LENDING CLUB CASE STUDY

Case Study Submitted By:

M SOMASEKHARA REDDY(Group Facilitator)

SABYASACHI RATH

Problem Statement

A consumer finance company specializes in lending various types of loans to urban customers. When the company 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:

- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company
- The data used in this case study contains the information about past loan applicants and whether they 'defaulted' or not.
- The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

Assumptions

- The charged off customers are mostly the ones who have employment length of less than 1 year
- The states mentioned in data are the states of US.
- We have assumed a 95% confidence interval for removing the outliers from Annual income column

Solution Approach

Data Loading and Data Understanding

- Load the Data to Jupyter File
- Understand columns from Data Dictionary

Data Cleaning

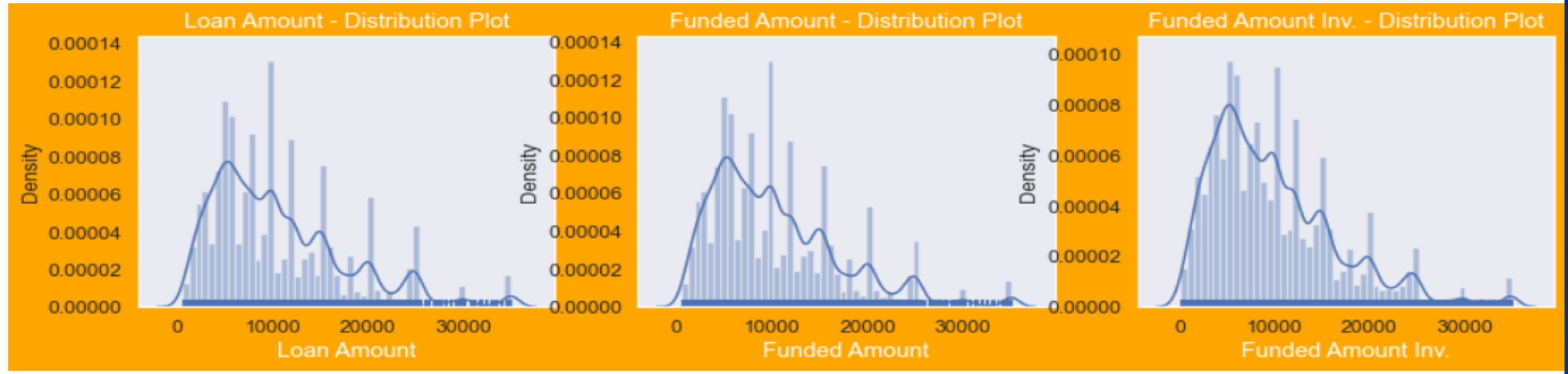
- Remove the current customers
- Remove duplicate rows
- Remove the blank rows and columns
- Remove columns with more than 30% blank values and single value
- Remove rows with more than 10 blank values
- Impute missing values for rest of columns
- Outlier Treatment

EDA

- Univariate Analysis
- Segmented Univariate Analysis
- Bivariate Analysis

Univariate Analysis

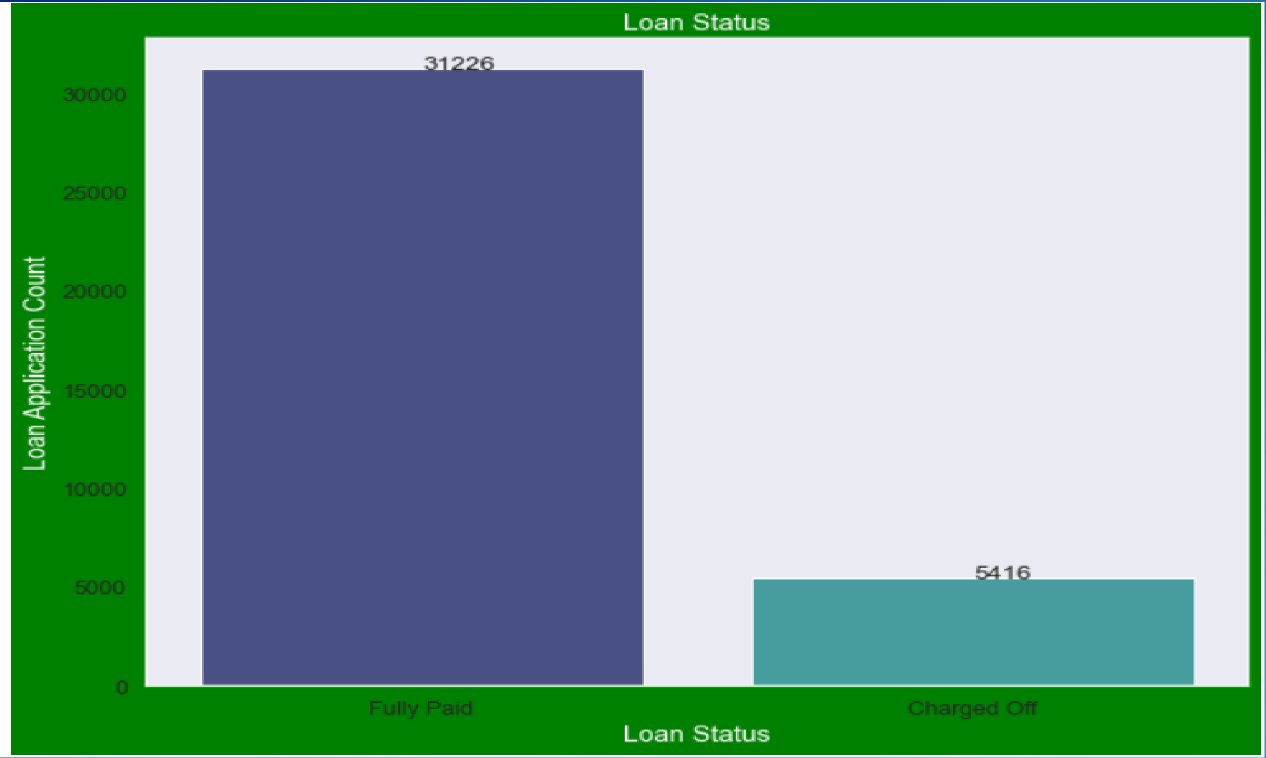
Univariate analysis using Distribution plot on Loan amount, funded amount and Funded amount by investor.

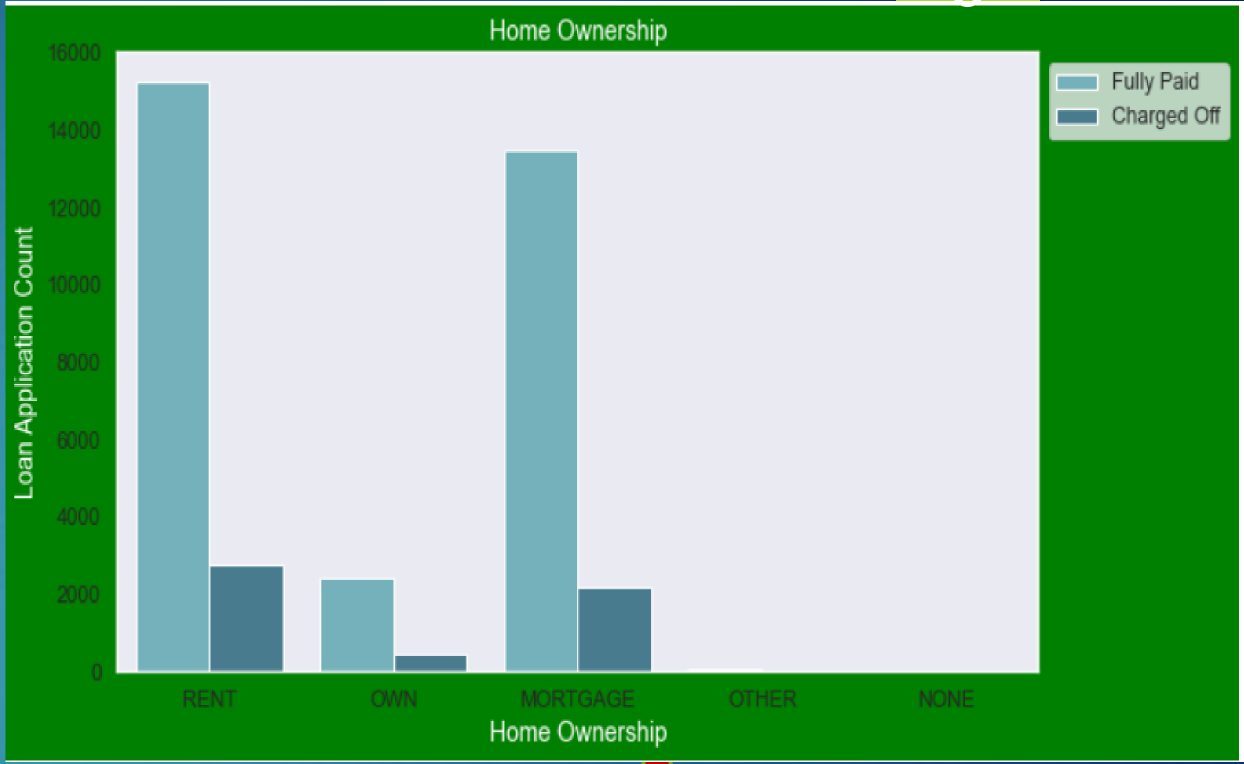


Observation:

- Distribution of amounts for all three looks close to similar.
- We have worked with only funded amount column for further analysis.

Segmented Univariate Analysis

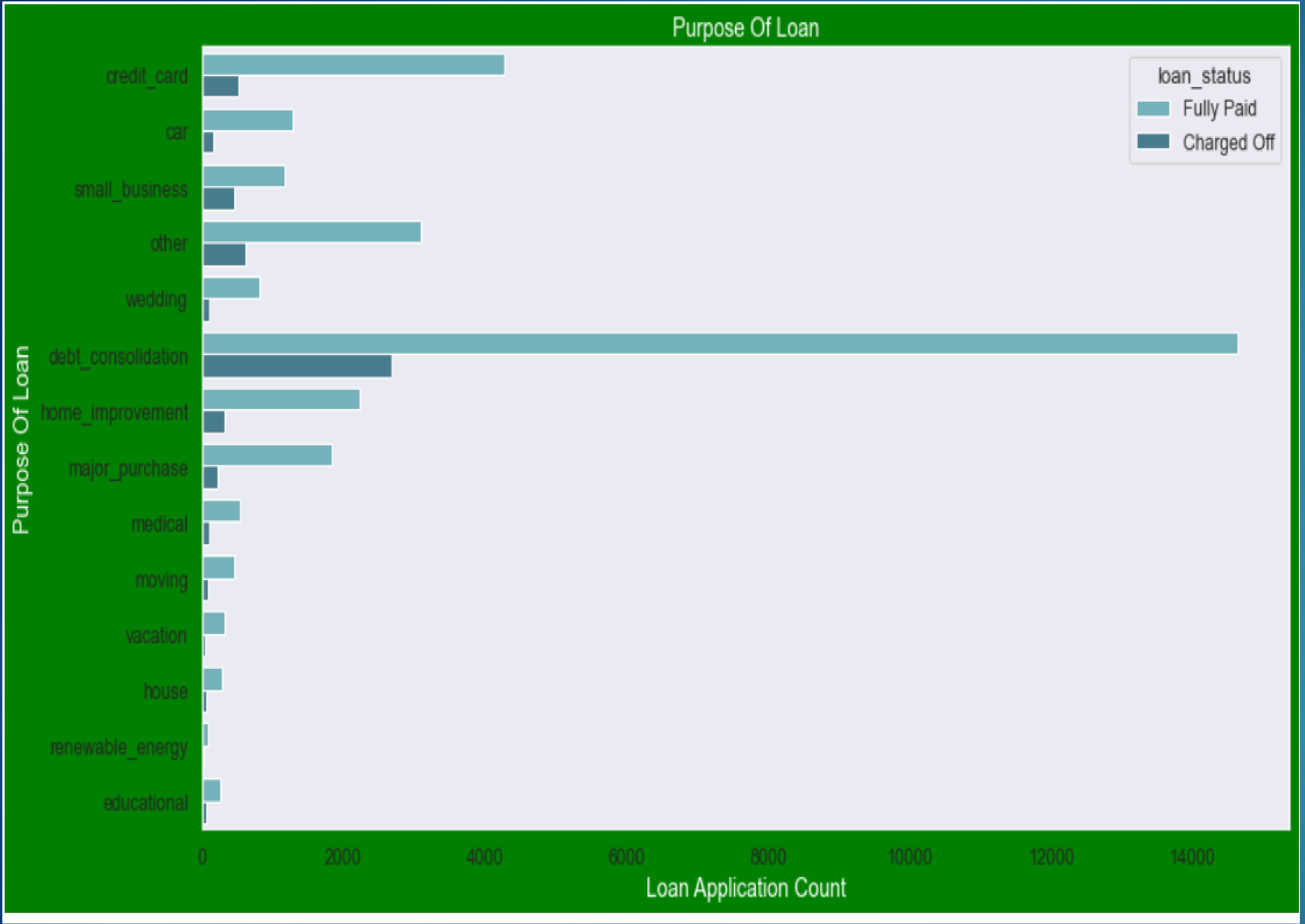




Close to 15% loans were charged off out of total loan issued

- Above plot shows that most of them living in rented home or mortgaged their home.
- Applicant numbers are high from these categories so charged off also high.

Segmented Univariate Analysis

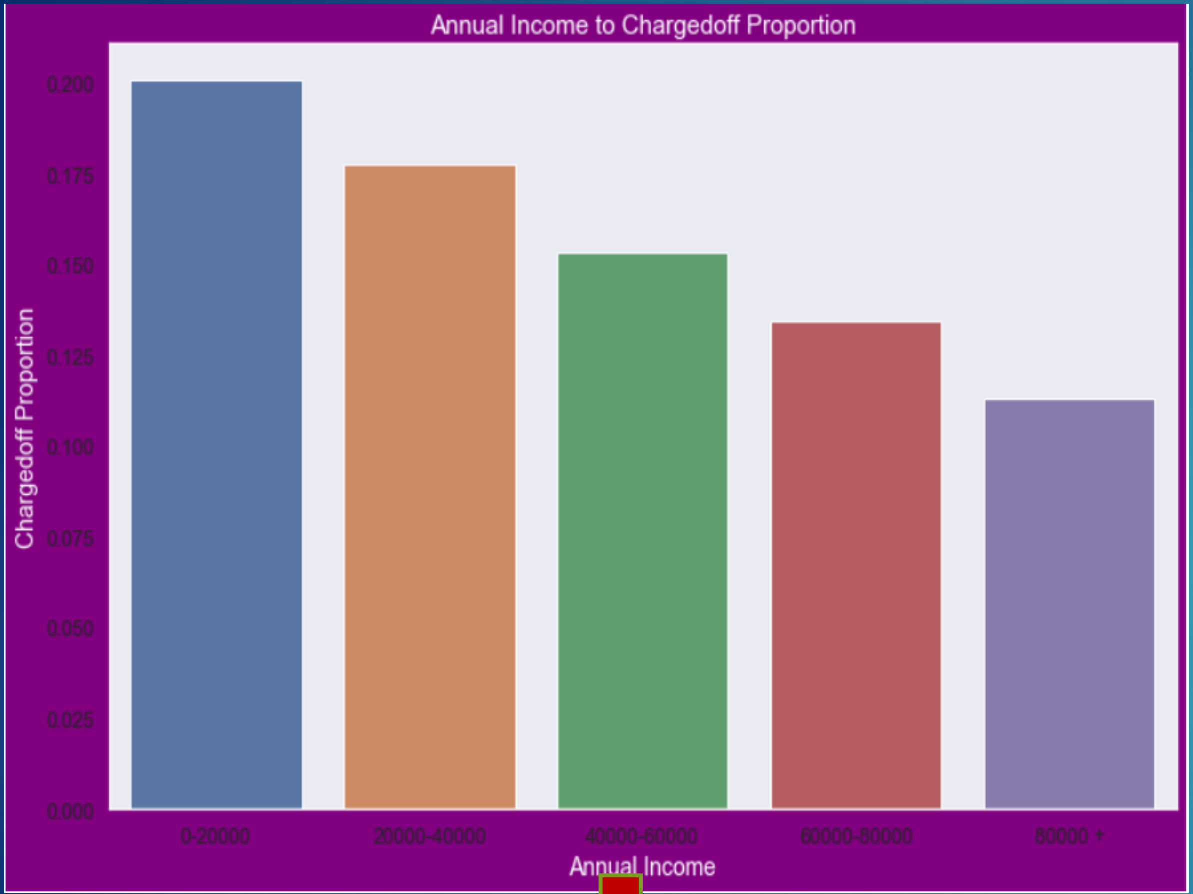


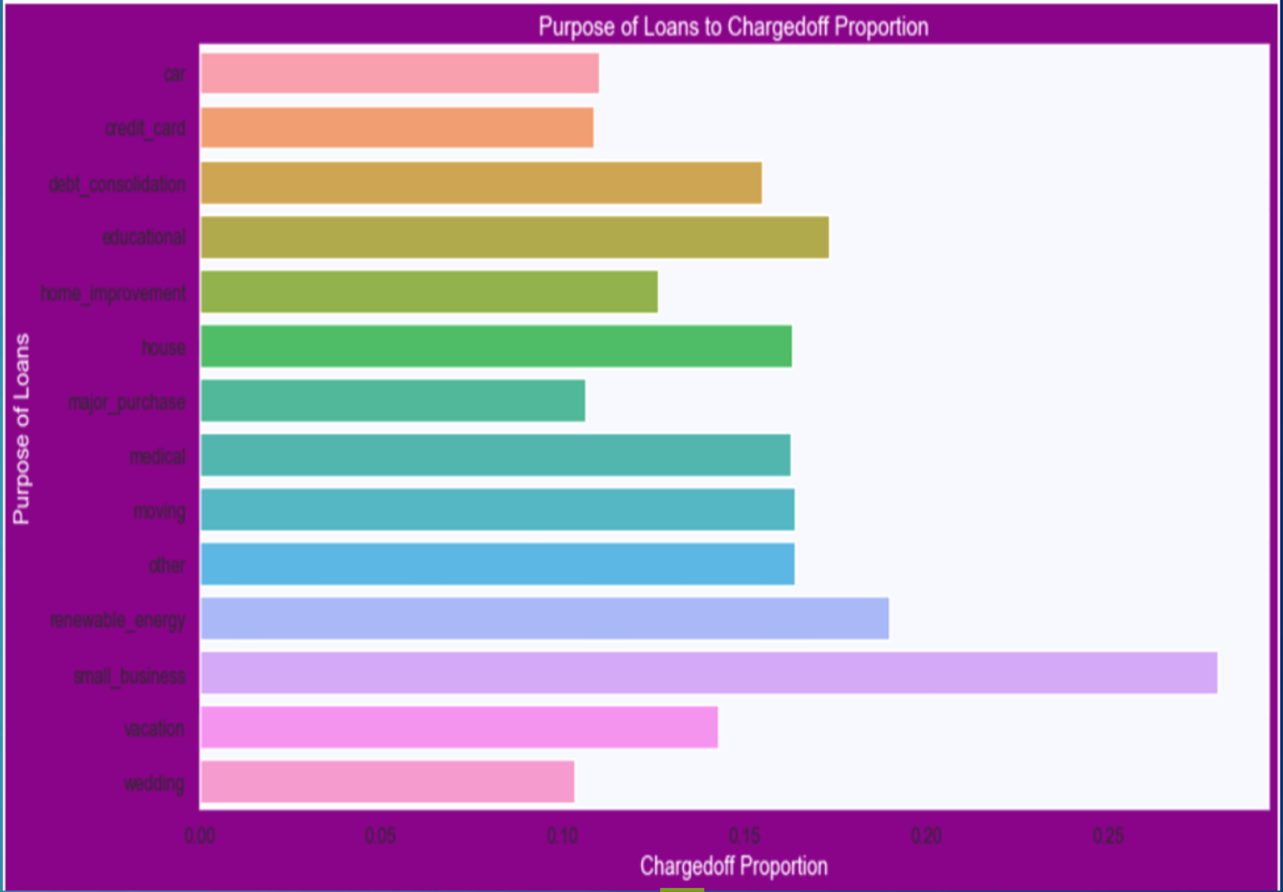
- Plot shows loans were taken for the purpose of debt consolidation
- Number of charged off count also high for these loans.

Segmented Univariate Analysis

- Count of loan application is increasing year to year.
- Increase in number of loan applications are causing to more number of charged off applications.

Bivariate Analysis

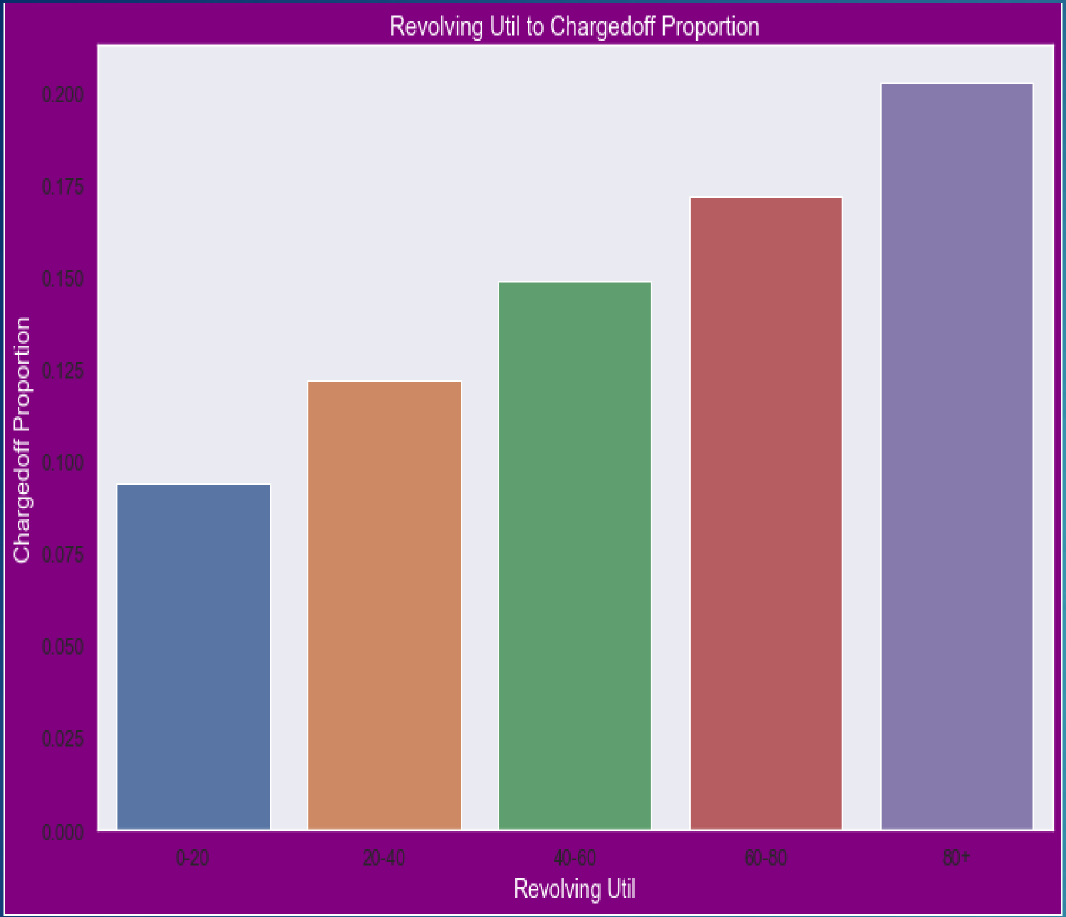


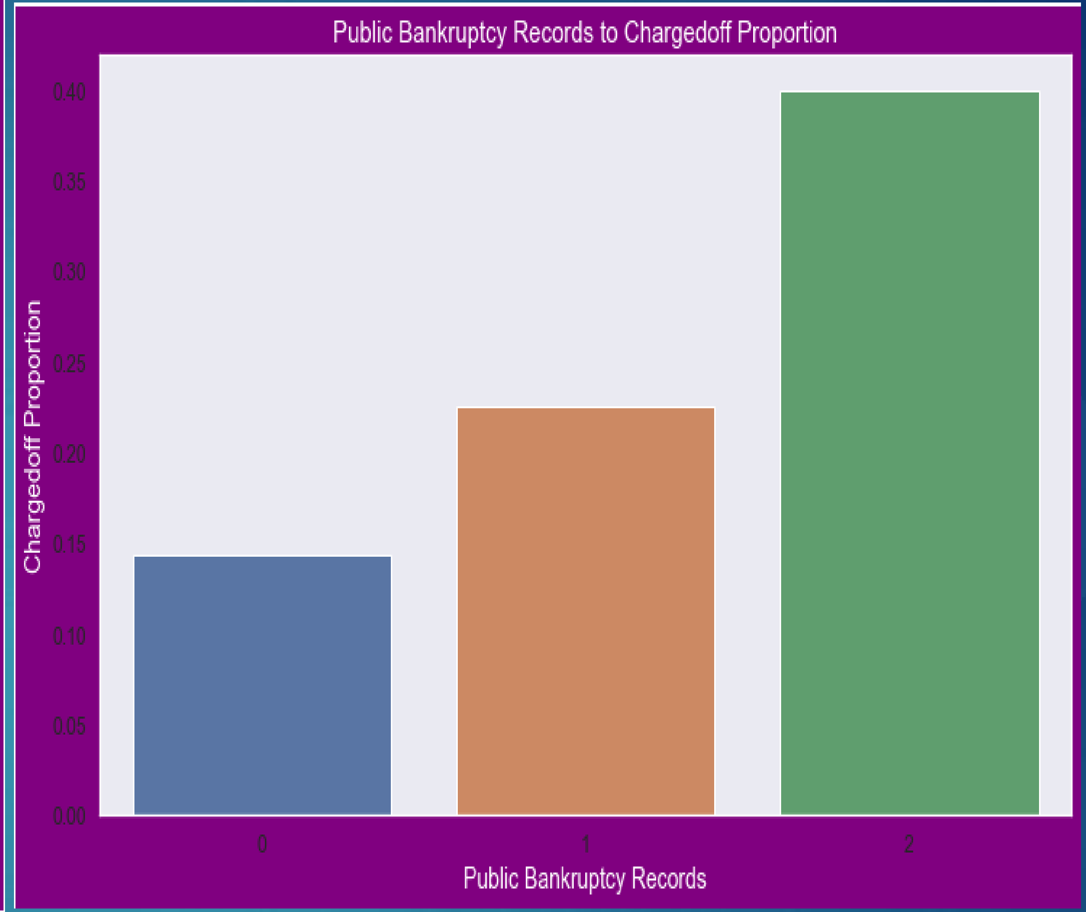


- With increase in annual income charged off proportion decreases.

- Small Business applicants have high chances of getting charged off.

Bivariate Analysis

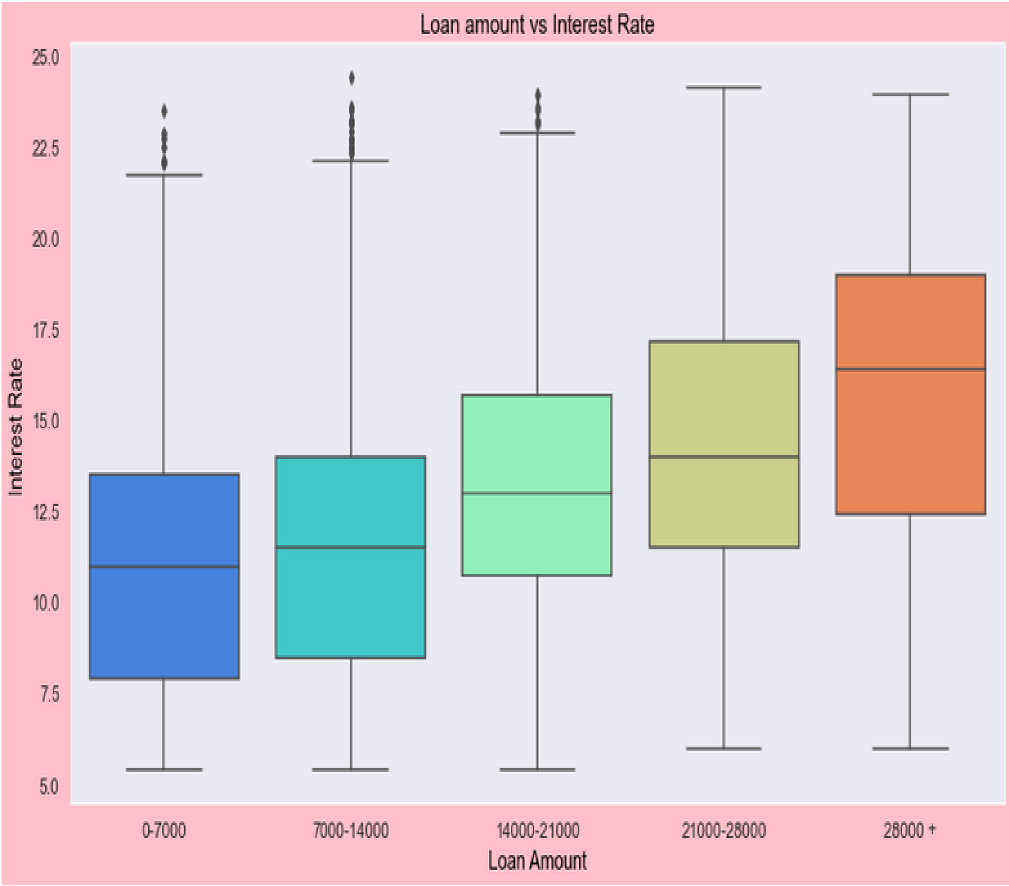


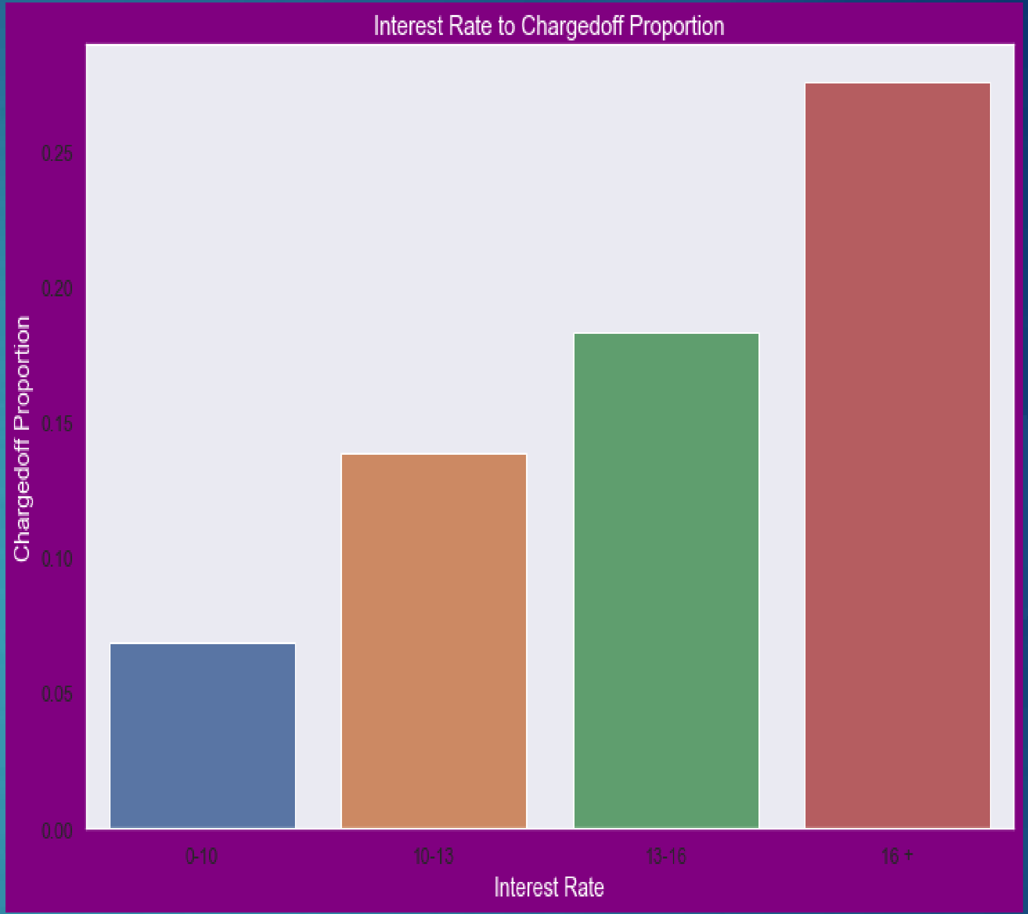


- **Chances of charged off is increasing linearly with revolving line utilization rate.**
- **Revolving utilization rate > 80% maximum possibility of charged off.**

- **People who have defaulted before are more likely to default in future as well.**

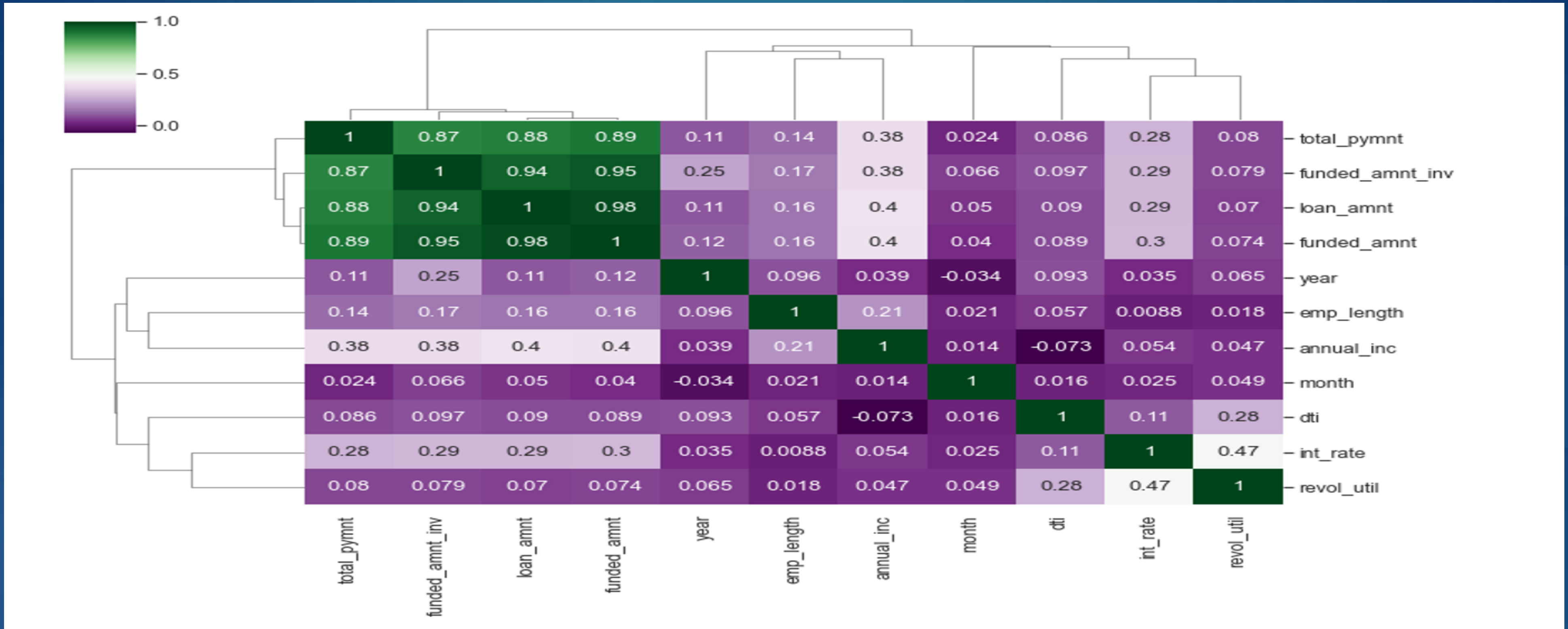
Bivariate Analysis





- Increase in Loan Amount increases the Interest Rate
- Increase in Interest Rate results in more charged off customers

Heat Map for Correlation



- Loan amount-funded amount, loan amount-funded amount investor and funded amount-funded amount investor have strong positive correlation.
- Positive correlation between annual income and employment years.
- DTI has limited impact on loan amount as per the dataset.
- Annual income and Loan amount are positively correlated.

Conclusion

- Increase in number of loan applications means increase of number of charged off
- Increase in annual income results in decreased charged off proportion.
- Interest rate charged by the company increases with increase in loan amount.
- Increase in Interest Rate increases the number of Charged off customers.
- The company can charge lower interest rate to higher income group to decrease the percentage of Charged off customers
- Purpose of Loan Debt Consolidation – High Charged Off, High risk, charge higher interest rate
- Bankruptcy record customers – High Risk, Lower Loan Amount and Higher rate
- Loan amount-funded amount, loan amount-funded amount investor and funded amount-funded amount investor – Strong Positive correlation.