LENDING CLUB CASE STUDY

Final Submission

Case Study Group:
- ➢ Samyatirtha Bhattacharya(Group Facilitator)
- ➢ Shanu Pataparambath Noyal Raj
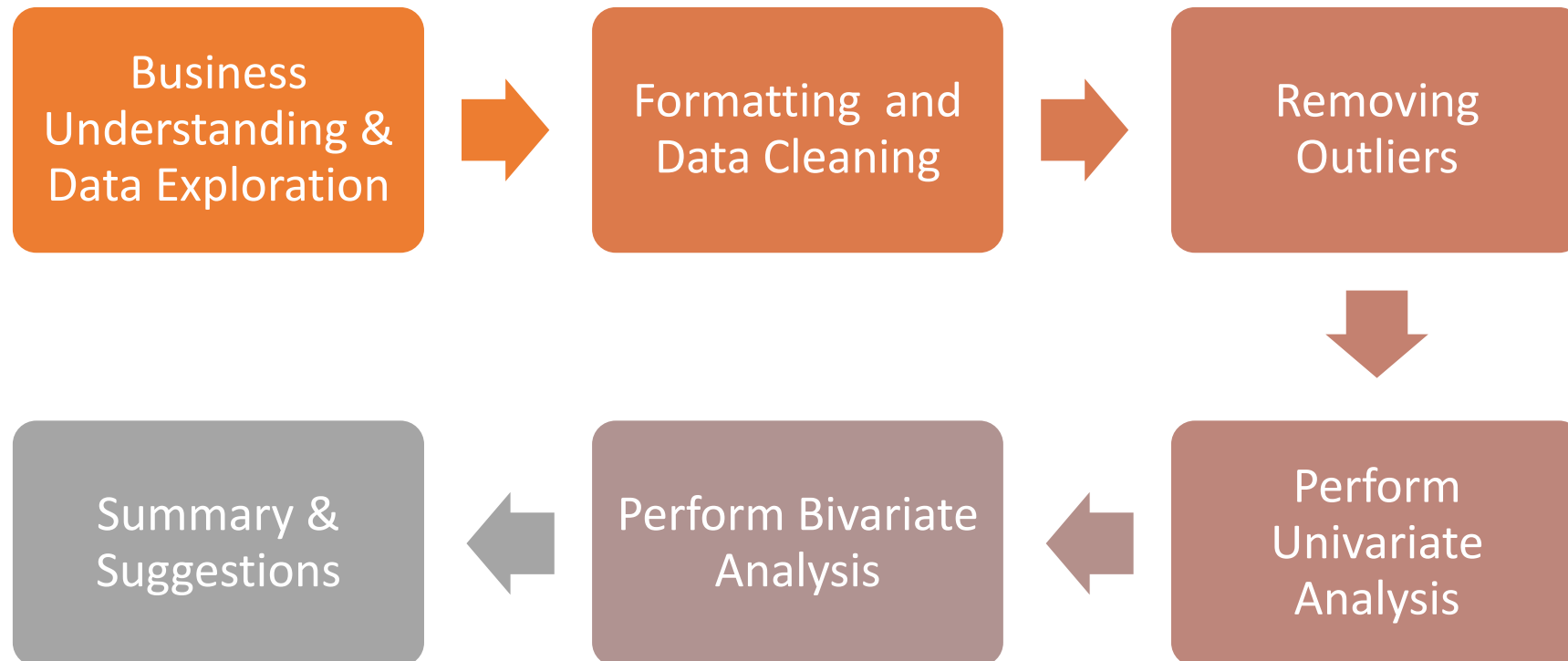
Problem Statement :

- You work for a **consumer finance company** which specializes 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:

- 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
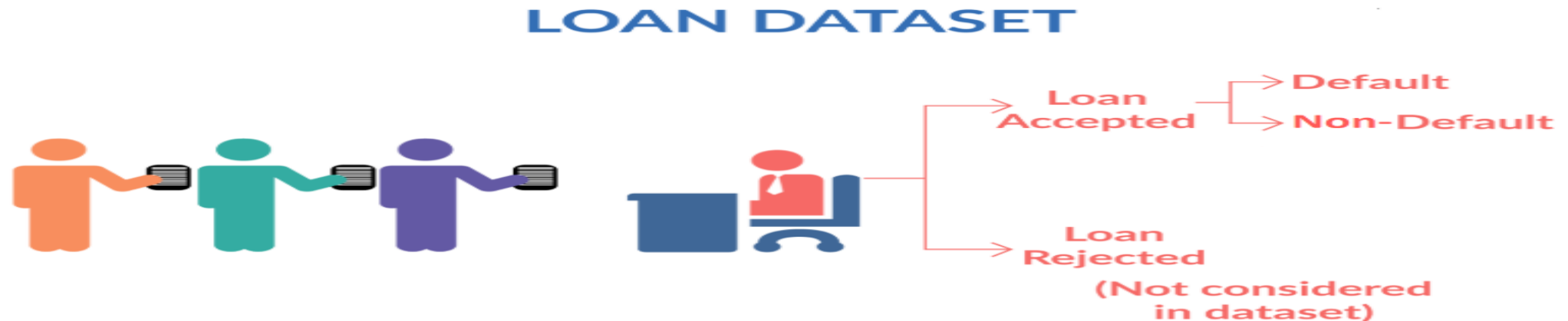
Overall Approach of the Analysis :

➢ As a Lending Club formed by investors, it is of paramount importance to have a well judgment about an applicant's profile.

➢ To prevent the loss to investors, one must perform conclusive EDA to identify the risks.

Business Understanding & Data Exploration	→	Formatting and Data Cleaning	→	Removing Outliers

Summary & Suggestions	←	Perform Bivariate Analysis	←	Perform Univariate Analysis

|	loan_amnt	funded_amnt_inv	int_rate	installment	emp_length	annual_inc	dti	inq_last_6mths	open_acc	pub_rec	
count	37497.000000	37497.000000	37497.000000	37497.000000	37497.000000	3.749700e+04	37497.000000	37497.000000	37497.000000	37497.000000	374
mean	11127.362856	10288.801951	11.958772	324.858107	5.063312	6.943653e+04	13.288131	0.872390	9.304264	0.053738	
std	7353.972102	7033.947315	3.682062	208.906836	3.402219	6.467208e+04	6.659208	1.071211	4.394866	0.234659	
min	500.000000	0.000000	5.420000	15.690000	1.000000	4.000000e+03	0.000000	0.000000	2.000000	0.000000	
25%	5500.000000	5000.000000	8.940000	167.530000	2.000000	4.100000e+04	8.160000	0.000000	6.000000	0.000000	
50%	10000.000000	8875.000000	11.830000	280.260000	4.000000	6.000000e+04	13.390000	1.000000	9.000000	0.000000	
75%	15000.000000	14100.000000	14.420000	429.350000	9.000000	8.300000e+04	18.550000	1.000000	12.000000	0.000000	
max	35000.000000	35000.000000	24.400000	1305.190000	10.000000	6.000000e+06	29.990000	8.000000	44.000000	4.000000	

After data cleaning, we got the data above that shows the presence of outliers as there is a significant difference between q75 and max values for some of the features. So, we will start with Univariate Analysis.
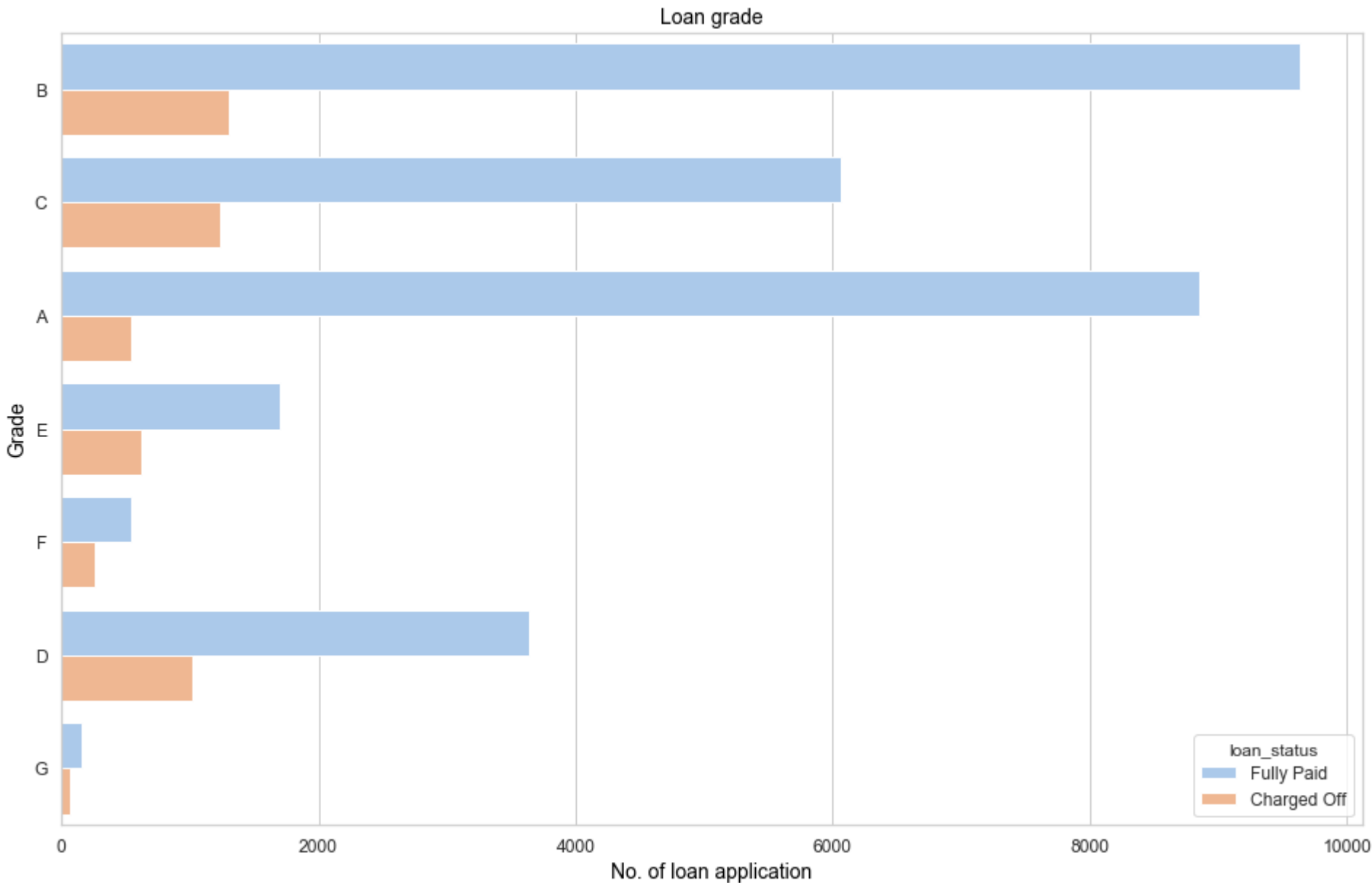
Correlation Matrix
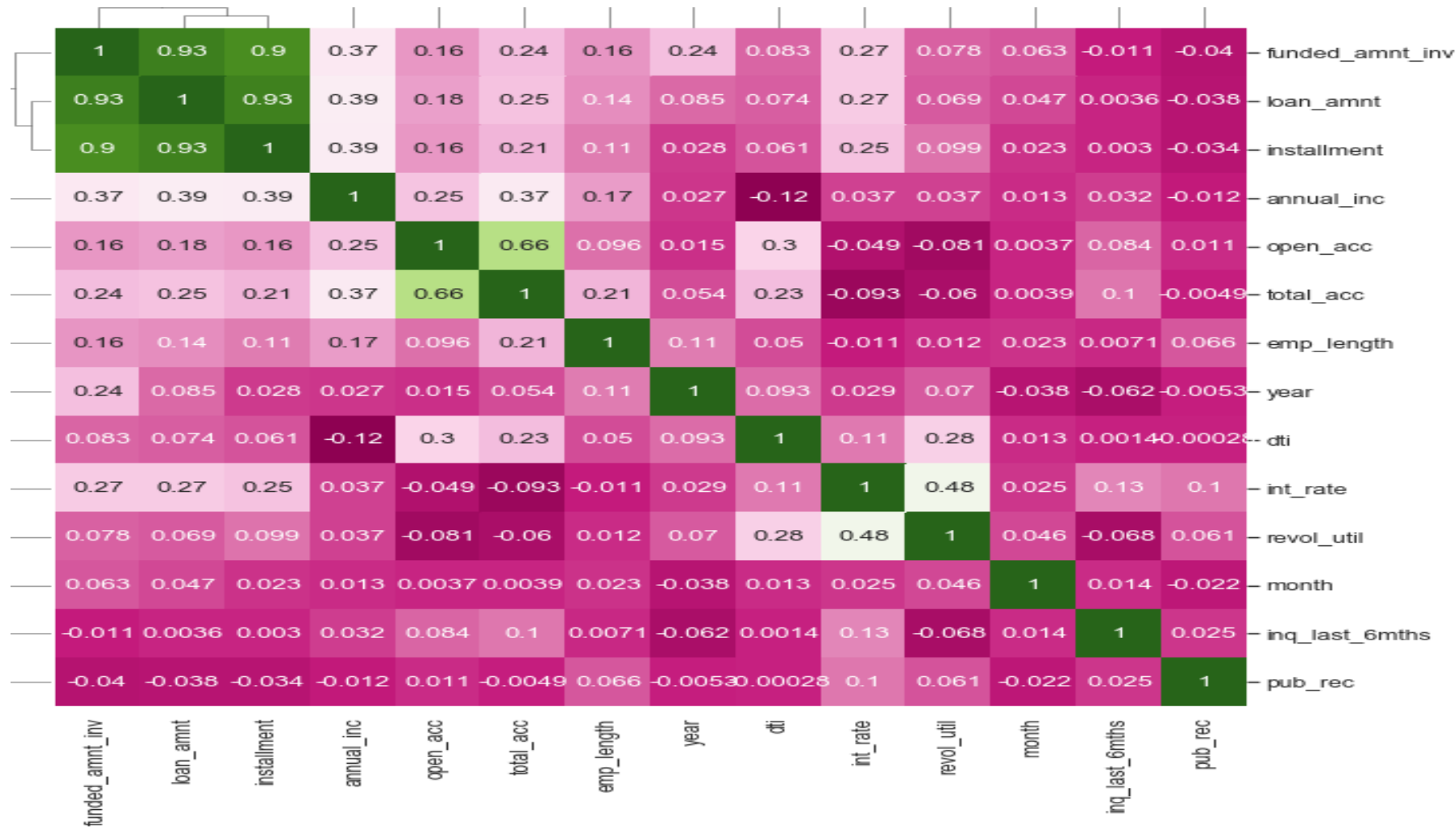
In the matrix, we see high correlation between loan_amnt, funded_amnt_inv, installment

Purpose of Loans:
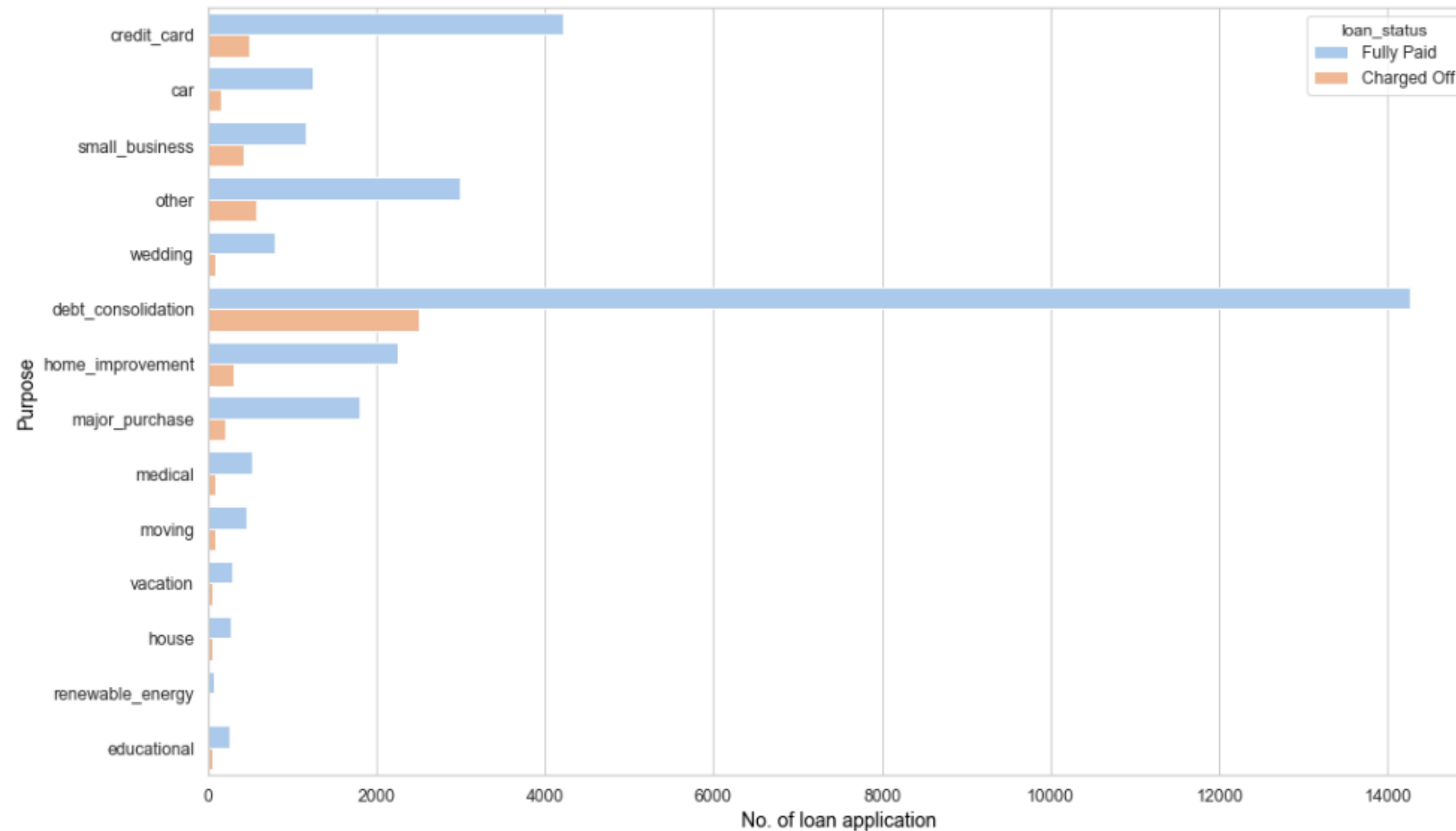
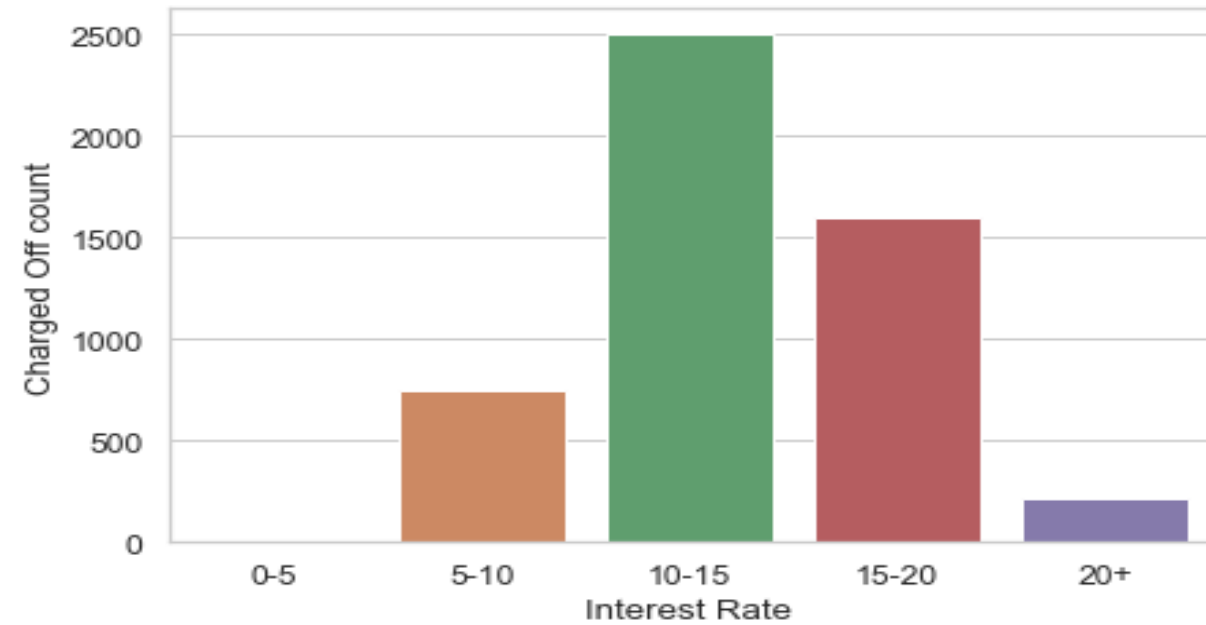**Purpose of Loans:**

➤ Most of the loans were taken for the purpose of debt consolidation

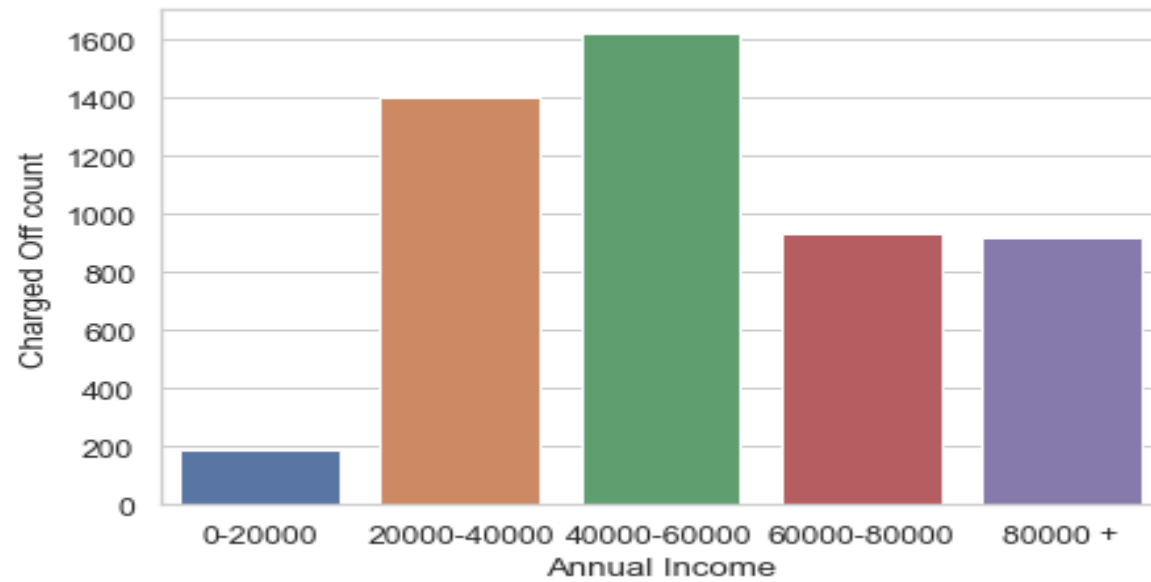➤ Number of charged off as well as fully paid count also high for this category loans.
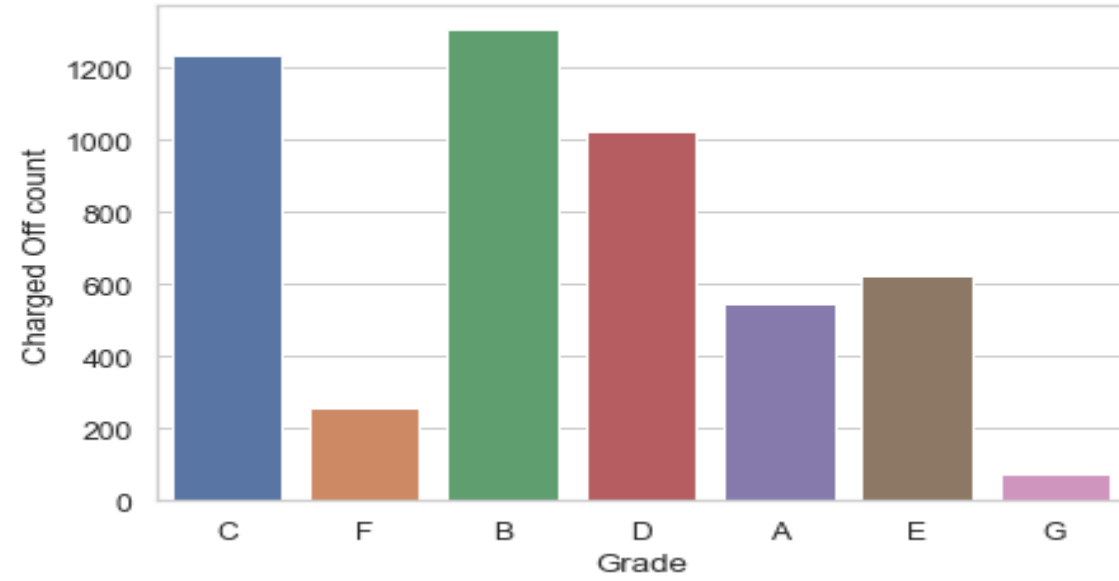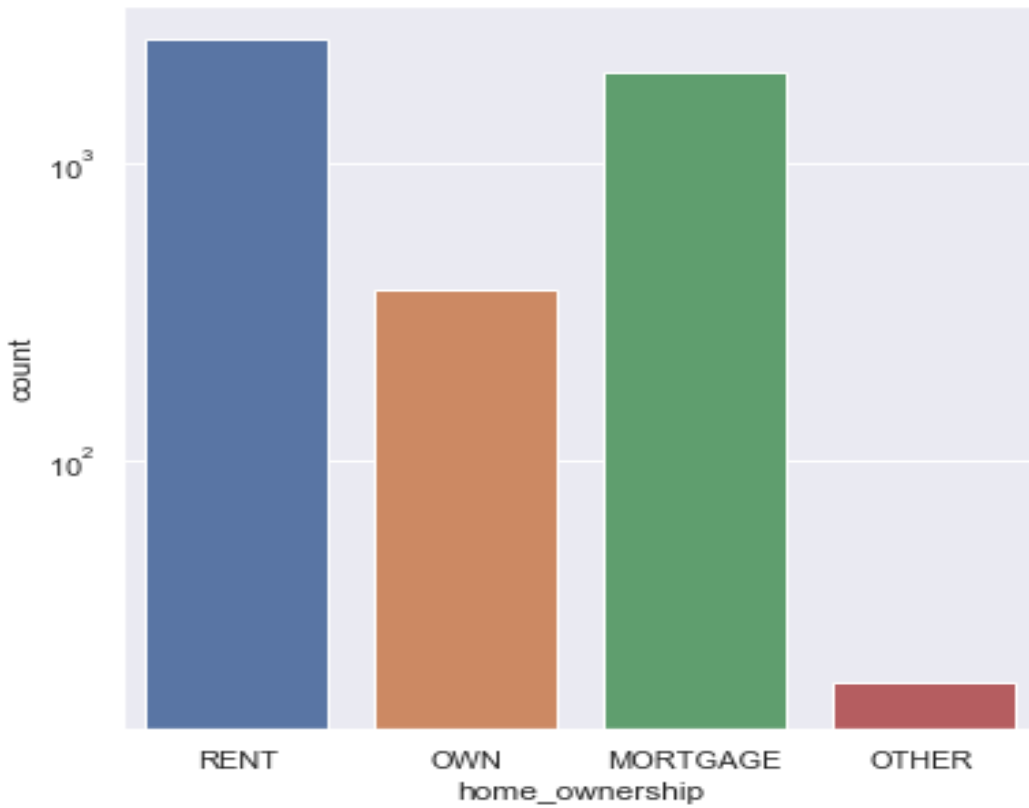
Grades v/s Loan charged off Proportion:

➢ Grade B loans has the most number of Fully Paid as well as Charged Off status.

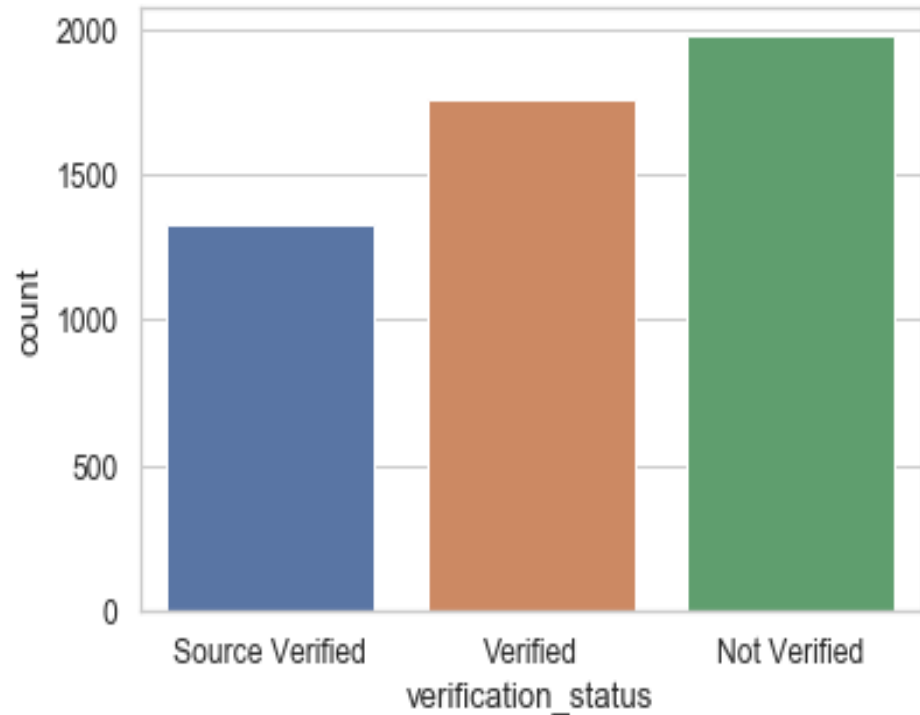➢ Grade "F" and "G" have very high chances of charged off.

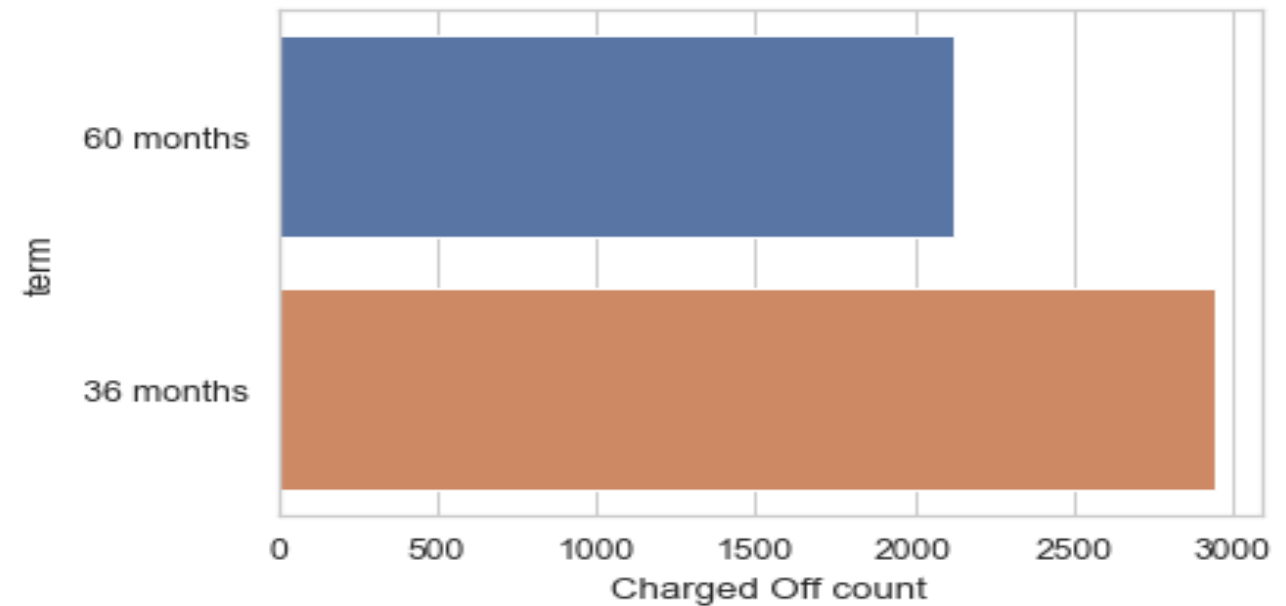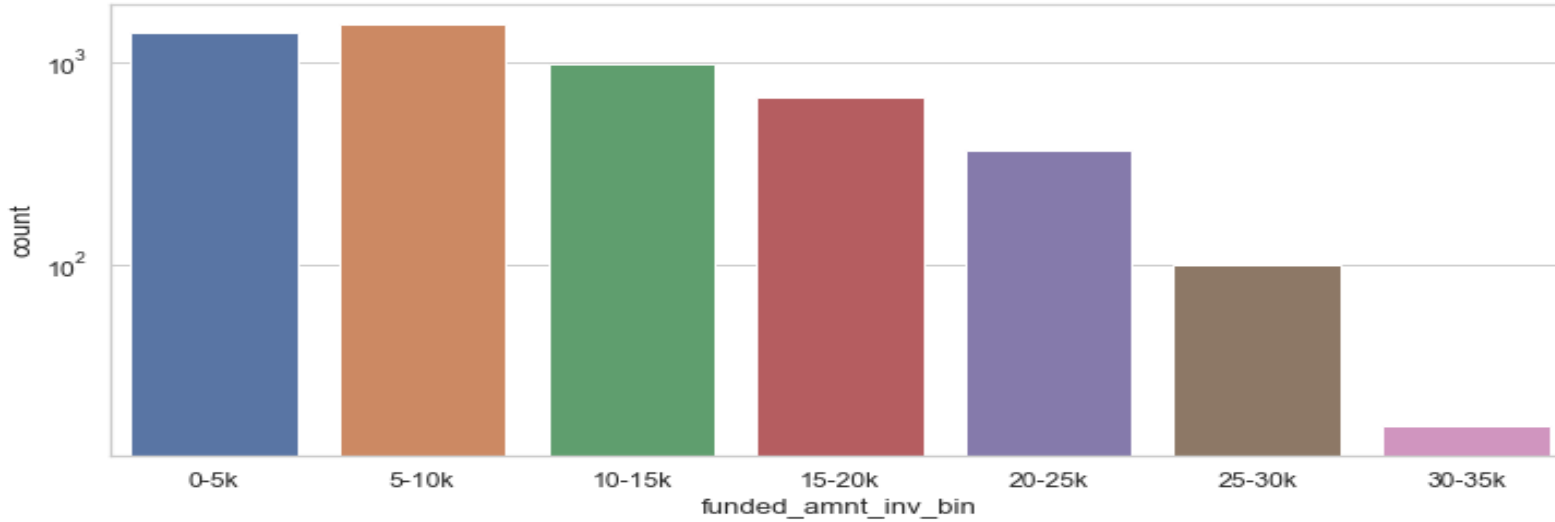**Most of the defaulters come in the bucket of 10-15% interest rate and 40-60k income.**

➢ Rented home owners defaulted the most.
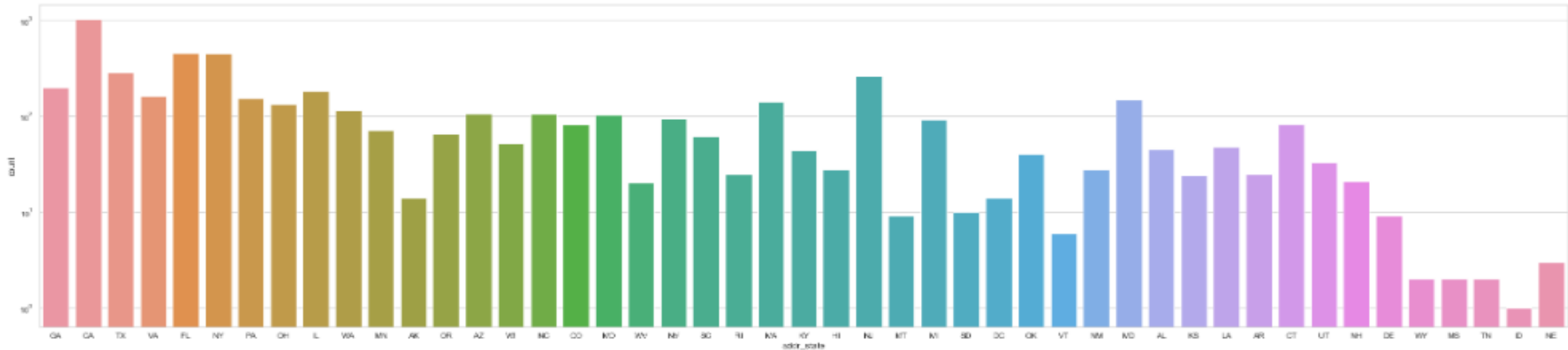➢ Grade B loans has the most number of Fully Paid as well as Charged Off status
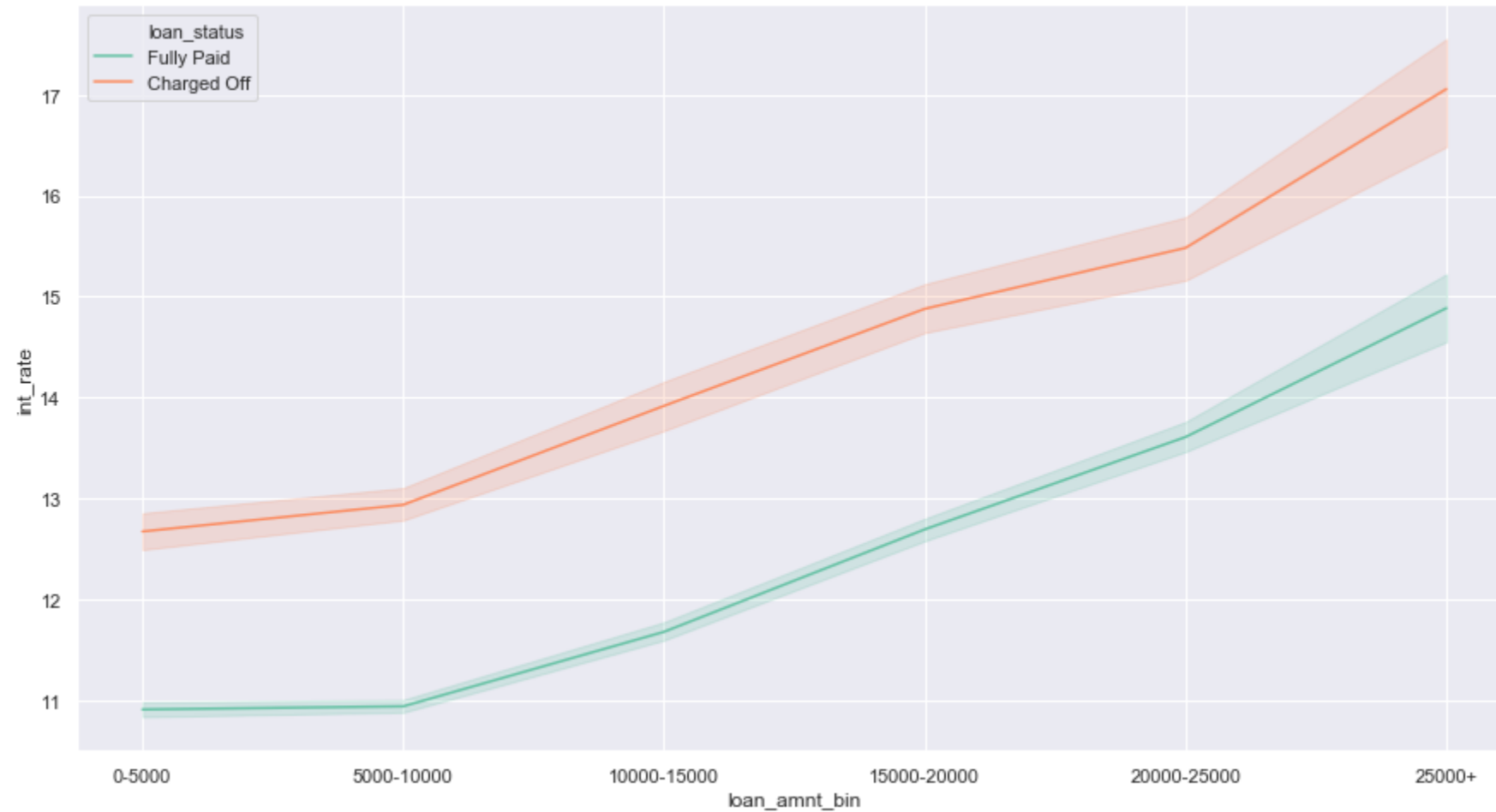
Grades v/s Interest Rate:

➢ Applicants are more likely to default when their credentials are not verified.
➢ 60 months tenure has maximum defaulters.

■Maximum loans defaulted when funded_amnt_inv is 5-10k.
■Maximum loans taken in CA state.

Interest Rate is maximum for the loans defaulted

Suggestions to Lending Club:

➢ All the credentials of applicants should be well verified.

➢ Applicants with Derogatory Public Records should be avoided.

➢ Investors should try to invest in lower tenure loans.

➢ Loans to be given to applicants according to their income.

➢ Loan should not be given to applicants who does not have a source of income.

➢ It is to be taken under consideration whether loan will be disbursed further for Debt Consolidation purpose.