# LENDING CLUB CASE STUDY

Contributors -

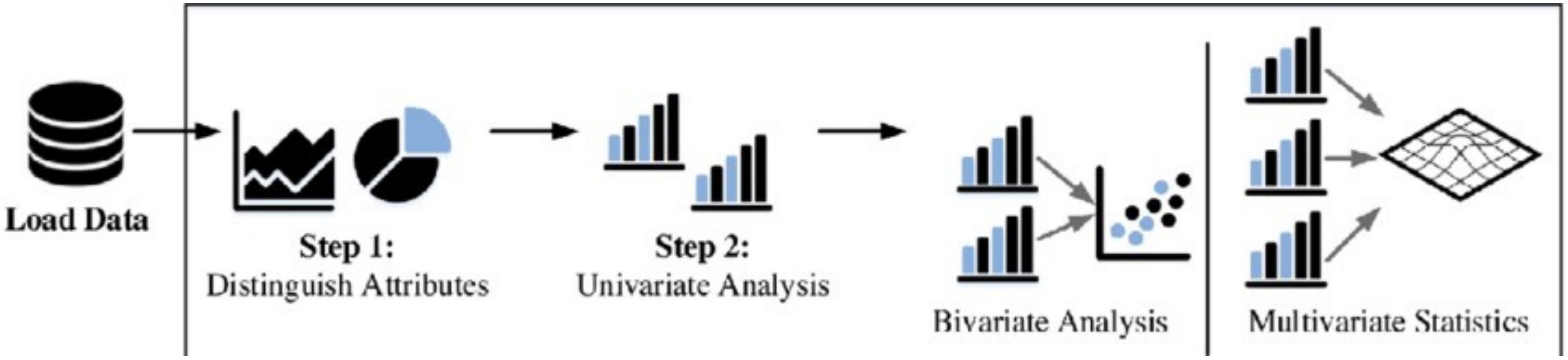
**Aparna Mehrotra**

**Pradeep Kumar Sharma**

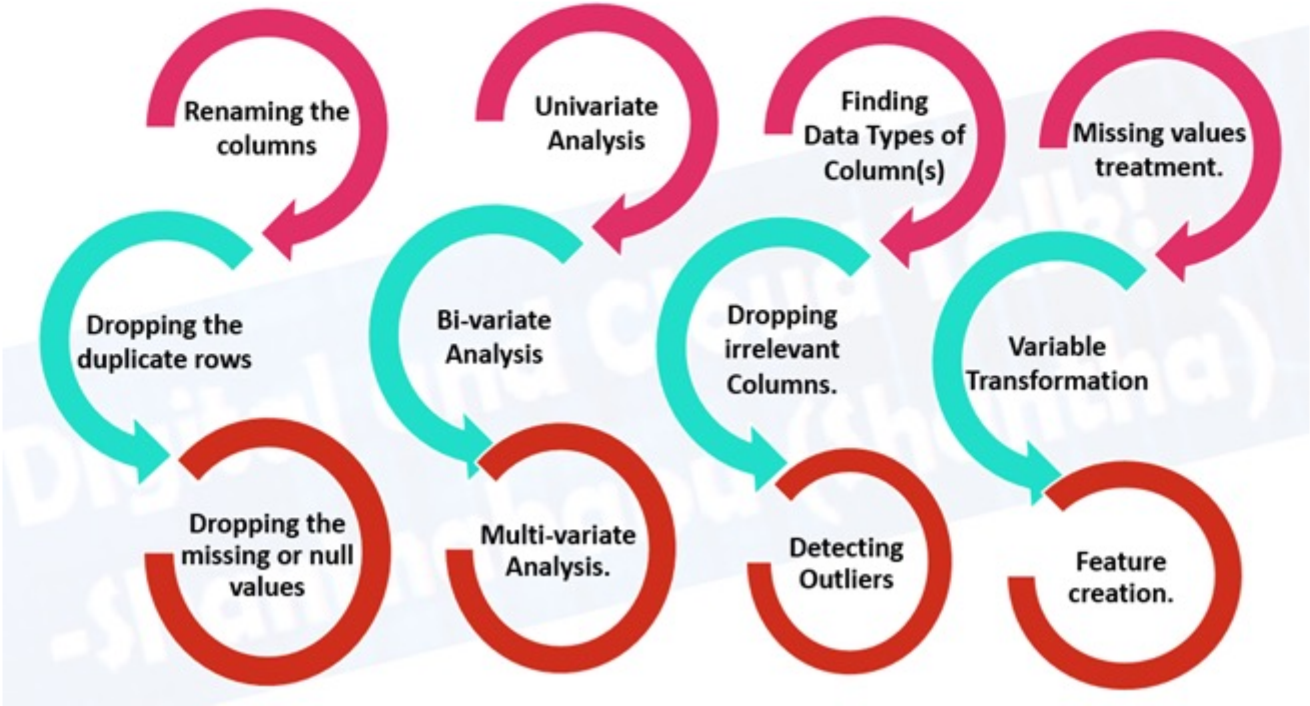
# PROBLEM STATEMENT

- A consumer finance company which specializes in lending various types of loans to urban customers, has to make a decision for loan approval based on the applicant's profile.
- The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.
- To identify the risky loan applicants, thereby cutting down the amount of credit loss considerably

# METHODOLOGY : Exploratory Data Analysis



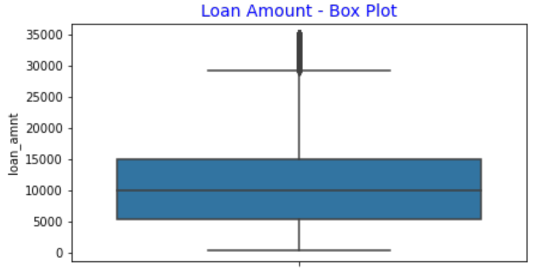
# APPROACH STEPS

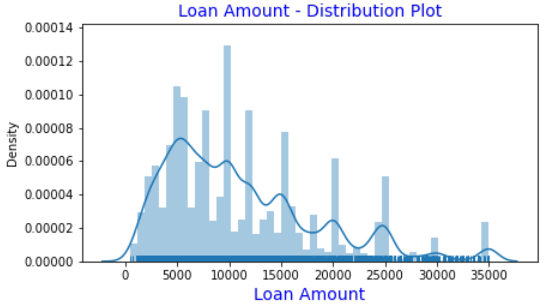


# DATA CLEANING

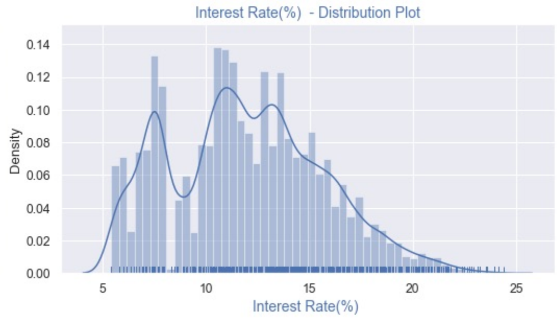
- Fix Missing values
- Fixing rows and columns
- Standardize values
- Fix invalid values
- Filter data
- Derived Columns

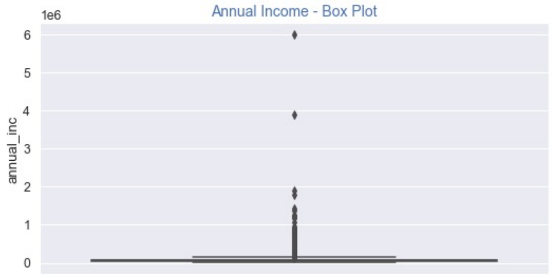
# UNIVARIATE ANALYSIS - Continuous



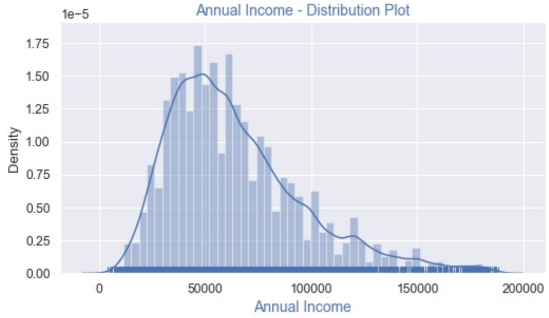







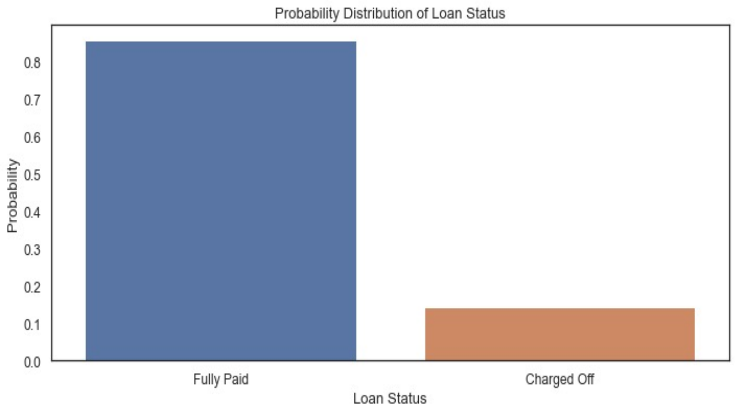


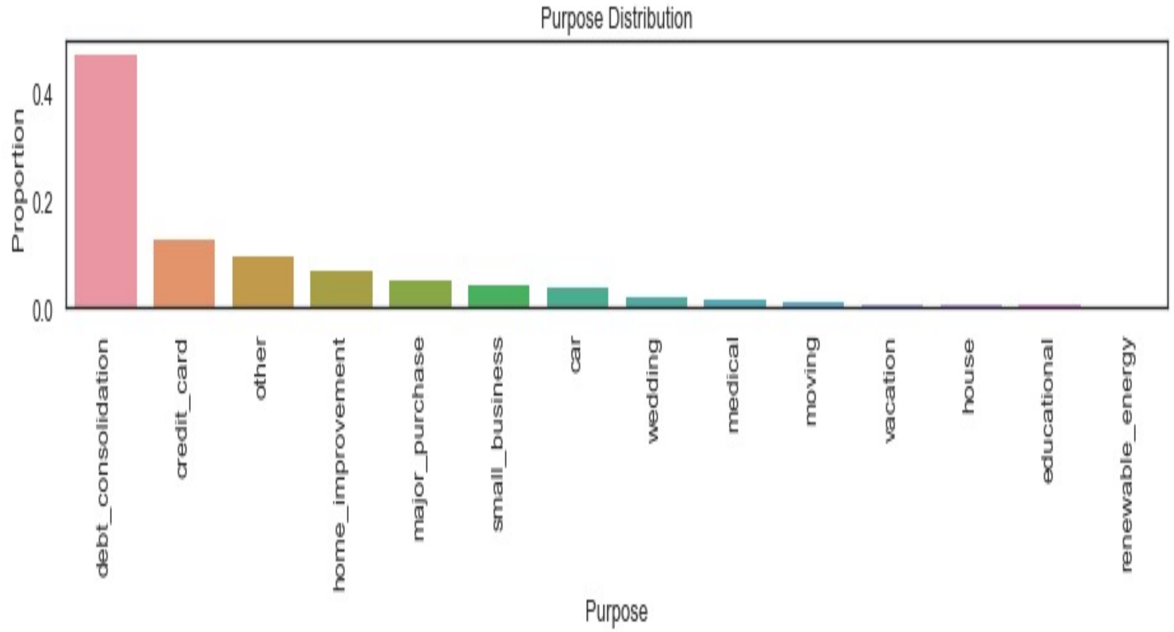



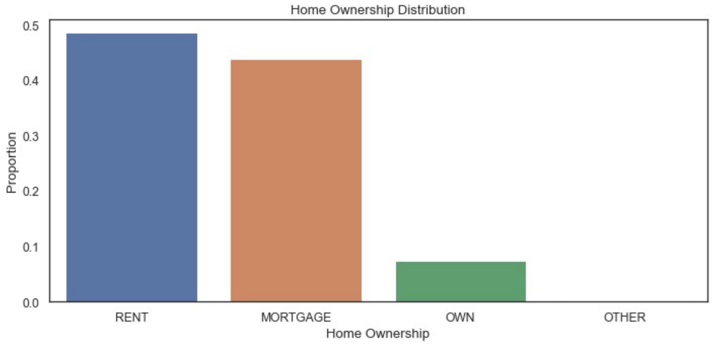

## <u>Observations:</u>

- Most of the loans are between 5000 to 10000 range
- Annual Income distribution plot shows that it needed an outlier treatment.
- Most of the loan applicants have an annual income in the range of 40K to 80K.
- Interest rates lies between 9% to 14.5%.
- A few borrower took loan at higher rates of interest greater than 22.5%

# UNIVARIATE ANALYSIS - Categorical



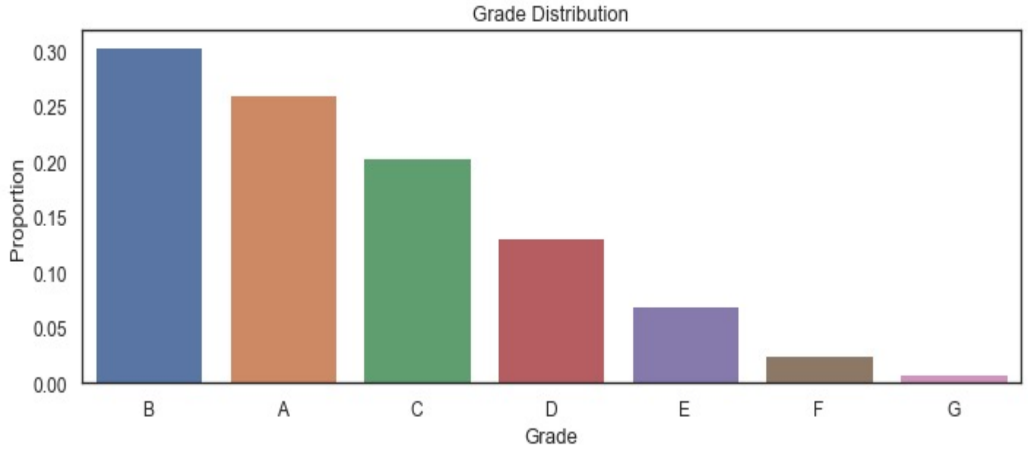



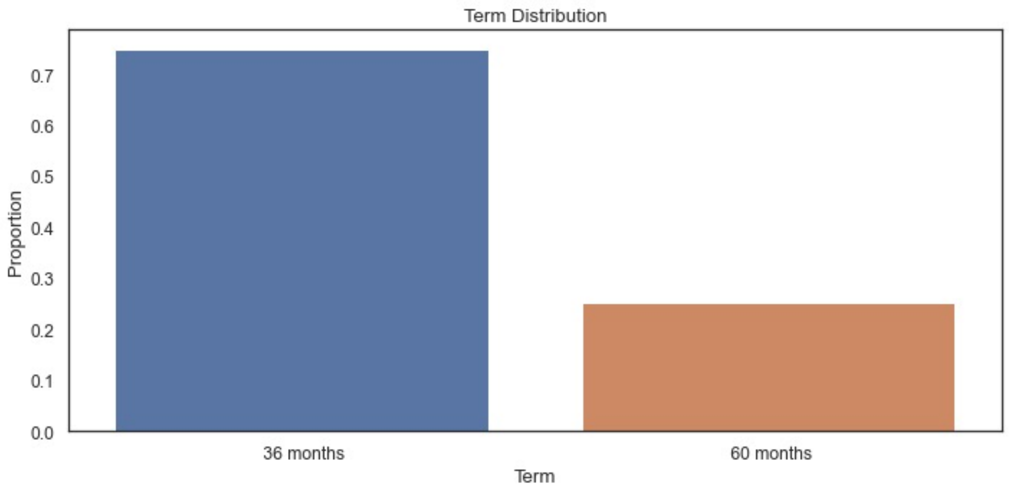



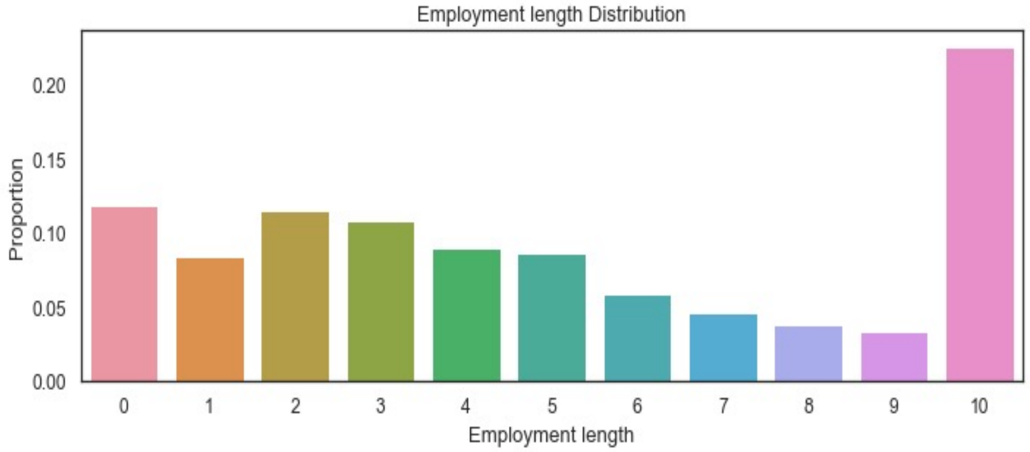

## Observations:

- We can see,  85% of loans are fully paid whereas 14% are defaulted.
- The lending club approve loans mostly for borrowers having rented and mortgage houses.
- Most of the loans are for the purpose of debt consolidation around 43% followed by credit card around 13%

# UNIVARIATE ANALYSIS – Ordered Categorical



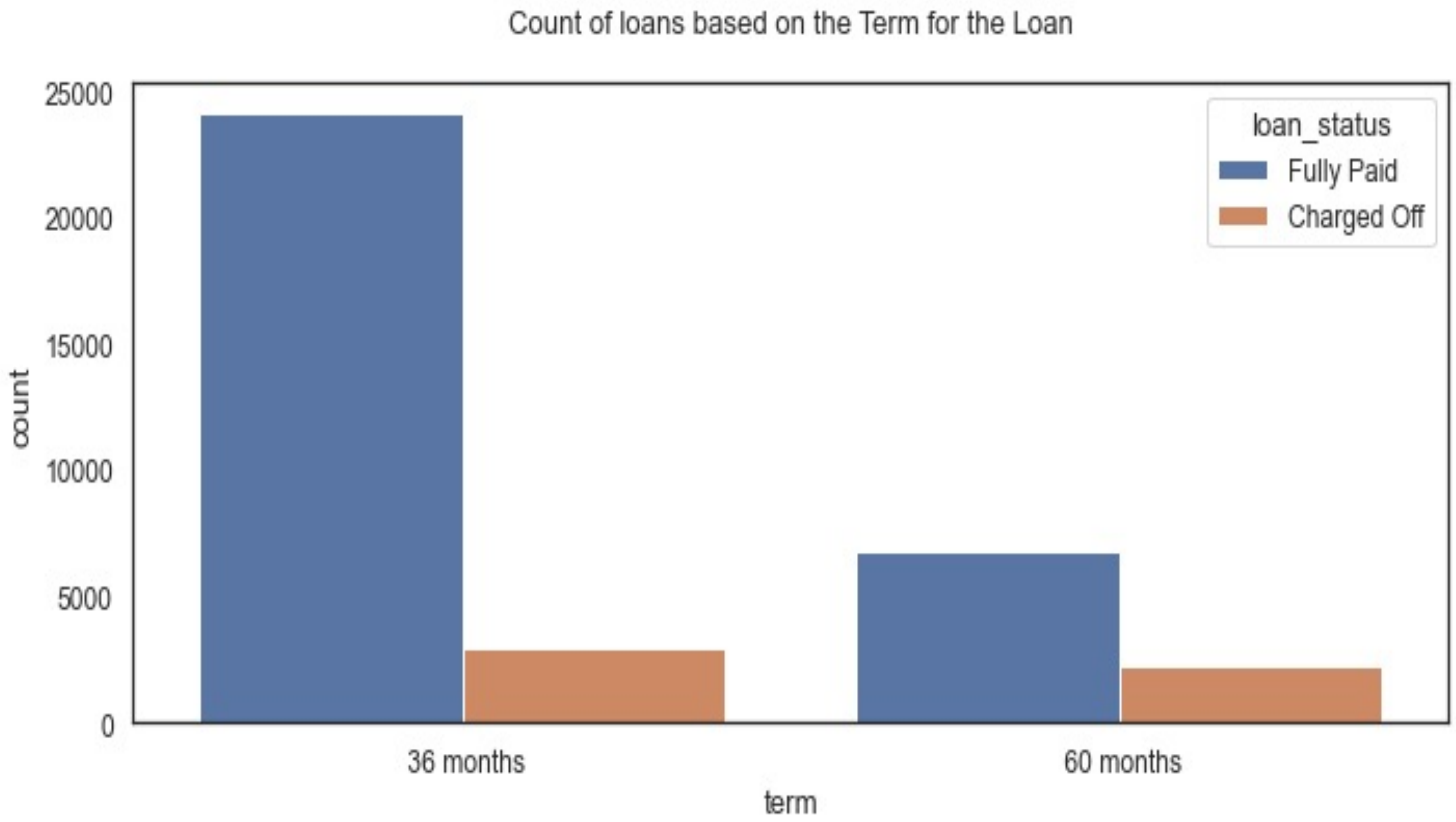




<u>Observations:</u>

- Grade A and B have more borrowers then other grades
- 10+ years has more borrowers then other employment length
- The lending club gives more loans for 36 months tenure than 60 months.

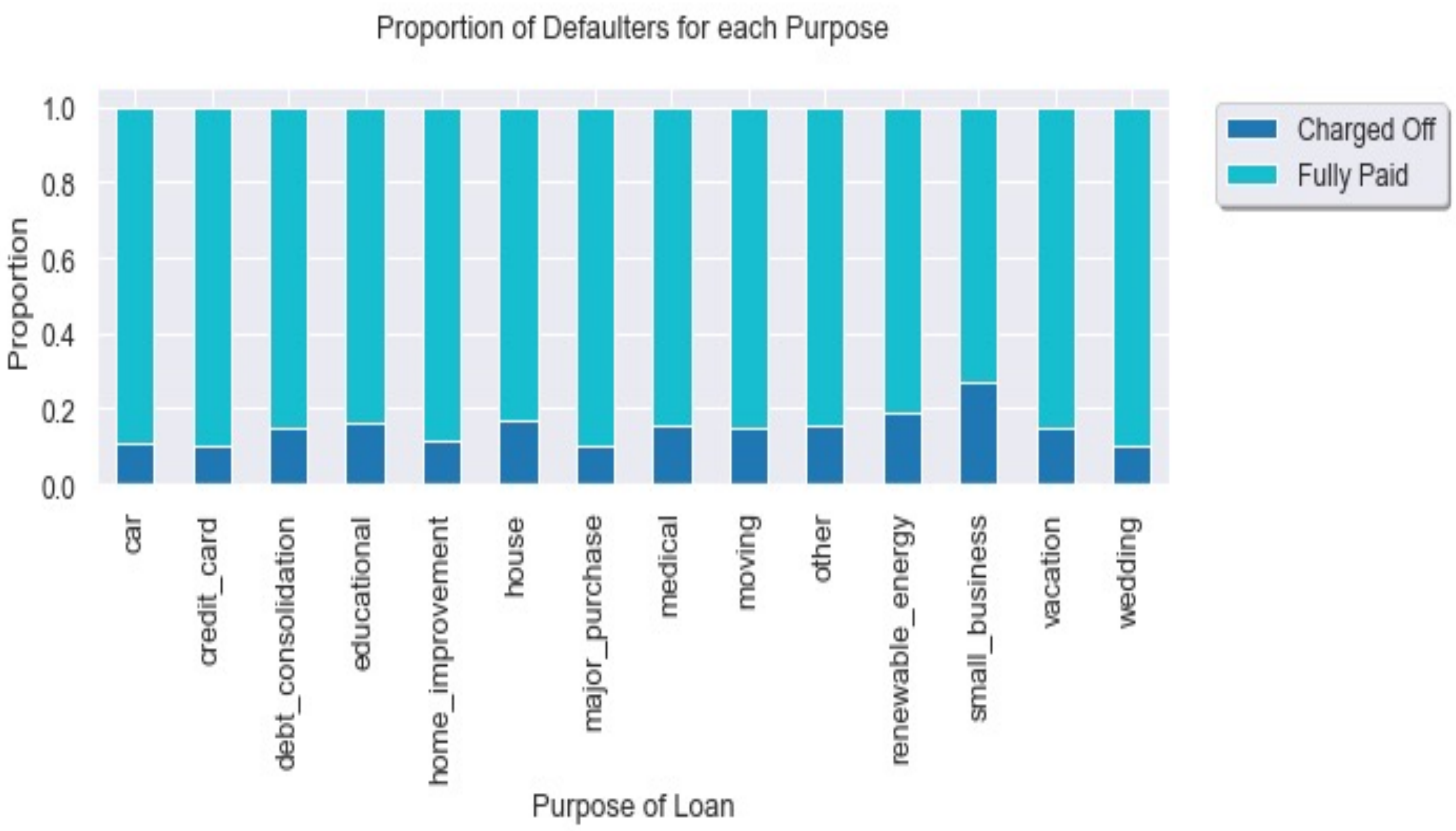
# SEGMENTED UNIVARIATE ANALYSIS




Observations:

- More number of loans are applied for 36 months term but the number of charged off for both categories are comparable

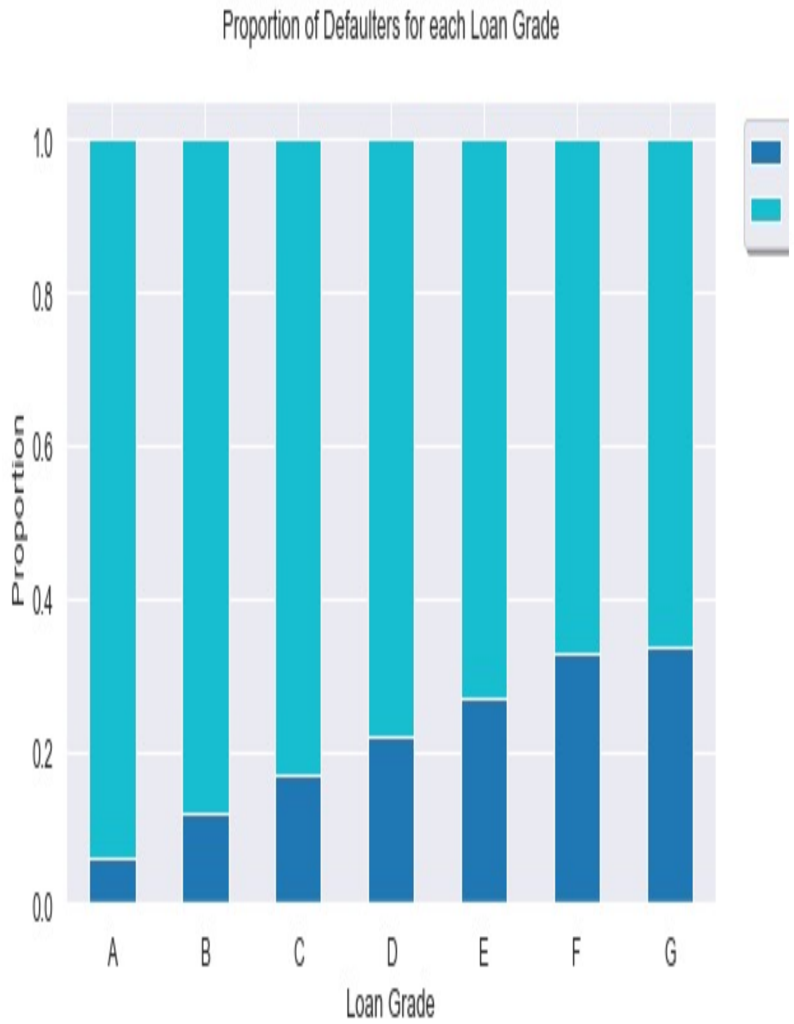
# Proportion of Defaulters for each Purpose

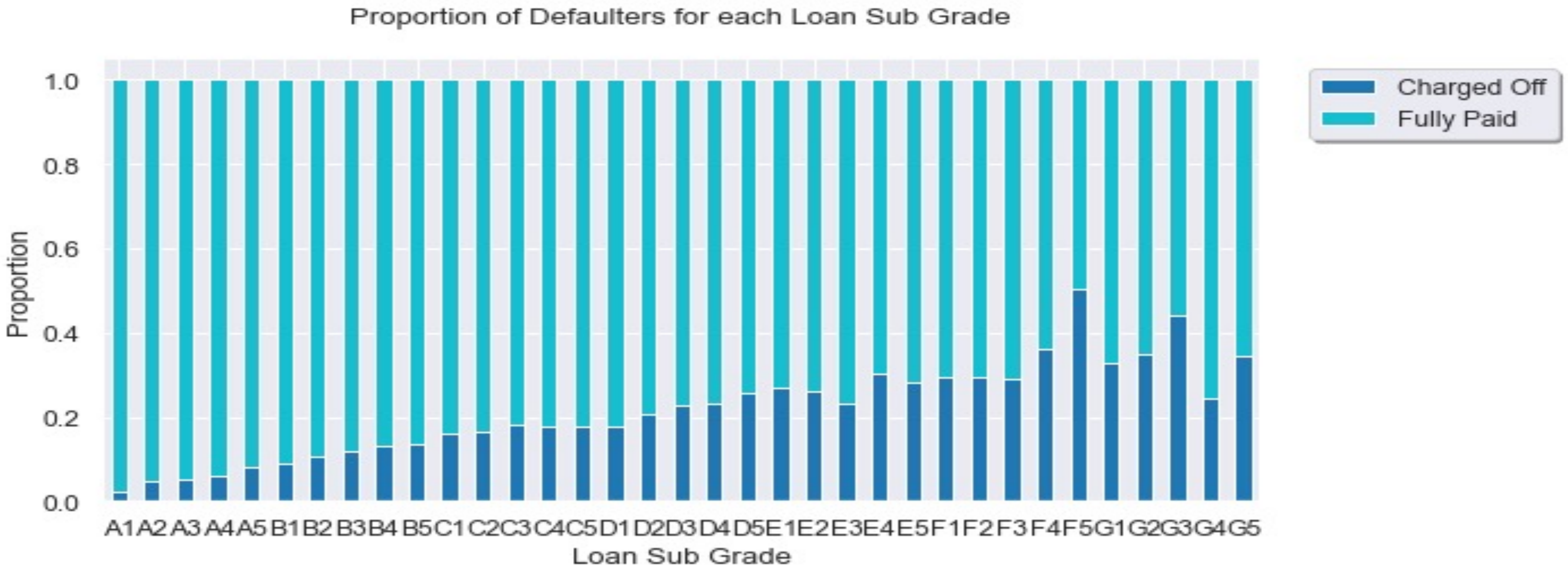



<u>Observations:</u>

Even though the number of Charged off loans are high under debt consideration, percentage-wise the loans taken for the purpose of 'Small Business' are around 23% charged off which is more than for any other purpose. Mostly these must be new business and must not be doing so well later.

# Proportion of Defaulters for each Grade/ Sub Grade

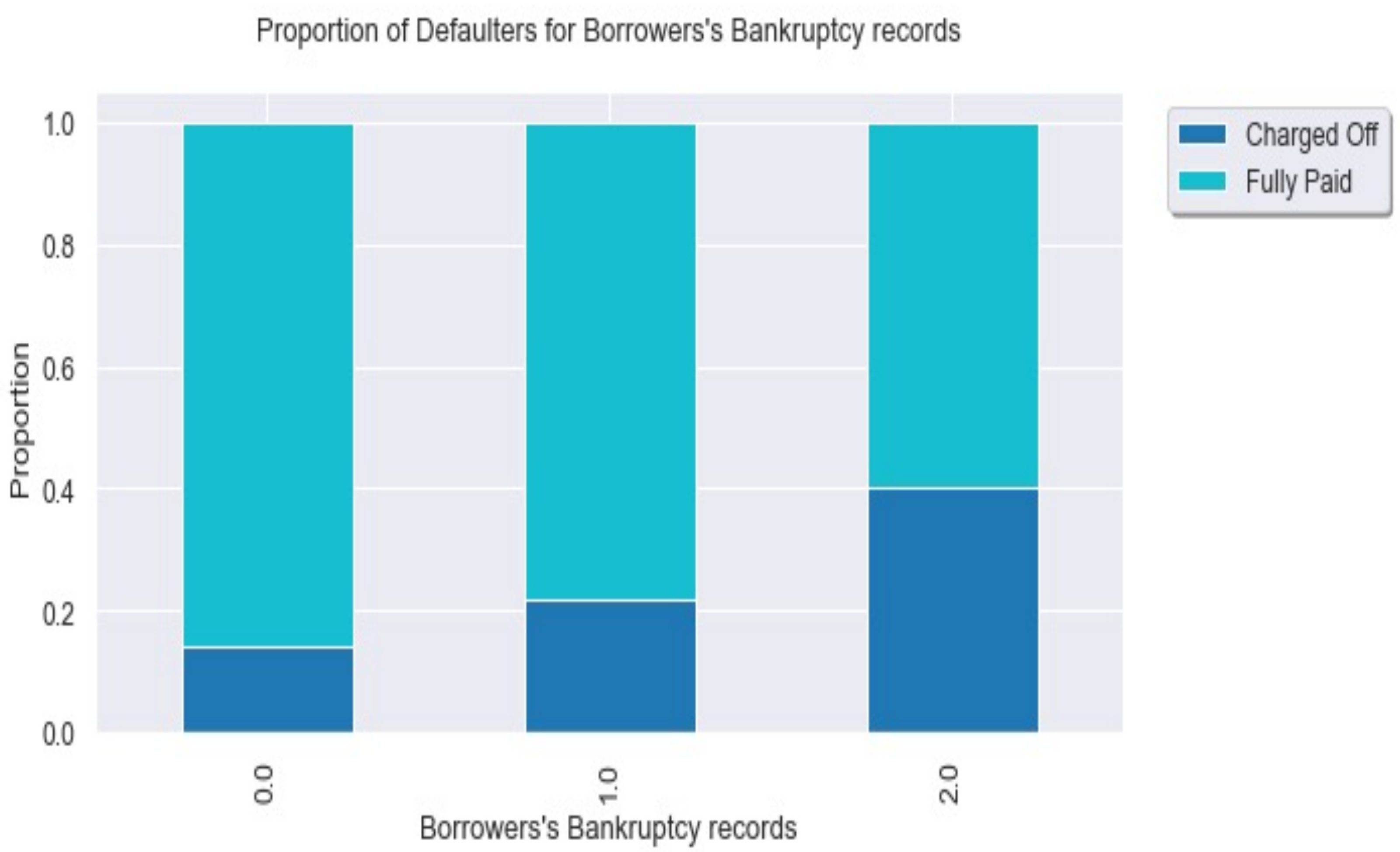



Observations:

- Percentage wise, grade F and G have more defaulters whereas grade A has the least. So, grade F and G type of loans are more likely to default
- Loan of Grade F5 has almost 50% defaulter

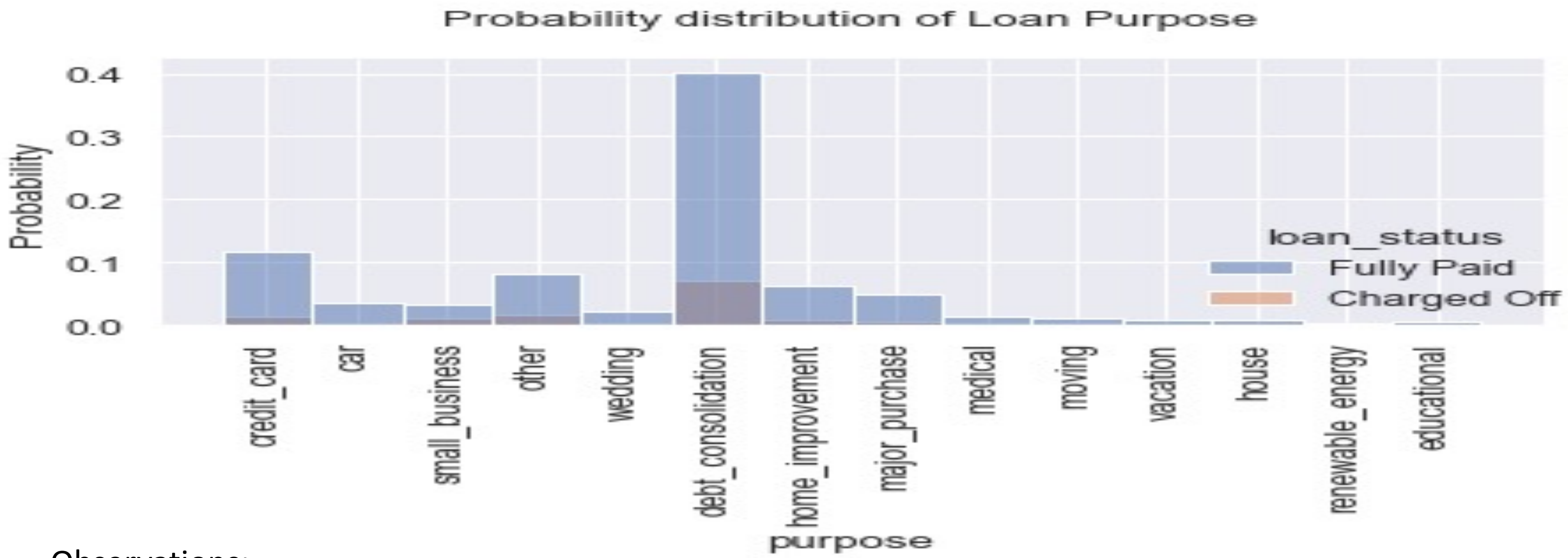
# Proportion of Defaulters with Bankruptcy Records




<u>Observations:</u>

Clearly, for higher number of Bankruptcy records, ratio of defaulters are high. Borrower's having 2 bankruptcy records are 40 % defaulters

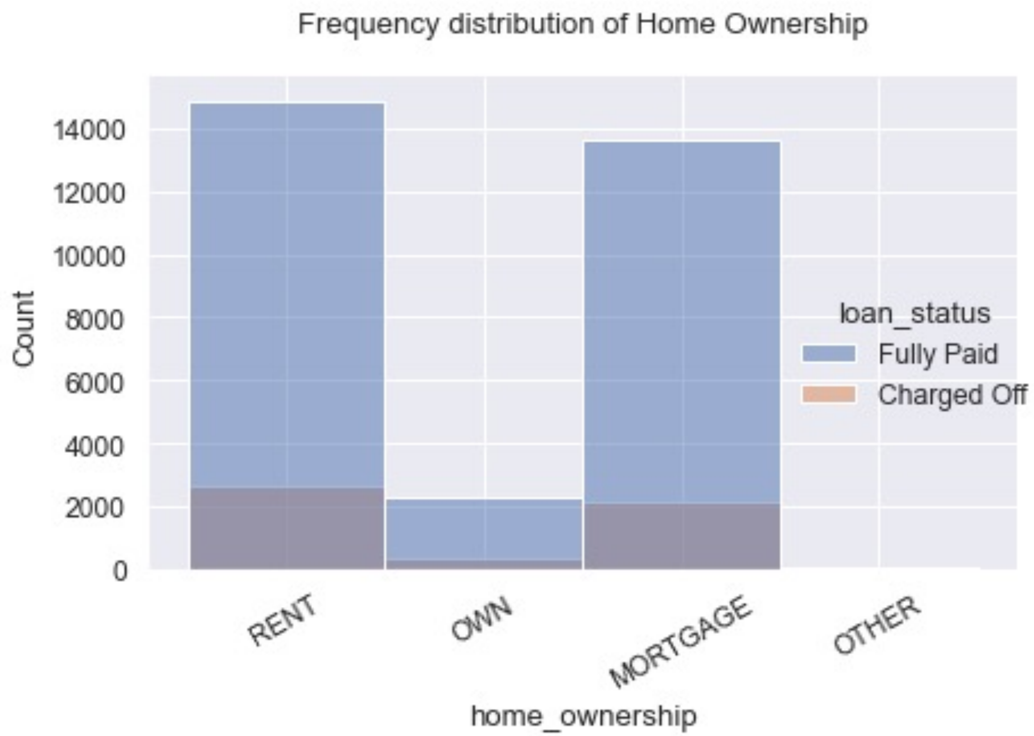
# Probability Distribution of Loan Purpose

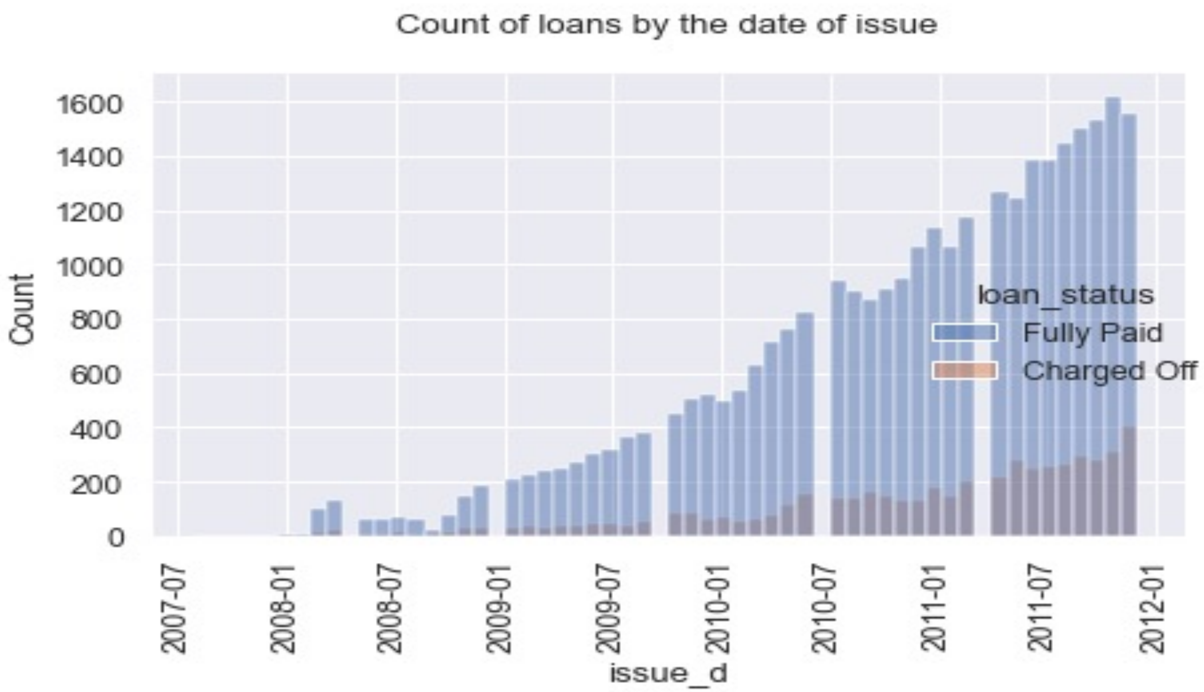



<u>Observations:</u>

- Most of the loans are for the purpose of debt consolidation around 43% followed by credit card around 13%
- The number of charged-off loans are also higher in category debt consolidation

# Count of Defaulters by Home Ownership and Date of Issue

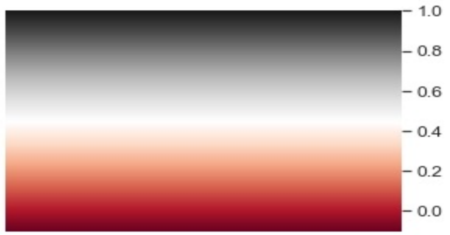



<u>Observations:</u>

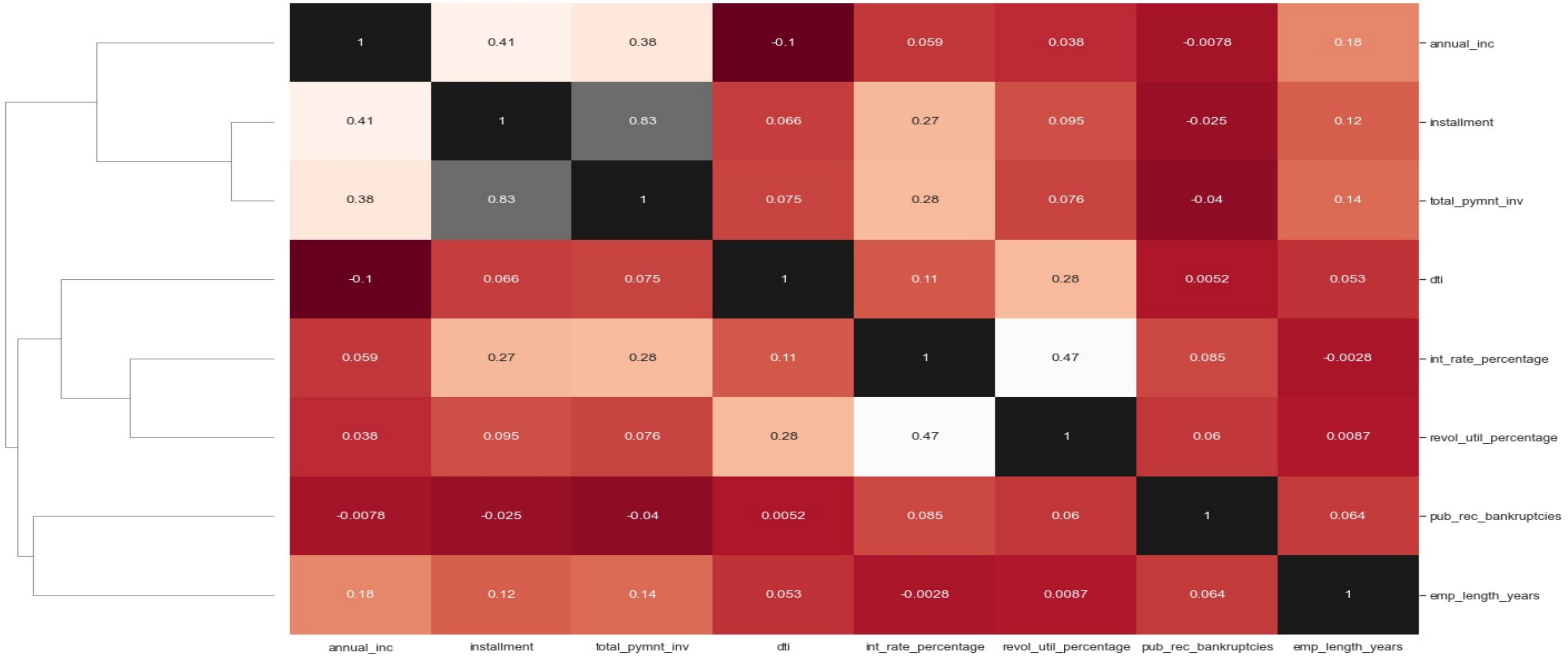
- The loan applicant are mostly the ones who have either rented or mortgaged and so the number of charged-off is also high under these categories
- The number of loans are increasing by every passing year and correspondingly the numbers of charged off applicants as well

# BIVARIATE ANALYSIS- Correlation

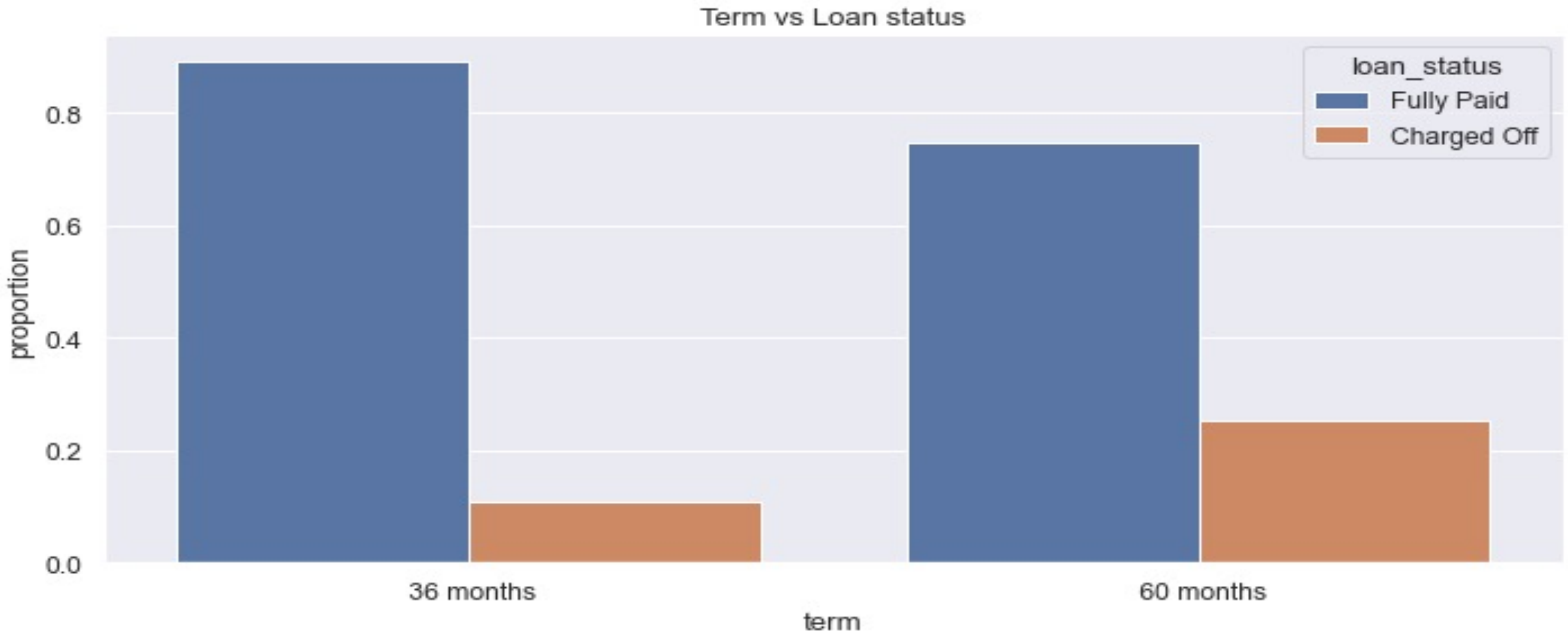



## Observation:

- Annual Income, installment, and total payment are strongly correlated.
- DTI is negatively correlated with annual income. So, the higher the income, less is debt to income ratio
- Annual income is positively correlated with employment years
- The higher the annual income, the less are the public bankruptcy records



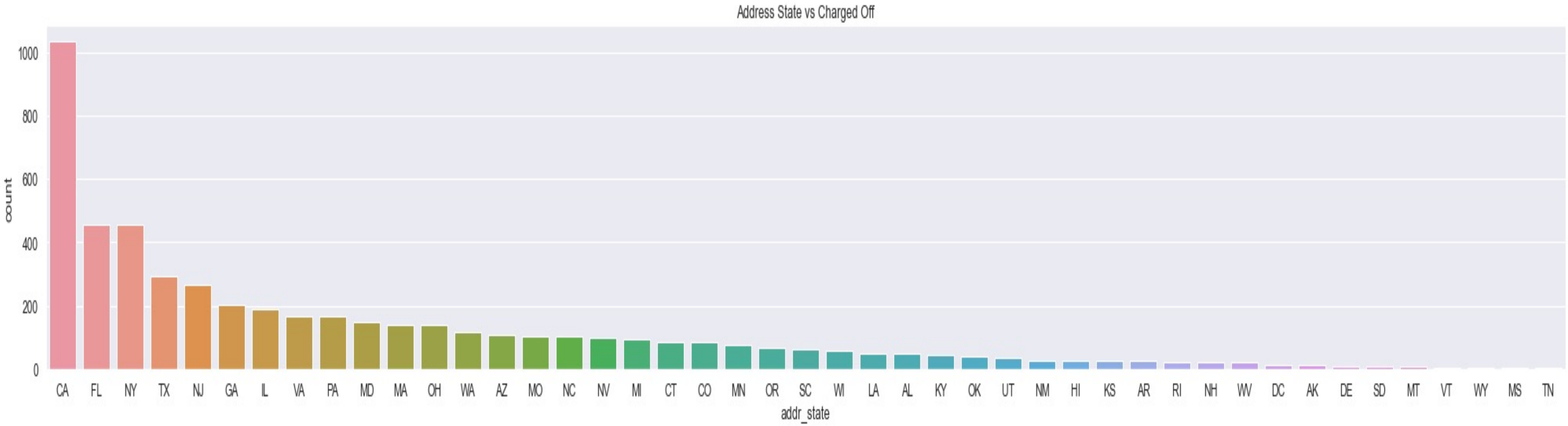
# BIVARIATE ANALYSIS: Term vs Loan Status




Observations:

- It seems, more proportion of borrowers have defaulted in 60 months tenure compared to 36 months.

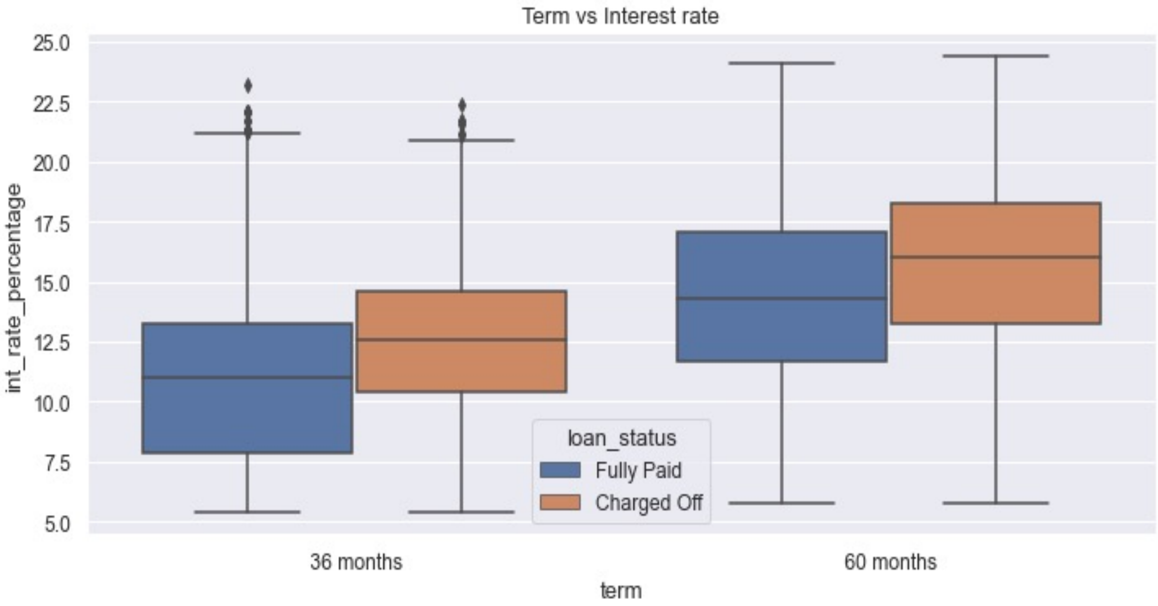
# Address State vs Loan Status

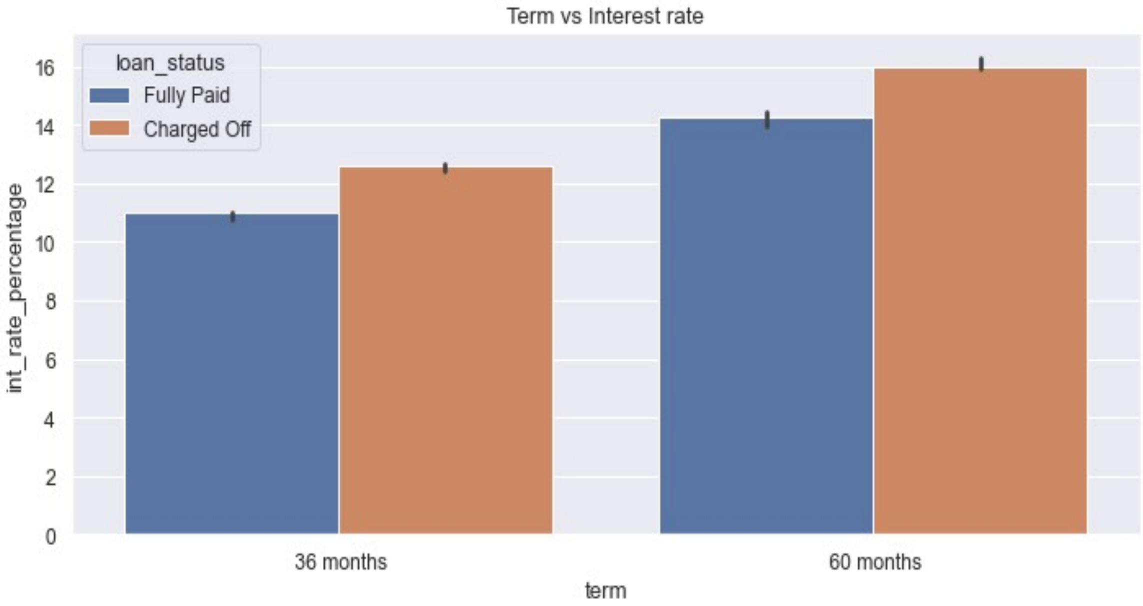



<u>Observations:</u>

- California , Florida and New York states have most number of defaulted borrowers.

# Term Vs Loan Interest Rate

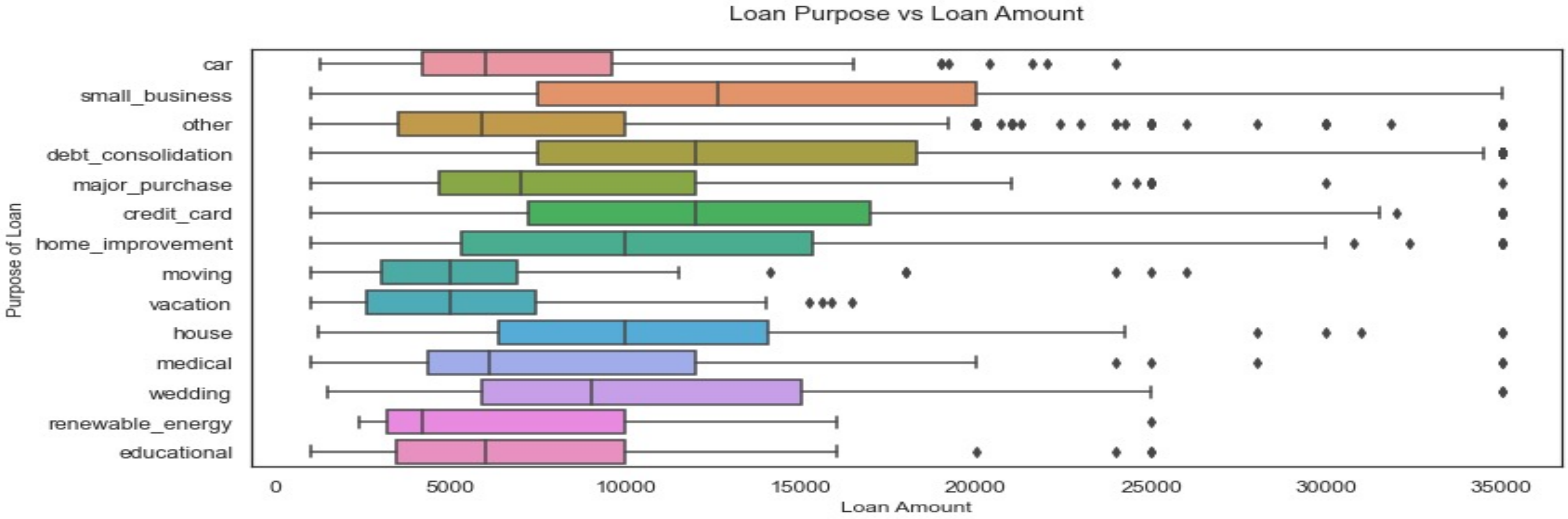



<u>Observations:</u>

- The interest rate for defaulters is higher for both the terms(36 and 60 months).

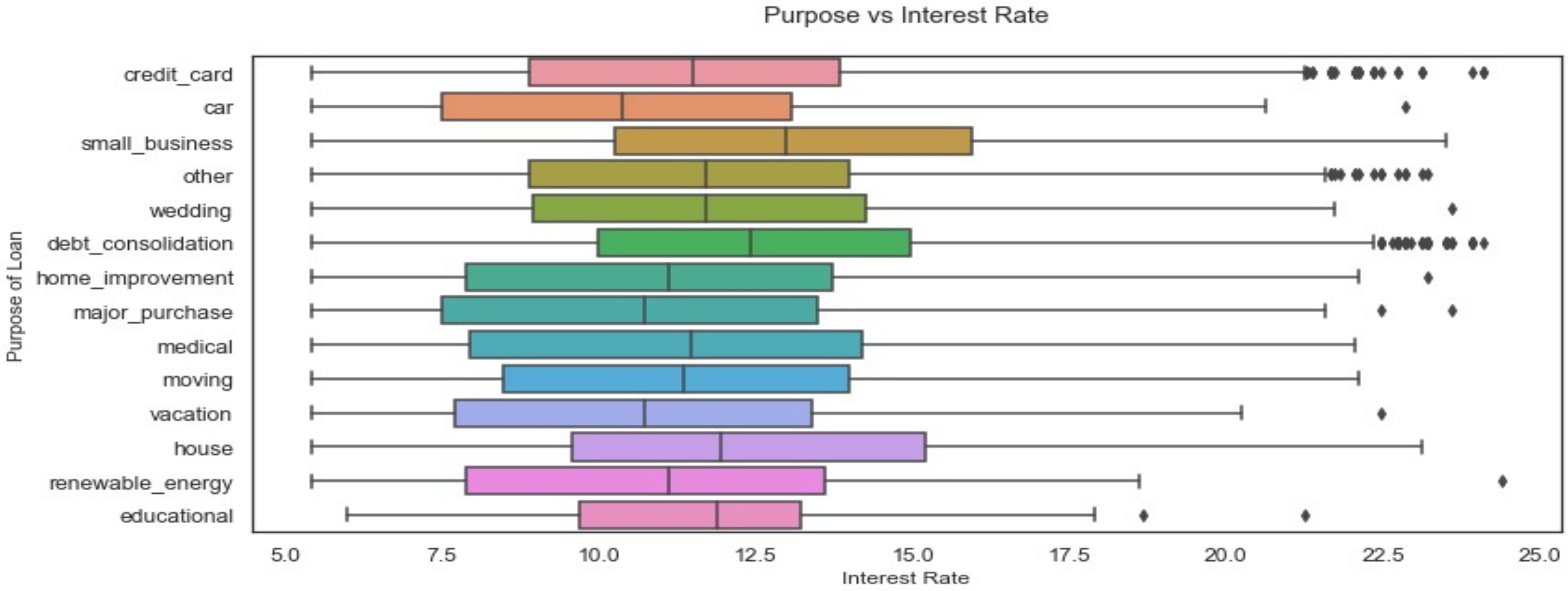
# Loan Purpose vs Loan Amount




Observations:

- Spread of Charged off loans is widest for Small Business Purpose and also it has highest median and 75th percentile value.

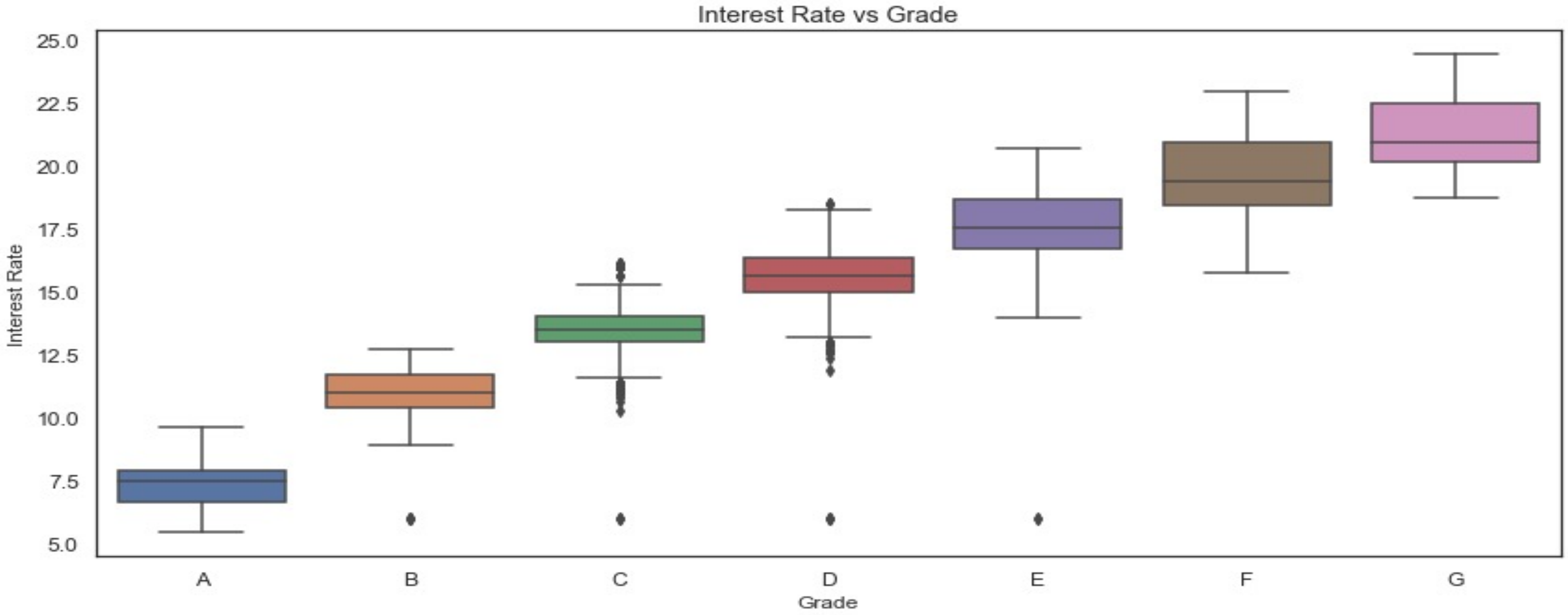
# Loan Purpose vs Loan Interest Rate




<u>Observations:</u>

- Loans taken for small business purpose have higher interest rates(median value), as they may be risky and so higher chances of defaulting

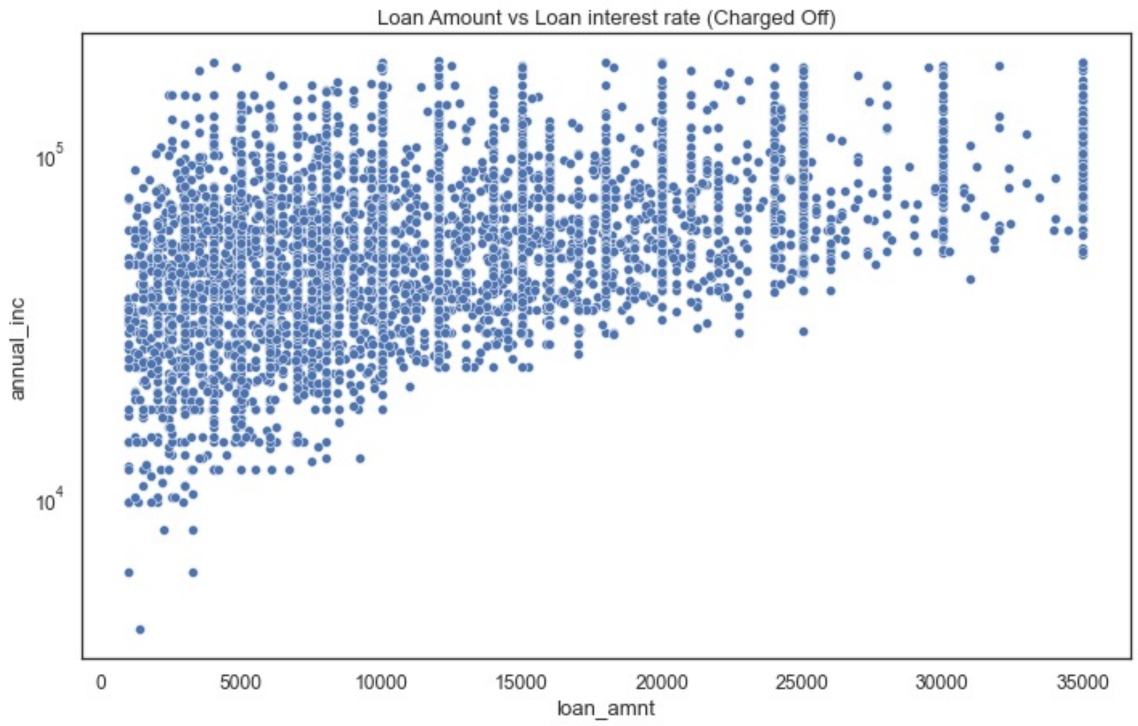
# Loan Interest Rate vs Grade

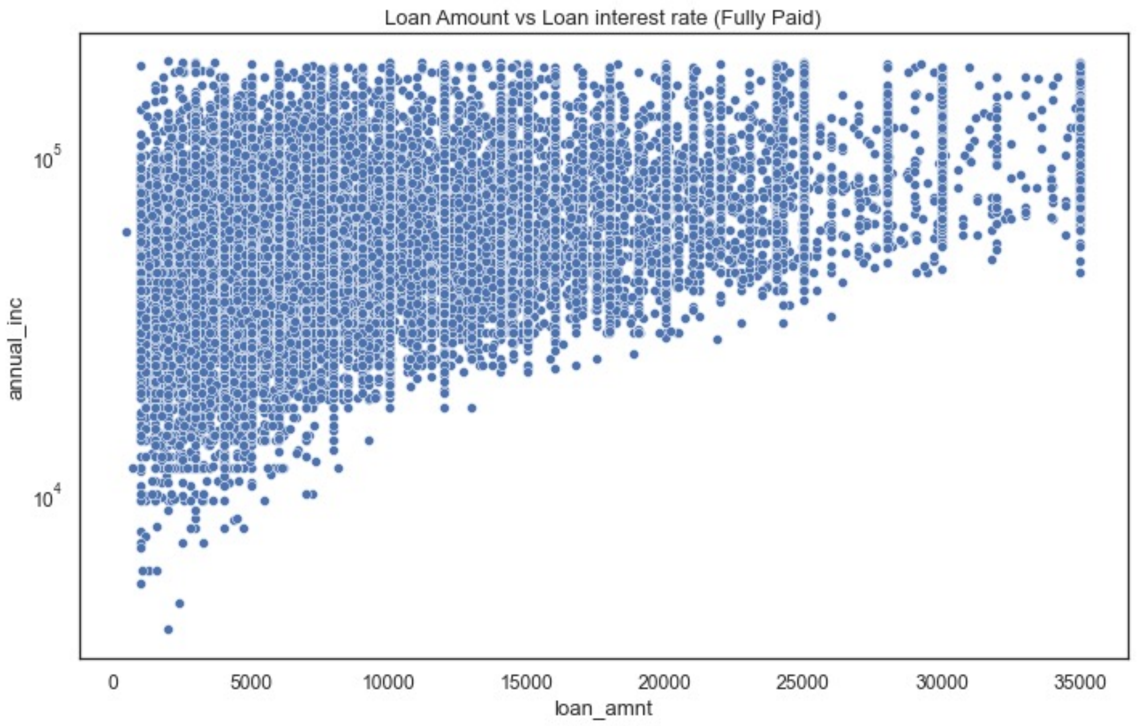



<u>Observations:</u>

- Loan Interest Rate increases as we move from Grade A to G.
- Clearly Grade F and G are risky loans, so higher interest rates

# Loan Amount vs Annual income

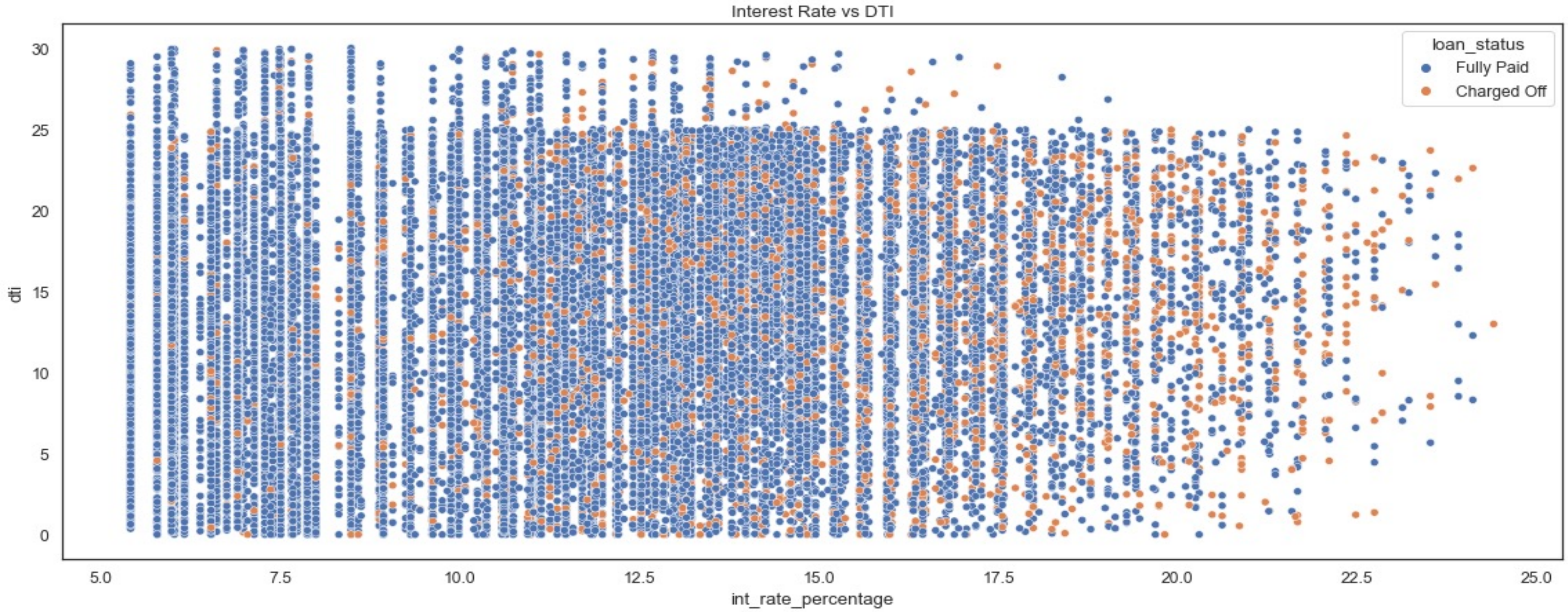



<u>Observations:</u>

- As compare to Annual income, both Fully paid and Charged Off loans are having similar pattern.
- Clearly, with higher income, the number of charged off loans are less

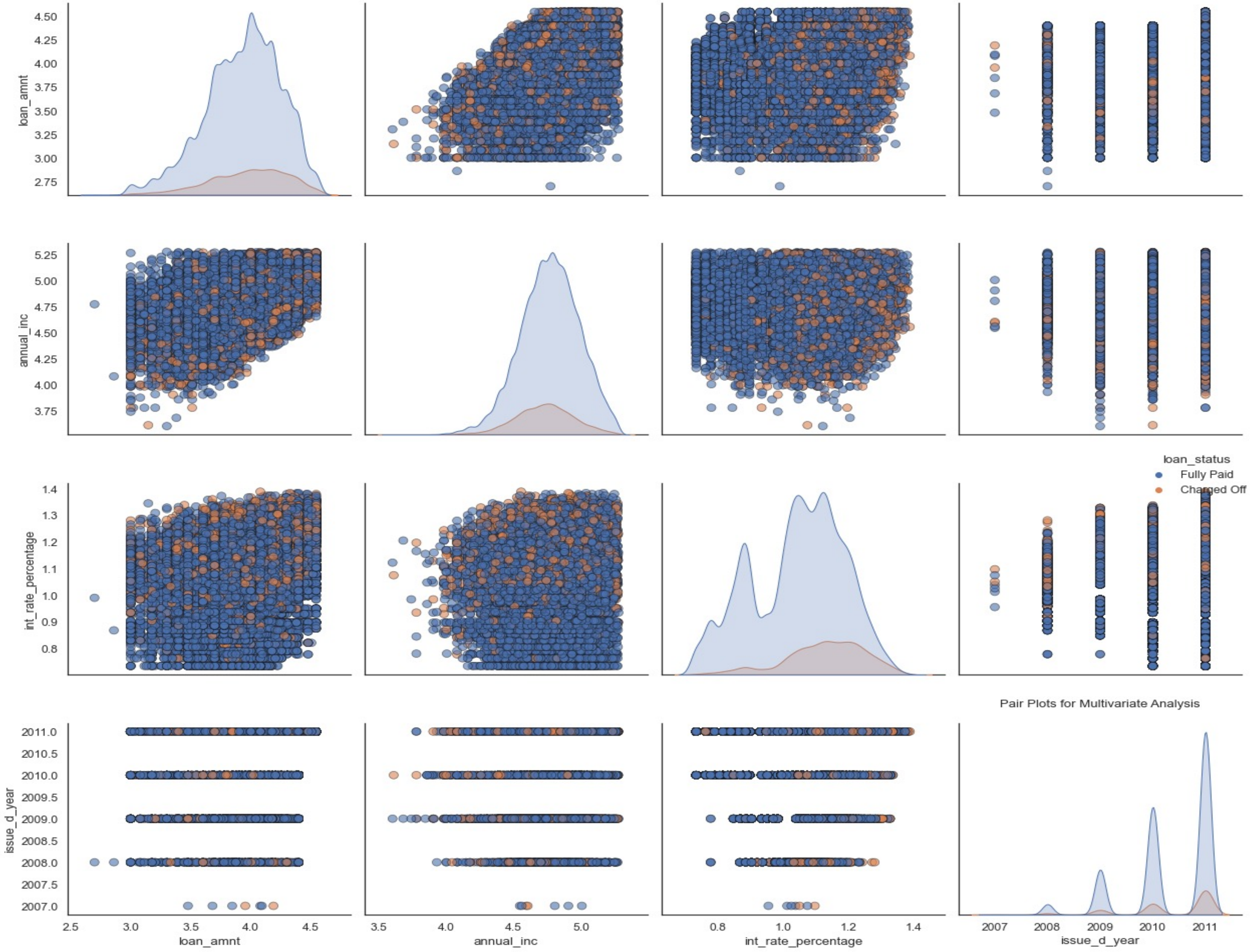
# Loan Interest Rate vs Debt to Income (DTI) Ratio




Observations:

- Charged off proportion is increasing with the increase in loan interest rate
- With lower interest rate, mostly the loans are fully paid

# MULTIVARIATE ANALYSIS - Pair Plots




Observations:

- As we can observe, there is higher charged-off ratio with higher interest rates.
- The loan amount slightly goes up with higher annual income.
- There is an increase in charged-off counts with the increase in year.
- There is an increase in Interest rate as loan amount increases

# IMPORTANT INFERENCES

- Loans given at a higher interest rates are more likely to default
- Lower Grades(F/G) Loans are more likely to be Charged-Off as compared to higher grades (A/B)
- Borrowers from states of California, Florida and New York are more likely to default. More background check is needed before loan approval.
- People with higher Bankruptcy Records have higher chance of defaulting
- Debt Consolidation and Small Business purpose loans must be scrutinized more as these can lead to higher financial loss.
- Borrowers having their own house or higher income are least likely to default
- Those having higher DTI ratio are also more likely to default.

thank you