Lending Club Case Study

Loan Default EDA

Business Overview

- The company specializes in lending loans to urban customers.
- 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, then approving the loan may lead to a financial loss for the company.
- The data provided contains information about past loan applicants and whether they defaulted or not.
- There are a total of 111 columns and 39717 records in the loan.csv file.

The aim is to identify patterns that indicate if a person is likely to default, 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, etc.
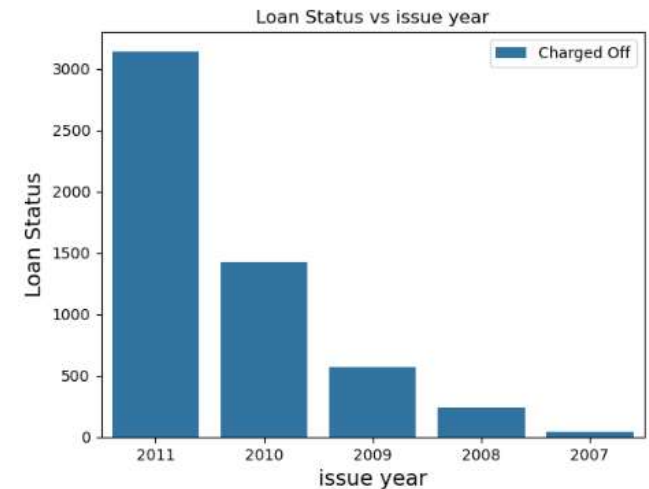
Univariate Analysis

Number of loans that have been charged off by month.

The month with the most charged off loans appears to be December, with above 600 loans charged off. There is a significant decrease in charged off loans in the following months, with February having the fewest at around 300.
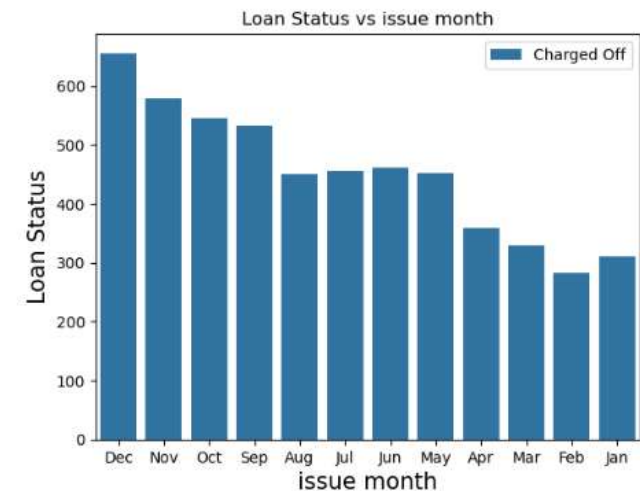

Total number of loans charged off by year of issuance.

It appears that a higher number of loans were charged off the more recently they were issued. For example, according to the graph, there were more loans charged off in 2011 than in 2007. It's important to note that may be data only goes up to 2011, so it is impossible to say for sure if this trend continues.
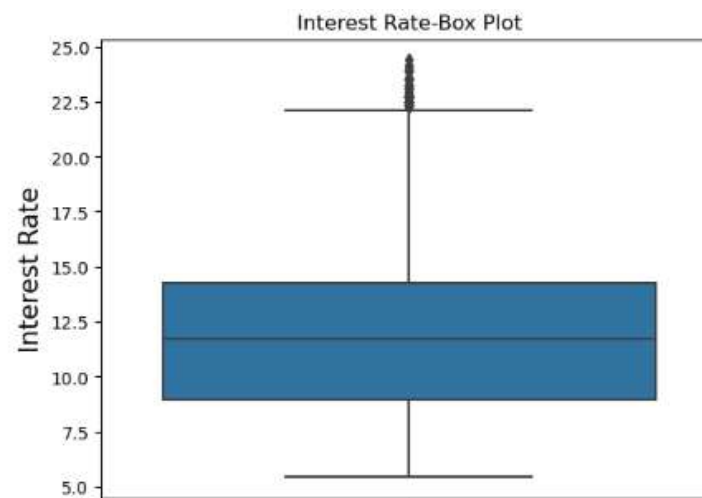
Univariate Analysis of Loan amount Variable

Majority of the loans fall within the range of 5000 to 15000
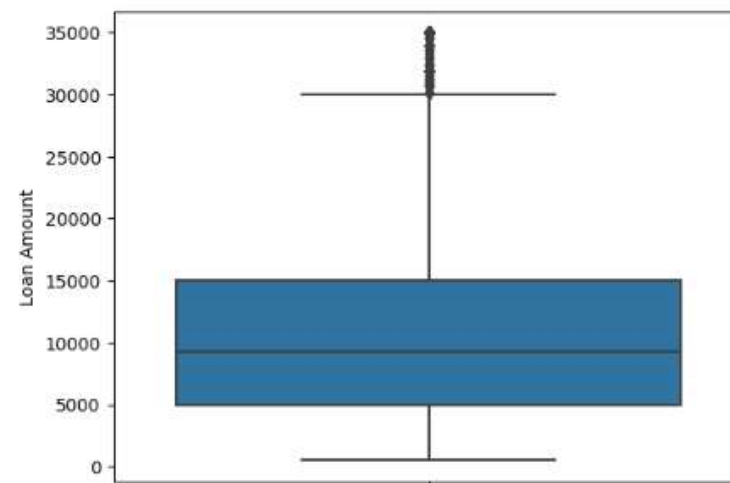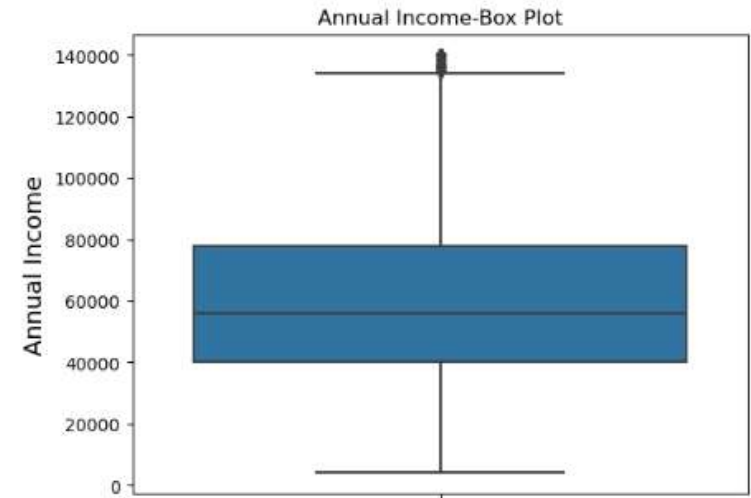

Univariate Analysis of Interest rate Variable

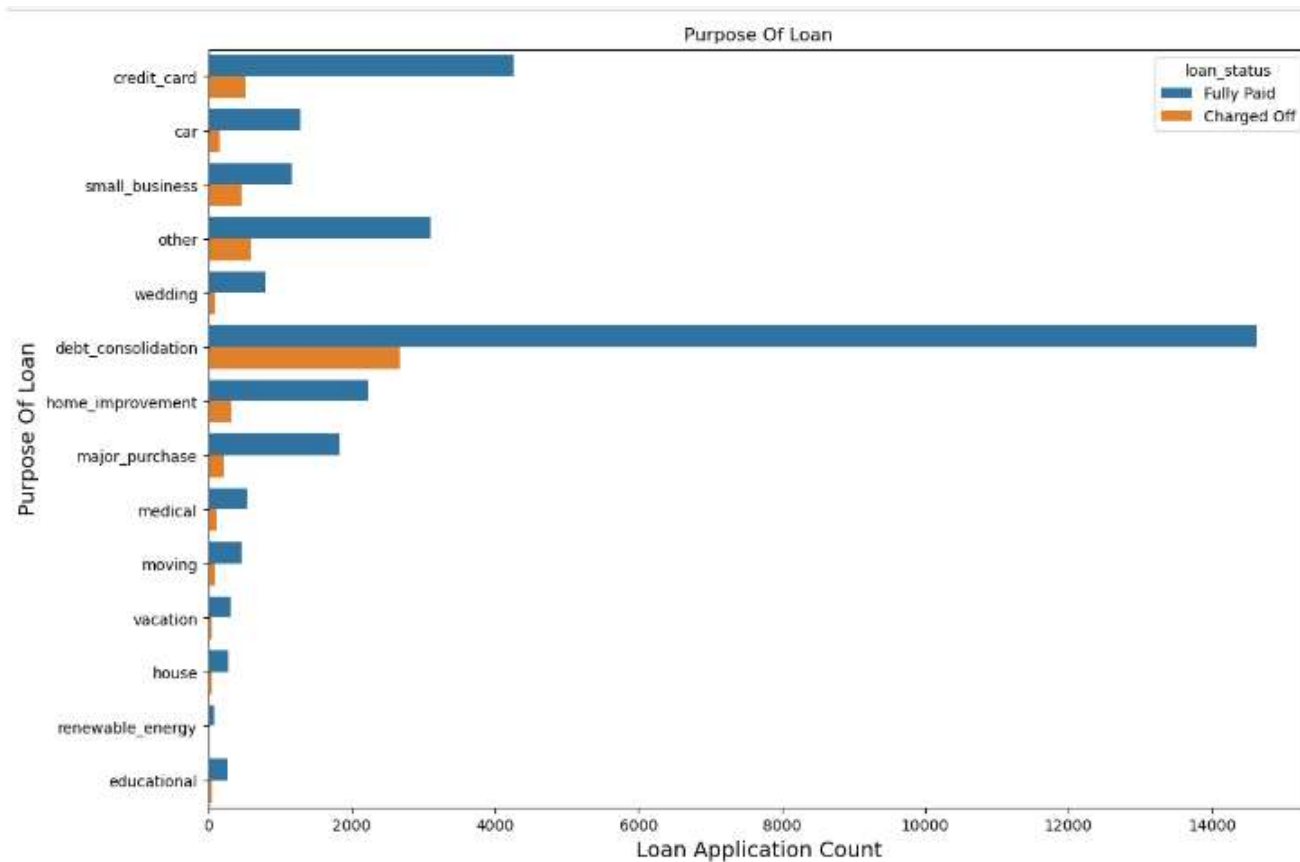Majority of the interest rates are in the range 10%-15%

Univariate Analysis for Annual Income Variable

Majority of the Annual Income are in the range 25000-75000

Annual Income-Box Plot

Univariate Analysis- Purpose of Loan

The Count Plot shows that most of the loans were taken for the purpose of debt consolidation & paying credit card bill. Number of charged off count also high too for these loans.
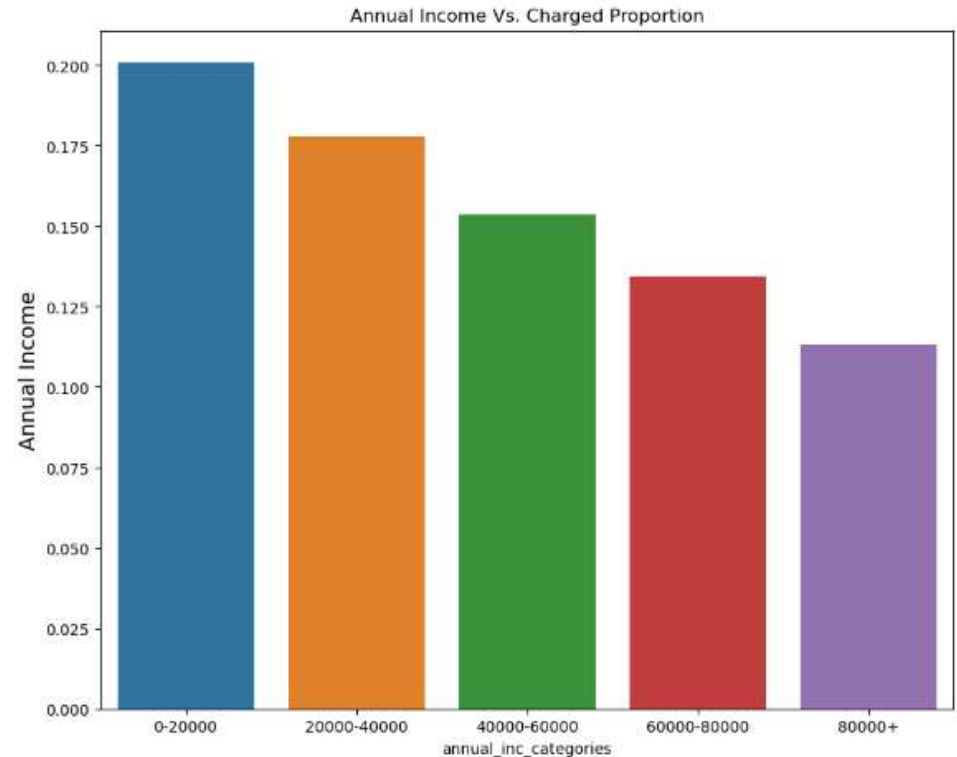
Bivariate Analysis

Bivariate analysis on employment length vs Charged Off Proportion

Applicants with less than a year of work experience are more likely to have a charge-off due to limited income to repay the loan. The rest of the applicants have a similar chance of being charged off.

Employment Length vs Charged Off Proportion
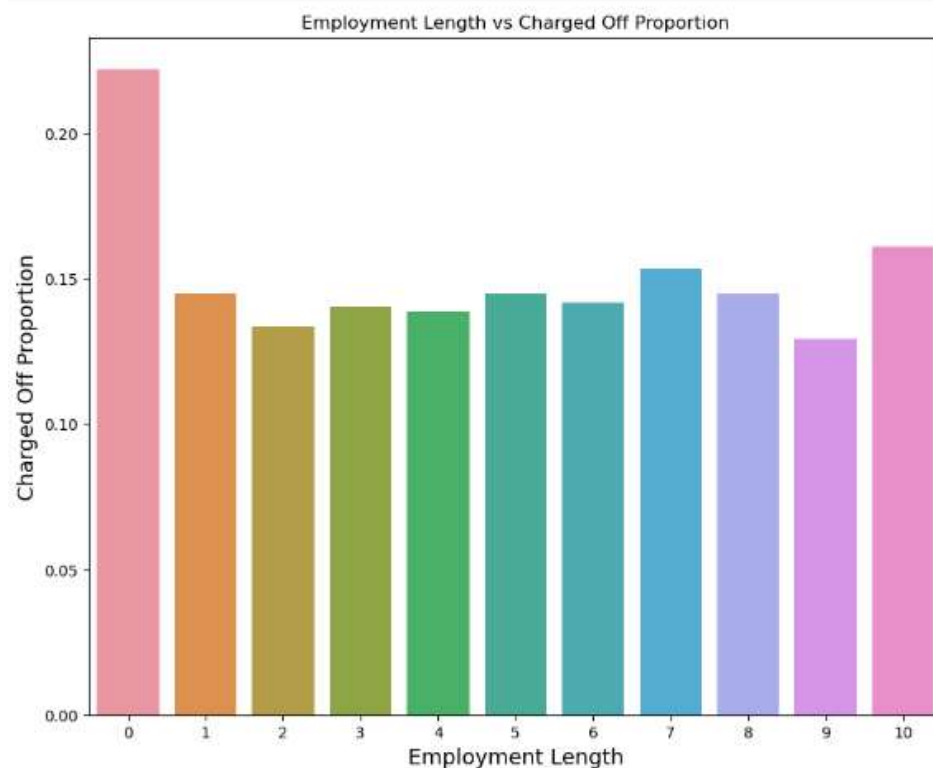
Bivariate Analysis on annual income on Charged Off Loans

- Income range 80000+ has a higher creditworthiness, resulting in a lower chance of chargedoff.
- There is a direct relationship between annual income and creditworthiness. Borrowers with higher income typically have a lower likelihood of having debt charged off.
- Income range 0-20000 has a lower creditworthiness, leading to a higher chance of charge-off.


Annual Income Vs. Charged Proportion

Bivariate analysis on pub_rec_bankruptcies against ChargedOff Proportion

- Those who already have pub_rec_bankruptcies value 1, have charged off proportion higher than who have no pub_rec_bankruptcies.
- pub_rec_bankruptcies count 2 has even higher charged off proportion but those numbers are not significant to decide.
- Not known is the column for which we don't have any information amount borrower.
- This also makes sense that who has defaulted before has more chances of defaulting in future as well.

Public Bankruptcy Records vs Chargedoff Proportion
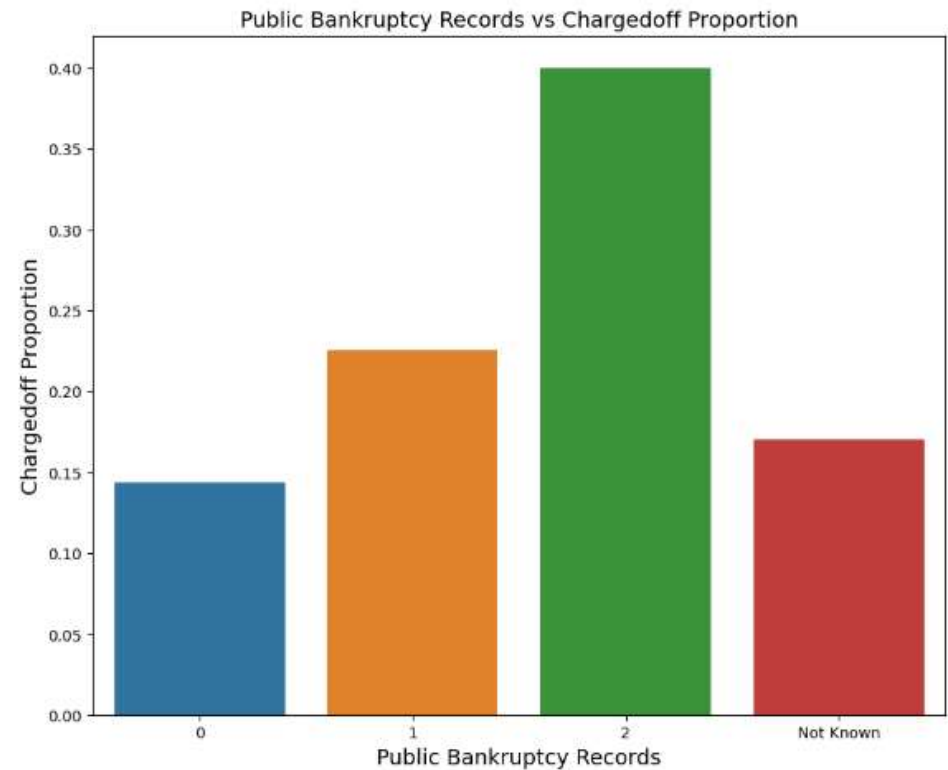
Bivariate analysis of grade against ChargedOff Proportion

- Borrowers with a Grade A rating have the lowest risk of default.
- Borrowers with Grade F or G ratings have a high risk of default.
- The risk of default increases as the creditworthiness grade decreases (A to G).

Grades vs Charged Off Proportion
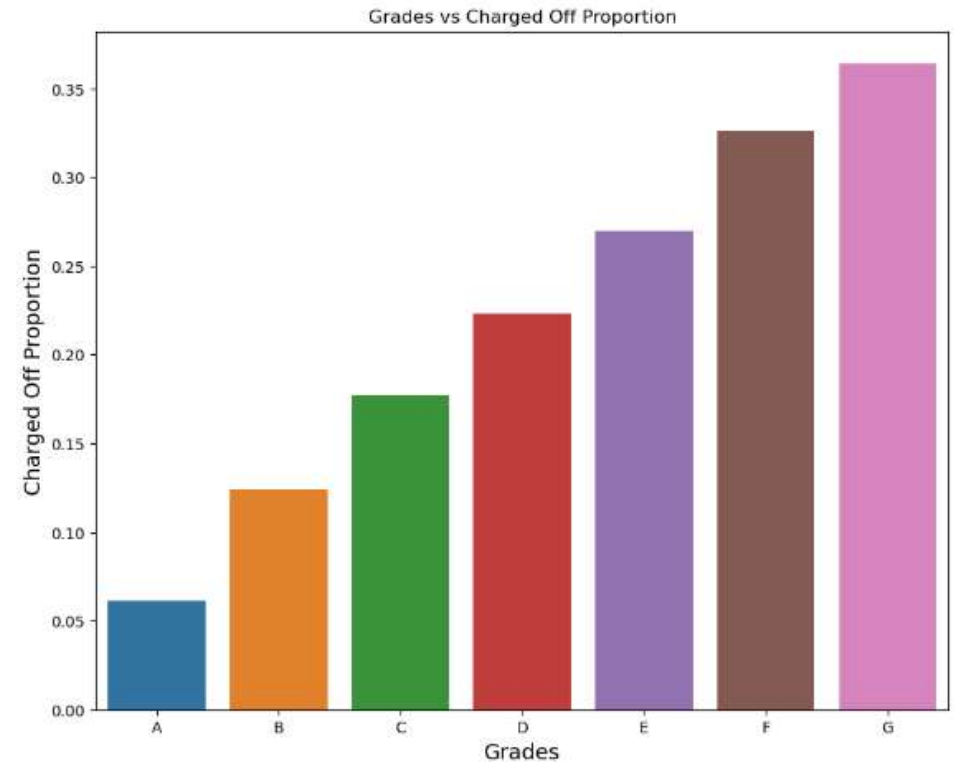
# Bivariate Analysis Purpose of Loan Vs Loan amount

Among loan purposes, small businesses have the highest median, 95th percentile, and 75th percentile loan amounts. Debt consolidation follows in second place, and credit cards come in third.
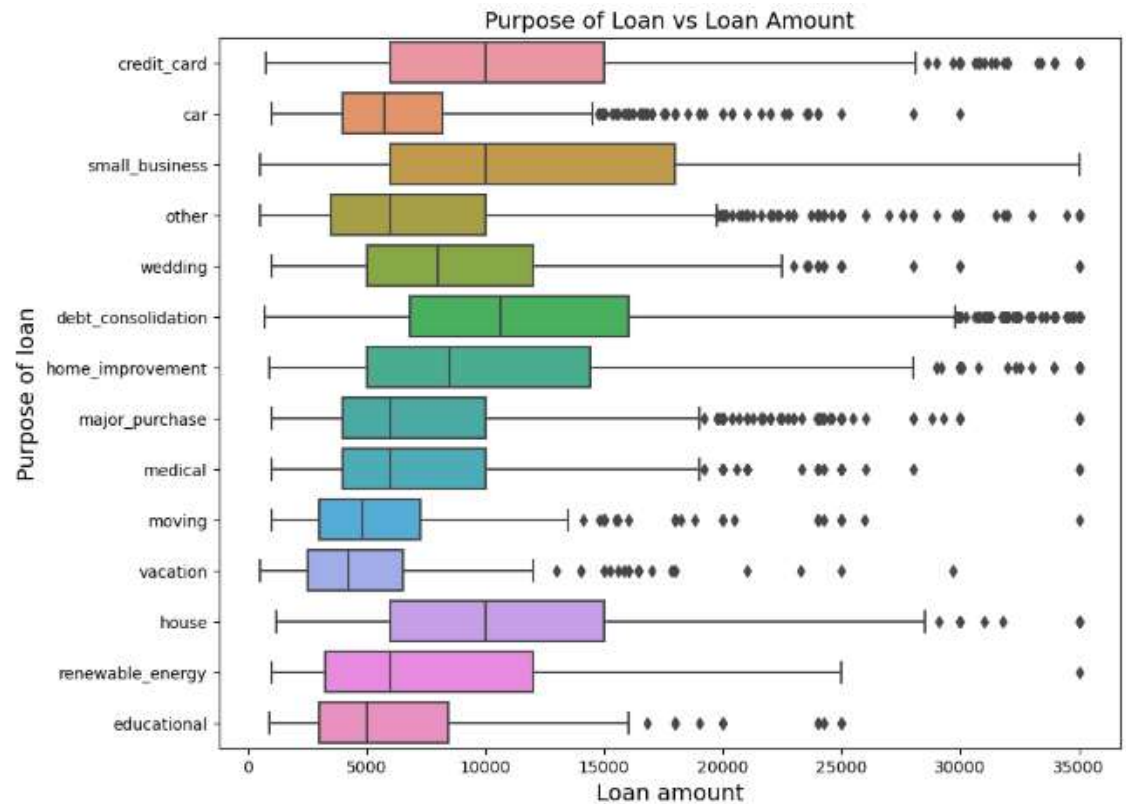
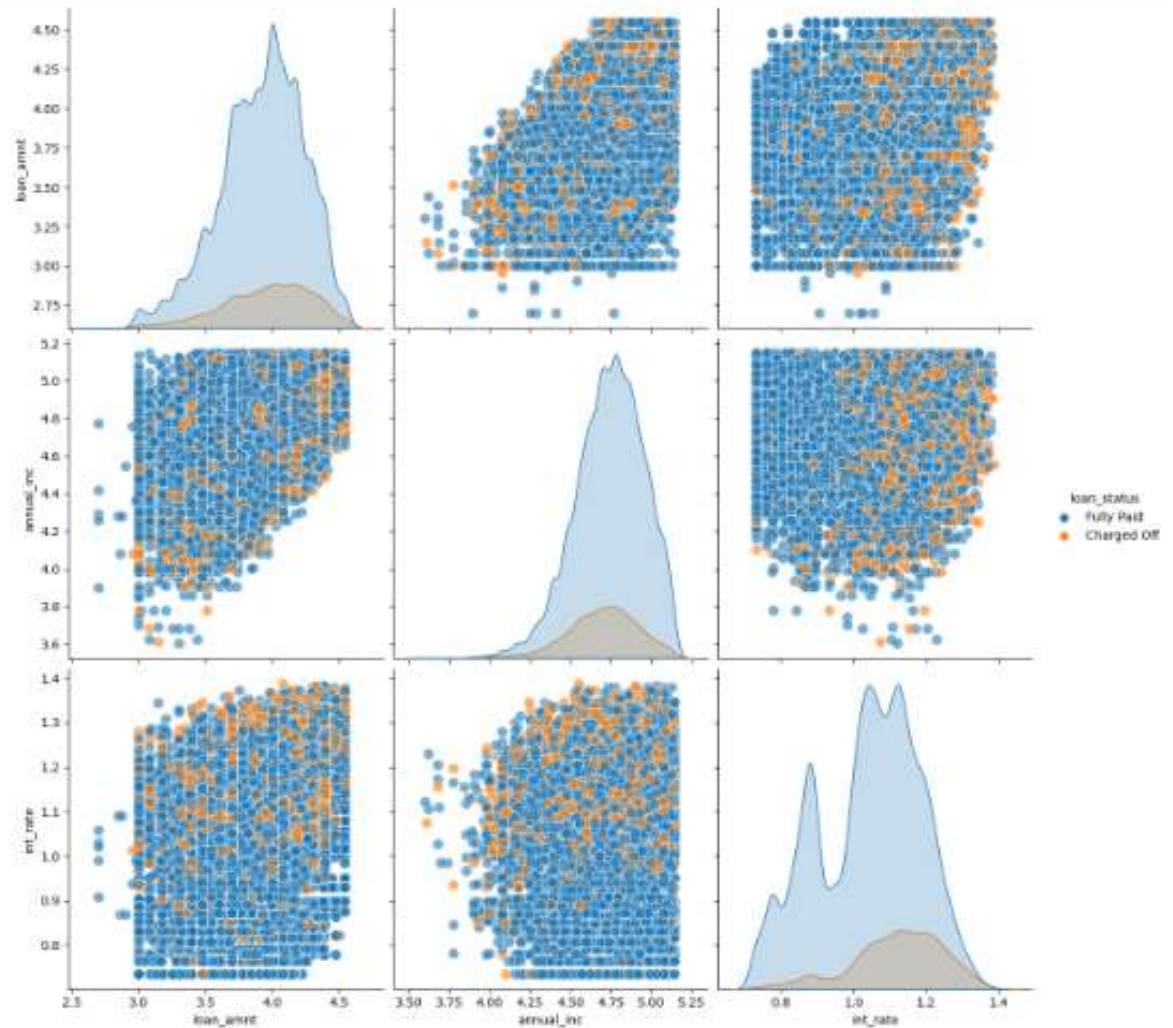**Bivariate Analysis Purpose of Loan Vs Interest rates**

Compared to other loan types, small business loans often carry higher interest rates.
Debt consolidation loans come in second for having high interest rates.
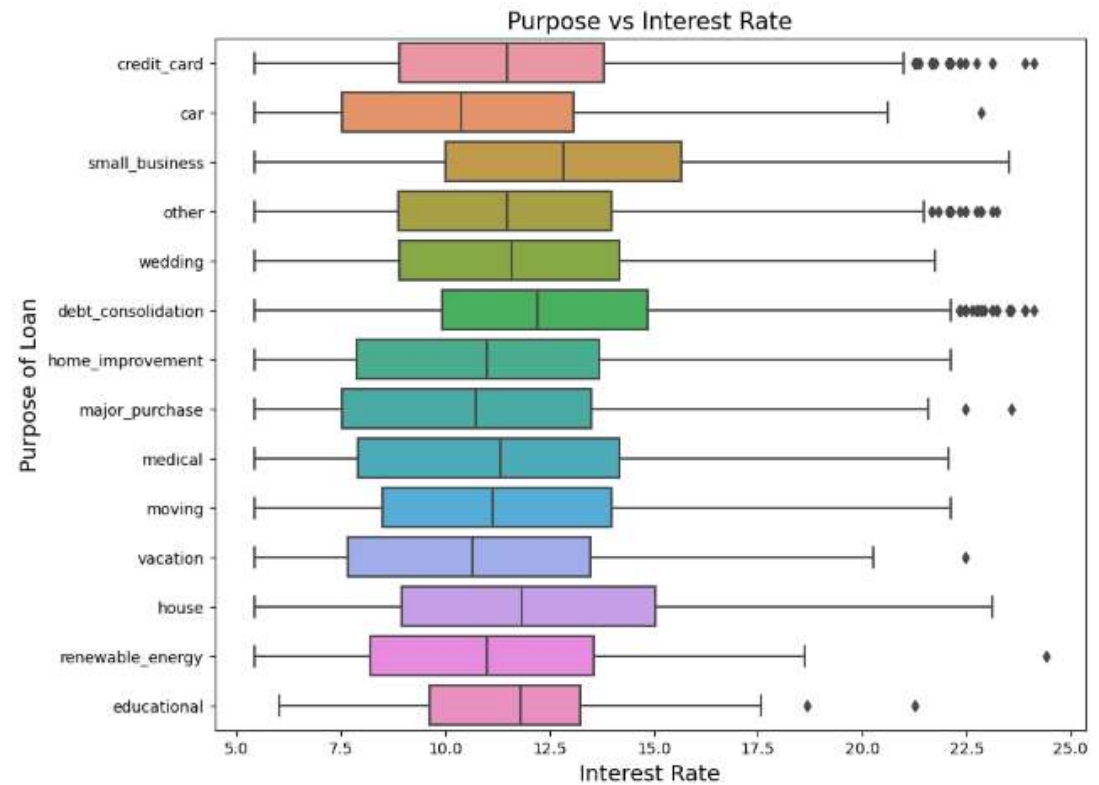
Multivariate Analysis

- Loans with higher interest rates are more likely to be unpaid.
- Yearly, there is an increase in number of charged off.
- Slightly Higher the loan amount for a higher the annual income.
- With increase in loan amount there is an increase in interest rate

Thank You