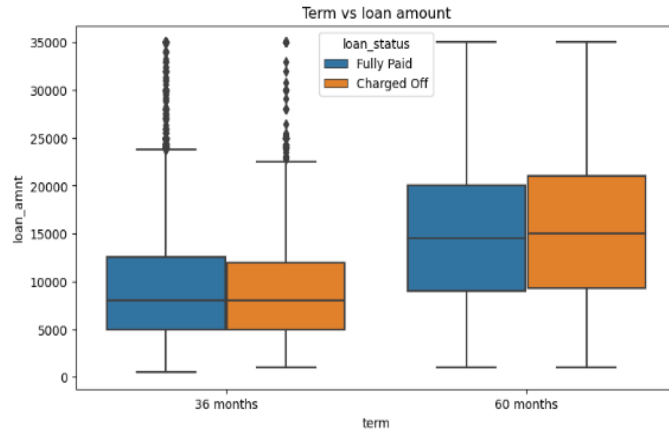


Segmented Univariate Analysis

- The interest rates are higher for higher tenure loans.
- Interest rates are increasing when grades are lowering (A to G).
- We can clearly see that, greater the interest rate more the chance of defaulting the loan.

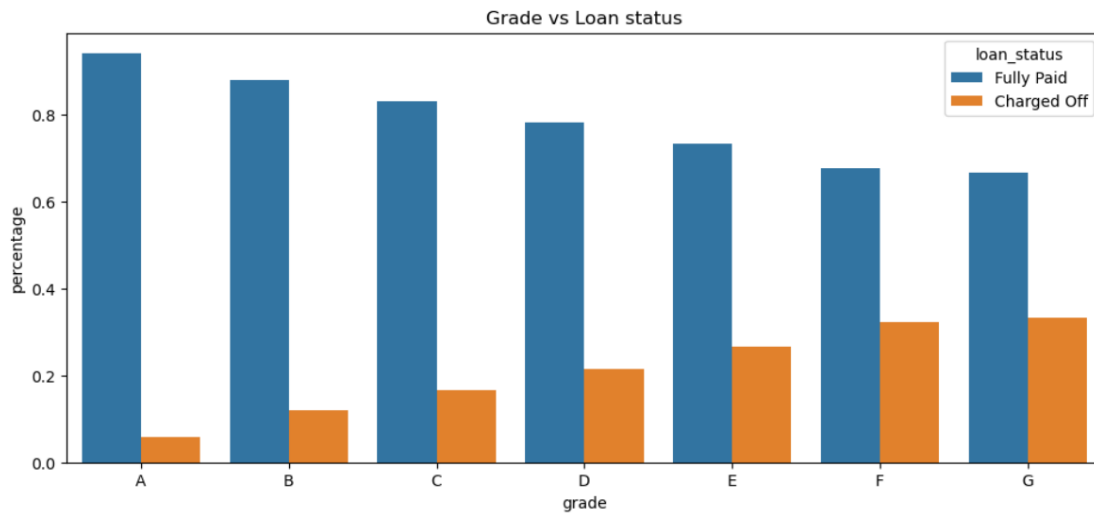
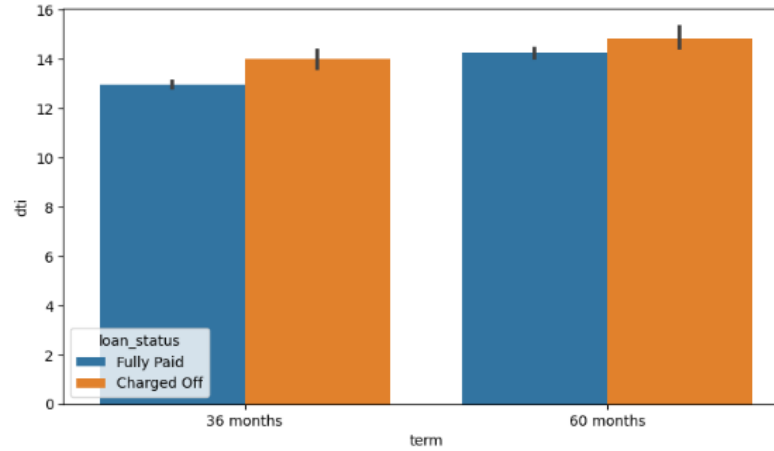
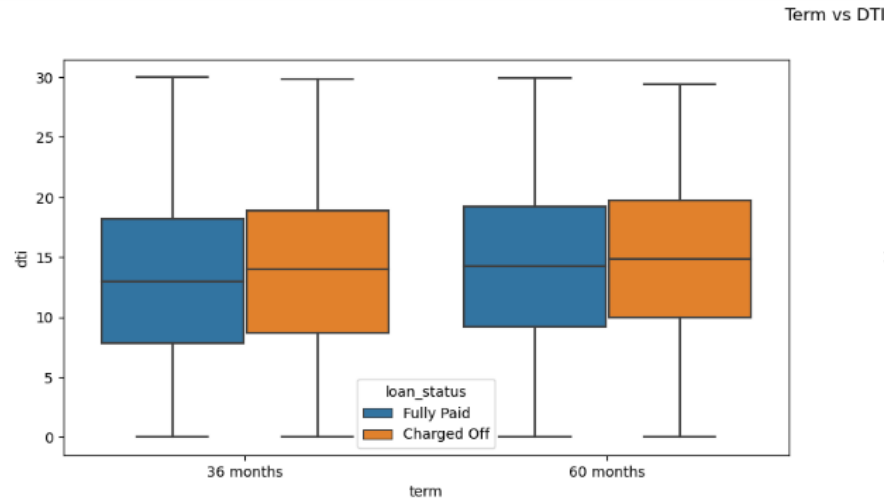


Bivariate Analysis



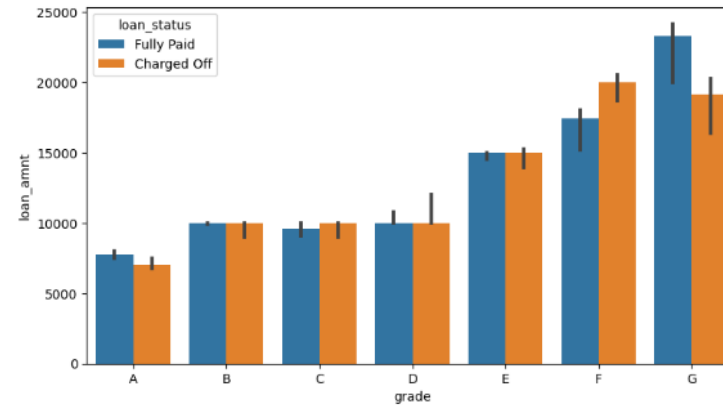
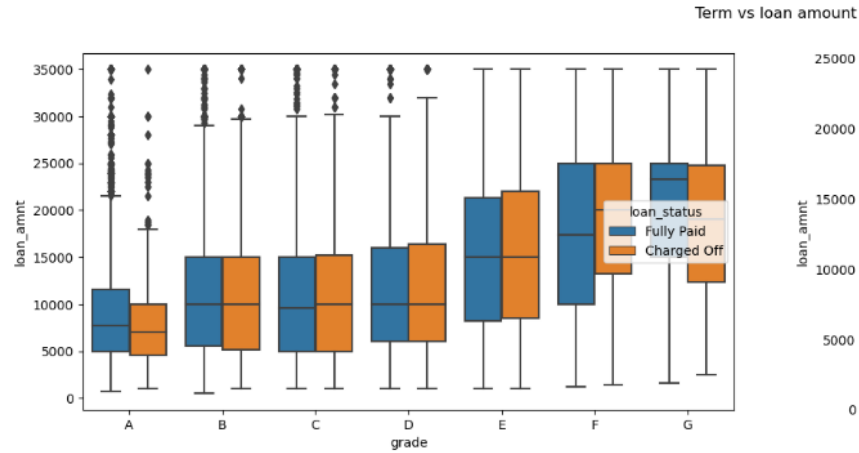
- Loan amount is not the deciding factor for defaults in both 36 and 60 months term.
- Borrowers have equal distribution in both default and non default for 36 and 60 months term.
- For higher interest rates the default rate is higher in both 36 and 60 months tenure.
- More proportion of borrowers defaulted loan in 60 months term than 36 months. Also, the Fully Paid rate is higher in 36 months tenure.

Bivariate Analysis



- Comparitively charge-off are higher when compared with fully-paid for the Debit to income ratio.
- The graph clearly shows that the Charged off increases as grades decreases.

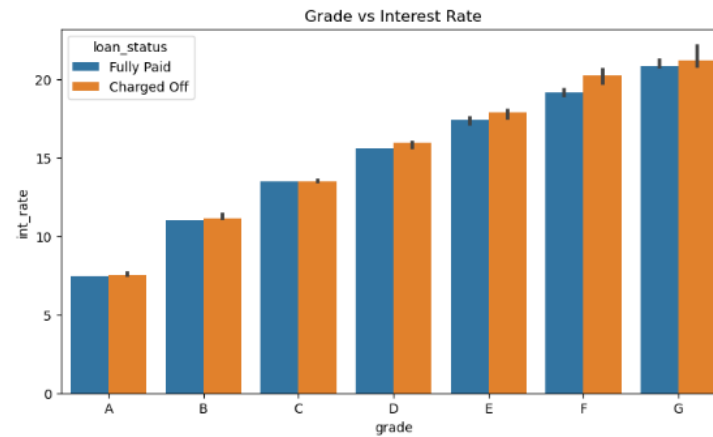
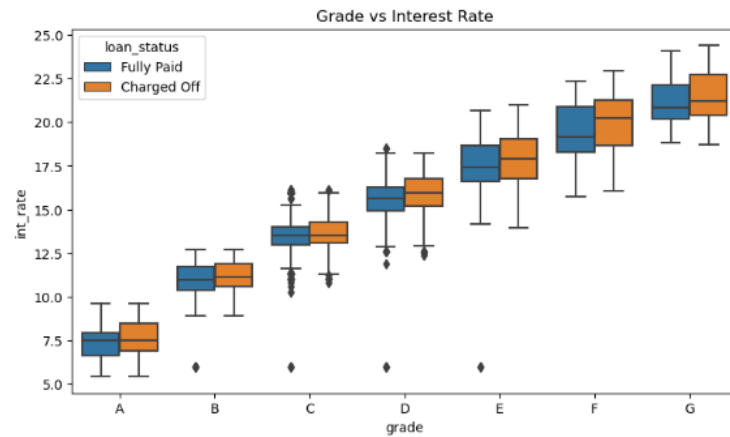
Bivariate Analysis



- For lower grades 'F' and 'G' there are more difference between charged-off and fully paid.

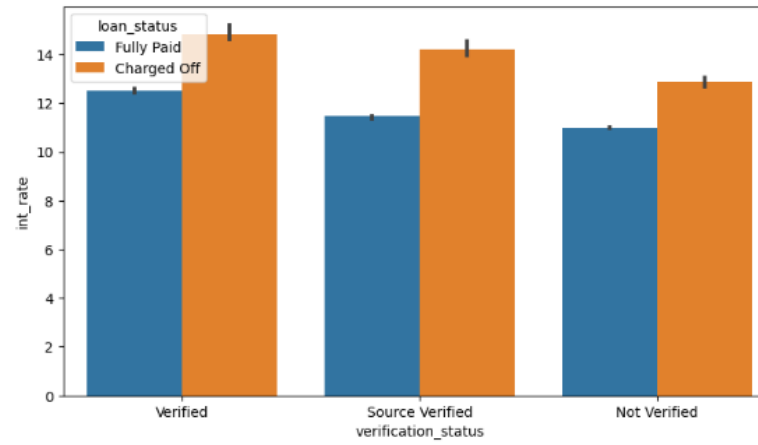
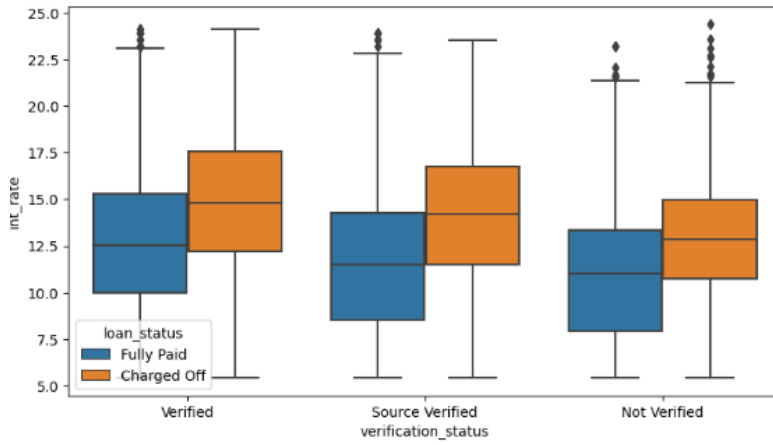
- The lower grade people has taken higher amount of loans and also they are more prone to default the loan.

- As grade decreases the interest rate gradually increases. and they are more and more prone to default the loan.

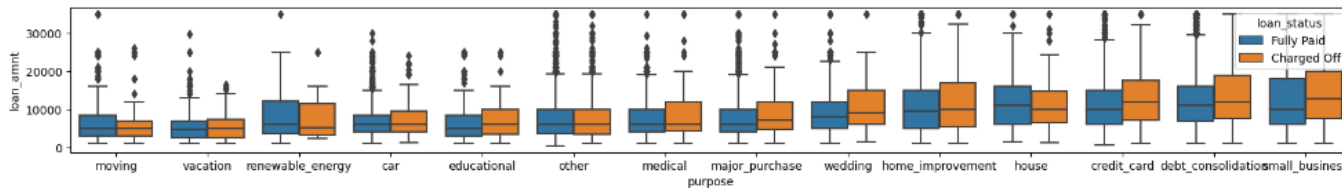
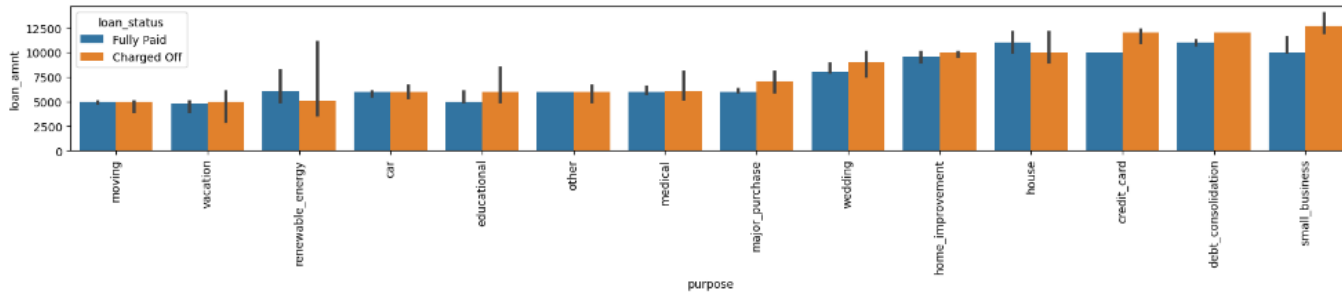


Bivariate Analysis

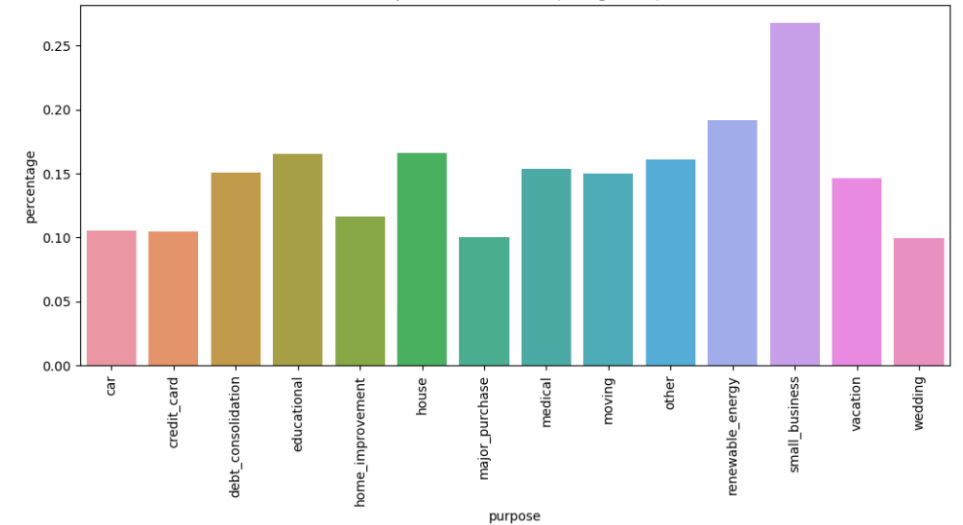
Verification Status vs interest rate



- Irrespective of verification status higher interest rates are incurring default of loan.
- Small Business has more defaults when the loan amount is also high.
- Charged-off are higher for small_business comparatively.



Purpose vs Loan status (Charged Off)



Final Observation

- The lower grade people has taken higher amount of loans and also they are more prone to default the loan.
- As grade decreases the interest rate gradually increases. and they are more and more prone to default the loan.
- Small Business has more defaults when the loan amount is also high.
- Irrespective of verification status higher interest rates are incurring default of loan.