Lending Club case Study Submission

Group:
Shilpa Mathur

Problem Statement

You work for a consumer finance company which specialises in lending various types of loans to urban customers. 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:

1. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
2. 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

The data given below contains the information about past loan applicants and whether they 'defaulted' or not. The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

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

When a person applies for a loan, there are two types of decisions that could be taken by the company:

1. Loan accepted: If the company approves the loan, there are 3 possible scenarios described below:
   - Fully paid: Applicant has fully paid the loan (the principal and the interest rate)
   - Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
   - Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan
2. Loan rejected: The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset)

Objective:
The company wants to identify the driving factors which influences the loan applicants defaulting the loan so that the company can utilize this information for risk assessment and cutting down the credit loss using EDA
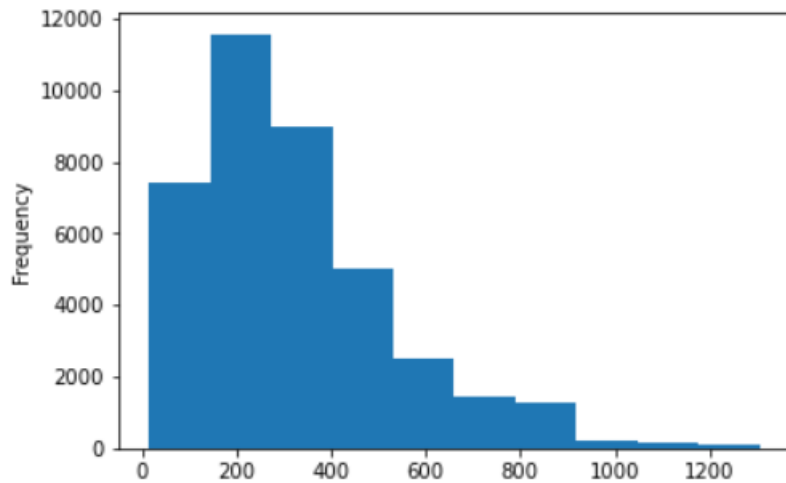
Exploratory Data Analysis

- Data Cleansing or Data pre-processing should be done.

- Do Univariate Analysis on the data.

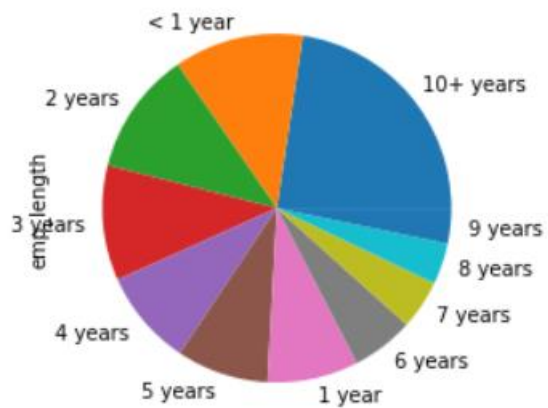- Next do Bivariate Analysis on the data.
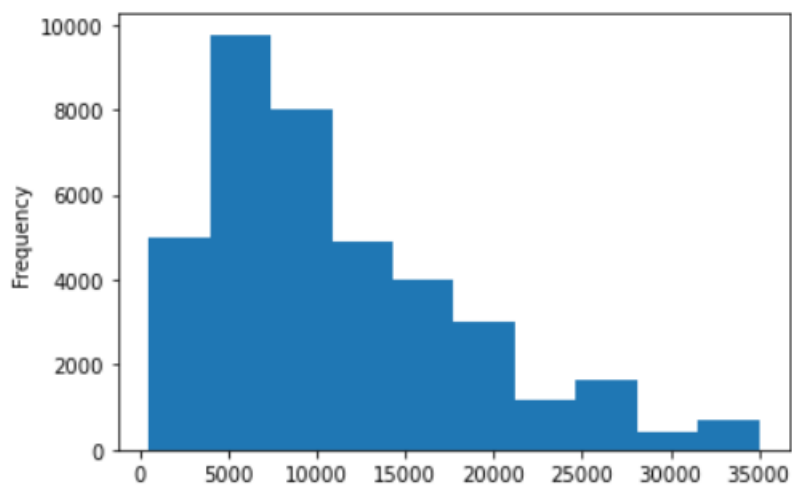
This histogram shows that most of the applicants loan amount ranges between 5000 to 10000
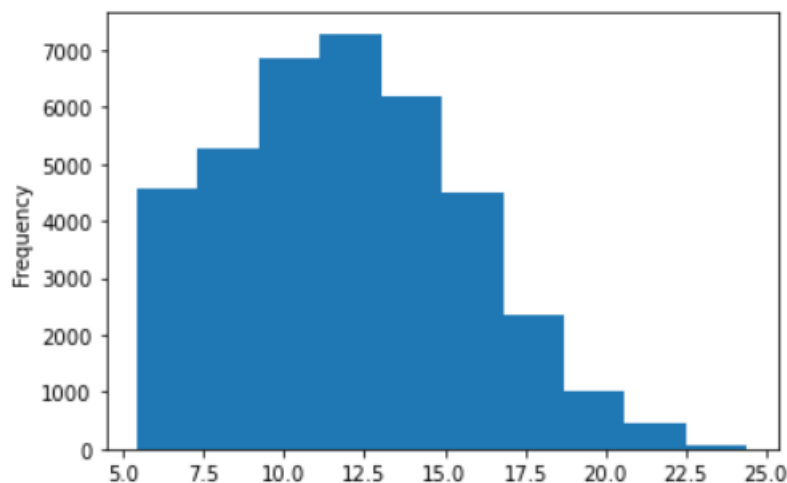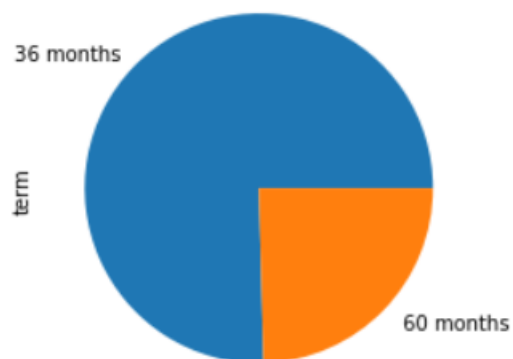

This histogram shows that most of the applicant's interest rate is between 10% to 15%
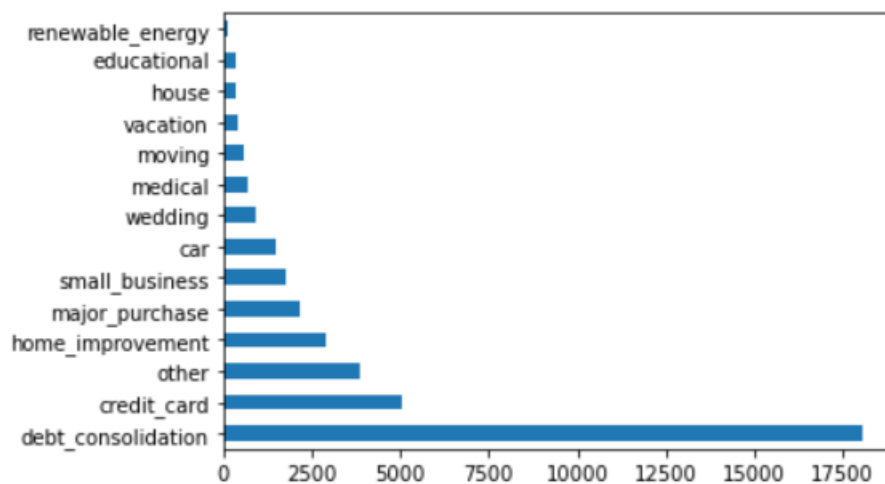
This histogram shows that for most of the applications monthly instalment ranges between 200 and 400.
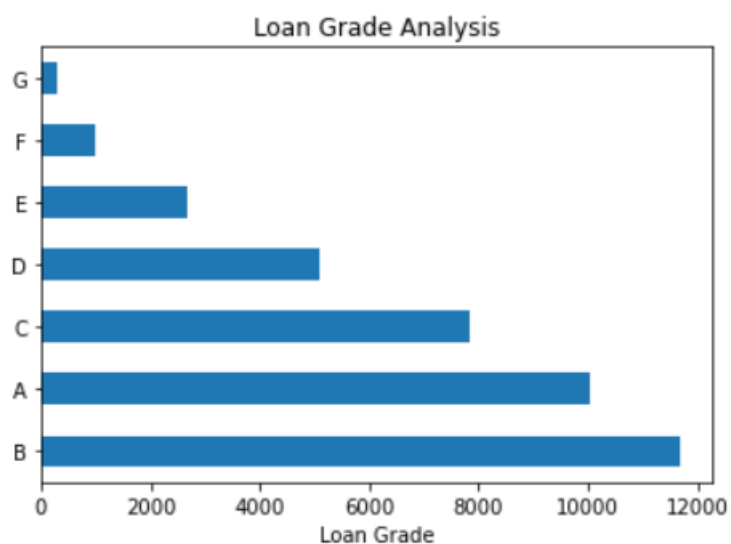

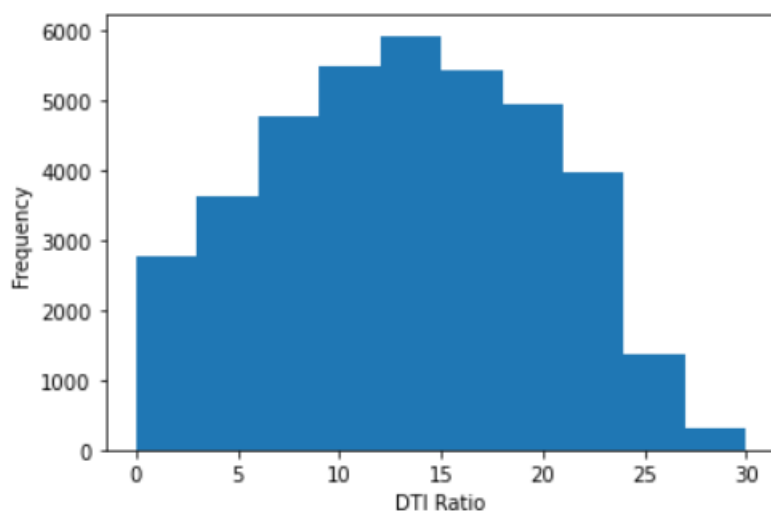Most of the applicants have 10+ years of experience

According to the illustration most applicants prefer 36 months of term.
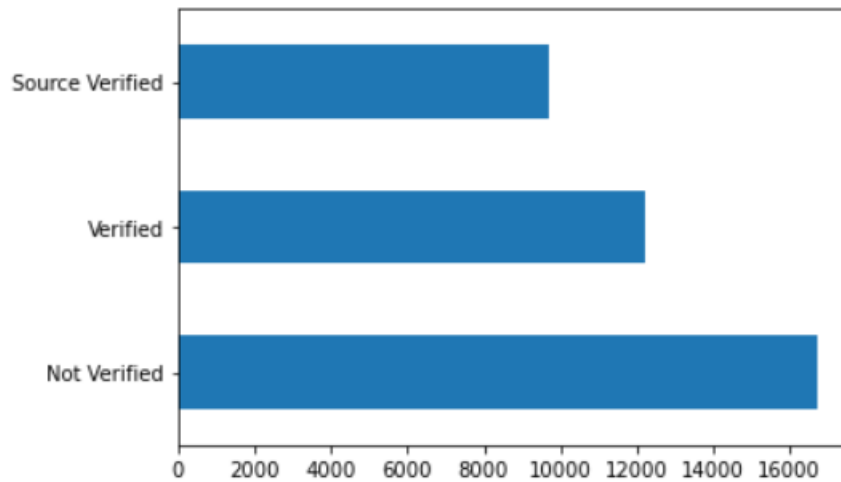
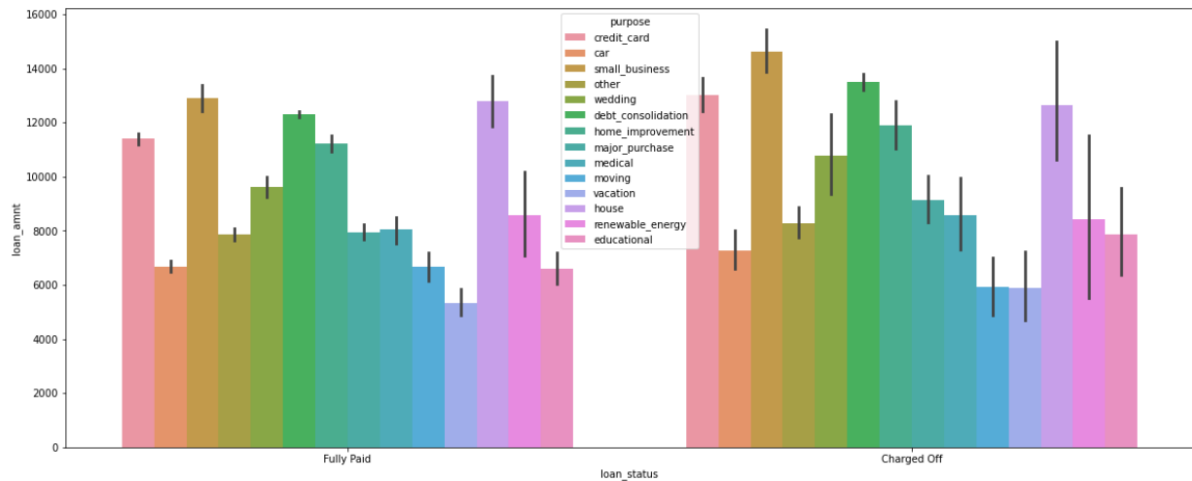Most of the applicants are taking loan for repayment of debt consolidation

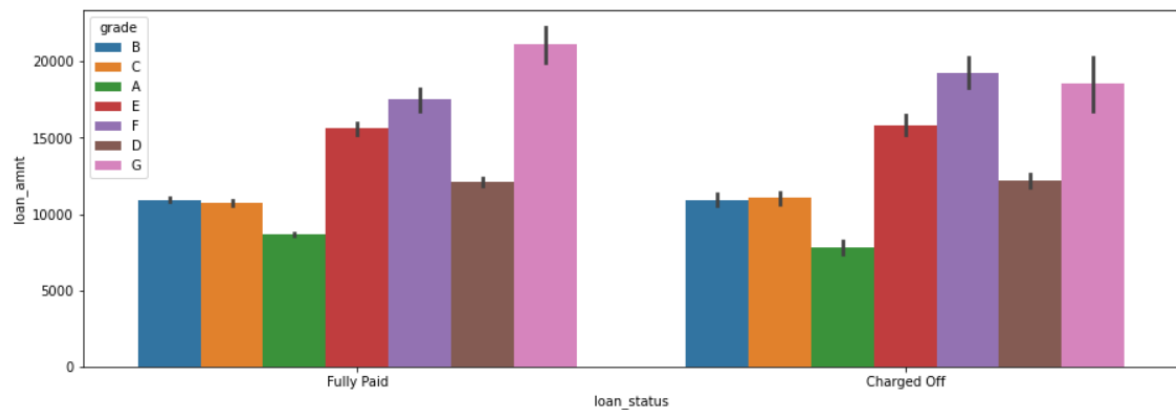Most of the applicants belong to B Grade

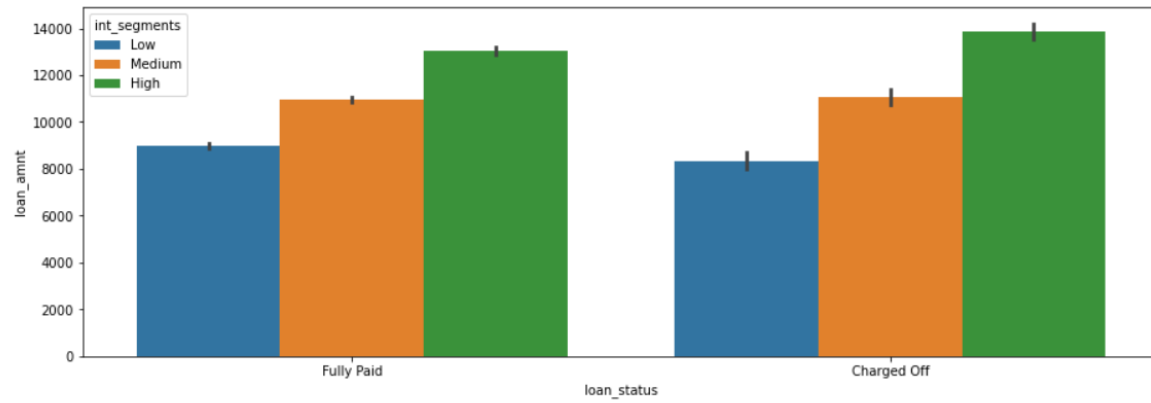Most of the applicants DTI ratio is between 10 and 20.

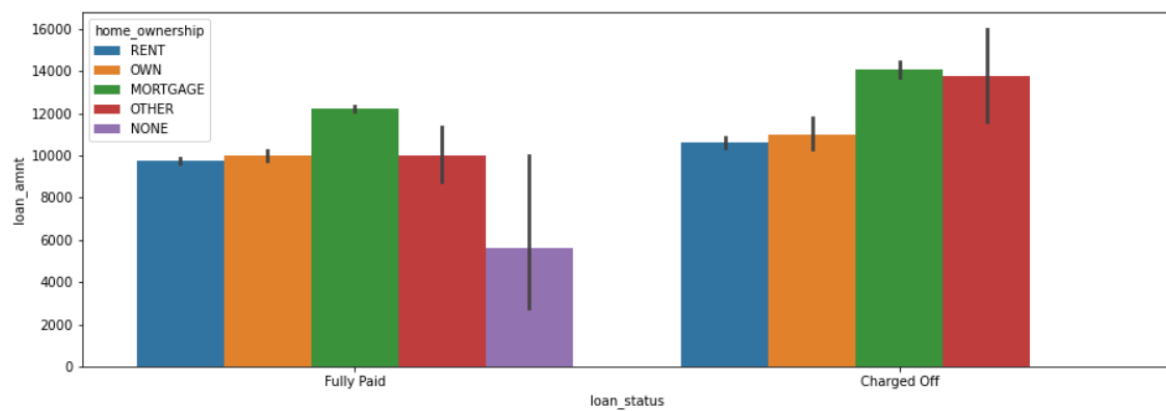Most of the applicants status is Not Verified.

This indicates that applicants take higher loan amount for Small business purpose.
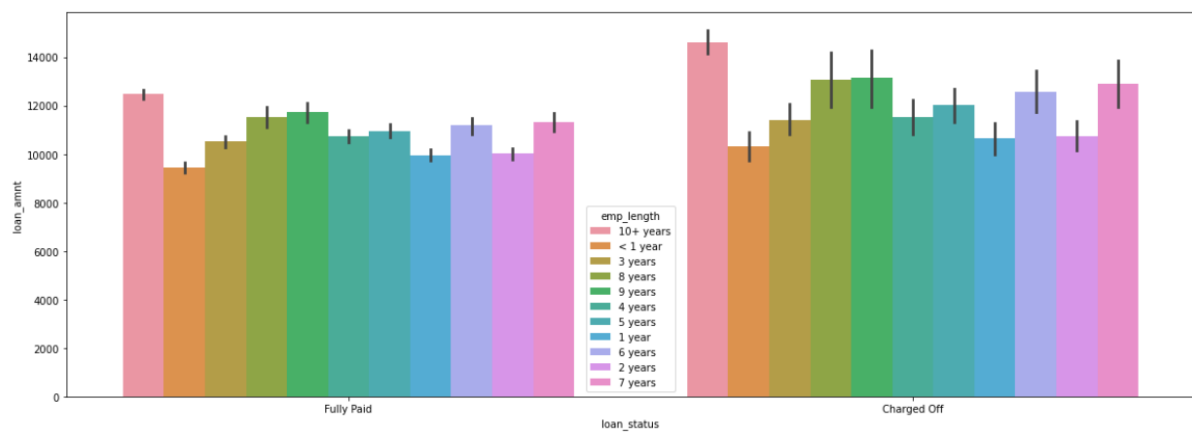

Maximum Loan applicant from charged off group belong to grade "F" and their loan amount also high
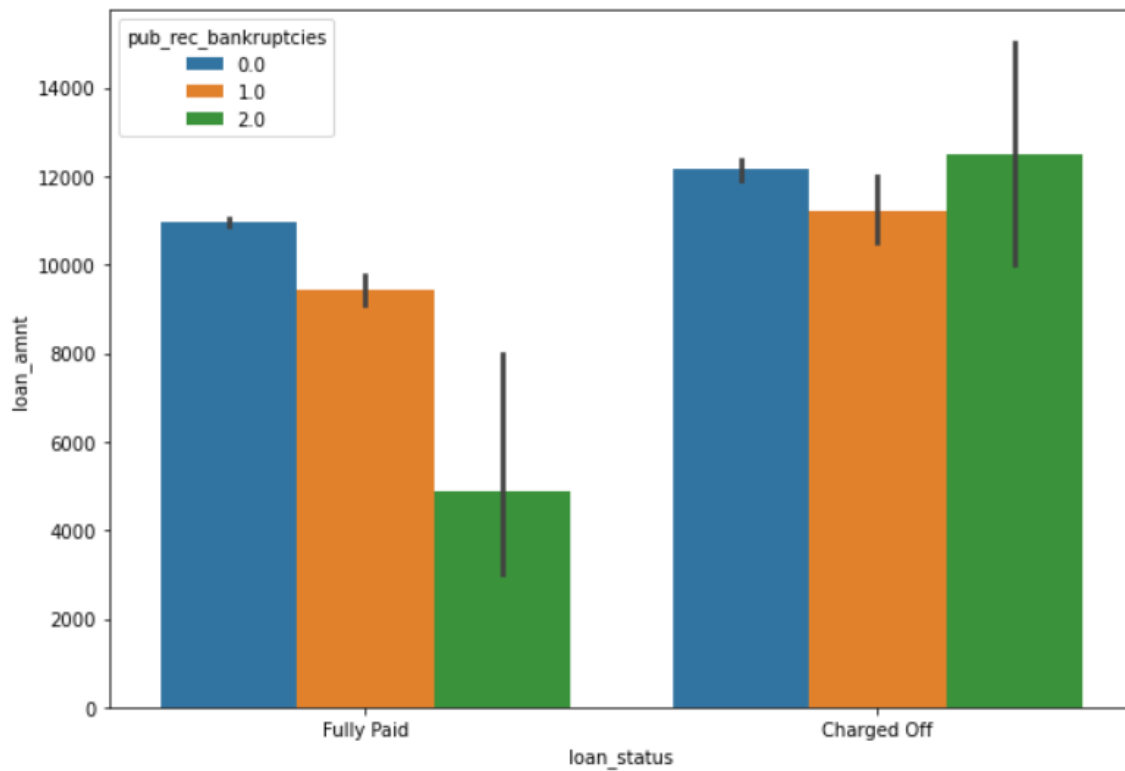
Maximum Loan applicant from charged off group have high rate of interest and their loan amount also high
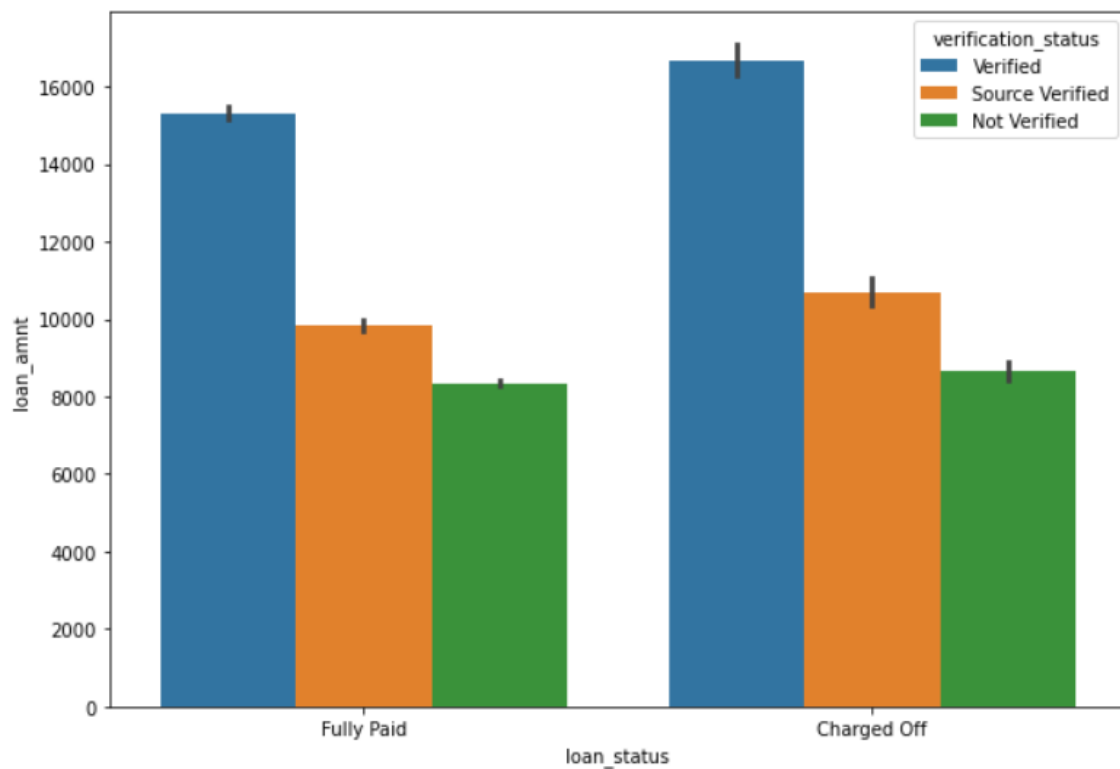

Maximum Loan applicant from charged off group have home ownership as mortage and their loan amount also high.

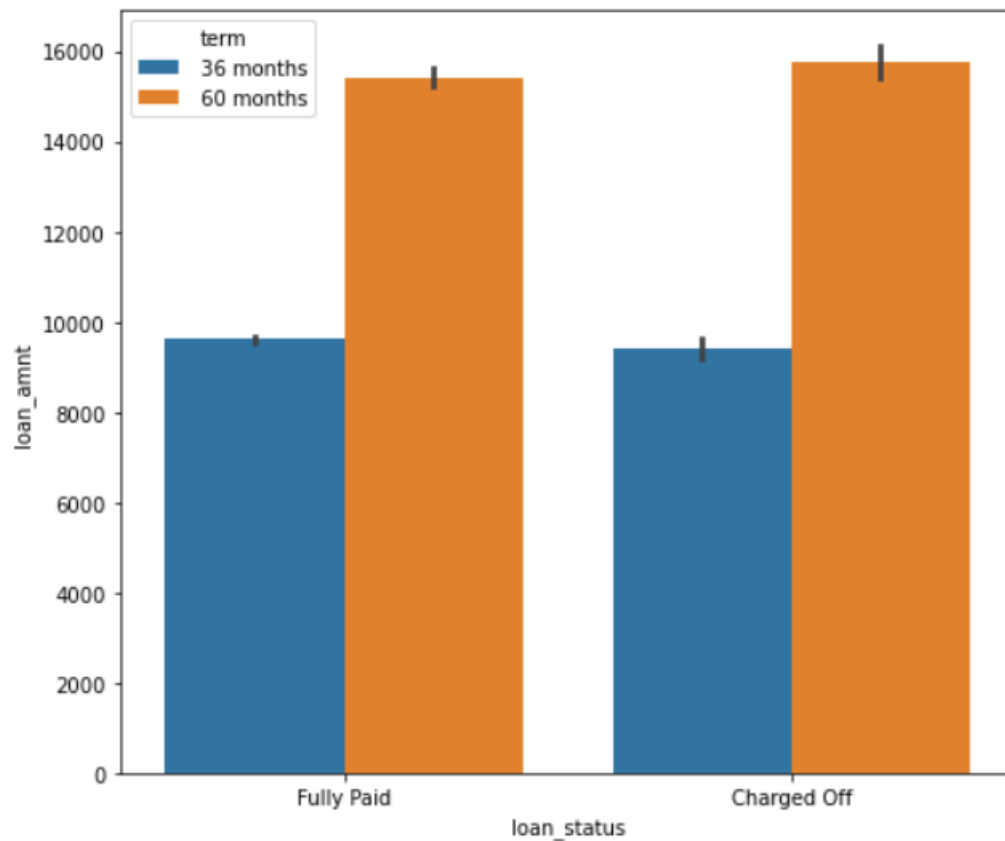Maximum Loan applicant from charged off group have employment length 10+ years

Most applicants from charged off group have high loan and 2 bankruptcies records

Most applicants from charged off group have high loan amount and their verification status id Verified.

Most applicants from charged off group have high loan amount and high term loan
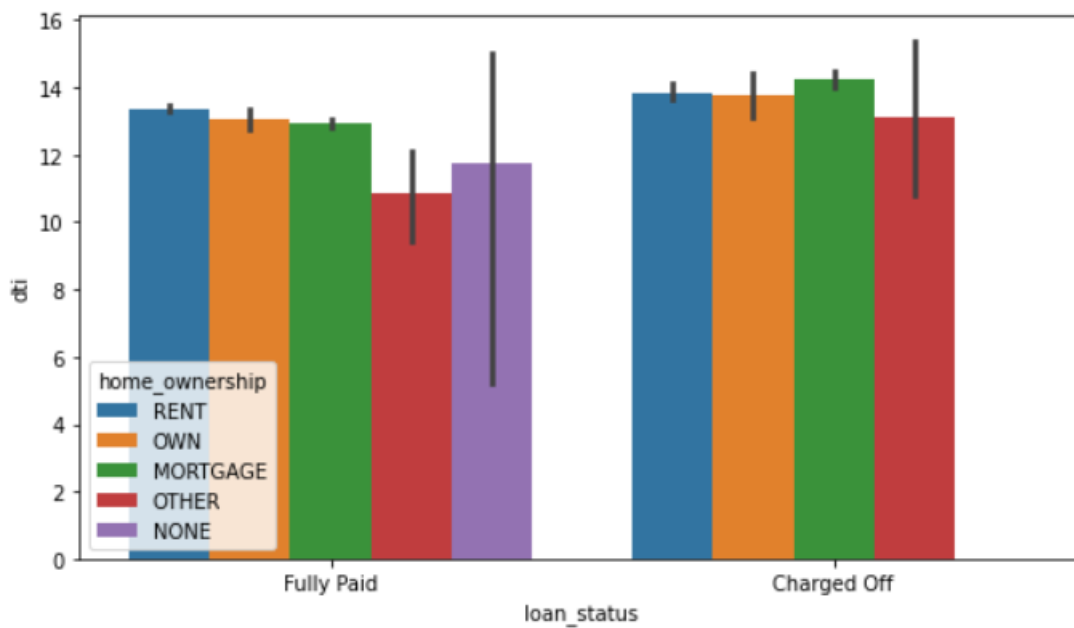
Most applicants from charged off group have dti and home ownership as Mortgage

There is no significant differentiation between dti and term of fully paid customers and Charged off customers

Most applicants from charged off group have dti and verification status as Verified

Conclusion

<u>Univariate Analysis Inferences:</u>

This indicates that the loan amount for most applications is between 5k to 10K

More loan applicants tend to go for 36 months term compared to 60 months

Most of the applications have around 200 to 400 as the monthly instalment

For most applications the interest rate is around 10% to 14%

For majority of the applications the DTI is around 10 to 18

We observe that a greater number of customer loan application have status as "Not Verified"

More number of loan applications belong to "B" Grade

More applications belong to applicants with 10+ years of employment

Most loan applicants stay in rented properties followed by applicants who pay mortgage.

Debt consolidation and credit card card repayment is major purpose of Loan

Maximum customers do not have any bankruptcies filed

<u>Bivariate Analysis Inferences:</u>

Maximum Loan applicant who default have higher loan amount and purpose seems to be 'small business'

Maximum Loan applicants belonging to charged off group have higher loan amounts and have interest rates too.

Maximum Loan applicant from charged off group belong to grade "F" and their loan amount also high

Maximum Loan applicants from charged off group have higher loan amount and belong to +10 Years of employment length category
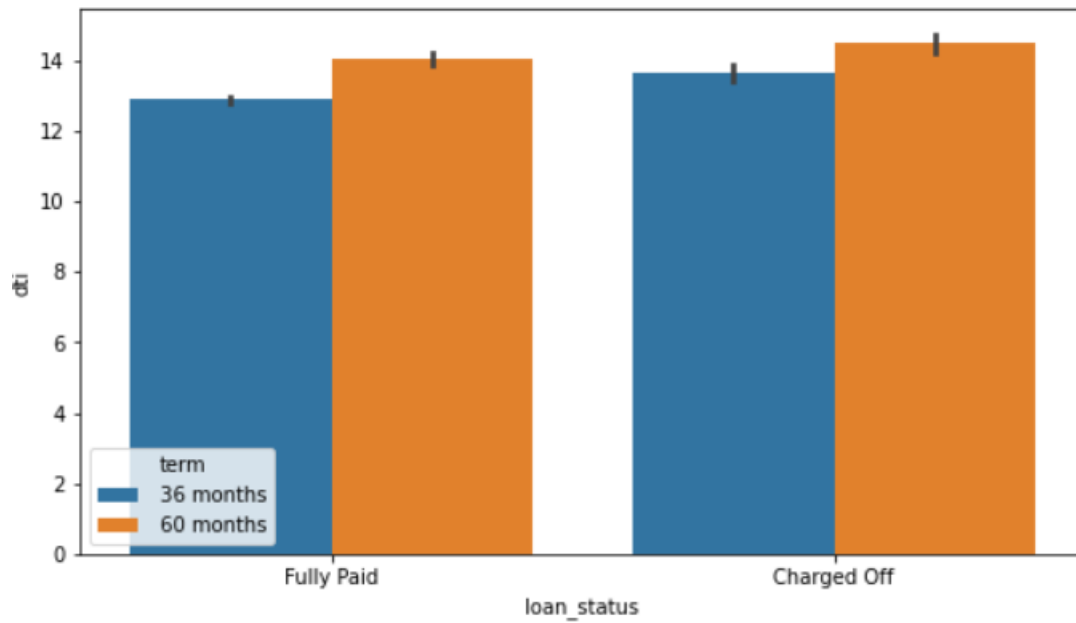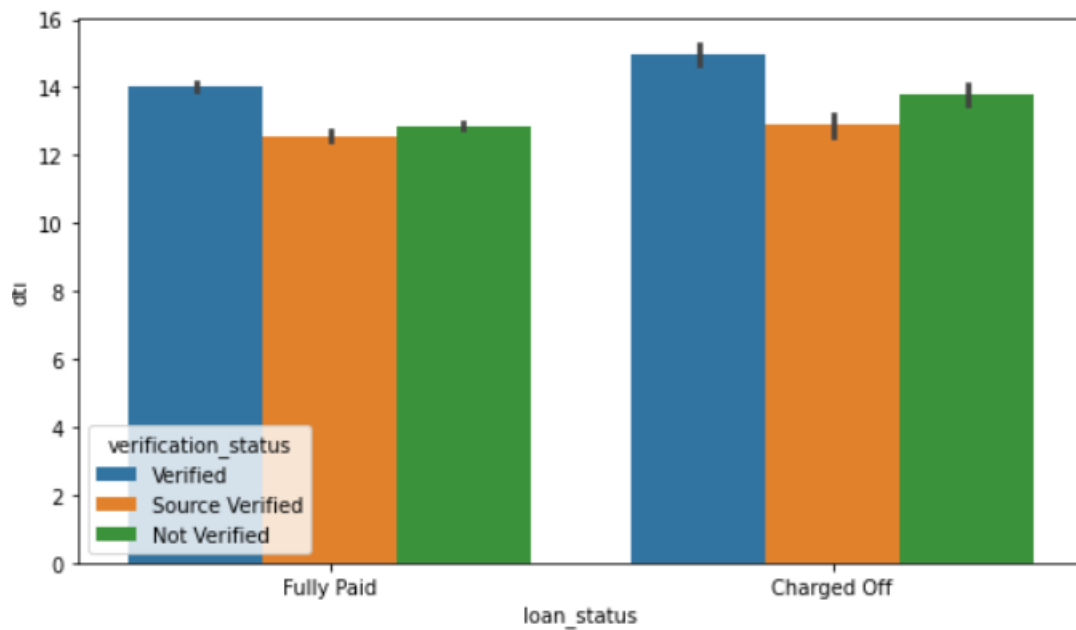
Maximum applicants from charged off group have high loan and more bankruptcies records on their name

Maximum applicants from charged off group have higher loan amount and their homeownership details are specified as Mortgage.

Maximum applicants from charged off group have higher loan amount and their verification status seems to "verified"

Maximum applicants from charged off group have higher loan amount and their loan term is high

Maximum applicants from charged off group with higher DTI have mortgage.

Most applicants from charged off group have high DTI and verification status as Verified.