

LENDING CLUB CASE STUDY

Submitted By- Namita Rani and Subhashis Pradhan

Introduction

Lending Club is the largest lending platform, facilitating loans loans. Investors are able to search and browse the loan listings on Lending Club website and select loans that they want to invest in based on the information supplied about the borrower, amount of loan, loan grade, and loan purpose. Lending Club wants to understand the driving factors behind loan default, i.e. the driver variables which are strong indicators of default. The company can utilize this knowledge for its portfolio, risk assessment and to minimize the risk of losing money while lending to customers.

PROBLEM STATEMENT

Identify patterns using EDA to understand how consumer attributes and loan attributes influence the tendency of default for the given dataset for loan applicants.

Problem Solving Approach

1. Analyzing the Default Rate w.r.t to one or more variables
2. Finding Correlation between one or more variables to see how they are related.
3. Analyzing the spread and variations of one or more variables.

Data Cleaning

1. Drop columns with more than 90% null values
2. Drop Columns with no significance for further data analysis
3. Dropping rows with null values.

Driving Metrics

1. Creating derived metrics
2. Binning the numerical columns.

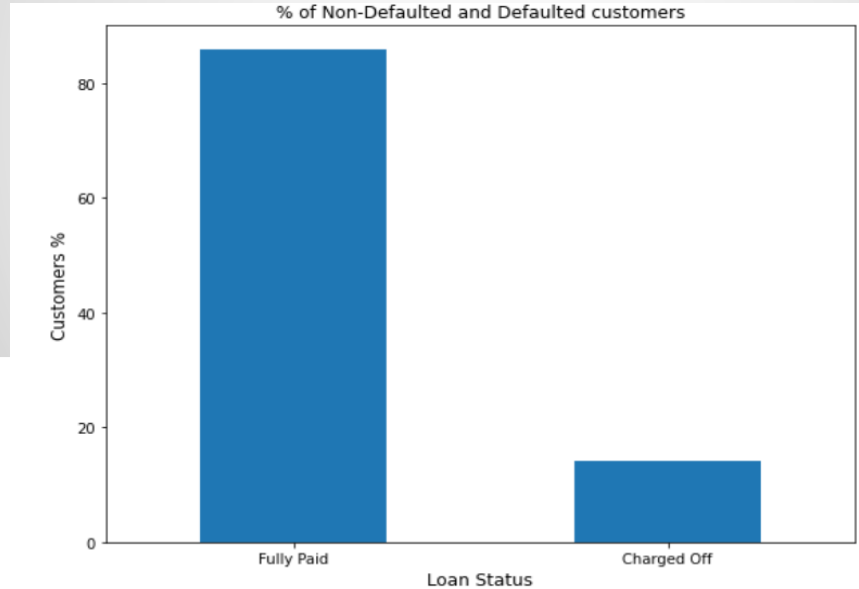
Univariate Analysis

Segmented Univariate and
Univariate Analysis

Bivariate Analysis

Conclusions

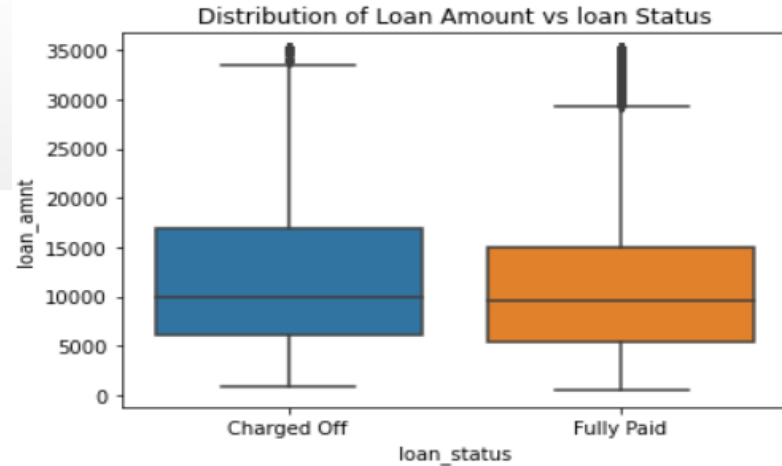
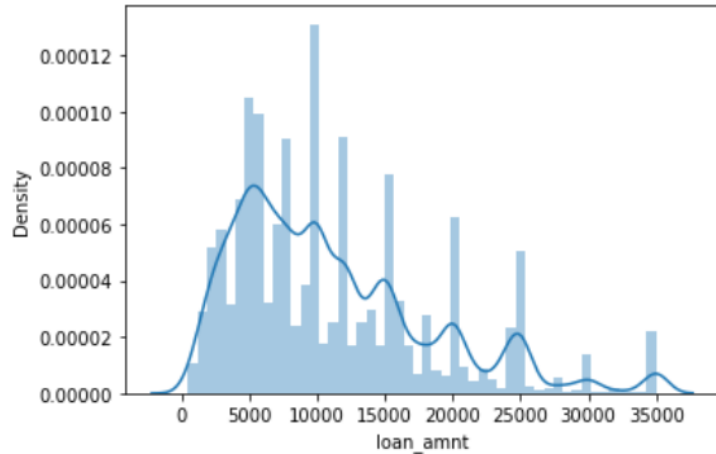
Loans Status Analysis



Around 14% loans being charged off. And 86% being Fully Paid.

Loan Amount Distribution and loan amount across Loan Status Plot

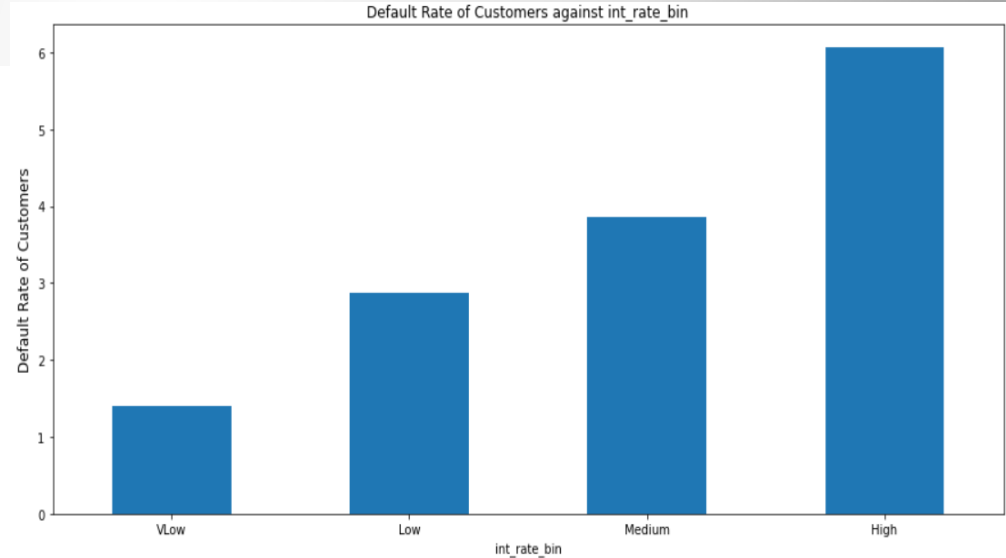
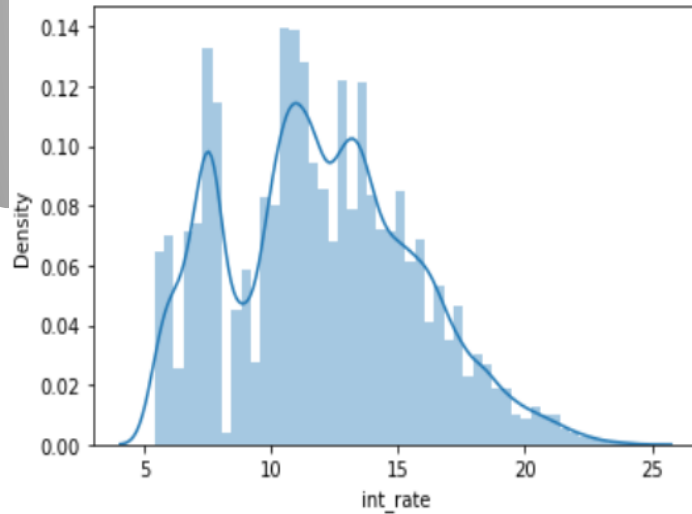
Distribution of Loan Amount



We can see that most of the loan in the charged off category has loan amount between 6k to 17k. Also, Charged Off Customers borrow more loan amount.

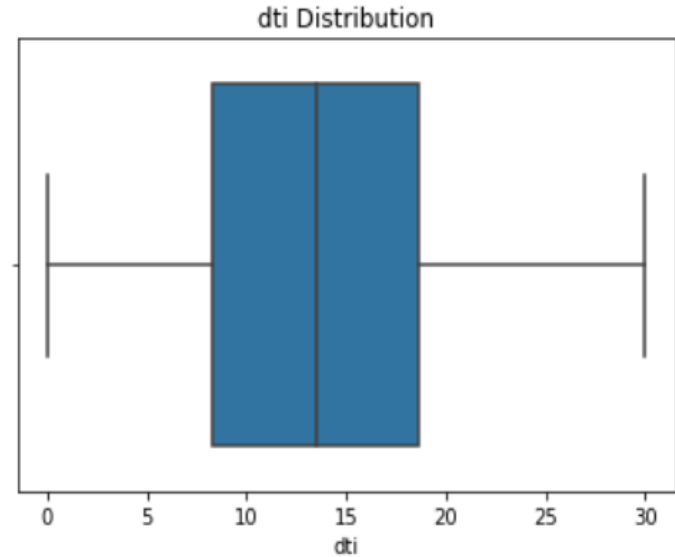
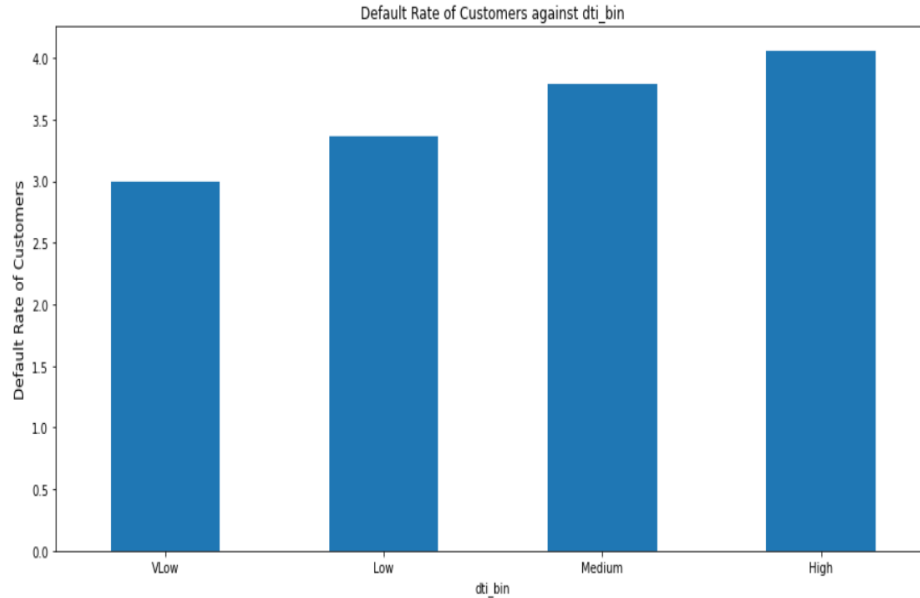
Distribution of Interest Rate and Impact of Interest Rate on Default Rate

Distribution of interest Rate



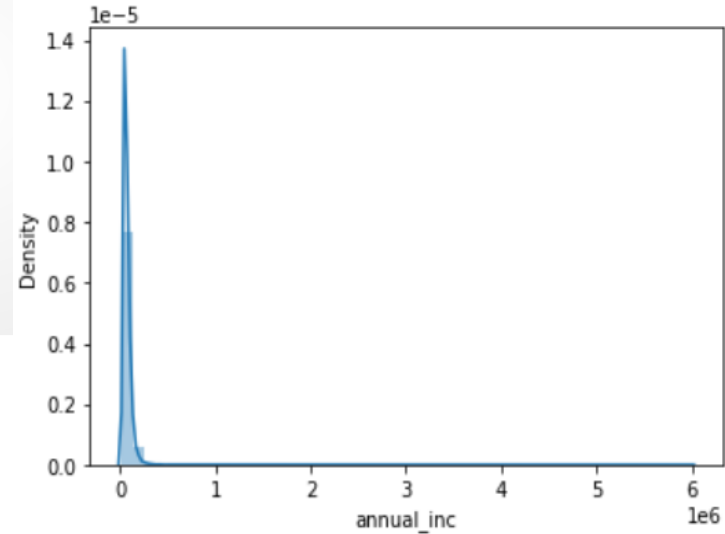
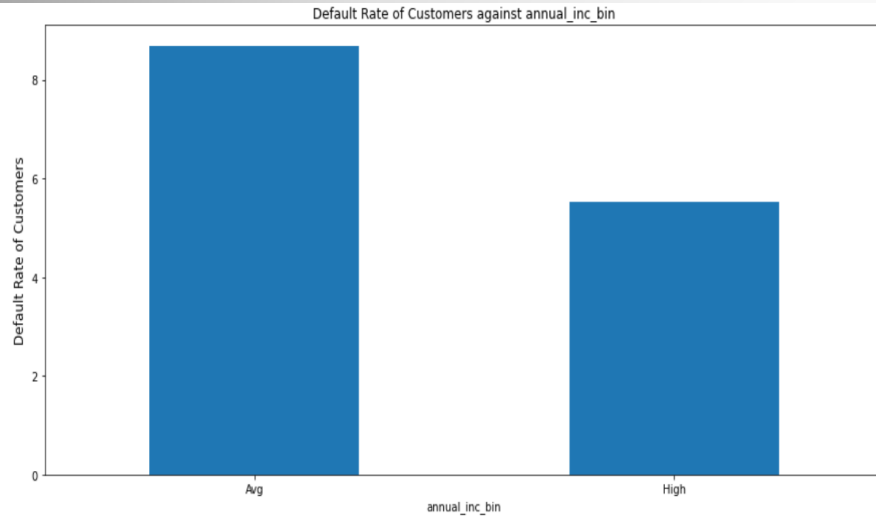
The distribution showing interest rate is in the range of 11% to 13% for most loans. It is evident from above plot that Default Rate of Customers is high for loans with higher interest Rate.

DTI impact on Default Rate and DTI Distribution



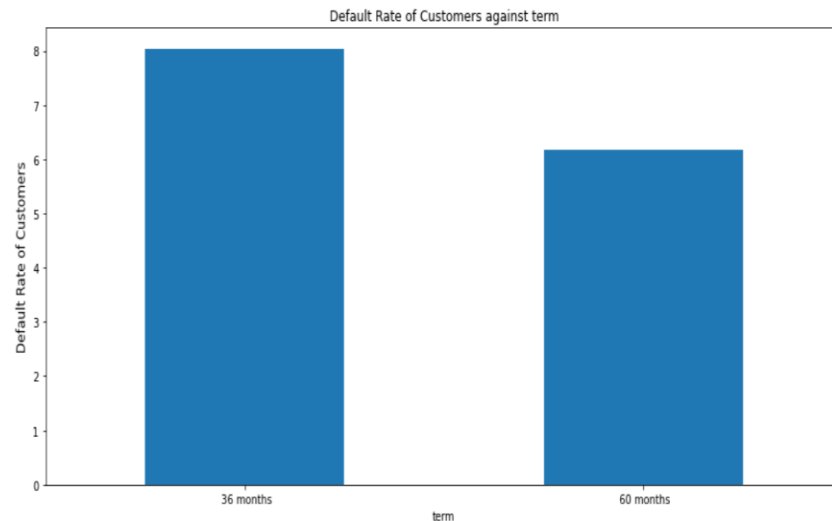
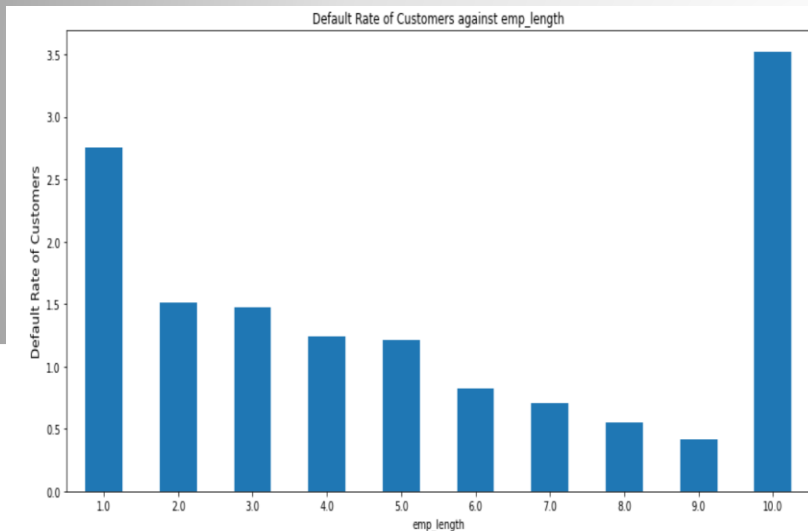
Customers who have high dti value tend to default more.
Most dti values for the customers lie between 7- 20

Impact of Annual Income on Default Rate and Annual Income Distribution



Customers with low annual income tend to default more than the customers with high average income
Also, it's clear from the graph that for most customers annual income lies between 50k-80k

Impact of Emp Length and term on Default Rate



Default Rate of Customers decreases as the emp length of Customers increases.

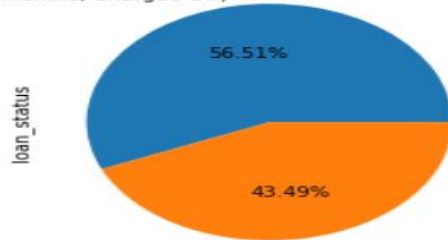
For Emp length > 10 years, there is max default Rate which can be because of 2 reasons:

1. Many of the customers having emp length > 10 can be Retired, implying no income.
2. There is a huge data for emp length > 10 category than for other categories, hence more default Rate.

Term of loan and Loan Status

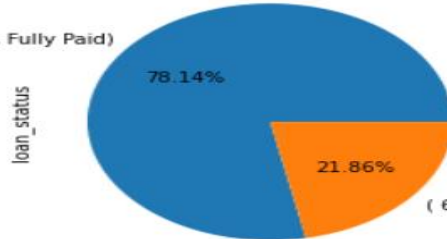
Comparison of fully paid and charged off loans across different term

(36 months, Charged Off)



(60 months, Charged Off)

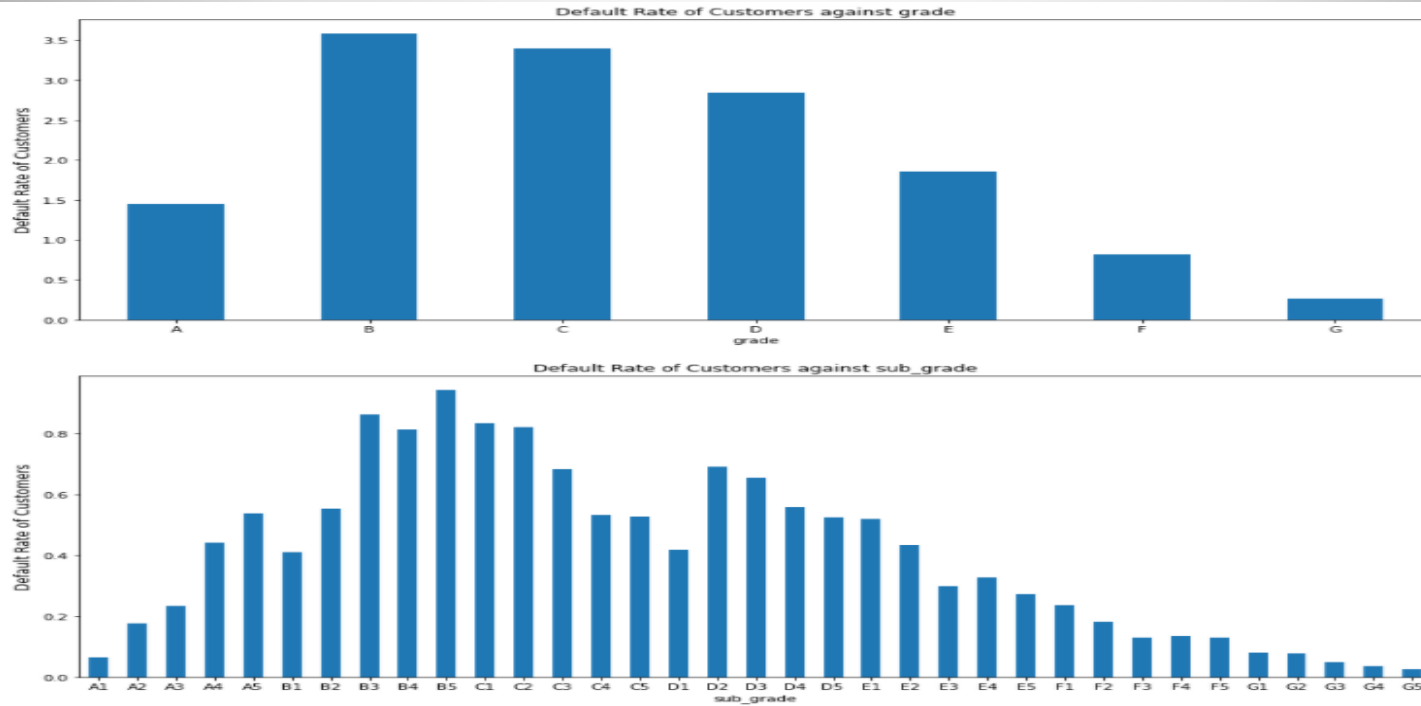
(36 months, Fully Paid)



(60 months, Fully Paid)

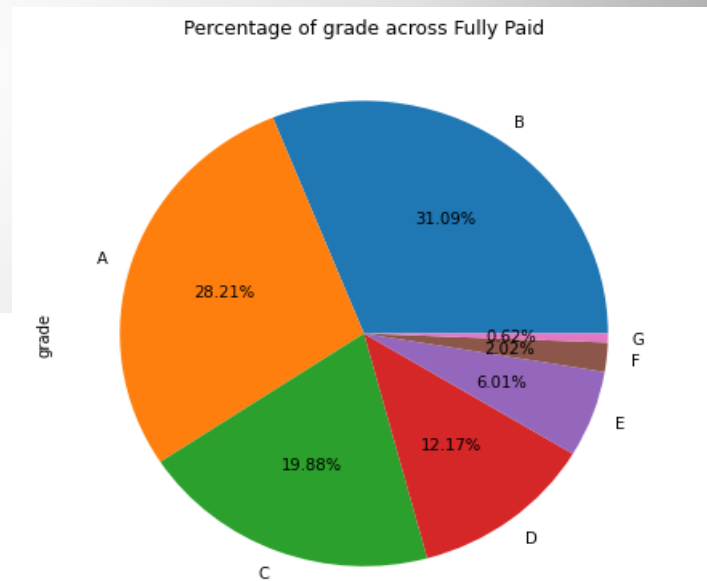
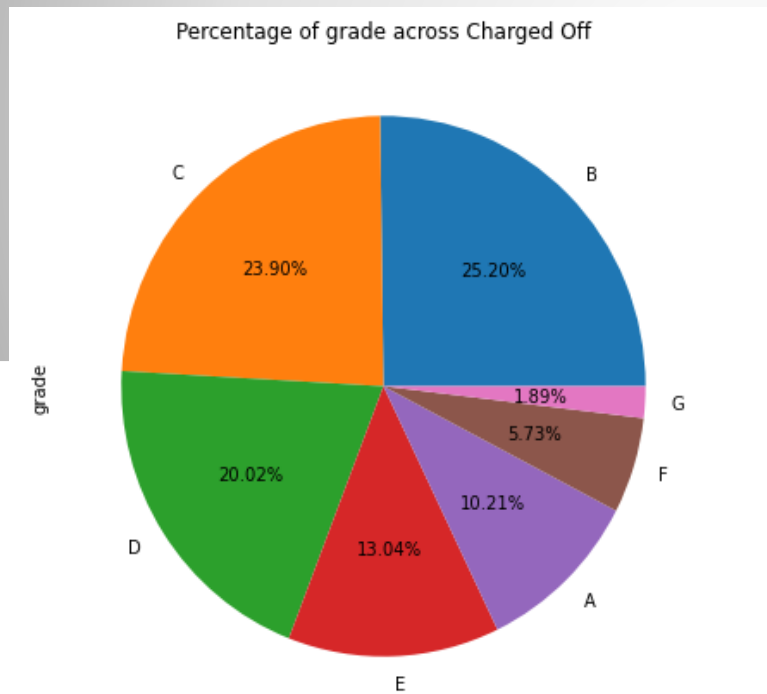
36 Months term customers tend to default more.

Impact of Loan grade and sub grade



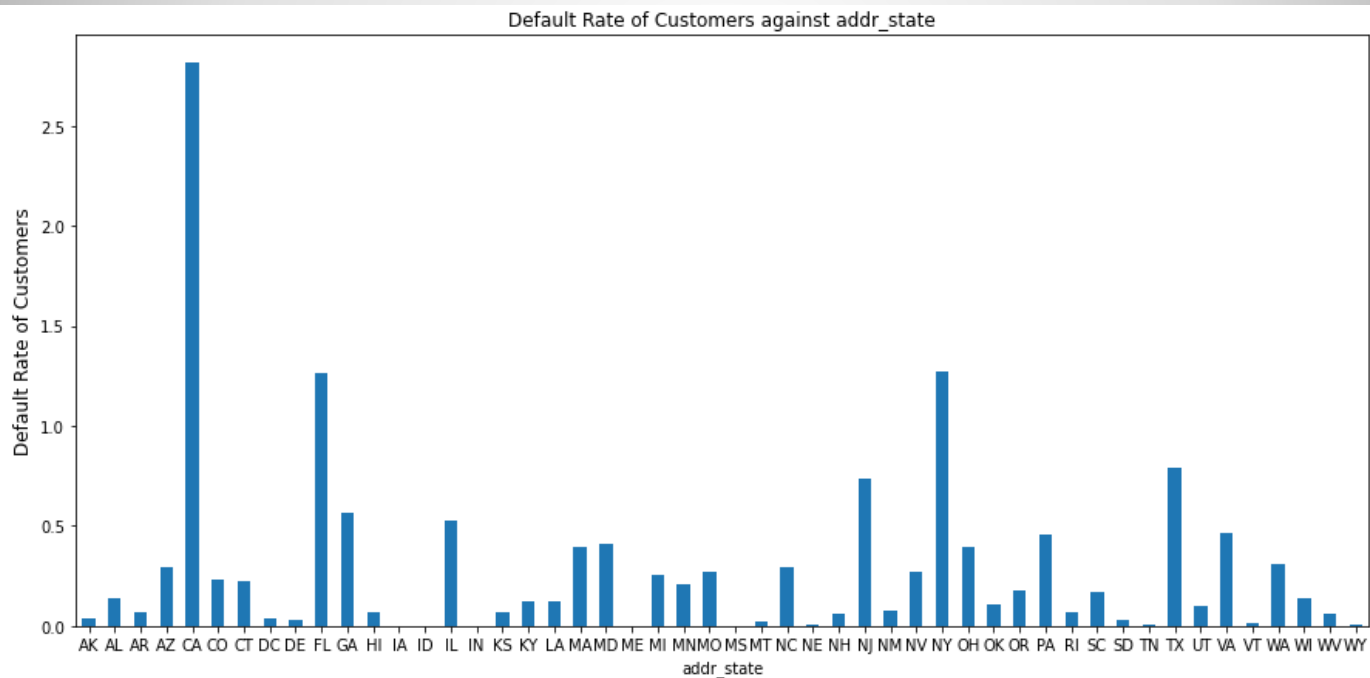
Most of the loans for defaulters are with a grade of B ,C and D.

Impact of loan grade on Default Rate



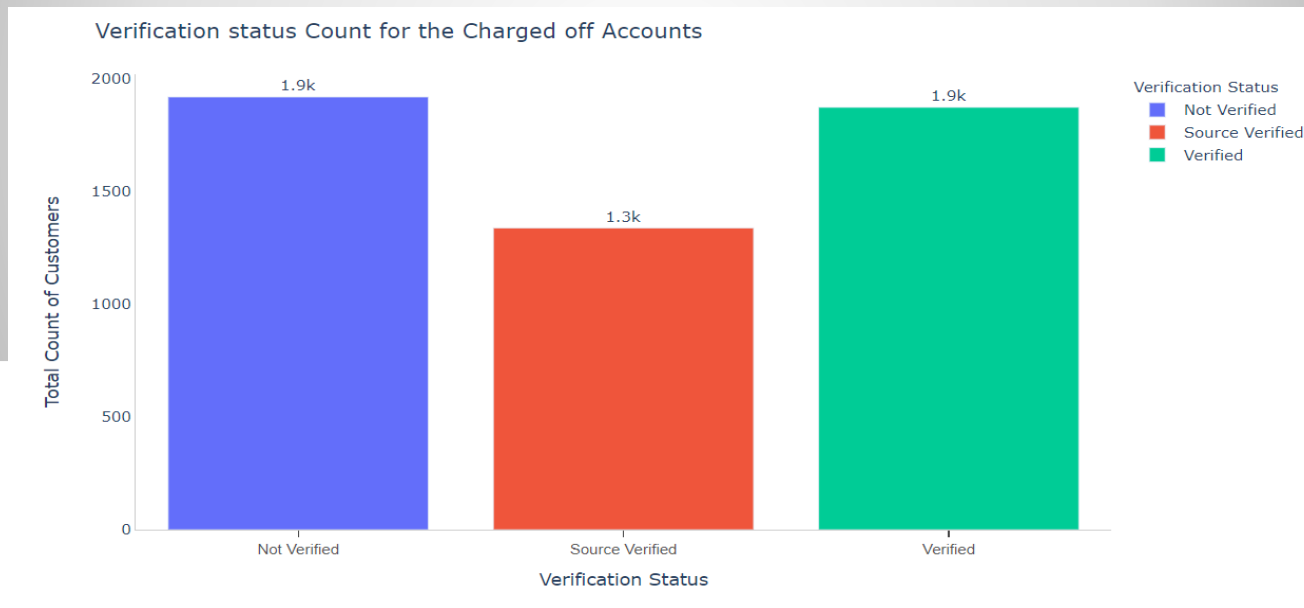
Charged off loans constituted higher percentage of the lower grades as compared to fully paid loans.

Impact of Address State on Default Rate



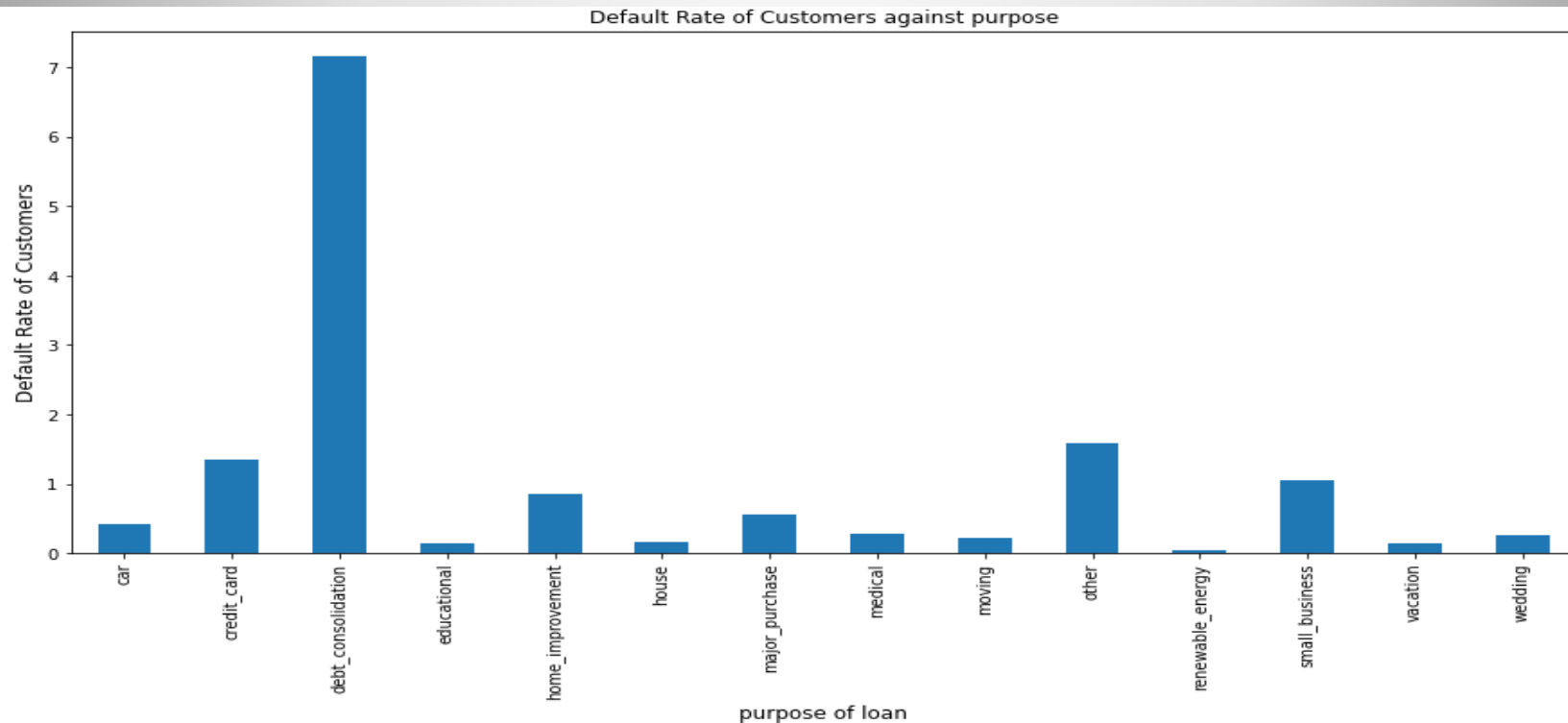
It can be observed that maximum Default Rate have customers from CA state.

Default Customer Count for Different Verification Status



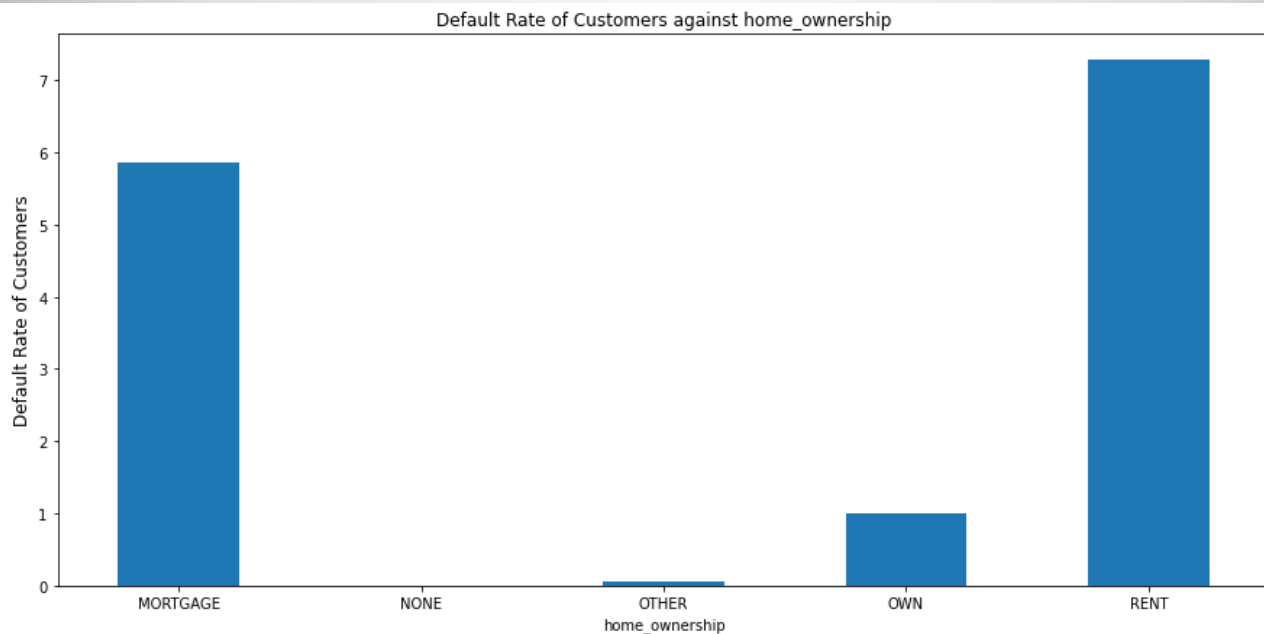
It can be observed that customers for whom income is source verified, tend to default less.

Impact of Purpose on Default Rate



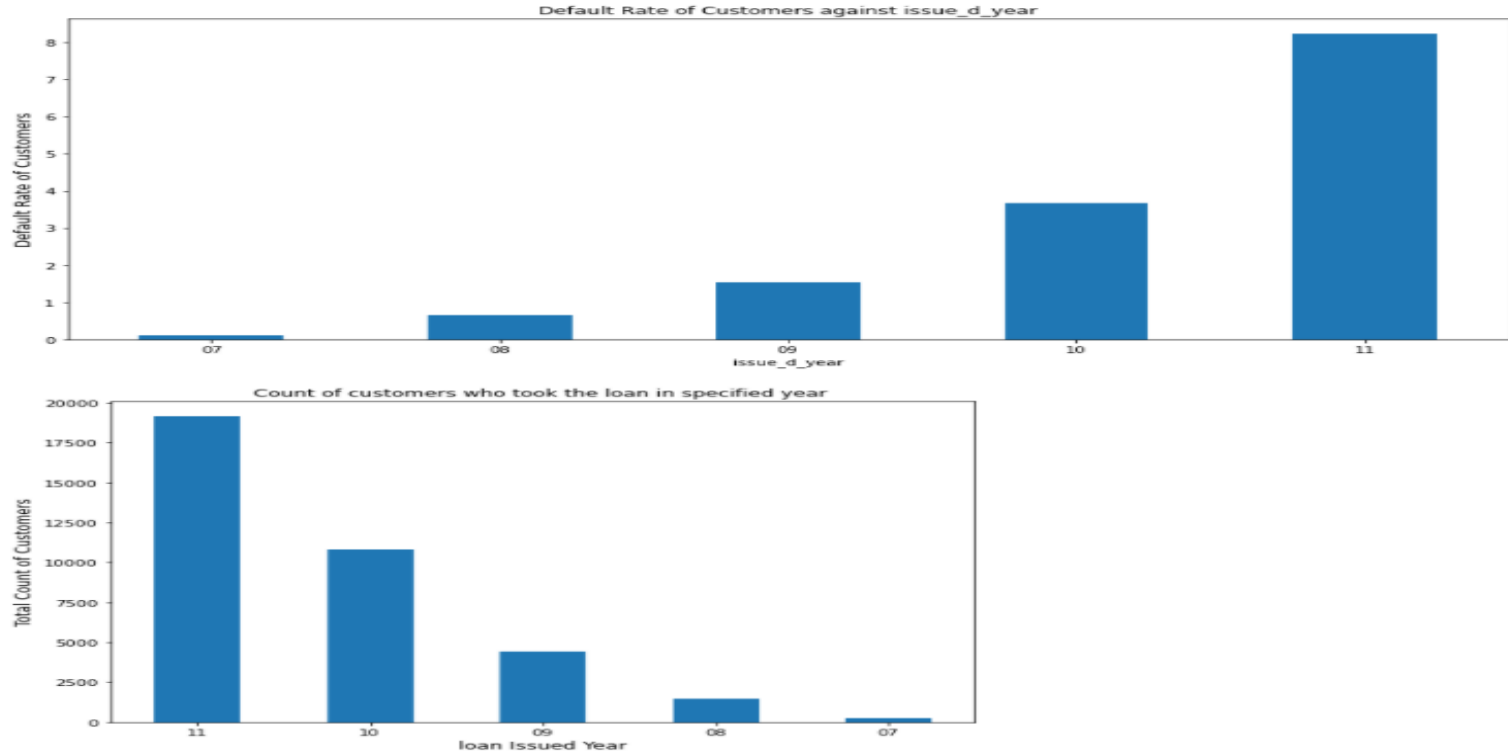
Charged off loans are maximum for the purpose of debt_consolidation.

Impact of Home Ownership on Default Rate



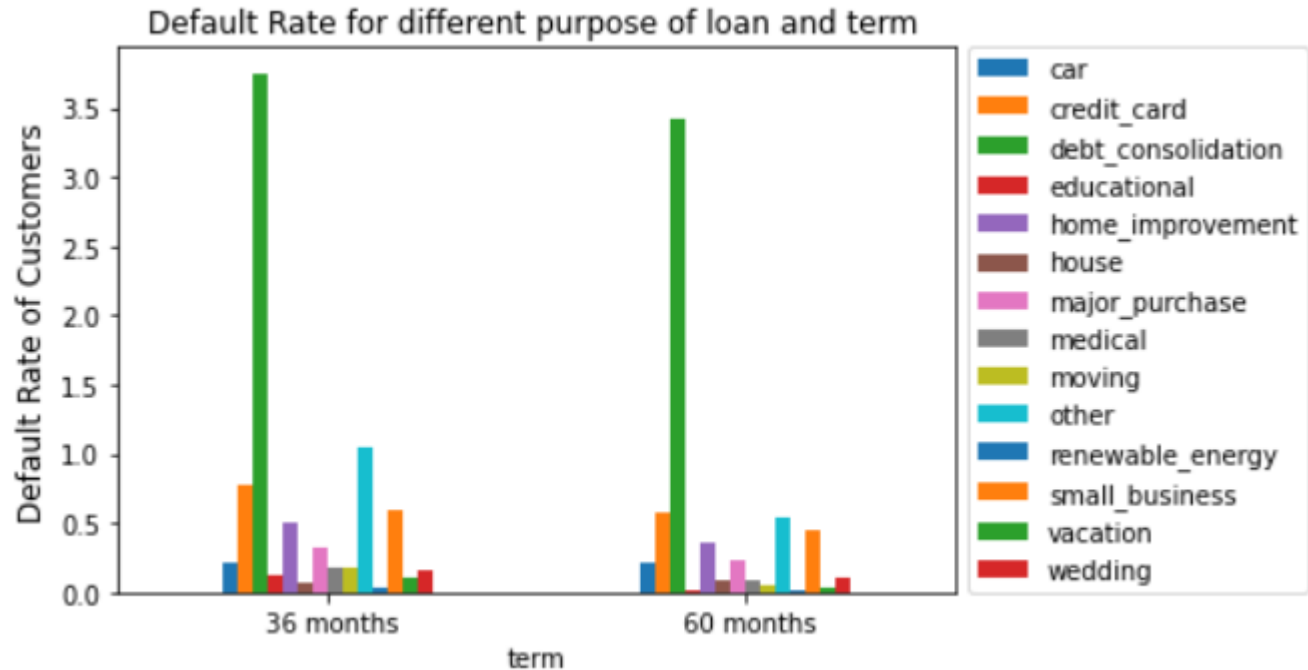
If home ownership of the customer is Rent, Customer tends to default more.

Impact of issued Date on Default Rate



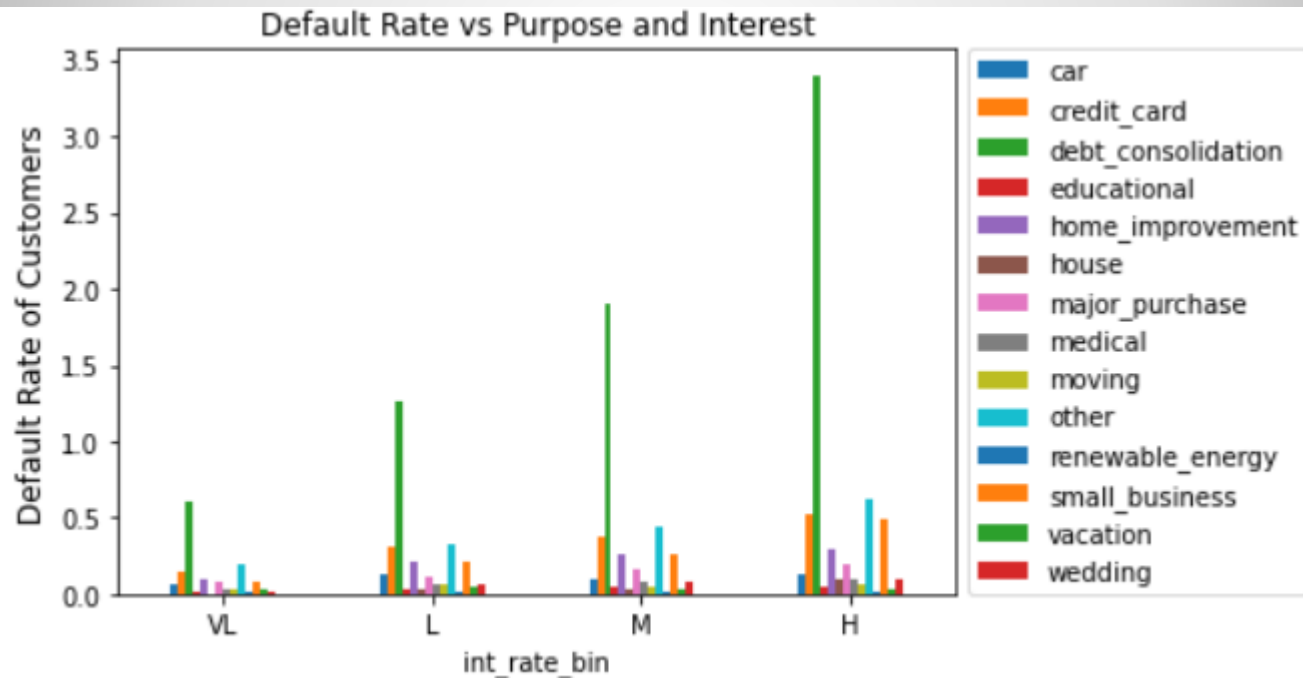
It can be observed that with each passing year, default Rate is increasing.
This is because more customers are opting for loans.

Impact of Default Rate on basis of term and purpose of loan



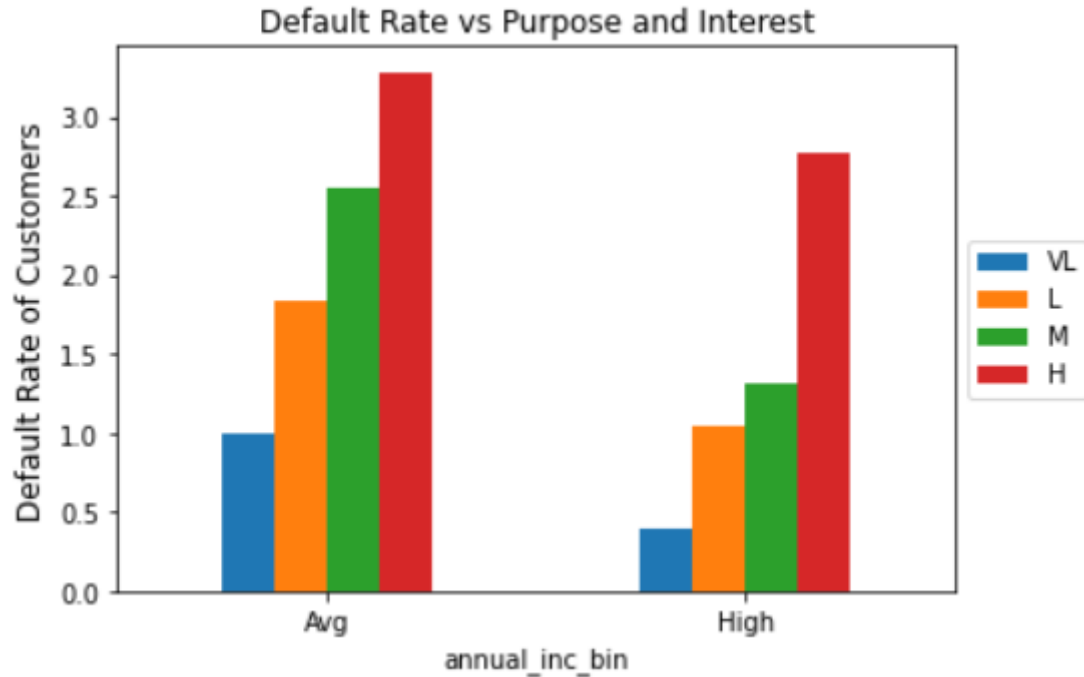
It can be observed that Default Rate is maximum for 36 Months term when the loan purpose is debt_consolidation.

Impact of Default Rate on basis of Purpose and Interest



As Interest Rate increases, Default rate is increasing and maximum default rate is for the loan purpose debt_consolidation

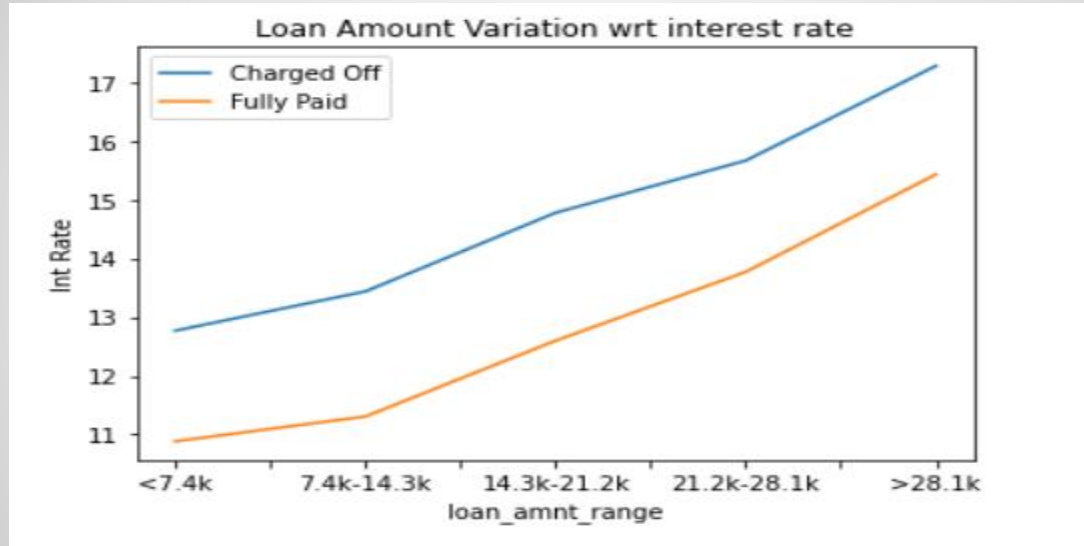
Default Rate for Different Annual Income and Interest Rate



It can be observed that for lower incomes and high interest rates , person tends to default more.

Loan amount Variation wrt Interest Rate

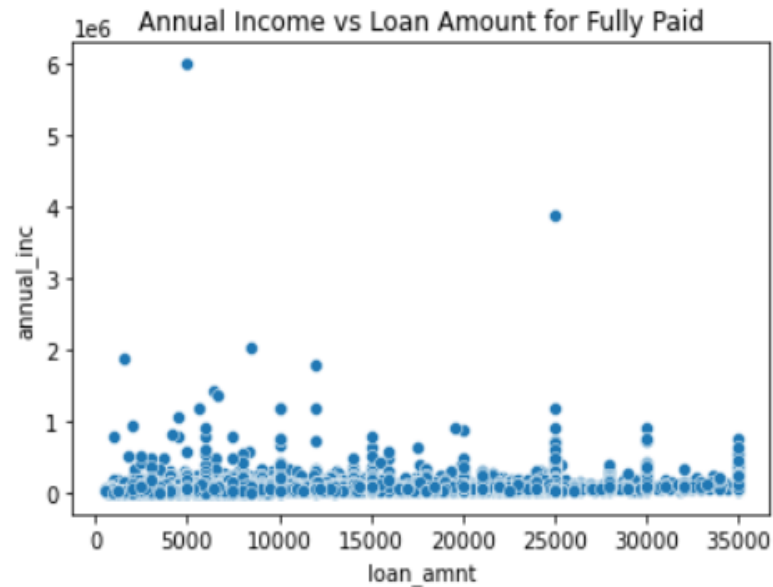
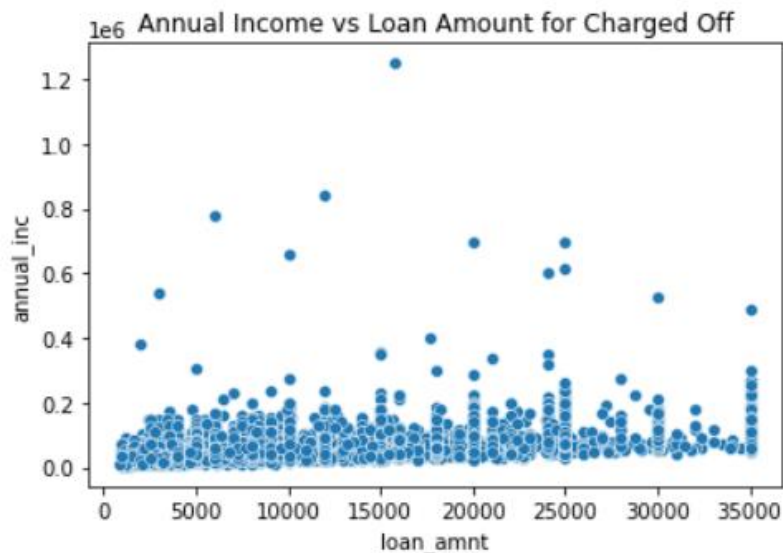
Loan amount Range wrt Interest Rate for charged off and fully paid loan



It is seen that charged_off customers were charged higher interest rates. As can be seen above, if the interest rate is more than 12.5 % for lower loan amount between 7-14k and more than 14.5% for higher loan amount between 14-28k then the person is likely to default.

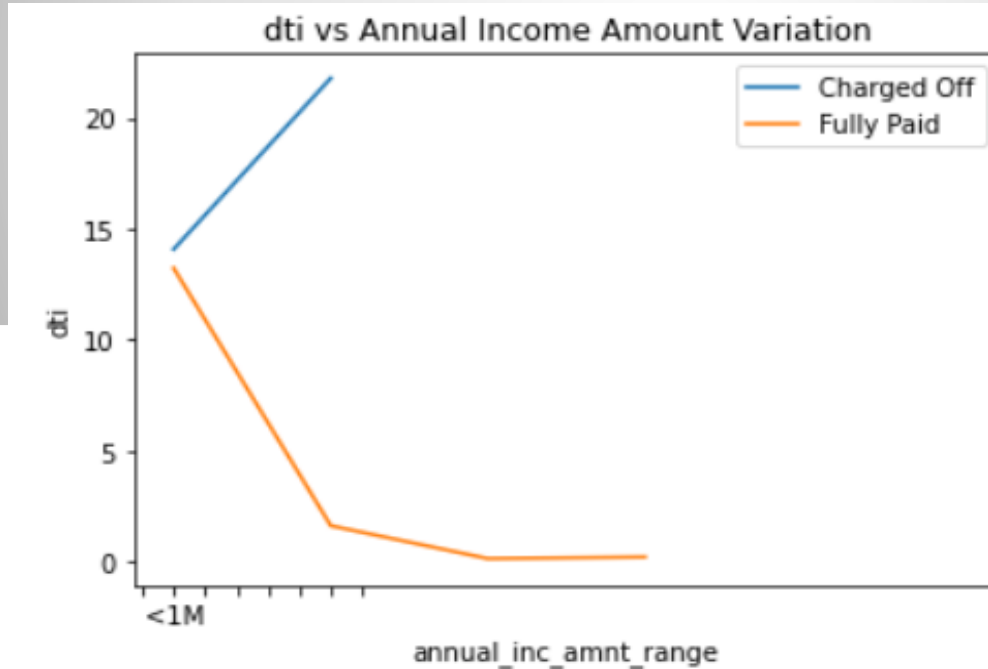
Annual Income vs Loan Amount

Annual Income vs Loan Amount for Charged off and Fully paid



Annual income range is 0- 2 lacs for Defaulters as compared to 0- 10 lacs for Non defaulters for the requested loan amount.

Dti vs Annual Income Amount



For DTI
>14%, there
are high
chances of
customer
default as the
annual income
increases

Conclusions

1. If the interest rate is more than 12.5 % for lower loan amount between 7-14k and more than 14.5% for higher loan amount between 14-28k then the person is likely to default.
2. For DTI >14%, there are high chances of customer default as the annual income increases.
3. Borrowers with less annual income tend to default more than the borrowers with high annual income.
4. Borrowers with increasing emp length tend to default less. However, for Emp length > 10 years, there is max default Rate which can be because of 2 reasons. First, Many of the customers having emp length > 10 can be Retired, implying less or no income. Secondly, There is a huge data for emp length > 10 category than for other categories, hence more default Rate.
5. As the term increases from 36 Months to 60 Months, customers tend to default less.
6. Default Rate is quite high for 36 Months term when the loan purpose is debt_consolidation.
7. Default Rate is more for Borrowers from CA state.
8. It can be observed that most loans for defaulters are with a grade of B ,C and D.
9. Charged off loans constituted higher percentage of the lower grades as compared to fully paid loans.
10. If home ownership of the customer is Rent, default Rate is more.
11. The customers for whom income is not source verified, tend to default more.
12. For Charged Off loan borrowers, annual income is less than the fully paid customers, which could be the possible reason to default.
13. Defaulted customers have asked for more loan amount.