Case Study: Lending Club

Meet Our Team

Abhishek Varun

Anish Nilesh Rane





Agenda



1. Introduction to Problem Statement

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:

1.If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company

2.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





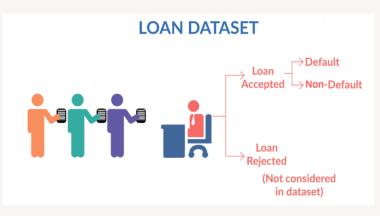
Applying the techniques we have learnt in EDA, this will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers

2. Abstract

- Lending club is the largest online loan marketplace, facilitating personal loans, business loans and financing for medical procedures.
- Borrowers can easily access lower interest rate loan through a fast online interface.
- The objective of analysis is to use the information about past loan applications and find whether they "default" or not

3. Understanding Data

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.



When a person applies for a loan, there are **two types of decisions** that could be taken by the company:

- **1.Loan accepted:** If the company approves the loan, there are 3 possible scenarios described below:
 - 1. Fully paid: Applicant has fully paid the loan (the principal and the interest rate)
 - **2. 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'.
 - **3.** 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

2.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)

4. Problem Solving Methodology

Data Cleaning

Data Understanding

Univariate Analysis Segmented Univariate Analysis

Bivariate Analysis

Recommendation

Remove 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

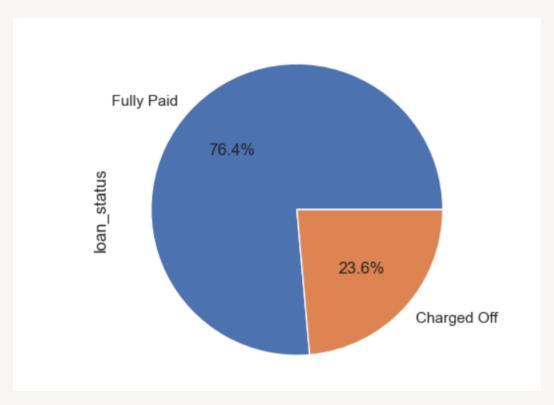
Analyzing each column, distribution of each column

Analyzing the continuous data columns with respect to categorical column

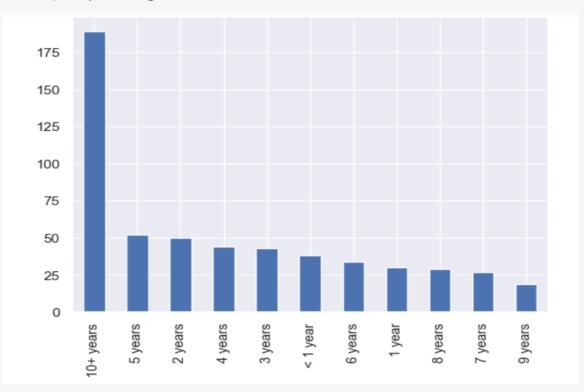
Analyzing the two variable behavior like term and loan status with respect to loan amount

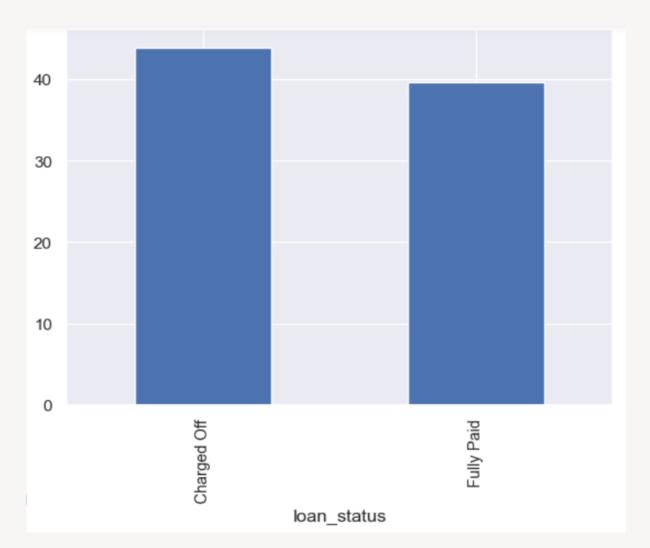
Analyzing all plots and recommendations for reducing the loss of business by detecting columns to loan defaulters

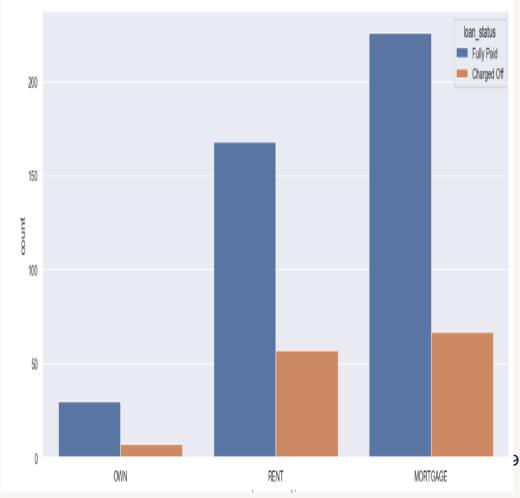
Distribution ratio of Loan status attributes



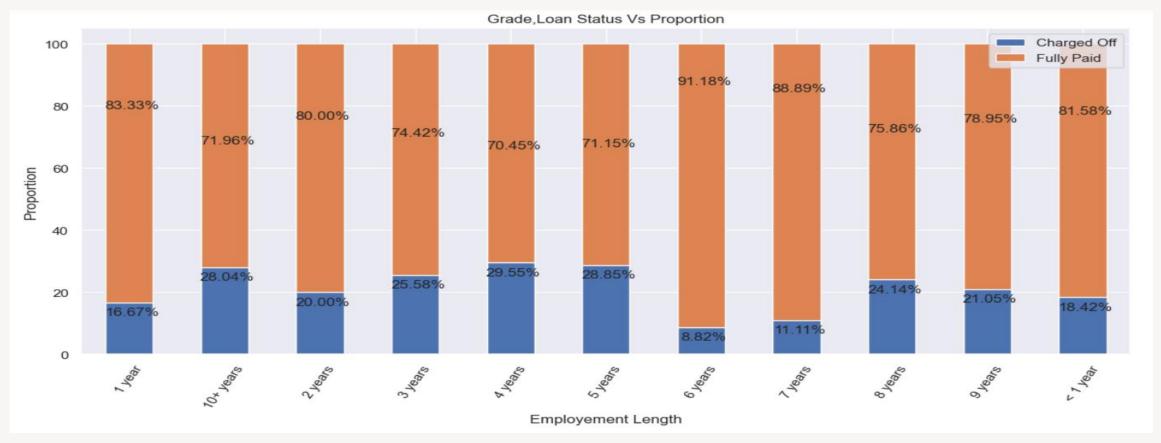
Employ length Distribution







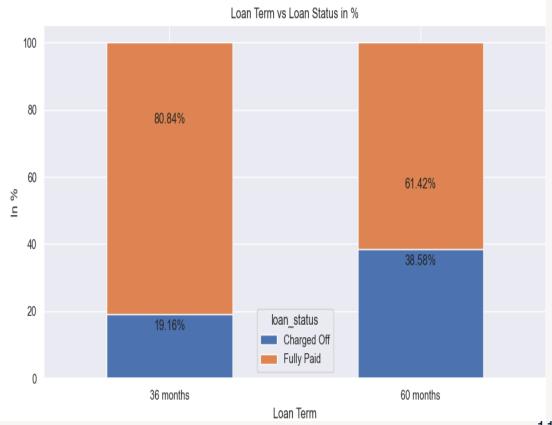
Employment length having 10+ Years are more likely to Charged off. Hence Lending club should examine carefully before approving the Loan to them



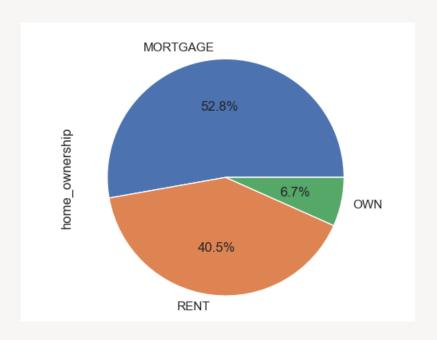
Checking the correlation using heatmap for more insights

Borrowers with 3 Years tenure are most likely to Fully Pay where as Borrowers with 5 Years tenure are most likely to Charged off

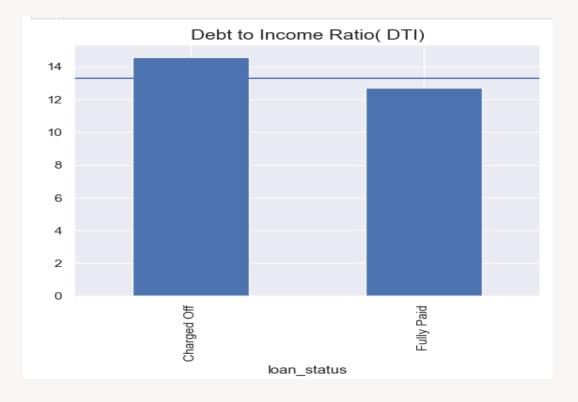




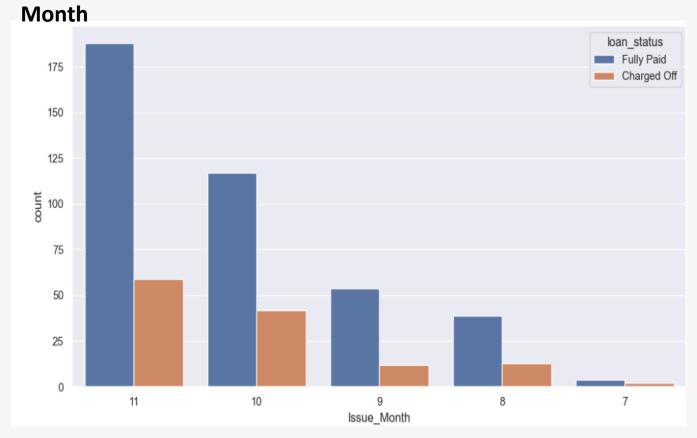
The feature Home Ownership reflect an interesting insights, that most of the borrower who stays in Rent or Mortgage are the loan applicants and out of these two who stays in Rent is having higher percentage of Fully Paid



If the DTI Ratio is more than 13.27 then Borrowers are most likely to Default.



Most of the Applicant has been issued the Loan in the December



Conclusion

- Lending club should reduce the high interest loans for 60 months tenure, as they are prone to loan default.
- Lending Club should control their number of loan issues to borrowers who are from CA, FL and NY to make profits.
- 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 more number of public derogatory are having more chance of filing a bankruptcy. Lending club should make sure there are no public derogatory records for borrowers.
- Grades are good metrics for detecting defaulters.



