

Lending Club Case Study

By –

- Ranjit Patra
- Kanchan Singh

Problem Statement and Objective

Problem Statement:

Lending Club is a marketplace for loans that matches borrowers who are seeking a loan with investors looking to lend money and make a return.

When Lending Club receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision :

1. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
2. If the applicant is not likely to repay the loan, then approving the loan may lead to a financial loss for the company

Objective:

The Objective of this case study is to use the provided data (loan.csv) to identify driving factors before making decision on loan application. So, the lending club can utilize these factors to make decisions before approving or rejecting an loan.

Data Cleaning

1. Removed fully empty columns.
2. Columns which were having more than 70% empty data were removed.
3. Columns those had one unique entry throughout were dropped.
4. Filled a few columns empty value with 0 or other available values.
5. There were no duplicate rows in the data set.
6. Removed columns which were not required for analysis.
7. Did not consider loan status with “Current” value.

Outliers Treatment:

- Loan Amount - See the flowing stats before and after outlier treatment for loan amount.

Before Outlier Treatment

	loan_amnt	funded_amnt	funded_amnt_inv
count	39717.000000	39717.000000	39717.000000
mean	11219.443815	10947.713196	10397.448868
std	7456.670694	7187.238670	7128.450439
min	500.000000	500.000000	0.000000
25%	5500.000000	5400.000000	5000.000000
50%	10000.000000	9600.000000	8975.000000
75%	15000.000000	15000.000000	14400.000000
max	35000.000000	35000.000000	35000.000000

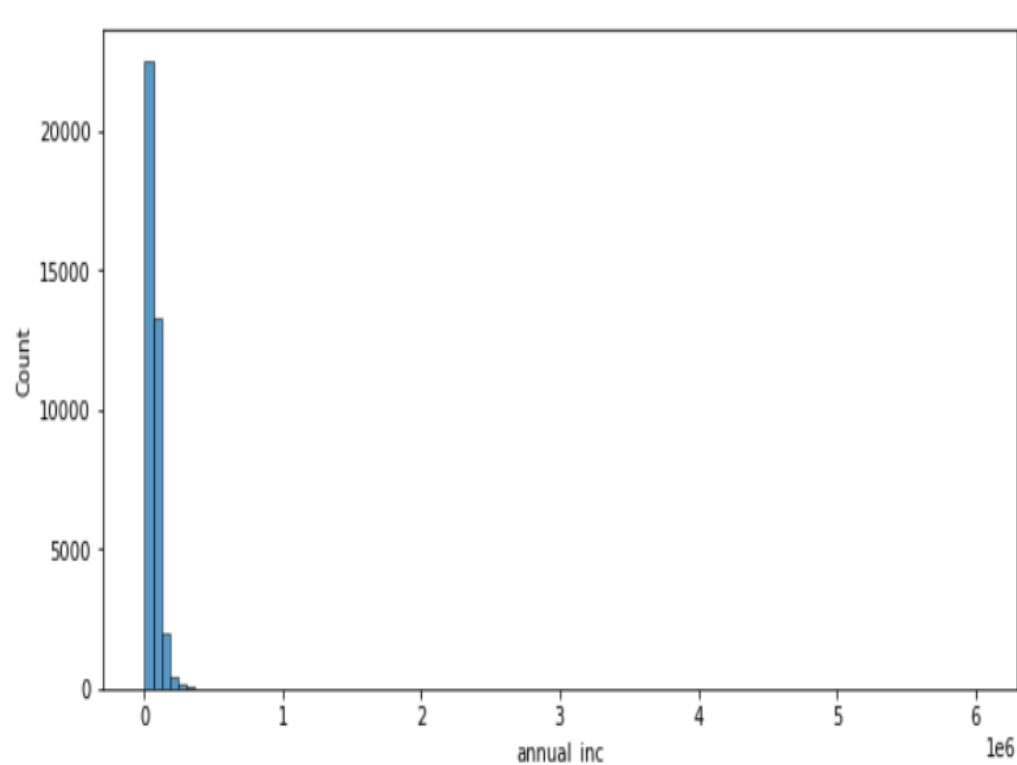
After Outlier Treatment

	loan_amnt	funded_amnt	funded_amnt_inv
count	38512.000000	38512.000000	38512.000000
mean	10534.148447	10305.254206	9768.830486
std	6457.350486	6251.480244	6202.386907
min	500.000000	500.000000	0.000000
25%	5125.000000	5000.000000	5000.000000
50%	9600.000000	9250.000000	8500.000000
75%	15000.000000	14400.000000	13750.000000
max	29900.000000	29850.000000	29825.000000

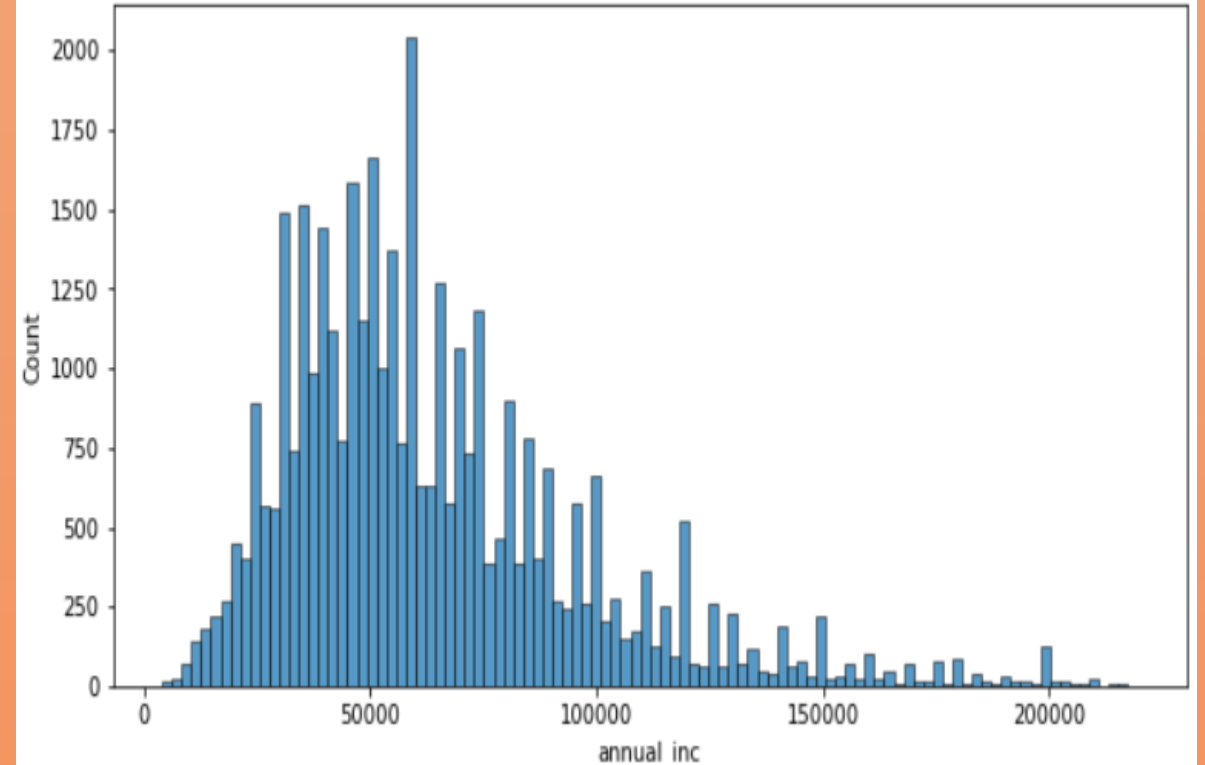
Outliers Treatment:

- Annual Income – See the flowing stats before and after outlier treatment for loan amount.

Before Outlier Treatment



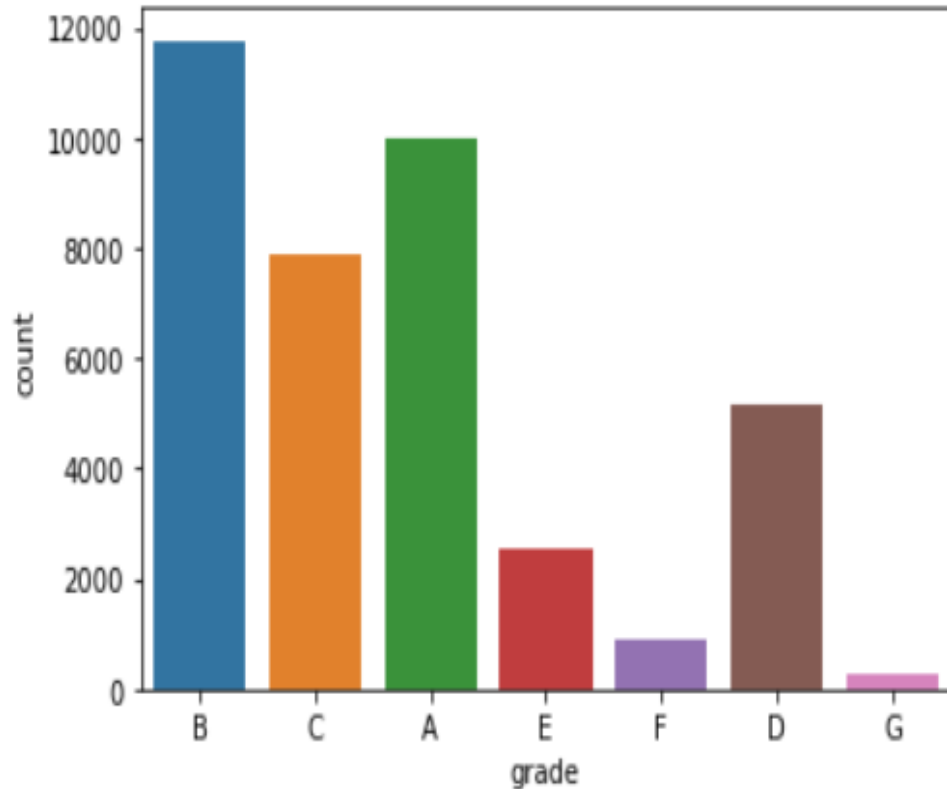
After Outlier Treatment



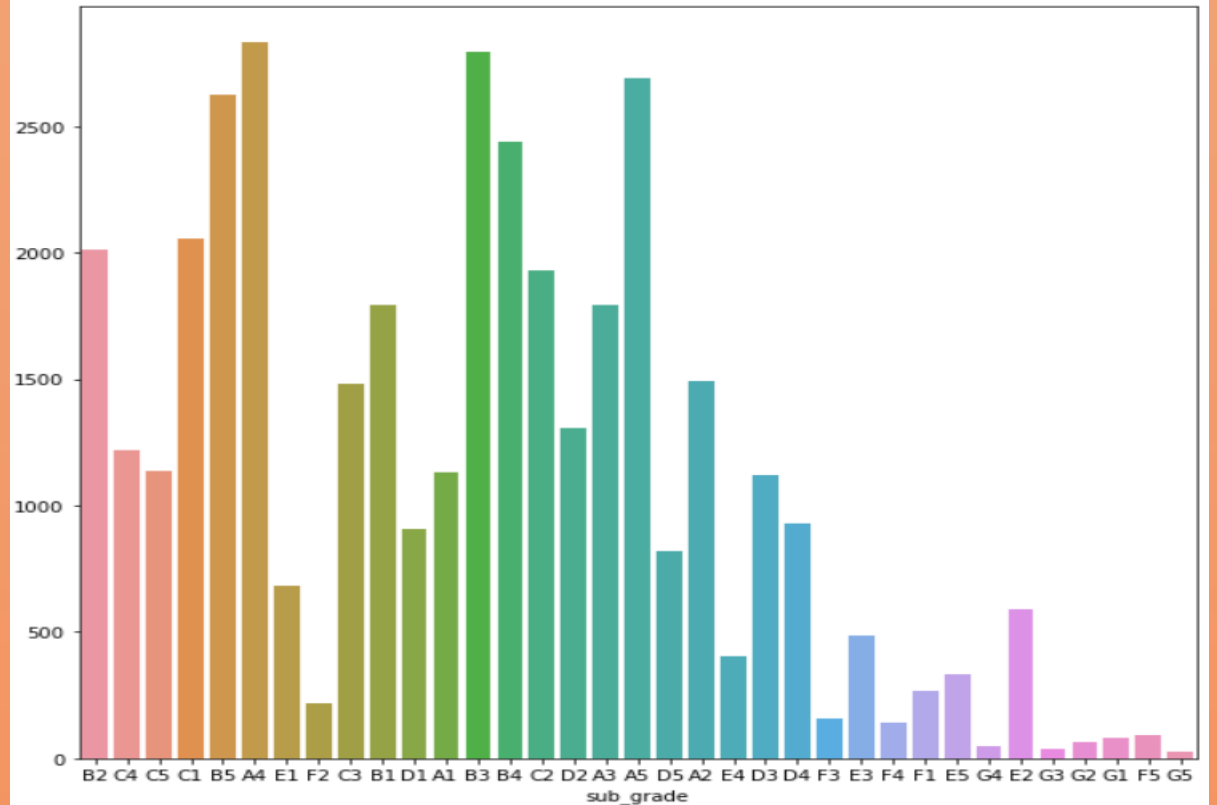
Univariate Analysis

- **Grade and Sub Grade** – We can see that most applicants are from Grade B and Sub Grade A4 and B3.

Before Outlier Treatment

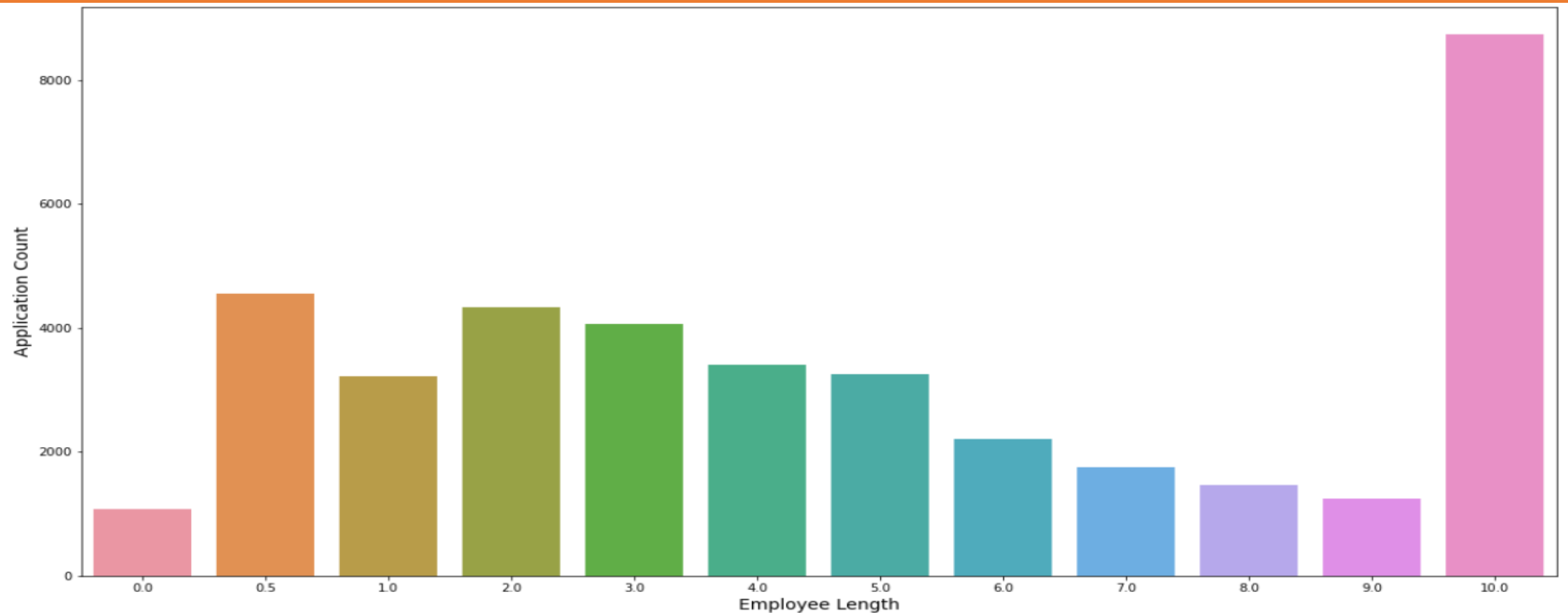


After Outlier Treatment



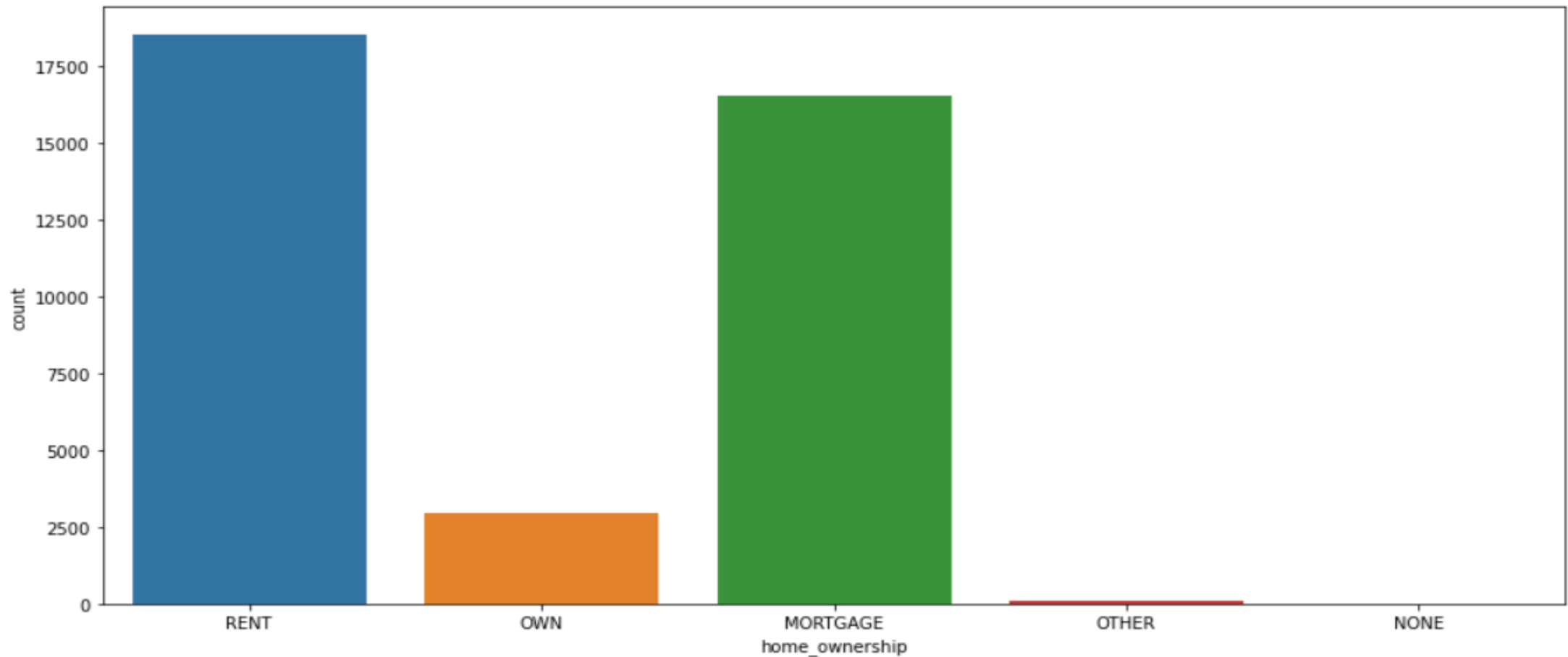
Univariate Analysis

- **Employment Length** – Loan applicants are getting decreased with employment length, until it reaches to 10 and applicants are high with 10+ employment.



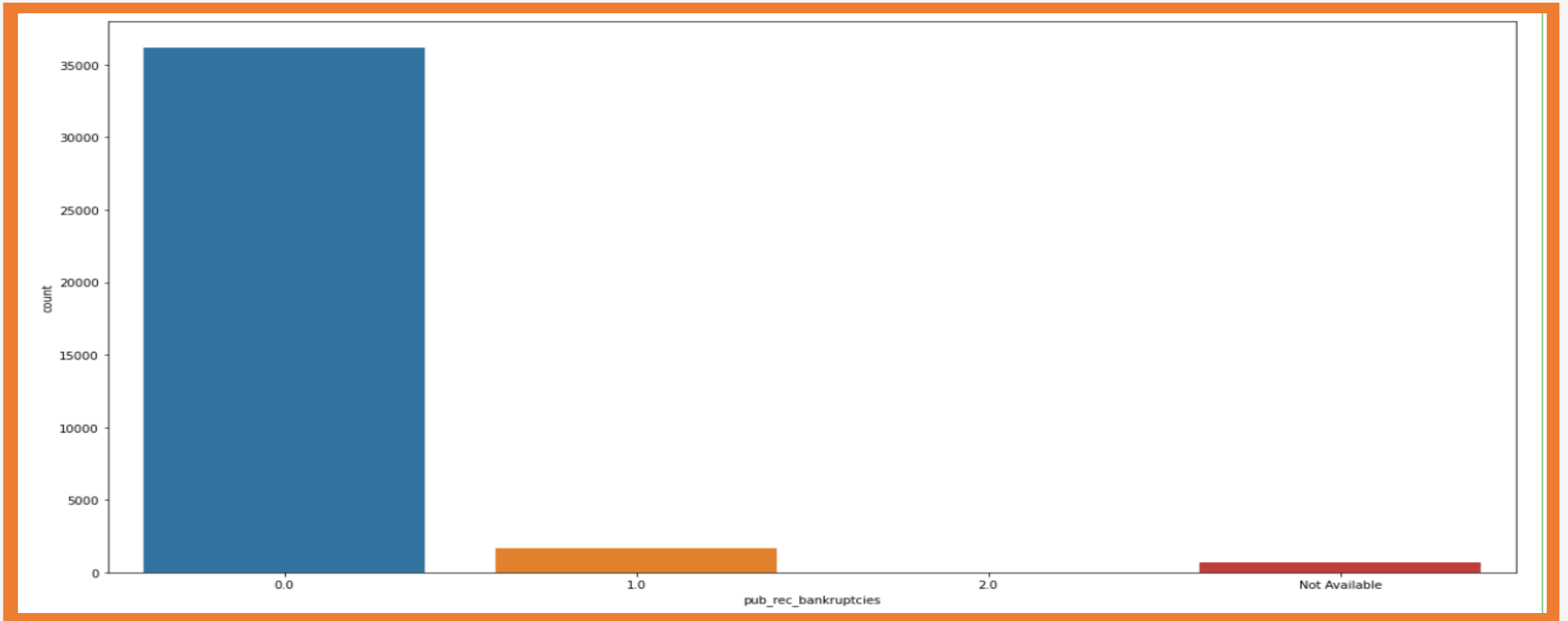
Univariate Analysis

- **Home Ownership** – The majority of the applicants belongs to RENT or MORTGAGE category.



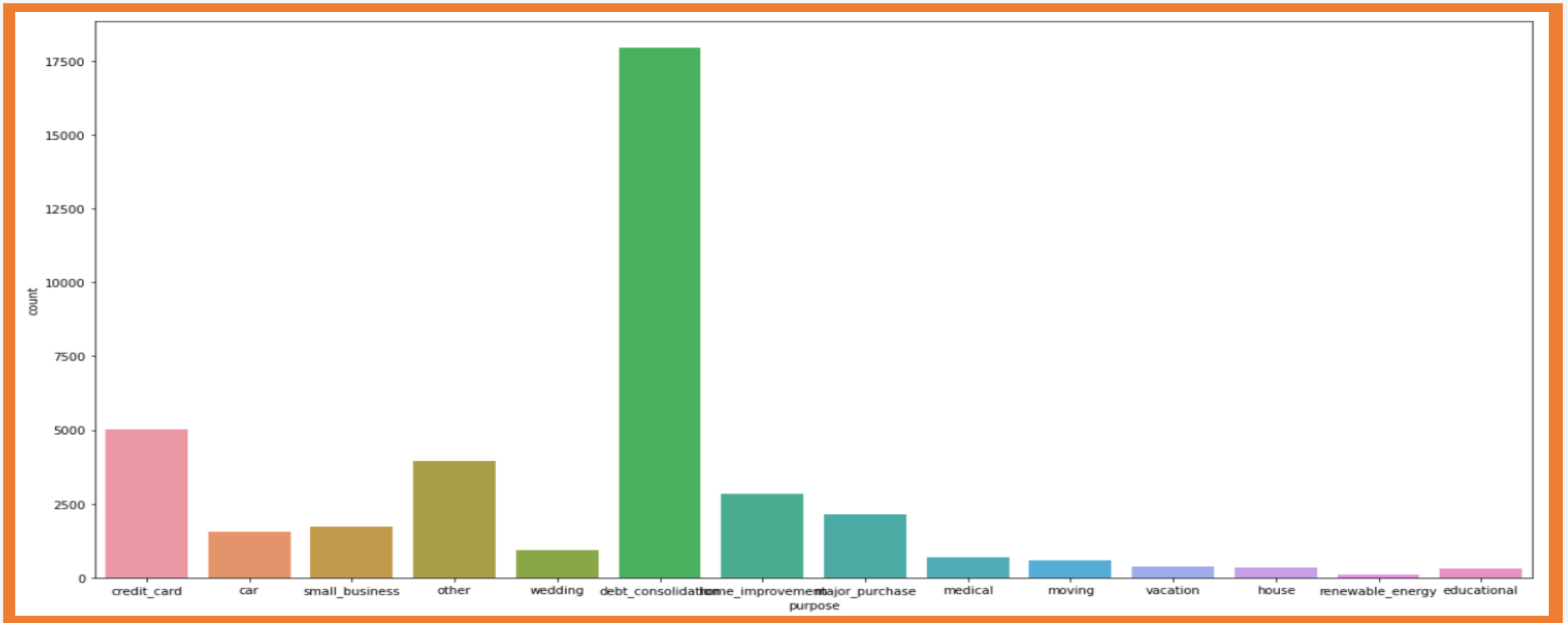
Univariate Analysis

- **Public Record Bankruptcies** – Most of the applicants has 0 public record bankruptcies.



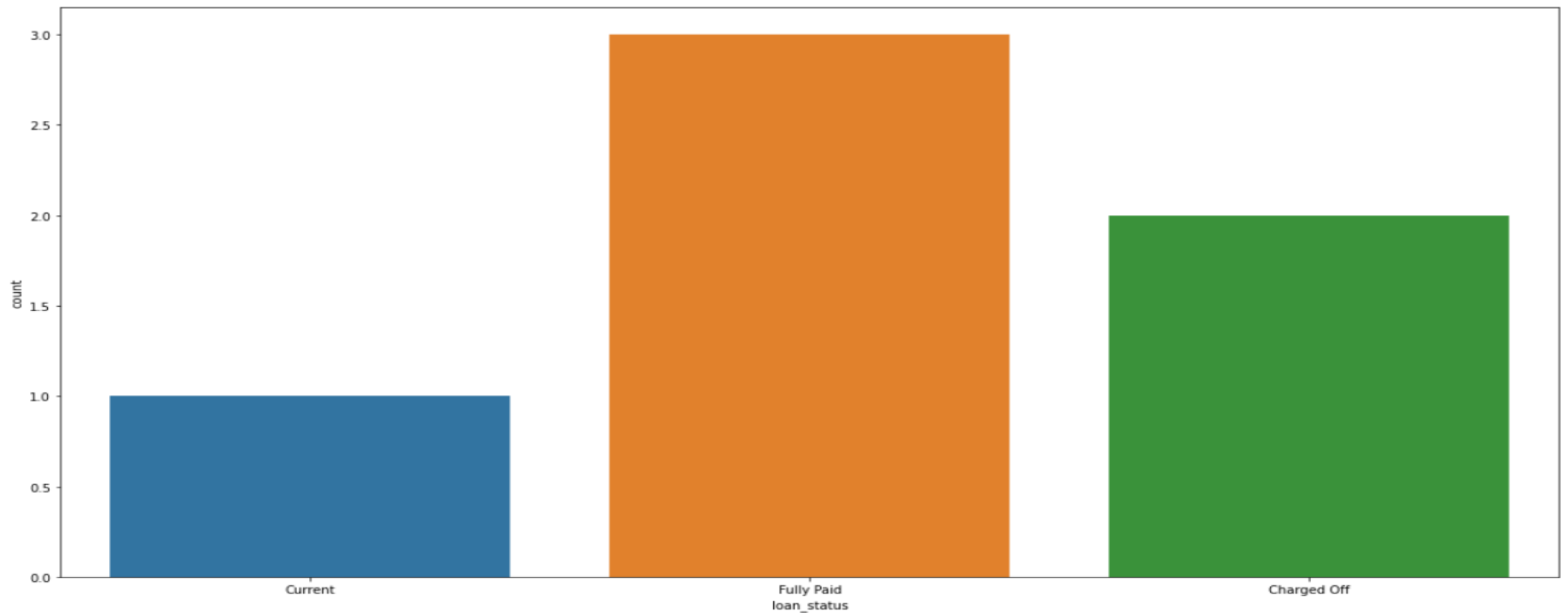
Univariate Analysis

- **Purpose** – Majority of the applicants applied loan with the purpose debt consolidation.



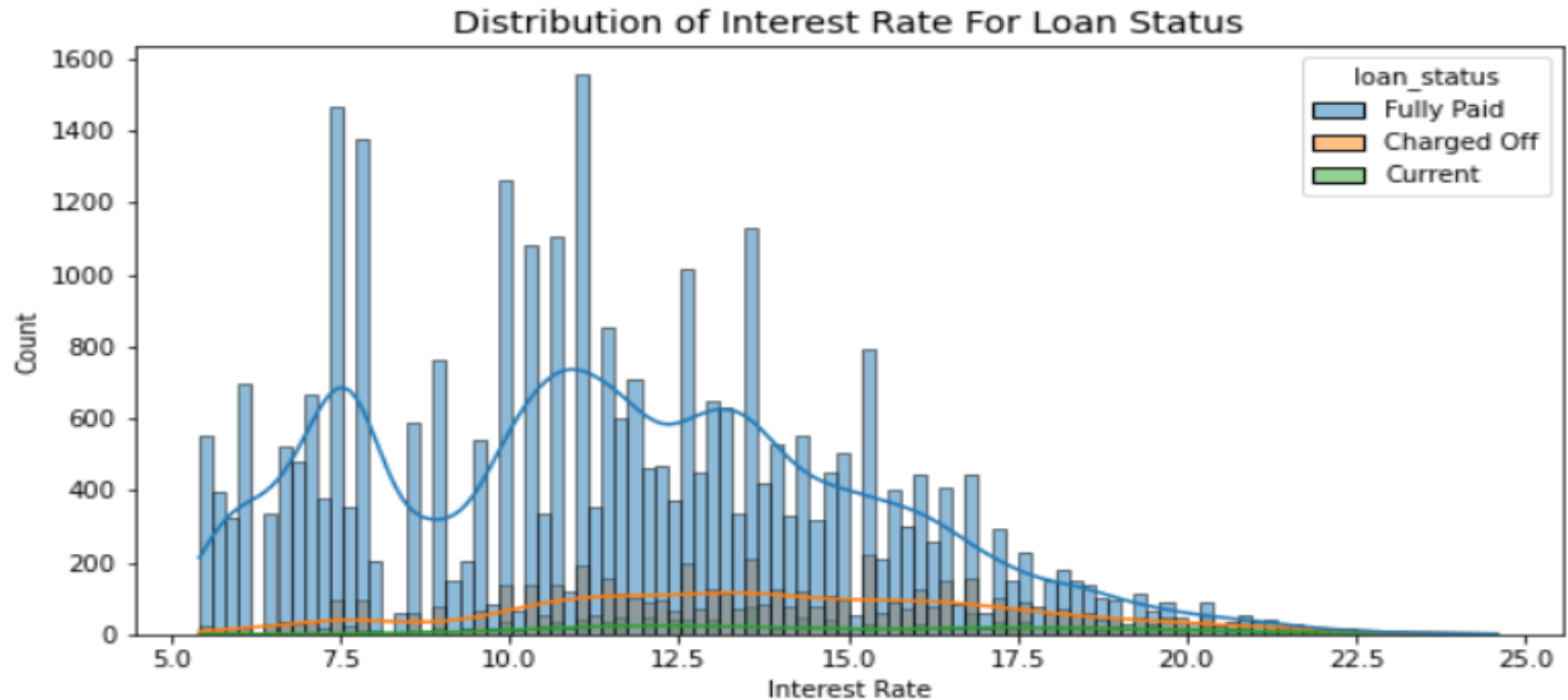
Segmented Univariate Analysis

- **Loan Status** – Charged Off getting increased for more Bankruptcies and it is 22.5% for 1 times bankruptcies and increased to 33% for 2 times bankruptcies.



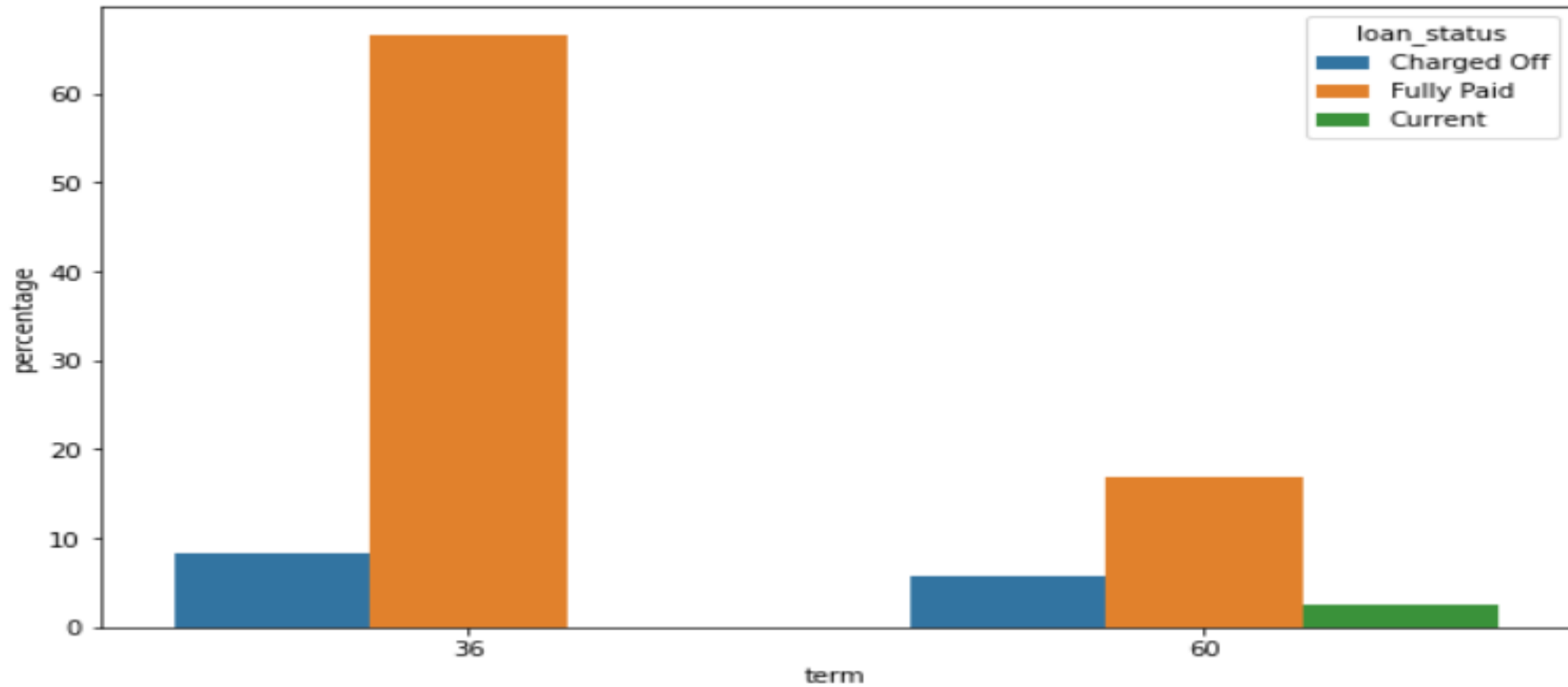
Bivariate Analysis

- Interest Rate and Loan Status – Charged Off increases with interest rate and shows a decline after 17.5%.



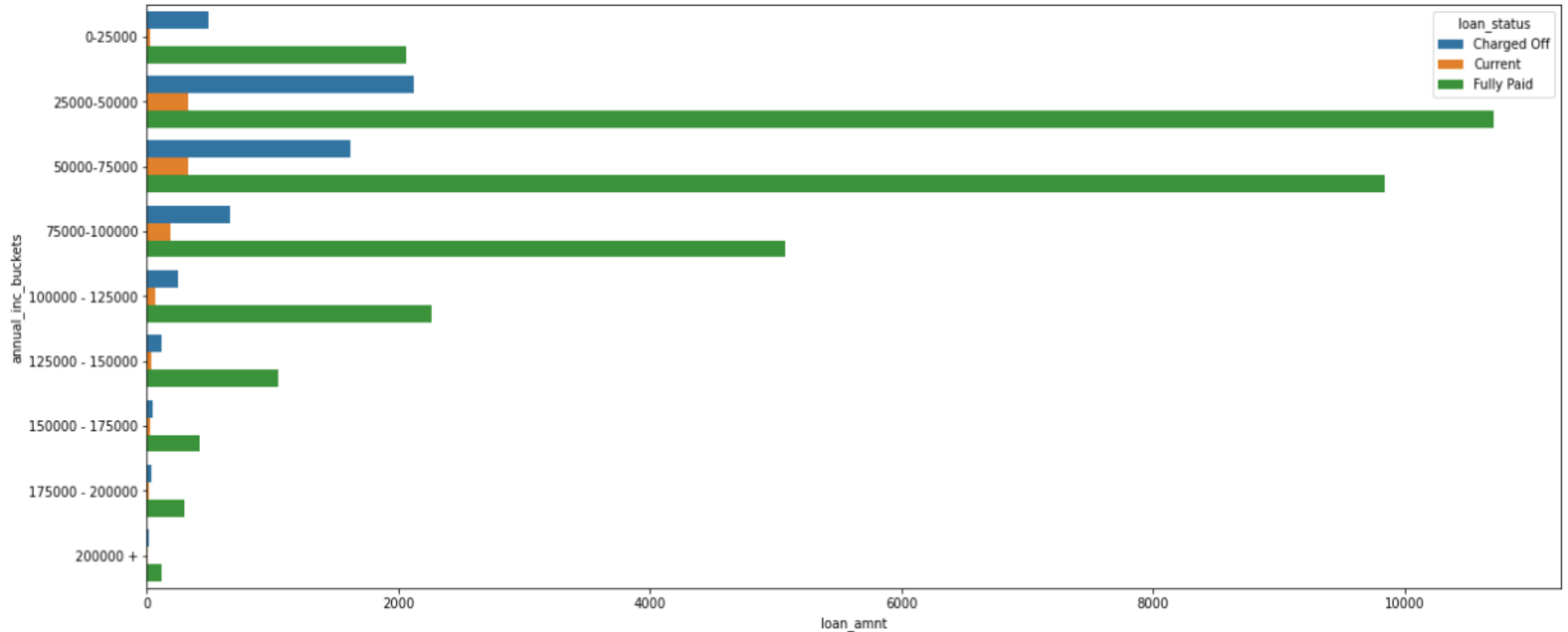
Bivariate Analysis

- Term and Loan Status – Longer term has a high charged off percentage.



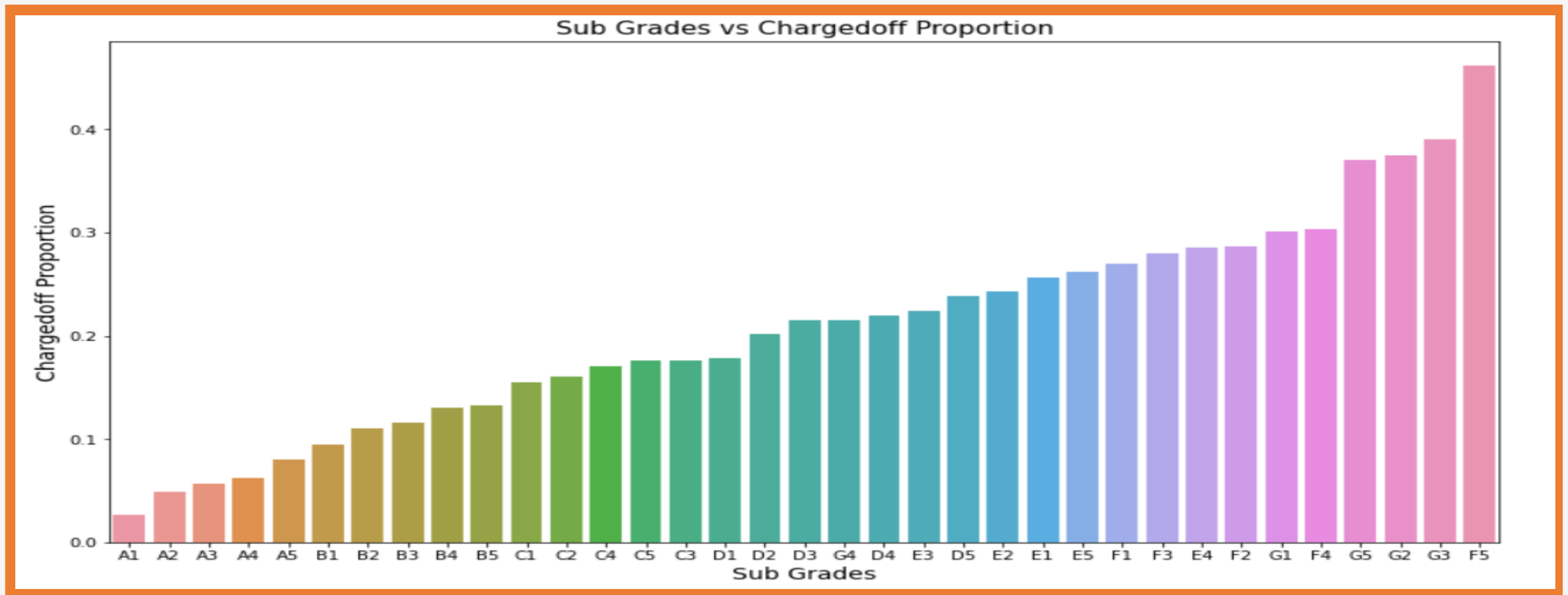
Bivariate Analysis

- Annual Income and Loan Amount – Longer term has a high charged off percentage.

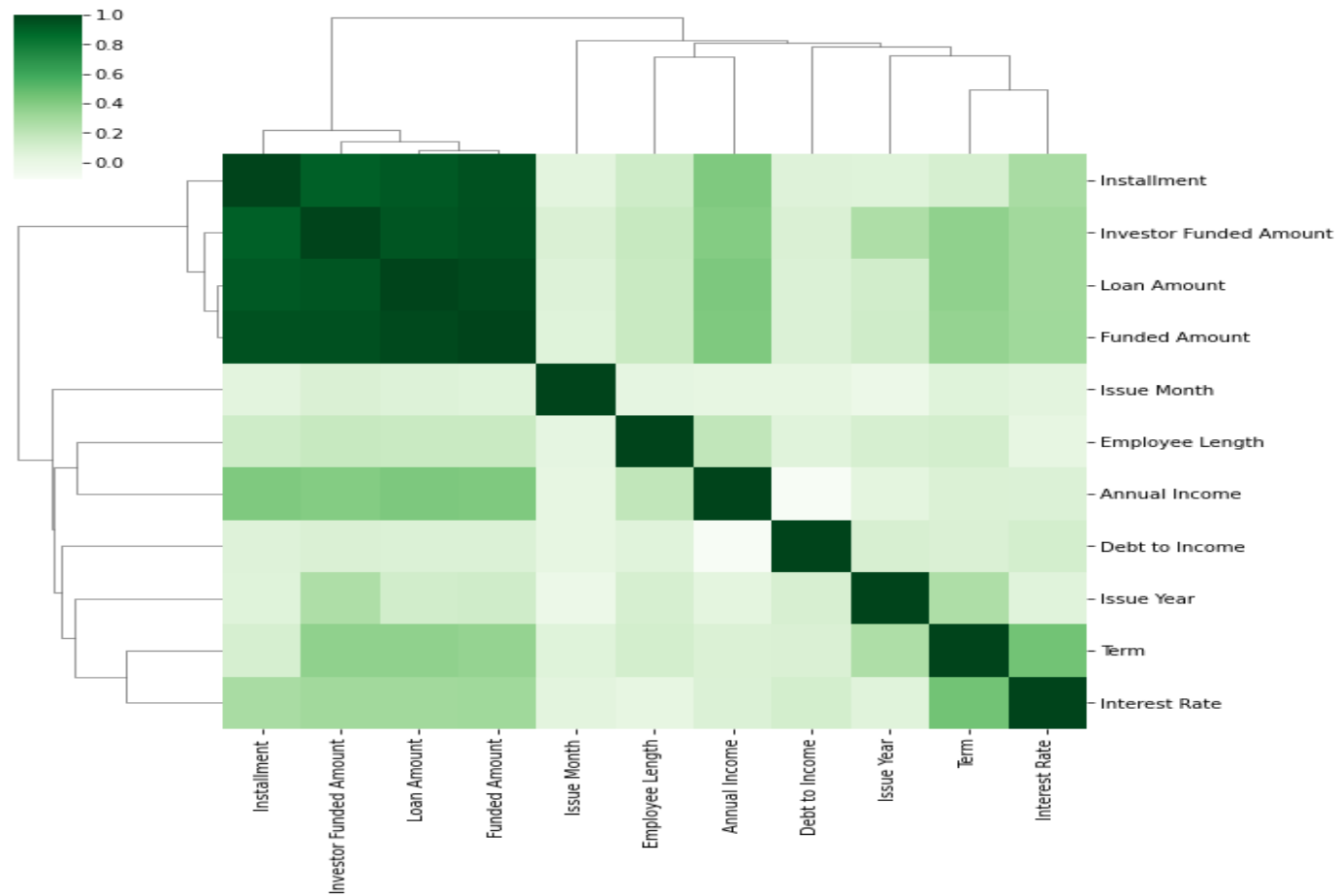


Bivariate Analysis

- **Sub Grades and Charged off Proportion** – Sub Grades of "A" has very less chances of charged off. Sub Grades of "F" and "G" have very high chances of charged off and proportion of charged off is increasing with sub grades moving from sub grades of "A" towards sub grades of "G".



- Co relation among numerical variables.



Recommendations

Driving Factors	
Public record bankruptcies	Applications with bankruptcies records should be rejected.
Term	Applicants with long term has high percentage of charged off and we should deny those applicants.
Annual income	There is high risk involved with applicants with high annual income.
DTI	Increase in DTI Charged off also getting increased. So, we should wisely consider these applicant.
Interest Rate	High rate are not defensive to become a default, so better strategy could be deny those particular applicants.
Address State	Applications from the state NV, FL, MO has high chance of charged off. Therefore denying loans to risky individuals may be better strategy.
Purpose and Funded Amount (Investor)	Investor approved amount in bucket (20000 - 25000) has high risk involved.