Lending Club Case study

S Kaviya   B Avinash

ML C57

Objectives

❑ **Problem Statement**

❑ **Data description and assumptions**

❑ **Defaulter's basic characteristics**

❑ **Distinguishing features of defaulters**

❑ **Commonalities among loan appliers**

❑ **Interesting insights on defaulter pattern**

❑ **Descriptive Analysis**

Problem Statement

- The problem statement defined here is to find the patterns in defaulter behaviour in the given dataset, which will be used in future while analysing a loan application and act to the lending club's best advantage.

- The analysis approach used here on the whole is the to first clean the data and separately analyse 'Fully Paid' and 'Charge Off' customers ,find differences in driving factors between both of them and leave off 'Current' customers as they don't bring much information to the table.

Data Description And Assumptions

Data Cleaning
- Identified the null columns and removed all the null columns from the data frame
- Removed the columns which doesn't add any value to the analysis
- Columns with same values are removed as they don't add any value

Assumptions

- Grades are not categorized , but referring from articles we assumed
  - **A is most credible and G is least credible**

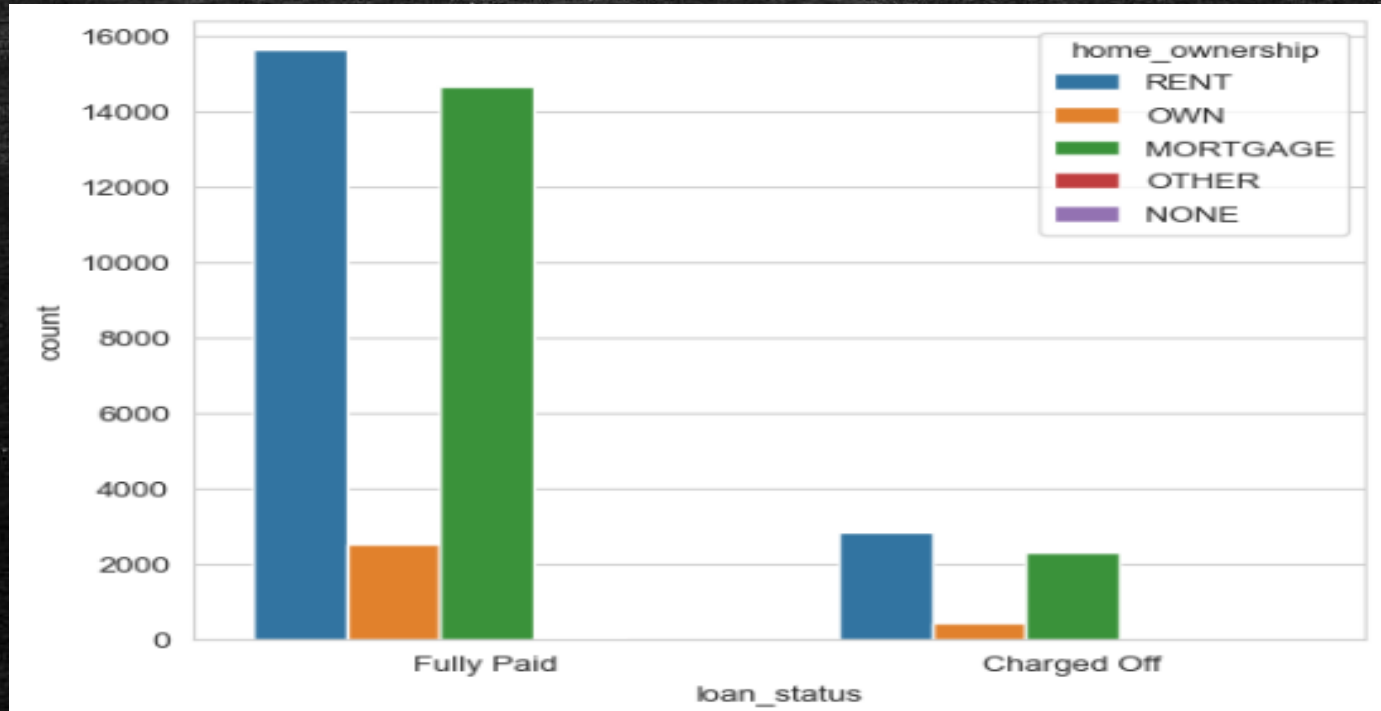- Thereby in subgrades
  - **1-most credible and 5-least credible**

Defaulter's basic characteristics

There are five key factors identified as drivers in defaulter's behaviour
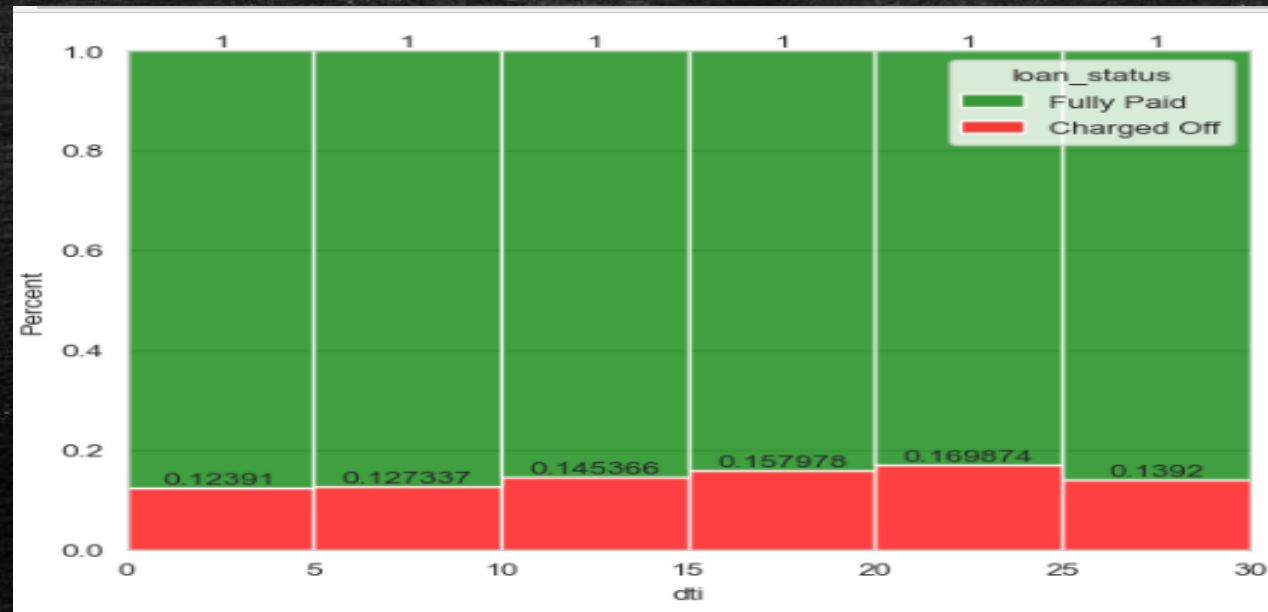
- Home ownership Status

- Interest rate

- Debt to Income ratio

- Grade

- Number of public record bankruptcies

Home Ownership Status

- People with RENT and MORTAGE tend to default more due to those added commitments, or maybe that is a explanation of their financial status to repay a loan.

- 91% of Charged Off people are either paying rent or Mortgage.
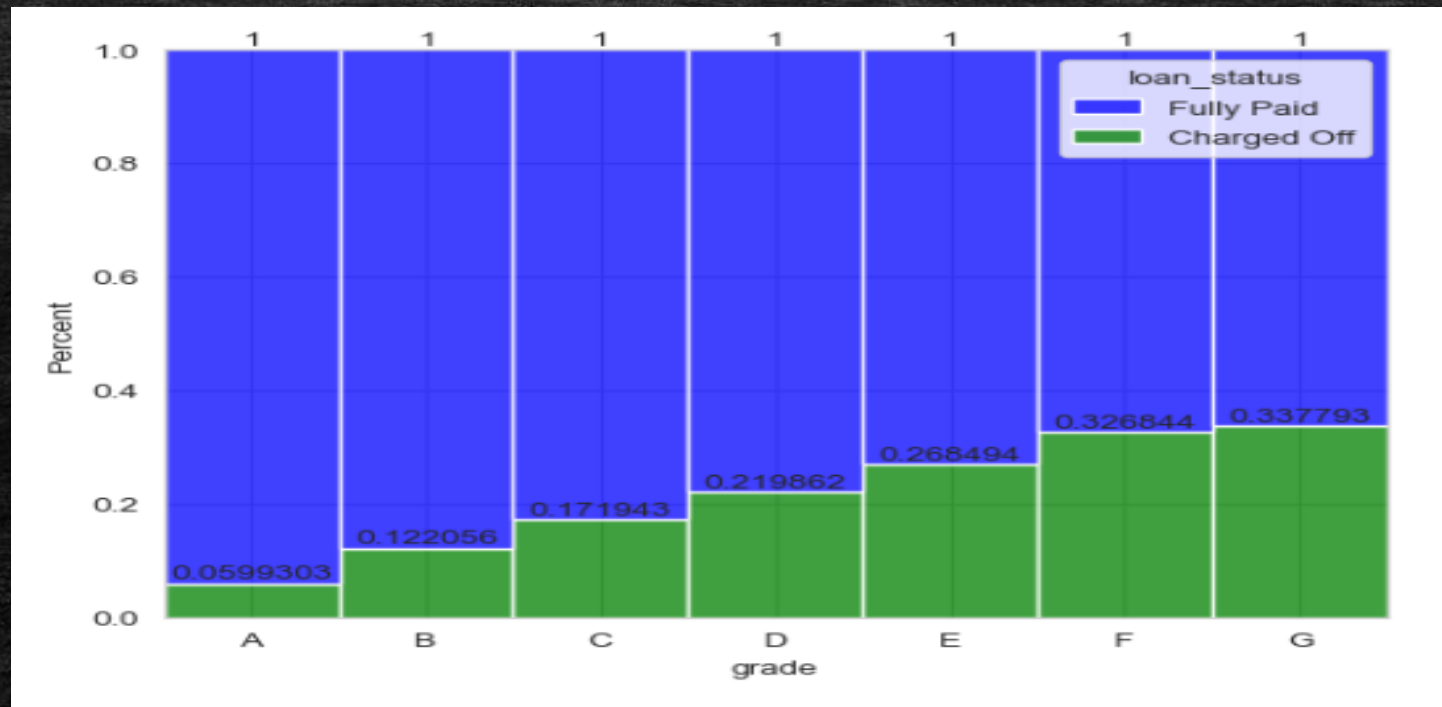
Debt to Income Ratio

- As we can see from the graph when the debt to income ratio of the customer increases, the ability to repay the loan reduces, so Charged Off percentage increases.
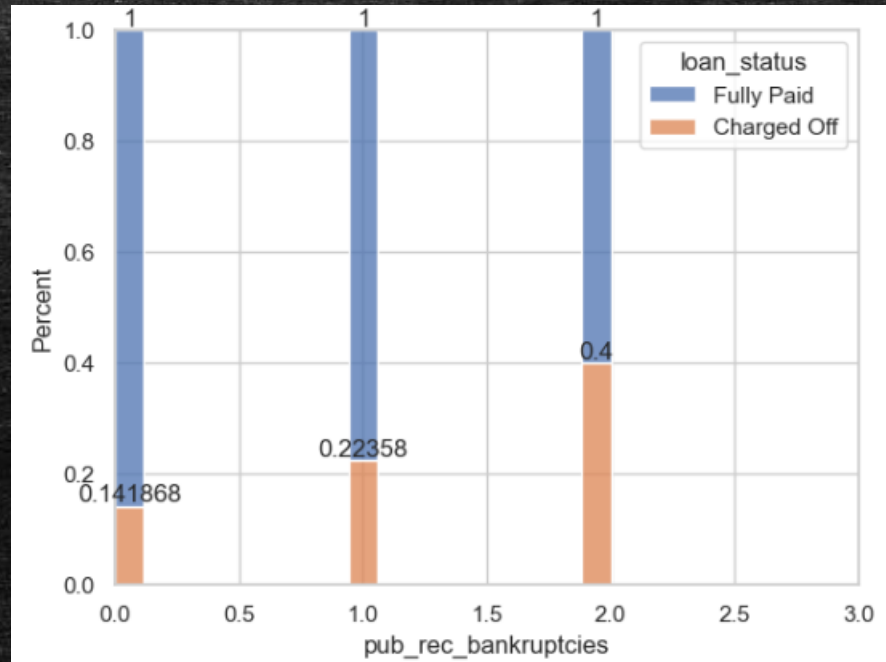
Interest Rate
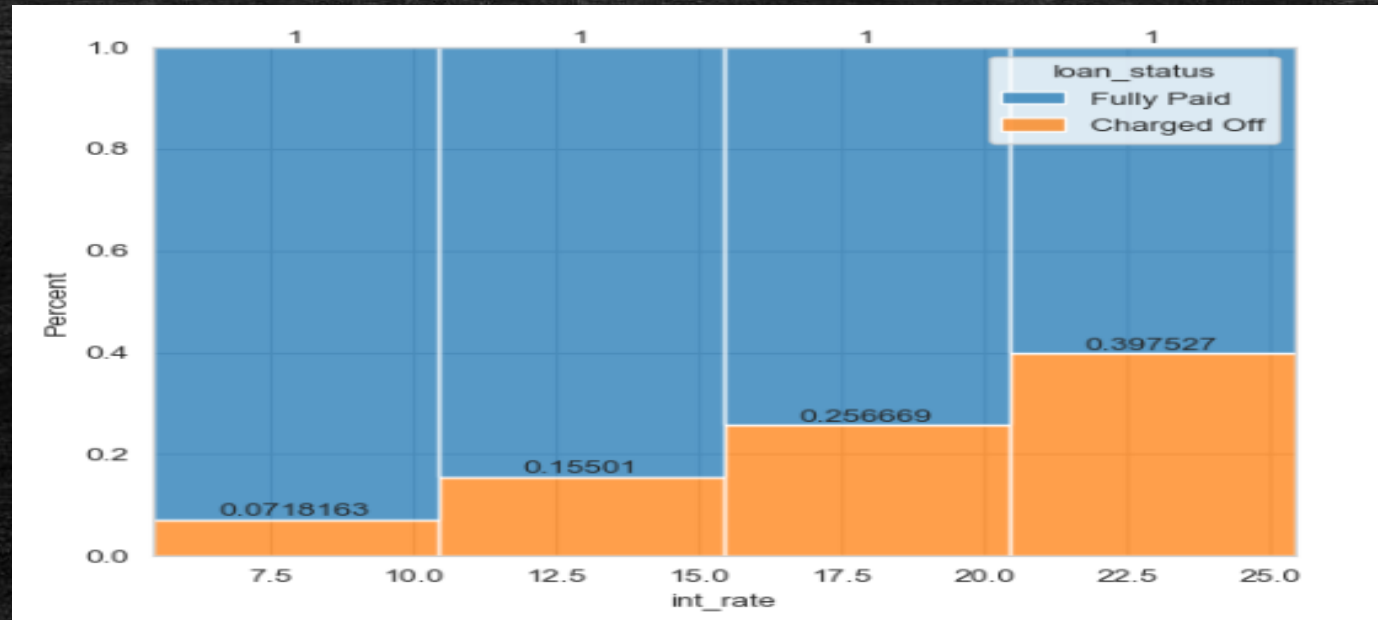
- When the interest rate increases ,that depicts a sure increase in Charge Off percentage.
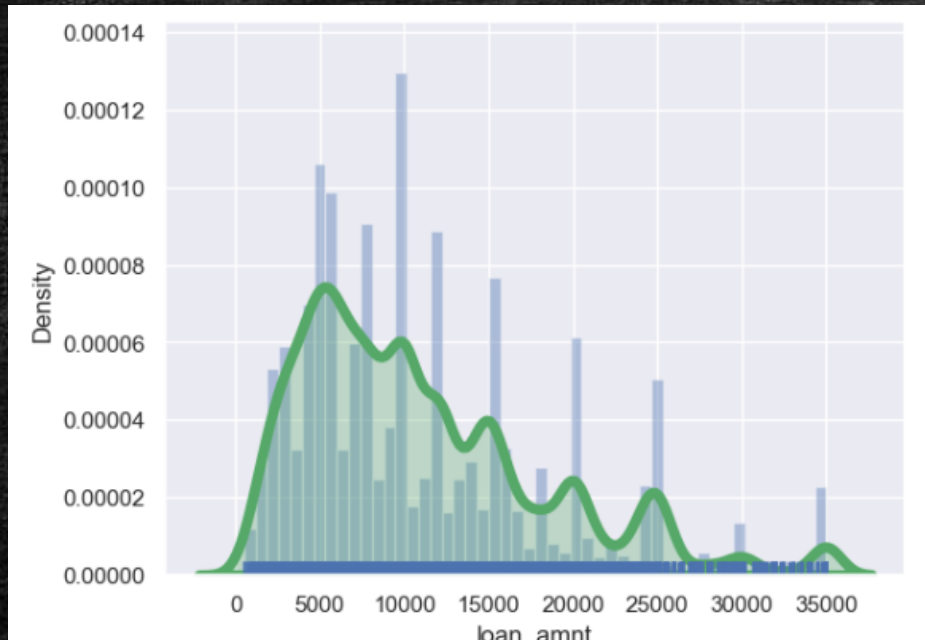
Grade of Customer

- When the grade of the customer goes from to A-G ,charge Off percentage increases depicting the loss of credibility.
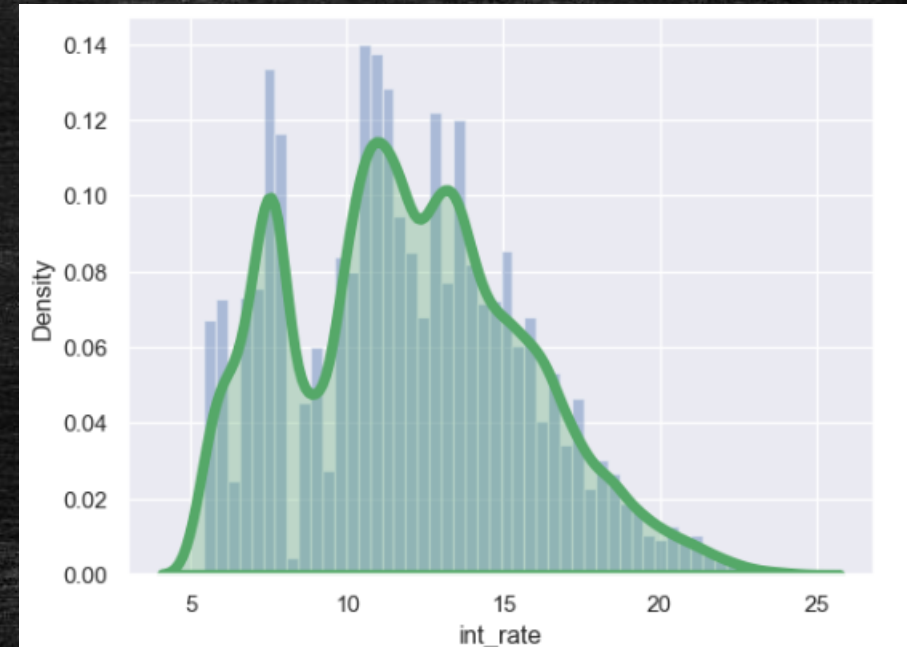
Number of Public Record Bankruptcies

- As observed in the above graph, the defaulters are increasing with the increase in public record bankruptcies
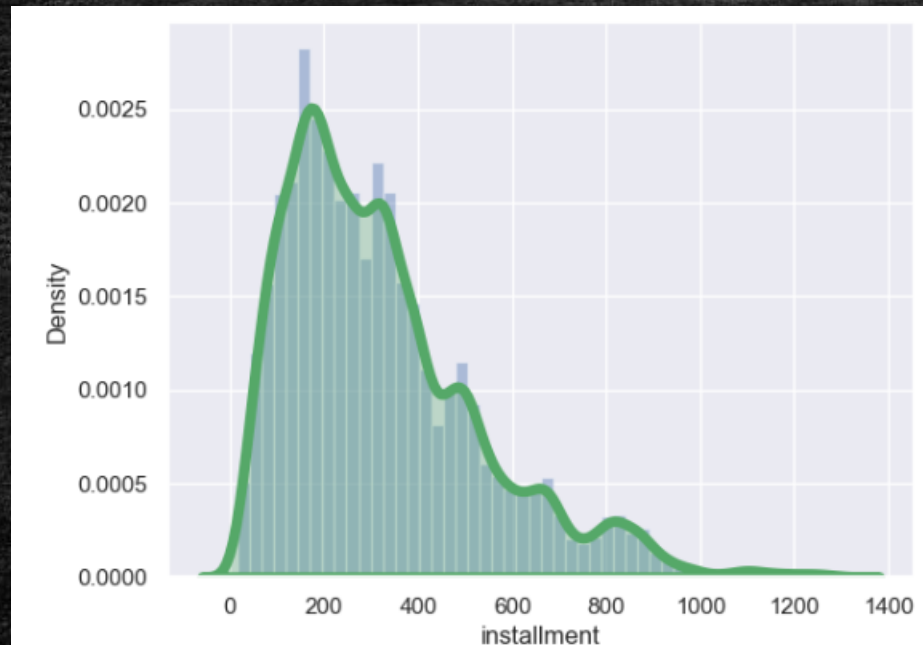
Loan Amount and Interest Rate

▪ The loan amount varies between 500-35000 with the median value of 9600 for all the 'Fully paid' and 'Charged-off' customers. Most of the customers have the loan between 5300 to 15000.
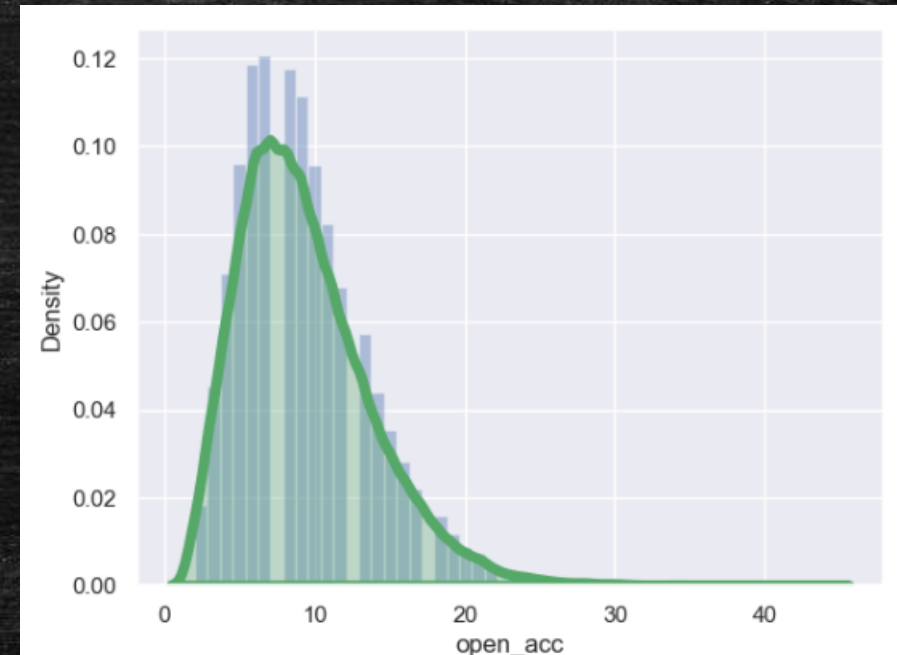
▪ The interest rate provided for the loans range between 5-24 and the majority of it lies between 8-14
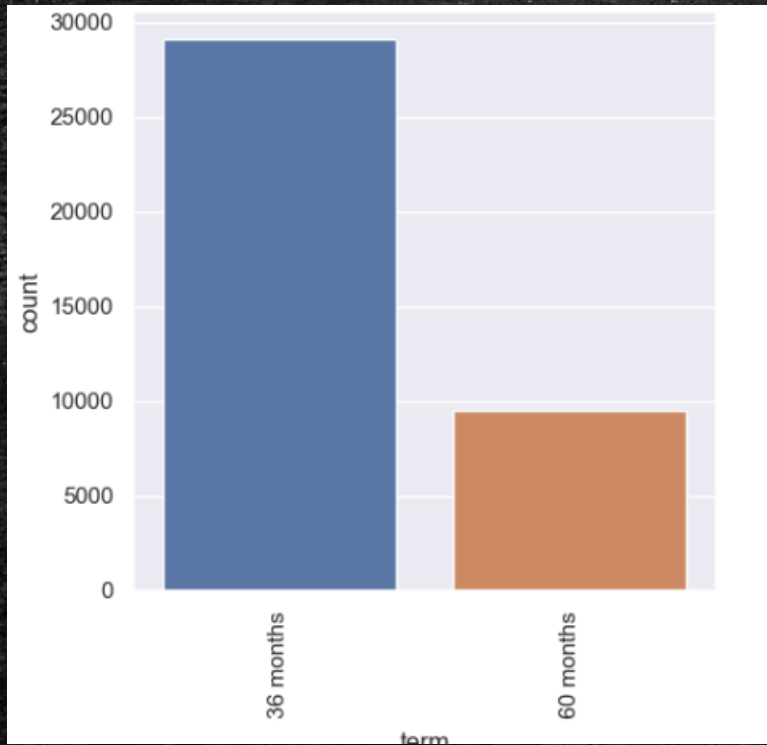
Installment and Total accounts

- ▪ Monthly installment varies from 15 to 1305, and the most of it is under 931.
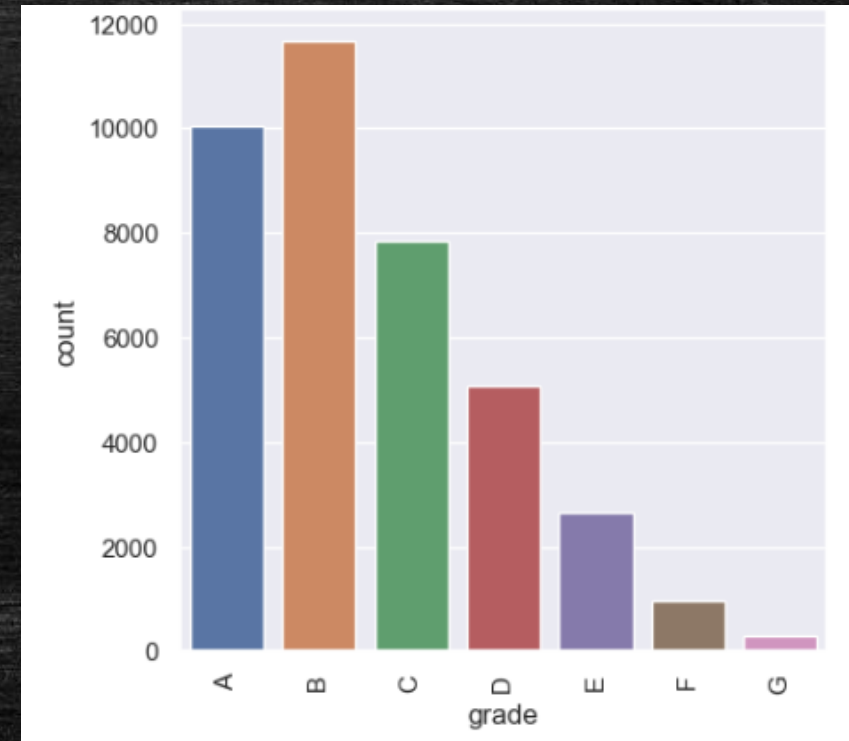
- ▪ 50% of customers have 6-12 open accounts, and 9 accounts on an average
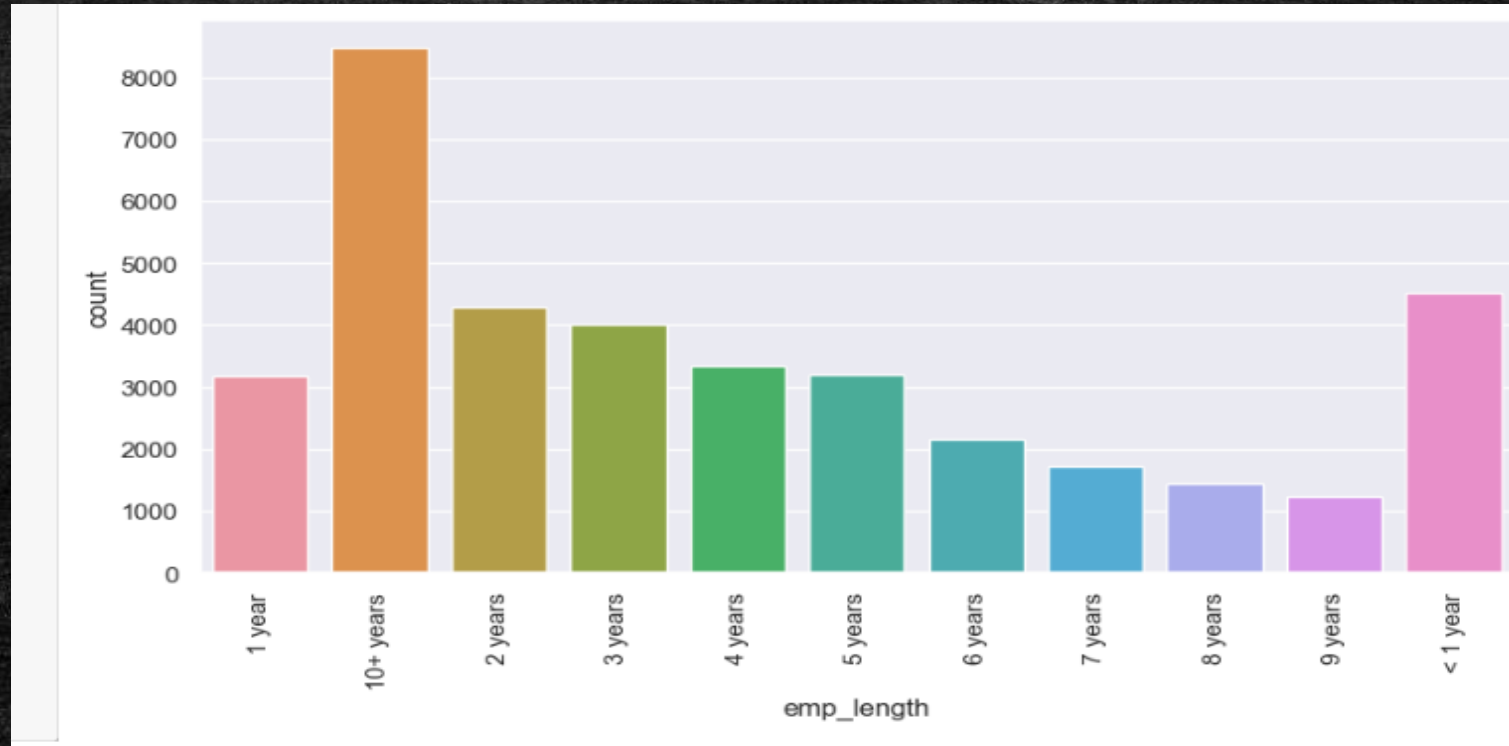
Tenure and Grades

- Most of the people applying for loan choose 36 months tenure term.
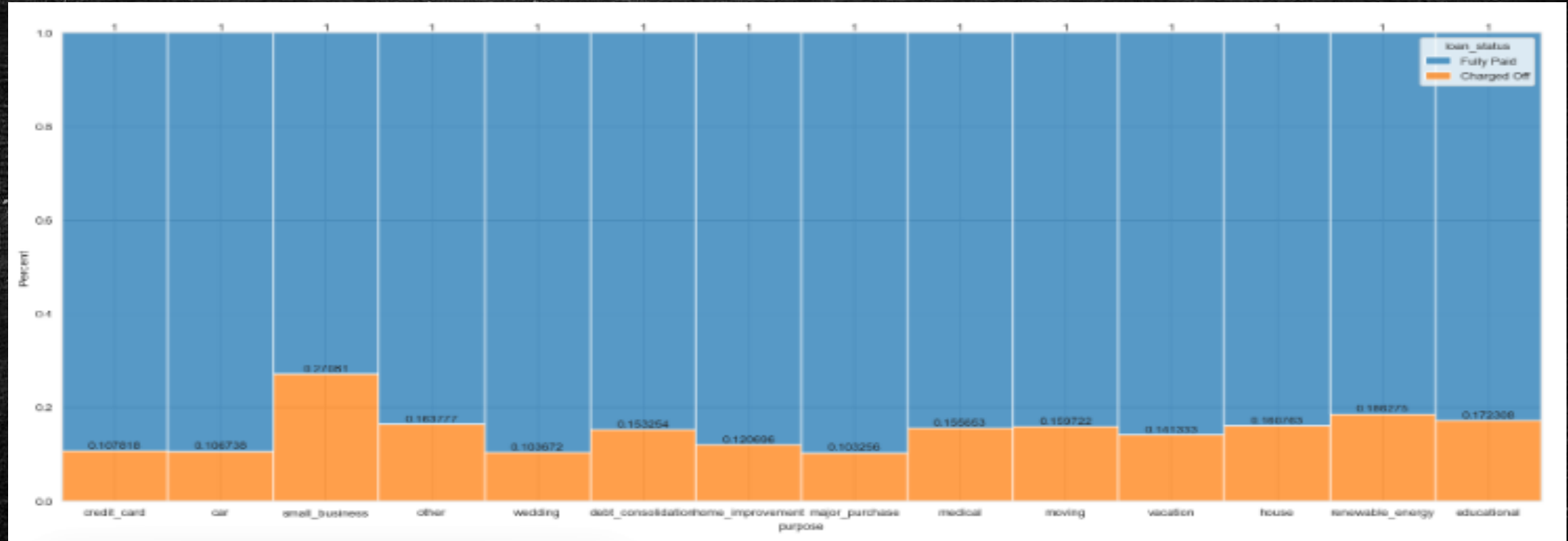
- Larger portion of candidates applying for a loan fall in Grade B

Employment Length

- Higher number of people applying for loan are employed more than 10 years , which is understandable . But their immediate next successor is people employed less than a year , which could be explained by the increased need ,sudden cashflow and a 'can' attitude.

Purpose of the loan

- Higher percentage of defaulters fall in the top 3 categories of small businesses, renewable energy and education which are mostly volatile to receive expected money.
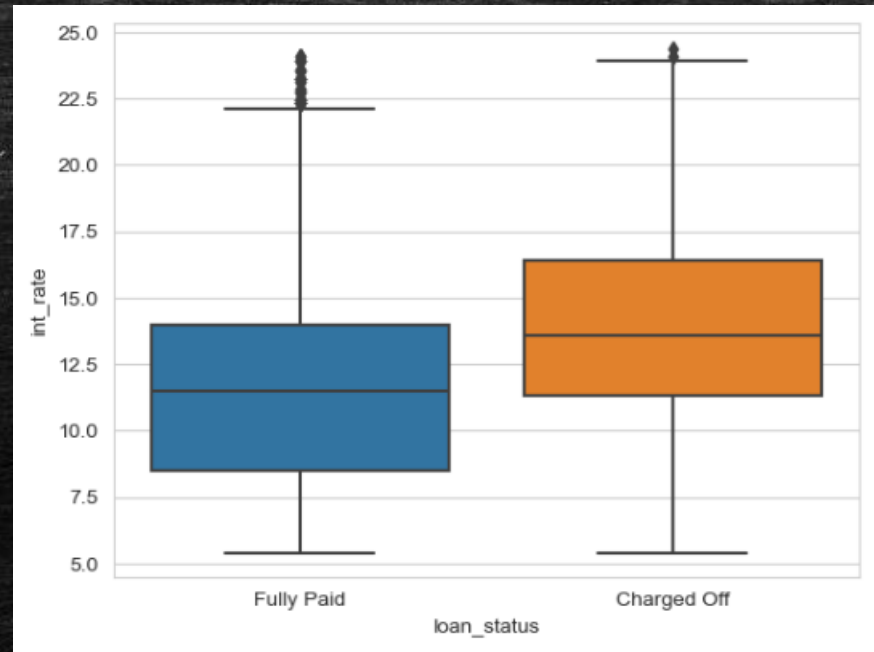
Distinguishing Features of Defaulters

Bivariate

Loan Status and Total Recovered Late fee

- Customers who are paying late fee tend to be defaulted as observed in the above chart the median varies between Full paid and Charged off
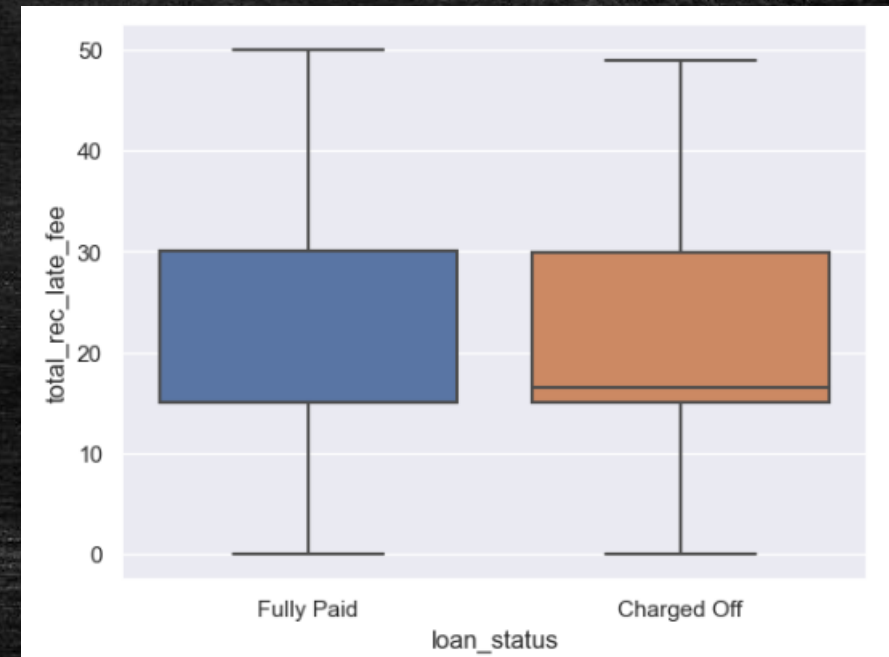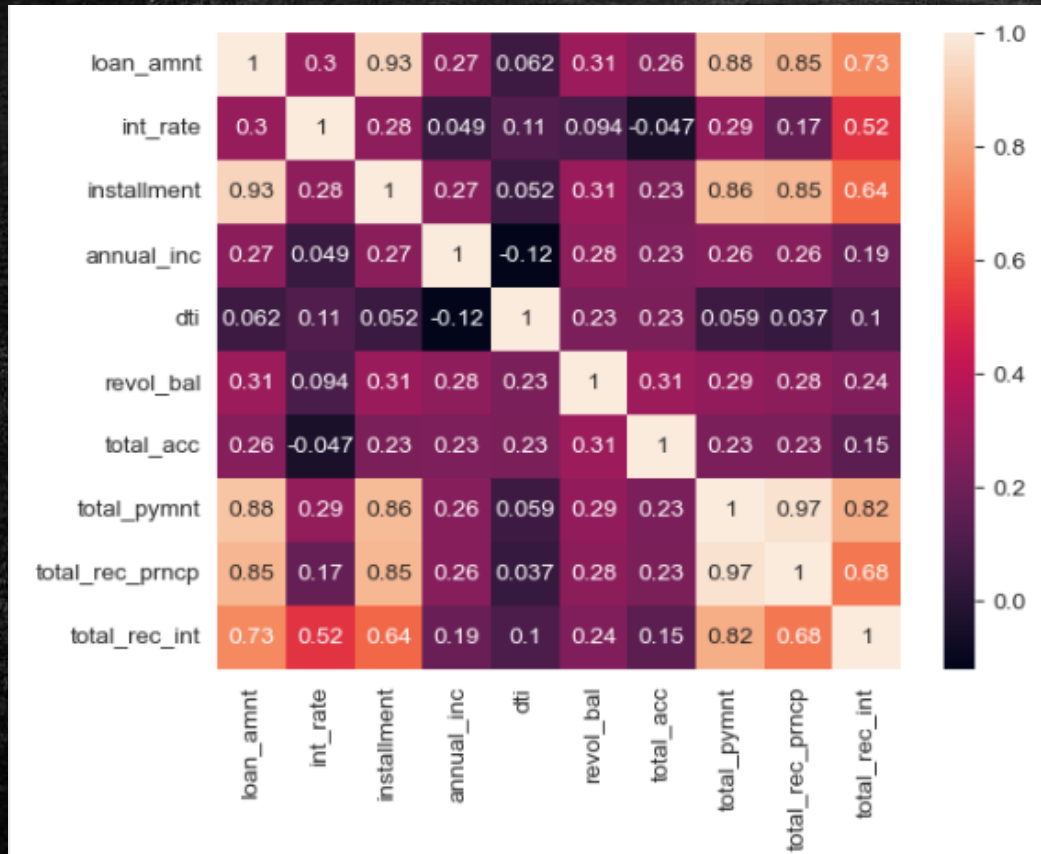
Loan Status to Interest Rate

- 50-75% of Charged Off people lie where 75% of fully paid interest rate lies.

- This denotes when interest rate increases there is a higher chance of defaulting.

Commonalities among loan appliers

Multivariate

From the correlation matrix certain observations were made,

- Loan amount and installment amount are highly positively correlated

- Total payment , Total recovered principle and Total recovered interest are as a group highly positively correlated

- Loan amount is well positively correlated to the above group

Interesting insights on defaulter pattern

Derived metric

Payment Difference calculated from loan amount and total payment
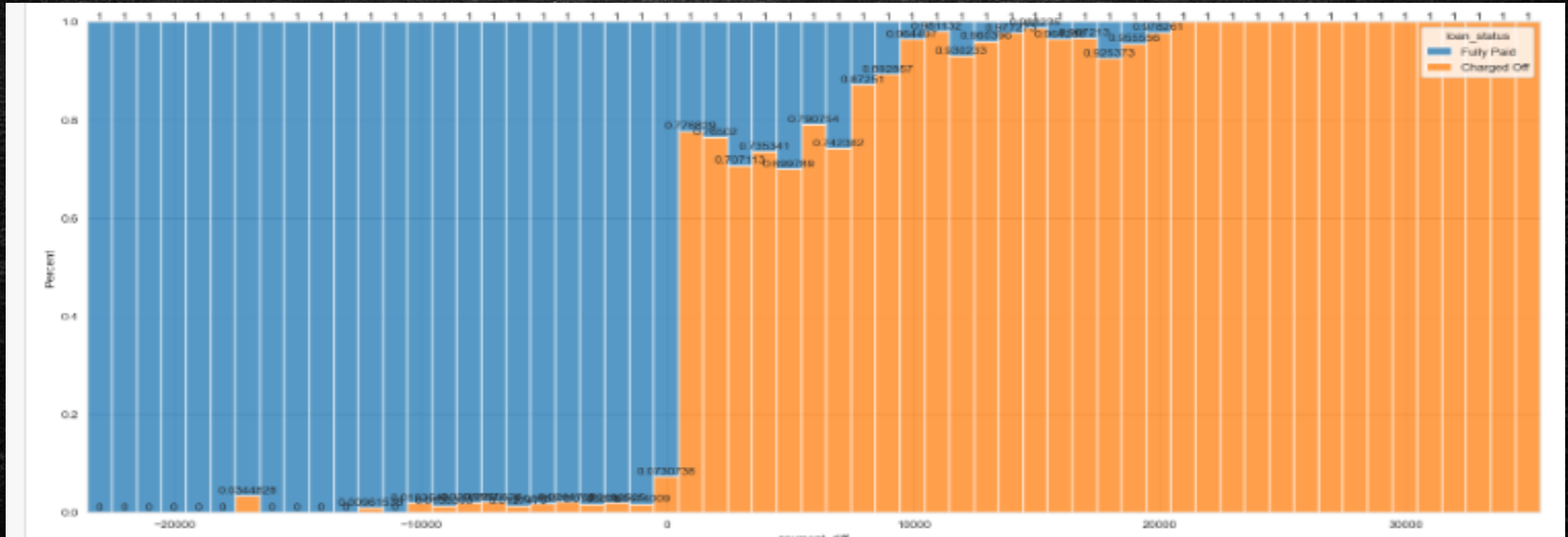
▪ **Payment_difference = Loan_amount - total_payment**
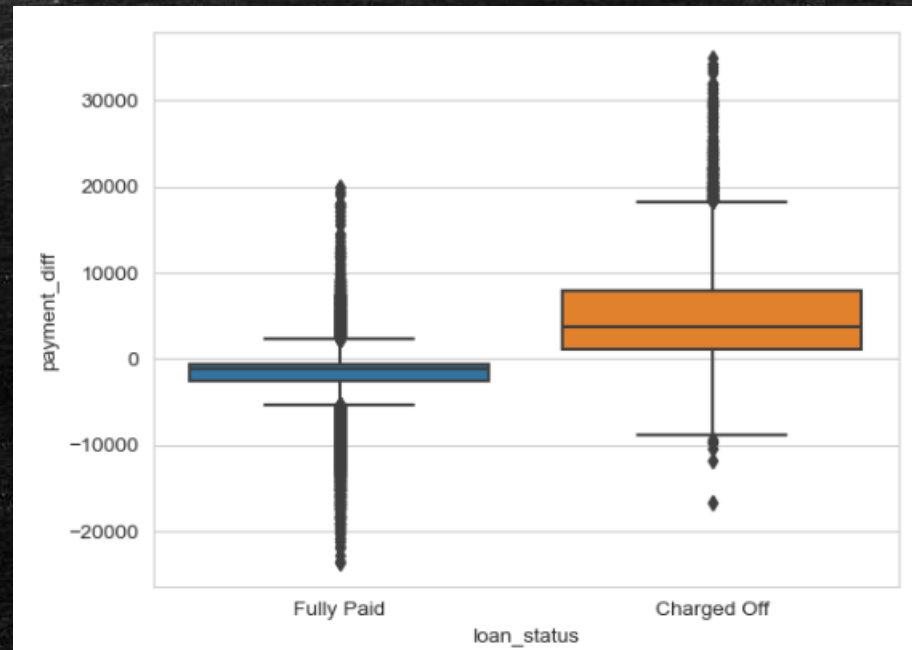
▪ 25 -50% charge off happens when total payment reaches loan amount, customer's may think they have paid enough morally, if they are not able to pay financially
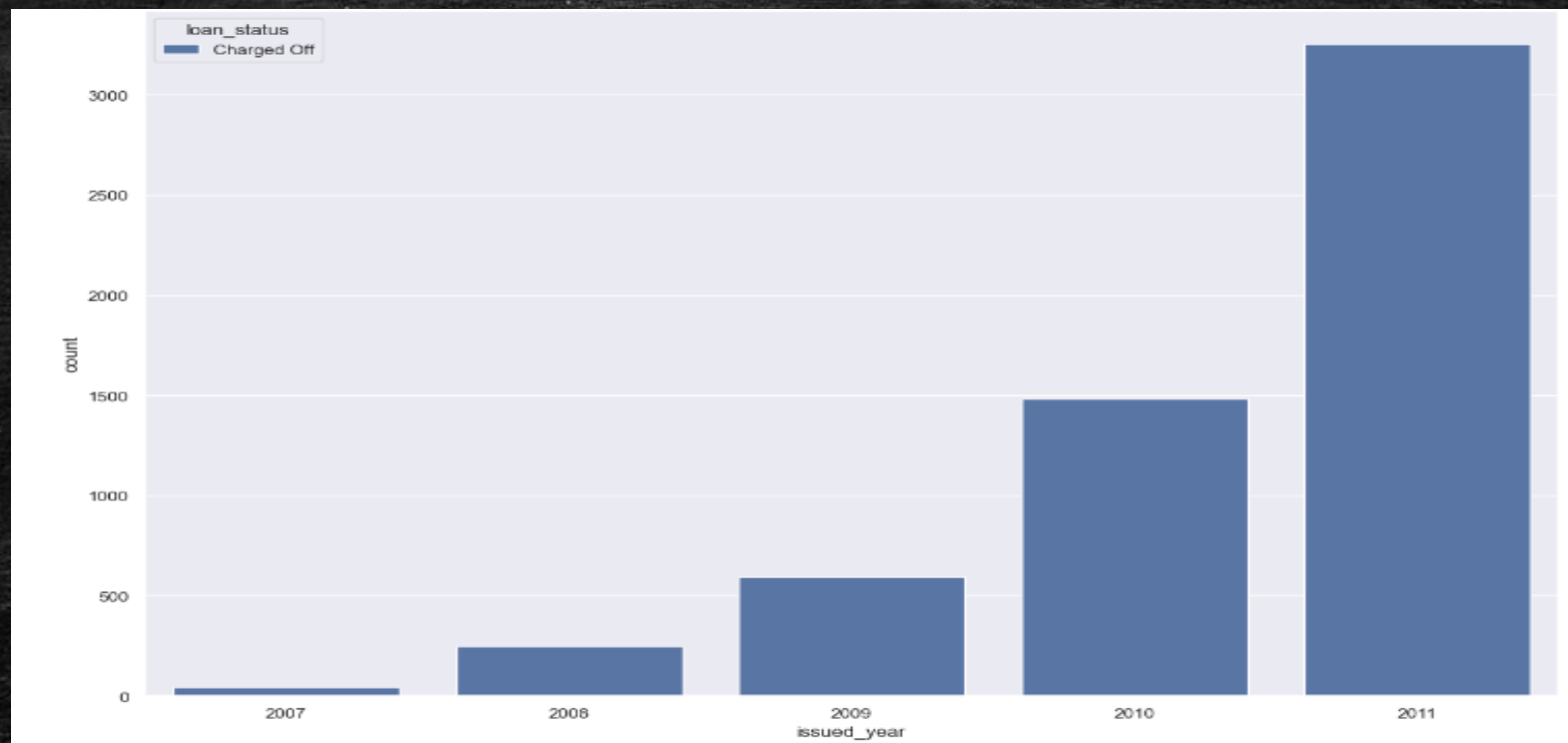
Payment Difference

Derived metric – Business Driven

- When the customer pays off more than the disbursed loan amount chances of them fully paying the loan increases.

Issued Year calculated from Issued Date
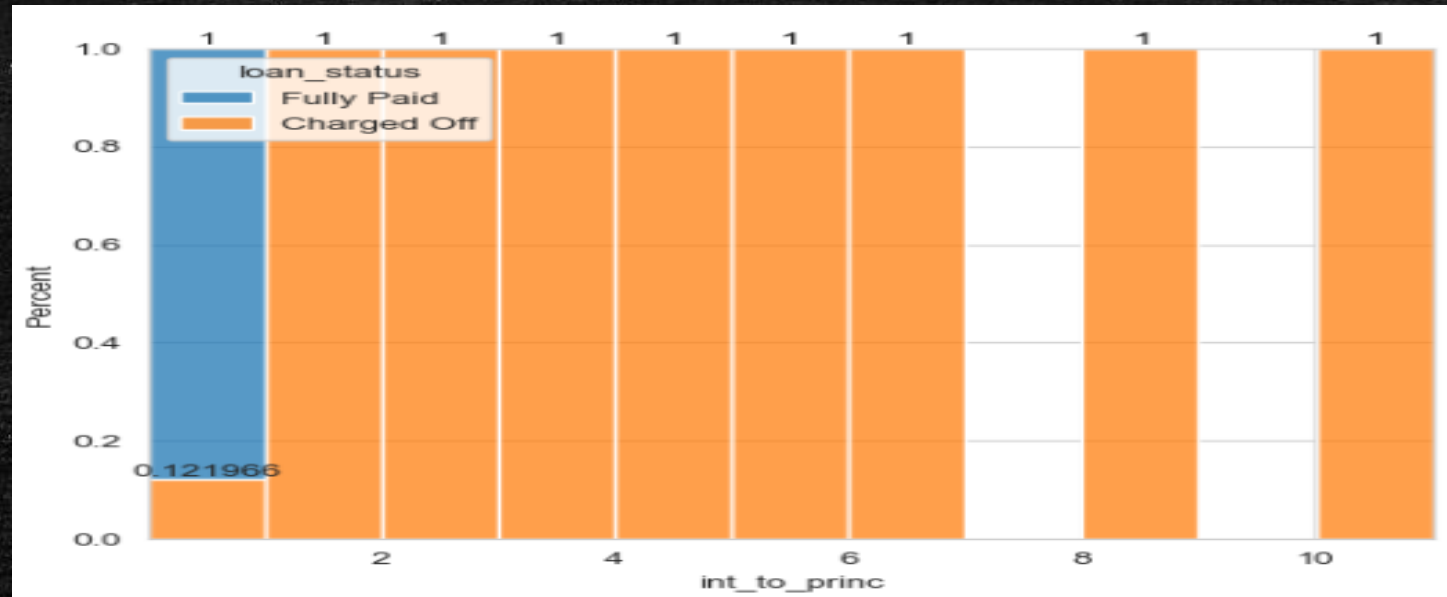
Derived metric – Type Driven

- Defaulters are increasing with increase in the issued year

Interest to Principle Ratio

Derived metric – Business Driven

- Total received interest to Total received principle ratio is derived to find where charge off's happen more.

- Int_to_princ = tot_rec_prncp / tot_rec_int

- When the total received interest to principle ratio is more than 1 ,customers totally chargeOff , considering a substantial part of amount they paid went to interest.

Recommendations

Based on the observations , certain recommendations can be made to reduce defaulter percentage , thereby reducing the money lost.

- Firstly , credibility of customers is very important ,so when customer's grade deteriorates , we should check for any other very advantageous factors before issuing a loan.

- One such advantage can be owning a home , because home owners tend to default less . Thereby issuing a loan to rent and mortgaged customers should be taken care with caution.

- When any of the expected criteria fails ,so in order to still issue loan and compensate other factors ,if we increase interest rate , there is a high possibility that would be paid off. So instead increase tenure , which doesn't have much impact on defaulting.