Lending Club

Analysis of Lending Club Case study

Objective

The Objective of this activity is to analyse the drivers for loan default and predict defaulting based on historical data. To achieve this, we will be performing Univariate, Bivariate and Multivariate analysis on the given loan dataset.

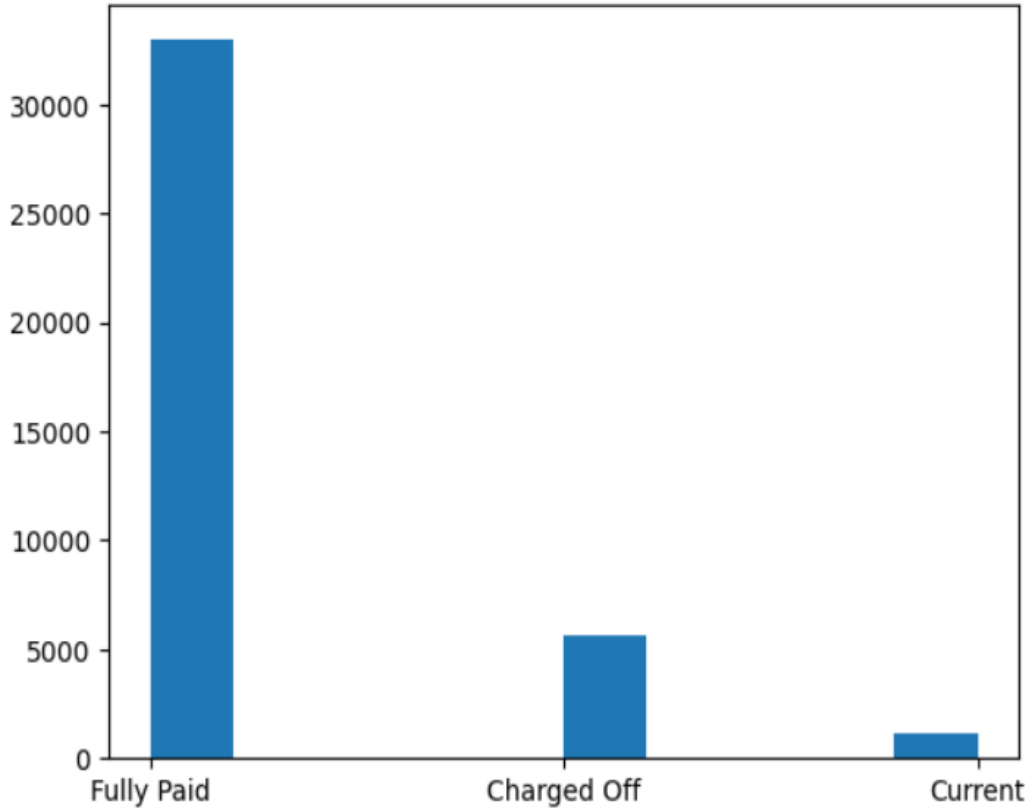
We will try to understand different attributes and their relationship and attribution to loan defaults.

Lastly we will try to list down our observations and recommendations.

16.5

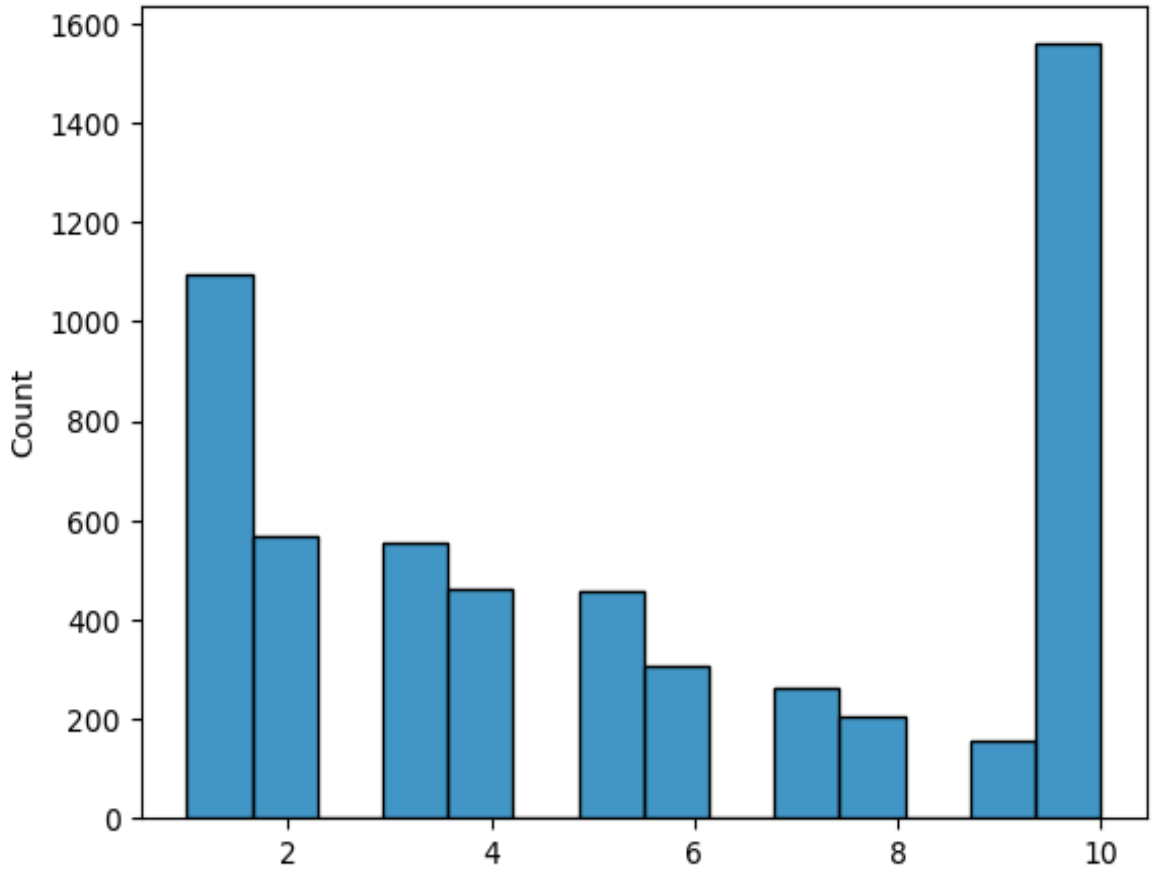
Overall default rate of loans

We see that out of the given dataset, 16.5 % percent of loans have defaulted

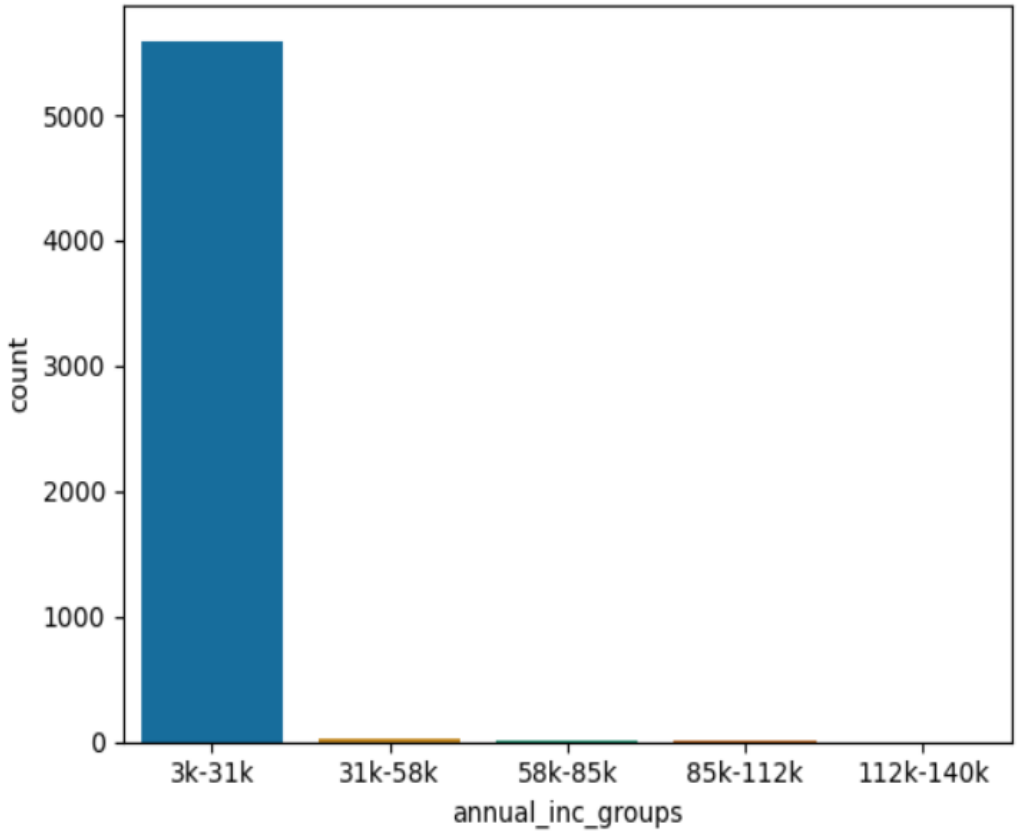


Employment length distribution of defaulters

Most defaulters lie between employment length of below 2 years and 9-10 years



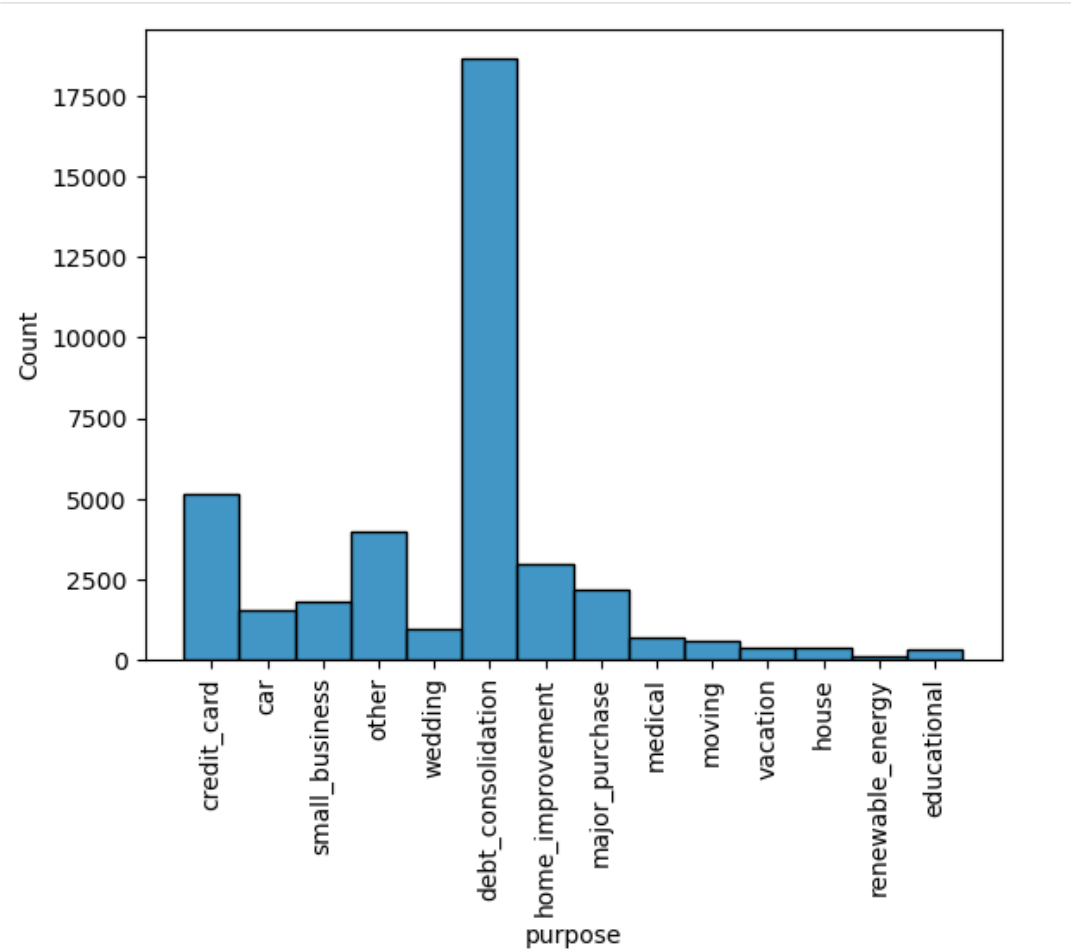
Annual Income distribution for Defaulters



Annual Income group of 3k-31k is having the highest default rate

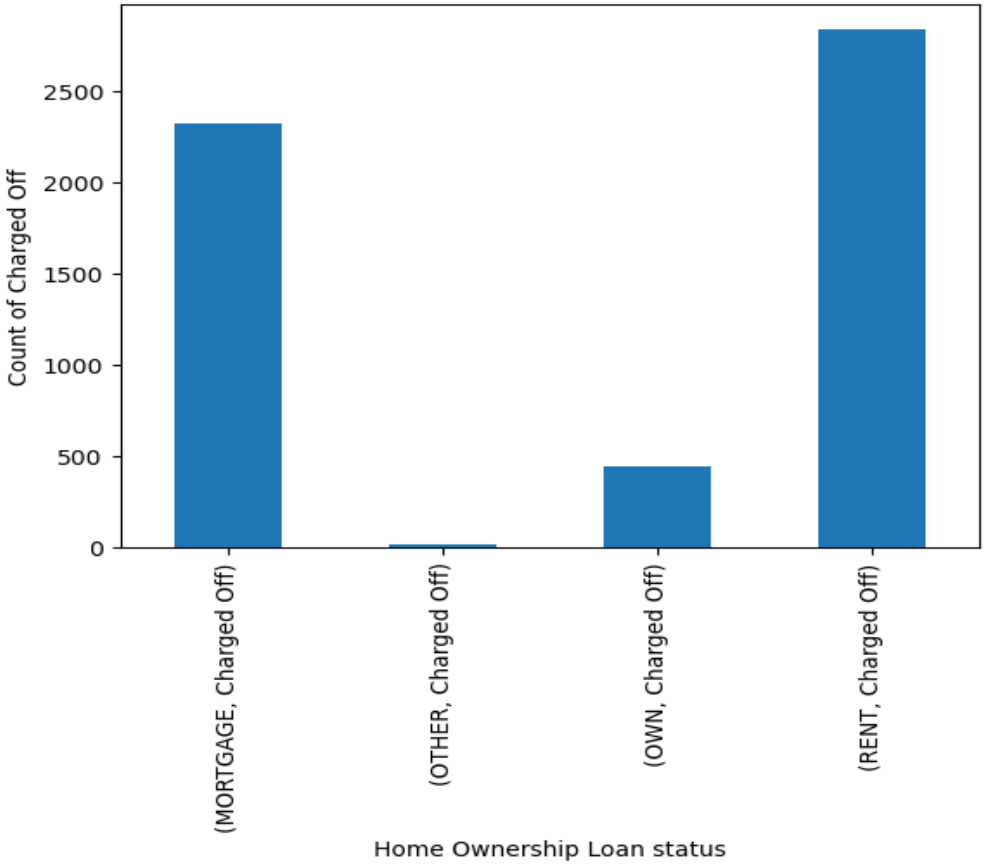
Distribution of loan purpose in the loan dataset

We can see that maximum loans were taken for the purpose “debt consolidation”

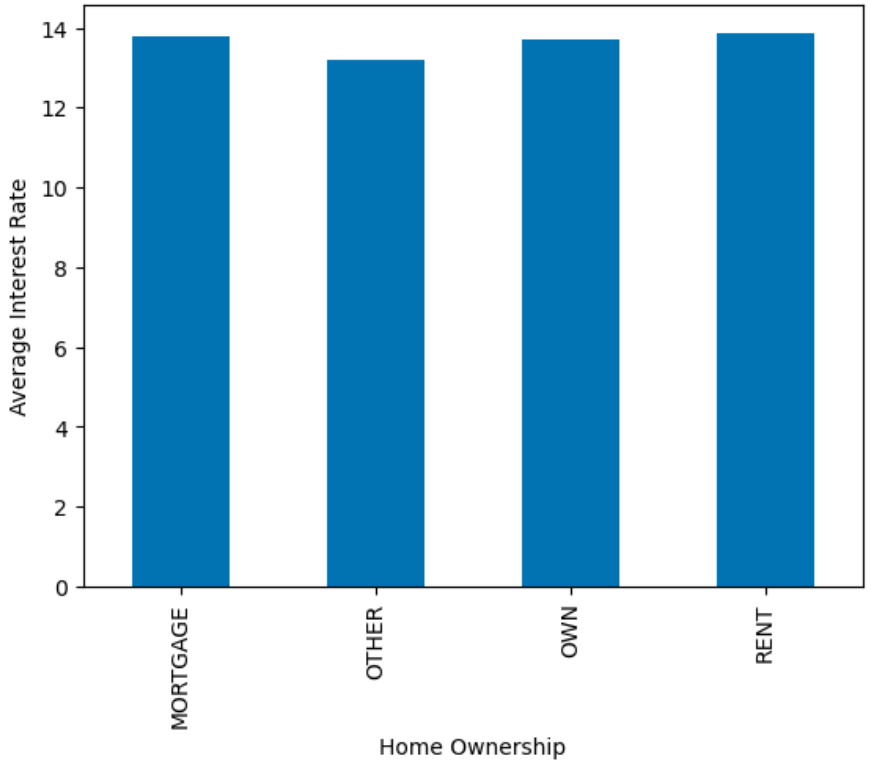


Relationship between house ownership and loan status

RENT home ownership have the highest Charge-off frequency

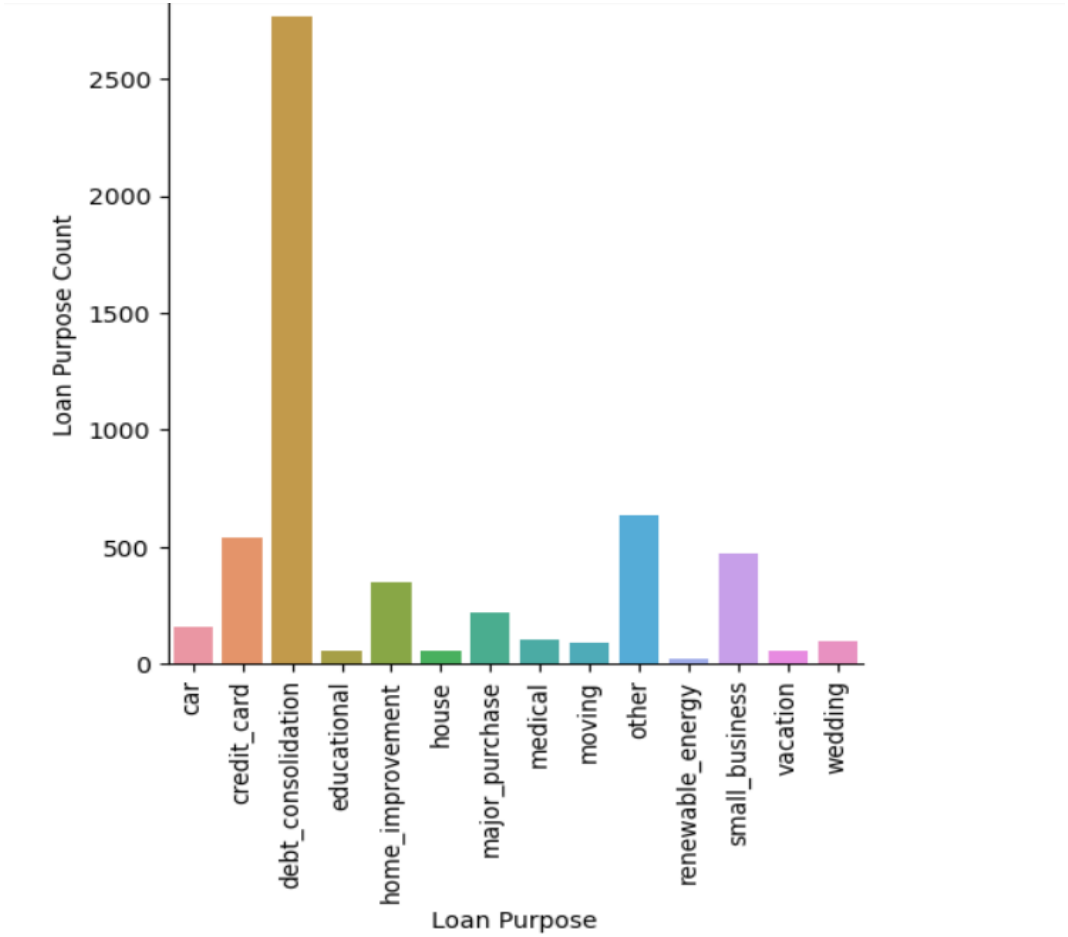


Highest Interest rate is for MORTGAE home ownership

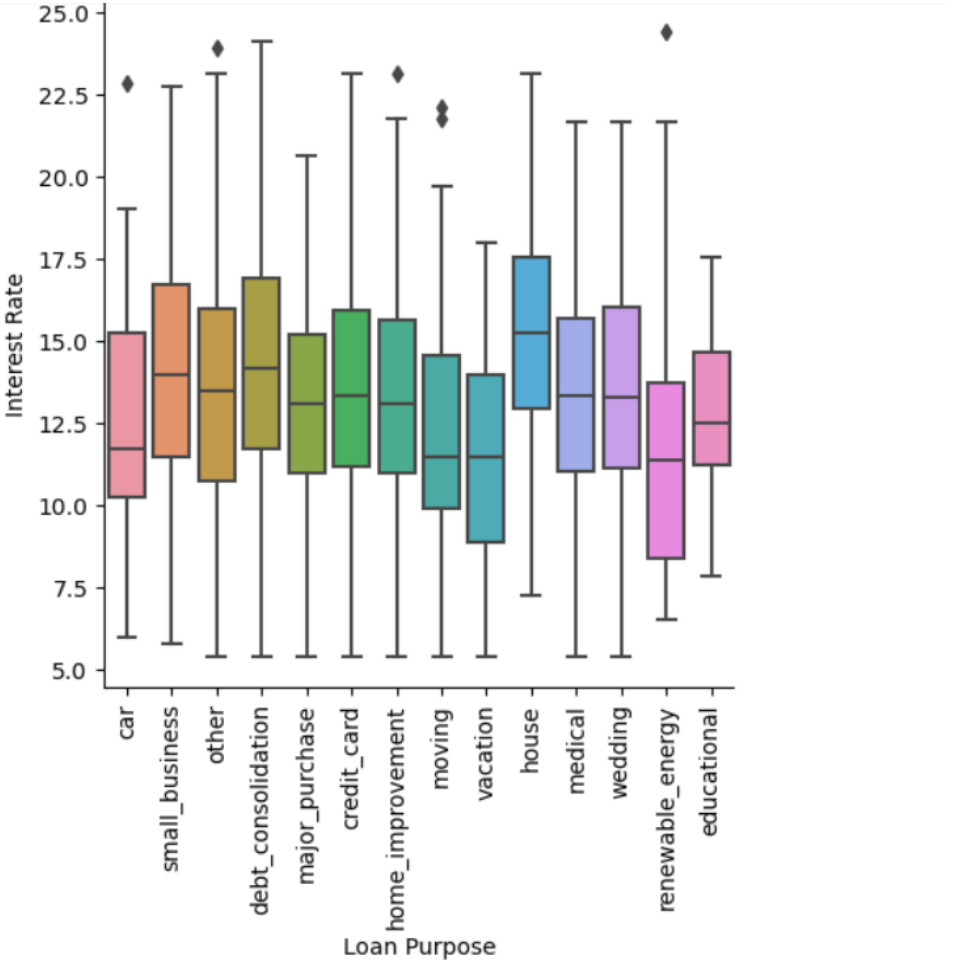


Comparing loan purpose and defaulting rate

Highest number of loans are for Debt Reconsolidation, but loans for house have the lowest defaulter rate

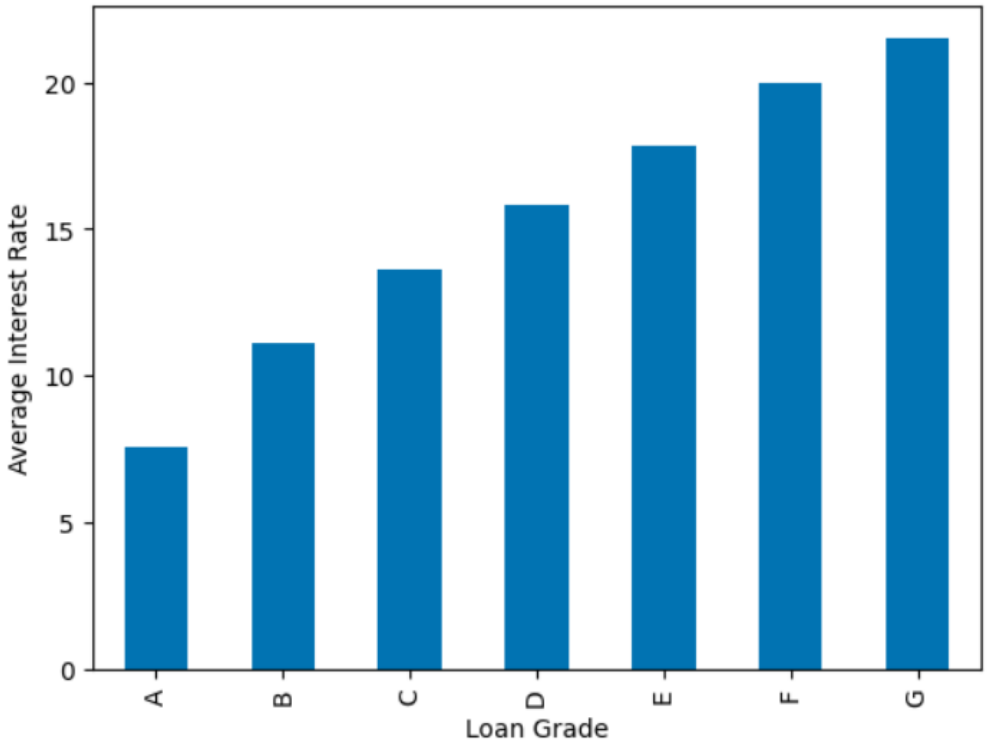


Highest Average Interest Rate is for house and a little lower for Debt Reconsolidation

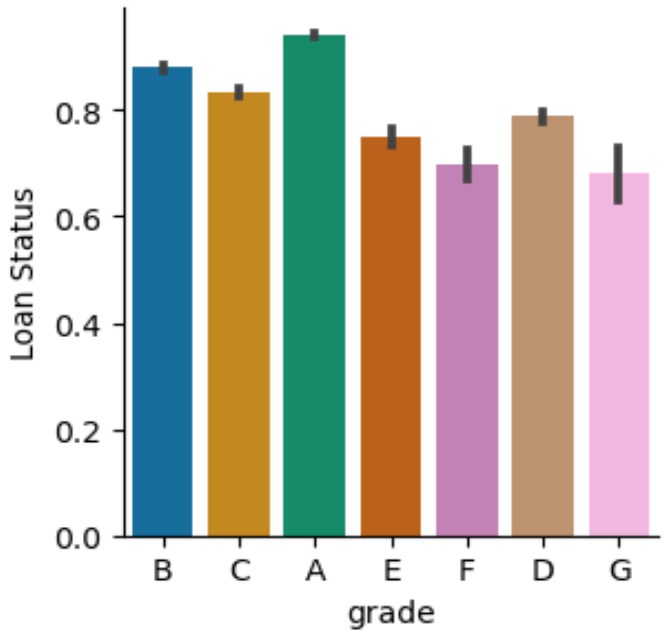


Distribution of Loan Grades and Their average interest rates

We observe that Loan Grade G has the highest Interest Rate.

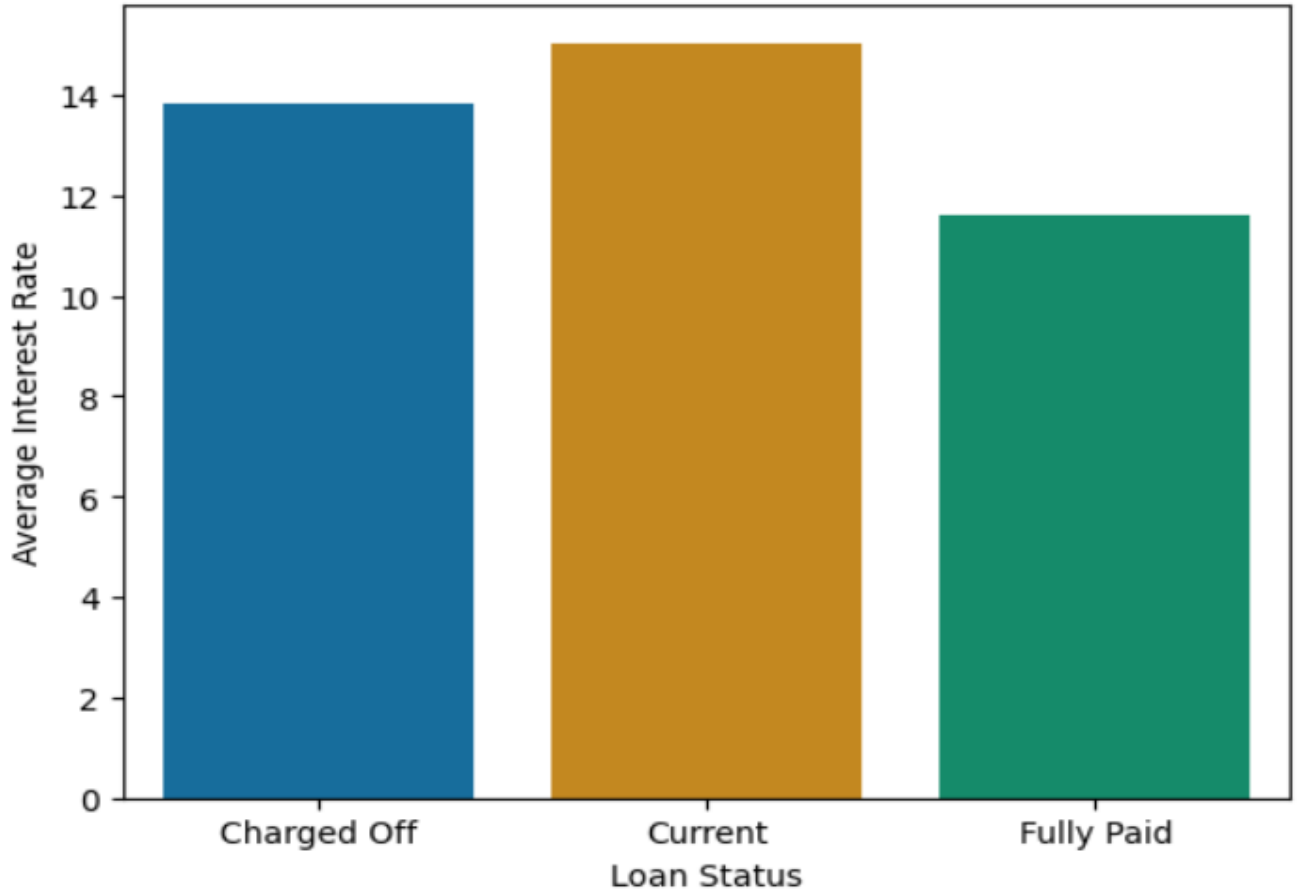


Also most defaulting loans are Grade A and B.



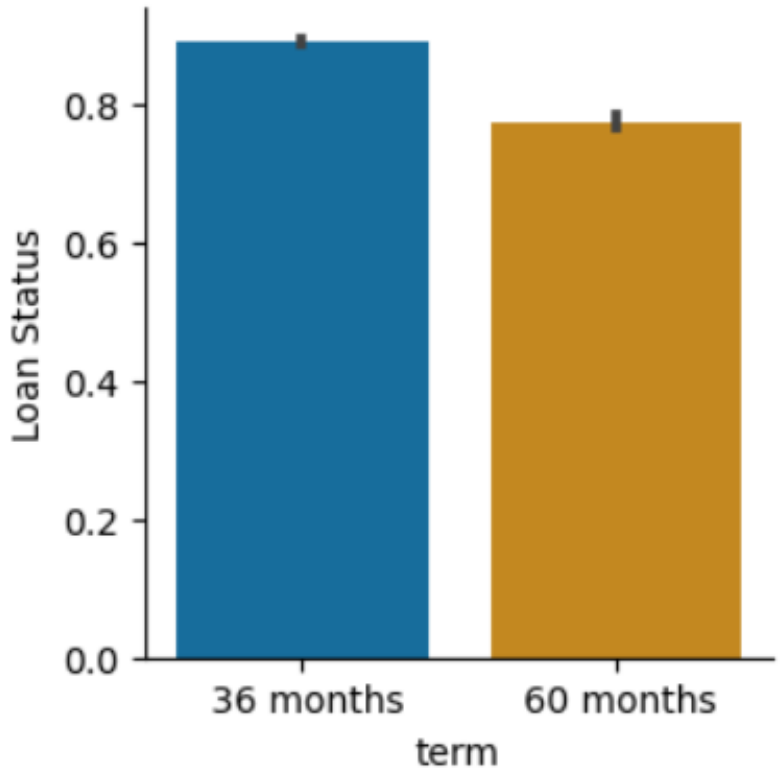
Relationship between Interest rates and Defaults

A higher interest rate has resulted in more defaults as we can see in the graph



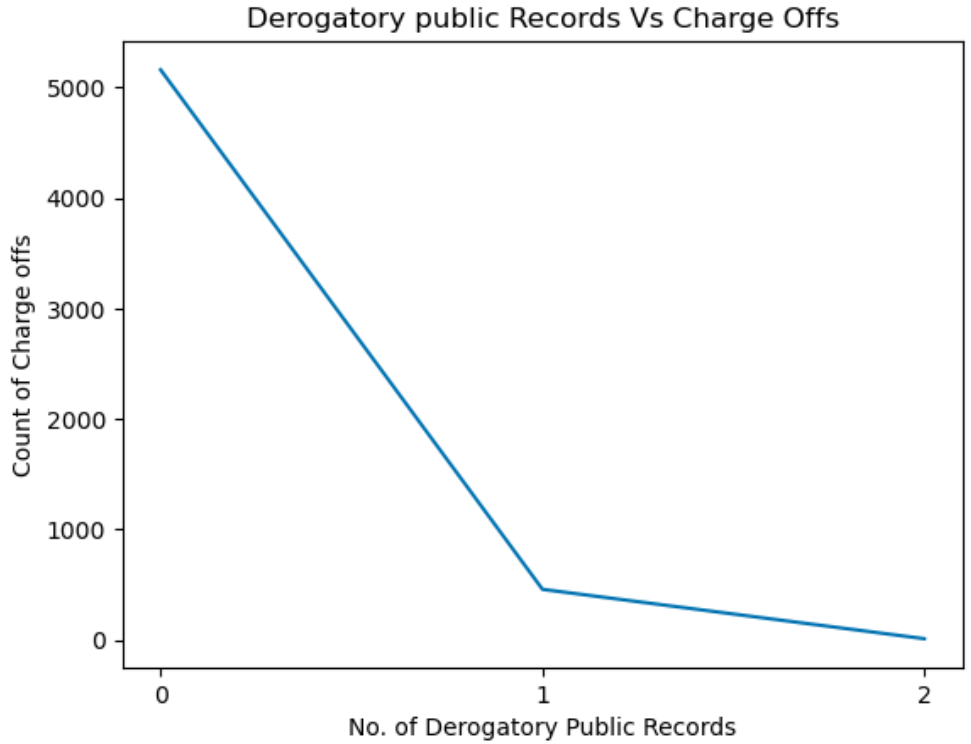
Loan status with Loan Term

Loans with term 36 months have defaulted more than 60 months

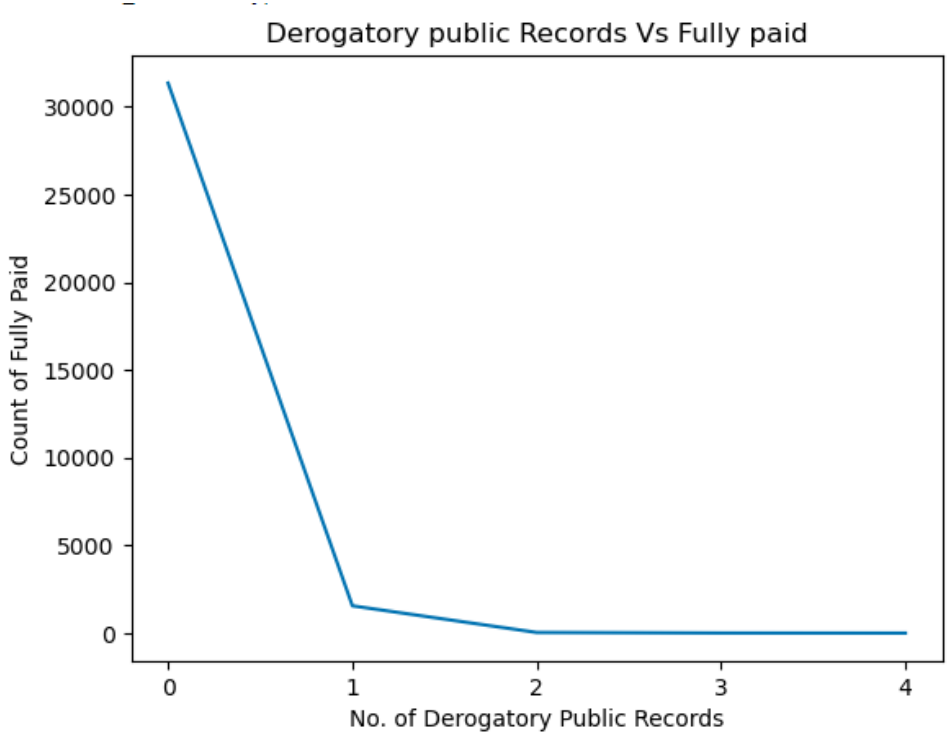


How public derogatory records affect loan payment

91% have charged off when public derogatory records were 0

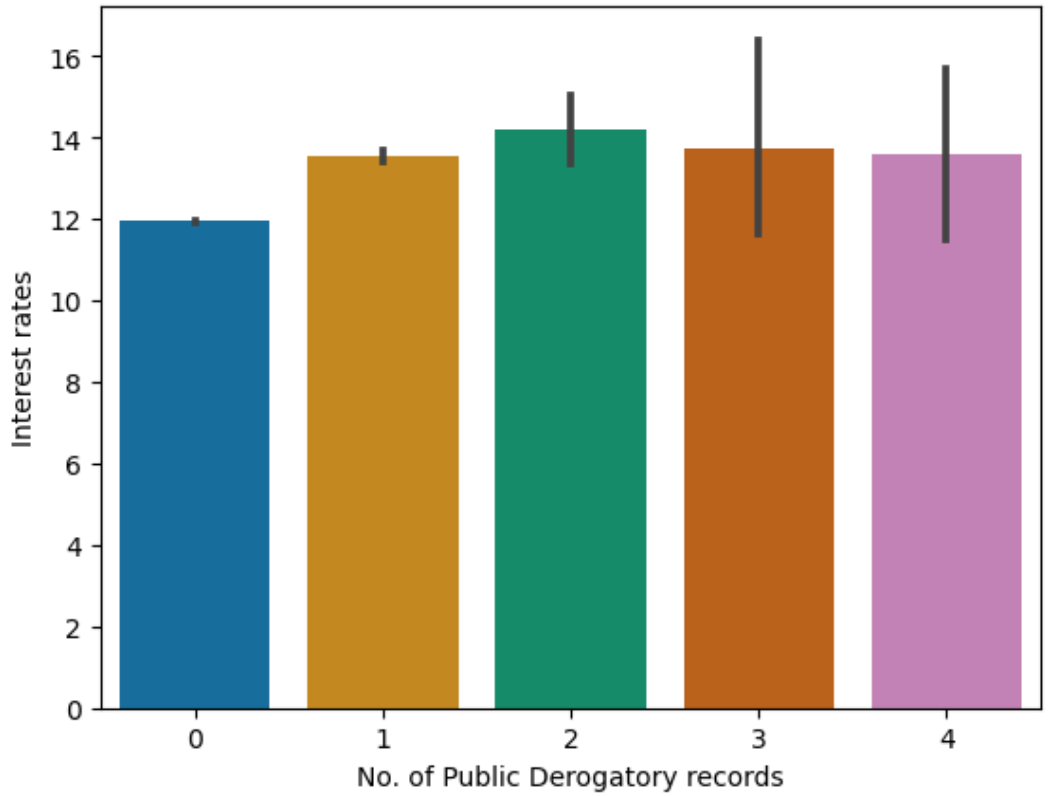


95 % have fully paid when public derogatory records were 0. But with increase in public derogatory records, there is not much difference in loan payment.



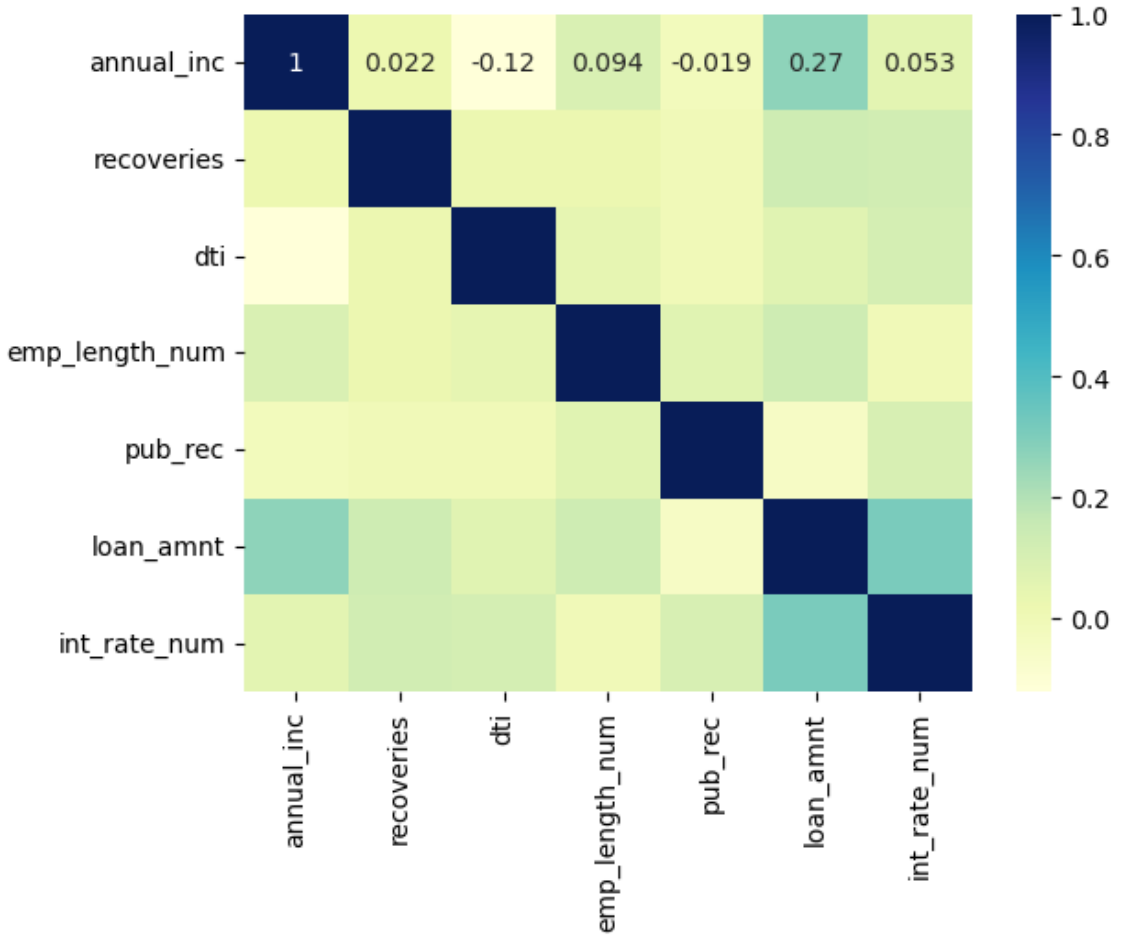
Interest Rate distribution with Public Derogatory records

So we see lowest interest rate for 0 public derogatory records.



Multivariate Correlation Analysis

- We see some co-relation of 0.3 between Loan Amount and Interest Rate
- Co-relation between public derogatory Records and interest rate is around 0.1.
- Loan Amount and Annual Income have a co-relation of 0.27
- DTI to Interest Rate is 0.1



Key Observations & Recommendations

- *Interest rates can be increased for RENT ownership and decreased for MORTGAGE ownership, as we see that MORTGAGE home owners have a lower defaulting rate.*
- *An average of lower interest rates have resulted in Full payment. Hence decreasing the interest rate can be considered to avoid losses.*
- *Interest rates should not be much different based on Public derogatory records, as we see a high rate of charge-offs (91%) when public derogatory records were 0.*
- *Annual Income group of 3k-31k have the highest default rate. But the interest rate for this group is lower than the higher income group. Low income groups can either be granted lower loan amounts or interest rate can be decreased to avoid defaulting.*
- *Most Loans have been taken for Debt consolidations but the interest rate is highest for housing Loans. This can be redistributed.*
- *Loan Grades A and B have the highest default rates and Loan Grade G has the highest Interest Rate. Loan grades A and B should be looked at more closely and interest rates be revised.*