Lending Club Case Study

Exploratory Data Analysis

By

Jyothsna Devi

Sujai Jain

Objective

▶ It helps developing basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers.

▶ It helps a Bank to make a decision for loan approval based on the applicant's profile.

▶ Using EDA we can understand how consumer attributes and loan attributes influence the tendency of default.

Constraints

▶ When a person applies for a loan, there are two types of decisions that could be taken by the company:

▶ Loan accepted: If the company approves the loan, based on 3 possible scenarios described below:

    ▶ Fully paid: Applicant has fully paid the loan (the principal and the interest rate)

    ▶ Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.

    ▶ Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan

▶ Loan rejected: The company had rejected the loan because the candidate does not meet their requirements etc.,

Data Cleaning and Preparation Process

- Import Data from csv

- Data cleaning is the process of removing data which are not contributing for analysis
  - Removing null , NA values
  - Remove loan status with value 'Current'
  - Remove unique values
  - Remove behavioural columns

Univariate Analysis

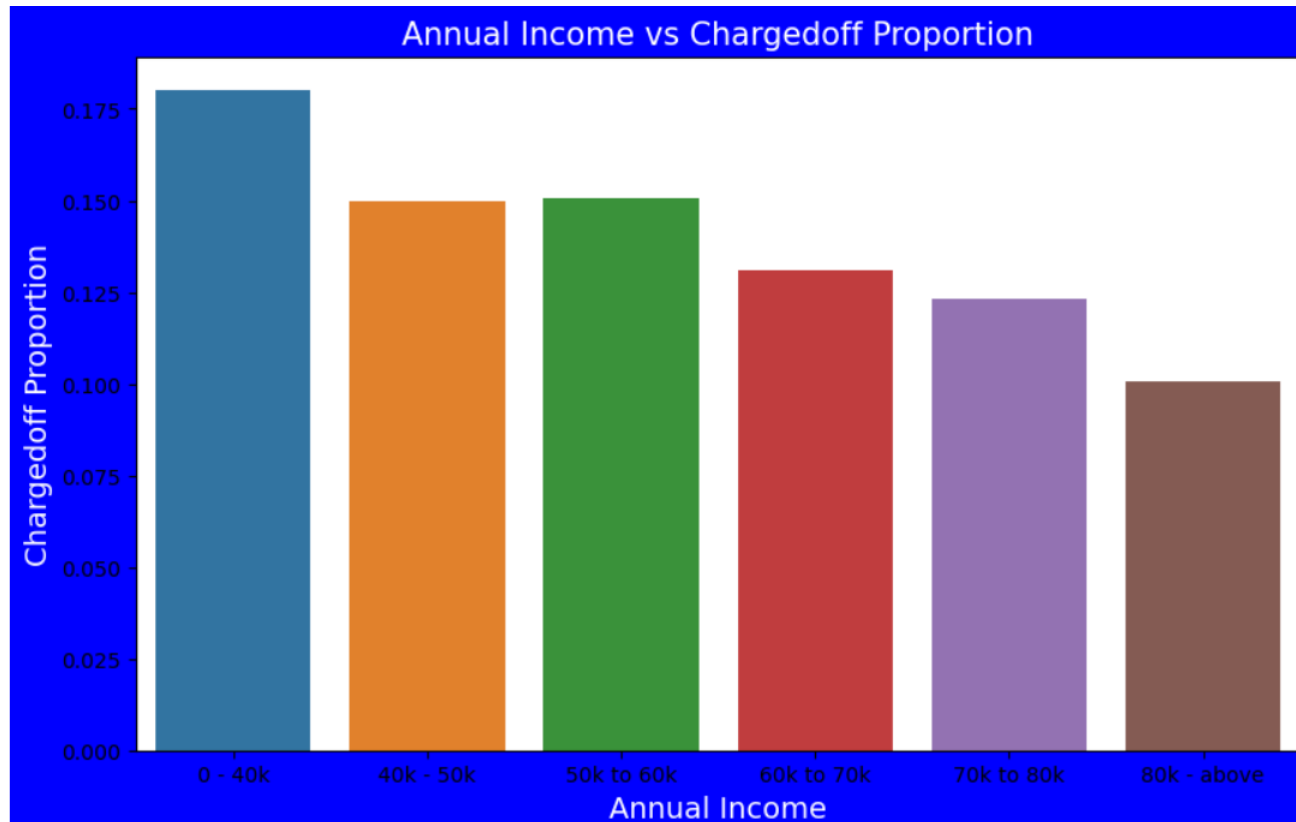▶ Most of the loan amount applied was in range of 5000-14000
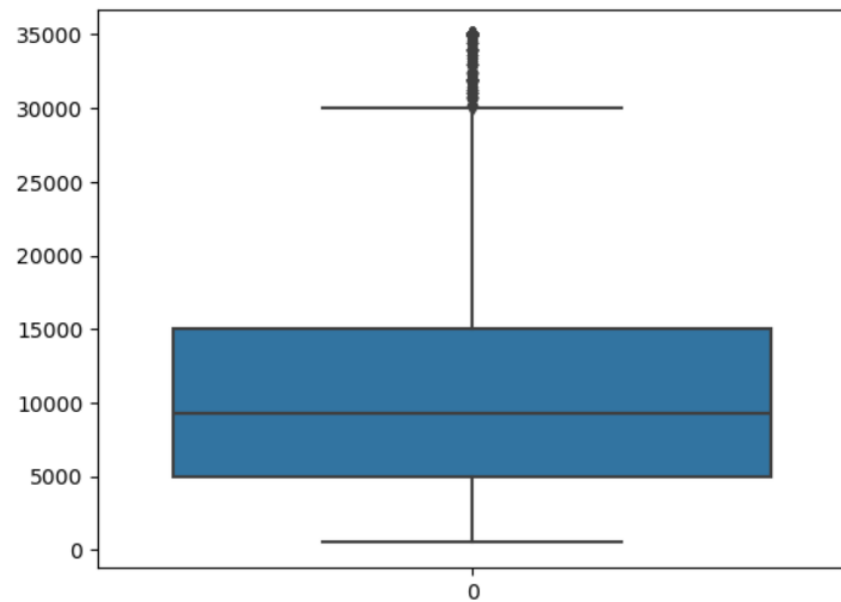
▶ Annual income of most of the applicants is in the range of 40000 – 750000
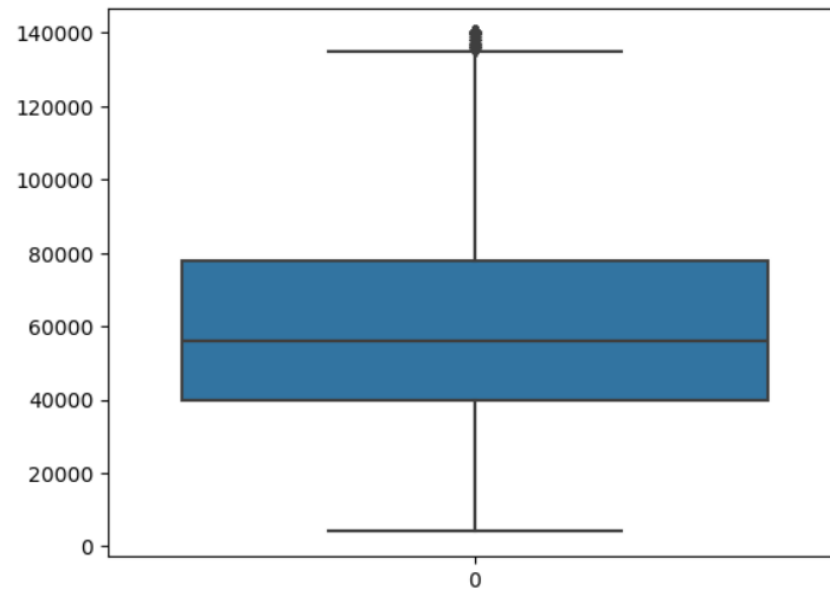
Bivariate Analysis

▶ Increase in annual income decrease Charged Off proportion

Interest Rate Vs Charged Off

▶ Charged off proportion is increasing with higher interest rates

▶ Interest rate less than 10% has very less chances of charged off.

▶ Interest rate more than 16% has high chances of charged proportion.
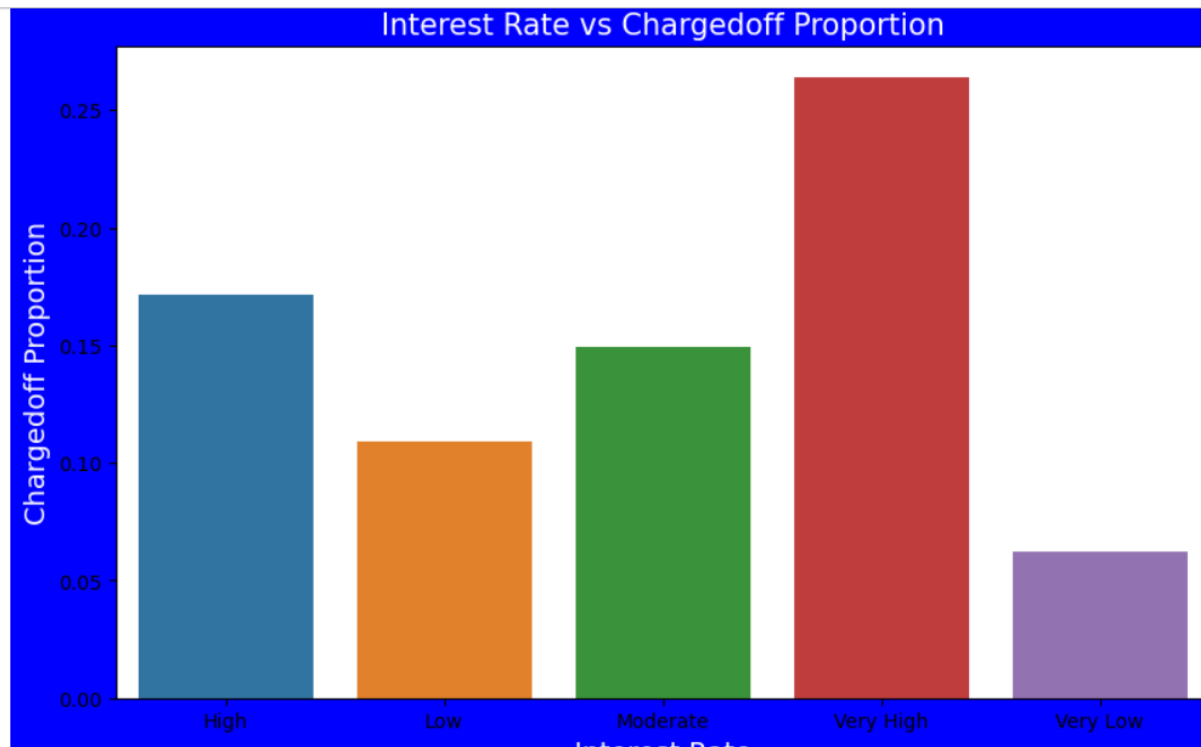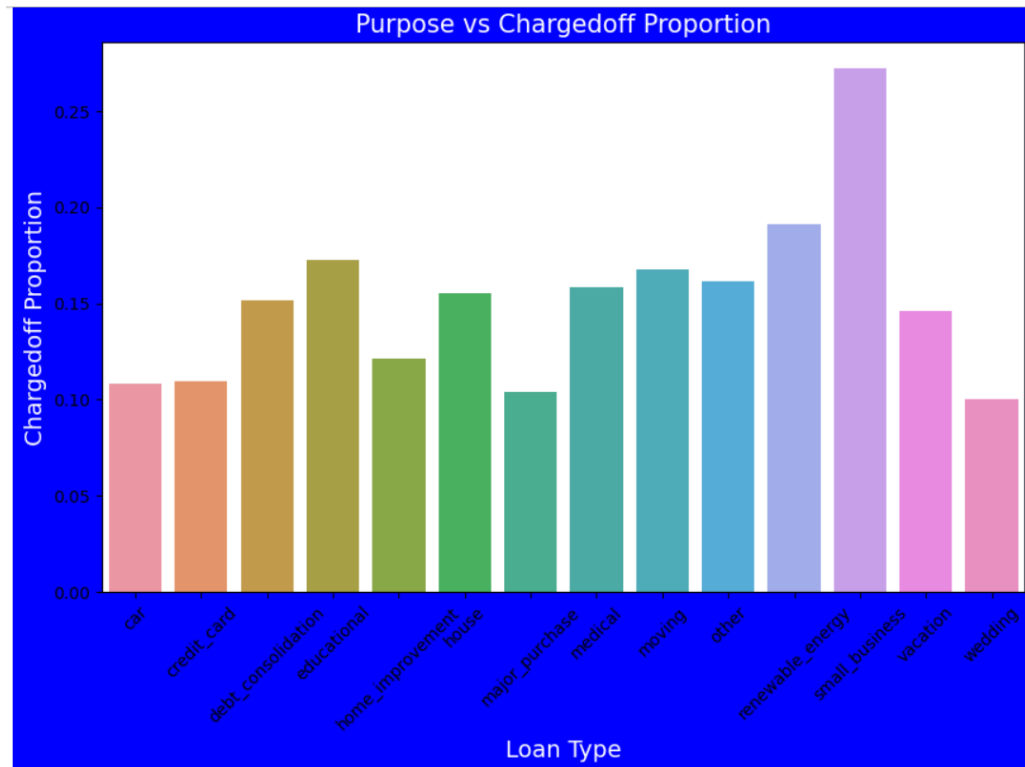
Home Ownership Vs Charged Off

► Those who are not owning home have high chances of loan defaulter.
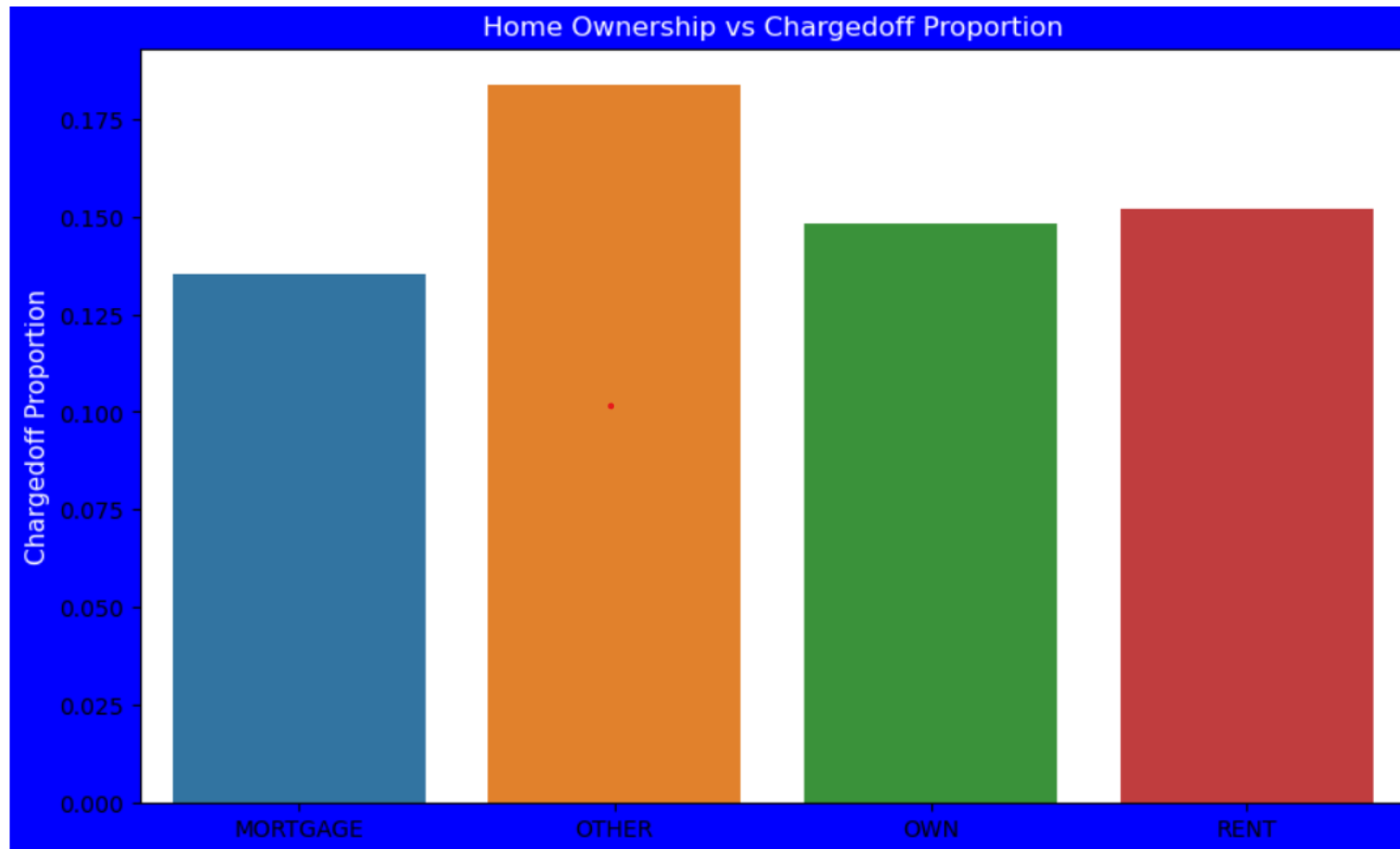
Loan Type Vs Charged Off

▶ Applicants who is having home loan is having low chances of charged off.

▶ Those who took loan for small business is having high chances of charged off.

Purpose vs Chargedoff Proportion

Bankruptcies VS Charged Off

▶ Bankruptcies with 2 is having high impact on loan defaults

▶ Bankruptcies Record with 0 is low impact on loan defaults
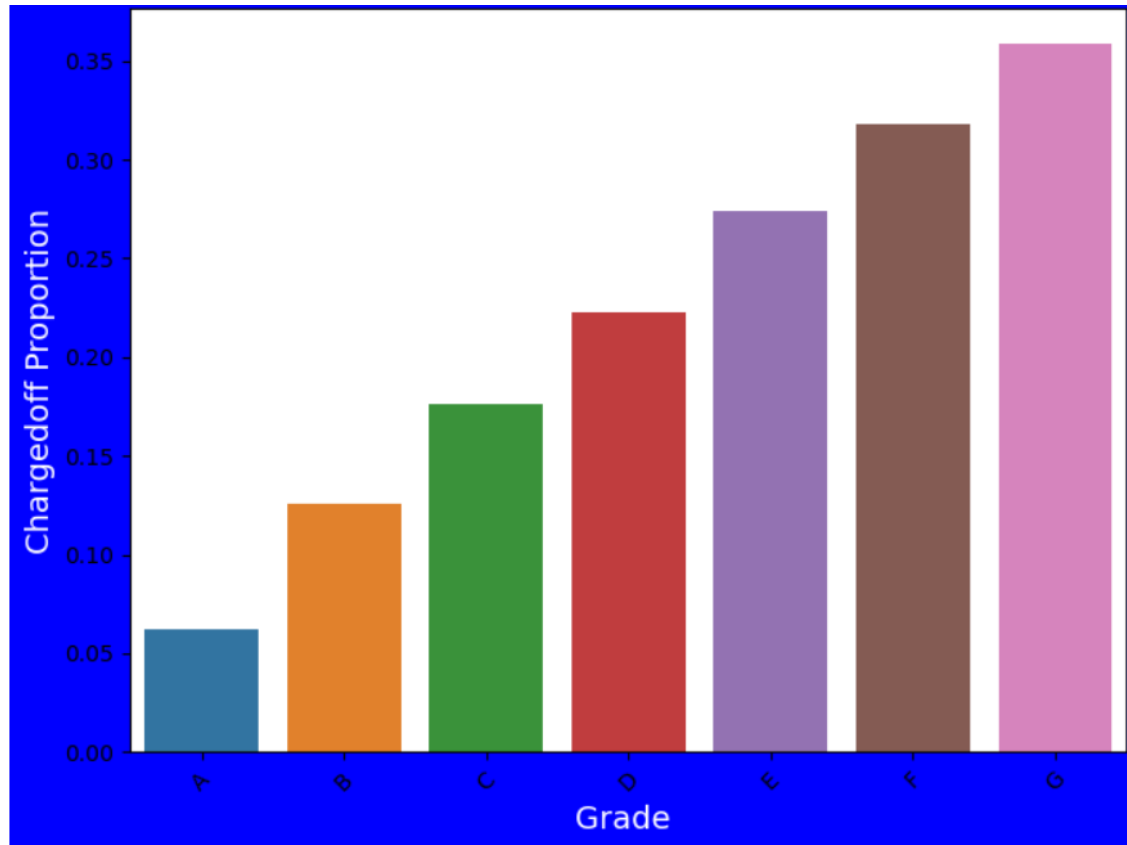
▶ Lower the Bankruptcies lower the risk.
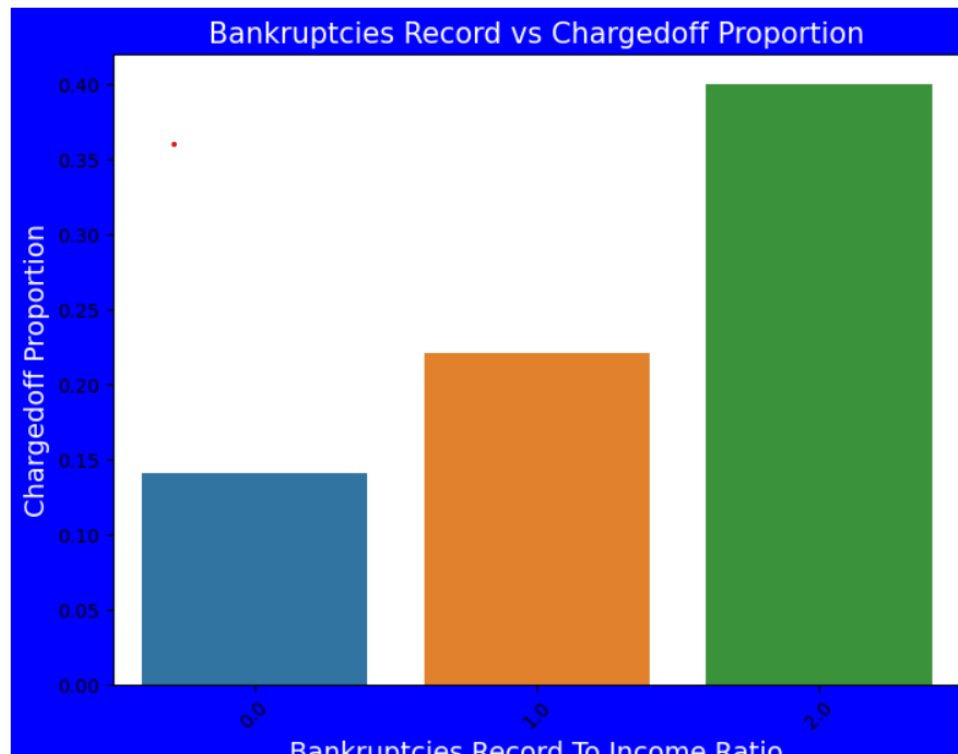
Issue Year Vs Charged Off

▶ Loans issued in 2007 has highest loan defaults
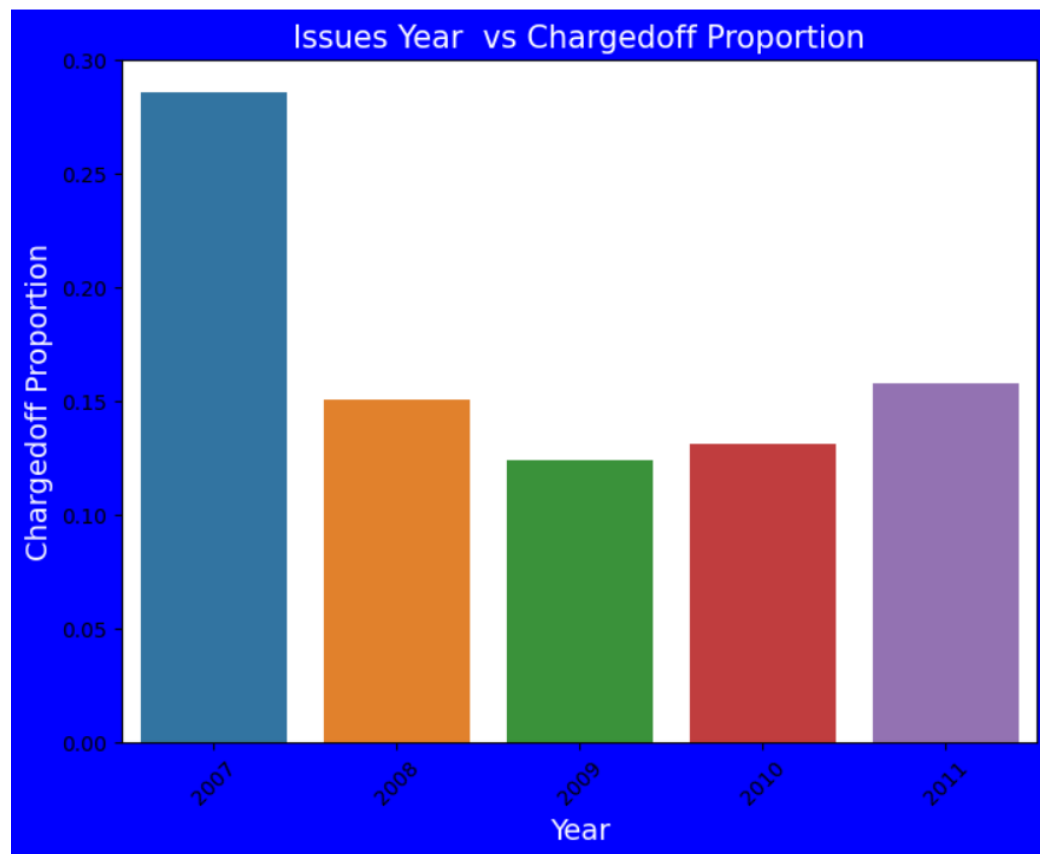
Grade Vs Charged Off

▶ Applicants with grade G has highest loan defaults

Conclusion

▶ Applicants whose income range is between 0-20000 has high chances of charged off.

▶ Interest rate more than 16% has high chances of charged off.

▶ Applicants who are on rent or mortgage has high chances of defaulters.

▶ Applicants who has bankruptcy record has high chances of defaulters.

▶ Loan applicants who got loan in 2007 has high chances of charged off.

▶ Loan applicants with Grade G has high chances of charged off.