

Lending Club Case Study

By Ramakant Shreshti



Background – Lending Club Case Study

Background

Lending club is the largest peer-to-peer marketplace connecting borrowers with lenders. Borrowers apply through an online platform where they are assigned on an internal score. Lenders decide 1) whether to lend and 2) the terms of the loan such as interest rate, monthly installment, tenure etc..

Business Objective

The business objective is to take a decision whenever they receive a loan application whether to reject or approve based on certain variables

Data understanding

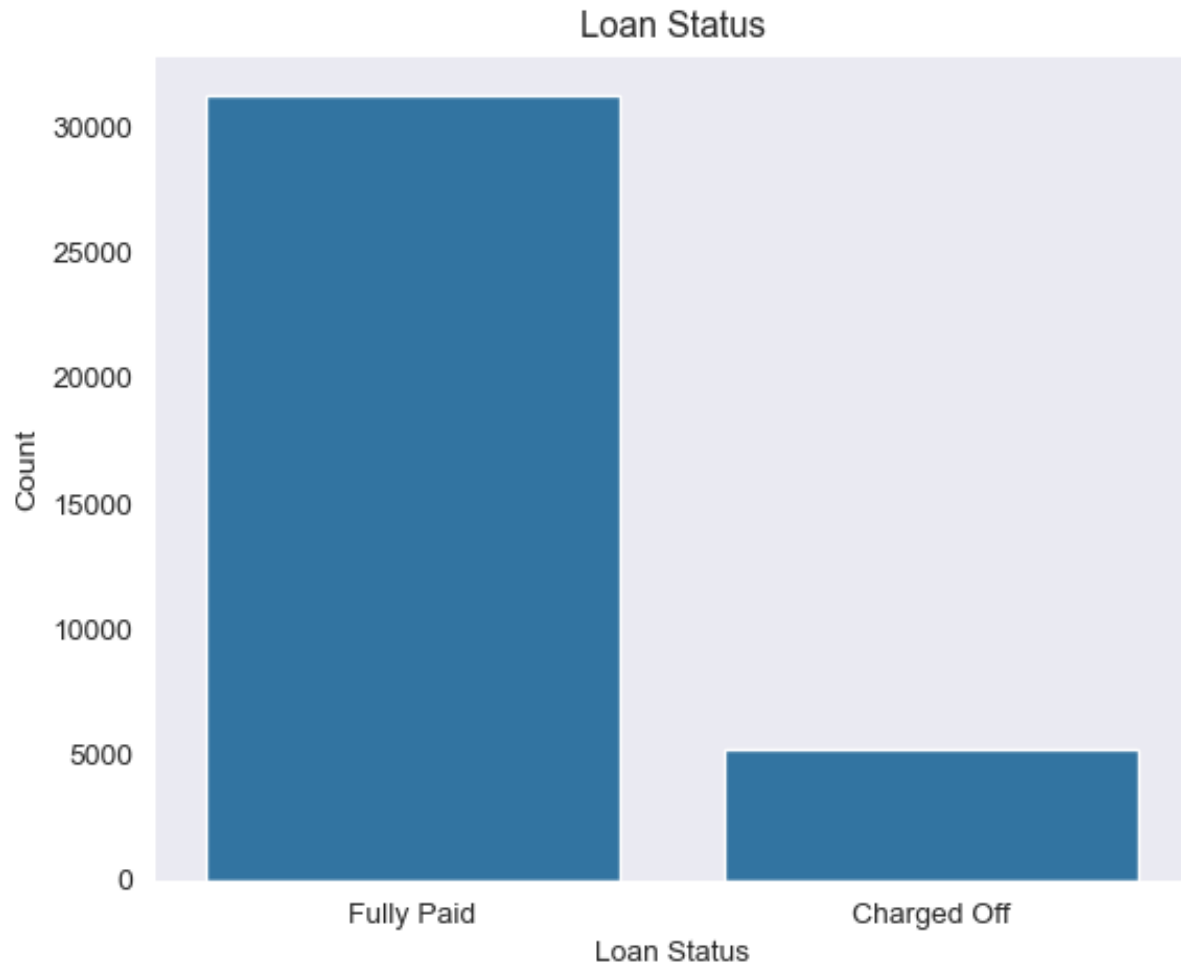
The data given contains information about past loan applicants and whether they 'defaulted' or not. Data has details regarding approved loan not the rejected ones. It has 3 status of loan which is Fully Paid, Current and Charged-Off.

Along with that it also has data related demographic like which state the applicant belongs to, what is his annual income and so on..

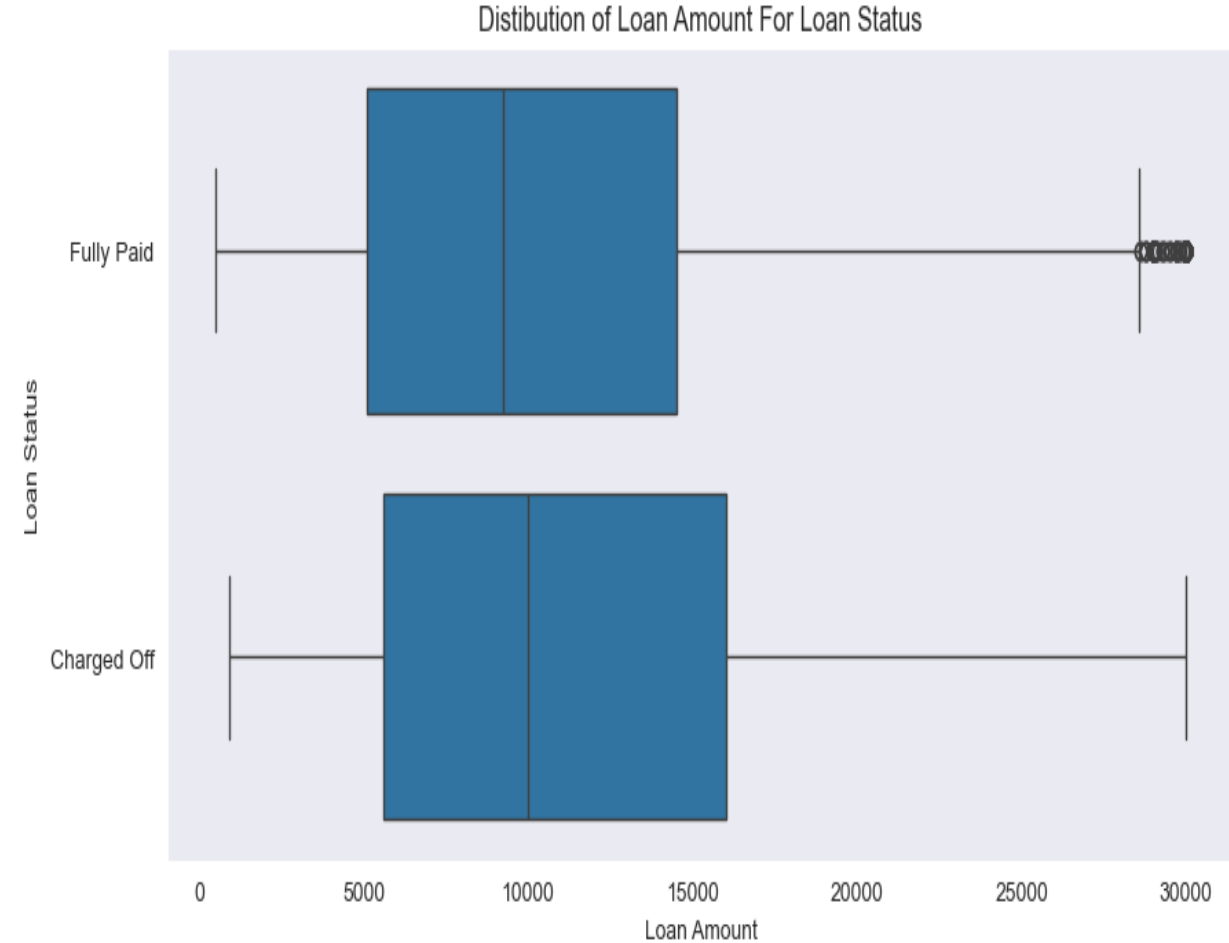
Data Clean-up and preparation:

Importing the data -> Removing null columns -> Removing duplicate rows -> Removing irrelevant columns -> Removing / fixing null values -> correcting data types and deriving new columns -> Removing outliers

Loan Status and Amount



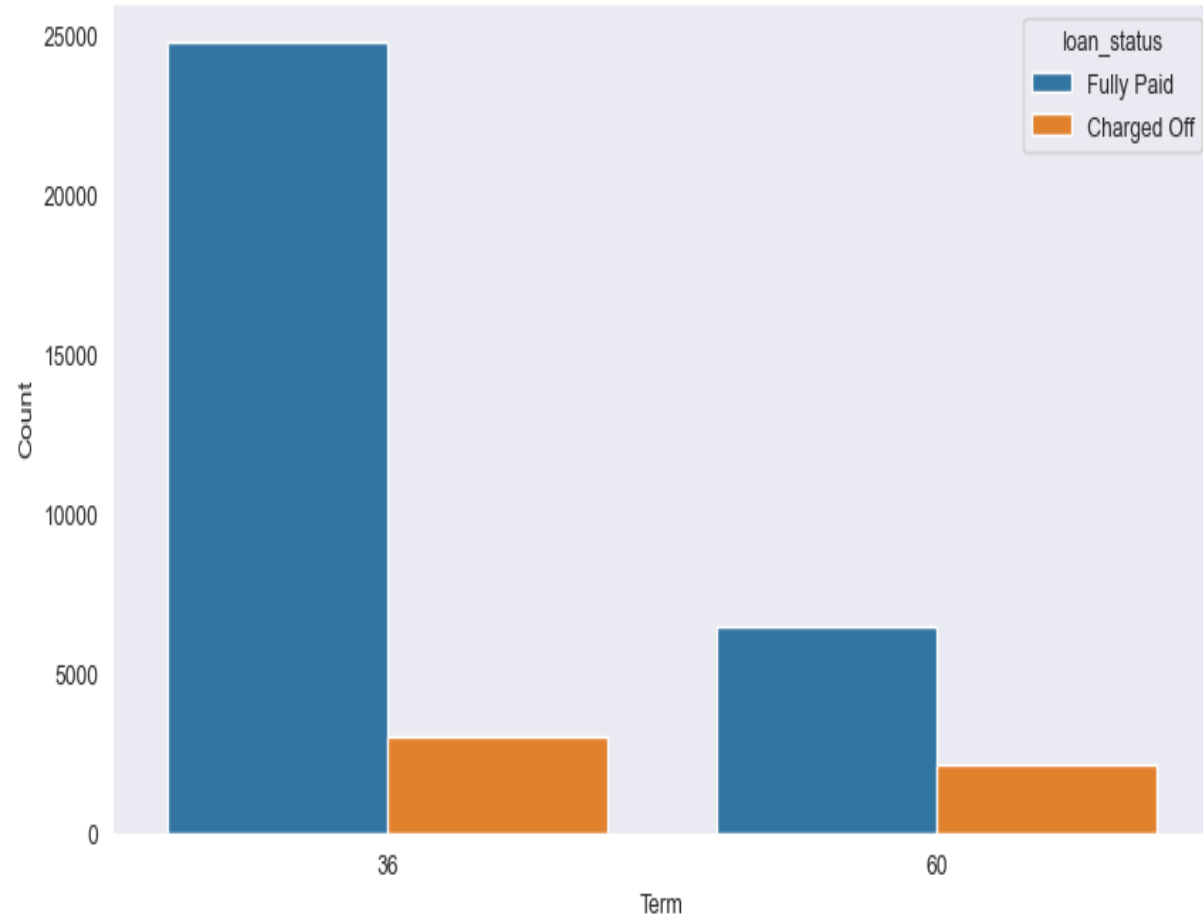
Loan Status: The number of charged off loan is much smaller(14.2%) compared to total count.



Loan Amount: It varies from 500 to 35000 with a median of 10605. Loan amount is majorly small and very few clients have taken large loans and larger it goes we have higher chance of defaulting

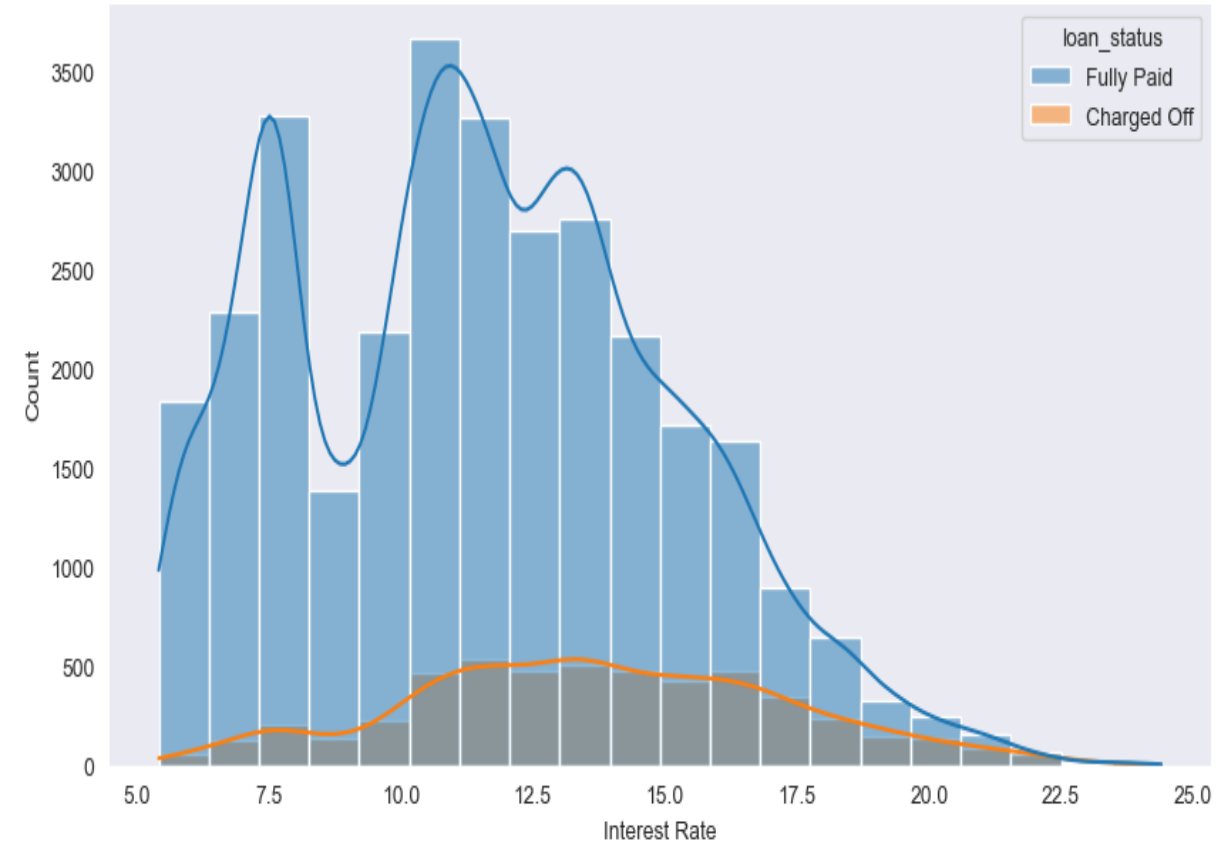
Term and Interest Rate

Distribution of Term For Loan Status



Loan Term: The Loans taken for 36 month term are much more than 60 months and have lower chance of defaulting.

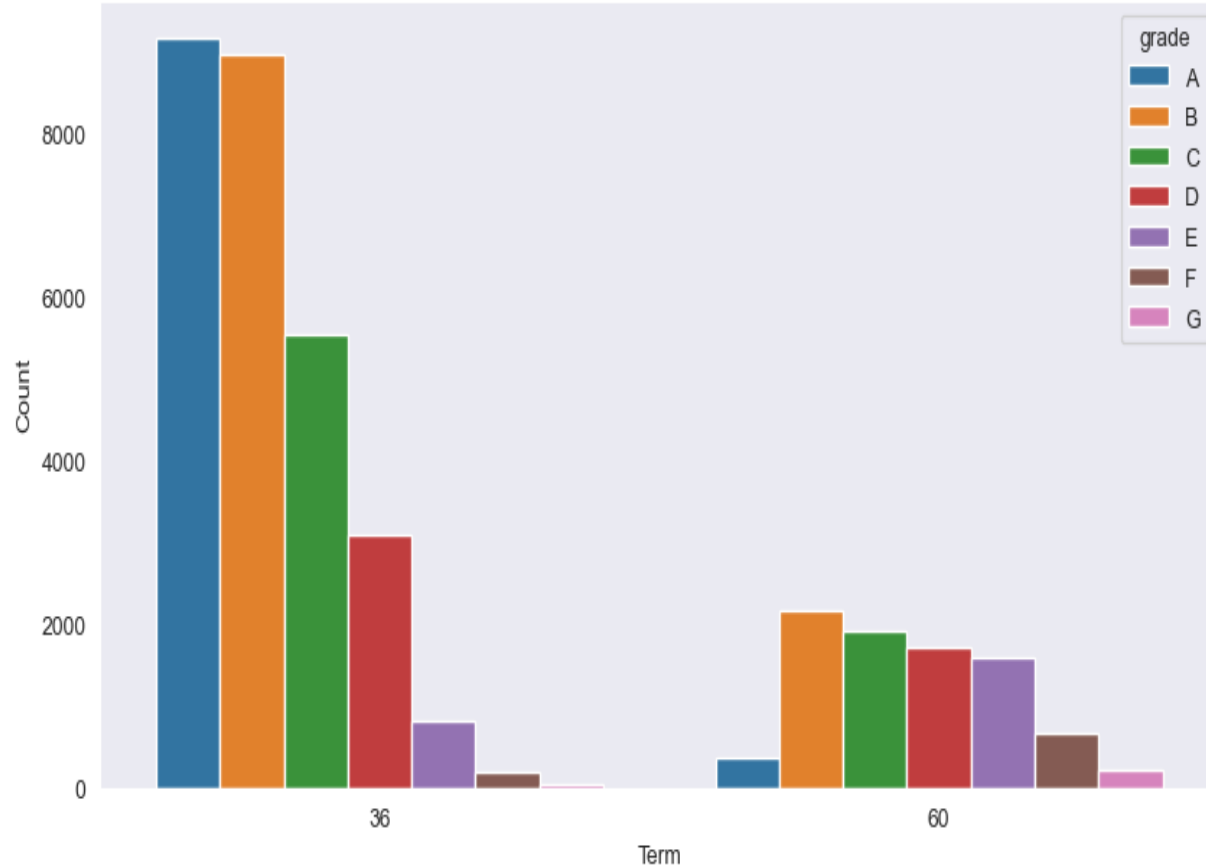
Distribution of Interest Rate For Loan Status



Interest Rate: Majority of the loan taken fall under the 5 – 15 % of interest bracket and from there onwards the count of loan taken starts decreasing as the interest rate increases. Also there is chance of defaulting as the interest rate increases.

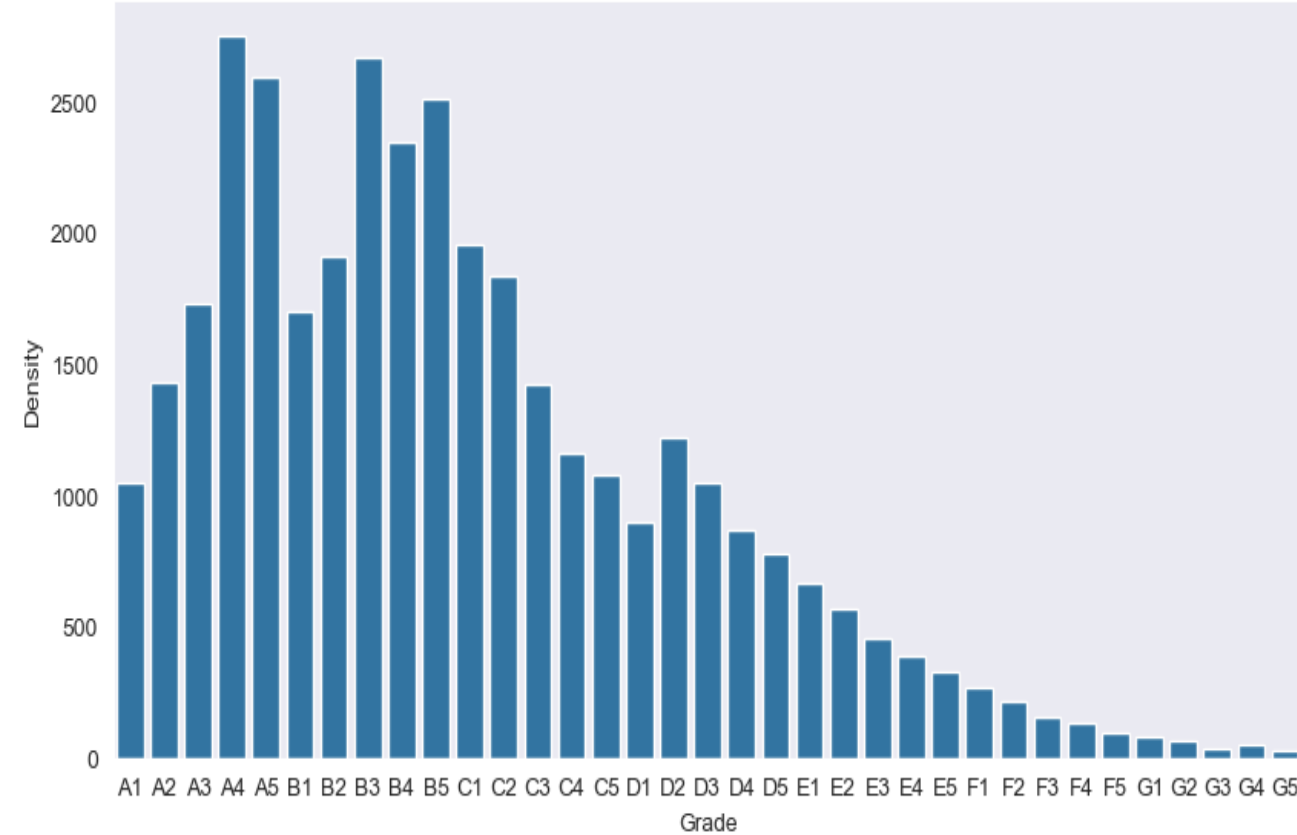
Grade and Sub-Grade

Distribution of Term For Grade



Grade: Large number of loan approved are of higher grade and for 36 month term, as they are of low risk thus low chance of defaulting. 60 month term loans have larger number of lower grade loans with high risk.

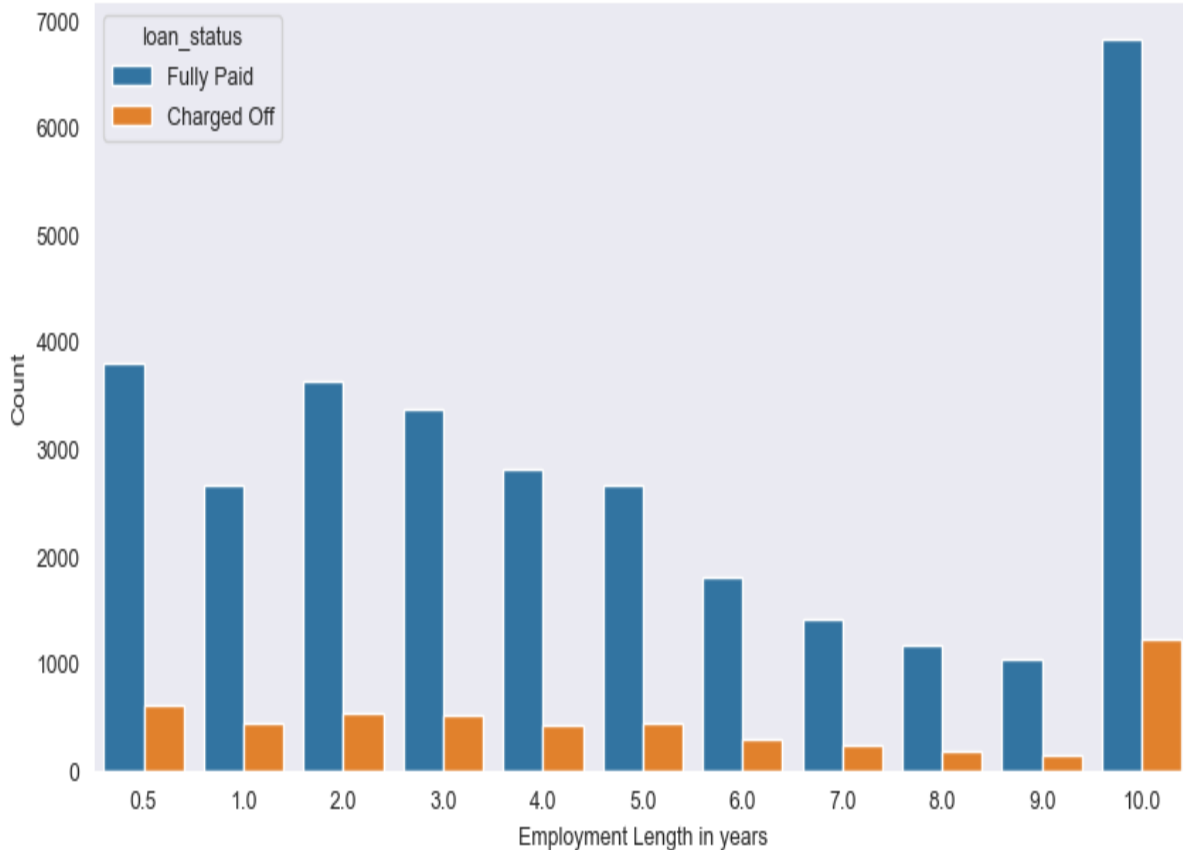
Distribution of Secondary Grade



Sub Grade: This provides more insight that the loans within grade are more skewed towards lowered sub grades.

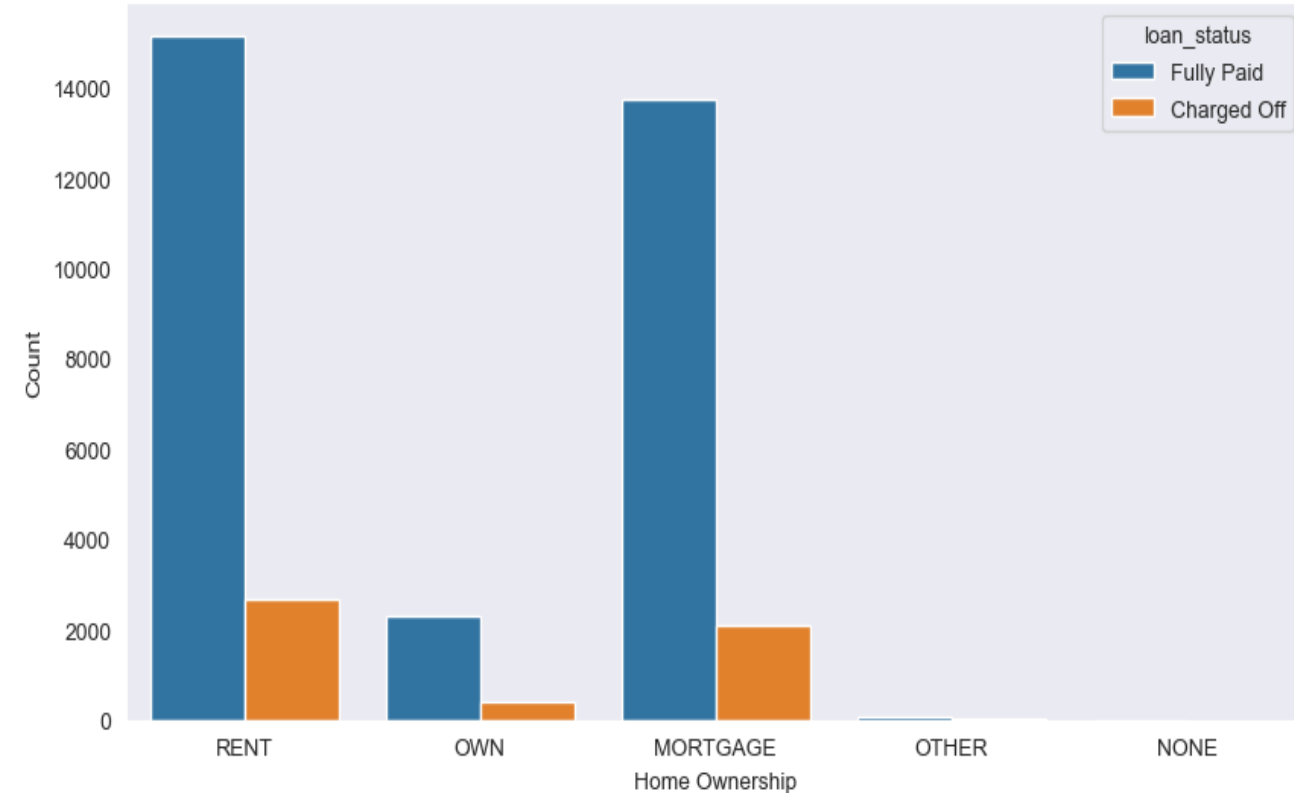
Employment Length & Homeownership

Distribution of Employment length for loan status



Employment Length: The Employees with 10+ years of experience are likely to default and have higher chance of fully paying the loan as well.

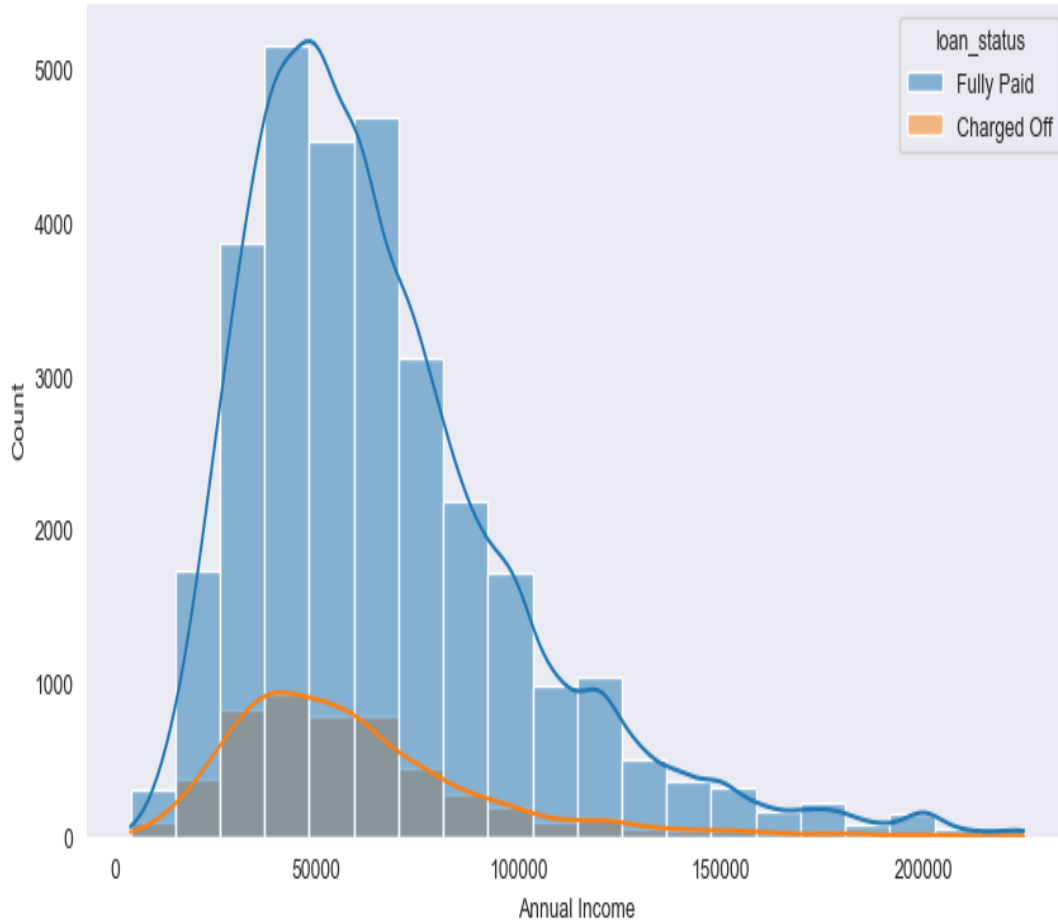
Distribution of Home Ownership For Loan Status



Home Ownership: The Defaulted loan are lower for the borrowers who own their property as compared to on mortgage or rent.

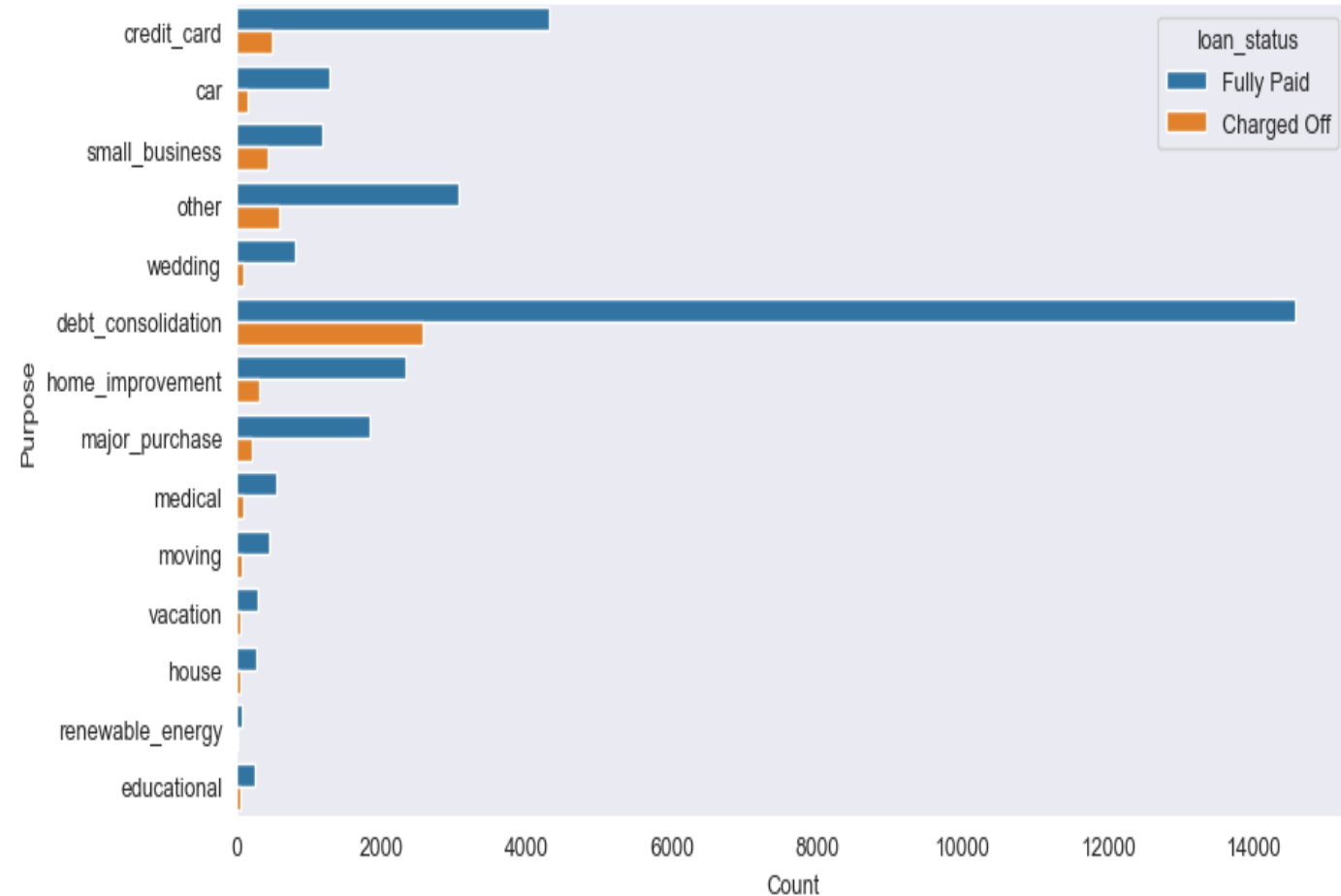
Annual Income & Purpose

Distribution of Annual Income For Loan Status



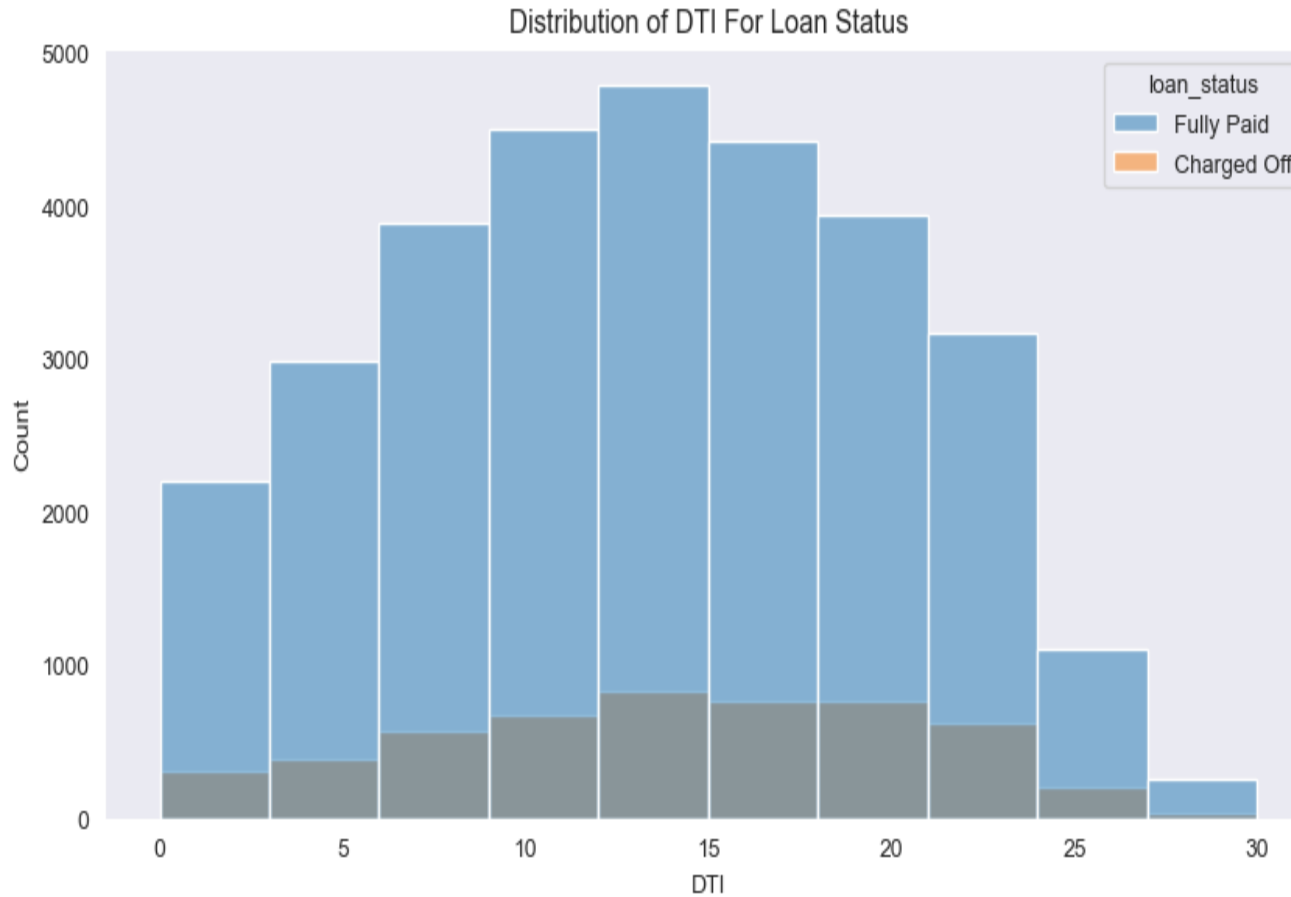
Annual Income : Most of the clients have low annual income compared to the rest and income lower than 50k has higher chance of defaulting.

Loan Taken For Different Purposes

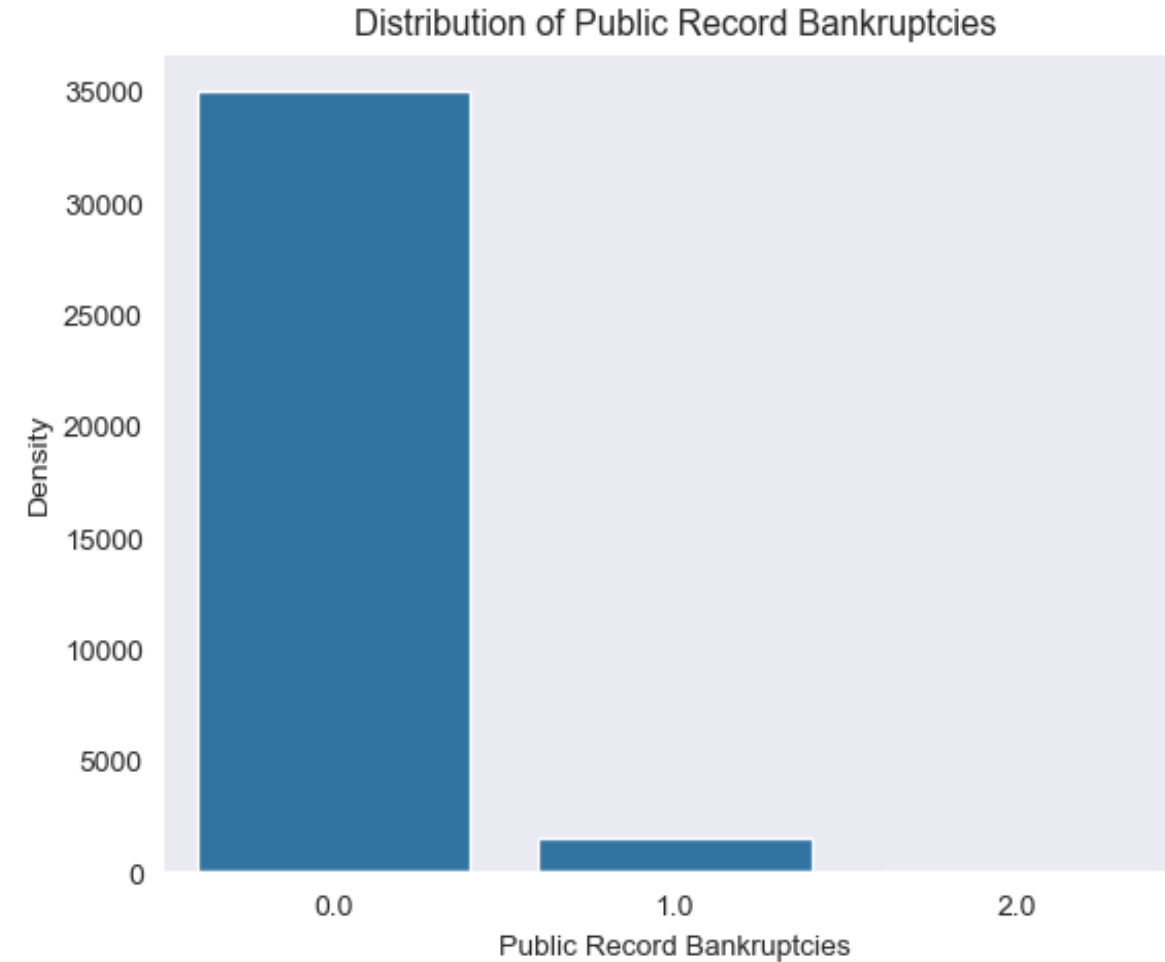


Purpose: Loans are taken mostly for debt consolidation followed by credit card payment. Whereas the debt consolidation has highest fully paid loan but also has highest defaulted loans as well

DTI ratio & Bankruptcy

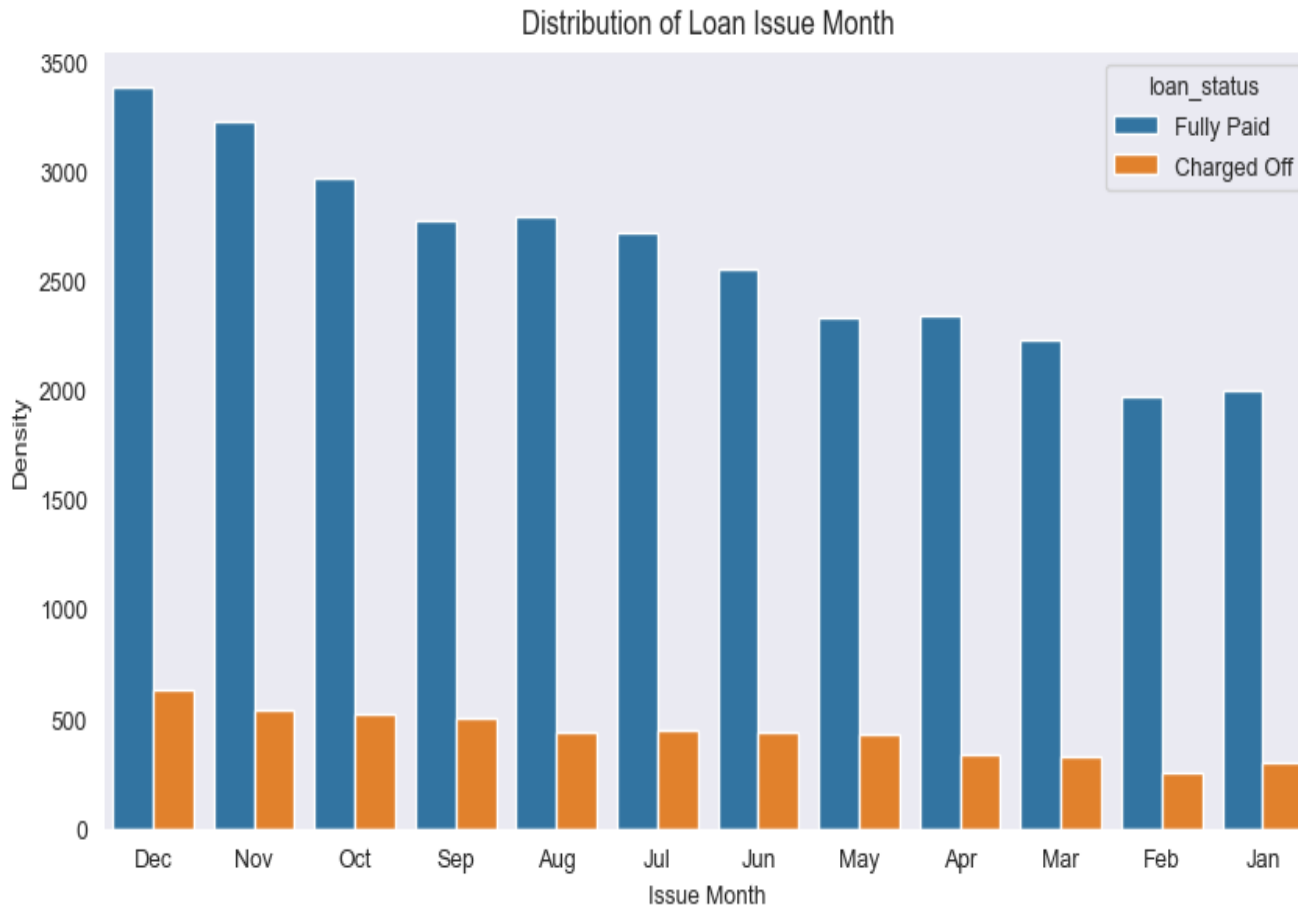


DTI: The large percentage of Clients have a large Debt to Income ratio which shows that lending to such clients can be very risky

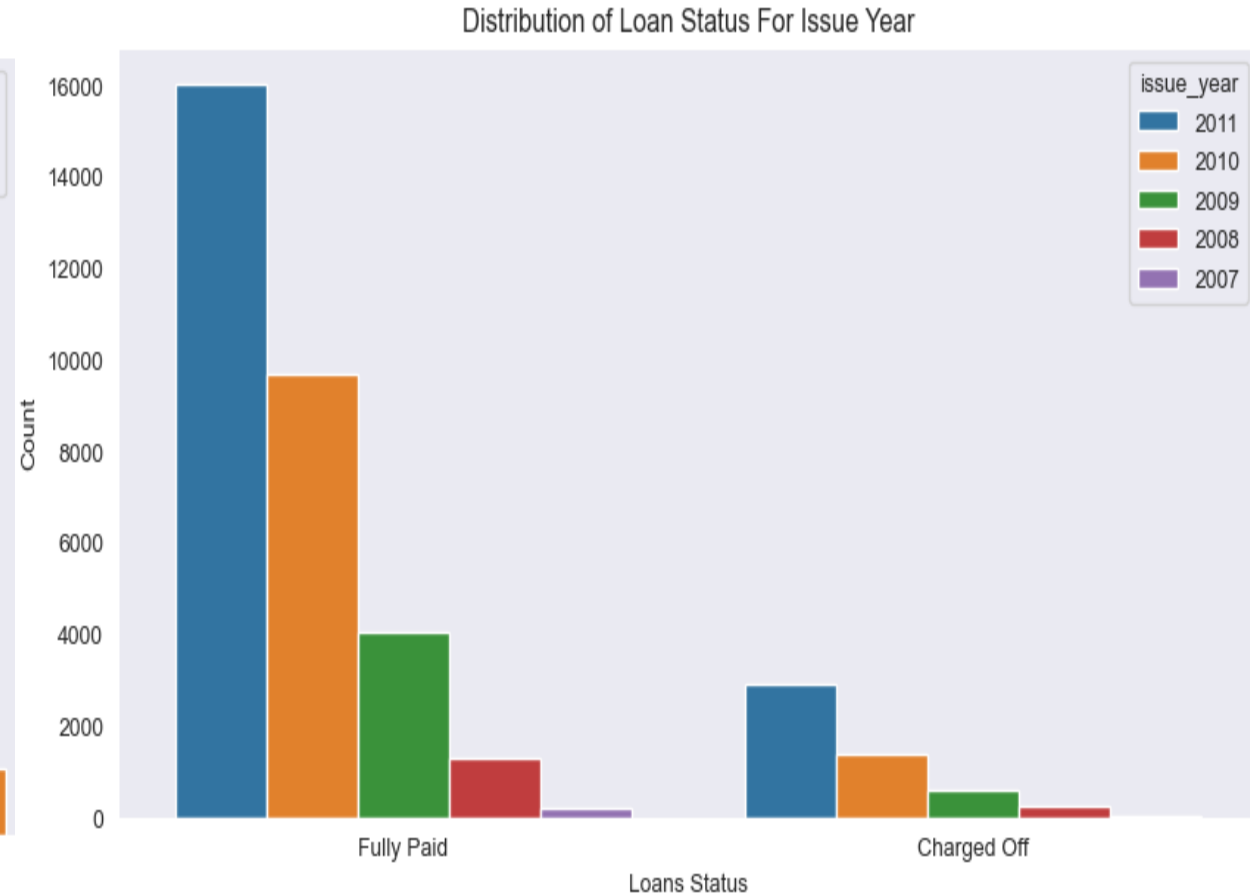


Public Recorded Bankruptcy: Majority of clients have no record of declaring bankruptcy.

Loan Trend over years

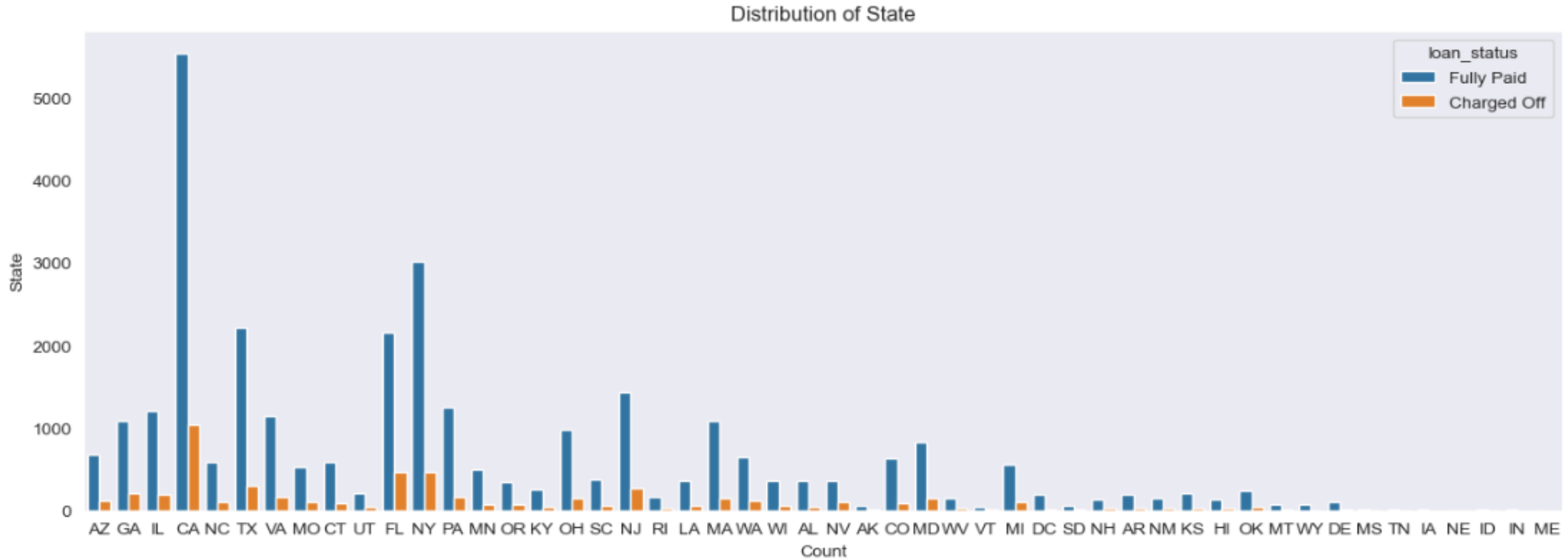


There is a gradual increase in loan taken throughout the year. Last quarter of the year shows majority of the loan taken. And also more defaults are seen in last quarter of the year.

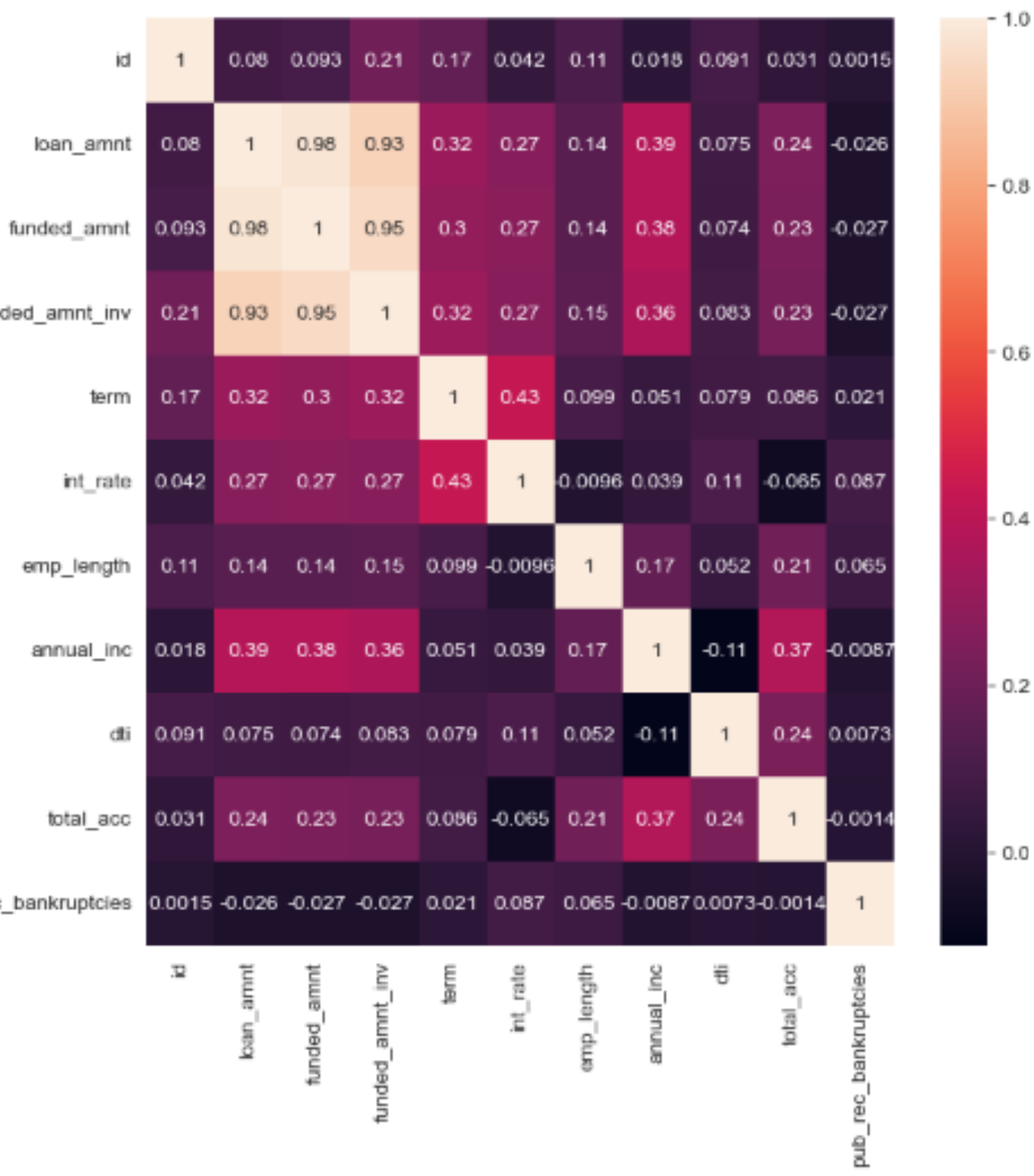


With each passing year loan taken are increasing exponentially which indicate large increase in DTI ratio and decrease in defaulting rate.

Location Based



Majority of the borrowers are from the large urban cities like California, New York, Texas, Florida etc. California has large number of customer who have defaulted.



Heat Map for numerical columns

- Observation is that Loan amount, investor amount, funding amount are strongly correlated. Annual income with DTI(Debt-to-income ratio) is negatively correlated. Debt income ratio is the percentage of a consumer's monthly gross income that goes toward paying debts. That means when annual income is low DTI is high & vice versa.

Recommendations

Recommendations

Major Driving factor which can be used to predict the chance of defaulting and avoiding Credit Loss:

- DTI
- Grades
- Verification Status
- Annual income
- Pub_rec_bankruptcies

Other considerations for 'defaults' :

1. Borrowers should not be from large urban cities like California, New York, Texas, Florida.
2. Borrowers having annual income less than 70k are risky.
3. Borrowers having Public recorded bankruptcy should be avoided.
4. Borrowers having least grades are risky
5. Borrowers with High DTI are risky
6. Borrowers with working experience more than 10 years are risky.