upGrad

Executive PG Programme in Machine Learning & AI - August 2023

Lending Club Case Study

Team Members

- **Vishal Kumar Choudhary**
- **VamsiKrishna Koundinya**

- **October 2023**

Contents:

Problem Statement

Introduction

Business Understanding

Business Objectives

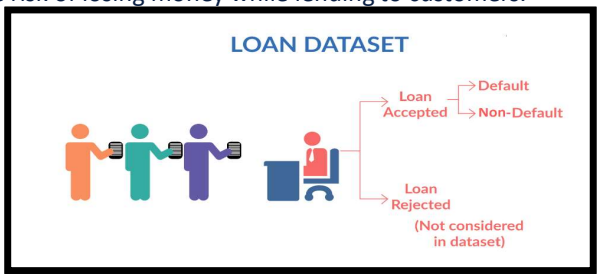
Analysis approach

Data Insights

Problem Statement

Introduction

Solving this assignment will give you an idea about how real business problems are solved using EDA. In this case study, apart from applying the techniques you have learnt in EDA, you will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimise the risk of losing money while lending to customers.



Business Understanding

You work for a **consumer finance company** which specialises in lending various types of loans to urban customers. When the company receives a loan application, 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, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company

The data given below contains information about past loan applicants and whether they 'defaulted' or not. The aim is to identify patterns which 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.

In this case study, you will use EDA to understand how consumer attributes and loan attributes influence the tendency of default. When a person applies for a loan, there are two types of decisions that could be taken by the company:

- Loan accepted: If the company approves the loan, there are 3 possible scenarios described below:
  - Fully paid: Applicant has fully paid the loan (the principal and the interest rate)
  - Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
  - Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan
- Loan rejected: The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset)

Business Objectives

This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

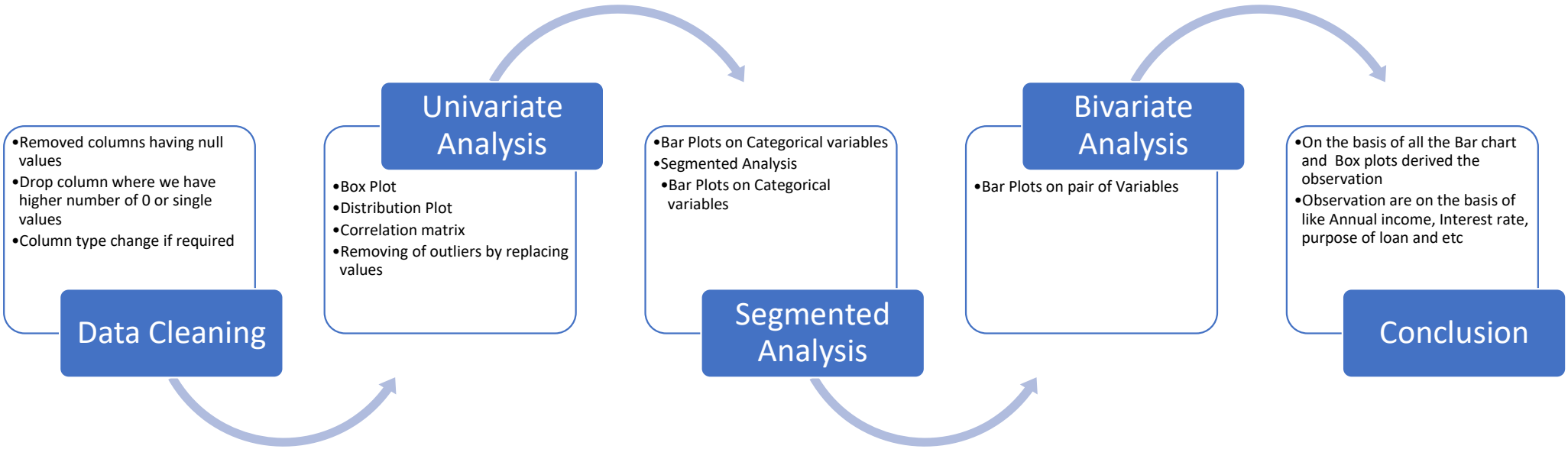
Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). Credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.

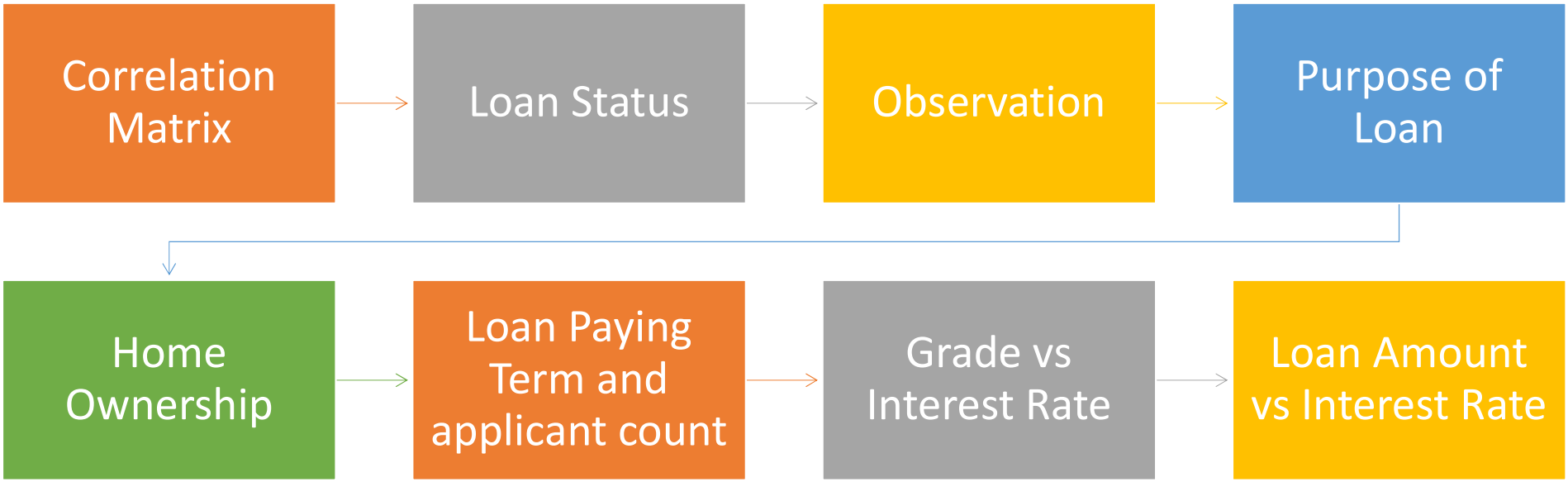
In other words, the company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.

To develop your understanding of the domain, you are advised to independently research a little about risk analytics (understanding the types of variables and their significance should be enough).

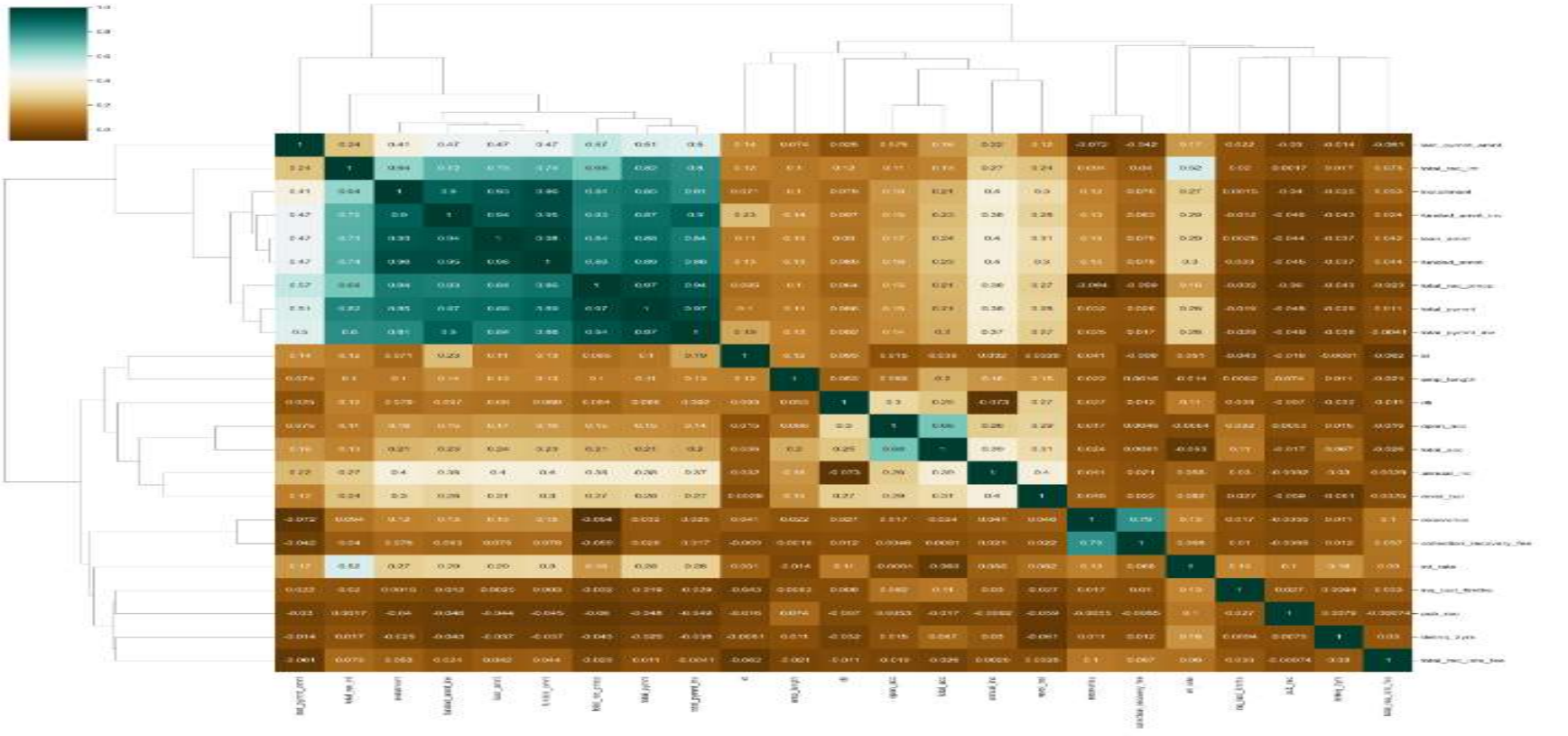
Analysis Approach



Data Insights:

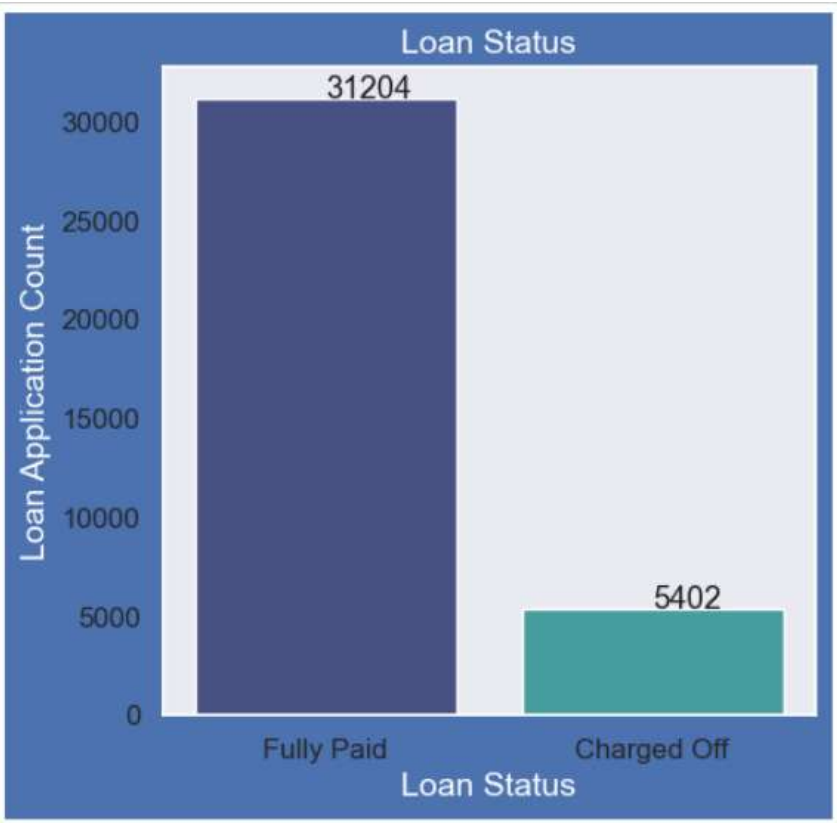


Correlation Matrix



- Basis of this analysis we can clearly see Loan amount, funding amount and installment are strongly correlated. Annual income with DTI is negatively correalted, so this means means when annual income is low DTI is high & vice versa.
- Positive correlation between annual income and employment years which means income increases with work experience

Loan Status



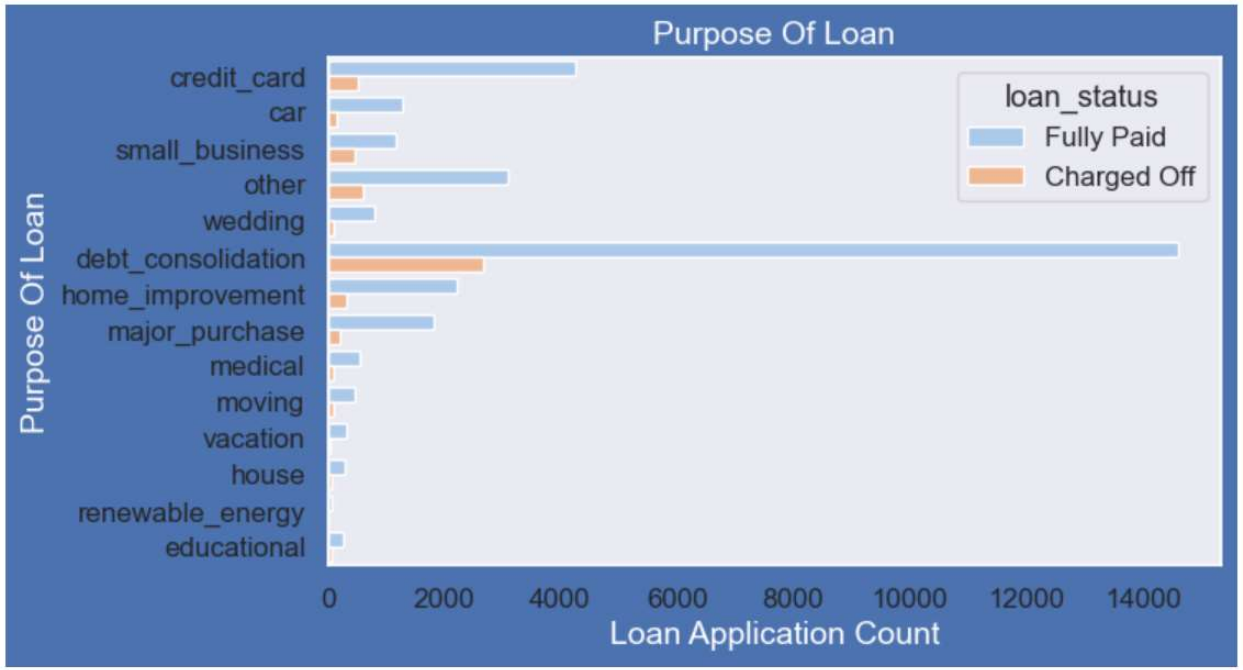
❖Using this graph we can say that almost to 17% loans were charged off out of total loan issued

Observation

Observation on the basis of all the charts prepared

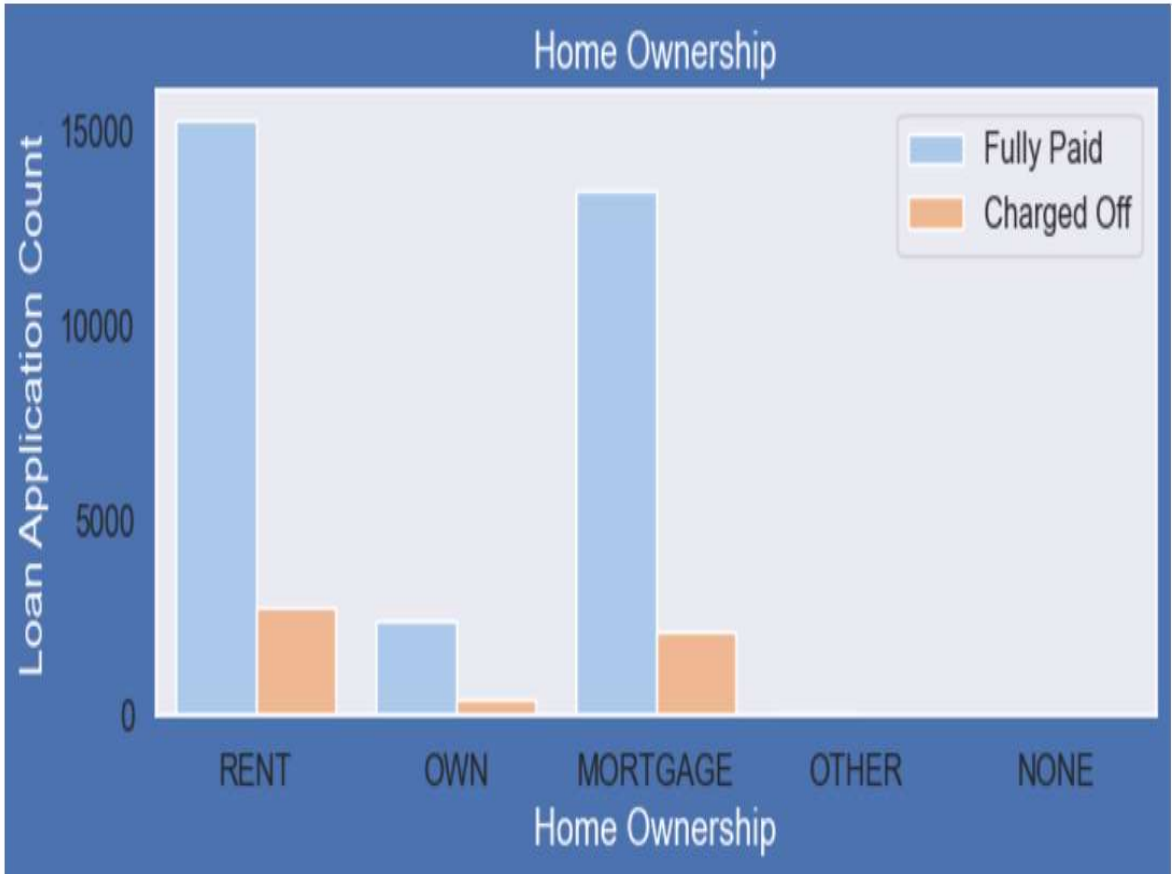
- Annual income vs loan purpose: Applicants taking loan for 'home improvement' and have income of 60k -70k
- Annual income vs home ownership: Applicants whose home ownership is 'MORTGAGE and have income of 60-70k
- Annual income vs int_rate: Applicants who receive interest at the rate of 21-24% and have an income of 70k-80k
- Loan Amount vs Interest Rate: Applicants who have taken a loan in the range 30k - 35k and are charged interest rate of 15-17.5 %
- Loan vs House Ownership: Applicants whose home ownership is 'MORTGAGE and have loan of 14-16k
- Loan amount vs Grade: When grade is F and loan amount is between 15k-20k
- grade vs interest rate: For grade G and interest rate above 20%

Purpose Of Loan



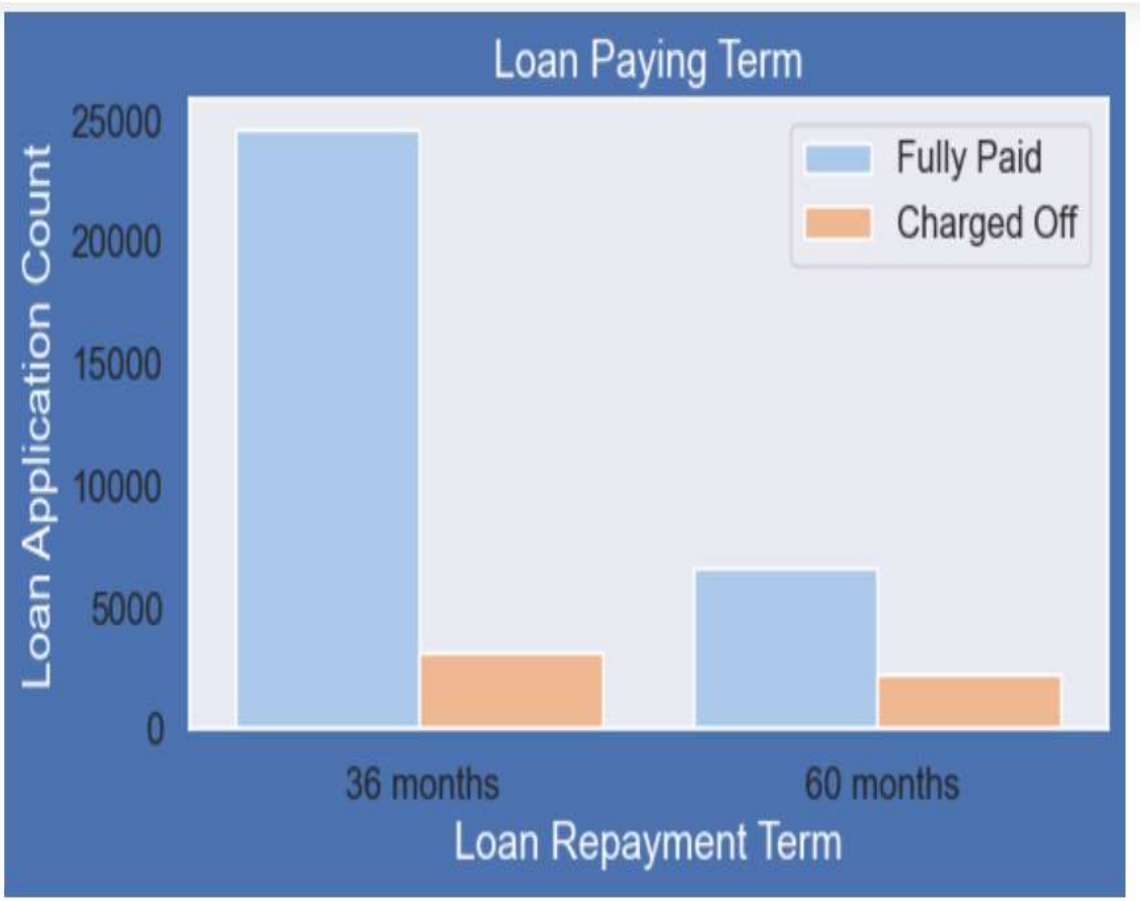
- Most of the loans were taken for the purpose of debt consolidation & paying credit card bill and number of charged off count also high too for these loans.

Home Ownership



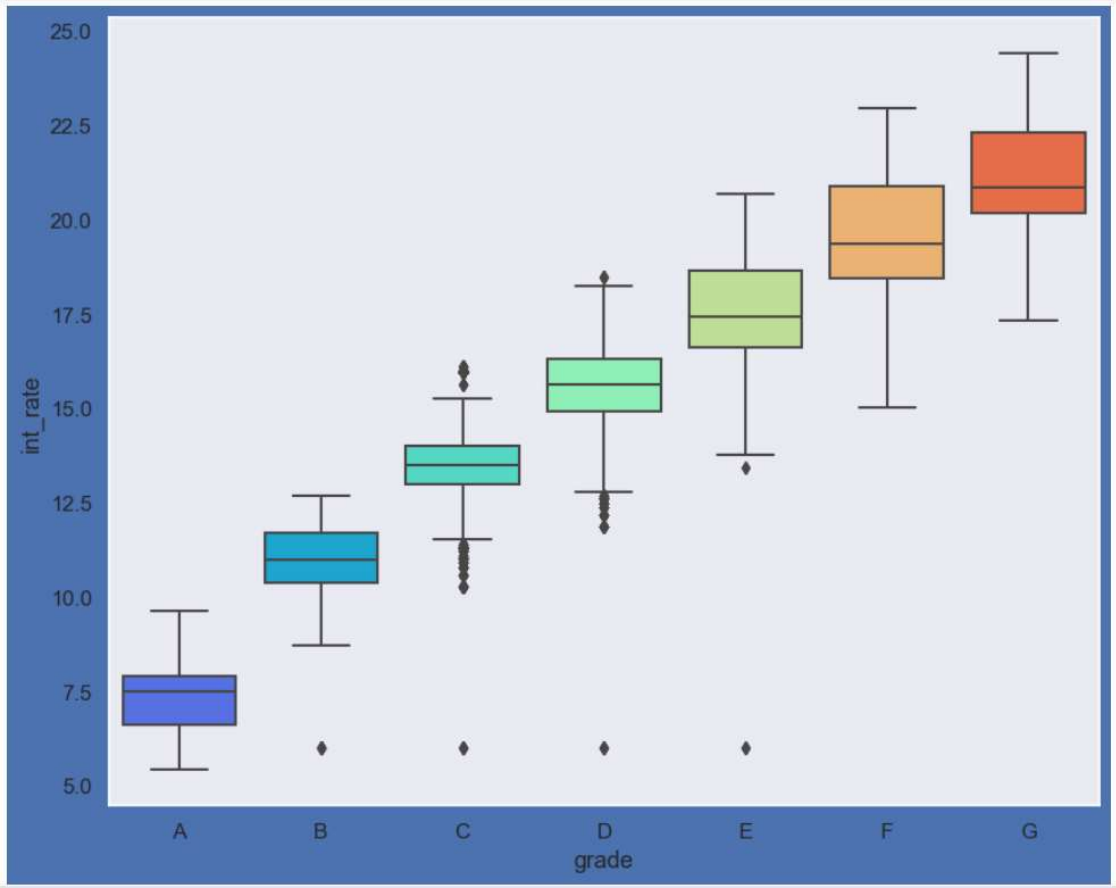
- Most of them living in rented home or mortgaged their home and Applicant numbers are high from these categories so charged off is high too.

Loan Paying Term and applicant count



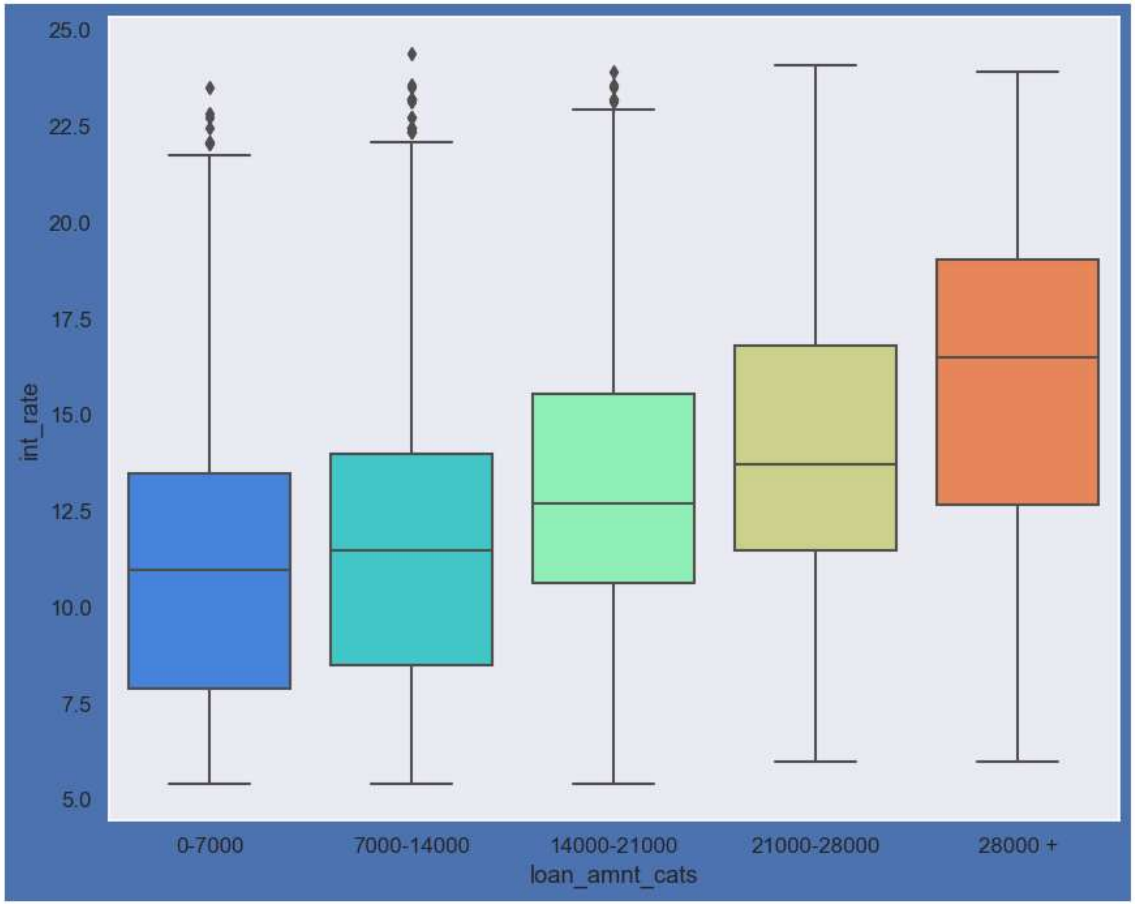
- who all had taken loan to repay in 60 months had more % of number of applicants getting charged off as compared to applicants who had taken loan for 36 months.

Grade vs Interest Rate



- This graphs clear that interest rate is increasing with grades moving from A to F. The higher the borrower's credit grade, the lower the interest rate offered to that borrower on a loan.

Loan Amount vs Interest Rate



- This graphs clear that interest rate is increasing with loan amount increase. This may be probably when loan amount is more it is taken for longer loan term, we saw earlier that longer the loan term more the interest rate.

upGrad

Executive PG Programme in Machine Learning & AI - August 2023

Thank You !!

https://github.com/VISHALKr-CHOUDHARY/LendingClubCaseStudy.git