Lending Club Case Study

Submitted By:

Bharath BP & Prakash Dakshina

Objective of the Case Study.

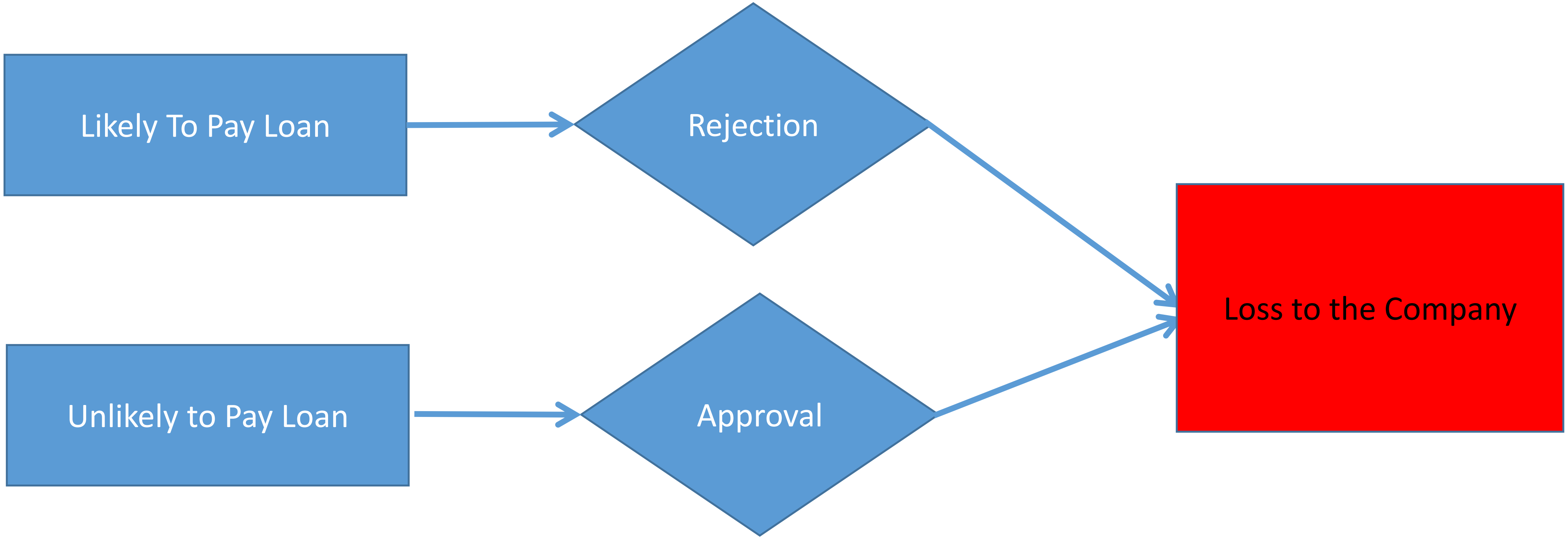
- To develop a basic understanding of risk analytics in banking and financial services
- To understand how data is used to minimize the risk of losing money while lending to customers.

Selection of Case Study.

- To Get an idea how real world problems are solved using Exploratory Data Analysis, the case study of **Lending Club,** a consumer Finance Company is taken up.
- To proceed with Case study of Lending Club, it is essential to understand the following
    - Business Understanding (Decisions & Risks)
    - Business Objectives

Business Understanding(Risks).

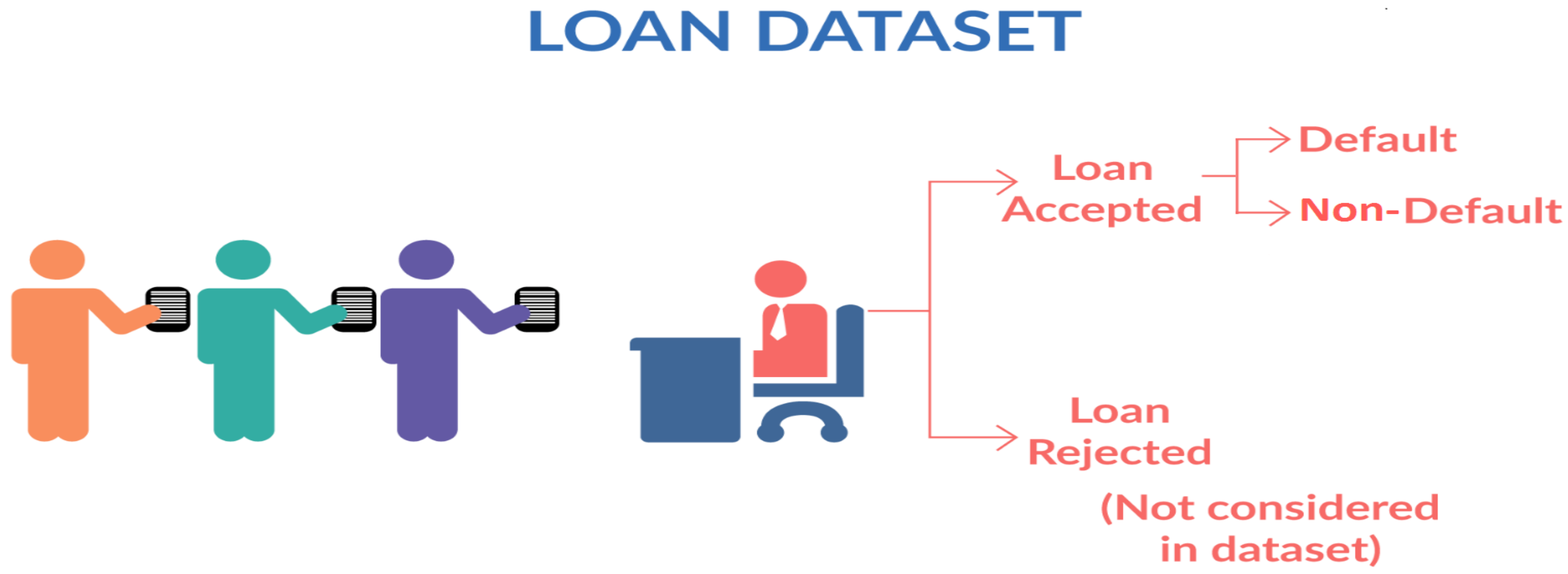
- Lending Club is a platform to provide Loans to various customers. The Two types of risks involved while evaluating a loan application



Business Understanding(Decisions)

Two types of Decisions are taken by the Lending Club

1) Loan Accepted
2) Loan Rejected.



Business Objectives.

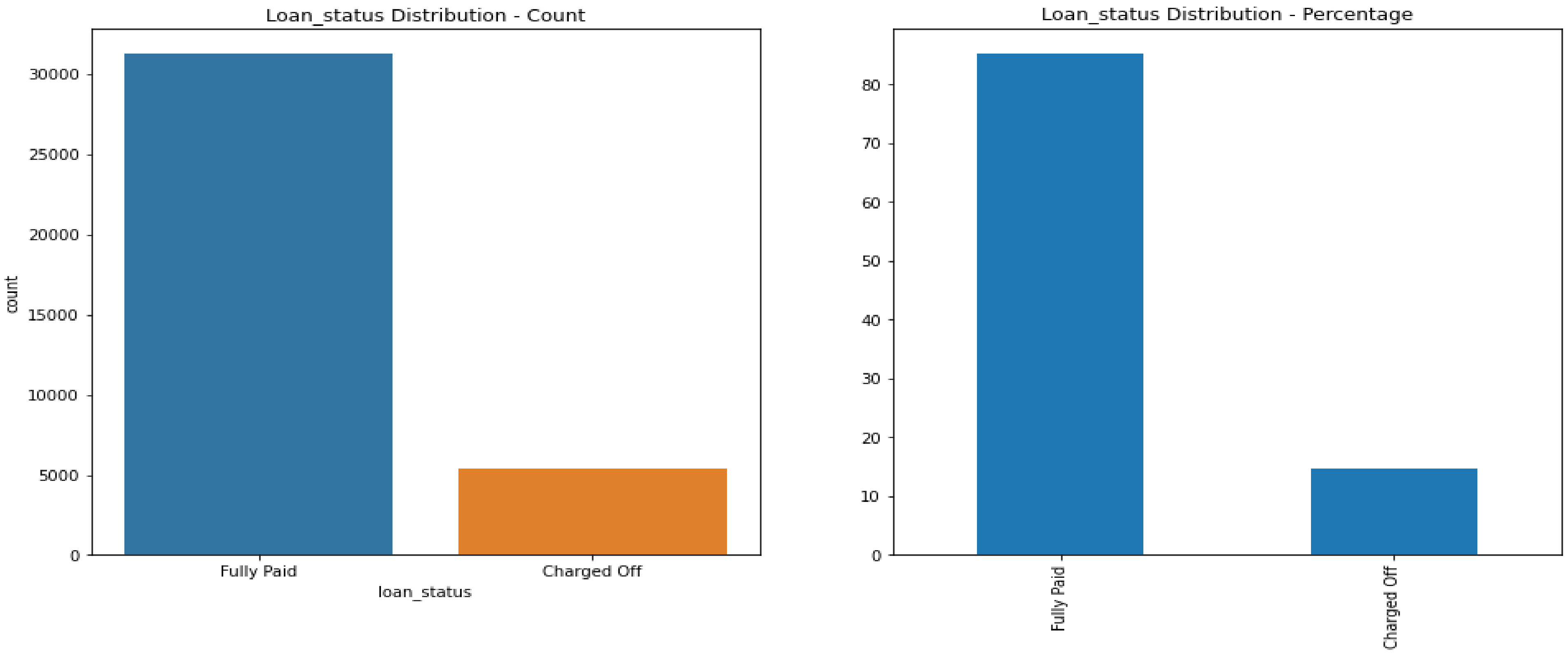
- To understand the driving factors/variables behind a loan default.
- Improve portfolio and risk assessment tools, identification of the risky loan applicants, will help in reducing such loans, thereby cutting down the amount of credit loss.

Approach taken for Analysis

1. Data Sourcing & Data Cleaning

- Load the Loan Dataset
- Remove unnecessary columns based on missing value percentage threshold
- Remove unnecessary columns which have few unique values.
- Handling of missing values
- Standardizing the data
- Outlier Treatment.

1. Data Sourcing & Data Cleaning (cont)



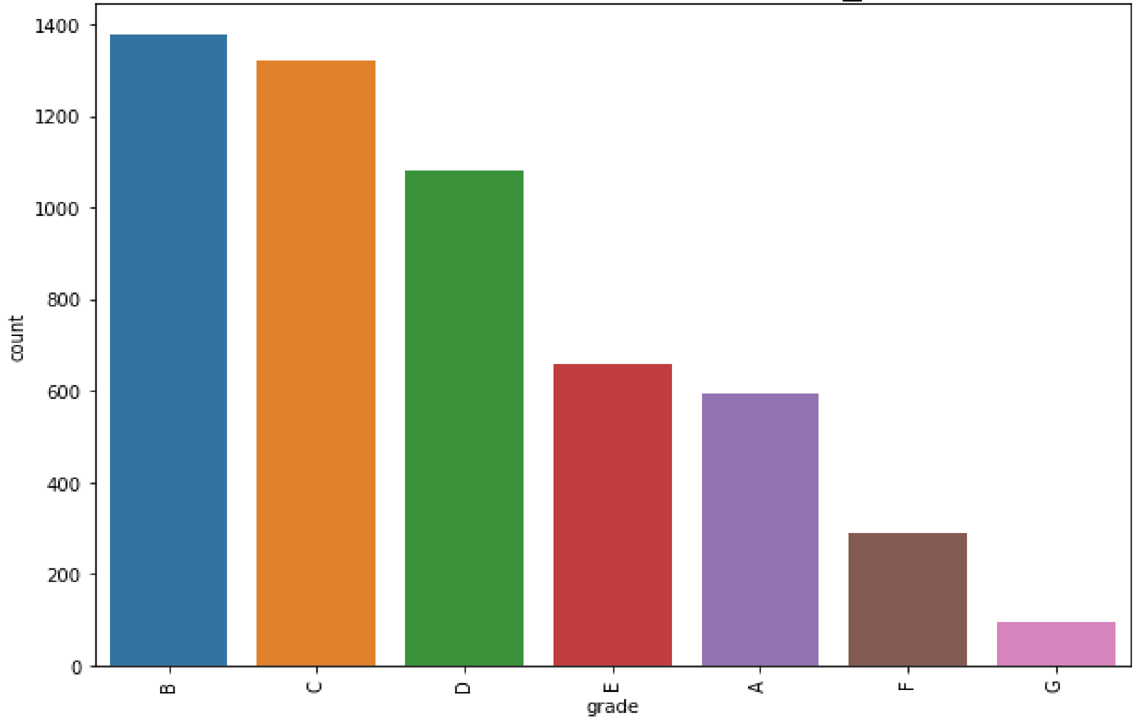
After performing all the steps, the final and Total Loan Status Distribution is given above.
Out of Total Loans,14.75 %(5413 Loans) of Loans have bee Defaulted/Charged off.

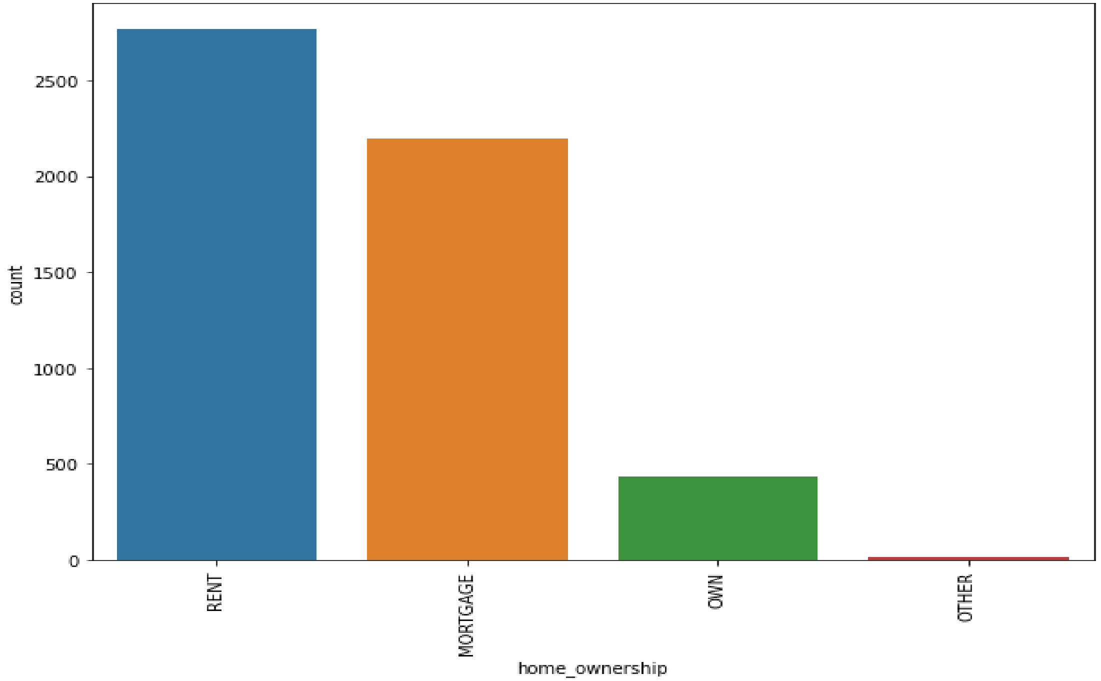
Note: Total Loans exclude current loans

2. Univariate analysis

- Grade-wise Default Rate
- Term-wise Default Rate
- Purpose-wise Default Rate
- Home-ownership wise Default Rate
- Issue Month-wise Default Rate
- Experience-wise Default Rate
- Income range Default Rate
- Loan amount Default Rate

Grade and HomeOwnership wise Default Rate

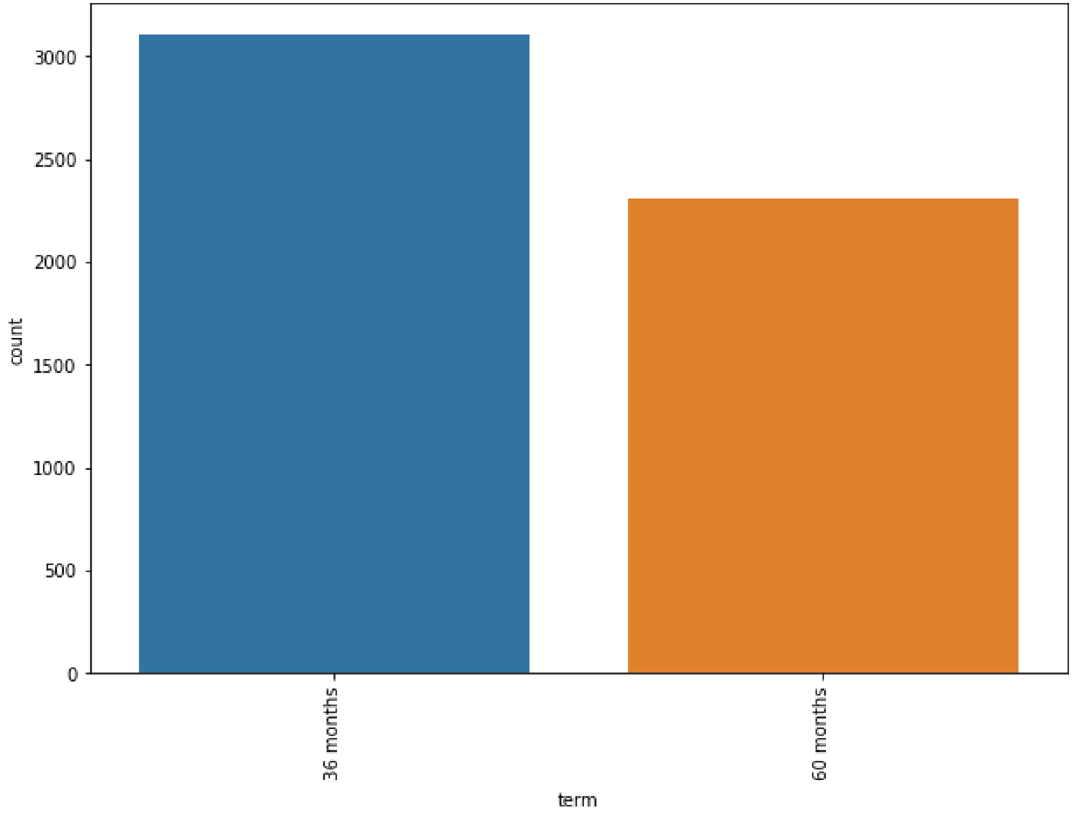


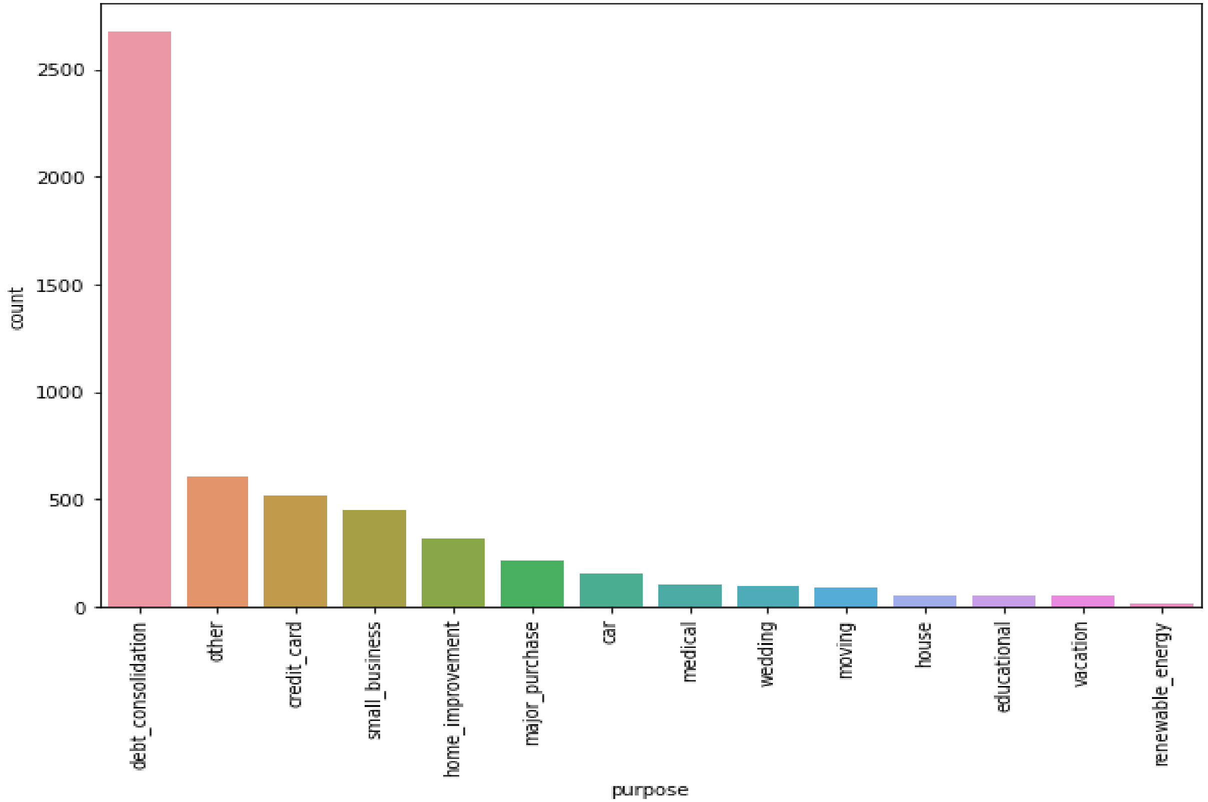


- **Based on the above graph, the applicants who fall under the Grade 'B' are most likely to default and the applicants from Grade 'G' are less likely to default the loan.**

- **Based on the above graph, Loan applicants with rented Homes are more likely to default. Home owners who own the home, are assumed to least likely for default (Others category seems negligible, please ignore)**

Term and Purpose wise Default Rate

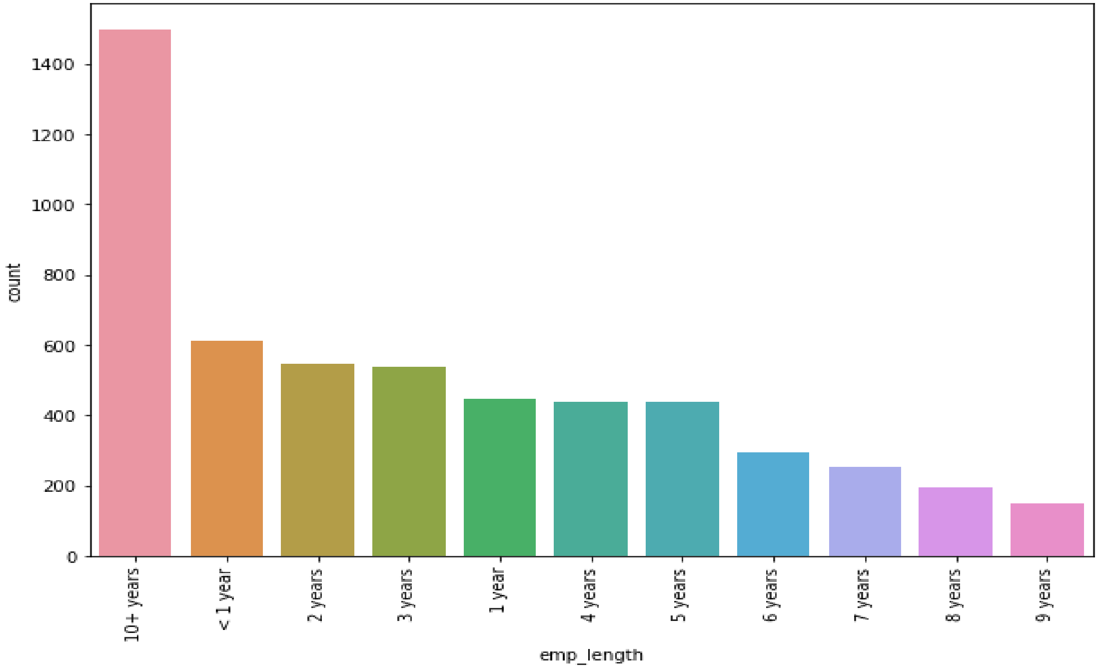


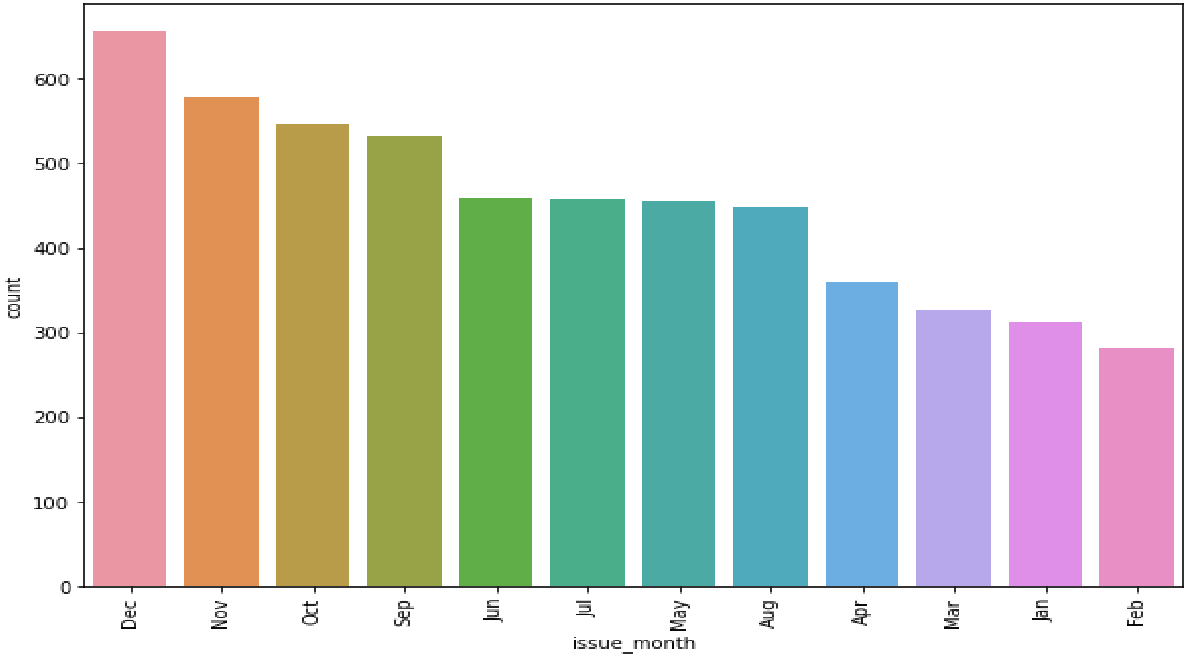


- **People who take loans with 36 Months Term are more likely to default. It might be due to the reason that shorter term duration which means that EMI component would be higher.**

- **Based on the above graph Loans taken for purpose of 'Debt_consolidation' seems to have the highest risk of defaulting on Loan.**

Experience and Month wise Default Rate

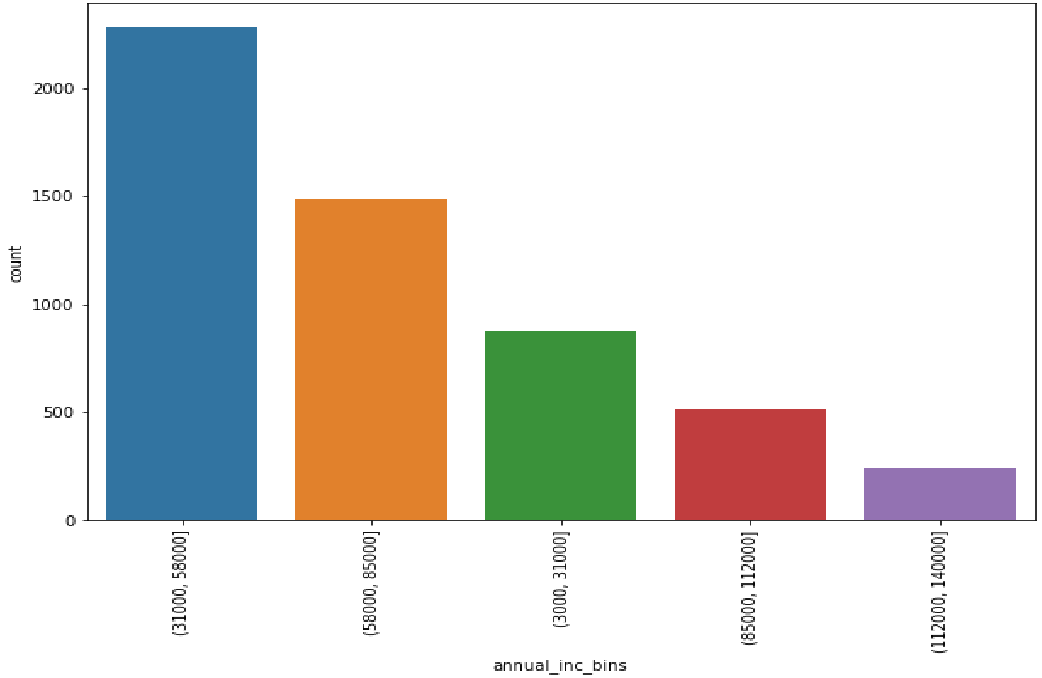


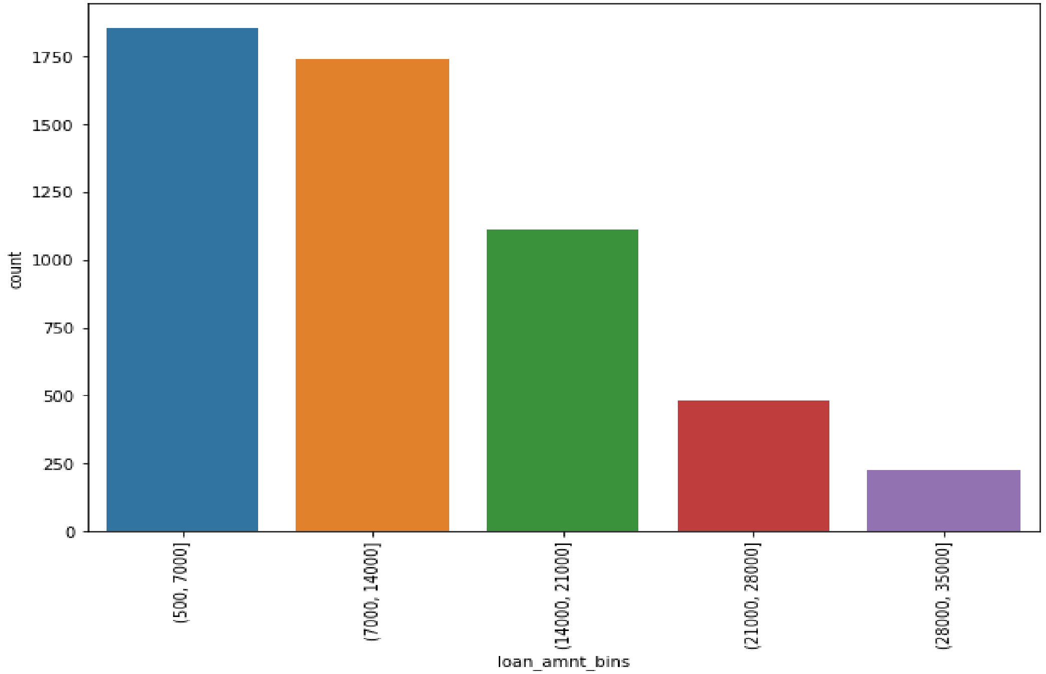


- **Employees with '<1 year experience'(in graph employees with 10+year having higher count this is because 10+ years will convert lot of applications people having experience in the range of 10-50 will also falls under this category, so we can assume <1 year experience having higher risk of defaulting)**

- **Based on the above graph the applicants who are issued loan in December month are more likely to default and the applicants who are issued loan in February are least likely to default the loan.**

Income Range and Loan Amount Range - Default Rate





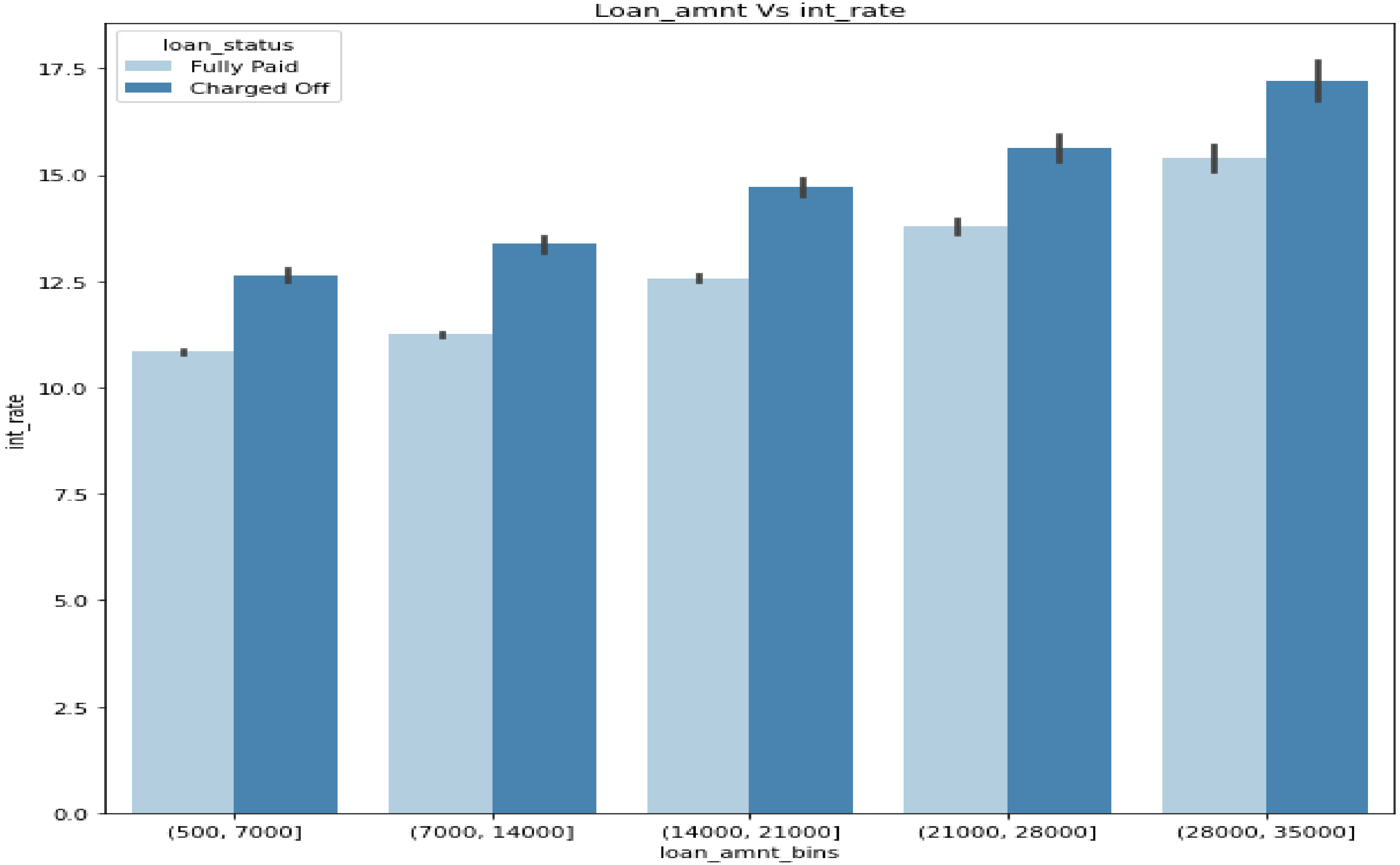
- **Based on the above graph, the applicants whose income range is of 31000- 58000 are more likely to default .**

- **Based on the above graph, the applicants whose loan amount range is of 500-7000 are more likely to default . When the loan amount range is low, Applicant tend to neglect. This is one of the possibility**

2. Bi/Multi variate analysis

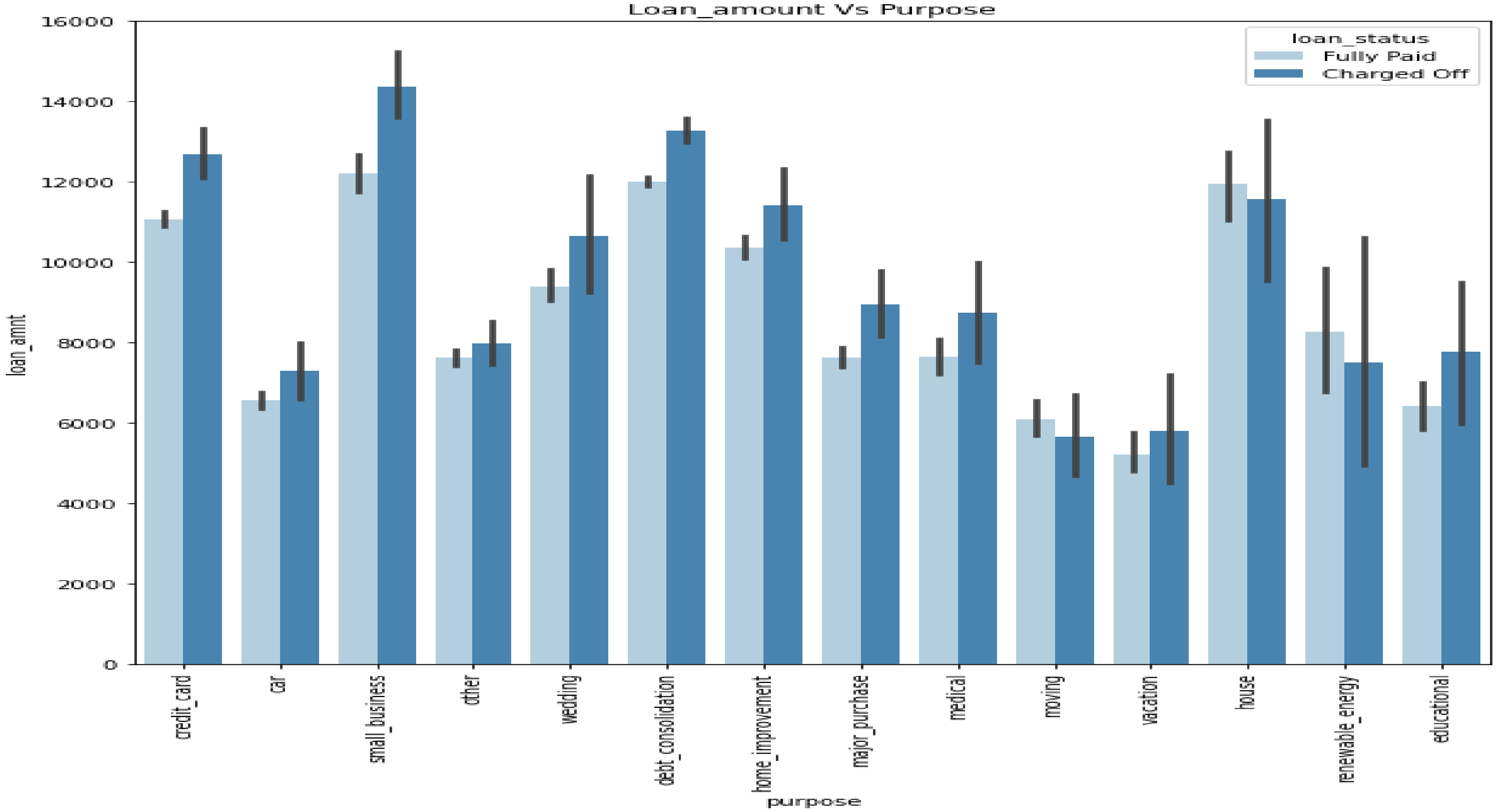
- Analyzing the Loan amount versus Interest Rate.
- Analyzing the Loan amount versus Loan purpose.
- Analyzing the Annual income versus Interest rate.
- Analyzing the Loan amount versus verification status.
- Heat Map.

Analyzing the Loan Amount versus Interest Rate



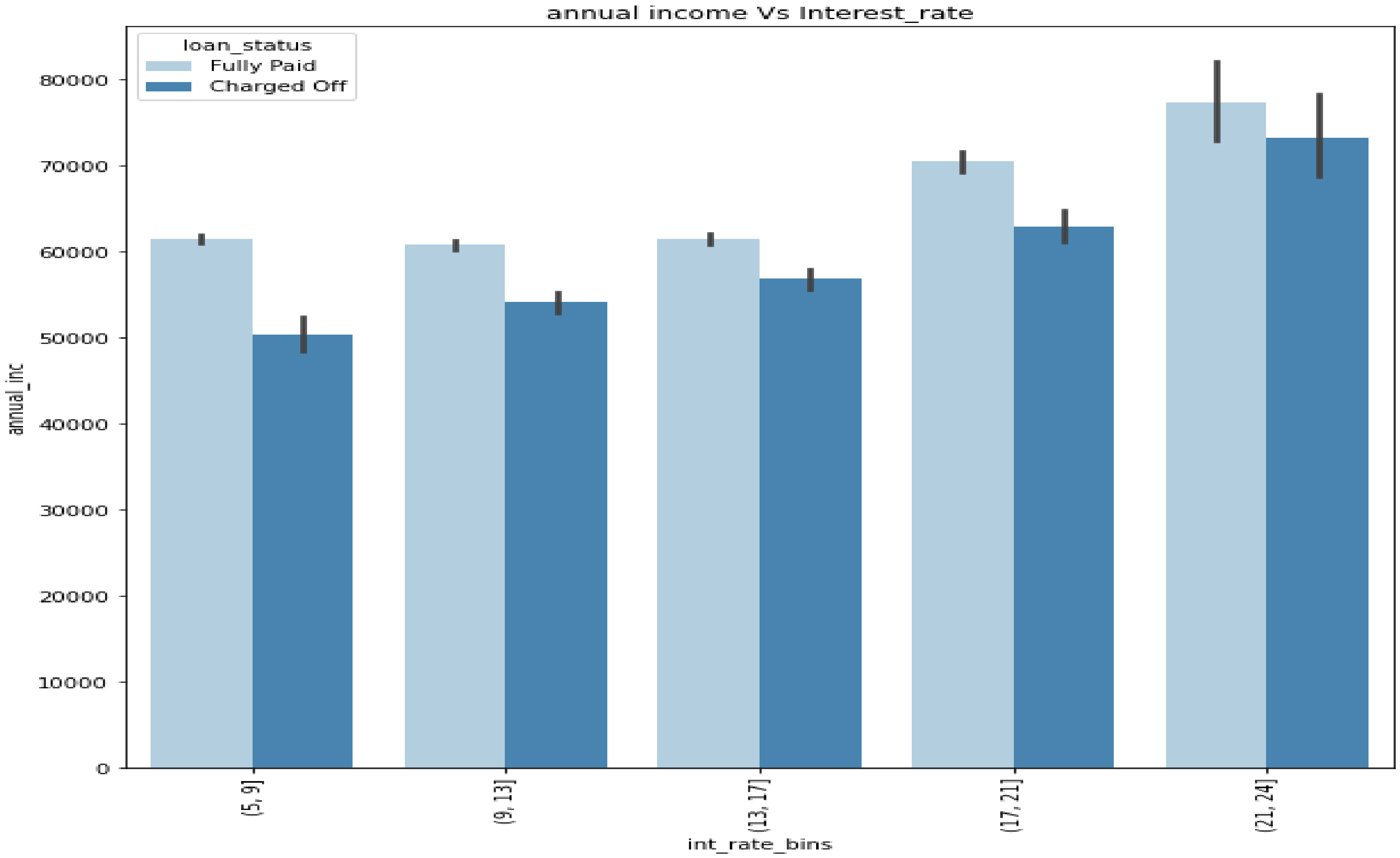
- **Based on the above graph, the applicants whose Loan amount in the range 28k - 35k and interest rate at 15-17.5 % are highly likely to Default .**

Analyzing the Loan amount versus Loan purpose.



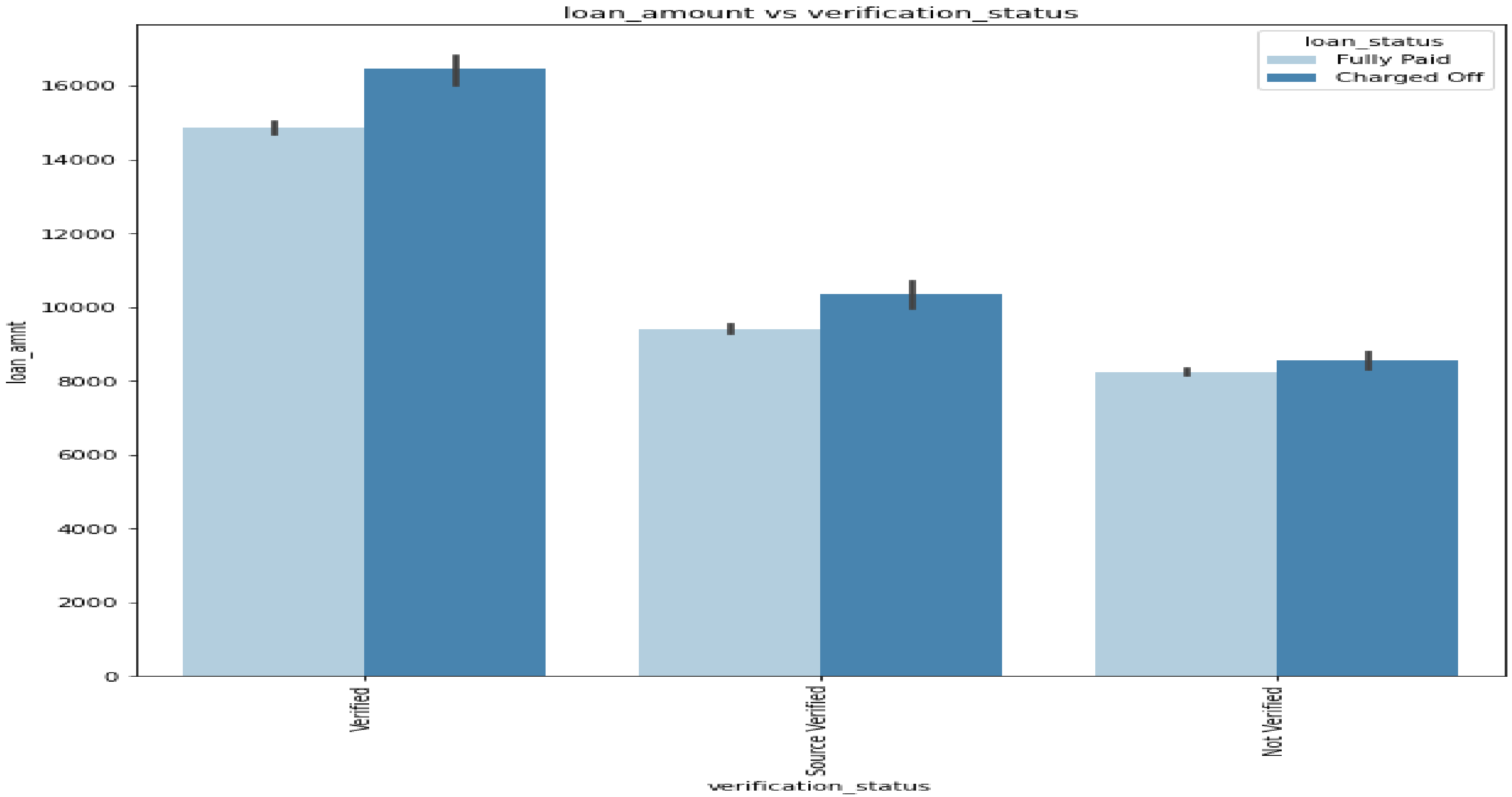
- **Based on the above plot, the loan applicants whose obtained loans for small business and the loan amount in between 14-16k are highly likely to default**

Analyzing Annual Income versus Interest Rate



- **Based on the above plot, the loan applicants having annual income between 70-80k and interest rate at 21-24% are highly likely to default.**

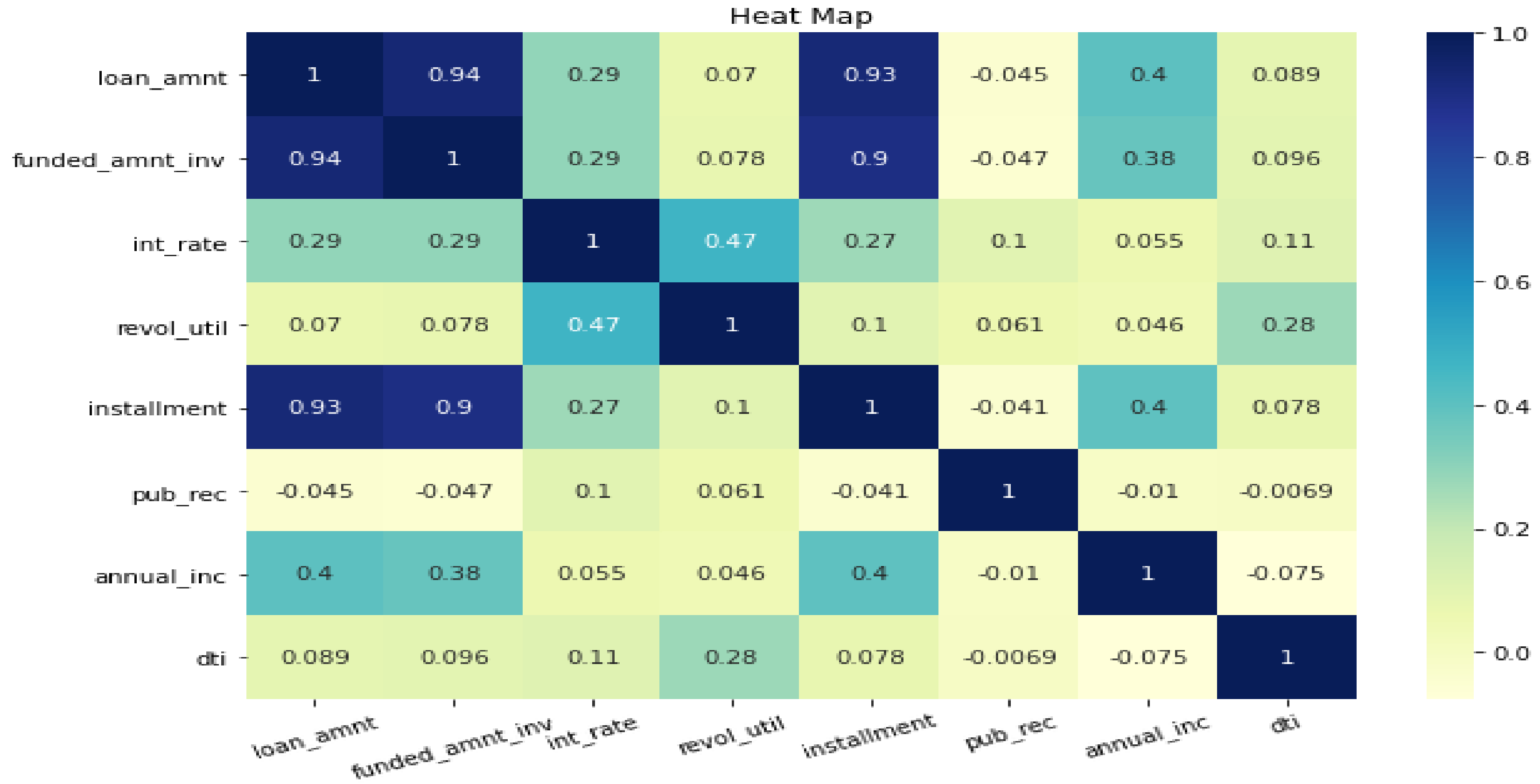
Analyzing the Loan amount versus verification status



Based on the above plot, the loan applicants having the loan status as verified and loan_amount between 14k - 16k plus are highly likely to default.

Heat Map

UpGrad



- **From Above Heatmap we can see that Interest rate has a slight positive correlation with Default rate. Loan Amount, funded_amount and Amount_funded_inv have high correlation among themselves and also with installment which is expected. Income has positive correlation with Amount funded. Income also has slightly negative correlation with Default rate which would imply Higher Income groups would be less likely to default. DTI has negative correlation with Income, As Income increases DTI will decrease.**

Conclusions

In Summary, there is a high chance of defaulting

when the following apply loan:

- People with Grade 'B' apply.
- Applications with tenure of '36 months‘
- Applicants who took loan for 'debt_consolidation'.
- People having house_ownership as 'RENT‘
- Applications submitted in the month of December and November.
- Employees with '<1 year experience
- Applicants having annual income between 70-80k and interest rate at 21-24%
- Applicants whose obtained loans for small business and the loan amount in between 14-16k
- The loan applicants having the loan status as verified and loan_amount between 14k - 16k plus are highly likely to default.

Suggestion:

The strong indicators of default are listed above. We suggest the Lending Club to utilize this analysis to assess and mitigate the risks and there by cutting down the overall credit loss.