Lending Club Case Study

Group Members

Rini Varghese

Case Study

Context :

A Consumer finance company that specializes 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

Problem Statement :

Using EDA 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 utilize this knowledge for its portfolio and risk assessment.

Analysis Approach

Data Cleanup

❑ Analyse the loan data set for it shape.
❑ Identify and remove all the null values columns that may not have significance in the Analysis
❑ Drop the Columns that do not have significance in the Analysis
❑ Update the correct Datatypes for Columns with date , float and int datatypes

Univariate Analysis

❑ Checked the distribution for Categorical Columns and the frequency of numerical columns

Bivariate Analysis

❑ Analysed the correlation relation between variables
❑ Checked how two variables affect each other

Final Summary

❑ Publish the insight and approach that the Lending club needs to take to avoid defaults

Observation

Parameters Affecting loan default as per the Analysis were

- ❑ Loan Status
  - ▪ Number of Fully paid was more than the charged off, so there is an imbalance in the dataset
- ❑ Grade:
  - ▪ The Grades for loans were assigned from A to E and the observation is that more loan defaults were happening from grade F to G. Lending club needs to restrict such loans to avoid default/
- ❑ Annual Income :
 - ▪ Loan defaults are higher for lower income, and progressively reduce as incomes go up.
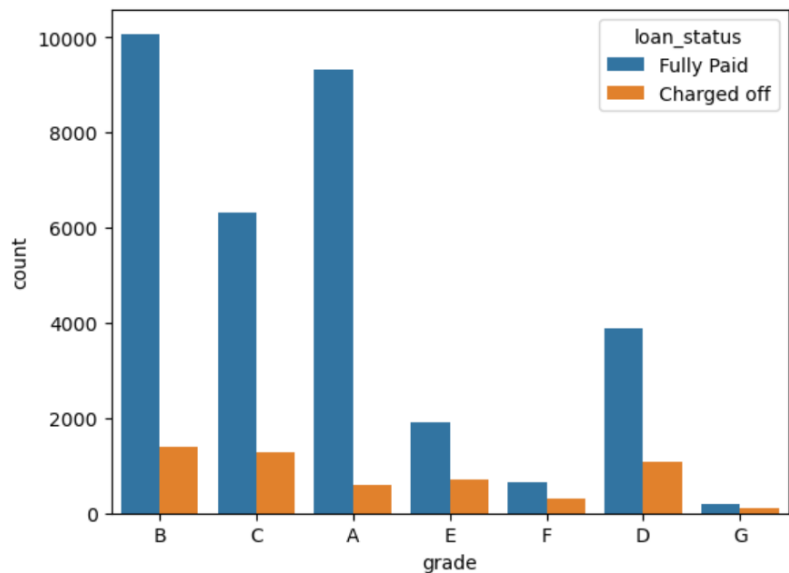- ❑ Verification Status
  - ▪ Higher loan amounts are verified more often. But we see a higher charge-off rate. So higher loan amounts are riskier and are also verified more often by Lending Club.
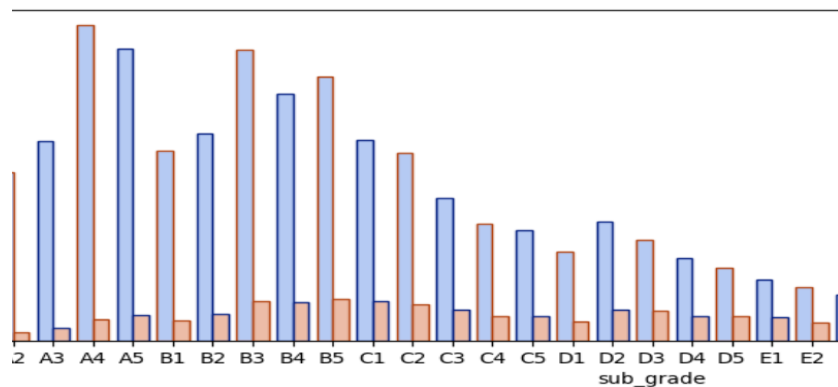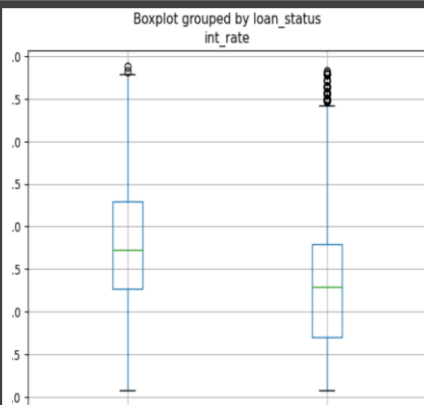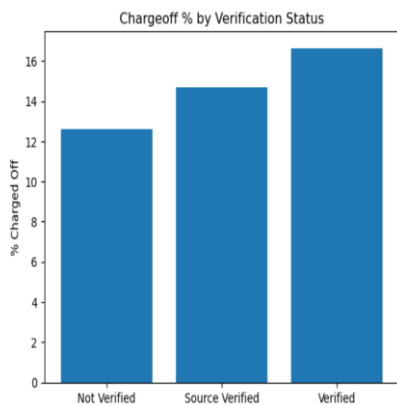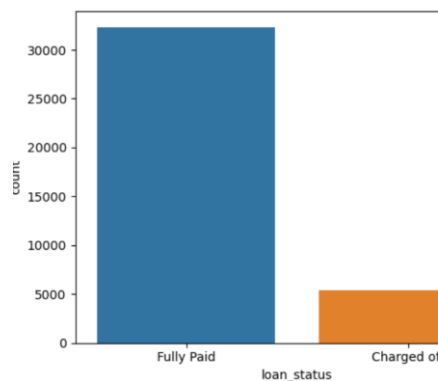- ❑ Purpose
  - ▪ The high defaults are where the purpose is debt consolidation
- ❑ Term – Lower term of 36 months were showing more default compared to longer-term loans . So the lending club needs to avoid higher loan with shorter term to prevent default
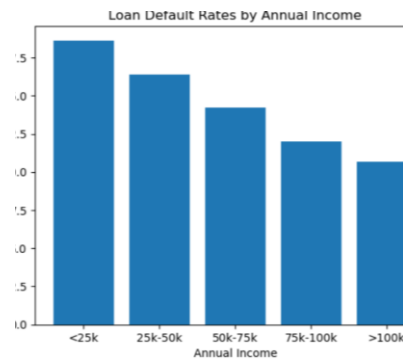
Variables affecting default

- Verification status
- Term
- Annual income
- Interest Rate

Thank You