Lending Club Case Study

Deepak Bhadoria

Ramandeep Mehra

Objective

Objective of this case study is implement Exploratory Data Analysis (EDA) techniques on loan data from a consumer finance company to get an understanding of risk analytics in banking and financial services and also to understand how data is used to minimize the risk of loosing money while lending to the customers.

By this case study we also improve our understanding of visualizations and how different charts can be used to understand the data better.

Business objective of this case study is to help company understand the driving factors behind loan default. The company wants to utilize this knowledge for its portfolio and risk assessment.

Identification of risky loan applicants using EDA is the aim of this case study.

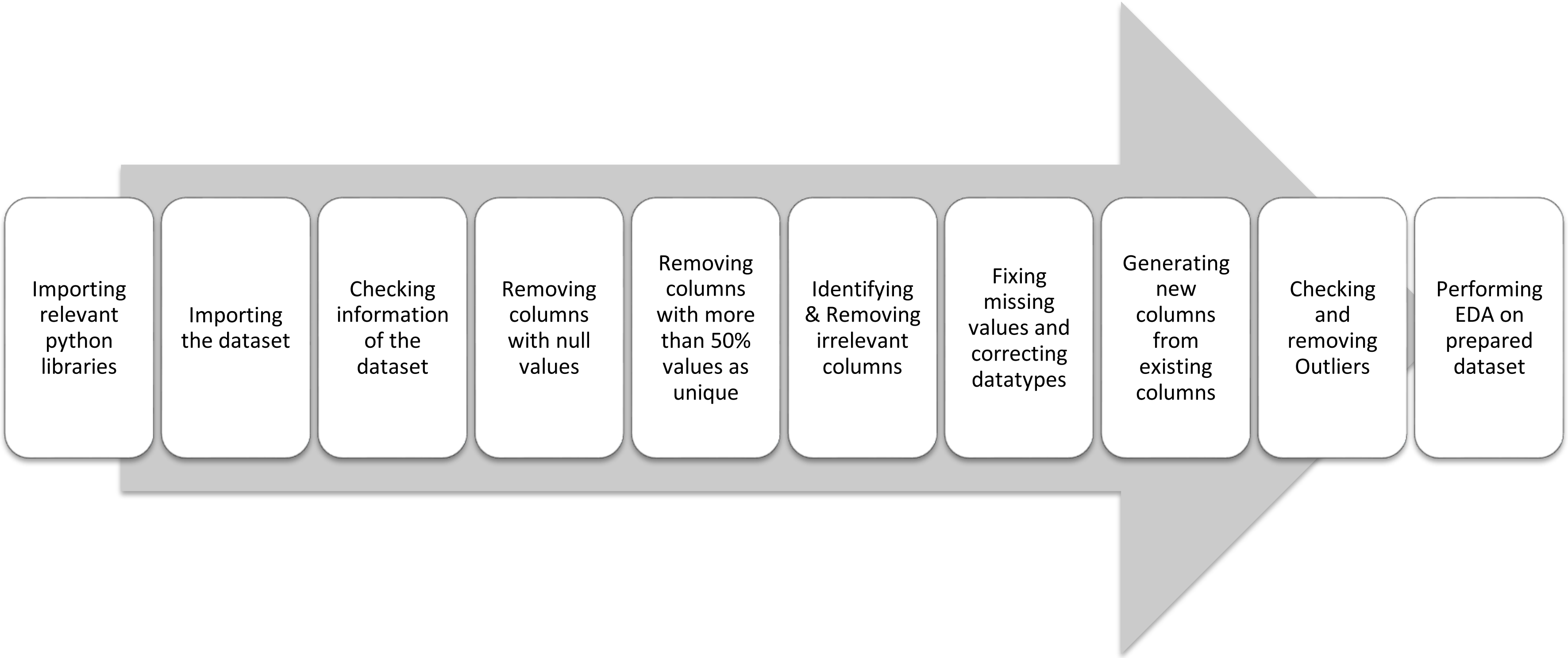
Business Understanding

When the company receives a loan application, the company has to make a decision for loan approval based on the applicant’s profile. Two types of risks are associated with the bank’s decision:

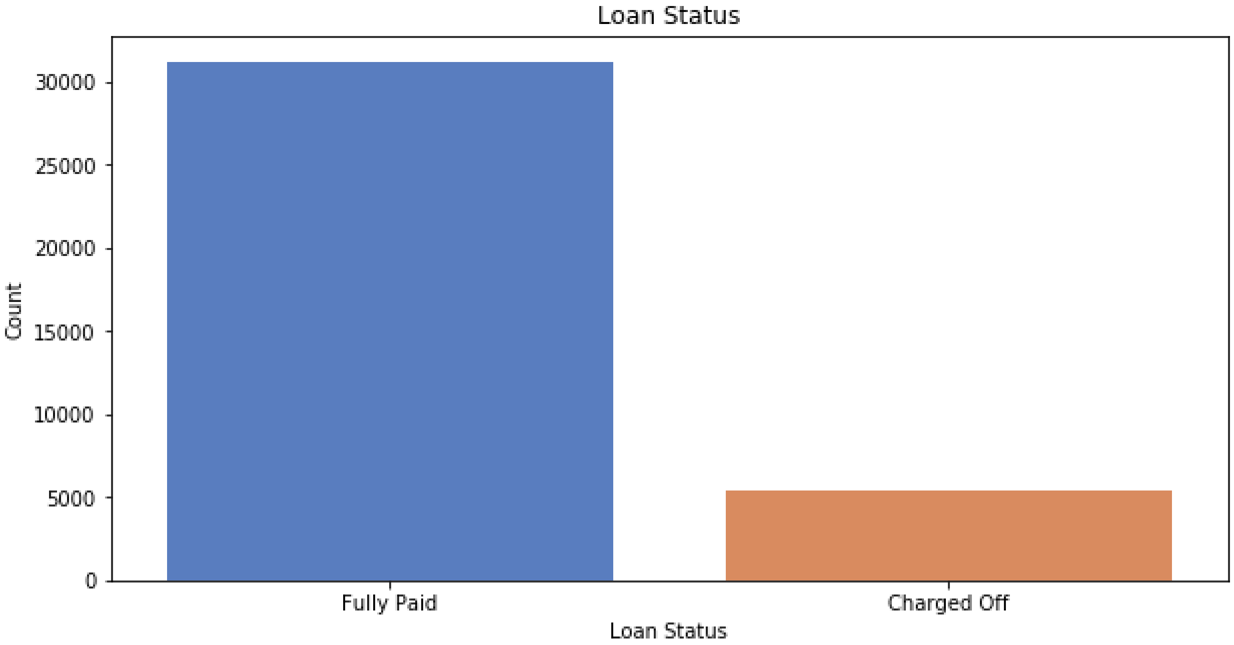
- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

In this case study, we use EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of default.

Data Cleanup and Preparation Steps/Process



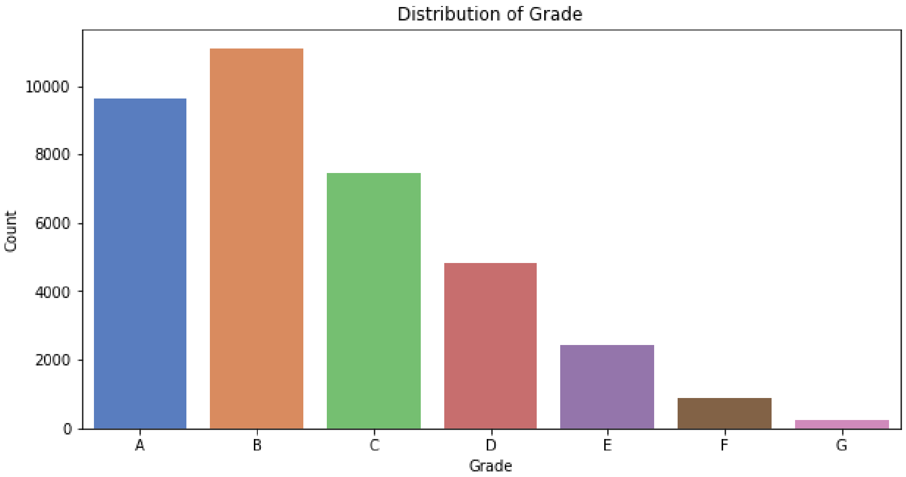
Loan Status

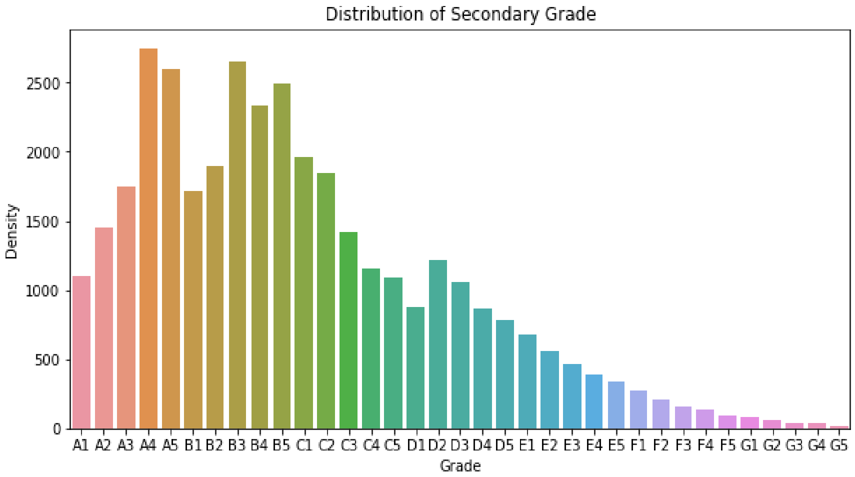


Observation:

- The count of defaulted Loans (Charged Off) is much small as compared to the Fully Paid loans.

Loan Distribution against grades & sub-grades





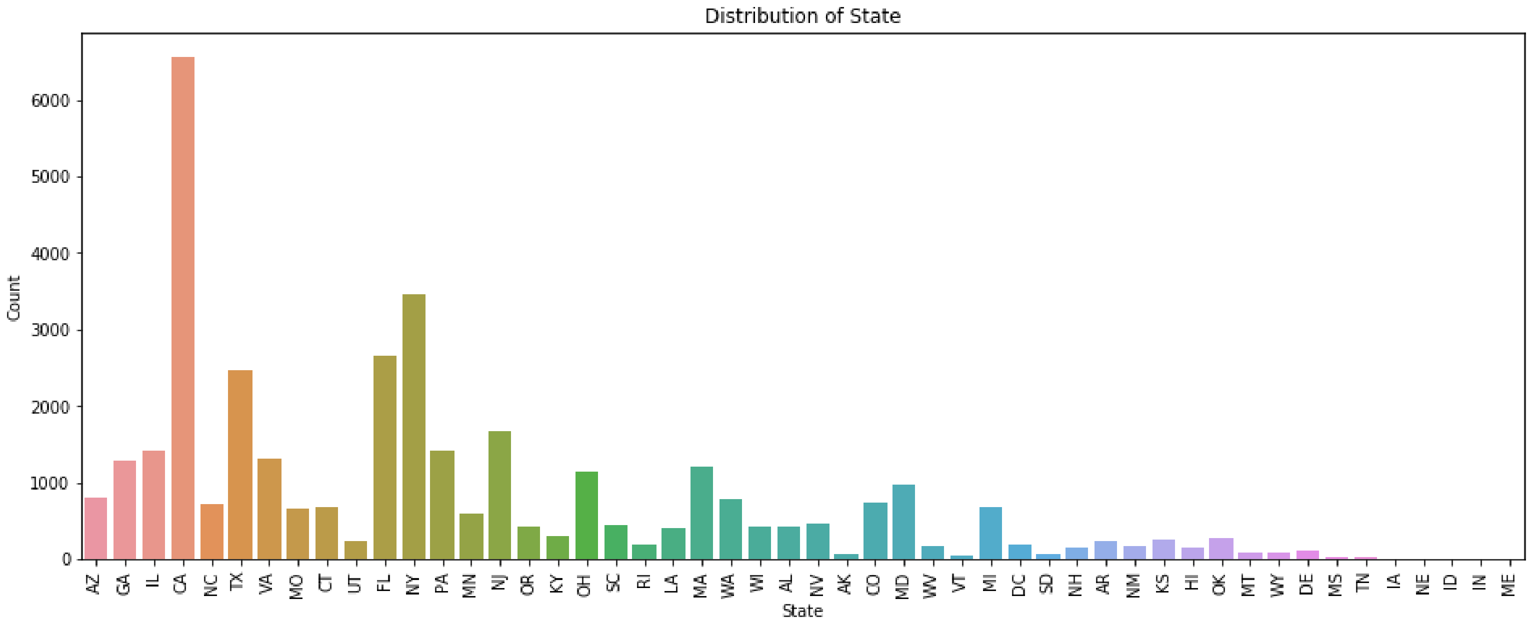
Observation:

- A large amount of loans are with **category A & B** as compared to other grades, concluding that most loans are high grade loans

Observation:

- A large amount of loans are with lower sub-grades in loan grade A & B showing most loans are high grade loans, concluding that sub grades are in sync with loan grade inferring most loans being high grade loans

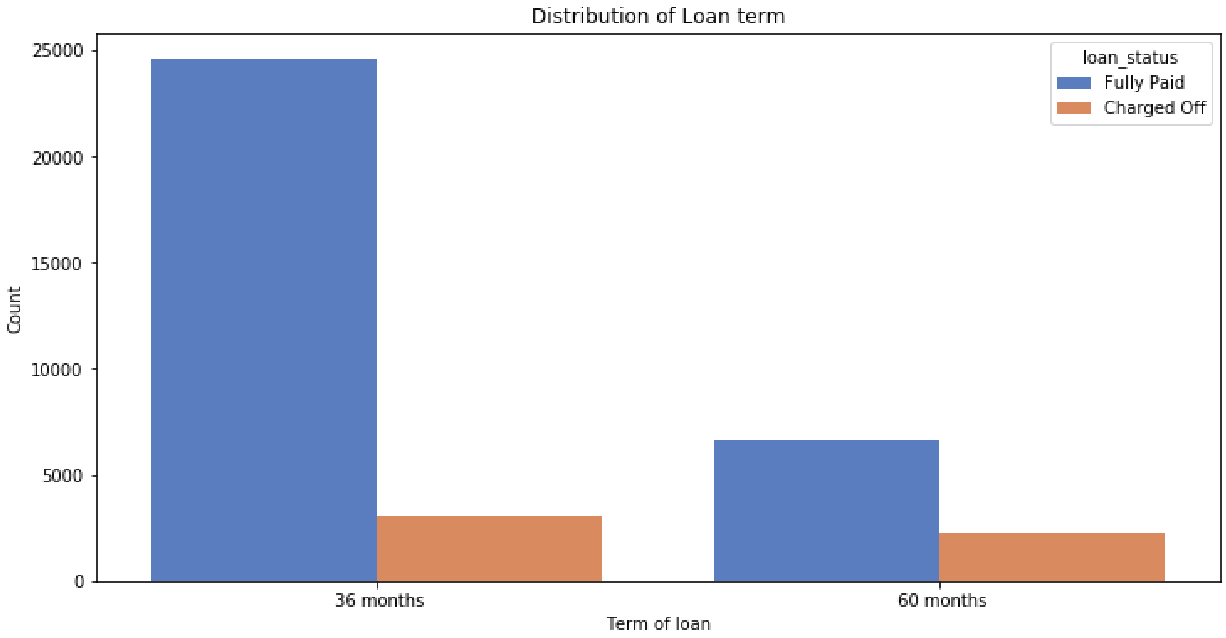
State wise Loan Distribution



Observation:

- Looking at the distribution of loans along with states, we can conclude that majority of the borrowers are from big urban states like California, New York, Florida, Texas.

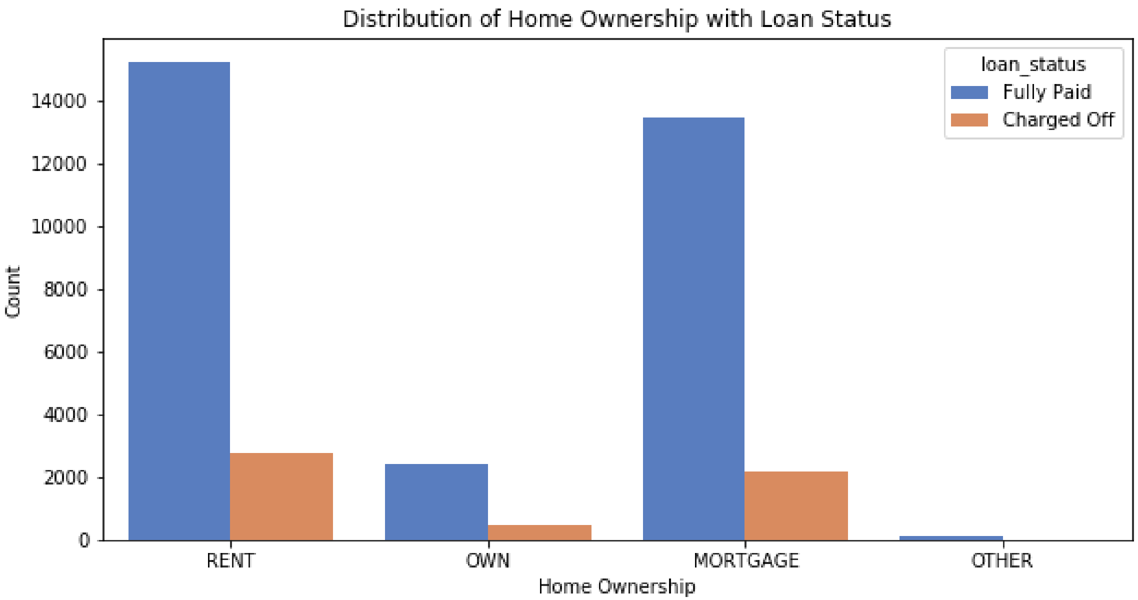
Loan Term Distribution



Observation:

- Majority of the loans are approved for 36 months period

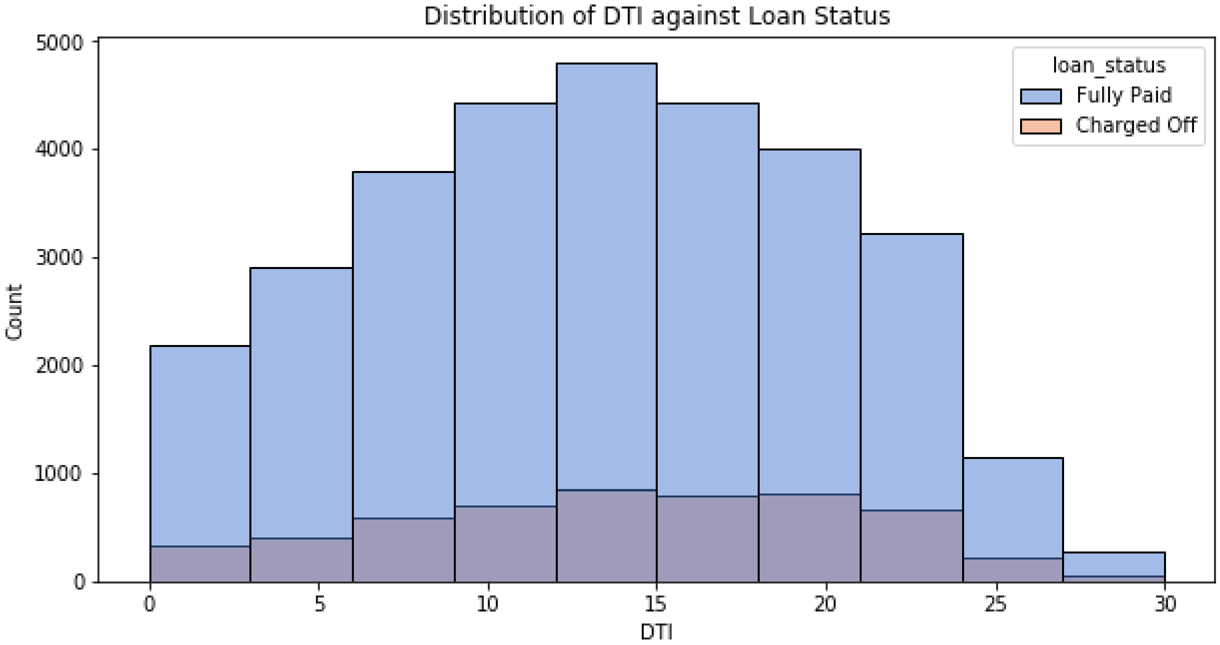
Home Ownership vs Loan Status



Observation:

- Most of the borrowers does not possess any property and are on either "RENT" or "MORTGAGE"
- The % of loan defaulters are more if the borrower is already having an ongoing "MORTGAGE"

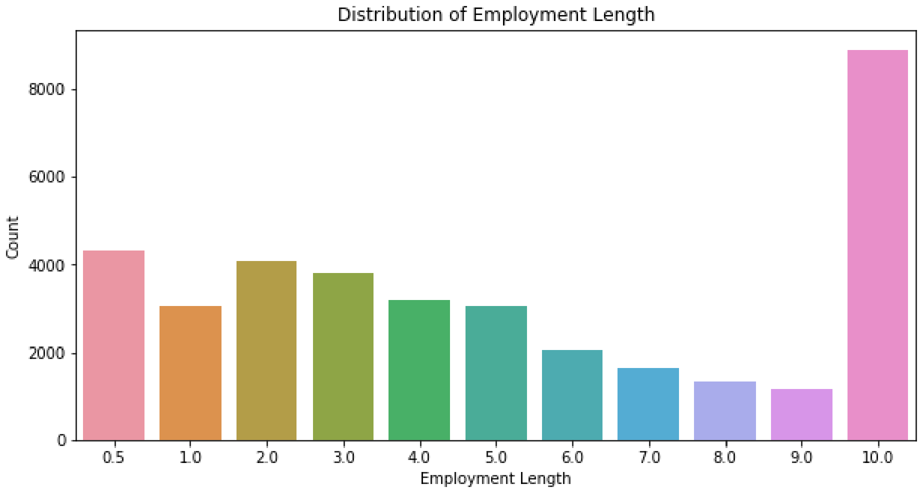
dti ratio vs Loan Status

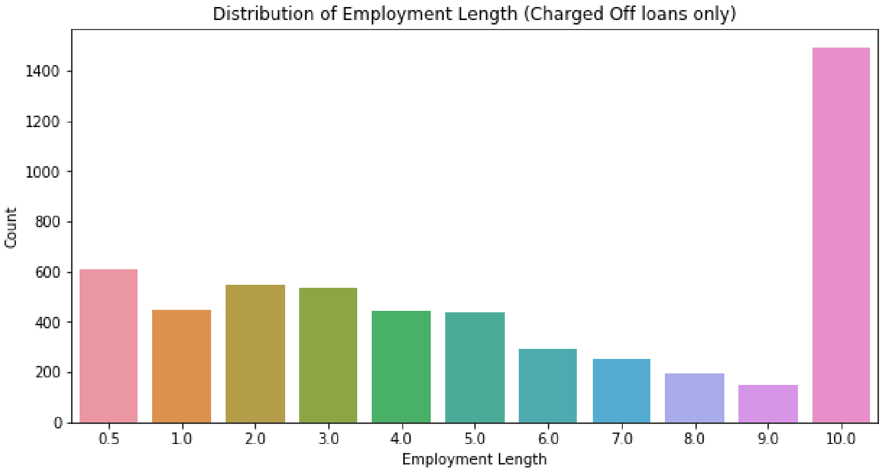


Observation:

- Looking at the distribution we can conclude that for majority of the borrowers, the debt to income ratio is very high, majorly concentrated at the range of 10-20 dti ratio.

Applicants Employment Length

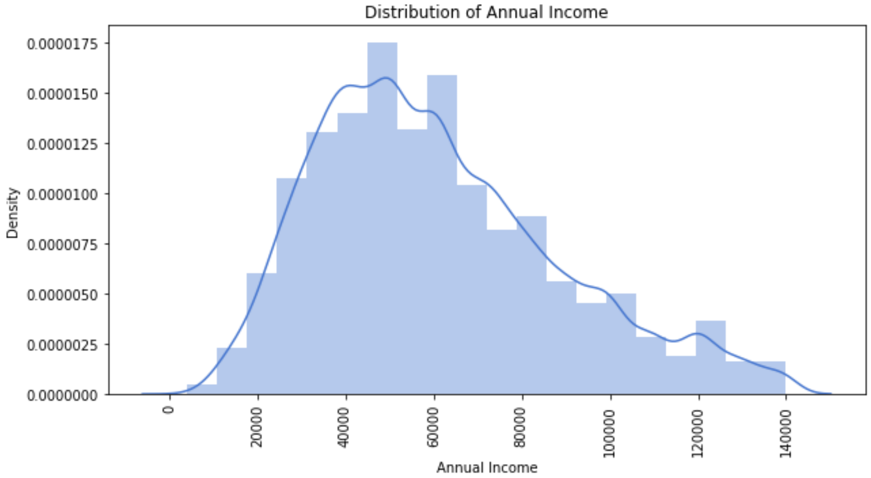


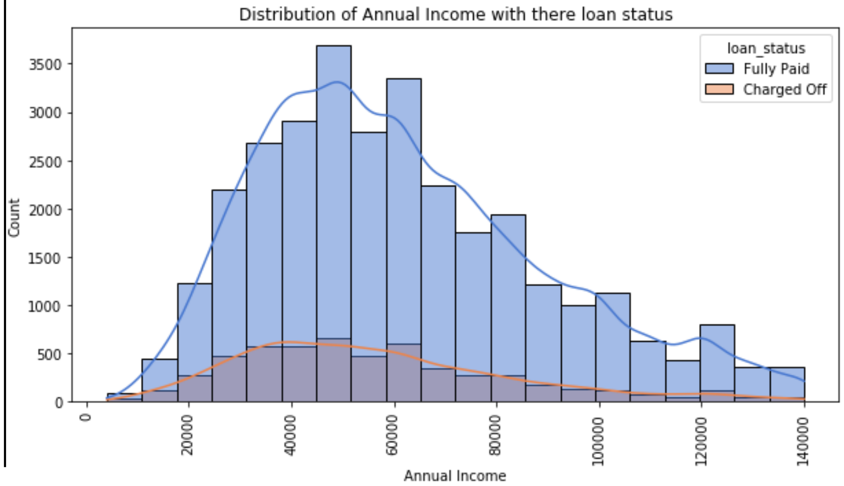


Observation:

- Most of the borrowers are having experience of 10+ years
- Most of the defaulters are having experience of 10+ years

Annual Income of Borrowers & there loan status

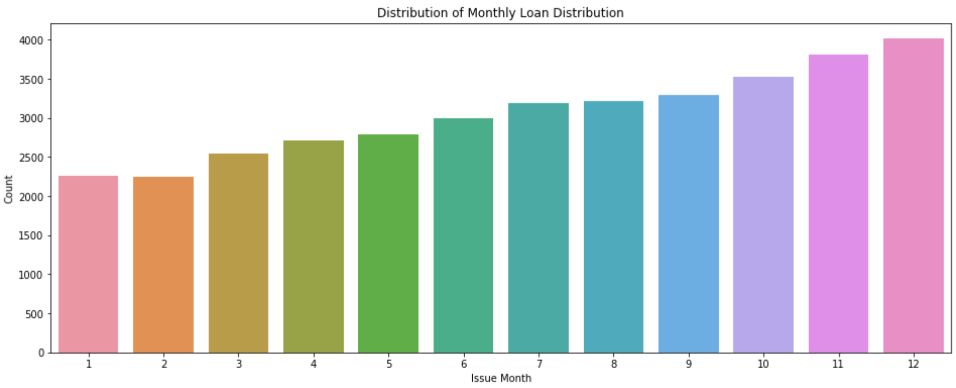




Observation:

- The distribution for Annual Income looks like a normal distribution but it is left skewed. So we can say that most of the borrowers have low annual income.
- Looking at the right hand chart: We can conclude that the borrowers with less than 50000 as the annual income are most likely to default the loan.

Loan Trends – Month of the year/Yearly



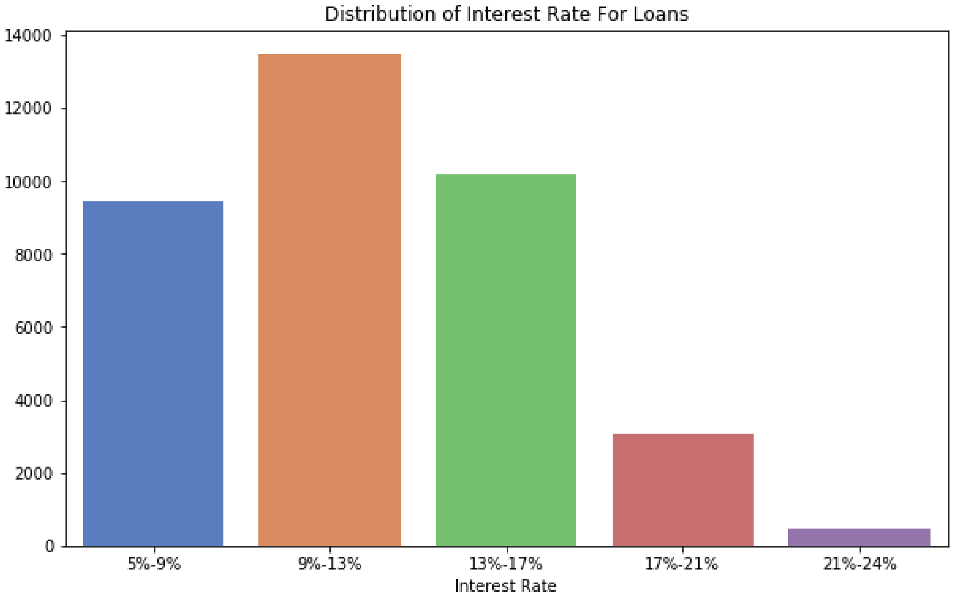
Observation:

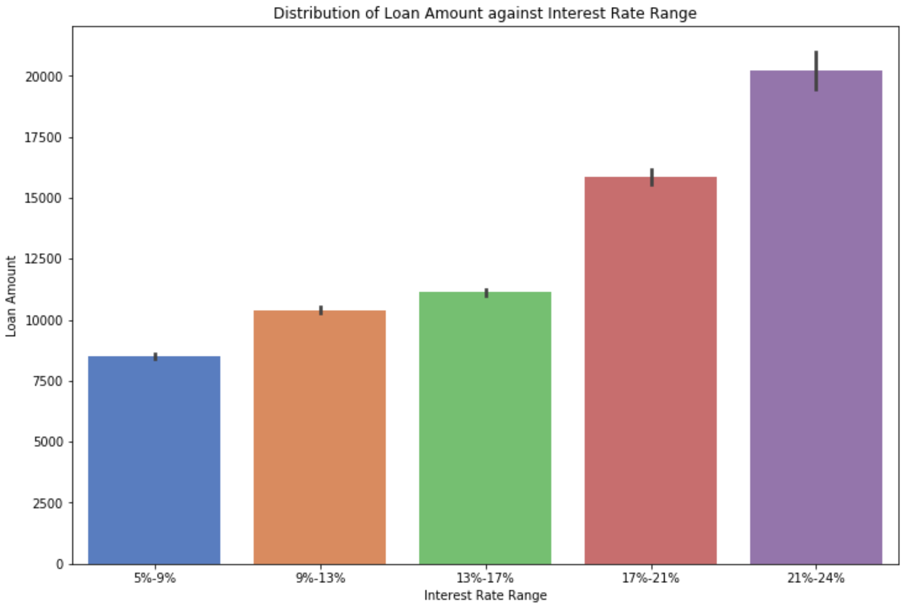
- Looking at the distribution we can conclude that majority of the loans are processed in the last quarter of the year
- As the loans processed in last quarter of the year are more, the defaulters are also more from the last quarter of the year

Observation:

- Looking at the distribution we can conclude that the loan approval numbers increases positively year on year and the increase with time is exponential
- With the increase in number of loans approvals, the loan defaulters have also increased positively year on year

Interest Rate Range and Loan Amount Distribution

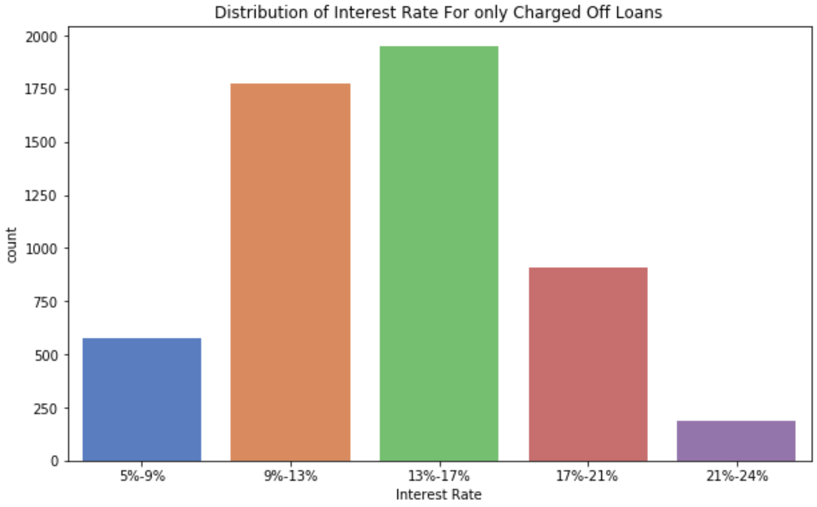


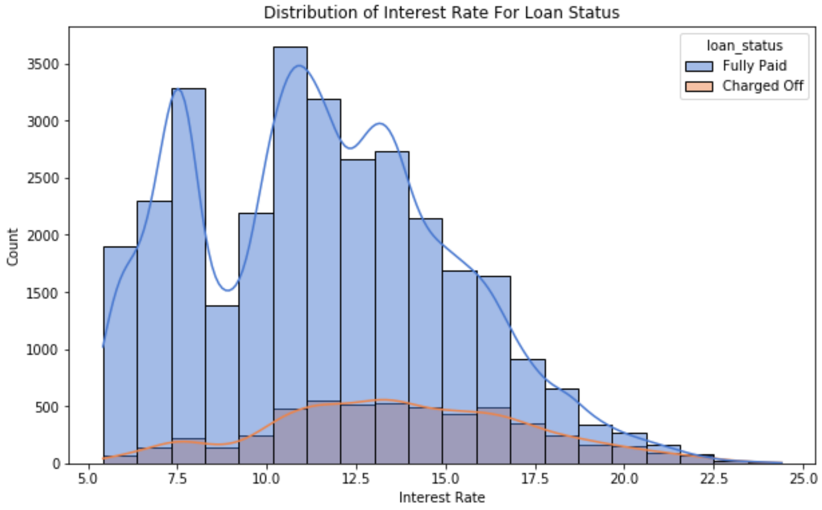


Observation:

- For most of the loans approved, the interest rate is between 9%-13%.
- Interest Rate increases as the value of loan amount increases.

Interest Rate

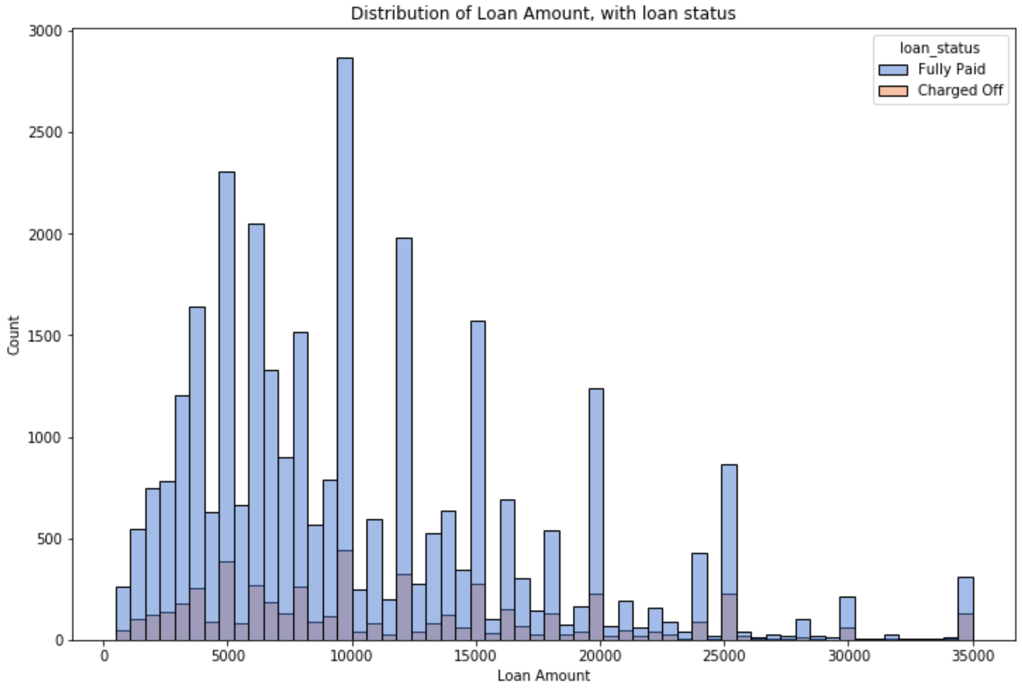




Observation:

- For the borrowers where the interest rate is between 13%-17% - the defaulters % is more as compared to borrowers with other interest rate percentage
- The rate of defaulters decreases and continues to decrease if the interest rate is more than 17%

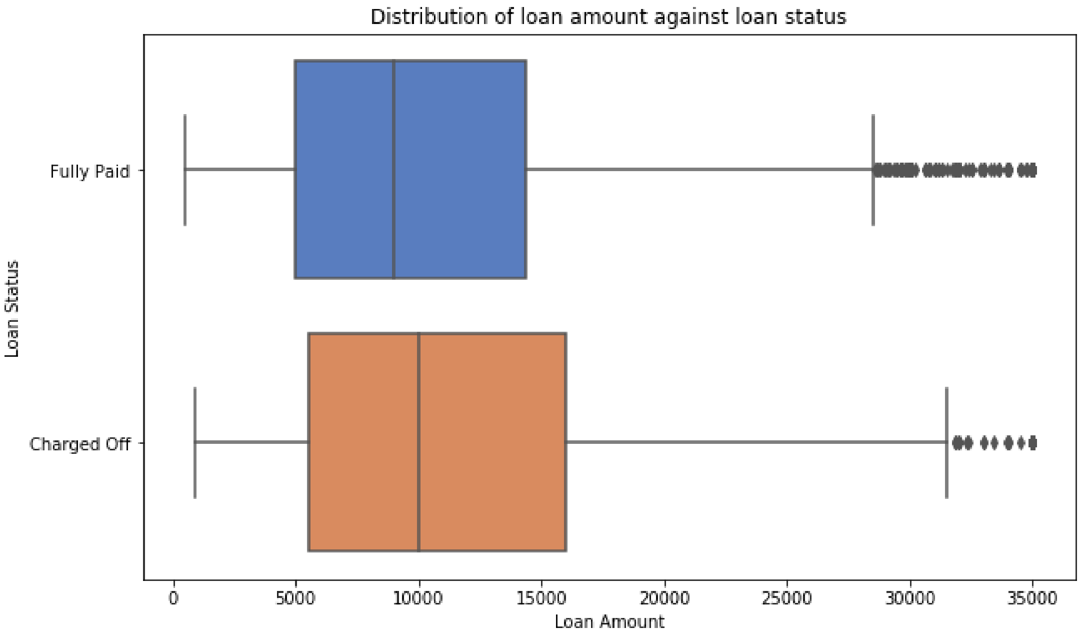
Loan Amount against Loan Status



Observation:

- Defaulter percentage is more when the loan amount is more than 10k

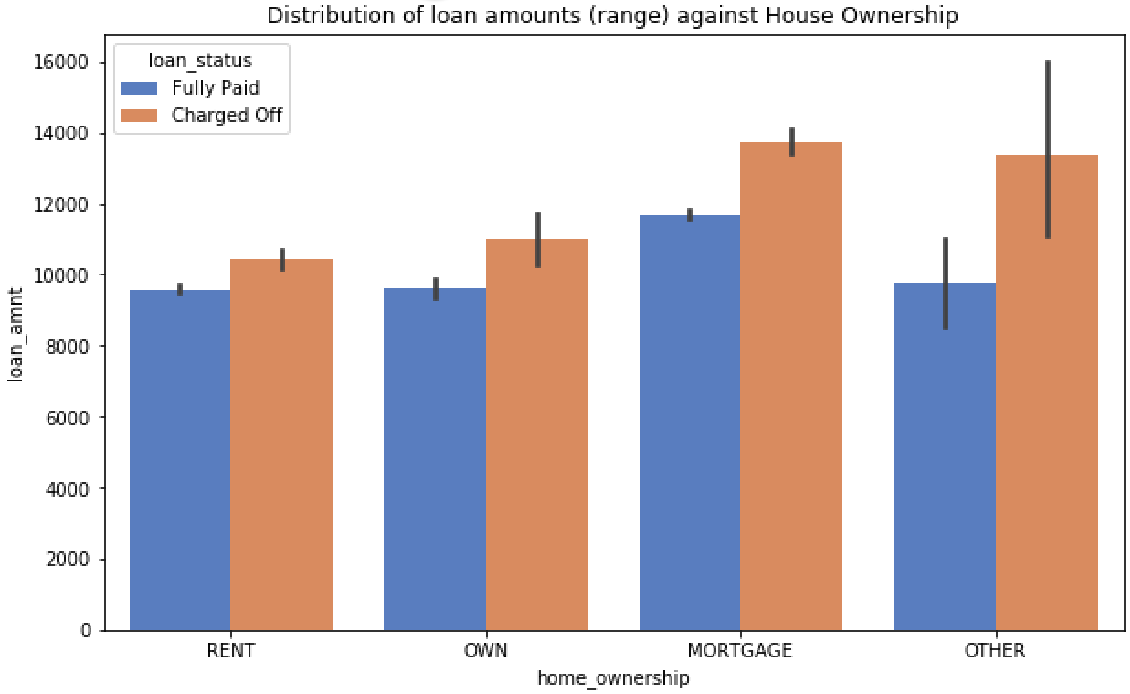
Loan Amount against Loan Status



Observation:

- The mean and 25th quantile is same for both Fully Paid and Charged Off loans
- We see a larger 75% quantile in "Charged Off" loans - indicating large amount of loans have high chance of defaulting.

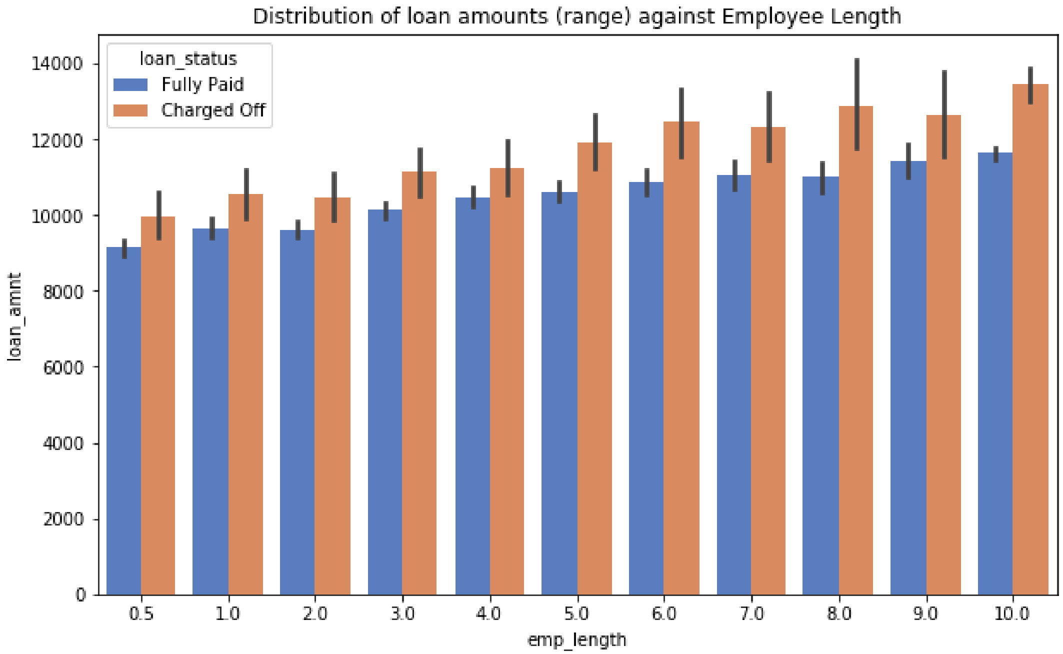
Loan Amount against House Ownership



Observation:

- Applicants with an existing MORTGAGE have higher rate for defaulting at the loan payments.

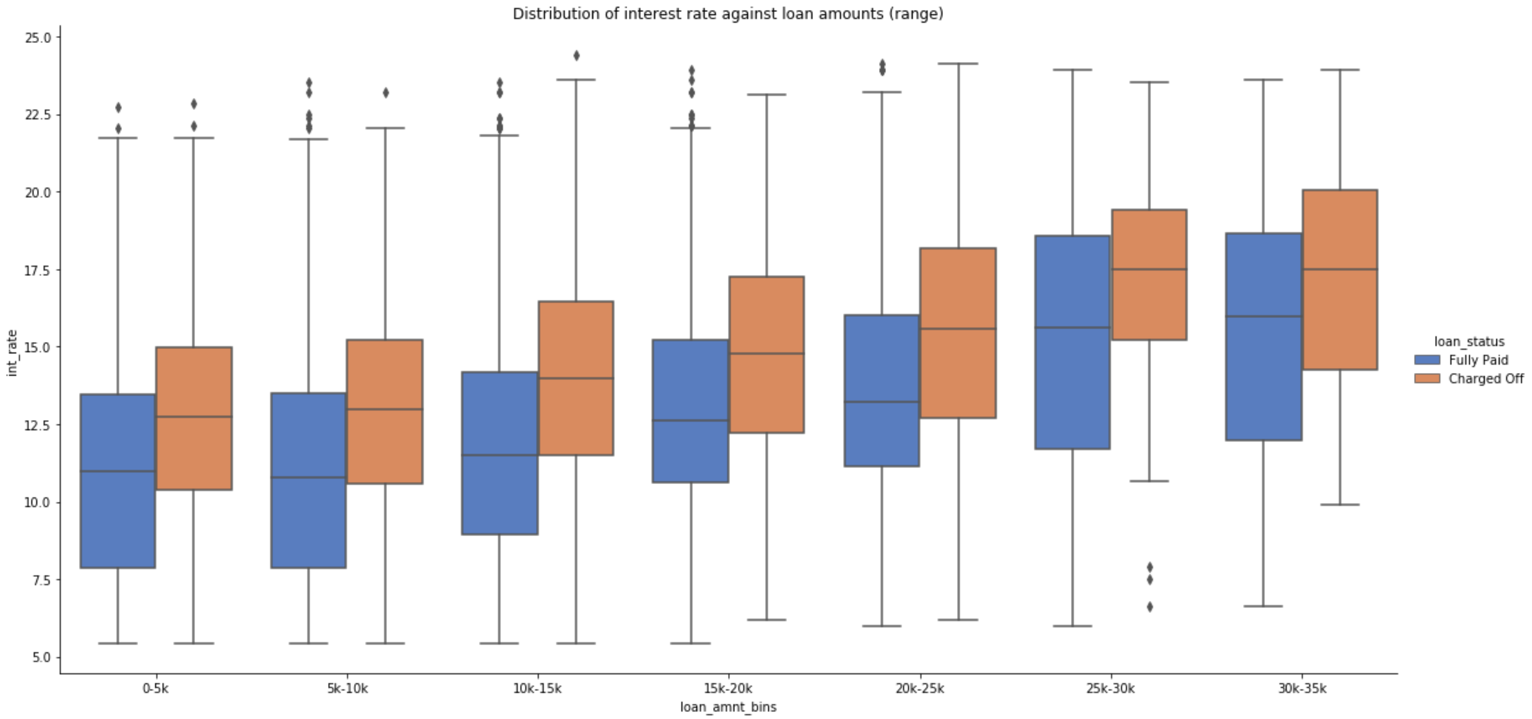
Loan Amount against House Ownership



Observation:

- For Applicants with employment length of 10+ years, chances of defaulting at the loan payment is more when loan amount is more than 12k

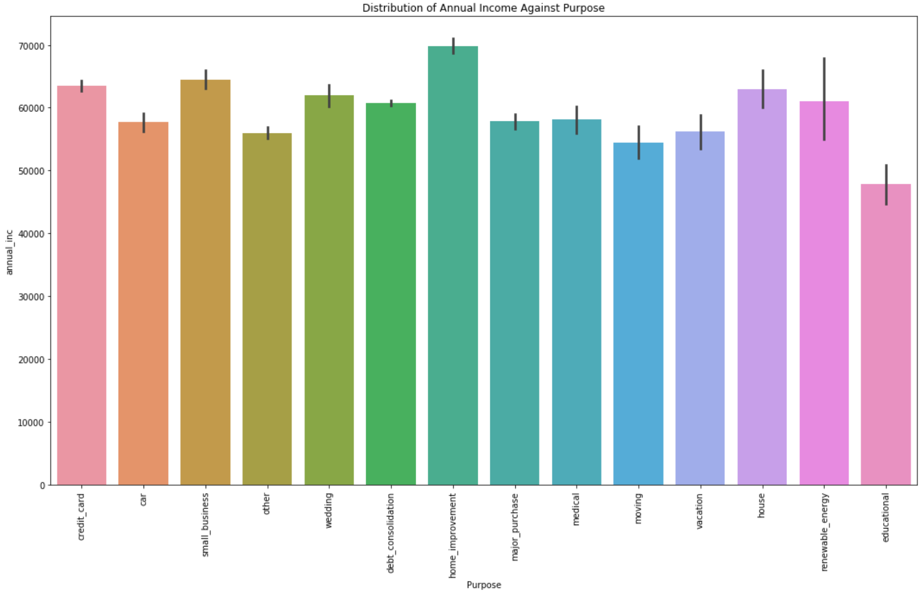
Interest Rate Distribution for loans

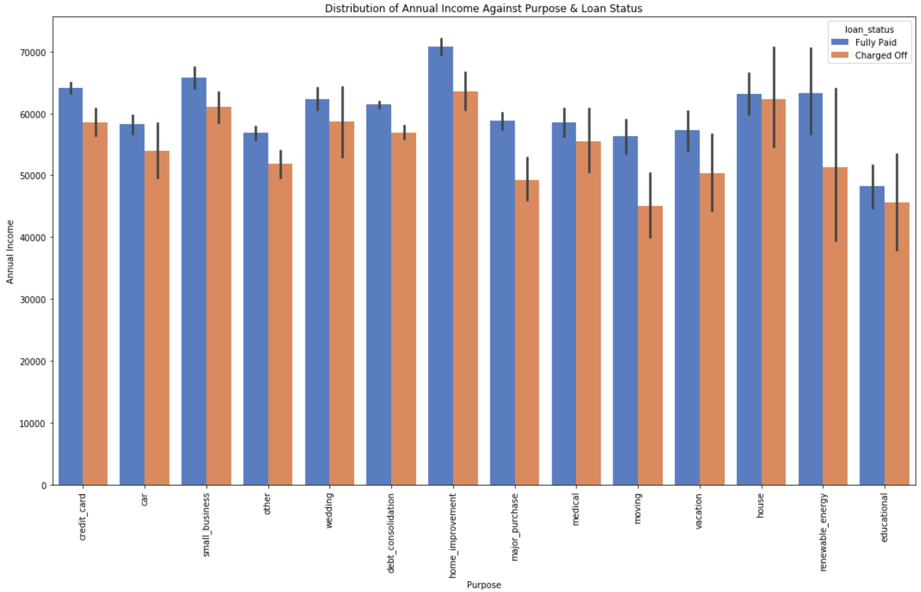


Observation:

- Interest rate for charged off loans is pretty high as compared to fully paid loans
- This can be a very strong factor for loan defaulting

Annual Income against Loan Purpose





Observation:

- The individuals with higher salary mostly applied loans for home improvement, small business and against credit cards
- We also see the highest number of loan defaulters against the debt consolidation, small business, credit card and home improvement category

Summary

Following are few observations with respect to the Charged Off loans.

There is more chances of defaulting when:

- Applicants who use the loan to clear other debts, start a small business, clear off credit card bills or home improvement.
- Applicants who receive interest at the rate of 13-17%.
- Applicants where the employee length is 10+ years.
- Applicants with income less than 50000.
- Applicants for whom the dti is ranging between 12 – 15.
- The Applicant is already having an ongoing "MORTGAGE"
- The loan amount is applied for a larger value (>10000)

Few more observations:

- The chances of defaulting increases for applicants:
  - Whose home ownership is "Mortgage" and loan value is in range of 14k - 16k
  - Employment length is 10+ years and loan amount is in range of 12k - 14k
  - When the interest is changed in the range of 15-17% and loan is in the range of 30k - 35k
- Interest rate for charged off loans is pretty high as compared to fully paid loans. This can be a strong factor for loan defaulting.