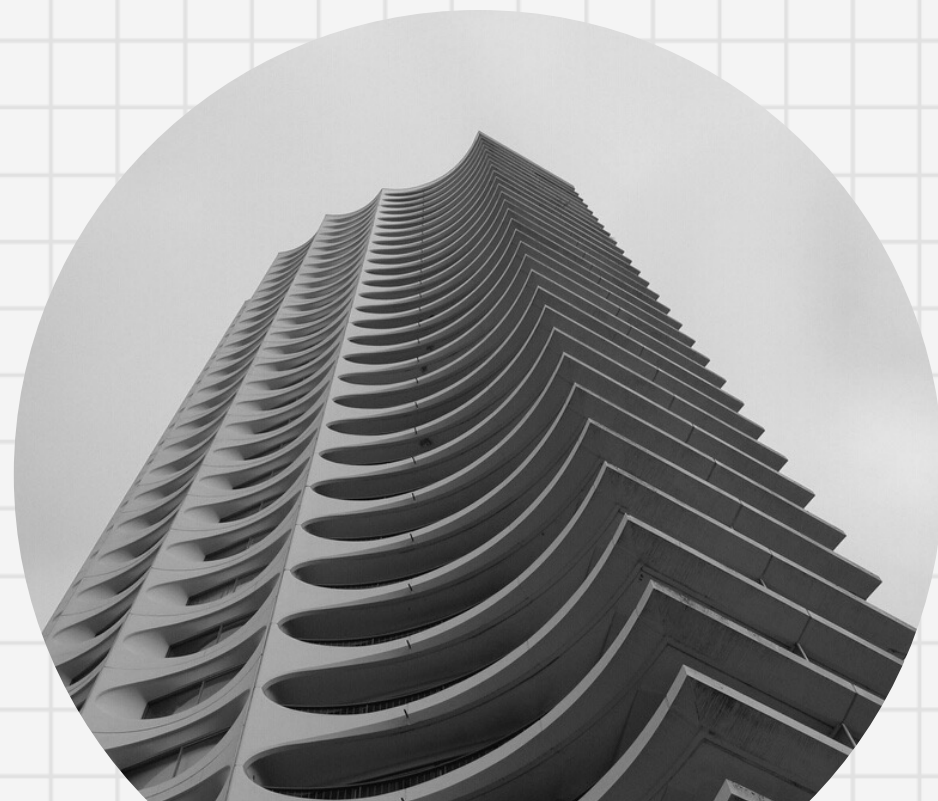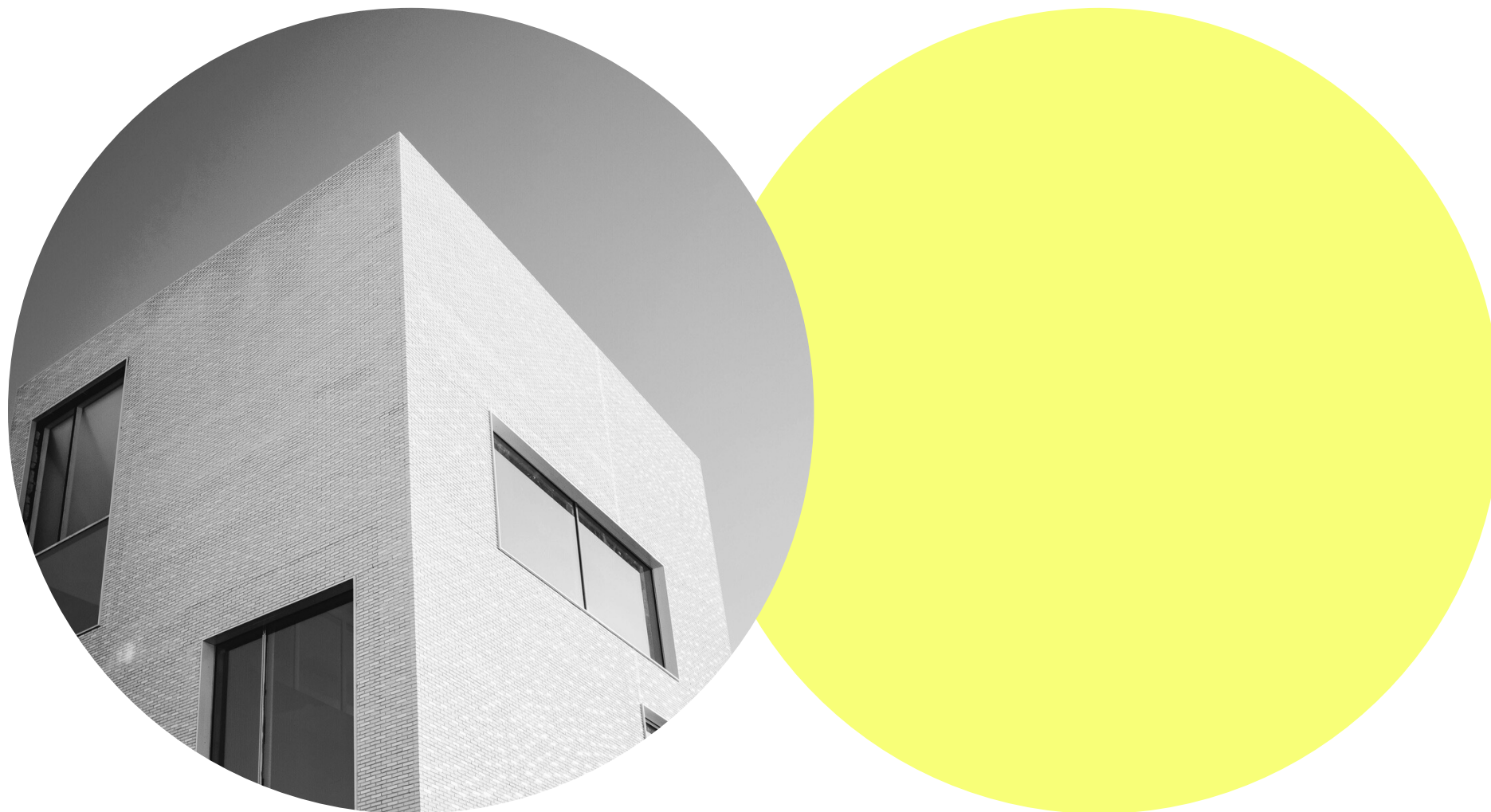# Lending Club Case Study

By Naveen V & Prabhat Kumar Singh
(Group – 100089)

Abstract

Business Objective

Understanding how consumer attributes and loan attributes influence the tendency of defaults. This analysis will help lending club for its portfolio and risk assessment.

Data Source

The data given contains the information about past loan applicants and whether they 'defaulted' or not.

Data Analysis Goals

Identify driving factors or variables which are strong indicators behind loan defaults.

Problem Solving Methodology

Step - 1

Step - 2

Step - 3

Step - 4

Step - 5

Understanding Business Objective

Data Understanding / Exploration

Data Cleaning

Data Analysis

**Conclusion**

Data Cleaning
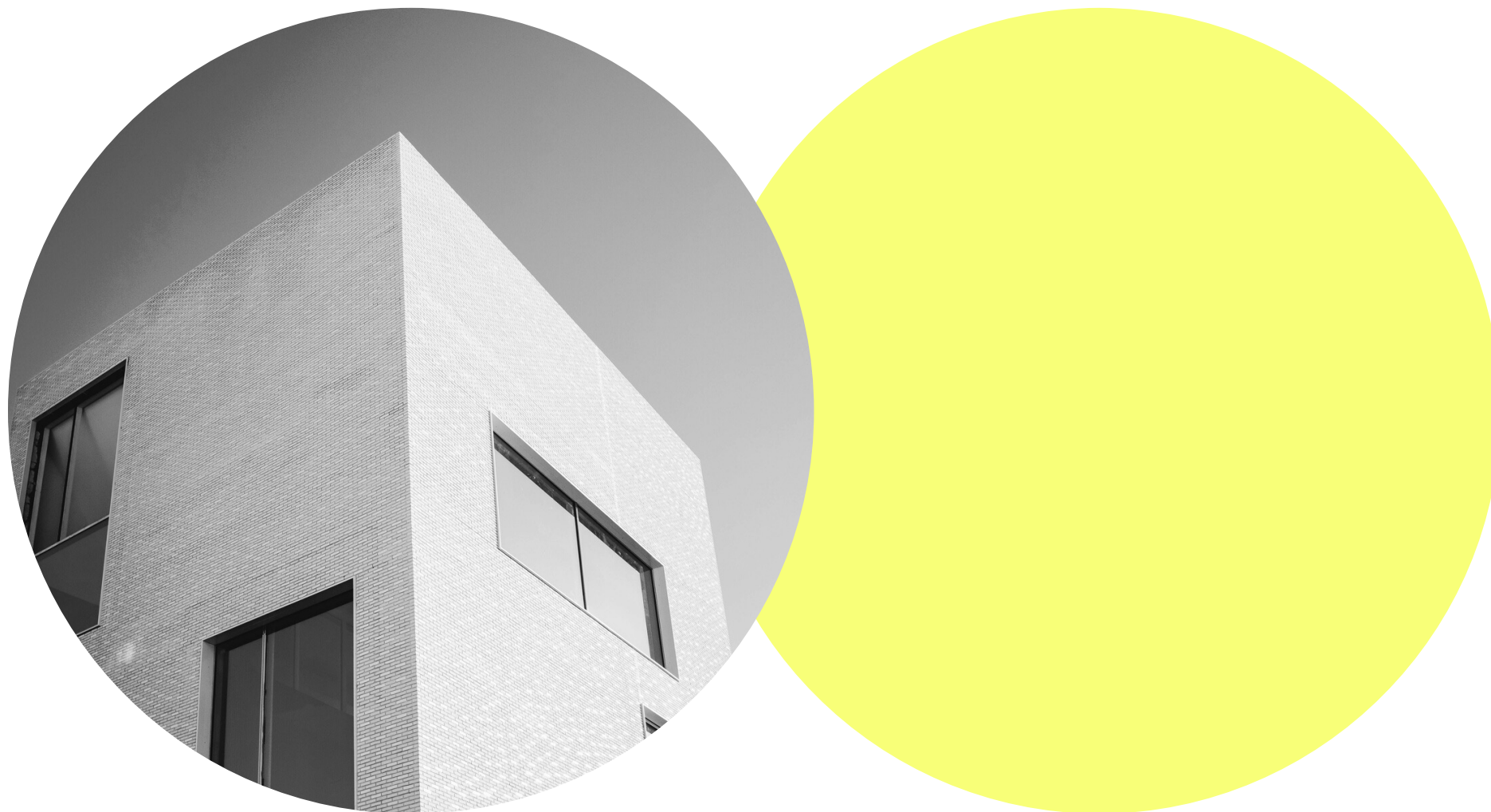
Steps

- All columns which have null values are dropped

- Dropped columns which have more than 60% of missing data

- Dropped column desc which contains similar information in columns purpose and title

- Dropped columns with single value

- Dropped rows which are with loan status as current as these entries will not give any info regarding variables for defaulters

- Dropped rows where bankruptcy details were empty.

- There are certain behaviour variables which are not available at the time of applying loans and hence these columns can be removed since they can't be used to predict loan approval.

Univariate Analysis

- From the plots we can conclude that majority of the cases loan amount(loan_amnt) requested are approved by approver(funded_amnt) and investors(funded_amnt_inv).

- **Hence we can consider Loan amount itself for further analysis**
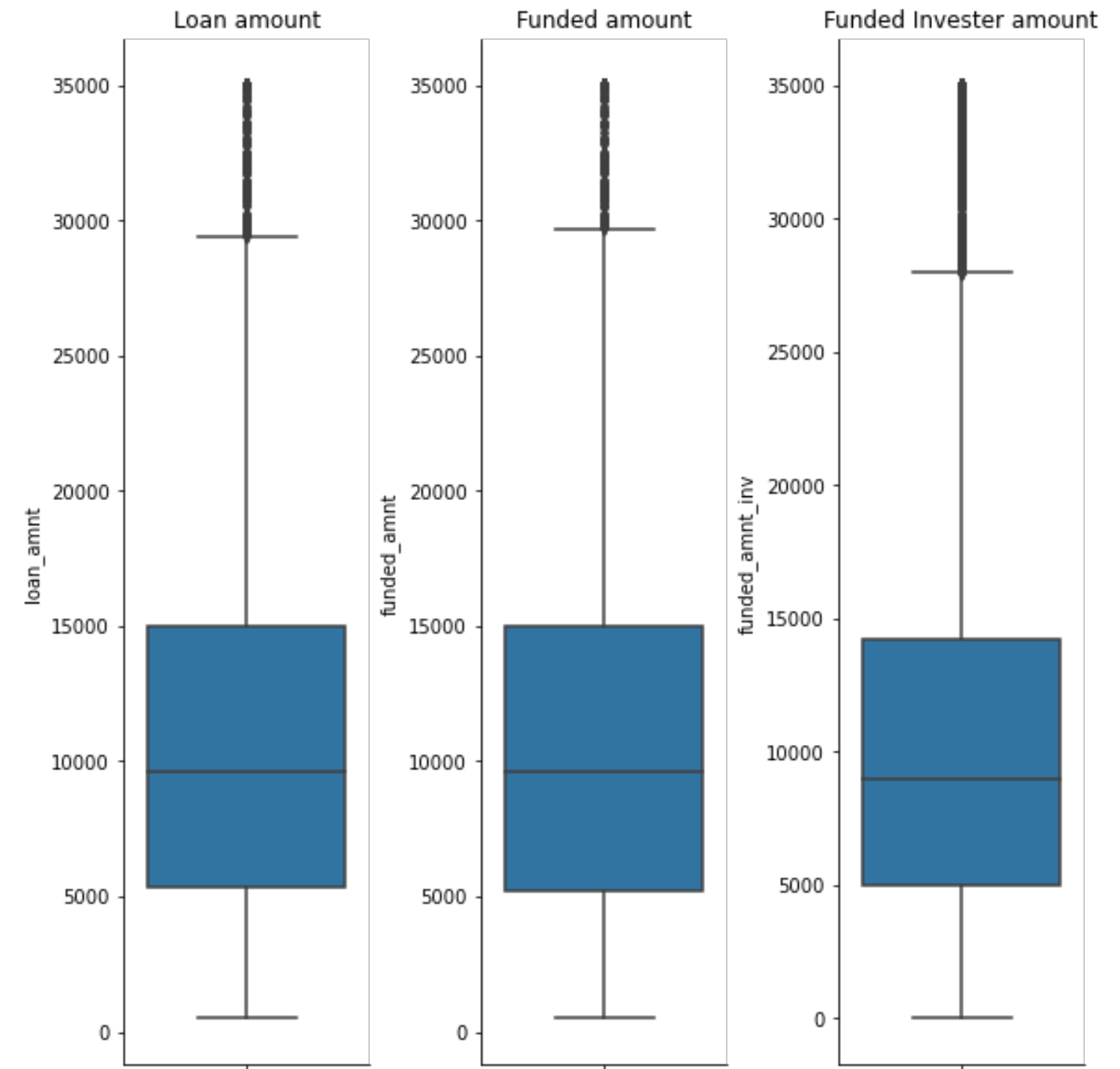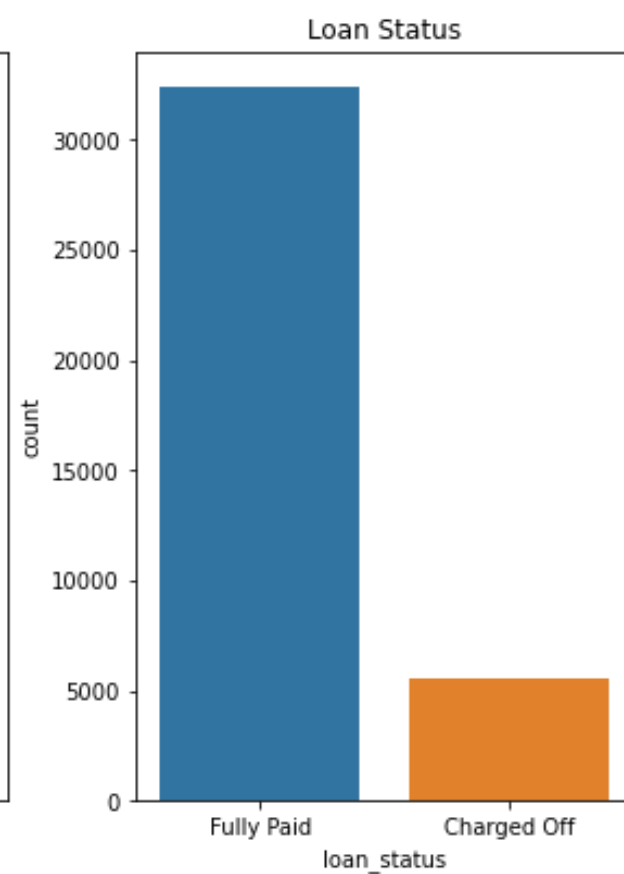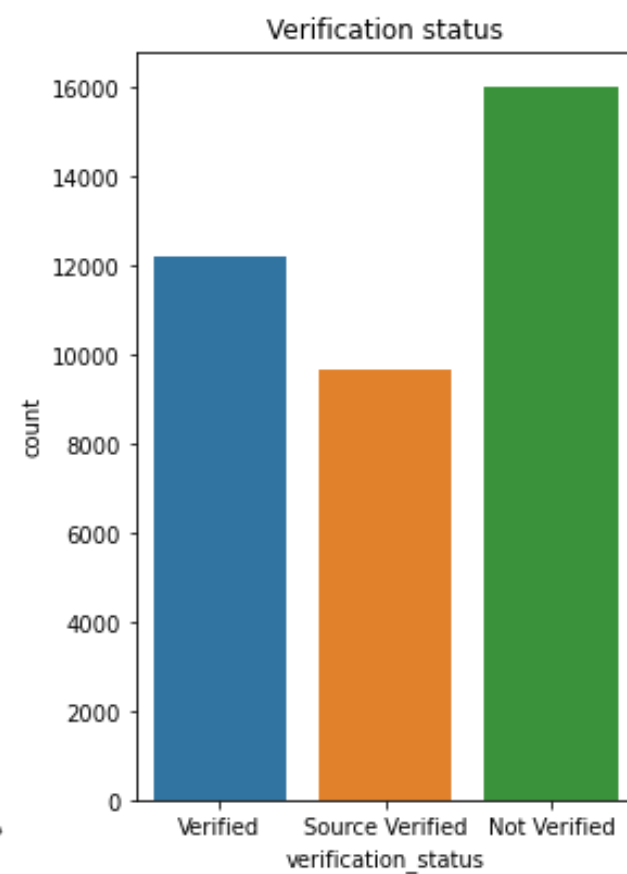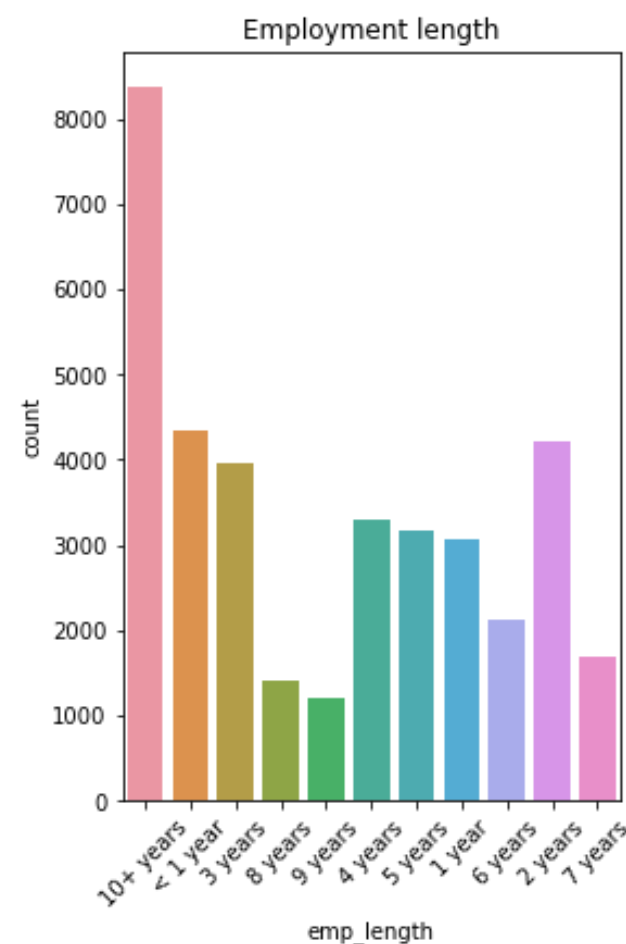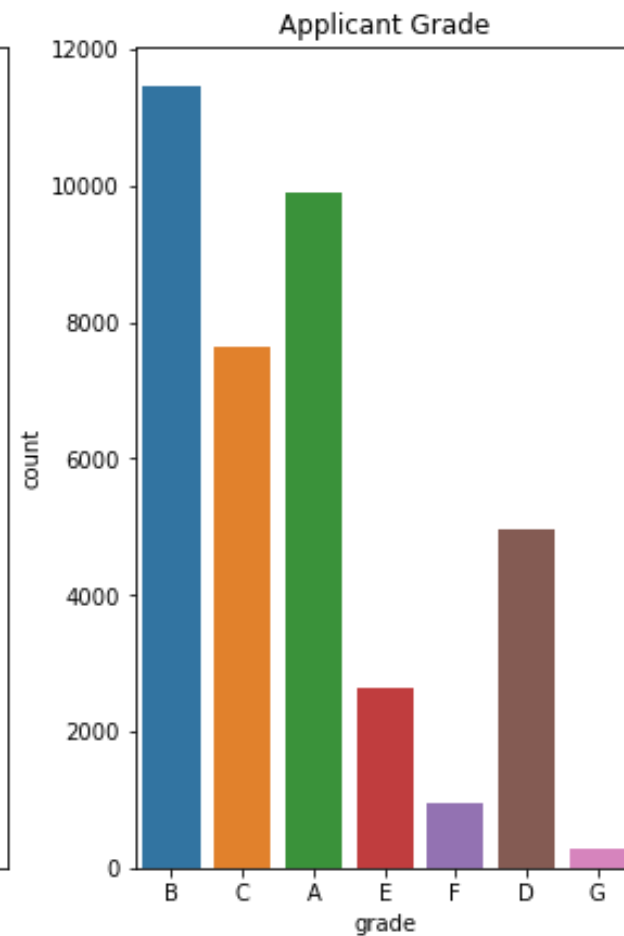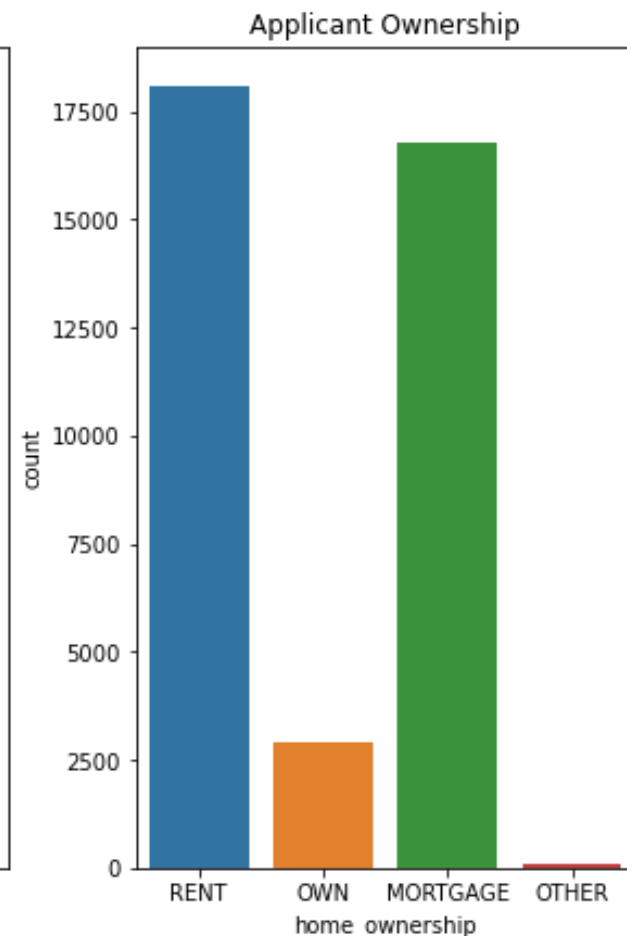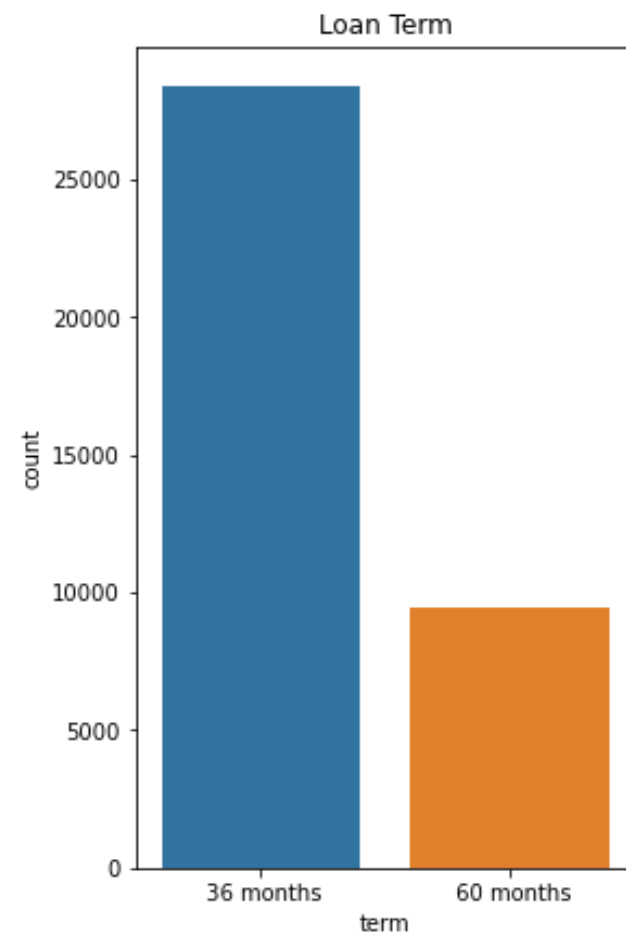
1. **Term** : People opting for short team loan (36 months) are relatively higher compared to long term loans (60 months)
2. **Home ownership** : Applicant who are having Rent or Mortgage property are more intended to take loan over others
3. **Grade** : People from grade A and B are among the most who take the loan
4. **Employment Length** - Among applicants employees having maximum group of people are having 10+ years years experience
5. **Verification Status** - For provided data we can infer majority of the loans are issued without verification
6. **Loan Status** : About 85% of loans are fully paid and 15% of loans are Charged off
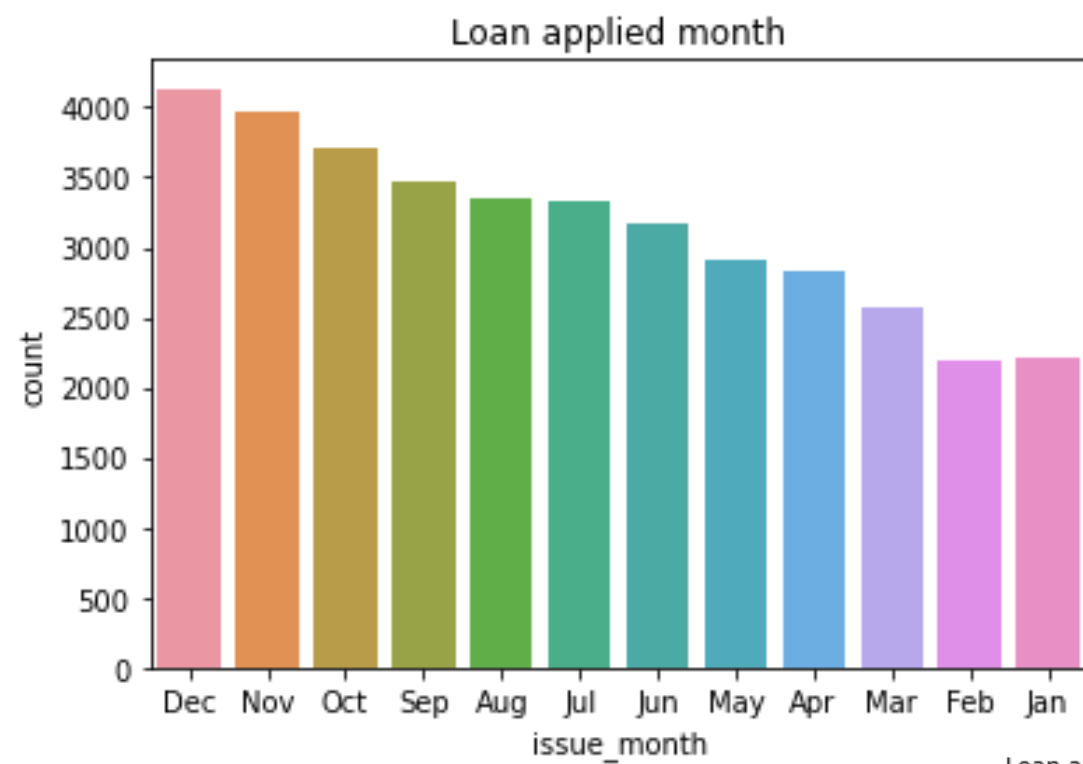
1. **Loan issue month** : Number of people opting for loan increases as month progresses from Jan to Dec.

2. **Loan applicant State** : People are from CA and NY are most who applied for the loan.

3. **Pupose** : Debt consolidation is the most purpose for people who are applying for the loans
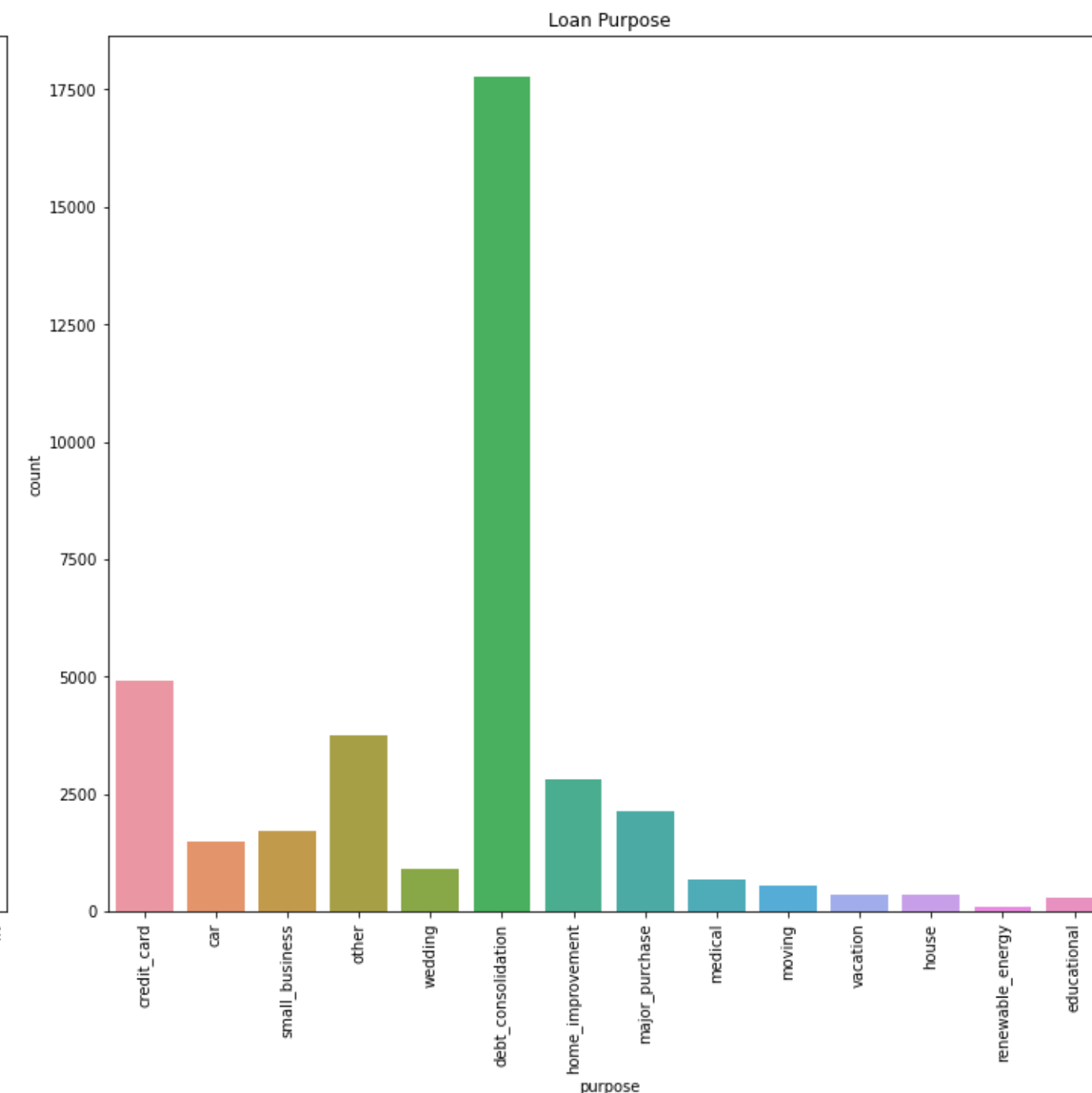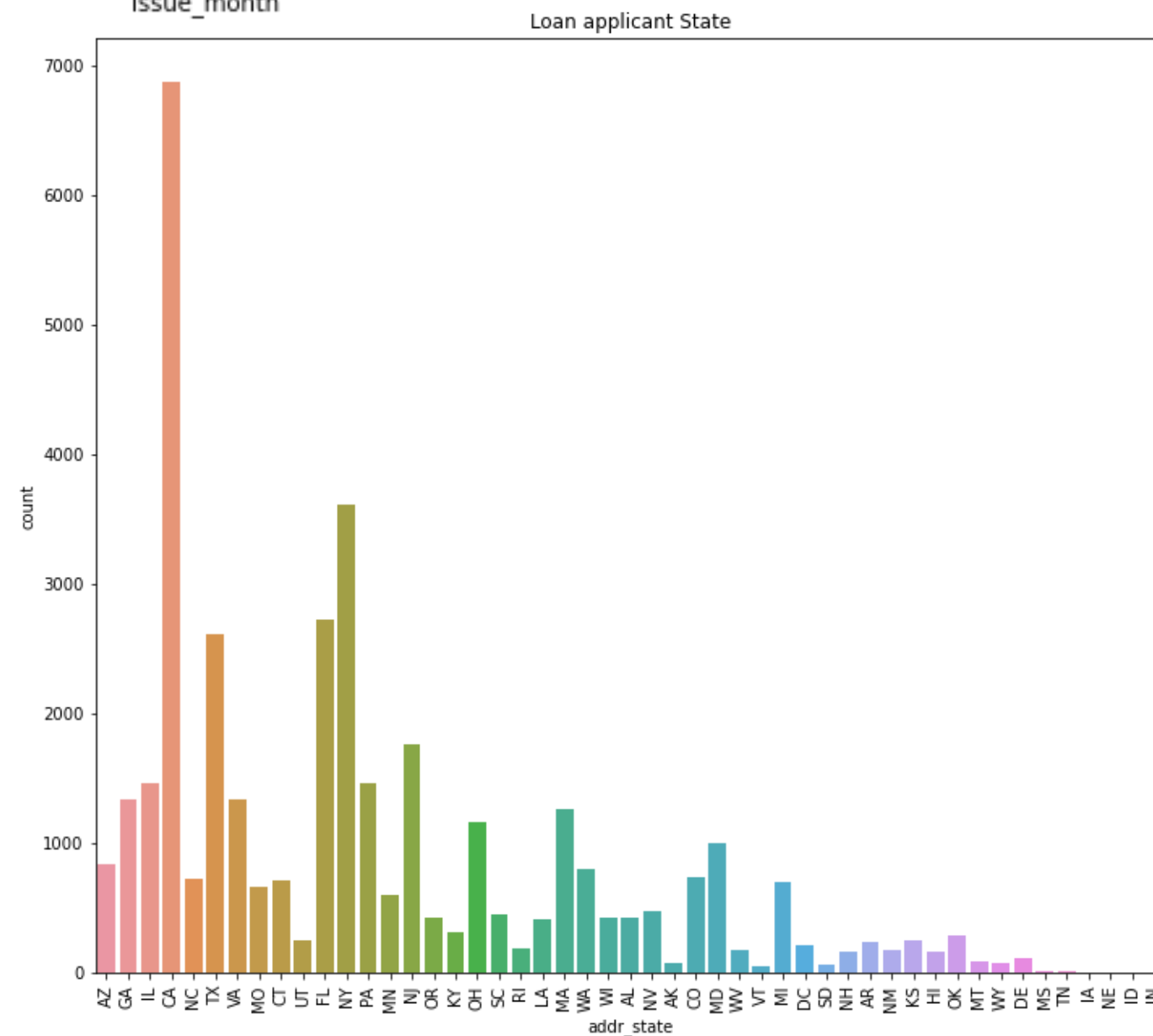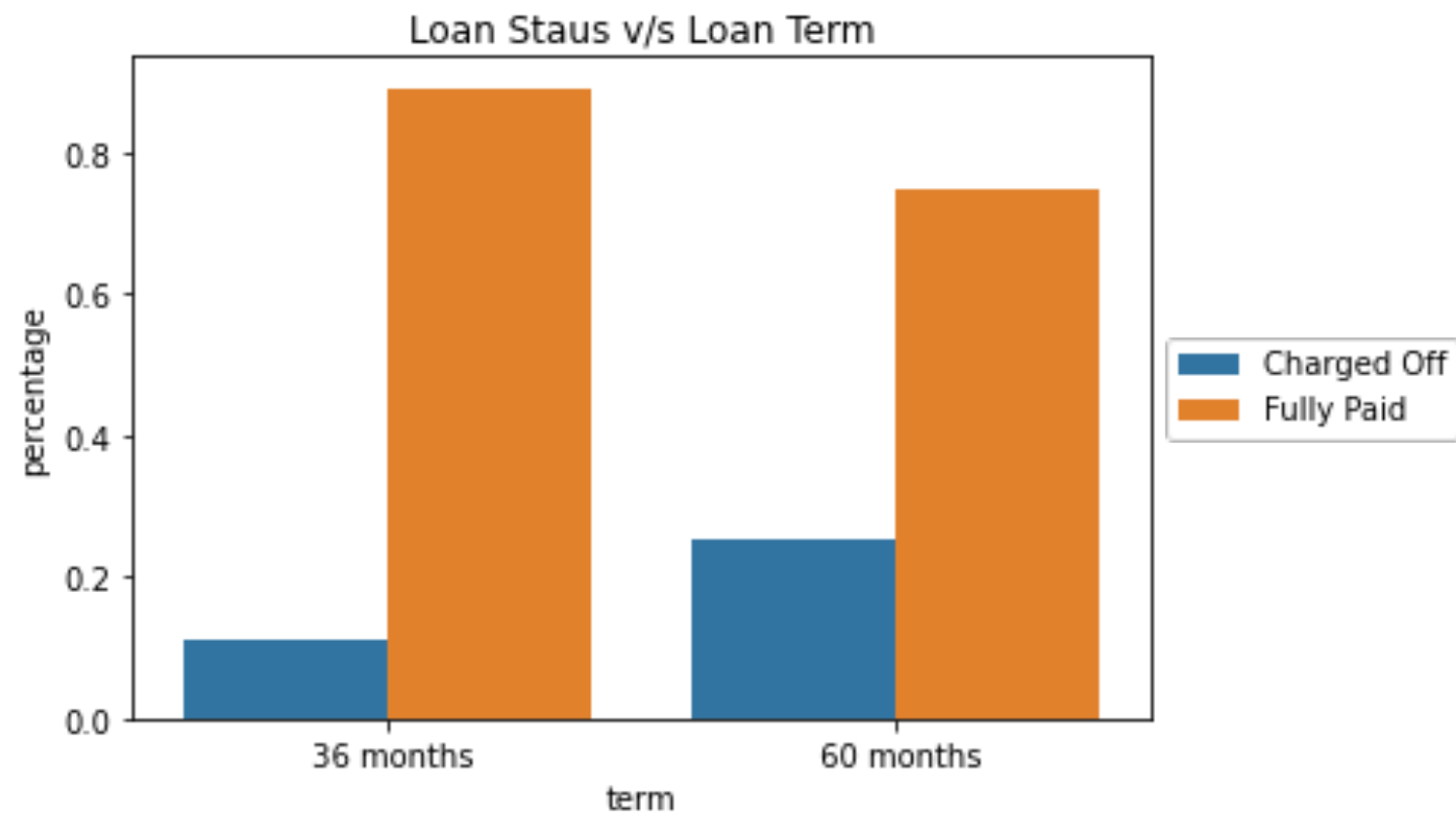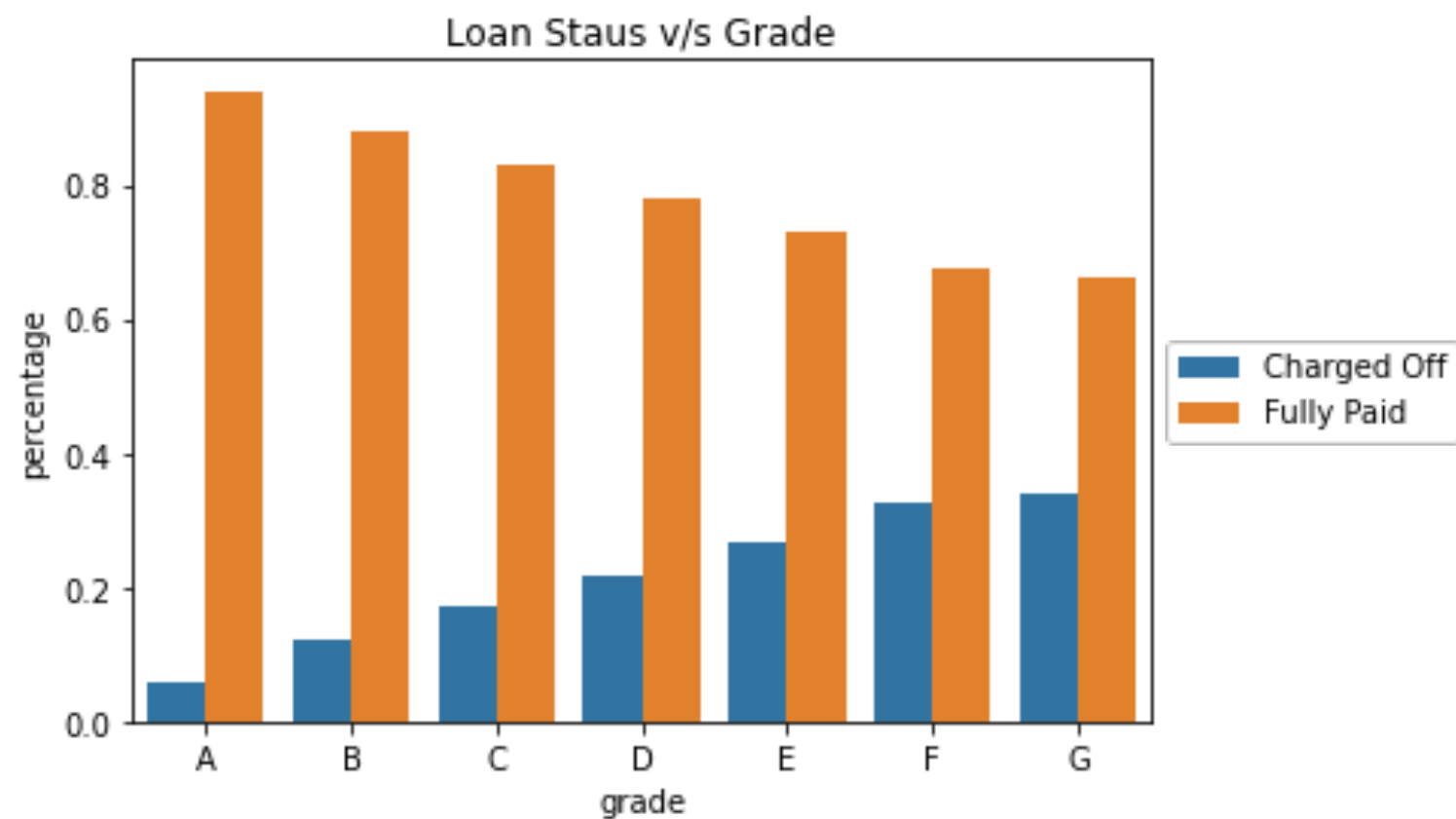
Bivariate Analysis

Loan Status vs Interest Rate

People belonging to Charged off category are having higher interest rate than fully paid ones
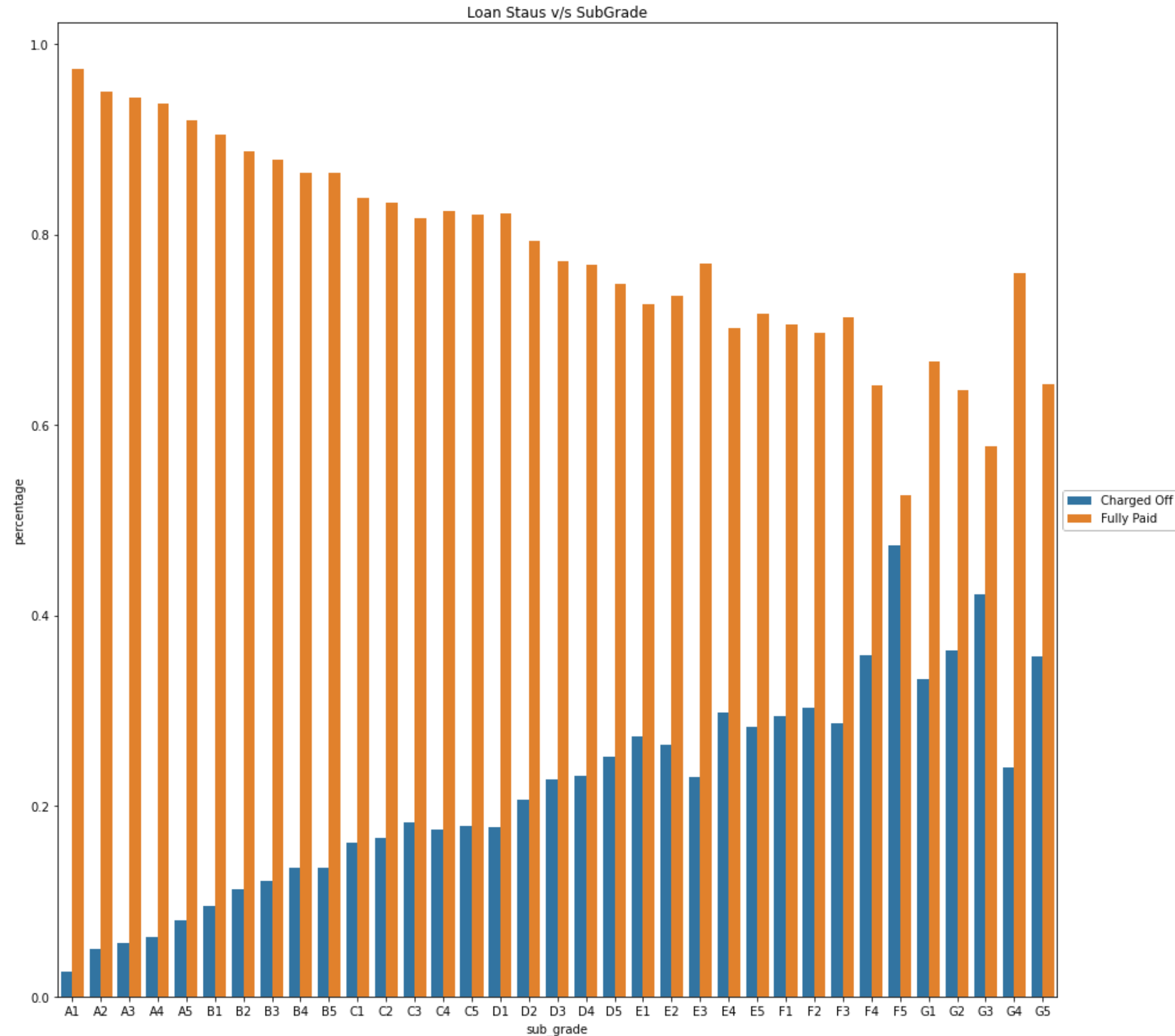
Loan Staus v/s Loan Term
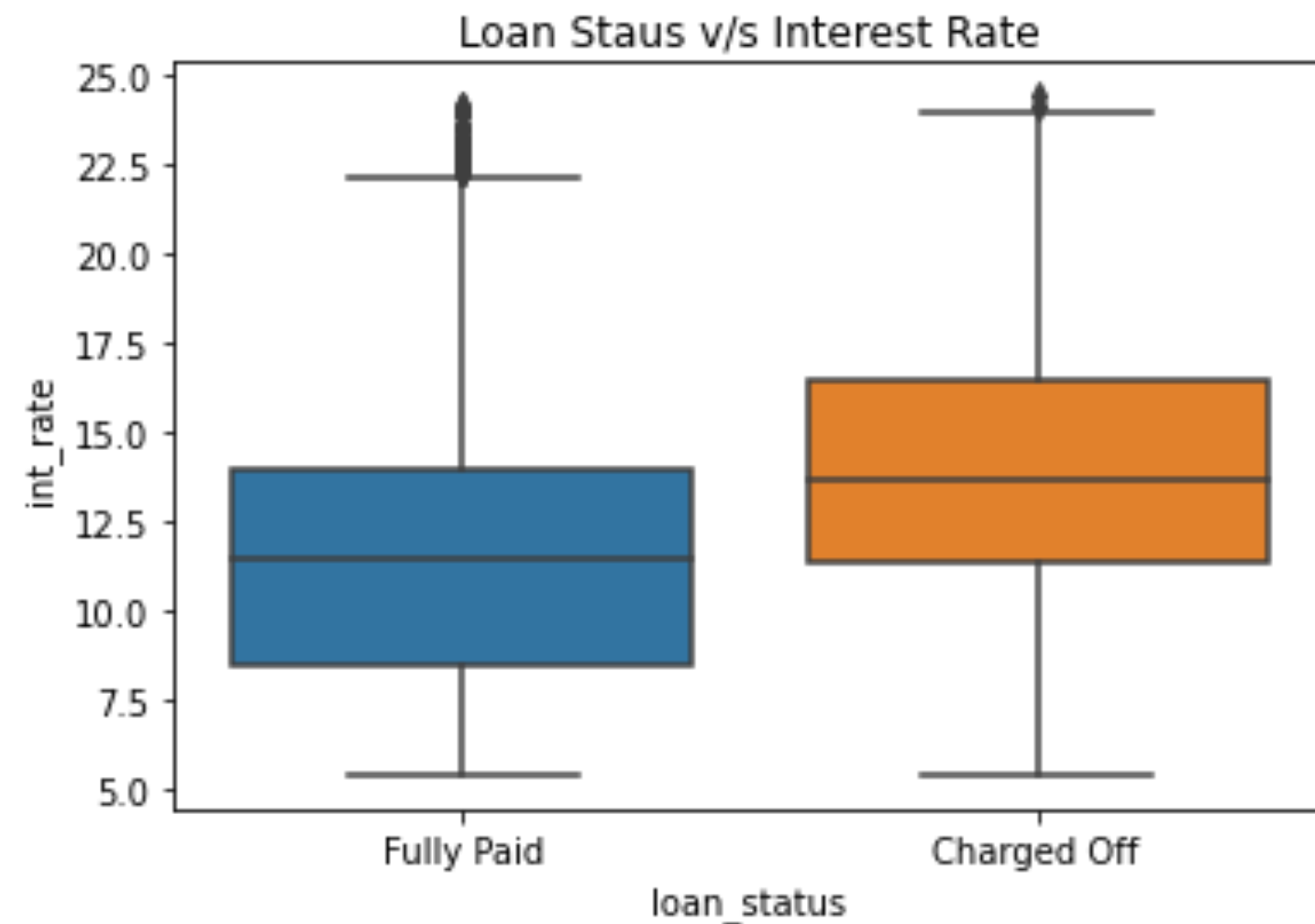
- Number of people Charged off are more in long term loan(60 months) compared to short term. Also number of people fully paid is more in short term loan(36 months).

Loan Staus v/s Grade

- Number of people belong to Group F & G are Charged off more compared to other group. Also from as we move from Group A to G proportion of charged off people increases.

As a proof of our observation from loan status vs group analysis, we see incremental curve for charged off people from subgroup A to G. Among the sub group there are more people from F4,F5,G2,G3 and G5 belongs to Charged Off category.

Loan Staus v/s purpose
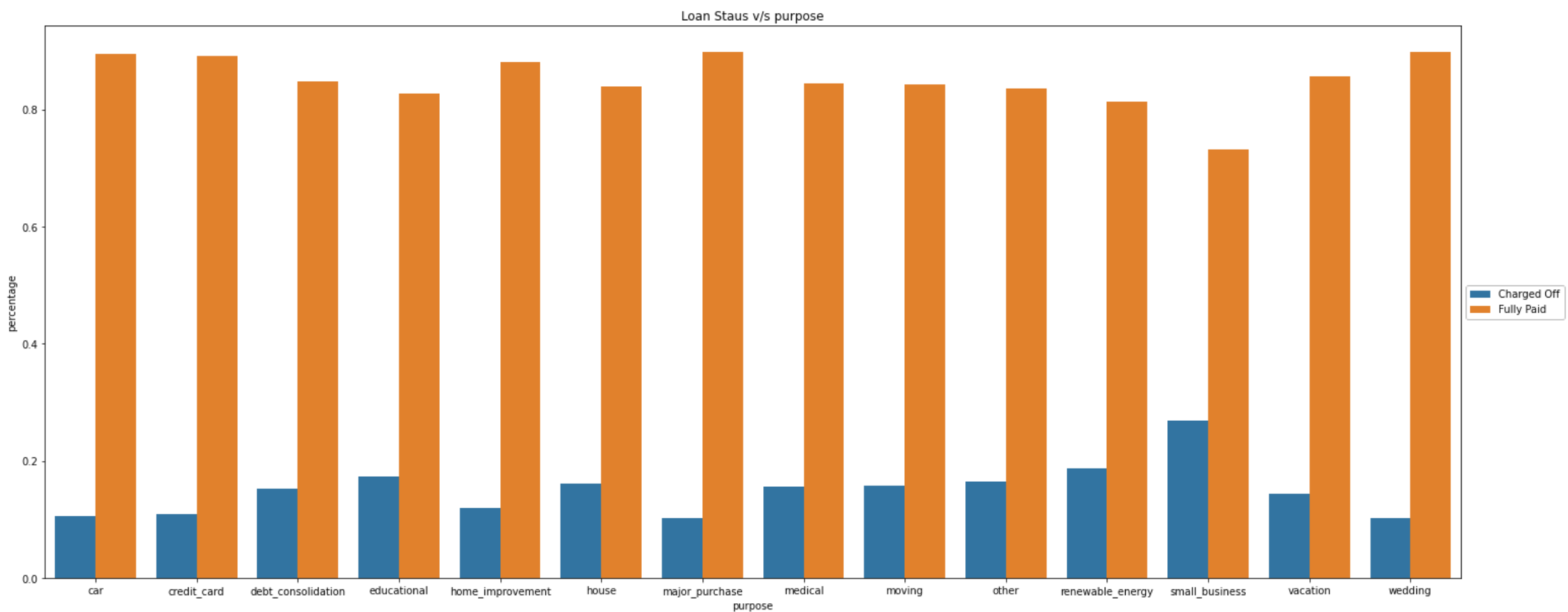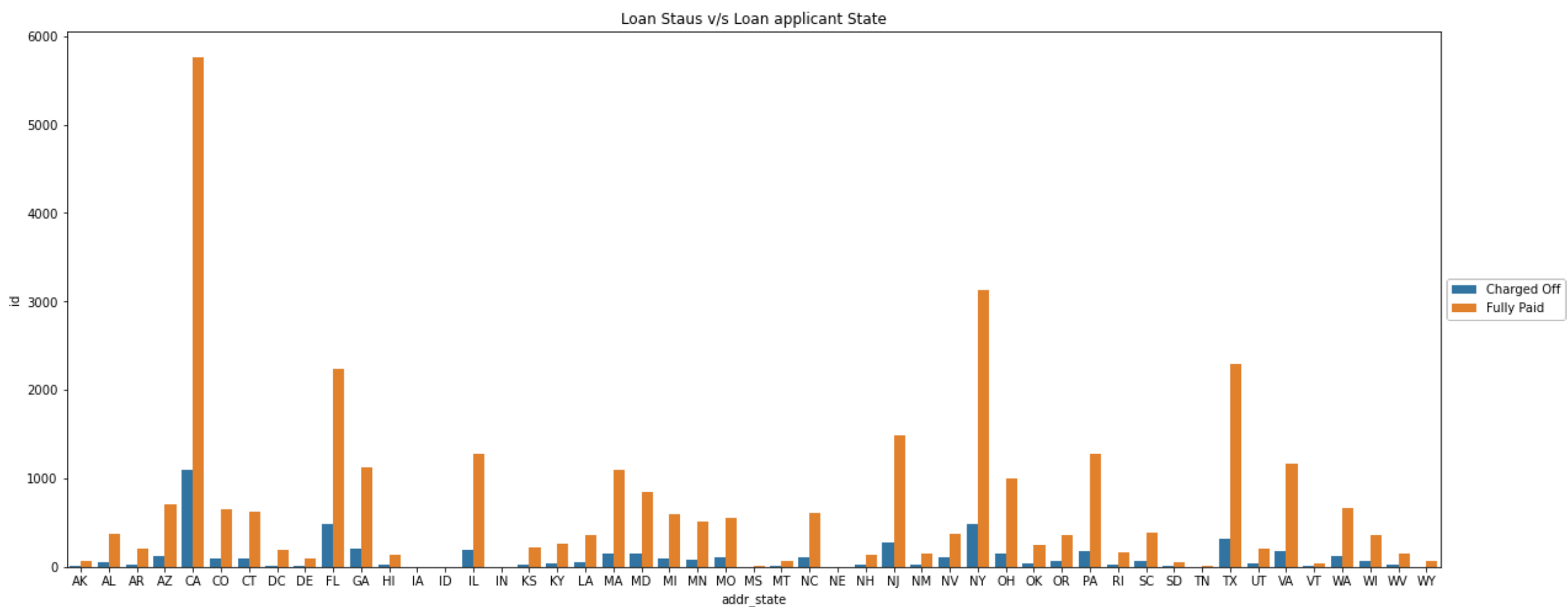
- People having purpose of the loan as **'small_business'** have higher proportion of Charged off people compared to others.

Loan Staus v/s Loan applicant State

- Among the people who belong to Charged off category, frequency of people belong to State 'CA' is more compared to other status. Also there are very less people under charged off category for states 'AK','DC','TN' and 'WY'.
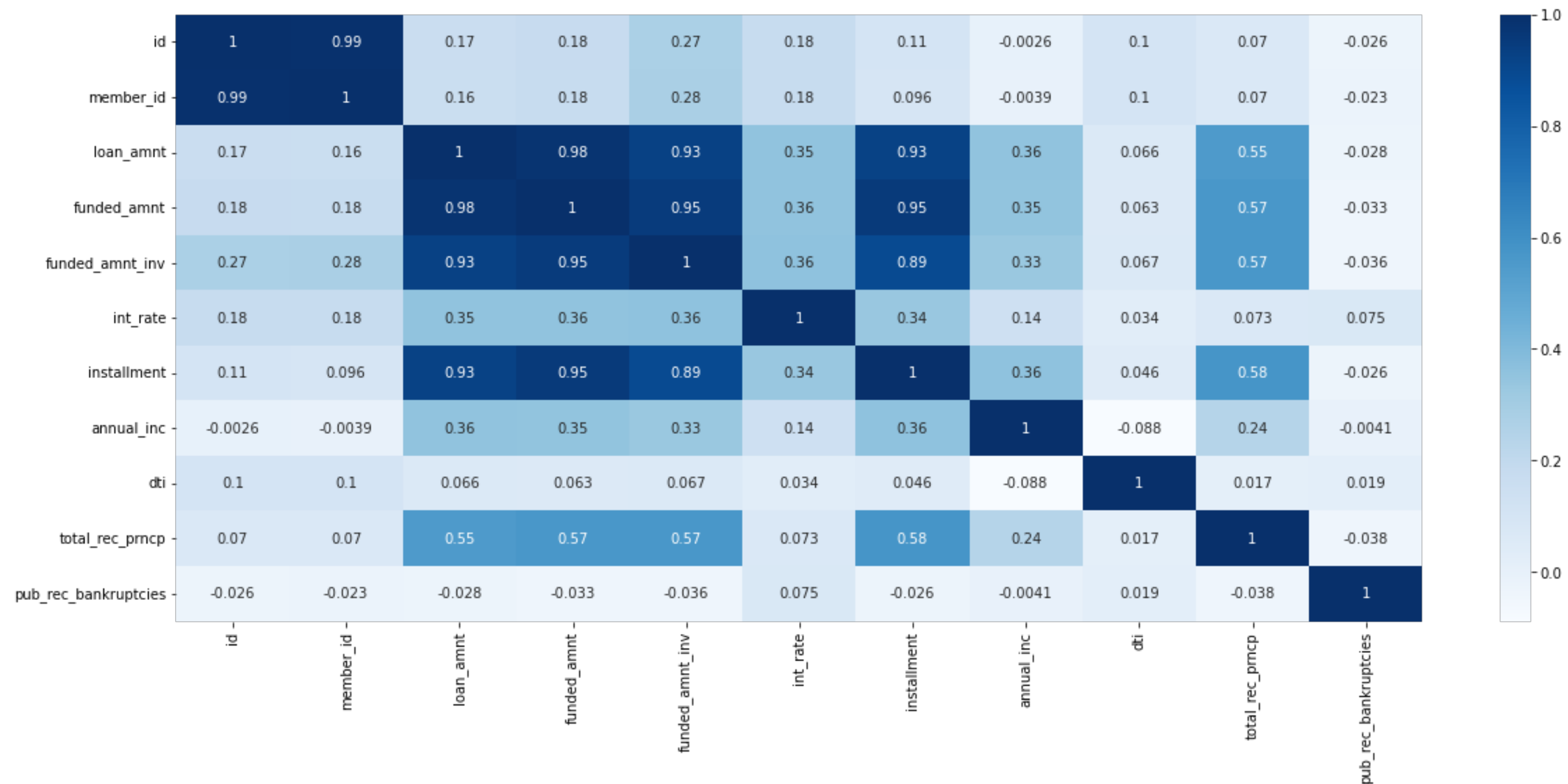
Below are the observations for relation between columns for Charged off loans
- There is strong positive correlation between loan amount with funded amount and funded amount is invested, this is because almost all loan amount request will be funded close to requested amount.
- Positive correlation between loan amount with instalments - instalment increases with increase in loan amount.
- There is negative correlation between annual incomes with DTI, increase in annual incomes decreases DTI.

Important Recommendations

1. Loans with high interest should be reduced as they are prone to defaults.

2. Reduce loans with higher tenures as they are charged off more.

3. More information to be collected before granting loans for borrowers with lower grades.

4. More attention should be paid while issuing loans towards the end of an year as chances of defaulting is high.

5. Small businesses tend to default loans more, recommendation is to stop providing loans to small businesses.

6. Higher DTI may lead to defaulters.

Thank you!

Feel free to reach out to us
if you have any questions.

Naveen V

Github UserId - @Naveenkv93

Prabhat Kumar Singh

Github UserID - @prabhats97