Lending Club Case Study

Dr. Atul Tiwari

Objectives

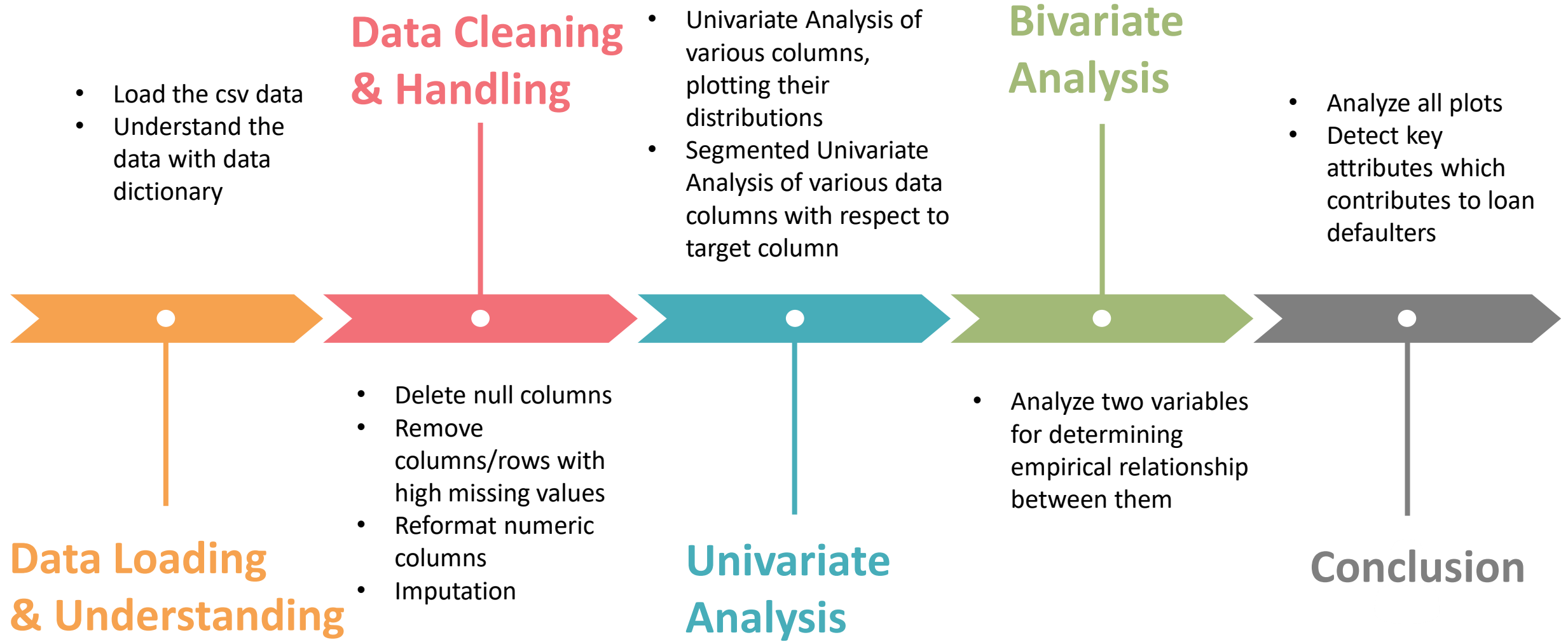The aim is to understand the pattern between consumer attributes and loan attributes

Using EDA, understand how consumer attributes & loan attributes influence the tendency to default
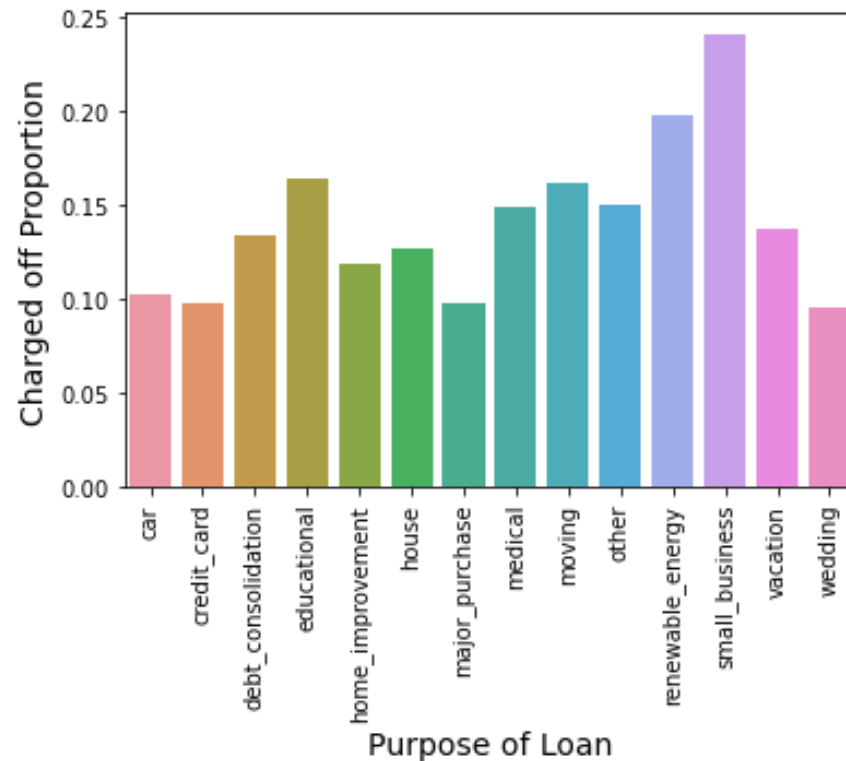
Identify the key driving factors (variables) associated with loan defaults (i.e., strong indicators)

METHODOLOGY

Data Cleaning & Handling

- Univariate Analysis of various columns, plotting their distributions
- Segmented Univariate Analysis of various data columns with respect to target column

Bivariate Analysis

- Load the csv data
- Understand the data with data dictionary

- Analyze all plots
- Detect key attributes which contributes to loan defaulters

Data Loading & Understanding

- Delete null columns
- Remove columns/rows with high missing values
- Reformat numeric columns
- Imputation

Univariate Analysis

- Analyze two variables for determining empirical relationship between them
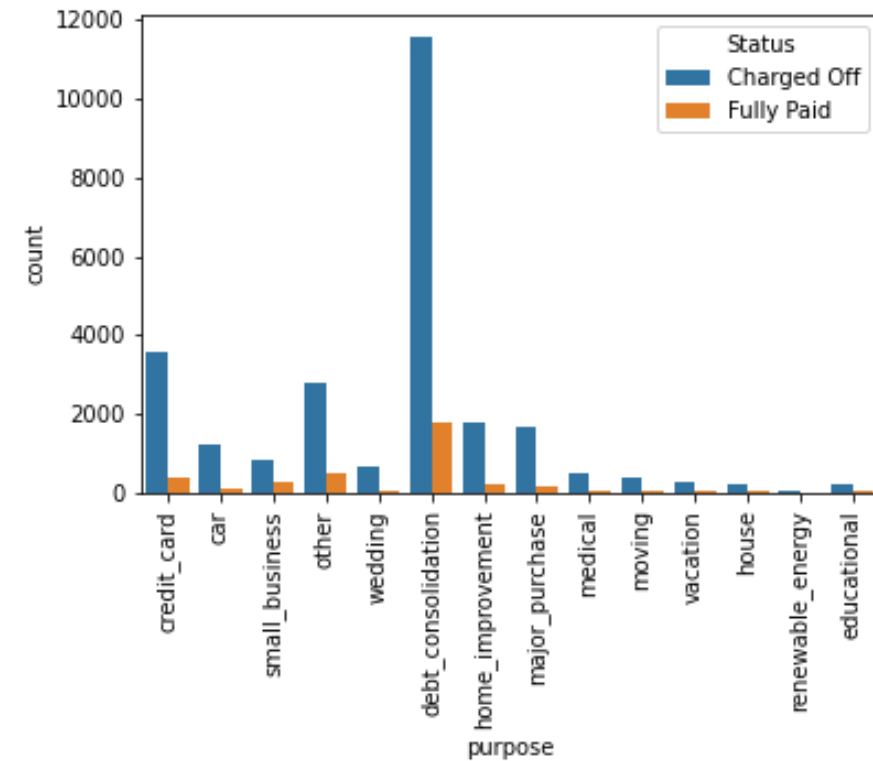
Conclusion

Purpose of Loan

Most of the loans are taken for debt consolidation and credit card bill payment

Most common purpose for loan defaults are – Small business, renewable energy and educational applicants
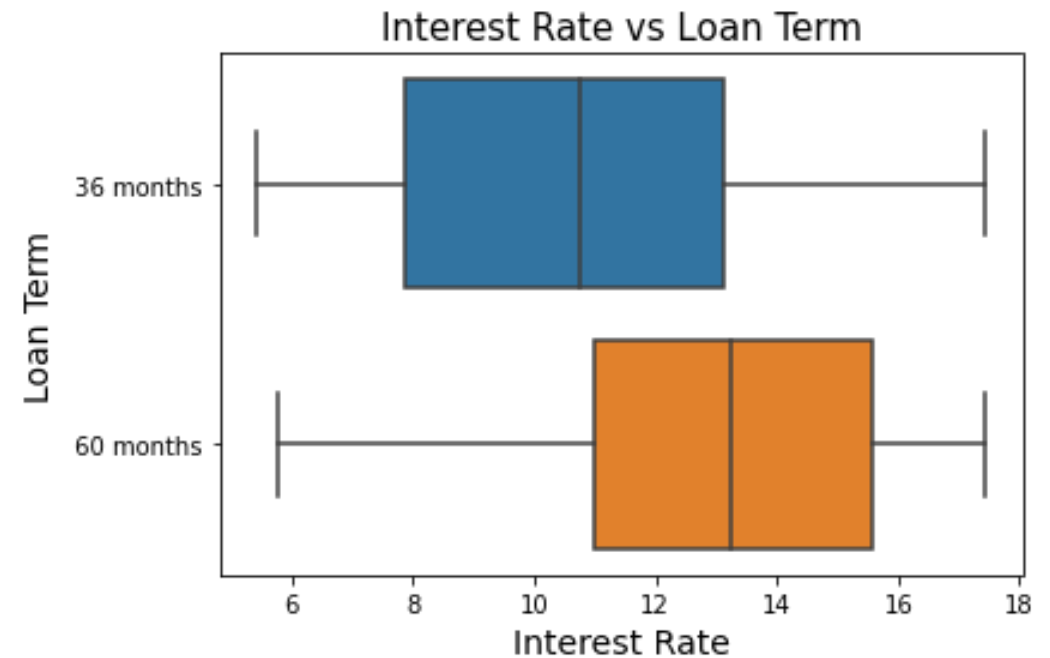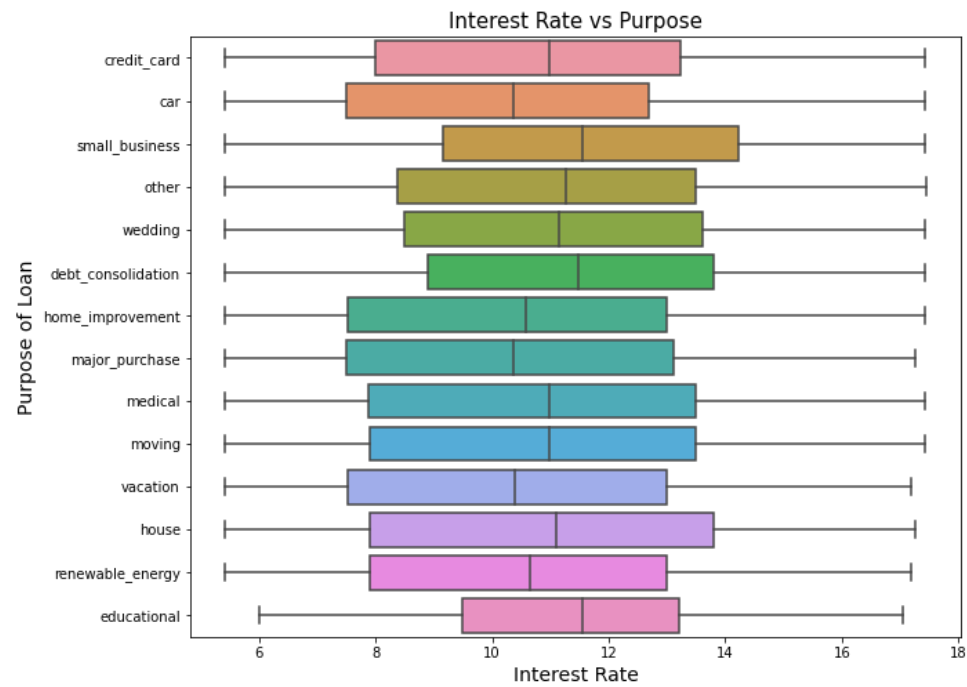
Loan amount

- Defaulters increase with the increase in loan amount and is maximum from the loan amount of 20000+
- Interest rate increases with the increase in loan amount
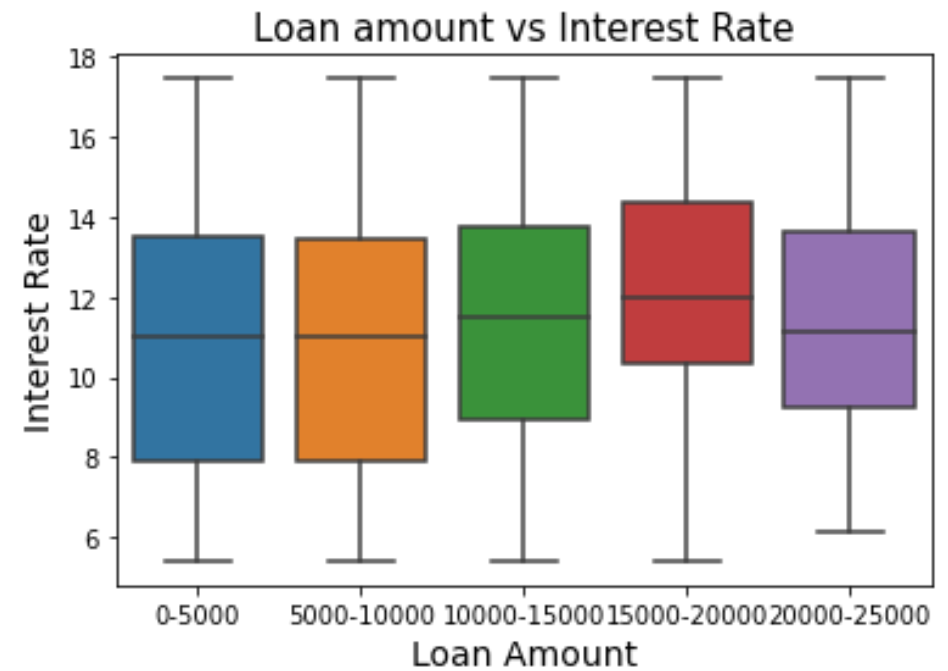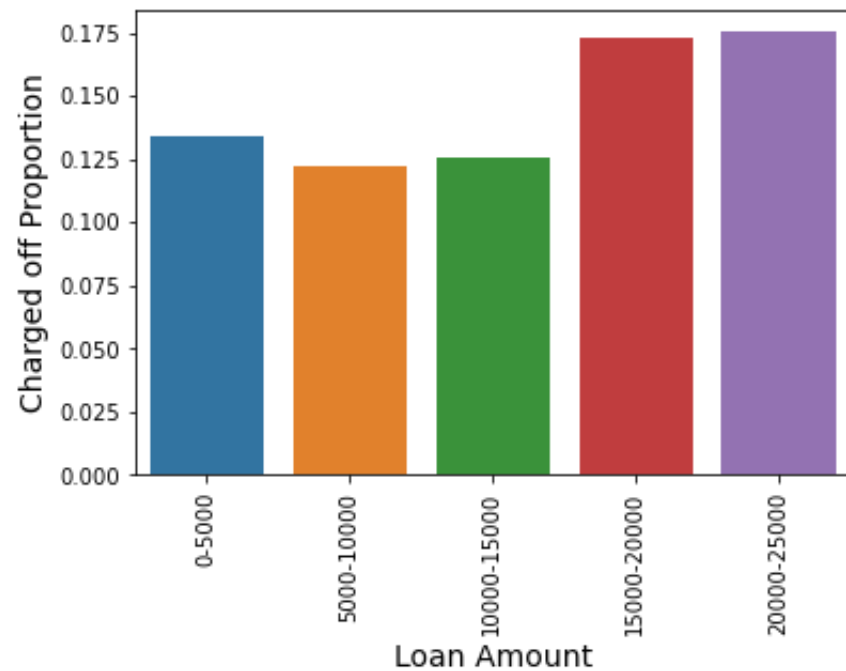
Interest Rate vs Loan Term & Purpose

- Interest rate is higher in case of 60 months loan term
- Top 3 purpose for high interest rate are - Small Business, Debt Consolidation, Educational
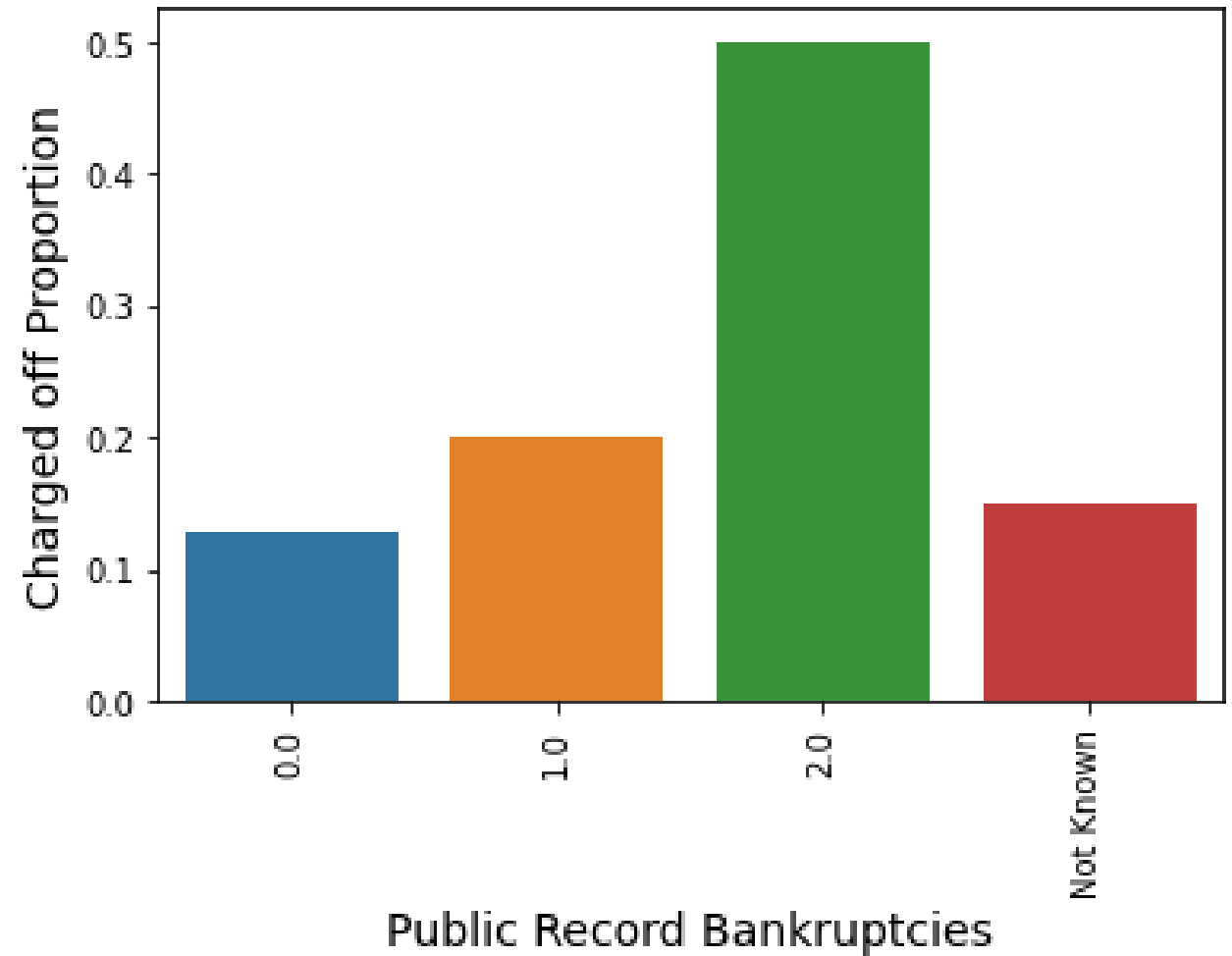
Grade vs Interest Rate

- Interest rate increases as grade change from "A" to "F"
- Chances of charged off (defaulters) is increasing with grade moving from "A" towards "F" ("A" being the lowest and "F" being the higest defaulter)
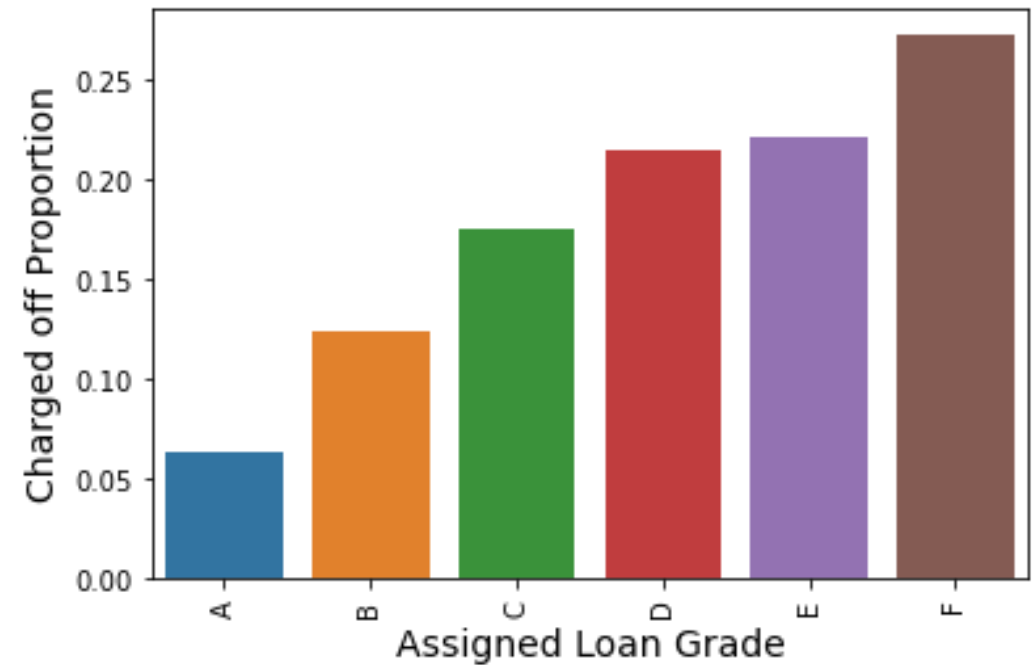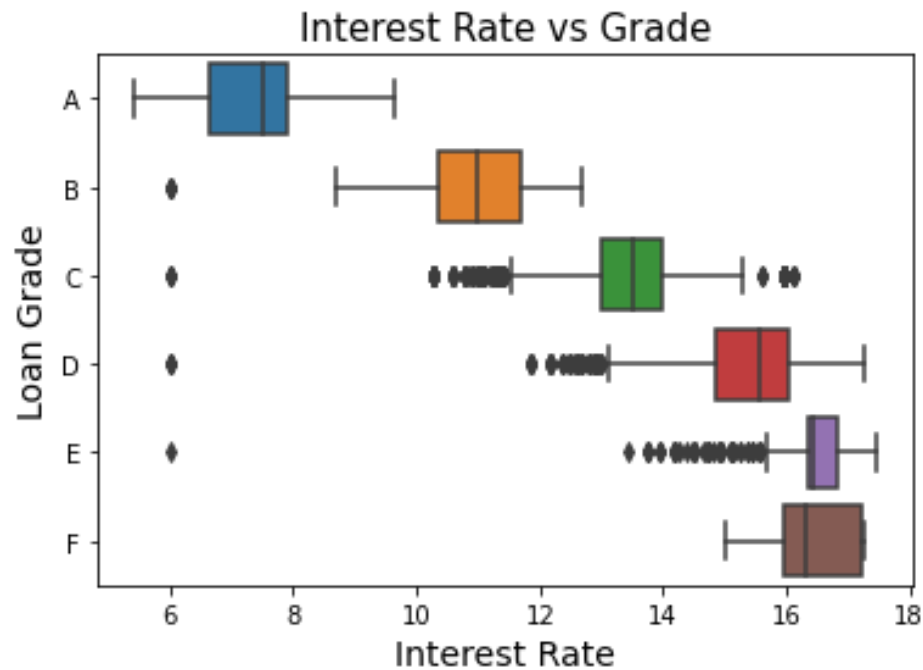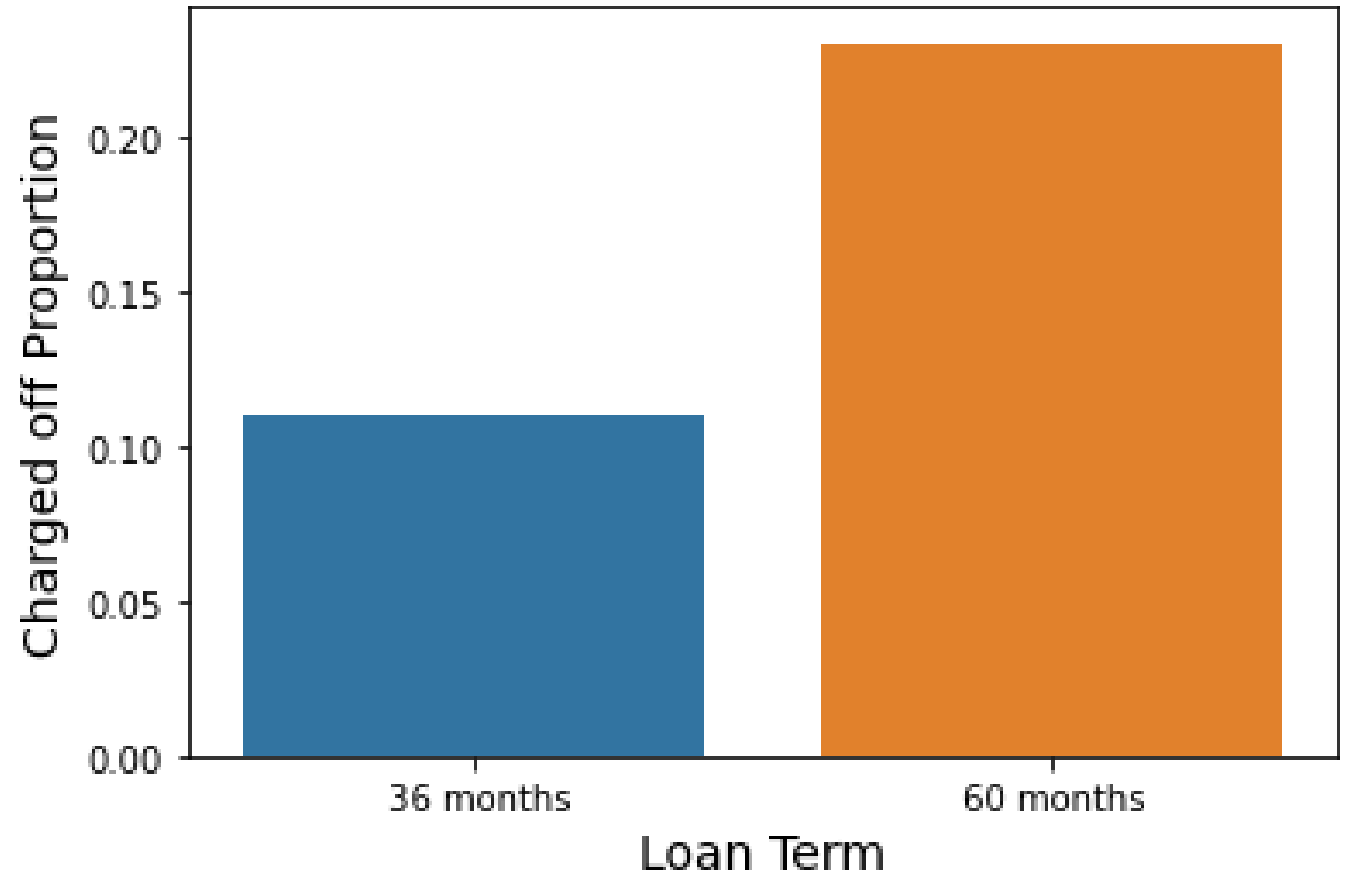
Public Record Bankruptcies

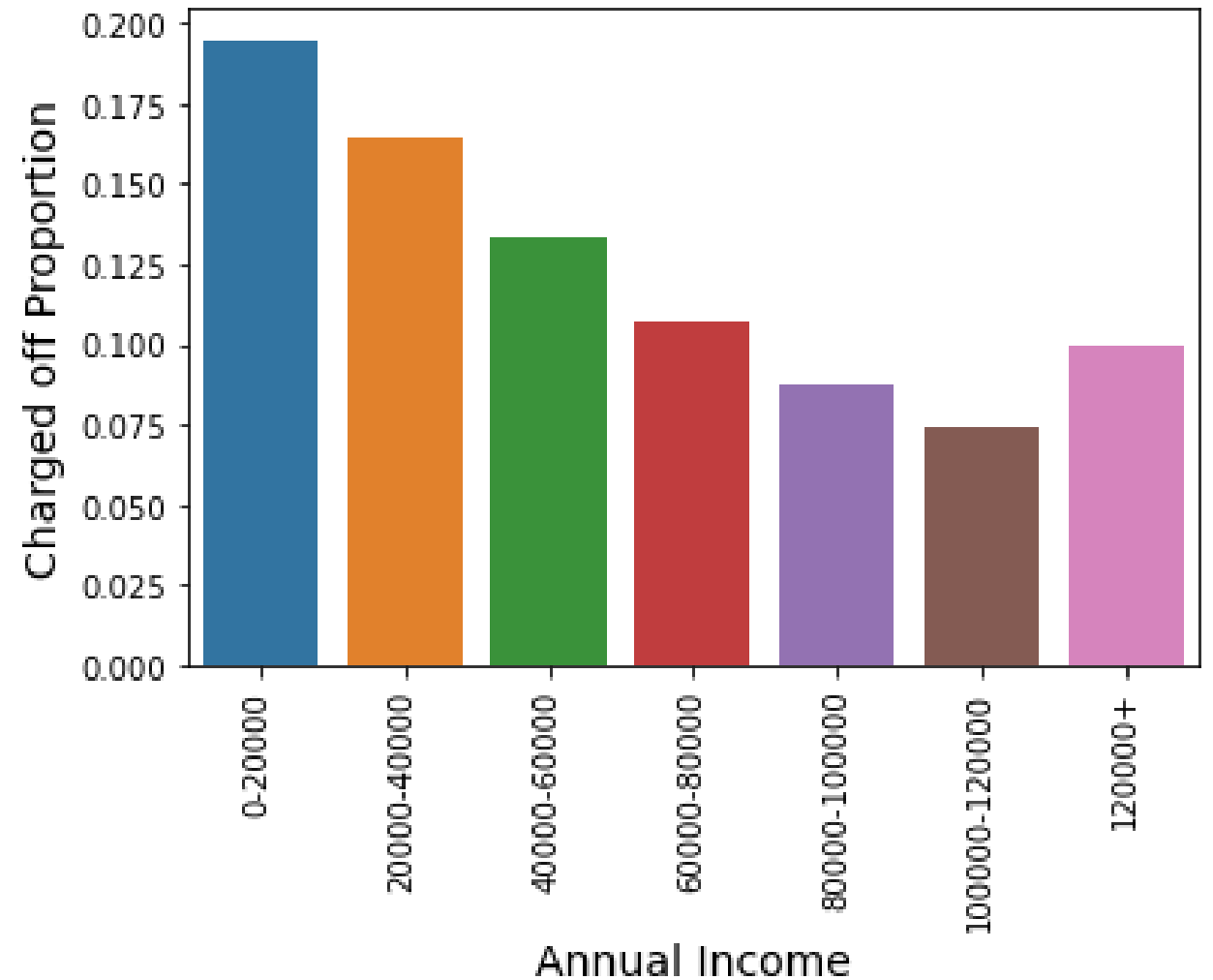- Higher number of public record bankruptcies results in high defaulters

Loan Term

- Chances of charged off is higher in case of 60 months loan term

Annual Income

- Chances of charged off (defaulter) is higher in lower income range (0-20000)

Conclusion (1/2)

Interest Rate

- Charged off (defaulters) proportion increases with increase in interest rates. Interest rate more than 16% has the highest chance of charged off.

- Interest rates increases with increase in loan amount.

- From the above two points it is justified why loan amount of 20000+ has high chances of charged off.

Purpose of Loan

- Top 3 defaulters categories with respect to purpose of loan are - a. Small Business applicants b. Renewable energy applicants c. Educational applicants

Annual Income

- Low annual income range people are more likely to default. Income range (0-20000) has high chances of charged off.

Grade

- Proportion of charged off applicants are more with grade moving from "A" towards "F".

- One of the reason for increasing proportion of charged off applicants while moving from grade "A" towards grade "F" can be attributed to increased interest rate as grade moves from A to F.

Conclusion (2/2)

Employment Length

- Employment length of less than one year has highest proportion of charged off loan applicants.

Home Ownership

- With respect to home ownership highest applicant numbers are for those living in rented home. Next highest loan applicants are those who mortgaged their home. But Charge off chances is less in case of mortgaged homes. Charge off chances are more in case of other categories.

DTI

- Chances of charged off applicants increased with the increase in DTI.

Loan Term

- Charged off applicants numbers are higher in case of 60 months loan term.
- Interest rate is higher in case of 60 months loan term, leading to increased Charged off applicants.

Public Record Bankruptcies

- Higher number of Public Record Bankruptcies results in higher rate of Charged off applicants.