LENDING CLUB CASE STUDY

GROUP MEMBERS:-

1. Nitesh Kumar Nishant
2. Sanjeev Prasad

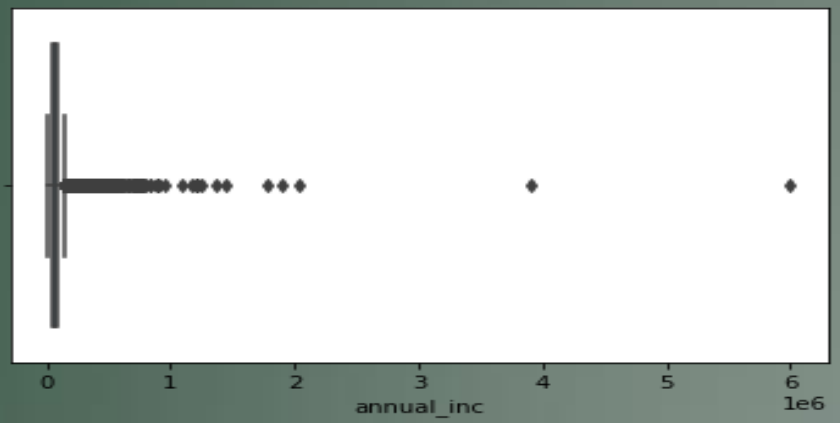
Brief Introduction

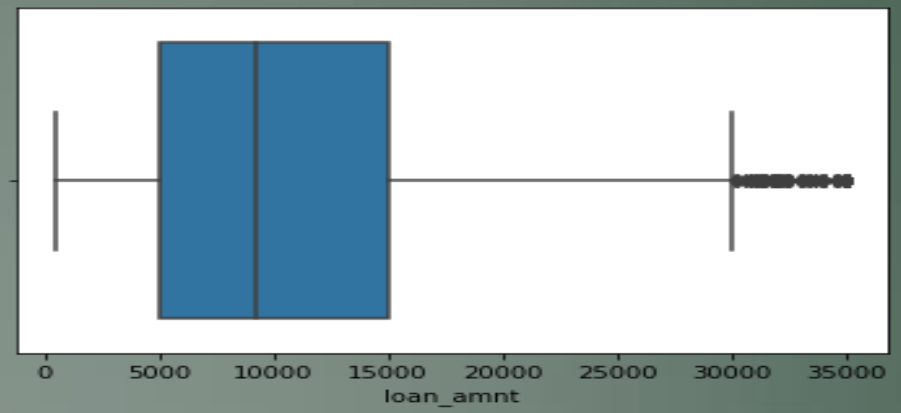
- Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and other lending services.
- Borrowers can easily access lower interest rate loans through a fast online interface.
- The aim of analysis is to find out the information about past loan applicants and find whether they ‘defaulted’ or not.

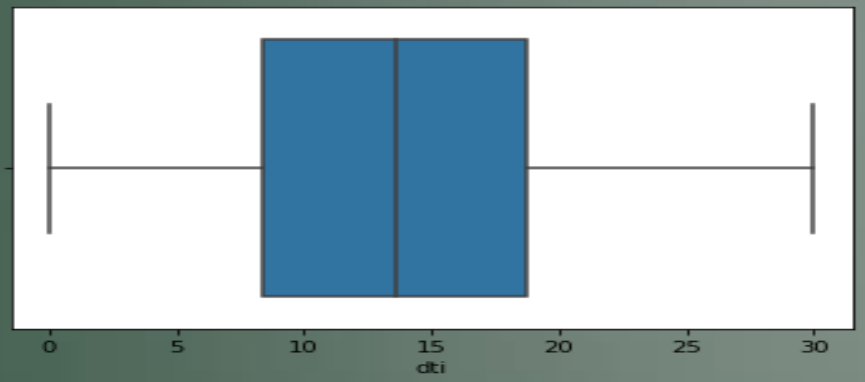
Method to be followed

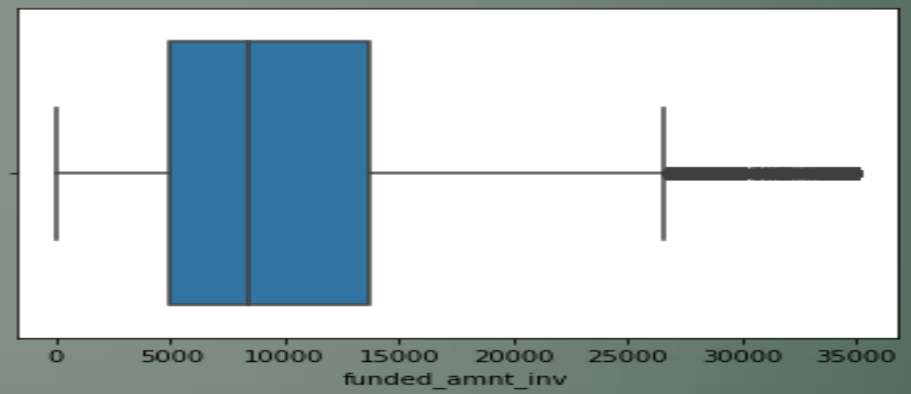
Data Cleaning	• Removing the null valued columns and unnecessary variables • Then we will check null value % and remove the respective rows
Understanding the data	• To have better understanding of columns we will refer data dictionary which have been provide to us
Univariate Analysis	• We will be analyzing each column and plotting required graph. • Distribution of each column
Bivariate Analysis	• Analyzing the two variables at same to get better understanding loan status • We will try to fetch as more information as we can
Segmented Analysis	• Analyzing the continuous data column with respect to the categorical column
Recommendation	• After Doing above operation, we will finally come to our objective

Finding Outliers using Box plot





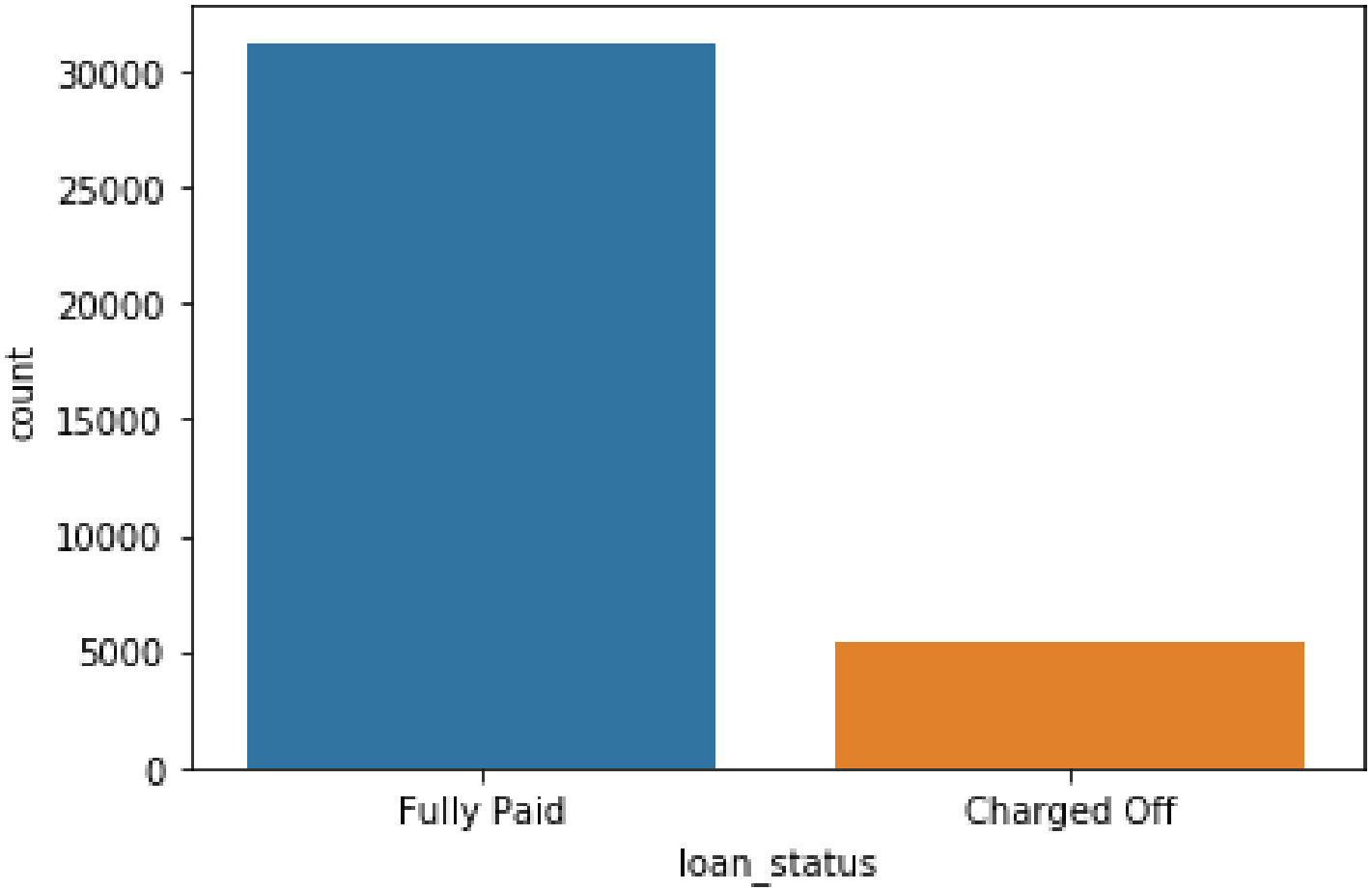




Removing outliers which are showing on the graph

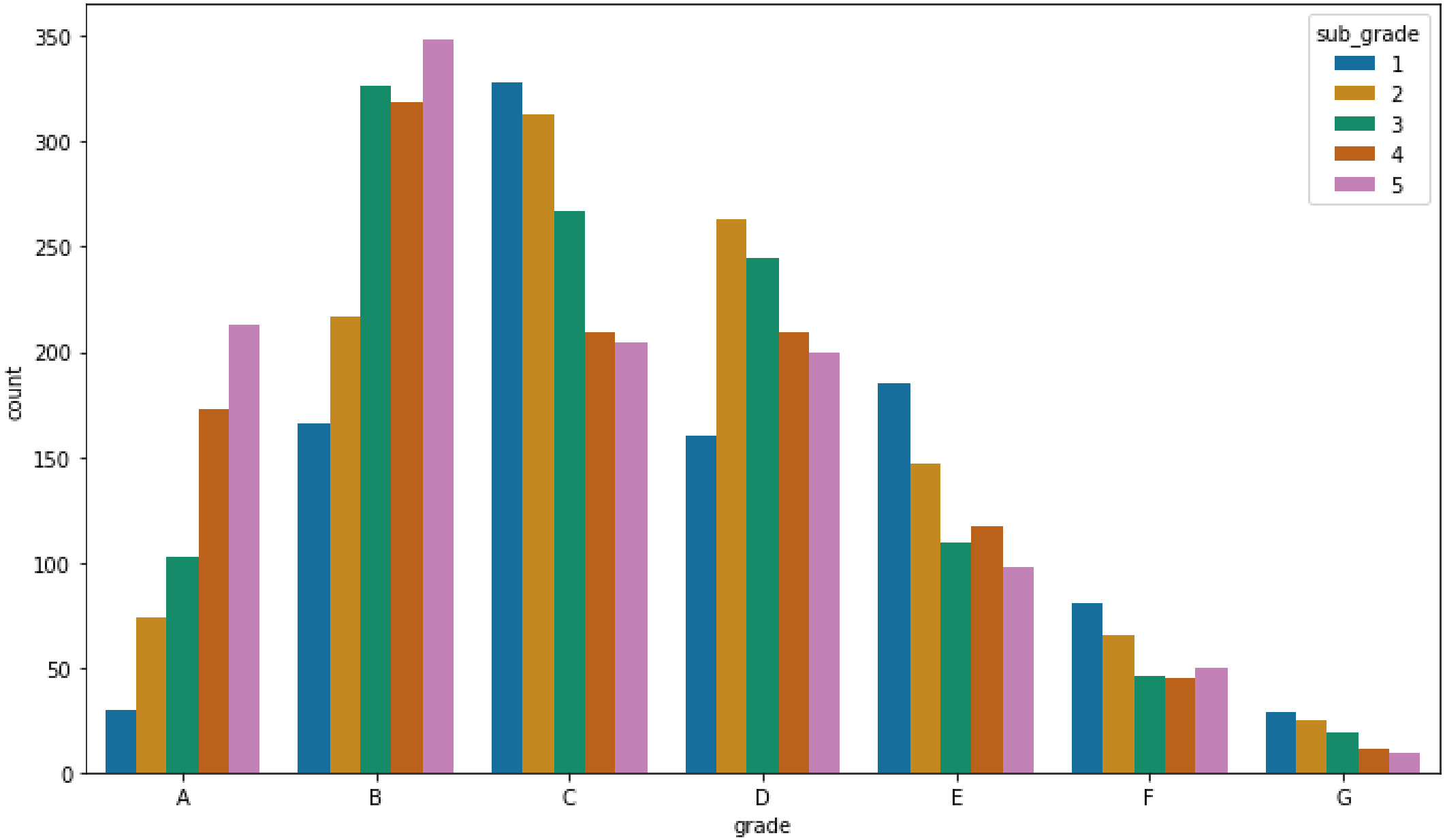
- ***The values after 95 percentile seems to be disconnected from the general distribution and also there is huge increase in the value for small quintile variation.***
- ***So, considering threshold for removing outliers as 0.95***

Visualizing Categorical data

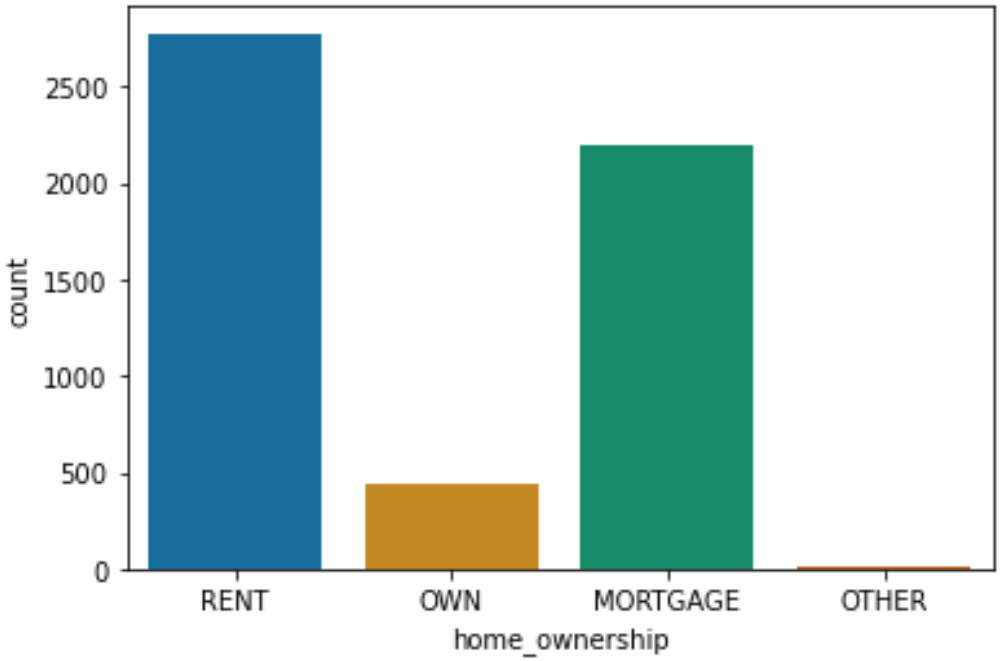


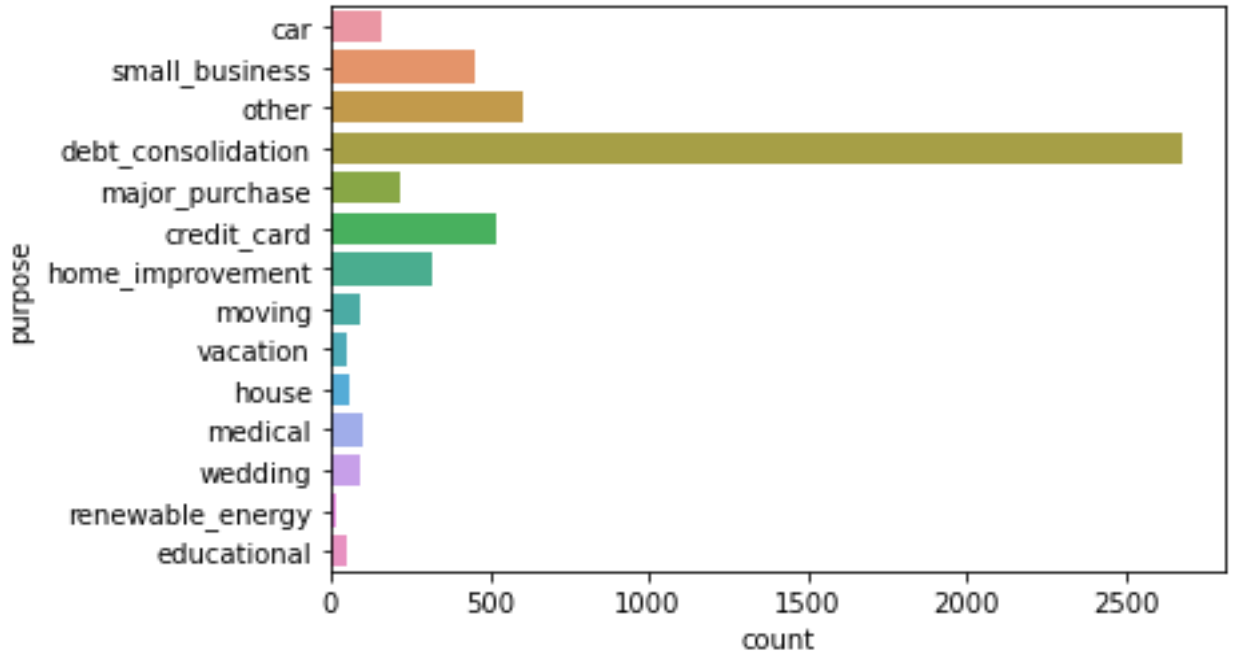
We are analyzing and visualizing only the defaulter data. So subsetting the data while plotting only for 'Charged Off' loan_status for below plots

Visualizing Categorical data



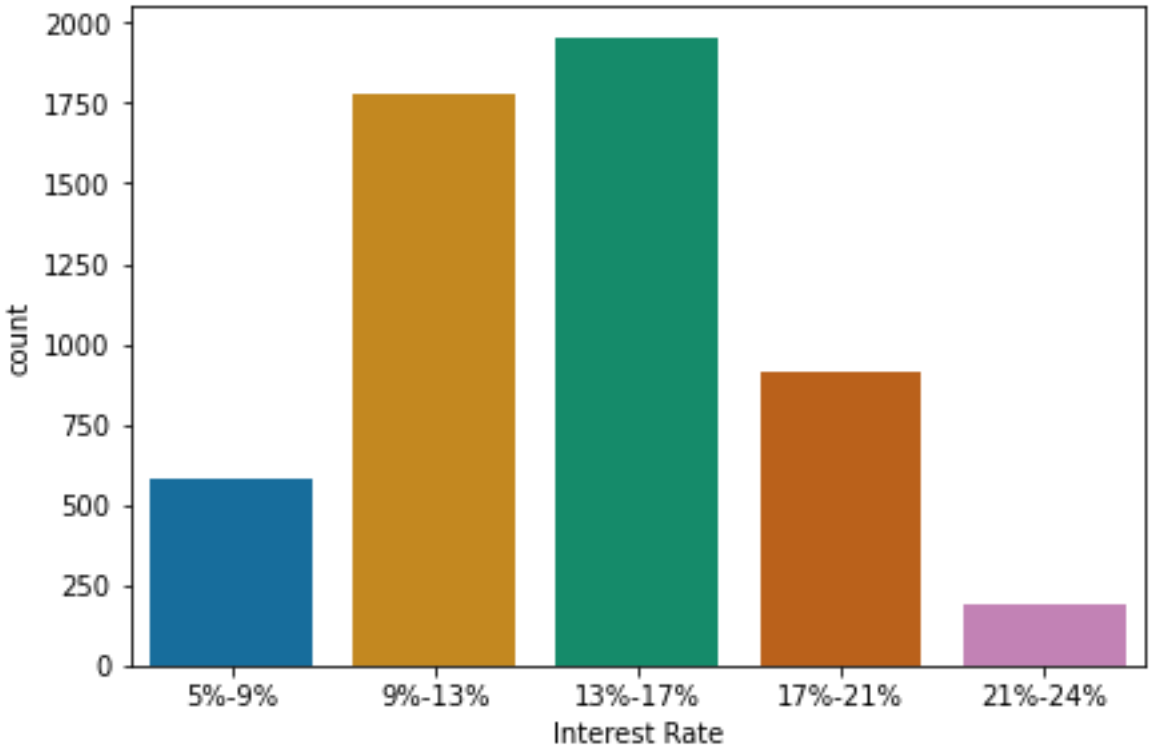
Analyzing home_ownership

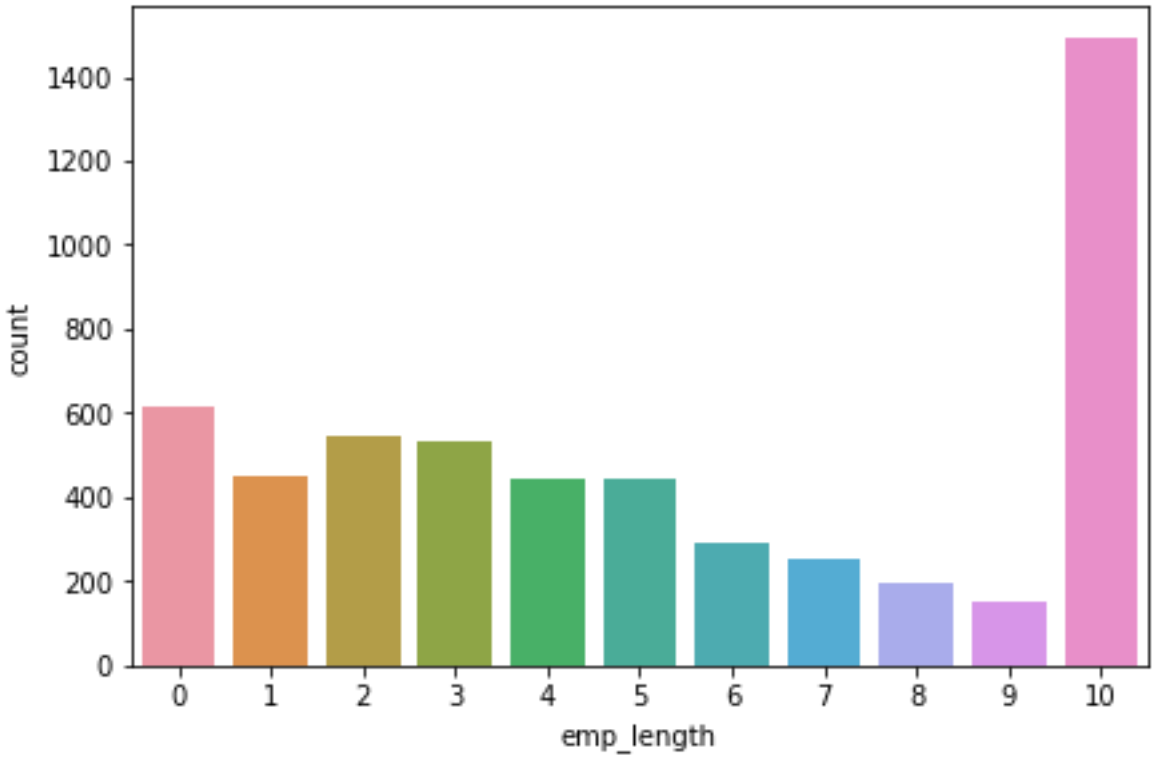




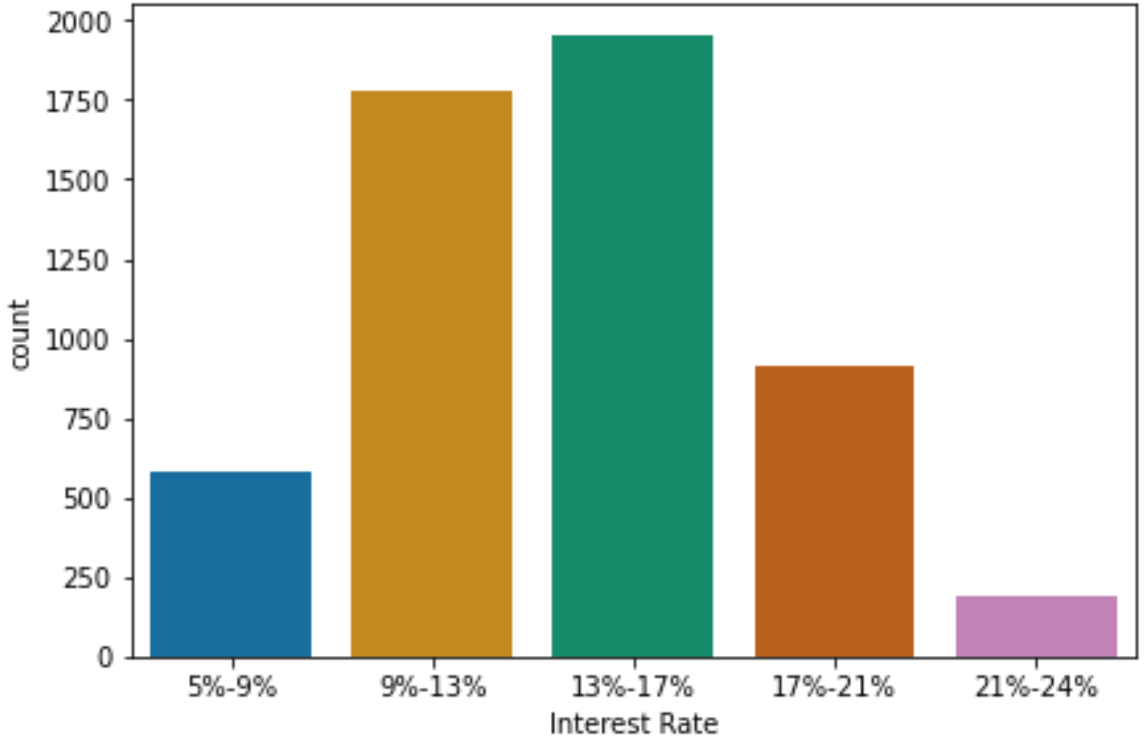
There were few values of "NONE" so we have replaced with others

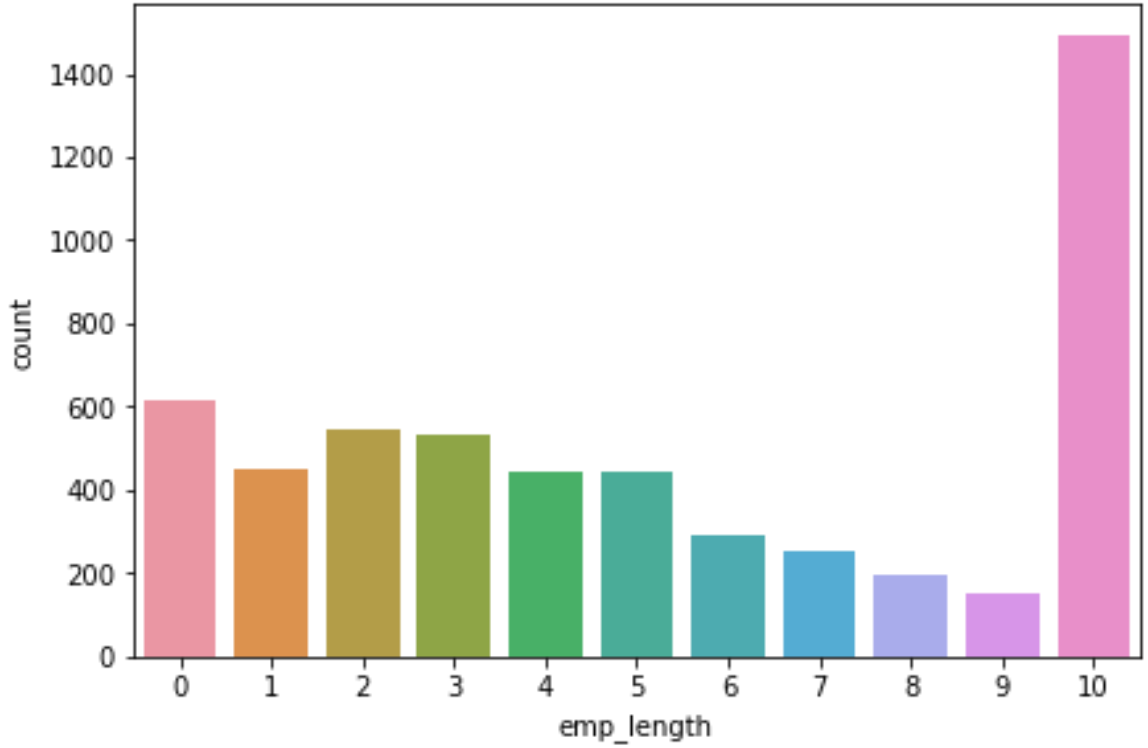
Analyzing rate of interest with respect to the rate of interest of bins



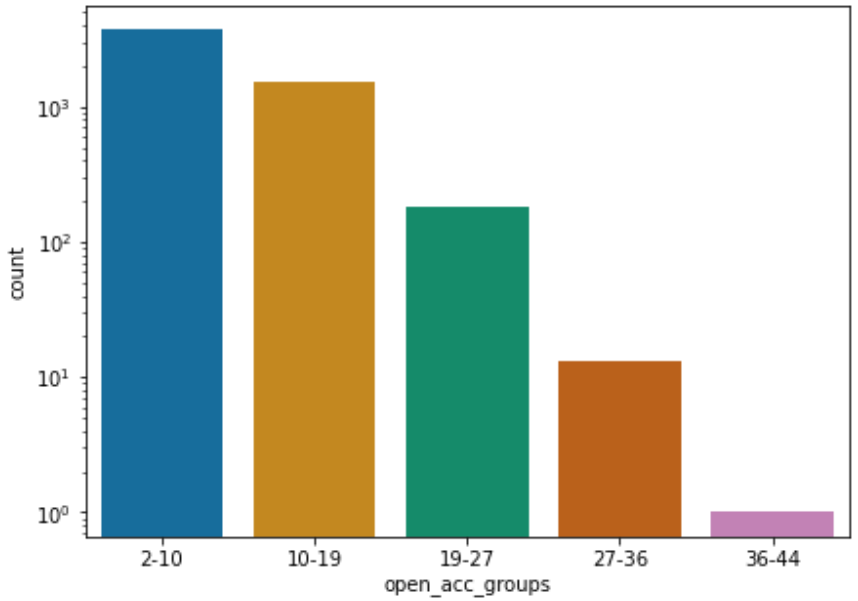


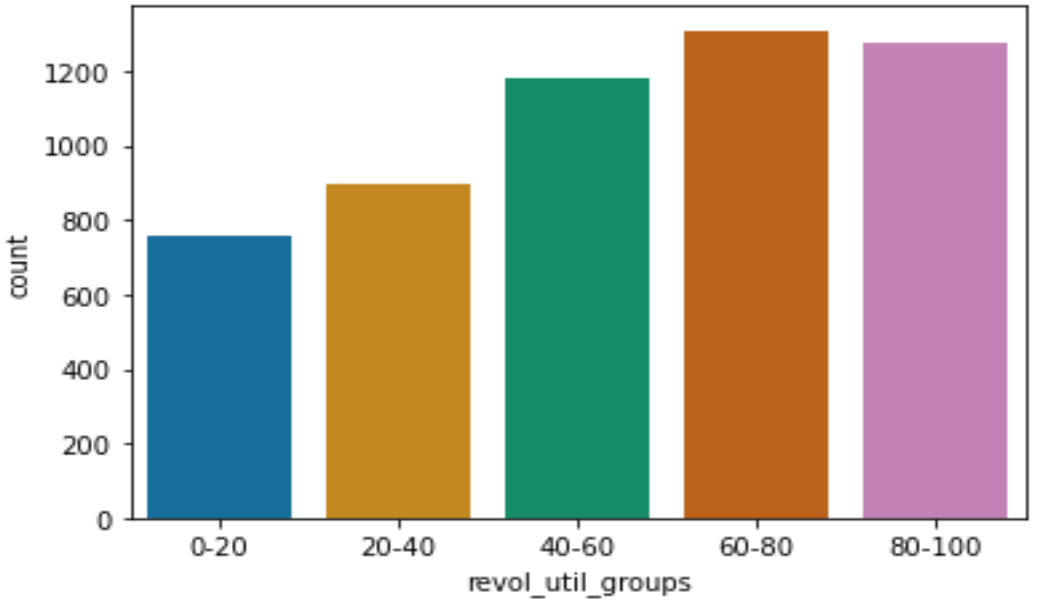
Analyzing columns b/w count & Interest Rate, and count & emp_lengtth

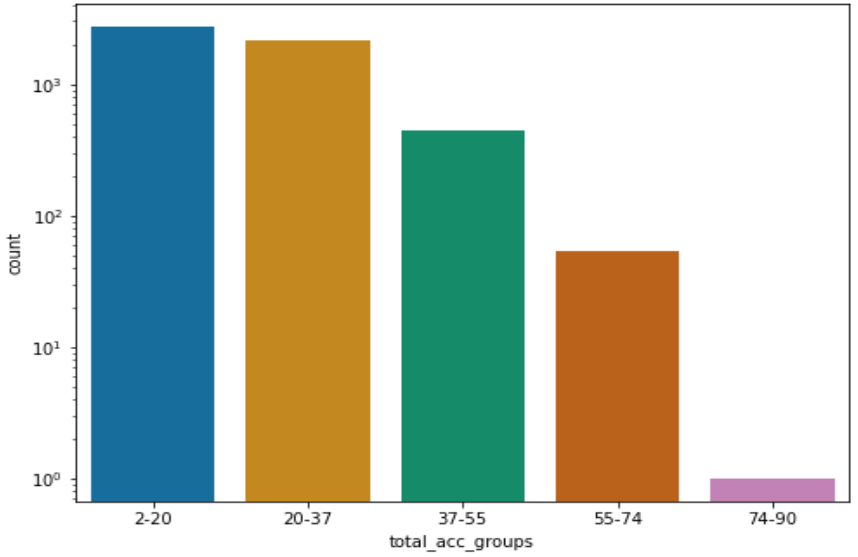


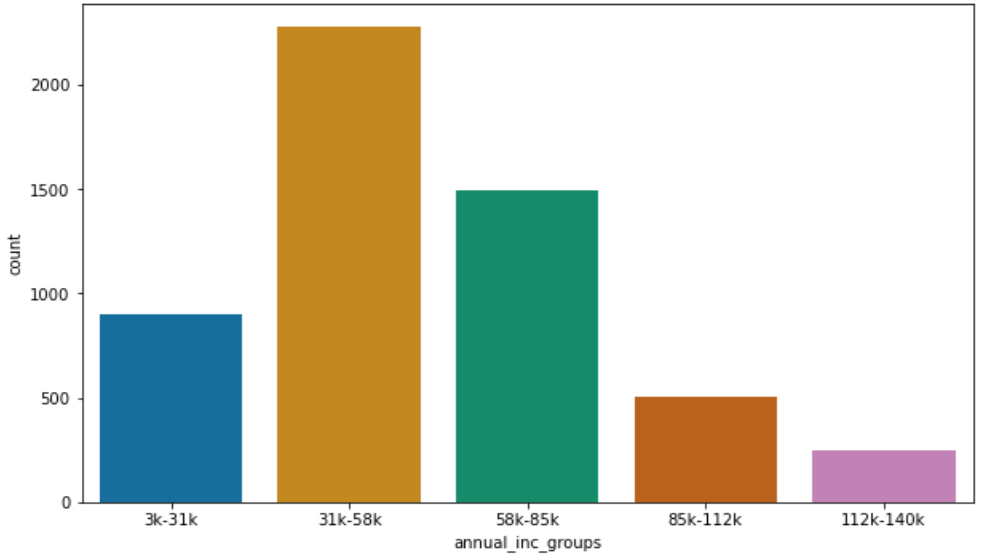


Analyzing columns like open_acc,revol_util,total_acc,annual_inc

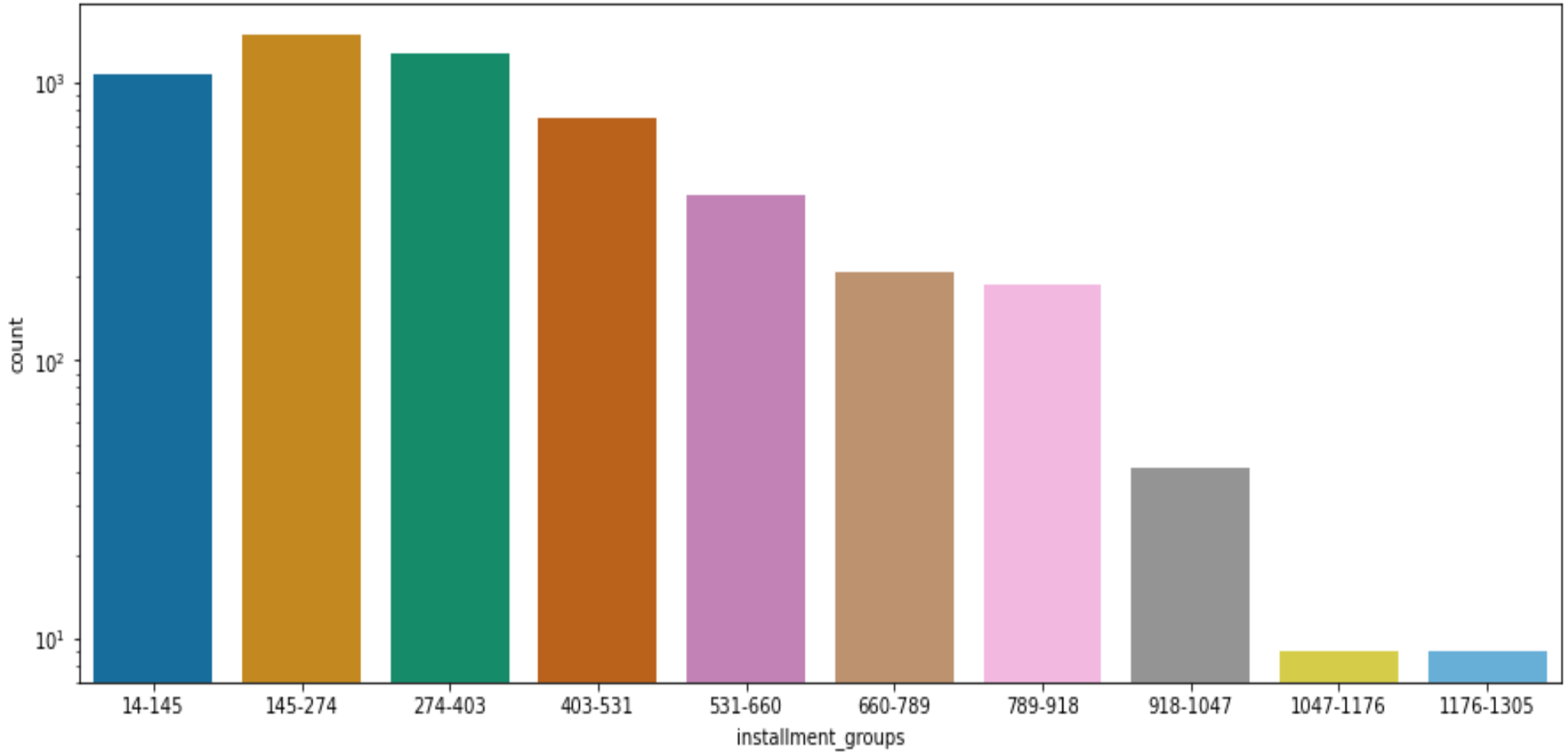






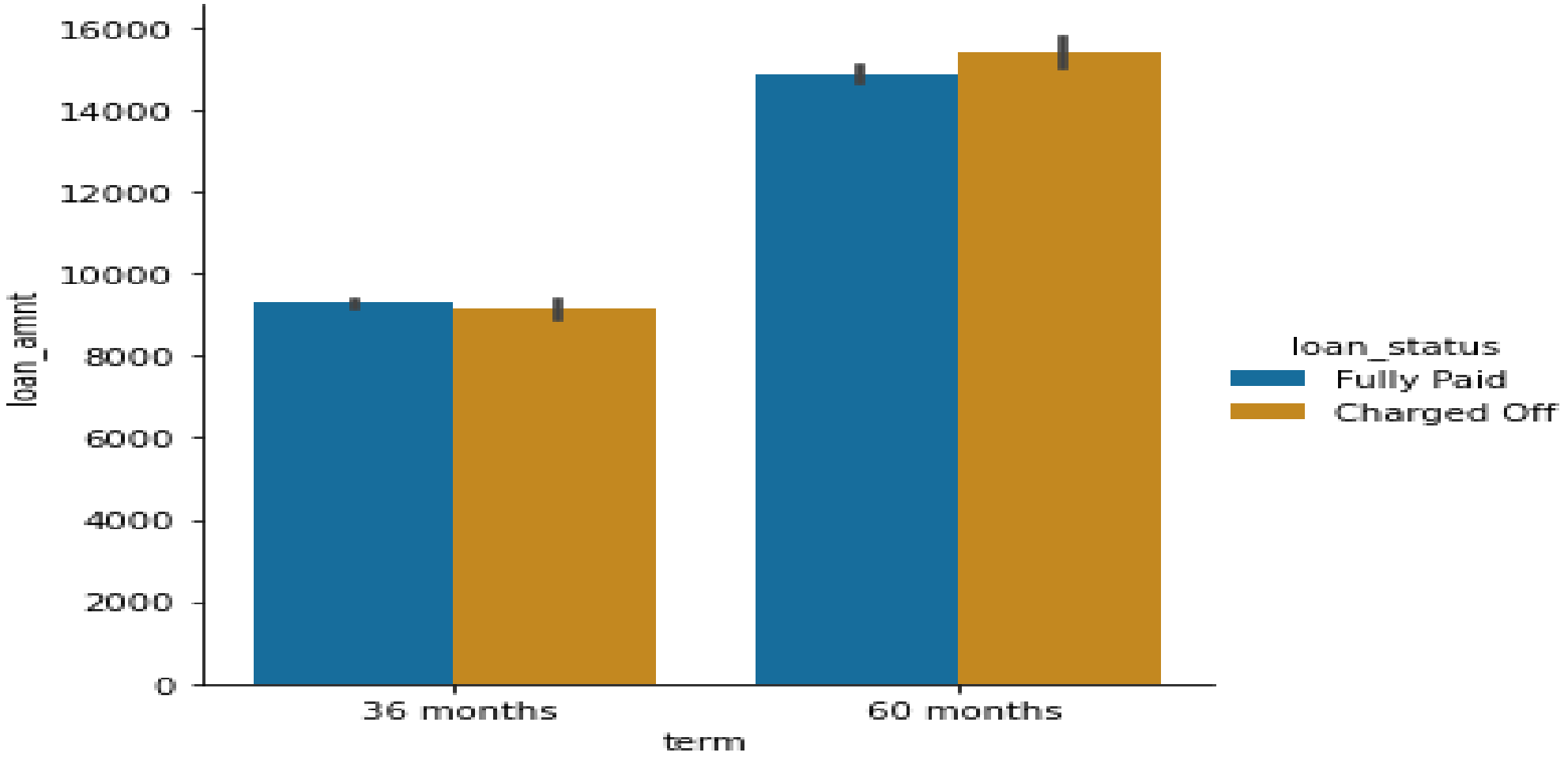


Analyzing columns between Count and Installment_groups



The above chart shows us the probability of getting default.

Analyzing columns loan_amnt between term



Conclusions

- Lending club should reduce the high interest loans for 60 months tenure, they are prone to loan default.
- Grades are good metric for detecting defaulters.
- 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.
- Borrowers whose income is between 70k-80k and received rate of interest 21%-24% ,consider as threat of defaulters.
- Loan status history should be verified before sanction any loan.
- Lending club should avoid granting loan to lower salary employees.
- Financial crisis is also one of the reasons why loan get defaulted , They must have aware from this situation.