Lending Club Case Study

Business Objectives

- Lending club is 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.

- Main objective is to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.

Problem Solving Approach:

- Data Loading and Data Understanding.

  Using data dictionary and dataset understand the variables and data
- Data Cleansing

  Check for missing values set a threshold and remove the variables having high percentage of missing values.
- Data Standardization

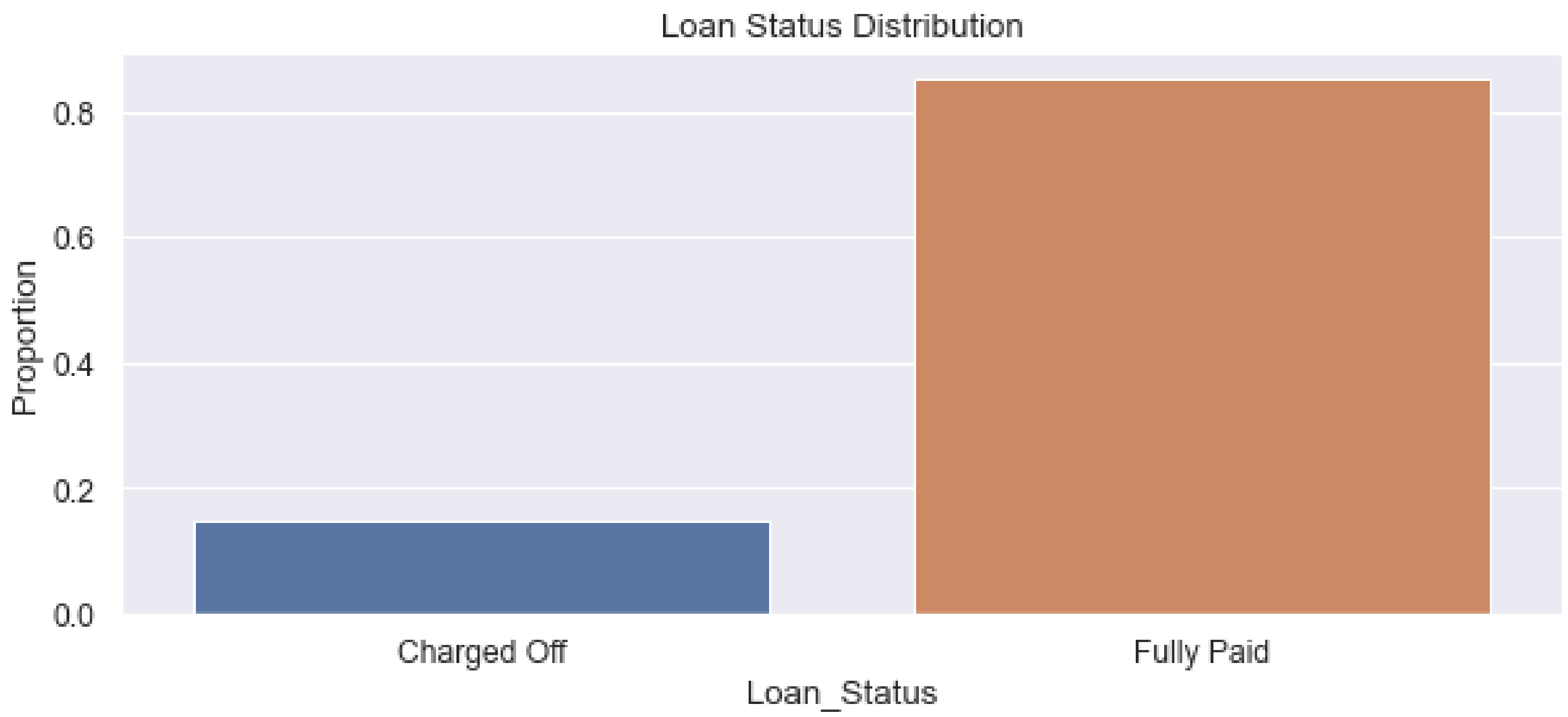
  Remove '%' , '>' from variables
- Univariant Analysis

  Analysing single variable and get insight from it
- Bivariant Analysis

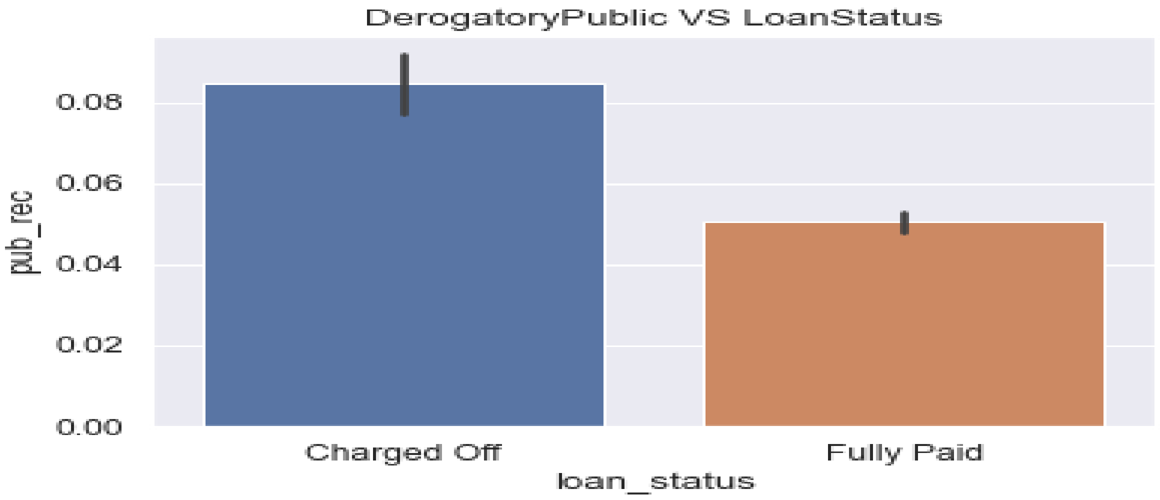
  Analysing two variables behaviours and get insight from it
- Segmented Univariant Analysis

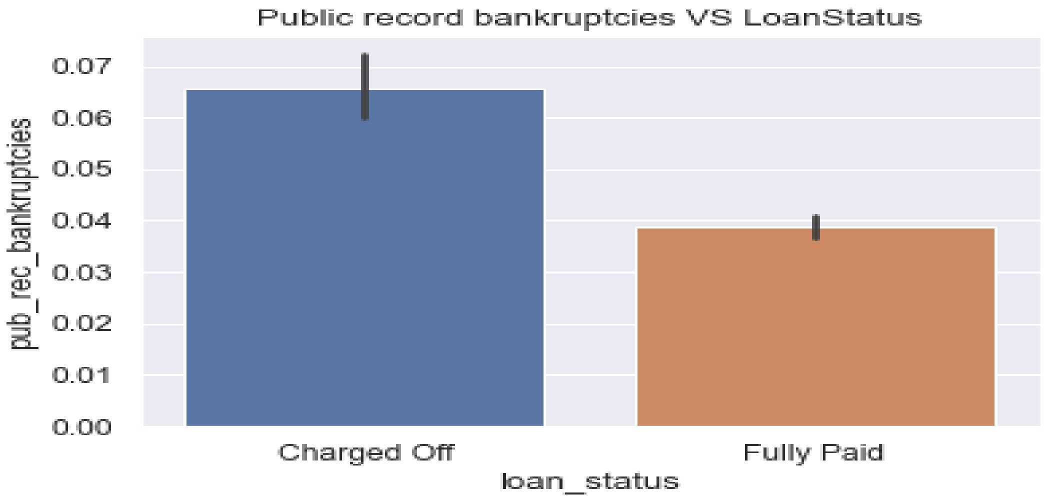
  Analysing continuous variable against categorical variable
- Observations/Recommendation

Analysis Result

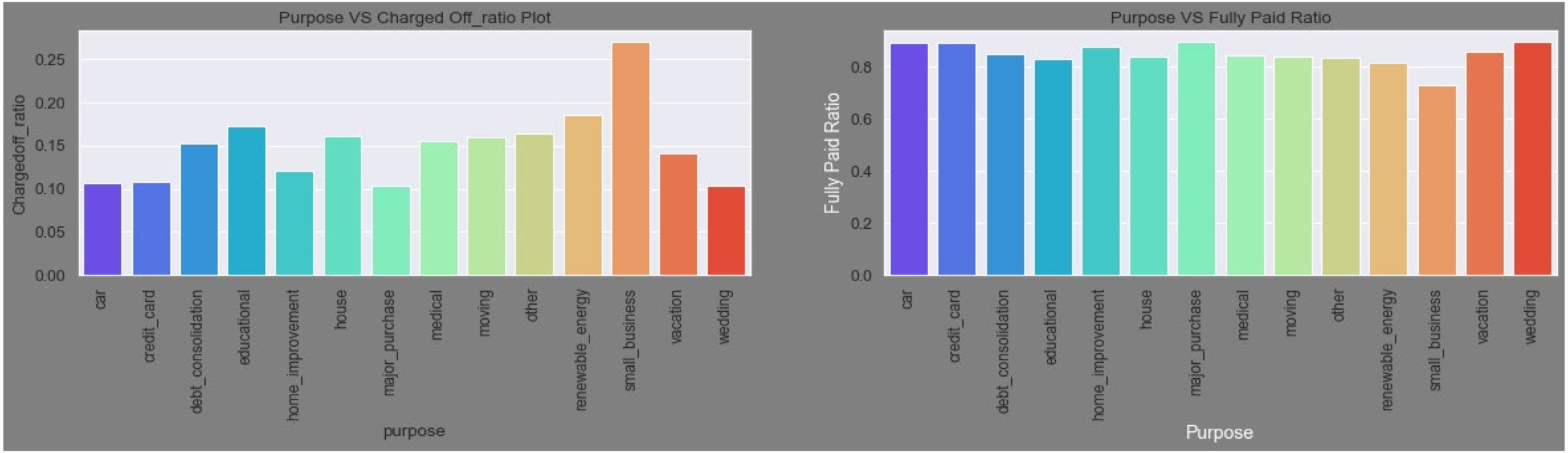


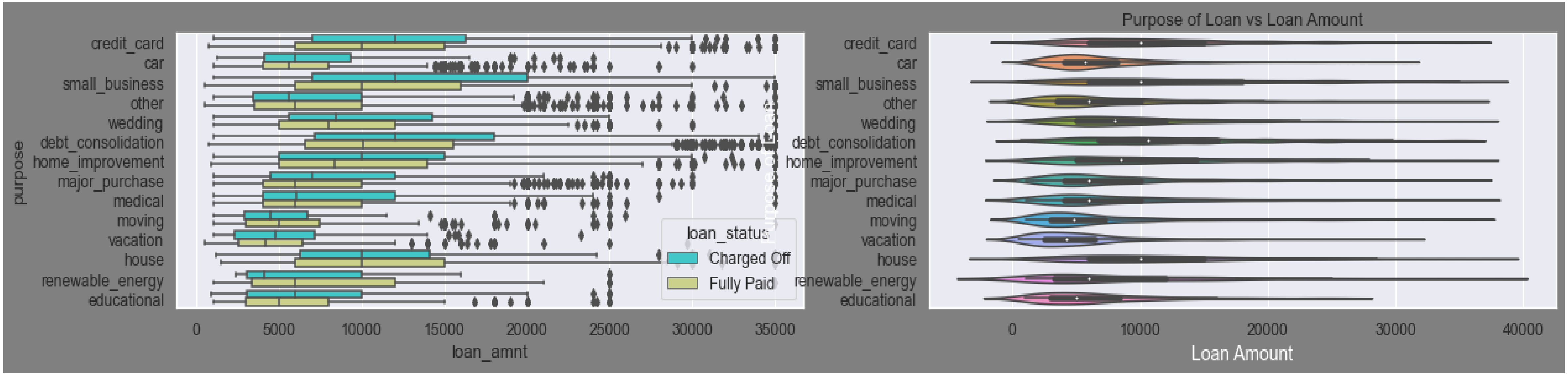
Dataset is biased towards Fully Paid loans. More then 80% is Full Paid loans and Charged Off Loan around 14.5%



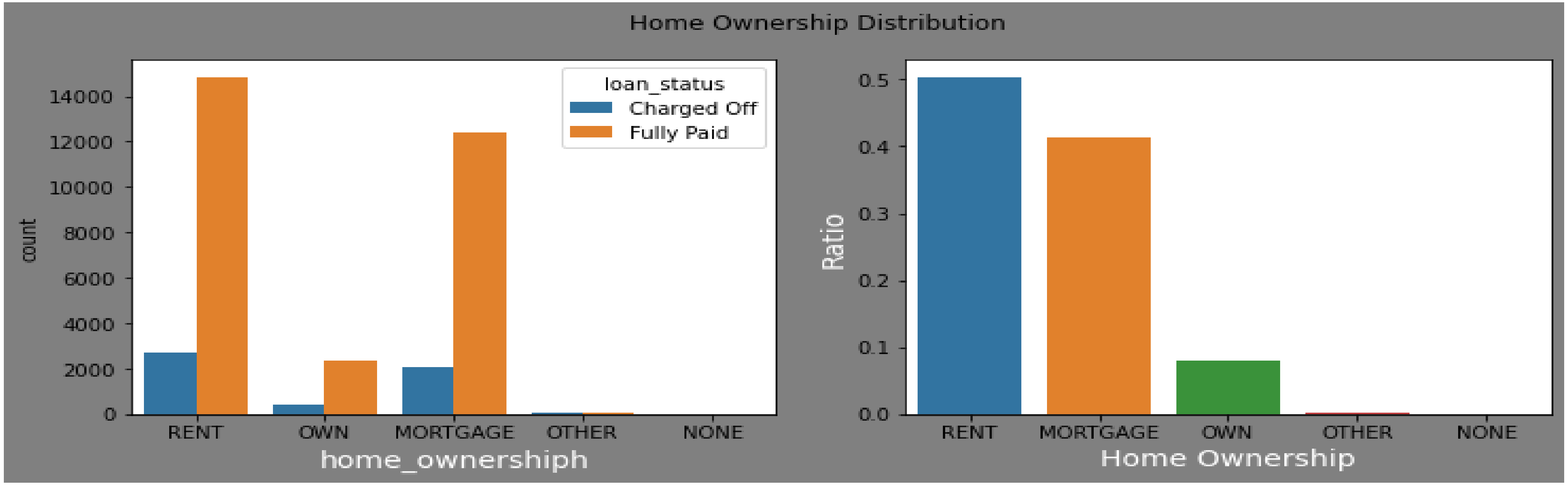


We can see for default loan Derogatory Public and Public Record Bankruptcies is high. This would be good indicator for loan to be default.

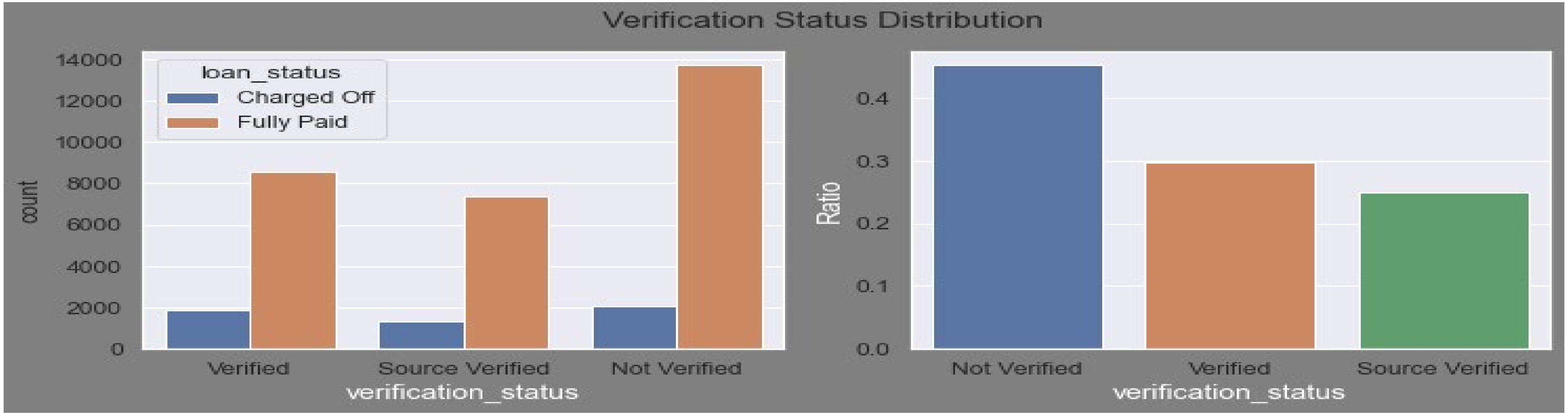


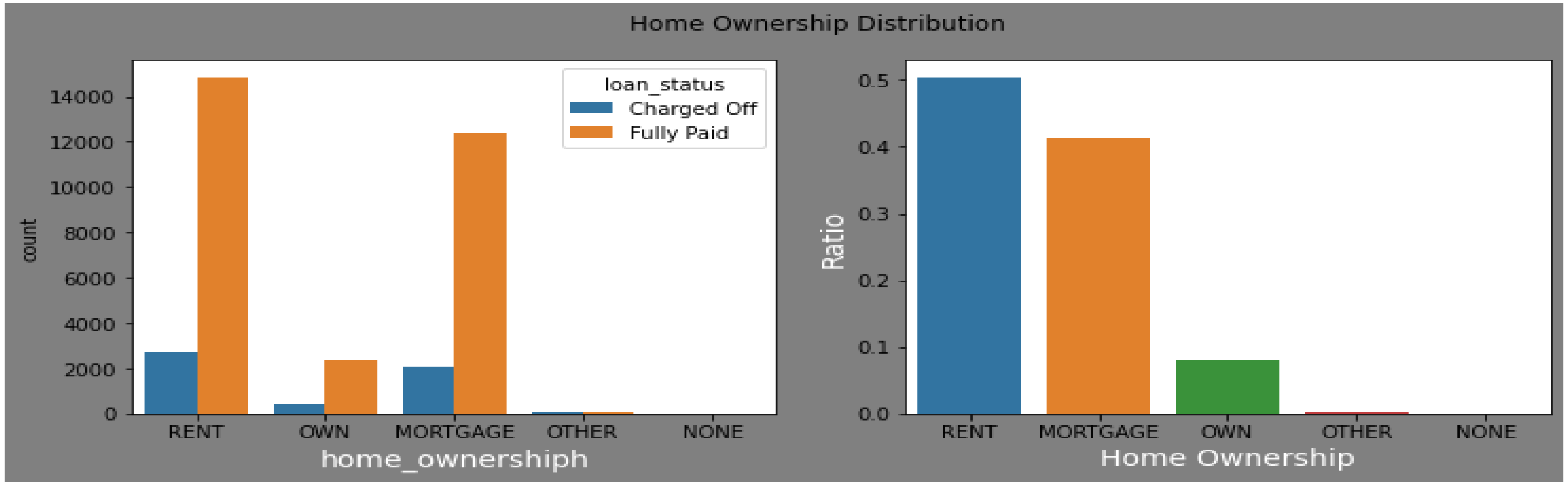


Chances of default is high if applicant took the loan for small_bussiness

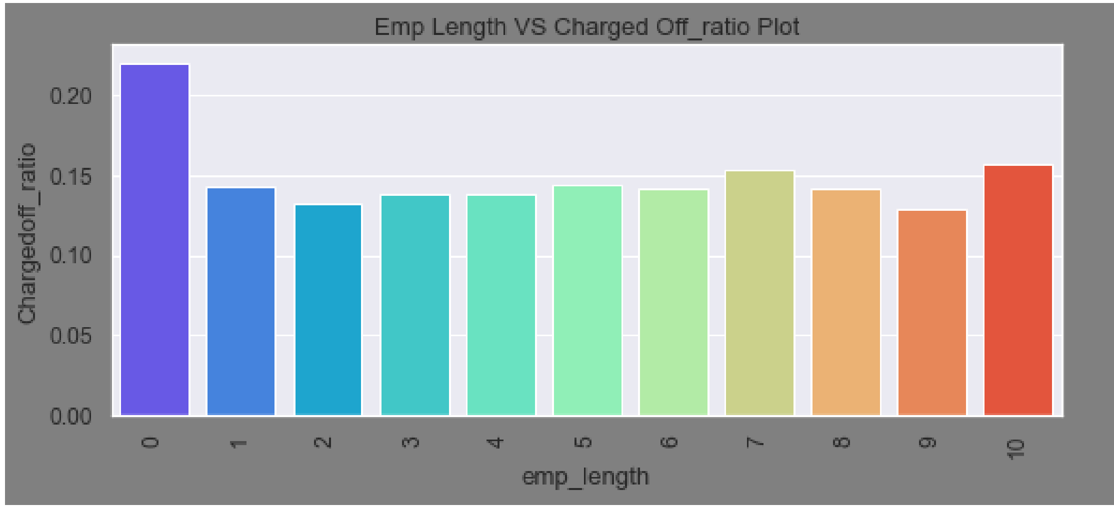


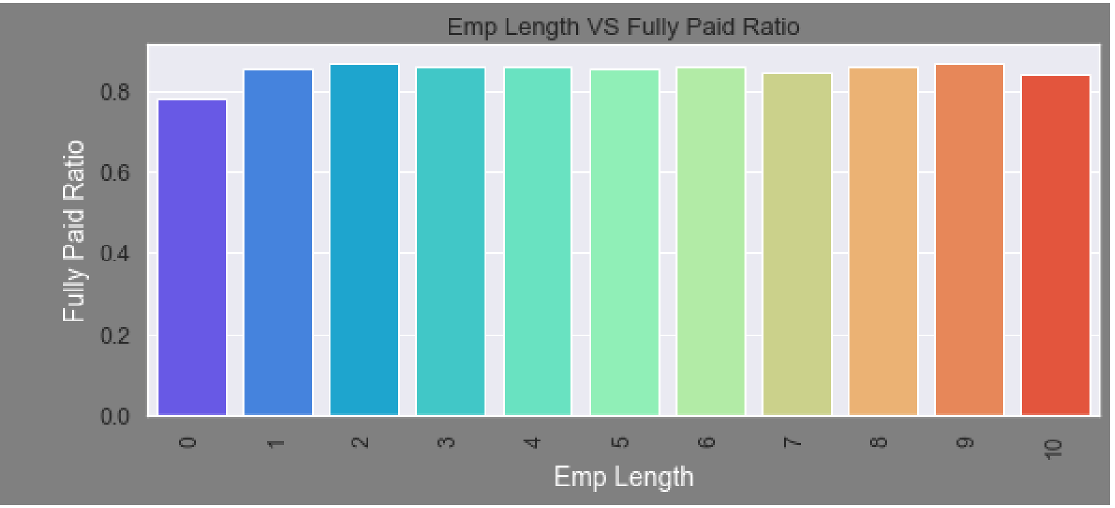
If Grade is low chances of default is high.

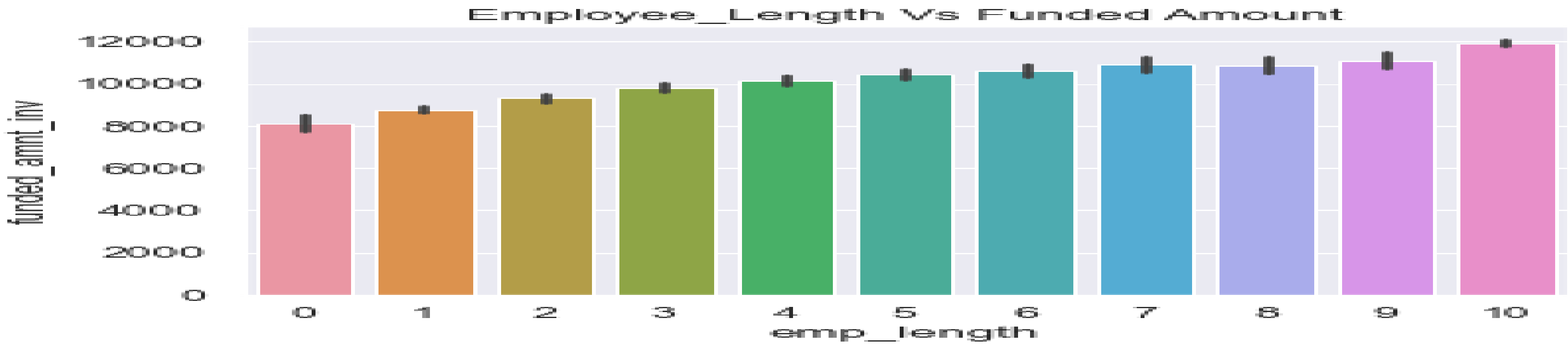




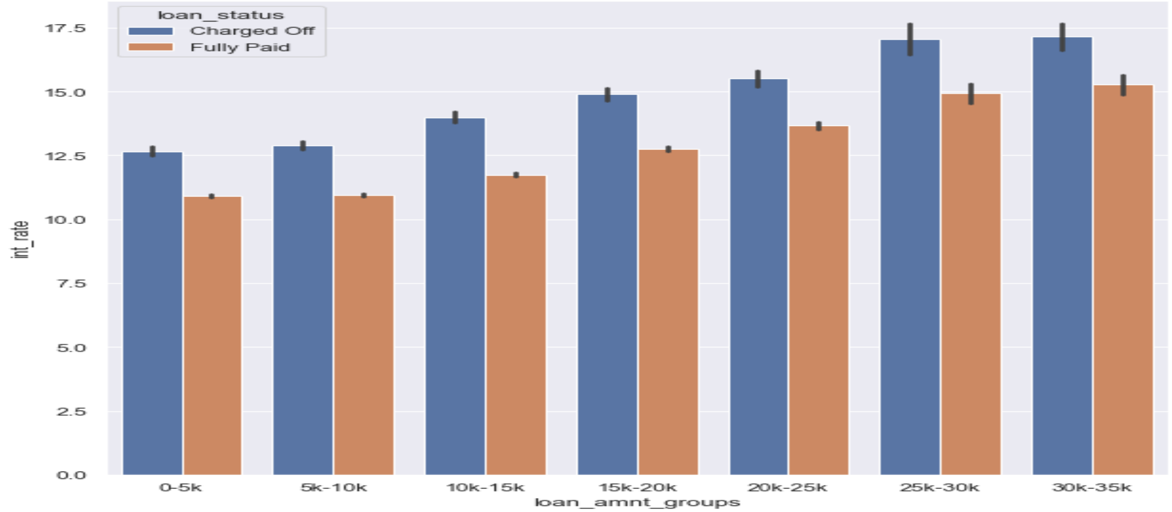
If income is not verified then there in high probability of change off loan. Also applicate having home ownership as a Rent and Mortgage is also

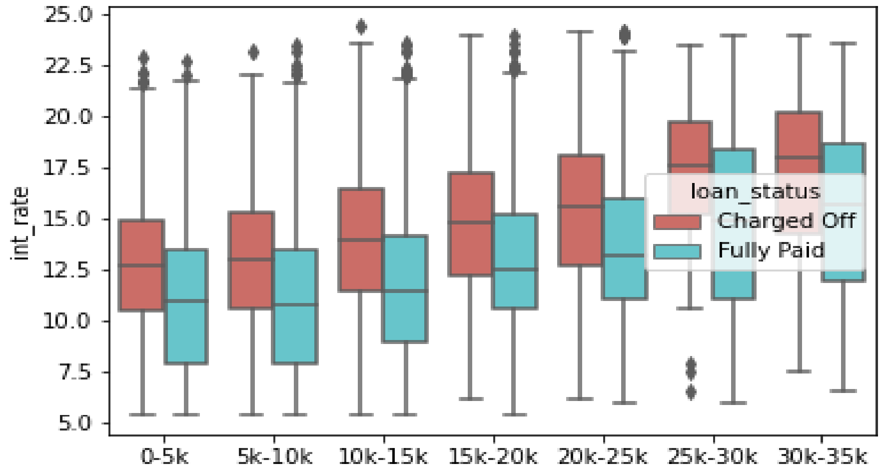


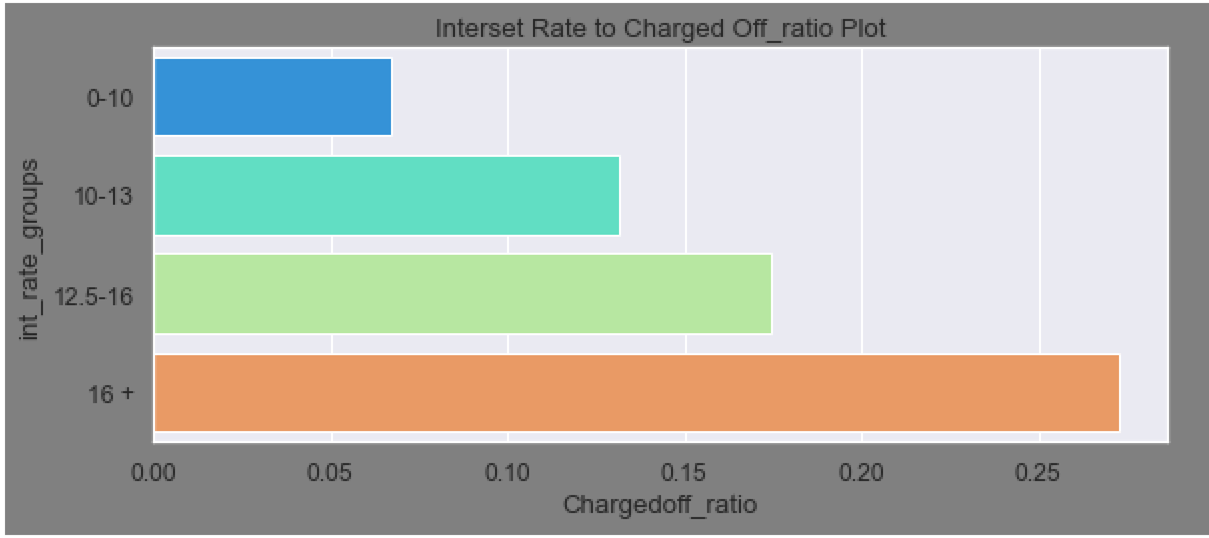


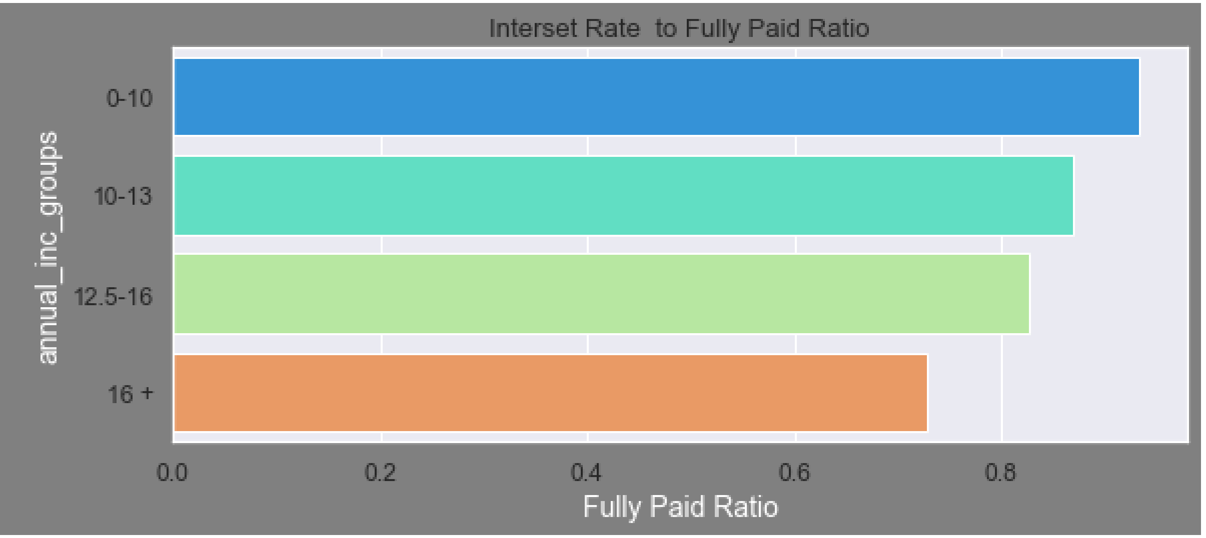


Chargeoff ratio is high for employee length <1 year and funded amount also corelated to emp_length

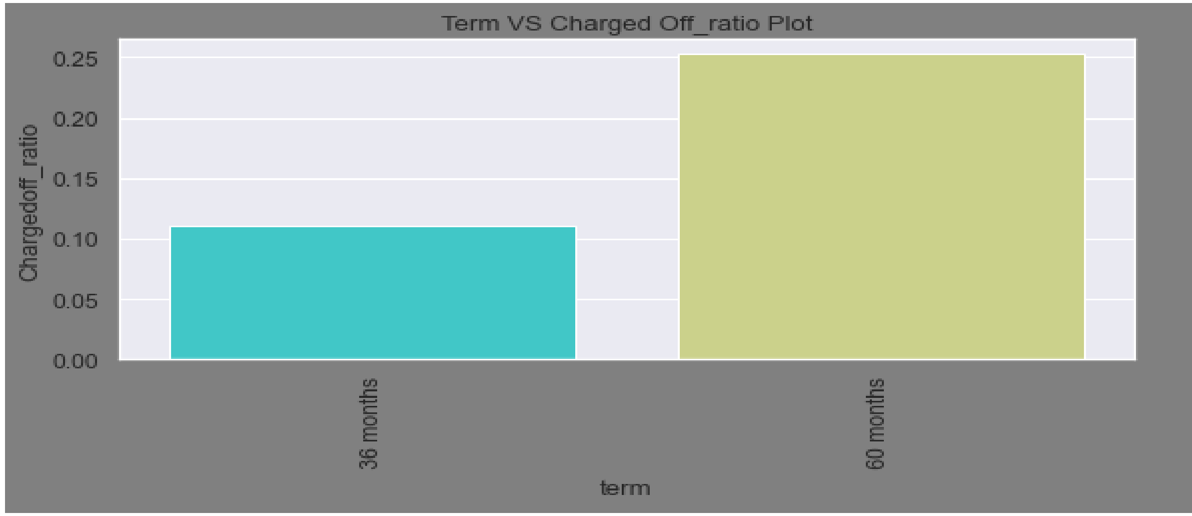


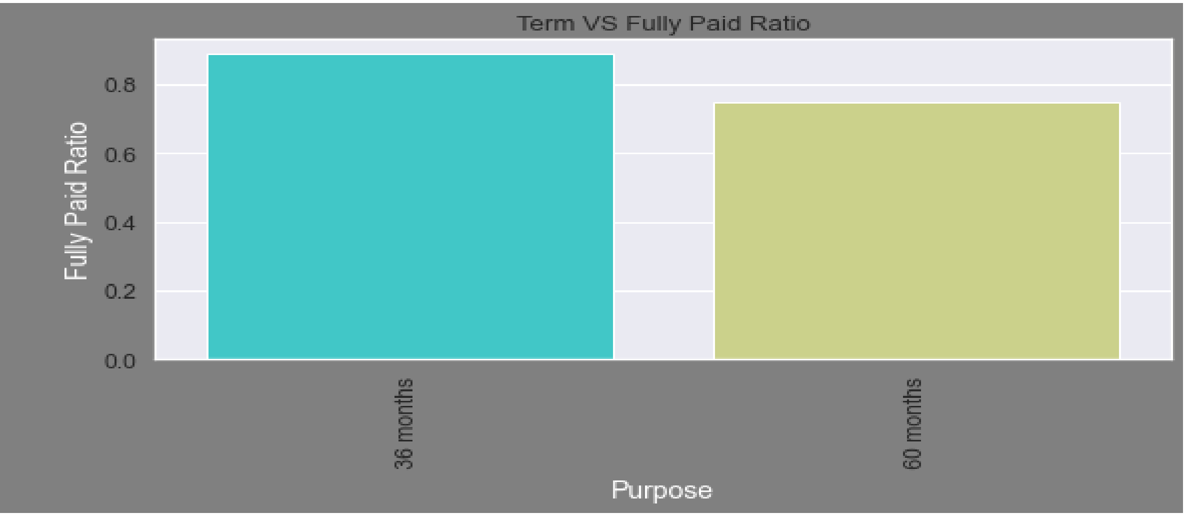


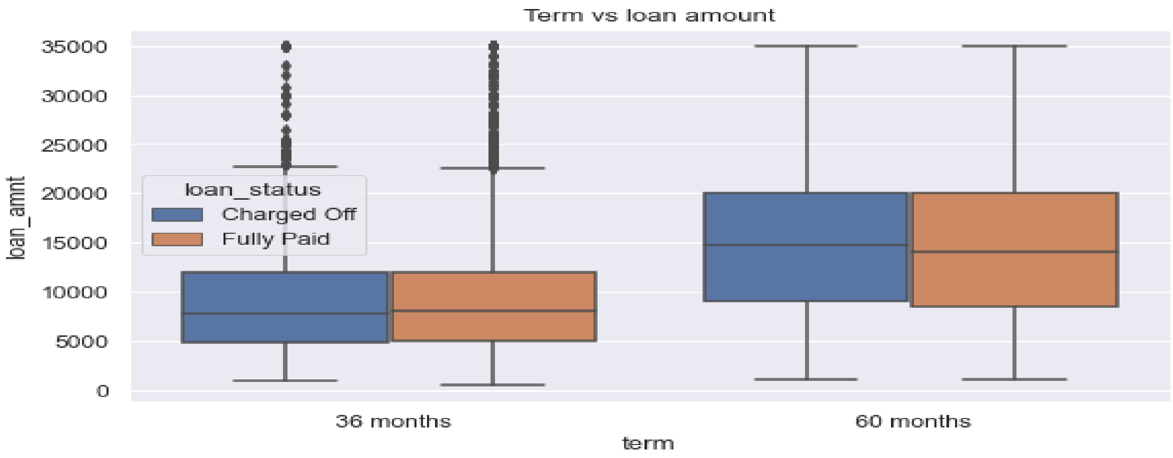


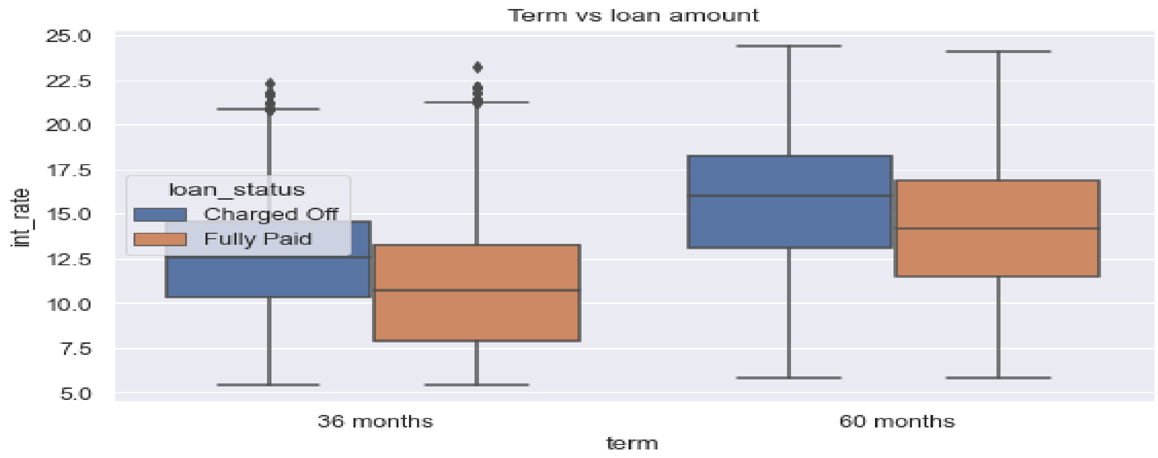


Main reason for default is high interest rate as we can see for Full Paid loans majority of loans fall under 0-10% and for charge off loans interest rate is >16%.

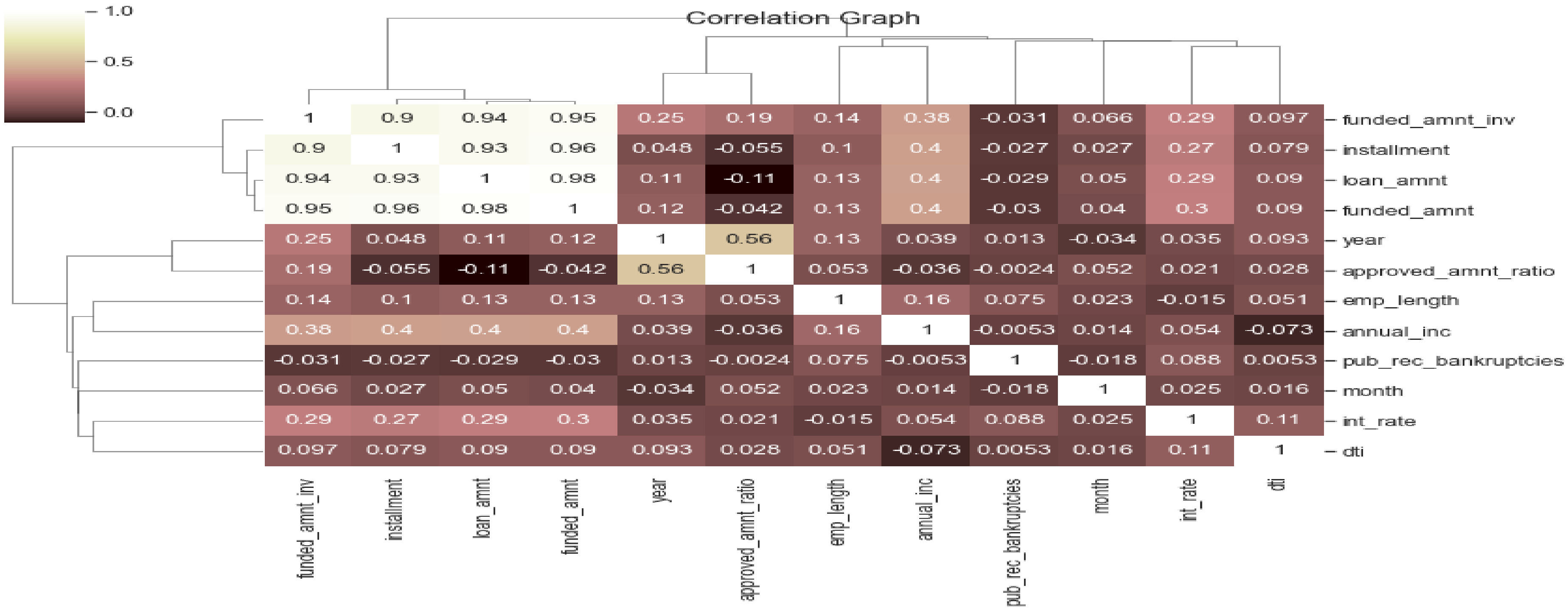




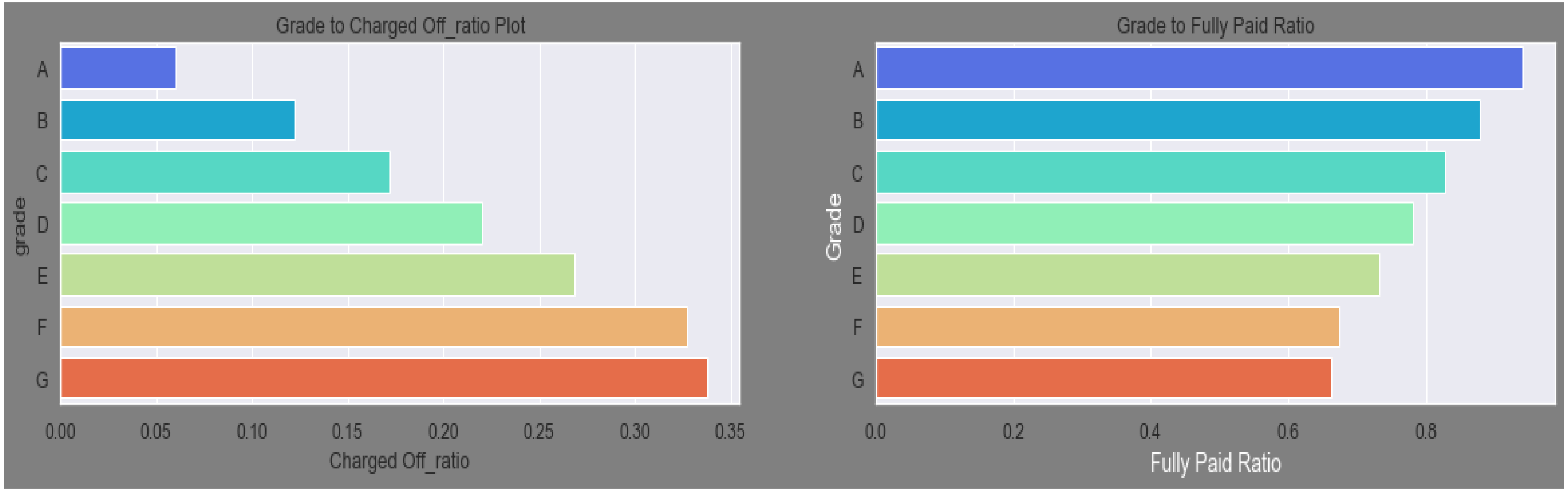


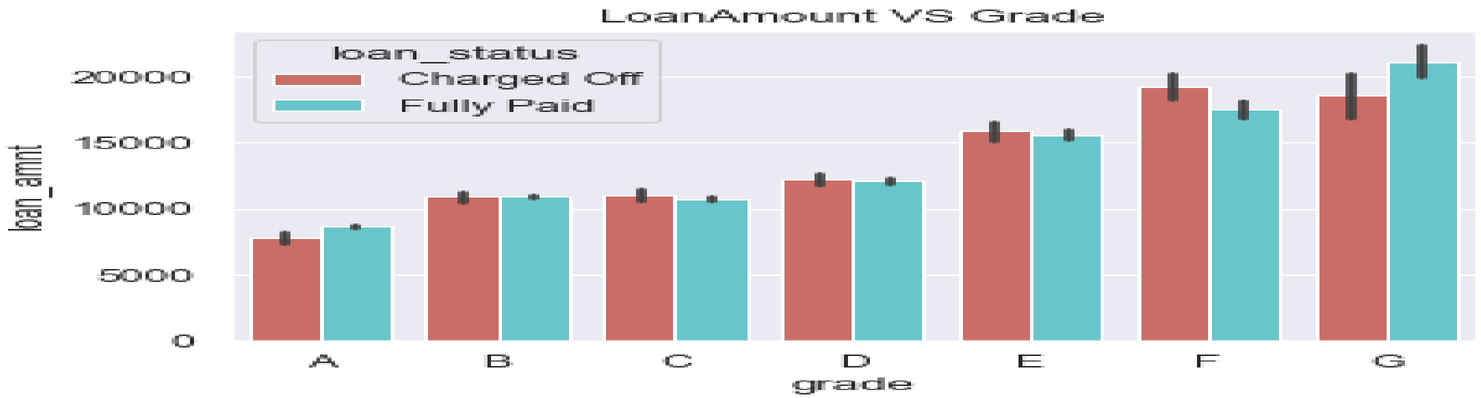


For 60 months term chances of charge off is high because the loan amount and interest rate is high for this term.

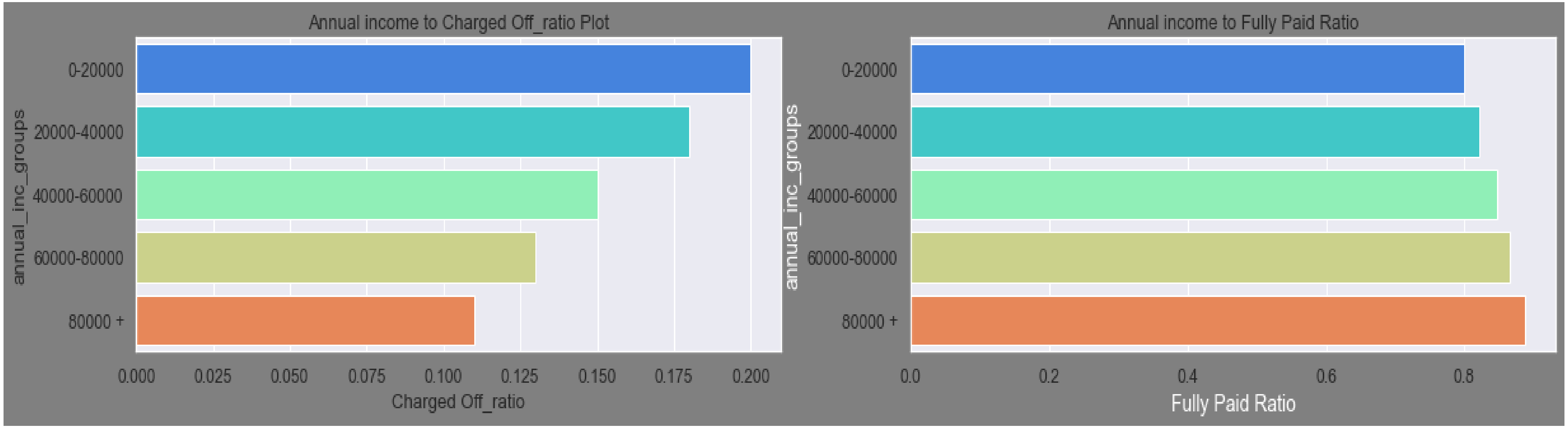


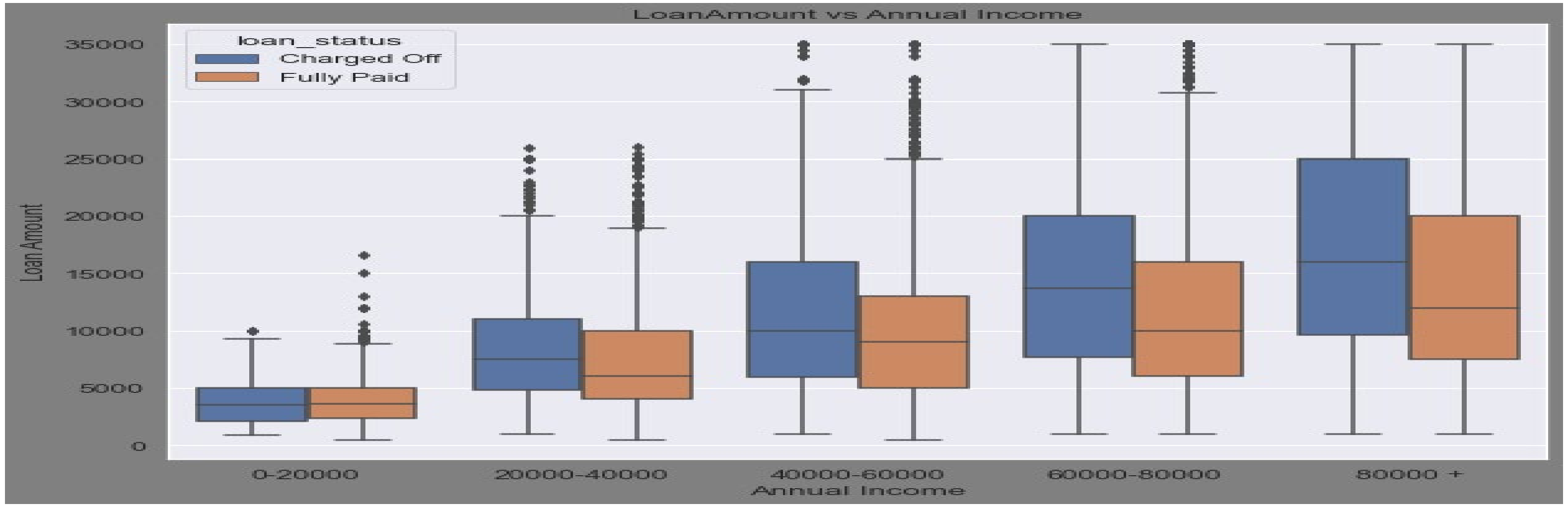
1. All three amounts (Loan amount, investor amount, funding amount) are strongly correlated.
2. As we can see Annual income and DTI are negatively correlated with each other. That means when Annual income is low DTI is high & vice versa.
3. there are positive correlation between annual income and employment years. As employment years increases annual income increases





As grade increases from A-G there is more chances of loan to be default. Which is opposite of Fully paid loans. We can see for Grade A there is more chances of loan to be paid.
And Loan Amount is to high if we move from grade A to G
So if loan amount is to high and grade is G there is high probability of default.





Annual income for charged off Applicate is low as compared to fully paid loan.
Annual income and loan amount is correlated as Annual amount increase loan amount increases.

Recommendation :-

- Applicate with low income group and applied for High Loan Amount at High Interest rate for 60 months tenure.

- Derogatory Public and Public Record Bankruptcies variables are good Metrix to decide loan to be defaulted. If value is >0 there is more chances to be default.
- Using Grade we can identify charged off loans as Grade . As grade increases from A-G there is more chances of loan to be default.
- To avoid default company should verified the income of applicant. Not verified customer have high chances to default.