Lending Club Case Study

Agenda

- For a consumer finance company which specialises in lending various types of loans to urban customers, two types of risks are associated with the bank's decision:

  1. Applicant likely to repay the loan, not approving the loan results in a loss of business
  2. Applicant not likely to repay the loan, approving the loan may lead to a finan

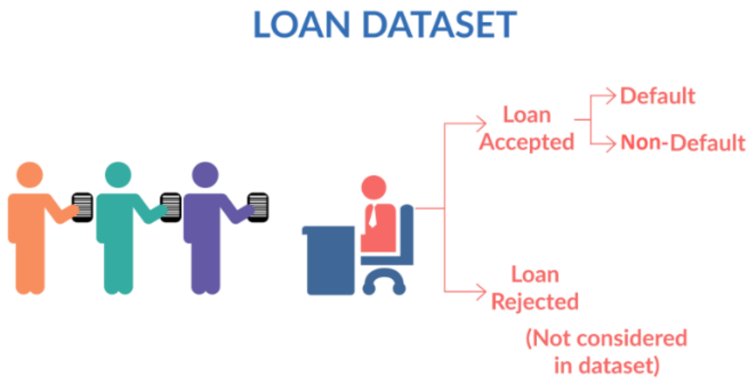


Figure 1. Loan Data Set

- In this case study we have identified patterns which indicate if a person is likely to default
- In this detailed study of consumer attributes and loan attributes that influence the tendency of defaulting is done, from previous Lending Club data to find if a person is likely to default or not .

Assumptions Made

- For applicants with employment length as >10 years we have considered as 10 years and for <1 year as 0
- For few applicants the annual income data seemed to be abruptly high, so we have considered threshold for removing outliers as 0.95 percentile.
- Home Ownership data for 3 applicants was missing and so considered them as other for the simplicity of analysis.

Approach

- Data Loading & Data Understanding
  - Read Data File Available
  - Understand Data Dictionary

- Processing Steps

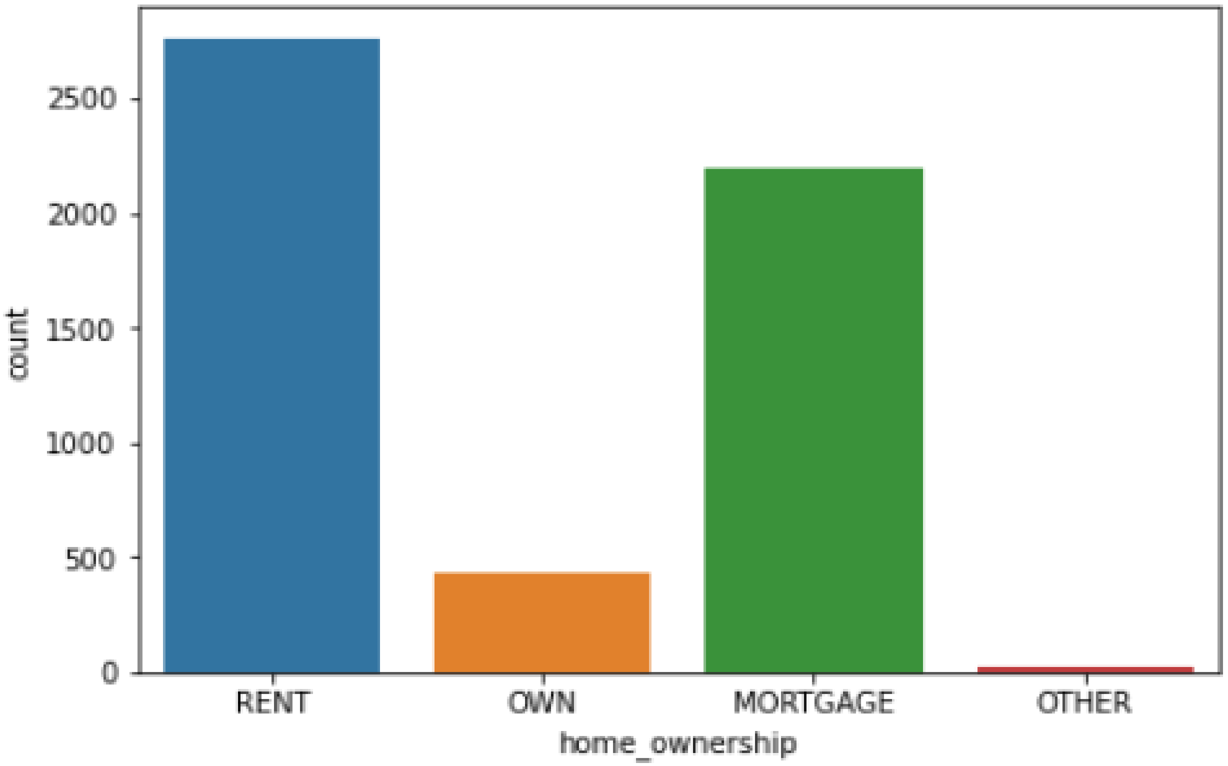
  Removal of unnecessary columns not important for the purpose of the current analysis.

- Missing Value Check
    - Decide a threshold of missing percentage
    - Drop the columns having missing values more than threshold
    - Imputation Techniques:-
        - Continuous
            - Mean(No Outliers)
            - Median(Outliers)
        - Categorial
            - Mode
- Outlier Treatment
    - Box Plot
    - IQR Method
- Exploratory Data Analysis
    - Univariate Analysis
    - Bivariate Analysis
    - Segmented Univariate

Exploratory Data Analysis

On comparing between applicants who have defaulted & applicants who have fully paid their loan below categories shows a higher trend of defaulting their loans:-

Categorical Data Analysis



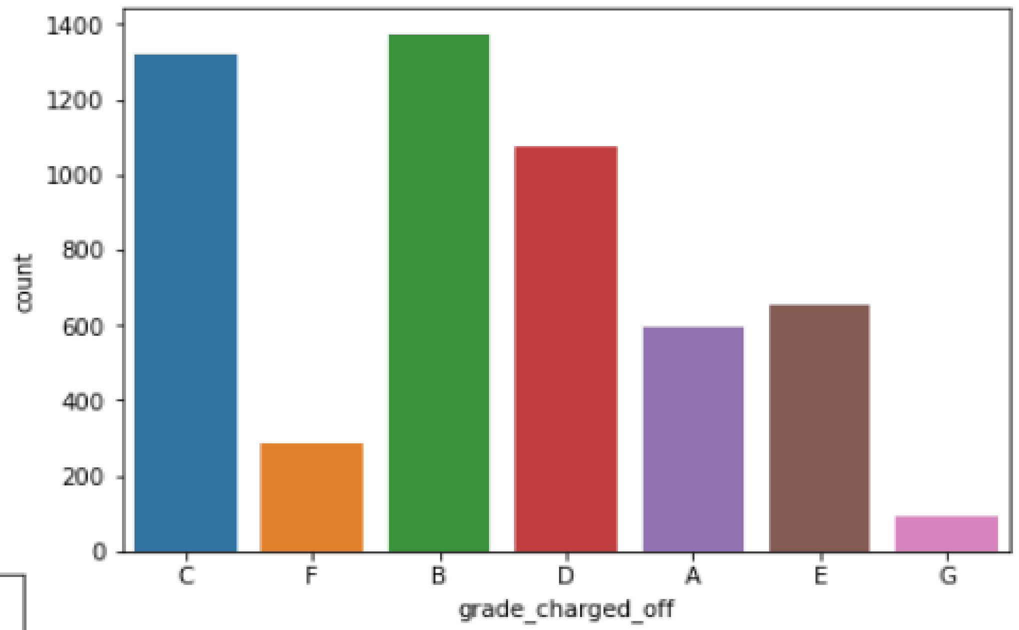
Home Ownership

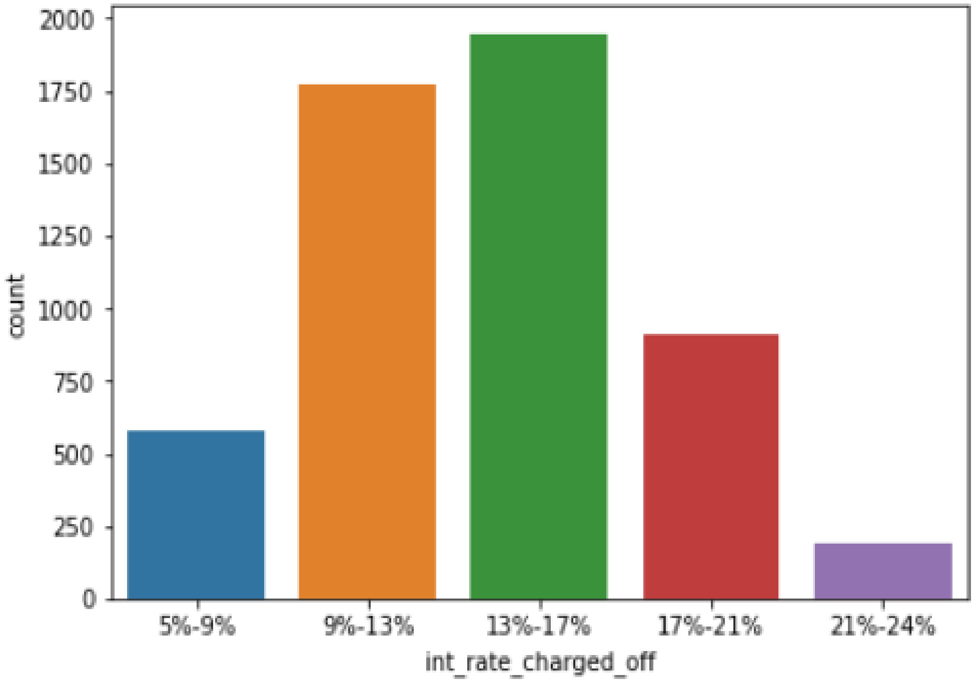
The applicants who mostly are on rent are found to default mostly.

Grade

Mostly the ‘B , C’ grade applicants are found to default mostly.

.

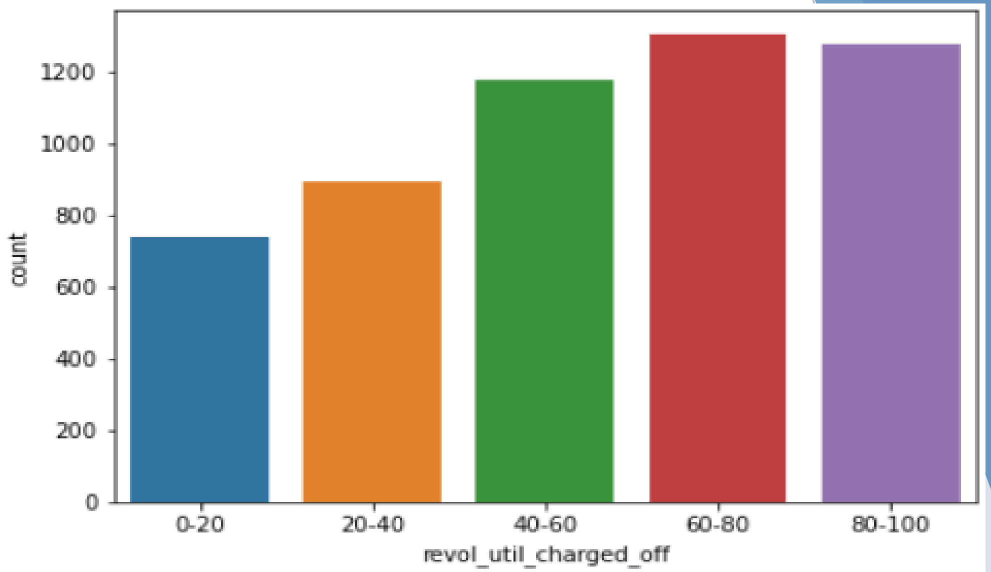




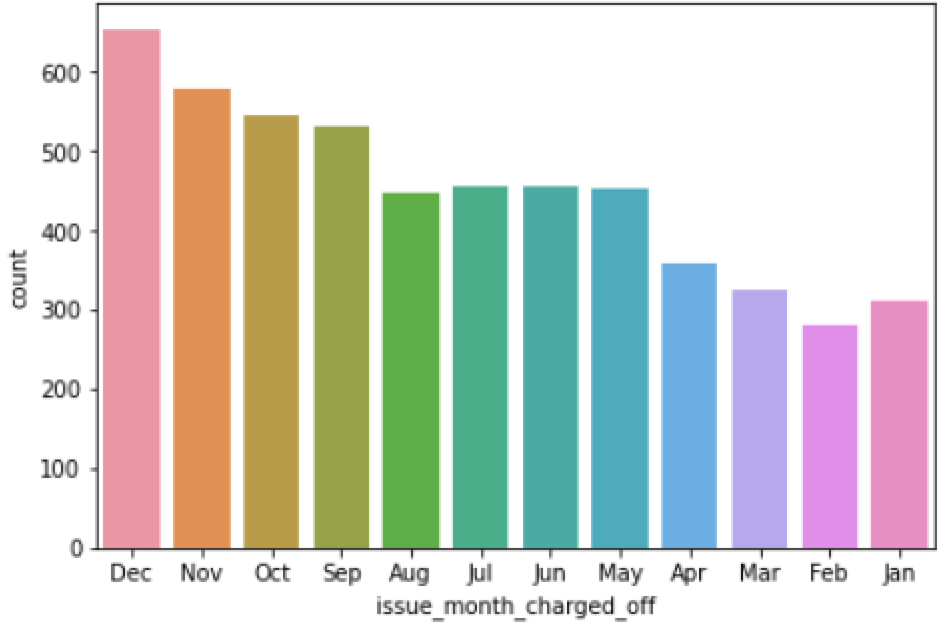
Interest Rate

Applicants who are charged 13-17% of interest rate are found to default mostly

Revolving line utilization rate

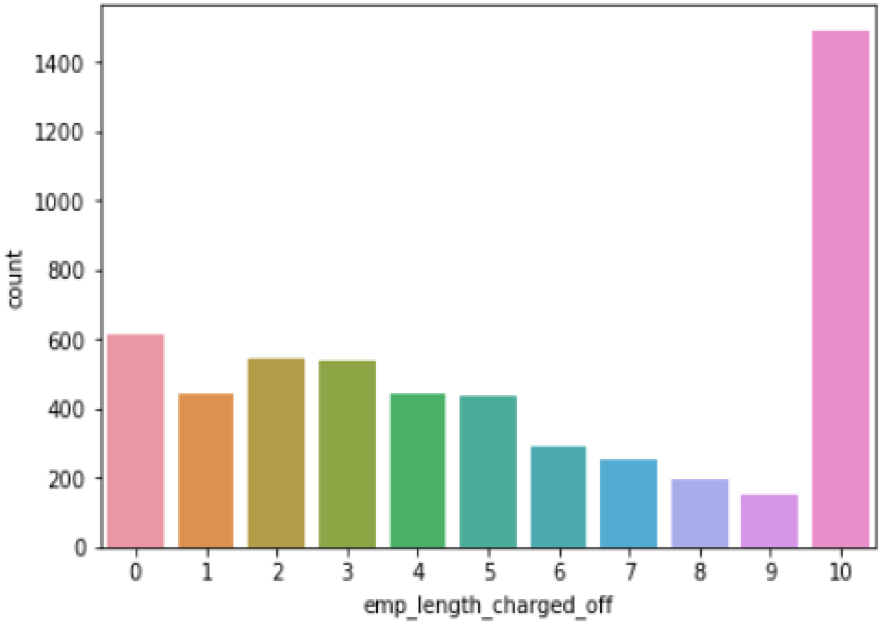


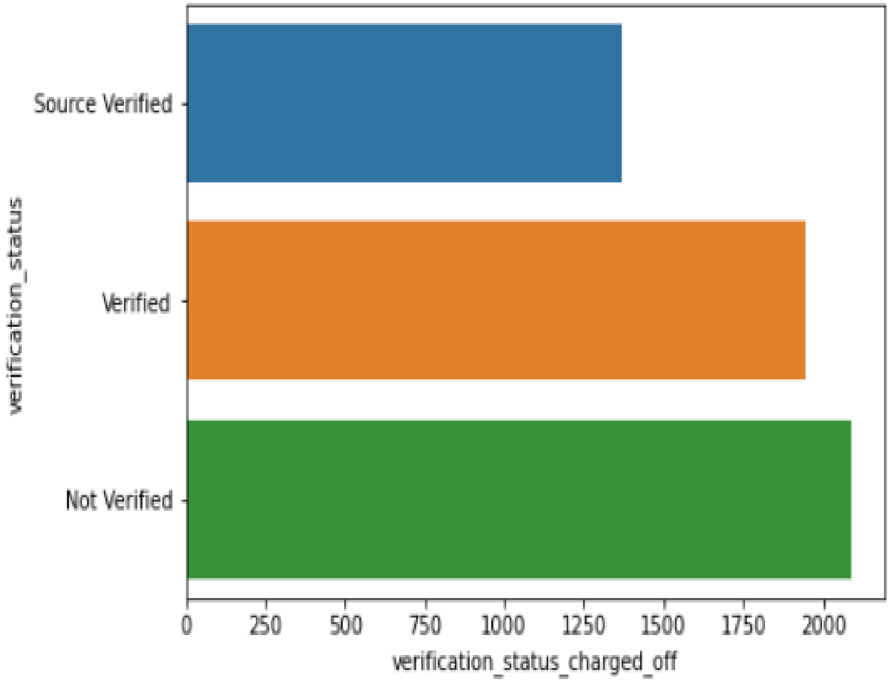
Revol_Util for applicants between range 60-80 shows a trend to default



Loan Issue Month

Loan applications mostly in Dec or ending of the year have defaulted mostly

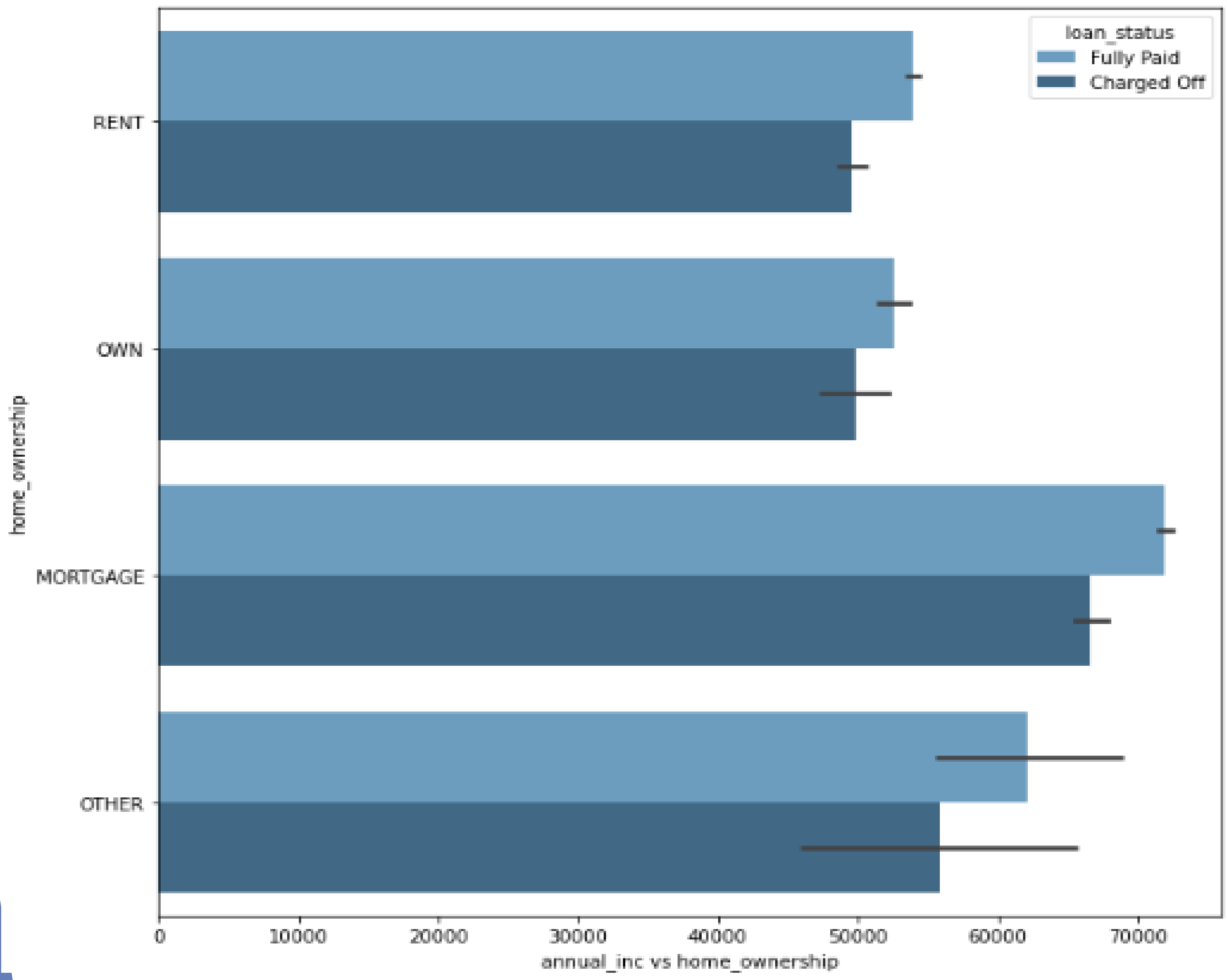




Also below are some trend which is visible both for defaulters and full paid applicants so there is not much concluding significance to note from them:-

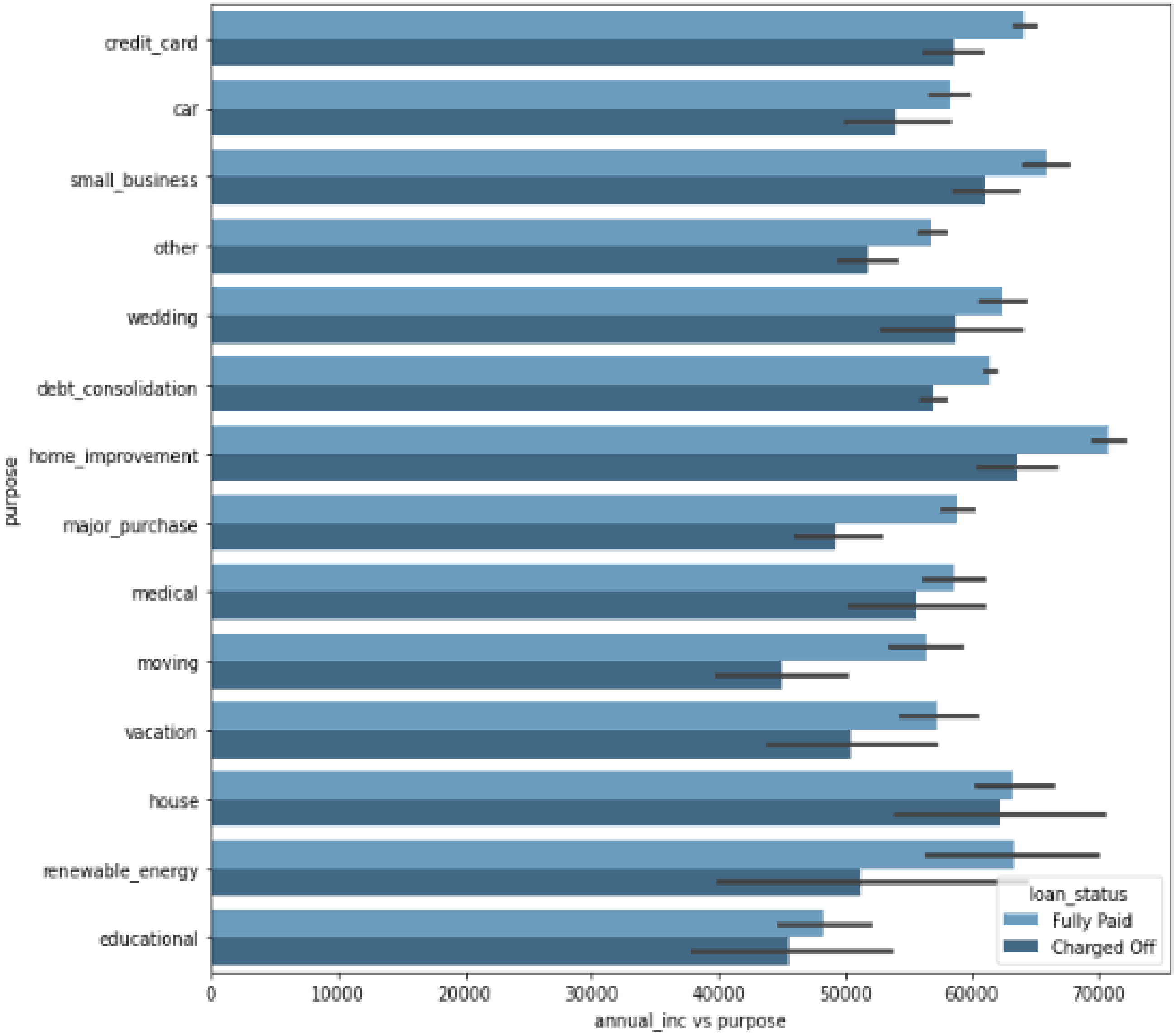
- The applicants who mostly apply for loan for debt consolidation.
- Applicants having <=10 yrs. of employment length
- The verification status is not verified
- Open account 2-10
- Annual income 31k – 58k
- Dti between 12-18

Annual Income vs Home Ownership



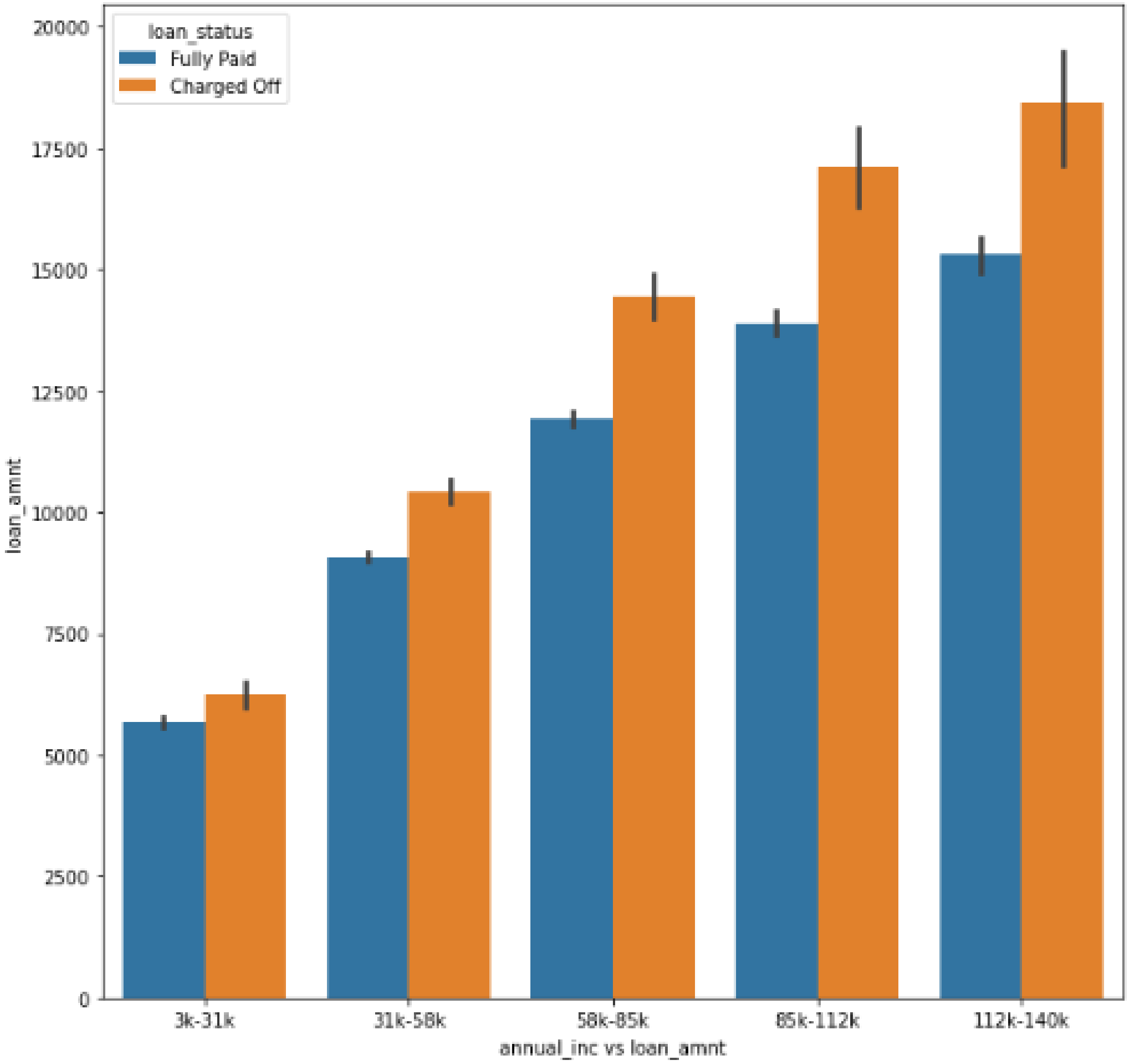
The applicants whose home ownership is 'MORTGAGE' and their income range fall between 60-70k has shown the highest trend to default their loans

Annual Income Vs Purpose



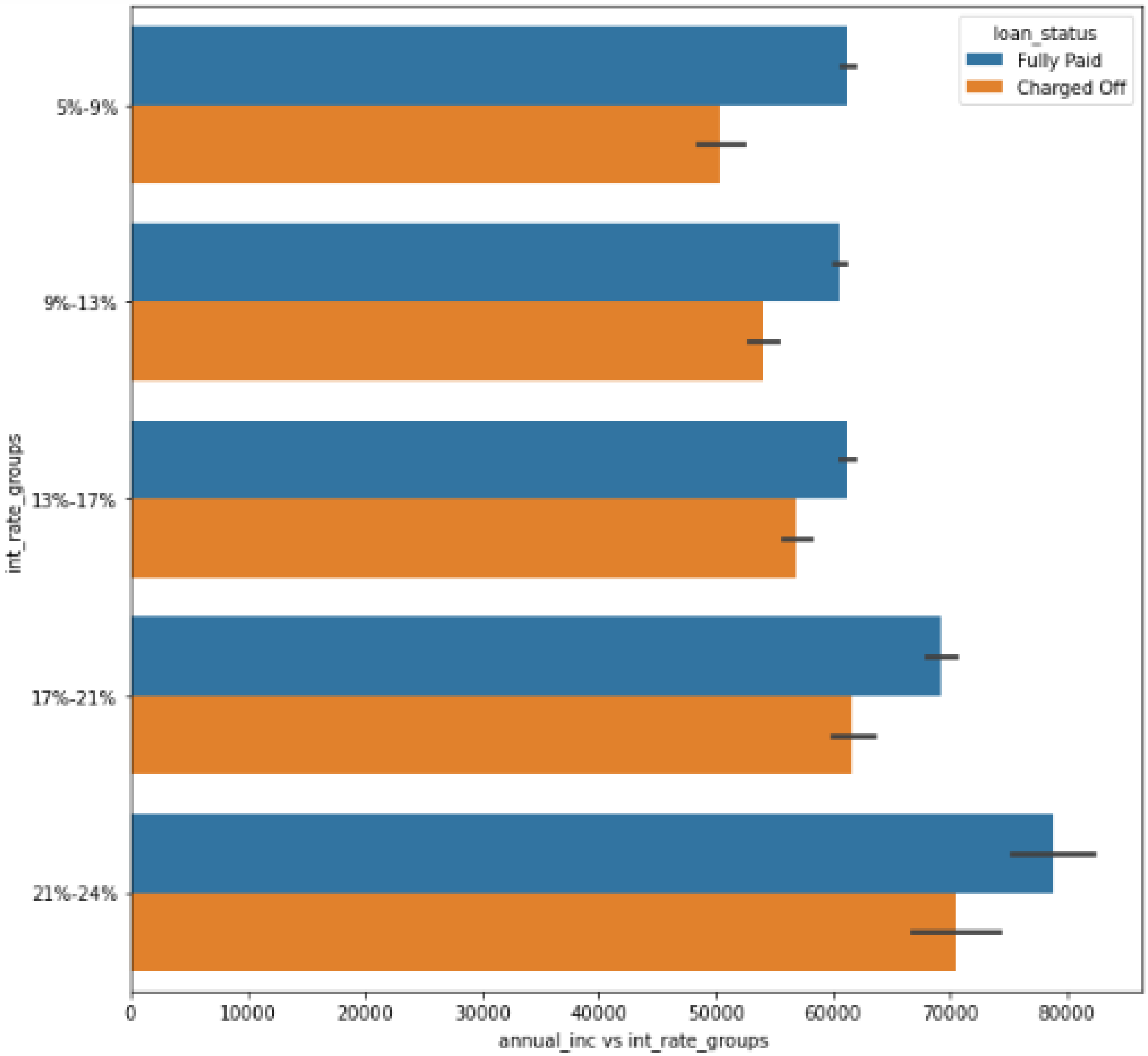
The applicants with high annual income tend to apply for loan with purpose as ‘home improvement’ , ‘house’ , ‘renewable energy’ , ‘small business’ also the defaulters are mostly found to be showing purpose as ‘home improvement’ , ‘house’ , ‘small business’.

Annual Income Vs Loan Amount



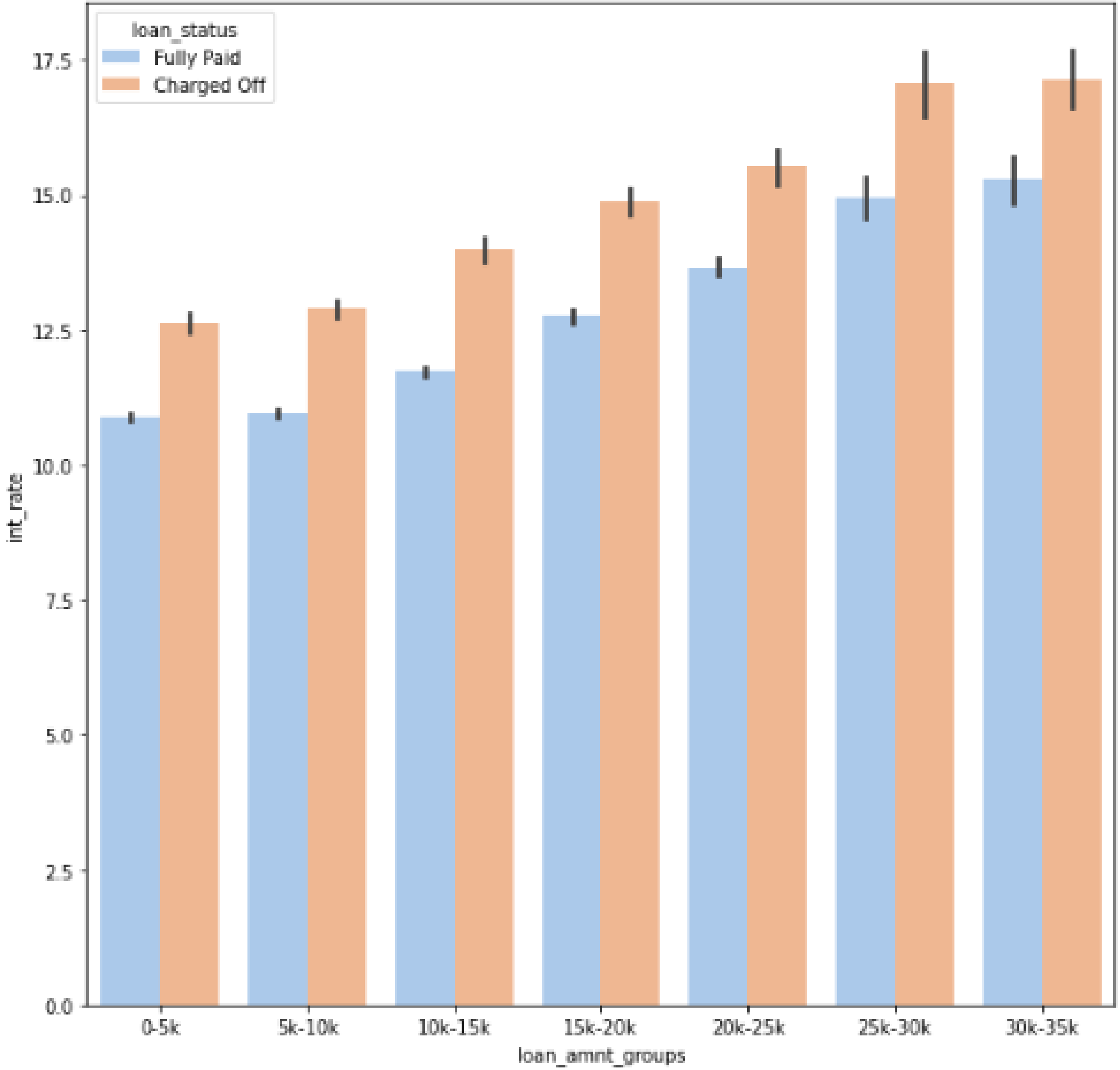
Applicants whose annual income is between 112k-140k and take up the loan amount ranging between 17.5k – 20k tend to default more

Annual Income Vs Interest Rate



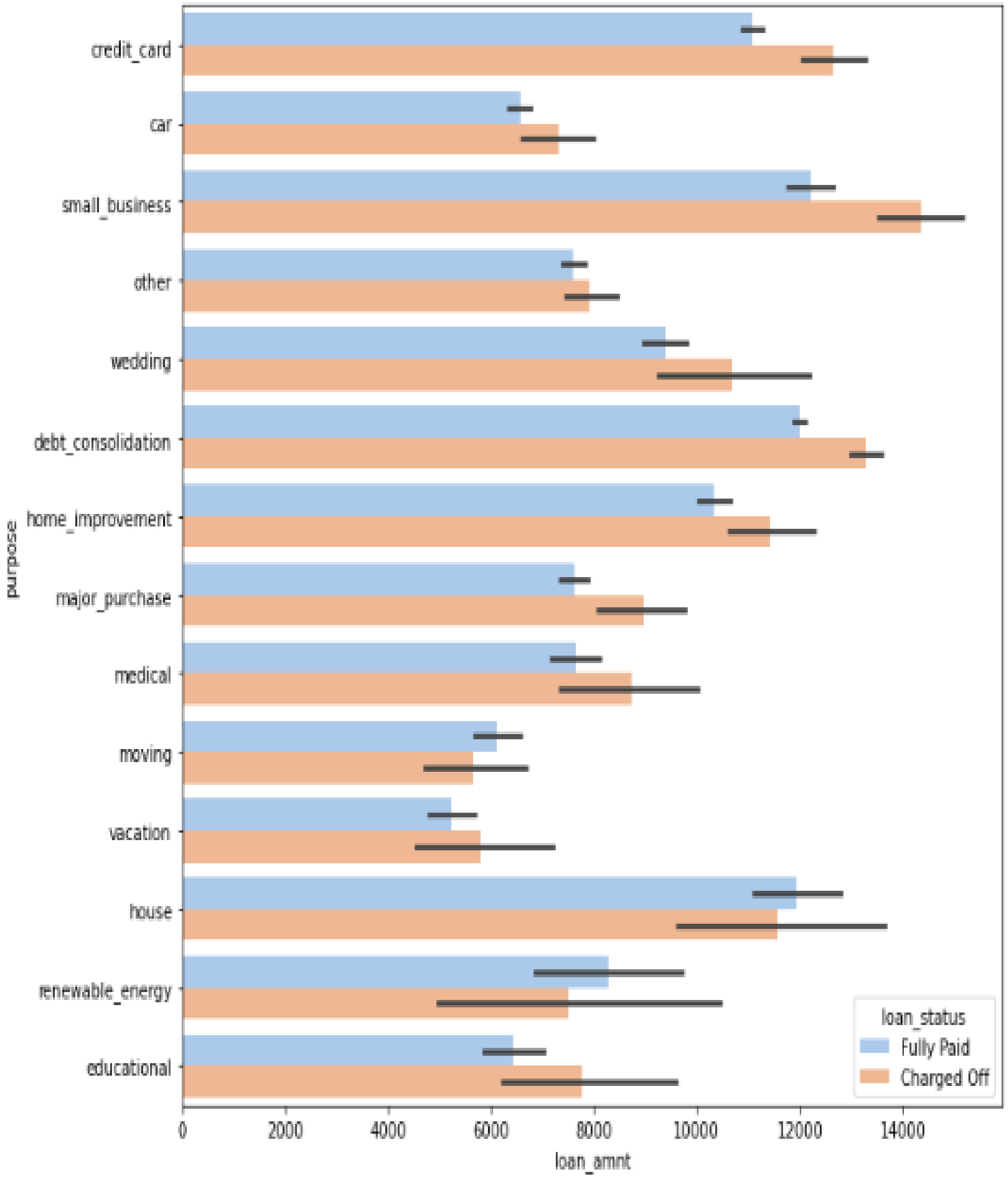
Applicants of annual income range between 70k – 80k and take up loan at an interest rate of 21-24% tend to default more.

Interest Rate Vs Loan Amount Group

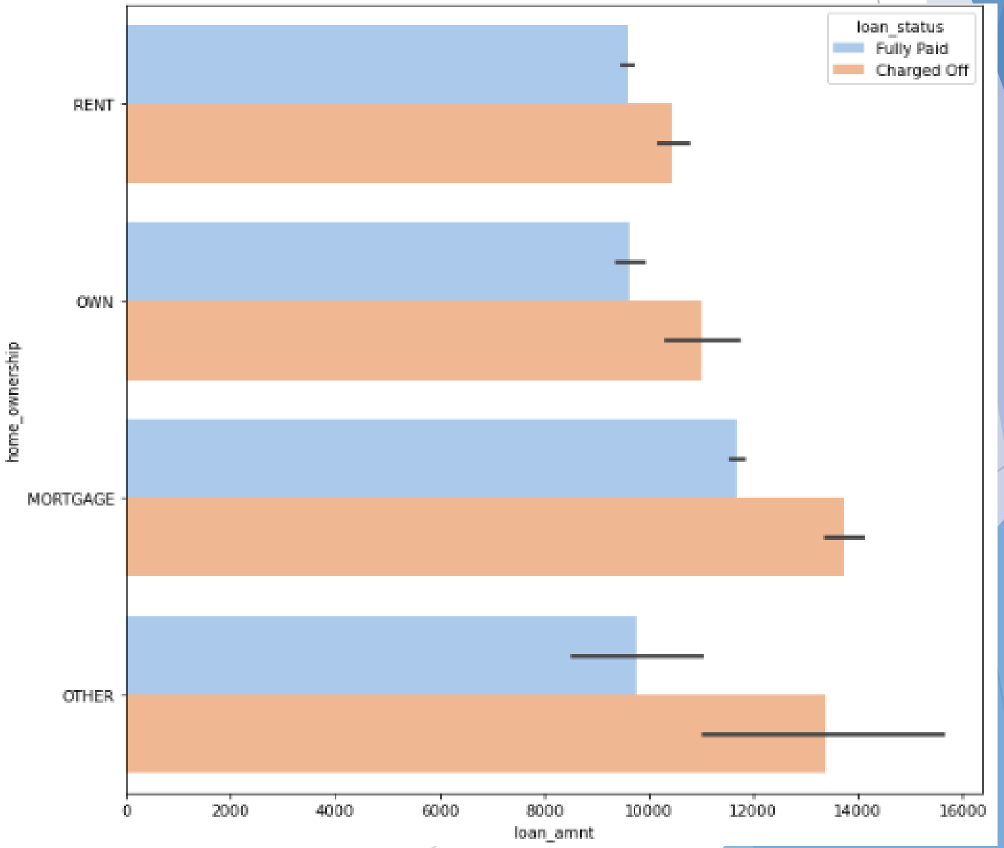


Applicants applying for loan amount between 25k-30k and 30k-35k and interest rate charged around 17.5% tend to default more

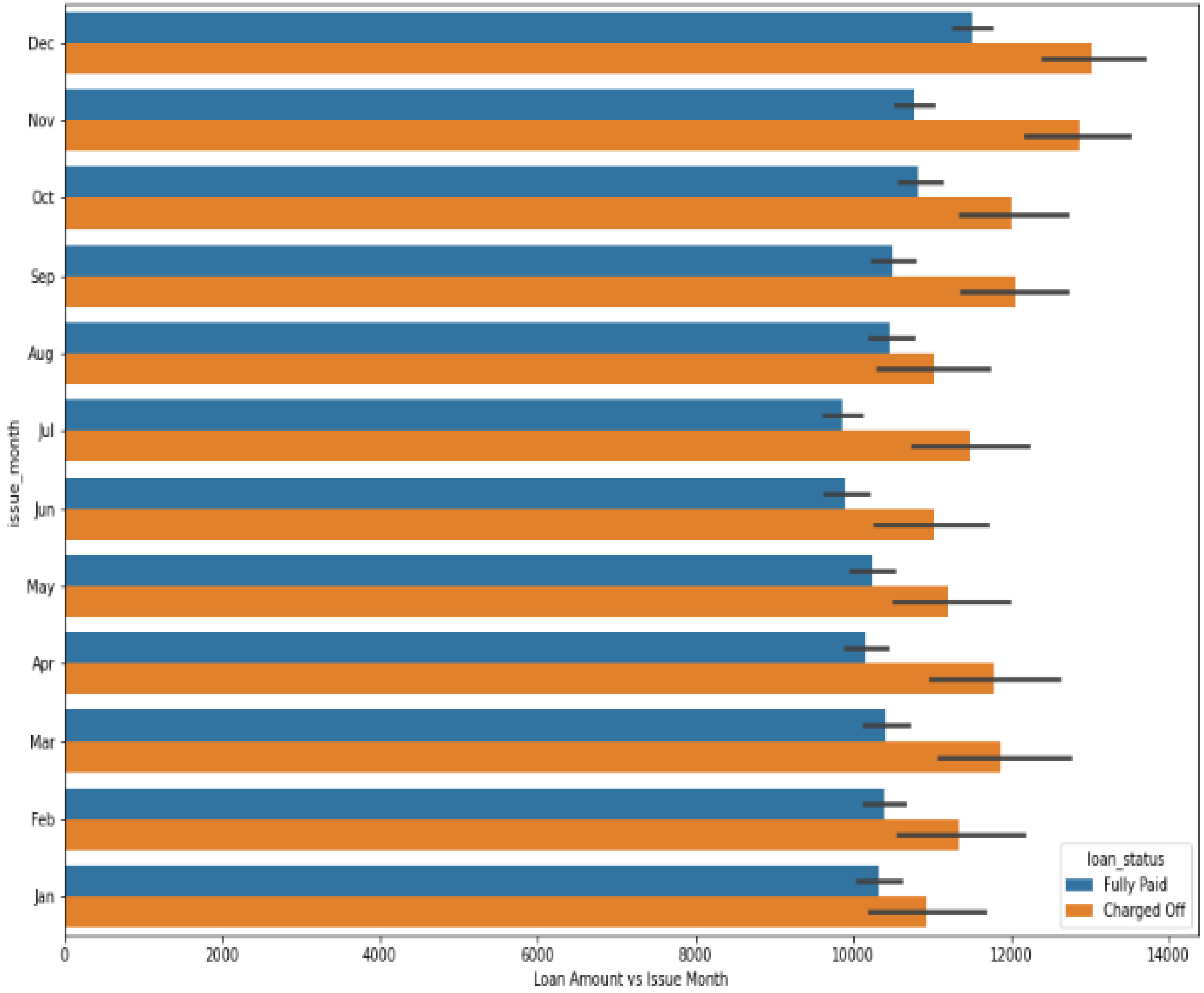
Loan Amount Group Vs Purpose / Home Ownership



Applicants who tend to default more apply for loan amount of around 14k for the purpose of small business or debt consolidation and are found to stay on Mortgage.

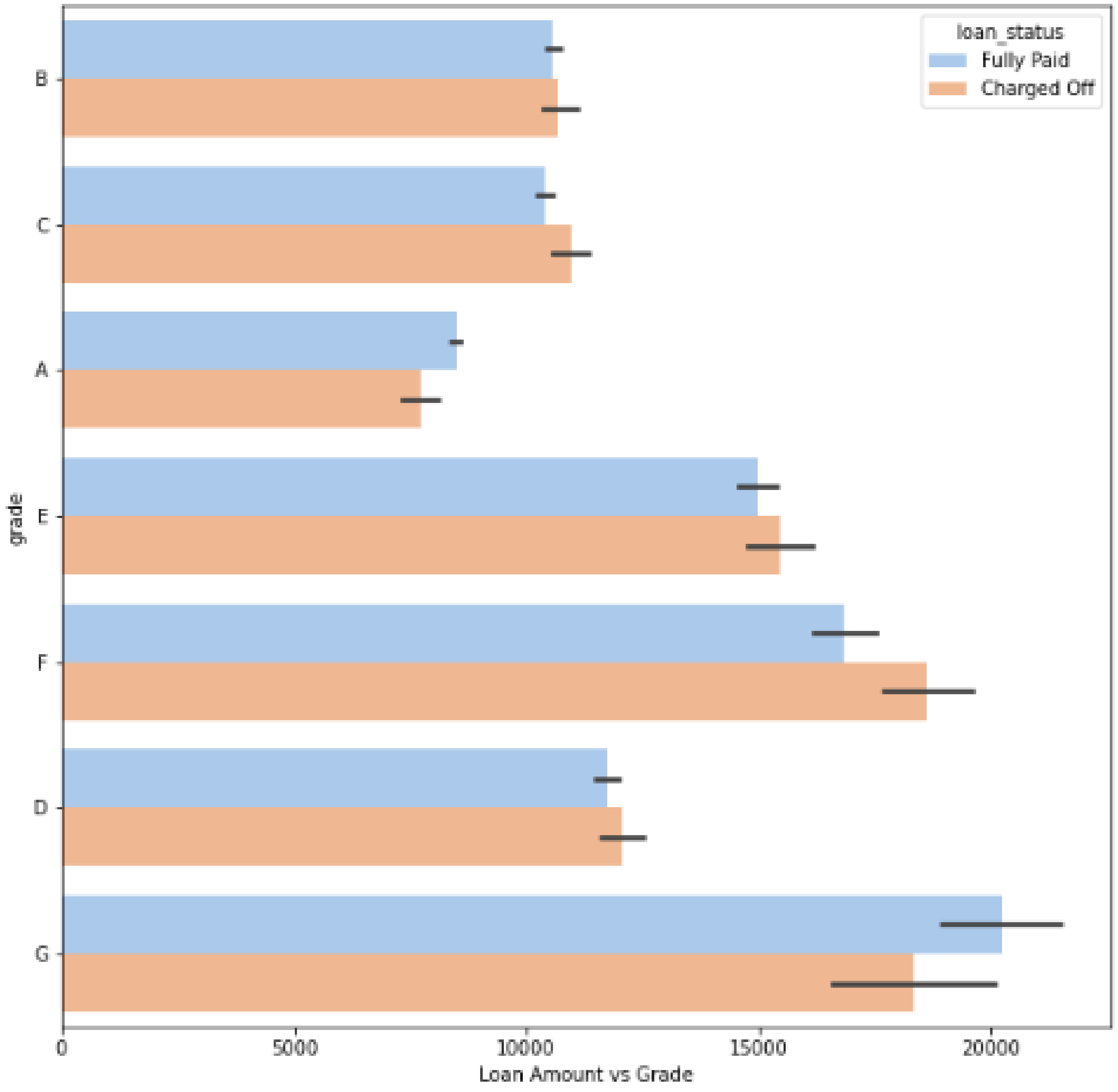


Loan amount vs Month Issued

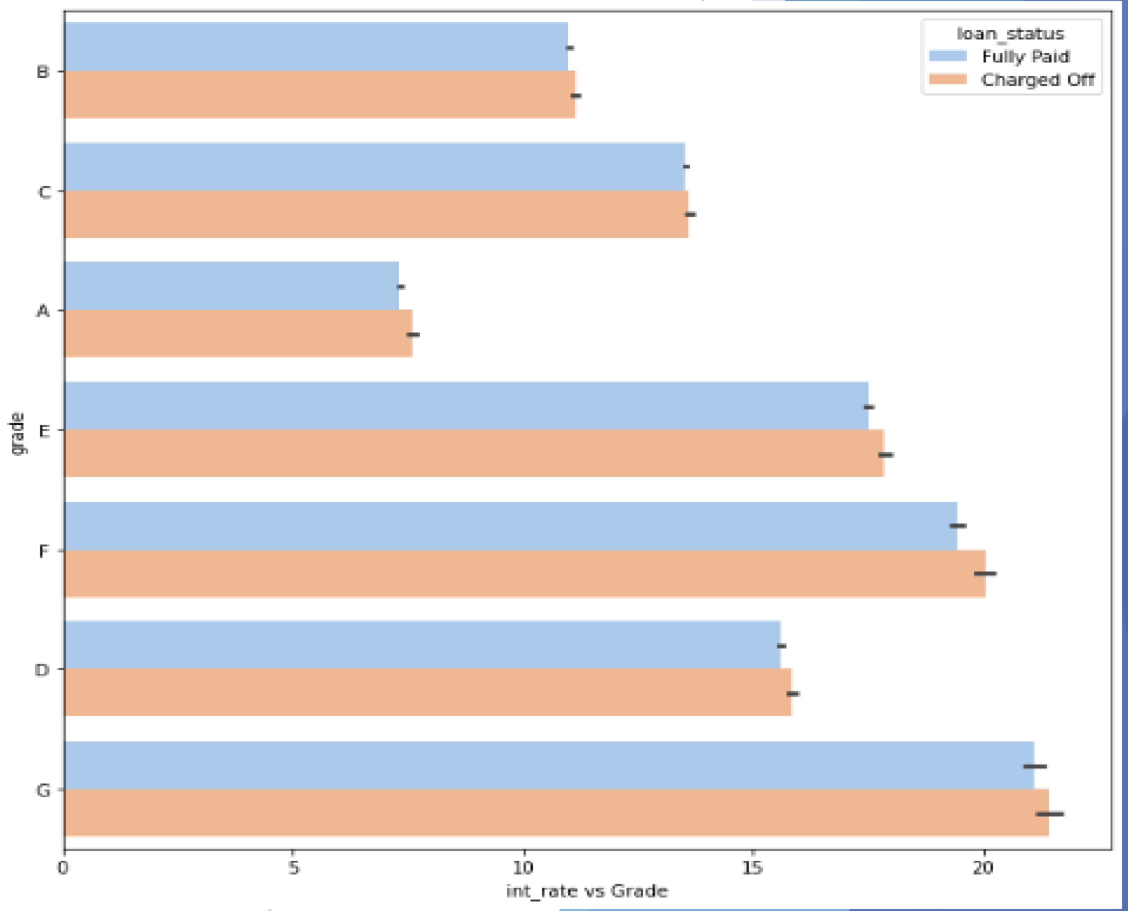


Applicants mostly defaulted when they applied for loan in the month of December during year ending and mostly for loan amount of around 12k – 14k

Grade Vs Loan Amount Group / Interest Rate



Applicants who defaulted mostly fall in Grade G,F category and the loan amount taken is 15k-20k with high interest rate of around 20%



Recommendation

- Profit of the organisation though earned through the interest the borrower pays, **the rate need not be too high (near or above 20%)** specially for G,F grade of applicants when they apply for high amount (15k-20k) loan. The same is also applicable for all others having high annual income.

- **MORTGAGE Home Ownerships** having annual income of around 70k and applying for around 14k loan, for the purpose of small business or debt consolidation, their expenditure should be scrutinised thoroughly as they show a high percentage of defaulters. They also tend to apply for loan at the end of the year.