Credit EDA Assignment

LEARNER – PRITAM SHETTY

Problem Statement

The Loan providing company is finding it hard to give loans to the people due to their insufficient credit history. Hence using the data from the past , and using it to develop patterns related to application who defaulted, so that business standard can be developed which will assist in taking decision on application in future.

Approach

1. Understanding the problem statement.

2. Understanding the data.

3. Dropping unwanted columns.

4. Checking for NULL values for individual values and if possible imputing it.

5. Univariate Analysis

6. Bivariate Analysis

7. Finding Correlation

Assumptions

Assumptions for Application Data

1. Considering XNA as Transgender in CODE_GENDER column.

Previous Application Data Assumption

1. XNA as all other types of loan in NAME_CONTRACT_TYPE

Column.

2. Only 2 Saturday working per month in WEEKDAY_APPR_PROCESS_START

Column.

3. XAP and XNA are NULL values. Assuming the values are MNAR type of missing values. Hence Not imputing it in NAME_CASH_LOAN_PURPOSE column.
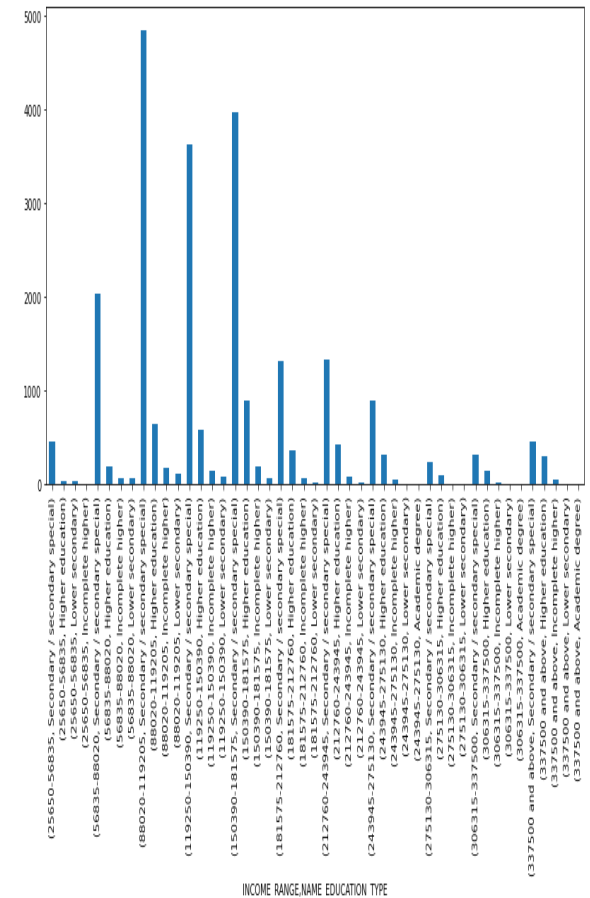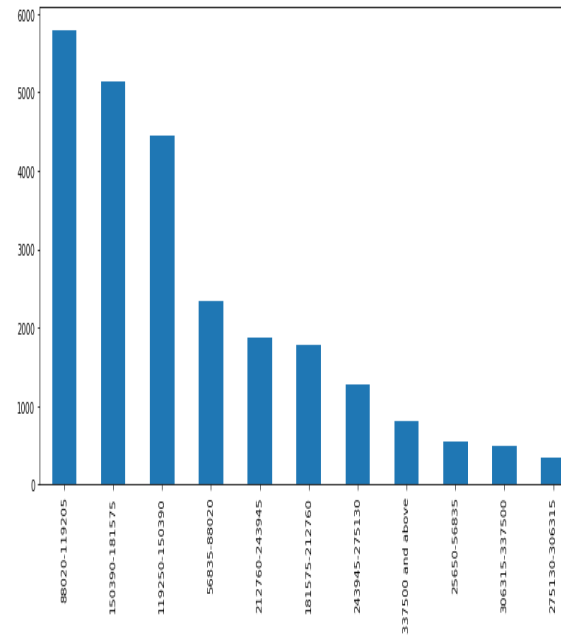
Observations

Plot 1

Distribution of applicant according to salary range.

Plot 2

Count of applicants based on Income Range and Education Type.

Observation

Loan providing company should extra precaution step for applicant with Income Range of 88020-119205 with Education Type as Secondary or secondary special.
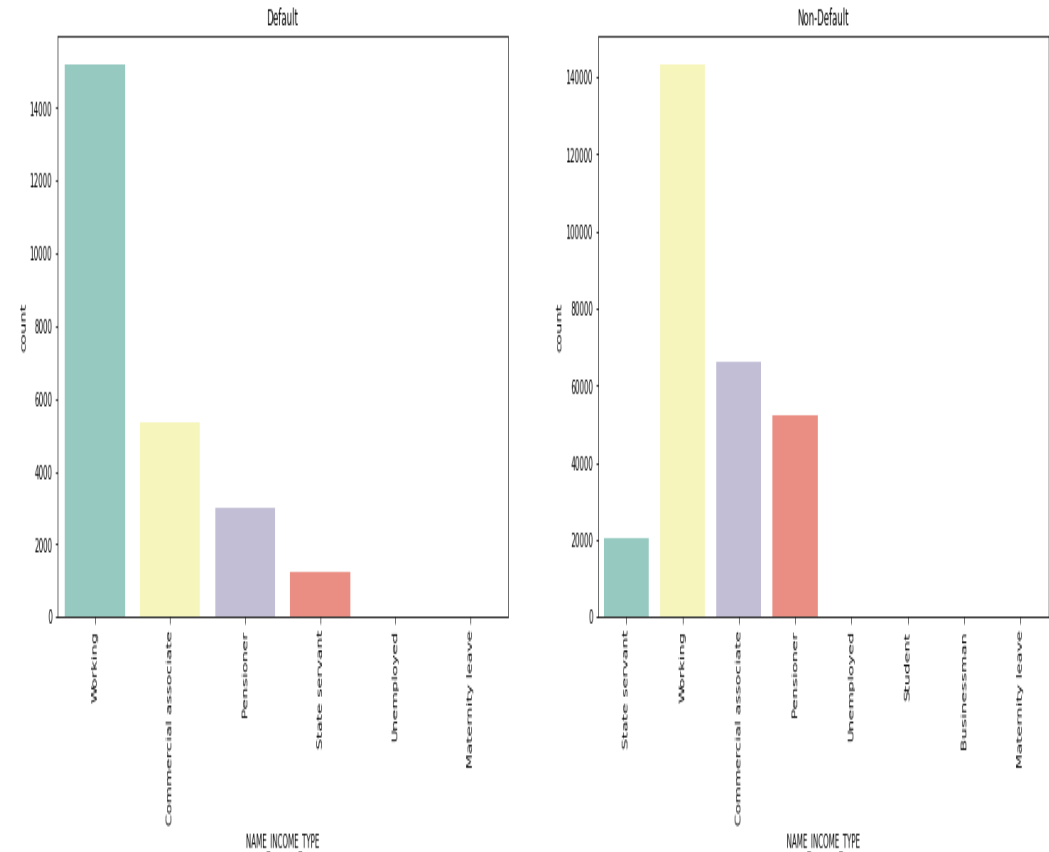
Contd.

Plot 1

Profession of applicants who have payment difficulty.

Plot 2

Profession of applicants who do not have payment difficulty.

Observation

There are high percentage of people from Working group who can default on loan. The Pensioner group has less people with payment difficulty as compared to non-payment difficulty.

Contd..

Plot

Relation between Loan application status and Type of Loan.
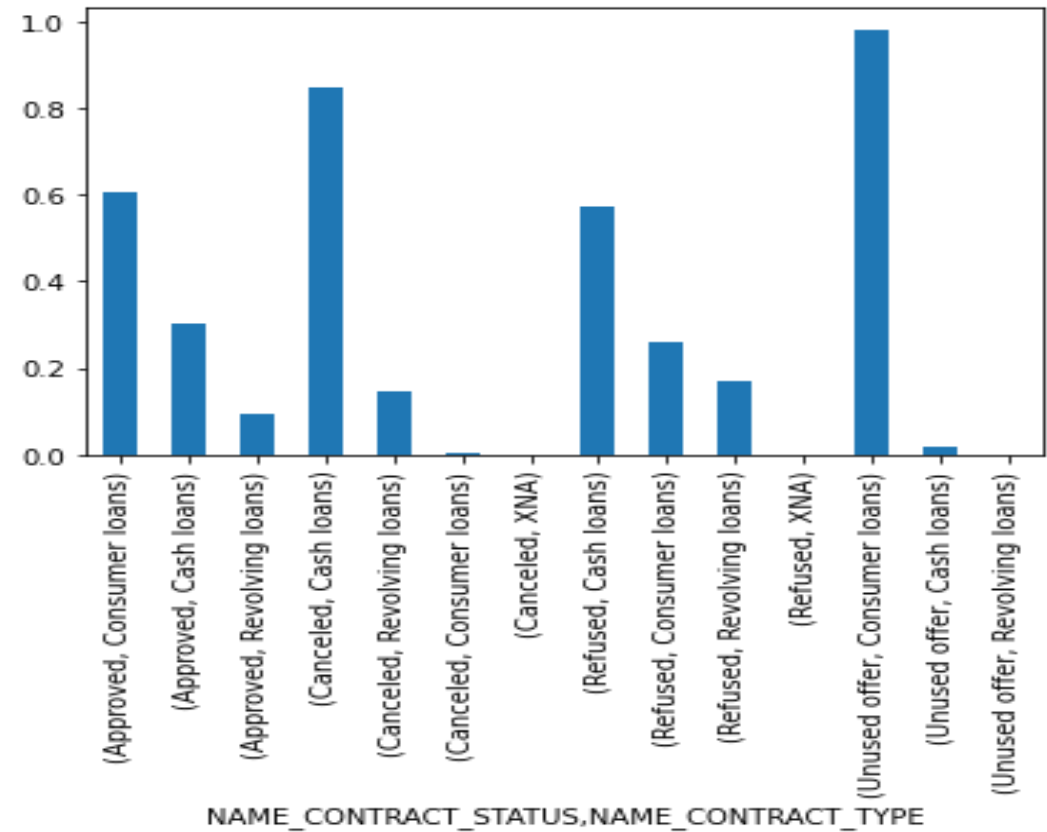
Observation

Highest Approval - Consumer Loan

Highest Cancellation - Cash Loan

Refused - Cash Loan

Unused - Consumer Loan

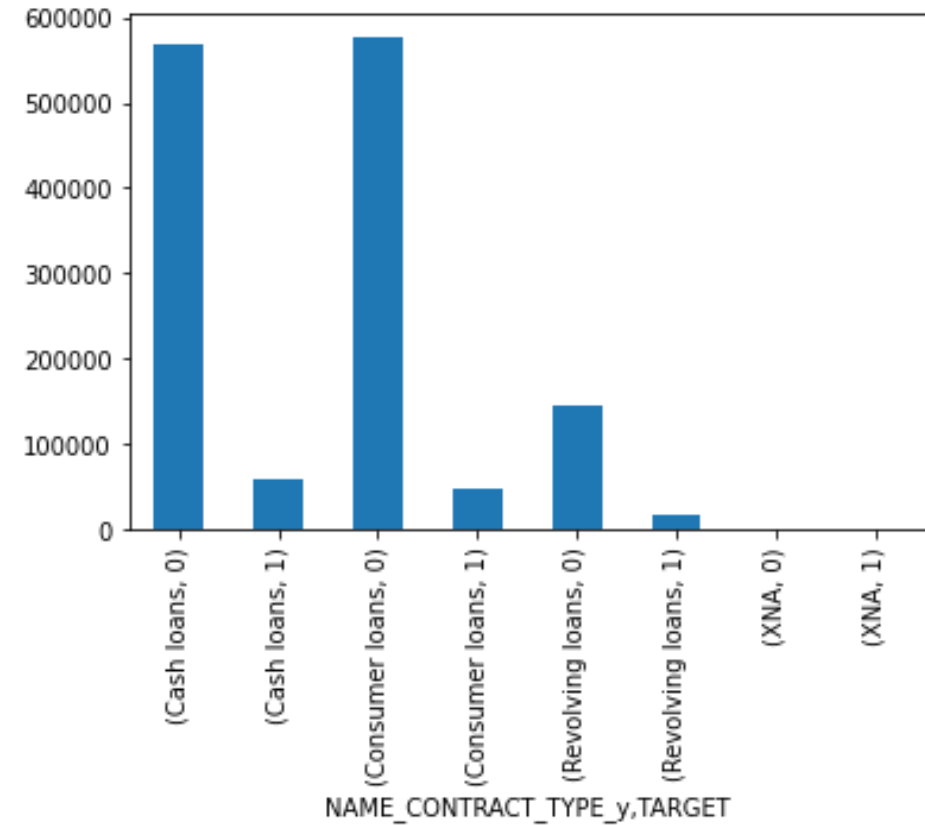Loan Providing company should give more preference to Consumer Loan as cancellation as more in Cash Loan.

Contd…

Plot

It describes the relation between TARGET and Type of Loan.

Observation

Loan providing company should be careful while lending Cash Loans.

Contd....

Plot
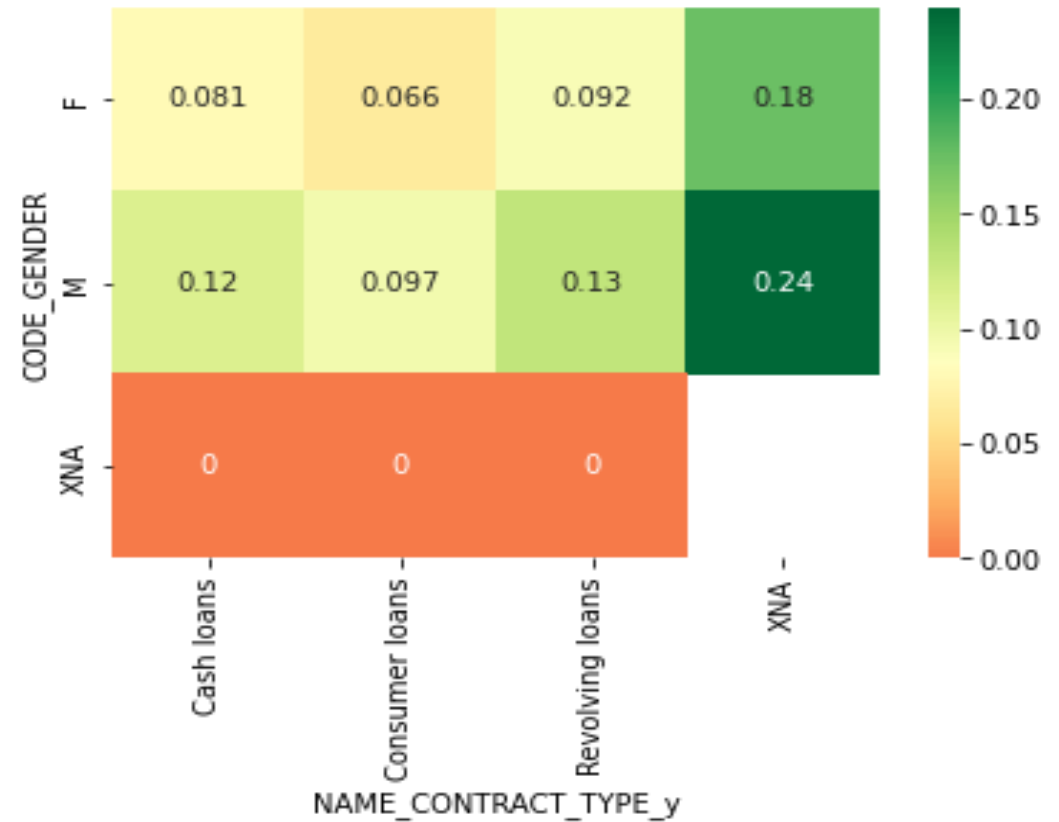
Index – Gender

Column – Loan Type

Value – TARGET

Assumption - XNA as other type of loan

Observation

Loan Providing Company should be caution in following situation

a. Male while taking other type of loan.

b. Female while taking other type of loan.

c. Male while taking cash loan.

Contd…..

Plot

Index – Gender

Column – Loan Approved Amt in Range

Value – TARGET

Assumption – XNA as Transgender

Observation

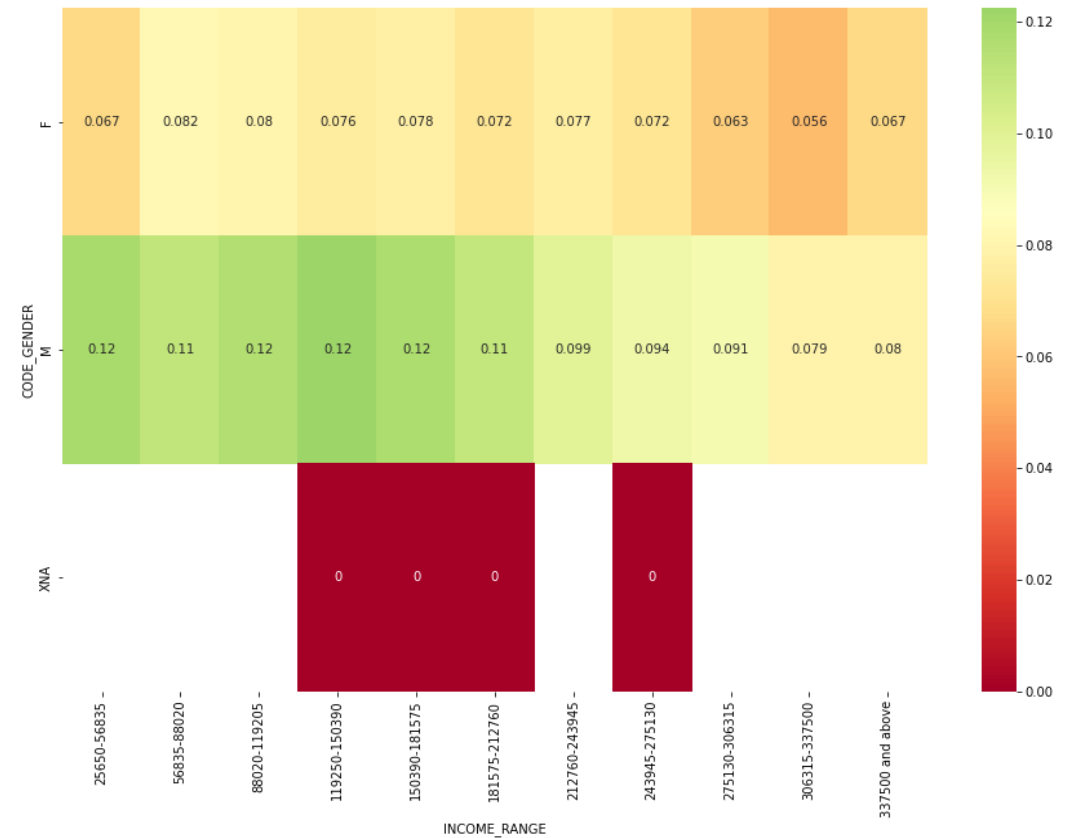Female applicant in the Credit range of 0-10000 have defaulted more.

Contd……

Plot

Index – Gender

Columns – Income Range

Value – TARGET

Observation

Male applicant at lower income range are more prone to default.
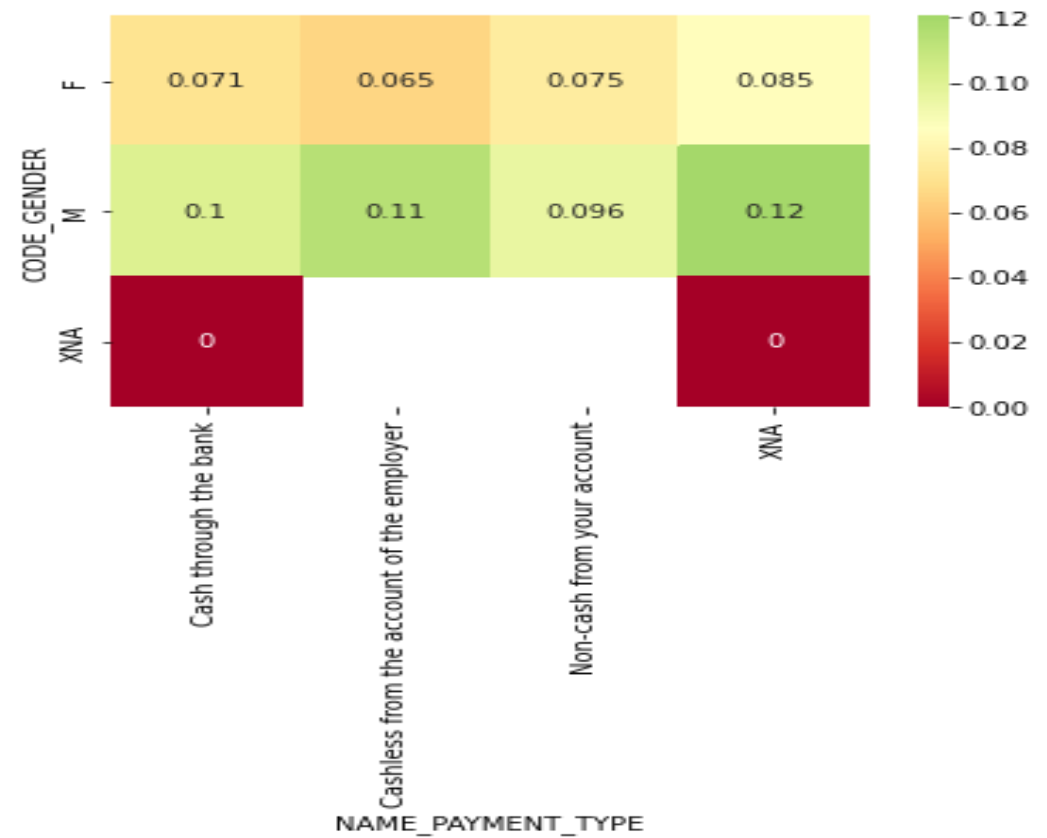
Contd…….

Plot

Index – Gender

Column – Payment Type

Value – TARGET

Payment – XNA as all other type of payment.

Observation

Loan Providing company should be caution when Male applicant Payment mode is not clear.

Conclusion

- Loan providing company should extra precaution step for applicant with Income Range of 88020-119205 with Education Type as Secondary or secondary special.

- There are high percentage of people from Working group who can default on loan. The Pensioner group has less people with payment difficulty as compared to non-payment difficulty.

- Loan Providing company should give more preference to Consumer Loan as cancellation as more in Cash Loan.

- Loan providing company should be careful while lending Cash Loans.

- Female applicant in the Loan Credit range of 0-10000 have defaulted more.

- Male applicant at lower income range are more prone to default.

- Loan Providing company should be caution when Male applicant's Payment mode is not clear.

Thank You