LENDING CLUB CASE STUDY

AN APPLICATION OF EDA TO DERIVE INSIGHTS ABOUT CONSUMER DEFAULT BEHAVIOUR

OVERVIEW OF LENDING CLUB

- In the Lending club case, we need to analysis financial loss for bank and lending company. We need to identify the risk from the loan applicant so that we could reduce the NPA (Non performing Asset) for the loan.
- We are looking at the Loan disbursal data for a large Online consumer finance company. The data contains ongoing loans, Fully paid and charged-off cases. Along with this various customer related data.
- The objective is to look at the existing customer data and using Exploratory Data Analysis find what attributes can be used to predict/discern final loan disposition i.e. will they be fully paid or charge off. This will help us in managing risk and taking pre-emptive steps where necessary.

THE CASE OF A DEFAULTING CUSTOMER

- We have consumer finance data of > 31K cases including Default, Fully Paid and Current customers
- Multiple loan related attributes, scores, customer attributes both categorical and numerical
- Find out using Exploratory Data Analysis if any variables identify Defaulting customers before it happens

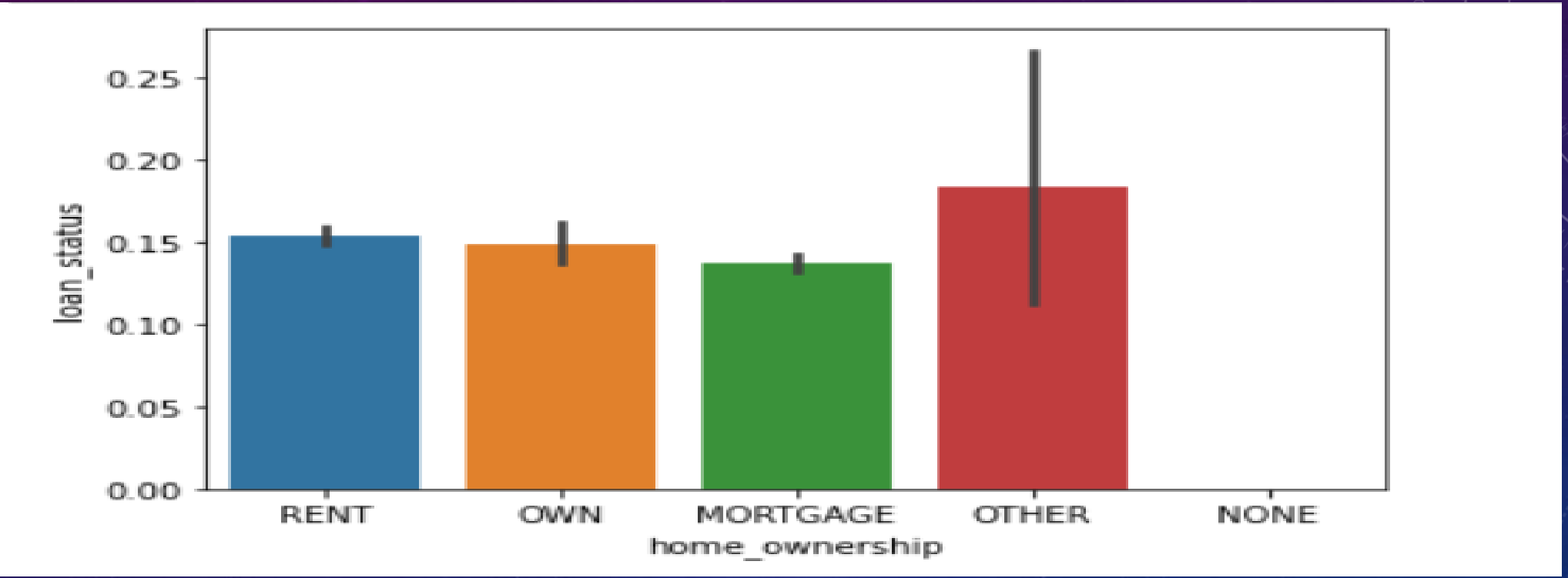
OUR APPROACH

- Data Understanding
- First clean the data to make it useful
    - Missing value treatment – removed variables>= 30% missing
    - Removed ID variables
    - Removed variable with constant values all across
    - Fixed columns with special characters(%, +)
- Univariate
- Bi-Variate

LOAN DATA DESCRIPTION

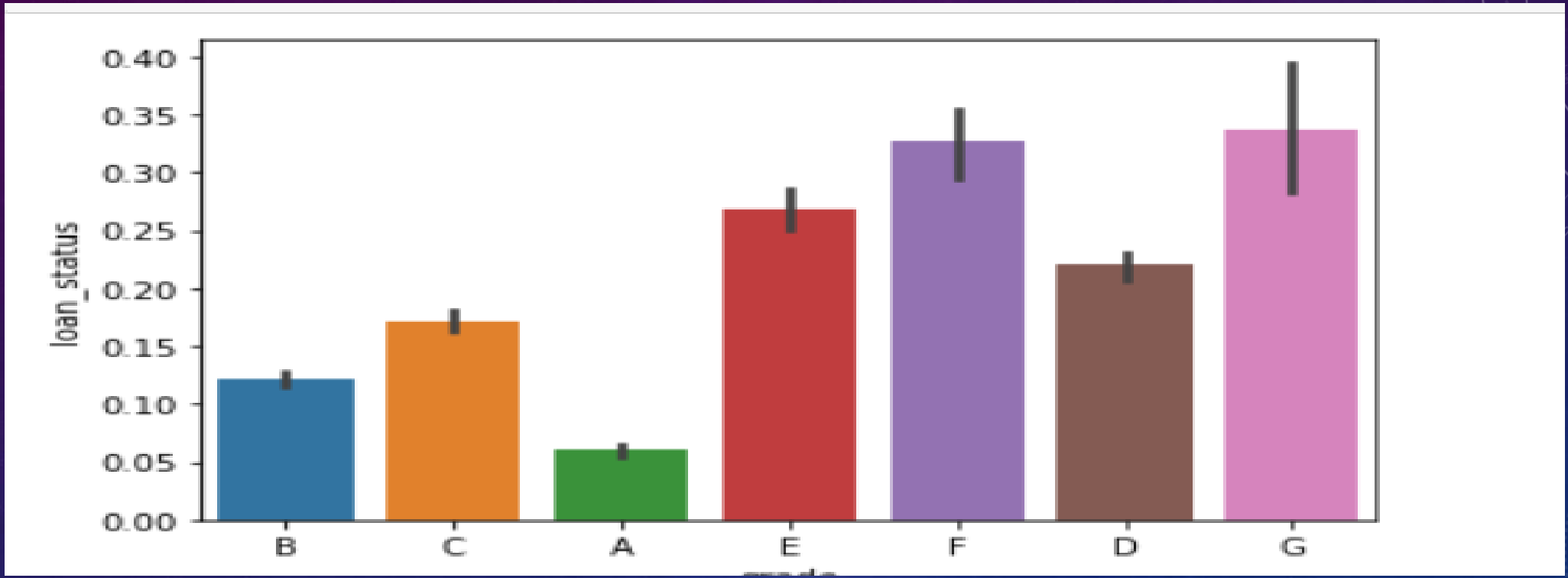
- Lending club study data provide the loan lending data from 2007- 2015.
- The data set provide the information about borrower and their past history of employment like income detail, grade and loan amount. It contains many features and records about the loan customer.
- The data set provide many attributes which help data scientistic to analysis the loan defaulters.
- The data set has many fields which have null values

DATA ANALYSIS



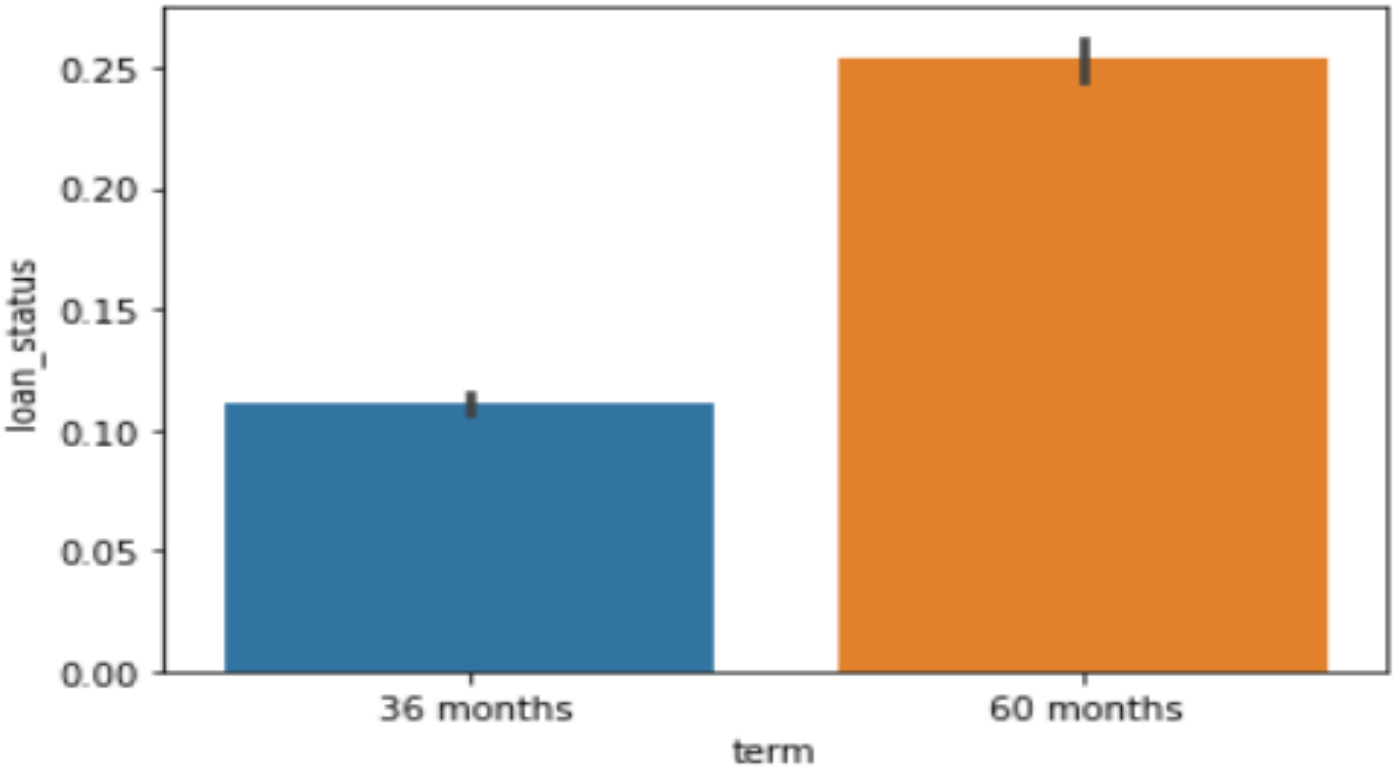
We could say that the customer who didn't mention the home owner ship having more number of defaulters.

DATA ANALYSIS



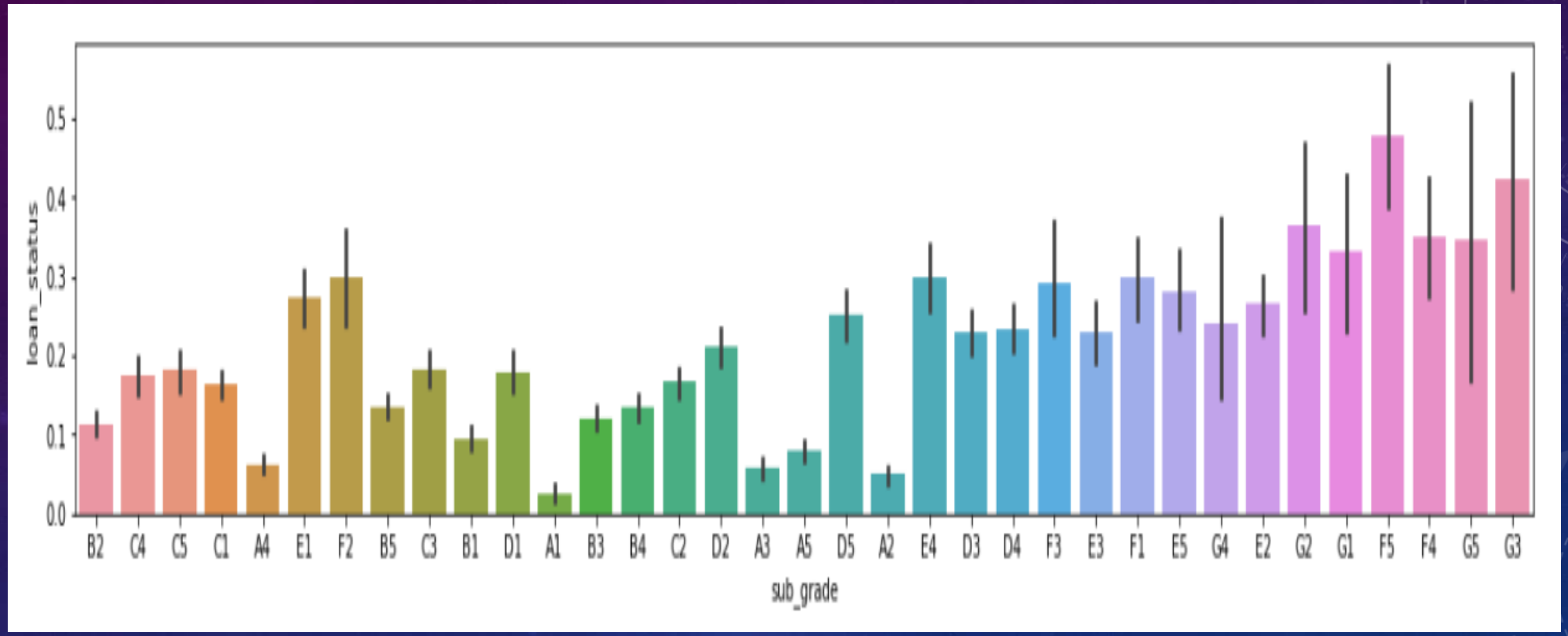
We could say the "G" Grade employee have more defaulter and "A" Grade employee get less defaulter

DATA ANALYSIS



We could say more tenure having more loan defaulter

DATA ANALYSIS



We could say the "F5" sub Grade employee have more defaulter and "A1" Grade employee get less defaulter

RESULT

- **As per the data Analysis, we can say if bank or lending company will give the loan to small business then its a high change that they could not return back the money on time.**

- **The loan amount which approve by the lending company, it should be low amount compare of the employee annual salary.**

- **While approving the loan, the lending company should always consider Grade and Sub Grade of the employee.**