



Lending Club Case Study

Submitted by: Prasanna Chandorkar

Contents

- Objective
- Business Understanding
- EDA Analysis
- Recommendation

The Objective of this case study is to implement the EDA technique on a real-world problem and understand the insights and present in a business-first manner via presentation.

Benefits of the case study:

- Gives an idea about how EDA is used in real-life business problems.
- It also develops a basic understanding of risk analytics in banking and financial services.
- How the data is used to minimize loss of money while lending it to clients.
- It improves our understating of visualization and what charts to use for real-life data.

Problem Statement :

- ✓ Lending Club 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 labeled as 'charged-off' are the 'defaulters'.

Goal

- ✓ How data can be used to minimize the risk of losing money while lending to customers.

Risk

- ✓ If the applicant is likely to repay the loan, then not approving the loan is a loss of business (rejecting loans for non-default).
- ✓ If the applicant is not likely to repay the loan, then approving the loan may lead to financial loss (approving loans for default). The given dataset contains information about past loans and each row represents the loan details of the applicants.

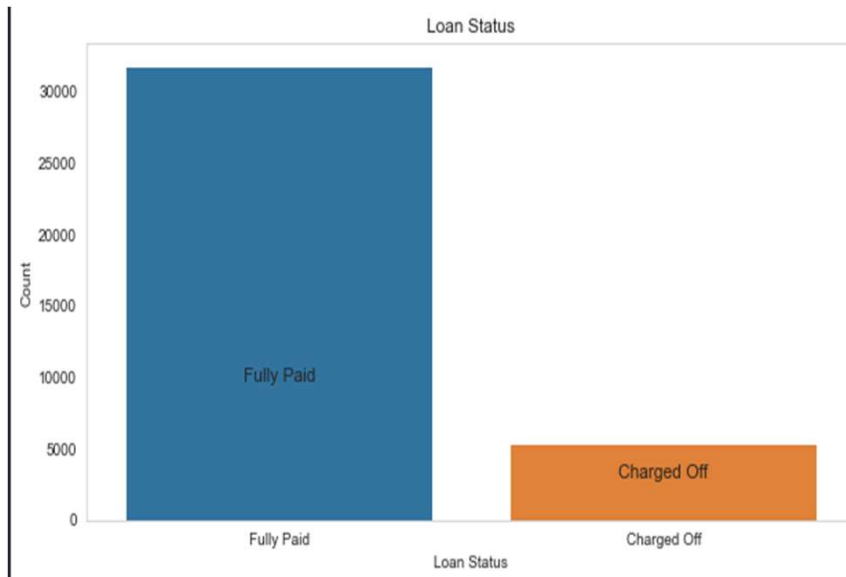
Dataset Details:

The data given below 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.

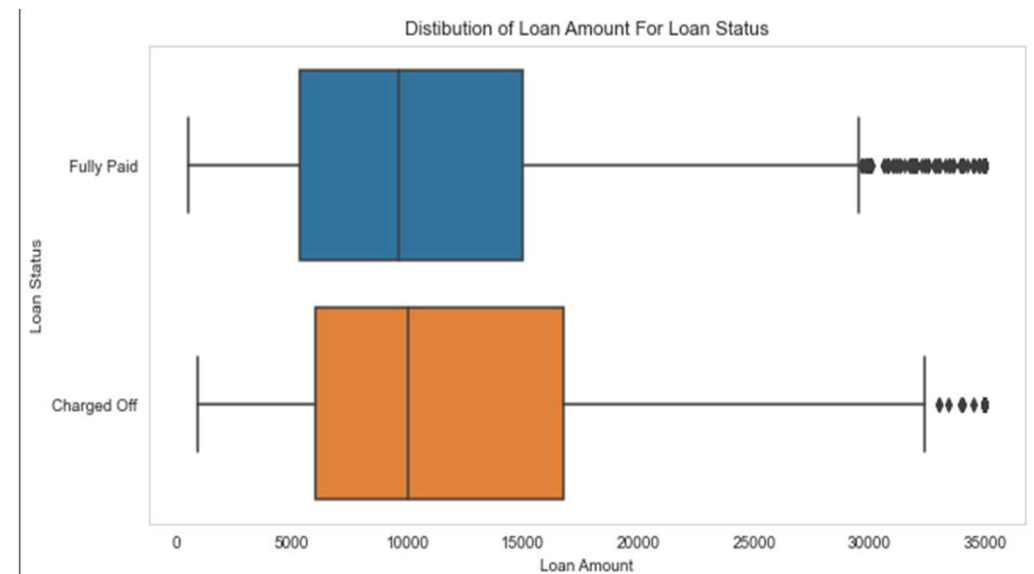
Data Clean-up and preparation process:



Loan Status and Amount

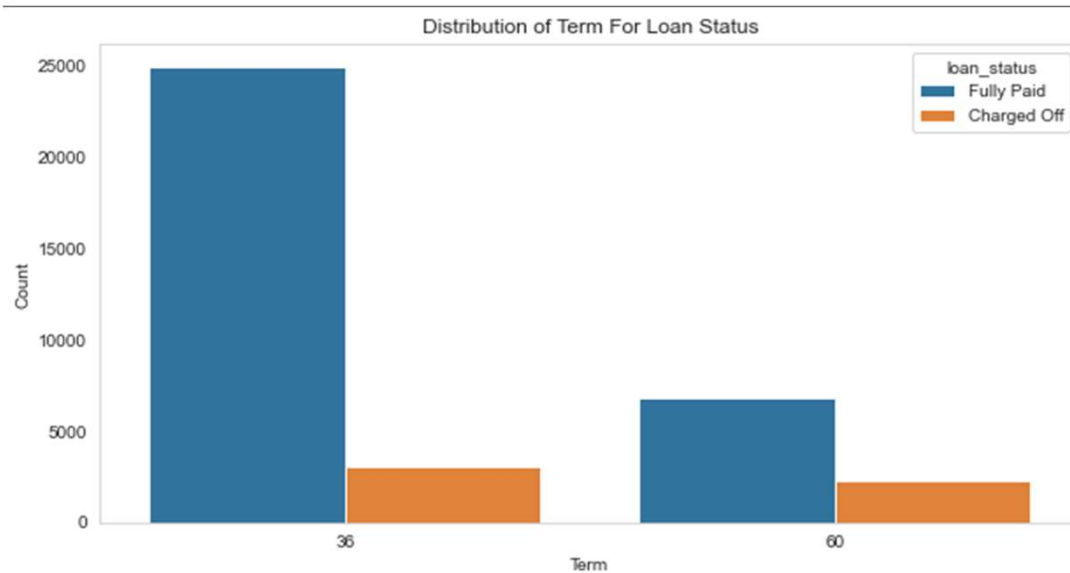


- **Loan Status:** The number of charged off loan is much smaller(14.5%) compared to total count.

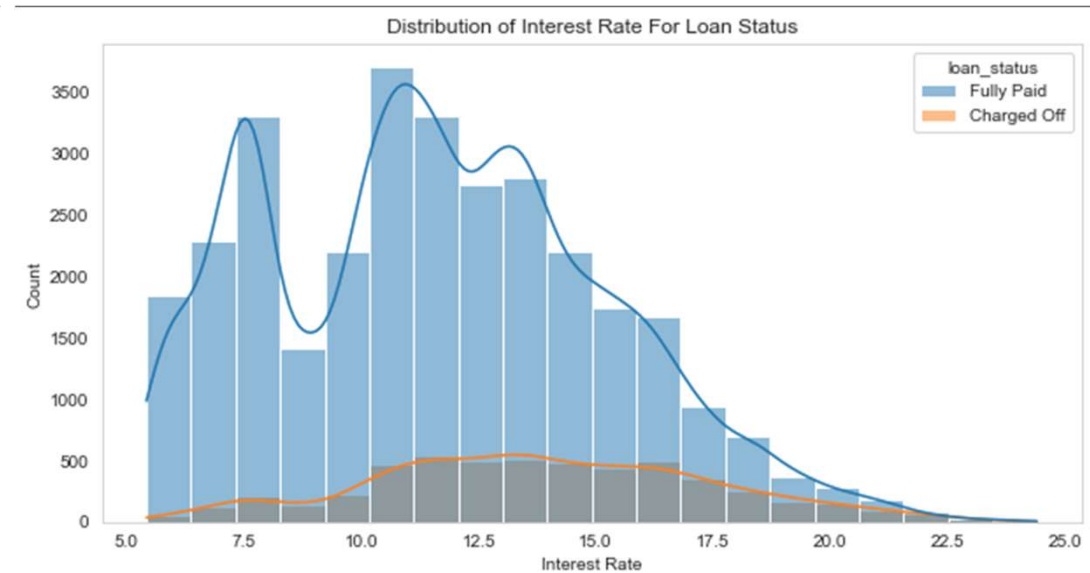


- **Loan Amount:** It varies from 500 to 35000 with a median of 10000. The loan amount is majorly small and very few clients have taken large loans. the larger it goes we have the higher chance of defaulting.

Term and Interest Rate

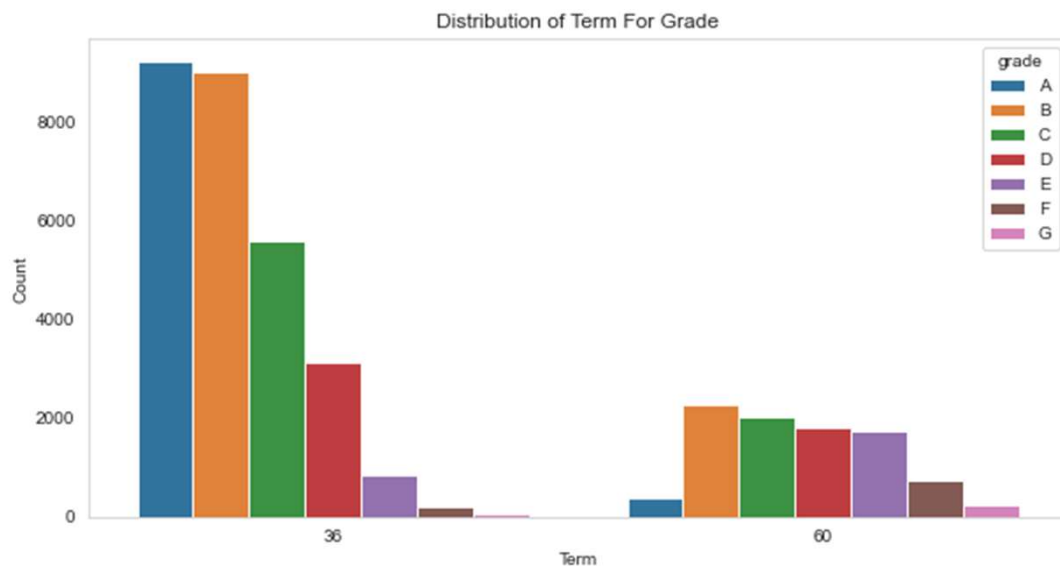


- **Loan Term:** The Loans taken for the 36-month term are much more than 60 months and have a lower chance of defaulting.

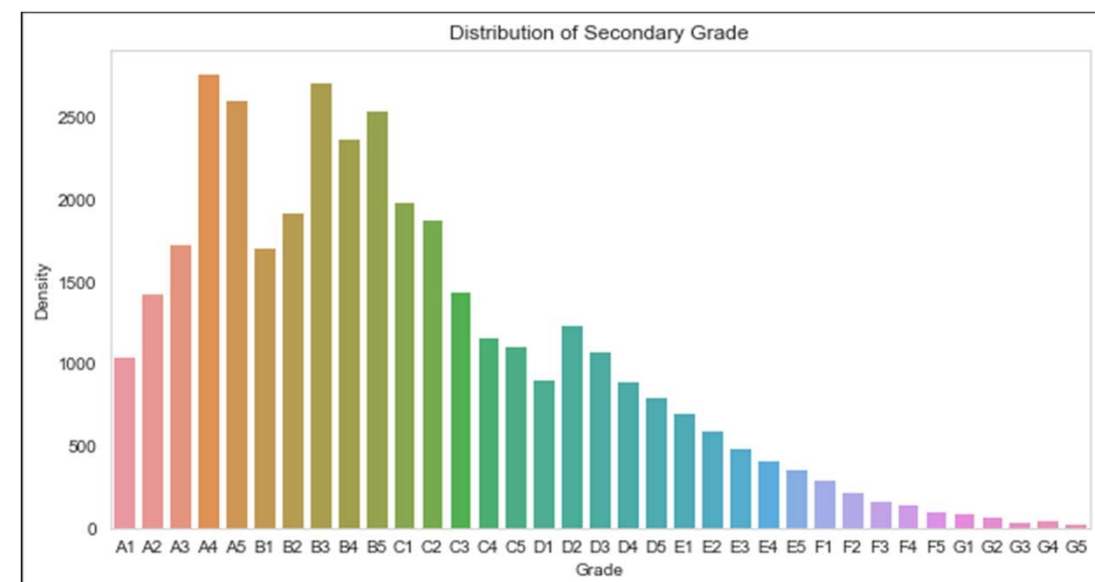


- **Interest Rate:** The count of loans taken varies with interest rate showing a peak around in 5-15 bracket and decreasing slowly where as the chance of defaulting increases with interest rate.

Grade and Sub-Grade

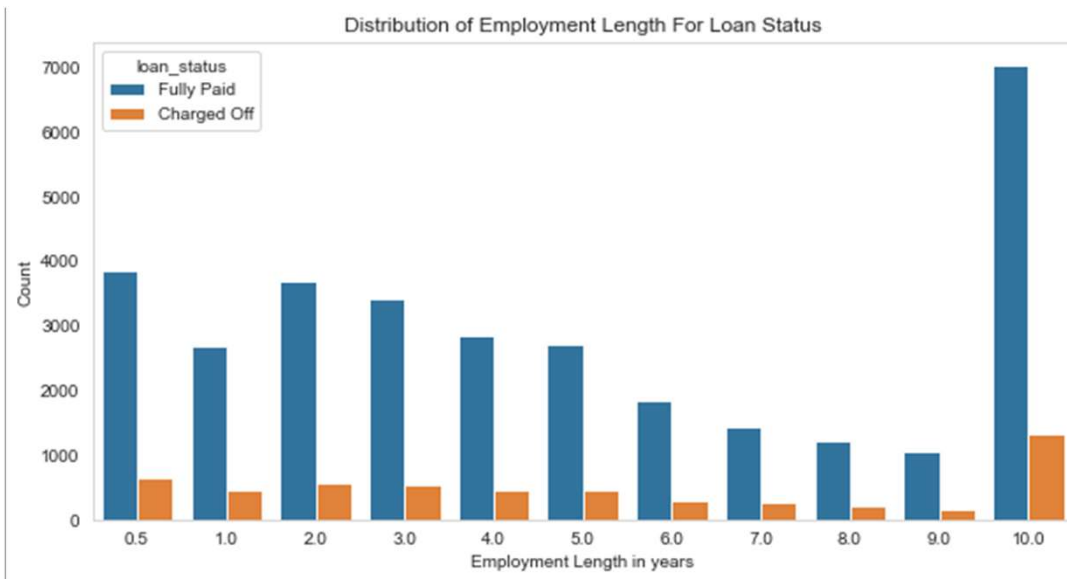


- **Grade:** The loan approved are major of a higher grade as they are of low risk and thus low chance of defaulting. 60-month term loans have a larger number of lower-grade loans with high risk.

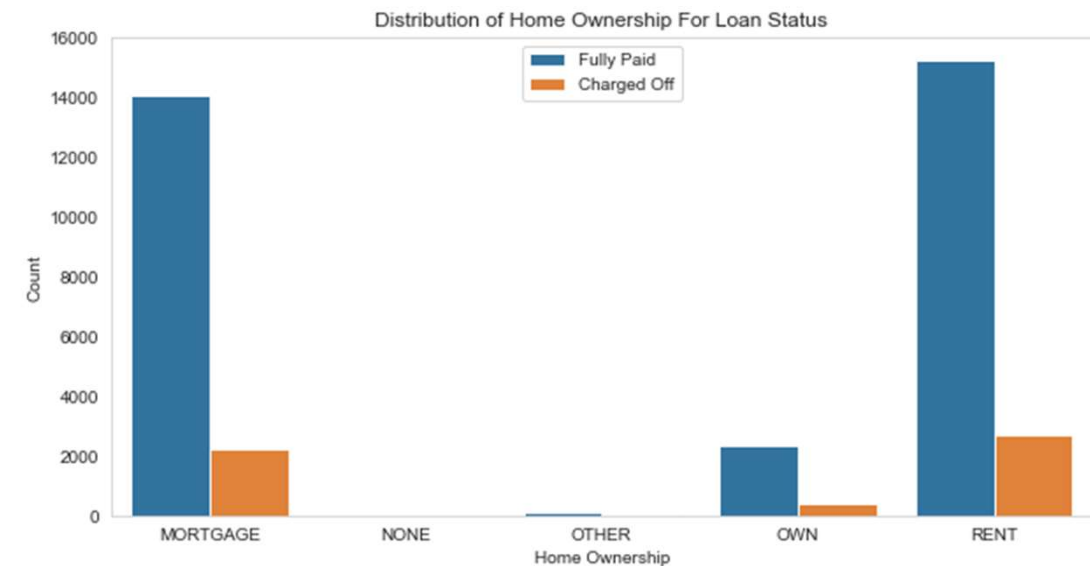


- **Sub Grade:** This provides more insight that the loans within a grade are more skewed towards lowered subgrades.

Employment Length & Homeownership

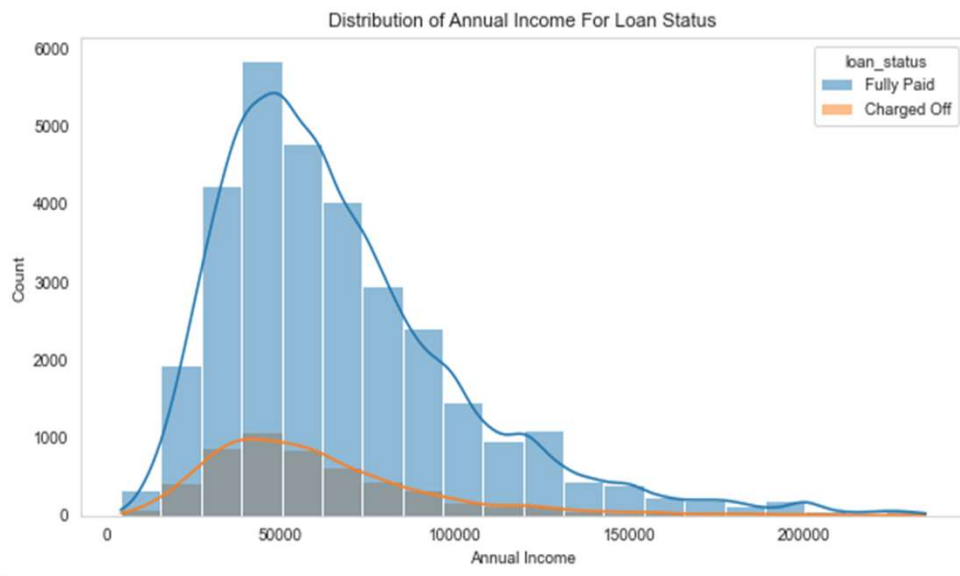


- **Employment Length:** The majority of clients have 10+ years of experience and has the highest number of the defaulted loan.

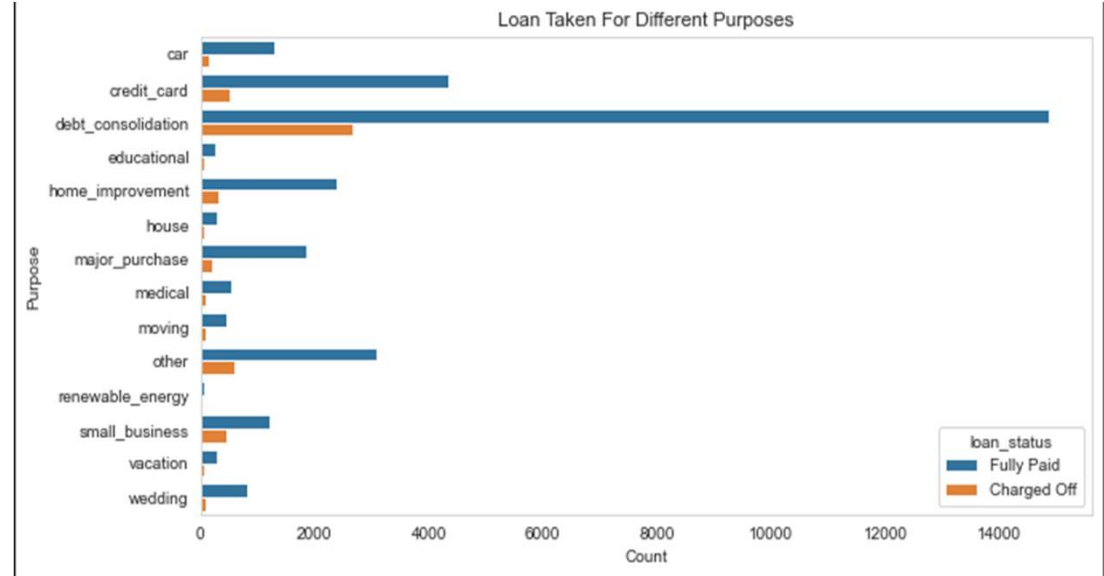


- **Home Ownership:** The majority of clients are lacking ownership of any property and are on rent or mortgage and have a higher chance of defaulting.

Annual Income & Purpose

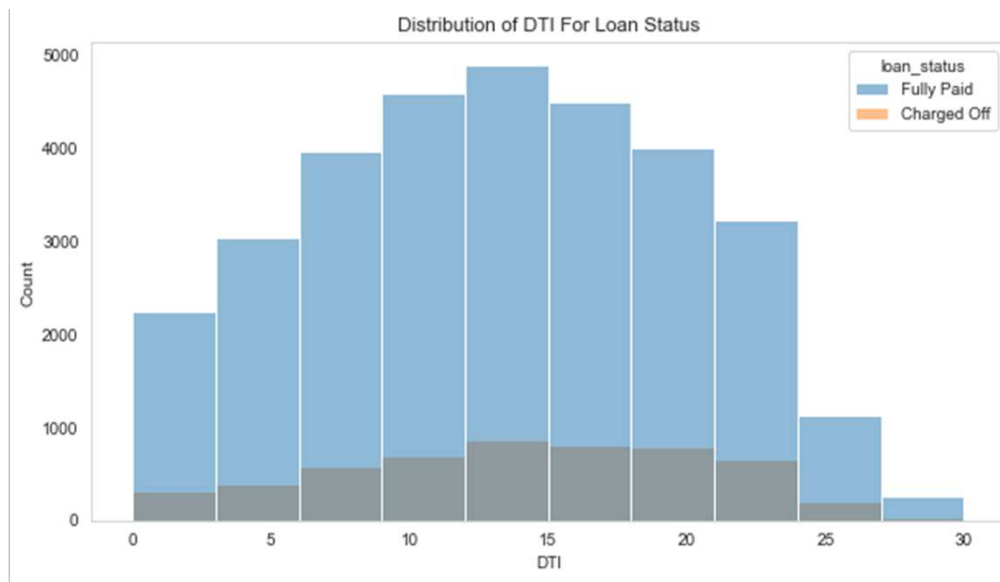


- **Annual Income :** The Majority of clients have low annual income compared to rest and income lower than 50k has higher chance of defaulting.

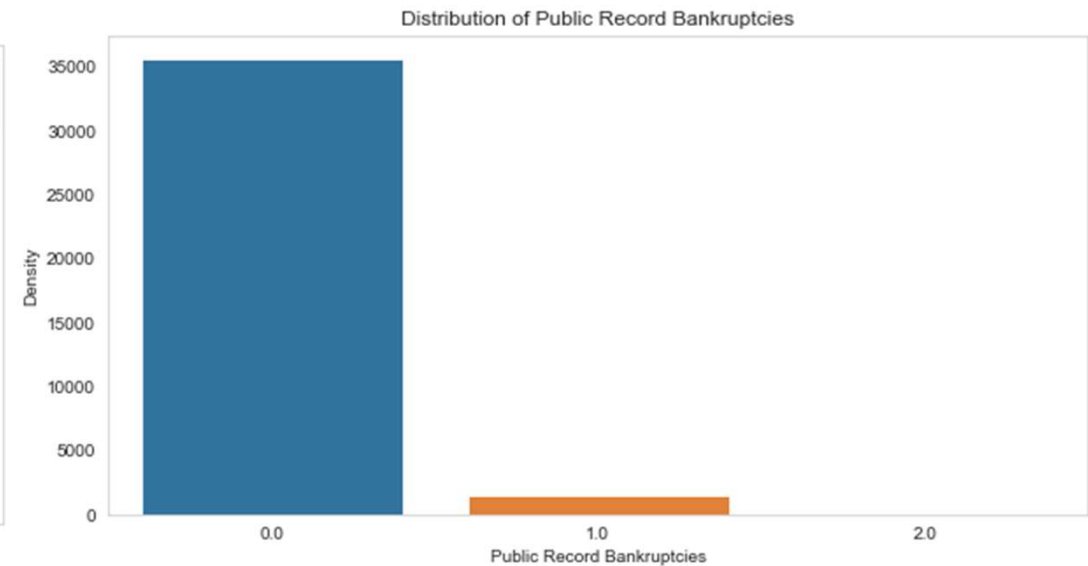


- **Purpose:** Loans are taken mostly for debt consolidation followed by credit card payment. Whereas debt consolidation has the highest fully paid loan but also has the highest number of defaulted loans as well.

DTI ratio & Bankruptcy



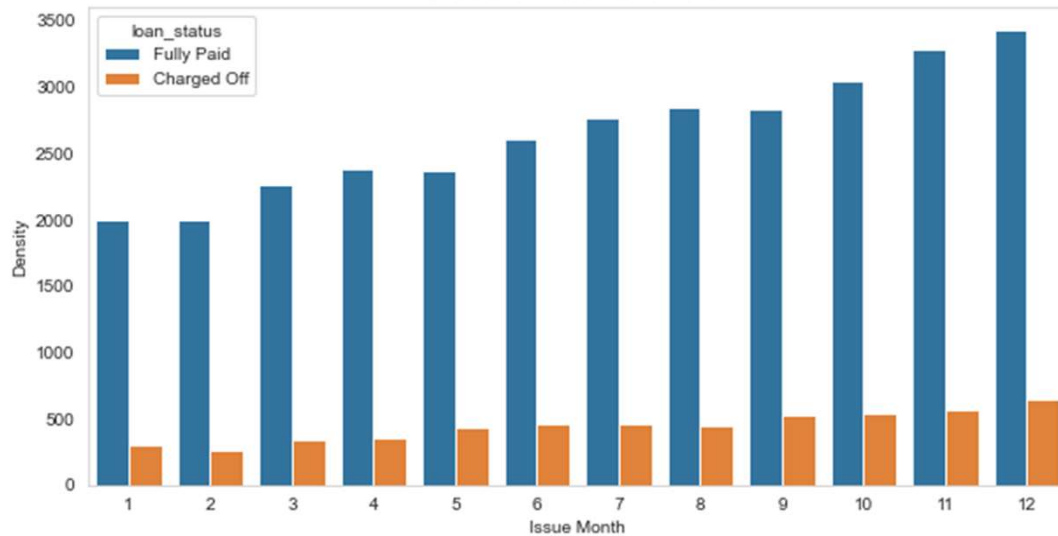
- **DTI:** The large percentage of Clients have a large Debt to Income ratio which shows that lending to such clients can be very risky.



- **Public Recorded Bankruptcy:** The majority of clients have no record of declaring bankruptcy.

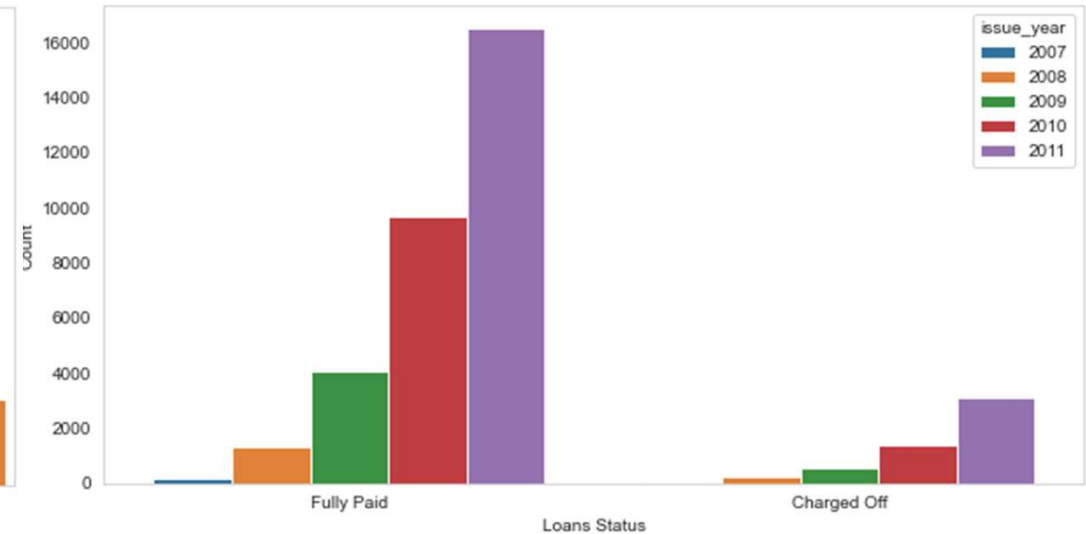
Loan Trend over the years

Distribution of Loan Issue Month



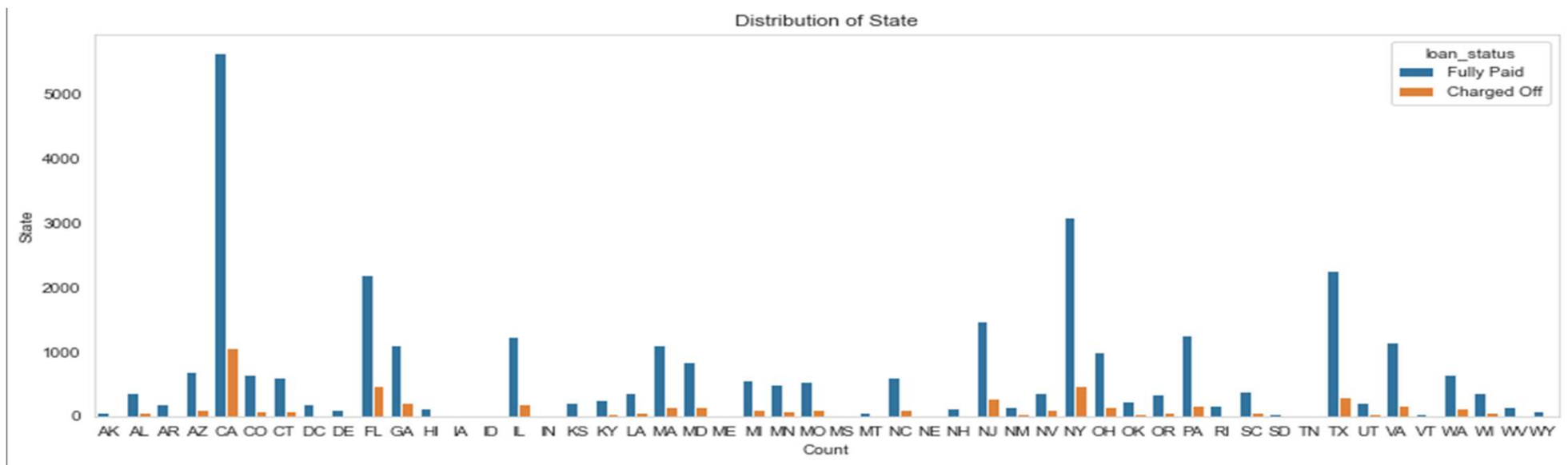
We see a gradual increase in loans taken through the year, with a lesser defaulting rate in April, August, and December quarter-wise and better late in the year.

Distribution of Loan Status For Issue Year



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

Location Based



For large metropolitan cities, we see a large number of loans, with a higher number of defaulted loans like California, New York, Texas, and Florida but have a lower chance of defaulting.

The major driving factor which can be used to predict the chance of defaulting and avoiding Credit Loss:

1. DTI
2. Grades
3. Verification Status
4. Annual income
5. Pub_rec_bankruptcies

Other considerations for 'defaults' :

1. Borrowers not from large urban cities like California, New York, Texas, Florida, etc.
2. Borrowers having annual income in the range of 50000-100000.
3. Borrowers having Public Recorded Bankruptcy.
4. Borrowers with the least grades like E, F, and G indicate high risk.
5. Borrowers with very high Debt to Income value.
6. Borrowers with working experience of 10+ years.

upGrad



Thank You!