Problem statement

The customer finance company wants to decide for loan approval based on the risk of the applicant's profile, For which based on the sample data available need to find out the driving factors behind the loan default. Company can use this information for conducting Risk Analysis

Analysis Approach

» Understand all the columns from the data dictionary

» Clean the data

  - Drop the columns if it has more than 70% rows are NA

  - Make Derived columns from existing columns for analysis

  - Remove the columns that are not useful for the analysis

  - If the columns has less than 5% NA , Replace it with mode of the respective column

  - Remove unwanted signs like %,years etc.. From the number columns and change the same to number format

  - Remove the outliers in data

» Analyze the data

  - Make a list for continuous column and categorical column separately

  - Perform univariate analysis (histogram , box plot ,count plot)

  - Perform Bivariate analysis(Scatter plot , boxplot , bar plot)

  - Perform multivariate analysis(heatmap)

» Understand the graphs and list the driving factors for Loan to become default

Univariate Analysis

» Interest Rate is between 10 to 15%

» Loan Amount is between 5000 to 15000

» Installment is between 200 to 400

» Annual income is between 40000 to 80000

» Dti is between 10 to 20

» Around 14.2% of Loan is charged off

» Around 5000 applicants are charged off

» Those who had taken loan to repay in 60 months had more % of number of applicants getting charged off as compared to applicants who had taken loan for 36 months.

» The home ownership graph shows applicant with rented home or mortgage has more charged off

» If the number of loans are increasing there is a more possibility of charged off number getting increased
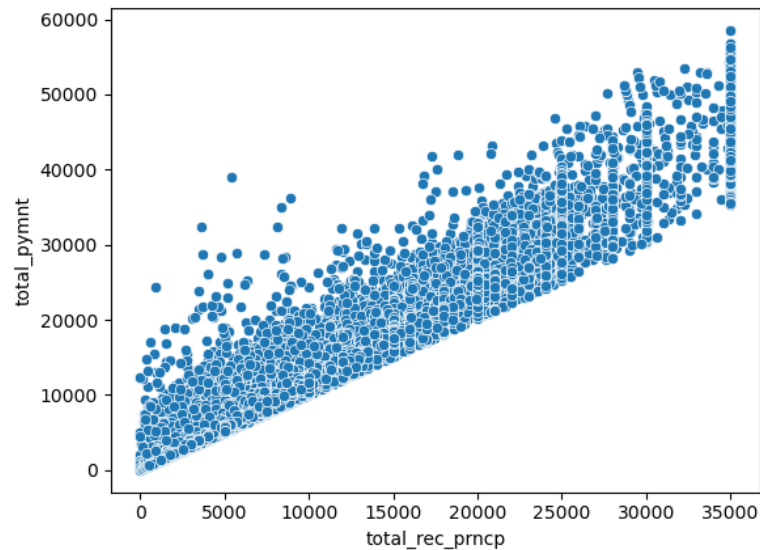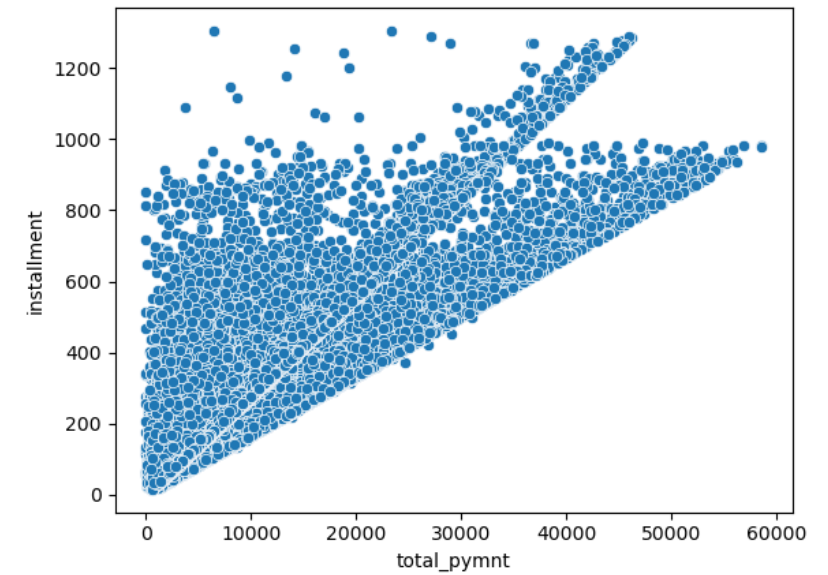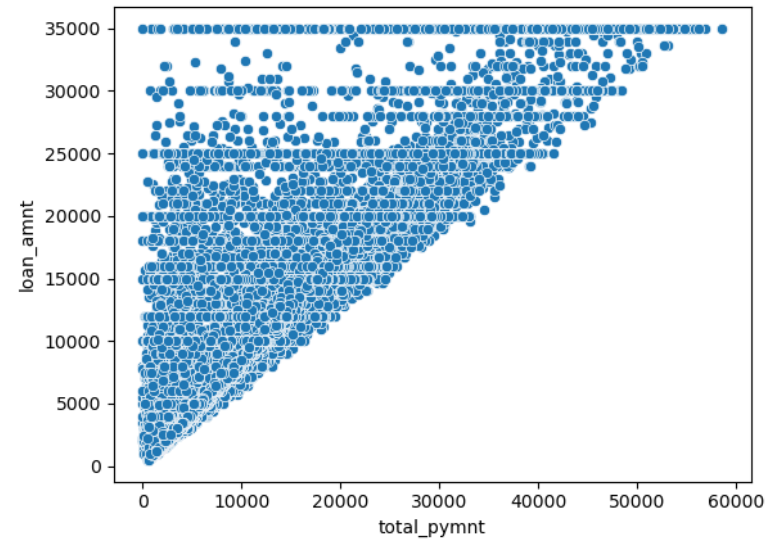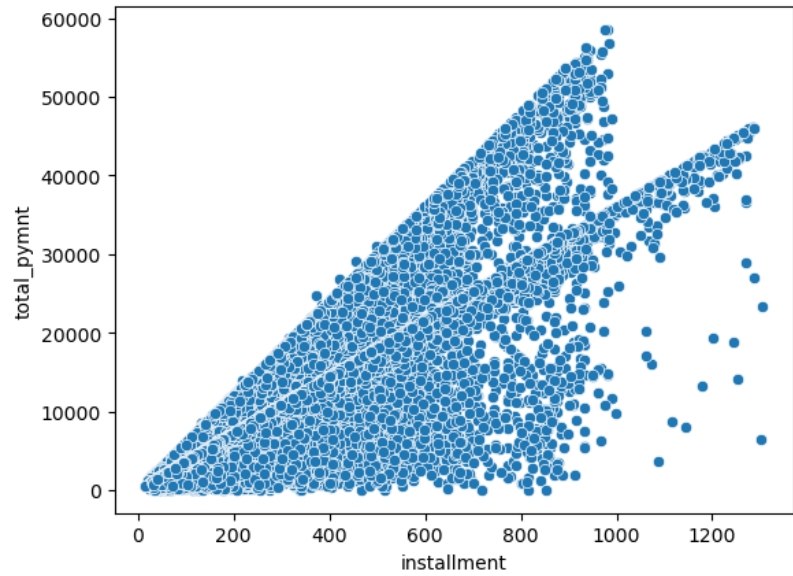
Univariate Plots
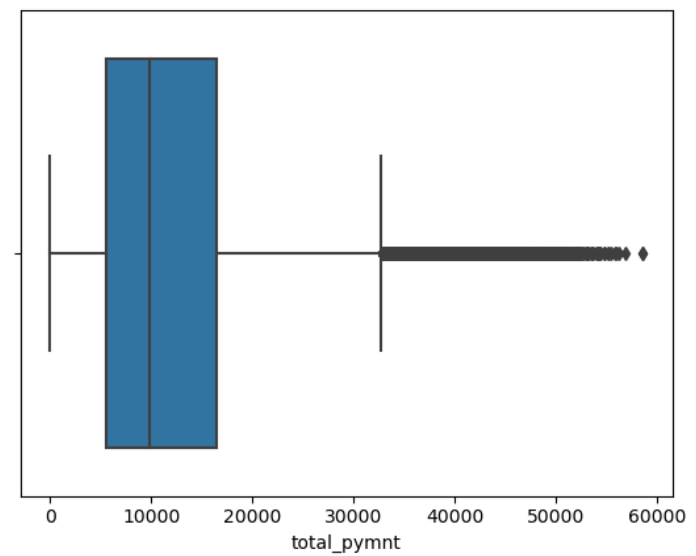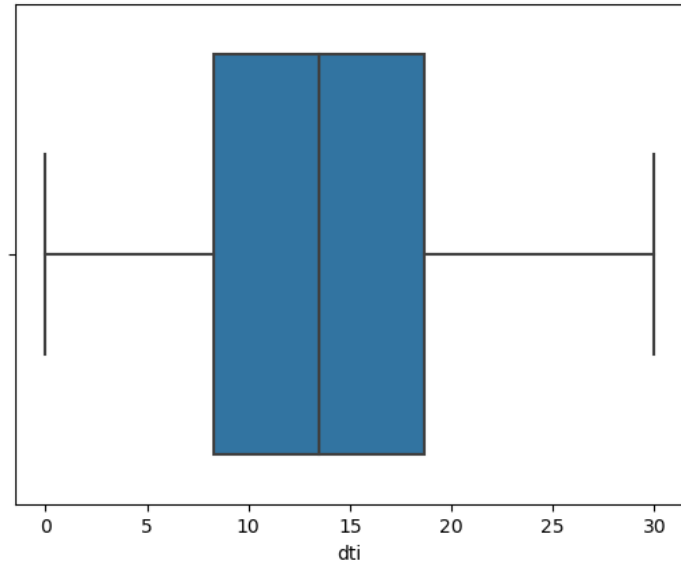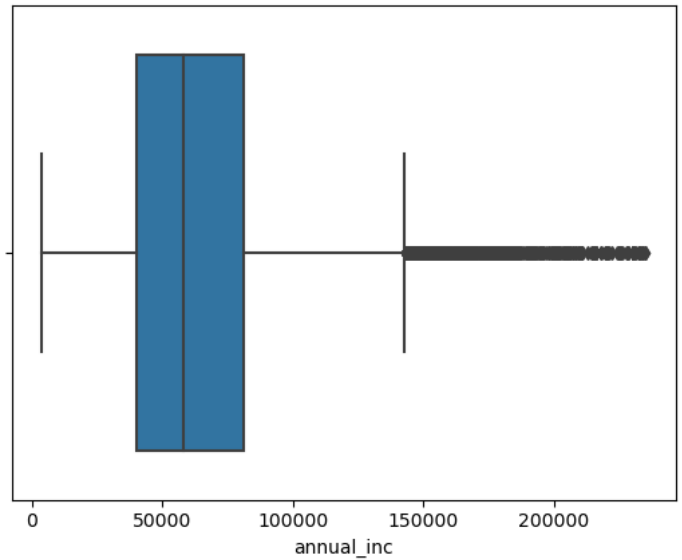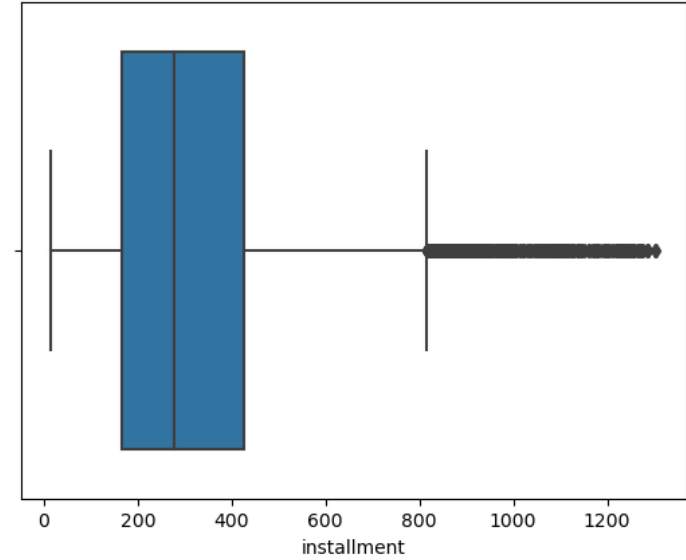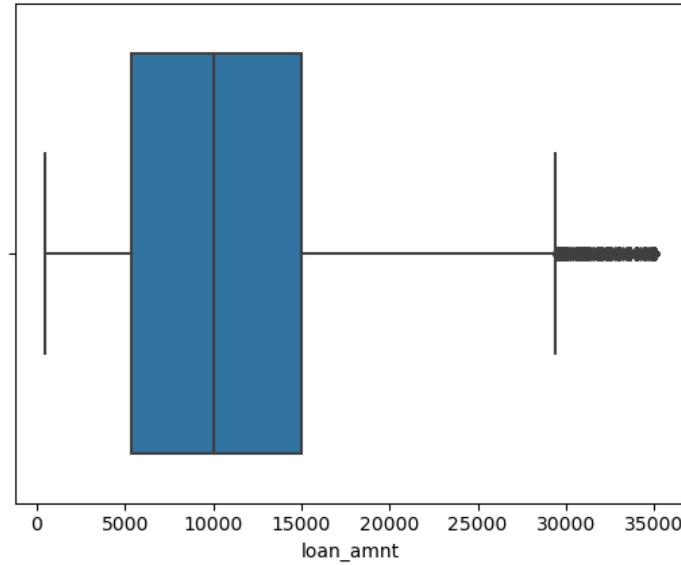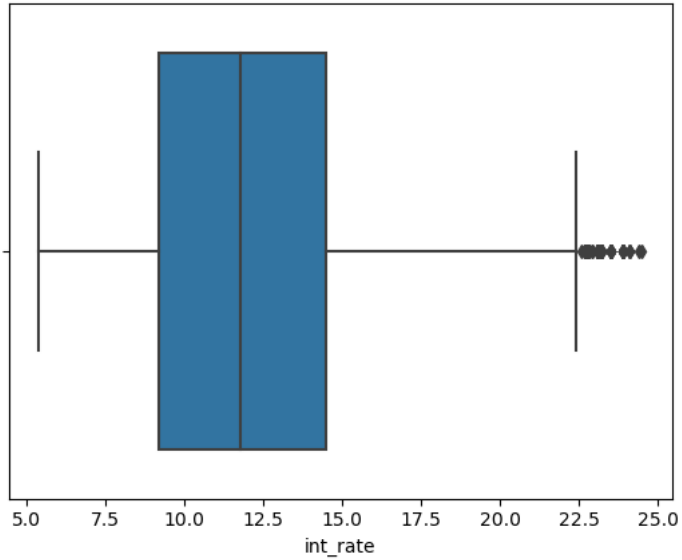
Univariate Plots
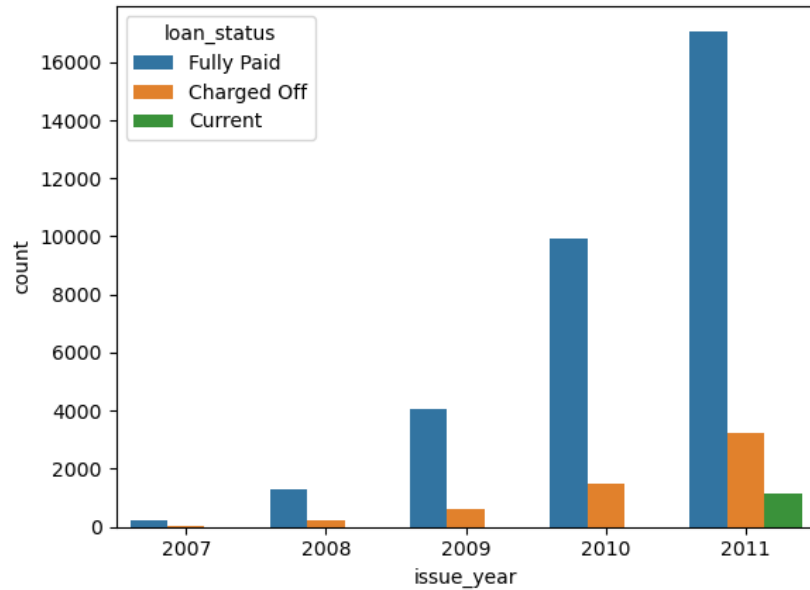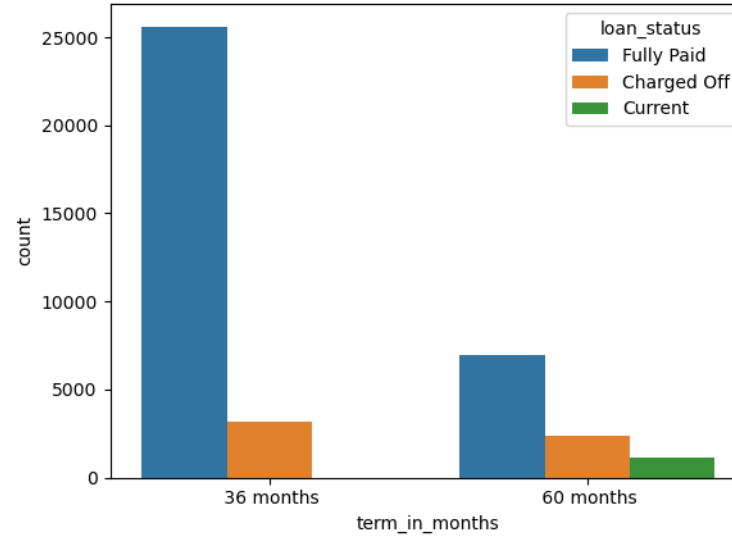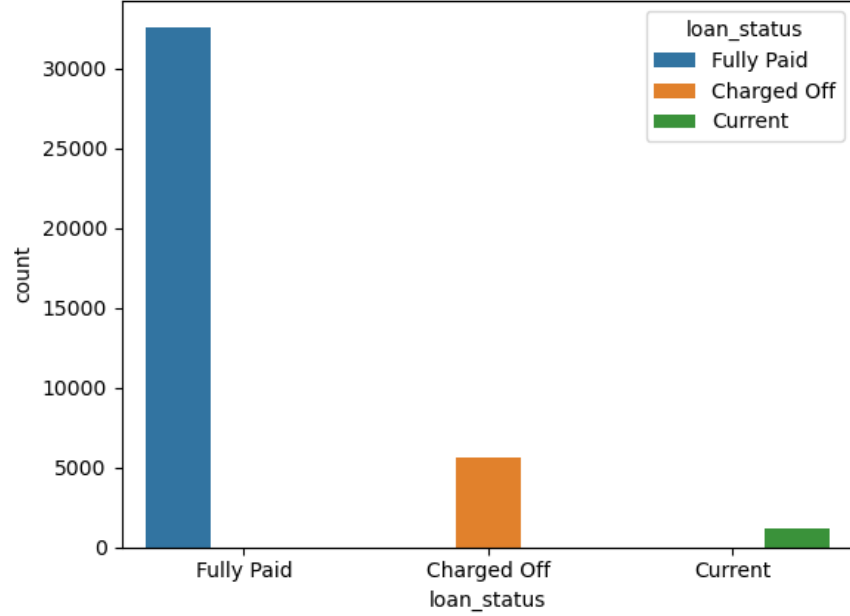
Bivariate Analysis

» Installment and Total Payment have a very good relationship

» Loan amount and Total payment have a very good relationship

» Total Payment and Installment have a very good relationship

» If annual income is more the charged of ratio is less

» Grade F&G has very high chance of getting charged off and the ratio increase with grades moving from A to G

» The Applicants who do small business has a very high chance of getting charged off.

» Charged off proportion is increasing with higher interest rates

» Employee whose experience is less than one year will be having a possibility to get charged off

» If the derogatory public records is 1 or 2 there is a more chance of getting charged off
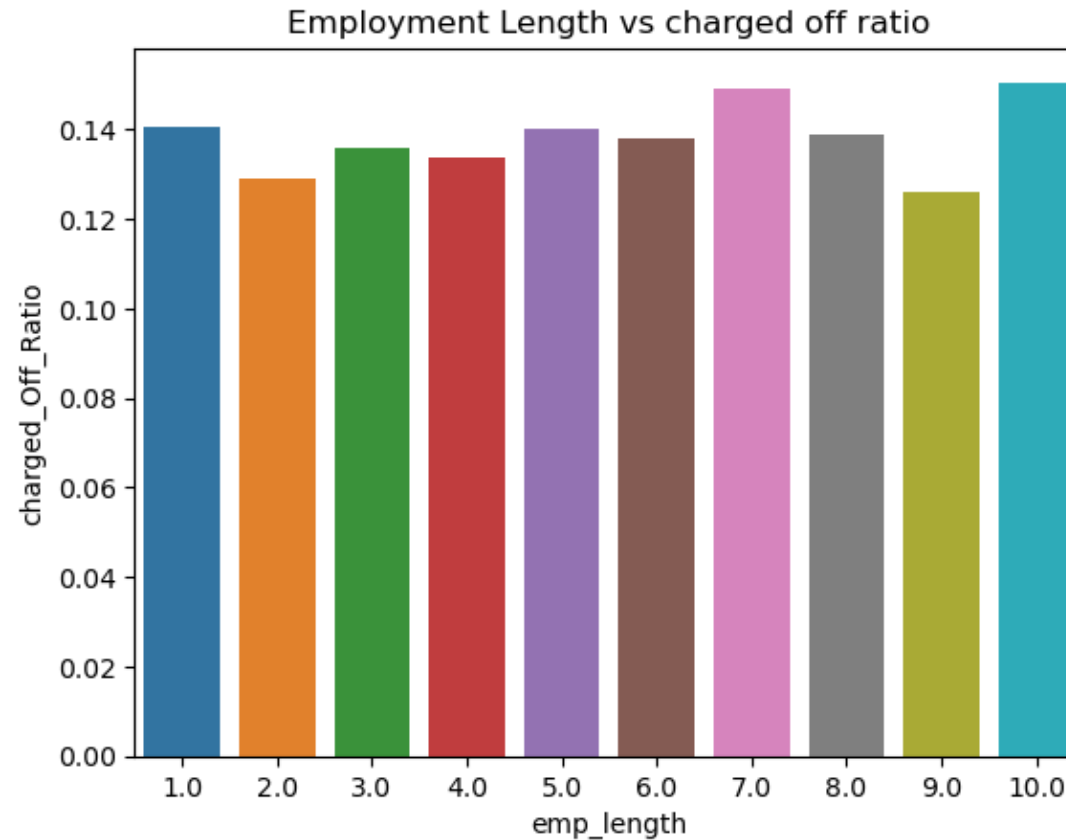
Bivariate Plots

1. Installment and Total Payment have a very good relationship
2. Loan amount and Total payment have a very good relationship
3. Total Payment and Installment have a very good relationship

Bivariate Plots

```
loan_new.emp_length.value_counts()
```
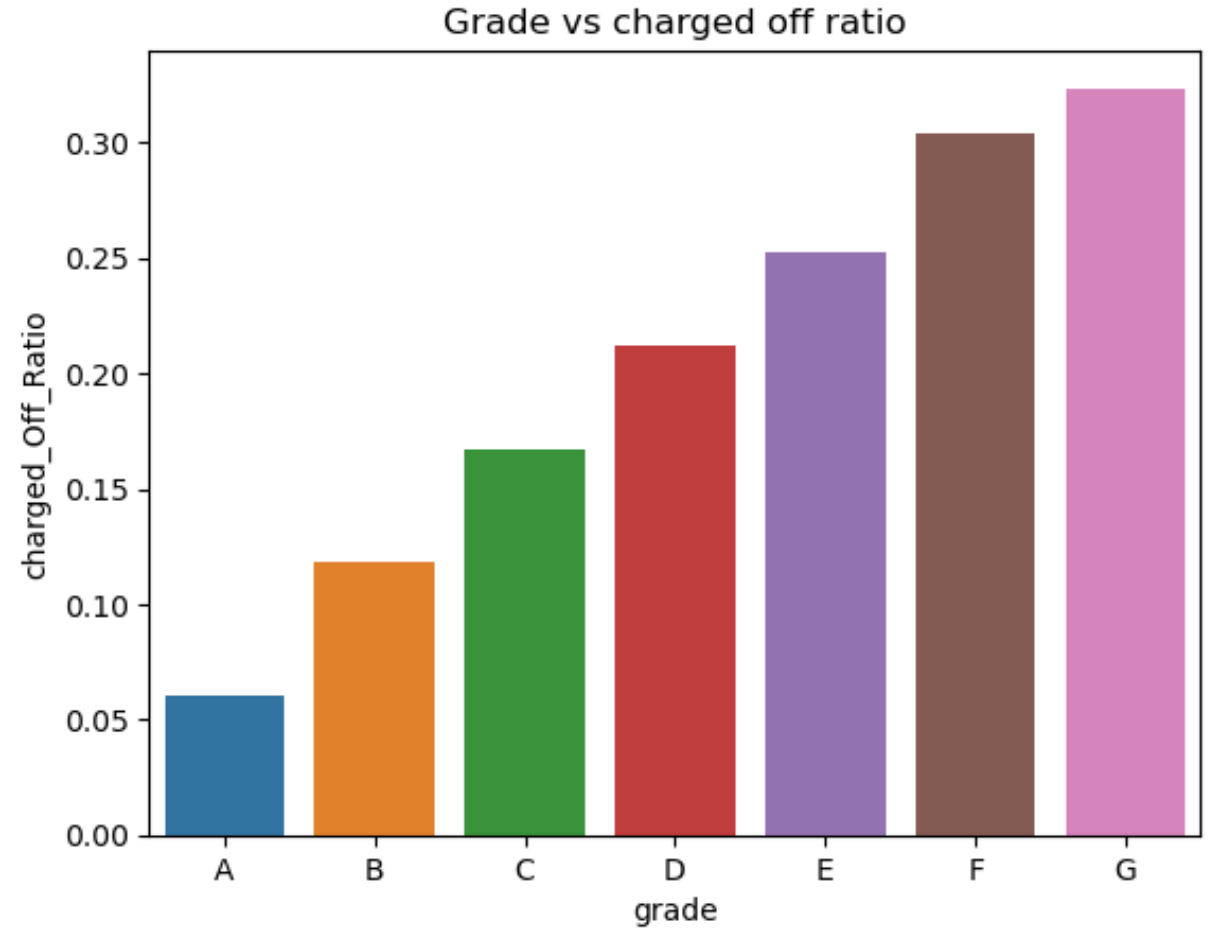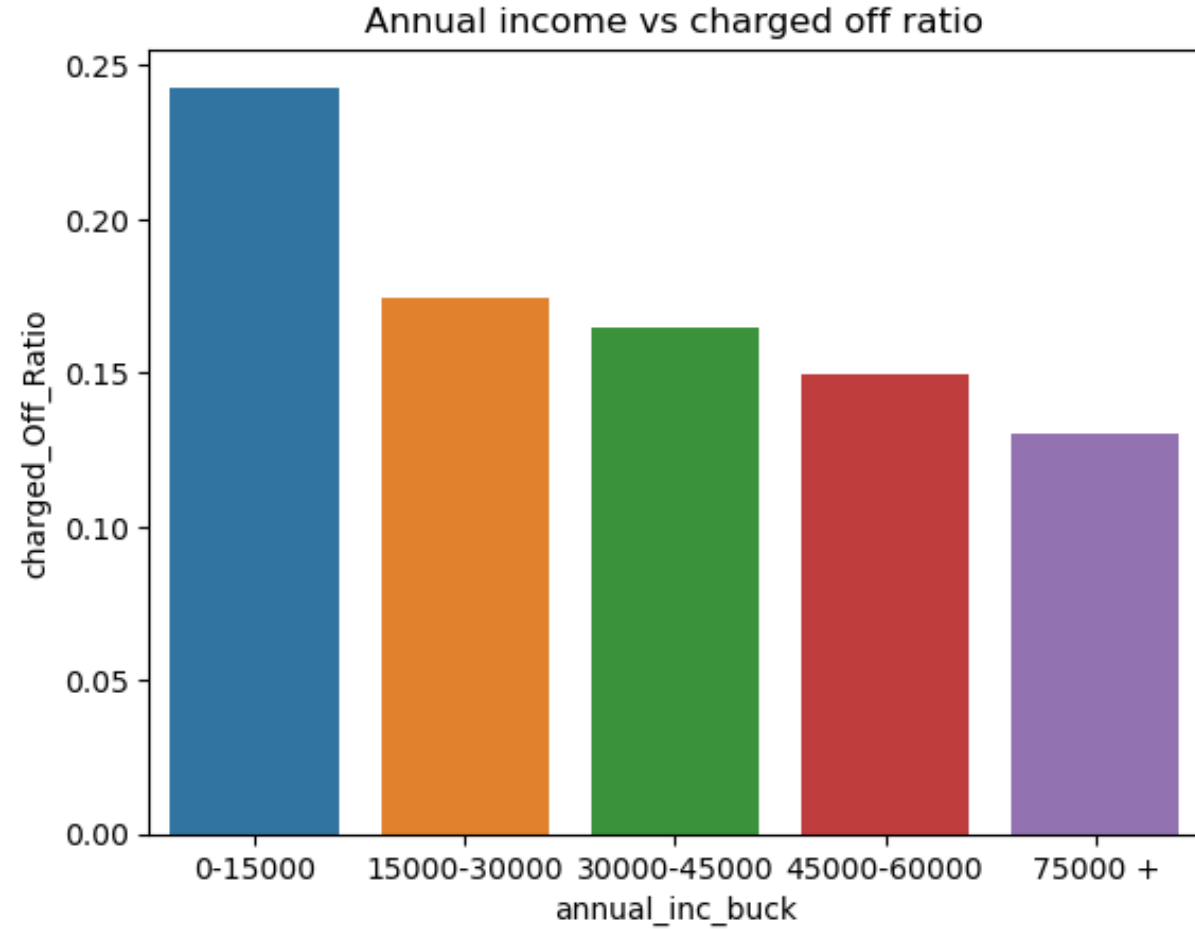
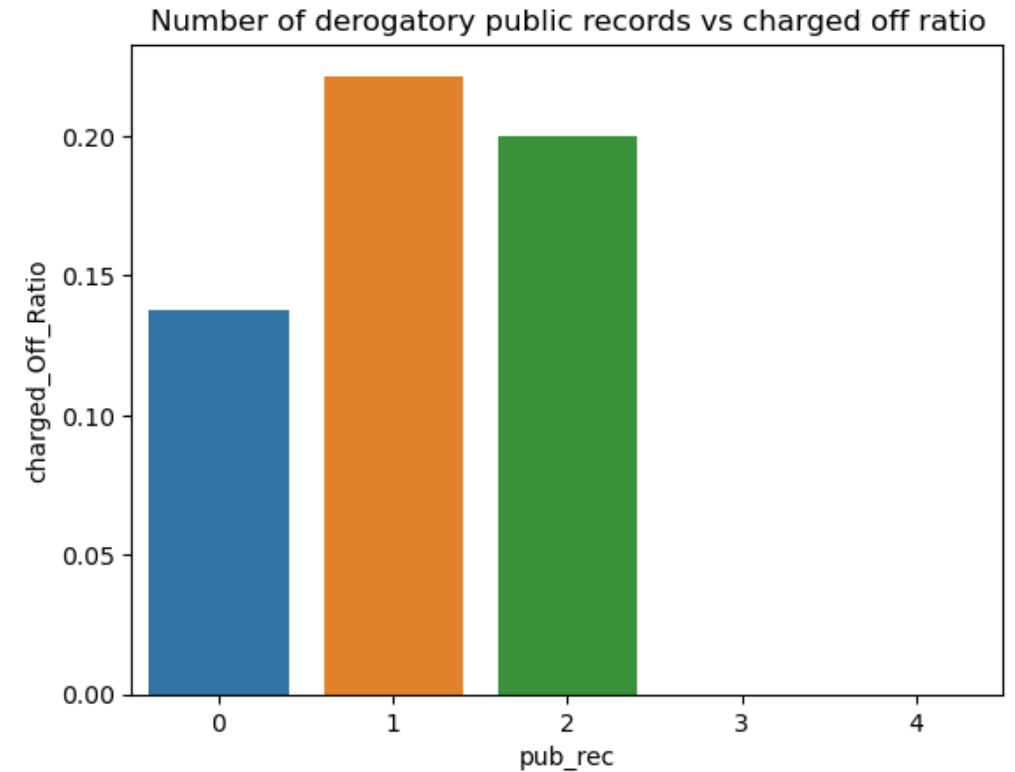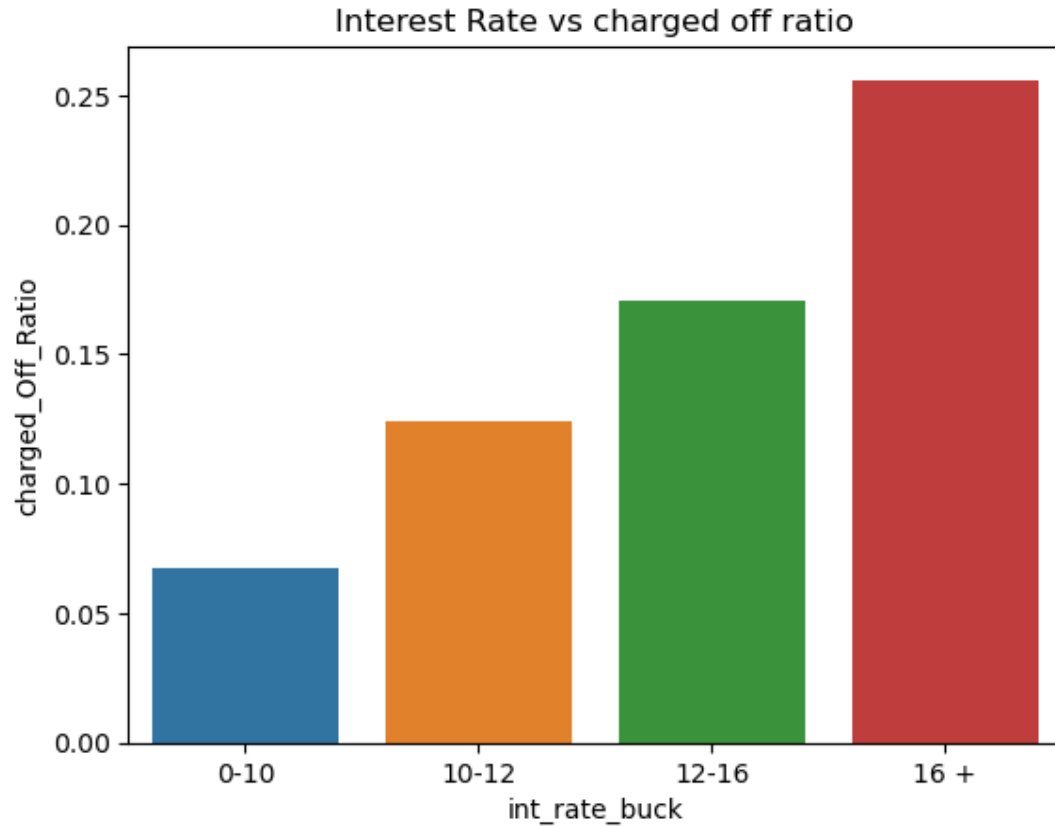|	
10.0	8740
1.0	7766
2.0	4342
3.0	4059
4.0	3410
5.0	3255
6.0	2209
7.0	1759
8.0	1462
9.0	1247

Employment Length vs charged off ratio

1. If annual income is more the charged of ratio is less
2. Grade F&G has very high chance of getting charged off and the ratio increase with grades moving from A to G
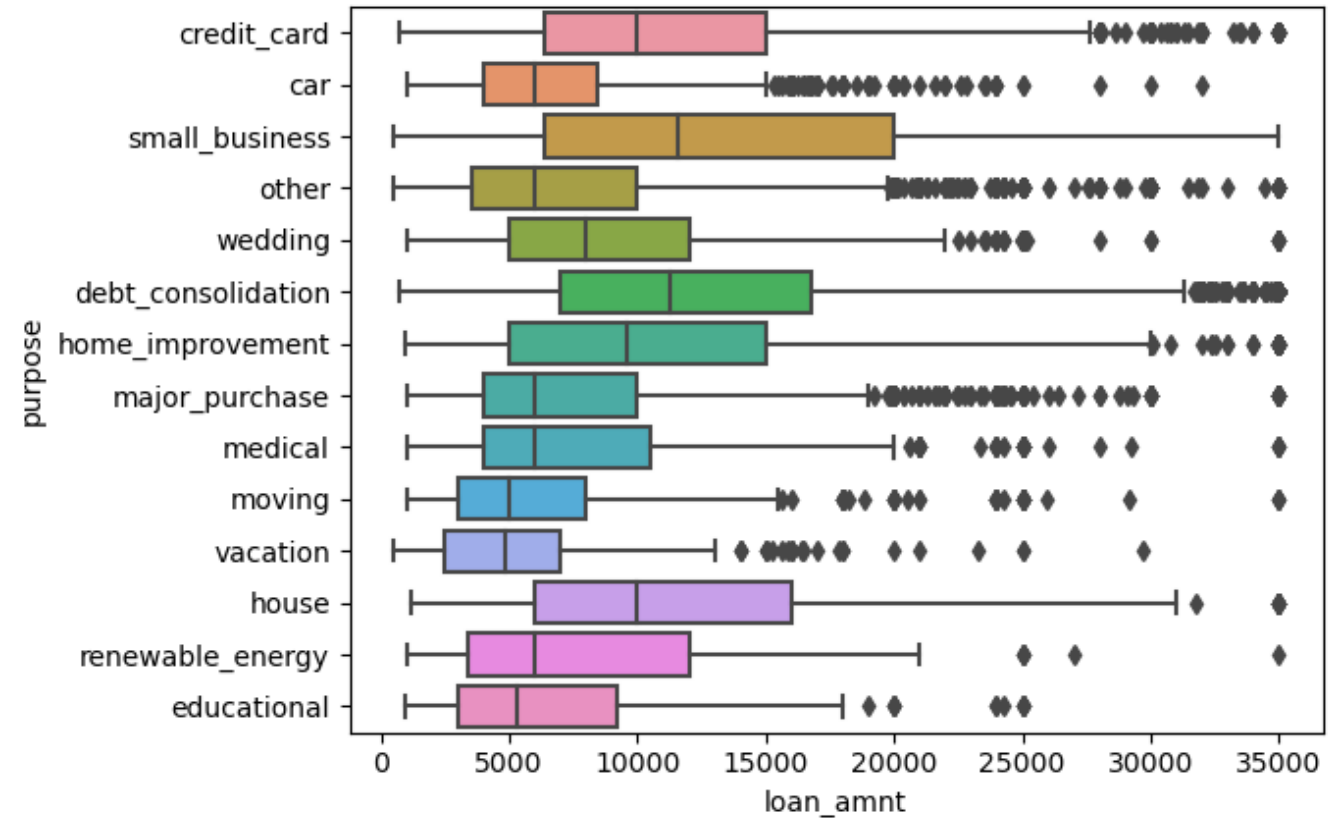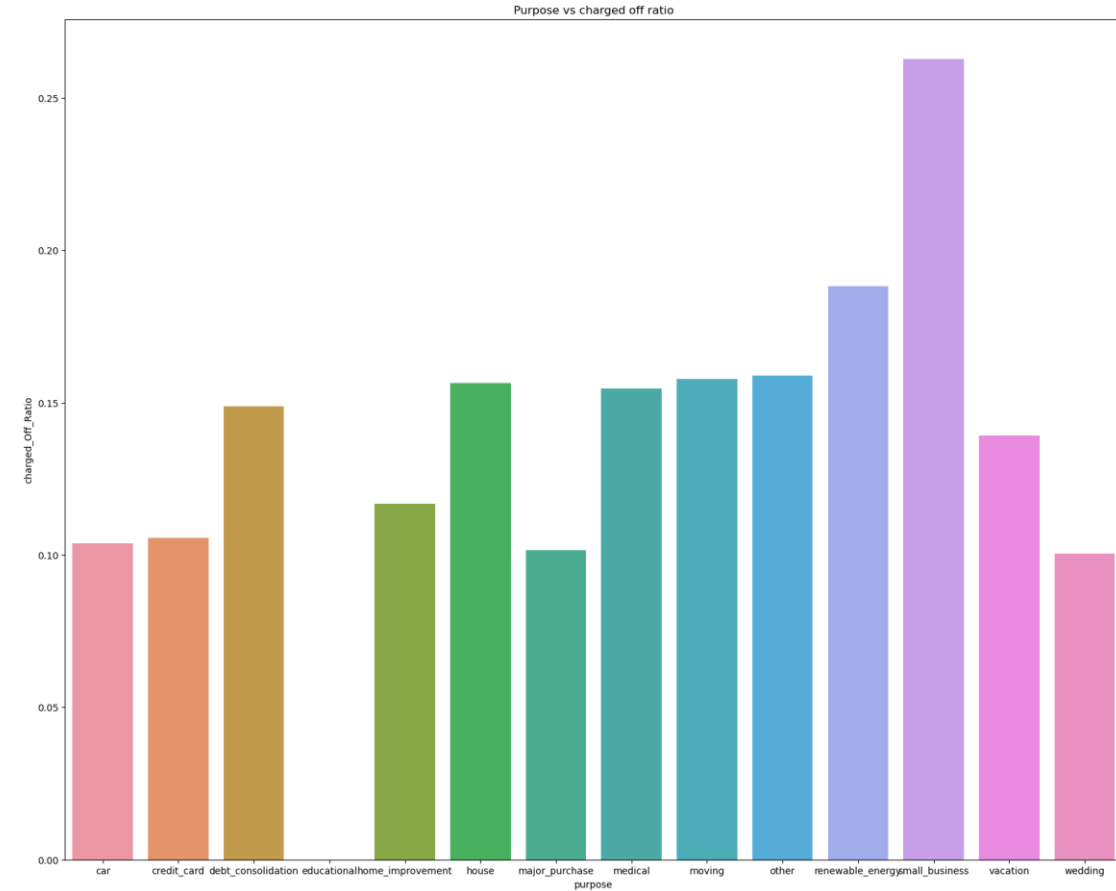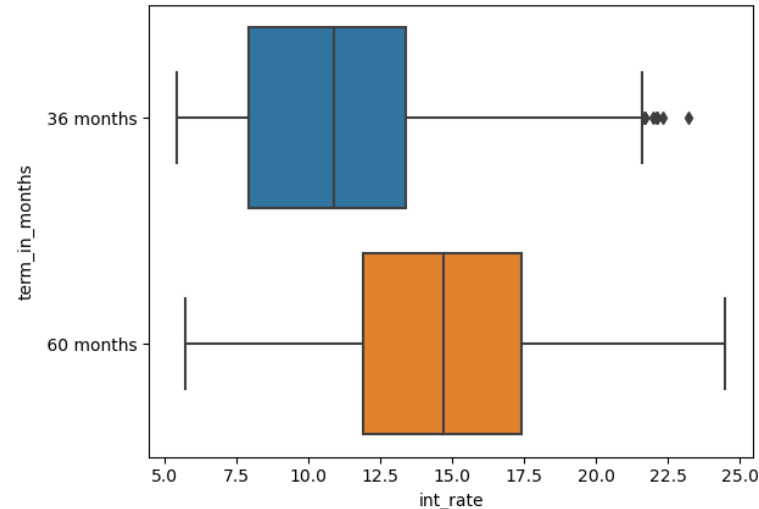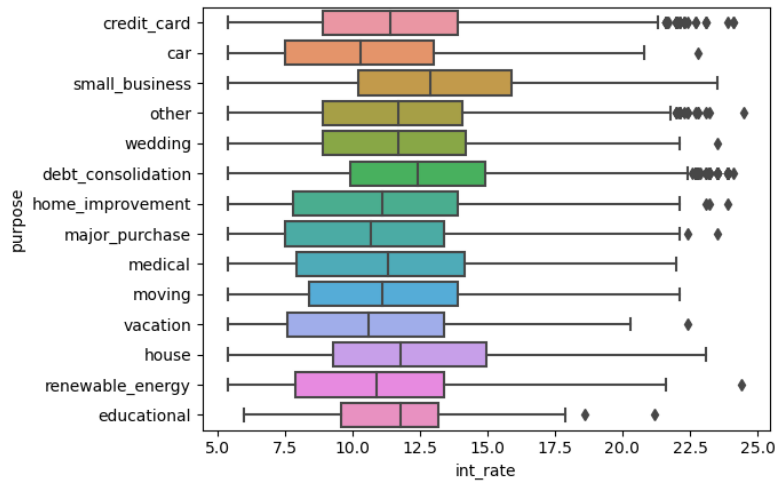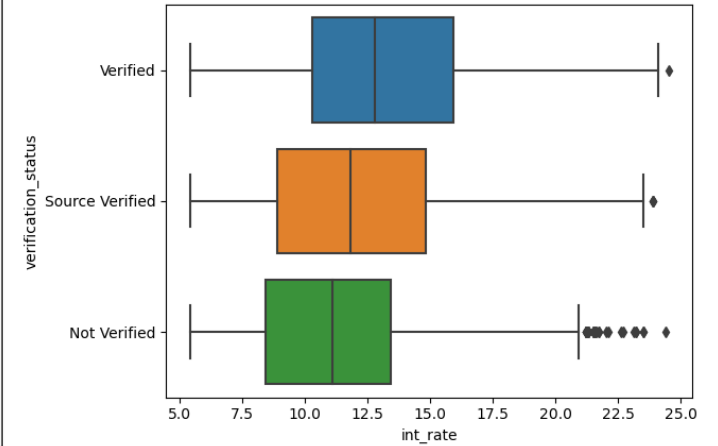
Bivariate Plots

Bivariate Plots

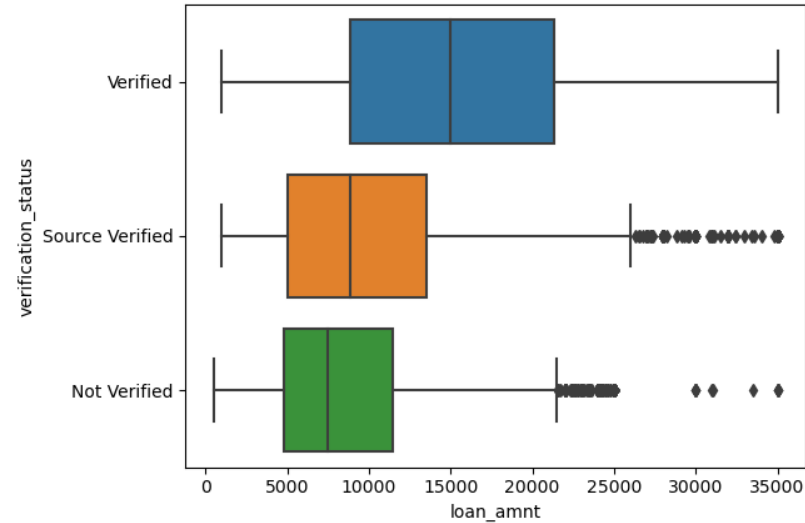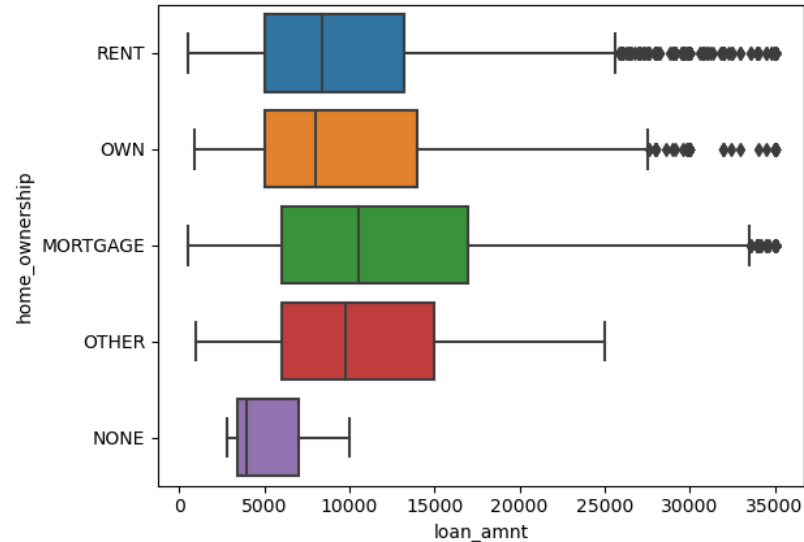

1.Interest rates with16 and more has a greater chances of getting charged off
2. If the derogatory public records is 1 or 2 there is a more chance of getting charged off

Bivariate Plots

The Applicants who do small business has a very high chance of getting charged off.

Bivariate Plots

1. Those who have Mortgage as home ownership has more loan amount
2. Those who have a verified status has more loan amount & interest rate
3. Applicant with small business & house loan has more interest rates
4. Interest rate is more for the term of 60 months.

Multivariate Plots
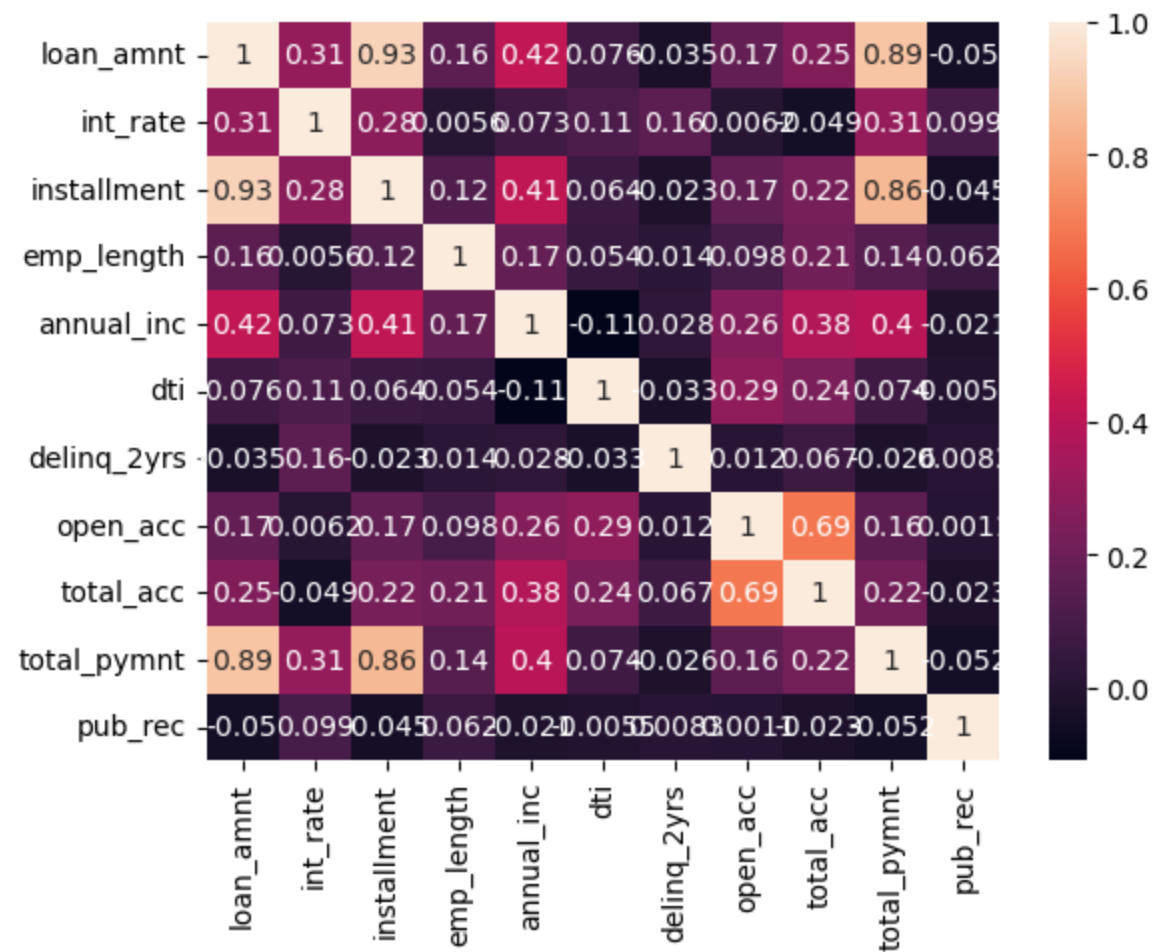
Results

» If annual income is more the charged of ratio is less

» Grade F&G has very high chance of getting charged off and the ratio increase with grades moving from A to G

» The Applicants who do small business has a very high chance of getting charged off.

» Charged off proportion is increasing with higher interest rates

» Employee whose experience is less than one year will be having a possibility to get charged off

» If the derogatory public records is 1 or 2 there is a more chance of getting charged off