
Lending Club Case Study

— Sudharshan Rajan —

Problem Statement

- Identify risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.
- Derive the driving factors (or driver variables) behind loan default. The company can utilize these insights for its portfolio and risk assessment
- Identification of such applicants using EDA techniques is the aim of this case study.

Analysis Approaches

1. Data Cleaning/Preprocessing
 - Cleanse/remove data that has more missing values
 - Remove columns that are not much relevant for analysis(Cols like member_id,installment)
 - Identify and remove Single Valued columns (Cols like pymnt_plan ,initial_list status)
2. Univariate Analysis
 - Categorical Features
 - Compare categorical features like Grade,Home Ownership vs Target variable (whether default or not)
 - Numerical Features
 - Correlation Between various numerical features (annual income vs dti ,annual income vs default).
 - Segmented Univariate Analysis
 - Categorize some of the numerical features into ranges for better insights like annual_inc range ,int_rate_range, loan_amount_range

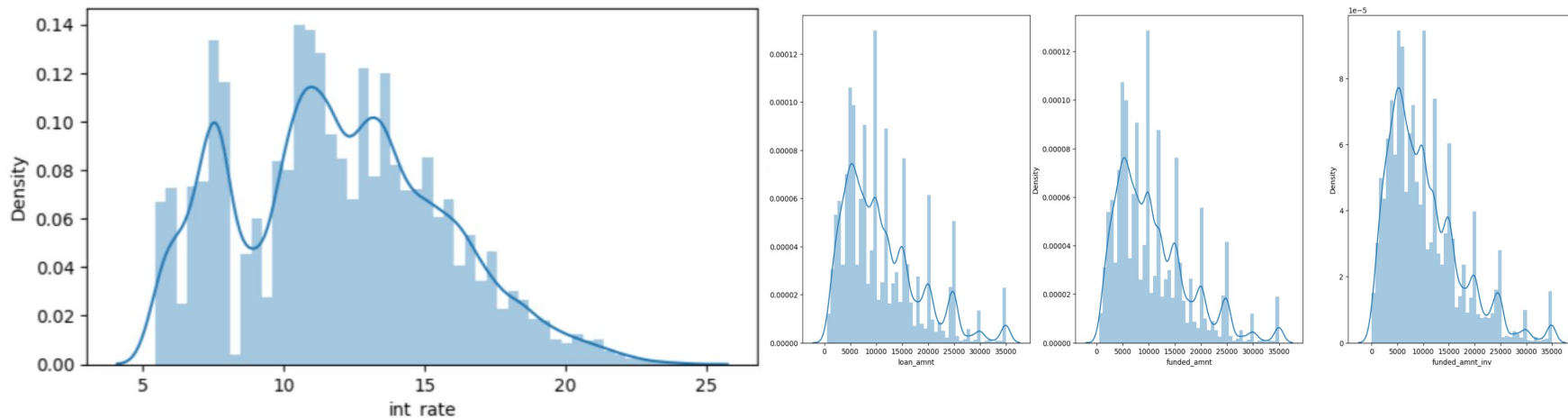
Analysis Approaches (Contd..)

Bivariate Analysis

Percentage of charged off loan under each categorical and numerical Features
Ex % of defaulters under each income category.

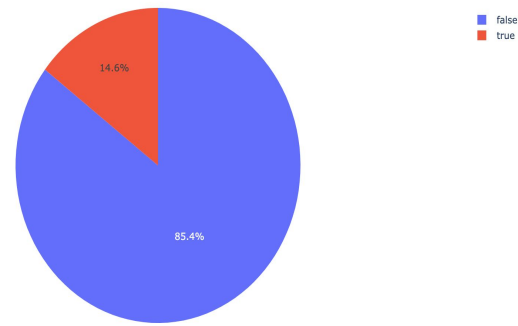
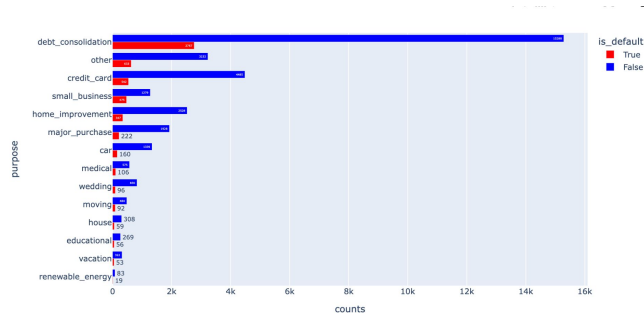
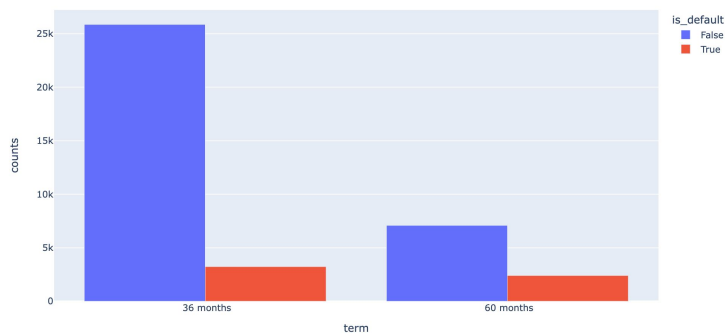
Data Overview

- Close to 57/111 columns had almost 100% missing values
- Most of the Loan Amount distribution was between 5000-10000
- Int rate was between 10-15%



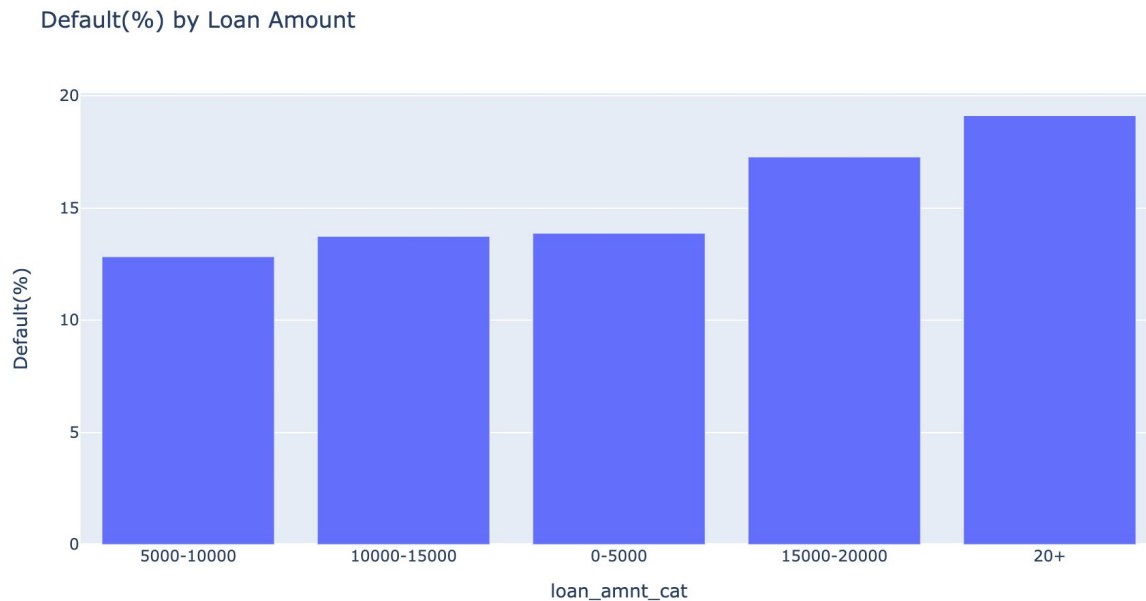
Data Overview (Contd..)

- Almost 15% loans are defaulted
- Most of the loans are for 36 Month term
- Most of the loans belong to purpose debt consolidation
- Most of the Loans belong to higher grades B& A



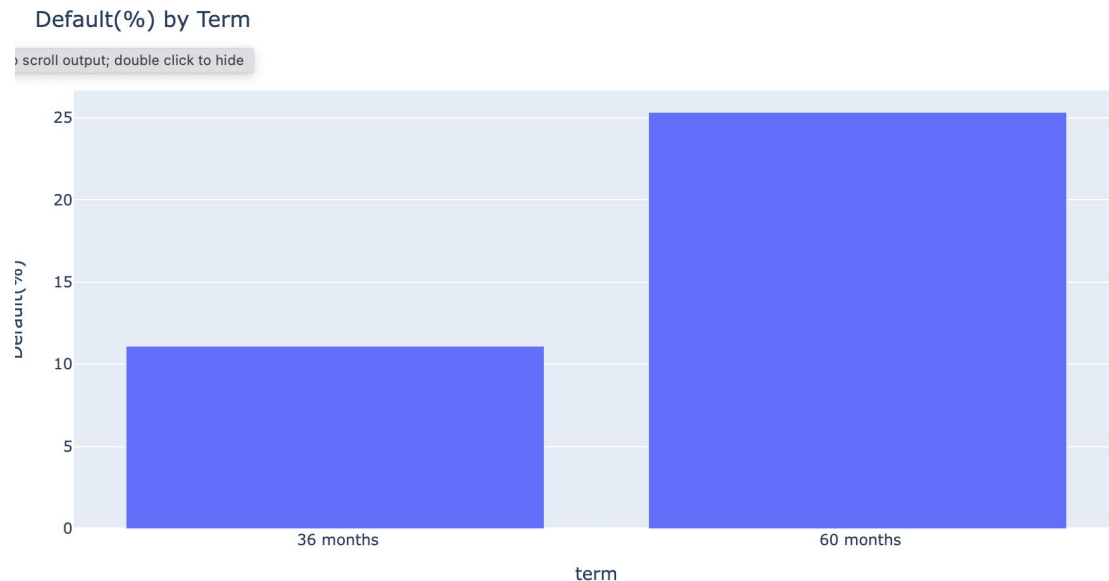
Default (%) by Loan Amount

No of defaulters increase if the loan amount is more than 20000



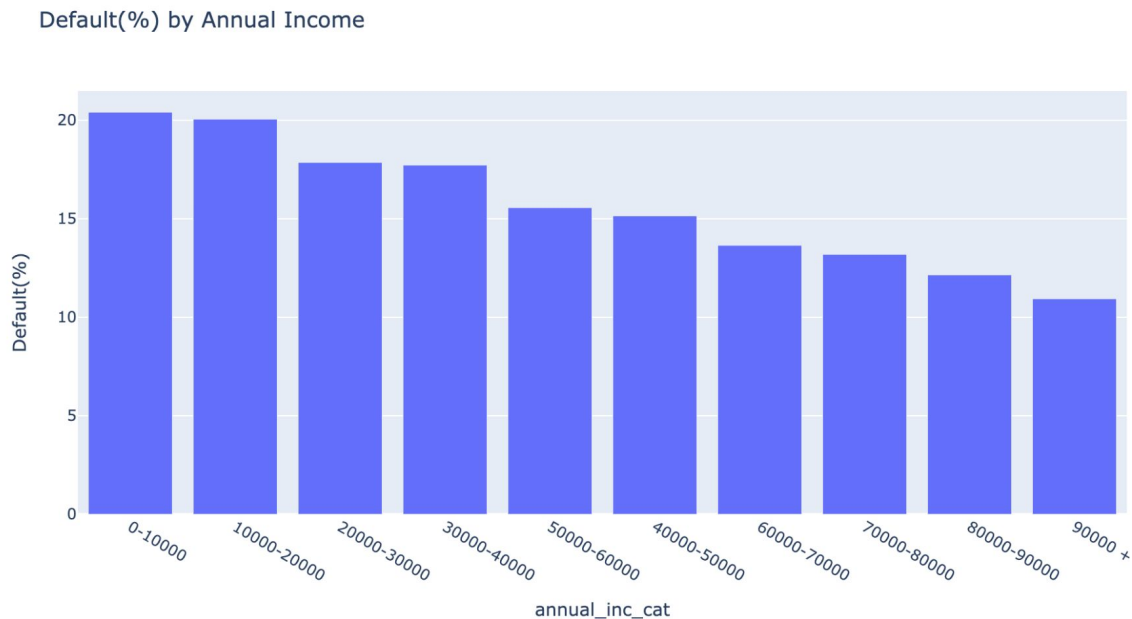
Default(%) by term

Defaulters are high when the term is 60 months



Default(%) vs Annual Income

Defaulters reduce when the annual income is more than 50K



Recommendations

- Approve Loan only if the annual income is more than 50000
- Reject Loans if the purpose is for small business (Defaulters(%) is close to 30)
- Reject Loans belonging to F or G grade(Defaulters are more than 30%).
- Reject Loans if the applicant has any record of bankruptcies.
- Double check annual income for loans with interest rate more than 20%
- Check the annual income to be more than 50000 if the term of the loan is 60 Months