LENDING CLUB CASE STUDY

AUTHORS/MEMBERS:

JITENDER KUMAR

BHARATHREDDY REDDIVARI

INDEX

PROBLEM STATEMENT

DIFFERENT PYTHON LIBRARIES USED

ANALYSIS DEFINITIONS

UNIVARIATE ANALYSIS

BIVARIATE ANALYSIS

RECOMMENDATIONS

PROBLEM STATEMENT

- Company

Lending Consumer Finance is one of the largest company which specializes in lending various types of loans to urban customers

- Problem Statement

With the provided dataset, the company has to make a decision for loan approval based on the applicant's profile.

There are two risks associated:

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, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company

DIFFERENT PYTHON LIBRARIES USED

Pandas	Numpy	Matplotlib	Seaborn
• Crosstab		• Heatmaps • Bar Plots • Histograms • Pie Chart	• Count Plot

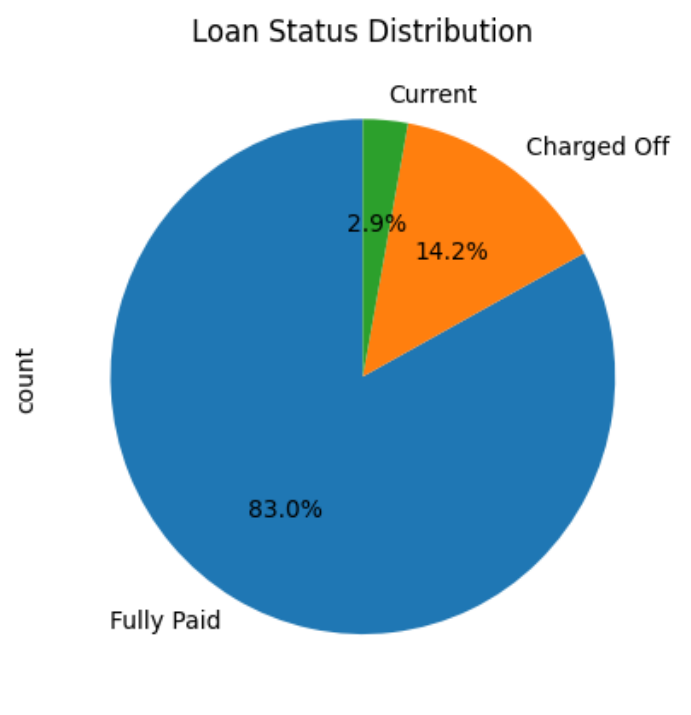
ANALYSIS DEFINITIONS

Univariate Vs Bivariate

Univariate analysis is the most basic form of statistical data analysis technique. When the data contains only one variable and doesn't deal with a causes or effect relationships then a Univariate analysis technique is used.

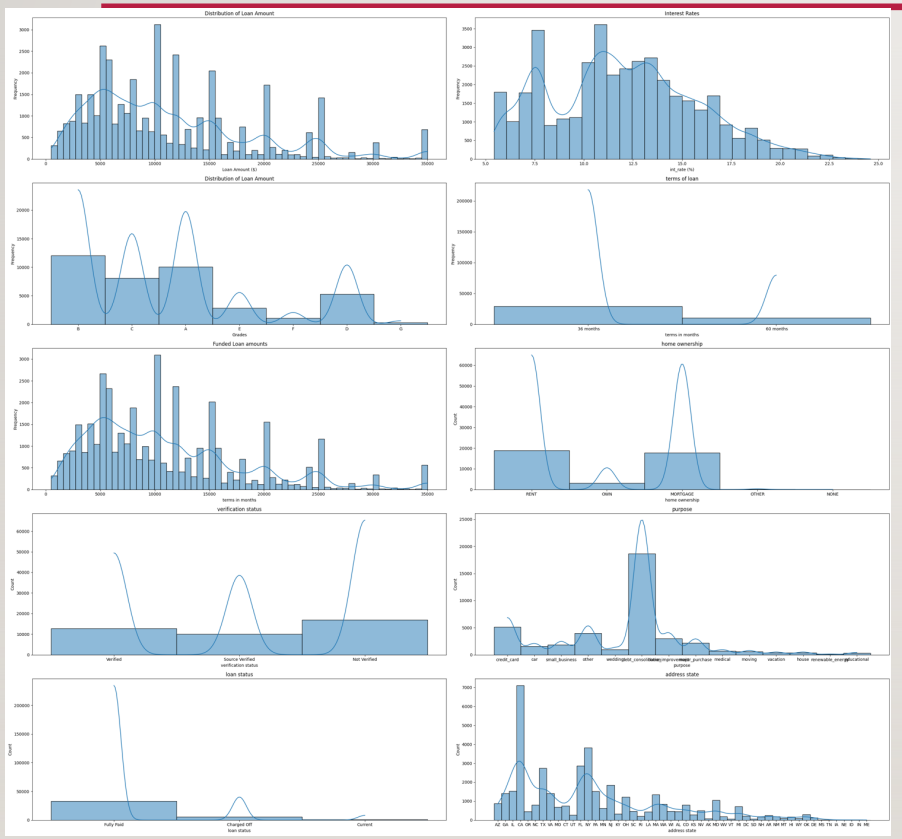
Bivariate analysis is slightly more analytical than Univariate analysis. When the data set contains two variables and researchers aim to undertake comparisons between the two data set then Bivariate analysis is the right type of analysis technique.

UNIVARIATE ANALYSIS



As per the Loan status 83 % of people fully paid the loan and around 14% are charged off which is very significant number
Along with this only 3 % have loans in current state which is a very small number

UNIVARIATE ANALYSIS



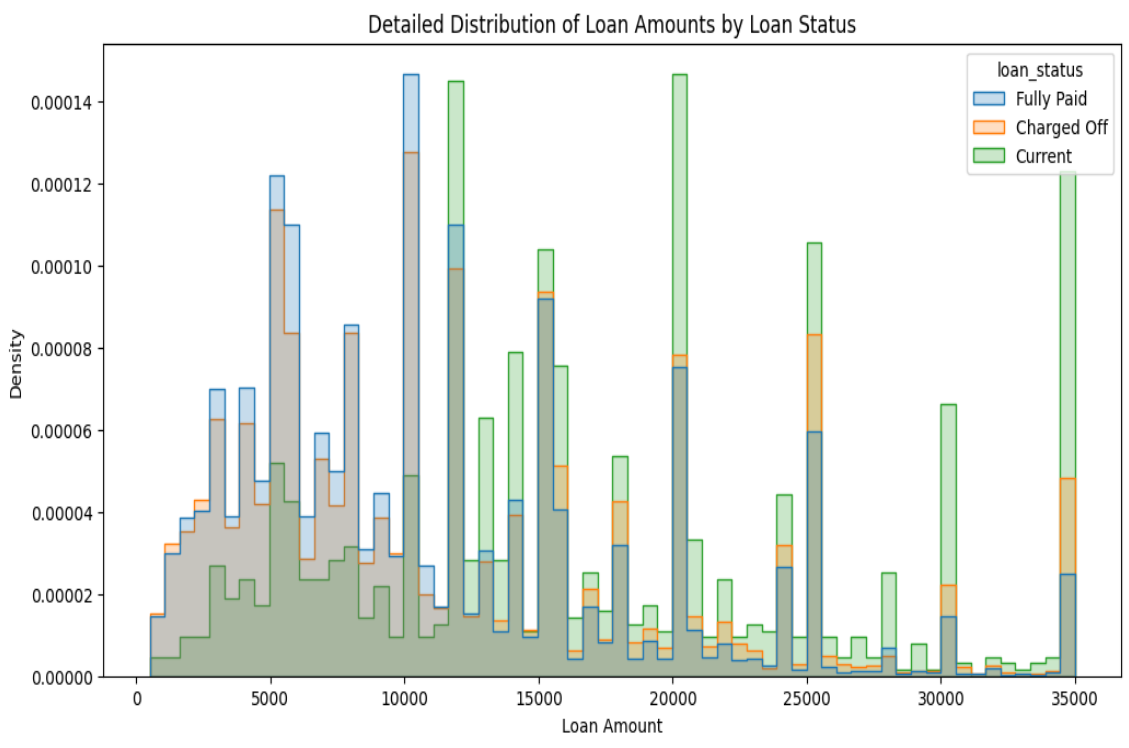
The Chart indicates that for loan amount most of the loans are in range of $5000-$15000, Bigger loans are less or declined as the loan amount increases.

The Chart shows that Interest rate falls in the range of 10% - 20% for most of the loans and there are very few loans with less interest rates and very few with high interest rates for bigger loan amount.

The Chart indicates that the grades B,C and D are most common which are moderate level of credit grades for most loans.

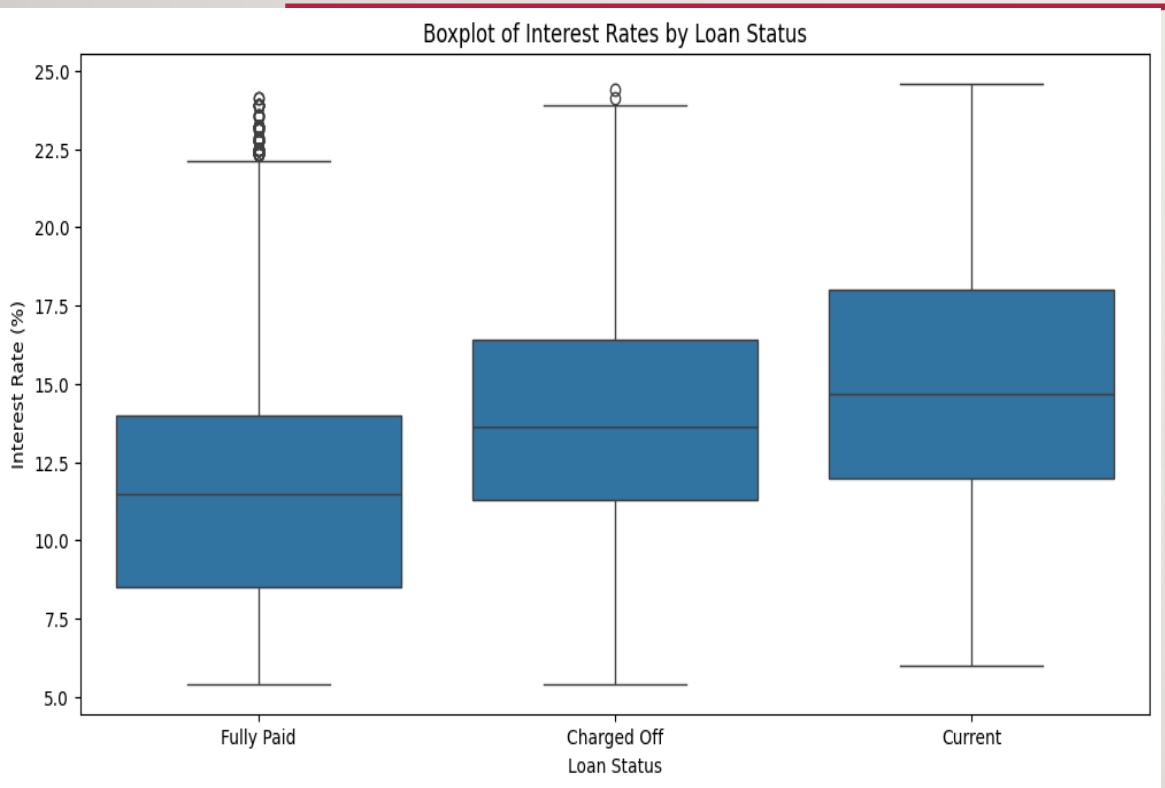
Chart indicates that most of the borrowers either rent or have mortgages and a smaller Number of borrowers own a Home.

BIVARIATE ANALYSIS



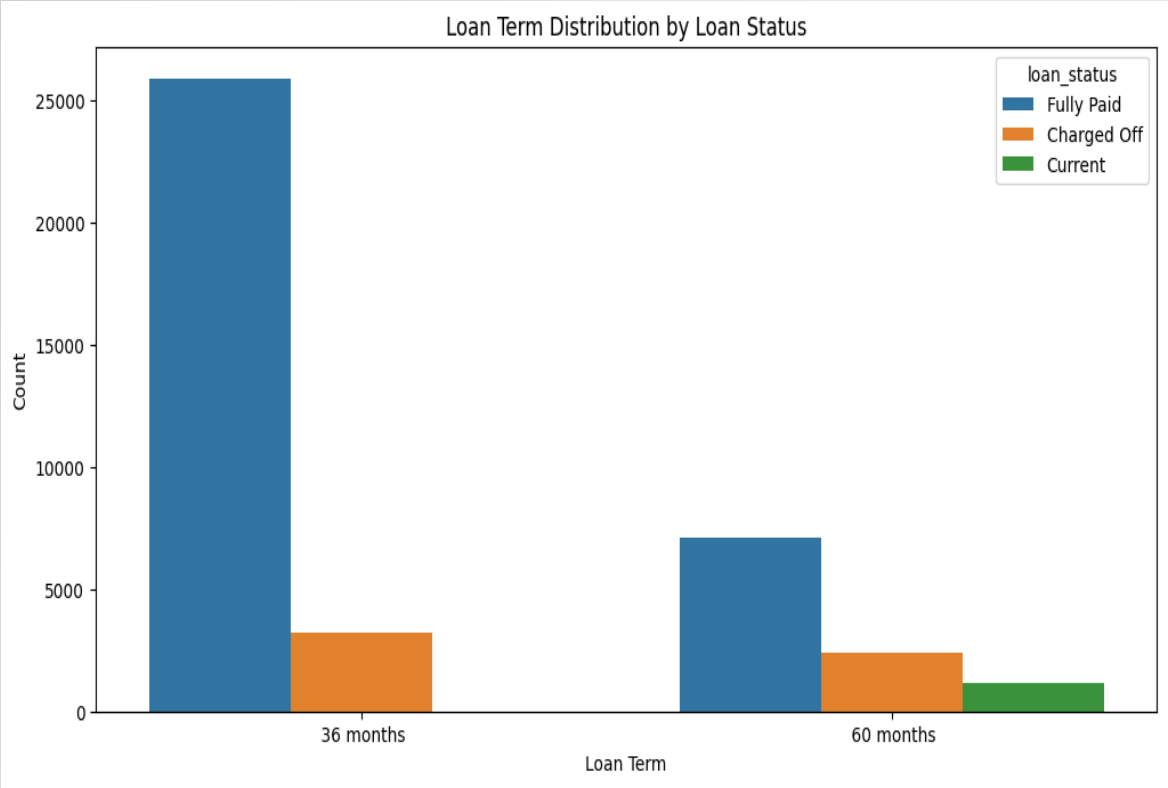
The Histogram shows a clear comparison of distribution Loan amounts for loan status and also observed that the loan amounts with lower amount have more fully paid borrowers while charged off tends to have with higher loan amount.

BIVARIATE ANALYSIS



Boxplot indicates that the interest rates for "Charged off" are high compared to "Fully Paid" loans.
Higher interest rated seems to be the key factor for the loans being "Charged off"

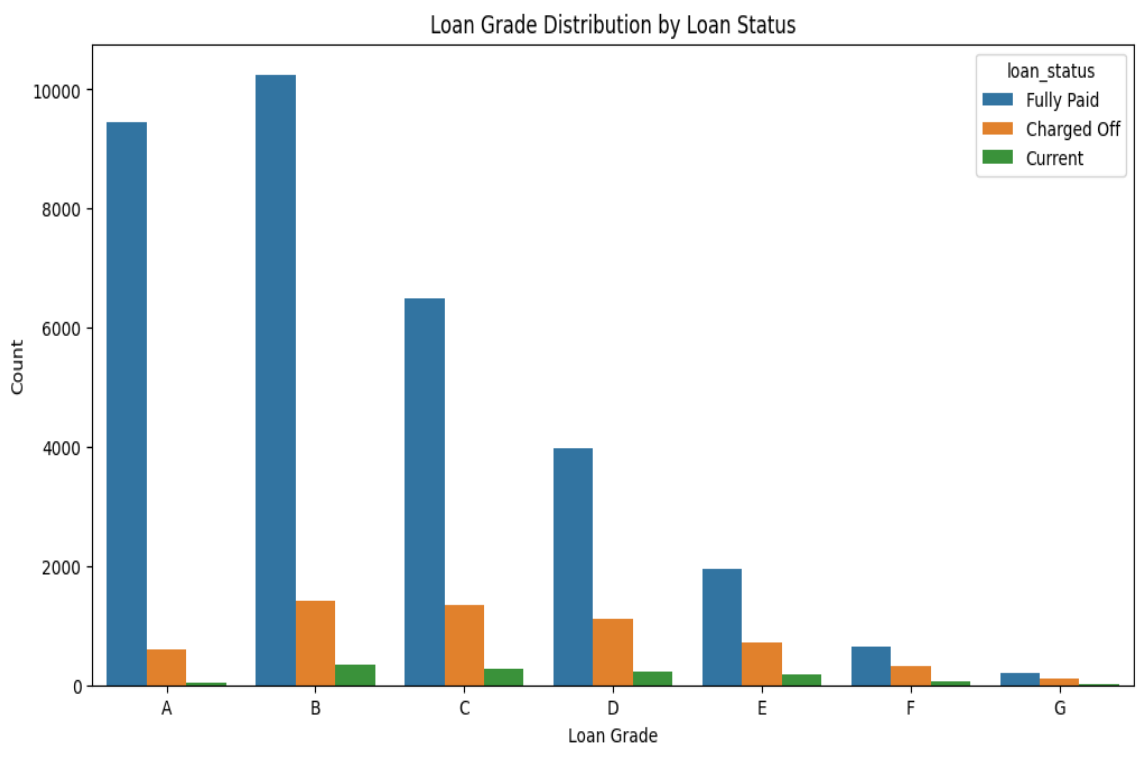
BIVARIATE ANALYSIS



Count plot indicates that the loan terms with 36 months and 60 months are common for "Fully Paid" & "Charged Off" status.

There is also a noticeable pattern where loans with higher loan term have higher count of being "Charged off" compared to "Fully Paid" which suggests longer loan term might be risk factor

BIVARIATE ANALYSIS



The Count plot indicates for Loan grades C,D & E have a higher count of "Charged off" compared to grades A & B.

It is also noticeable that loans with lower grades have high risk of being "Charged off"

RECOMMENDATIONS