Bank Loan CASE STUDY

SUBMISSION
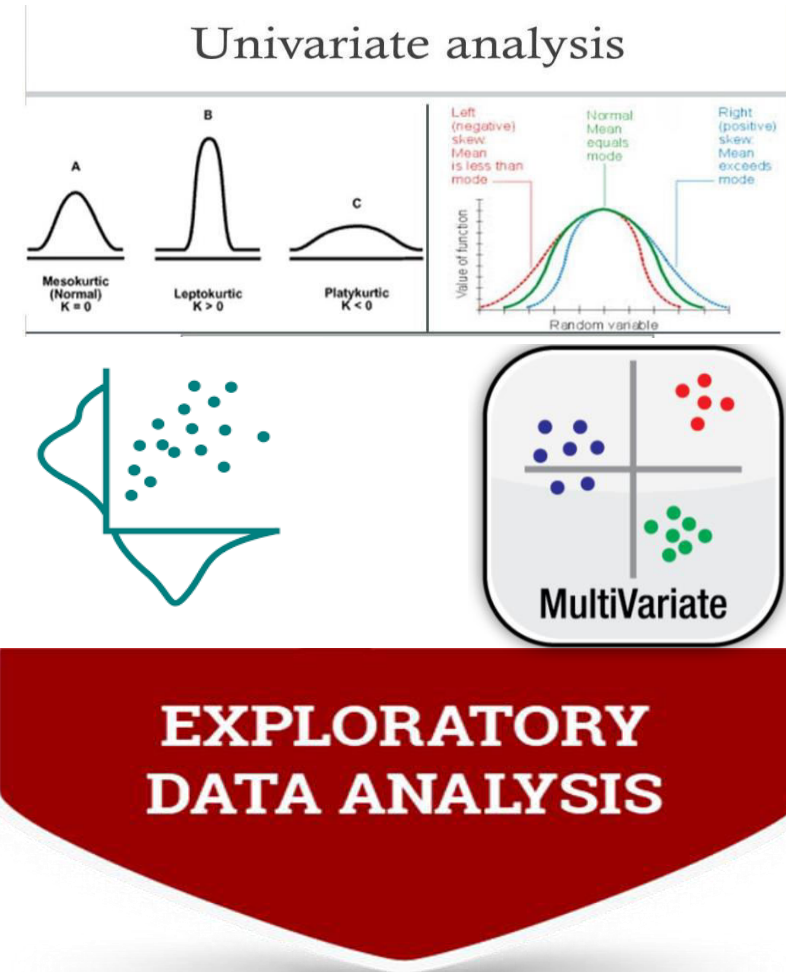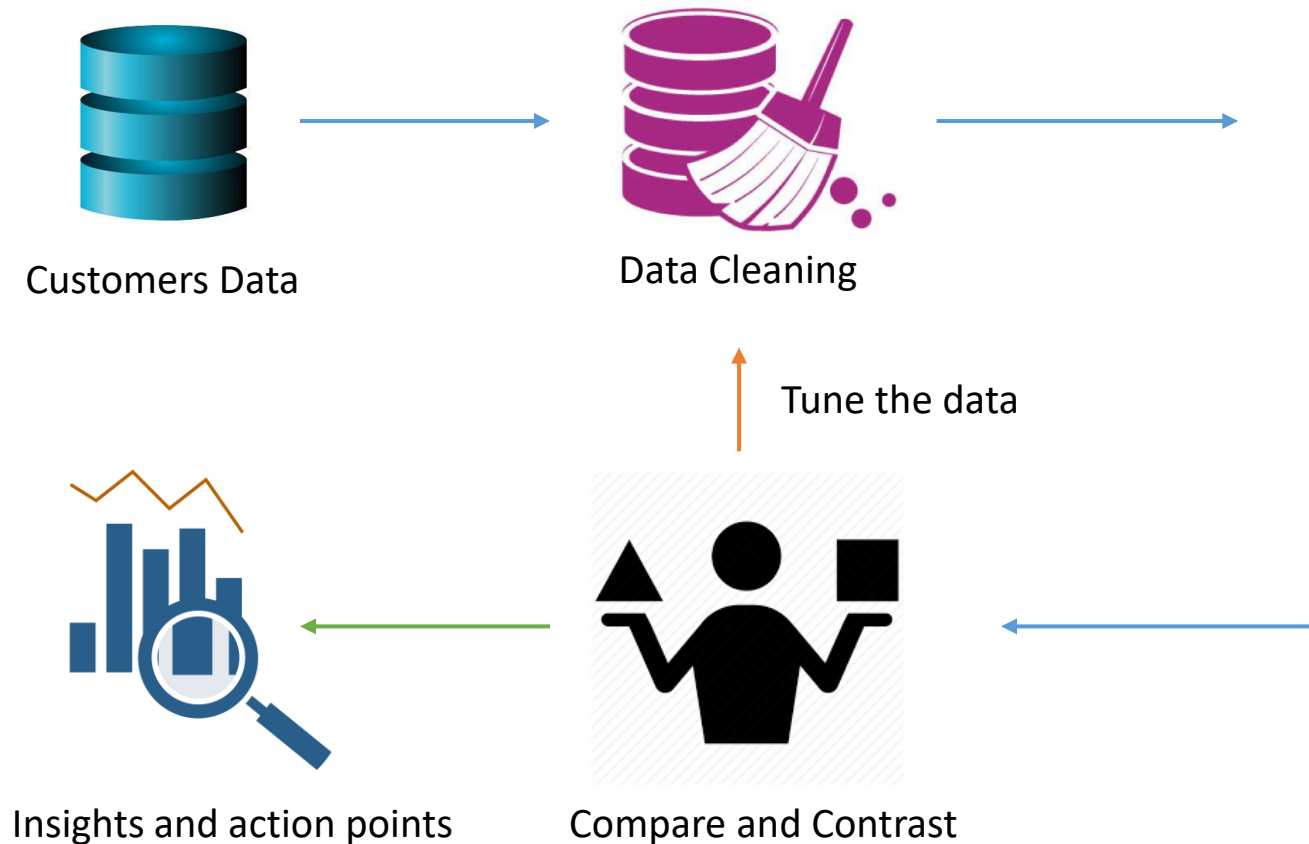
Name: Midhun M Thomas

Abstract

When the Bank receives a loan application, the bank has to make a decision for loan approval based on the applicant's profile. As a Bank risks are associated with the bank's decision and that will really affect the business directly.

Using univariate, segmented univariate, and bivariate analyses to derive business-relevant insights, supported by visualizations, to explain the importance of variables in distinguishing clients with payment difficulties from others.

The outcome of this study will provide a clear understanding of which customers are likely to repay their loans and which ones may default.
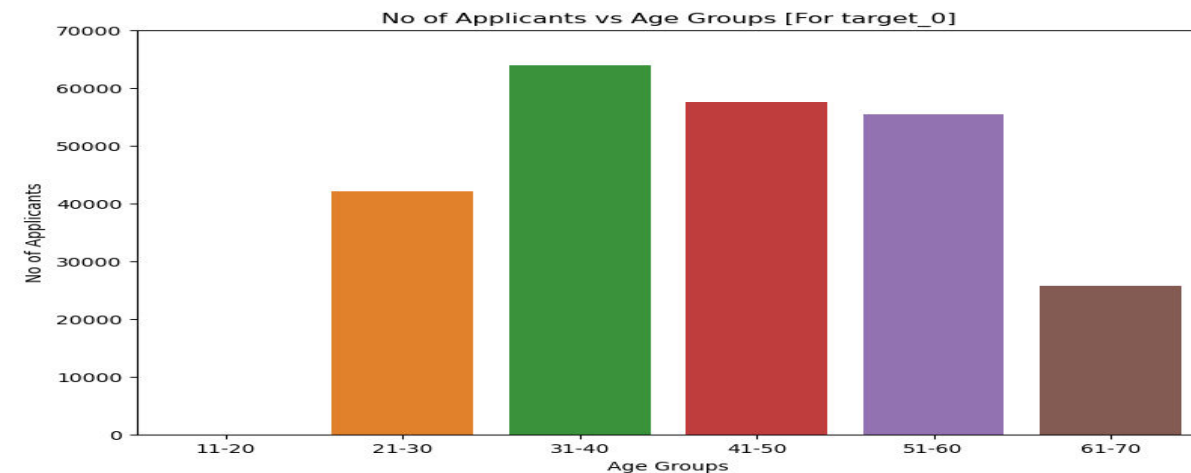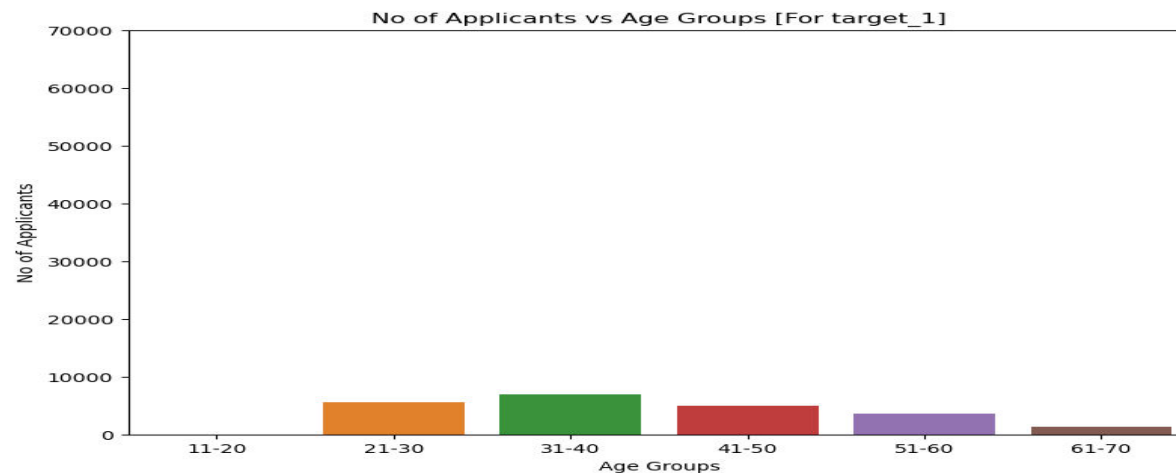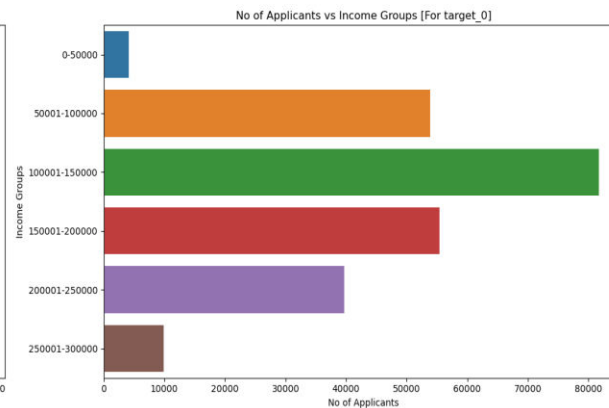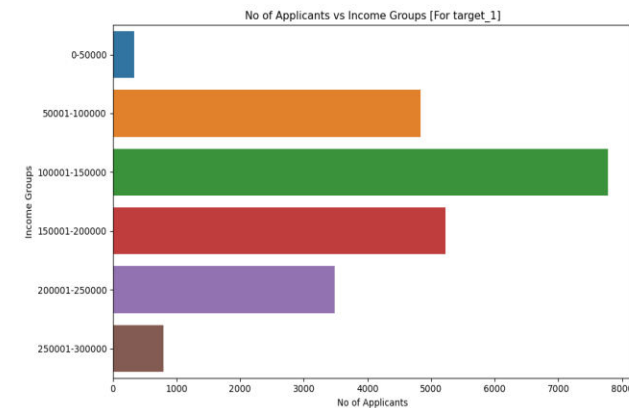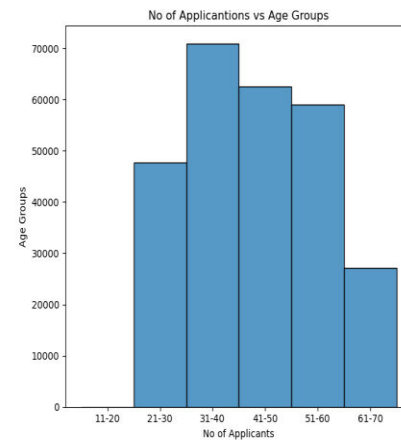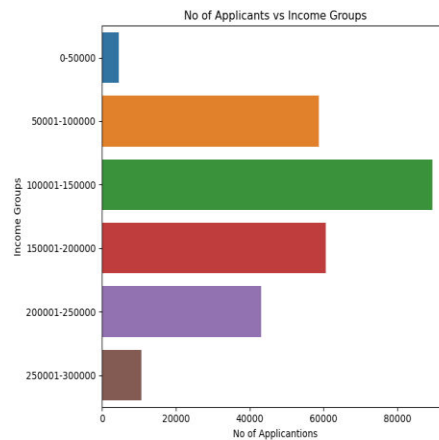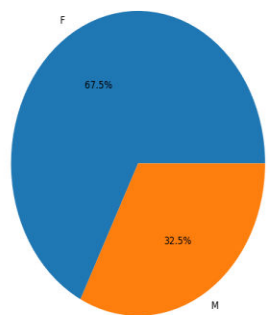
Problem solving methodology

Customers Data

Data Cleaning

Tune the data

Insights and action points

Compare and Contrast

Univariate analysis

MultiVariate

EXPLORATORY DATA ANALYSIS

Analysis Approach

Univariate Analysis

Univariate Analysis on segmented Data

Multivariate Analysis
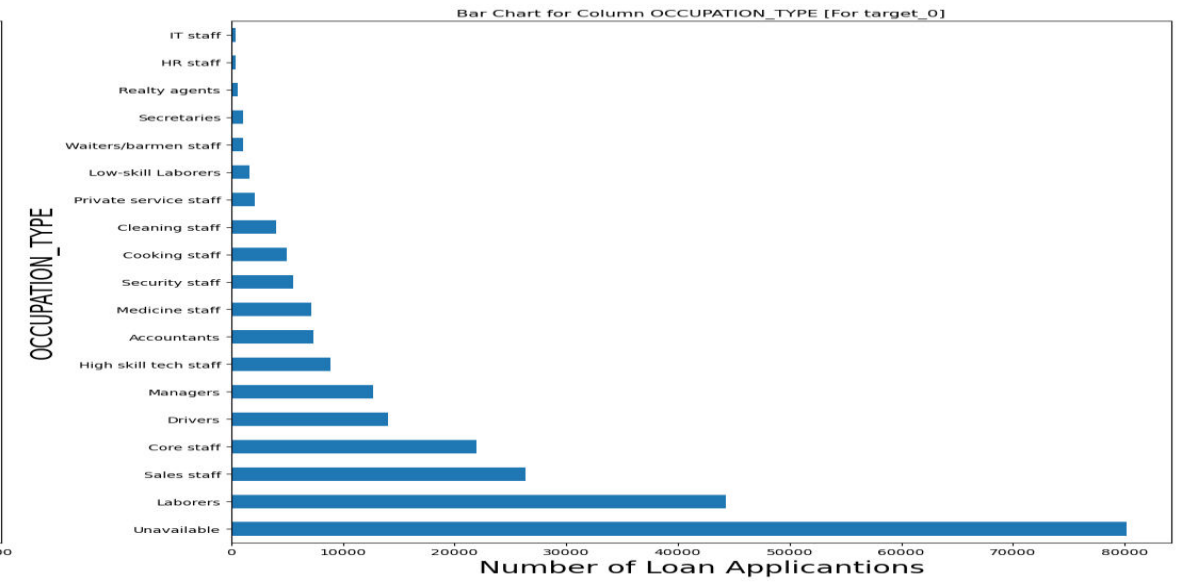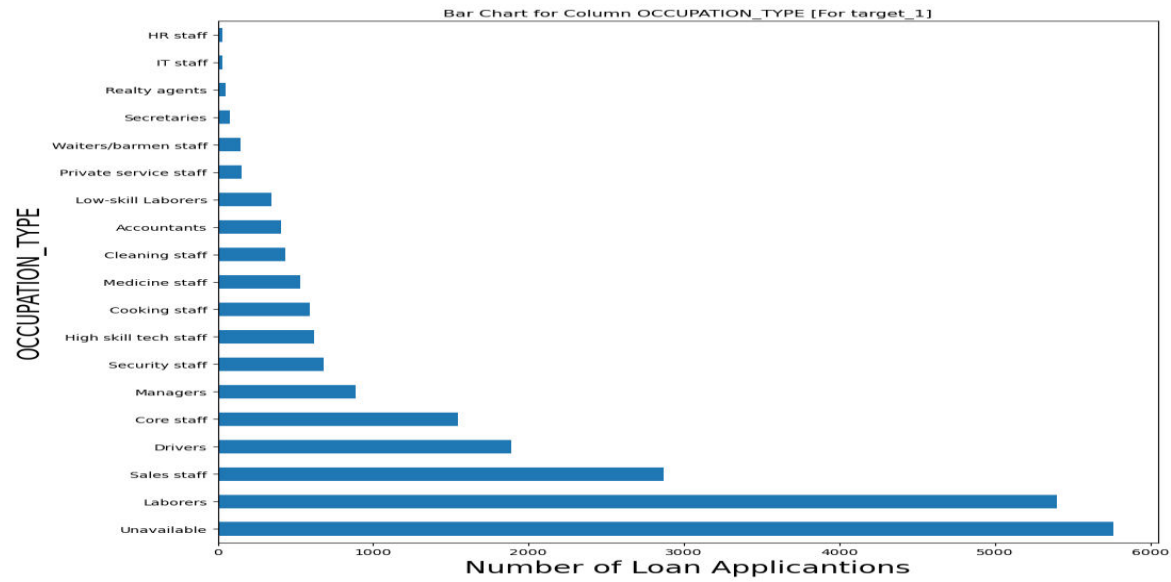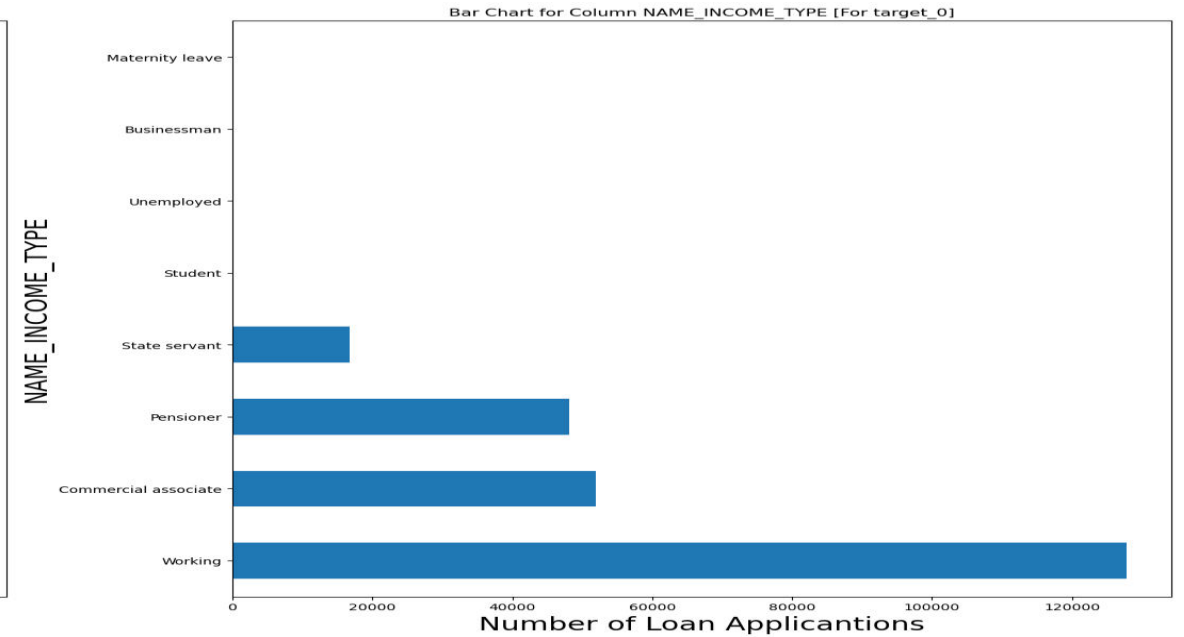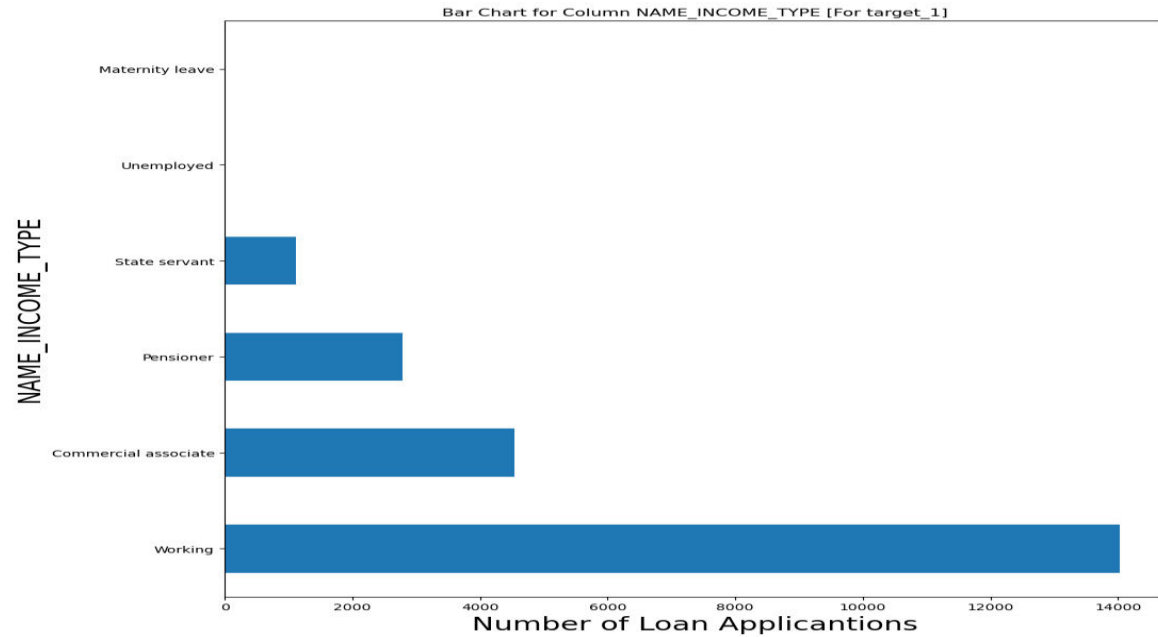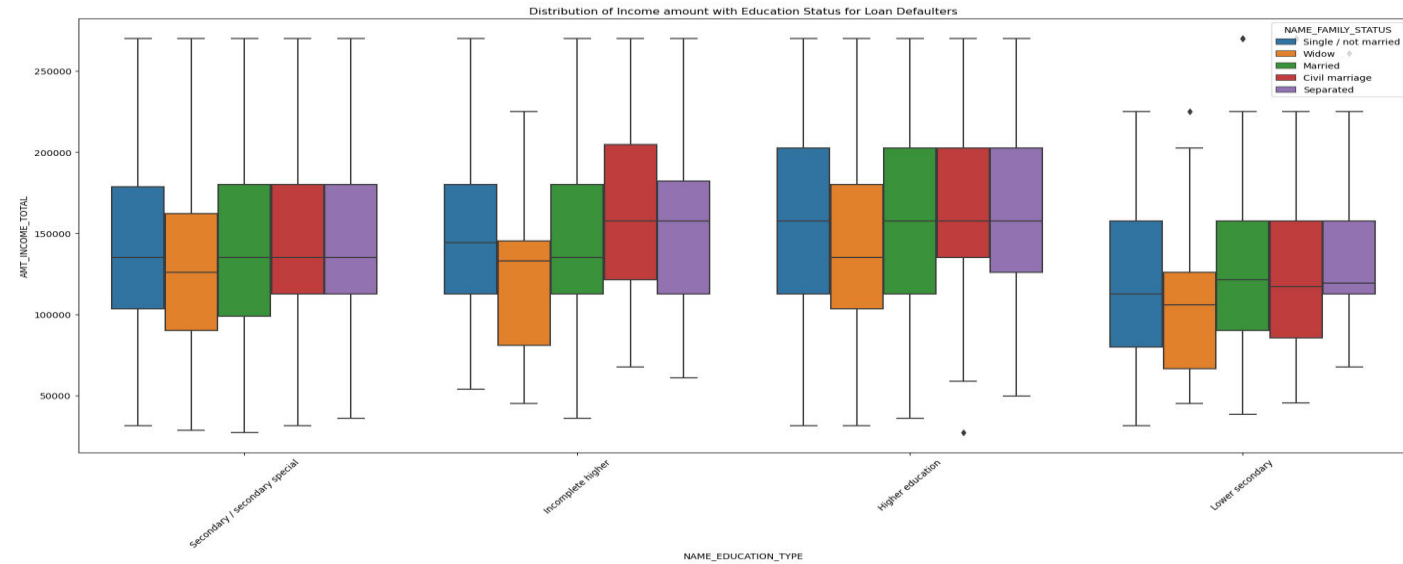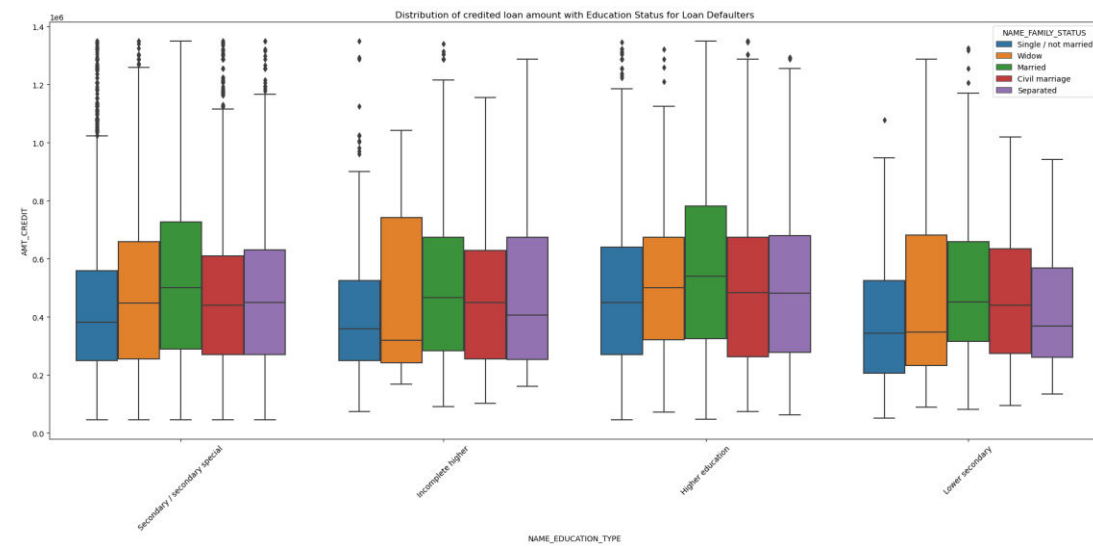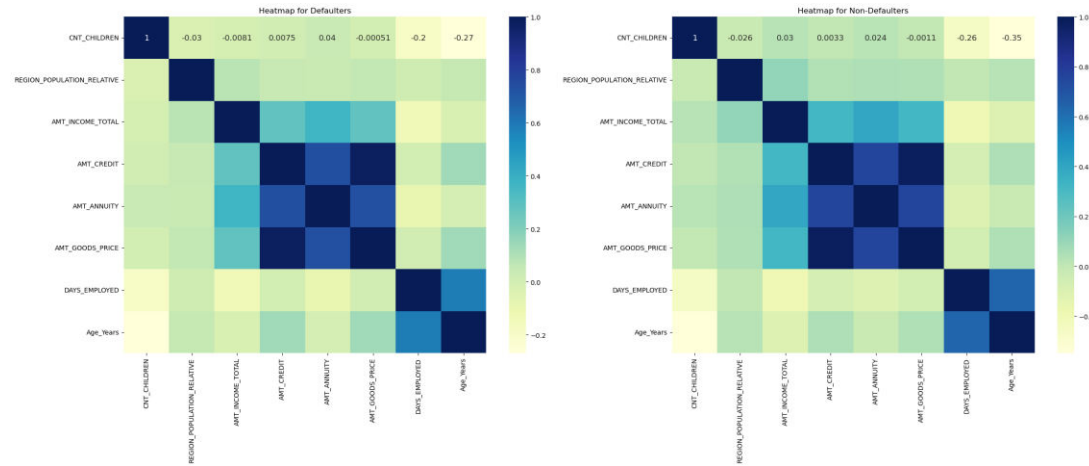
Analysis Approach

What is the approach on the avail be data :

     a. **Univariate Analysis** : Find the pattern from any given parameters defined in the previous slide mark any applicant profile has risk factors of loan default.

     b. **Bivariate Analysis :** In this technique, instead of considering a single parameter for the analysis, multiple variables will be used. Some variables can impact on multiple variables and this is very curtail for the business.

     c. **Multivariate Analysis:** In bivariate only two variables relations are measured and when multiple variables changes in same direction or different with any of the given variable, this could be a flag for the business to intervene.

     d. **Categorical Analysis :** Any categories in the applicants data can be analyzed to get a glance it.

Observations

Insights Application Data

- 1. Married individuals with higher education backgrounds have successful loans with higher ranges of credited loan amounts.

- 2. Married individuals with higher education backgrounds have successful loans with higher ranges of income amounts.

- 3. Individuals with academic degrees have the highest success rate in loans.
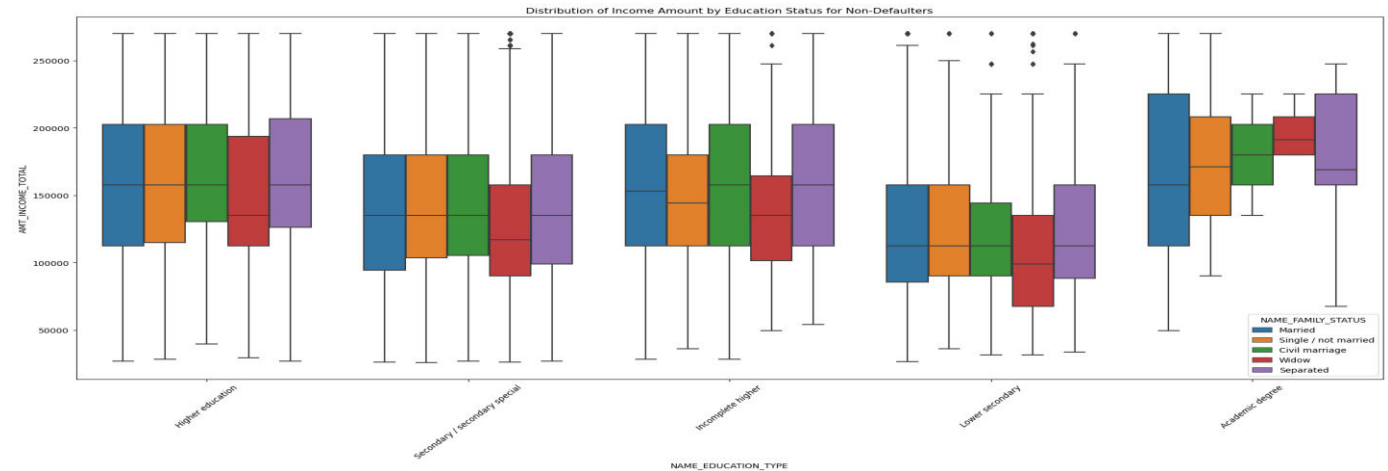
Insights Merged Data

1. When customers do not specify the purpose of their loan, the bank approves fewer applications.

2. The bank rejects more loan applications than it approves for all stated purposes.

3. The data on loan purposes is unclear, as there are a significant number of applications that were both approved and rejected for unspecified

Income amount vs Loan Status

Prev Credit amount vs Housing type
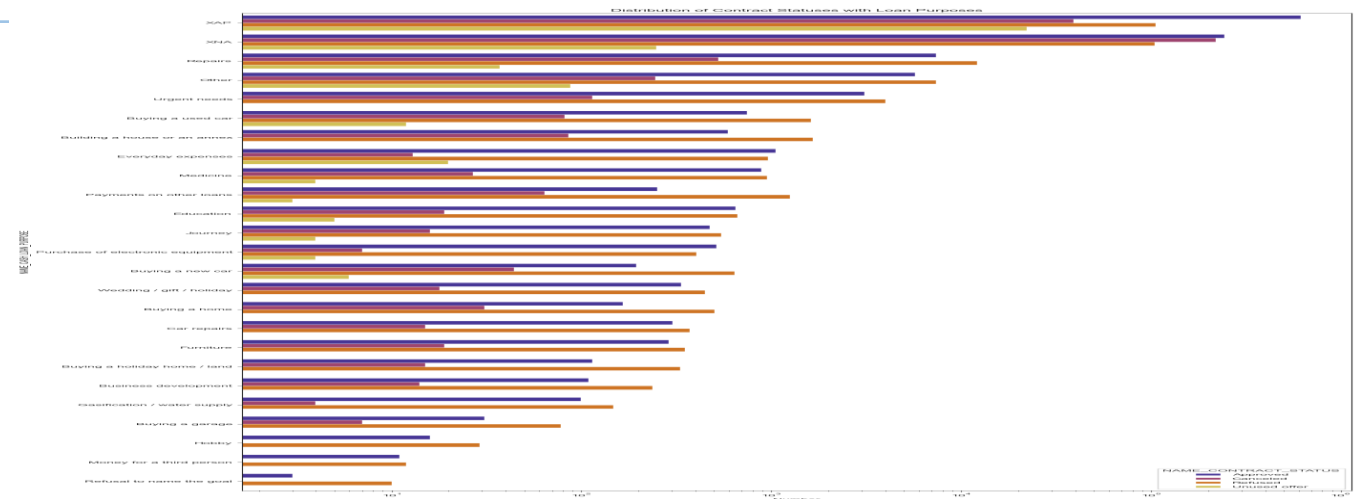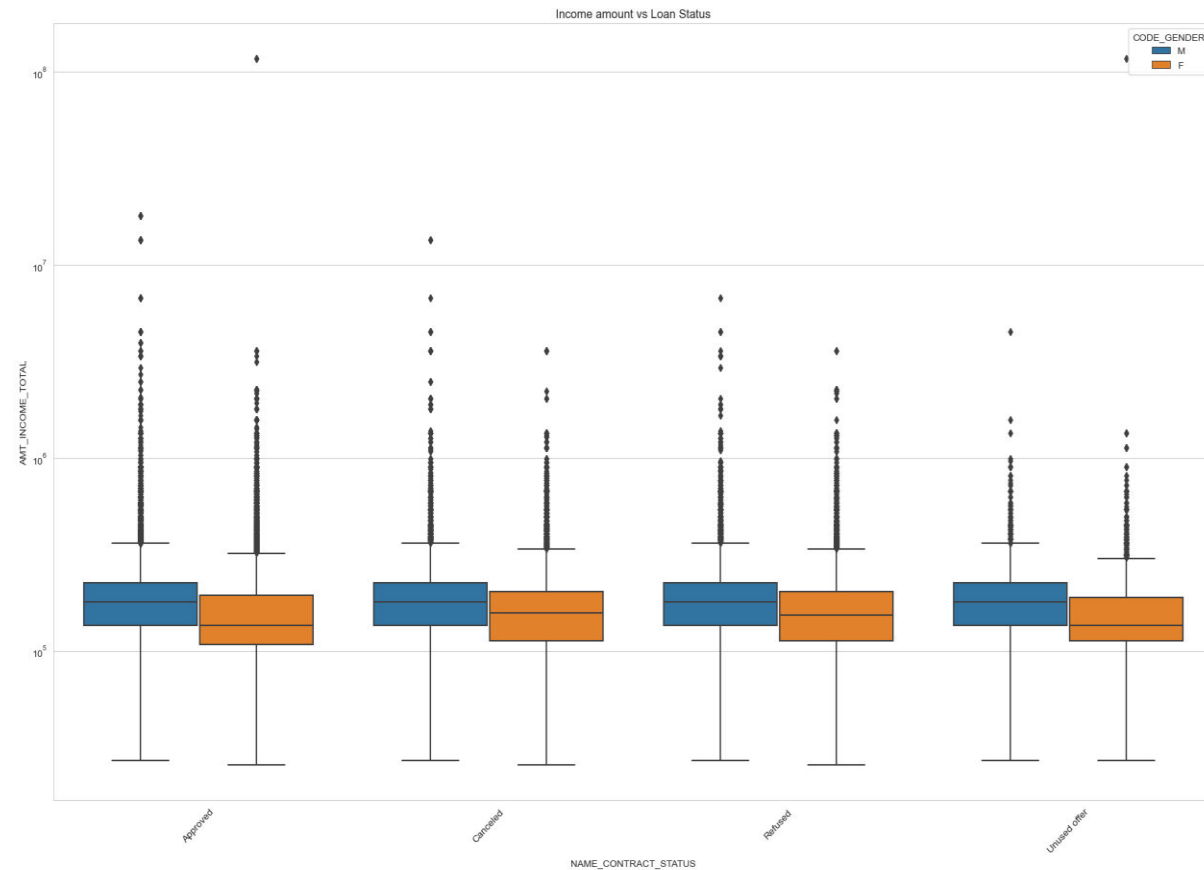
Insight: Boxplot

1. For male customers, the median income amount is the same across approved, canceled, refused, and unused loans.

2. Male customers have slightly higher incomes than female customers across all loan contract statuses.

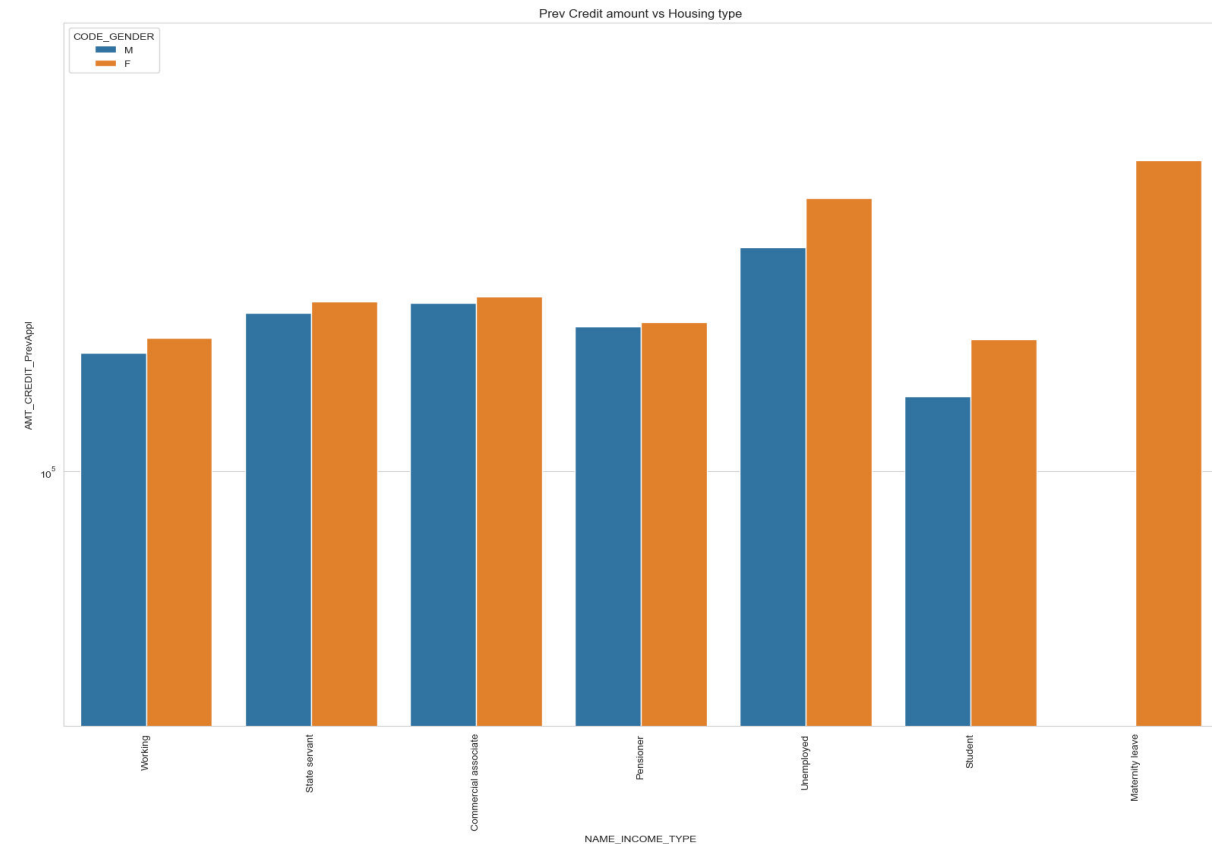3. The distribution of income amounts is nearly identical across all loan contract

Insight: Bar Chart

1. Females received higher loan amounts than males across all occupation types.

2. Unemployed females had the highest loan amounts in their previous loan applications

3. Maternity leave and student income types also show that females received higher previous credit amounts compared to males.

Heatmap for Defaulters
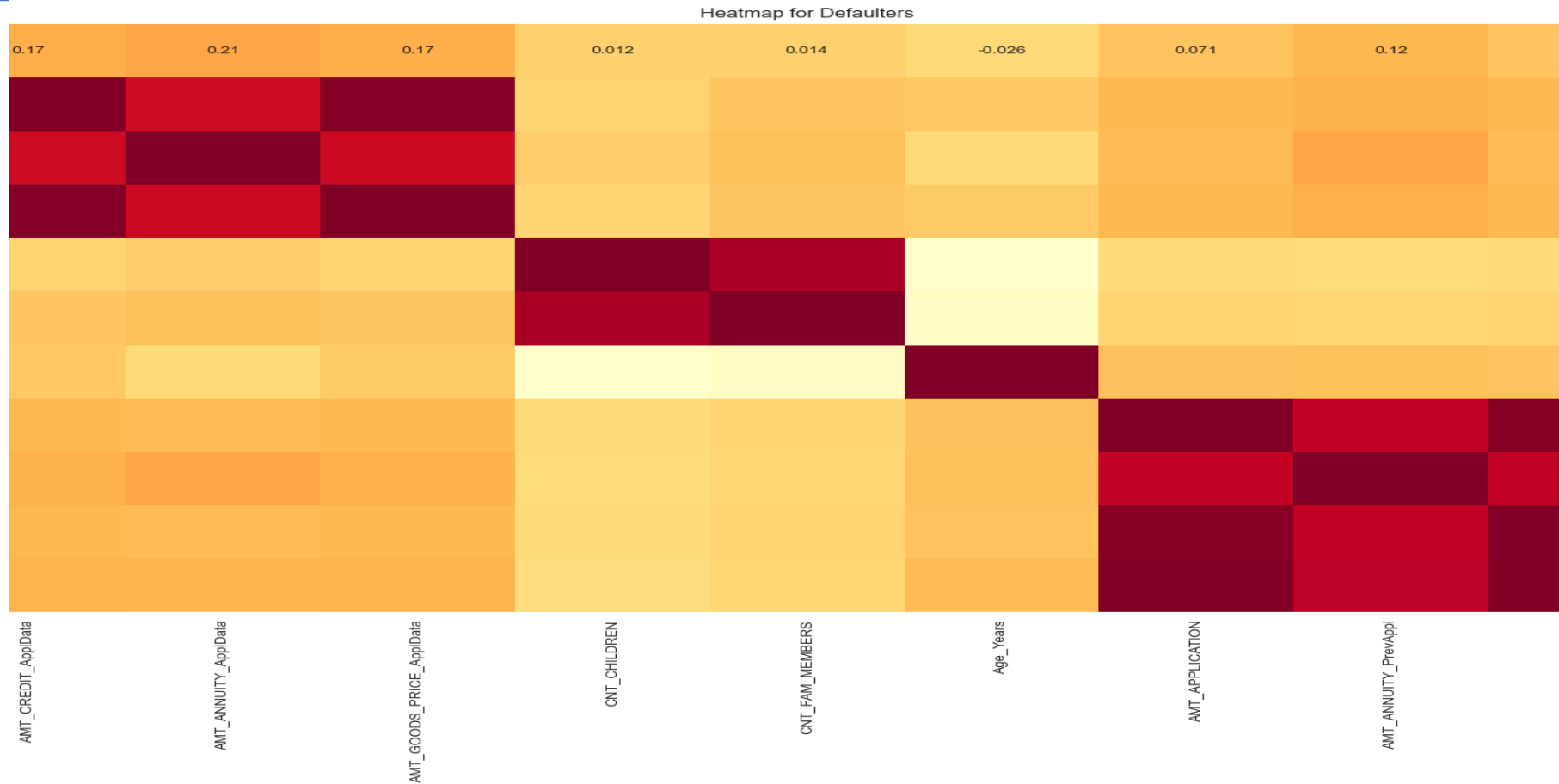
Insight :

1. The heatmap shows a strong correlation between the Goods Price and the Loan Amount of the current credited application. There is also a strong correlation between the Loan Amount requested by the applicant and the Loan Amount credited in the previous application.

2. Conversely, there is a weak correlation between age and the number of children, as well as between the Goods Price, the loan amount credited, and the loan annuity of both current and previous applications.

- There were a number of insights gained while analysis the given dataset some of the insights from the analysis are,

- The percentage of females is higher compared to males in loan applications.

- Higher loan applications have come from applicants with income group 100001-150000.

- People in age group 31-40 have applied the highest number of loans.

- Proportion of females is higher than males in both defaulters and non-defaulters.

- Proportion of both defaulters and non-defaulters not having a car is higher than those who have it

Findings and Recommendations

The following paraments with values are the most riskiest for considering a loan:

Action: Do not approve loan before further investigation

a. Since there are more defaulters in XAP and XNA, the bank should concentrate more on loans purpose.

b. Instead of lending to the unemployed, the bank should target the employed and pensioners

c. Banks should decrease their emphasis on the 'Working' income type, given its highest frequency of unsuccessful payments, in order to minimize financial losses for the organization.

Banks should prioritize targeting the contract types 'Student', 'Pensioner', and 'Businessman' for more profitable business opportunities.

Action: Eligible for the Loan

a. All other values are good to go