Lending Club Case Study

Group Members:

Sumit Baghel

Akshay Jain

Abstract

- Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.

- They offer lending services by connecting borrowers with investors through an online marketplace.

- The objective of analysis is to use the information about past loan applicants and find whether they 'defaulted' or not.

Business Understanding

Dataset Details:

- The data given in the dataset contains information about past loan applicants and whether they 'defaulted' or not. Data has details regarding sanctioned loan with loan status as is Fully Paid, Current and Charged-Off.
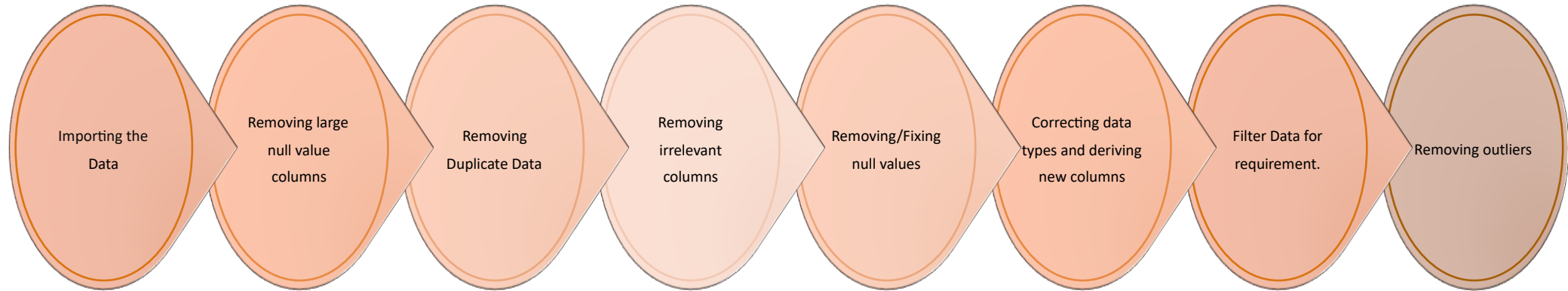
Data Analysis   Data Analysis   Data Analysis

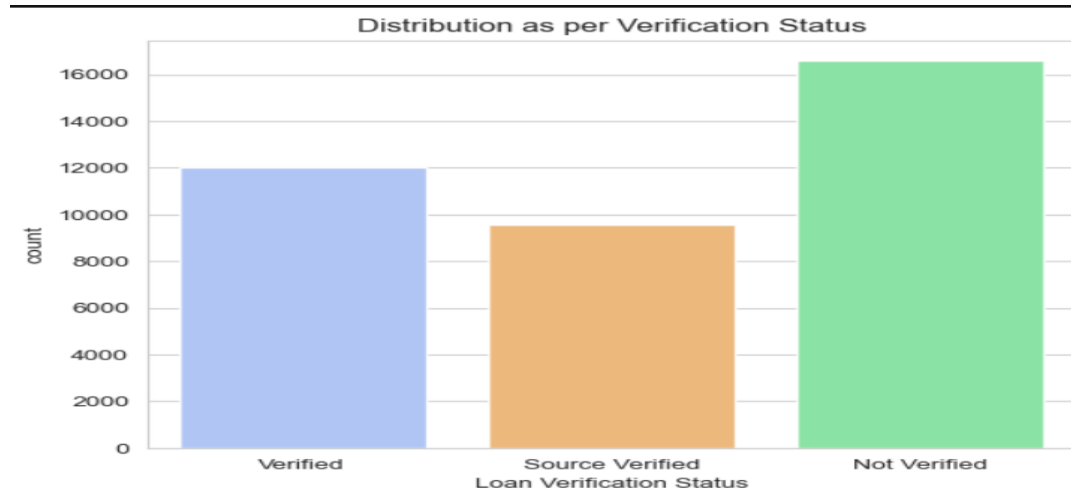| Data Cleaning	Data Understanding	Univariate Analysis	Segmented Univariate Analysis	Bivariate Analysis	Recommendations
Removing the null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.	Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses	Analysing each column, plotting the distributions of each column.	Analysing the continuous data columns with respect to the categorical column	Analysing the two variable behaviour like term and loan status with respect to loan amount.	Analysing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

Data Cleaning and Data Preparation

- Below are the steps followed for Data Cleaning and preparation

Importing the Data

Removing large null value columns

Removing Duplicate Data

Removing irrelevant columns

Removing/Fixing null values

Correcting data types and deriving new columns

Filter Data for requirement.

Removing outliers

Data analysis

Univariate analysis on Laon Status and verification status
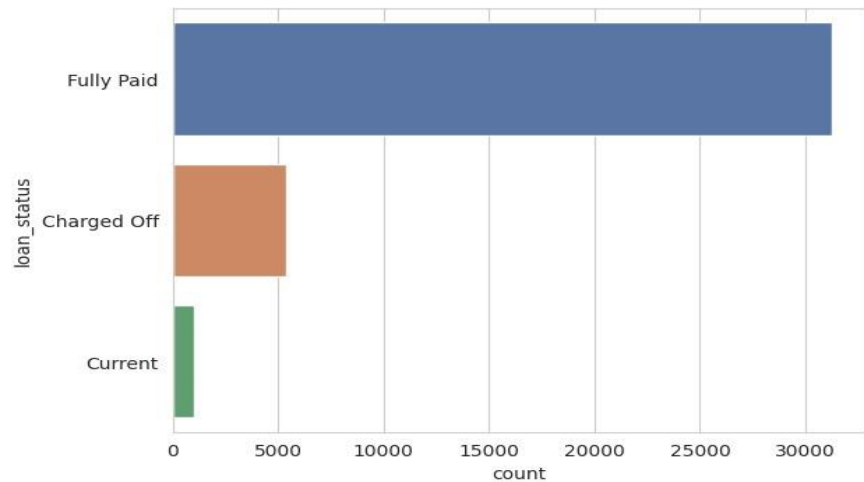

- As we can see the number of charged off loan is much smaller compared to total loan entries count
- Most of the loans are not verified that also can be one of the reasons for defaulting the loan

# Univariate analysis as Per Loan Term and home Ownership
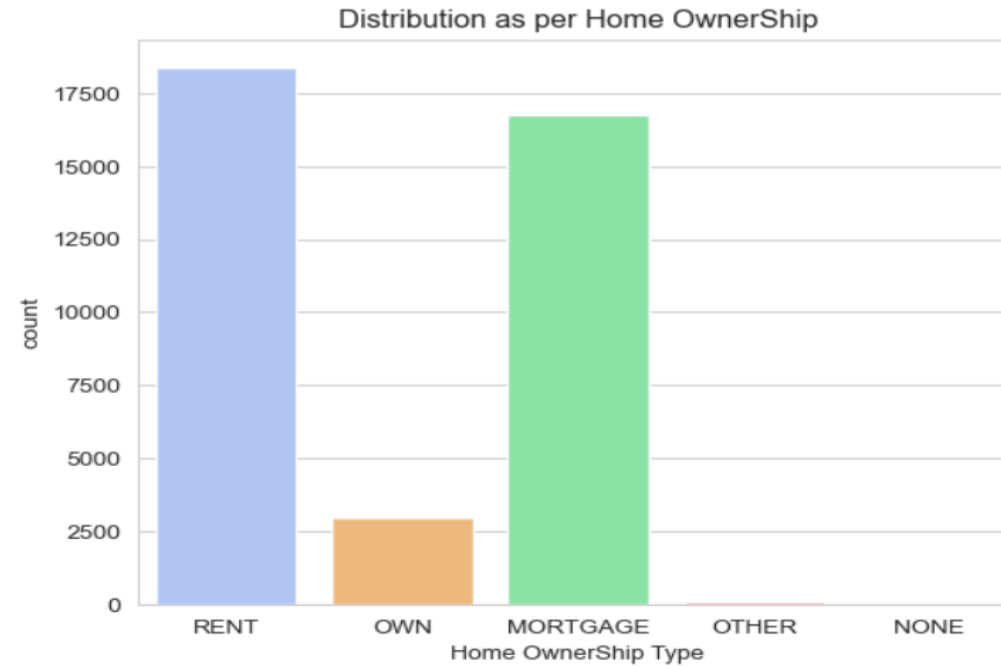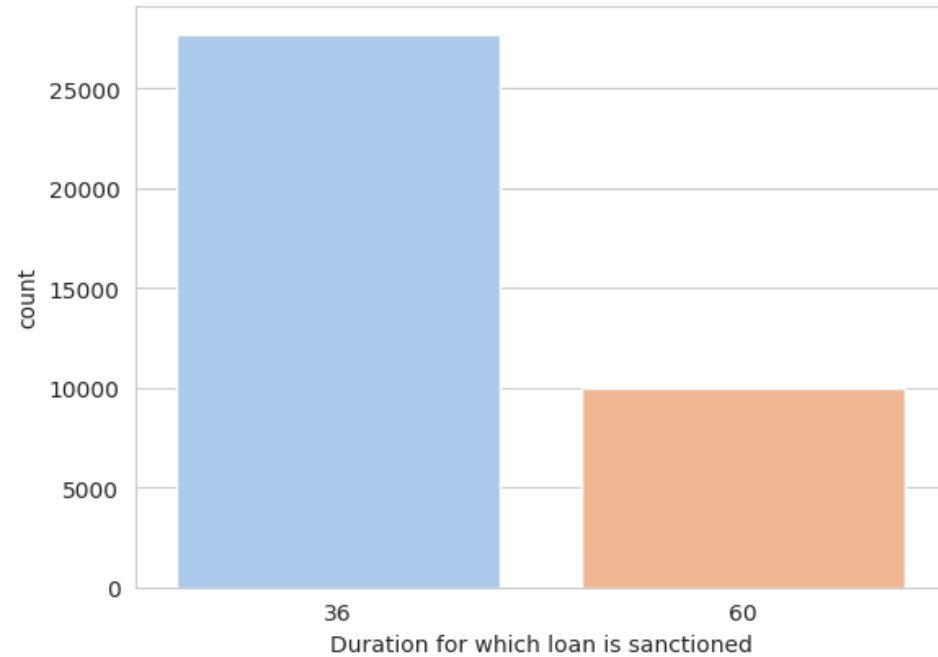
- Majority of the loan are approved for short duration of term i.e. 36 months

- Most of the borrowers who has taken a loan are living in Rented or mortgage property.

Univariate analysis as Per Loan Grades
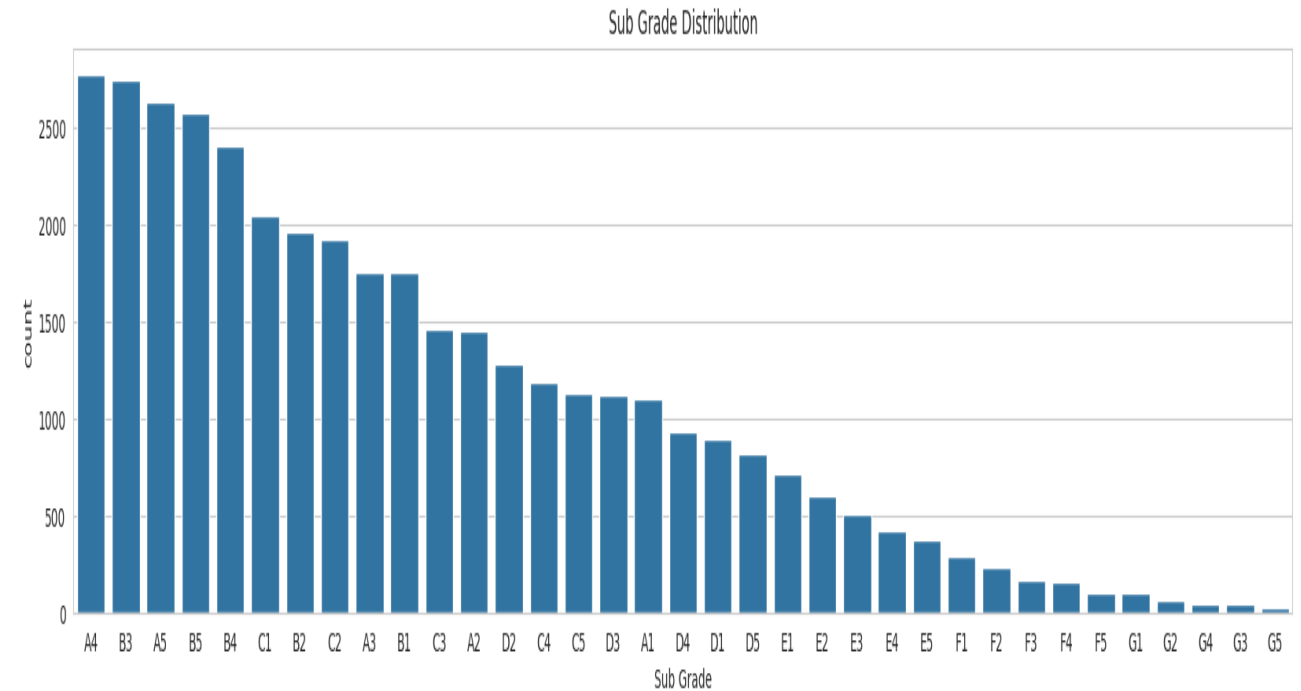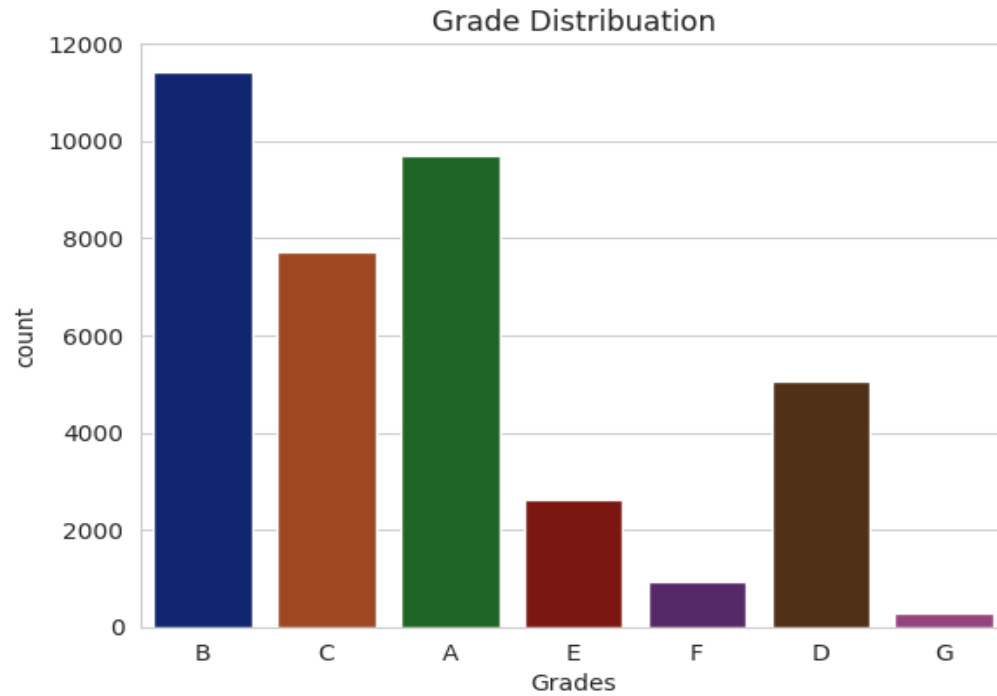
- Most of the loan is approved for Grade B followed by A

- This is also One of the important category To identify the defaulters

Univariate analyses on Annual Income and Purpose
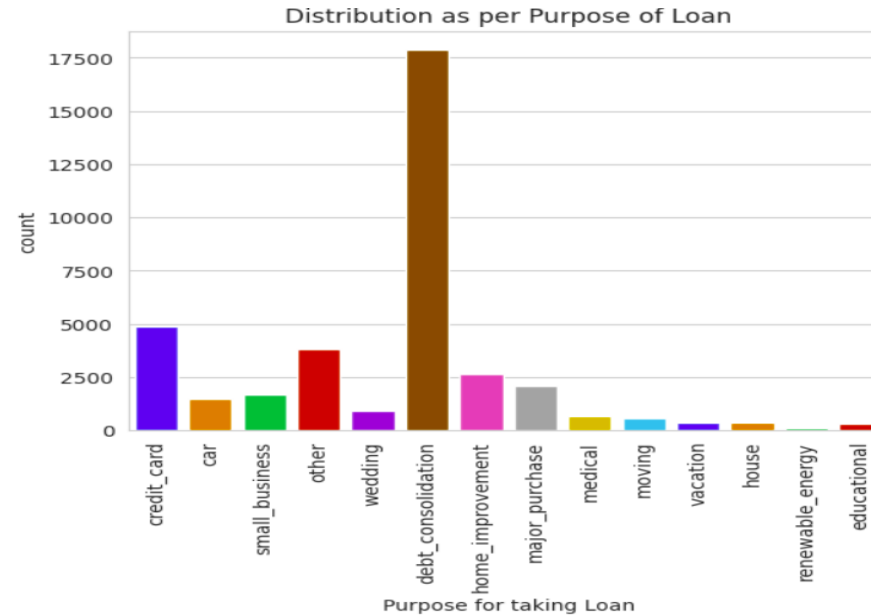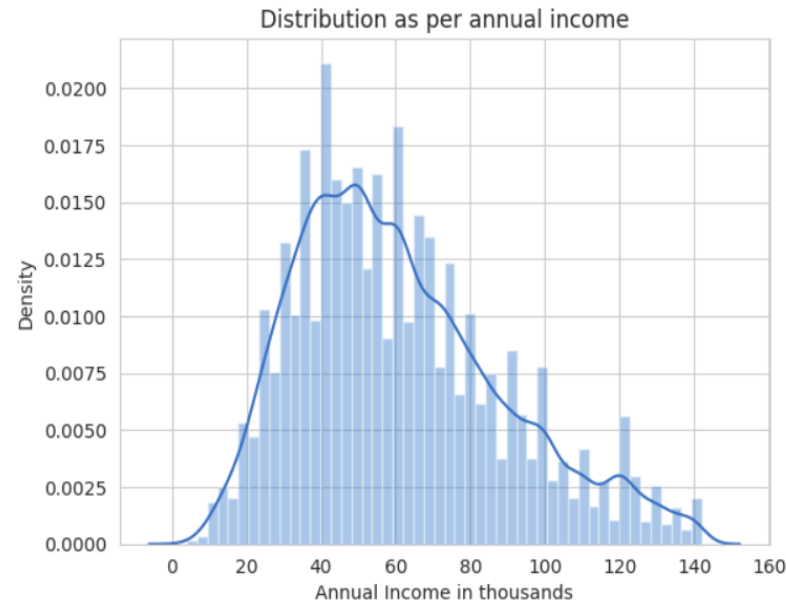
o The Majority of clients have low annual income compared to rest and income lower than 50k has higher chance of defaulting.

o Loans are taken mostly for debt consolidation followed by credit card payment. Whereas the debt consolidation has highest fully paid loan but also has highest defaulted loans as well.

Bivariate Analysis and Segmented Univariate

Bivariate analysis on loan status vs Interest rate


Distibution of Intrest rate For Loan Status
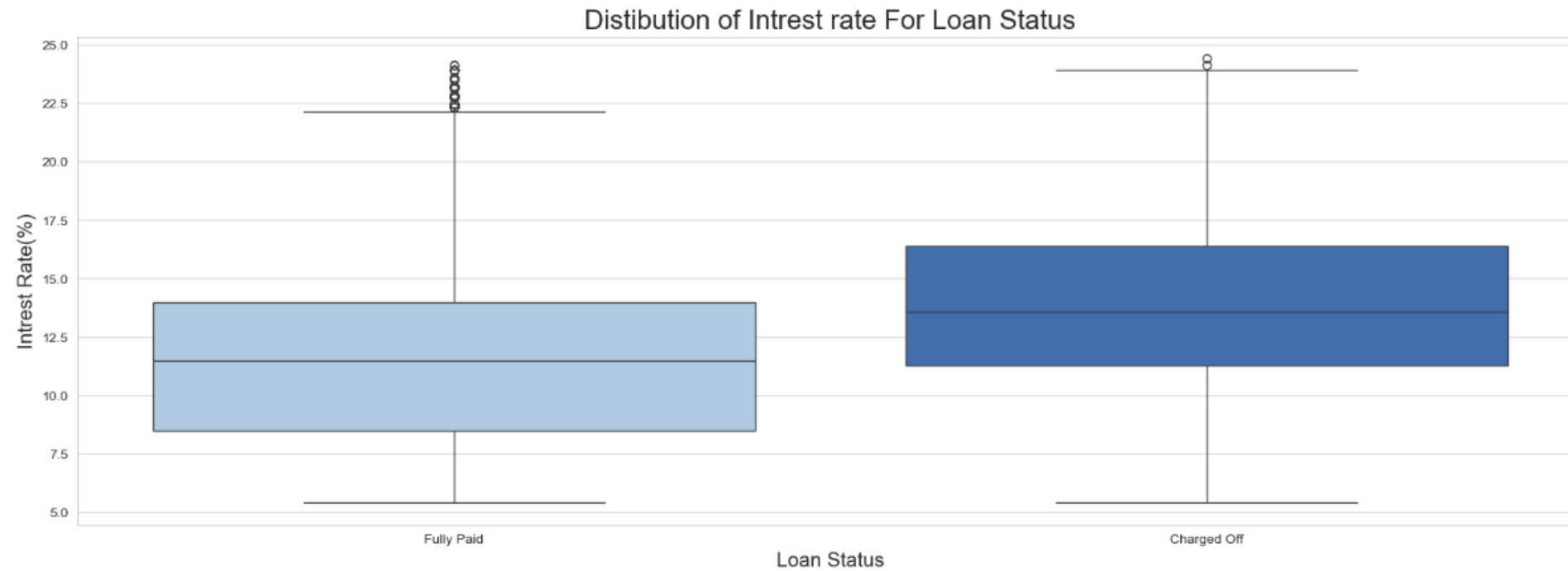
- As we can see the Higher the interest rate Higer the chance that borrower can default the loan
- Same applies for Loan amount as well higher the loan amount, Higher the chance of Defaulter

Bivariate Analysis of Loan Status vs Annual Income


Distribution of Annual Income For Loan Status
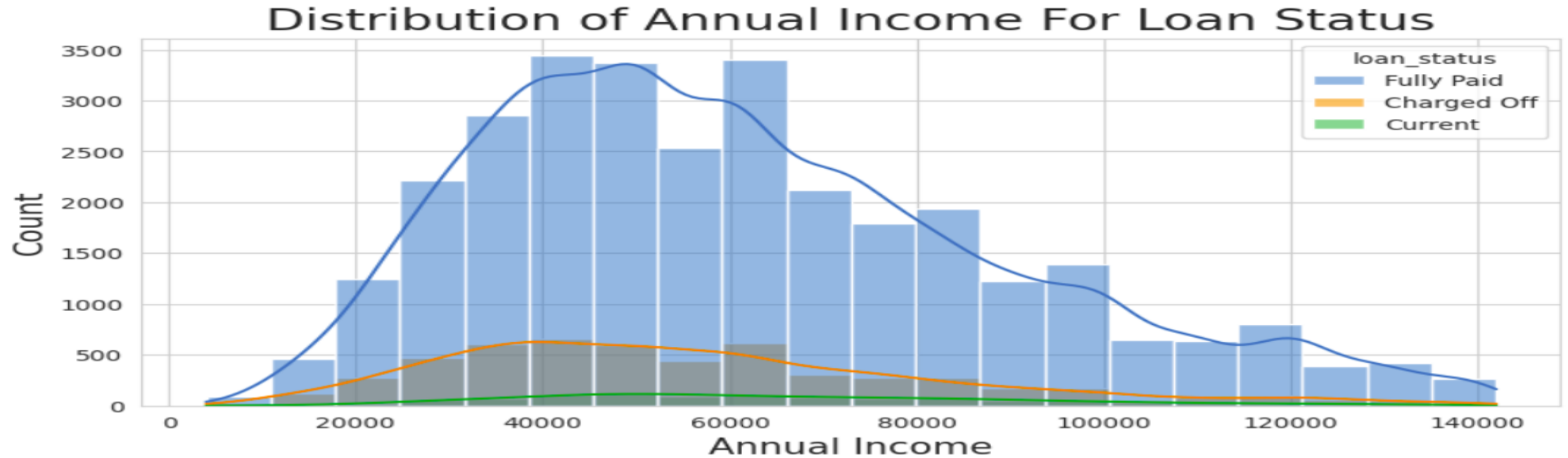
- Graph shows that lower the income of Borrower Higher the chance of defaulting the loan
- We can see there is clear spike between annual income below 60-65 k

Bivariate Analysis of Loan Status and Loan Amount vs Home Ownership
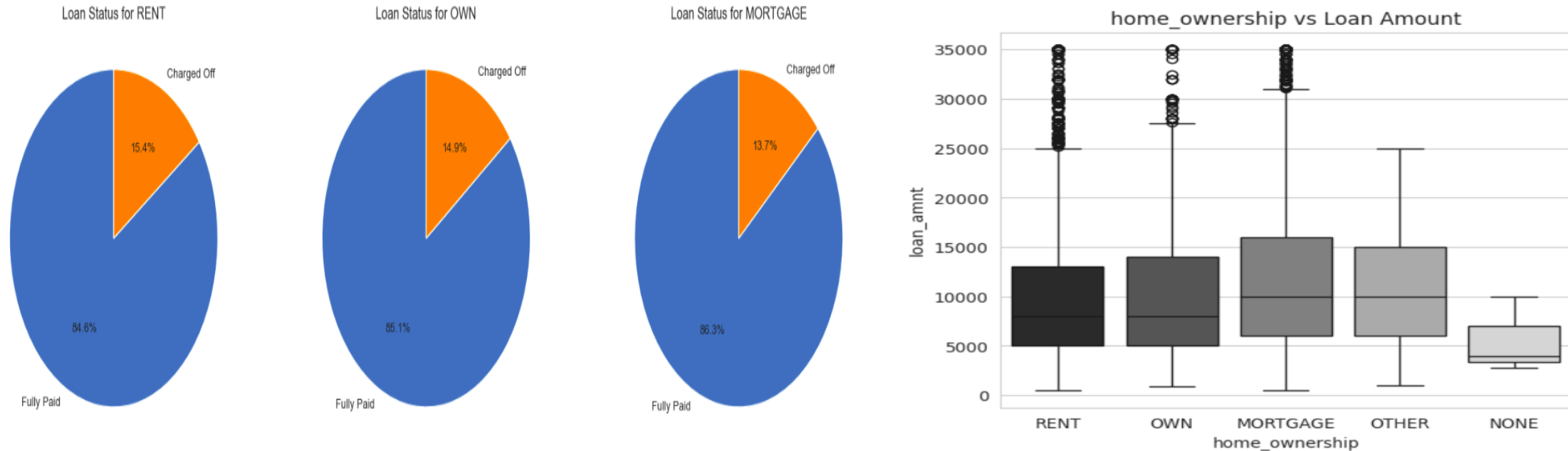
- As the Pie chart shows that borrower with a Rental home have the higher chance of defaulting a loan.
- Where as people with mortgage property taking higher amount of loan which is most likely to get defaulted

Bivariate Analysis of Interest rate vs term and Grade
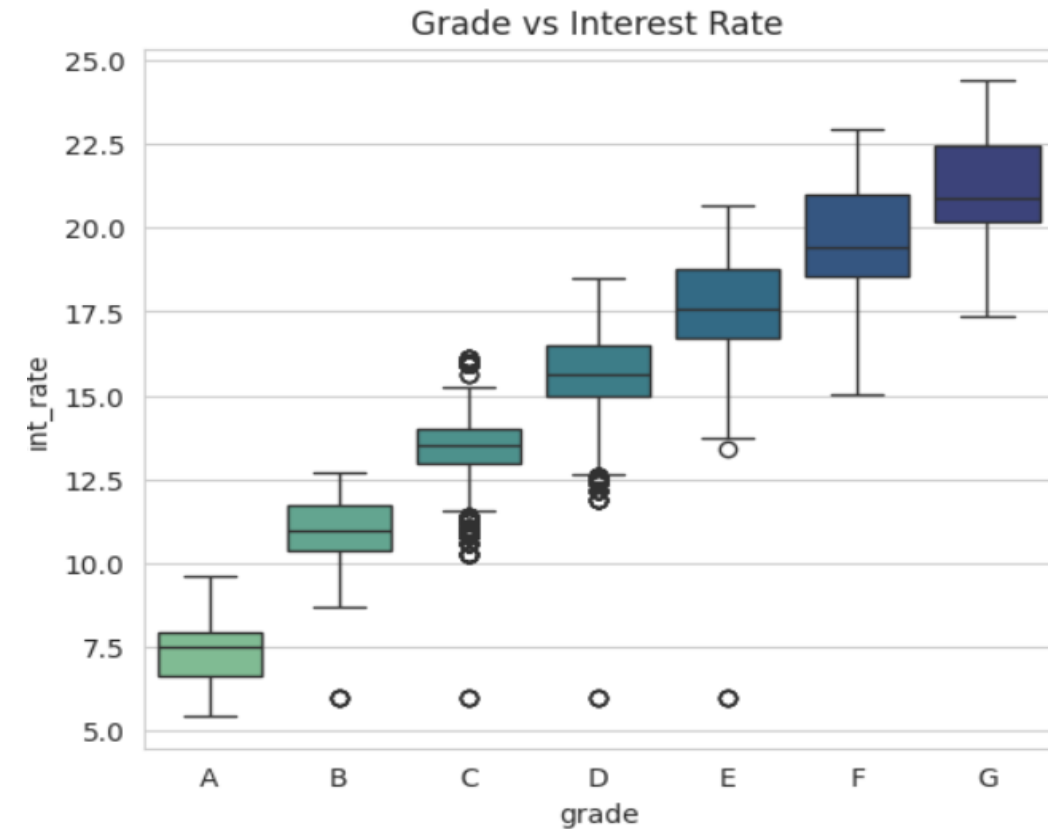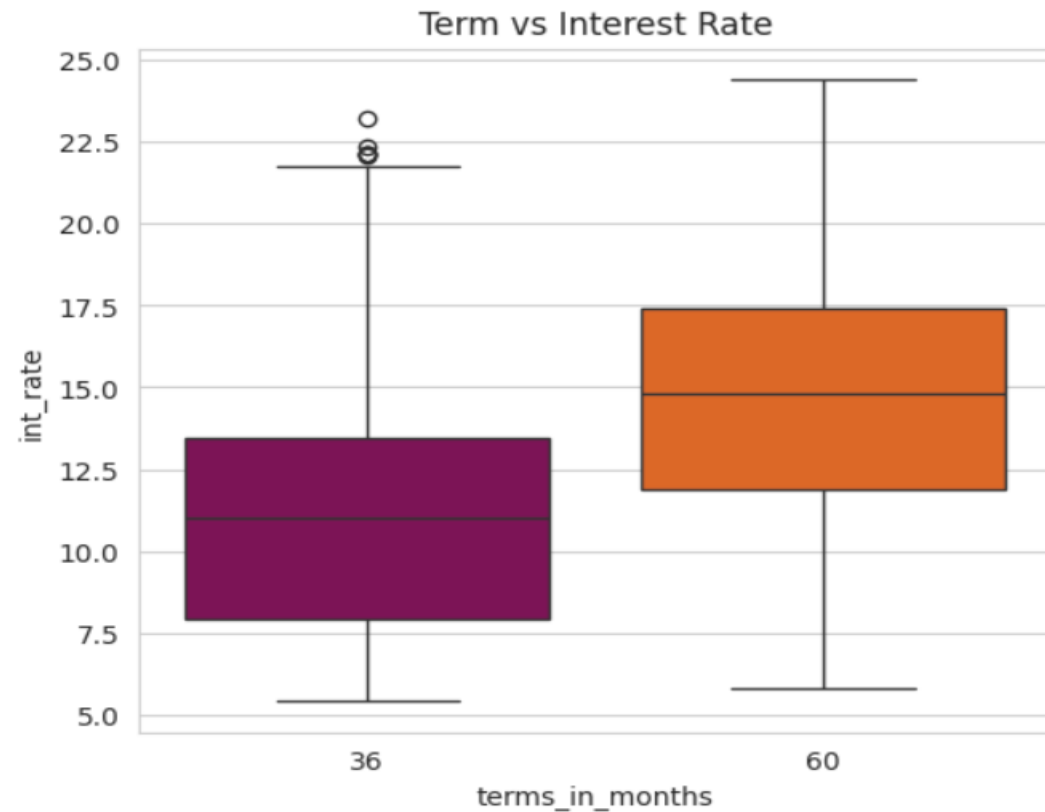
- As Observed Longer the Term Higer the Interest Rate
- Lower the Grade Higher the interest rate
- And as we said Earlier Higher the Interest rate higher the chance of defaulting the loan

Conclusions

- Lending club should reduce the high interest loans for 60 months tenure, they are prone to loan default.
- Lending club should examine more information from borrowers before issuing loans to Low grade (G to A).
- Small business loans are defaulted more. Lending club should stop/reduce issuing the loans to them.
- Borrowers with mortgage home ownership are taking higher loans and defaulting the approved loans. Lending club should stop giving loans to this category when loan amount requested is more than 12000.
- People with a greater number of public derogatory records are having more chance of filing a bankruptcy. Lending club should make sure there are no public derogatory records for borrower