Lending Club Case Study

KUNDAN TRIPATHI

MAY, 2023

Objective Of This Analysis Is To Study The Lending Club Loans Dataset And Come Up With Insights On What Drives Defaults

Lending Club is an online marketplace for personal loans where borrowers can apply for loans and interested investors can fund those loans.

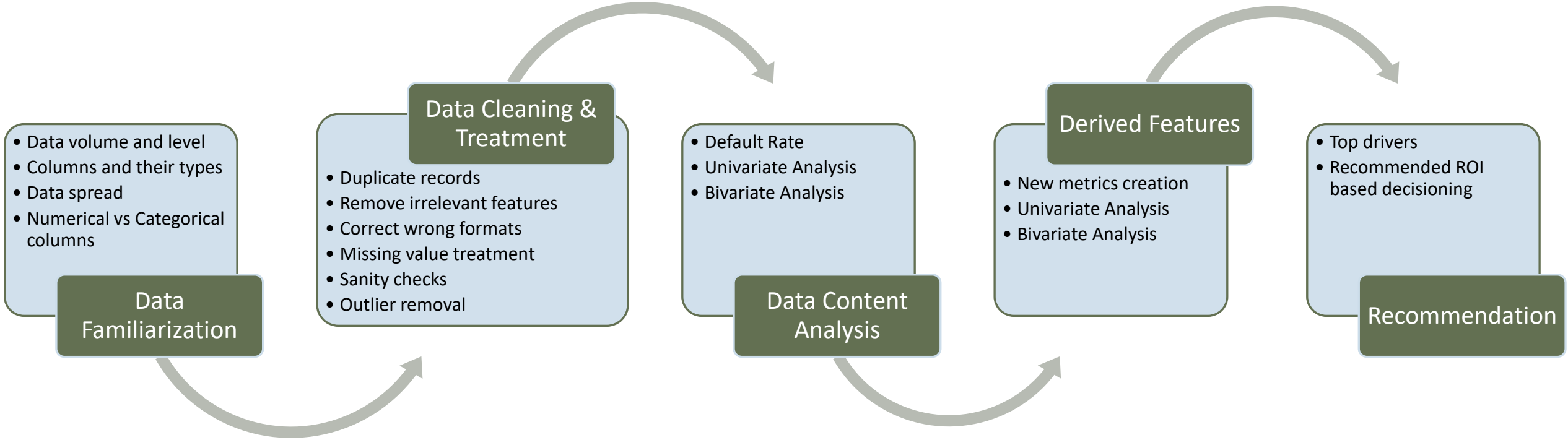
Data Description

We have been provided data from the Lending Club which is an online marketplace for loans. The data contains information on past loans and whether they have been paid vs defaulted. The data contains various consumer attributes and loan attributes and other variables related to the loans.

Business Problem

We have to conduct an exploratory data analysis on this dataset. 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. We need to isolate the driving factors behind the default i.e. variables that are strong indicators of default.

Process Followed To Conduct The Analysis Includes The Below Outlined Steps (not necessarily in strict order as EDA is an iterative process)



Few Top Variables Are Identified Which Are Biggest Predictors Of Default. Company Can Use These Fields And Create A Probability Based ROI Decisioning System.

Top variables which are predictors of default

- Loan term; Grade of loan; Number of delinquencies; Number of open accounts; Number of public record bankruptcies; Annual income; Debt-to-income ratio

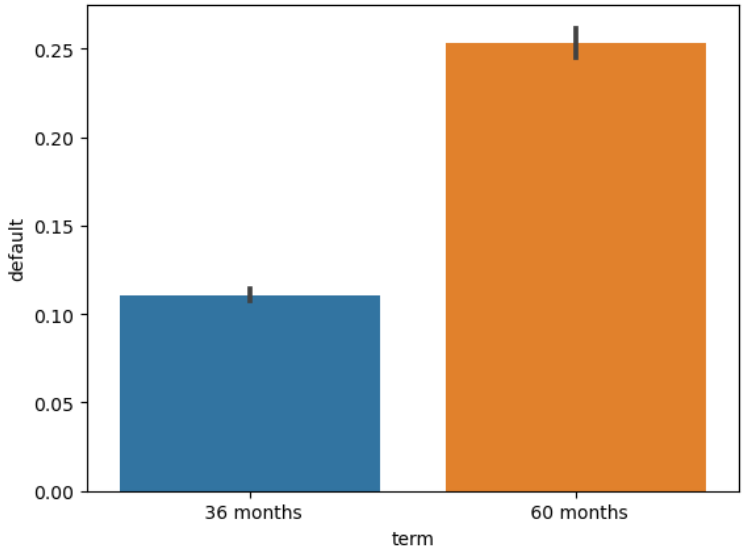
Recommendations (use ROI based approval system)

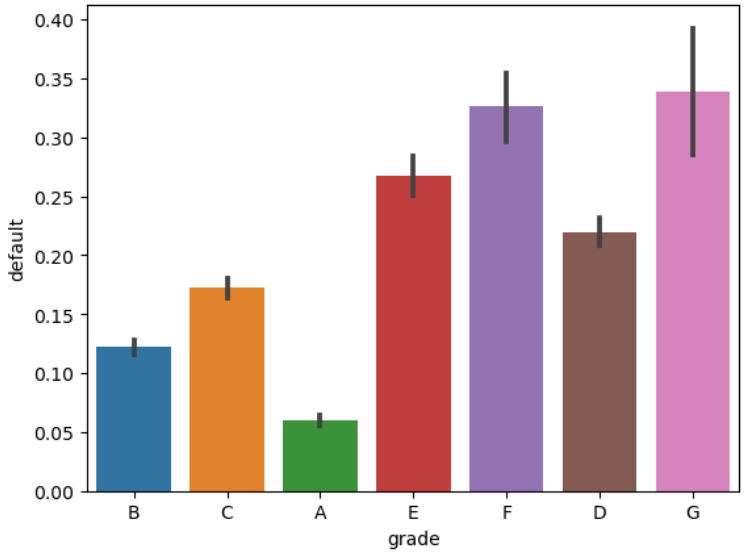
- The company should be especially careful about its longer term loans.
- Loans with Grade E onwards are risky an should be reduced.
- The more someone has been delinquent in past 2 years, more are they likely to default; especially 7 and above.
- The more inquiries in last 6 months, more are they likely to default; especially 7 and above.
- Higher number of open accounts suggest more risk. Anyone above 30 open accounts is very risky.
- More derogatory or bankruptcy public records show more riskiness.
- Very high loan amounts are risky along with lower annual income and high debt-to-income ratio.

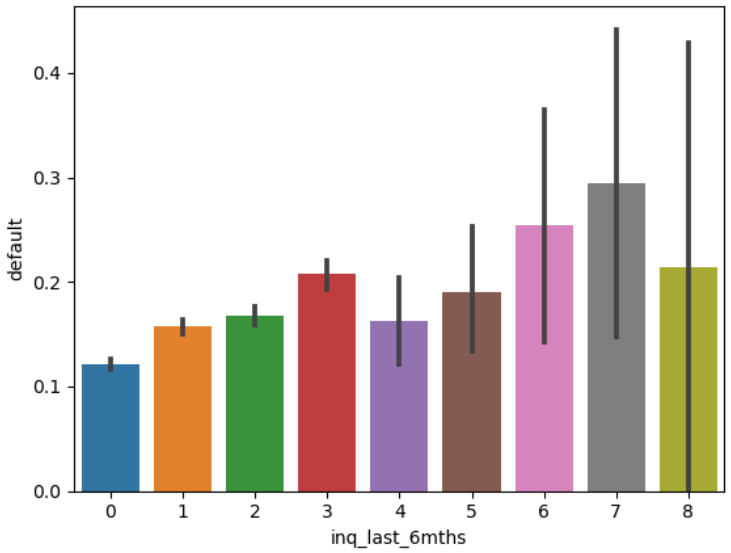
These recommendations need to be weighed against business profits and an ROI based system should be created to approve loans. The above qualified risk factors should be combined with one another and the probability of default should be calculated (same as the default rate by segments).

Using this probability of default, the return on investment of every loan judged and some criteria used for approval. We would recommend using an ROI cut-off of 3 to approve loans i.e. approve loans with ROI >= 3.

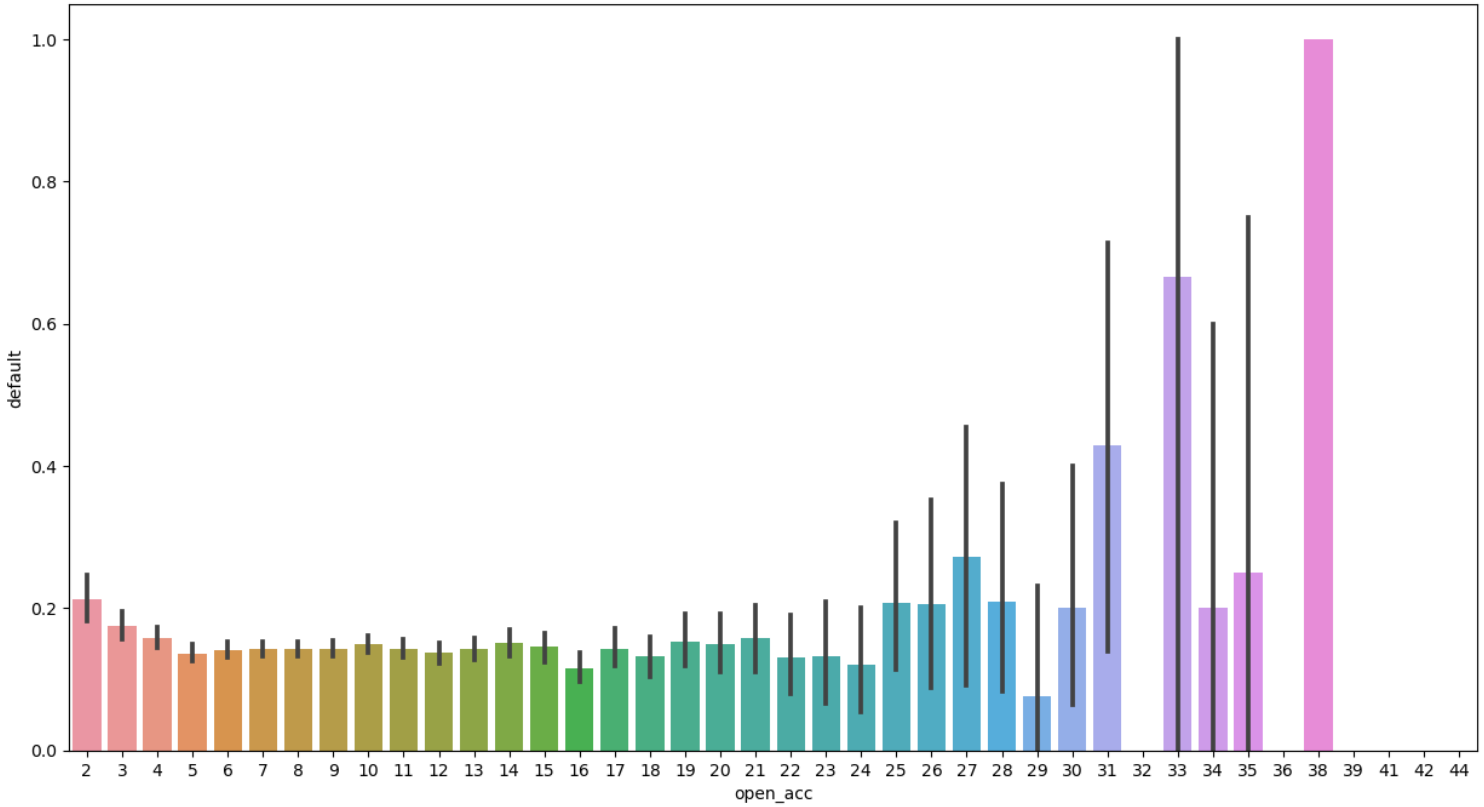
Important Charts From EDA (indicating default rates across segments)

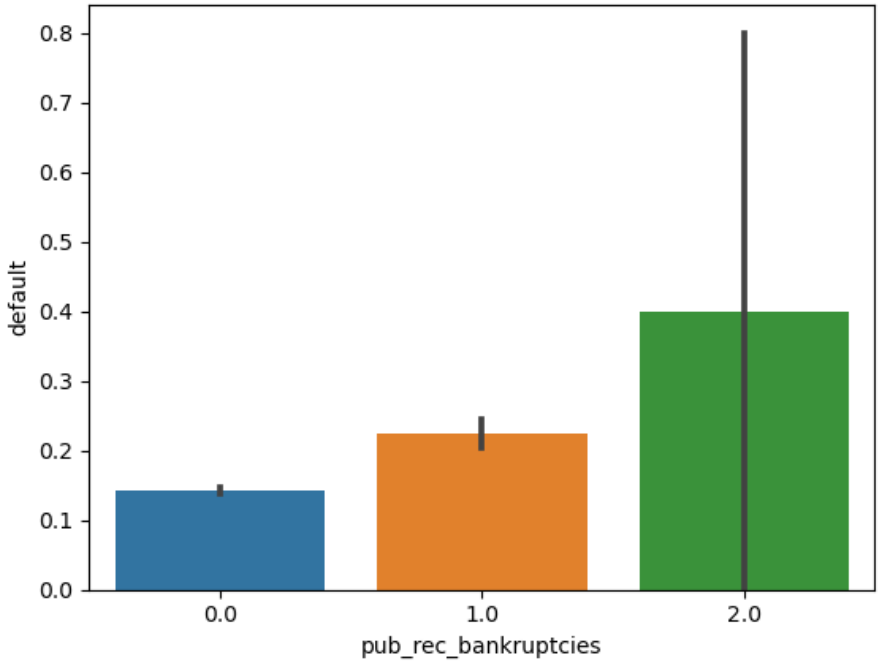






Important Charts From EDA contd. (indicating default rates across segments)





Important Charts From EDA contd. (indicating default rates across segments)