Credit EDA _ Case Study

BHARATHISESHAN NAGARAJAN

(Email: nbseshan@gmail.com)

Problem Definition

▶ The difficulty lies in the loan approval process for individuals with insufficient or nonexistent credit history. This makes it challenging for loan providers to accurately assess the creditworthiness of applicants.

▶ In certain cases, individuals exploit this gap in credit history assessment by deliberately defaulting on their loans. This poses a significant financial risk for lending institutions, as approving loans for such applicants may lead to potential financial losses.

Data Cleaning

▶ **Dealing with missing values:** Columns with missing values more than 50% are dropped.

▶ **Finding the number of unique values in each column:** Converting the data type of columns having less than 58 unique values to 'object'. Converting the data type of column SK_ID_CURR, REGION_RATING_CLIENT_W_CITY to object and converting CNT_CHILDREN to int

▶ **Dropping unwanted columns**

▶ **Dealing with XNA values by either imputation or creating a separate category.**

▶ **Removing the missing values in certain columns having missing values less than 1000 or slightly more**

Outliers Handling
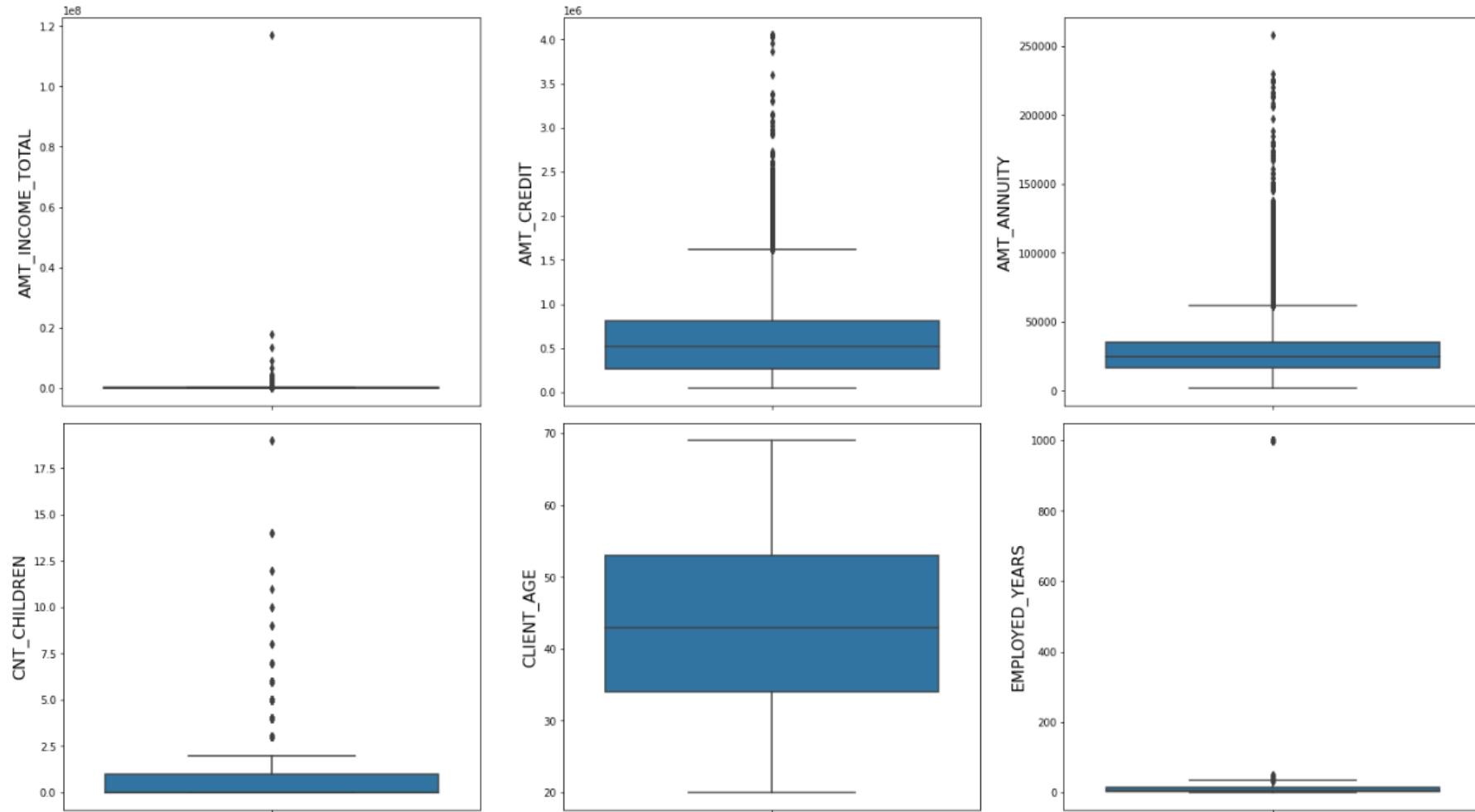

Box Plots for checking outliers

Defaulting Clients by Gender

Percentage of Defaulting Clients by Gender

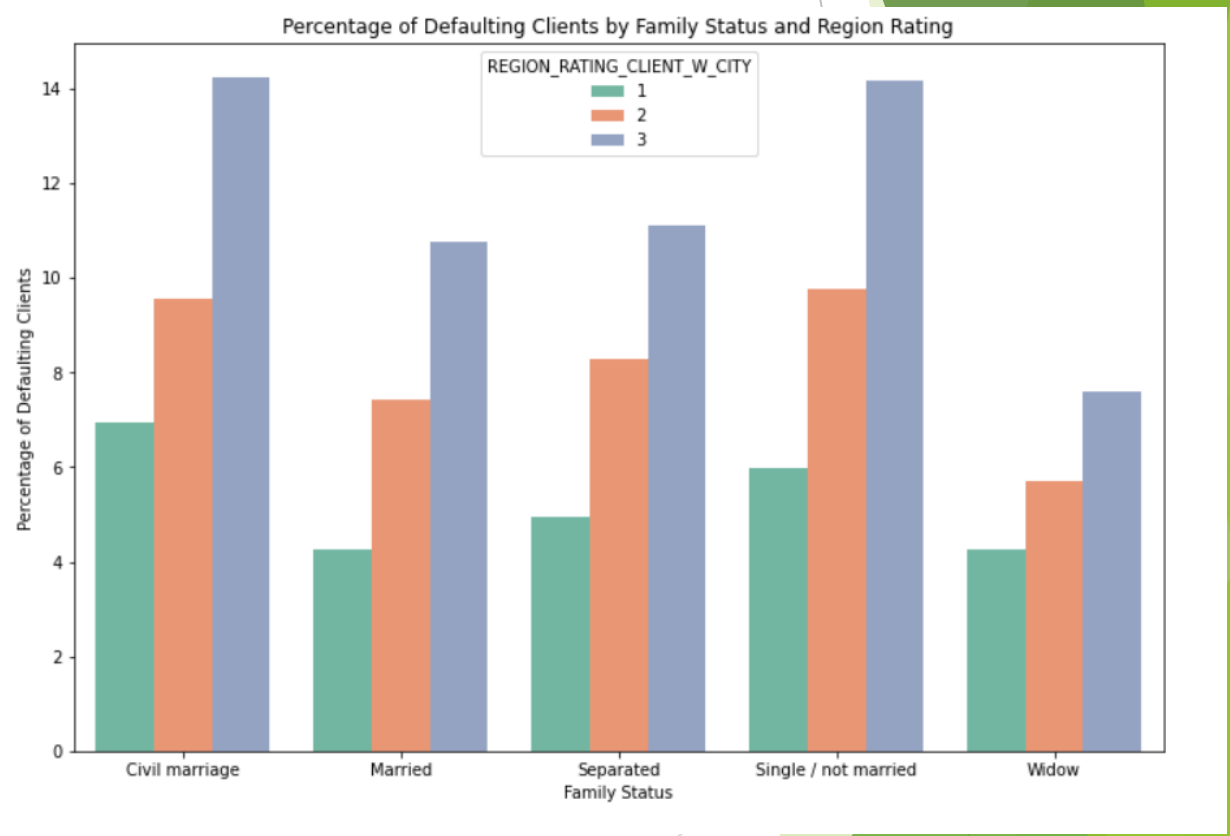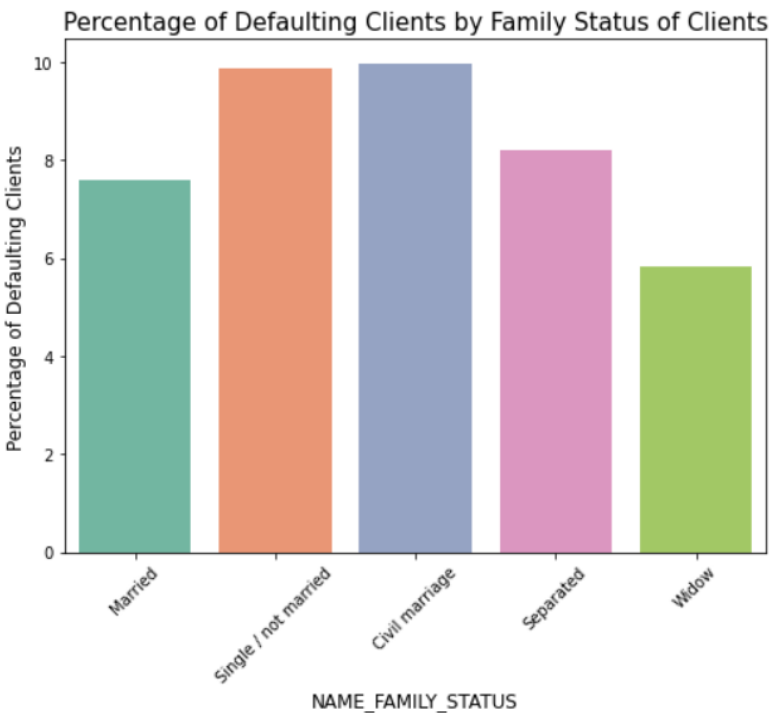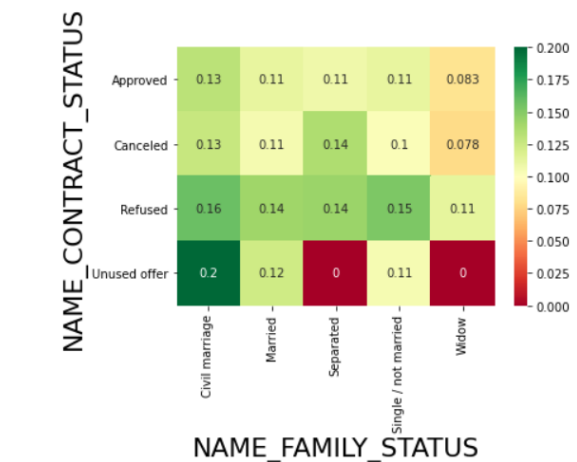Percentage of Defaulting Clients by Salary Category Genderwise
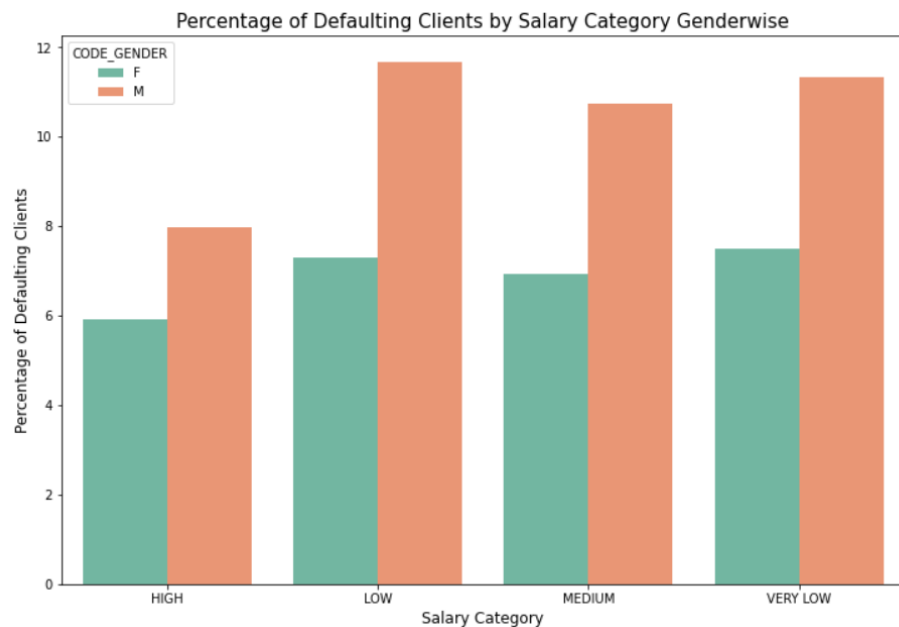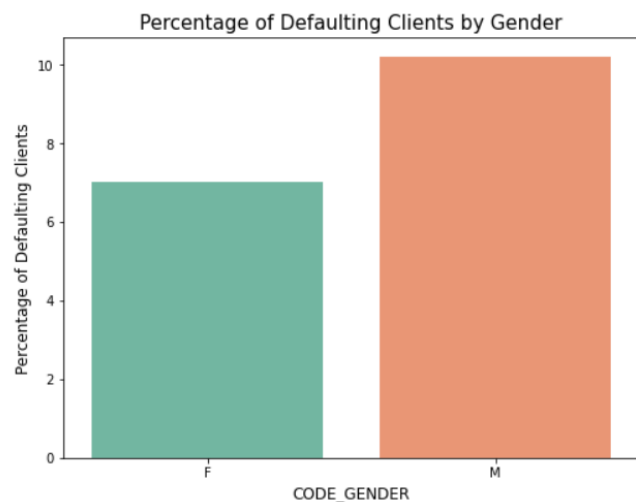
▶ ***The data indicates the following key inferences:***

1. Female clients generally pose a lower risk for the bank in comparison to their male counterparts.

2. Clients with a history of previous refusals are a highest risk for the bank.
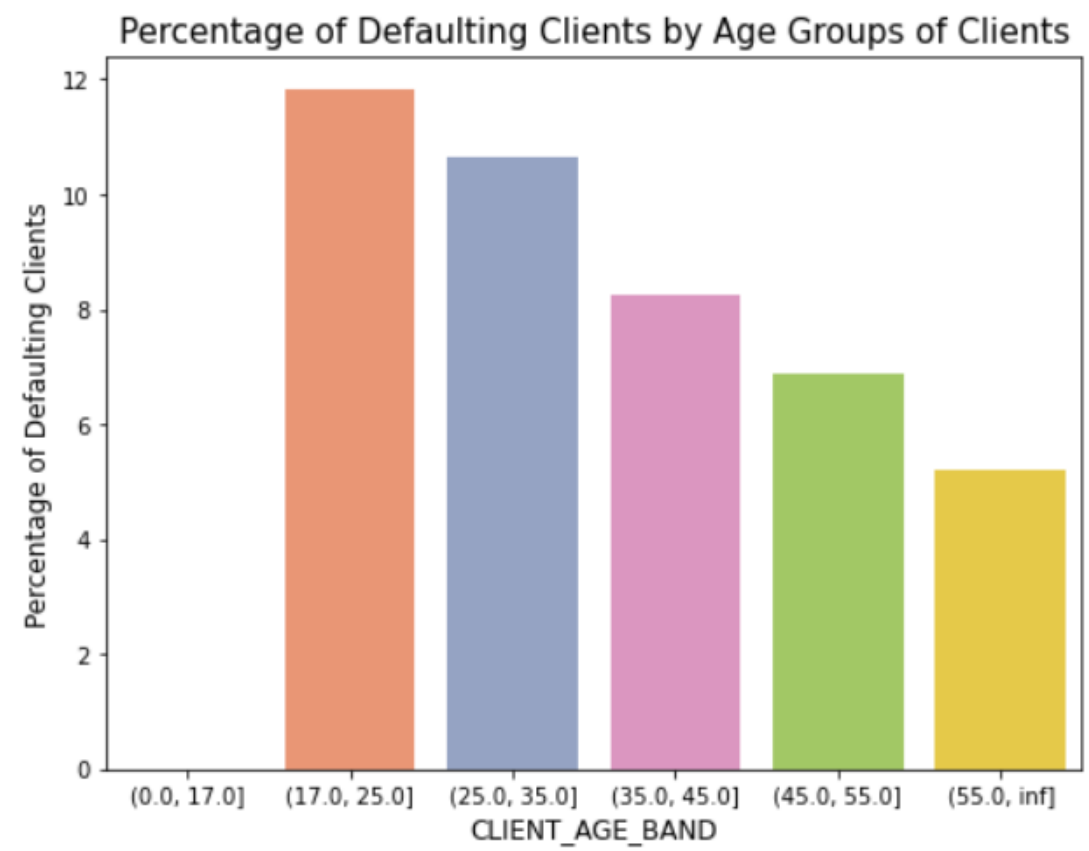

Heatmap of Gender and Contract Status based on Target
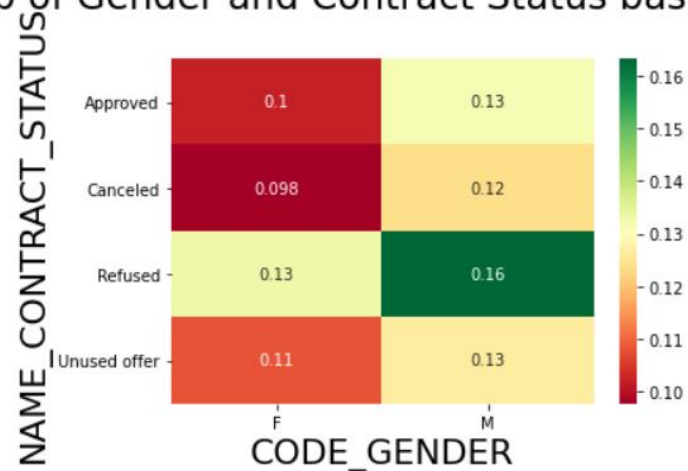
Defaulting Clients by Family Status

Percentage of Defaulting Clients by Family Status and Region Rating

Percentage of Defaulting Clients by Family Status of Clients

Defaulting Clients by Age Band

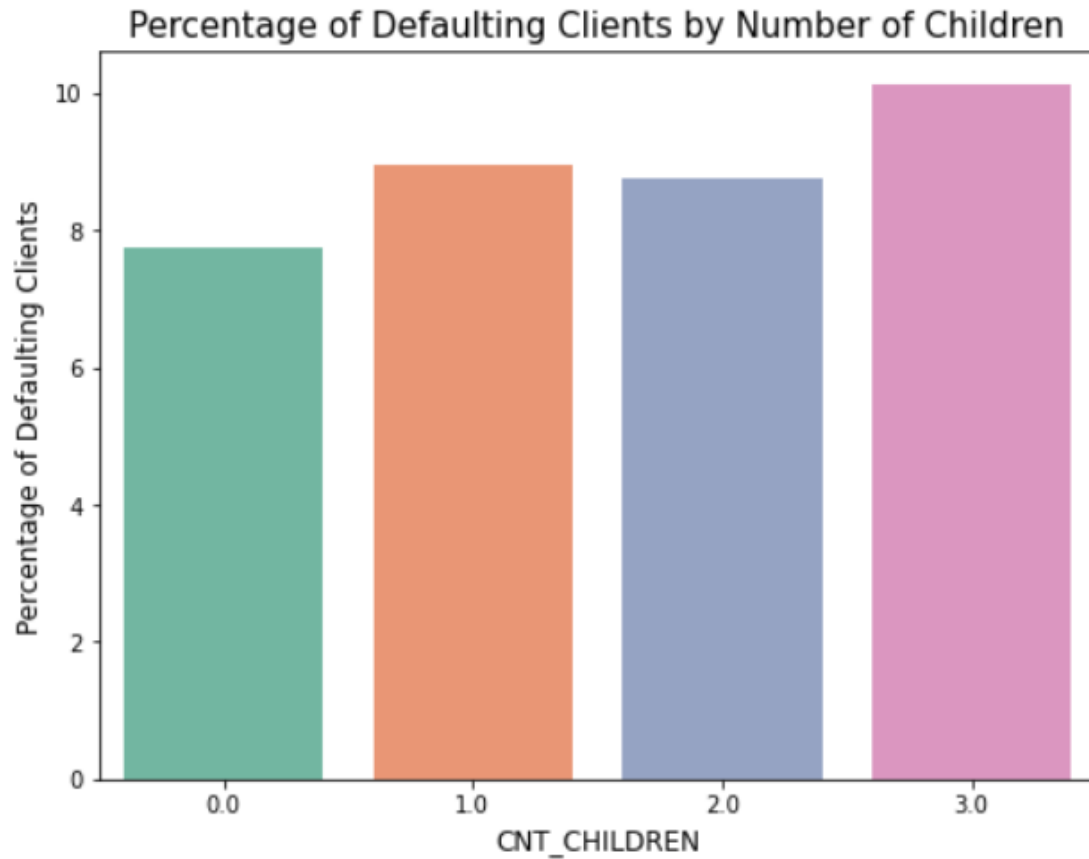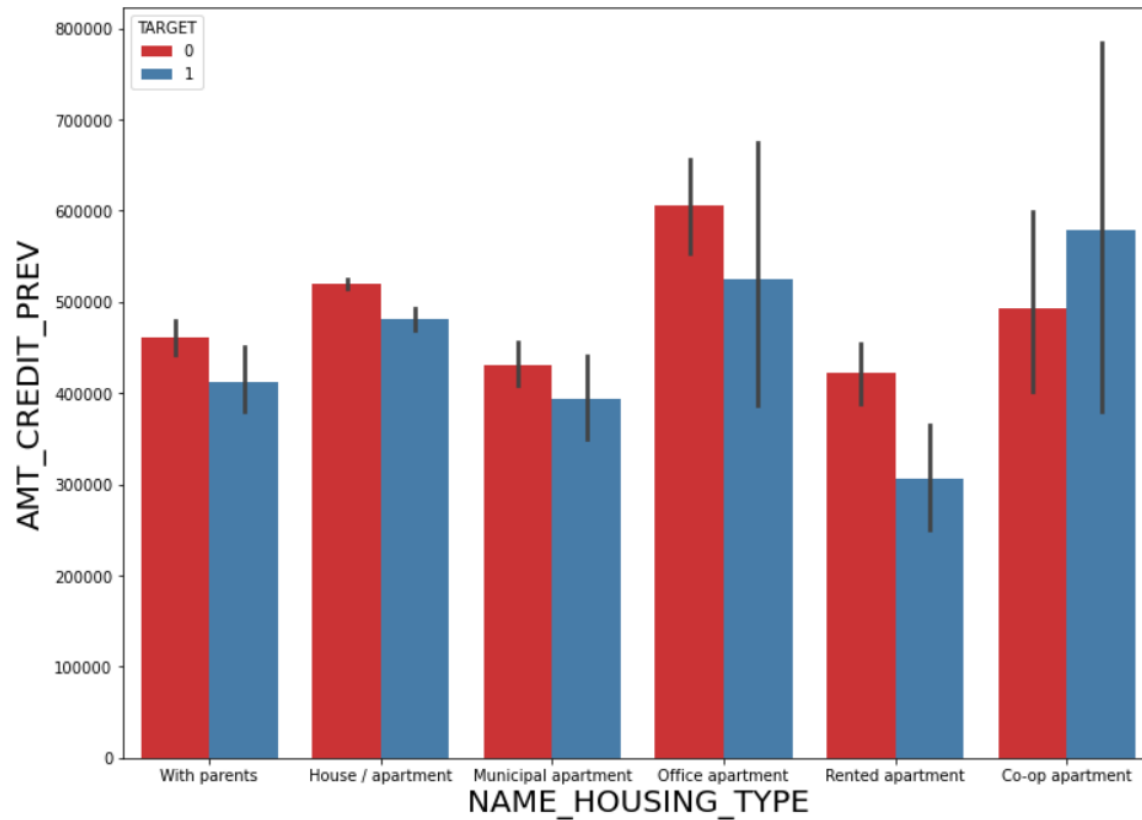Percentage of Defaulting Clients by Age Groups of Clients
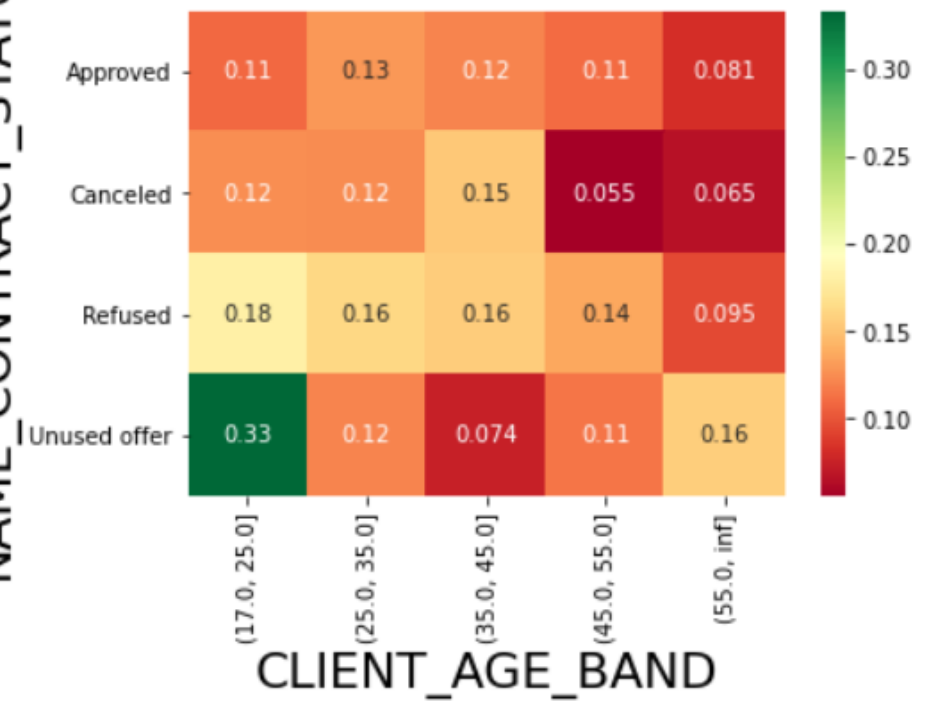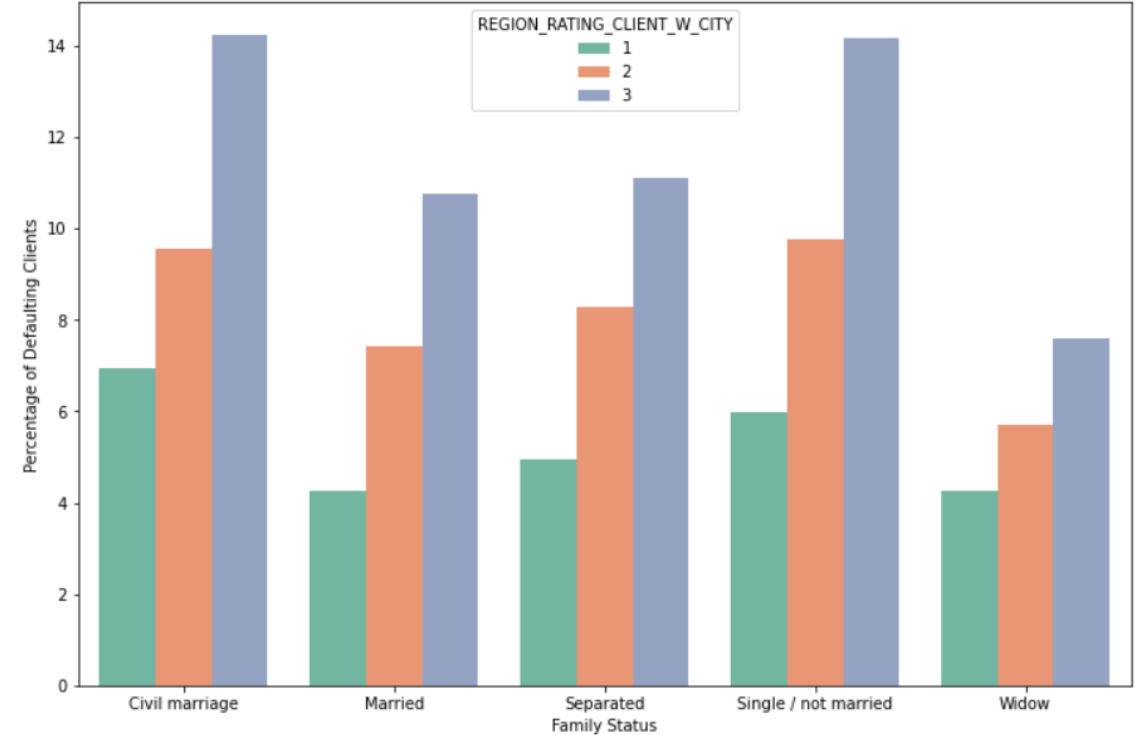
Defaulting Clients by Number of Children
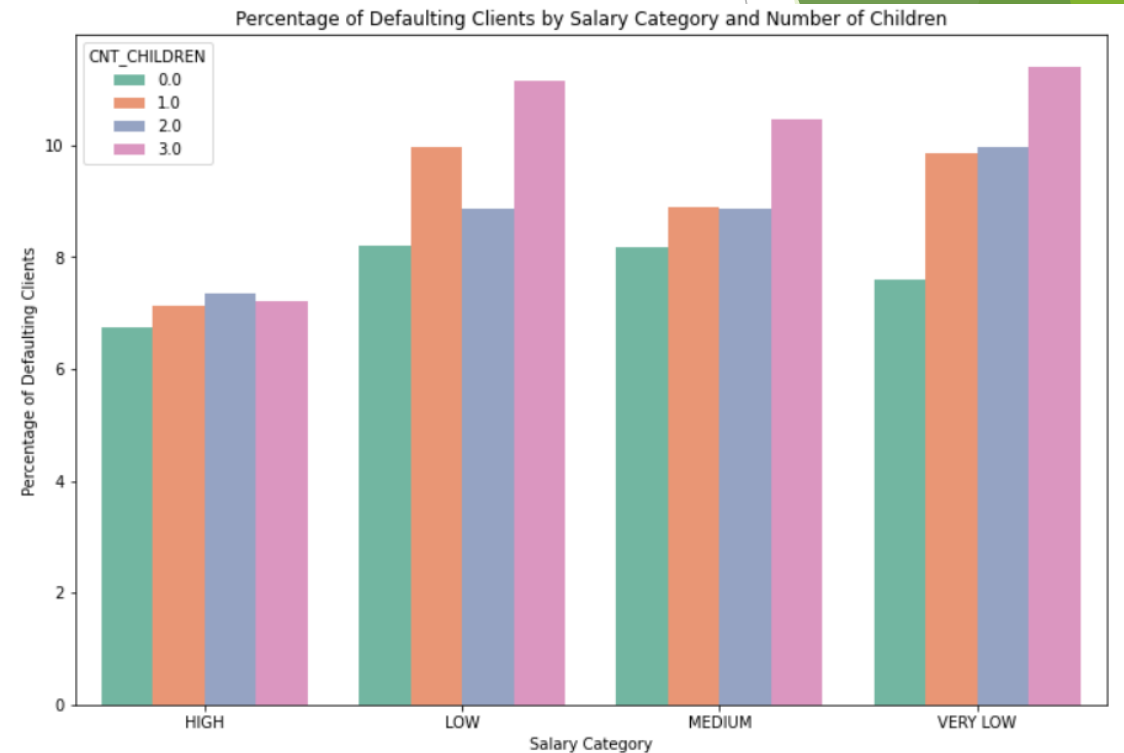
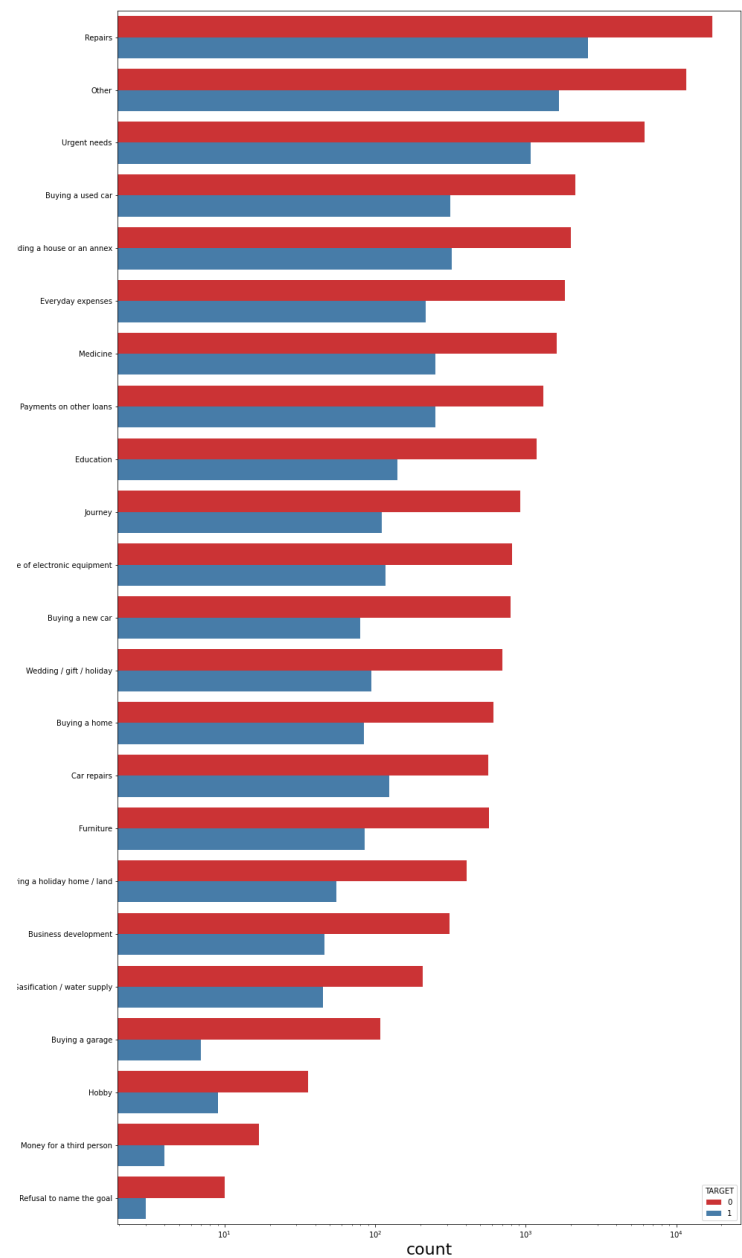# Defaulting Clients by Region Rating

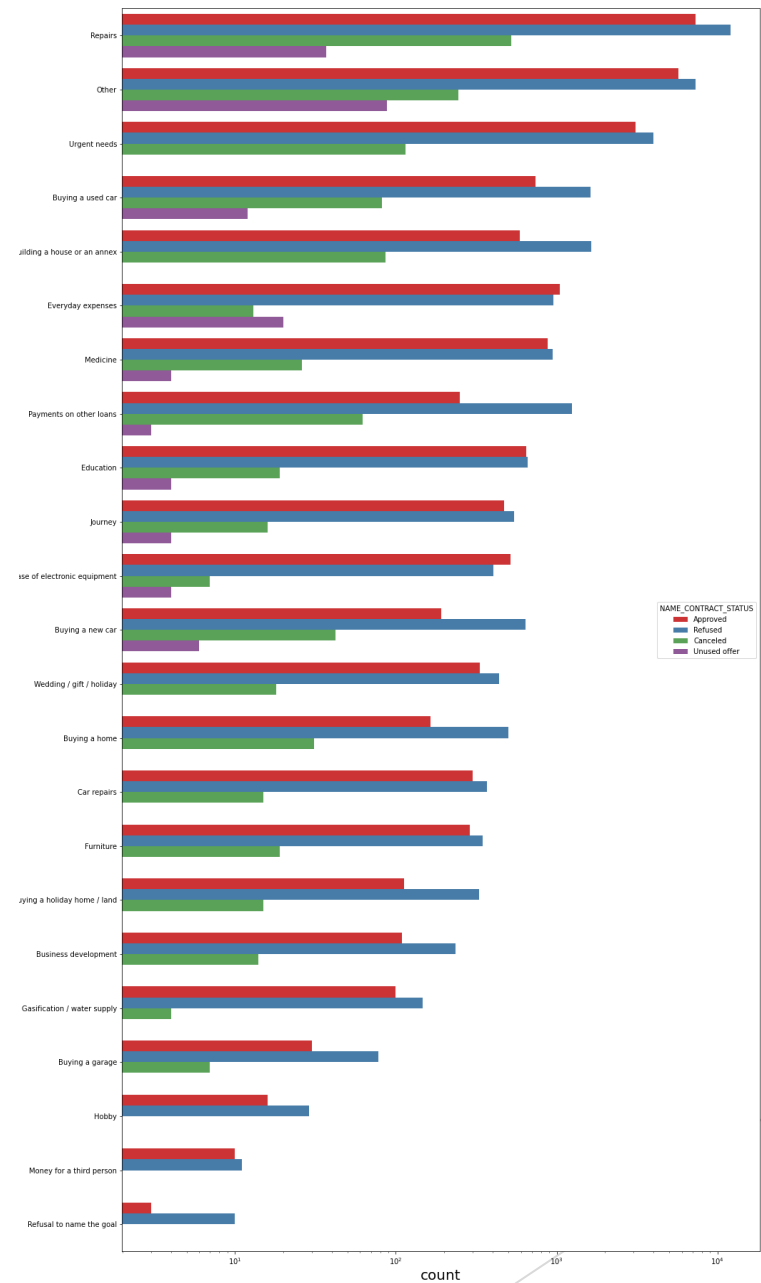Prev Credit amount vs Region Rating of Client City

Percentage of Defaulting Clients by Family Status and Region Rating

Distribution of purposes with target

Distribution of contract status with purposes

Percentage of Values by Name Client Type
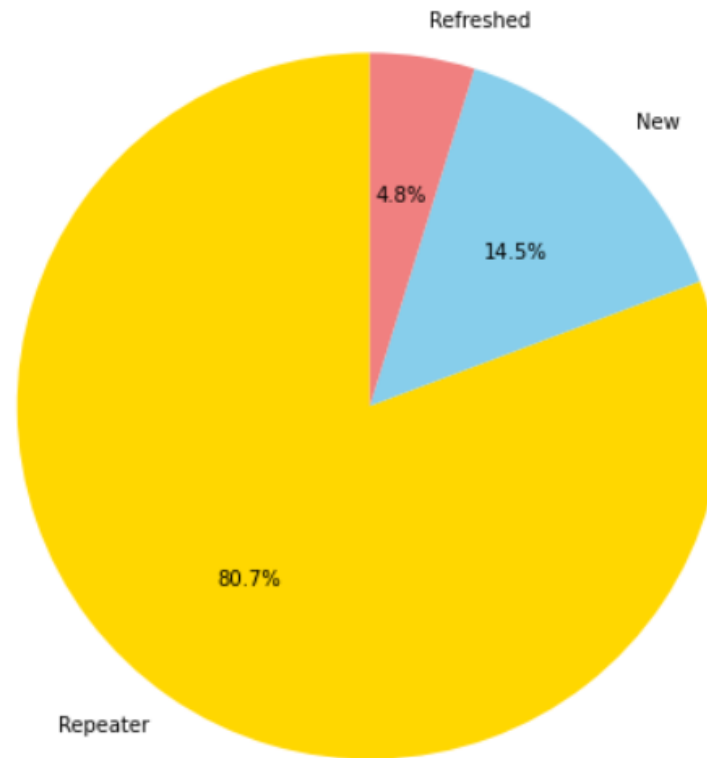
Observations

▶ We can see that clients living in cities with rating 3 have the highest percentage of defaulters. Also, we can see that people falling in the low income groups have the highest number of defaulters.

▶ The greater number of defaulters are people with 3 children.

▶ Females exhibit a lower proportion of defaulters compared to their male counterparts.

▶ Clients with family status Civil Marriage and Unmarried living in tier 3 cities have the have the highest percentage of defaulters.

▶ 'Repairs' as the loan purpose had the highest number of loan rejections.

▶ The category 'education' exhibited a nearly equal distribution of loan approvals and rejections.

▶ Notably, applications associated with 'paying other loans' and 'buying a new car' faced notably higher rejection rates compared to approvals.

Observations

▶ Applicants with loan purposes related to 'Repairs' encounter more challenges in meeting timely payments, suggesting a higher probability of payment difficulties in this category.

▶ Certain loan purposes, such as 'Buying a garage', 'Business development', 'Buying land', 'Buying a new car', and 'Education', exhibit notably higher rates of successful loan repayment compared to instances of payment difficulties. Hence, these specific purposes seem to be associated with minimal payment challenges.

▶ Widows demonstrate the most reliable payment behaviour and are the best risk overall.

▶ Civil married pose most risk

▶ The age bracket of 17-35 years poses the highest risk for the bank.

▶ Conversely, the age group of 55 years and above represents the least risky for the bank.

▶ We can see that only 14.5% of the clients are new and 80.7% are repeaters

▶ It is apparent that individuals residing in office apartments have a greater credit likelihood associated with target 0, indicating a higher chance of timely payments. Conversely, co-op apartments demonstrate a higher credit likelihood linked to target 1, suggesting an elevated risk of payment difficulties.

Suggestions for the Bank

▶ Implement stringent risk assessment protocols for clients residing in cities with a rating of 3, and design tailored loan products that match their financial capacity.

▶ Develop specialized financial products, such as flexible repayment schemes or insurance-backed loans, for individuals in the low-income bracket to reduce the risk of defaults.

▶ Introduce family-specific loan programs, including customized repayment options and financial counseling services, for clients with three children to enhance their loan repayment capabilities.

▶ Design gender-specific financial products and services that cater to the unique financial circumstances of females, ensuring equitable access to credit while mitigating the risk of default.

▶ Implement focused strategies for loan purposes such as 'Repairs', 'Paying Other Loans', and 'Buying a New Car' to streamline the approval process, offer financial counseling, and potentially introduce risk-mitigating measures.

Suggestions for the Bank

▶ Develop tailored loan products with flexible repayment options for applicants seeking loans for 'Repairs' to reduce the risk of payment difficulties in this category.

▶ Implement customized risk management strategies, including thorough credit assessments and personalized financial counseling, for clients with specific marital statuses, such as Widows and Civil Married, to minimize the risk of defaults.

▶ Tailored Financial Services for Specific Age Groups

▶ It is advisable for the bank to exercise caution when granting loans for co-op apartments, considering the potential payment challenges identified within this housing category. Instead, the focus should primarily be on housing types such as apartments with parents, house/apartments, or municipal apartments, as these categories exhibit a more promising track record of successful payments.