RISK ANALYTICS IN BANKING AND FINANCIAL SERVICES
(EXPLORATORY DATA ANALYSIS)

By-
Mohd Yasoob Haider
Amruta Kale

Objective

- To analyze the data of Borrowers and required attributes using EDA.

  Based on the observation of data analysis,

- Identify risky loans and non-risky loans.
- Identify driving factors for loan approval.

Approach

- Import the required libraries and read the CSV file for analysis

- Data cleaning and Manipulation:
    i. Removed all the columns with null values.
    ii. Dropped all the unnecessary columns and rows which is not required for analyzing the Loan status of Borrowers.
    iii. Correct the datatype of required attributes\

- Use of seaborn and pandas library to analyze various data elements for risk analysis
- Univariate analysis to observe single attribute
- Bivariate analysis of each attribute with loan status
- Multivariate analysis to observe the correlation between each parameter.

Attributes Analyzed

- Below attributes are used for analysis:
  - Loan amount
  - Loan Status
  - Interest rate
  - Term
  - Purpose
  - Number of Delinquency in 2 years
 - Number of inquiries in last 6 months

Results

- Percentage of regular and charged off loan status:
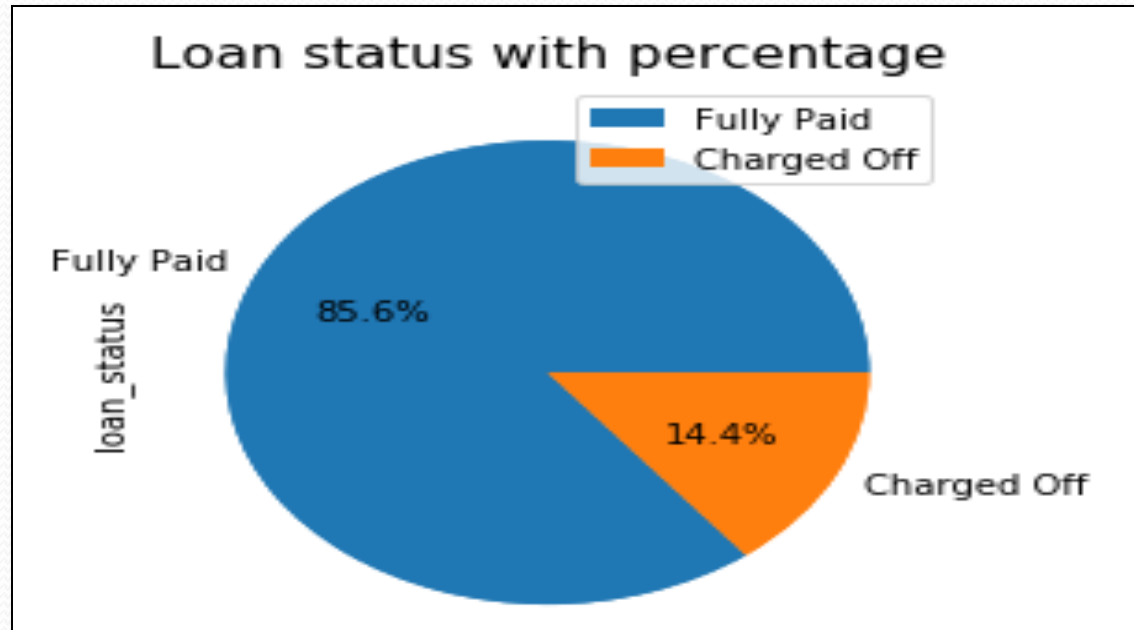
Observation:

- The given pie chart shows that the 85.6% Borrowers are likely to be regular whose loan can be accepted and called as non-risky loan.

- The remaining 14.4% Borrowers are likely to be defaulter and whose loan can be rejected and called as risky loan.

Results

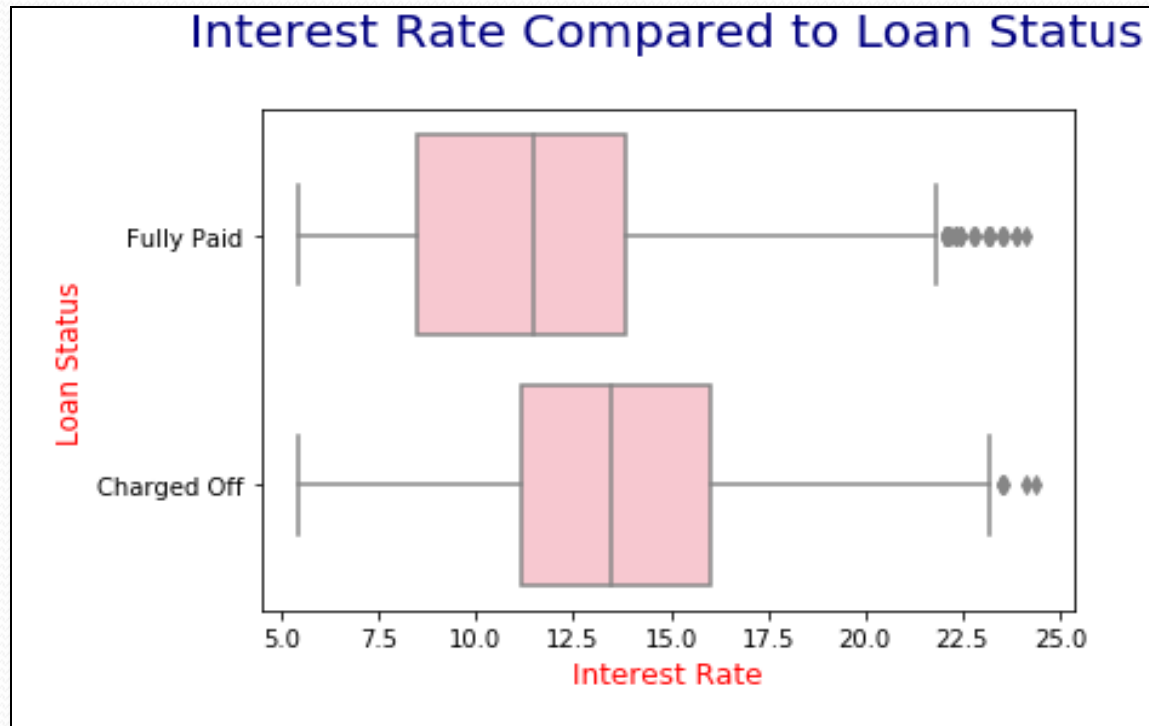Frequency plot For Customer's Address State in conjunction to Loan Status
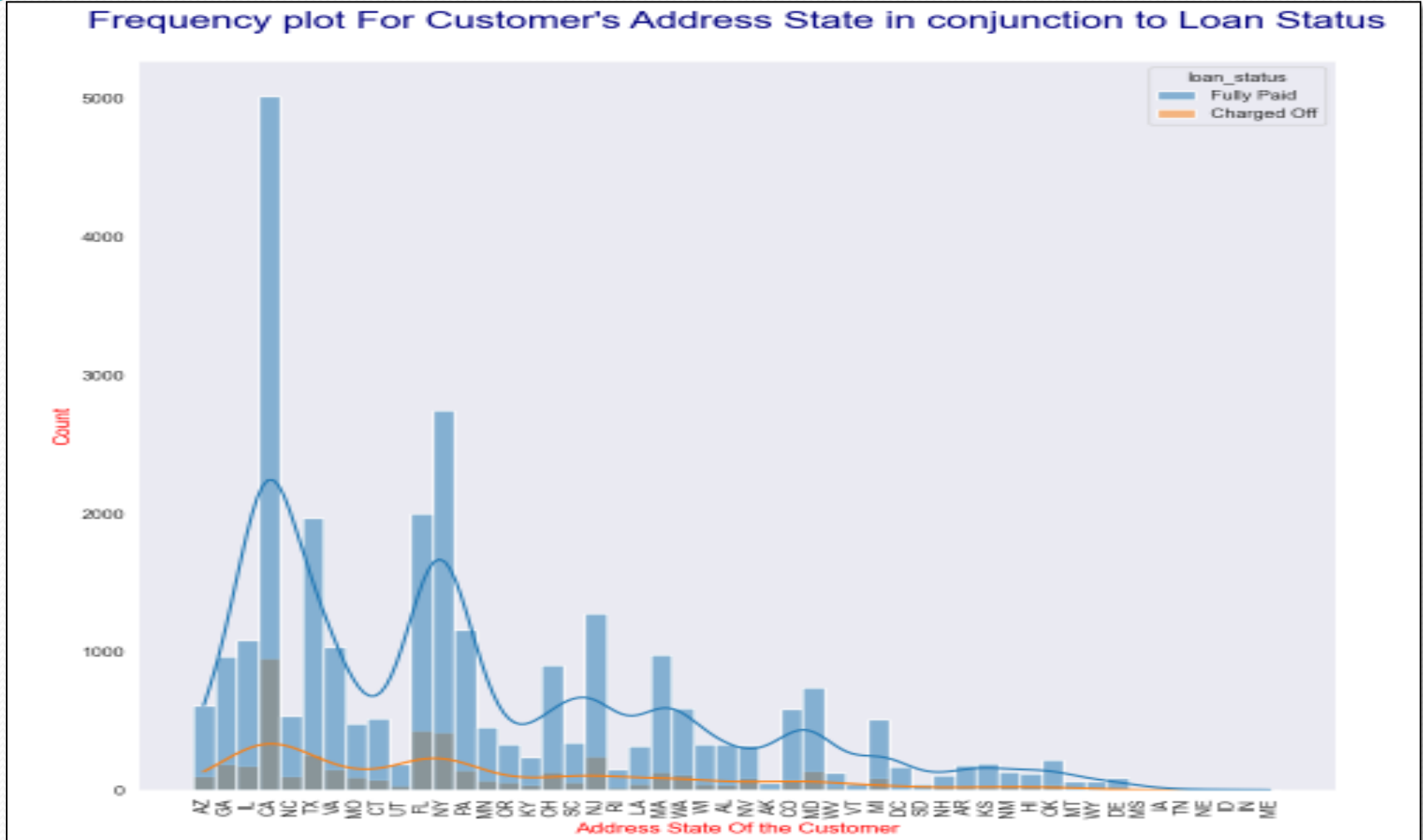
Observation:

- California, Texas, Florida, New York and New Jersey having more frequency of loan applicants and they are likely to be regular as compared to other state borrowers.
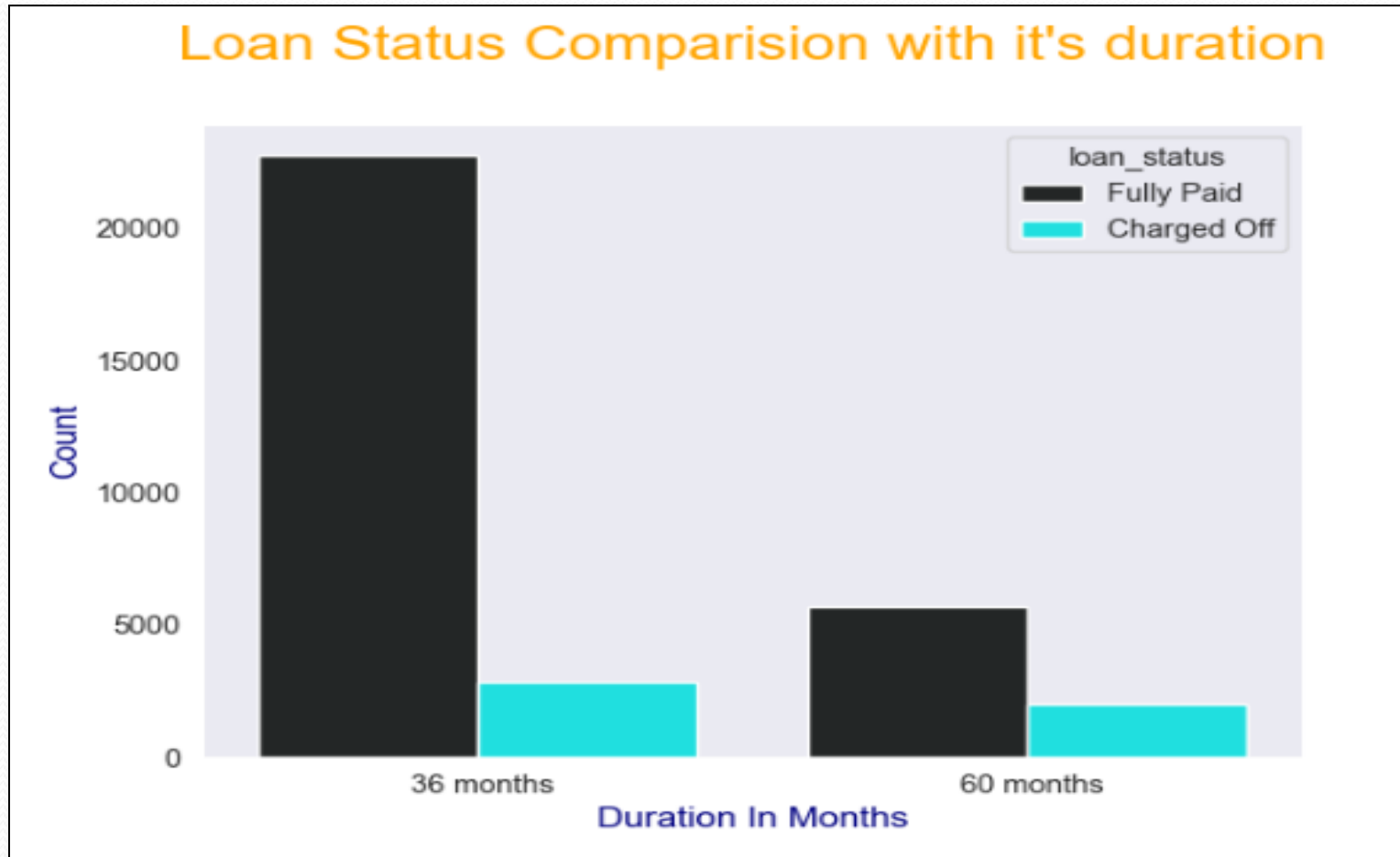
Results

- Interest rate compared with loan status :

Observation:

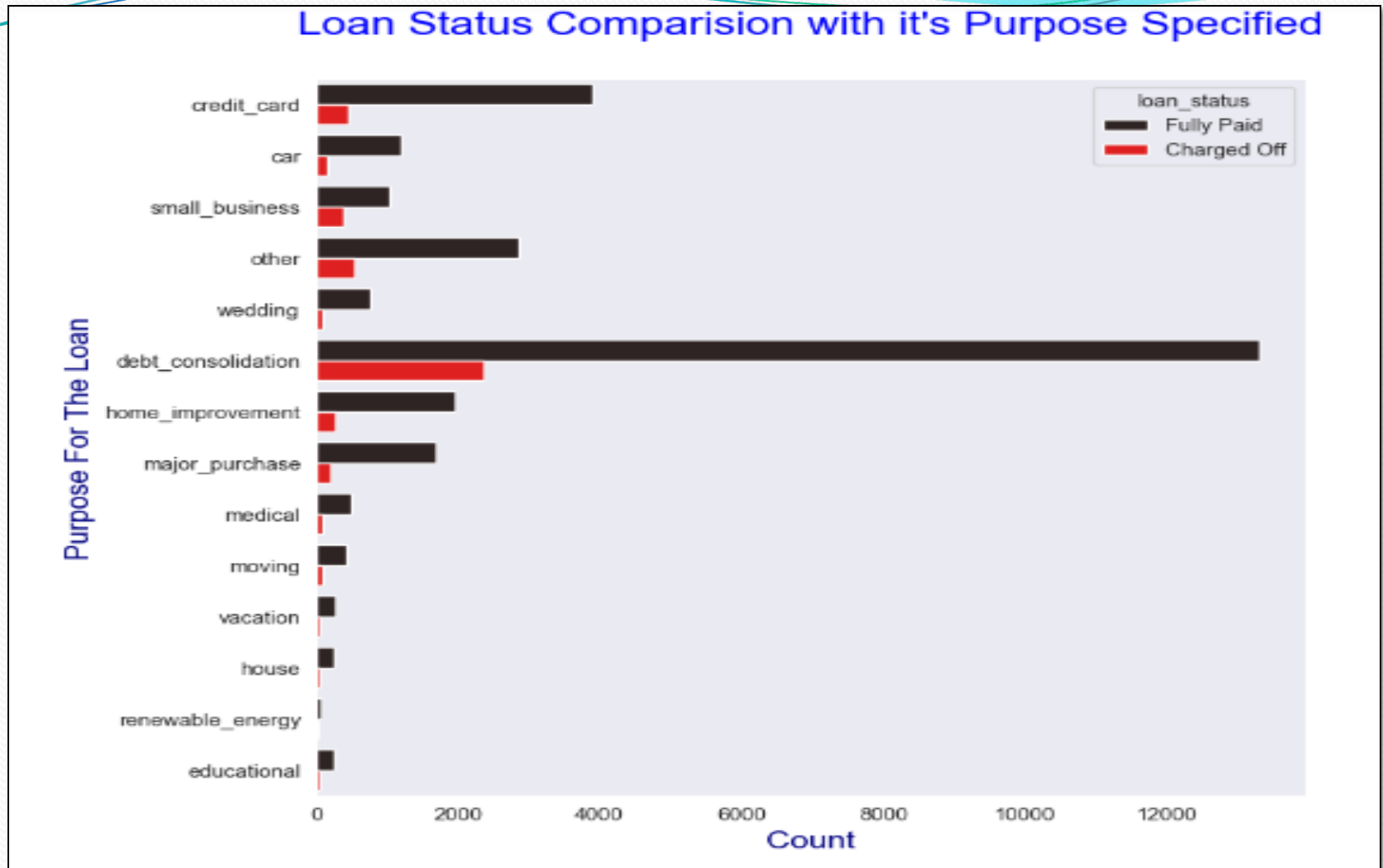- Borrowers whose interest rate is less are likely to be regular one.

Results

Observation:

- From the above Bar plot it is observed that the Borowers whose number of payments on the loan are less, those are likely to regular.

Results

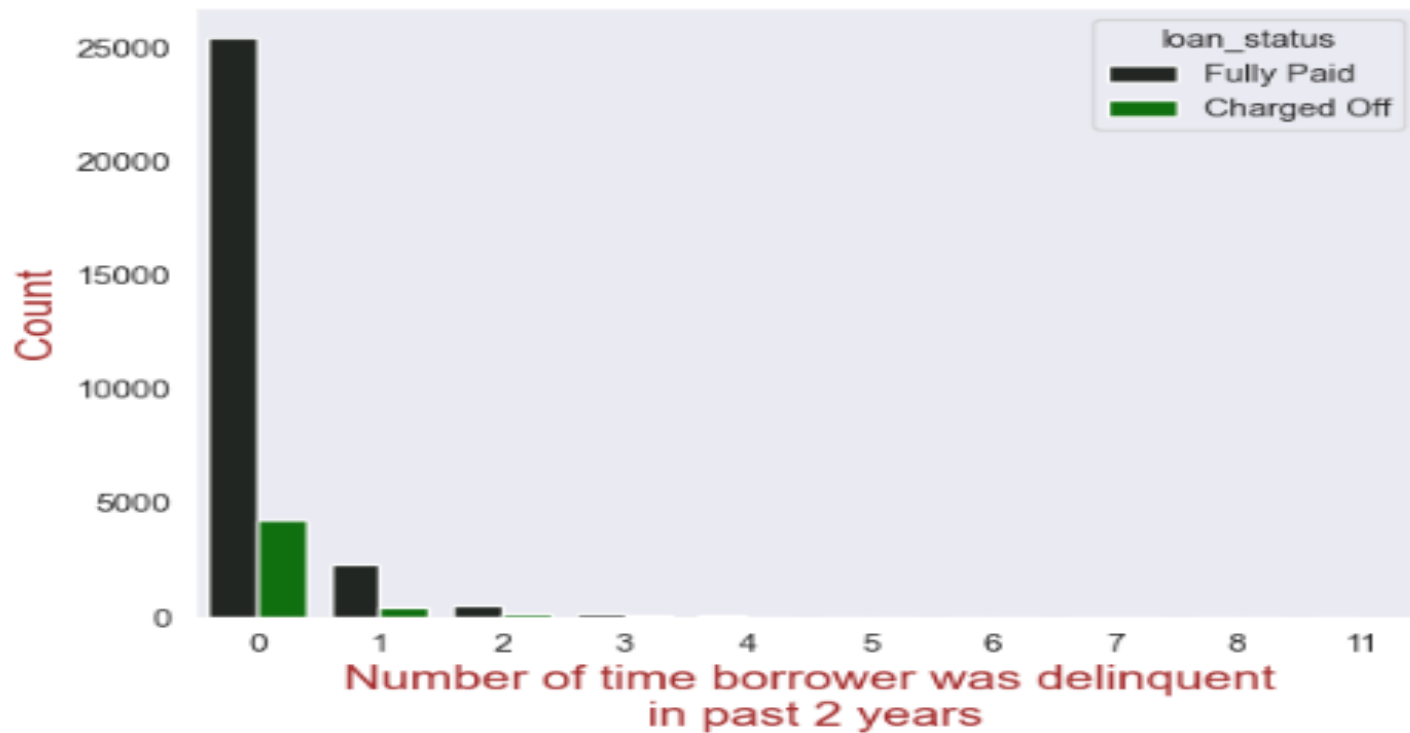Loan Status Comparision with it's Purpose Specified

Observation:

- It is observed that the Borrower who apply for loan for debt consolidation and credit card purpose are regular in greater no. as compared to borrower who apply for loan for other purpose.

Results

Loan Status Comparision with Number of Times Borrower Has Been Delinquent In Past 2 Years

Observation:

- The number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years are more, they are likely to be defaulter. Here it is observed that there are less number of borrowers are likely to be defaulter than regular one.

Results

Correlation Matrix Between Various Numerical Columns
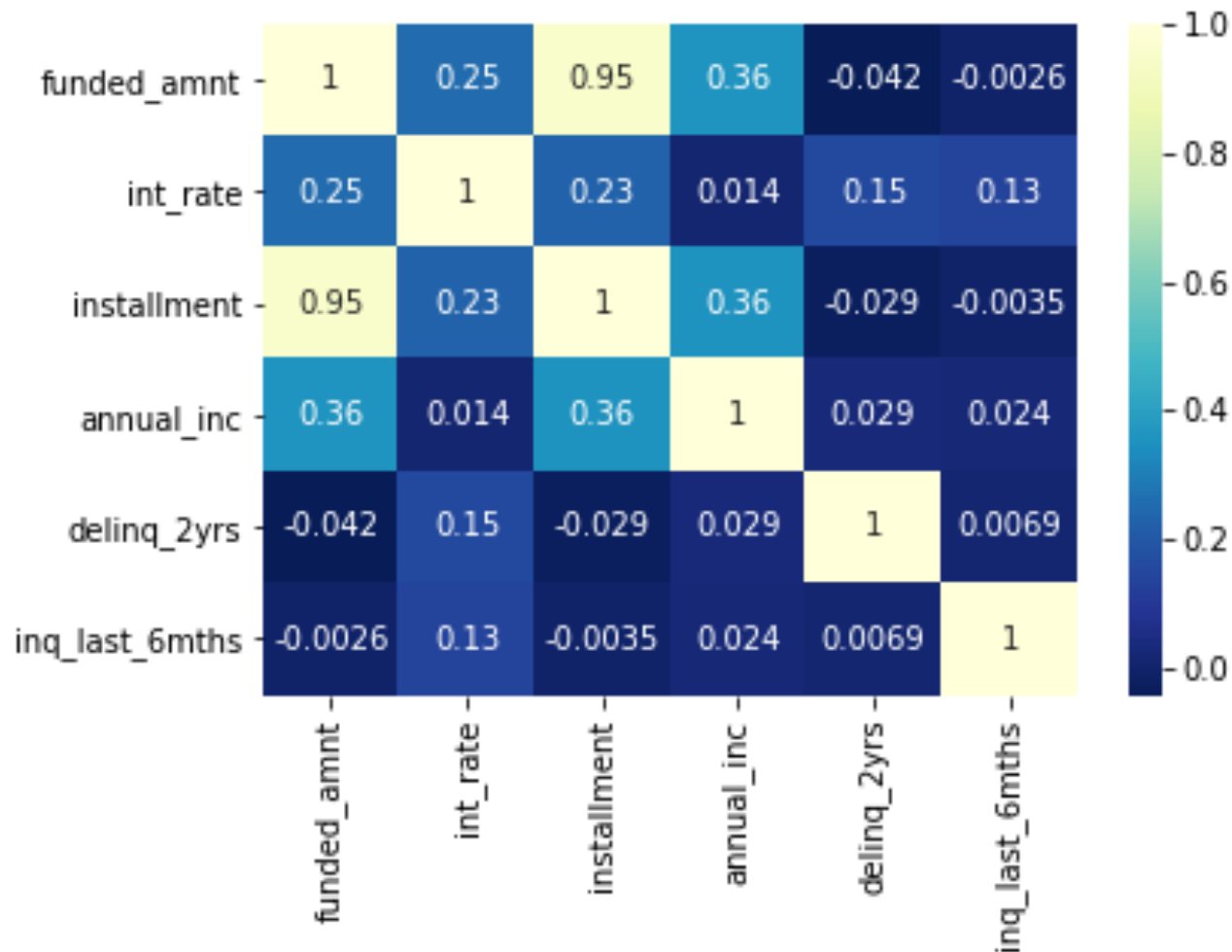
Results

Observations of correlation matrix:

1. Funded amount is in positive correlation with interest rate, installment and annual income and vice versa. If funded amount more then interest rate, installment and annual income will be more and vice versa.

2. Funded amount is in negative correlation with delinquency in 2 years and number of inquiries.

3. Interest rate is in positive correlation with installment, annual income, delinquency in 2 years and inquiry in last 6 months and vice versa. Thus if any one of the parameter increases then other will increase in the ratio given above.

4. Installments of each borrower are in positive correlation with annual income. If annual income of Borrower is more then he/she will apply for more interest.

5. The interest rate is in negative correlation with delinquency in 2 years and number of inquiries in last 6 months.

6. Annual income of borrower is in positive correlation with interest rate, installment, annual income and funded amount as shown in above heatmap

7. The number of delinquency in 2 years and number of inquiries are in positive correlation with each other. If there are more number of inquiries for funded amount in last 6 months, then such loan will be risky loan.

Results

Grade wise distribution of Loan Status
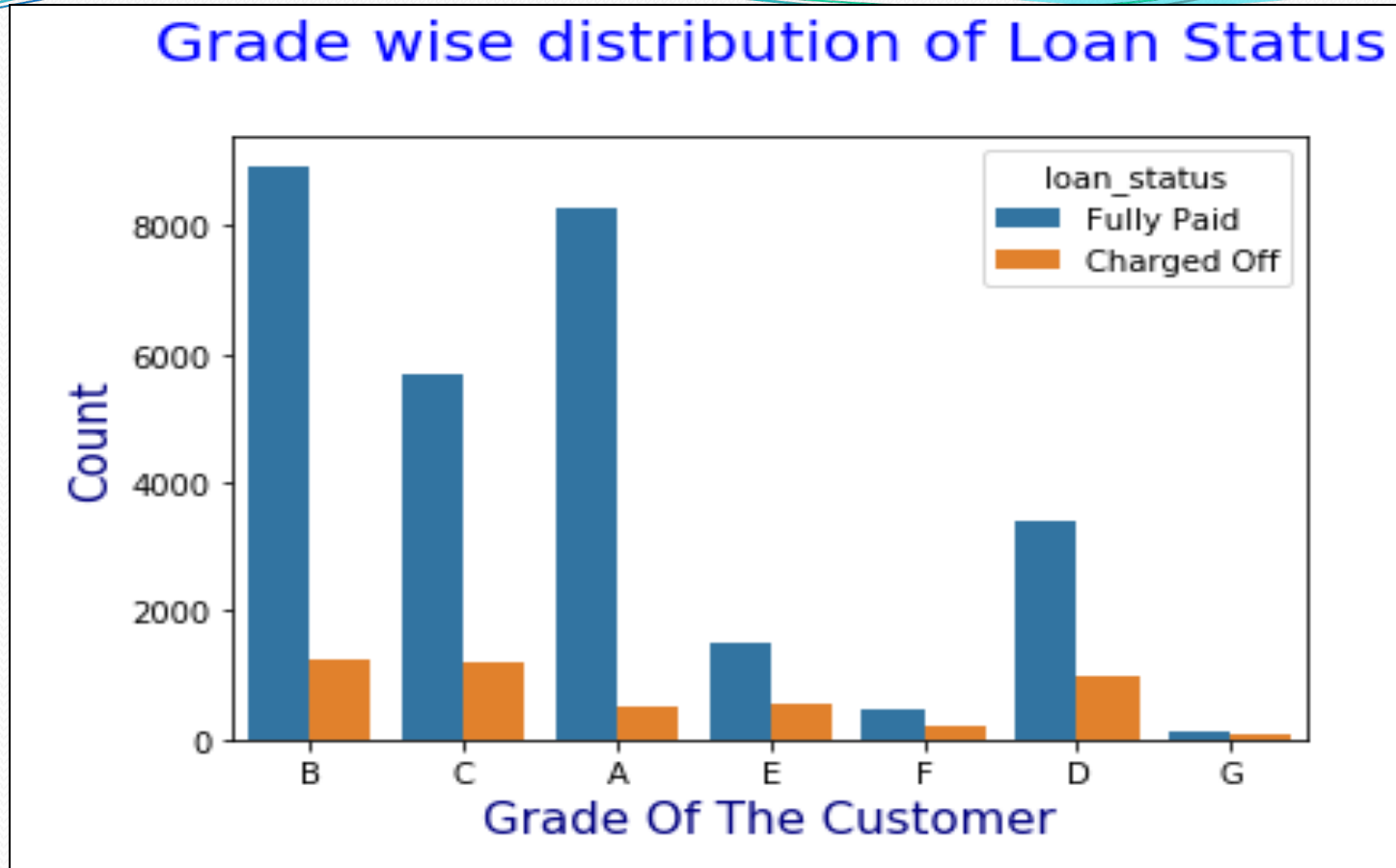
Observation:

- LC assign a Grade to Borrowers based on a borrower's credit history, and the likelihood of repayment of the principal and interest. Thus it is observed that the more number of Borrowers are having good grades, so there are more number of Borrowers are regular one as compared to defaulter one.

Results

Funded Amount of Loan versus Grade Of the Customer in context of Loan Status
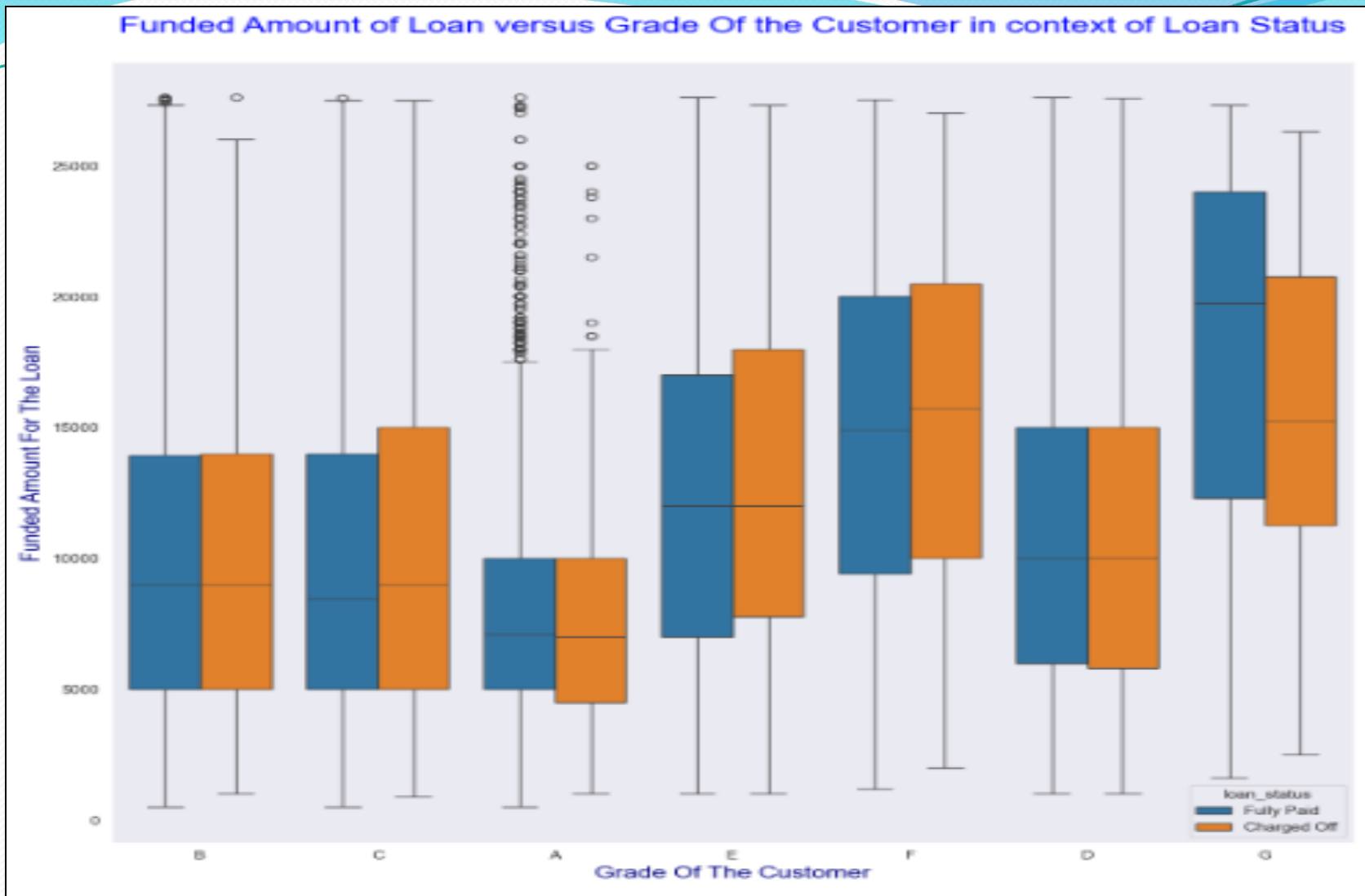
Observation:

- Borrowers having LC Grade as A, B and C applied for less loan amount and they are likely to be regular whereas Borrowers having grade E, F and G applied for large loan amount and as they paid for the same, so they can be considered as regular.

Conclusion:

- Analyzed the parameters which is to be considered for loan status of borrower.

- Thus by using EDA, it is observed that the 85.6% applicants are likely to be regular and 14.4% applicants are likely to be defaulter by understanding the required attributes and by correlating them using correlation matrix.

- Interest rate is driving factor for loan repayment.

- Borrowers having grade E, F and G applied for large loan amount and as they paid for the same, so they can be considered as regular.

- Loans related to credit card repayment or debt consolidation has loans which are large in numbers and higher probability for repayments.