Lending Club Case Study

CREATED BY:

UTKARSH MAHULIKAR

NAGENDRAN

Introduction

This case study aims to give an idea about how real business problems are solved using EDA.

Benefits of the case study:

➢Gives an idea about how EDA is used to solve real life business problems.

➢It also helps to develop a basic understanding of risk analytics in banking and financial services.

➢How the inferences are used to minimize credit loss.

Business Understanding and Objective

➢Business Understanding:

◦ This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

◦ Like most other lending companies, lending loans to 'risky' applicants is the largest source of financial loss (called credit loss). Credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.
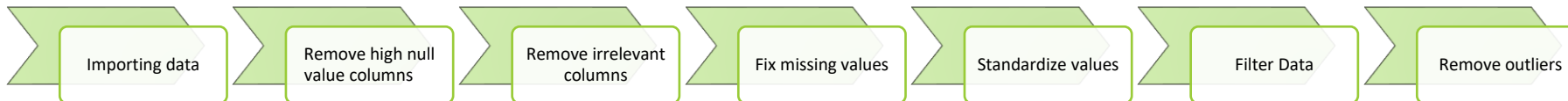
➢Business Objective:

◦ The business objective is to understand the **driving factors** (**or driver variables**) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

Data Cleaning Process

➢ Dataset Information:

The data given contains information about past loan applicants and whether they 'defaulted' or not. Data has details regarding approved loan not the rejected ones. It has 3 status of loan which is Fully Paid, Current and Charged-Off.

| Importing data | Remove high null value columns | Remove irrelevant columns | Fix missing values | Standardize values | Filter Data | Remove outliers |

Loan trend over time

➤ Loans approved have been increasing over the years.

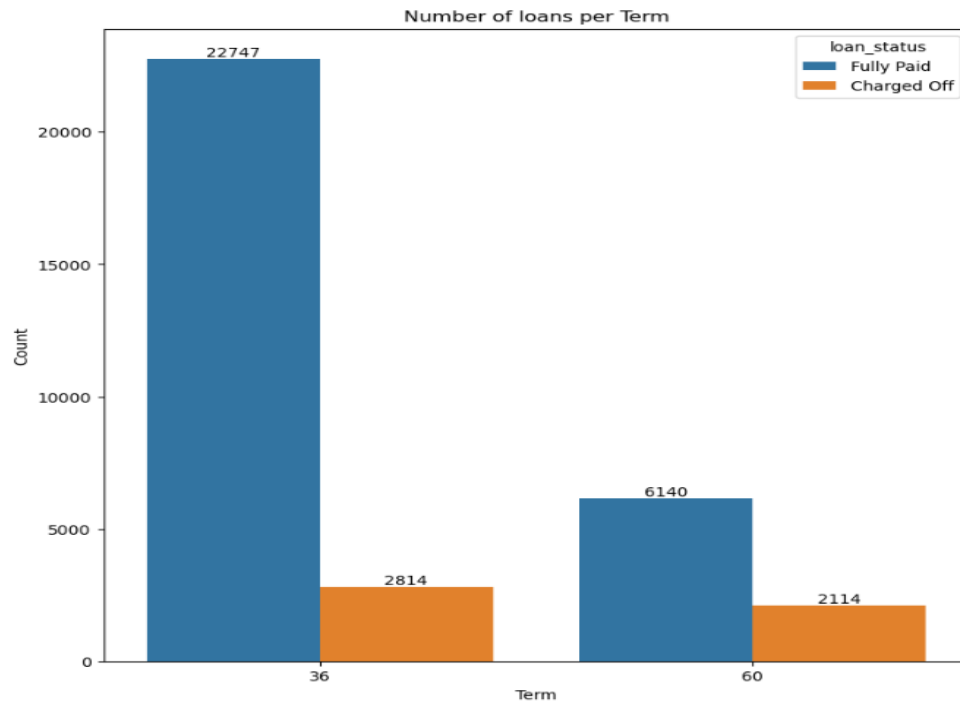➤ More loans are getting approved in the last quarter. This may be due to fiscal targets.

Distribution of Loans per issue year

Distribution of Loans per issue month
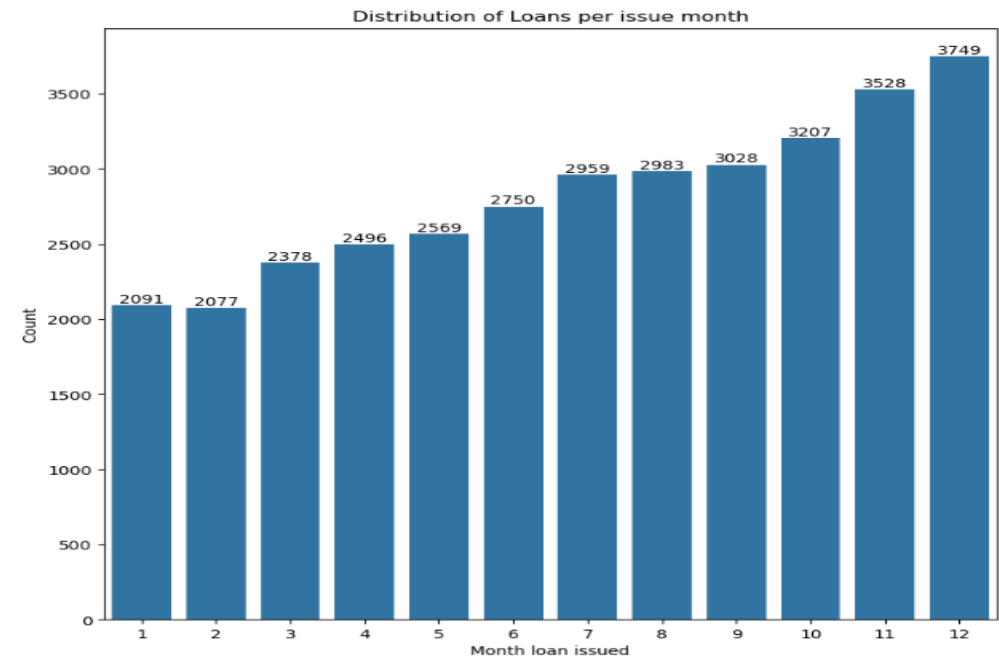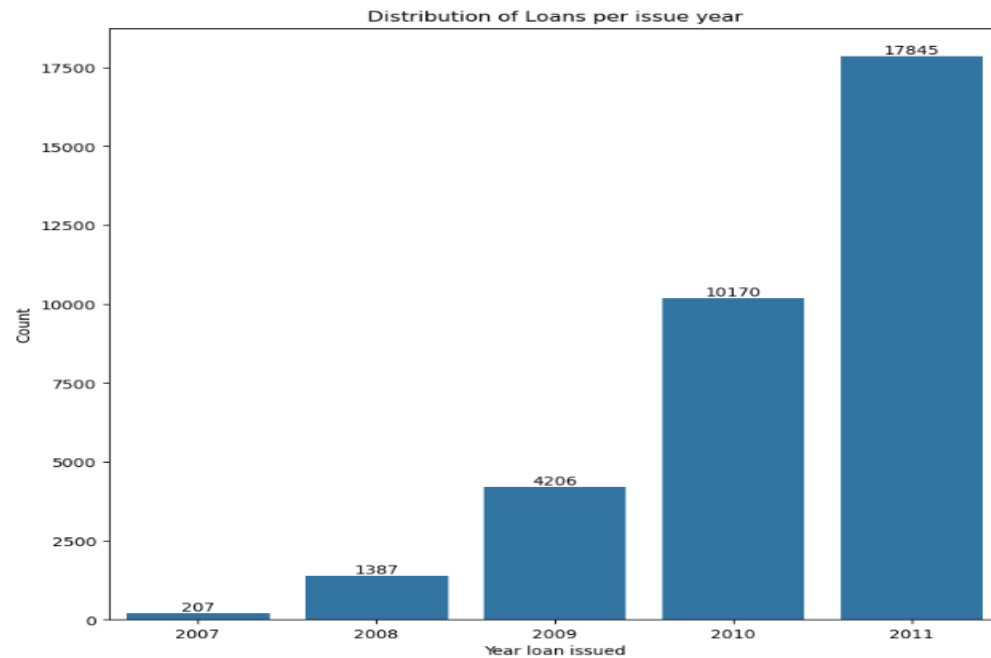
Loan Status

➢ The number of 'Fully Paid' loans is around 6 times more than the number of 'Charged Off' loans.
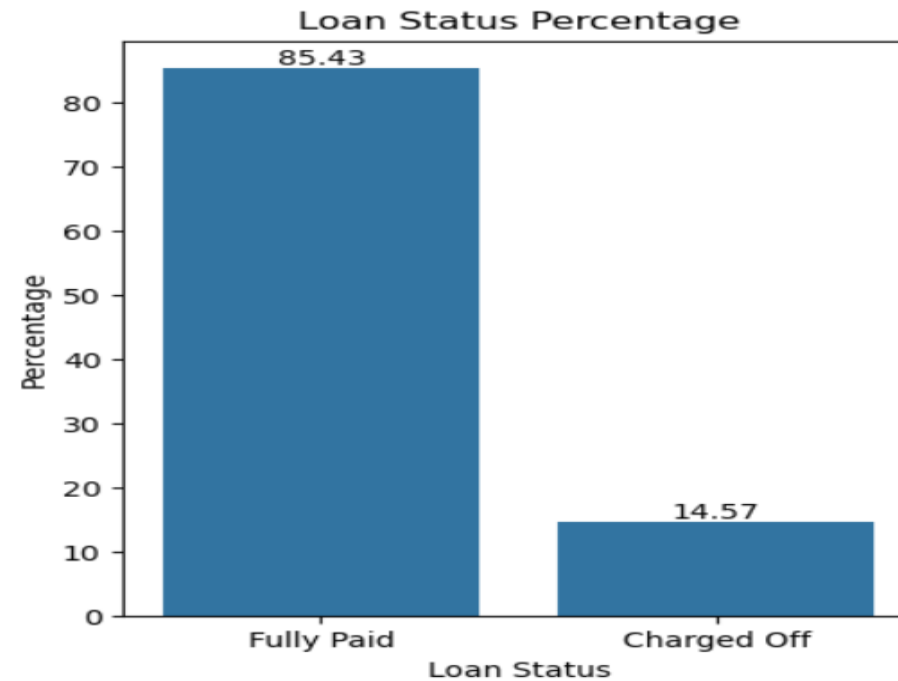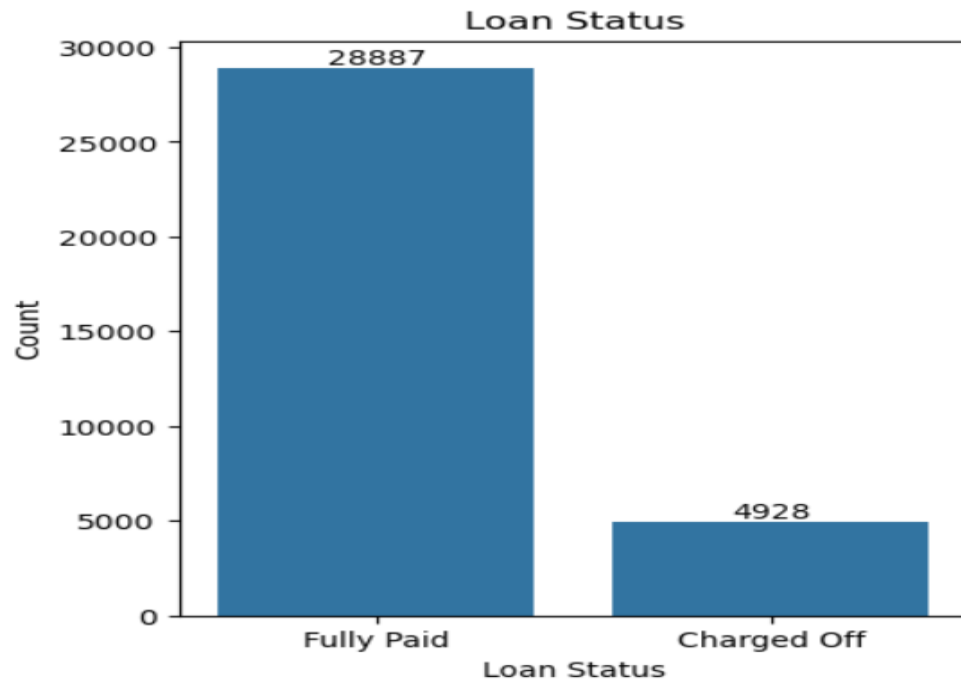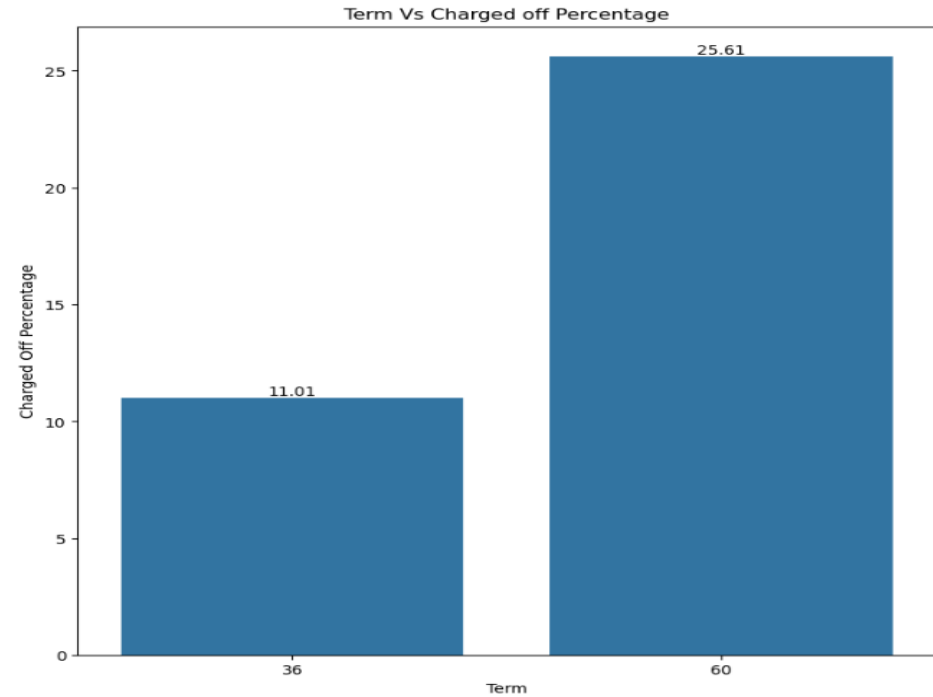
Loan Term

➤ Number of loans with 60 months term are significantly lesser than 36 months term loans
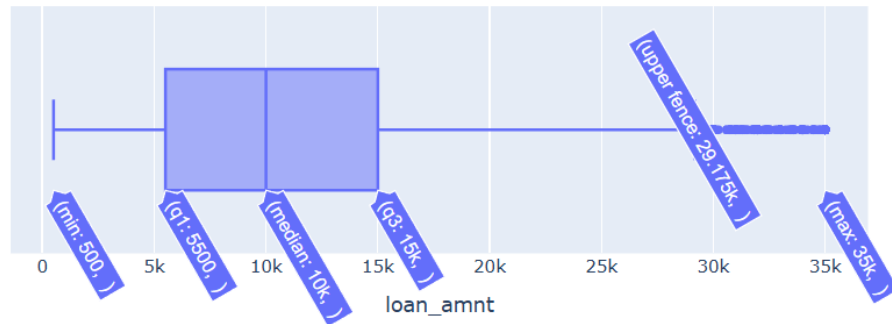
➤ Loans with 60 months term has more than twice the percentage of defaults as compared to 36 months term loans
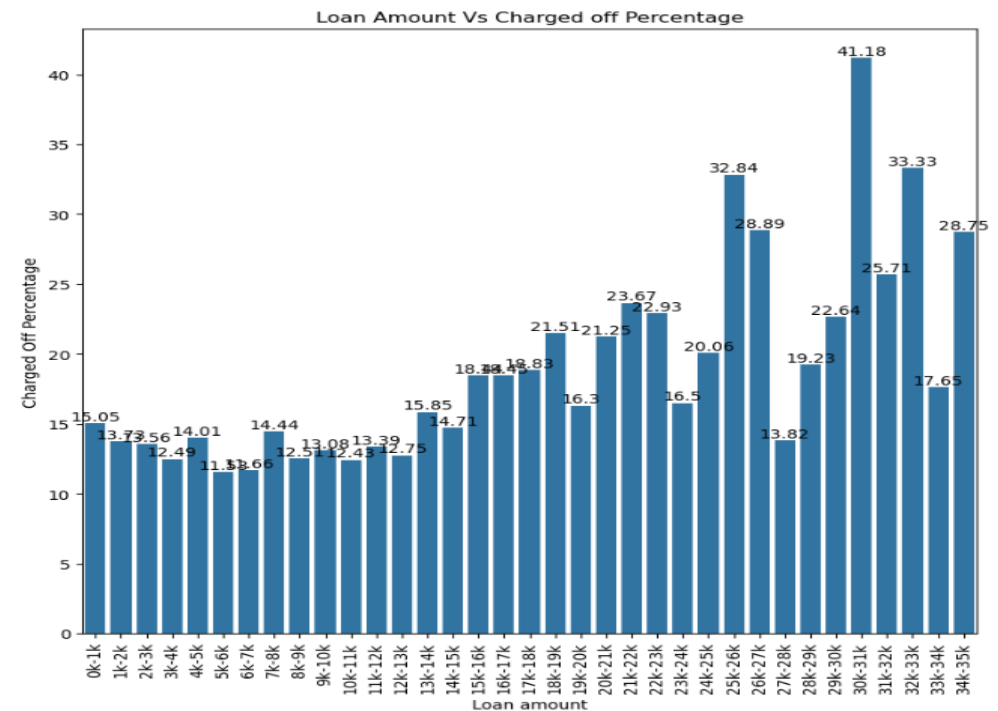
Loan Amount

➤ Loan amount varies between 500 and 35K with a median of 10K
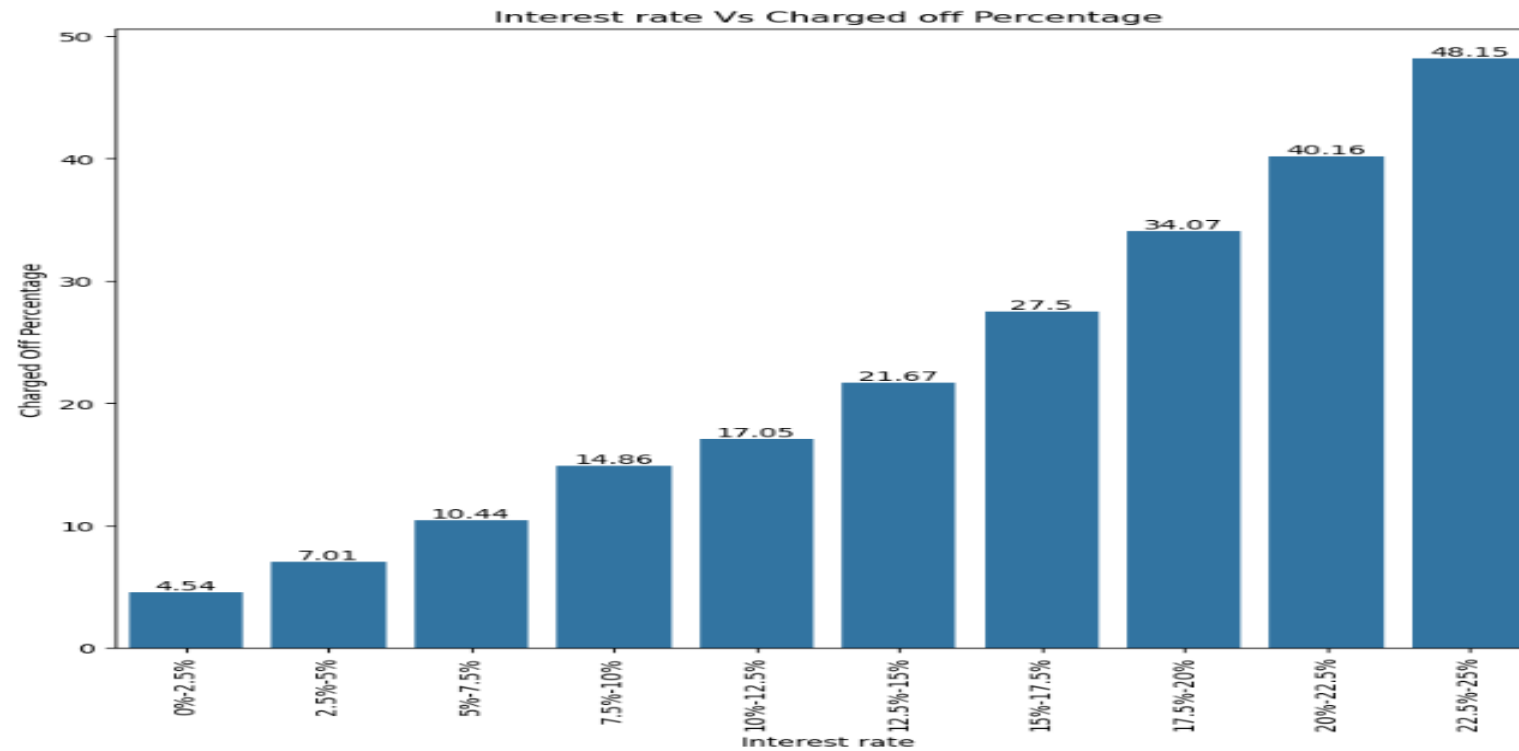
➤ Default Percentage is increasing with increasing loan amount, but its an abrupt curve

Interest Rate

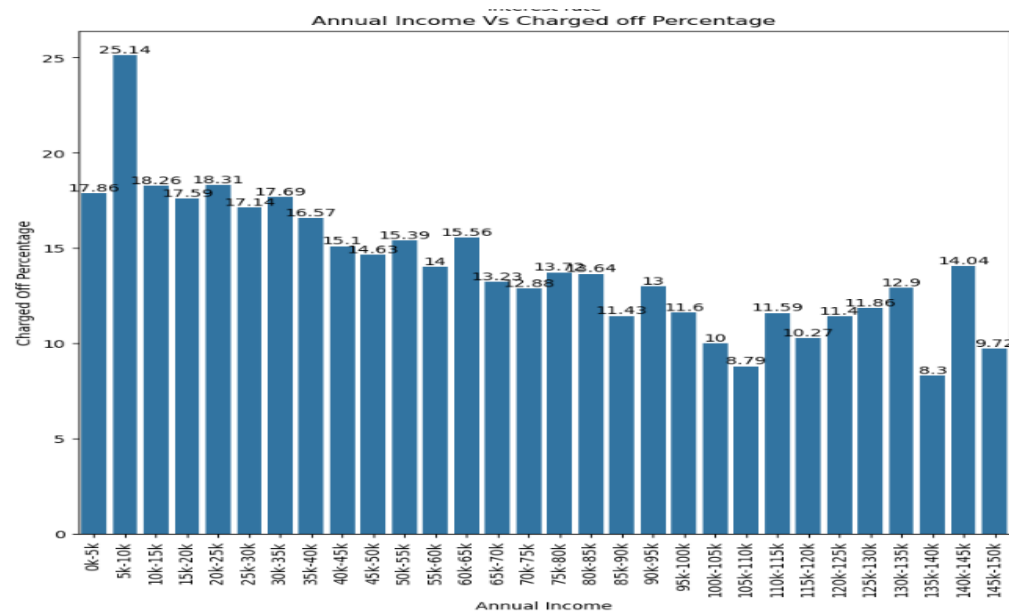➢ There is an exponential increase in default percentage with increase in interest rates

Interest rate Vs Charged off Percentage

Annual Income and DTI

➤ Lower Annual income has higher default percentage and the trend decreases up to annual income of 110k and then default percentage again increases.
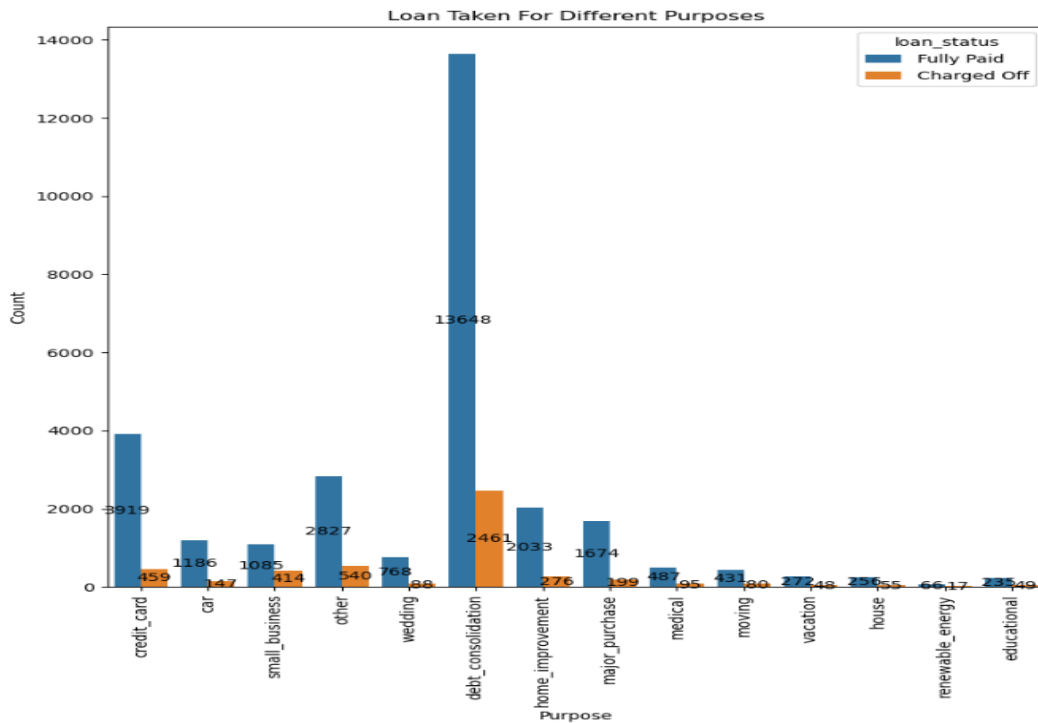
➤ Mid to high DTI has higher default percentage with the default percentage falling in the highest DTI range. This may be because the highest DTI range comes from high income group.
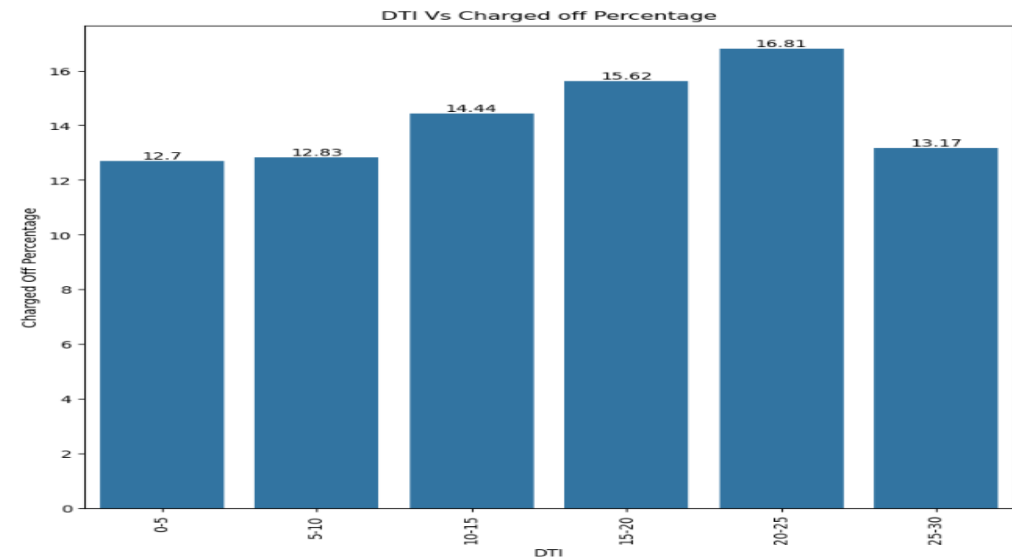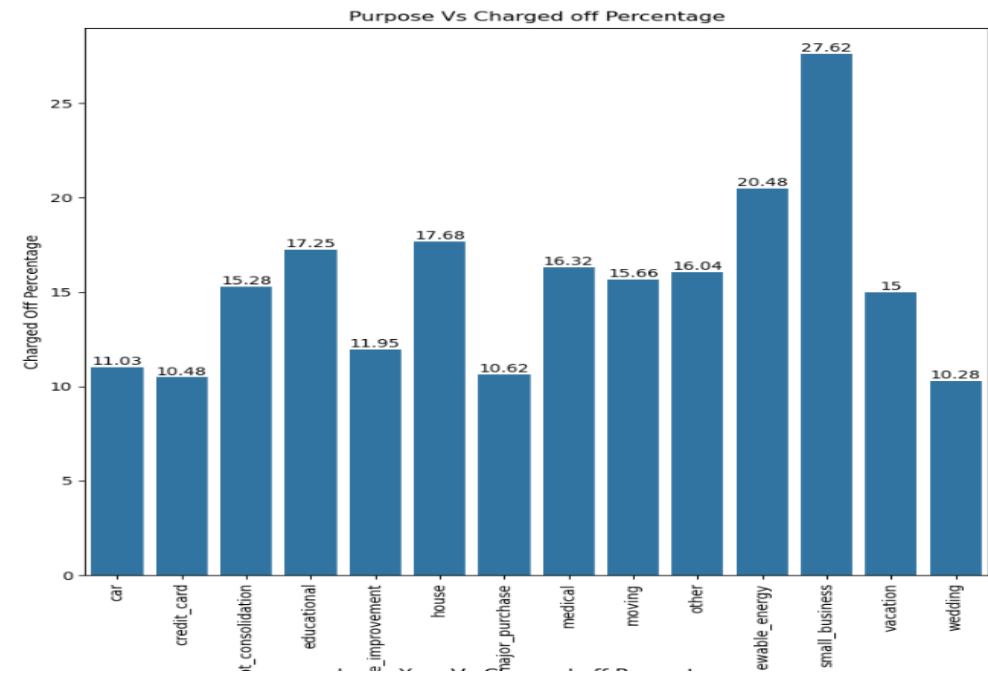
Loans by Purpose

➤ Loans for Debt consolidation are highest in number followed by credit card
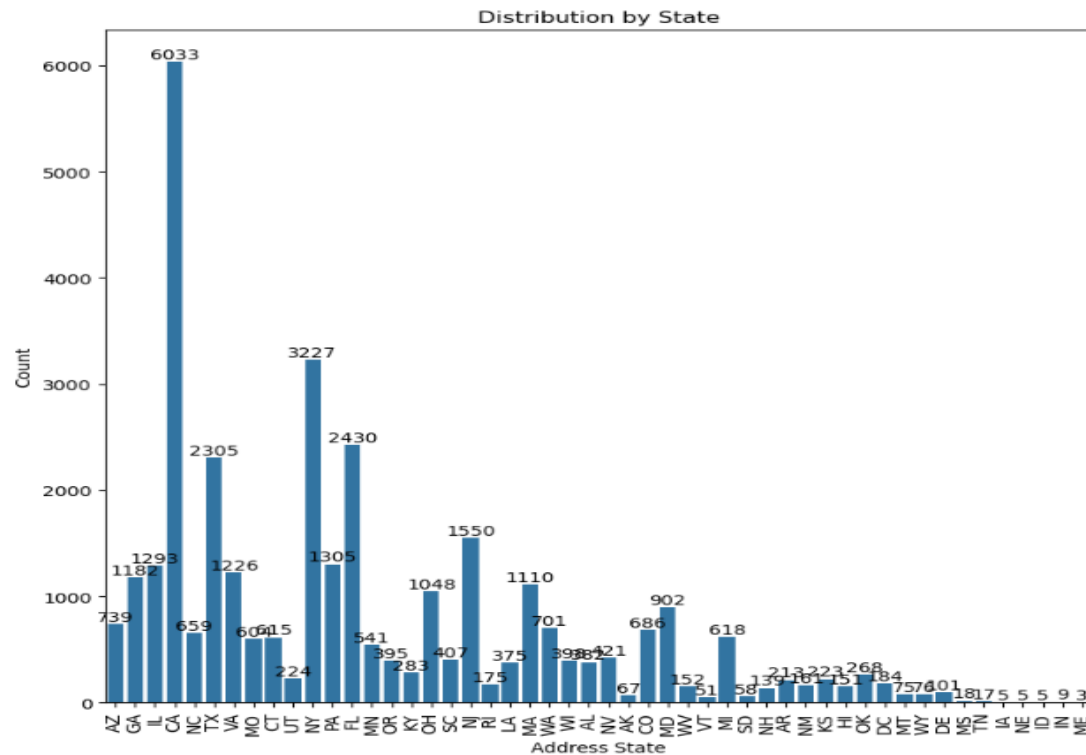
➤ Small business loans have highest percentage of defaults followed by renewable energy and housing loans

Loan Taken For Different Purposes

Purpose Vs Charged off Percentage

Loans by Address State

➢ Major states like California, New York, Florida and Texas have higher number of loans
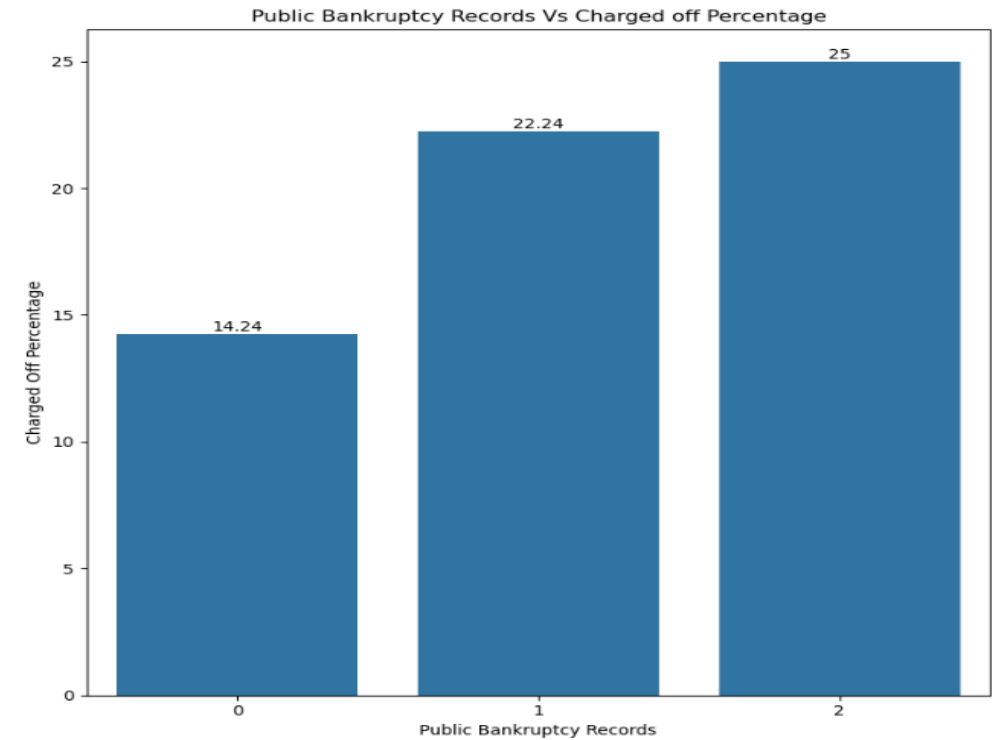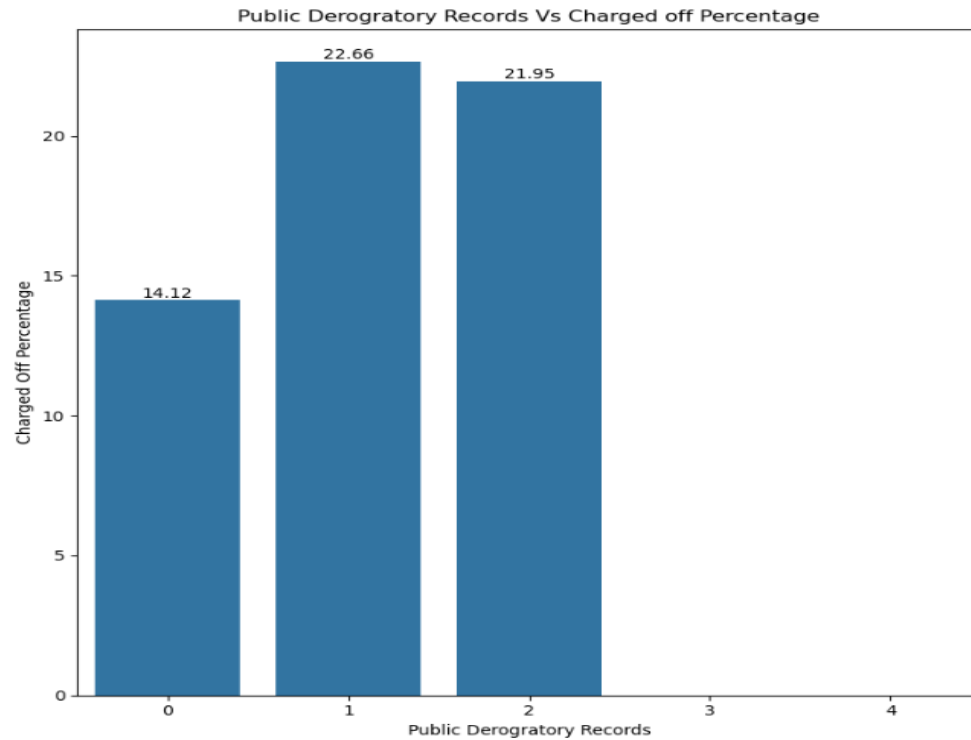
➢ Nebraska has an alarmingly high default percentage followed by Nevada, Idaho and Alaska
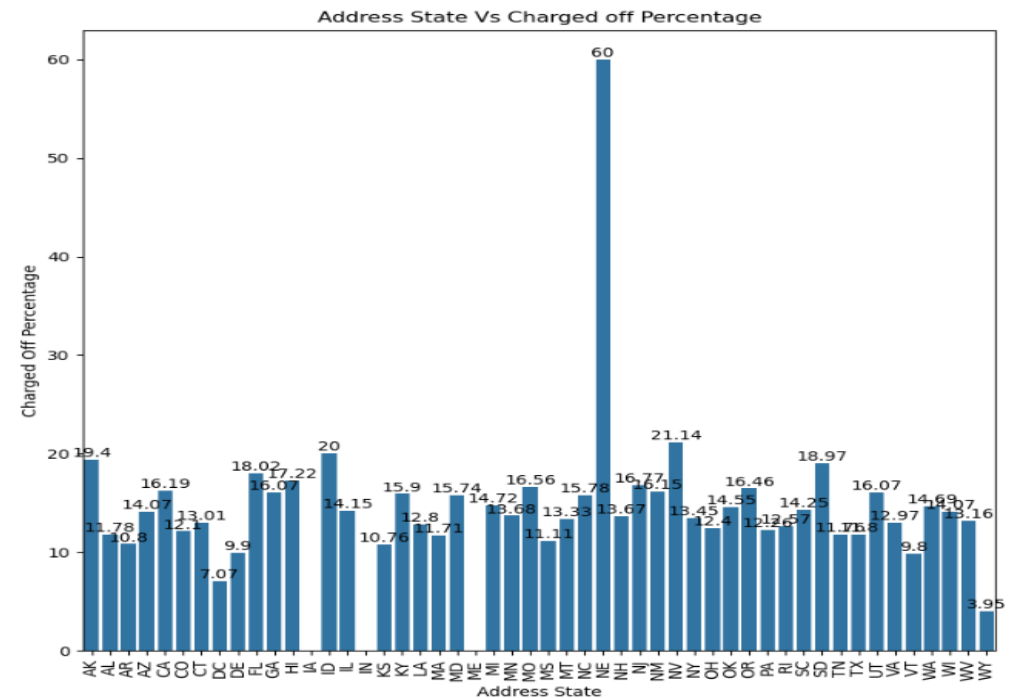
Effect of Public records

➢ The tendency of default is high for borrowers having public derogatory or bankruptcy records

Driving Factors for loan approval

➤ Major driving factors to facilitate loan approval decision are:

- Term
- Interest Rate
- Debt to Income ratio
- Annual Income
- Purpose
- Public derogatory/bankruptcy records

Recommendations

➢ LC should focus on 36 months term loans more than 60 months term loans.

➢ LC should think about reducing interest rate for their very high interest rate loans as they are showing high default tendency.

➢ LC should take special care in verifying borrowers from Nebraska, Nevada, Idaho and Alaska, as they have a higher percentage of defaults.

➢ LC should review their verification process since verified loans are having slightly more default percentage which is not expected.

➢ LC should check the number of open credit lines and the utilization of the revolving balance for the credit lines.

➢ Small business and renewable energy loans should be scrutinized more as they have high default tendency.

➢ LC should review the sub-grading for F5, G3 and G4 loans. G3 and F5 sub-grade can be shifted to the lower sub-grade in G and G4 can be shifted to higher sub-grade in F.

Thank You