Lending Club- A Case Study

Group Members:
- Neha Sawadekar
- Neeraj Tripathi

Problem Statement

Lending Club, a marketplace for personal loans, faces a challenge in decision-making when evaluating loan applications. The company must decide whether to approve or reject each application based on the applicant's profile. This decision is fraught with two significant risks:

1. Not approving a loan for an applicant who is likely to repay results in a loss of potential business.
2. Approving a loan for an applicant who is likely to default leads to a financial loss.

Goal

The goal is to identify the driving factors or variables that are strong indicators of loan default. Understanding these factors will enable Lending Club to enhance its portfolio and risk assessment processes, thereby minimizing financial risks and maximizing business opportunities.

Approach

01   **Data Exploration**

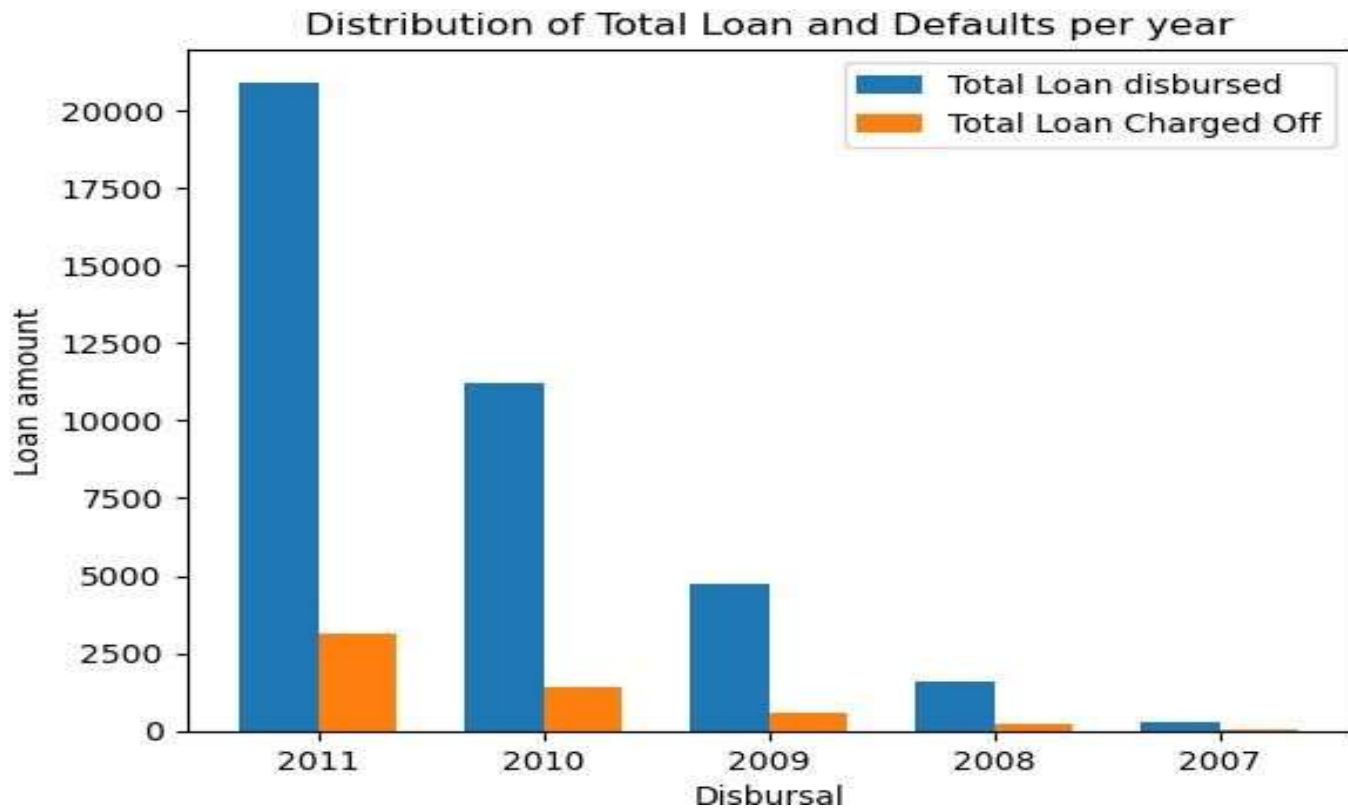**02**   **Data Transformation**

03   **Actionable Insights**

04   **Conclusion**

## Distribution of Total Loan and Defaults per year

Data Cleaning steps

- Imputing or removal of columns with Null values

- Removal of unwanted characters that interfere with calculations viz. %,>

- Dropping rows where imputing with mean/median/mode introduces bias

- Converting data types to facilitate Categorical / Quantitative Analysis

- Deriving new columns from existing for additional insights

- Binning continuous variables to plot as categorical variables

Loan Amount:

We noticed that 75% of the Loans are in the range of 15000 and below. Plotting the spread above and below the 75th percentile (15000)
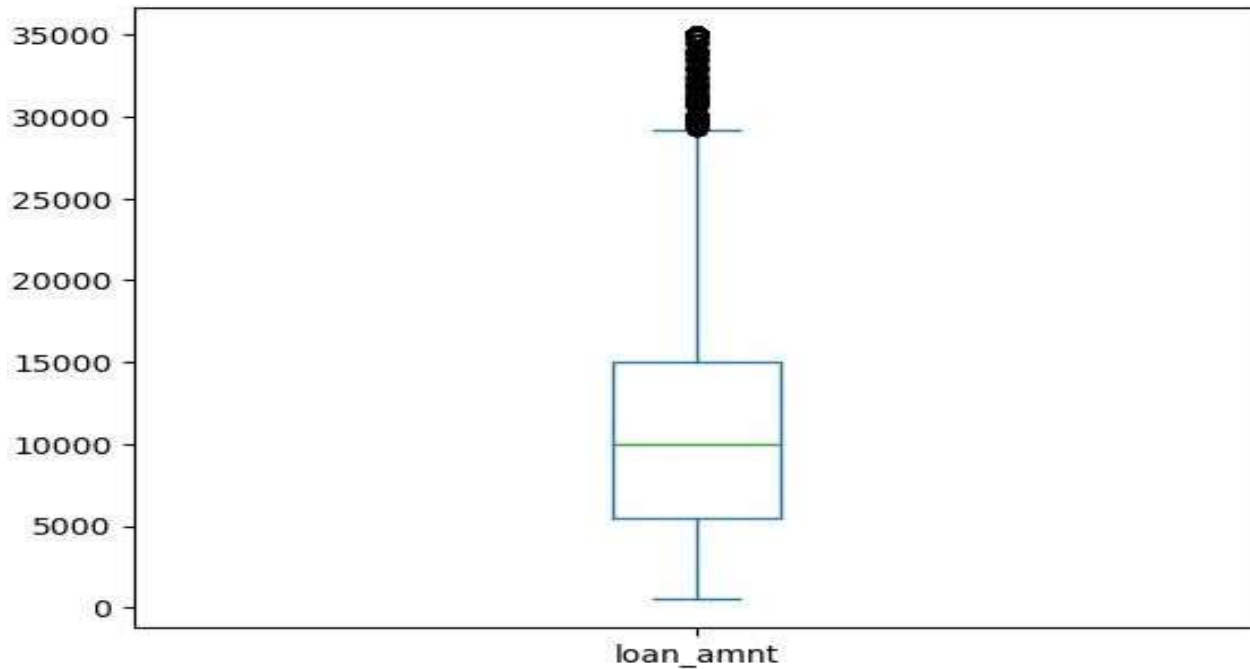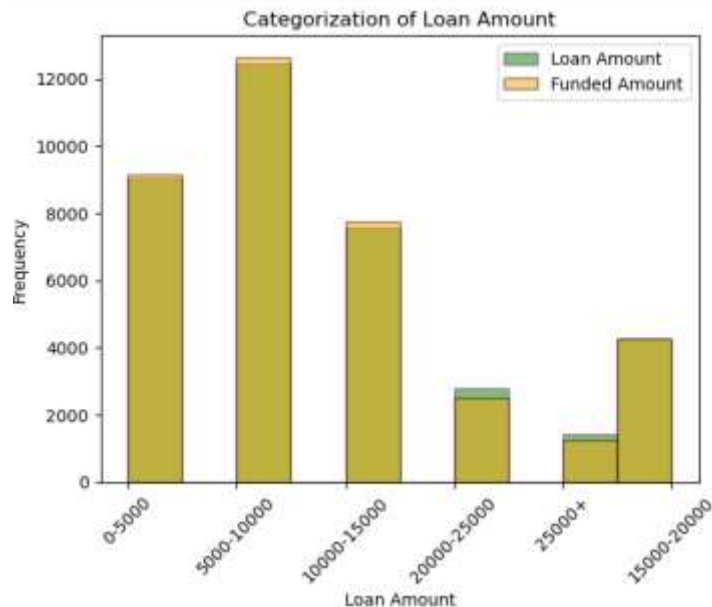
Loan Amount:

On processing the outliers in a separate plot we see that removing outliers will lead to missing out on insights that we will be able to derive against other variables
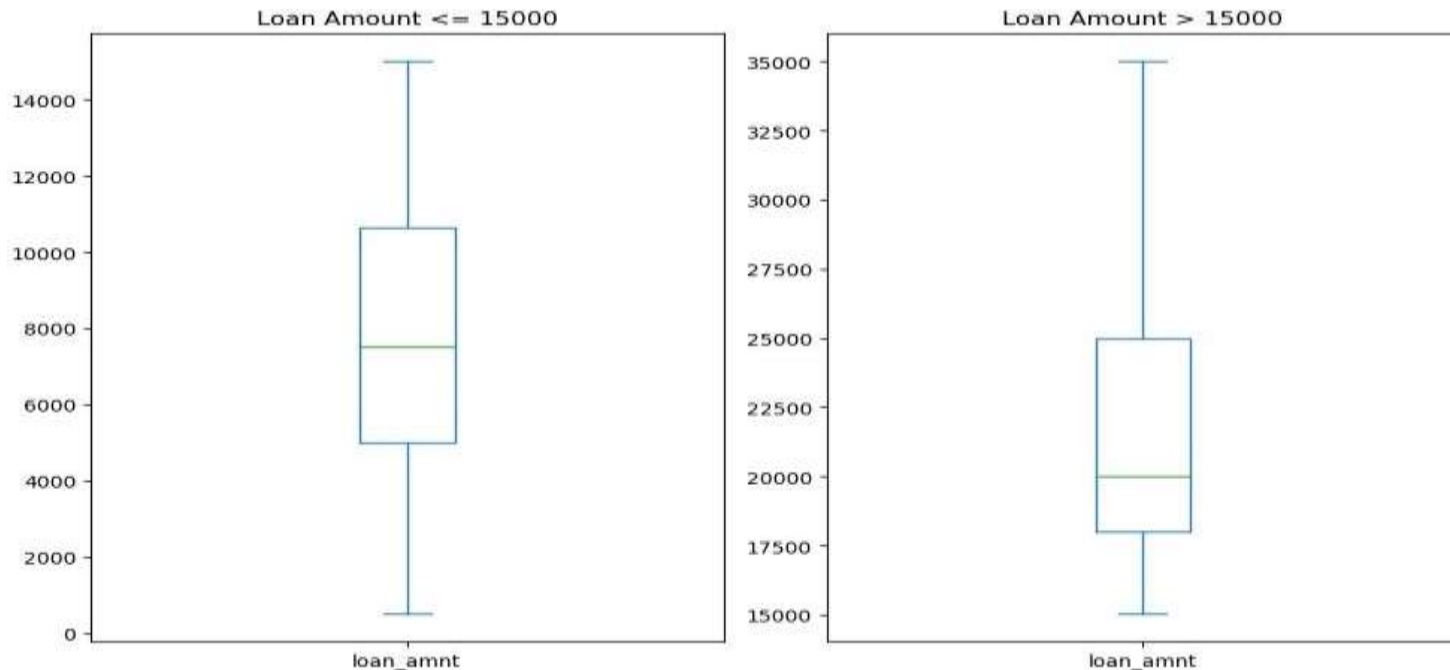
Loan Amount requested vs Loan Amount funded

We discovered that loan amounts had a very high density around the 20000 mark, with the largest number of loans in the 8000-10000 range. The pattern also indicates that in most situations, the entire loan amount requested was funded. We do see a few examples on the upper end of the loan magnitude spectrum when the funded amount is less than the loan amount.
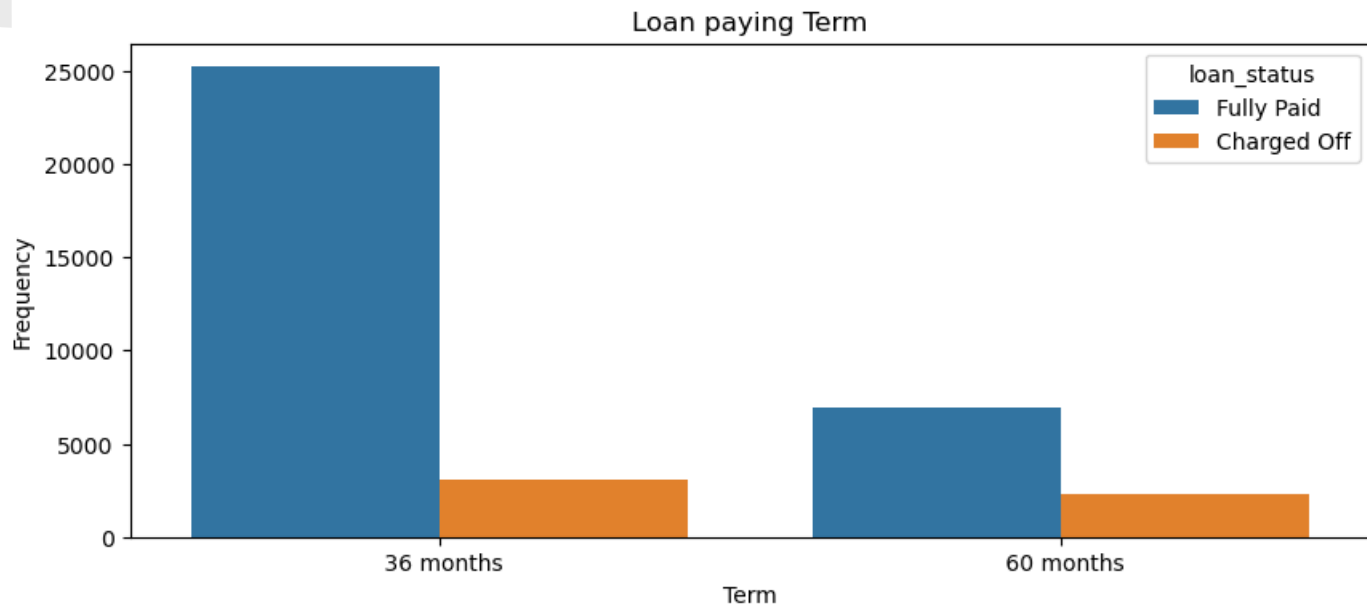
Ratio of Total Loan Amount Paid-Off to Loan Amount Defaulted
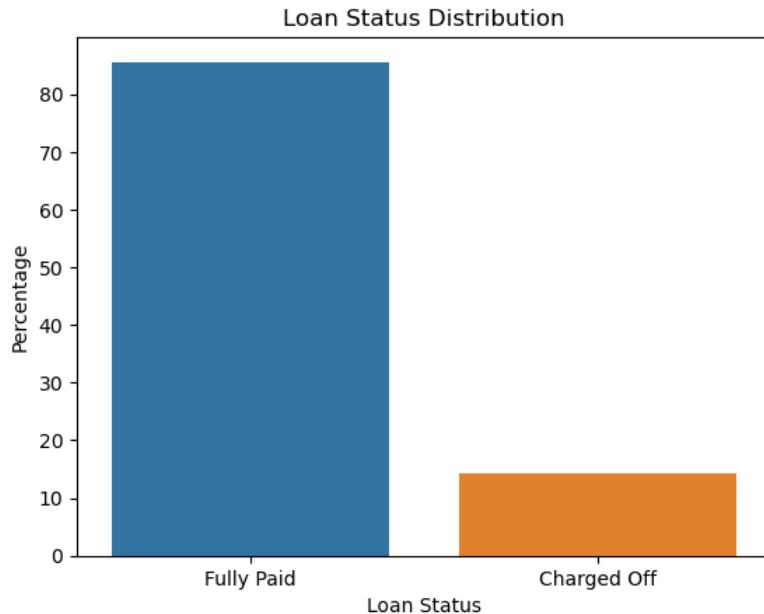
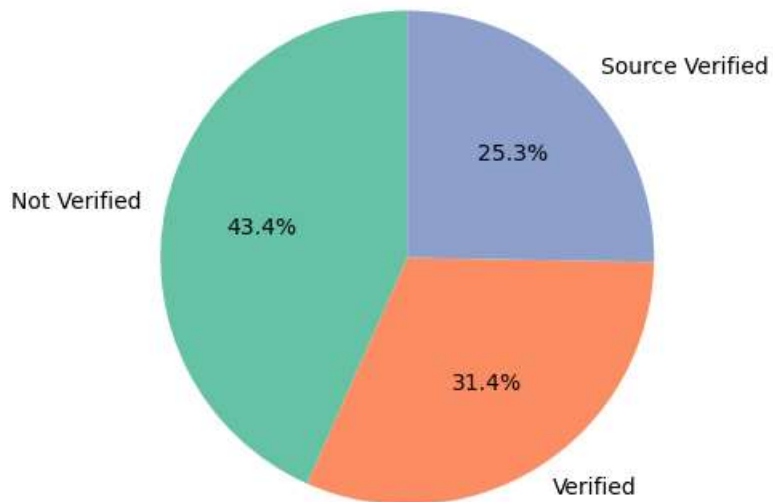**we can say that 85.7% have fully paid whereas 14.3% are charged off.**

Ratio of Total Loan Amount Defaulted per Term

Most of the loan taken for 36 months as compared to 60 months and Charged off loan status is high for 36 months.
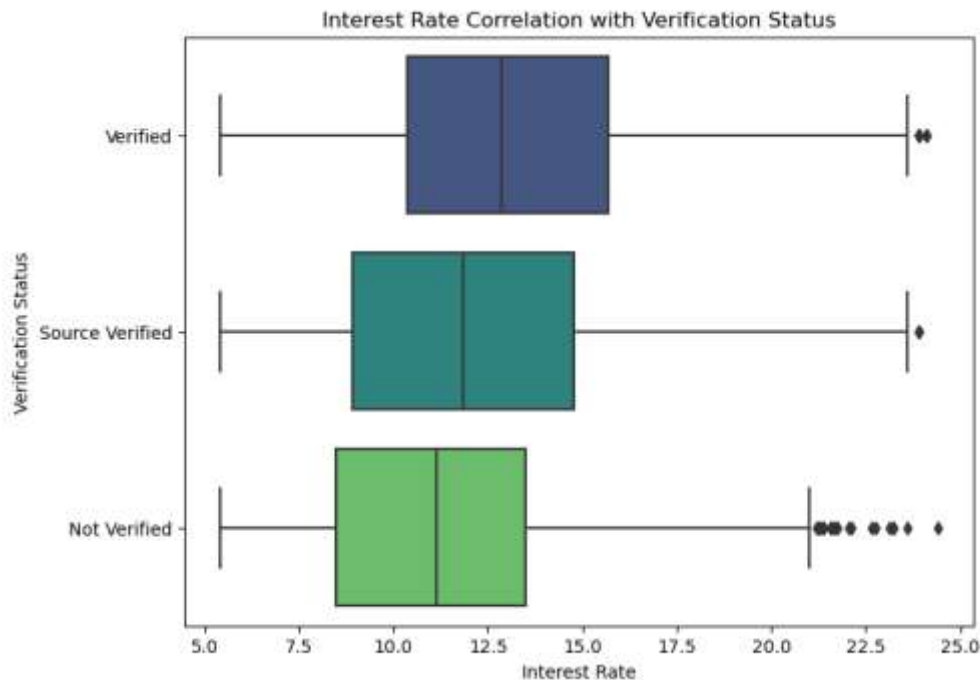
# Ratio of Verification Status

Verification Status Distribution

There are a more Not-Verified borrowers than Verified or Source Verified borrowers

Interest Rate Correlation with Verification Status

Interest Rate Correlation with Verification Status

There is no evidence that Not Verified Borrowers are charged higher interest

Loan Amount Vs Loan Status and Grade
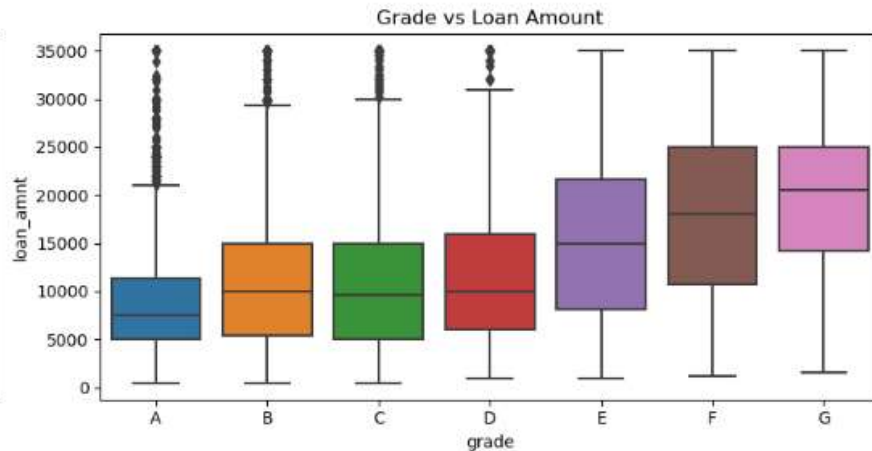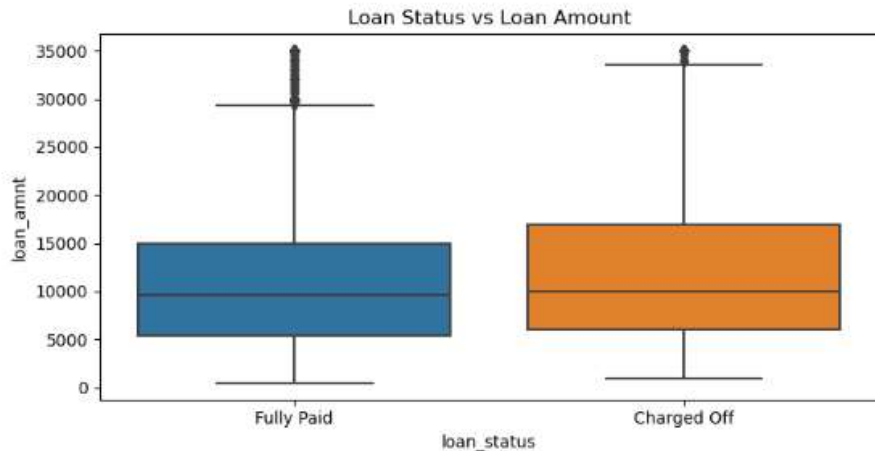
1. Charged off were having higher loan amount than fully paid.
2. In Grade vs loan amount we can say that, grade F & G are having max amount of loan. As the grade decreases amount of loan is increasing.
3. From this we can say that higher the grade more is the risk of default.

Interest Rate Vs loan status and Grade
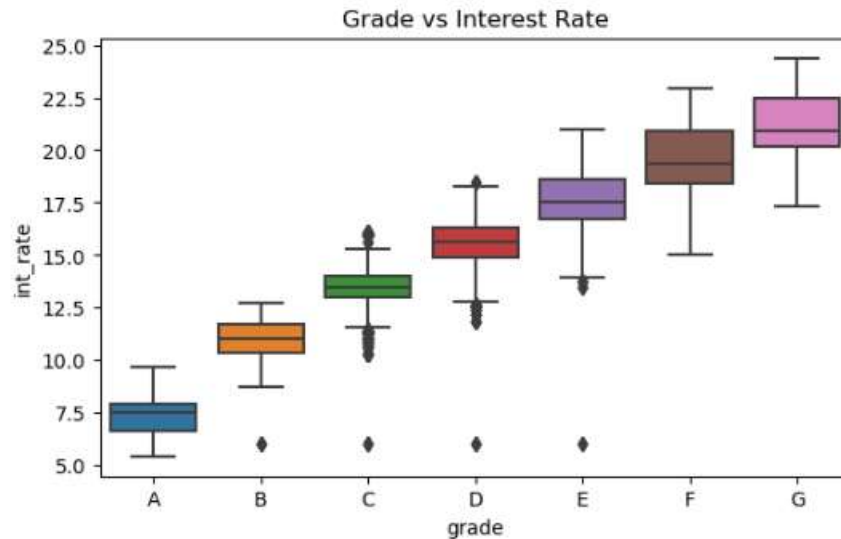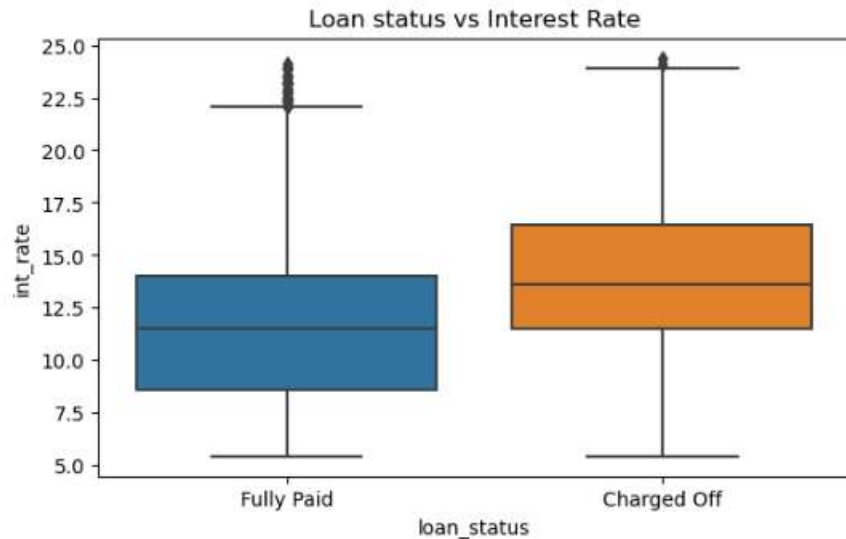
1. In loan status vs interest rate variable we can say that, the borrowers who has high interest rate are mostly defaulters.
2. In grade vs interest rate, as the grade decreases, interest rate is also increasing

Revolving Utilization Rate vs Loan Status
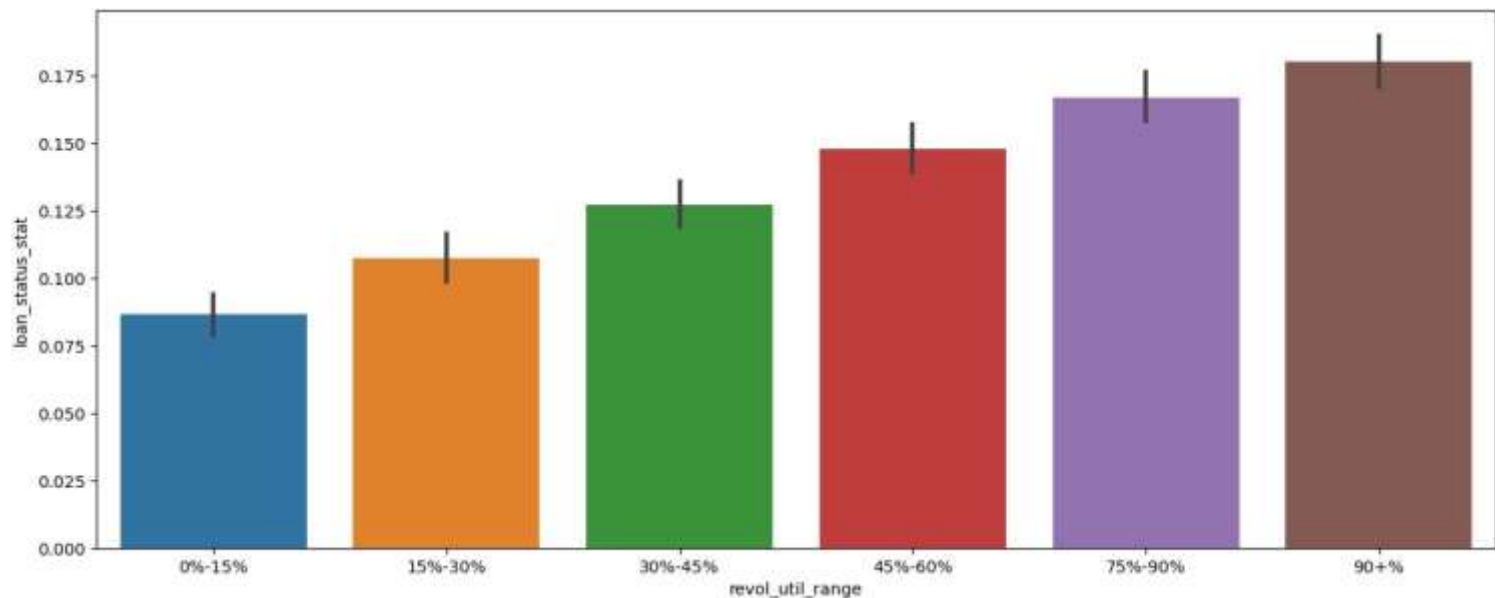
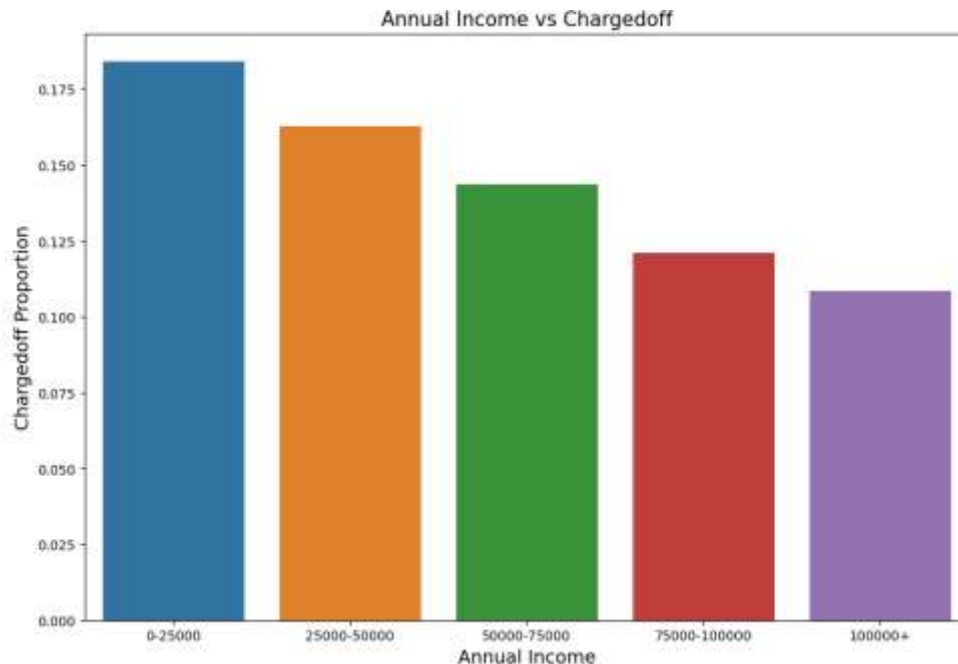**Higher the Revolving Utilization rate, higher the chances of default**

Public derogatory records vs Loan Status
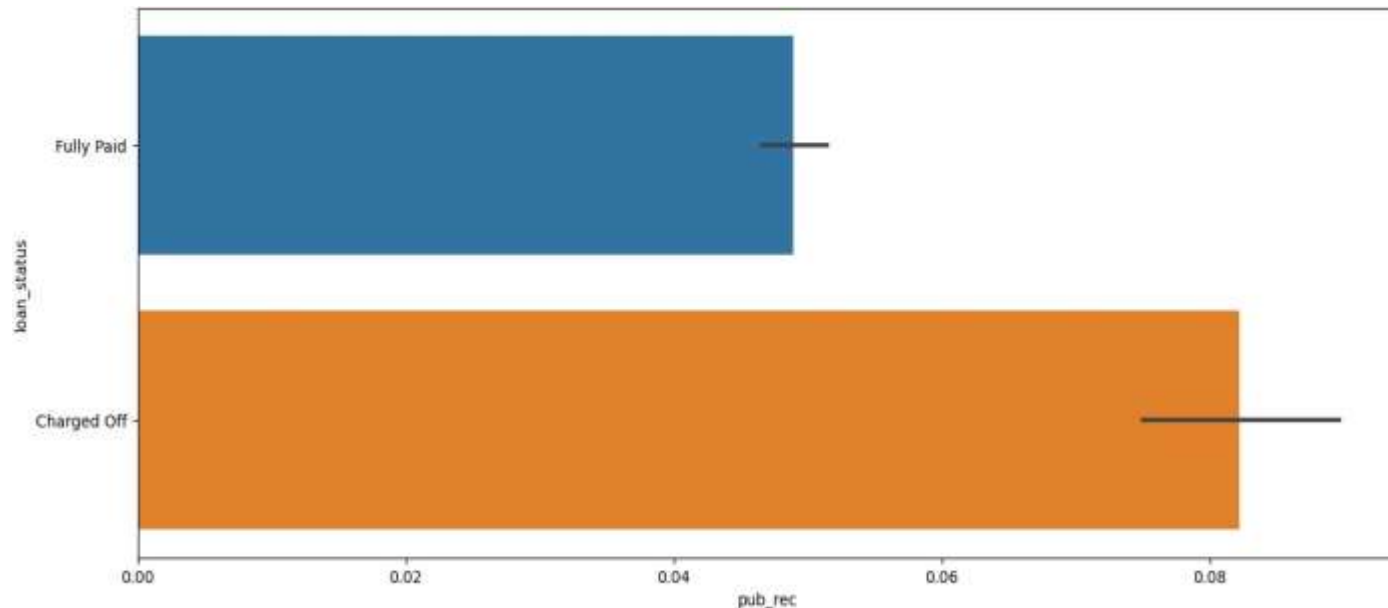
Borrowers who have existing public derogatory records have higher chances of default
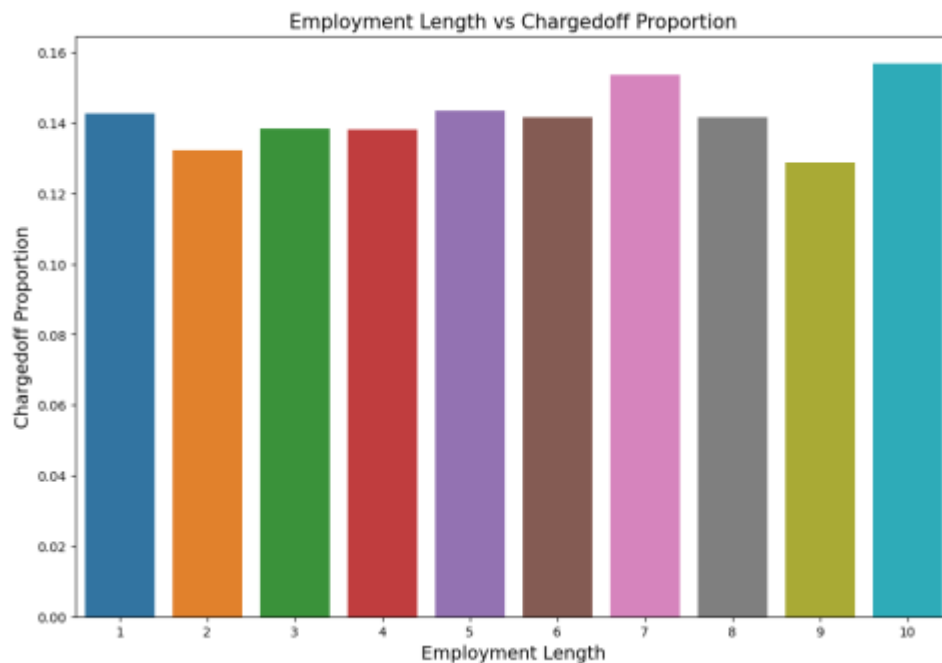
**Annual Income vs Charged off**

Annual Income vs Chargedoff

As annual income is increasing charged off proportion is decreasing. Highest charged off proportion are in the range 0 to 25k annual income

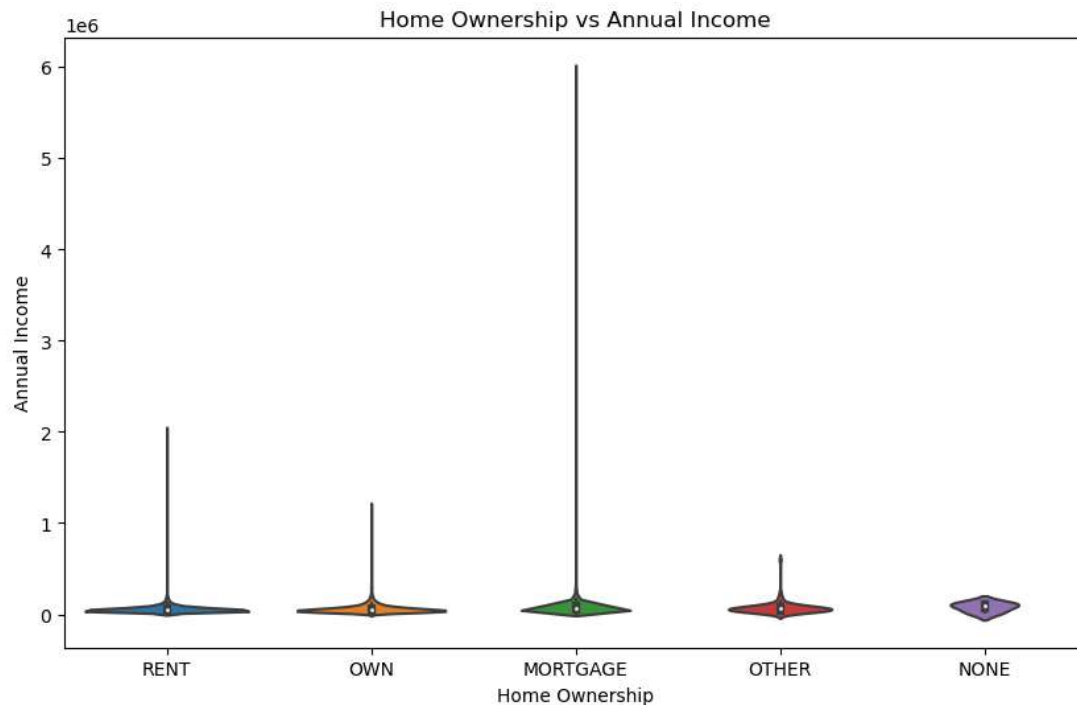**Employment Length vs Charged off Proportion**

The borrowers whose employee length is less than 1 year, 1 year and 10+ year are mostly getting charged off.
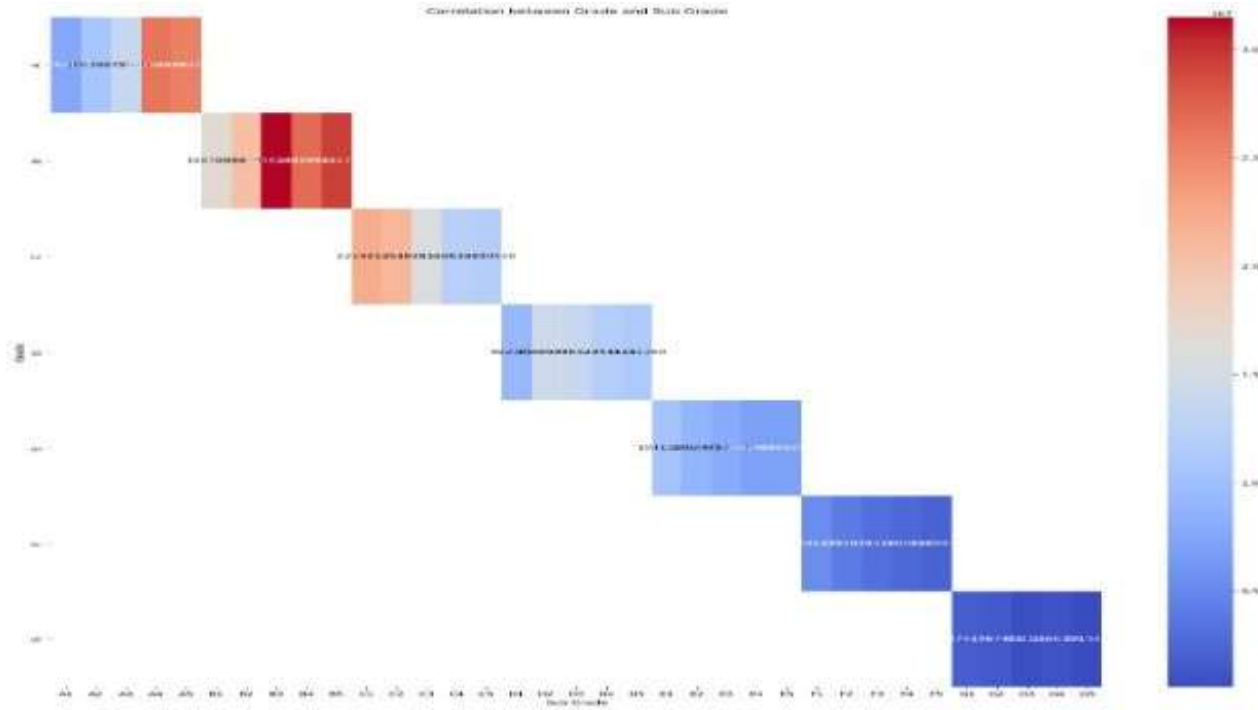
Annual Income vs Loan Default
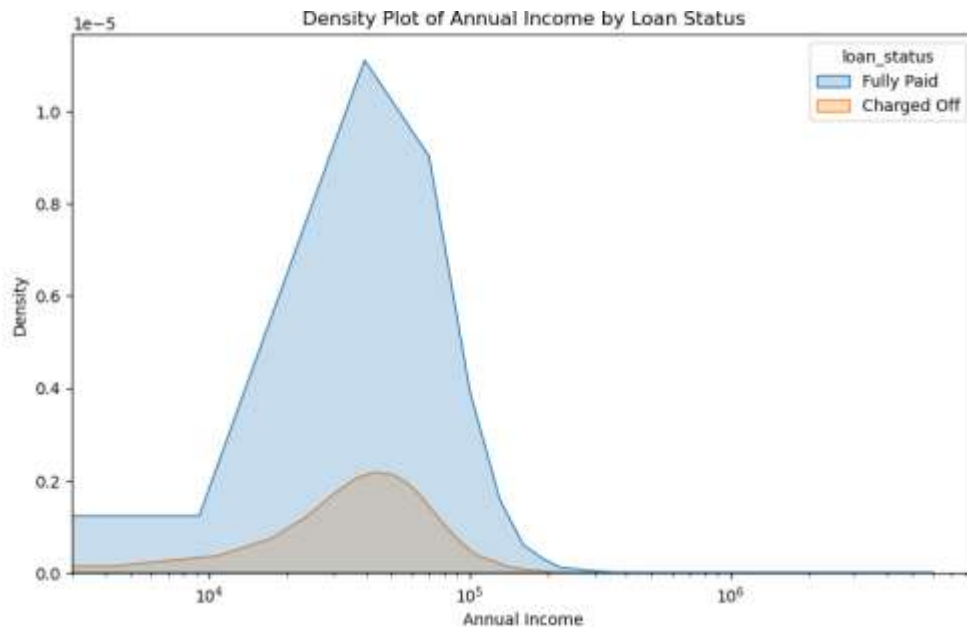
The higher the annual income, lesser the default chances

Home Ownership vs Annual Income

The borrowers who have the status as Mortgage are having high annual income than others.

Grade and Subgrade and their relation to Loan amount/interest rate

Grades B, A, C, D have the highest amount of loans and in particular subgrade B3, B5, B4 and B2 constituting the most loan amounts

From above analysis we can say that, there is more probability of defaulting when

- ➢ High interest rates
- ➢ The grade decreases, interest rate is also increasing
- ➢ Low annual income of borrower
- ➢ Loans of Grade B, A, C, and B
- ➢ High installments
- ➢ Previous records of bankruptcies
- ➢ Higher the Revolving Utilization rate, higher the chances of default
- ➢ Higher the loan amount, higher the chances of default

Thank you