Lending Club Case Study

Vijay Bhaskar

Problem Statement

Lending Club, a marketplace specializing in consumer finance and offering various loans to urban customers, is confronted with a crucial challenge in managing its loan approval process.

The company faces the task of making sound decisions to minimize financial losses, particularly from loans provided to applicants classified as "risky.“

Credit losses, which are the main concern, occur when borrowers fail to repay their loans or default. In simpler terms, borrowers identified as "charged-off" are responsible for the most significant losses to the company.

Objective

The main goal here is to help Lending Club reduce credit losses. This challenge is underscored by two potential scenarios:

2.1. It is essential to identify applicants who are likely to repay their loans as they can bring profits to the company through interest payments. Turning away such applicants could mean losing out on potential business.

2.2. Conversely, approving loans for applicants who are not likely to repay and are at risk of default could result in substantial financial losses for the company.

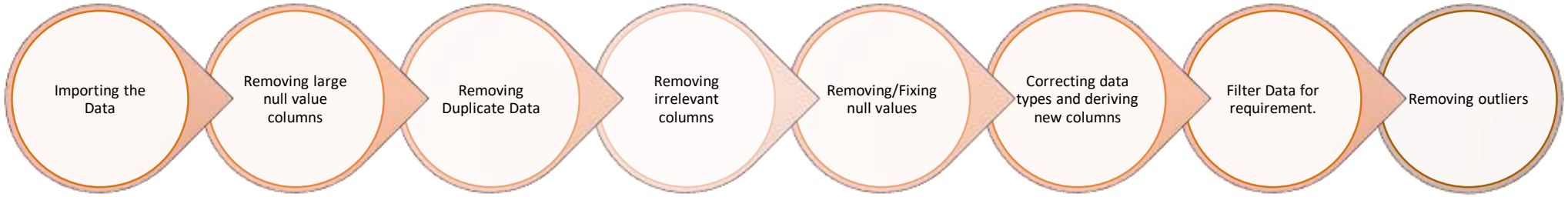
Business Understanding

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

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:



Variables

Following variables are indentified for univariate analysis

1. Categorical variables

   1.Ordered categorical data

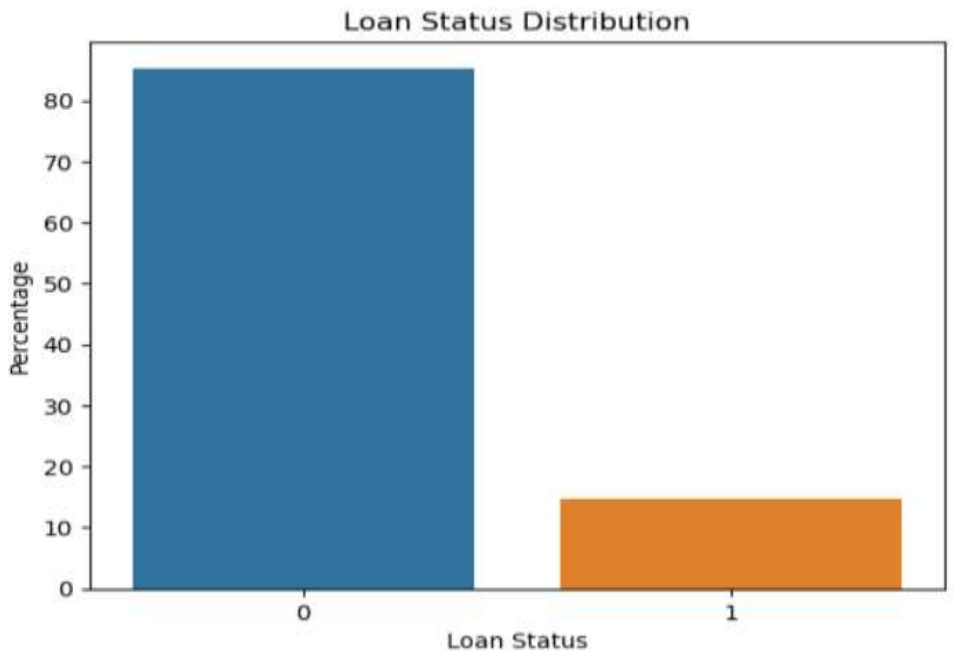
      1. grade
      2. sub_grade
      3. term
      4. emp_length
      5. issue_d_year
      6. issue_d_month
      7. issue_month_bin

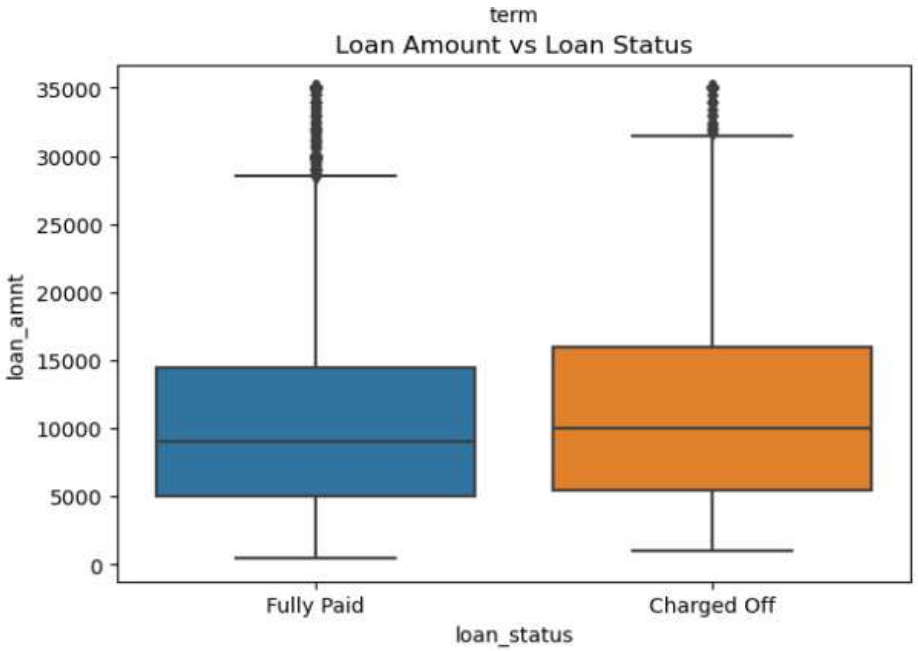
   2. Unordered categorical data
      1. addr_state
      2. purpose
      3. home_ownership
      4. loan_status
      5. loan_status_flag
      6. verification_status

2. Quantitative and Derived Variables

      1. approved_loan_amount_percent
      2. loan_amnt_bin
      3. annual_inc_bin
      4. int_rate_bin
      5. dti_bin
      6. installmets

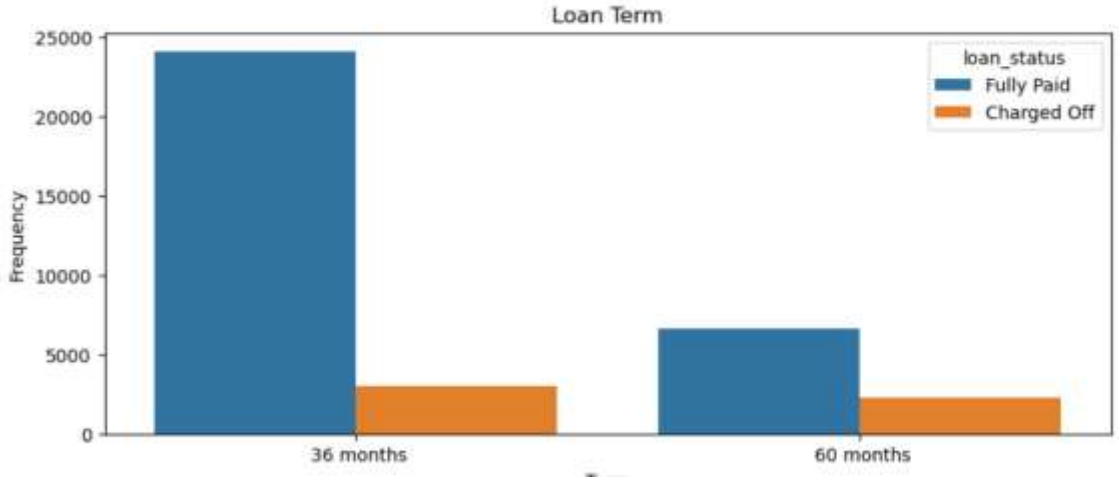
Loan Status and Amount

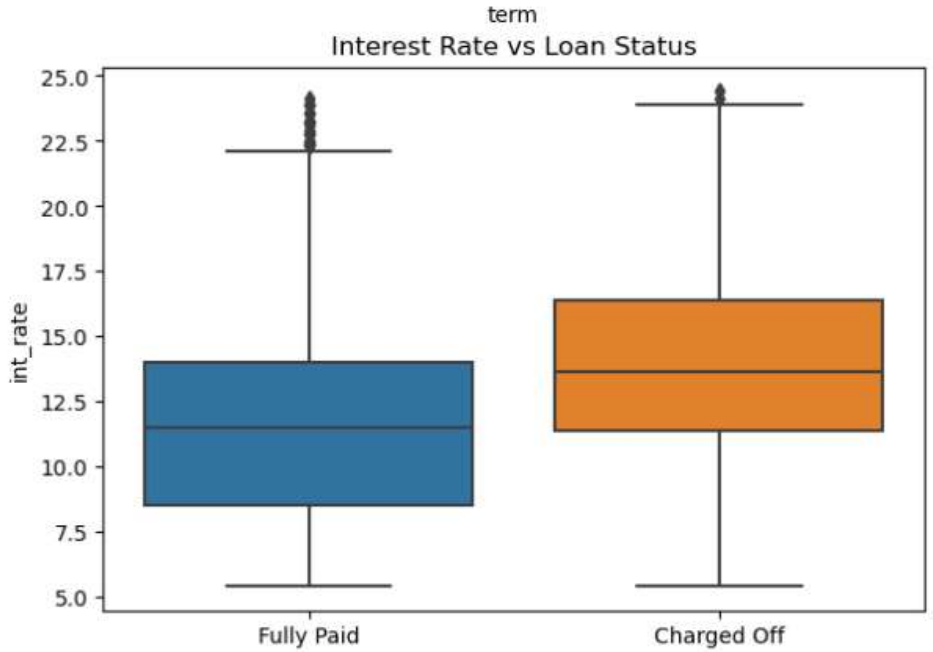




- **Loan Status:** The number of charged off loan is much smaller(14.3%) compared to total count.
- **Loan Amount:** It varies from 500 to 35000 with a median of 10000. Loan amount is majorly small and very few clients have taken large loans and larger it goes we have higher chance of defaulting.

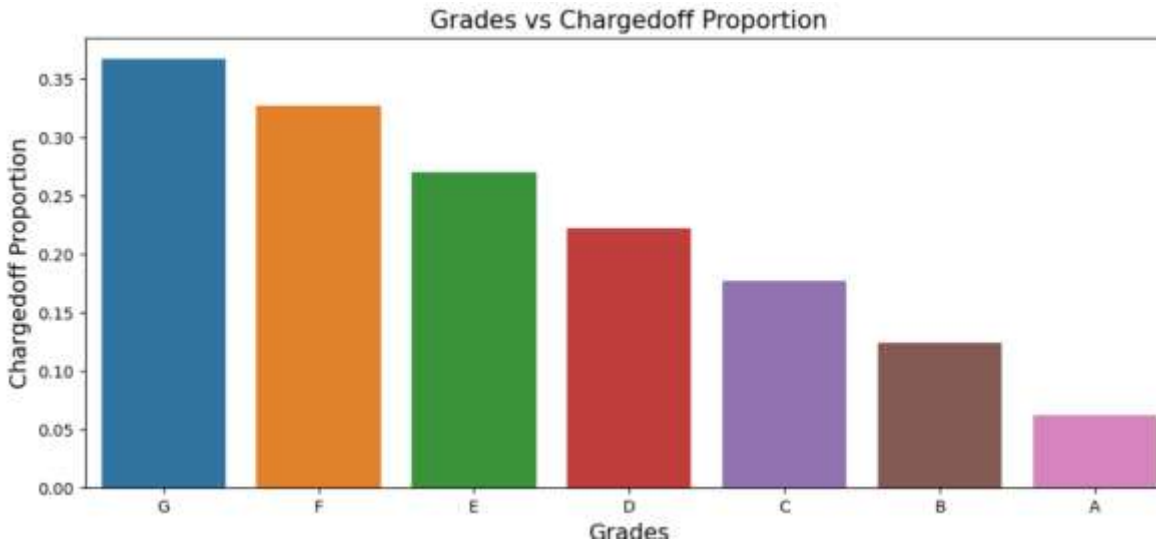
Term and Interest Rate

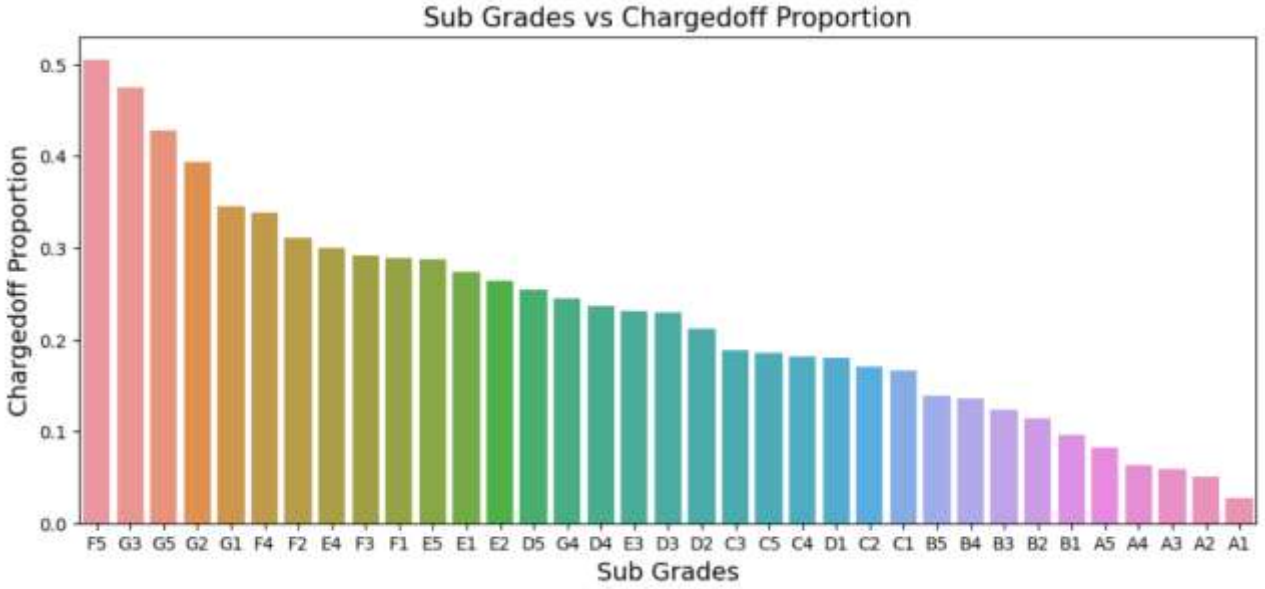




- **Loan Term:** The Loans taken for 36 month term are much more than 60 months and have lower chance of defaulting.
- **Interest Rate:** The count of loan taken varies with interest rate showing peak around in 5-15 bracket and decreasing slowly where as the chance of defaulting increases with interest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

Grade and Sub-Grade

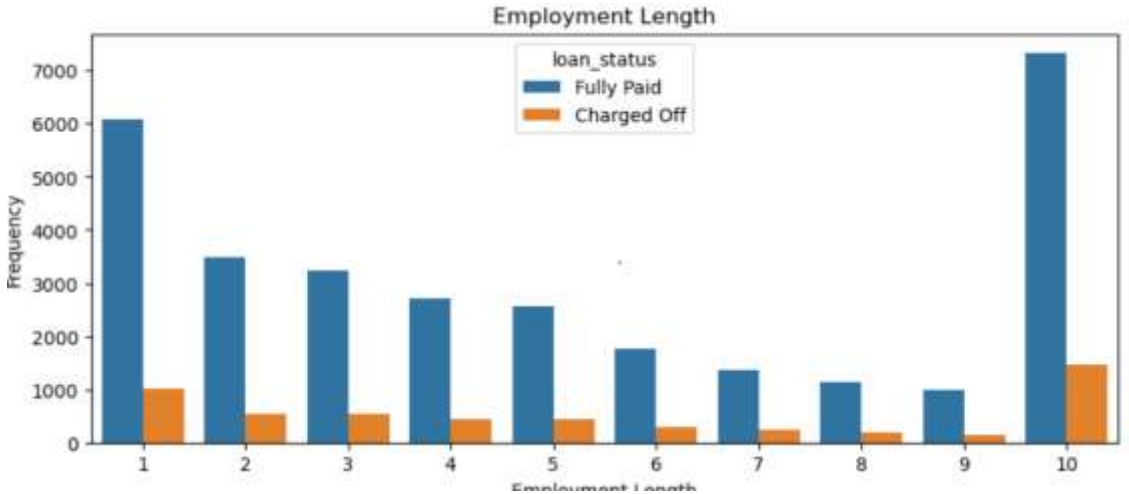


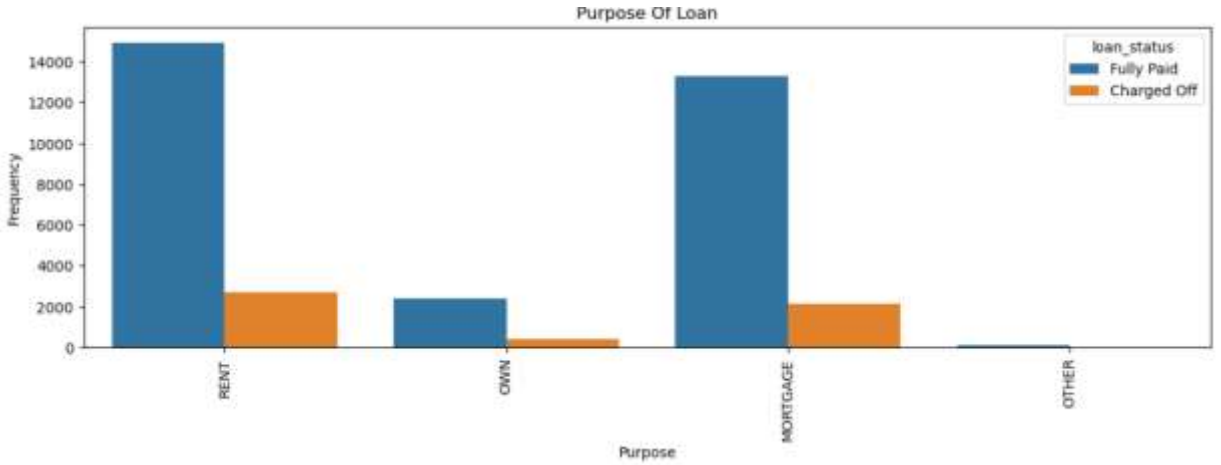


- **Grade:** The loan approved are majorly of higher grade as they are of low risk thus low chance of defaulting. 60 month term loans have larger number of lower grade loans with high risk.

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

Employment Length & Homeownership

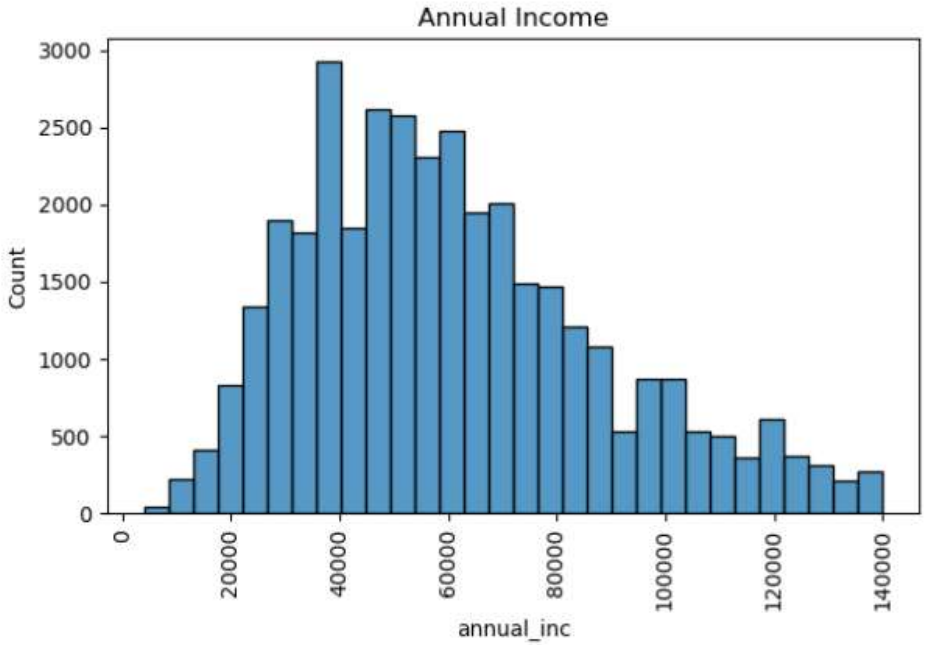


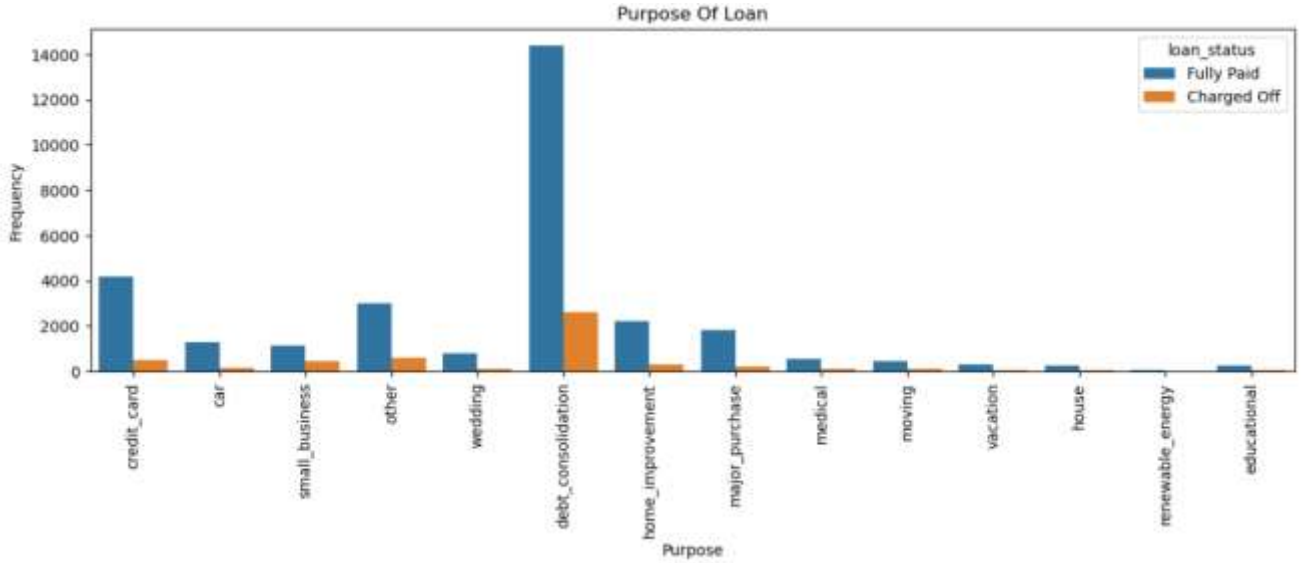


- **Employment Length:** Majority of clients have 10+ years of experience and has highest number of defaulted loan.

- **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

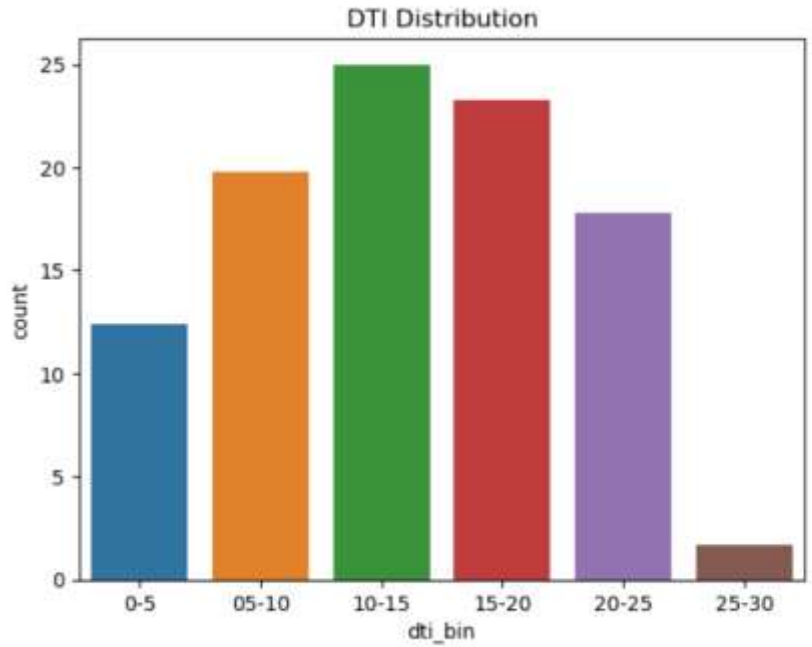


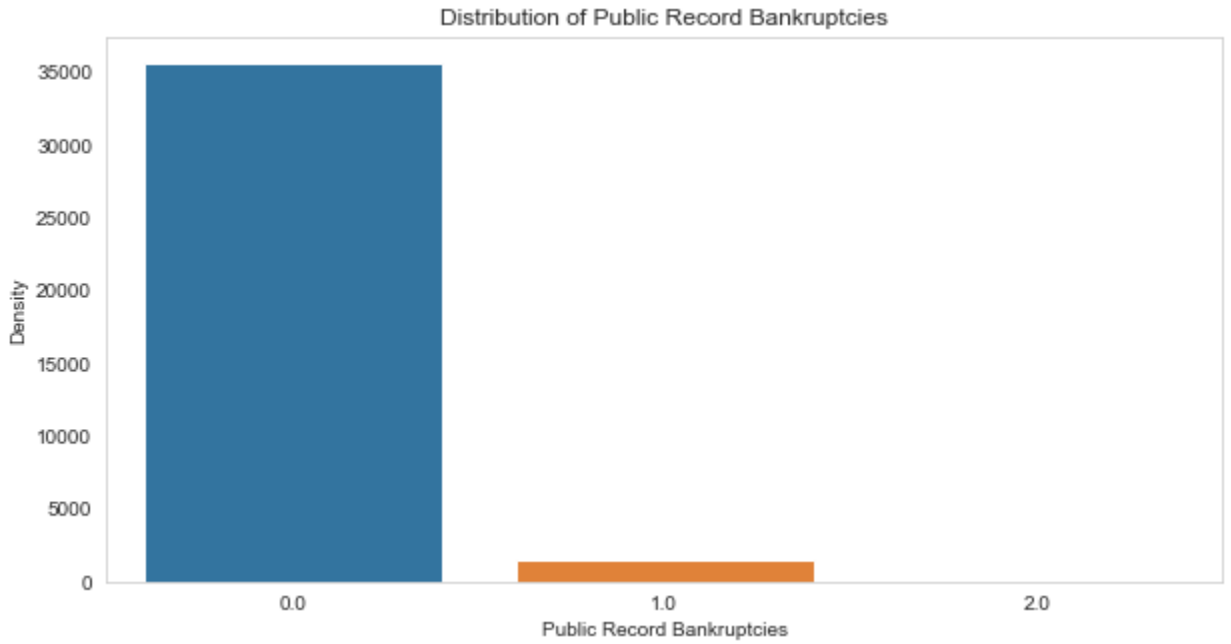


- **Annual Income :** The Majority of clients have low annual income compared to rest and income lower than 50k has higher chance of defaulting.

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

DTI ratio & Bankruptcy

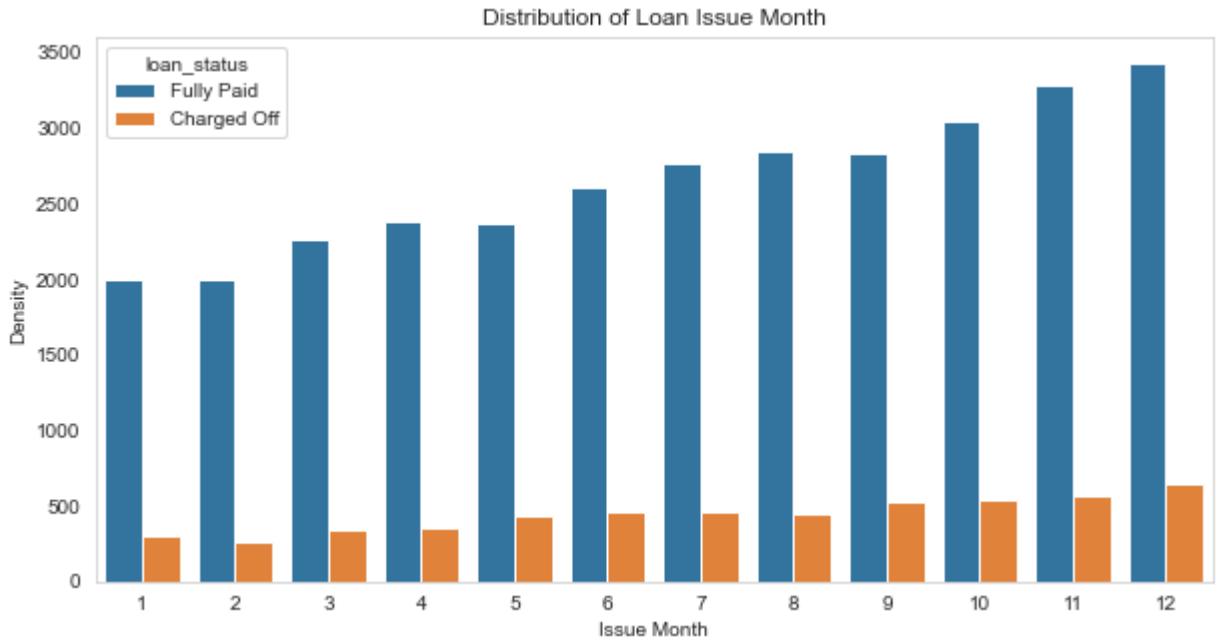


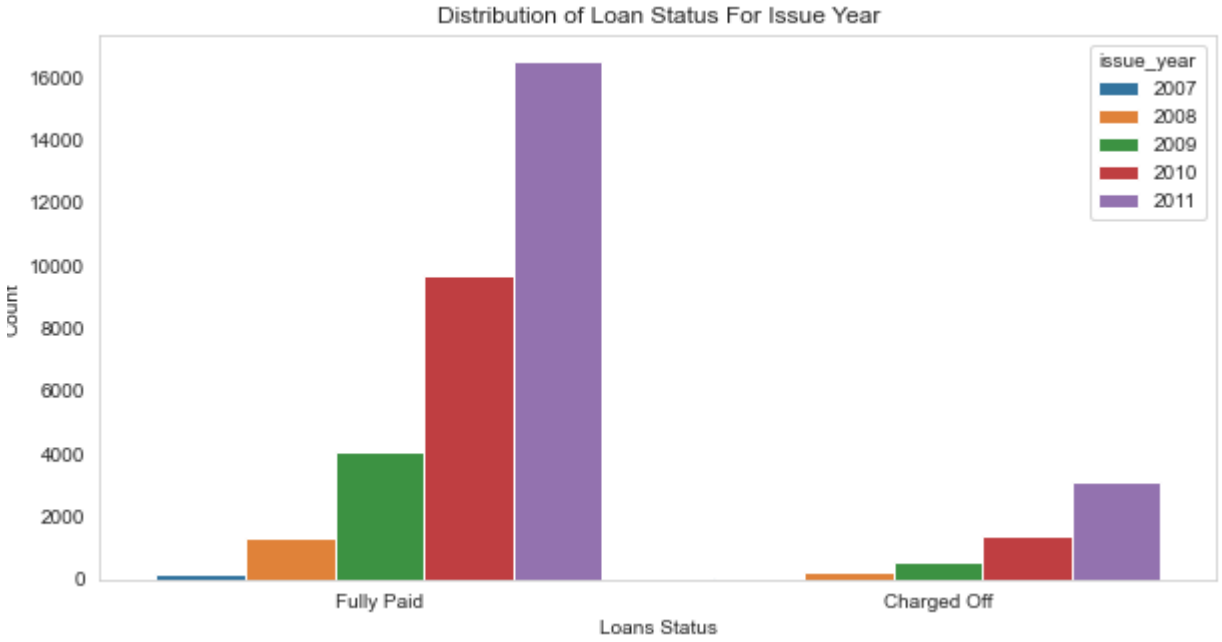


- **DTI:** The large percentage of Clients have a large Debt to Income ratio which shows that lending to such clients can be very risky.

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

Loan Trend over years

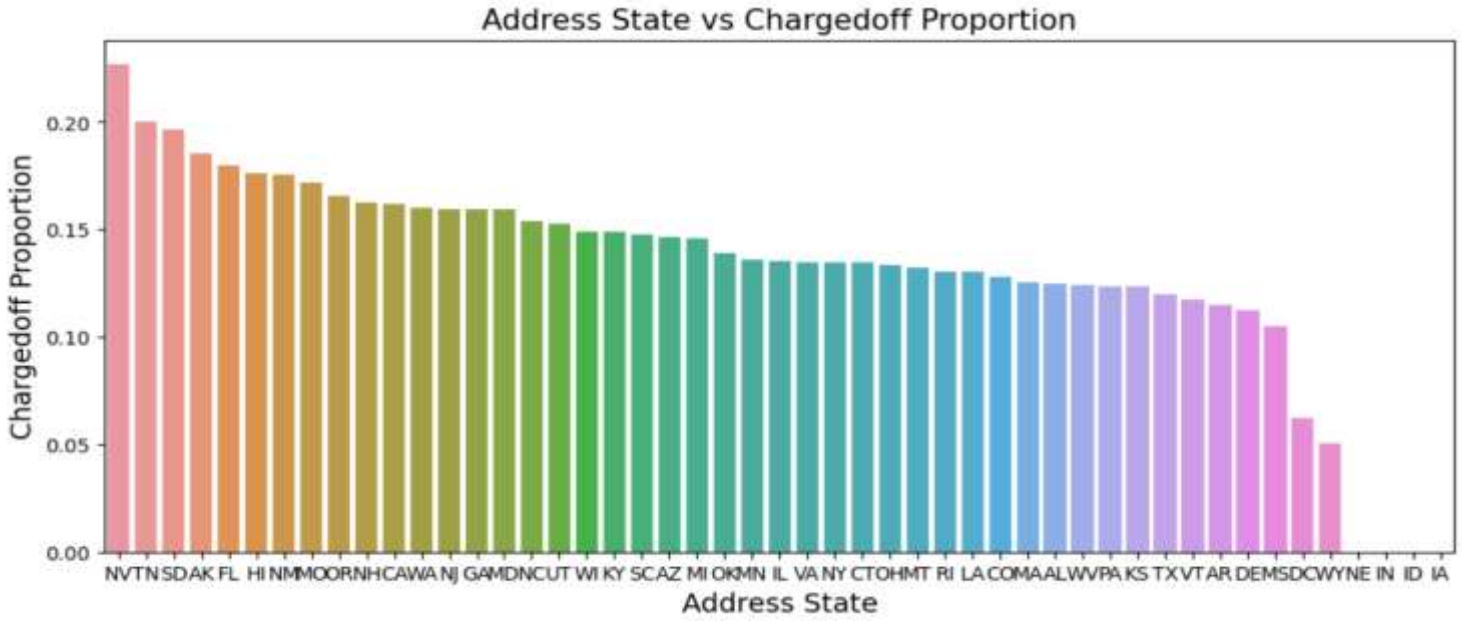




We see a gradual increase in loan taken through the year, with lesser defaulting rate in April ,August, December quarter wise and better more late in year.

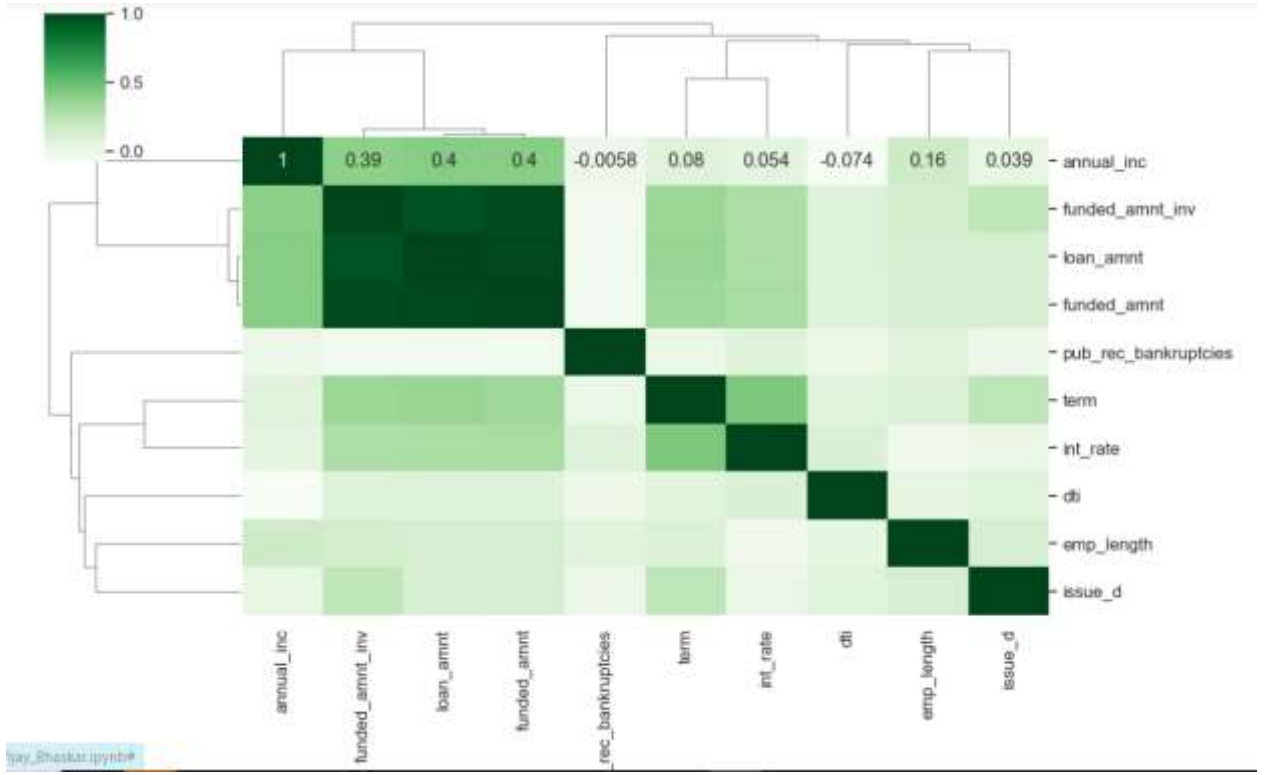
With each passing year loan taken are increasing exponentially which indicate we are seeing large increase in DTI ratio and decrease in defaulting rate.

Location Based



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.

Heat Map


1.0
0.5
0.0
1
0.39
0.4
0.4
-0.0058
0.08
0.054
-0.074
0.16
0.039
annual_inc
funded_amnt_inv
loan_amnt
funded_amnt
pub_rec_bankruptcies
term
int_rate
dti
emp_length
issue_d

Inference from Heat Map

Strong Correlation

- `installment` has a strong correlation with funded_amnt, loan_amnt, and funded_amnt_inv
- `term` has a strong correlation with interest rate
- `annual_inc` has a strong correlation with loan_amount

Weak Correlation

- `dti` has weak correlation with most of the fields
- `emp_length` has weak correlation with most of the fields

Negative Correlation

- `pub_rec_bankrupticies` has a negative correlation with almost every field
- `annual_inc` has a negative correlation with dti

Recommendations

1. Borrowers are more likely to default when they take loans for a 60-month term.
2. Borrowers with a 'Verified' loan status tend to take higher loan amounts with a 60-month tenure, increasing the risk of default.
3. Borrowers who rent their homes and take loans for debt consolidation purposes have a higher probability of default.
4. Borrowers with an annual income in the range of 0 to 20000 are more prone to defaulting.
5. Borrowers who take loan amounts in the range of 0 to 14000 have a higher default rate.
6. Borrowers receiving interest rates between 15-20% are at a higher risk of default.
7. Borrowers taking loans for small businesses are more likely to default.
8. Borrowers with lower grade (e.g., F < G) are at an increased risk of default.
9. Borrowers with subgrades F5, G3, or G5 are more likely to default.