Lending Club Case Study

PRESENTERS:

KALLAL NATH

PRIYADHARSHINI V

Case Study Objective

Expected outcome from the case study:
➢ Identify variables which are potential indicators of loan
➢ Share key observations around range of customer attributes and loan attributes which can add as useful insights for business.

Business Benefits:
➢ Contributes towards reducing credit loss of Lending Company
➢ Create a working strategy to efficiently handle prospective customers
➢ Helps in taking the right and effective decision on whether to reject or approve based on certain variables.
➢ Deriving data-driven decisions such as rejecting a loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.
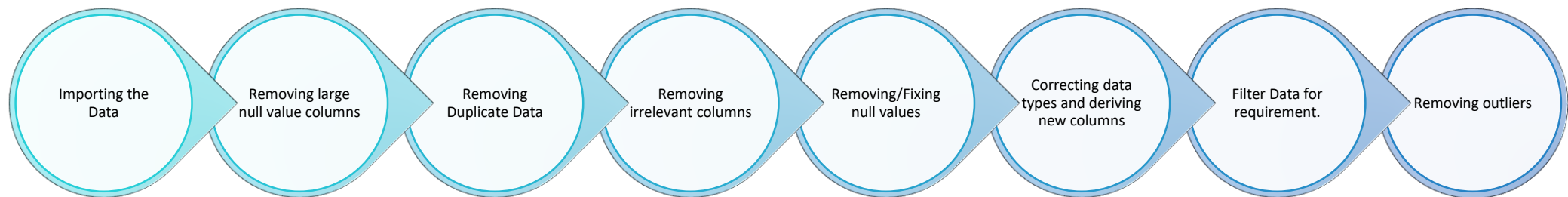
Analysis Approach

Documents and Artefacts provided for analysis:

➢ **Loan Dataset** containing 111 columns and 39717 rows

➢ **Data Dictionary** explaining the definition and business context of each field in the dataset

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 |

Data Exploration and Data Cleaning

Findings of data standardizations are mentioned below.

➢ 54 columns having 100% null values removed. 3 columns with more than 60% null values are removed

➢ 9 columns with only 1 unique value is removed

➢ 6 columns irrelevant to the analysis are removed

➢ Removed rows with null values to standardise data

➢ Removed '%' symbol and converted around 5 columns to appropriate datatype columns

➢ Identified and removed outliers in loan amount and annual income columns

Univariate Analysis

Observations in univariate analysis are not closely related to our objective. Presenting the facts derived out of the analysis-

➢ Around 14% of the borrowers during the period 2007 – 2011 turned out to be defaulters.

➢ Borrowers frequently opt for loan between 5%-7% and 10% -15% interest rate

➢ More than 50% of the loans granted in 5 years were with high grade 'A', 'B' and 'C'.

➢ Borrowers in rental house or mortgaged house are more interested in acquiring a loan.

➢ Majority of the borrowers have no record of Public Recorded Bankruptcy.

➢Majority of the borrowers have very large debt compared to their income, concentrated mostly in the 10-20 DTI ratio.

➢ Majority of borrowers have fairly low annual income as compared to rest who are less in numbers.

➢ Majority of the loans were taken for debt consolidation followed by credit card. Such borrowers already have running debt to cover. And they take further loan to pay for their credit card outstanding.
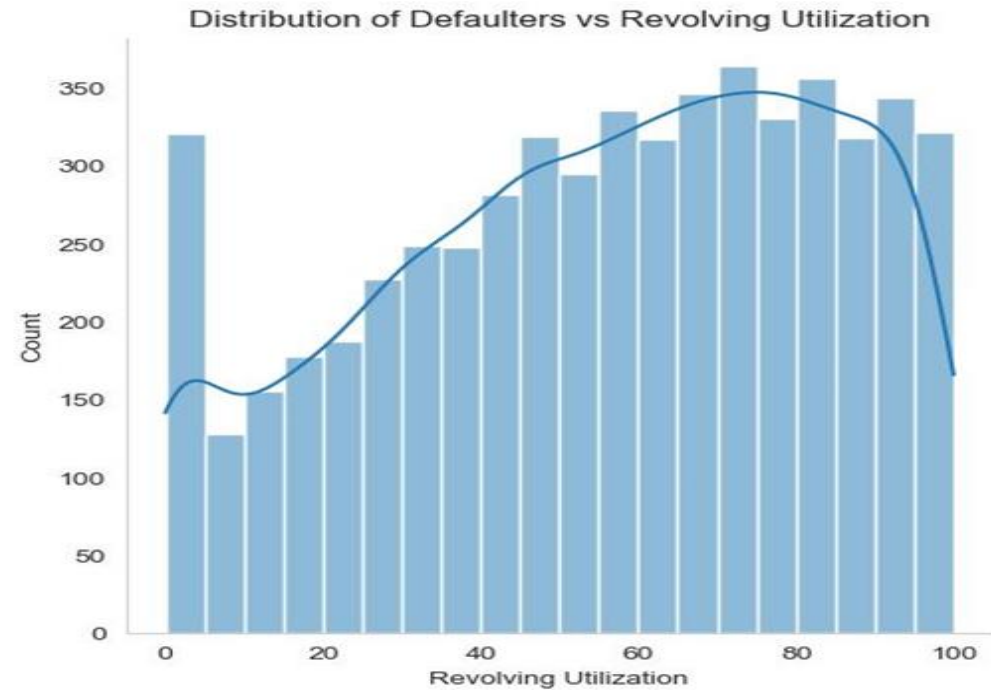
Segmented Univariate Analysis

Revolving Utilisation

Distribution of Fully Paid vs Revolving Utilization

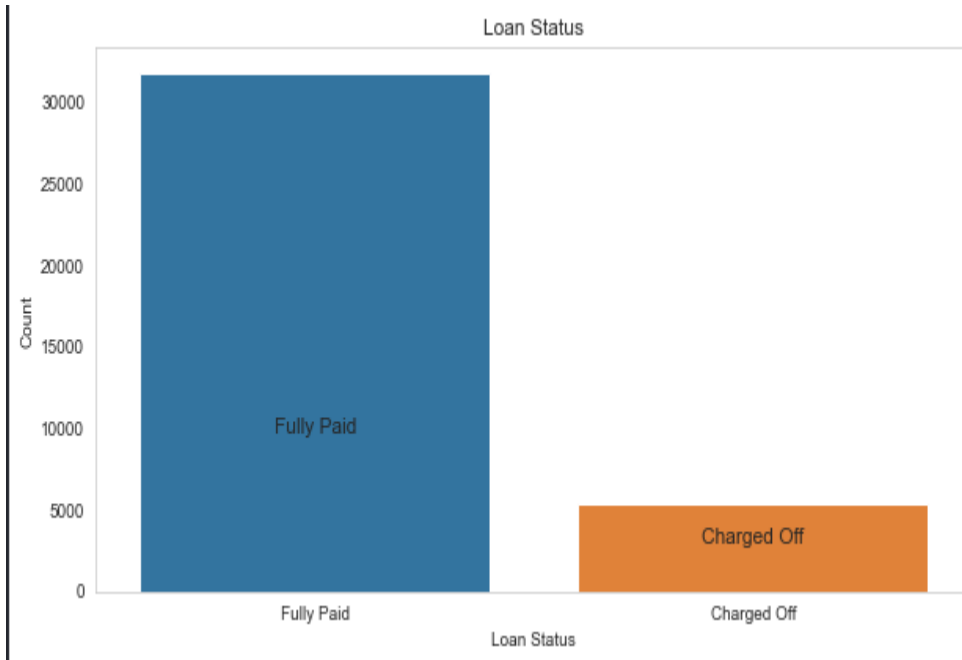Distribution of Defaulters vs Revolving Utilization
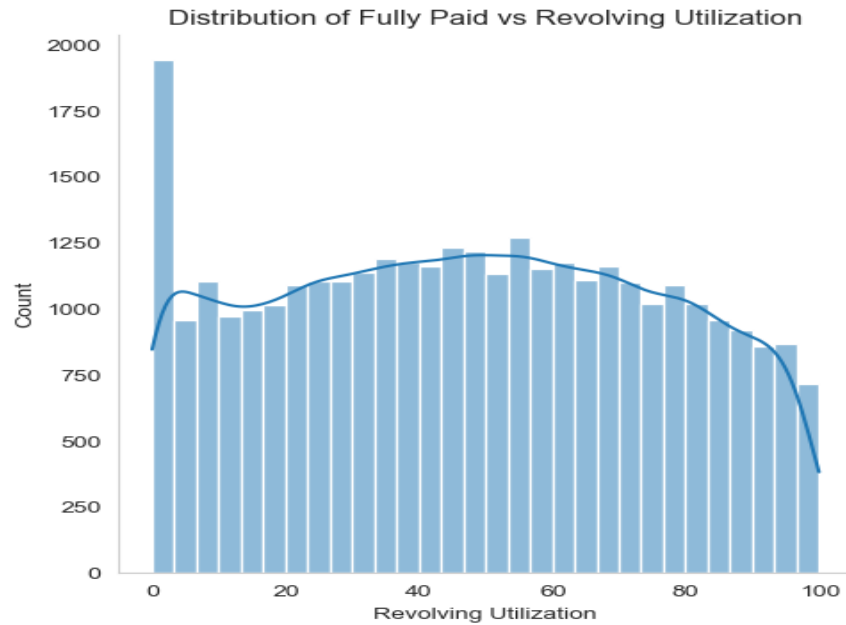
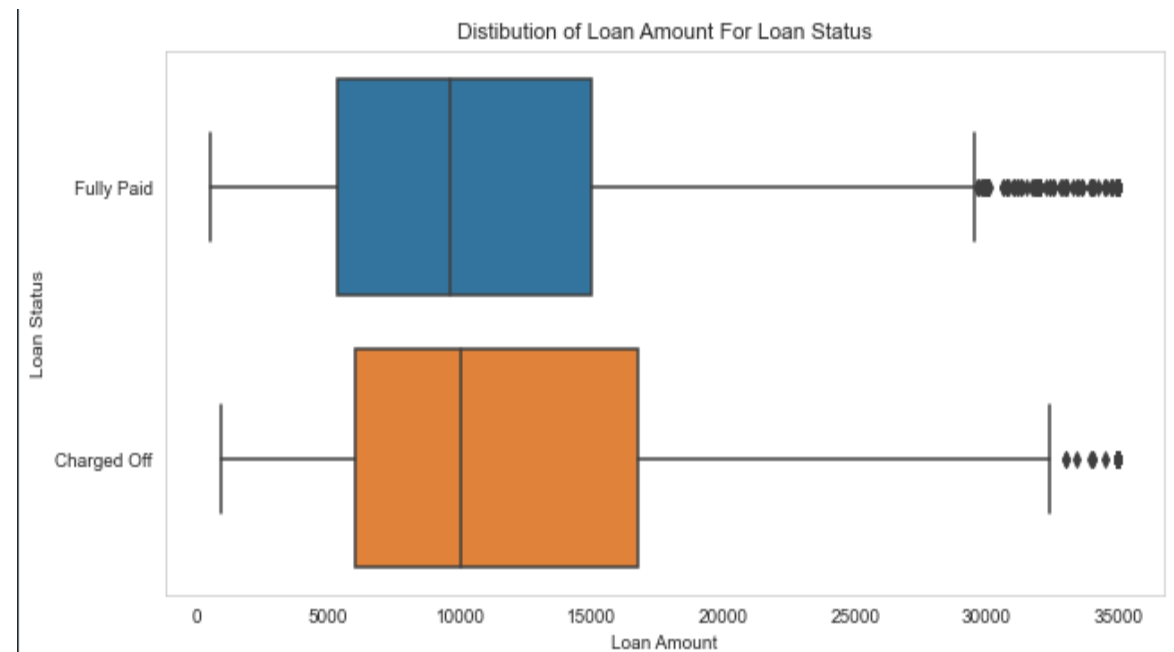- **Fully paid customers** maintain lesser revolving utilisation

- Defaulters are more likely to have more than 50% of revolving utilization.

Loan Status and Amount

**Loan Status:** The proportion 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. Loan amount is majorly small and very few borrowers took large loans and larger it goes there is higher chance of defaulting.

Term and Interest Rate

Distribution of Term For Loan Status

Distribution of Interest Rate For Loan Status
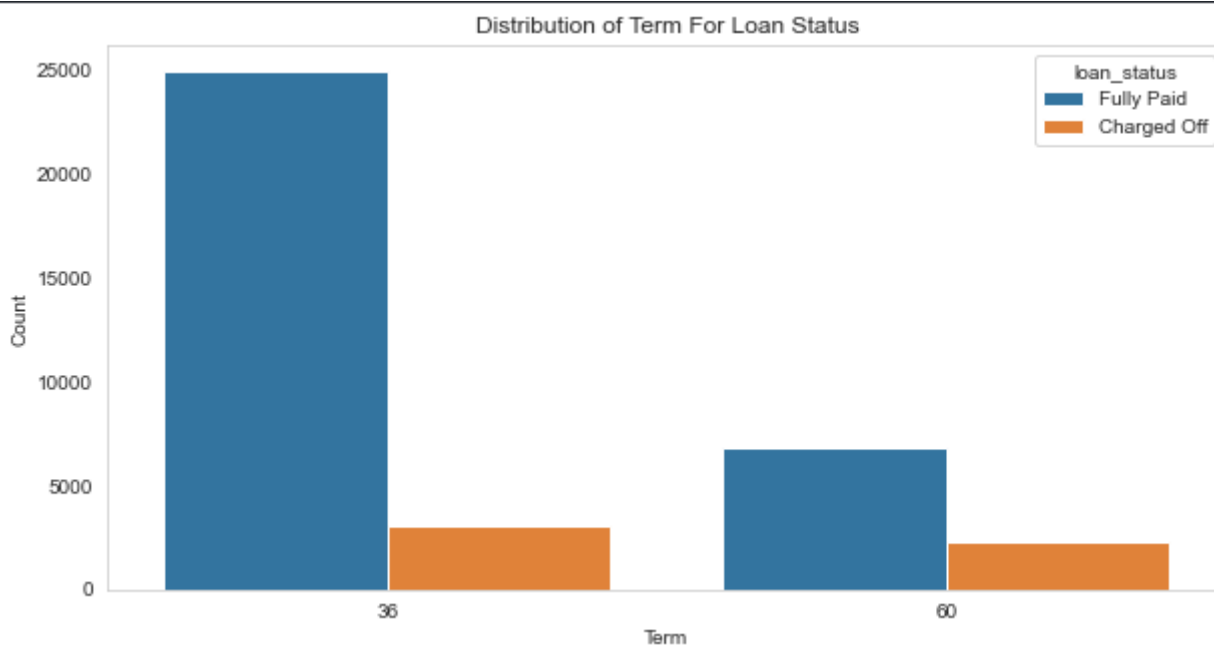
▪**Loan Term:** The Loans taken for 36 month term are much more than 60 months and have lower chance of defaulting.
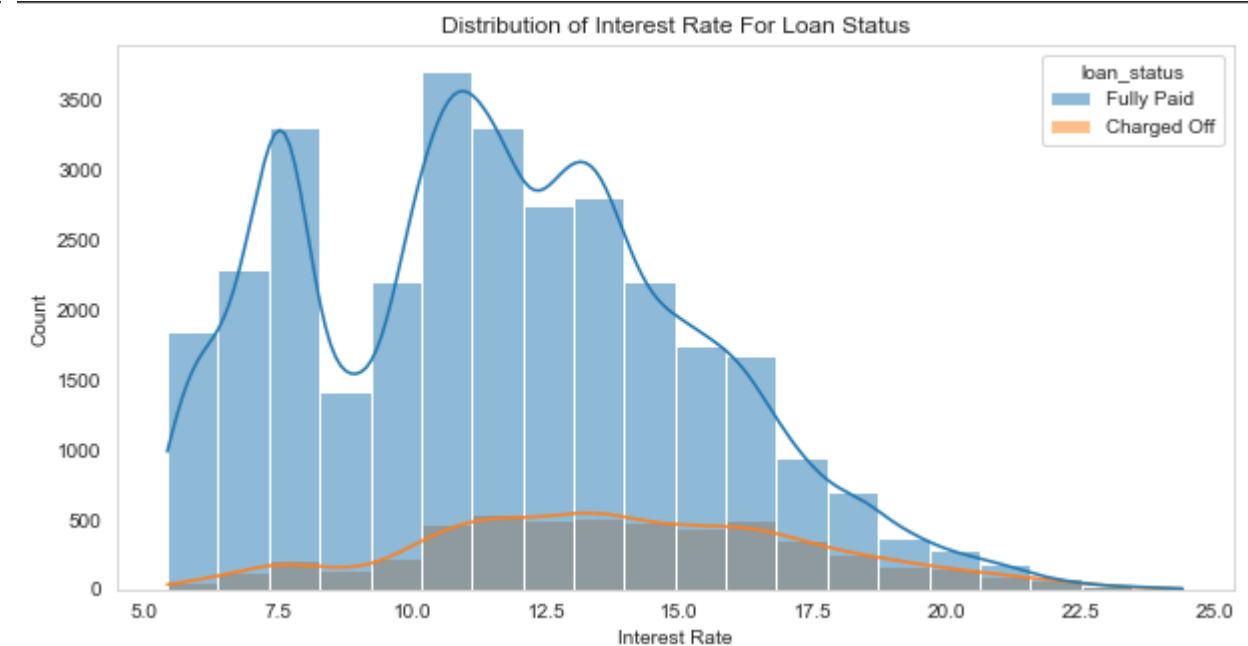
▪**Interest Rate:** The count of loan taken varies with interest rate showing peak around in 5-15 bracket and decreasing slowly where as the chance of defaulting increases with interest rate.

Grade and Sub-Grade

Distribution of Term For Grade

Distribution of Secondary Grade
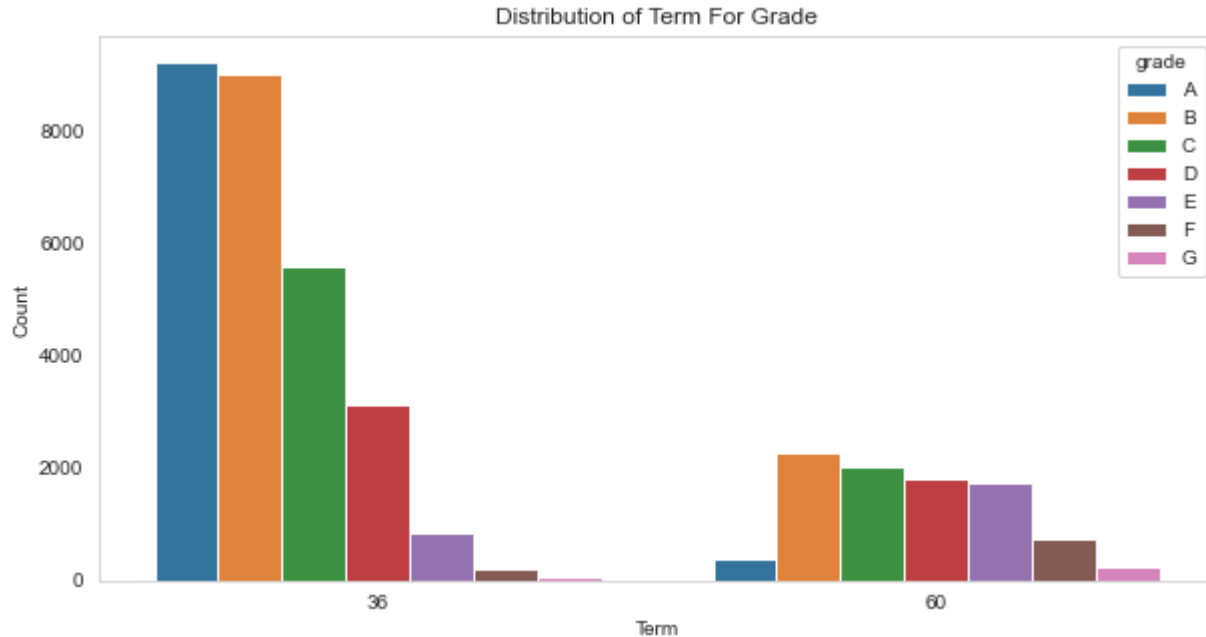
▪**Grade:** The loan approved are majorly of higher grade as they are of low risk thus low chance of defaulting. 60 month term loans have larger number of lower grade loans with high risk.
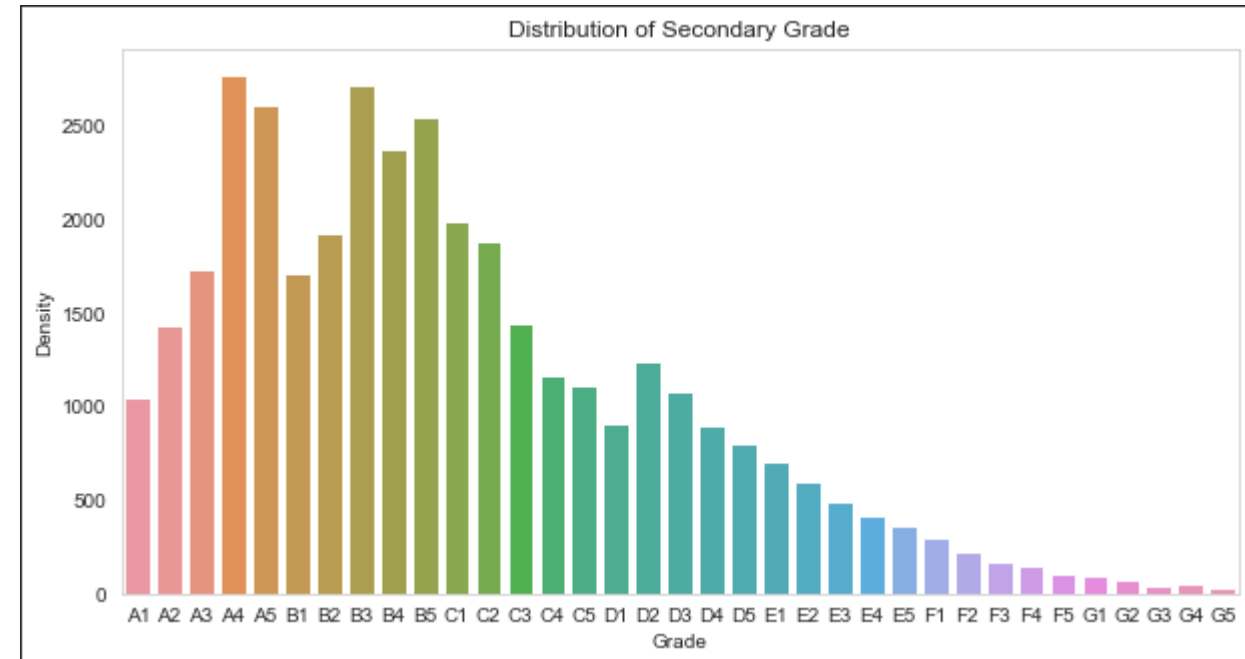
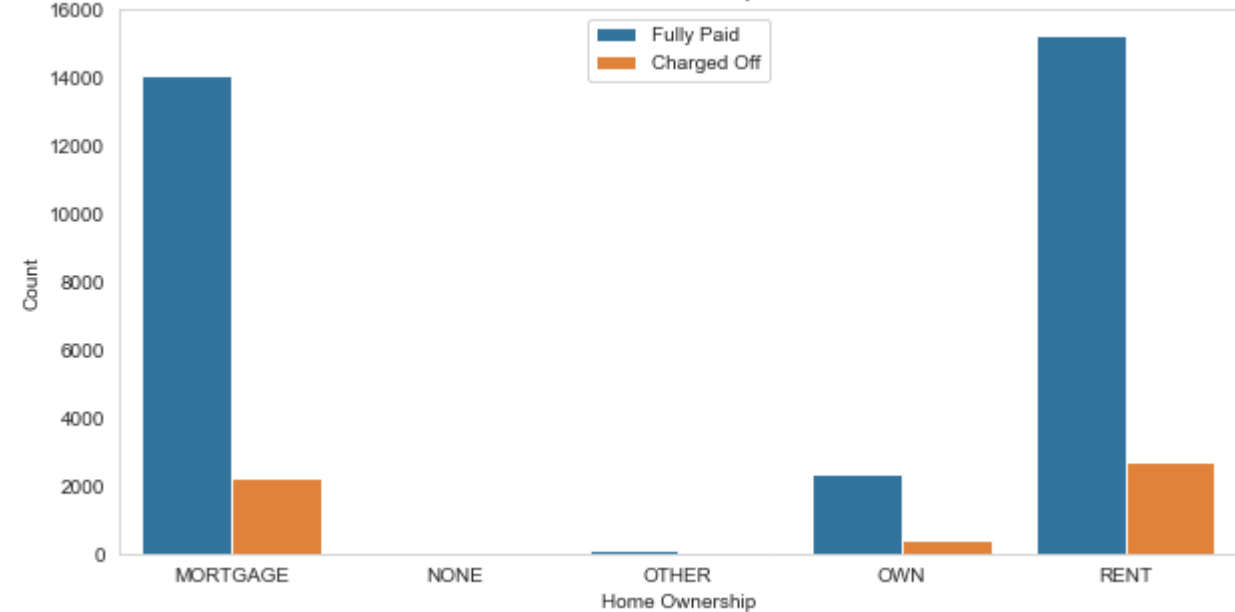▪**Sub Grade:** This provides more insight that the loans within grade are more skewed towards lowered sub grades.

Employment Length & Homeownership

Distribution of Employment Length For Loan Status

Distribution of Home Ownership For Loan Status
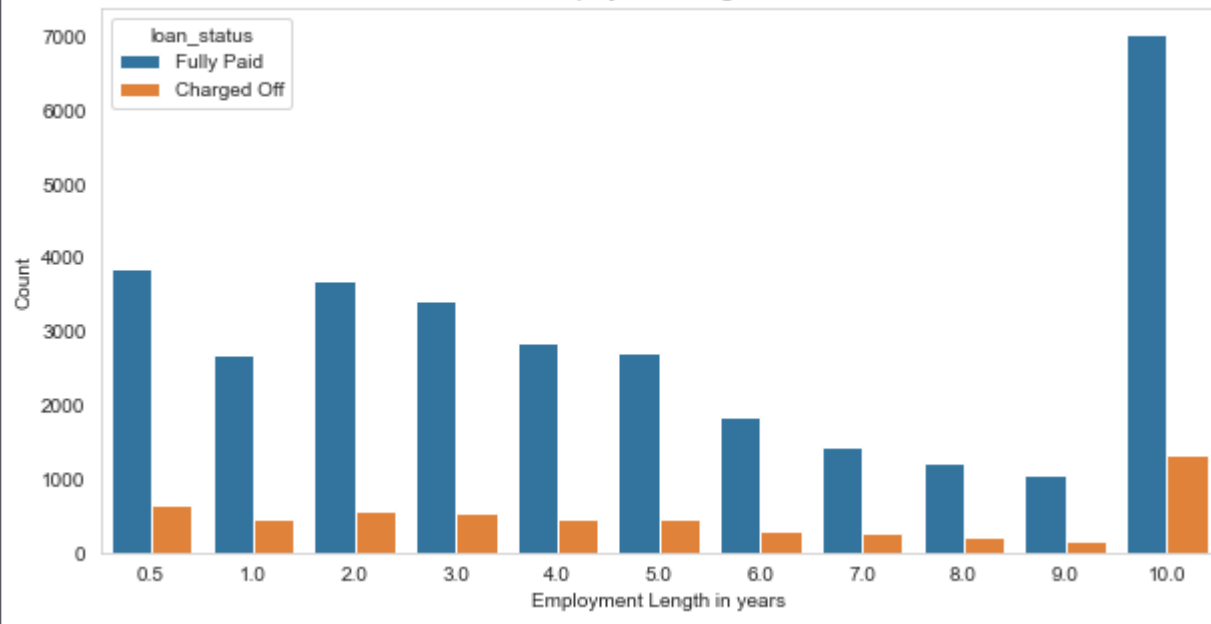
▪**Employment Length:** Majority of clients have 10+ years of experience and has highest number of defaulted loan.

▪**Home Ownership:** 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

Distribution of Annual Income For Loan Status

Loan Taken For Different Purposes
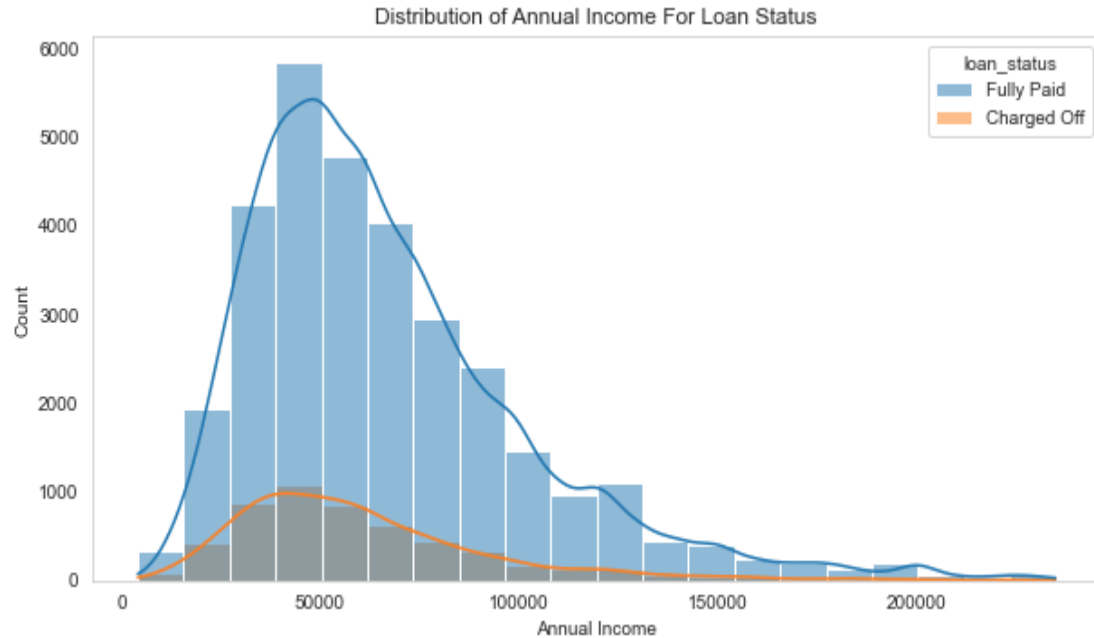
▪ **Annual Income :** Majority of borrowers have low annual income compared to rest and income lower than 50k has higher chance of defaulting.
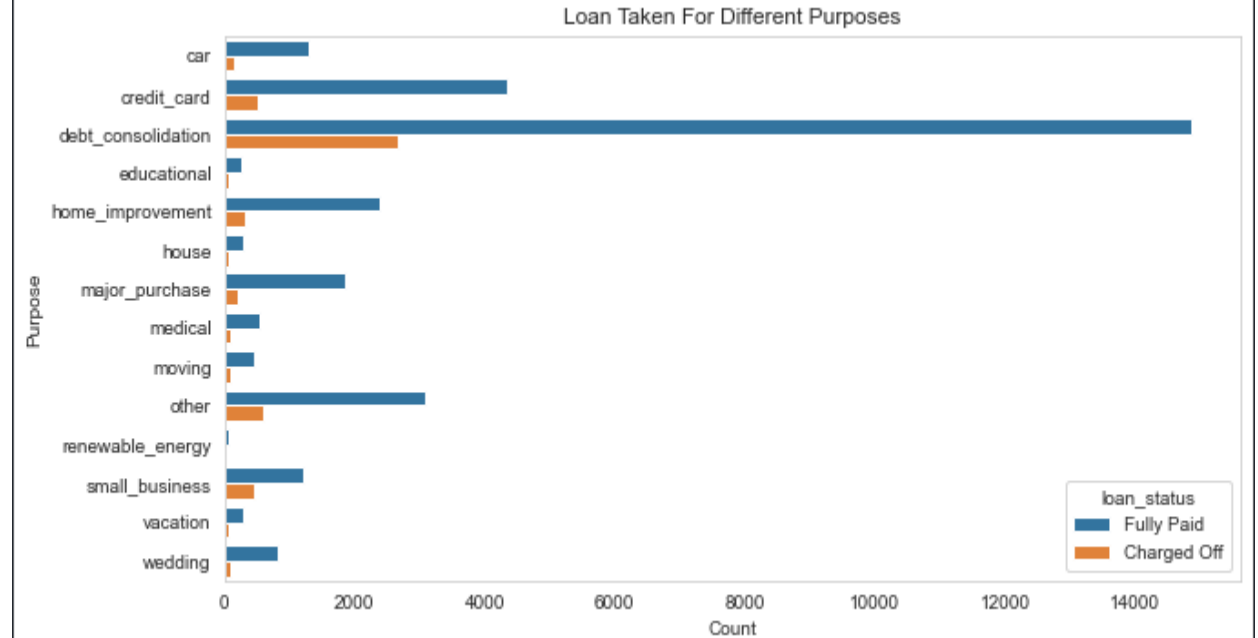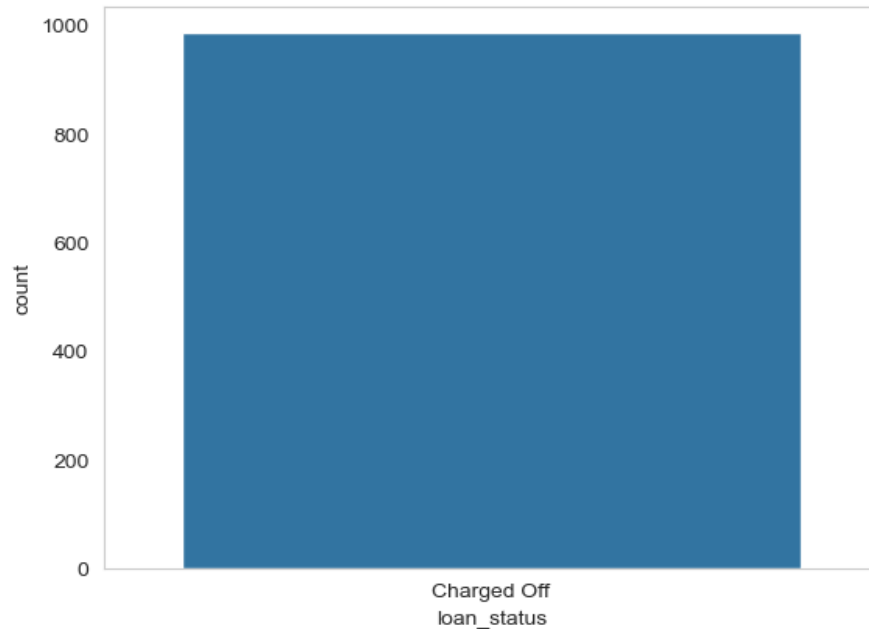
▪ **Purpose:** Loans are taken mostly for debt consolidation followed by credit card payment. Whereas the debt consolidation has highest fully paid loan but also has highest defaulted loans as well.
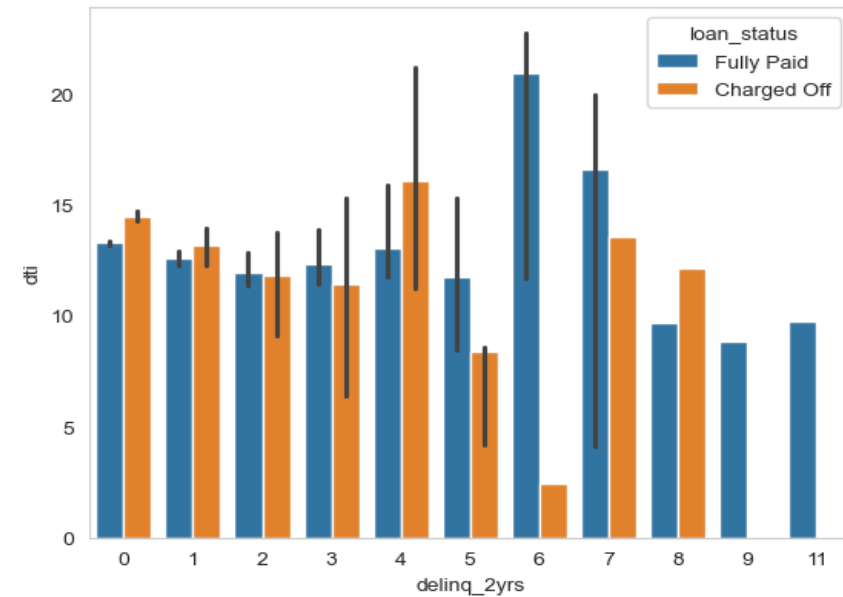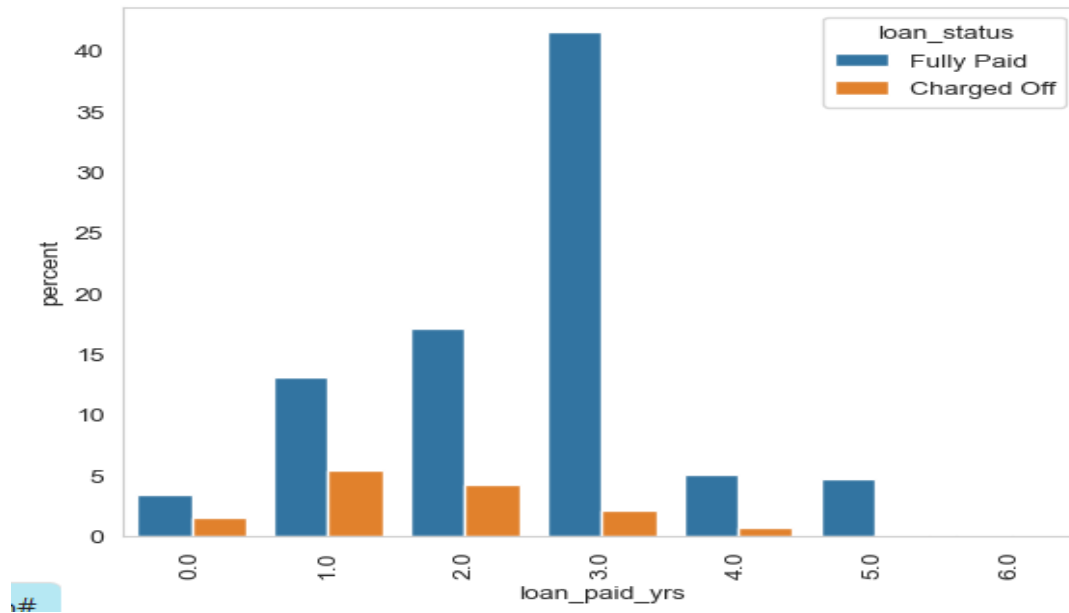
Bivariate Analysis

Total Interest and Principal paid

DTI & Delinquency

- Borrowers who paid interest exceeding principal amount are 100 % defaulters

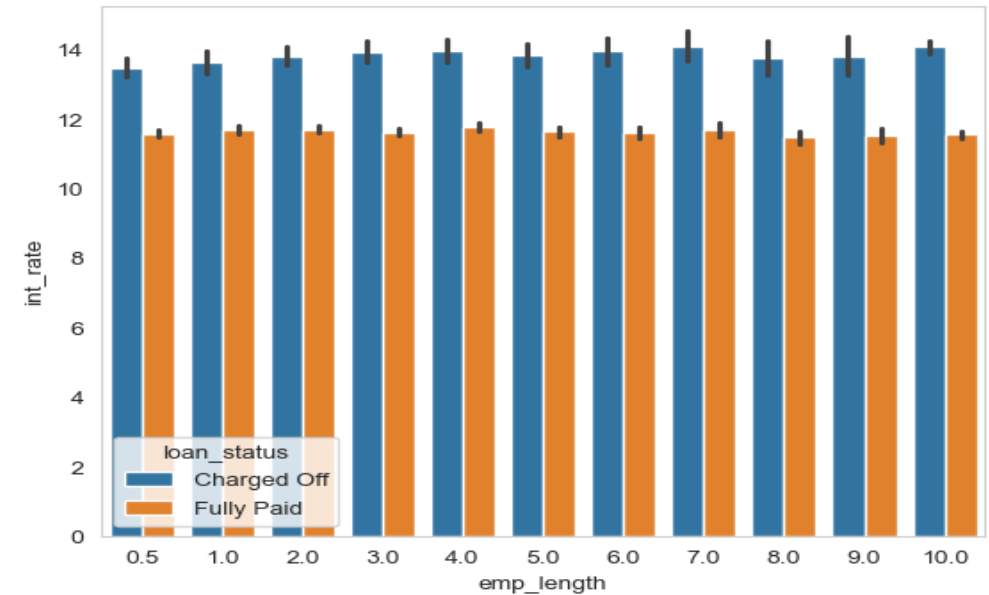- More than 12 DTI have higher chances for default. Delinquency have a very weak connection in judging defaulters
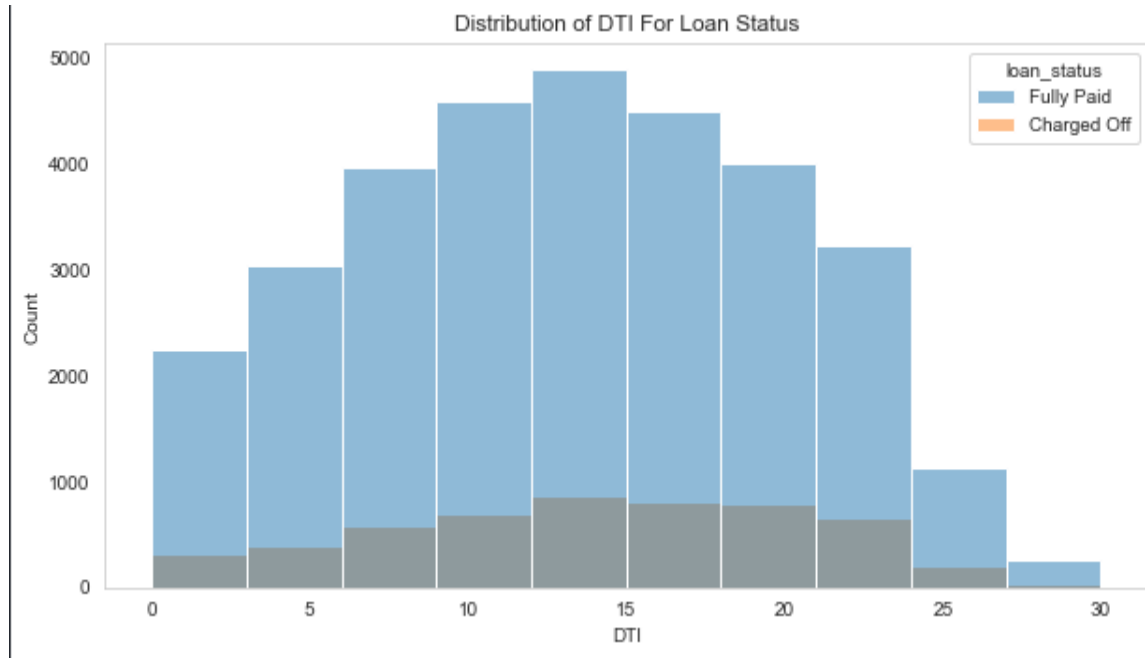
Loan paid duration

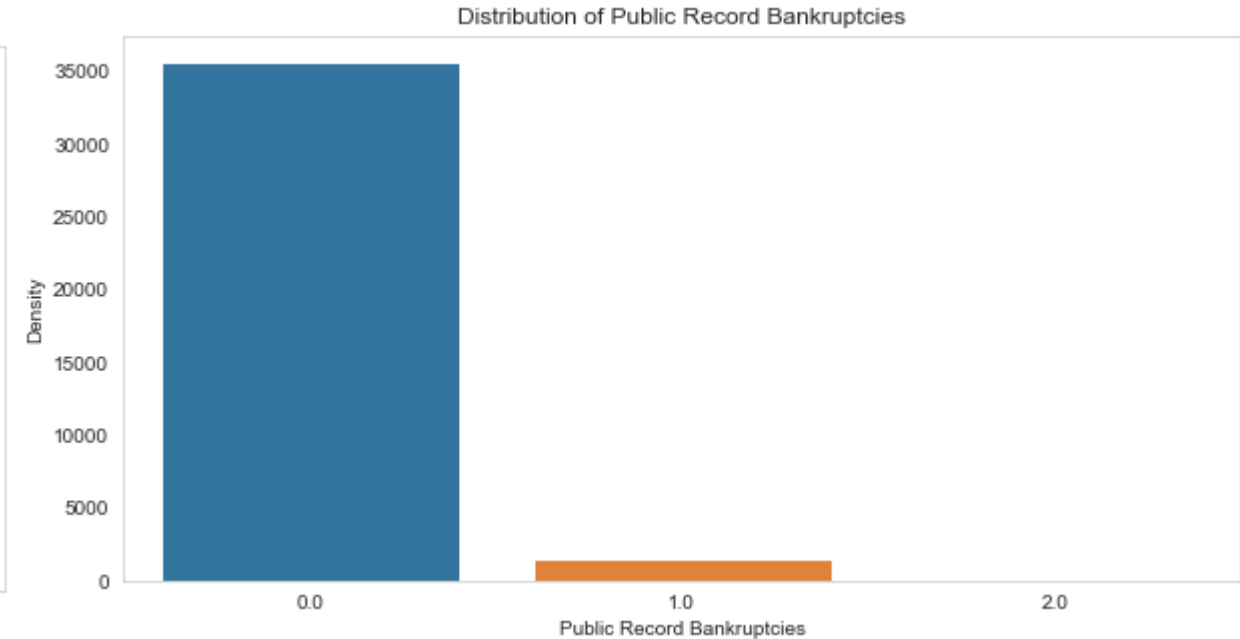- Loan duration of less than 3 years have high chance for defaulting

DTI & Delinquency

- More than 12% interest rate are 100% defaulters

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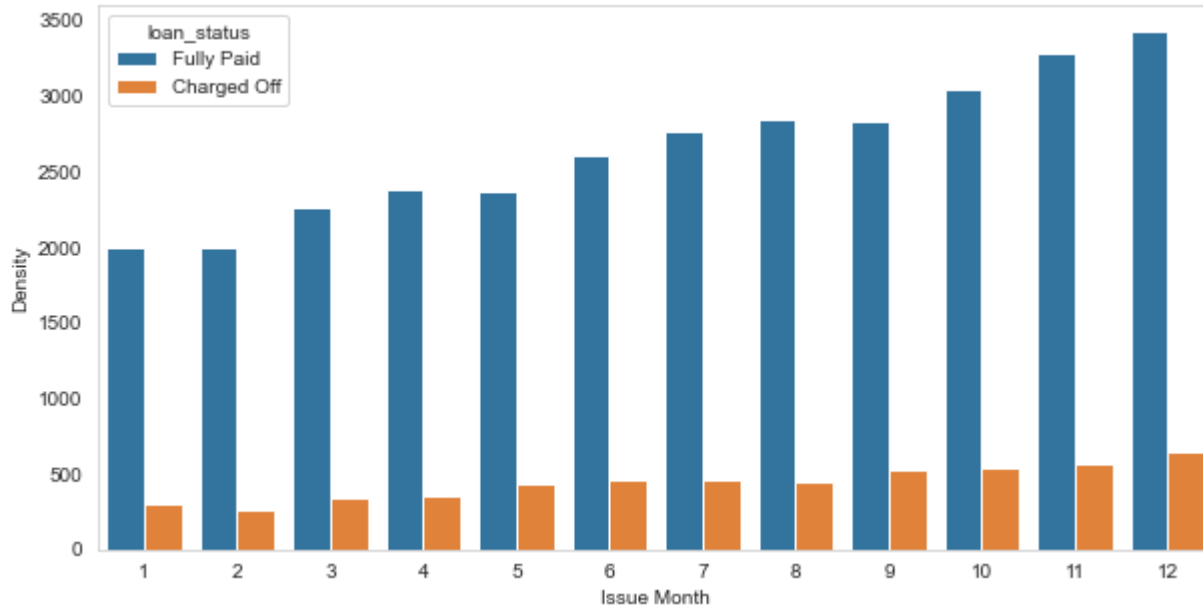
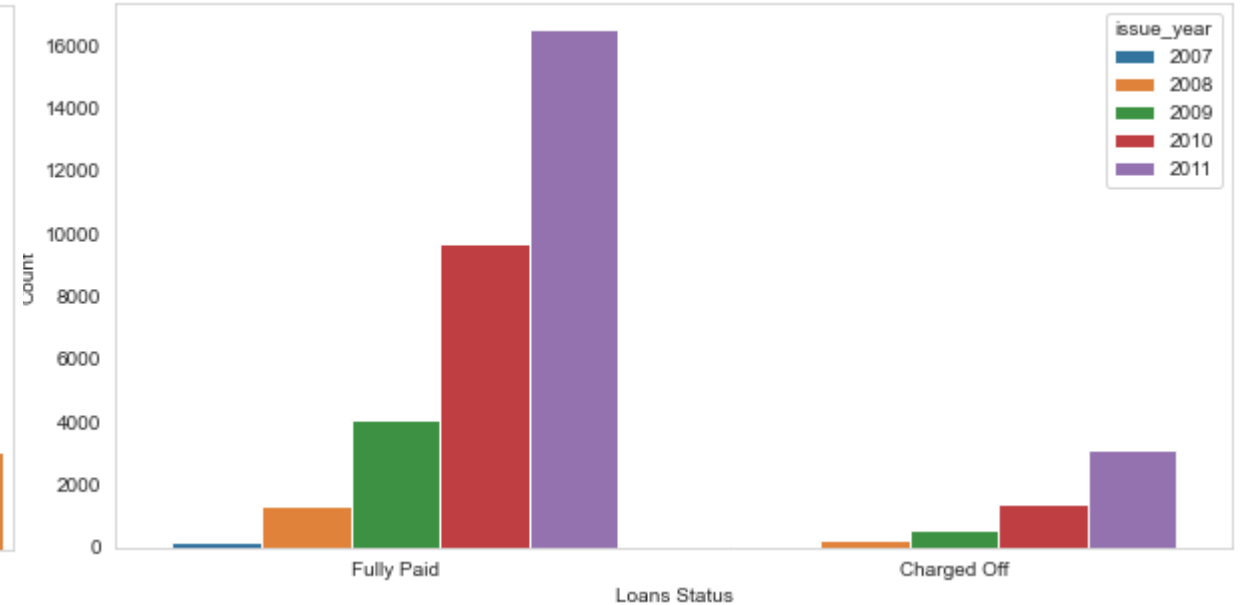■**Public Recorded Bankruptcy:** Majority of clients have no record of declaring bankruptcy.

Loan Trend over years

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

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

Location Based

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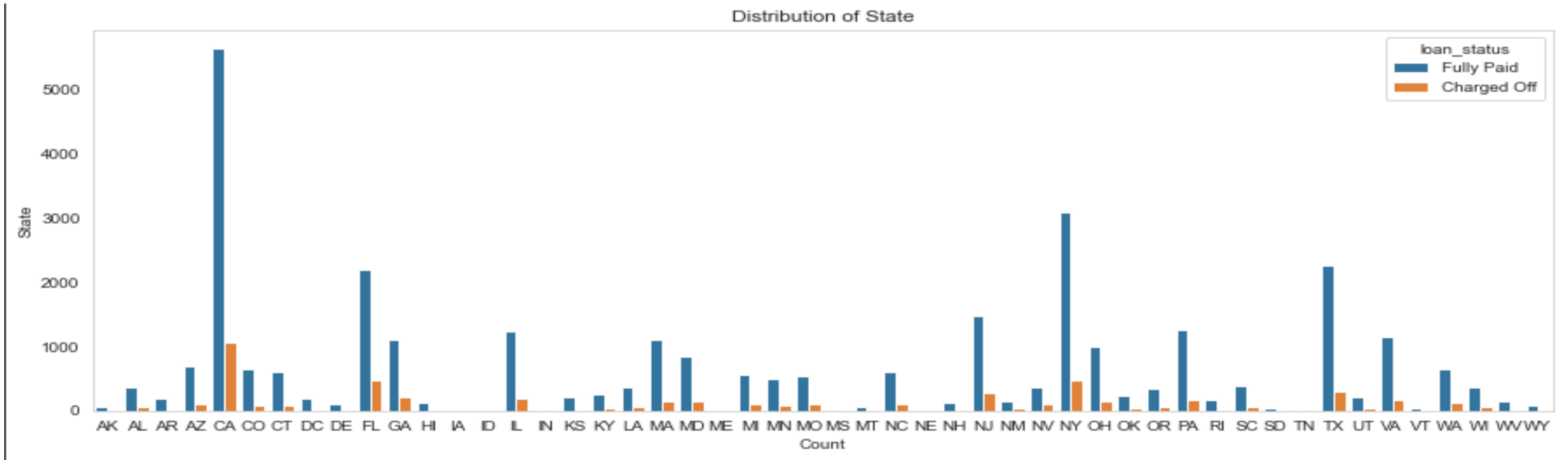
For large metropolitan cities we see large number of loans, with higher number of defaulted loans like California, New York, Texas, Florida but have a lower chance of defaulting.

Recommendations - Driving factors

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

1. DTI

2. Grades

3. Verification Status

4. Annual income

5. Public record bankruptcies

6. Interest rate

7. Purpose

Recommendations - Checkpoints for Defaulters

1. From large urban cities like California, New York, Texas, Florida etc.

2. Annual income in the range 50000-100000.

3. Borrowers having Public Recorded Bankruptcy.

4. Borrowers with least sub grades like E,F,G which indicates high risk.

5. Loan purpose is debt consolidation or credit card.

6. Borrowers having working experience of more than 10 years.

7. Borrowers with Debt-to-Income ratio greater than 13

8. Interest rate higher than 12

9. Paid interest exceeding paid principle

10. Revolving Utilization is more than 50%

11. Loan repayment duration is less than 3 years