

Lending Club Case Study

ANKUSH K

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.

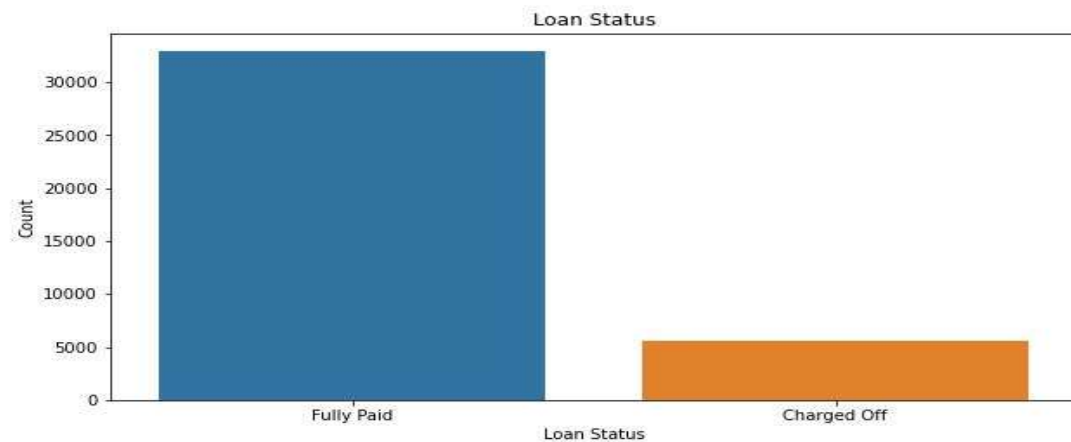
The objective is to take a decision whenever they receive a loan application whether to reject or approve based on certain variables.

Data Clean-up and preparation process:

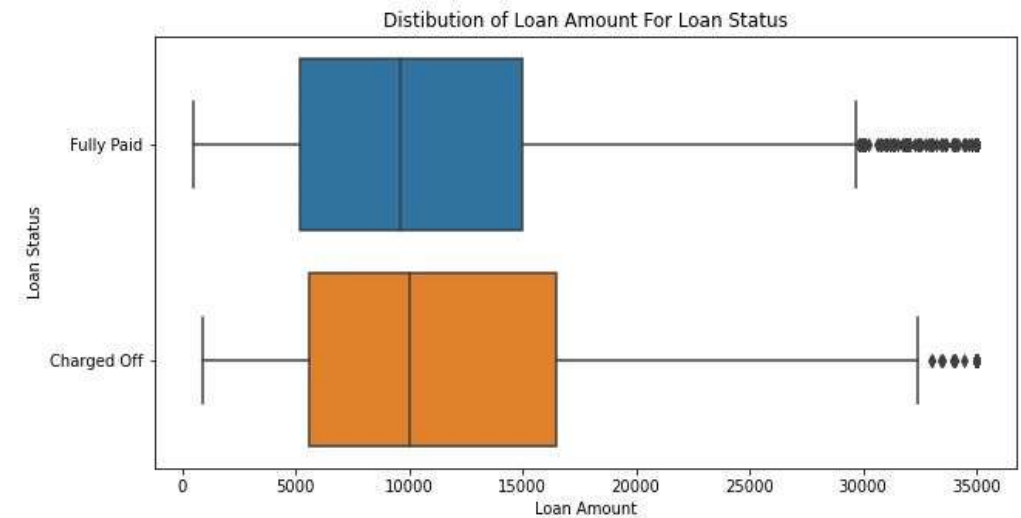
- ☐ Importing the Data
- ☐ Removing large null value columns
- ☐ Removing Duplicate Data
- ☐ Removing irrelevant columns
- ☐ Removing/Fixing null values
- ☐ Correcting data types and deriving new columns
- ☐ Filter Data for requirement
- ☐ Removing outliers

Loan Status and Amount

```
Fully Paid      85.413588  
Charged Off     14.586412  
Name: loan_status, dtype: float64
```



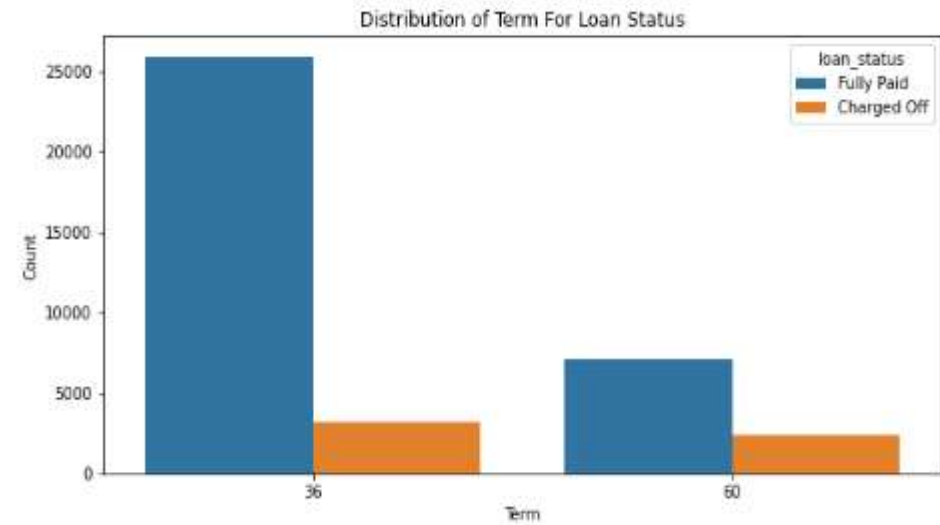
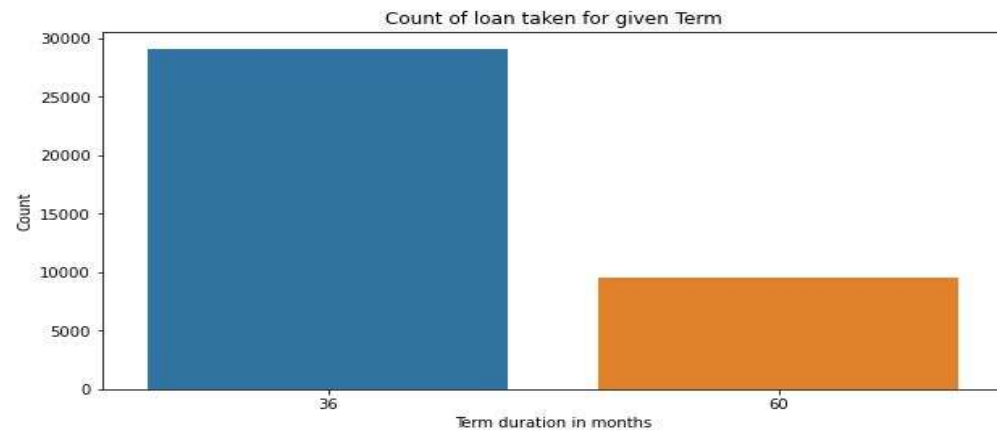
- **Loan Status:** The number of charged off loan is much smaller (14.5%) compared to total count.



- **Loan Amount:** It varies from 500 to 35,000 with a median of 10,000. Loan amount is majorly small and very few clients have taken large loans and larger it goes we have higher chance of defaulting.

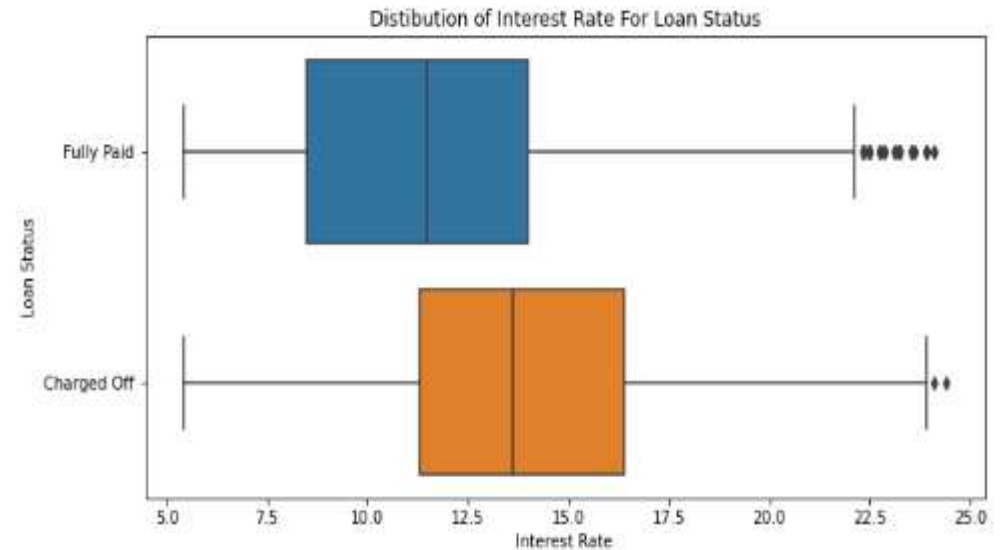
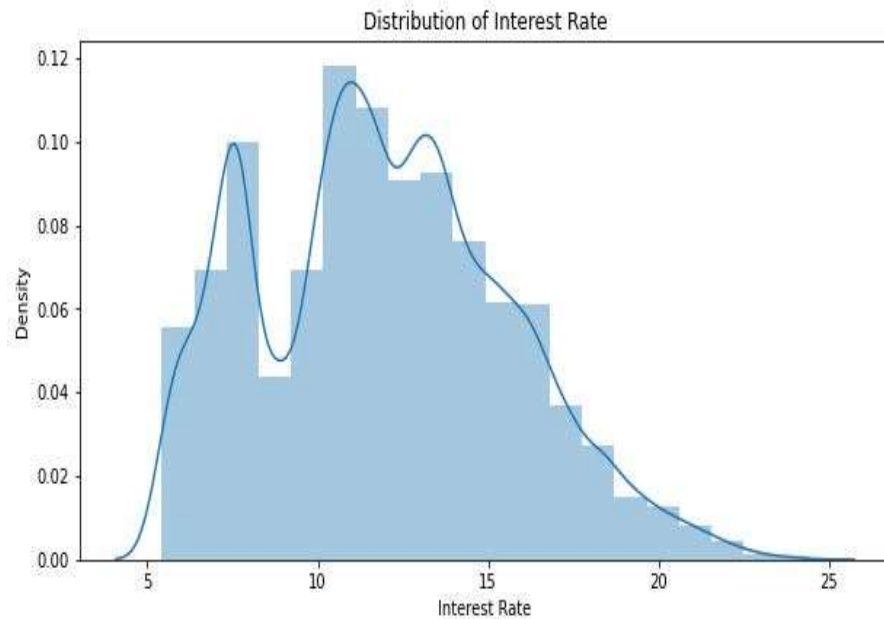
Loan Term

```
36    75.42318
60    24.57682
Name: term, dtype: float64
```



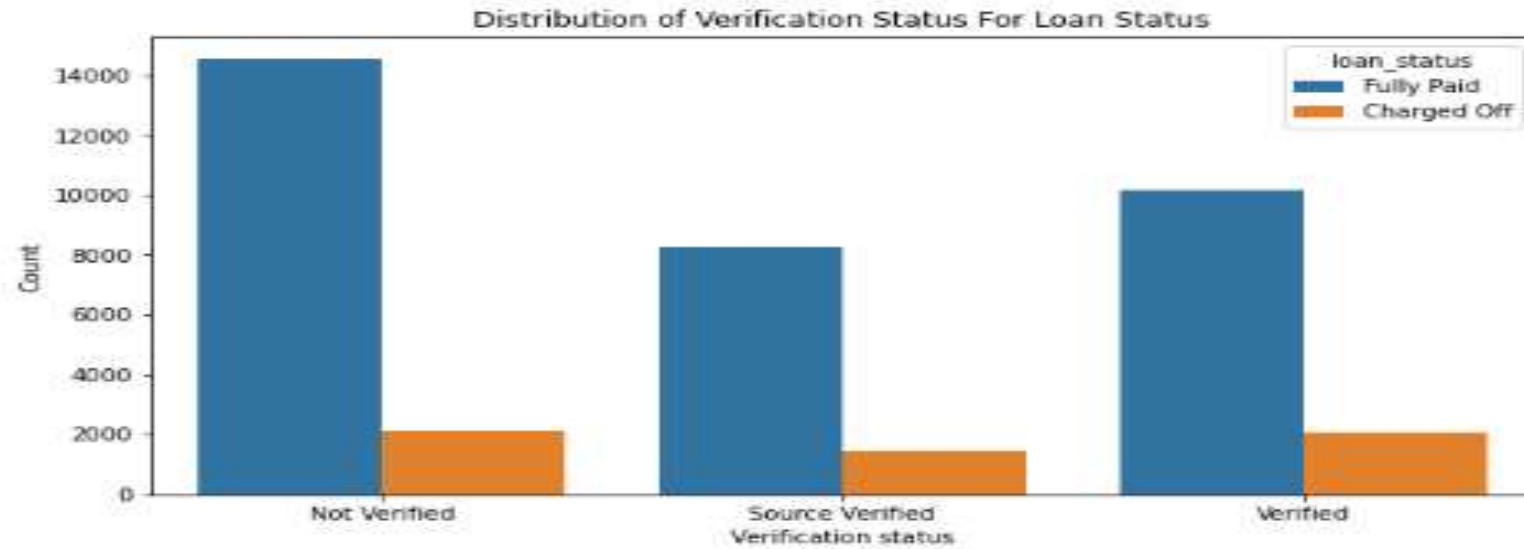
- **Loan Term:** The Loans taken for 36 month term are much more than 60 months and have lower chance of defaulting.

Interest Rate



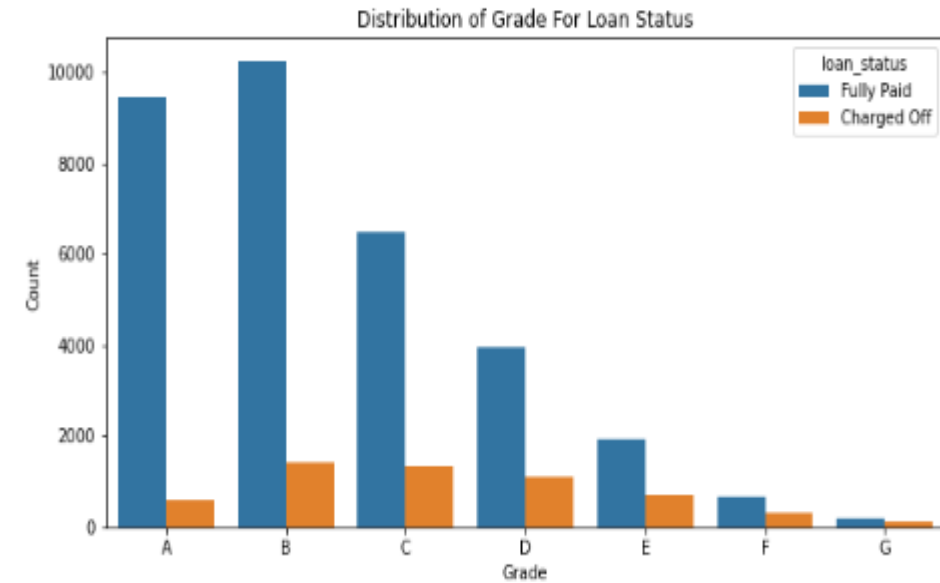
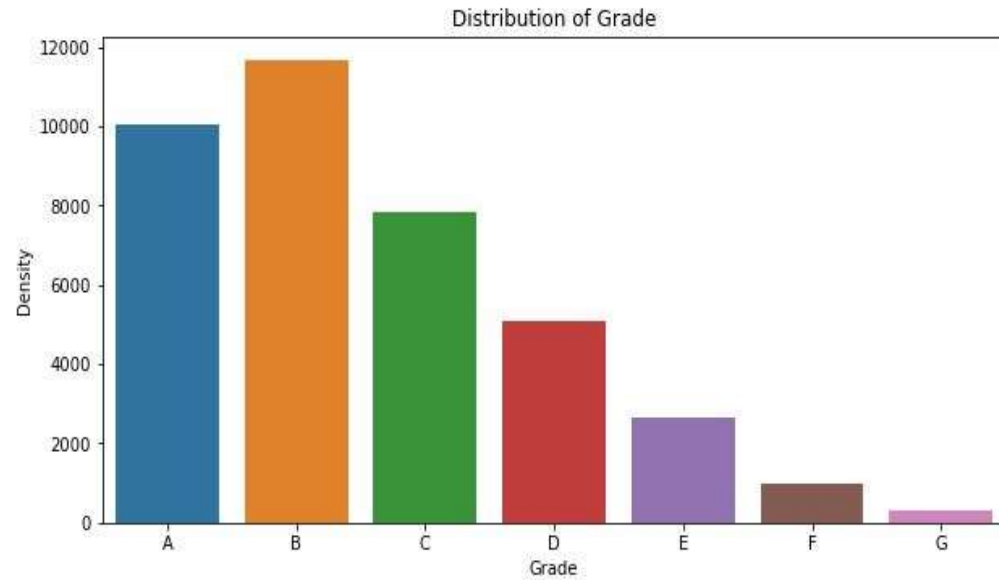
- **Interest Rate:** The count of loan taken varies with interest rate showing peak around in 5-15 bracket and decreasing slowly where as the chance of defaulting increases with rise in interest rate.

Verification Status



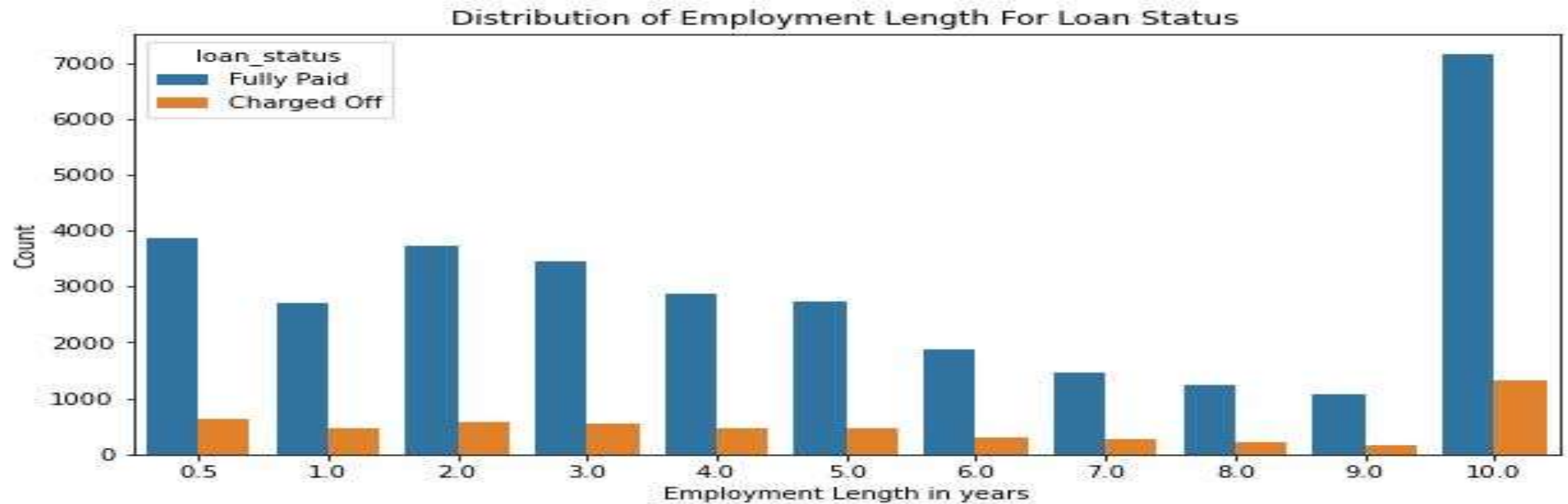
- **Verification Status:** The count of loan taken by verified accounts less than of not verified and thus defaulted loan is more

Grade



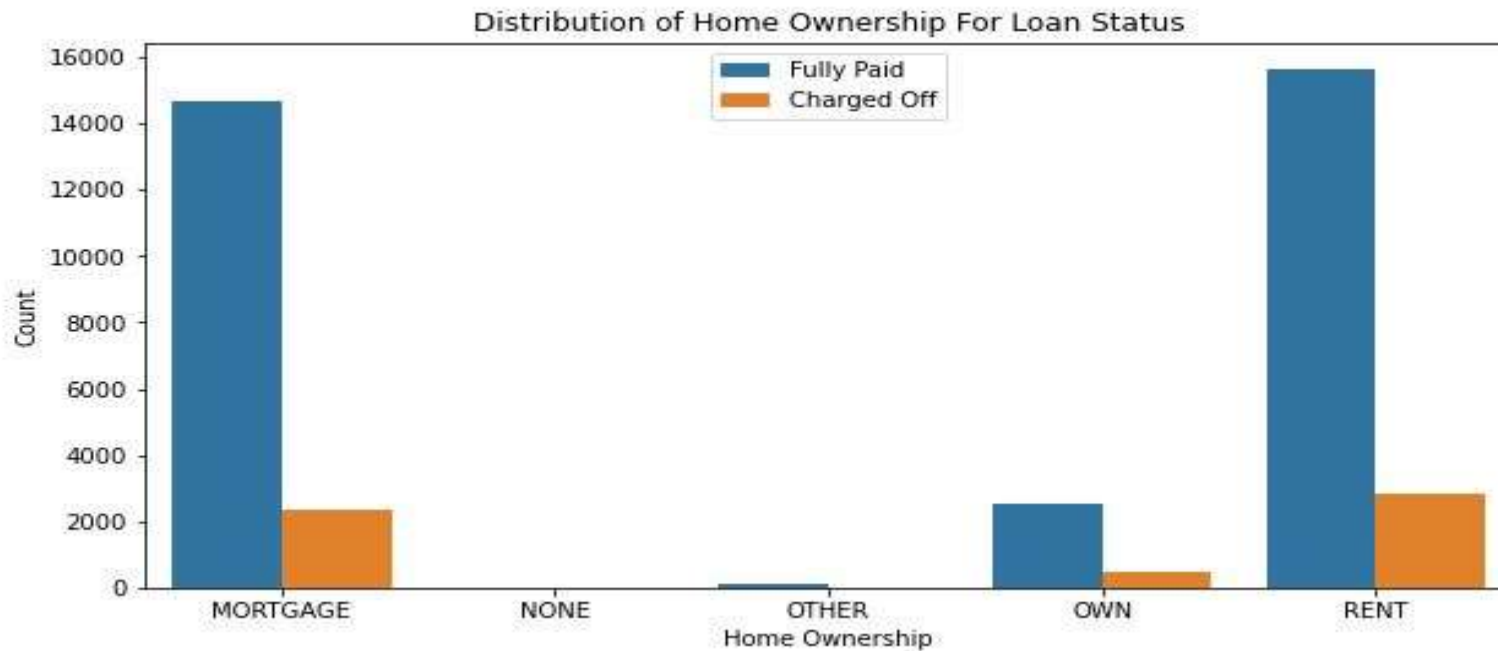
- **Grade:** The loan approved are majorly of higher grade as they are of low risk thus low chance of defaulting.

Employment Length



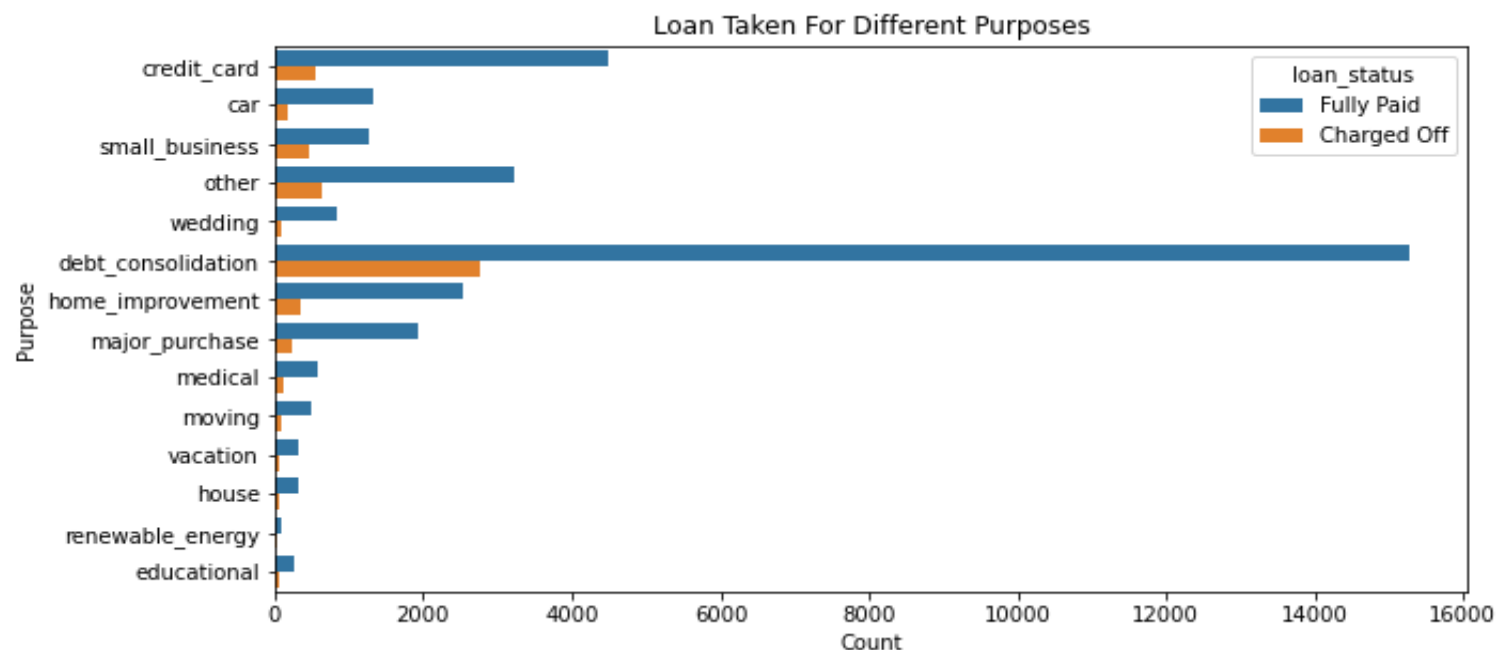
- **Employment Length:** Majority of clients have 10+ years of experience and has highest number of defaulted loan.

Homeownership



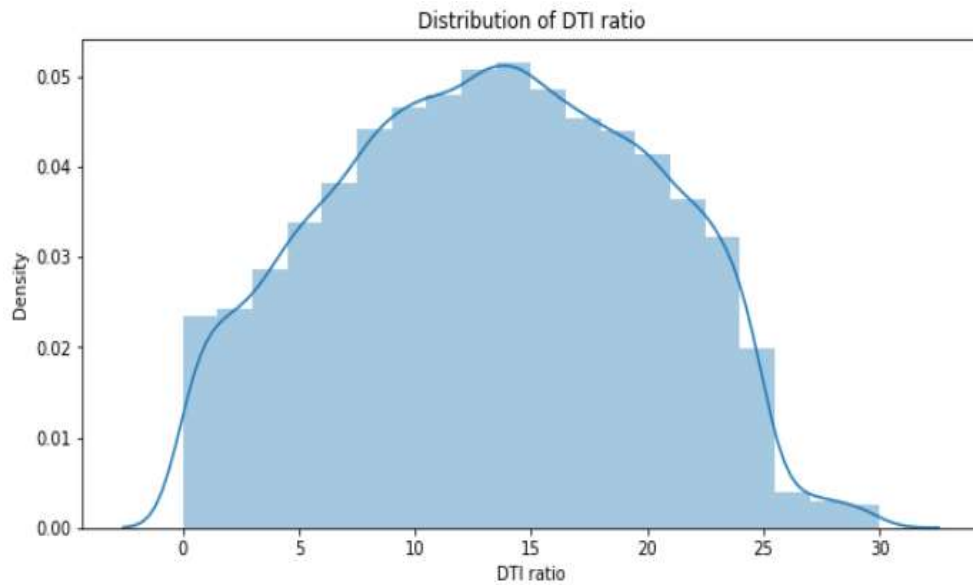
- **Home Ownership:** Majority of clients are lacking ownership of any property and are on rent or mortgage and have a higher chance of defaulting.

Purpose of loan taken

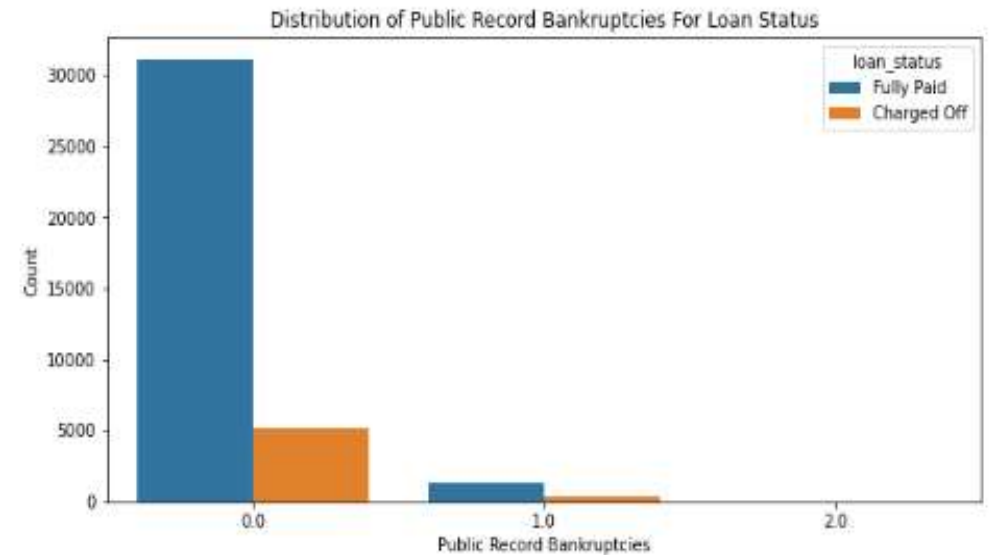


- **Purpose:** Loans are taken mostly for debt consolidation followed by credit card payment whereas the debt consolidation has highest fully paid loan but also has highest defaulted loans as well.

DTI ratio & Bankruptcy

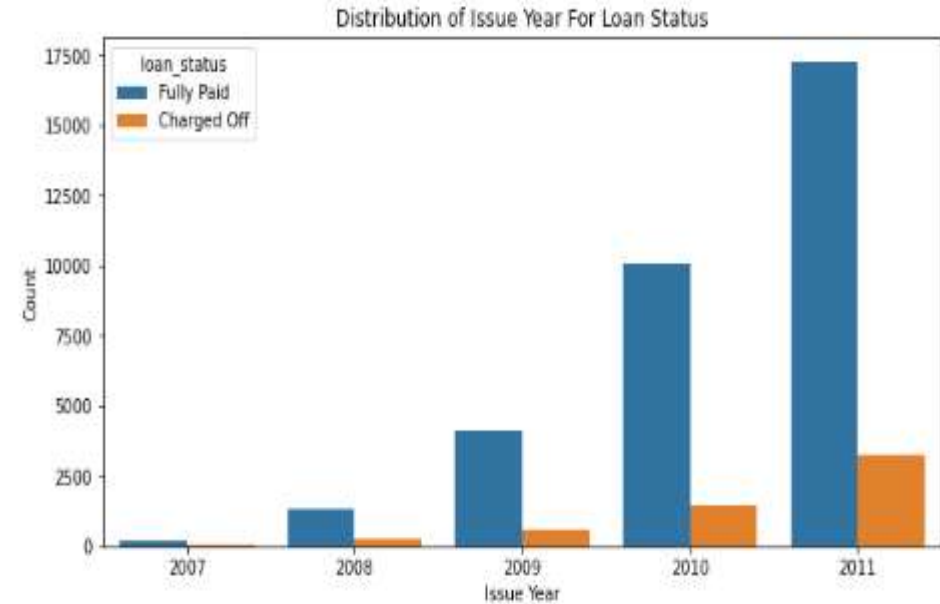
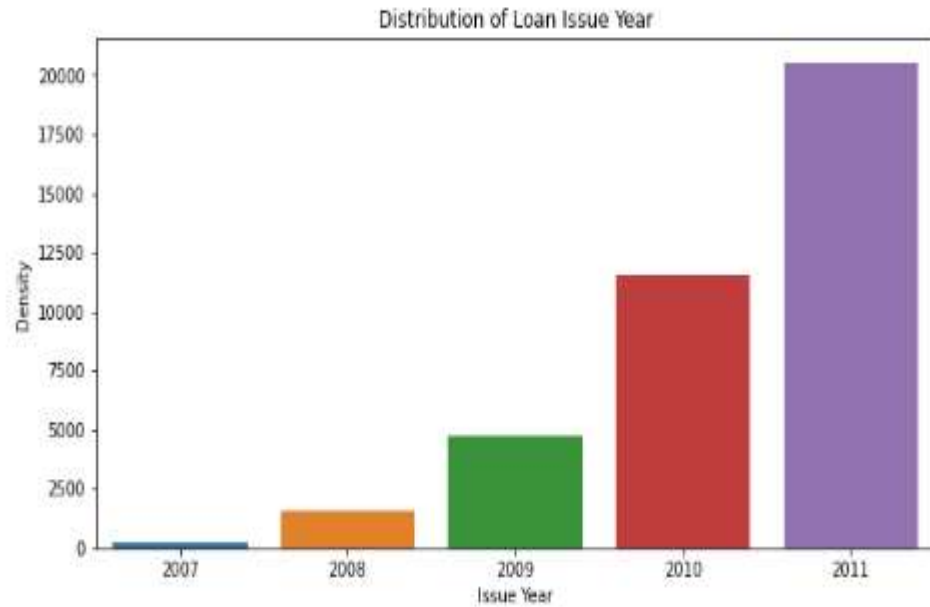


- **DTI:** The large percentage of Clients have a large Debt to Income ratio which shows that lending to such clients can be very risky.



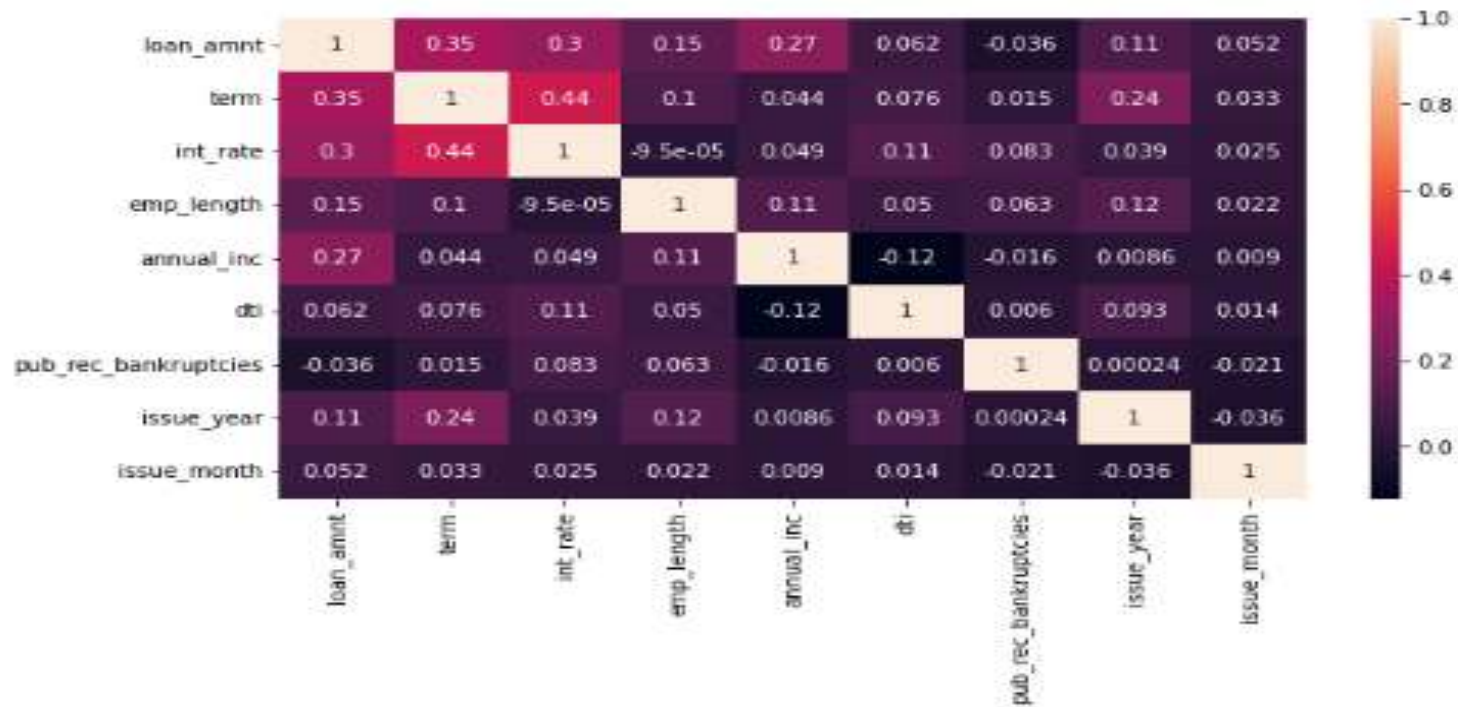
- **Public Recorded Bankruptcy:** Majority of clients have no record of declaring bankruptcy.

Loan Trend over years



With each passing year loan taken are increasing exponentially.

Correlation of Variables



We can check how the variables are correlated to each other and based on that can figure out the possible defaulters.

Recommendations

Recommendations:-

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

- Grades :when they are lower
- Loan Term: when they are of 60 months
- Interest Rate: when they are greater than 15%
- Purpose of Loan Taken : when they are for credit card and debt consolidation

➤ Considering a strict verification in these field would lower the chance of defaulting

Other considerations for 'defaults' :

- Borrowers having Public Recorded Bankruptcy.
- Borrowers with very high Debt to Income value.
- Borrowers with working experience 10+ years.