

Lending Club Case Study

Group Facilitator : Munirathinam Duraisamy
Group Member : Sreedhar K

Table of Contents

- Problem Statements
- Business Objective and Aim
- EDA Steps Followed
- Analysis Results / Insights
- Recommendations
- Conclusion

Problem Statement

Working 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

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

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

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)

Business Objective & Aim

Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

- The company wants to understand the driving factors (or driver variables) behind loan default, that is, the variables which are strong indicators of defaulters.
- The company would be able to utilize this knowledge for its risk assessment and portfolio.
- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicant's using EDA is the aim of this case study.
- Our aim is to identify the loan application which can be defaulted or deliquated 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

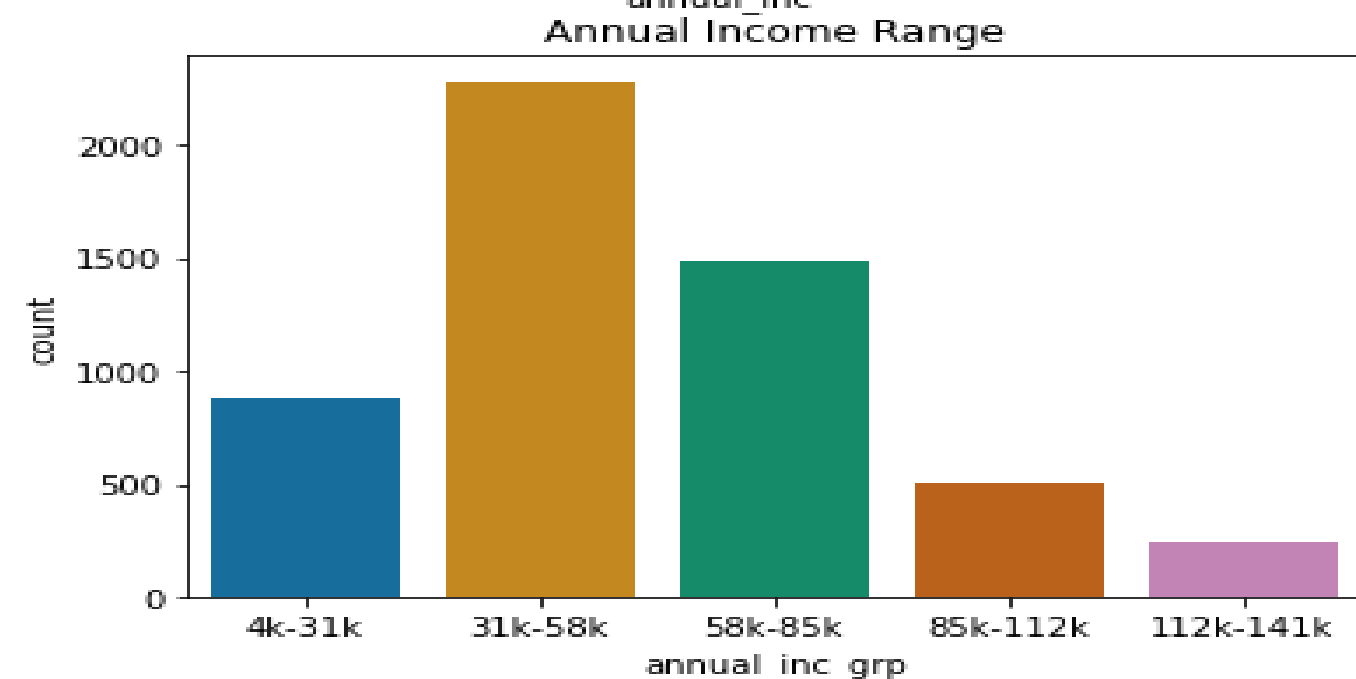
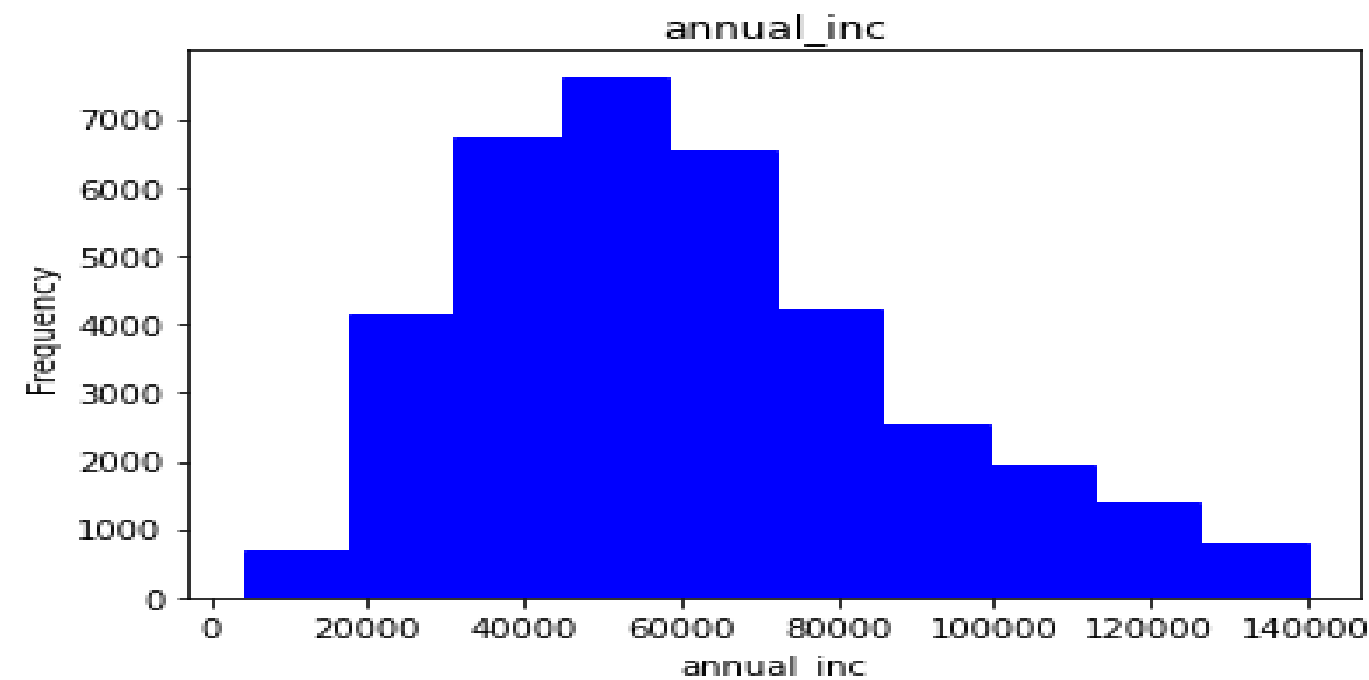
Exploratory Data Analysis Steps Followed

- Data Overview
- Import Required Libraries
- Data Cleaning
 - a. Missing Data Treatment
 - b. Standardizing Values
 - c. Remove Irrelevant Variables
 - d. Outliers Analysis and Treatments
 - e. Derived Metrics & Binning
- Data Analysis
 - a. Univariate Analysis
 - b. Bivariate Analysis
 - c. Multivariate Analysis
- Insights

Analysis Results / Insights

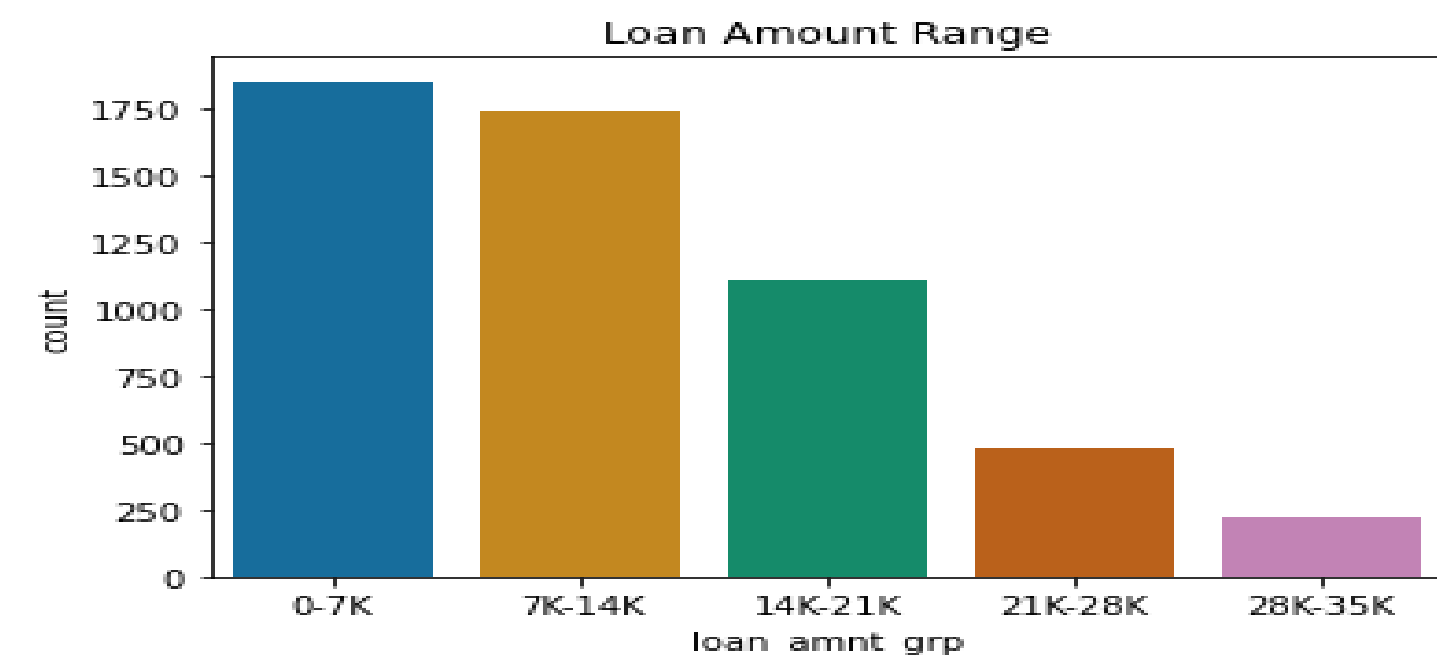
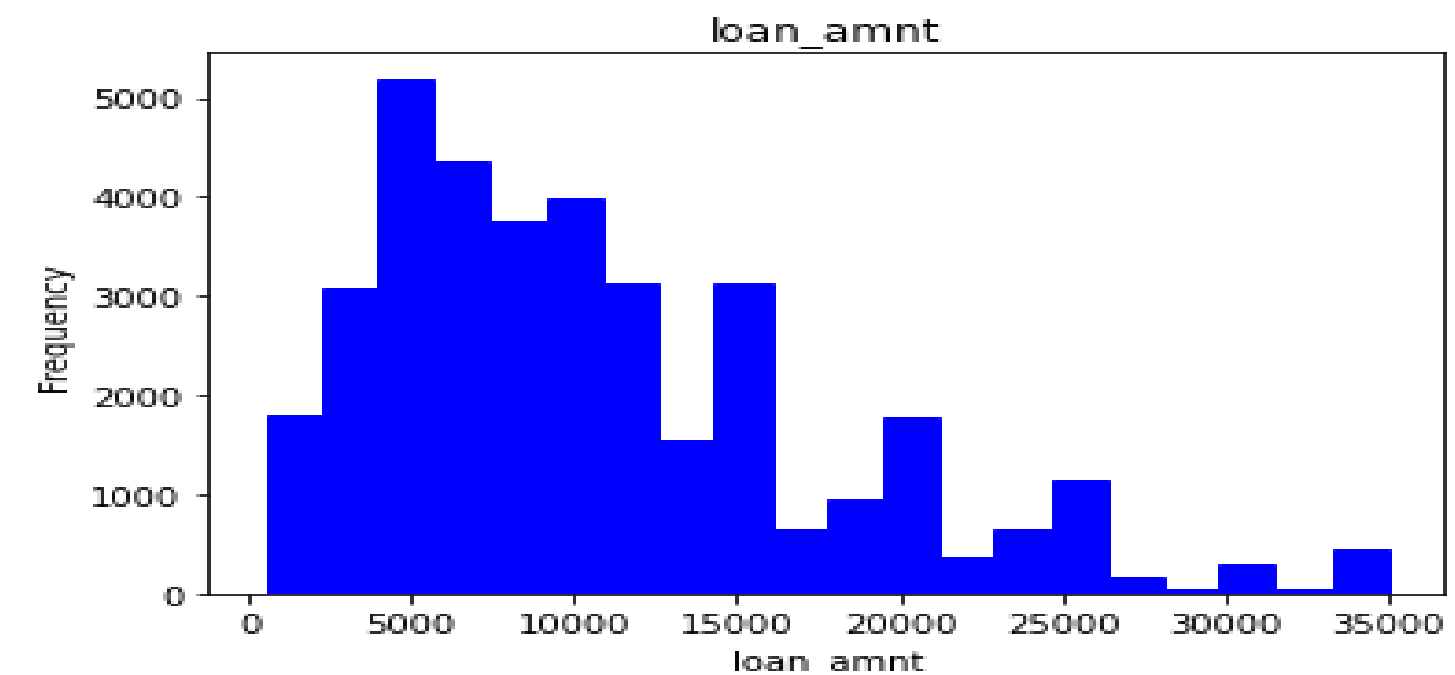
Annual income

- Maximum number of loan applications are under the slab between \$50000 and \$60000.
- The clients with Charged Off Accounts whose annual income is in the range of \$31K to \$58K are having the highest defaulters.



Loan amount

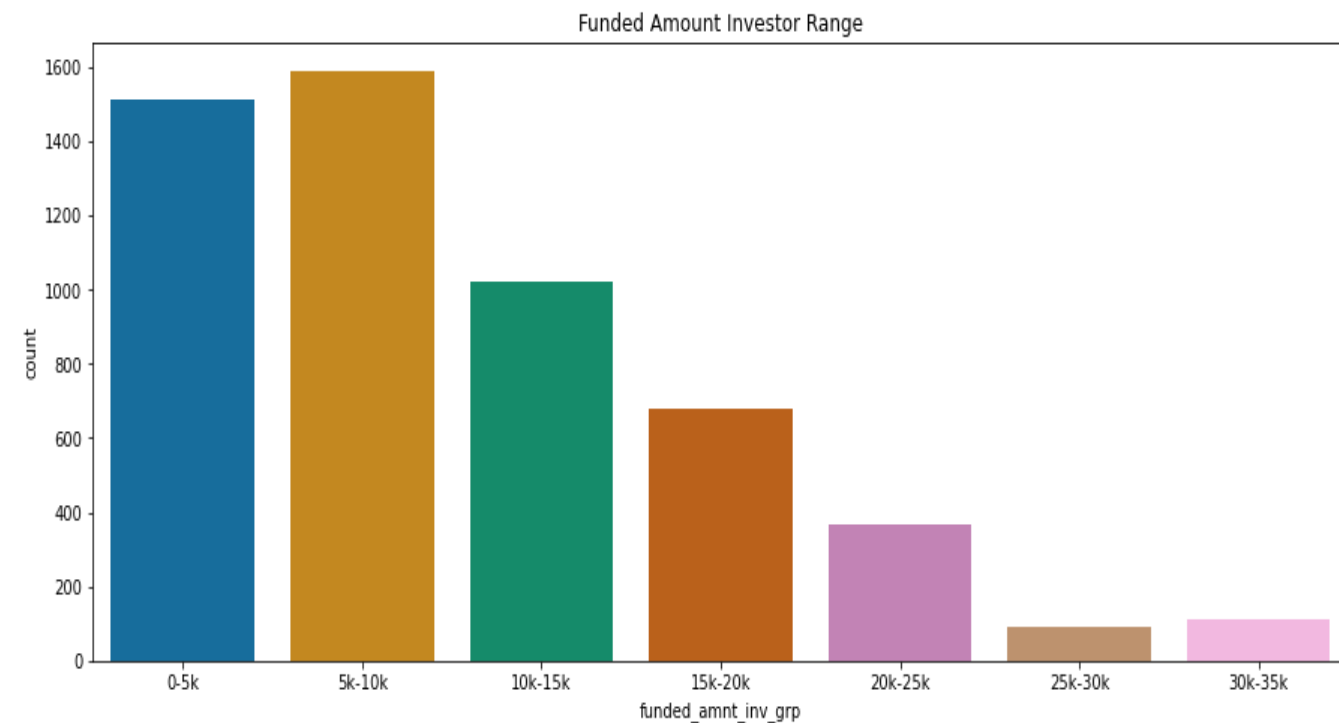
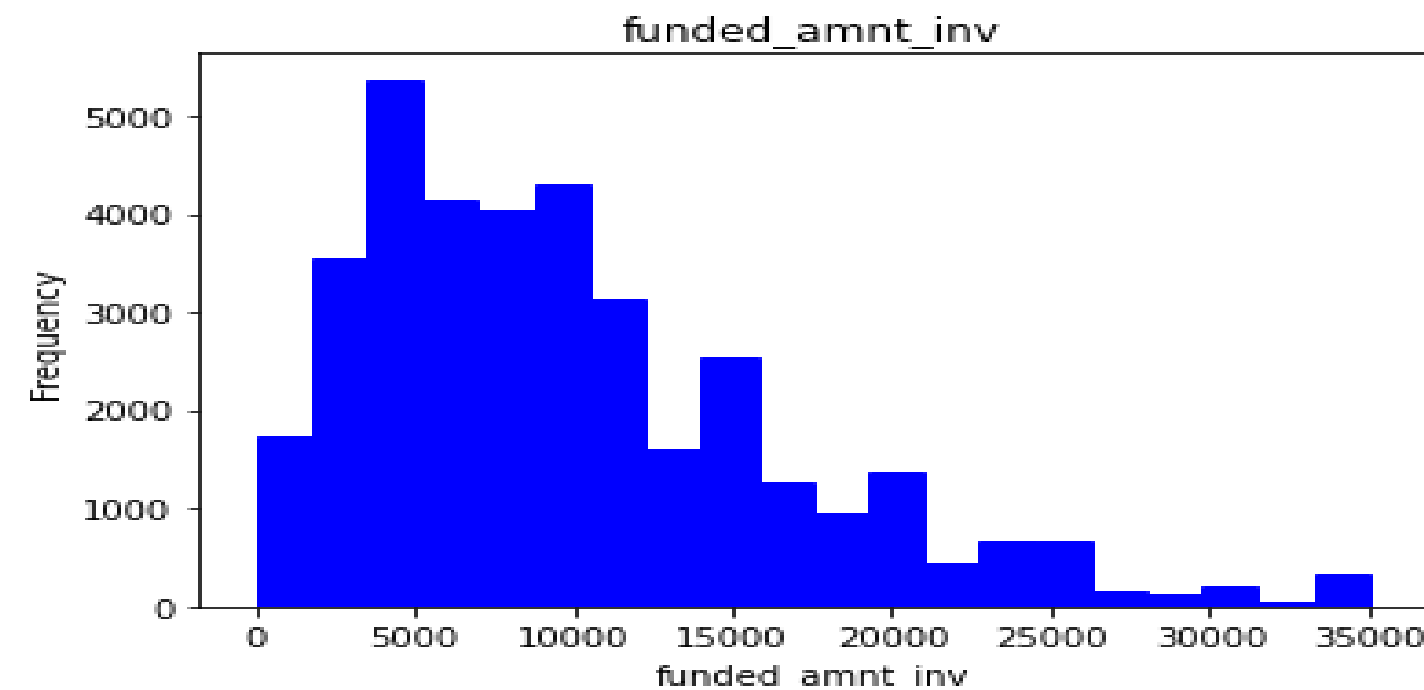
- Maximum loan applications are of the loan amount around \$5000.
- Within charged-off accounts, the loan amount range between \$0 and \$7000 has highest number of defaulters, and the loan amount range between \$7000 and \$14000 are also having quite high defaulters



Analysis Results / Insights cont...

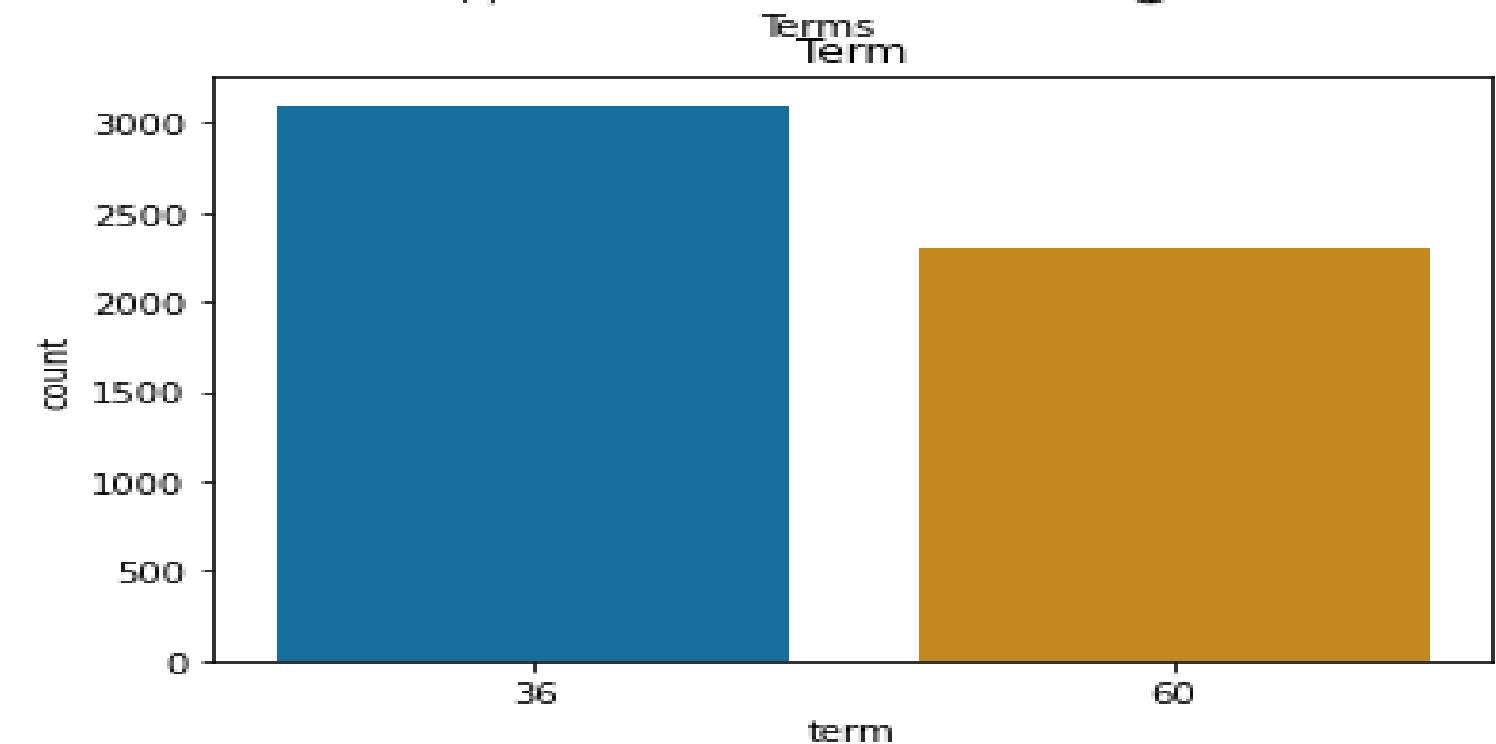
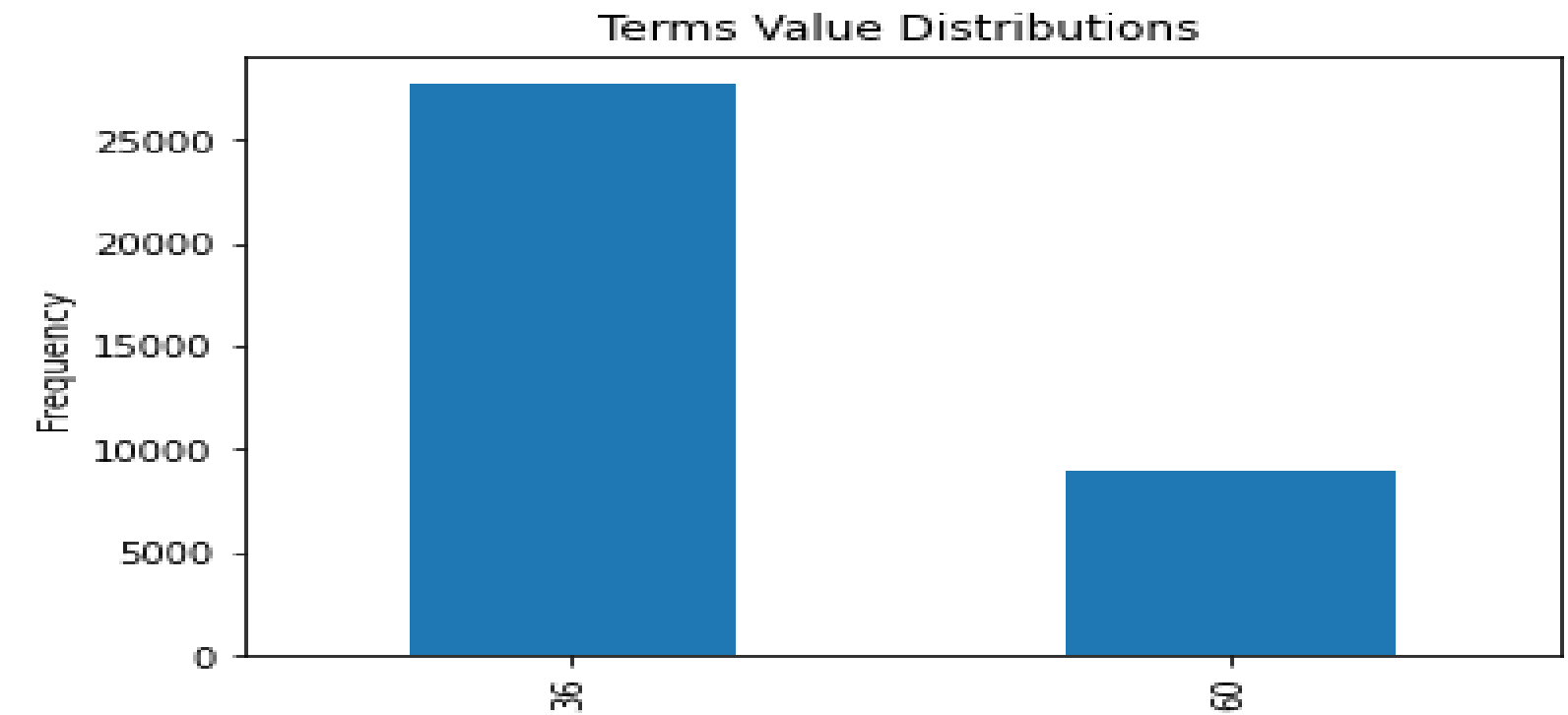
Funded amount investor

- ✓ Maximum loan applications are of the funded loan investor amount around \$5000.
- ✓ The clients whose funded amount by investor is in the range of \$5K to \$10K are having highest defaulters. Range of \$0-\$5K also have quite high defaulters



Term

- ✓ Here, we have only two terms of 36 months & 60 months.
- ✓ Maximum number of loan applications are in the term of 36 months, in case of both data set of charged-off accounts alone and charged-off + fully paid accounts.



Analysis Results / Insights cont...

Instalment

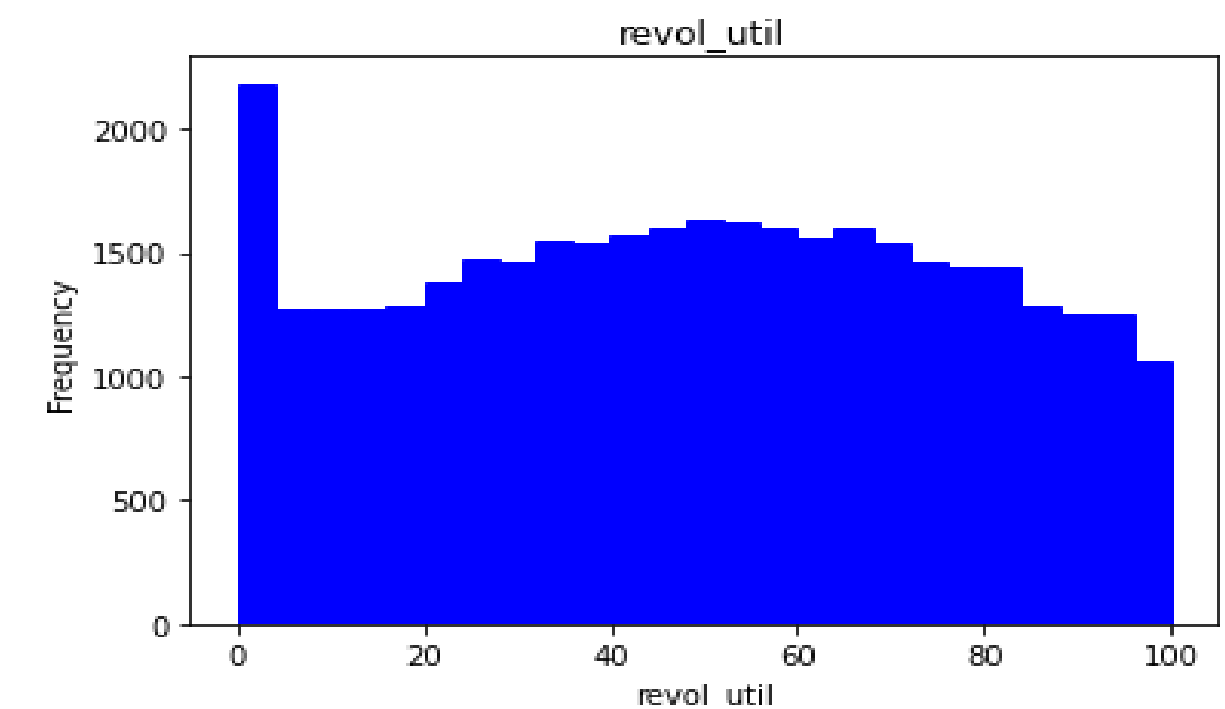
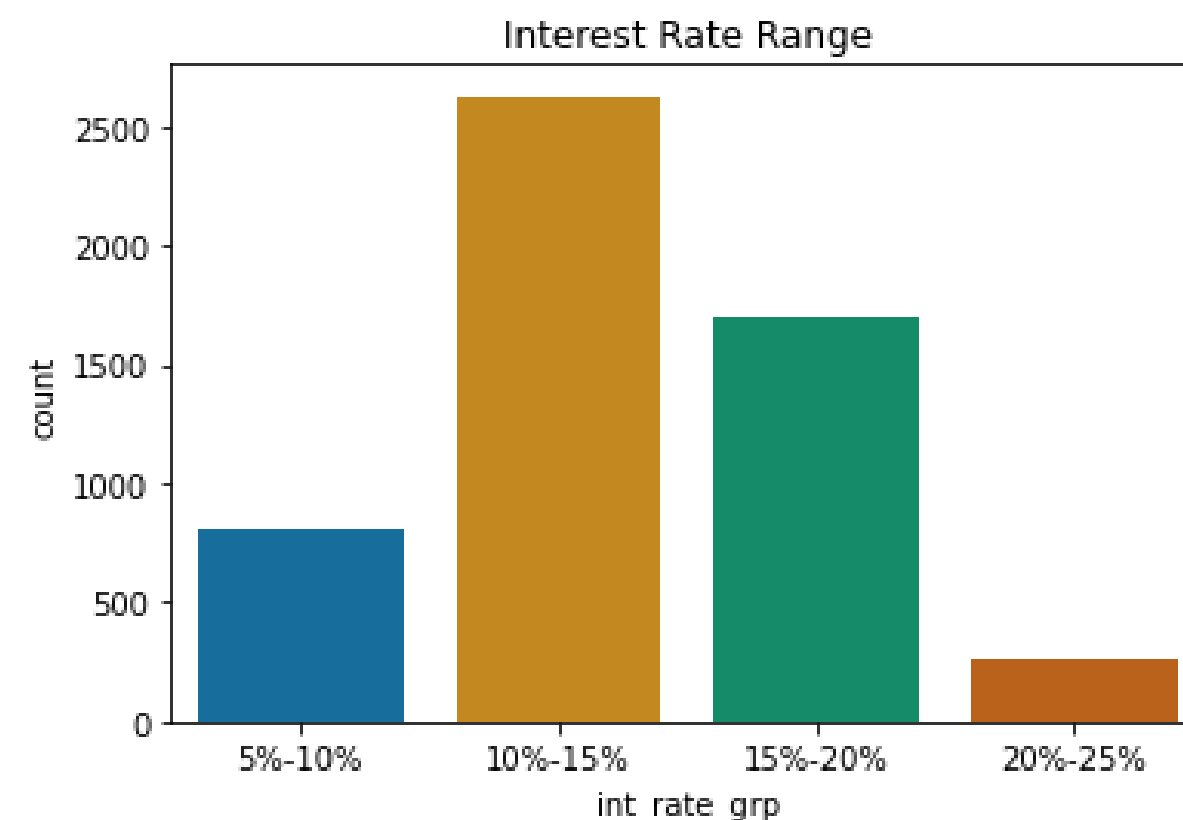
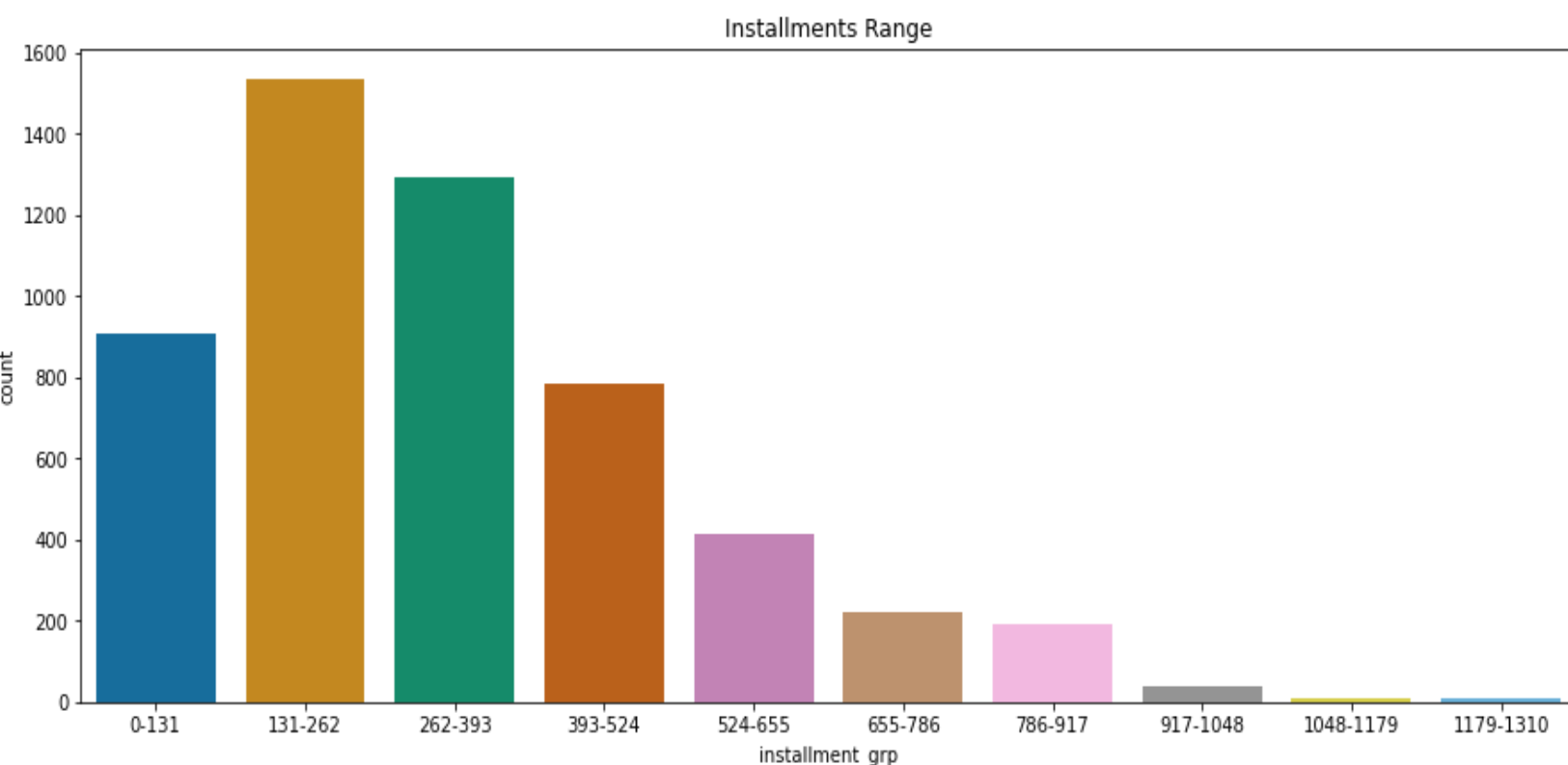
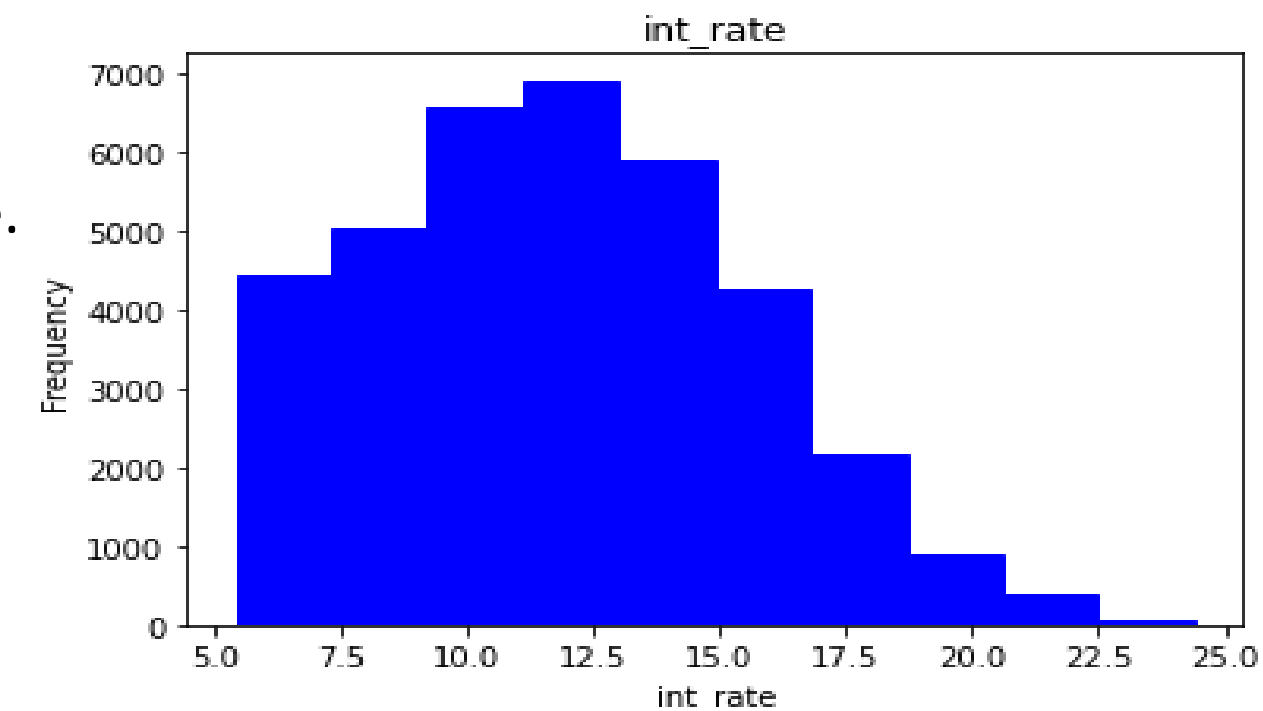
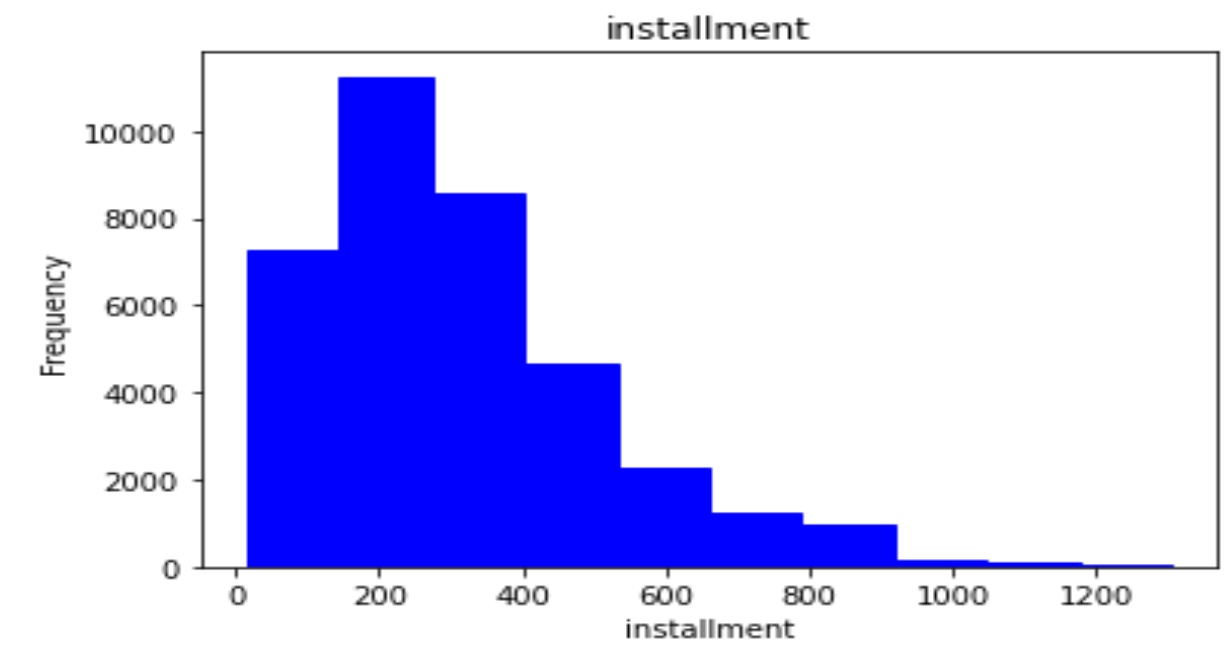
- Maximum number of loan applications are roughly between 150 and 275 installments.
- Within the charged-off accounts, the clients whose installments are between 131-262 are having the highest defaulters.

Interest Rate

- Maximum number of loan applications are in the range of roughly between 11.75 %and 14.25% interest rate.
- The interest rate range between 10% and 15% has highest number of defaulters within the charged-off accounts.

Revolving Credit

- There is no major difference in the utilization of revolving credit , except for in the range of 0-4%.
- At the same time there are good number of users utilizing the maximum credit which could be risky.



Analysis Results / Insights cont...

Debt-to-Income ratio (DTI)

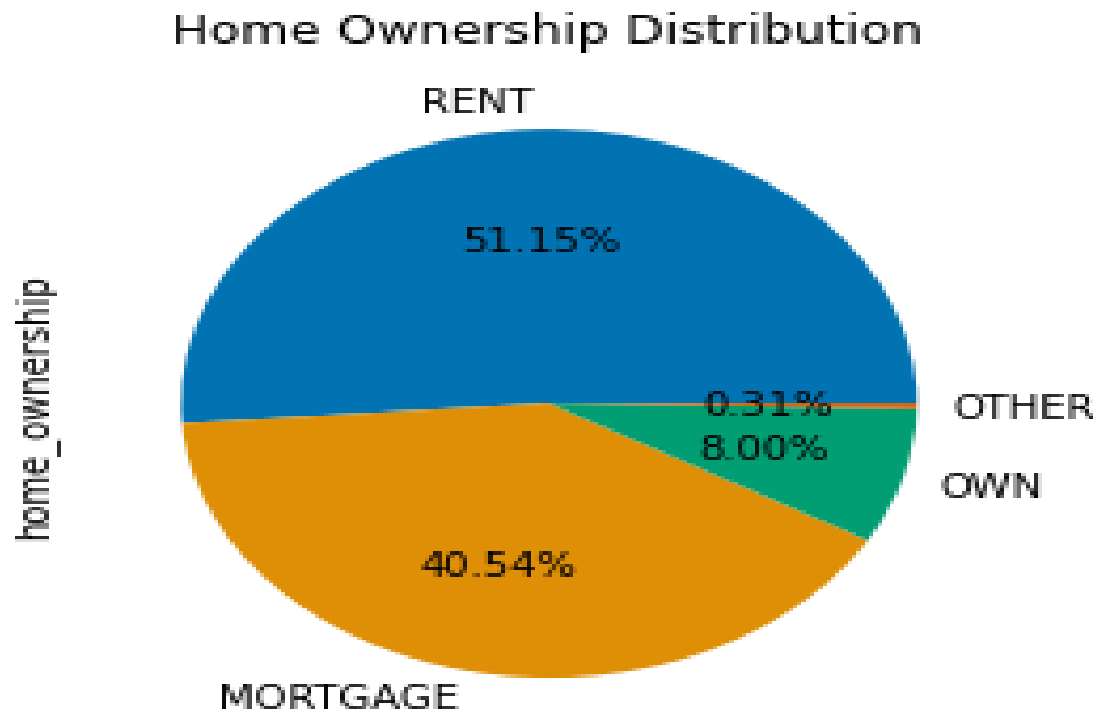
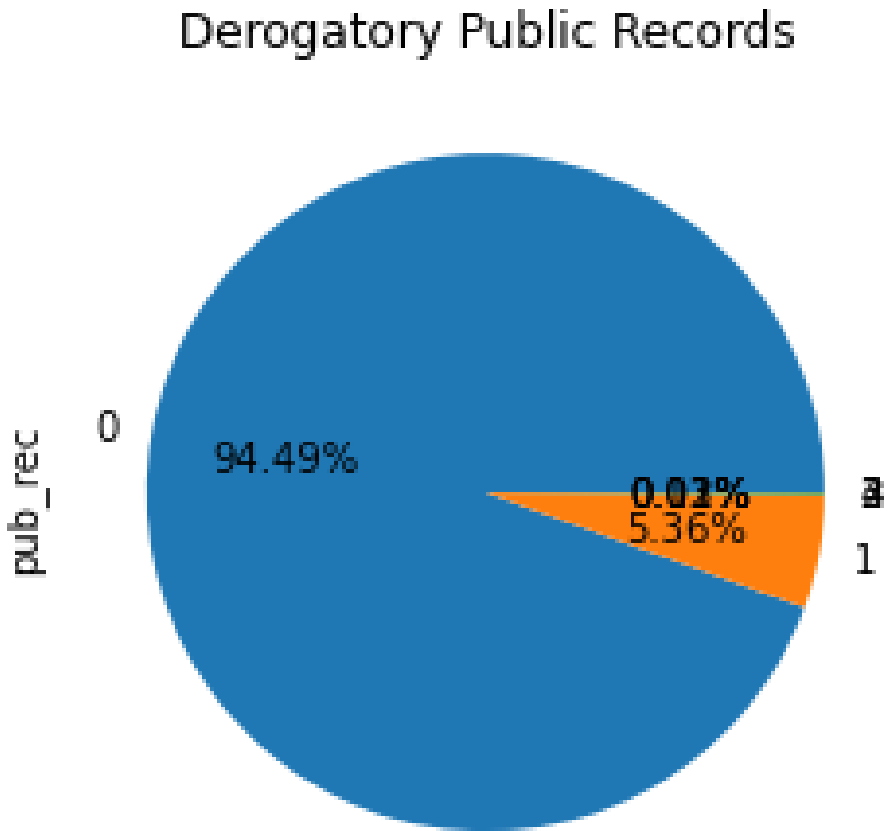
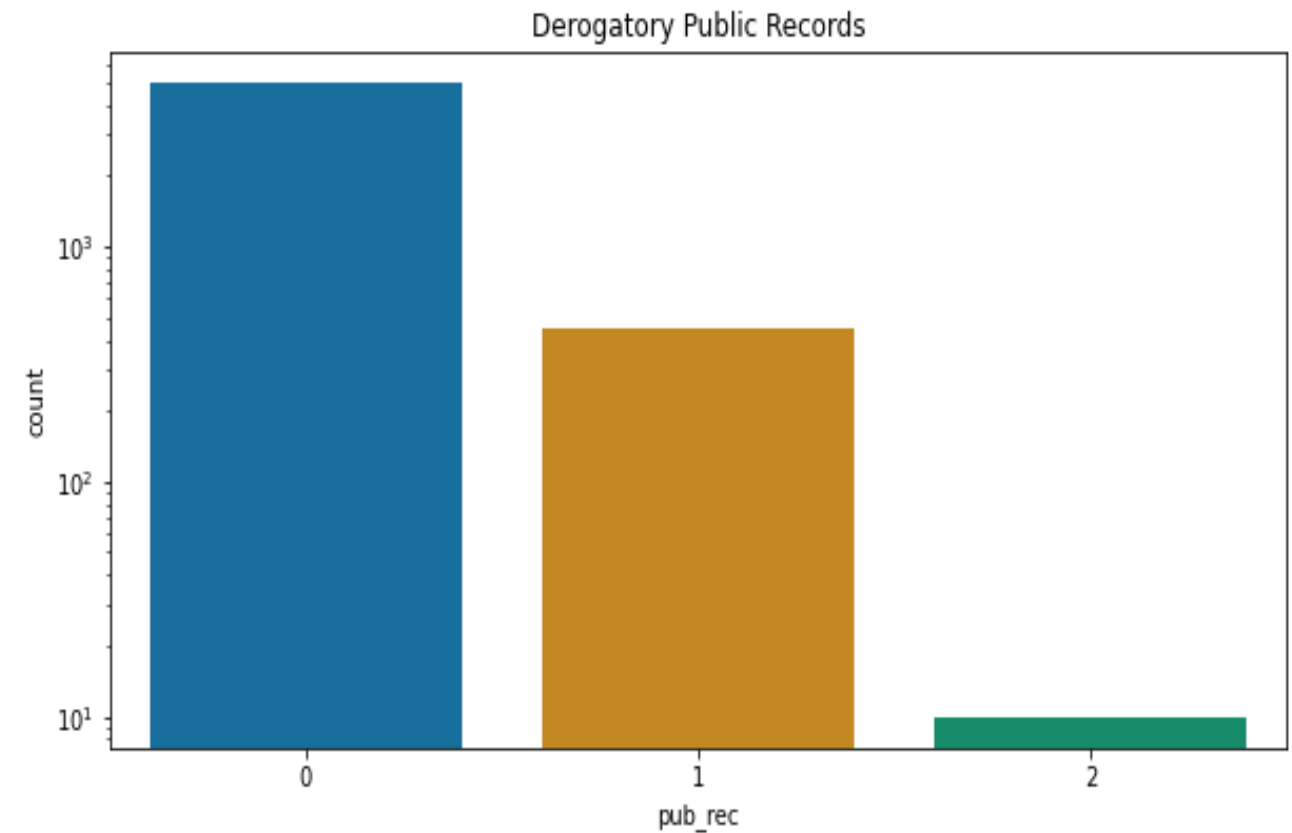
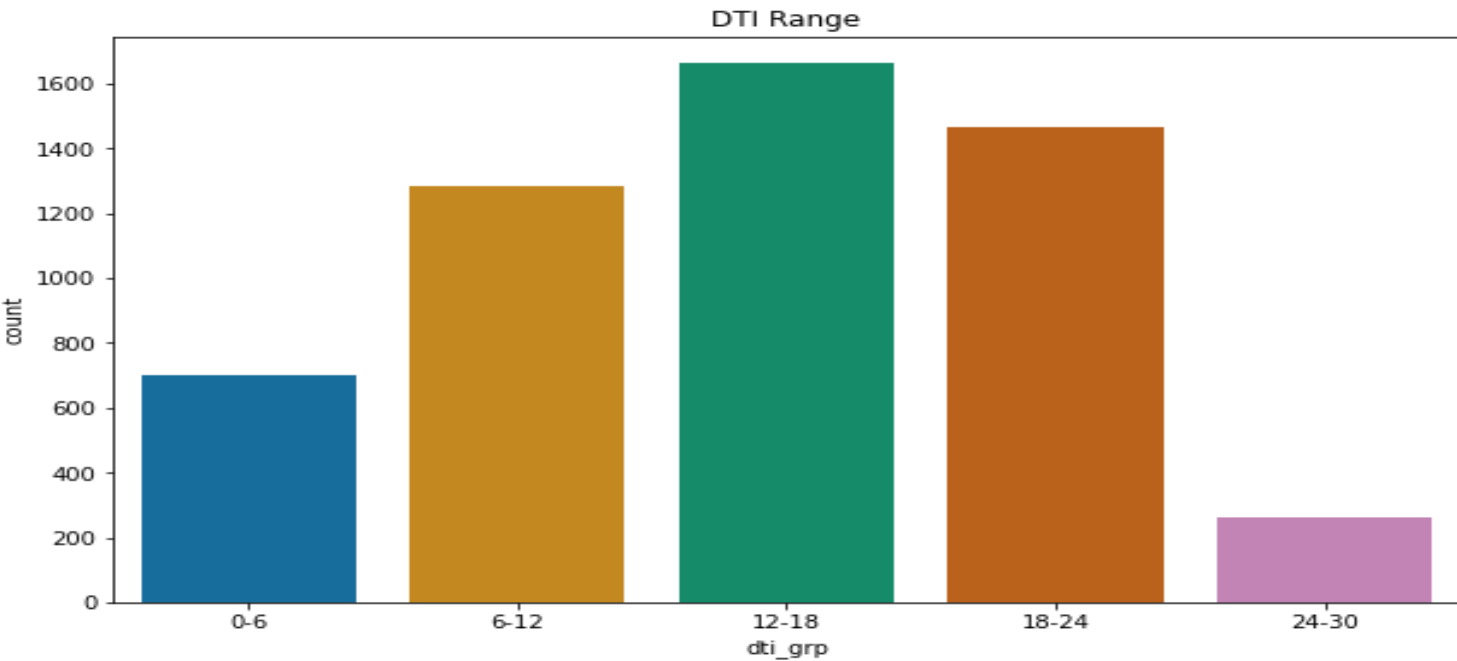
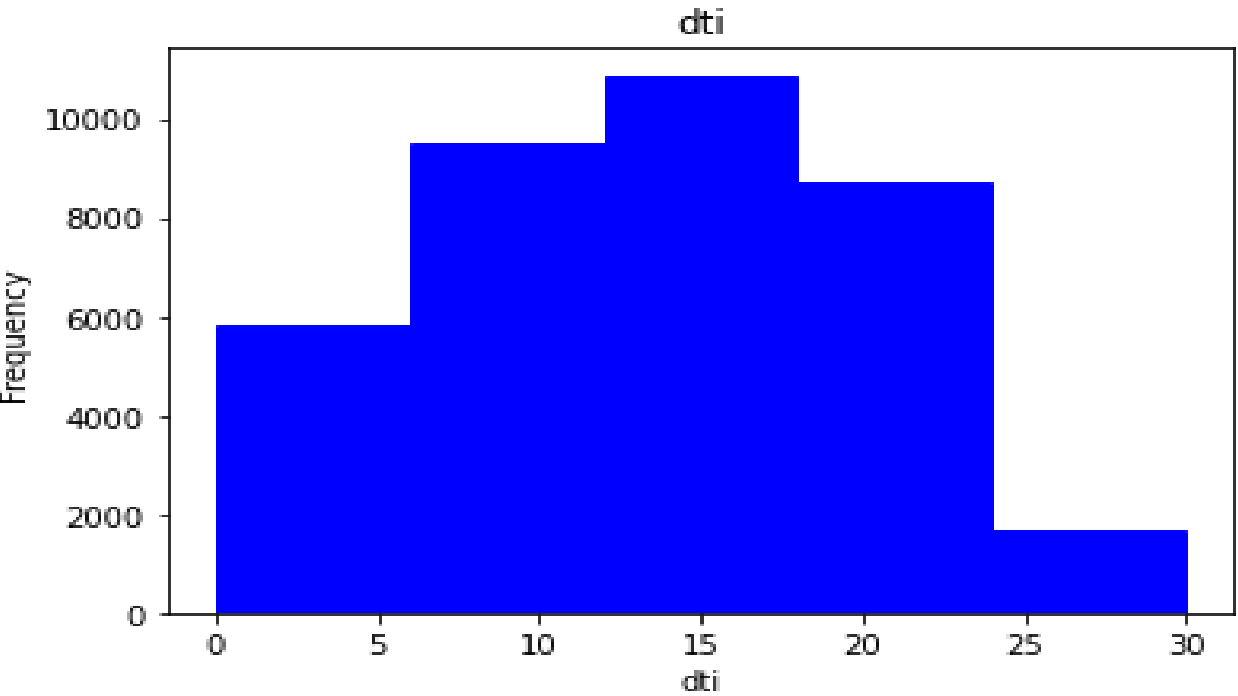
- Maximum number of loan applicants have their DTI roughly between 12 to 18 and they are moderately risky clients.
- Around 32% of loans approved for clients having DTI more than 18. This is more risky

Derogatory Public Records

- There are around 5.5% of loan applicants who has derogatory public record's, and they are risky applicants.
- More number of defaulters exists in 0 number of derogatory public record

Home ownership

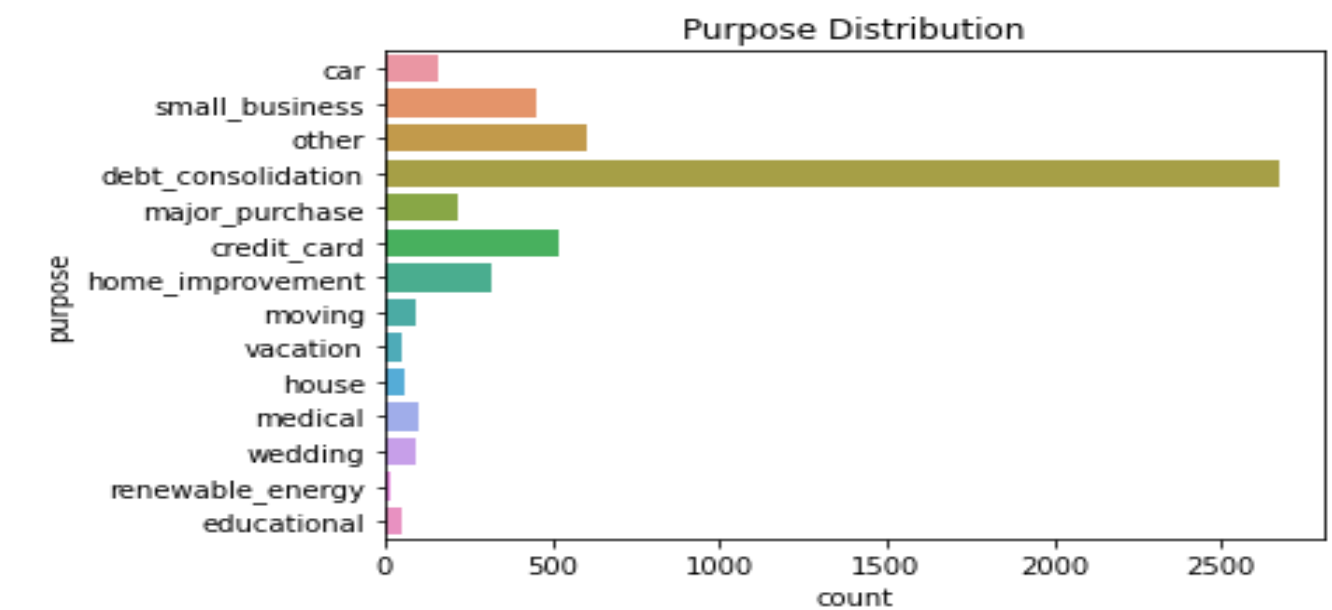
- Rent and Mortgage home ownership clients are riskier when compared to other categories.
- Rented home clients are the riskiest of all.



Analysis Results / Insights cont...

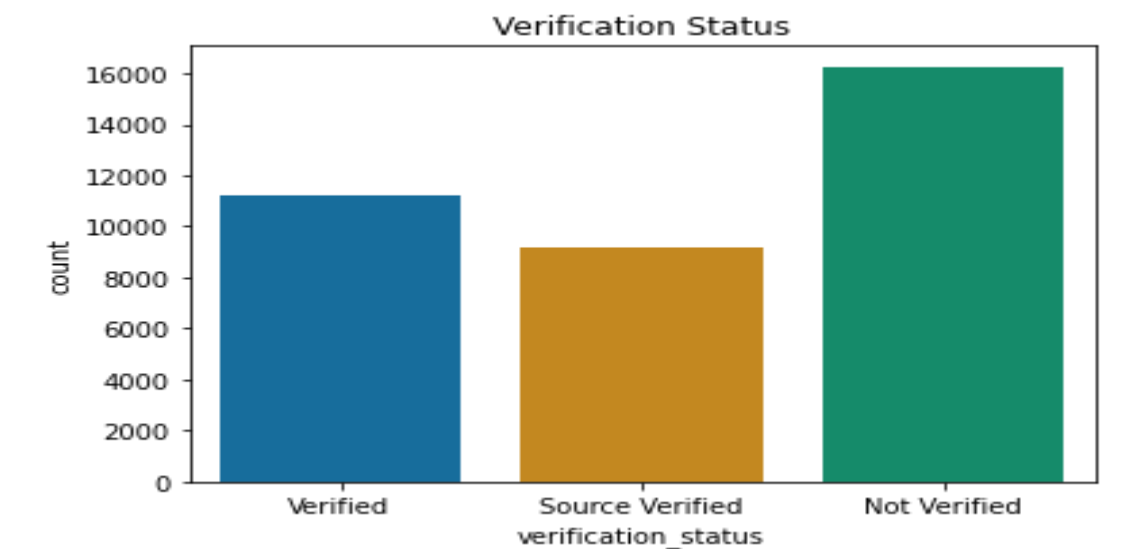
Purpose of Loan insights

- 47.24% of clients bought the loan for Debit Consolidation purpose
- Within the charged-off accounts also, applicants who bought loans for Debt-Consolidation purpose are more defaulters.



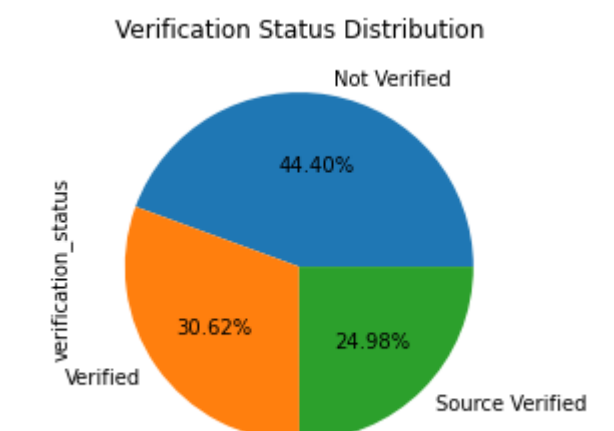
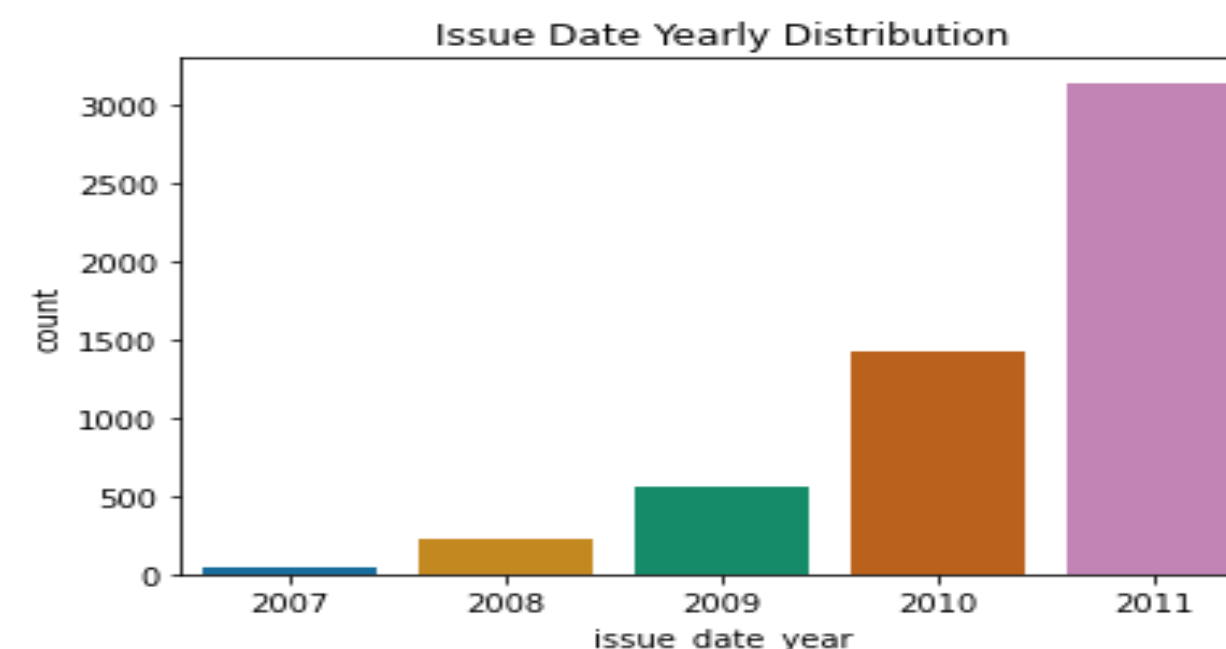
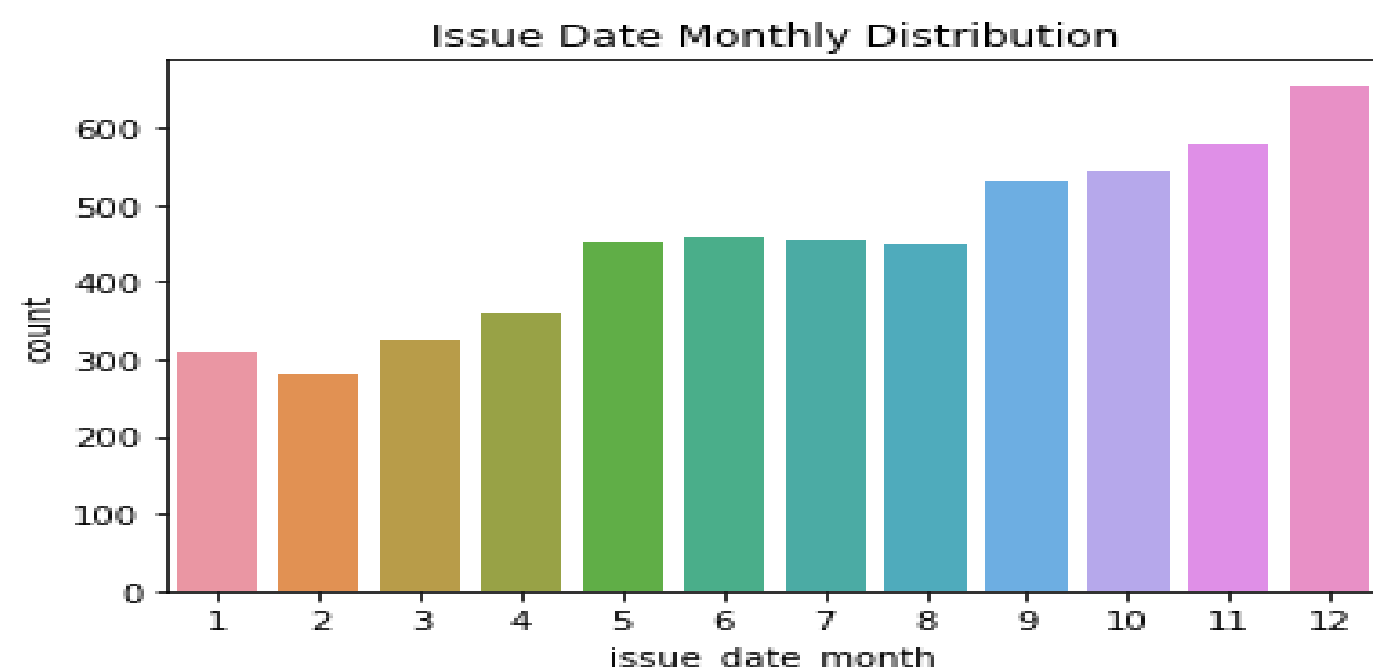
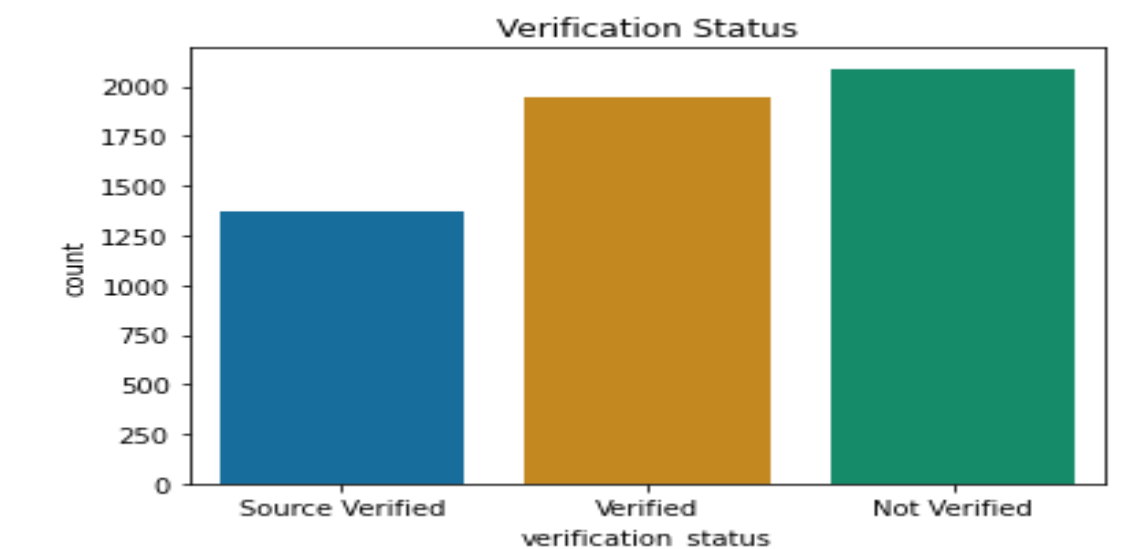
Verification status

- 44.4% of clients got loan approval without being verified. This is very risky
- Higher number of defaulters are also in Not Verified status only



Issue date distribution

- The loans which are issued in the month of Dec are more defaulted.
- The loans which are issued in the year 2011 are more defaulted. It might be because of recession.



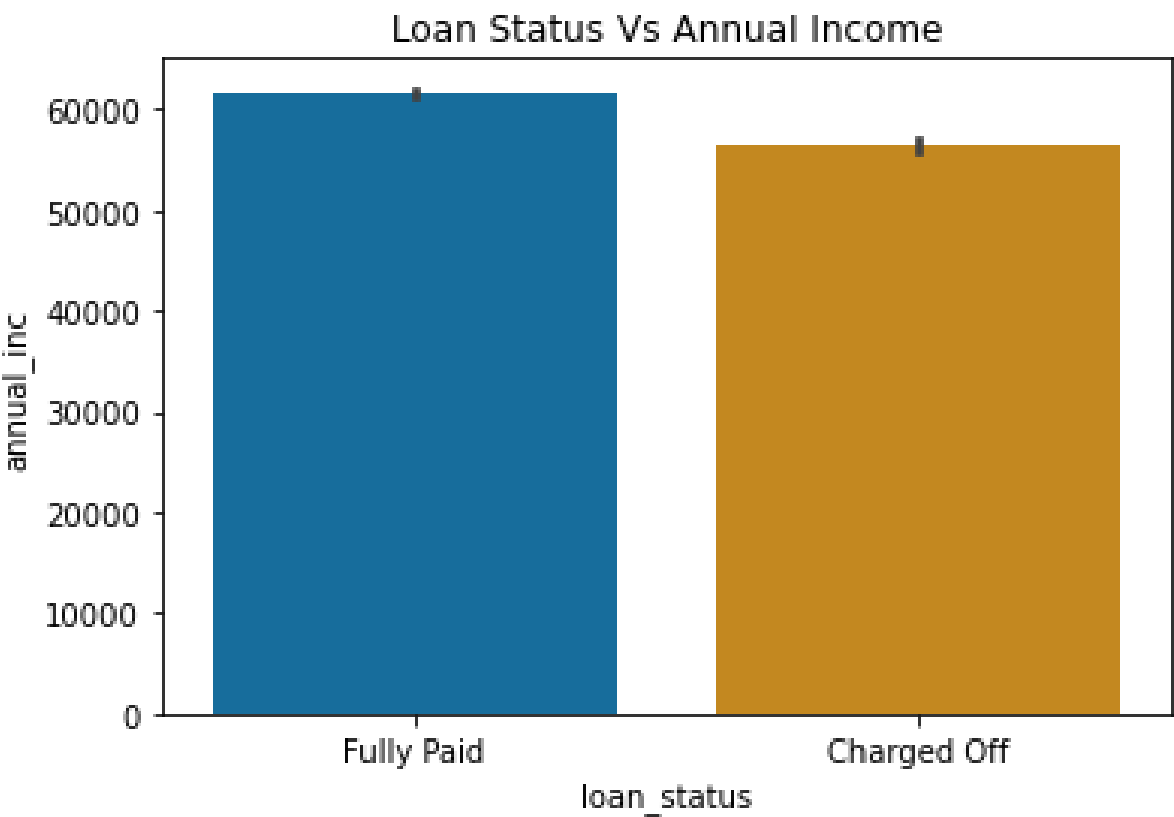
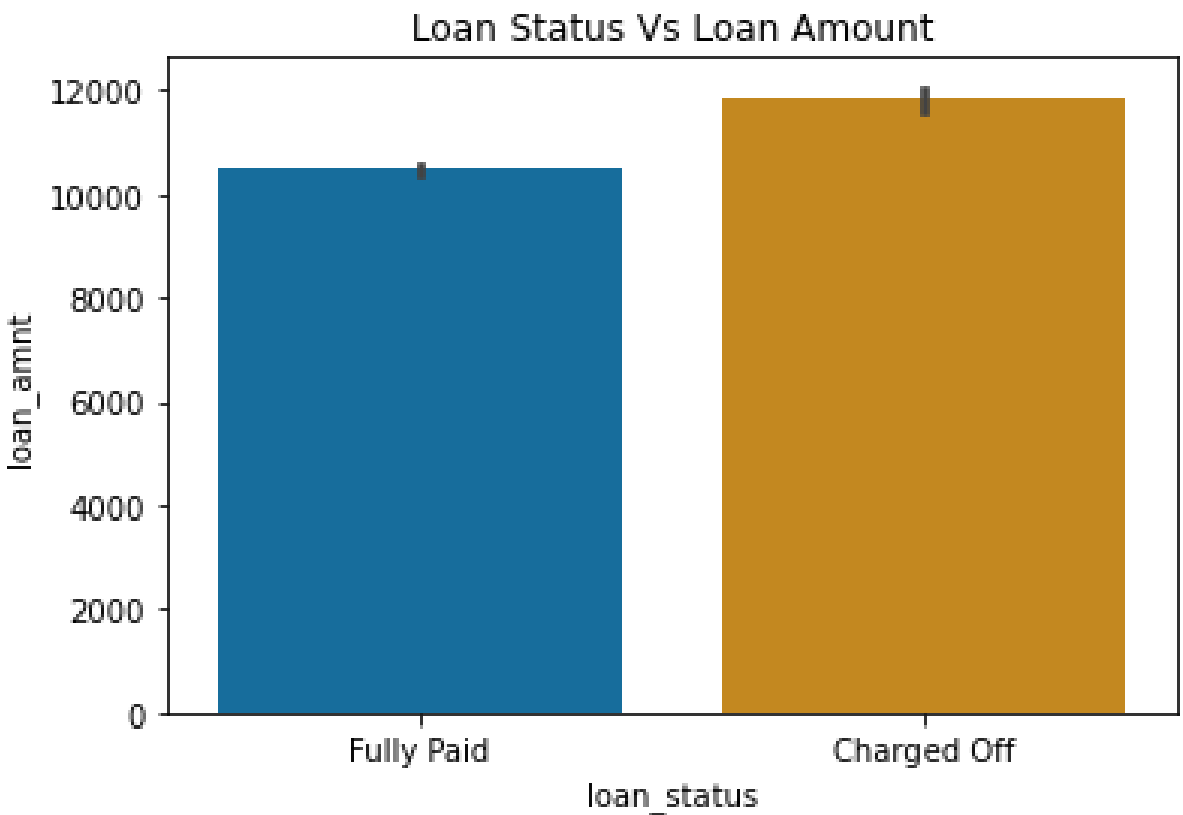
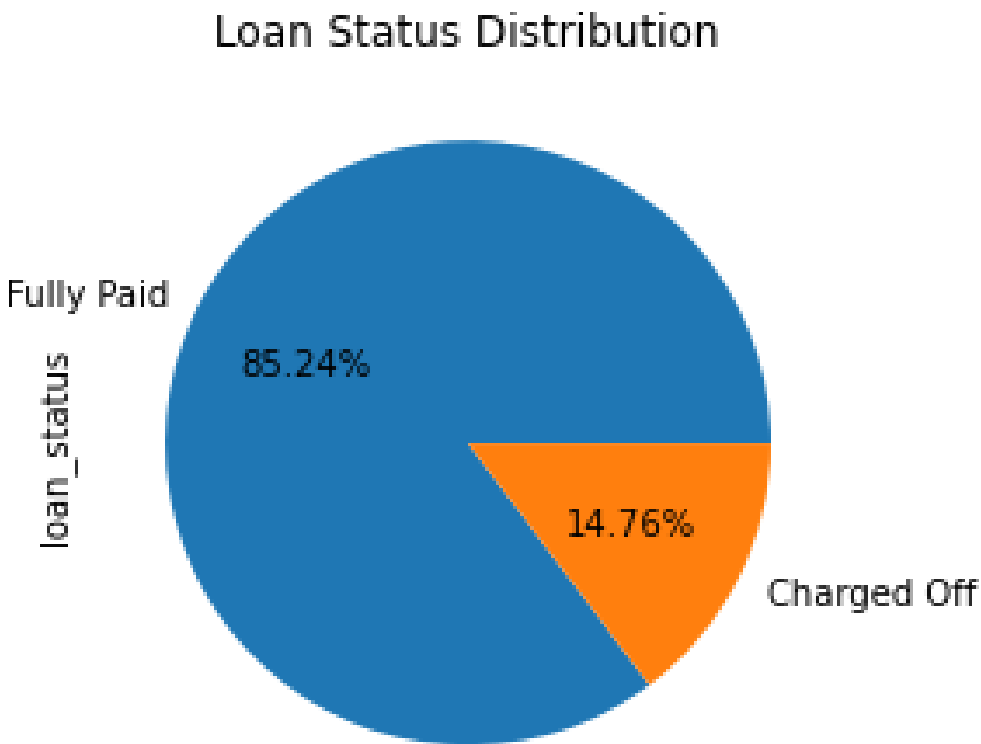
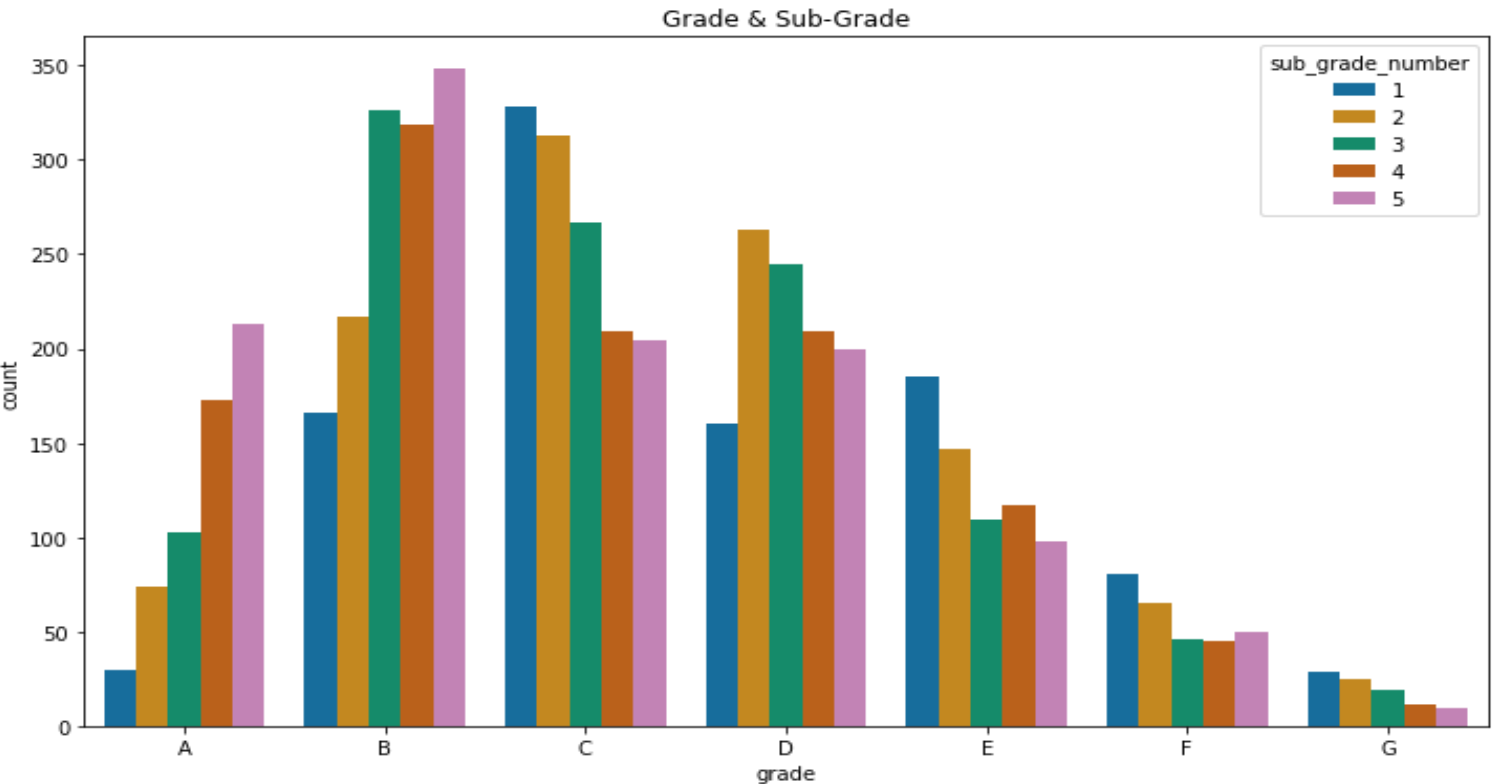
Analysis Results / Insights cont...

Grades and sub-grades

- Highest number of defaulters are in Grade B irrespective of sub-grades, of which the highest number of defaulters are in sub-Grade B-5

Loan status

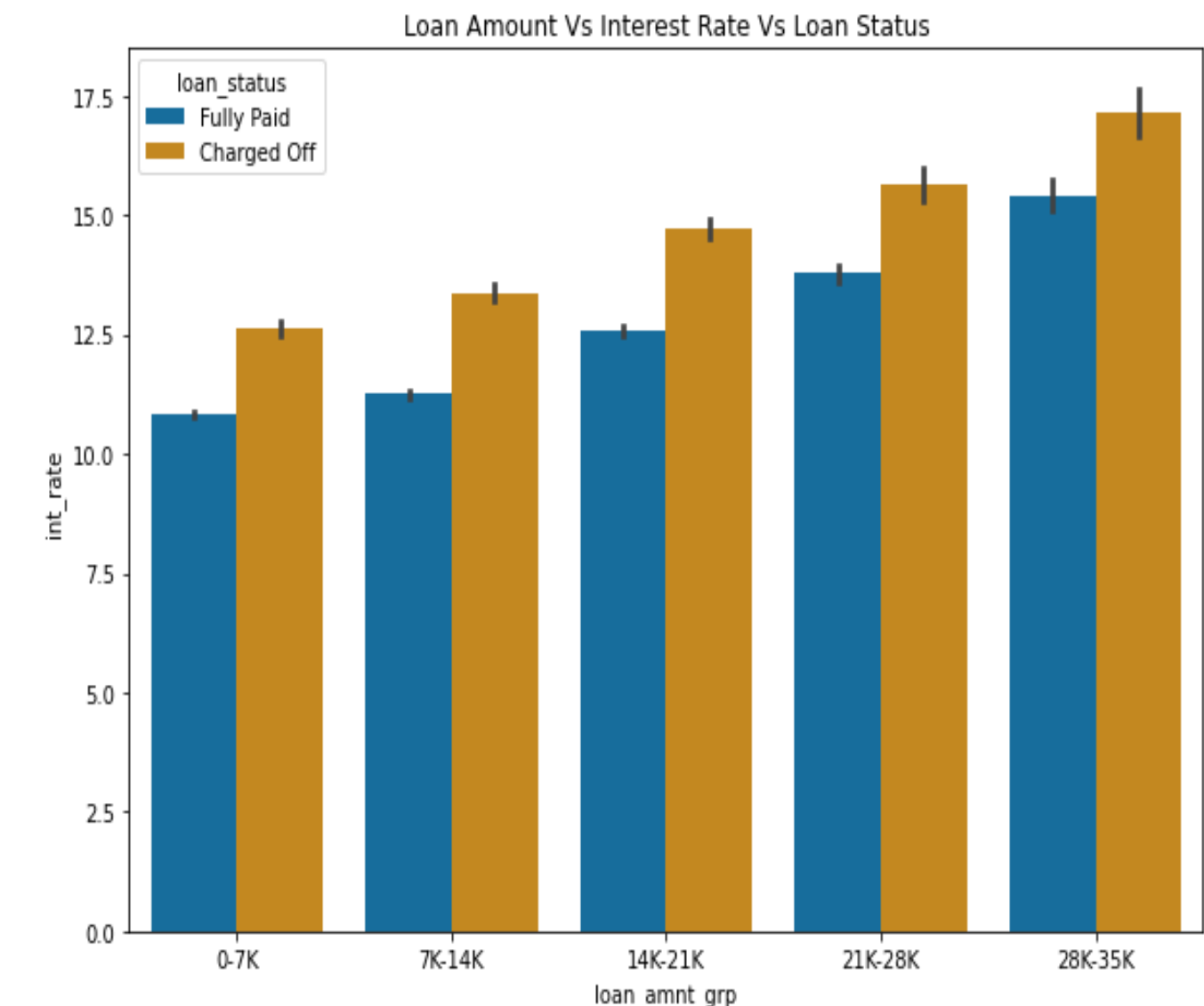
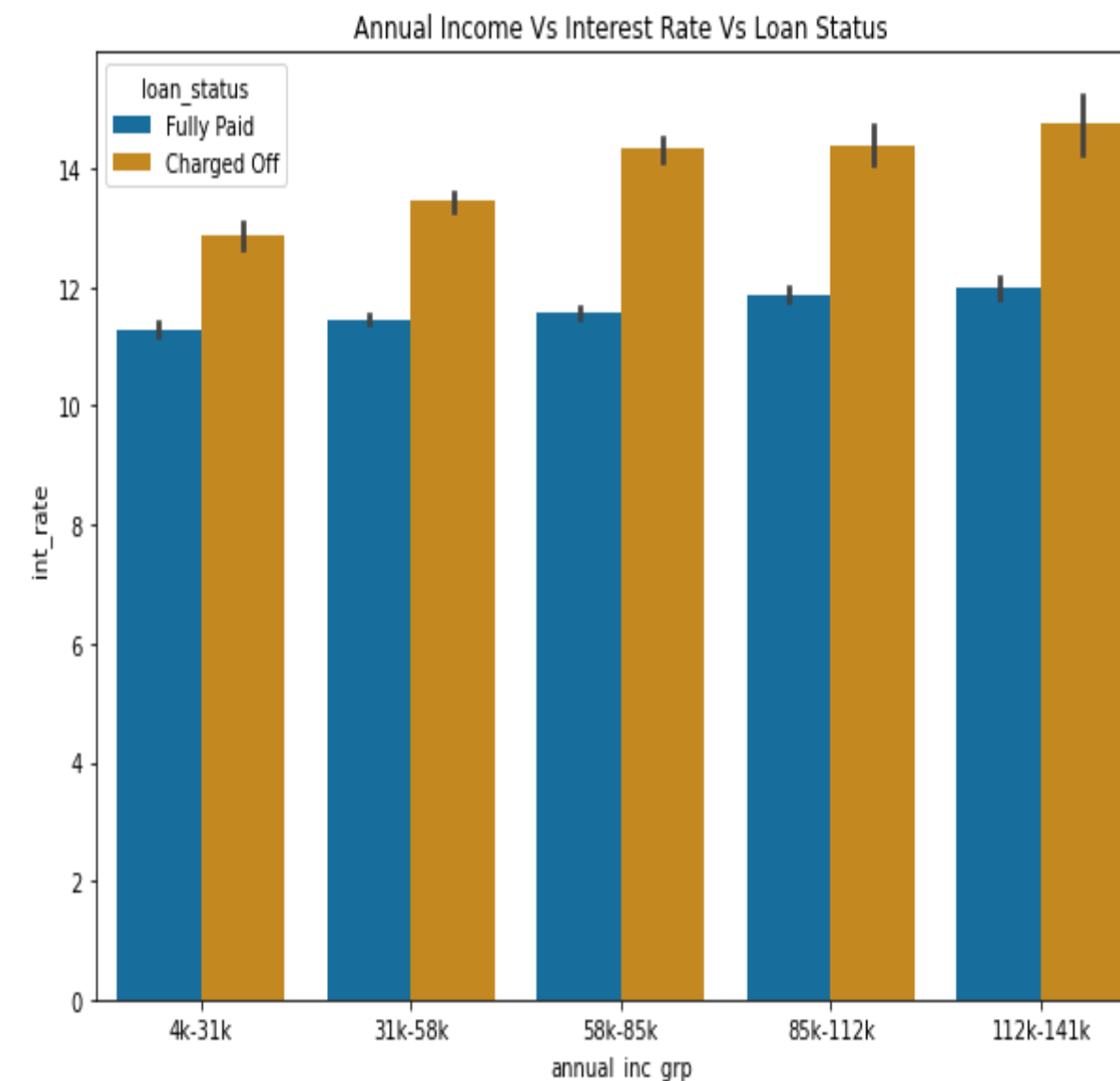
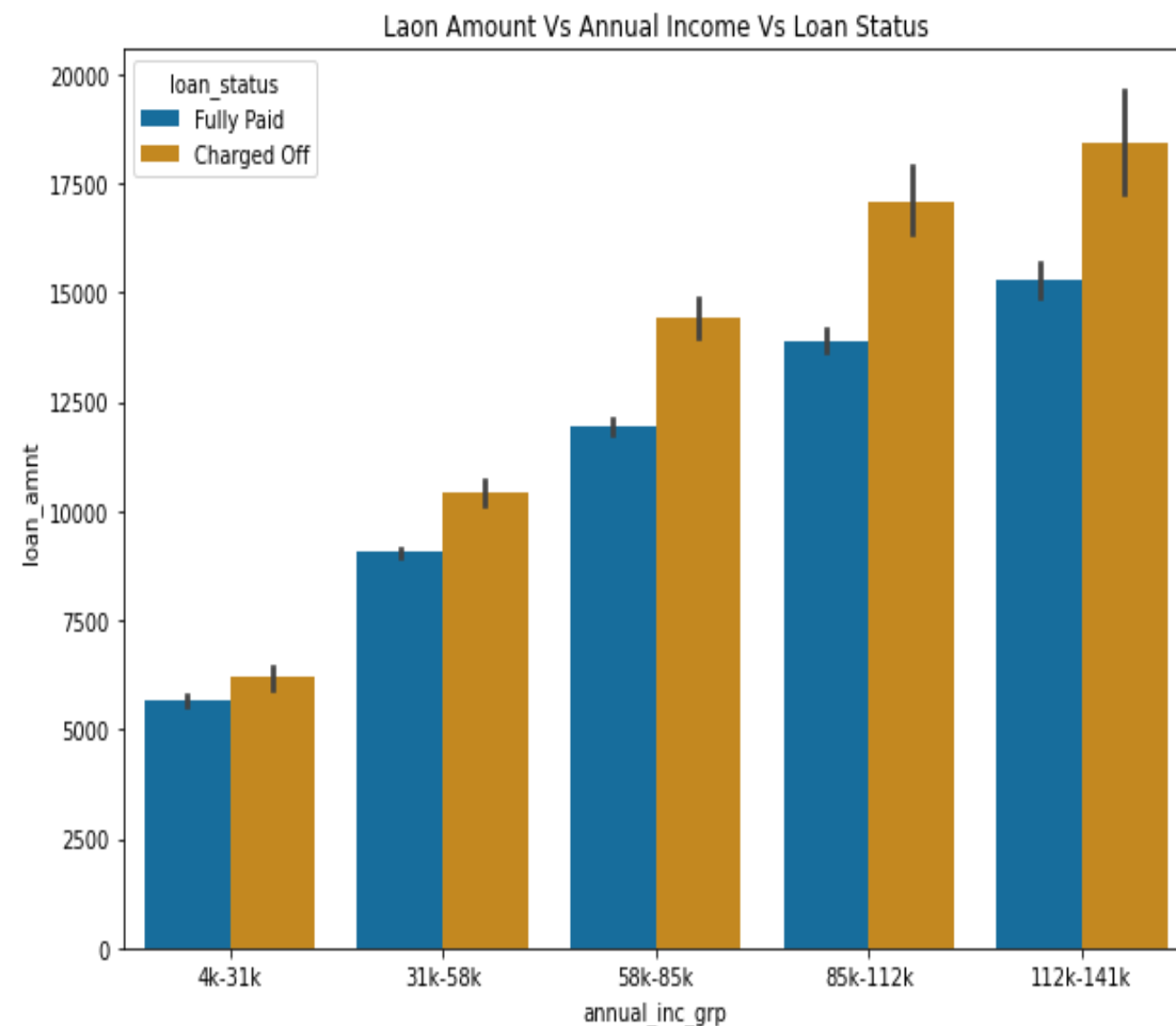
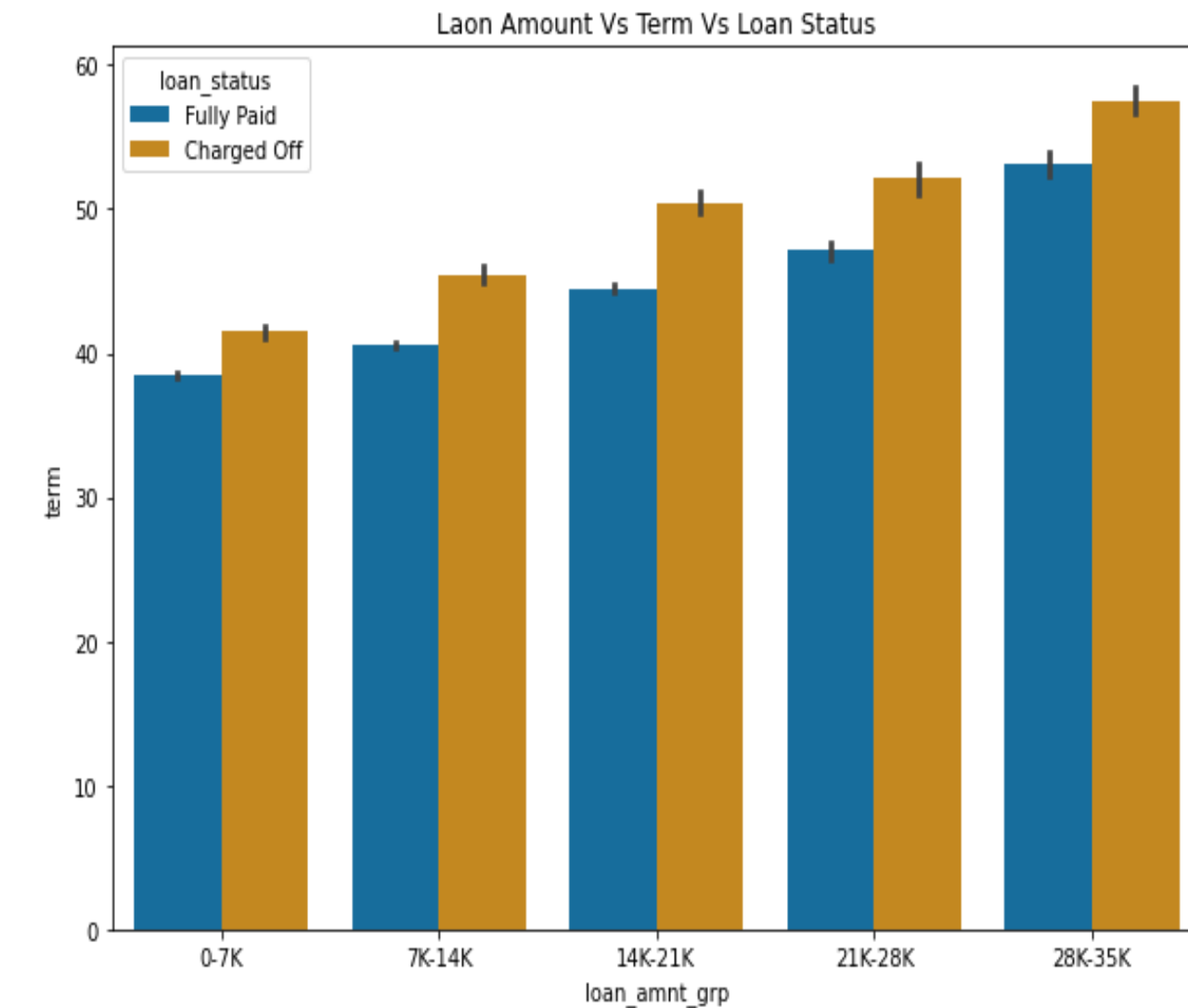
14.76% (5402 applicants) of the entire applicants are charged off.



Insights of Variable Relationships

The probability of defaulters could be because of below reasons:

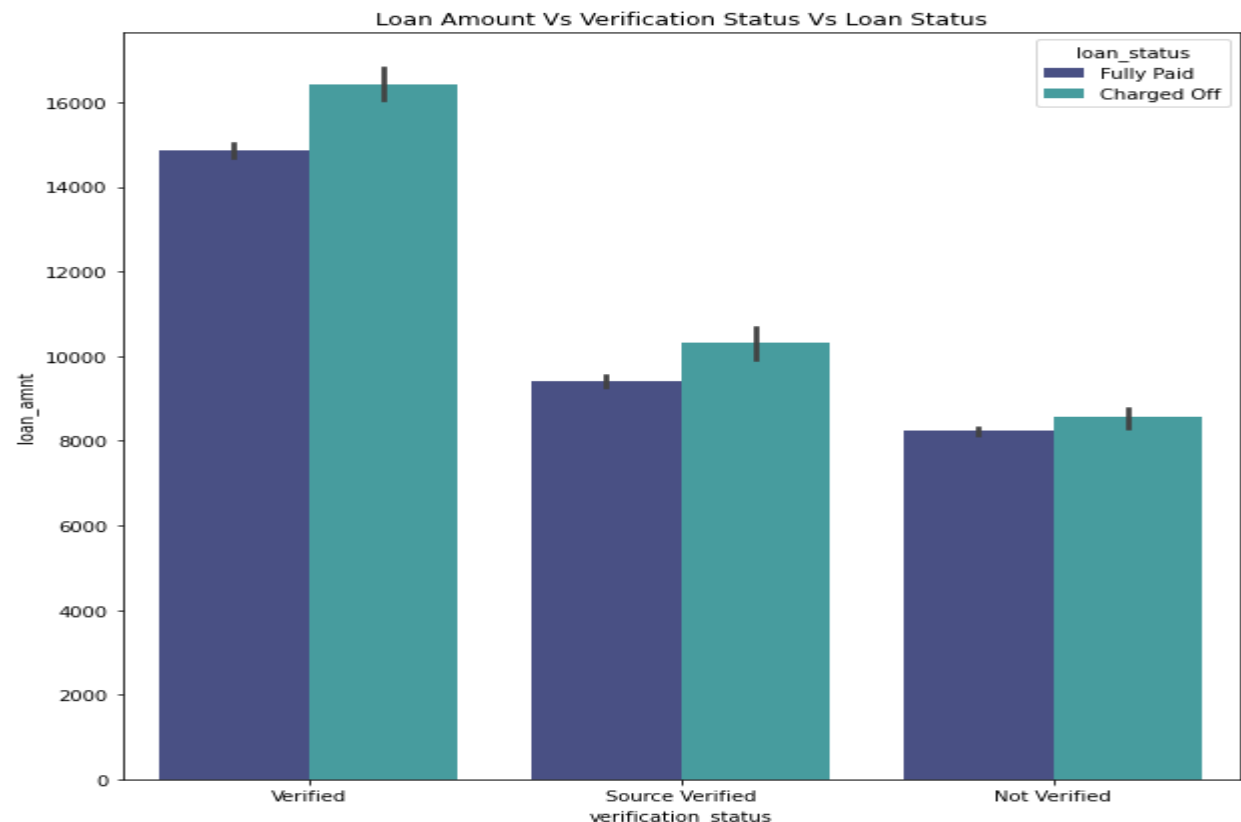
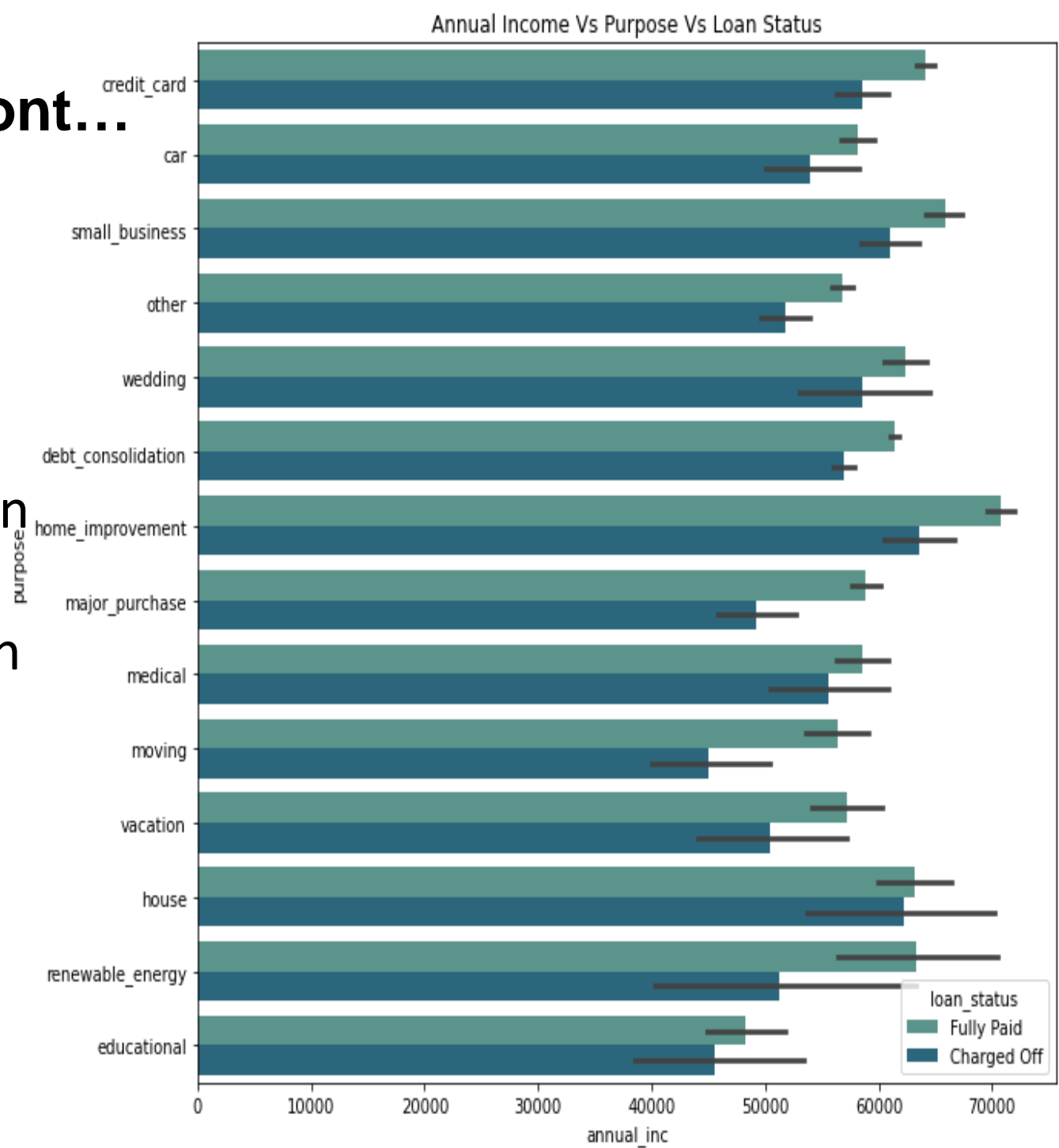
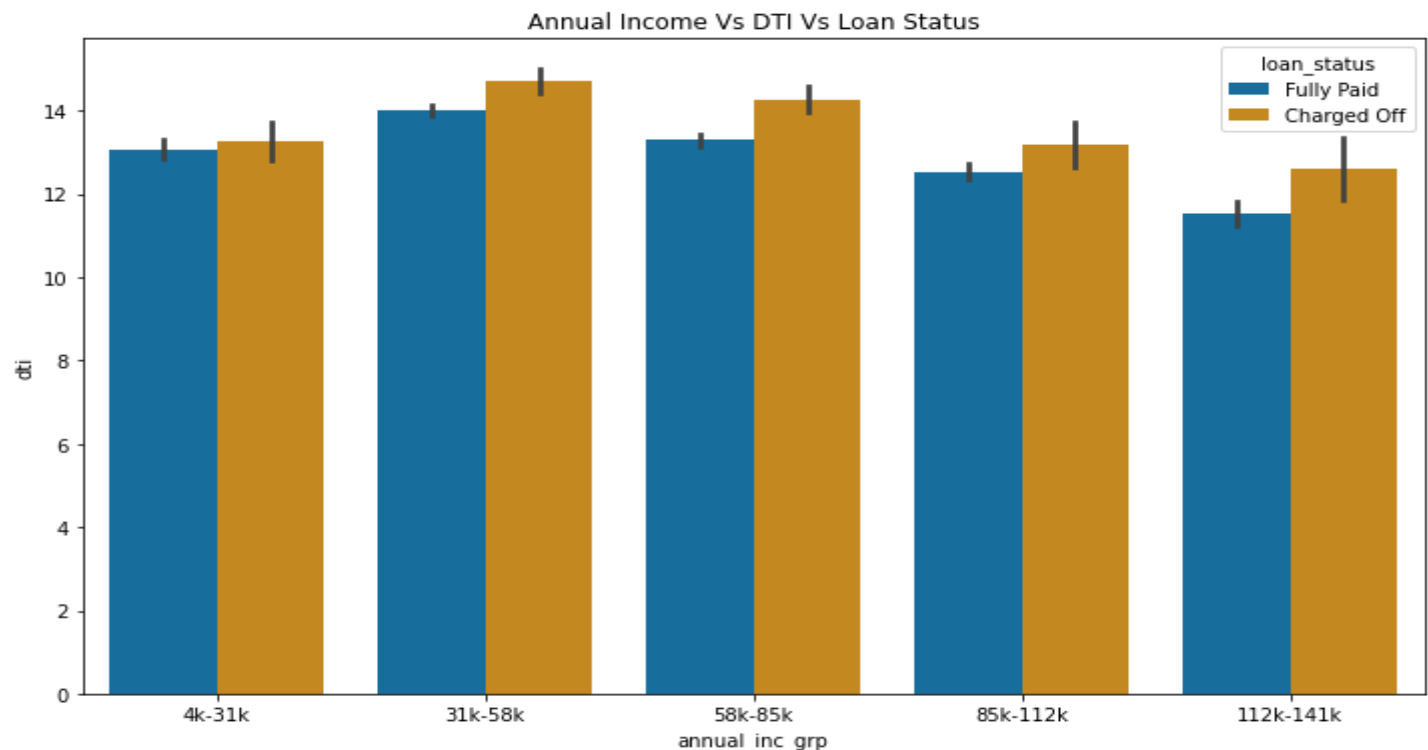
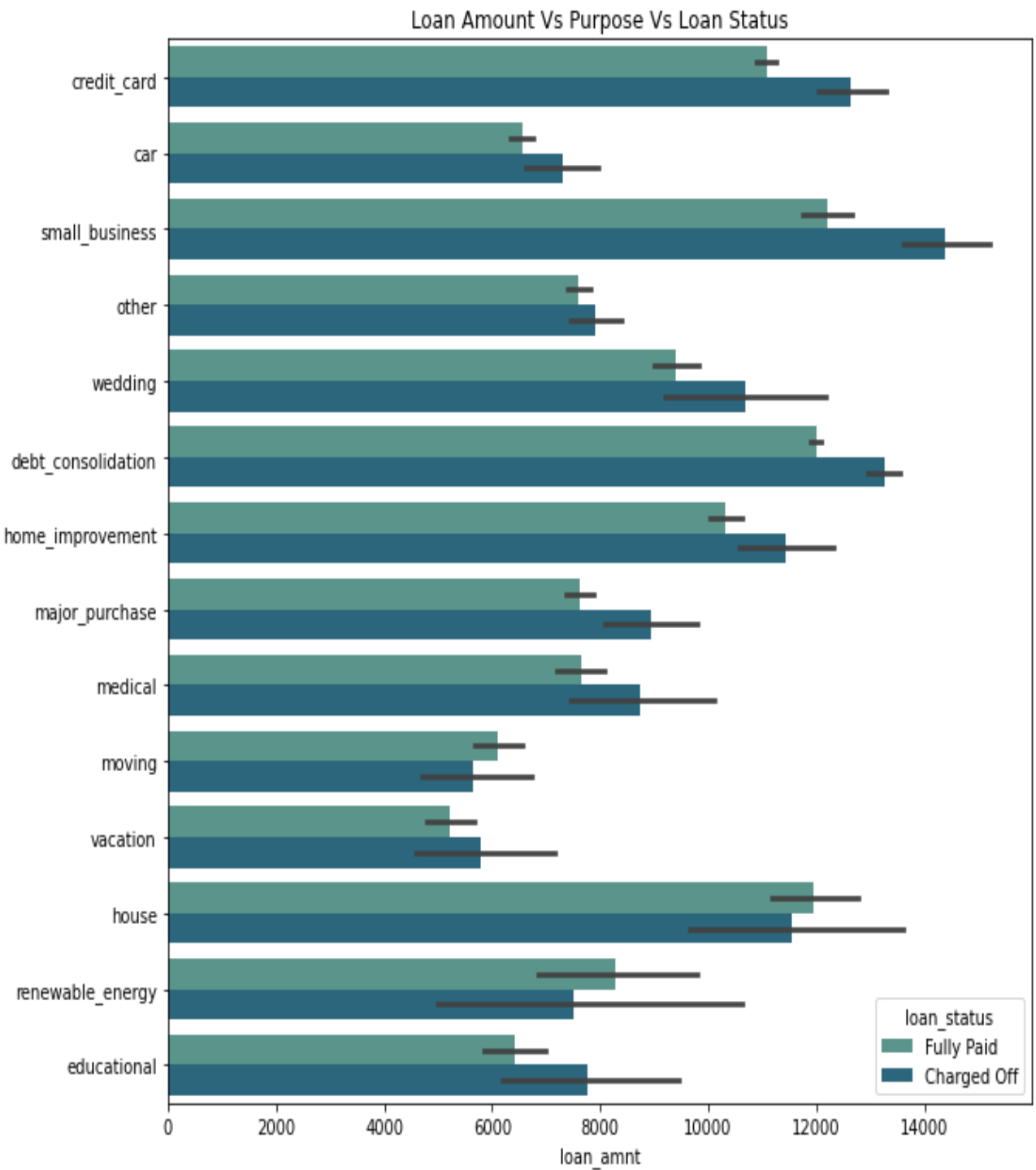
- ✓ Client has loan amount in the range of \$28K to \$35K with term more than 50 months
- ✓ Client has annual income in range of \$112K-\$141K with loan amount of more than \$15K
- ✓ Client with rate of interest more than 11% irrespective of annual income
- ✓ Client with loan amount in the range of \$28K to \$35K with interest rate more than 15%
- ✓ Interest Rate is more
- ✓ Client who applies for more loan cross all income group



Insights of Variable Relationships cont...

The probability of defaulters could be because of below reasons:

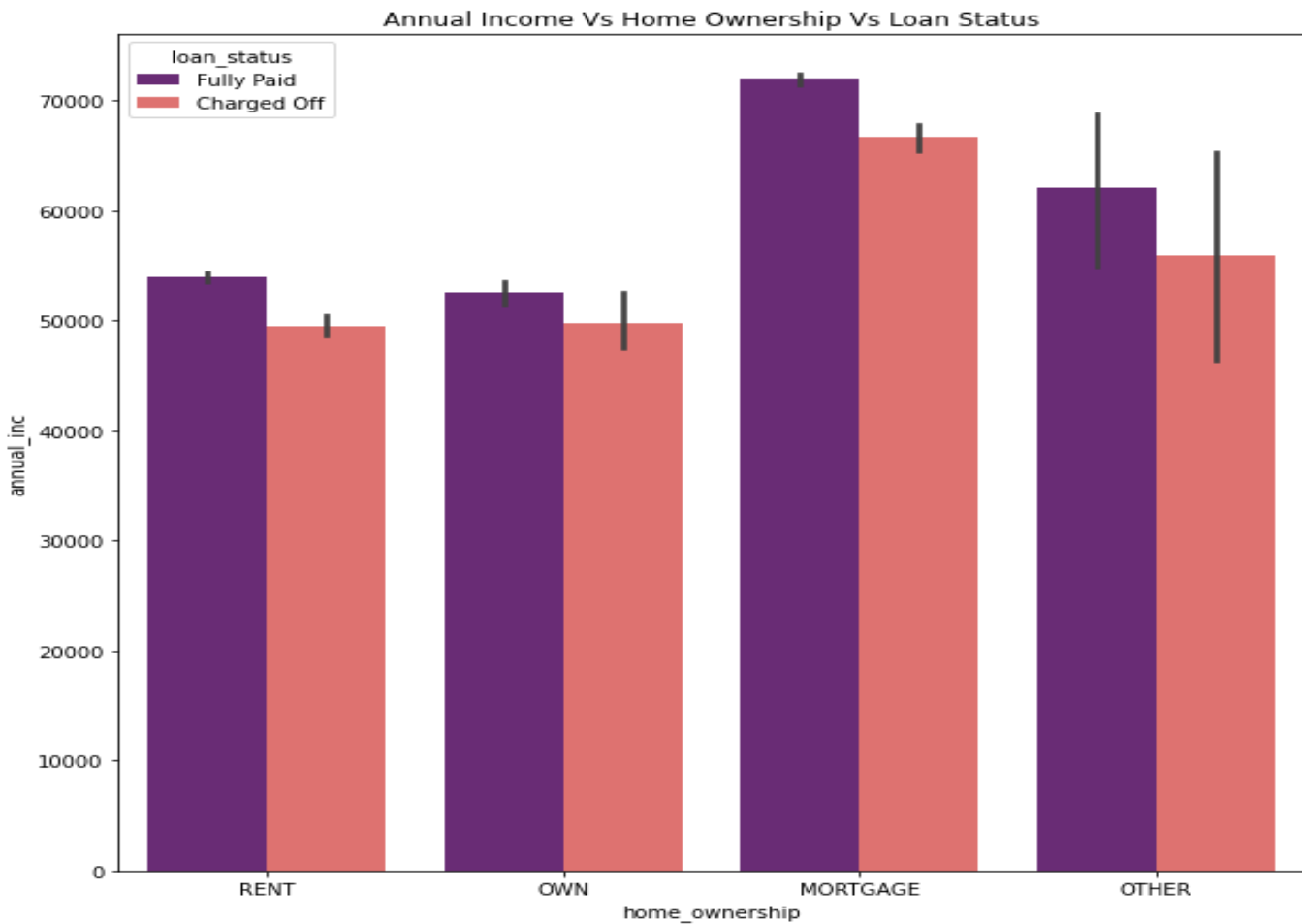
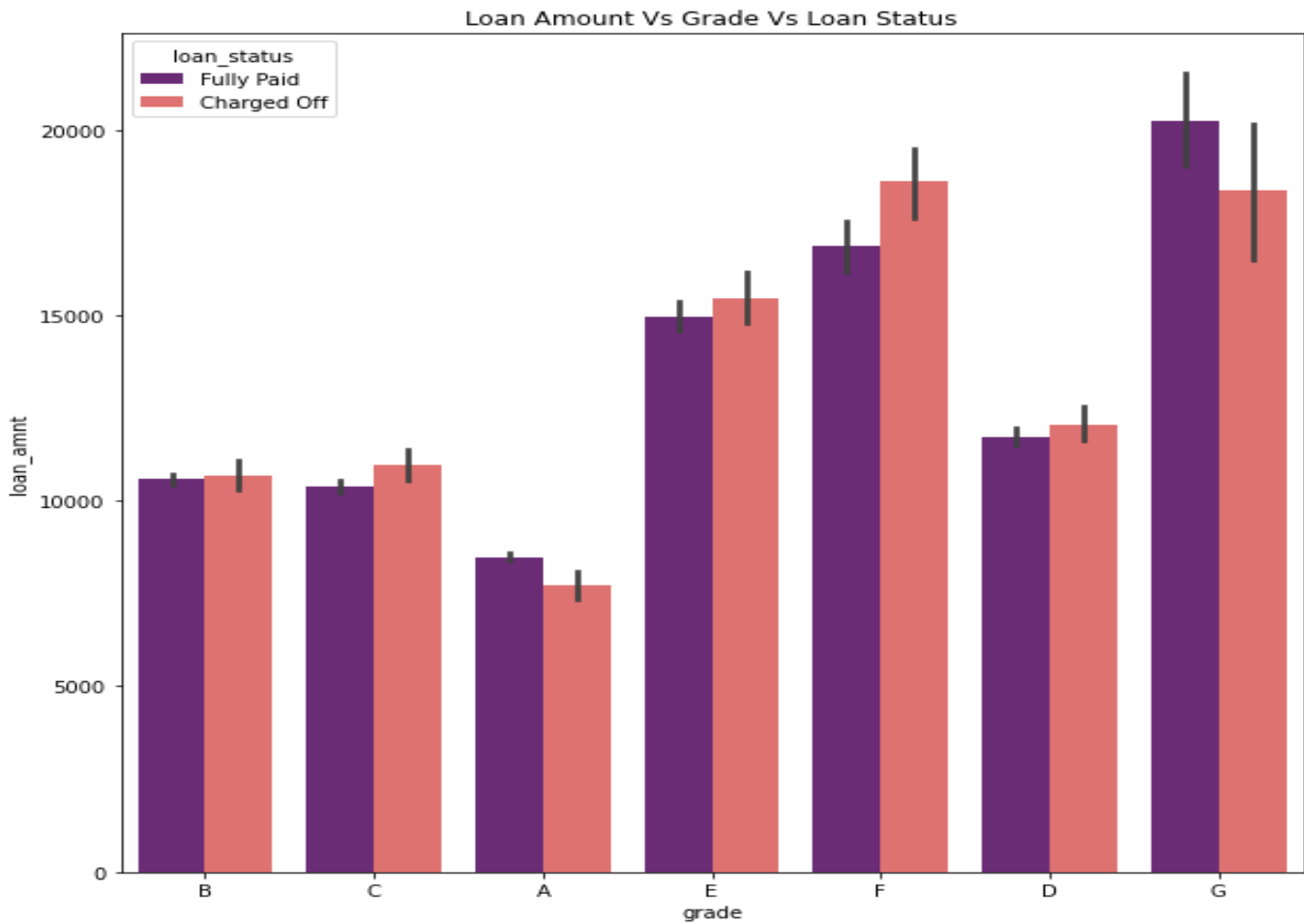
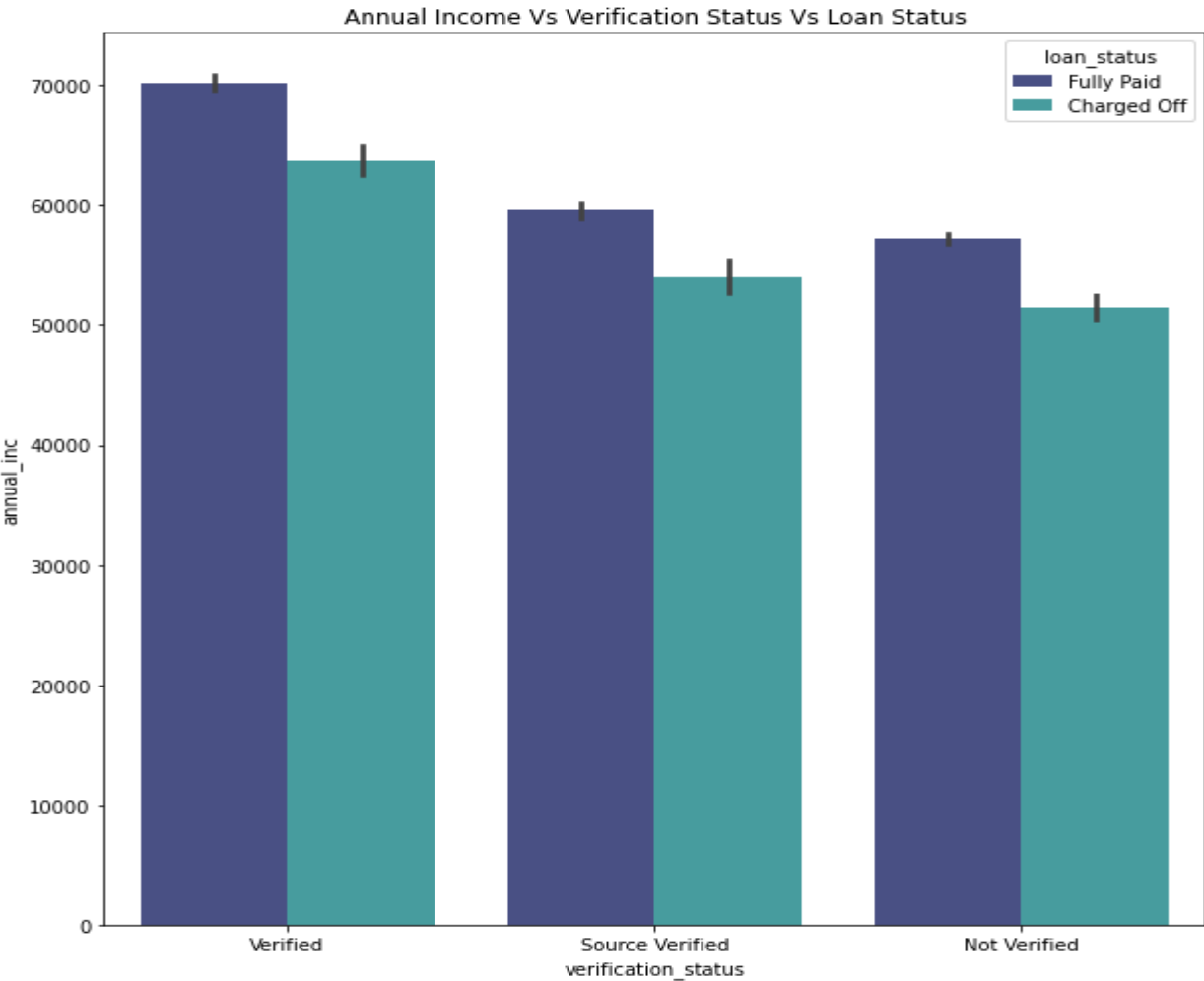
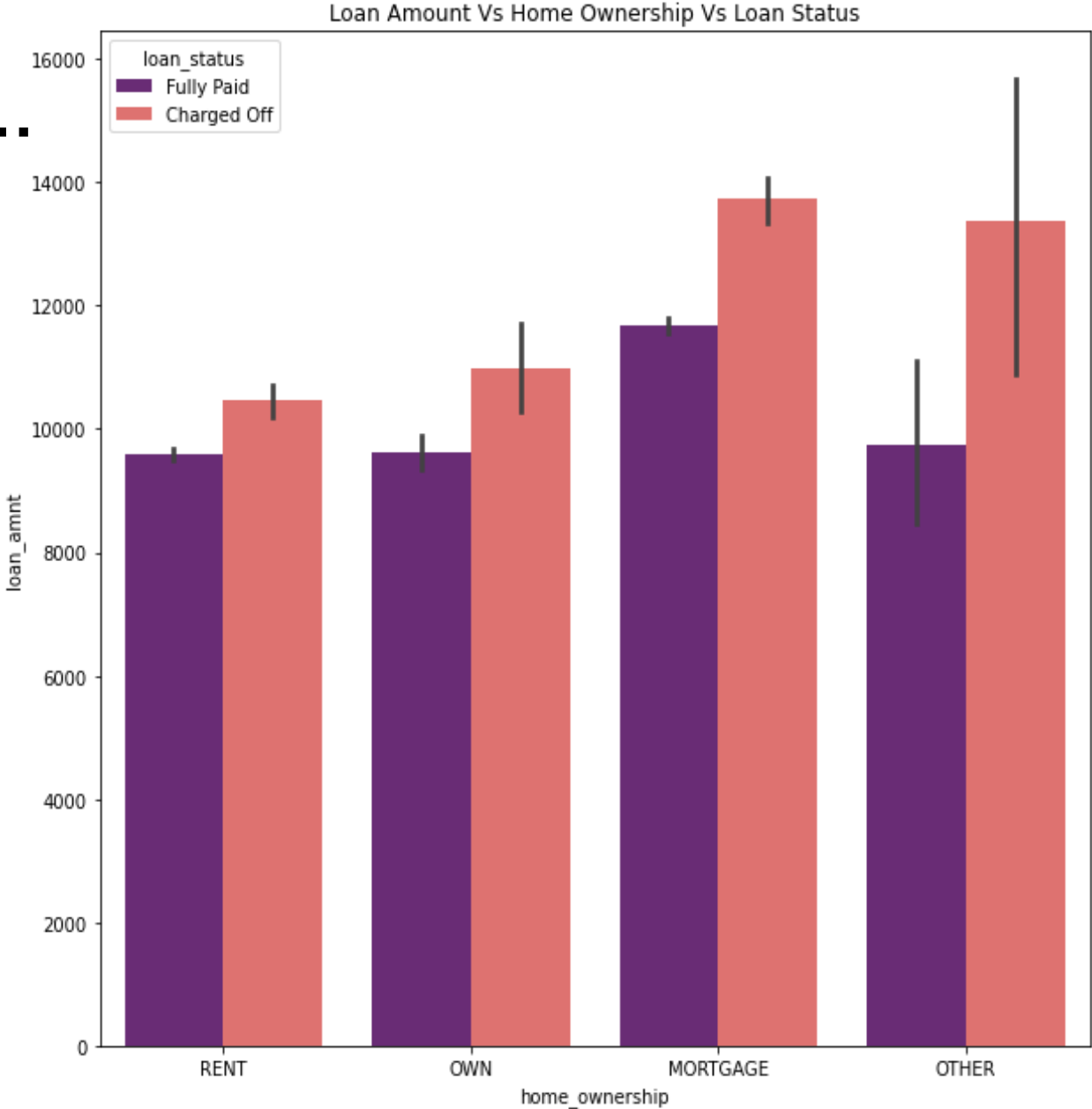
- ✓ Clients requesting small business loans with loan amount above \$12K
- ✓ Clients requesting home improvement loan with annual income of above \$65K
- ✓ Client with annual income in the range of \$31K to \$58K with DTI more than 14
- ✓ Verified loan account client requesting loan amount more than \$15K



Insights of Variable Relationships cont...

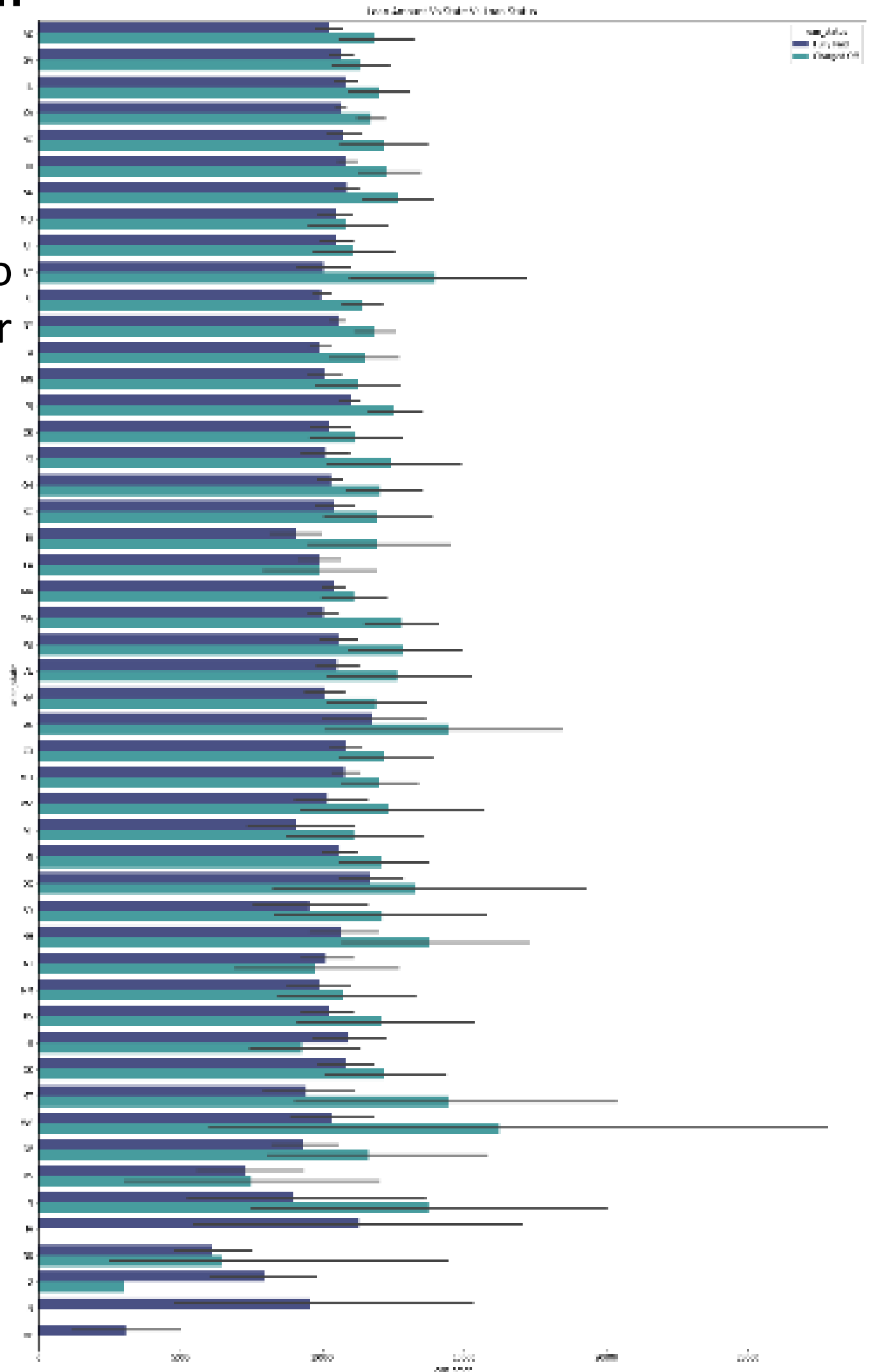
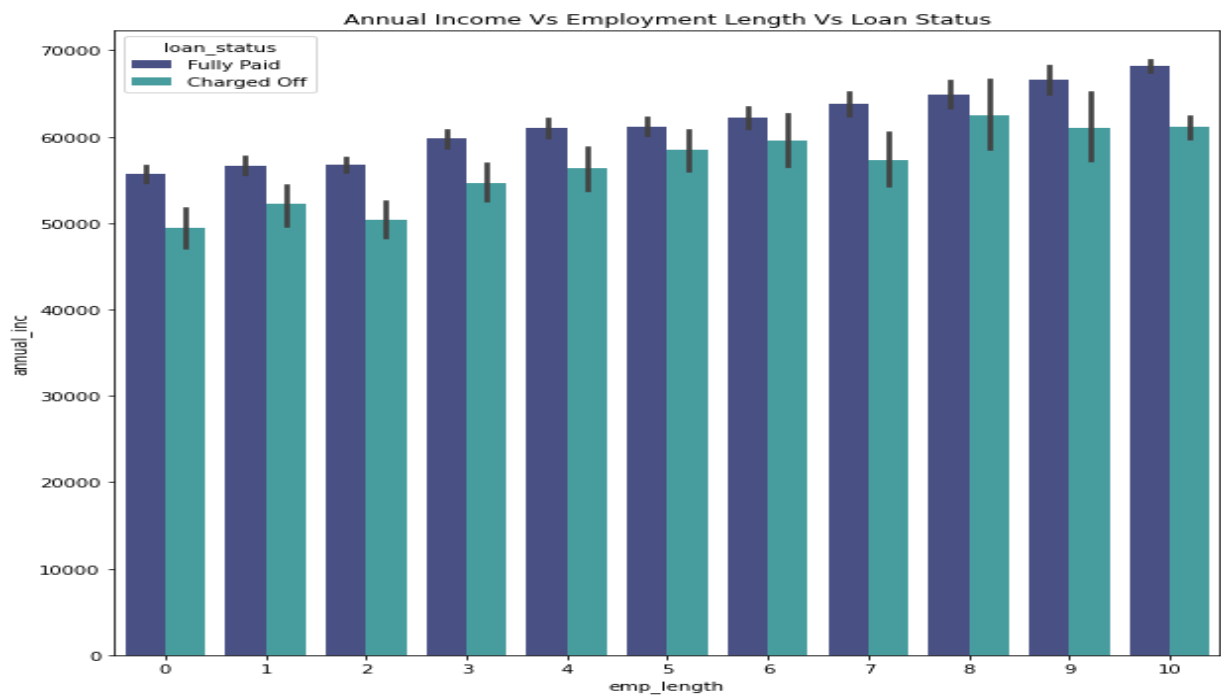
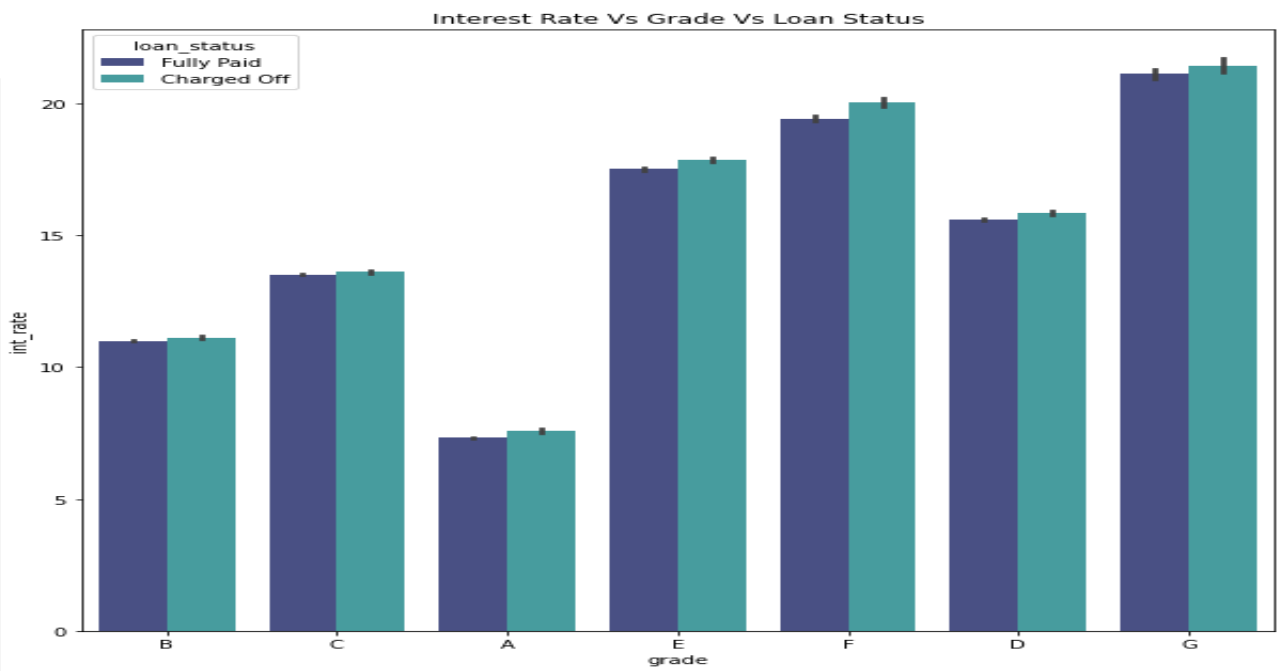
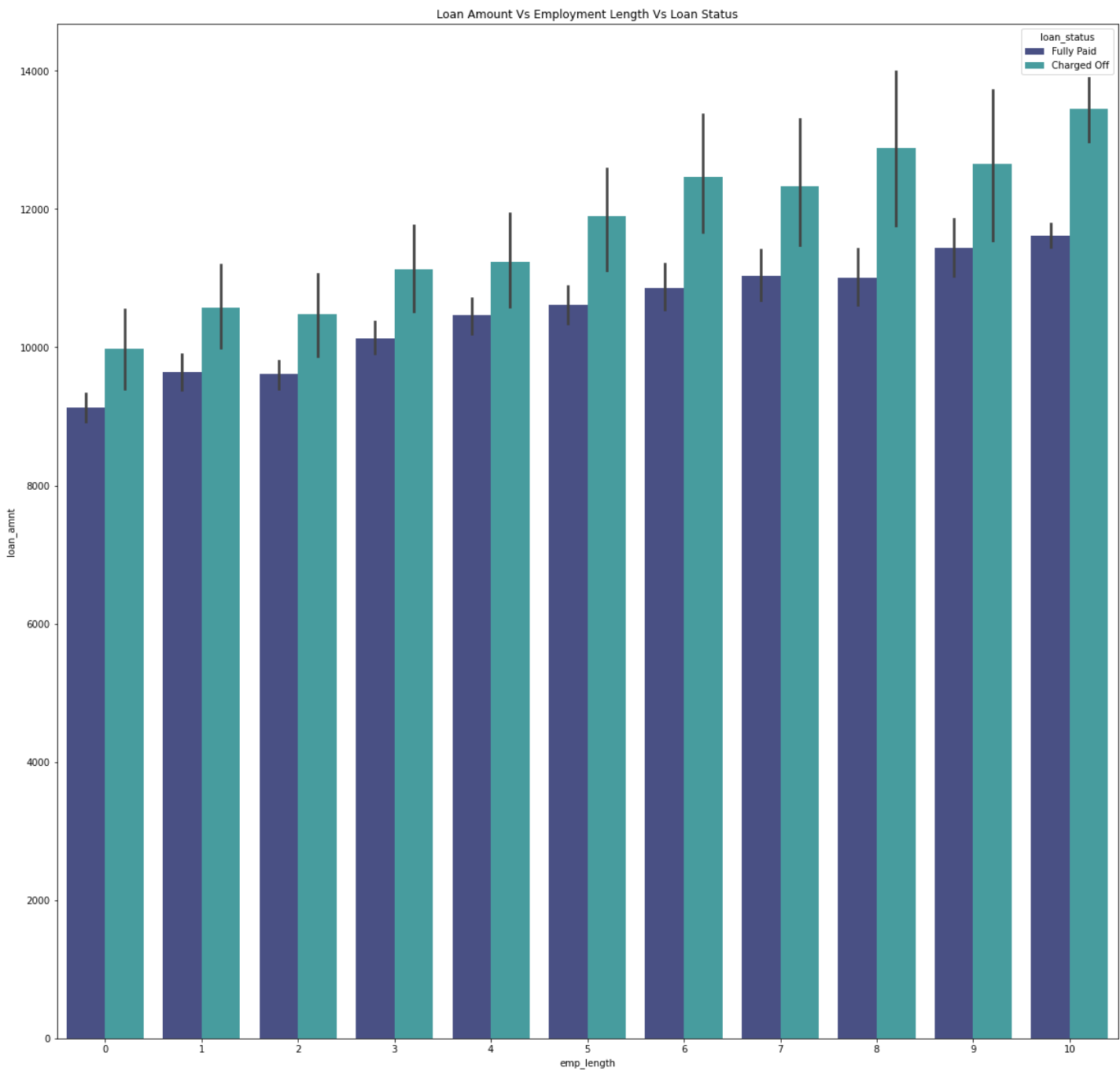
The probability of defaulters could be because of below reasons:

- ✓ Client whose home ownership is Mortgage or Others and has loan amount more than \$12K
- ✓ Client whose home ownership is Mortgage has more defaulters compare to other home ownership wrt annual income
- ✓ If Annual Salary is less, defaulters exists across all verification status
- ✓ Client with grade F & G and has loan amount more than \$15K. This could be because of high interest rate and loan amount disposed



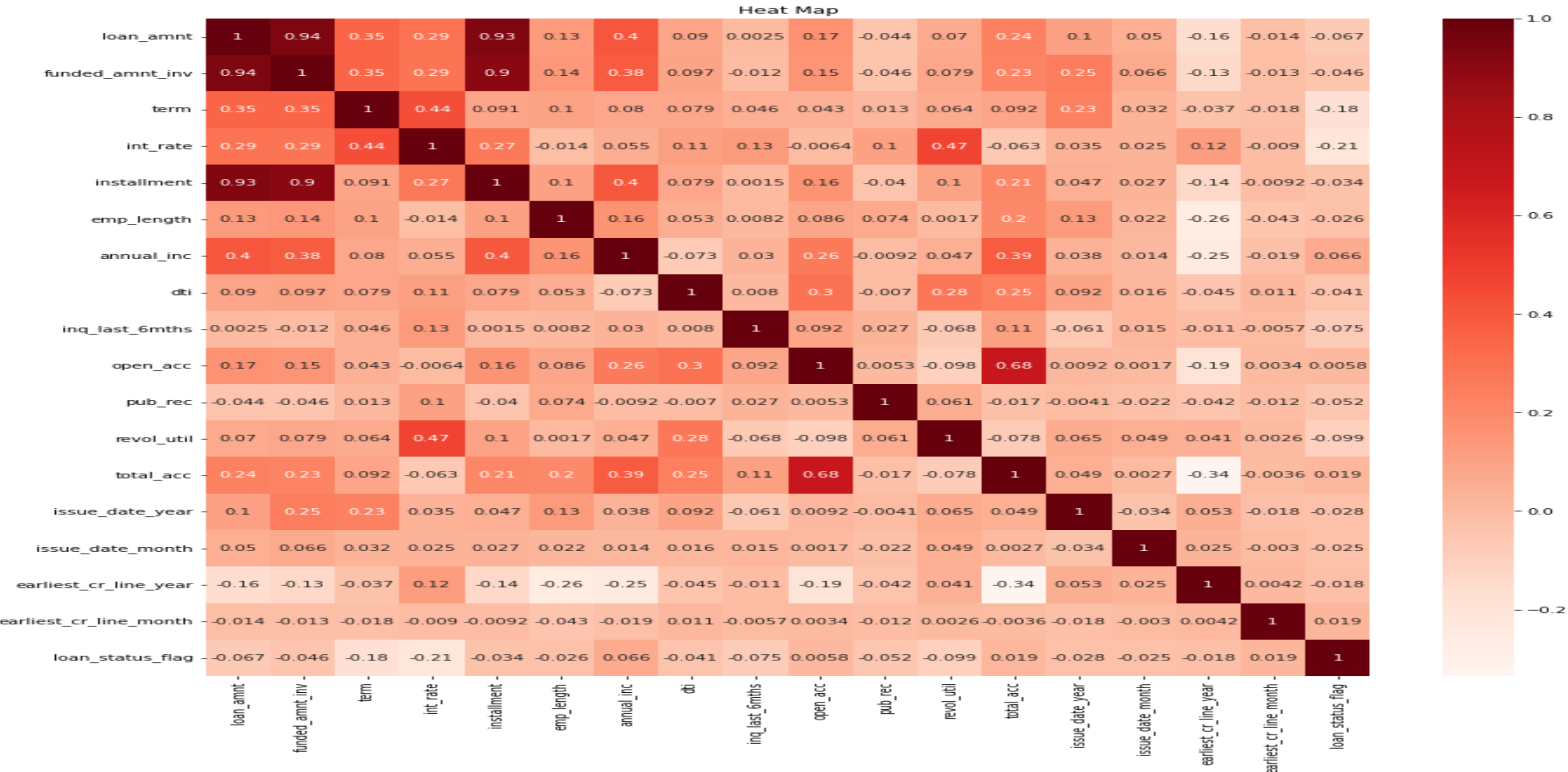
Insights of Variable Relationships cont...

- The probability of defaulters could be because of below reasons:
- ✔ More the employment experience, loan amount keeps increasing and defaulters are also high in all experience level. Ratio between defaulters and fully paid is not much. If lender provides loan amount around 1K-2K less across the experience level, defaulters will be less.
 - ✔ Clients in address state “WY” with loan amount more than 15000



Heat Map

Loan Amount, Funded value investor, annual income, Installments, interest rate, revolving utilization, and term are correlated



Recommendations

- Higher the loan amount, rate of interest and installments in all income range will leads to high probability of defaults
- More the DTI, the applicants are riskier
- All the loan applications must be verified
- Lower grade applicants has more probability of default. Especially, B and C grade needs to be considered for more verifications
- Home Improvement loan, Small Business loans and Debit Consolidation loan has more probability of defaults. Hence, recommend to reduce focus on these purpose loans or investigate more before approving.
- Reduce the number of terms
- Recommend to monitor the client's revolving utilization rate
- Investigate more before approving the loan for applicants having more derogatory public records
- Reduce the loan amount or number of applications approved for the applicants with home ownership type Rent and Mortgage

Conclusion

With the available Lending Club data shared, we had performed various steps of EDA, found the variables impacting the loan status and provided supporting analysis reports for the same. We also provided the recommendations to finance company to overcome the loss from loan defaults and credit loss.

Thank You