Lending Club Case Study

Name:

Akash Nande

Deepu Malik

Summary and Case Study Description

The data given contains 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, we will use EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of default.

The objective of the study is to find risky candidates for loan based on past data provided.

Solutioning , Analysis and Approach
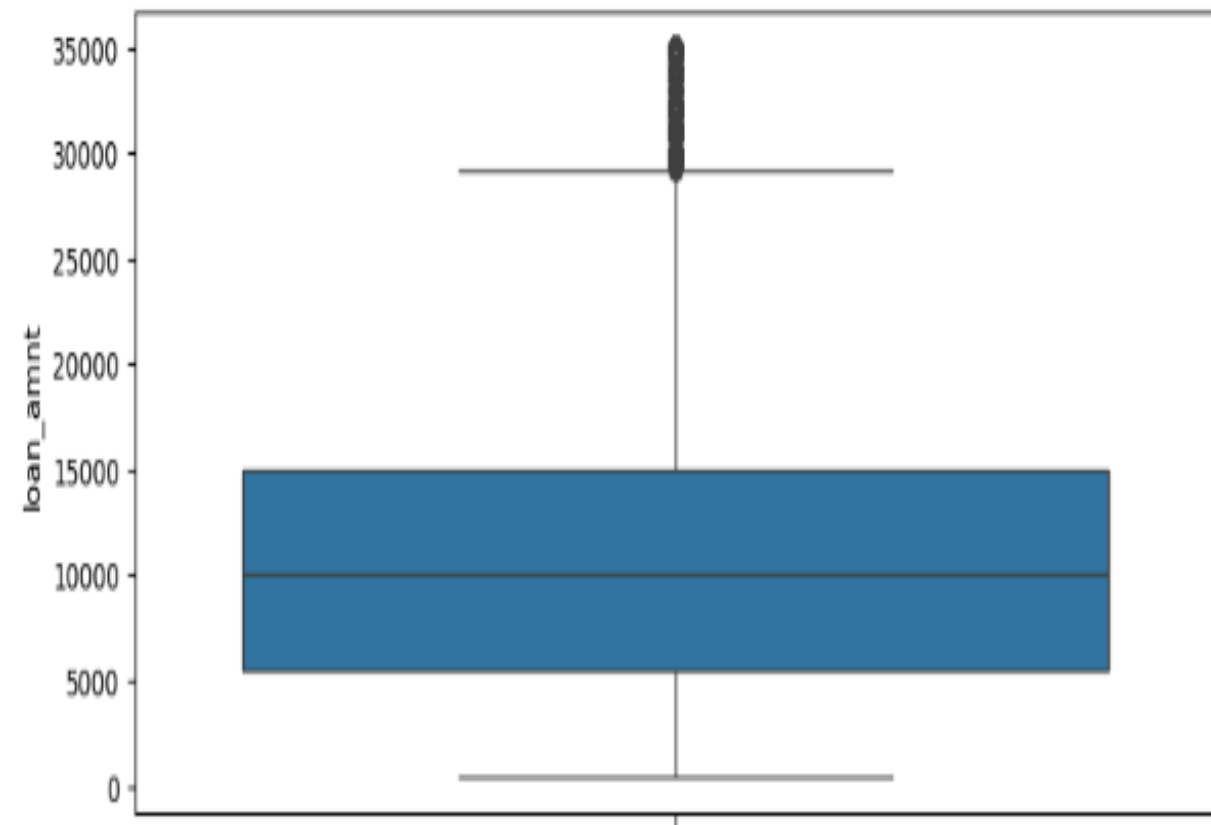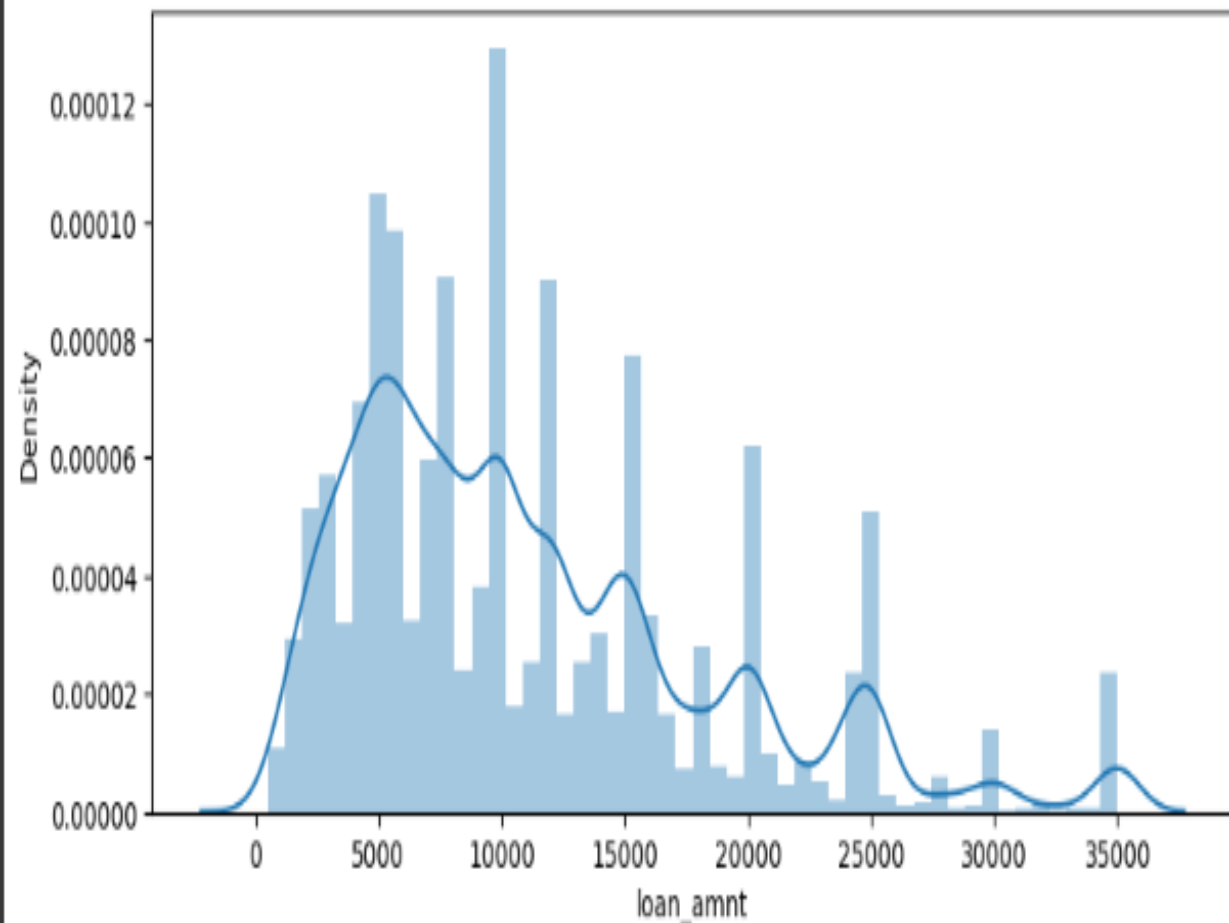
IMPORTING
PYTHON LIBRARIES

MANAGING THE
NULL VALUES BY
EITHER REMOVING
OR IMPUTATION

ASCERTAINING
OUTLIERS

IDENTIFYING DATA
IMBALANCE

UNIVARIATE,
SEGMENTED
UNIVARAITE AND
BIVARIATE ANALYSIS

FINDING
CORRELATION

- From the graph we can say that:-

1.10000 was the most loaned amount by each person
2.The highest amount disbursed is 35000.
3.The count of people who took 30000 was the lowest.
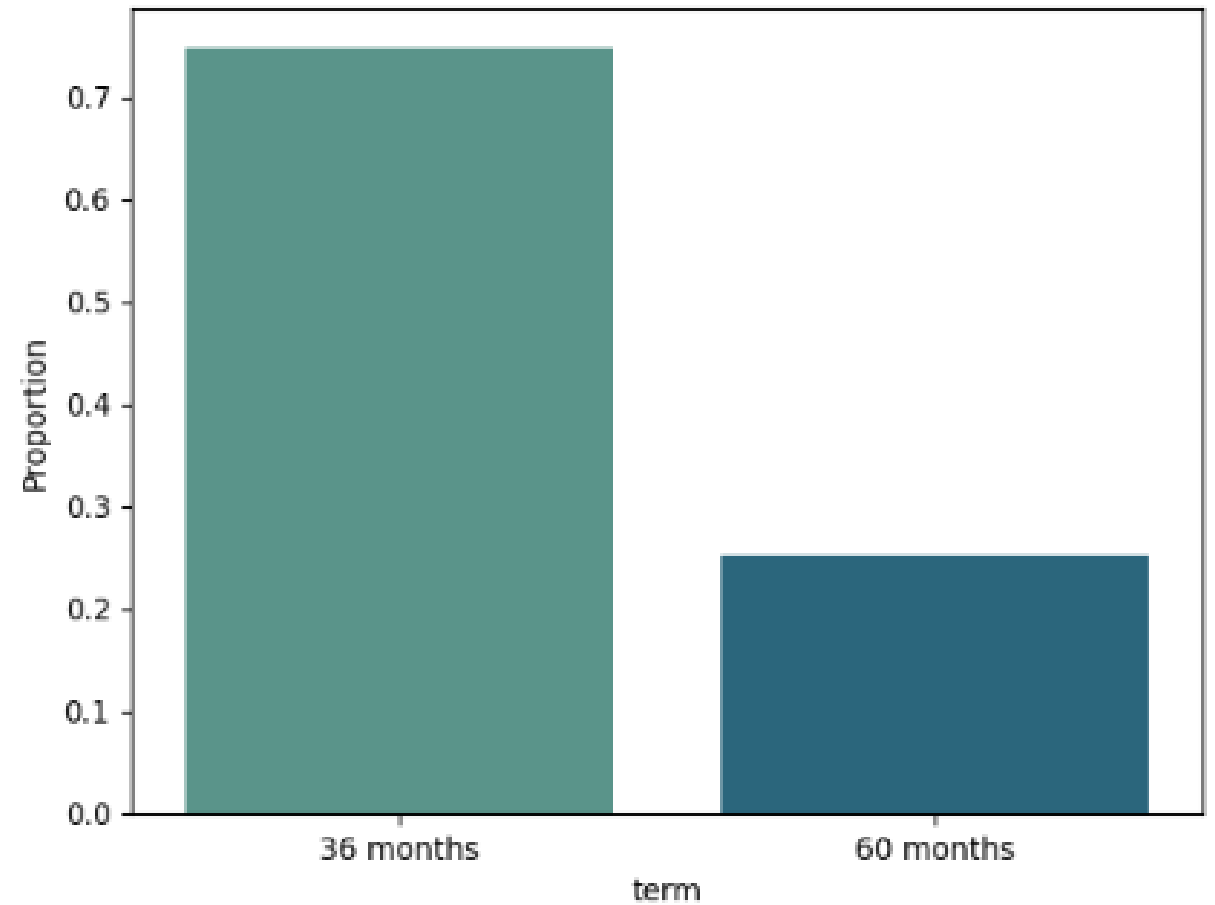4.Majority of loan amount is between 5000 and 15000

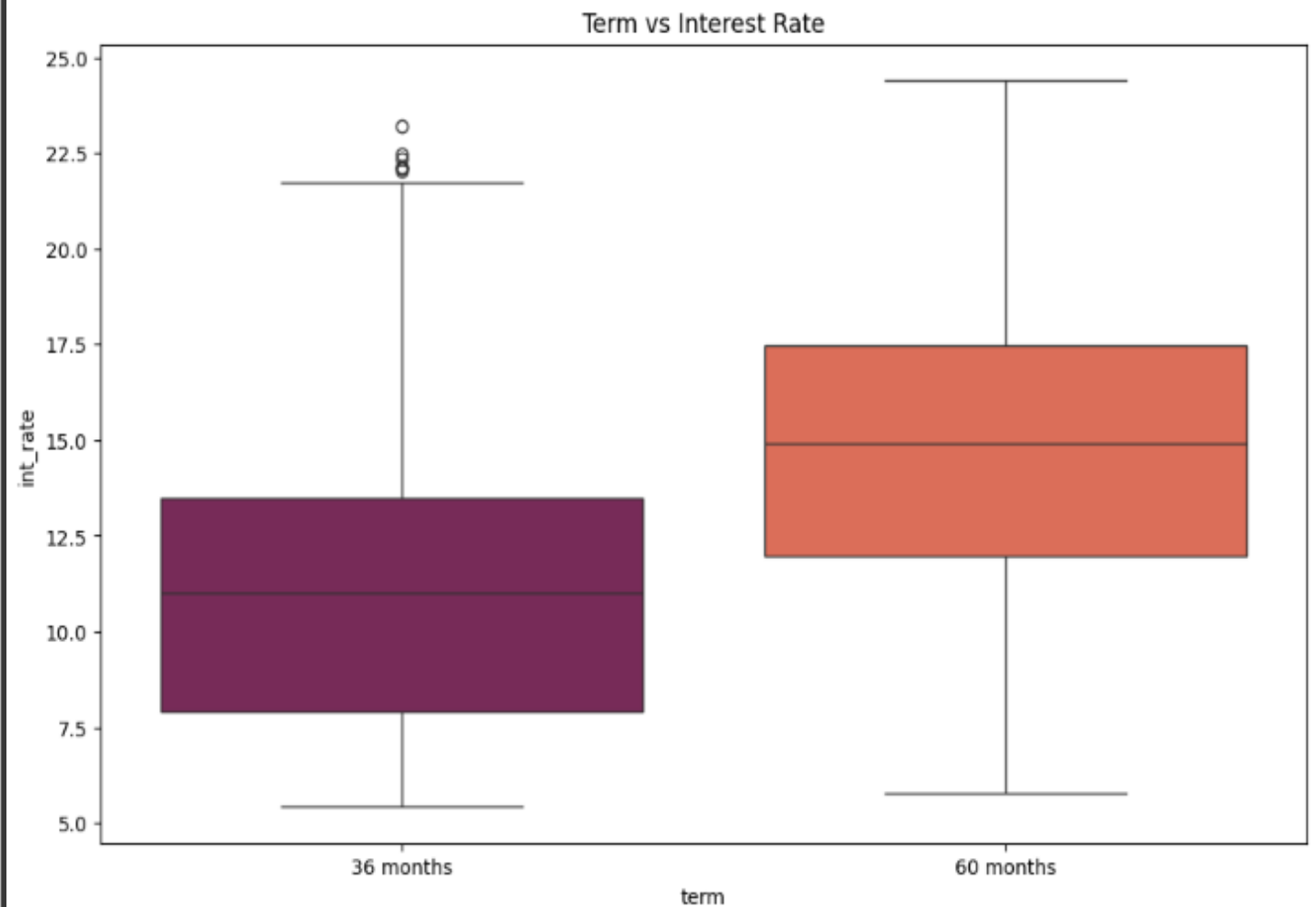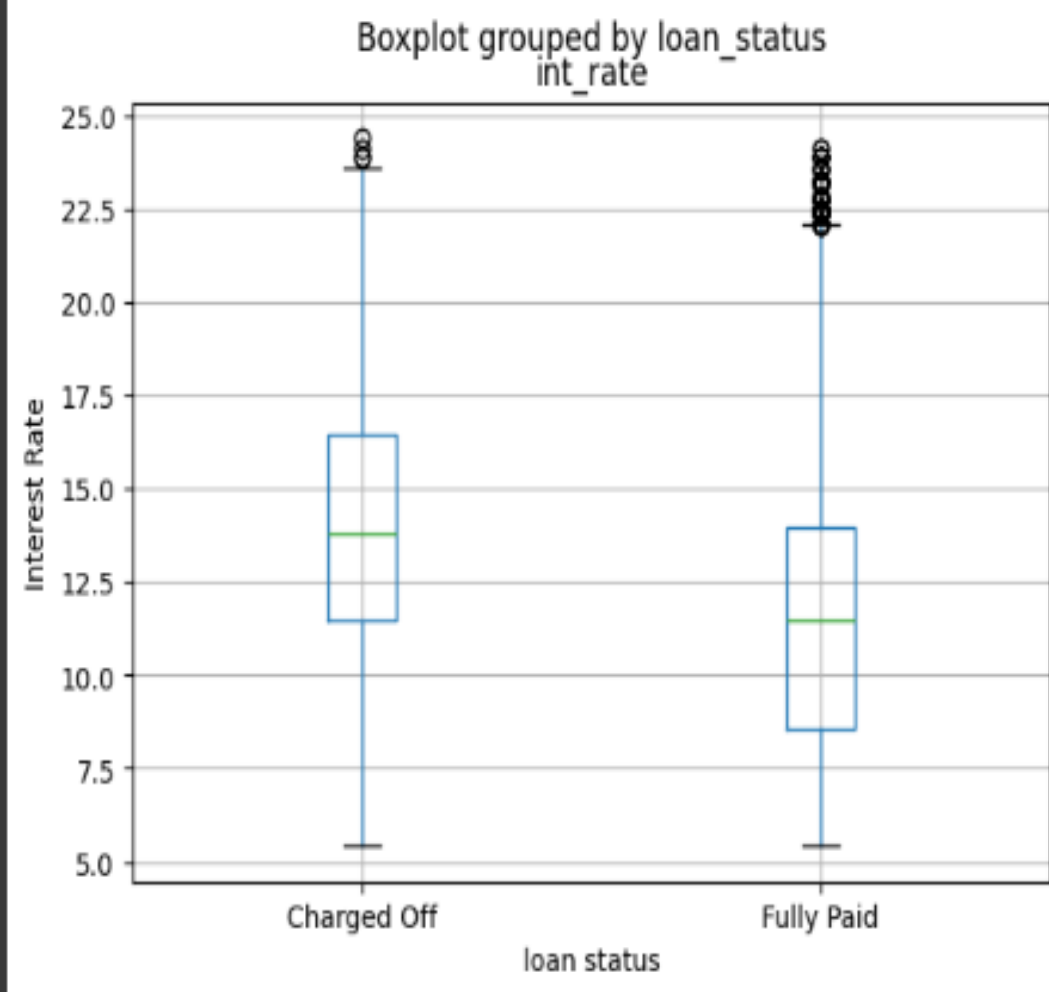From the analysis we found that , the loan amount disbursed has increased and nearly doubled year on year
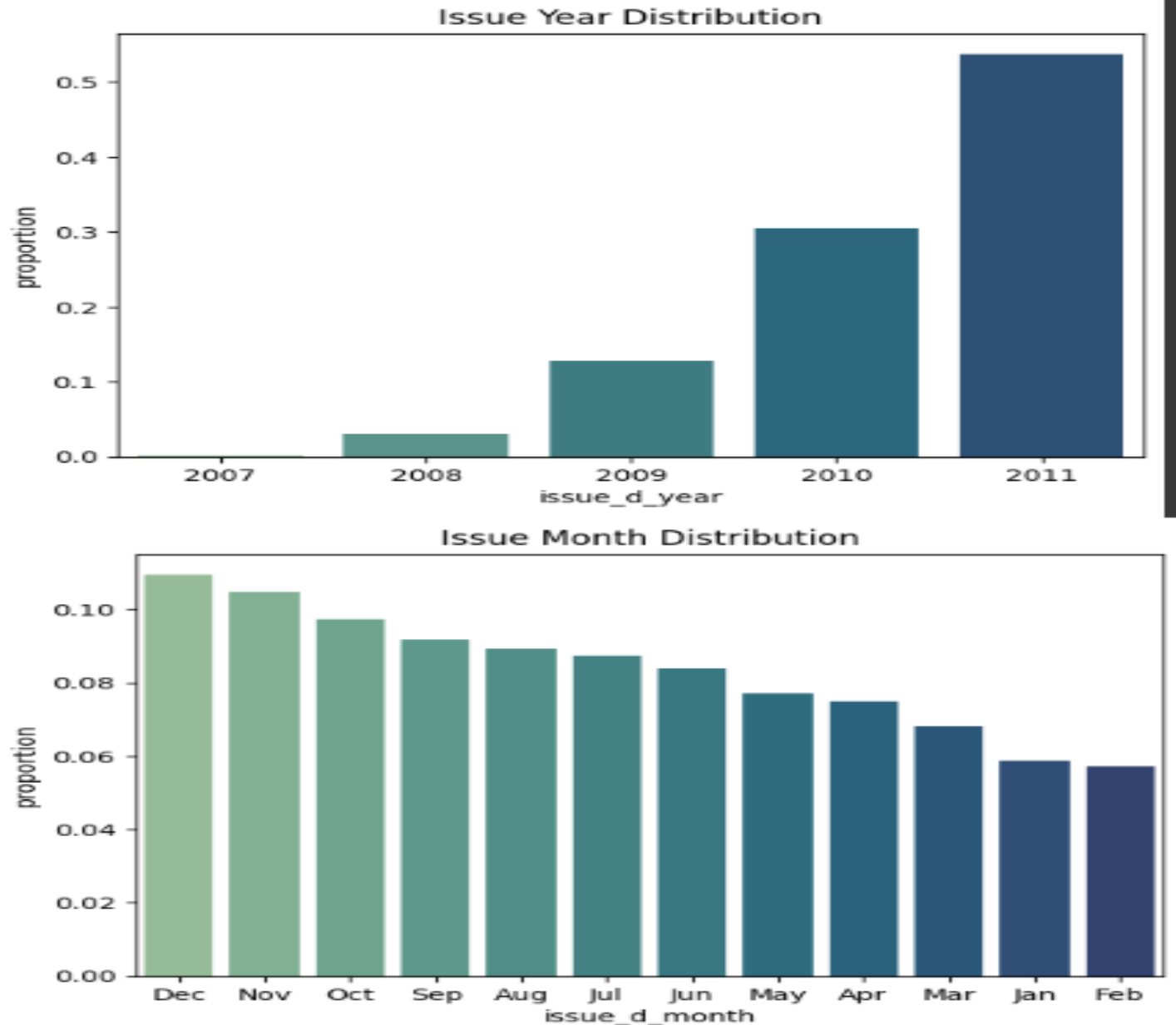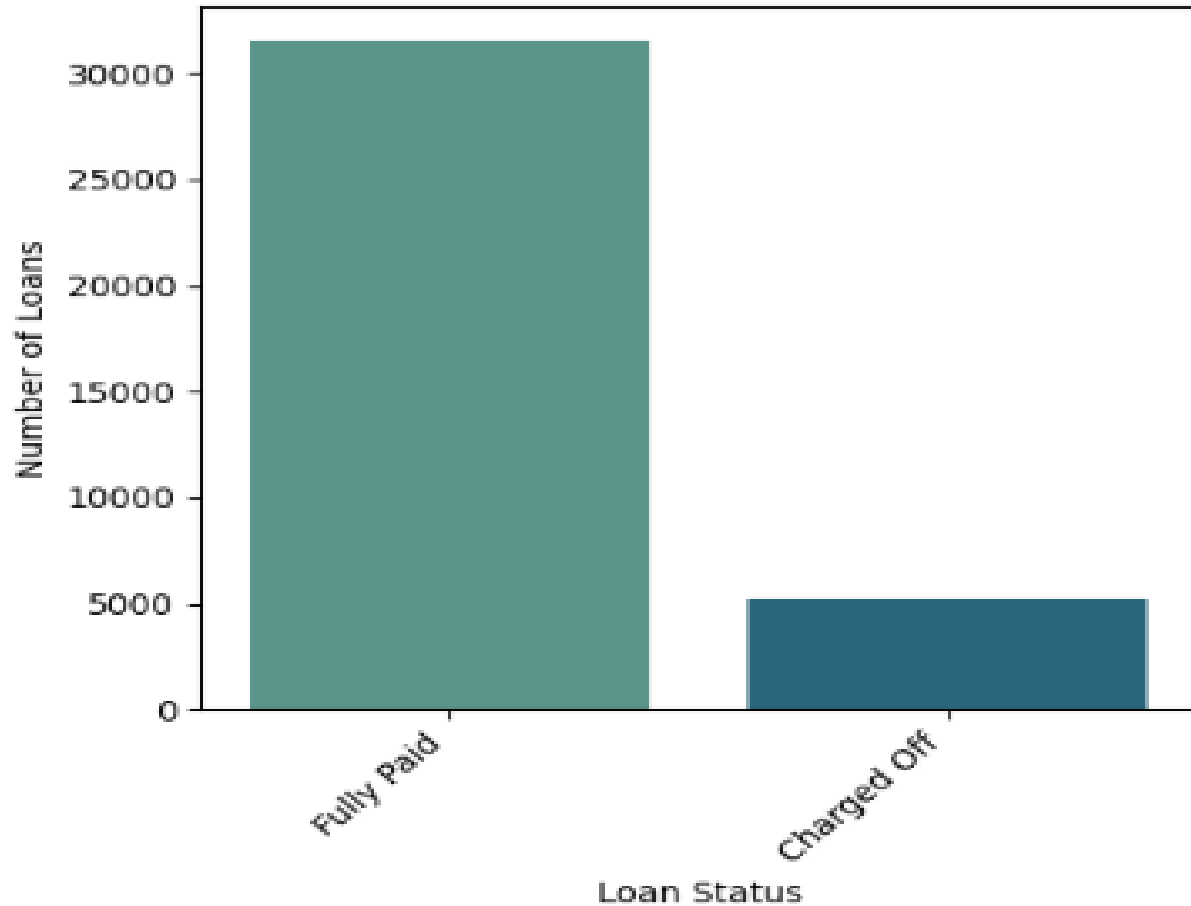
- From the above graph , we can say that most of the loans were taken for 36 months.
- Reading the graph ,we can also deduce that most loans were fully paid and nearly 5000 loans were charged-off

- From the above graph , we can say that higher interest rates are having more chance to get charged-off more often .
- Reading the graph ,we can also deduce that most loans with 60 months have higher interest rates.

Bar plot of loans by Term grouped by Loan Status
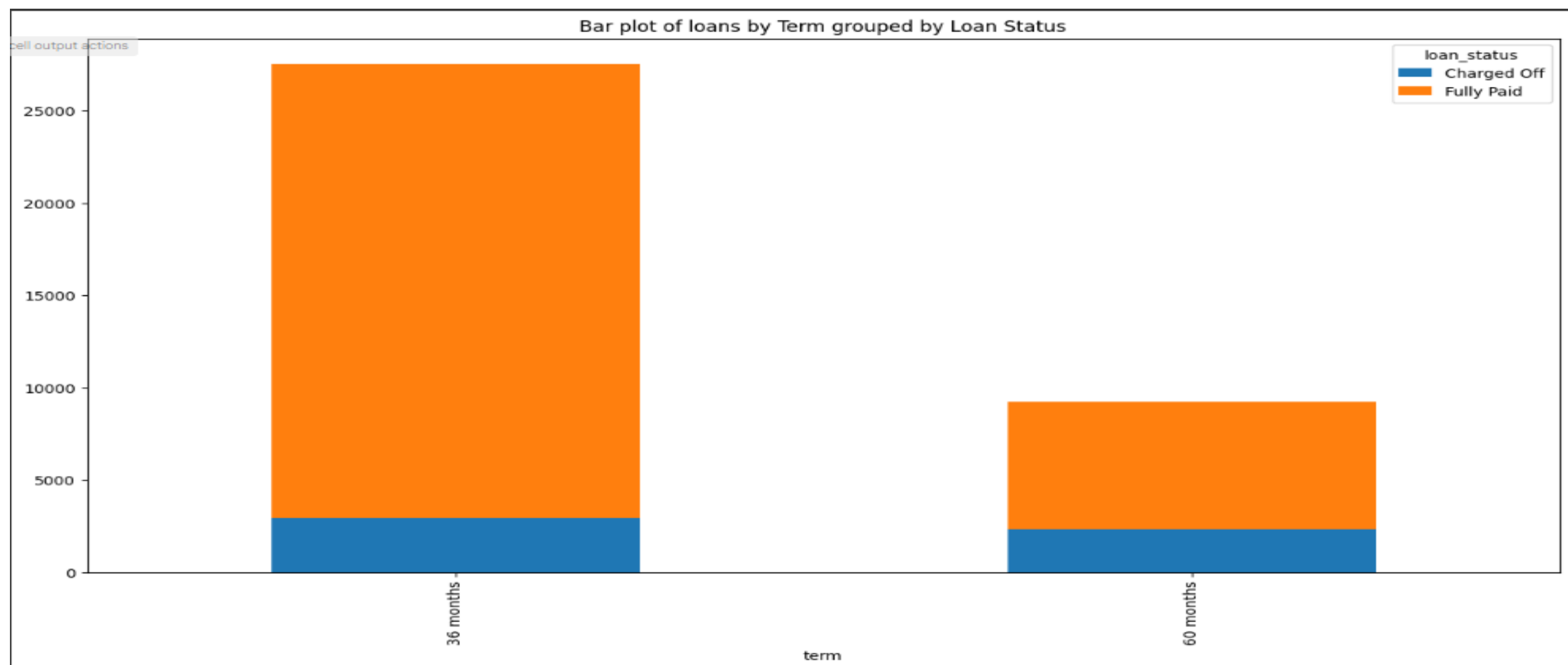
- Thus , post analysis we can say that Charged-off /Defaulters are greater with 60 months term with high interest rates
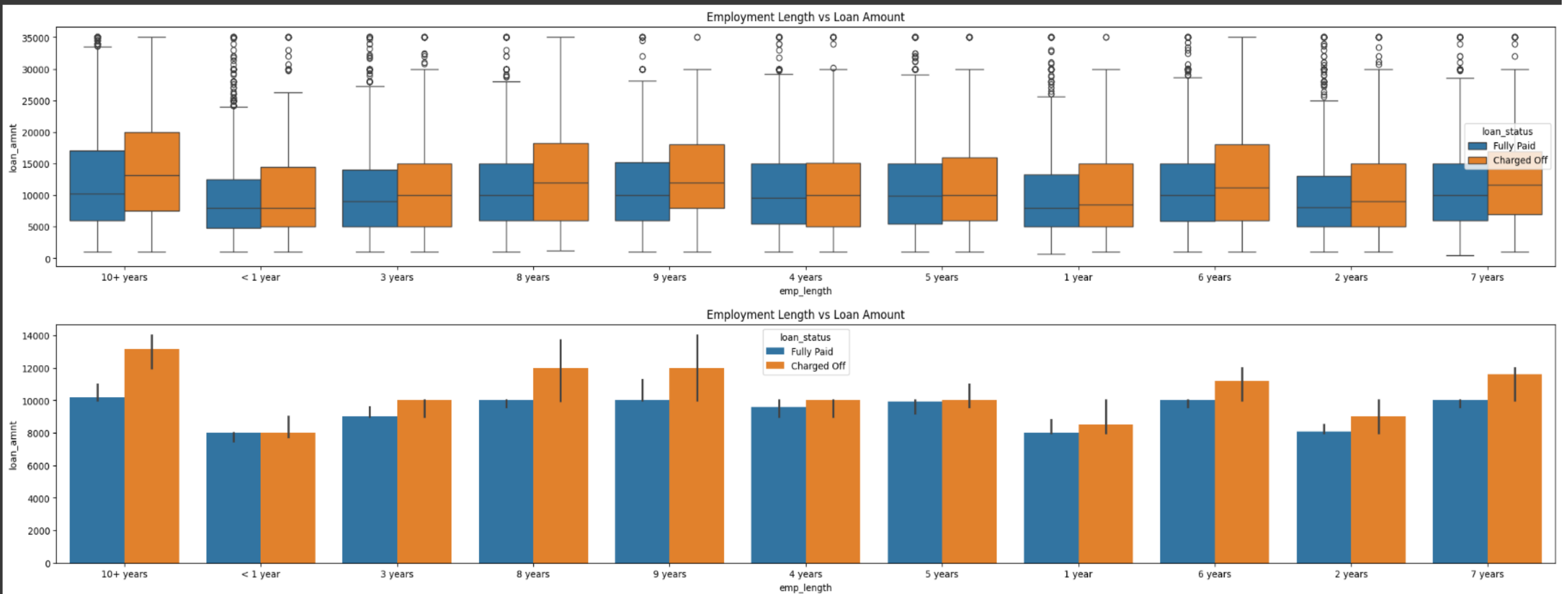
Grade vs Loan Amount
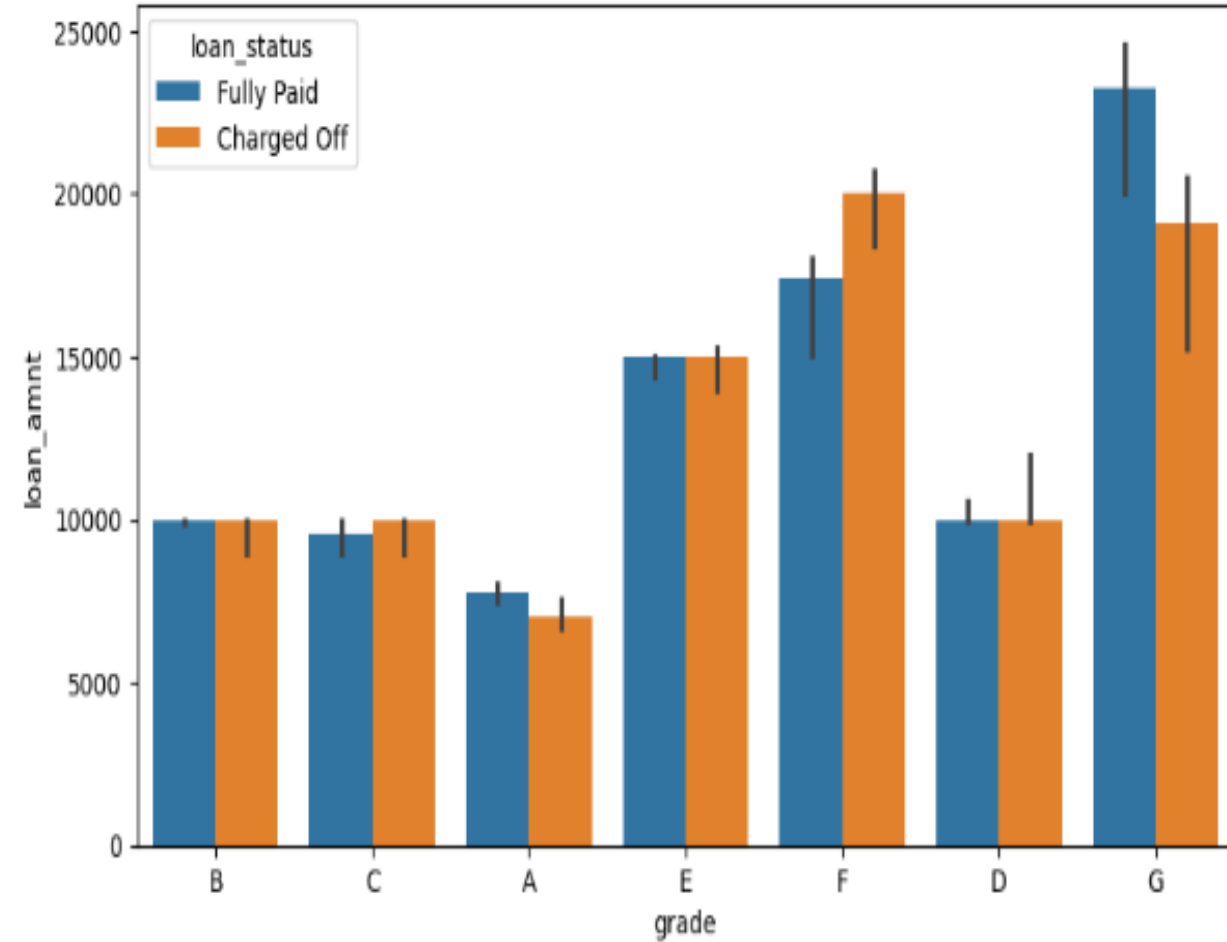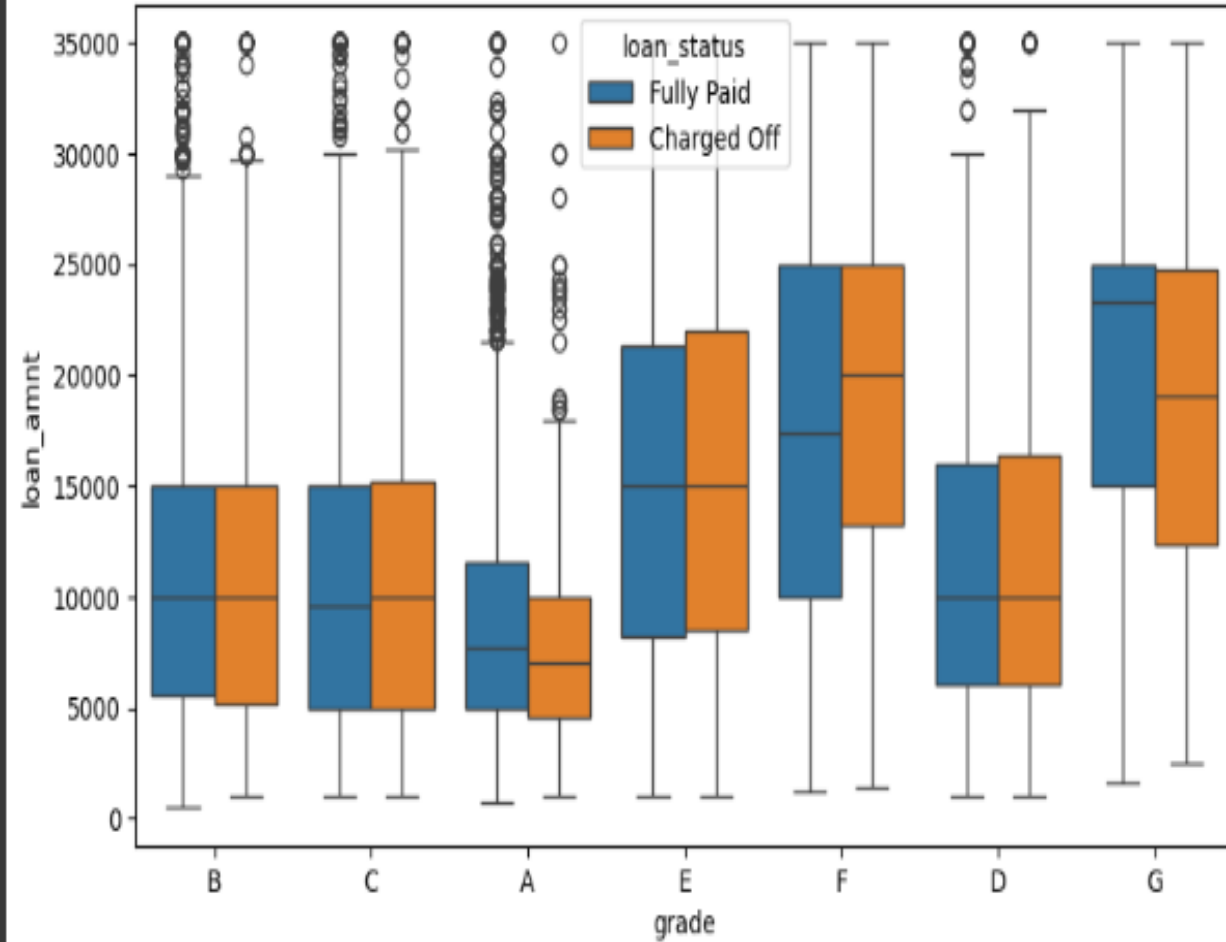
- People with E,F,G are more likely to be defaulters/Charged-off
- We can also deduce that the E,F,G category people take the most amount of loan and have a high probability to get charged-off
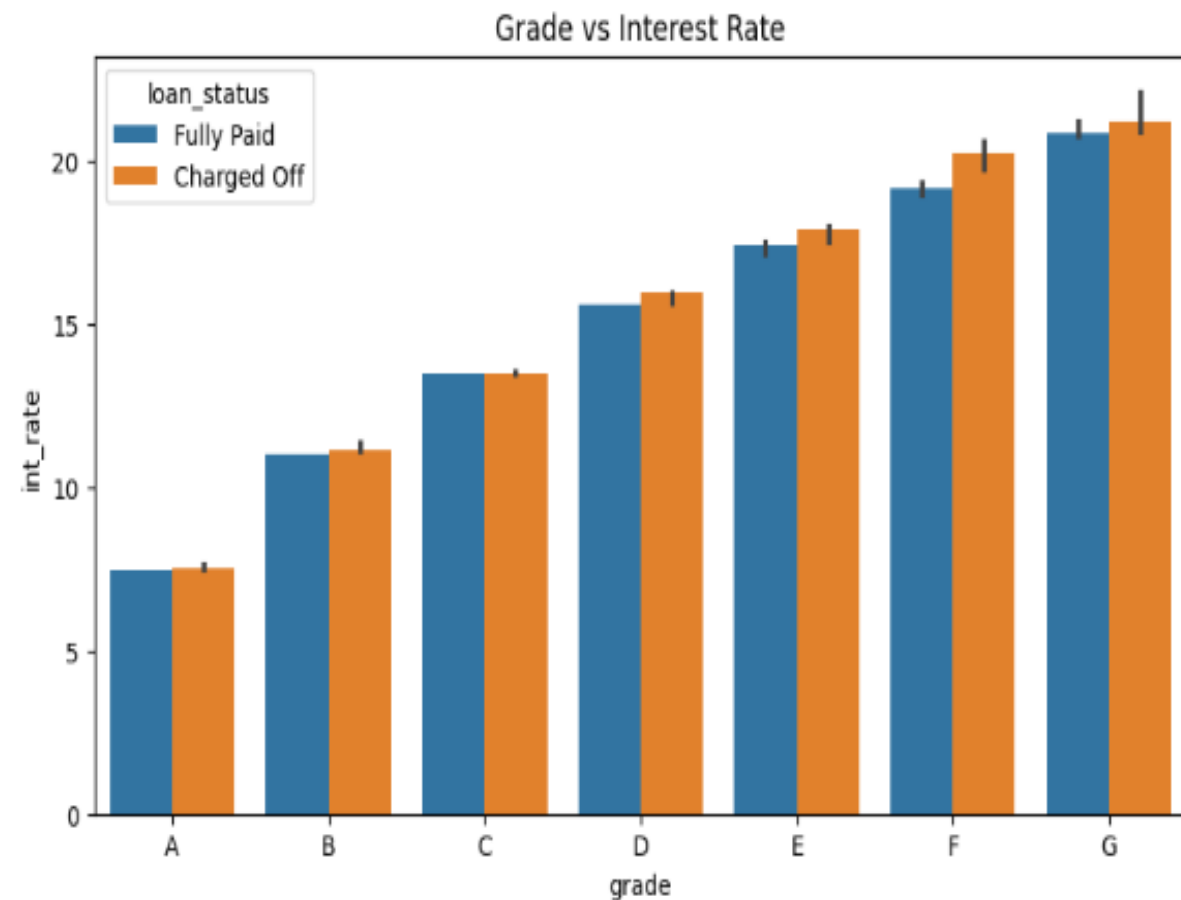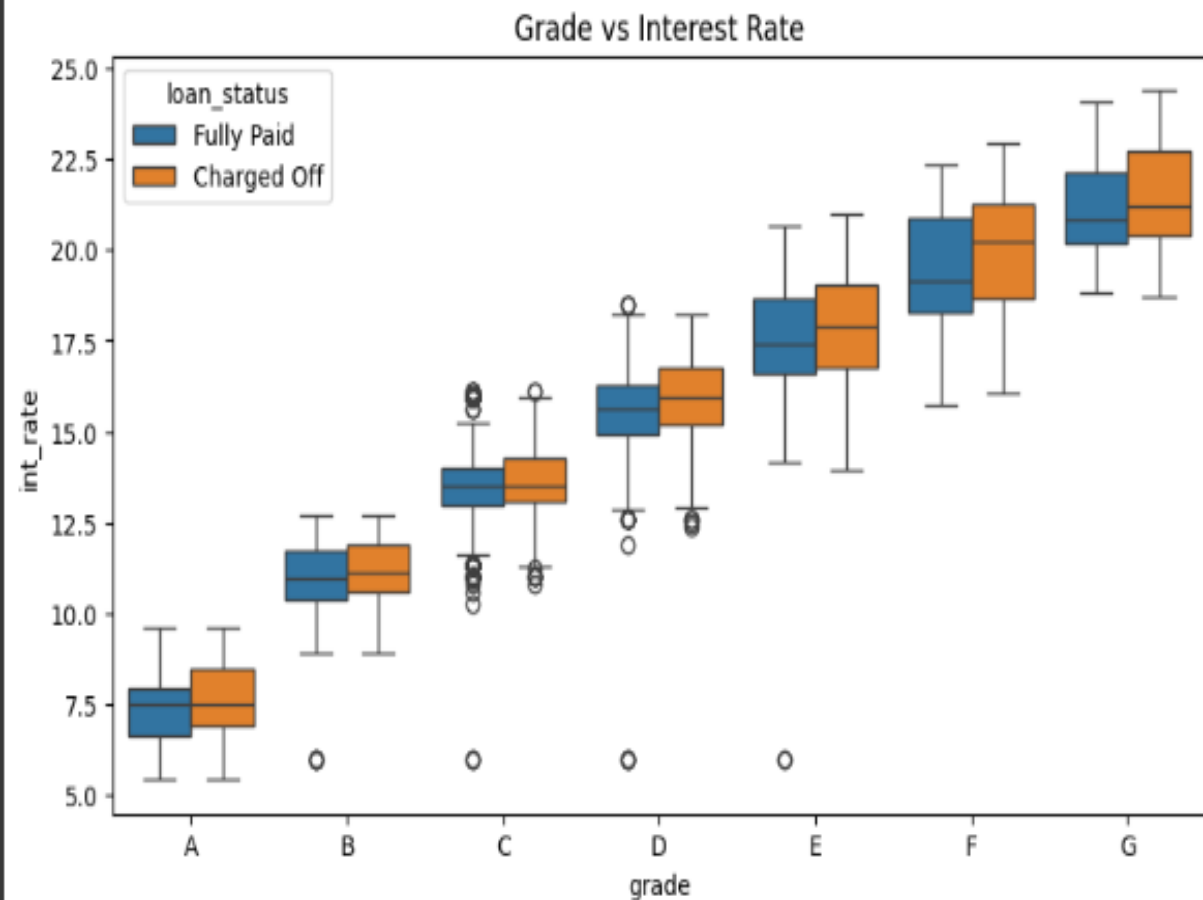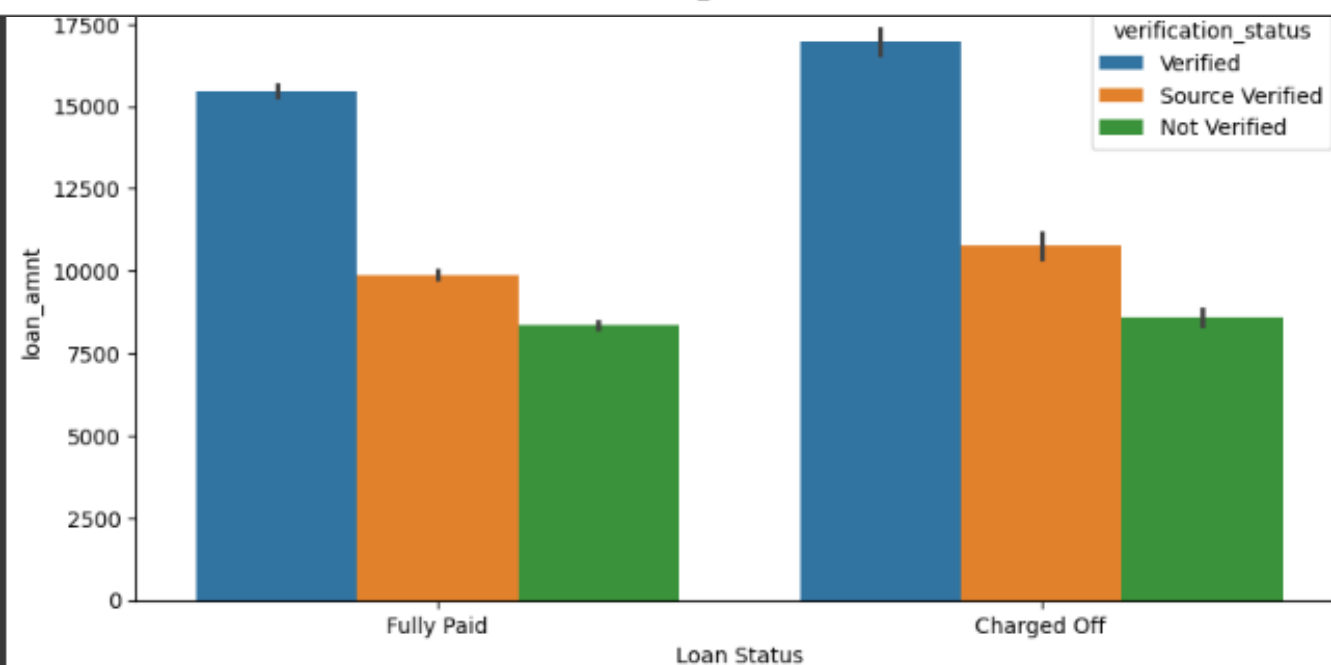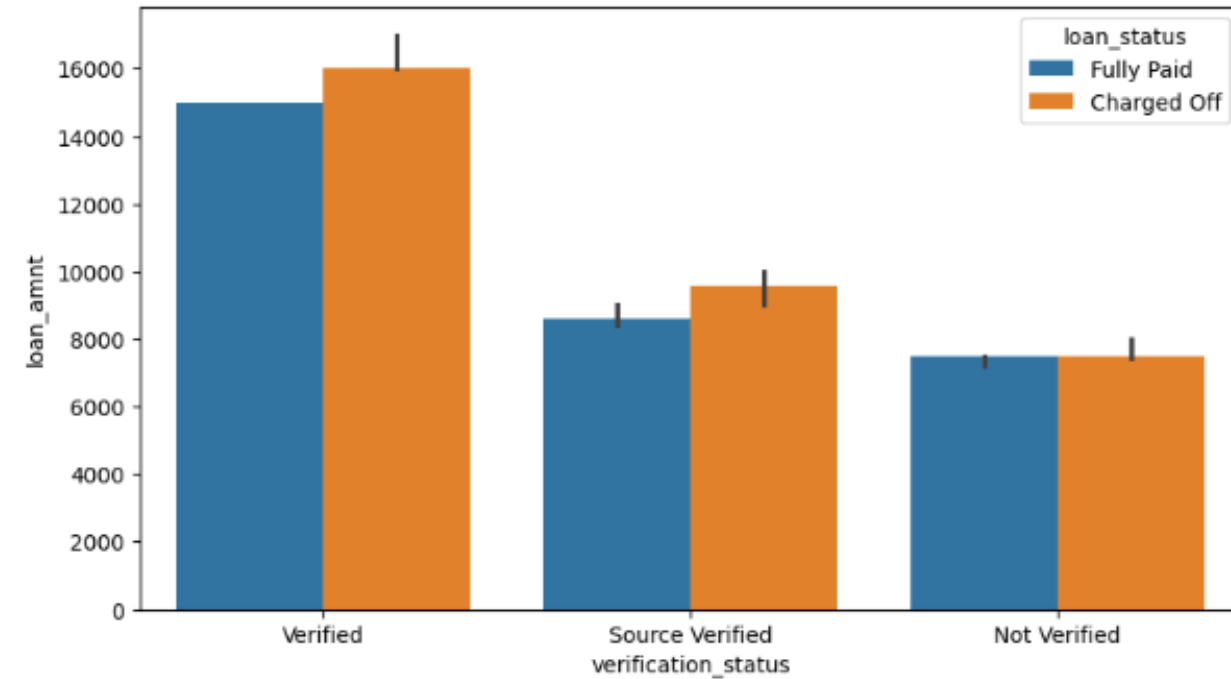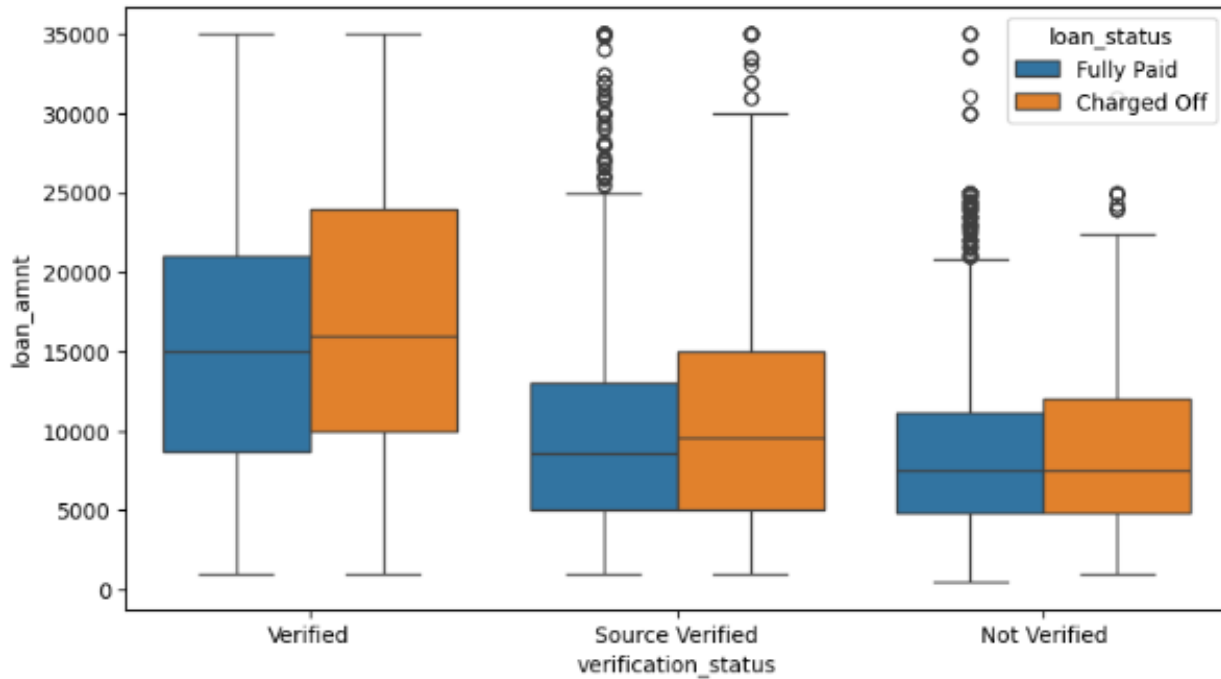
- From this graph we can confirm that as interest rate are higher for lower grades and they mostly take huge amount and gets charged-off

- From this graph we can say that most amount of loans are taken by people with 10+ years of experience
- And they are most likely to take big amounts and get charged-off

Verification Status vs Loan Amount

• From the given graphs ,we can say that there is a high percentage of data which are verified and source verified , thus we can assume that there is a case of corruption involved within the premises.

- Most people default on home loans that have high interest rates. The same observation can be made for small business loans and debt consolidation.

Address State vs Charged Off

Address State vs Fully Paid
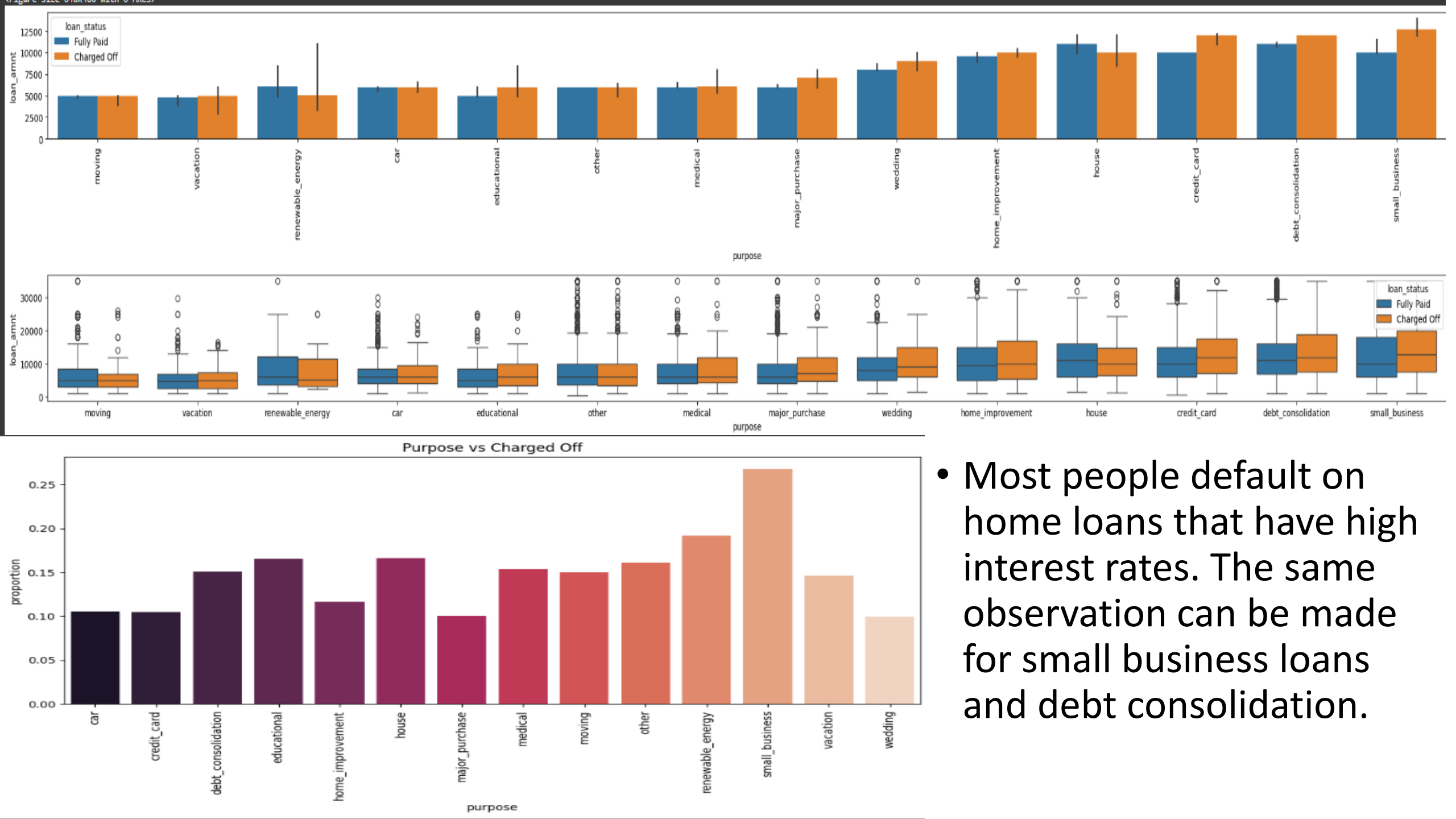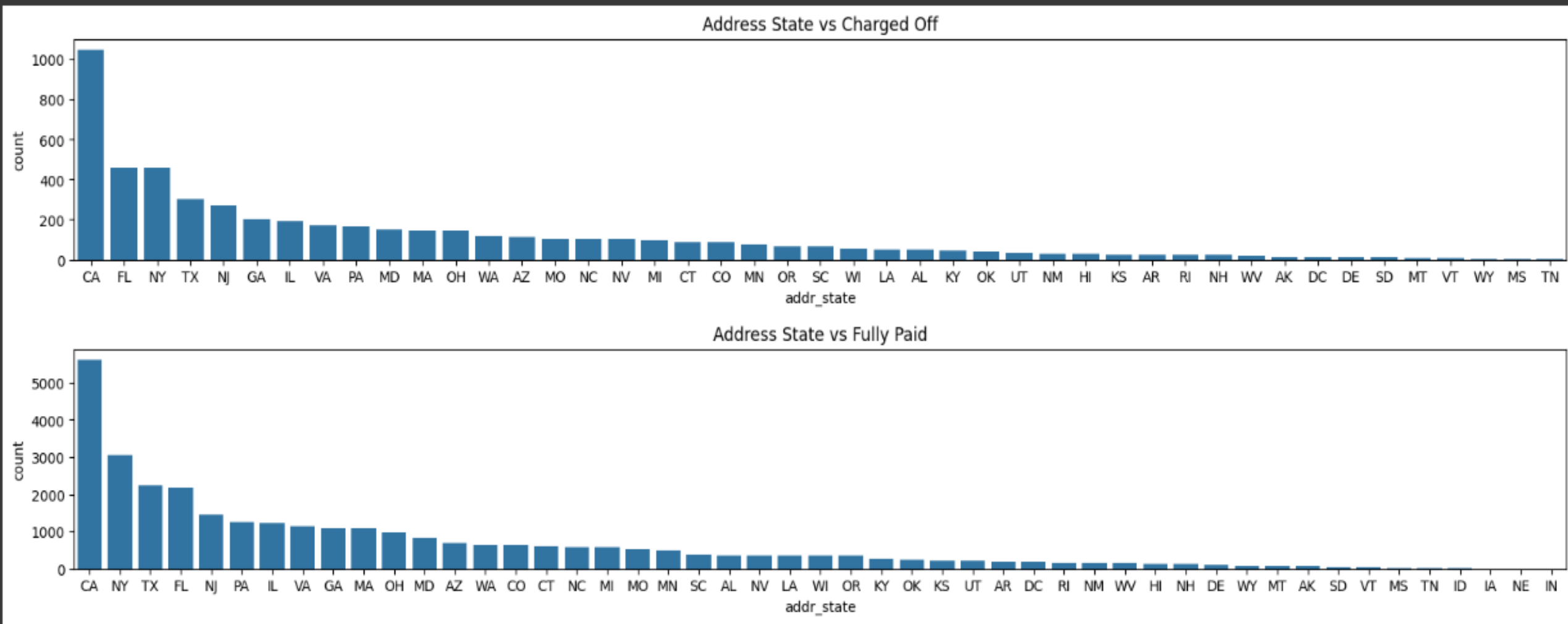
- Most Defaulters are from CA,FL and NY, so more precaution needs to be taken for CA as its more than double from FL,NY

Conclusion

- Loan amount with tenure of 60 months with high rates must not be disbursed as they are likely to be defaulted, for focus should be given on grades if disbursed.

- Lower Grade and High Interest Loans must be avoided as they have major chance to get defaulted.

- There is a high possibility of corruption within the organization as big amount loans were given after the borrower is verified.

- There should be a decrease in loan disbursement for small businesses, debt consolidation and credit card as they have a chance to get defaulted

- There must be less loan given to CA,NY and FL, so that the defaulters are less

- More experienced people with high amount of loan must be verified before giving loan