

LENDING CLUB CASE STUDY

BY - ROHIT PANDEY & SAGARIKA BHUYAN

PROBLEM STATEMENT

Business Overview - As a Finance Company specialized in Loan lending to customers. When a company receives a loan application company must perform assessment on the applicant's profile. The risk associated with the decision making is as follows

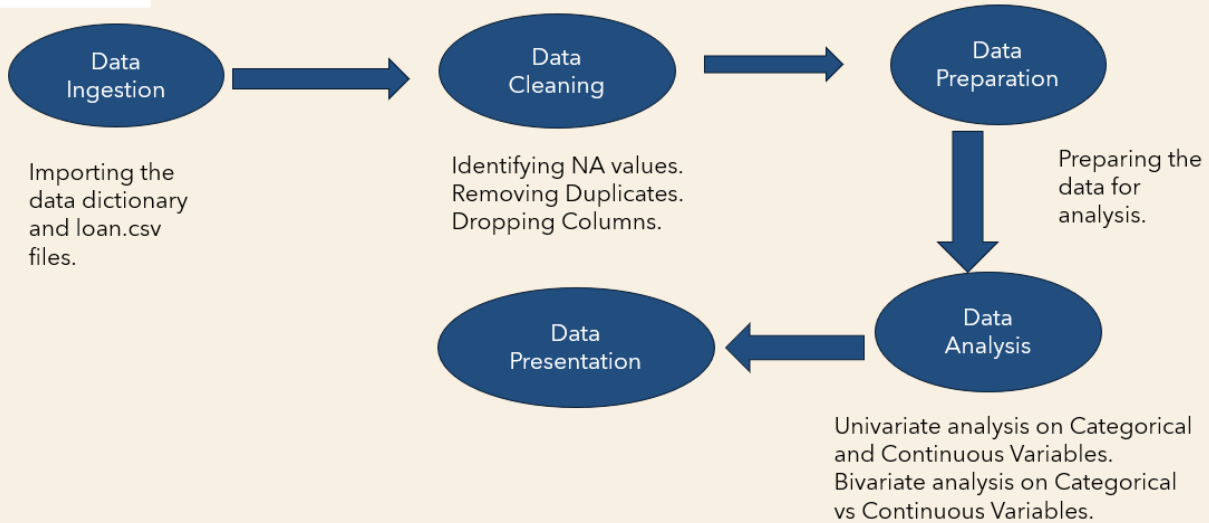
- i). If customer profile is assessed to likely to pay loan, then, not approving the loan results in loss of business to company.
- ii). If customer profile is assessed to not likely to pay loan, then approving the loan result in financial loss to the company.

Solution Context - For this business problem we need to understand the data provided. The data needs to be divided into Ordered, Continuous, Ordinal. The data contains information about the past loan applicant and whether they 'defaulted' or not. The aim is to identify 'customer attributes' to determine whether customer will default i.e. Not pay the loan Or not.

When a Loan gets accepted there are three possible outcomes

- i). Fully Paid - Applicant has fully paid the loan.
- ii). Current - Applicant is in process of paying the loan.
- iii). Charge off- Applicant has not paid the installment for loan for a very long time. Hence, defaulted the loan.

APPROACH



3

ANALYSIS

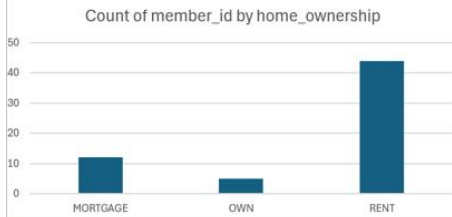


| | id | member_id | loan_amnt | funded_amnt | funded_amnt_inv | term | int_rate | installment | grade | sub_grade | ... | next_pymnt_d | last_credit_pull_d |
|---|---------|-----------|-----------|-------------|-----------------|-----------|----------|-------------|-------|-----------|-----|--------------|--------------------|
| 0 | 1077501 | 1296599 | 5000 | 5000 | 4975.0 | 36 months | 10.65% | 162.87 | B | B2 | ... | NaN | May-16 |
| 1 | 1077430 | 1314167 | 2500 | 2500 | 2500.0 | 60 months | 15.27% | 59.83 | C | C4 | ... | NaN | Sep-13 |
| 2 | 1077175 | 1313524 | 2400 | 2400 | 2400.0 | 36 months | 15.96% | 84.33 | C | C5 | ... | NaN | May-16 |
| 3 | 1076863 | 1277178 | 10000 | 10000 | 10000.0 | 36 months | 13.49% | 339.31 | C | C1 | ... | NaN | Apr-16 |
| 4 | 1075358 | 1311748 | 3000 | 3000 | 3000.0 | 60 months | 12.69% | 67.79 | B | B5 | ... | Jun-16 | May-16 |

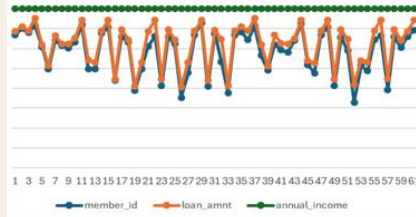
| | | | | | |
|---------------------|------------|-----------------|--------------|---------------------|--------------------------|
| int_rate | 10.65% | 15.27% | 15.96% | 13.49% | 12.69% |
| installment | 162.87 | 59.83 | 84.33 | 339.31 | 67.79 |
| grade | B | C | C | C | B |
| sub_grade | B2 | C4 | C5 | C1 | B5 |
| emp_title | NaN | Ryder | NaN | AIR RESOURCES BOARD | University Medical Group |
| emp_length | 10+ years | < 1 year | 10+ years | 10+ years | 1 year |
| home_ownership | RENT | RENT | RENT | RENT | RENT |
| annual_inc | 24000.0 | 30000.0 | 12252.0 | 49200.0 | 80000.0 |
| verification_status | Verified | Source Verified | Not Verified | Source Verified | Source Verified |
| issue_d | Dec-11 | Dec-11 | Dec-11 | Dec-11 | Dec-11 |
| loan_status | Fully Paid | Charged Off | Fully Paid | Fully Paid | Current |

4

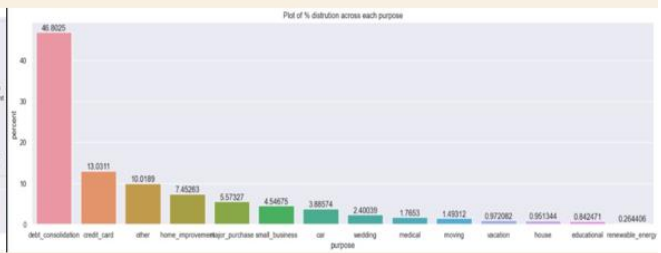
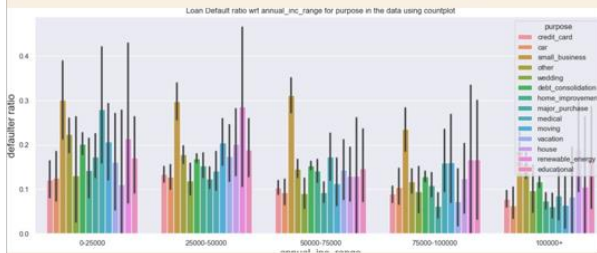
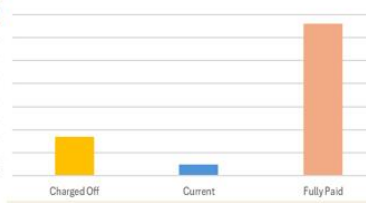
Count of member_id



Members Loan Vs Annual Income



Count of member_id by loan_status





CONCLUSION



- i). Data ingestion and analysis on the Data attributes to determine category of data such as 'Ordered, Categorical etc'.
 - ii). Data cleaning by removing duplicates, dropping column, removing NA values.
 - iii). Performing Univariate and Bivariate analysis on the data.
 - iv). Analyzing the data segment provided to determine risky customers and non risky customers.
 - v). If Non-Risky customer has current loan, then rejecting the loan depending upon the annual income.
 - vi). If a customer has defaulted the loan in past, then rejecting the loan for the customer.
- Performing Data Cleaning, Data Preparation for the Data Analysis.



THANK YOU



Rohit Pandey & Sagarika Bhuyan
rohpan86@gmail.com and
sagarikabhuyan08@gmail.com