

LENDING CLUB CASE STUDY

By:

Ashish Chaudhary

Uma Hyndavi Simhadri



PROBLEM STATEMENT AND DATASET

- A consumer finance Company lending various types of loans to urban customers.
- When company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile.
- Given past loan data from 2007 to 2011, comprises of Current(still paying loan), Charged off(Defaulters) and fully paid data
- The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.
- Aim is to Implement EDA on this real world problem

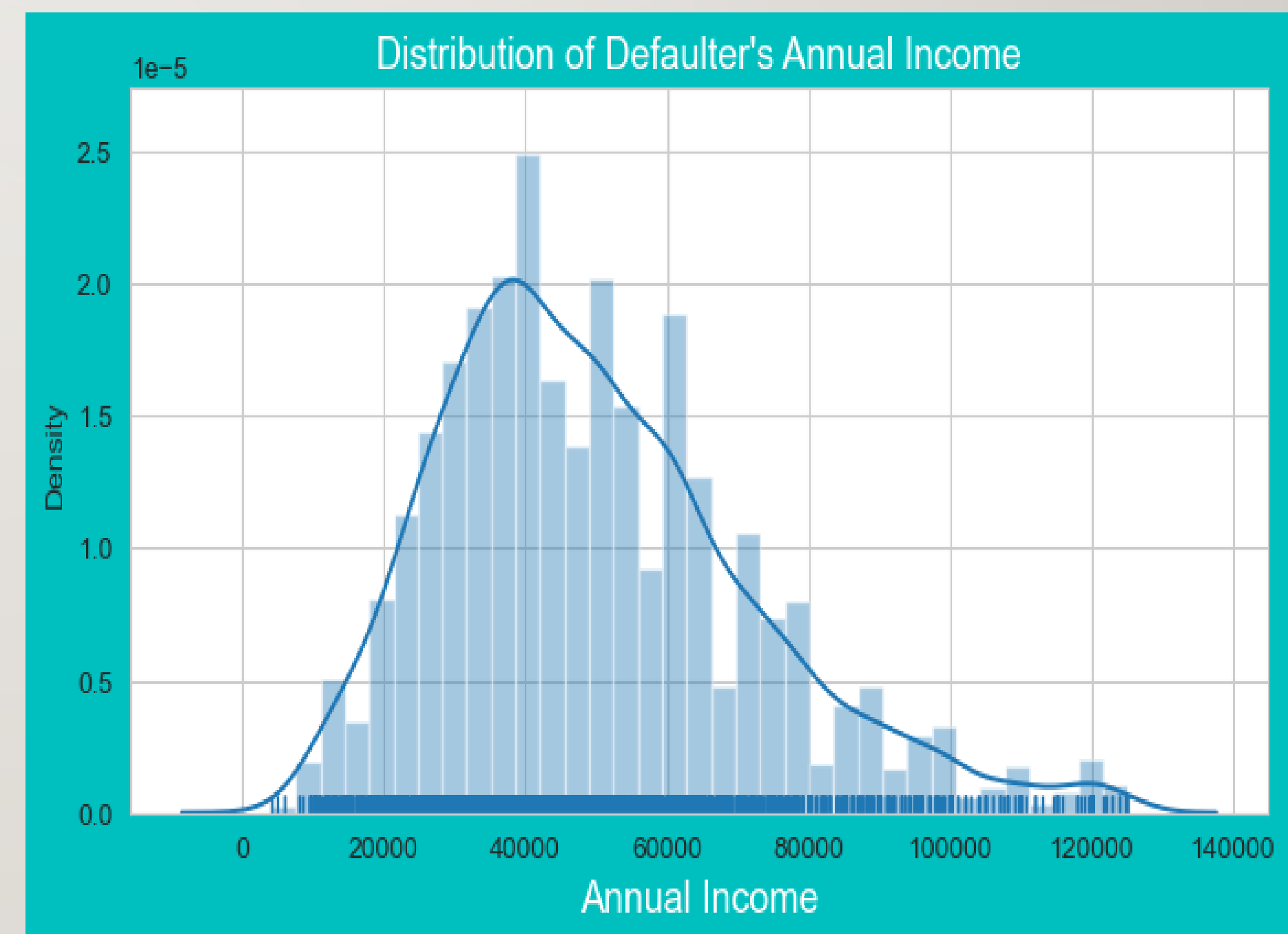
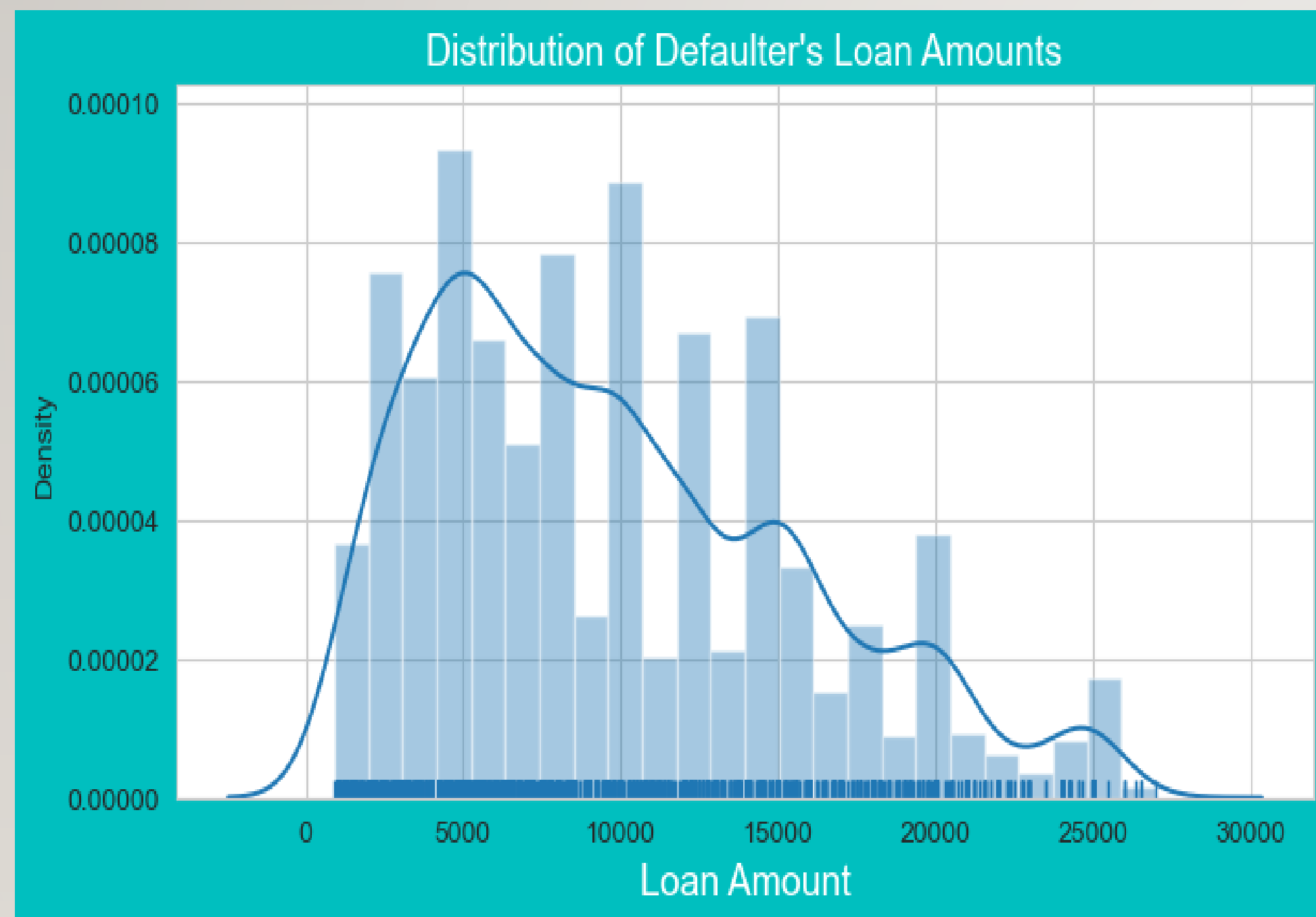


DATA CLEANING AND ANALYSIS

- Loading the Data set using pandas library
- Dropping Null columns
- Converting data type of column according to analysis
- Split columns accordingly
- Treating missing values
- Removing Outliers from numerical columns.
- Performing Univariate analysis on each column
- Bivariate Analysis
- Multivariate Analysis

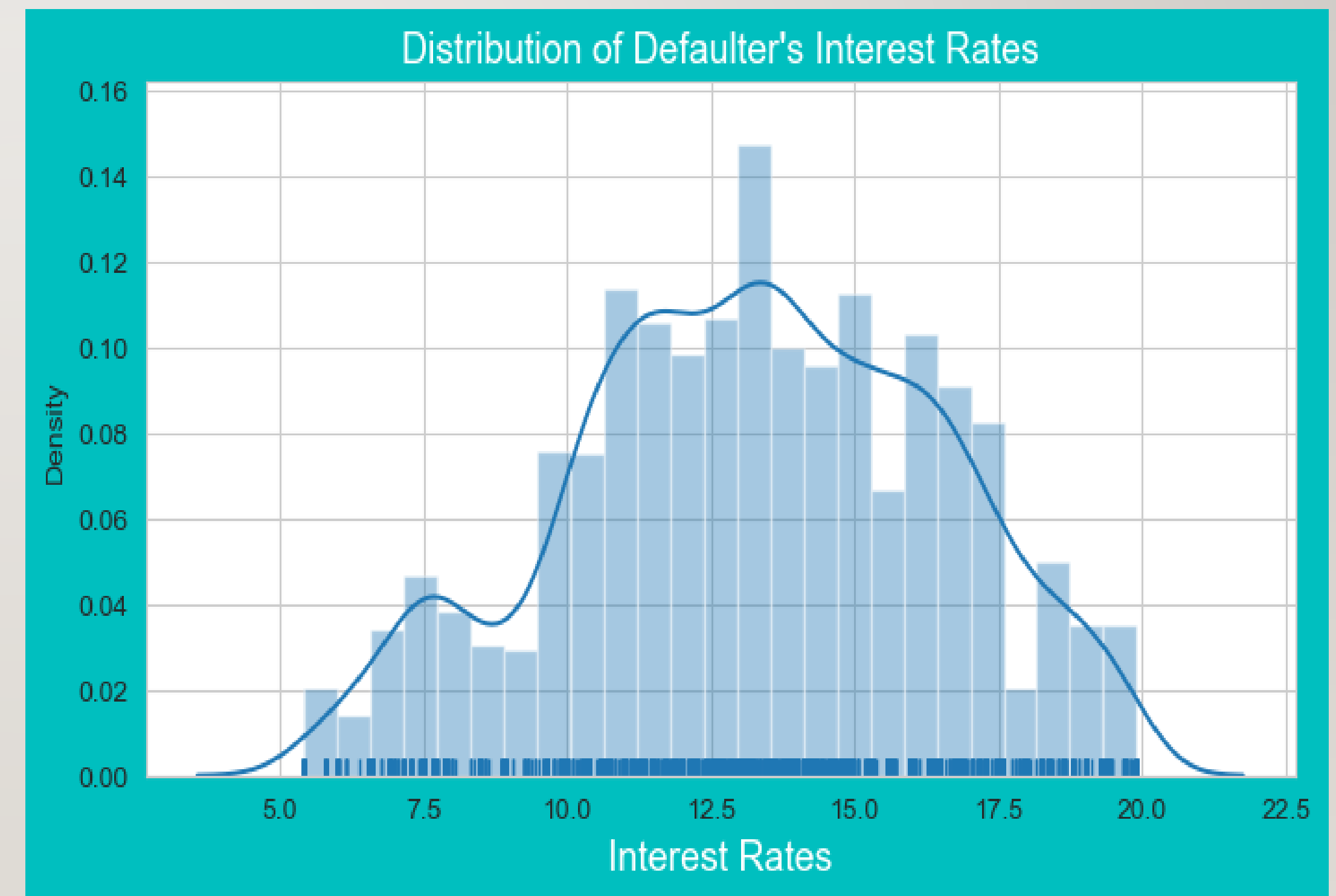
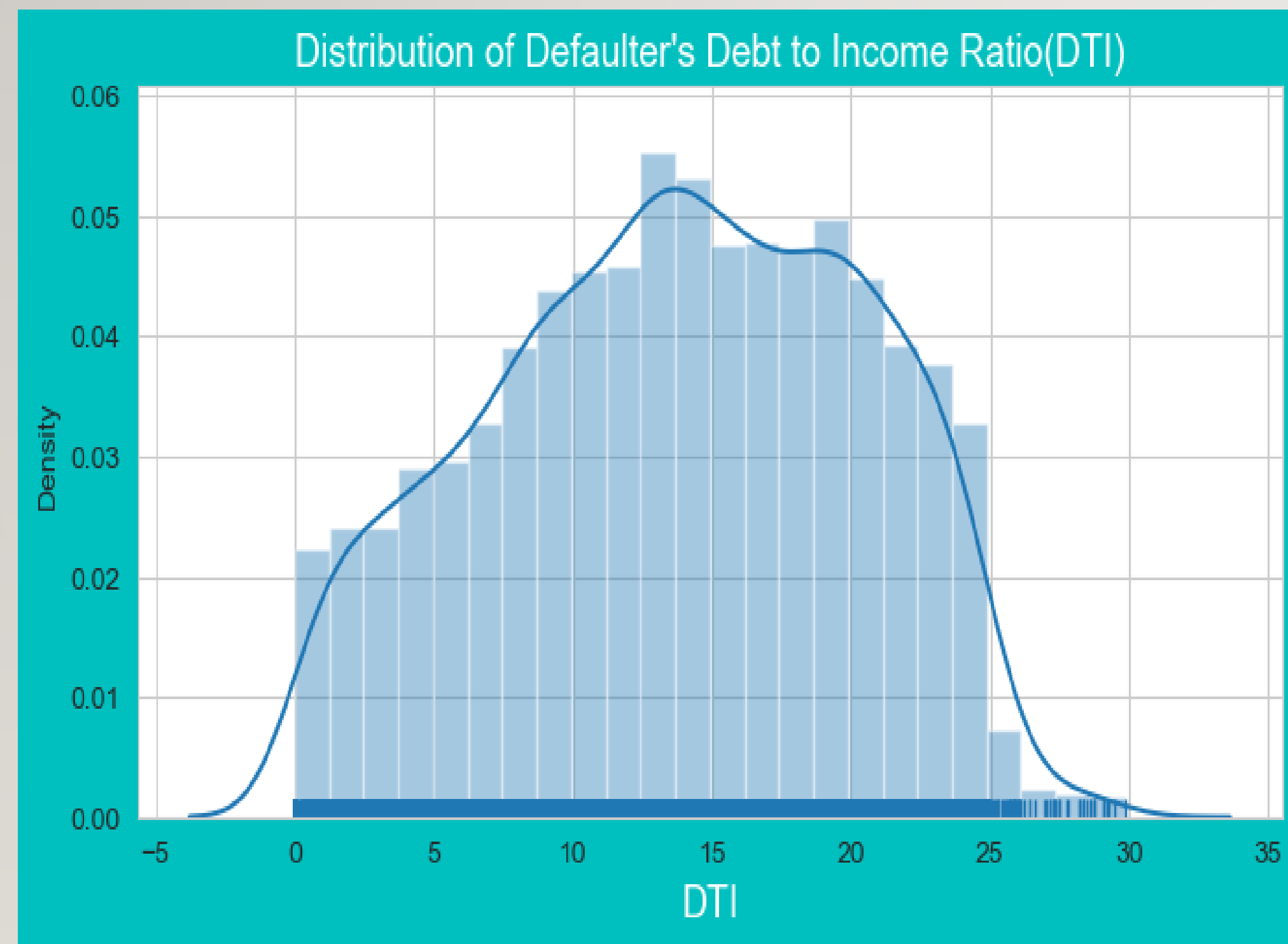
UNIVARIATE ANALYSIS

- ***Loan amount and Annual Income:*** Considering only charged off data, Borrowers who applied for Less loan amount with less annual income. Likely to become defaulters



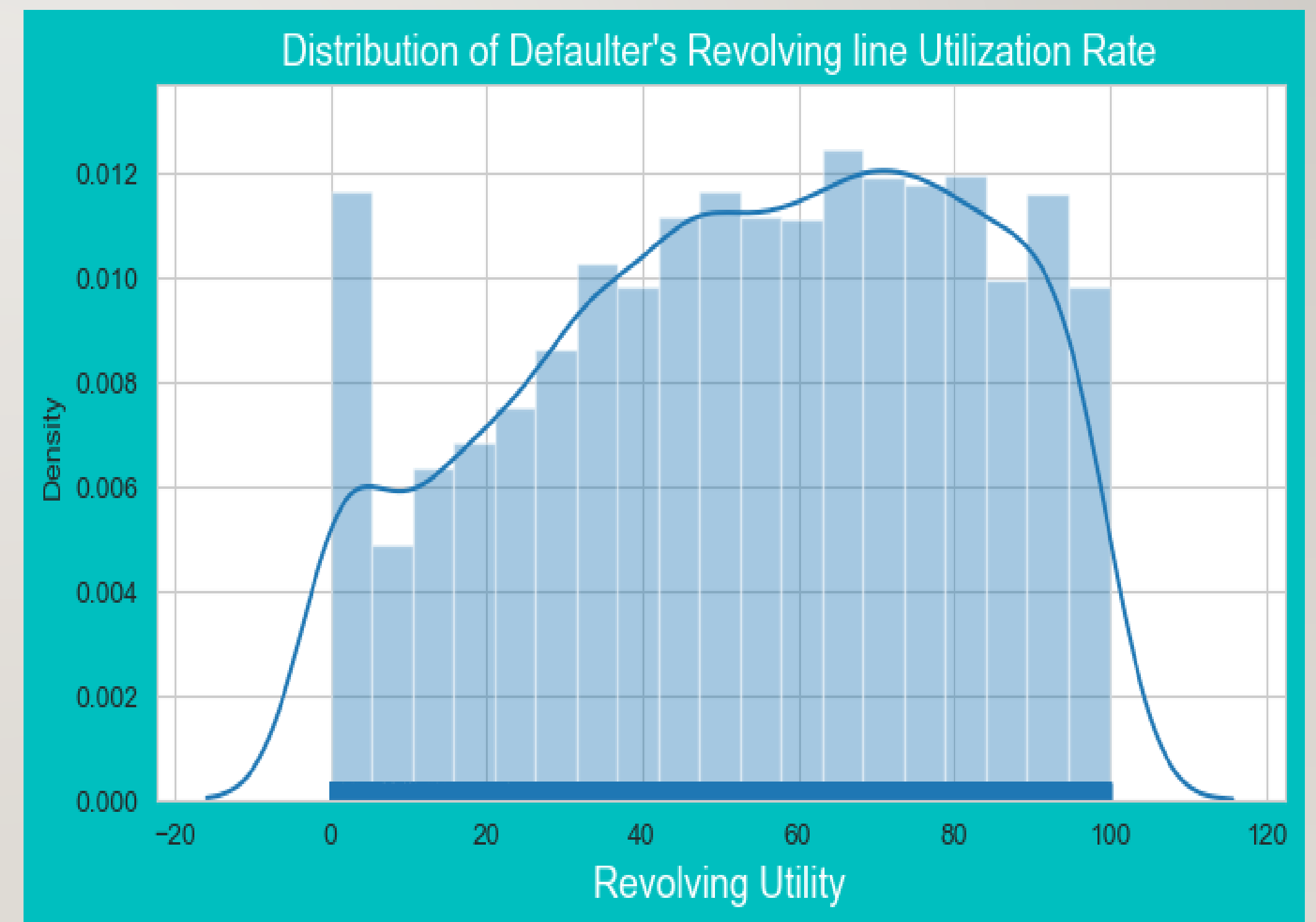
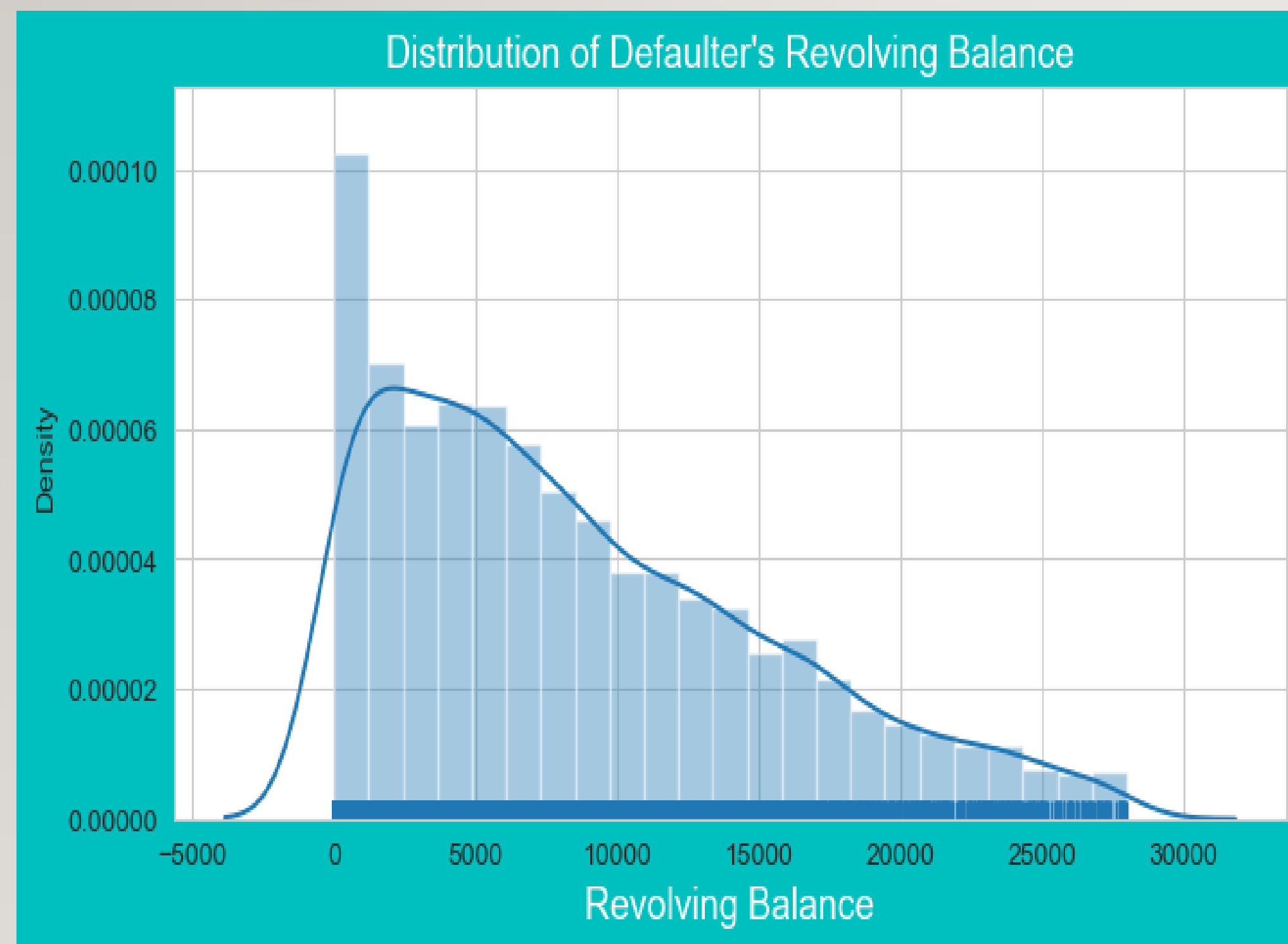
UNIVARIATE ANALYSIS

- ***DTI ratio and Interest Rates:*** Considering only charged off data ,Borrowers with High DTI ratio and high interest rates are likely to become defaulted.



UNIVARIATE ANALYSIS

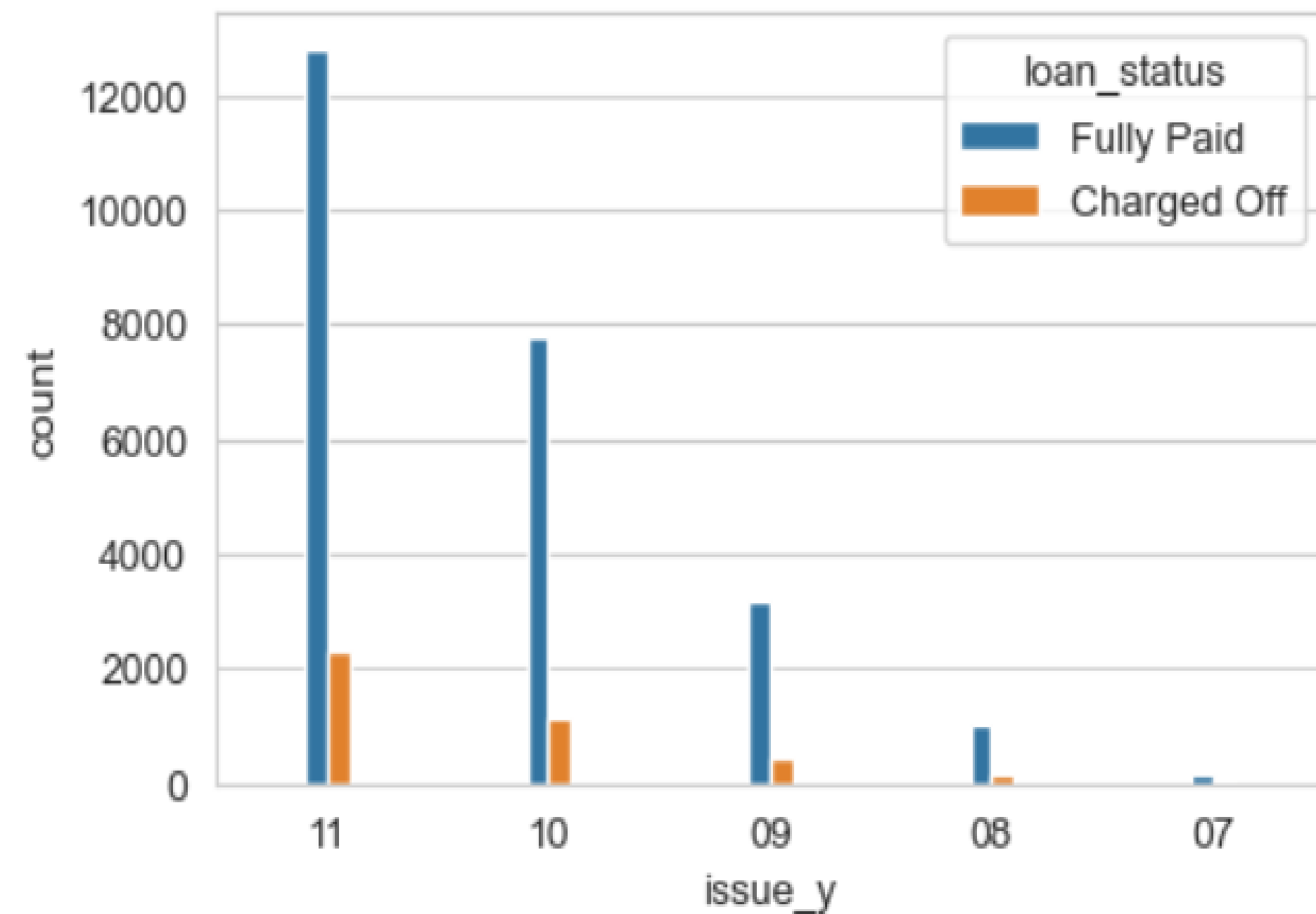
- **Revolving balance and Revolving Utility:** Considering only charged off data , Majority of Defaulters are from, Borrowers whose Revolving utility is high with Less Revolving balance.



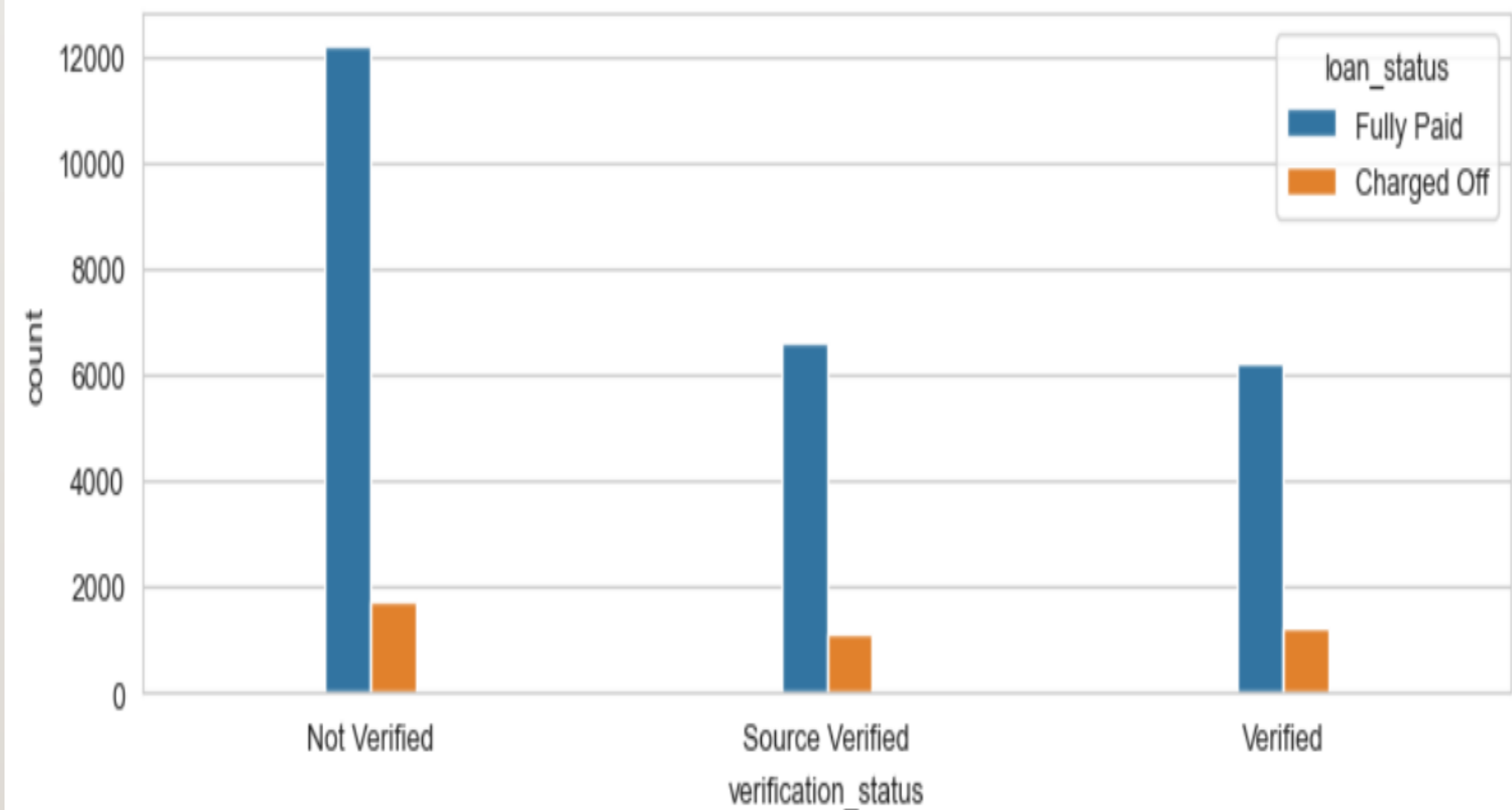
UNIVARIATE ANALYSIS

- ***Loan Issued Year and verification status:*** Considering only charged off data, Defaulter's rate keep on increasing from 2008 to 2011 and Year ending months having more defaulters. Majority of defaulters falls under not verified status.

CountPlot of issued year



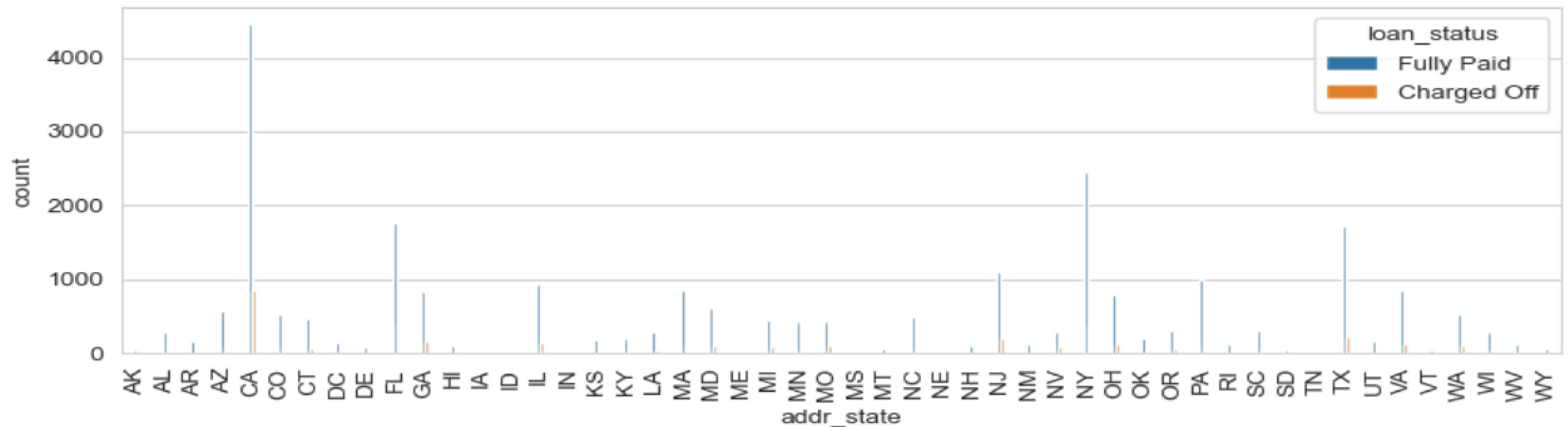
CountPlot of verification_status



UNIVARIATE ANALYSIS

- **Address state:** Considering only charged off data , Majority of Defaulters are from, California . Georgia and Texas having moderate number of defaulters. so, mostly Defaulters are from metropolitan cities.

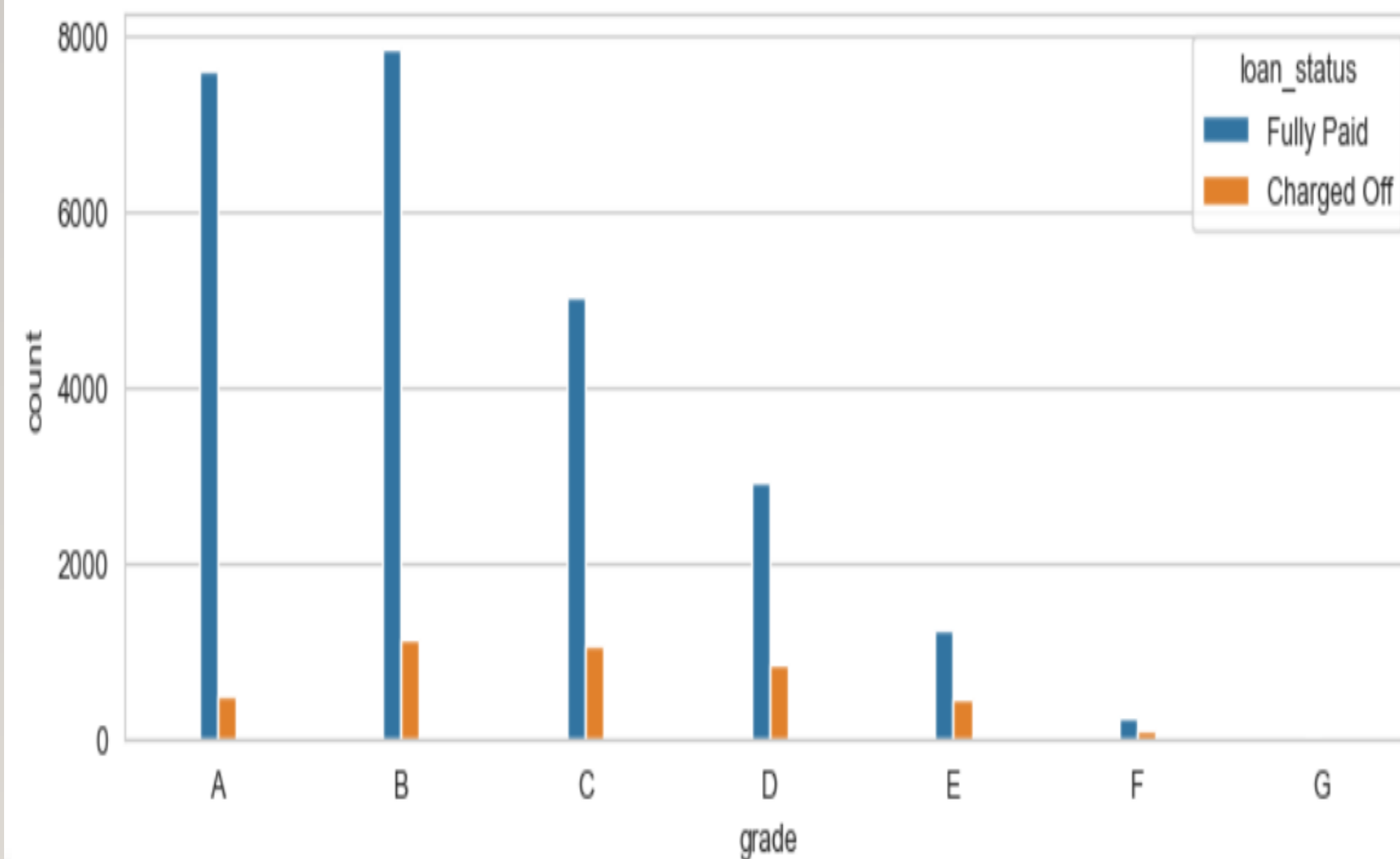
CountPlot of addr_state



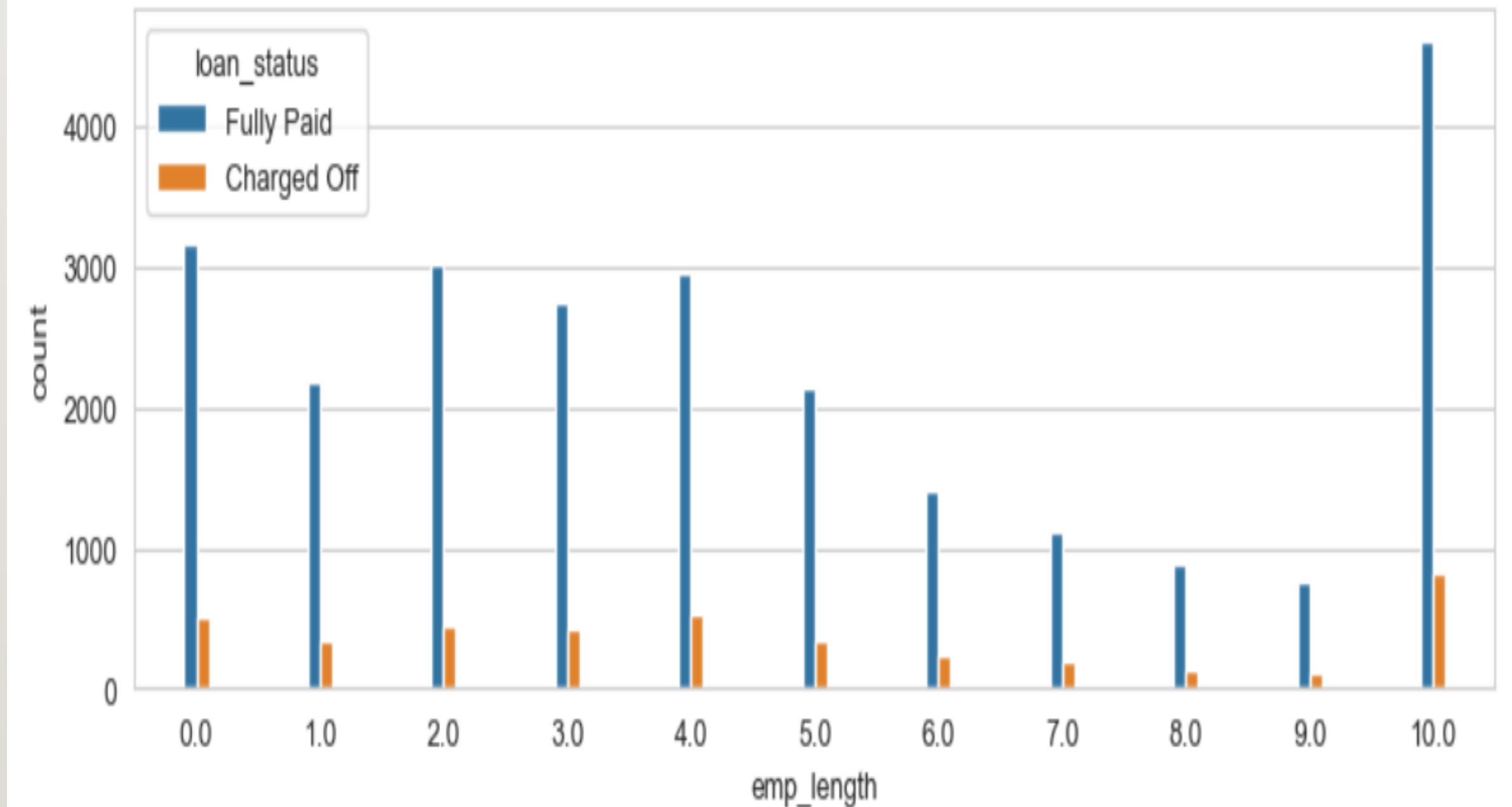
UNIVARIATE ANALYSIS

- **Grades and Employment length:** Considering only charged off data , Majority of Defaulters are from, Grade B and C having employment length 10years or above.

CountPlot of grade



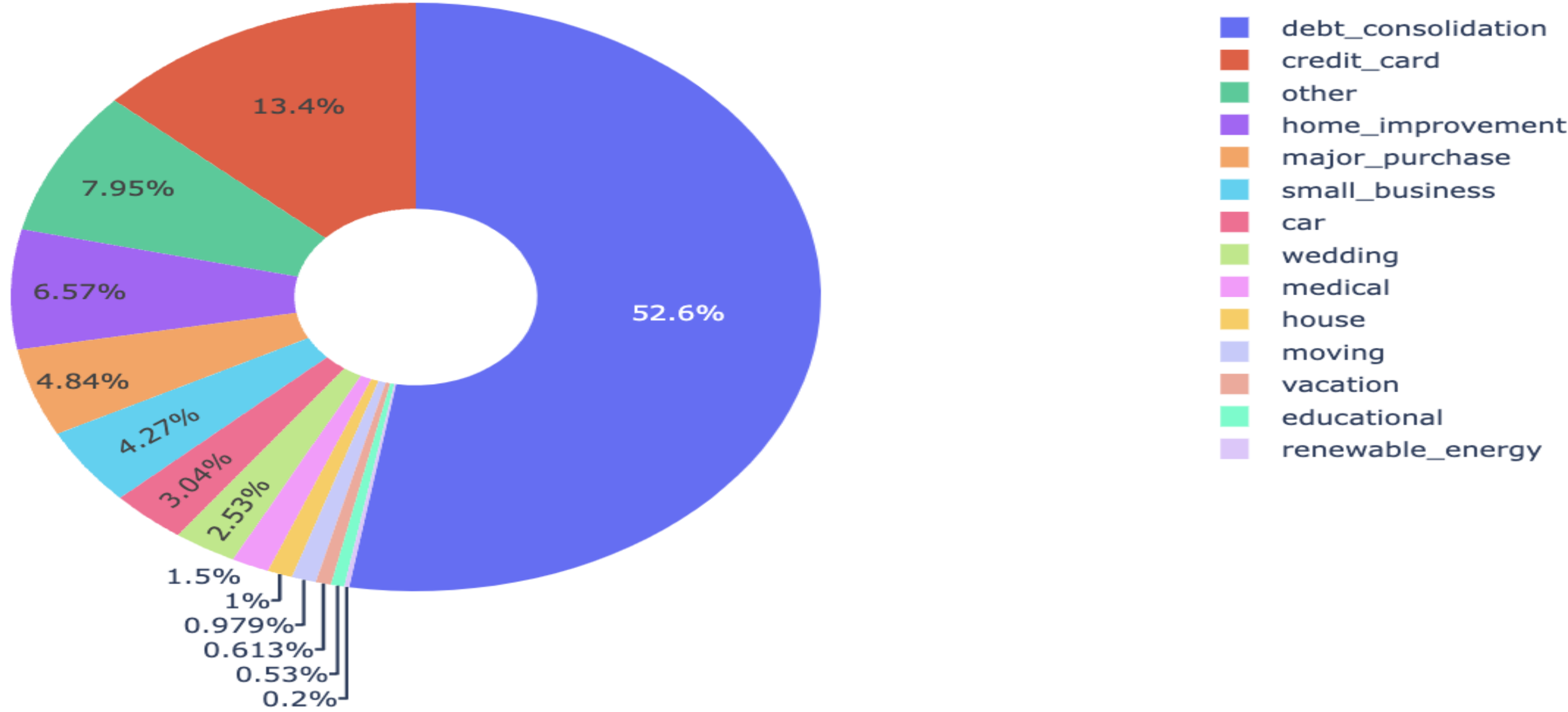
CountPlot of emp_length



UNIVARIATE ANALYSIS

➤ **Loan purpose :** *C* Considering only charged off data , Majority of Borrowers taken loan for purpose as debt consolidation.

Loan purpose Aggregate Information



UNIVARIATE OBSERVATIONS:

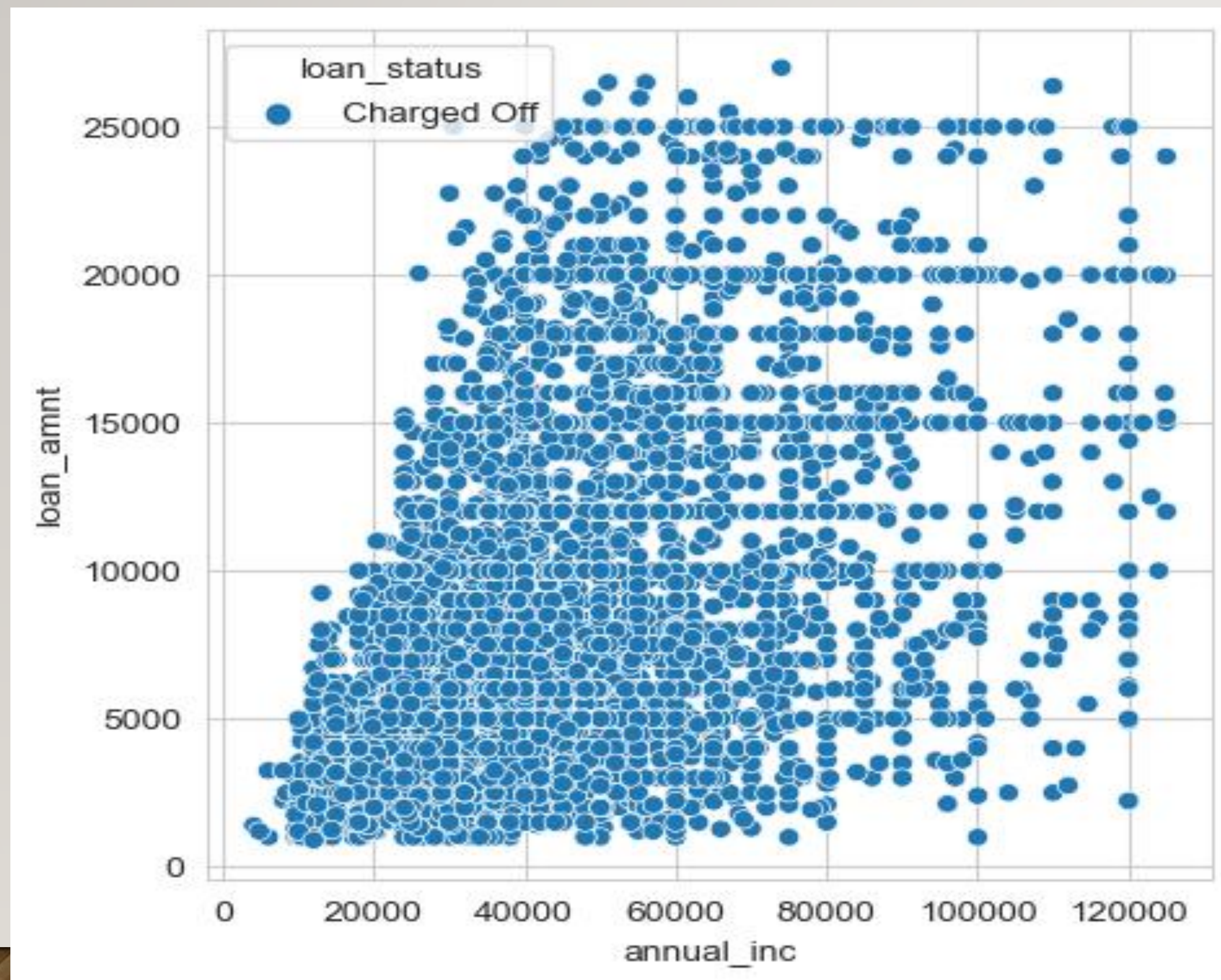
- Grade B has Higher Defaulters
- Sub grade B5 has higher Defaulters
- Rented Home borrowers are likely to become Defaulters
- No Inquiries in last 6 months likely to become Defaulters.
- Borrowers with No Source/Annual Income verification likely to become Defaulters.
- Higher the Revolving utility with loan amount under 15k has higher number of defaulters
- Year End has Higher Defaulters.
- Lesser Annual Income/Loan amount, Higher chance of Defaulters
- Higher Interest rates, Higher chance of Defaulters.
- Higher dti, Higher chance of Defaulter.
- Lesser Revolving Balance, Higher chance of Defaulters.
- Higher Revolving line utilization rate, Higher Chance of Defaulter.
- Year 2011 has High number of Defaulters.



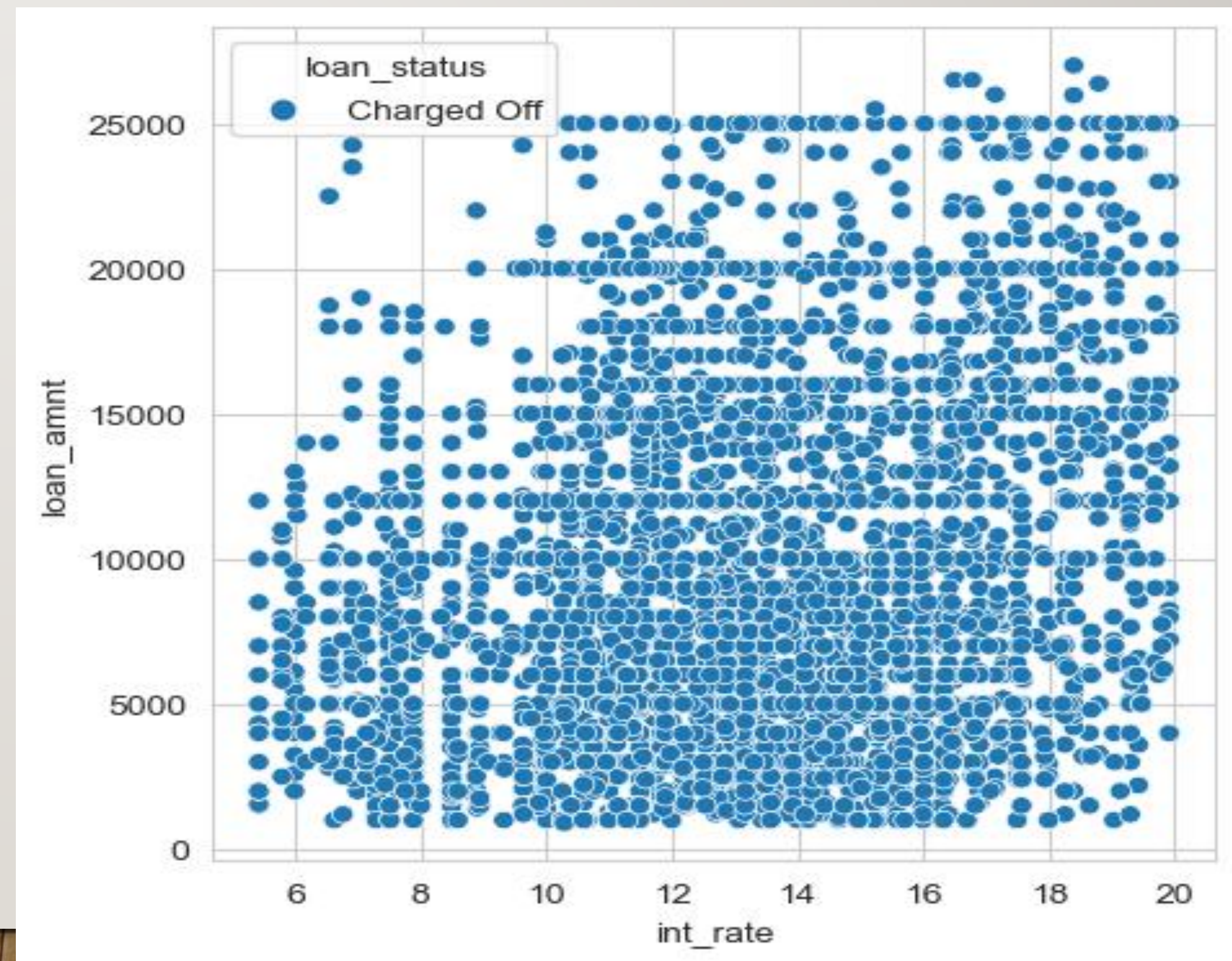
BIVARIATE ANALYSIS:

- *Loan amount* under 17k with *annual income* under 72k comprises majority of defaulters.
- *Loan amount* under 15k with *interest rates* between 10 to 17 comprises majority of defaulters.

Loan amount vs Annual Income



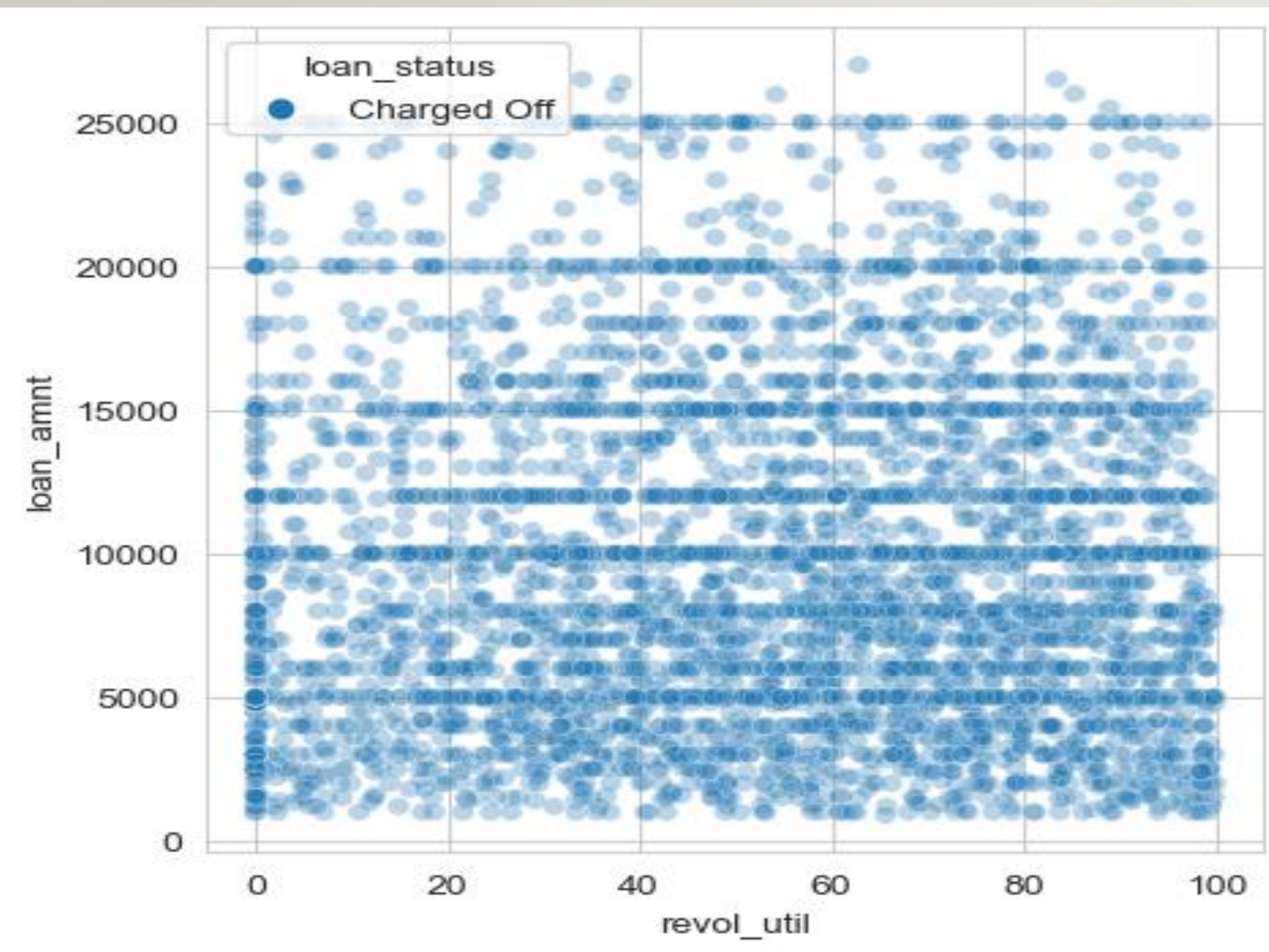
Loan amount vs Interest rate



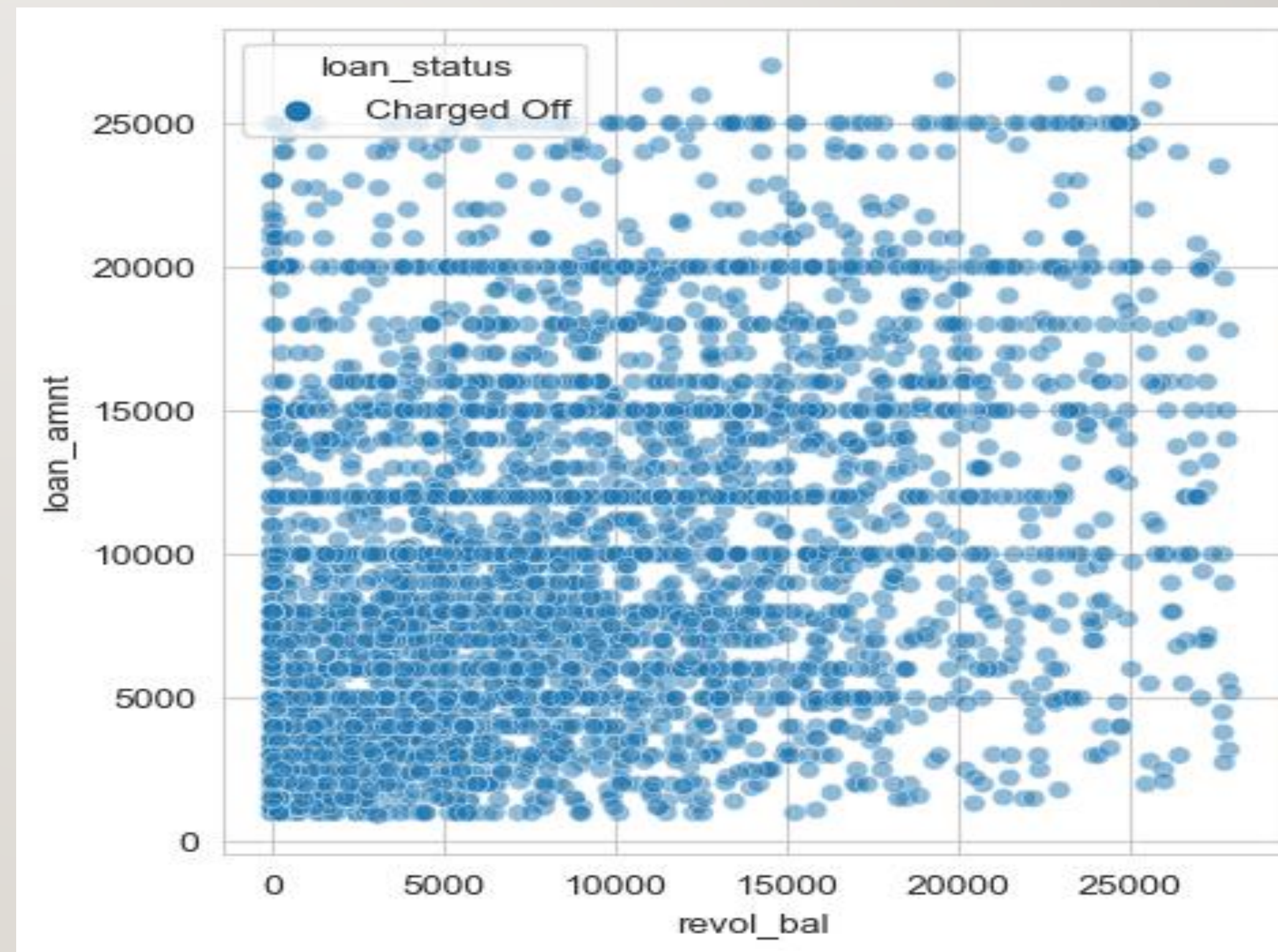
BIVARIATE ANALYSIS

- Loan amount under 11k with High *revolving utility*, comprises majority of defaulters.
- Loan amount under 16k with less *revolving balance*, comprises majority of defaulters.

Loan amount vs Revolving utility



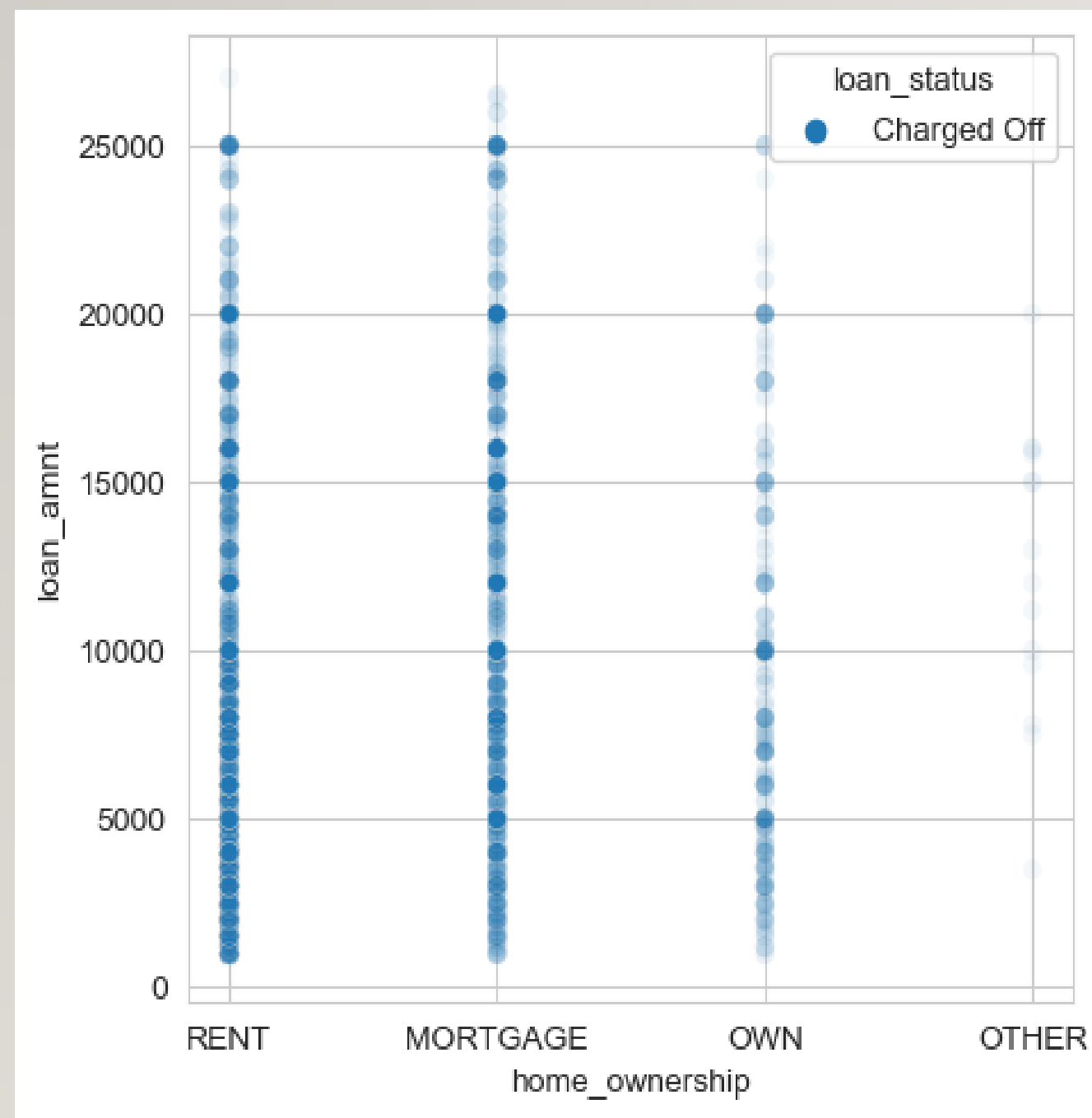
Loan amount vs Revolving Balance



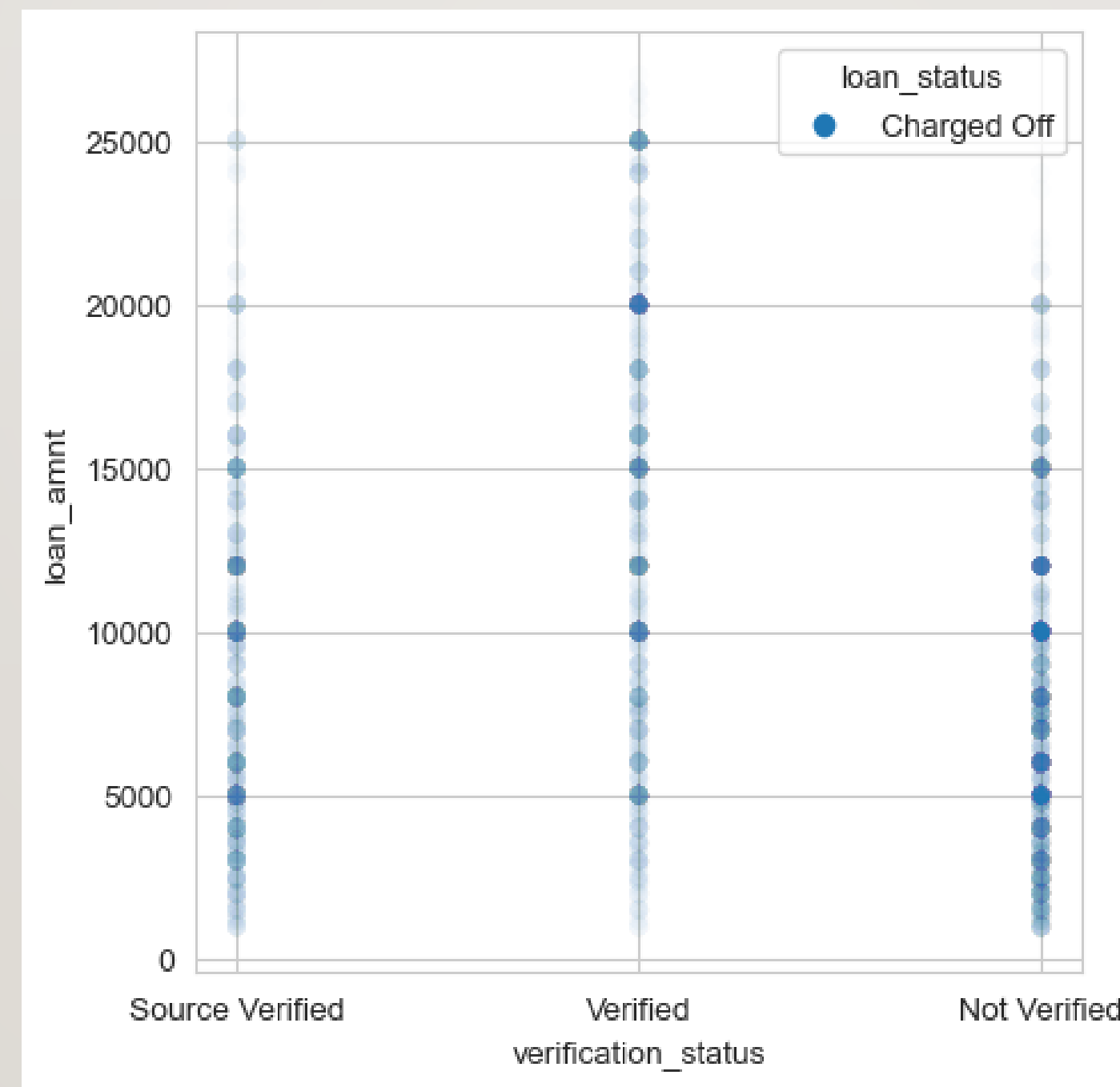
BIVARIATE ANALYSIS

- Borrowers whose Loan amount under 16k having rented home with 10 years of employment length are more likely to become defaulters.
- Borrowers whose Loan amount under 11k with no source or income verification are more likely to become defaulters.

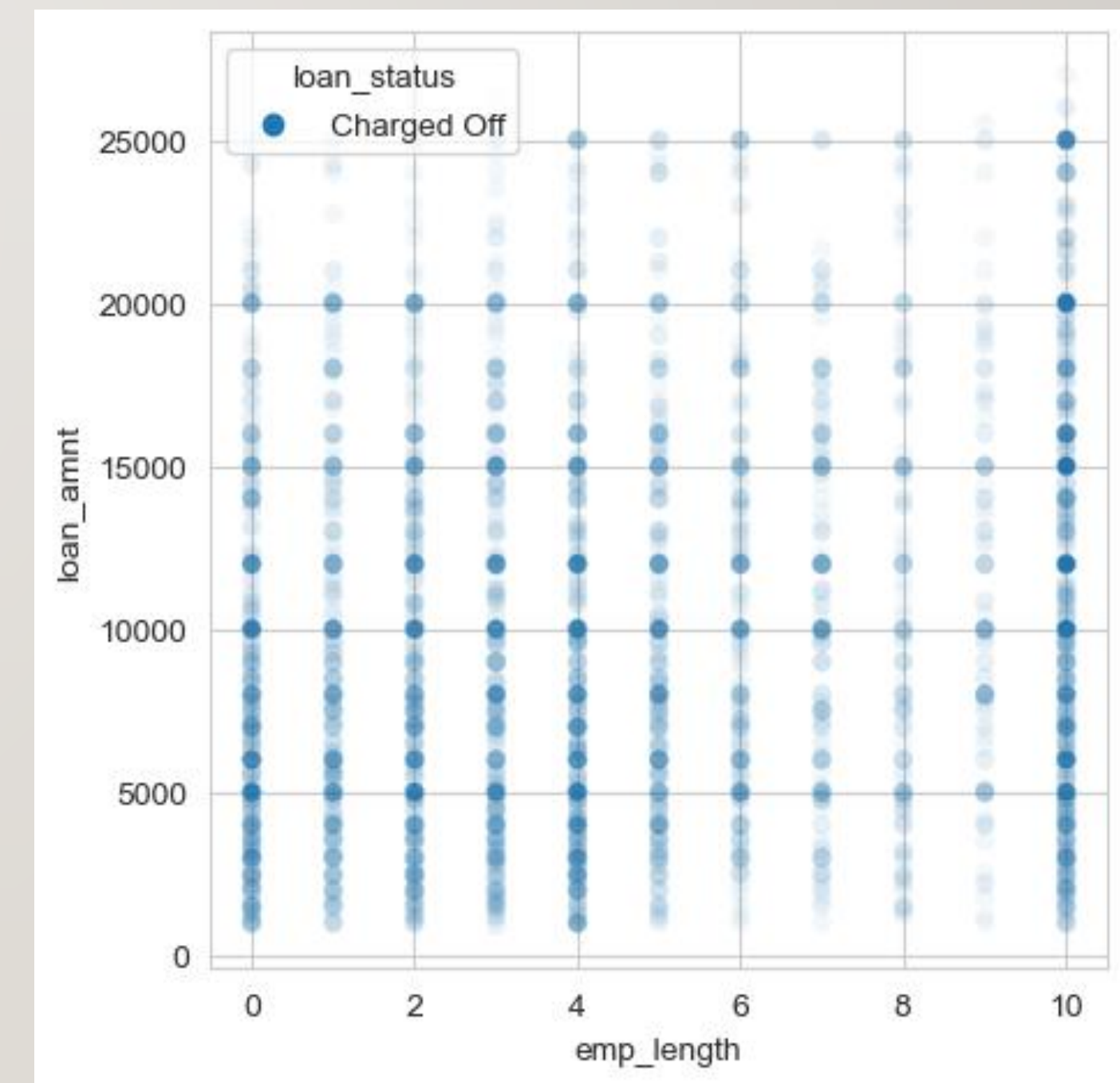
Loan amount vs Home Ownership



Loan amount vs Verification Status



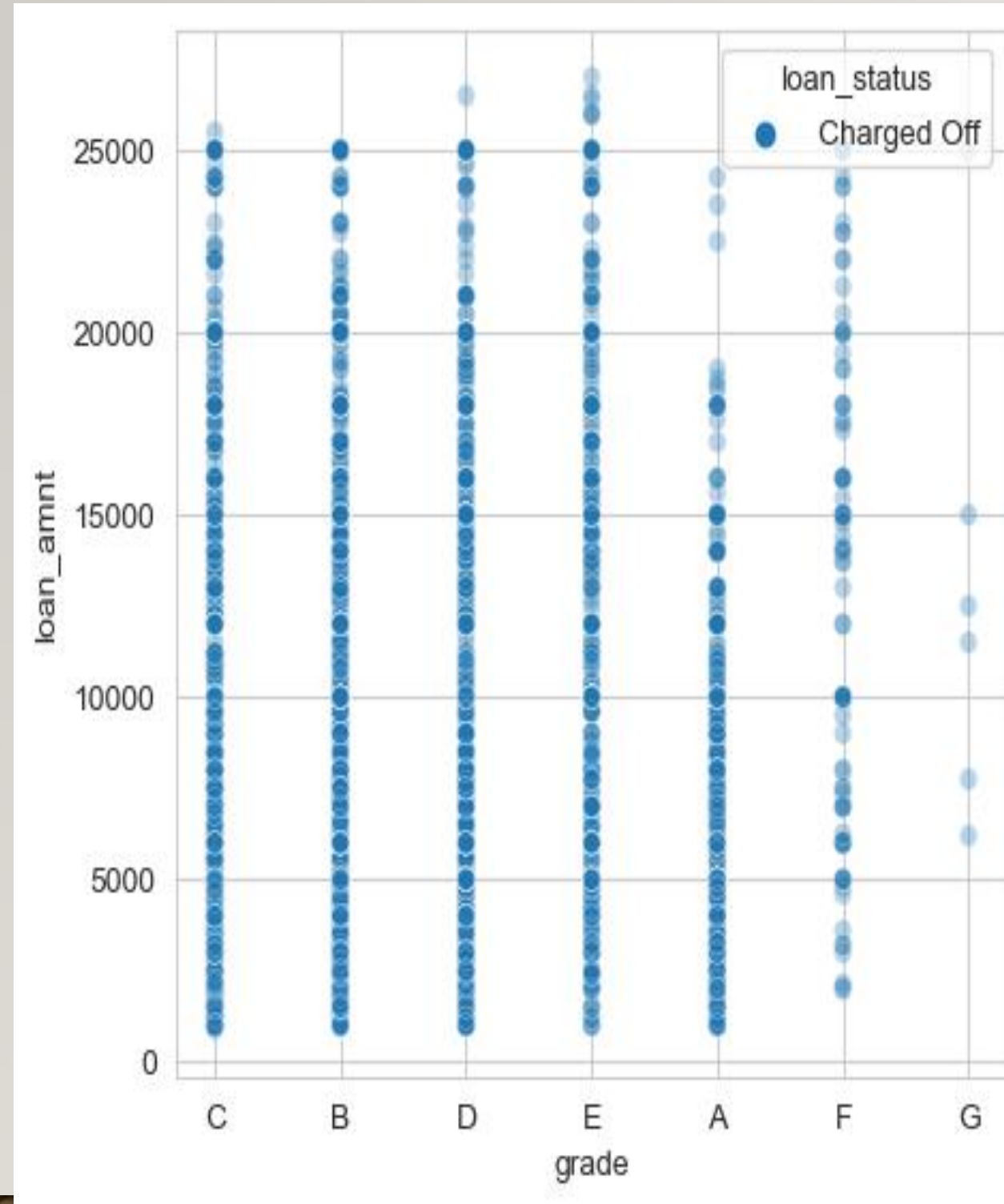
Loan amount vs employment length



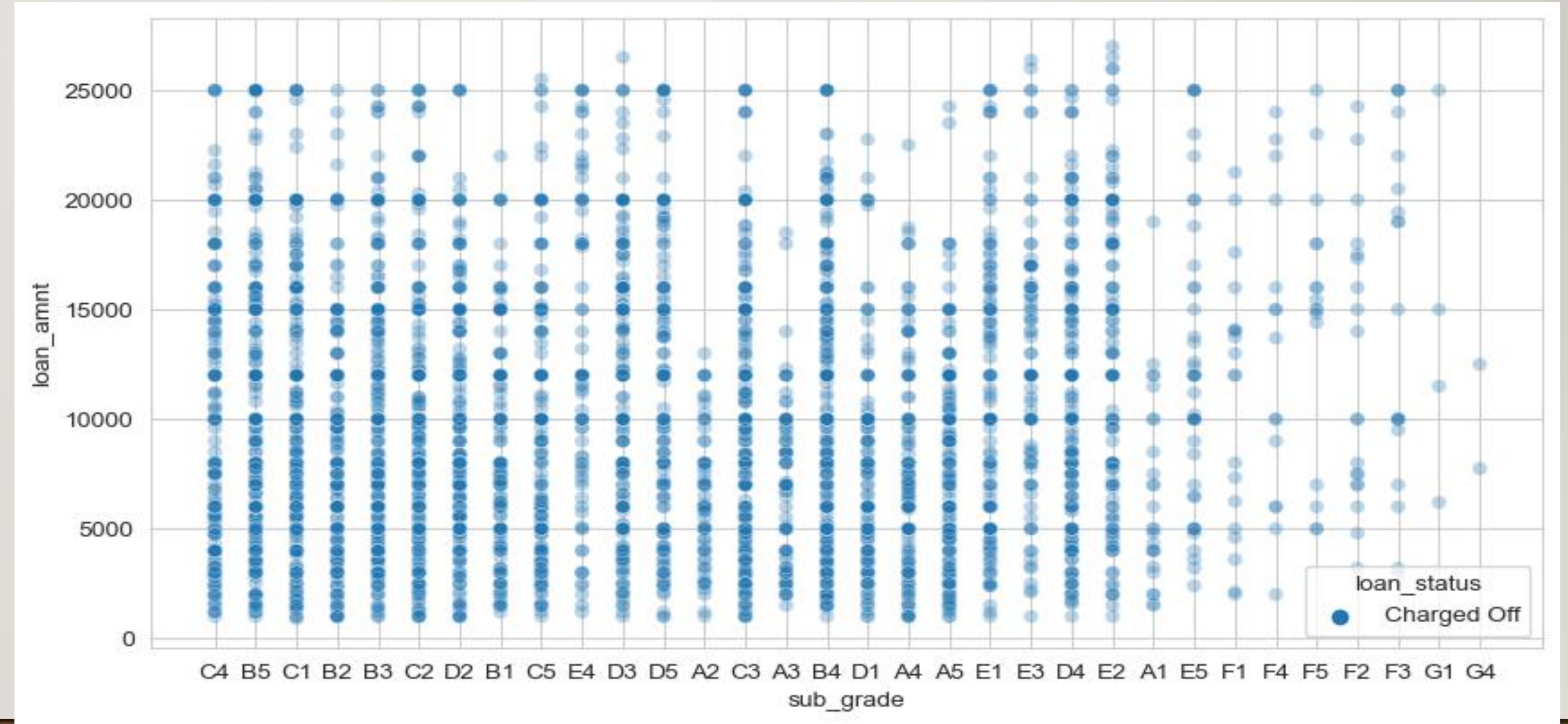
BIVARIATE ANALYSIS

- Loan amount under 20k with LC assigned grades as B,C and D, comprises majority of defaulters.
- Loan amount under 16k with LC assigned sub grades as B3,B5,C1,B4, comprises majority of defaulters.

Loan amount vs Grade



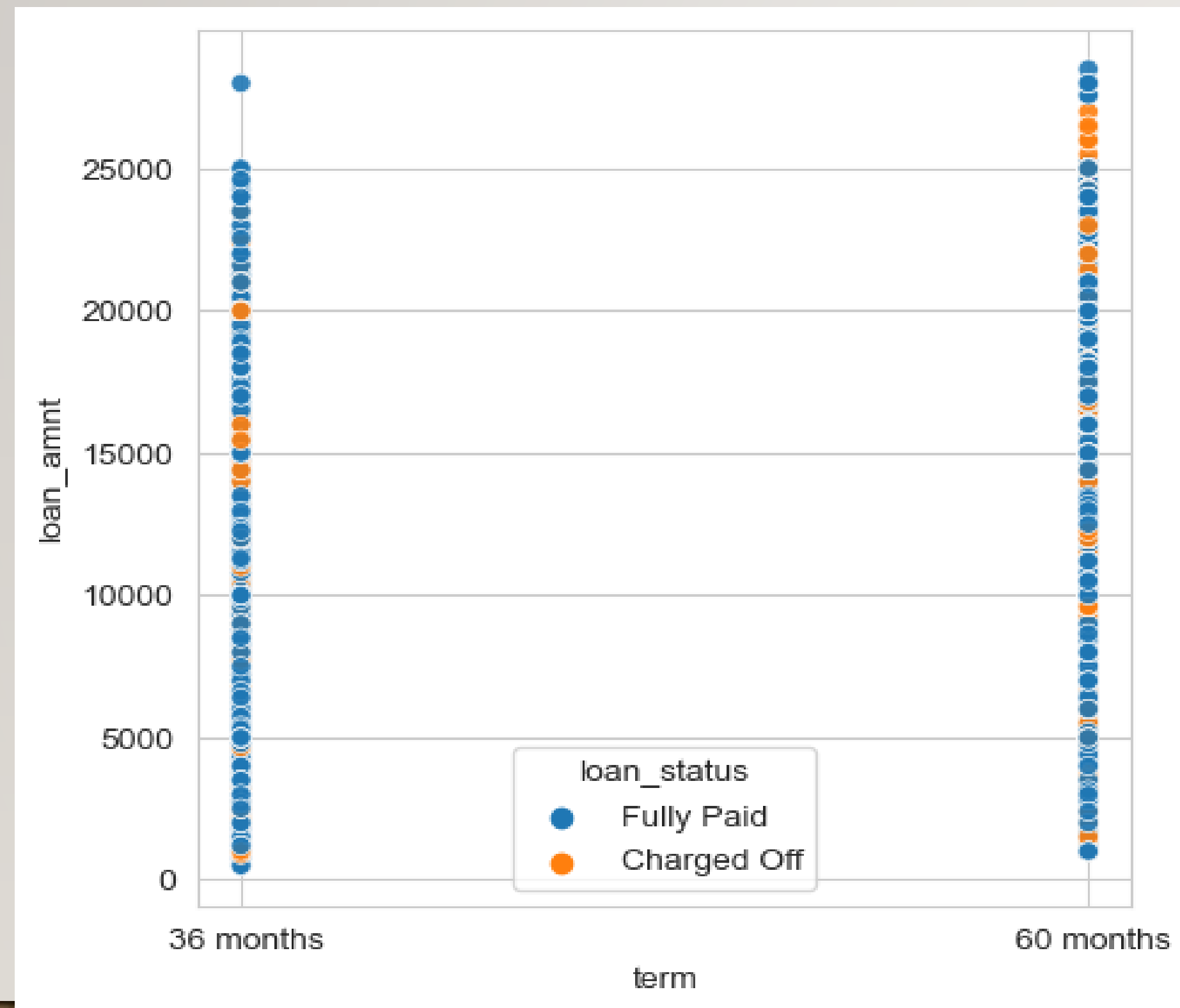
Loan amount vs sub-grade



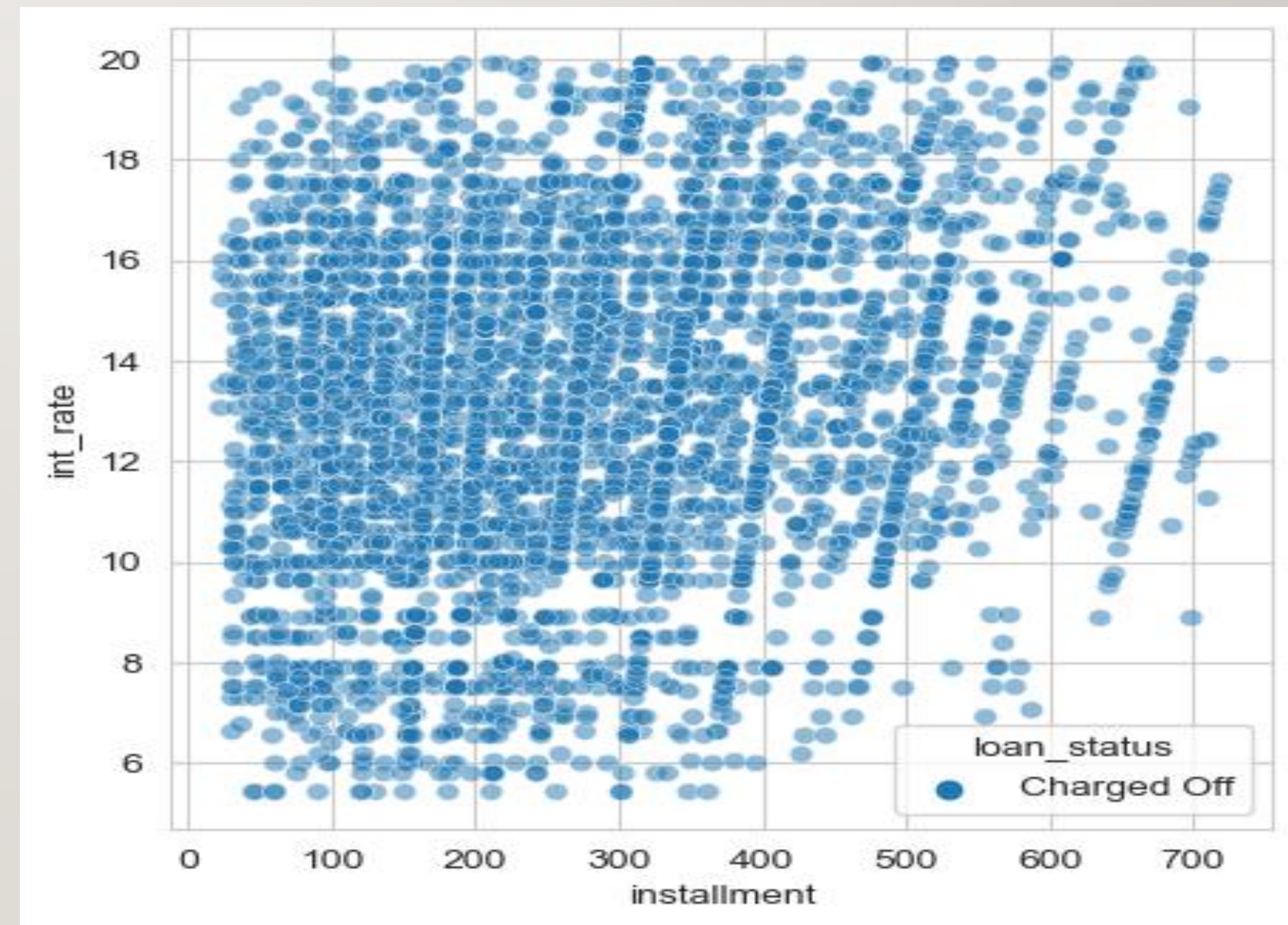
BIVARIATE ANALYSIS

- Regardless of loan amount, Defaulters are present in both term plans.
- loan installment from 156 to 296 and interest rates from 8 to 17, Comprises majority of Defaulters

Loan amount vs term



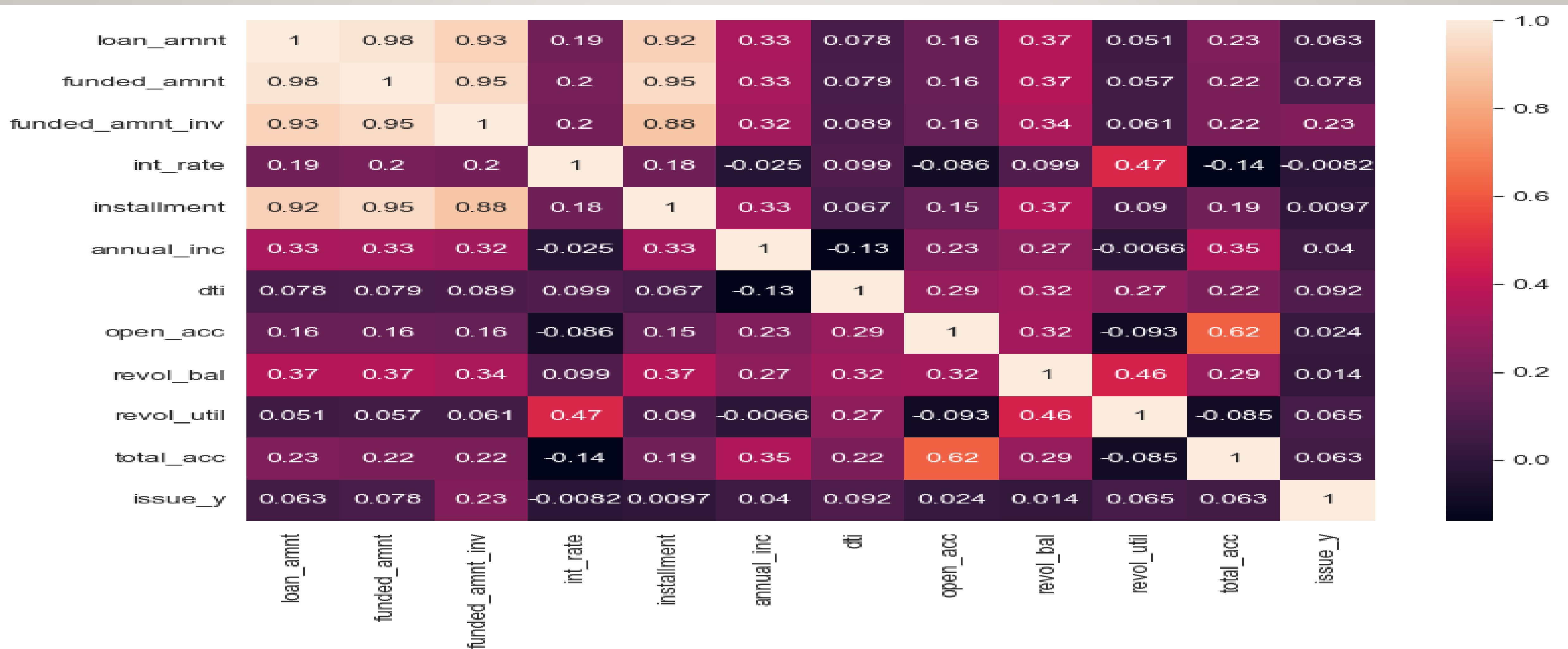
Interest rate vs installment



BIVARIATE OBSERVATIONS:

- Grades B,C & D Has high number of Defaulters
- Sub grades B3,B5,C1,B4,C2,A5 has higher number of Defaulters
- No source/Income verification has high Defaulters
- Employment length within 10/4/0yrs has high number of defaulters.
- Lower Revolving balance with loan amount under 10k has higher Defaulters
- Higher the Revolving utility with loan amount under 15k has higher number of defaulters
- Rented house Borrowers under 17.3k loan amount are likely to become of Defaulters.
- Borrower's under loan amount under 17k, Interest rates from 8 to 13.49 are likely to become defaulters
- Borrower's Loan purpose as Debt consolidation with loan amount under 23k are likely to become defaulters
- Borrowers from California State with loan amount under 17k are likely to become defaulters
- Regardless of loan amount Defaulters are high in both 60months and 30 months tenure
- More Defaulter's DTI ranges from 10 to 25 and loan amount between 20k to 60k
- Borrower's Annual income between 28k to 52k with B or C grades are likely to become defaulters
- Majority of Defaulters comprises under 17.3k loan amount and annual income from 28.2k to 52.4k.
- Borrower's Annual income from 28k to 53k with 10yrs of employment experience are likely to become defaulters
- Borrowers from California and Florida states with annual income under 80k are likely to become defaulters
- Borrower's with Installment from 156 to 296 with interest rates from 8 to 17 are likely to become defaulters.

MULTIVARIATE ANALYSIS



MULTIVARIATE OBSERVATIONS:

Highly Correlated:

1. Loan amount is highly correlated with funded amount, funded amount by investors, installment.
2. Funded amount and funded amount by investors are highly correlated with installment and loan amount.

Less Correlated:

1. Loan amount is less correlated with interest rates, annual income, dti, open accounts, total accounts, revolving balance and revolving utility.
2. Interest rate is less correlated with dti ratio, revolving balance, revolving utility
3. Installment is less correlated with interest rates, dti, open accounts, total accounts, revolving balance and revolving utility.
4. Annual income is less correlated with other numerical columns.
5. Revolving balance, open accounts, total accounts and dti are less correlated with others.

Negatively correlated:

1. Interest rates negatively correlated with annual income, open accounts and total accounts.
2. Annual income and dti are negatively correlated.
3. Open accounts and total accounts are negatively correlated with revolving utility.

STRONG INDICATORS OF DEFAULT:

- Grade as B
- Sub grade as B3
- Verification status as Not verified
- Employment length of 10years
- Revolving balance less
- Revolving utility high
- Home ownership as rented house
- Interest Rates from 8 to 17
- Loan purpose as debt consolidation
- Address state from metropolitan cities
- Debt to income ratio from 10 to 25
- Annual income from 28k to 52k
- Loan amount under 15k