CREDIT EDA Case Study

By Syed Ahemed Zahran P

Introduction

This assignment plans to provide you with a thought of applying EDA in a genuine business situation. In this task, apart from applying the methods that you have learnt in the EDA module, you will likewise foster a fundamental comprehension of chance examination in banking and monetary administrations and comprehend how information is utilized to limit the gamble of losing cash while loaning to clients.

Business Understanding

The loan providing organizations find it hard to give credits to individuals because of their lacking or non-existent record of loan repayment. Thus, a few clients use it for their potential benefit by turning into a defaulter. Assume you work for a buyer finance organization which has practical experience in loaning different sorts of credits to urban clients. You need to utilize EDA to dissect the examples present in the information. This will guarantee that the clients capable of repaying the loan are not rejected.

Business Understanding

- When the company receives a loan application, the company has to decide for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:
- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

Business Understanding

- The data given below contains the information about the loan application at the time of applying for the loan. It contains two types of scenarios:
- **The client with payment difficulties:** he/she had late payment more than X days on at least one of the first Y instalments of the loan in our sample,
- **All other cases:** All other cases when the payment is paid on time.

Business Understanding

- When a client applies for a loan, there are four types of decisions that could be taken by the client/company):
- **Approved:** The Company has approved loan Application
- **Cancelled:** The client cancelled the application sometime during approval. Either the client changed her/his mind about the loan or in some cases due to a higher risk of the client, he received worse pricing which he did not want.
- **Refused:** The company had rejected the loan (because the client does not meet their requirements etc.).
- **Unused offer:** Loan has been cancelled by the client but at different stages of the process.
- In this case study, you will use EDA to understand how consumer attributes and loan attributes influence the tendency to default.

Business Objectives

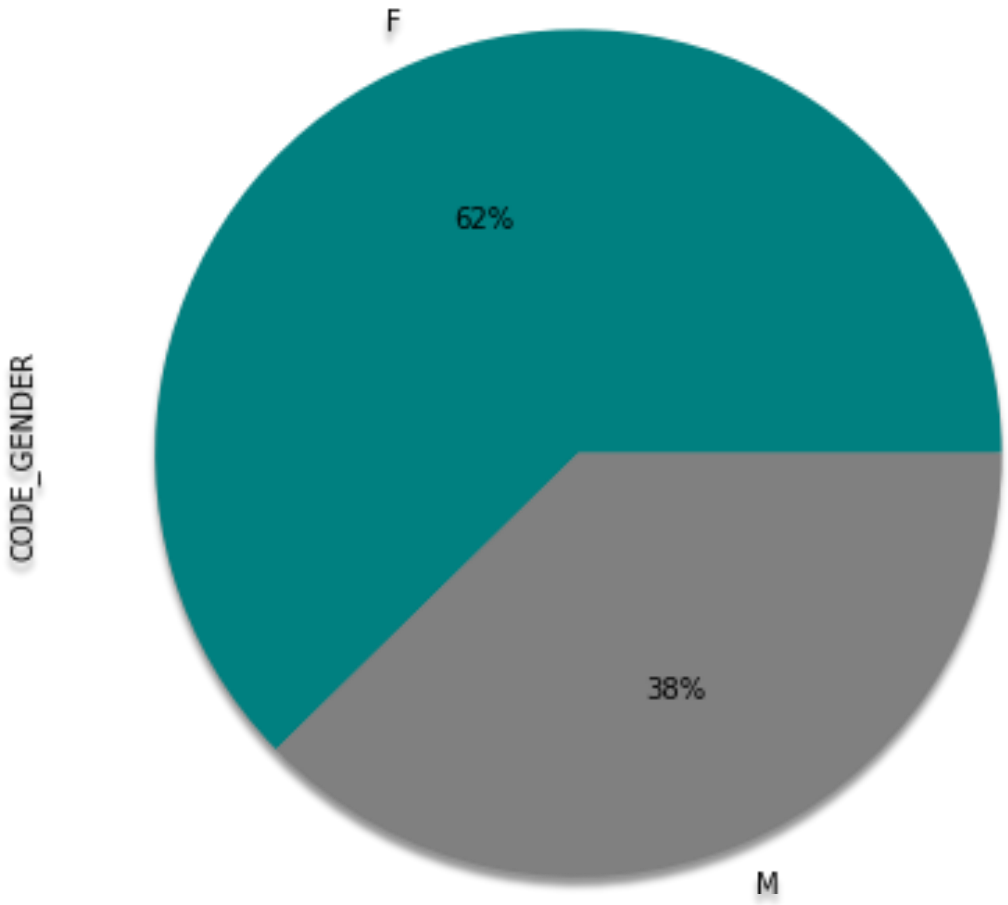
- This case study aims to identify patterns which indicate if a client has difficulty paying their instalments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. This will ensure that the consumers capable of repaying the loan are not rejected. Identification of such applicants using EDA is the aim of this case study.
- In other words, the company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.
- To develop your understanding of the domain, you are advised to independently research a little about risk analytics - understanding the types of variables and their significance should be enough

Data Understanding

- There are 3 datasets for this case study:
  1. '*application_data.csv*' contains all the information of the client at the time of application.
     The data is about whether a **client has payment difficulties.**
  2. '*previous_application.csv*' contains information about the client's previous loan data. It contains the data on whether the previous application had been **Approved, Cancelled, Refused or Unused offer.**
  3. '*columns_description.csv*' is data dictionary which describes the meaning of the variables.

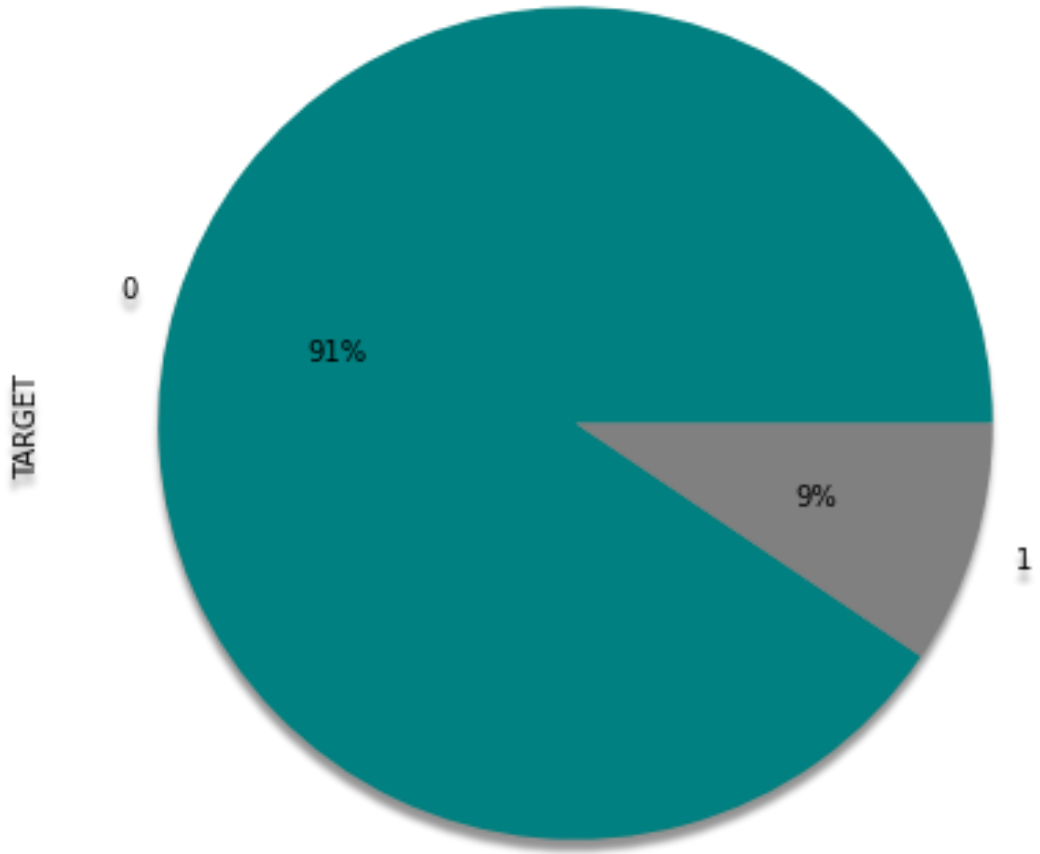
ANALYSIS OF GENDER DISTRIBUTION

- 62% of the total clients are female and 38 % are male



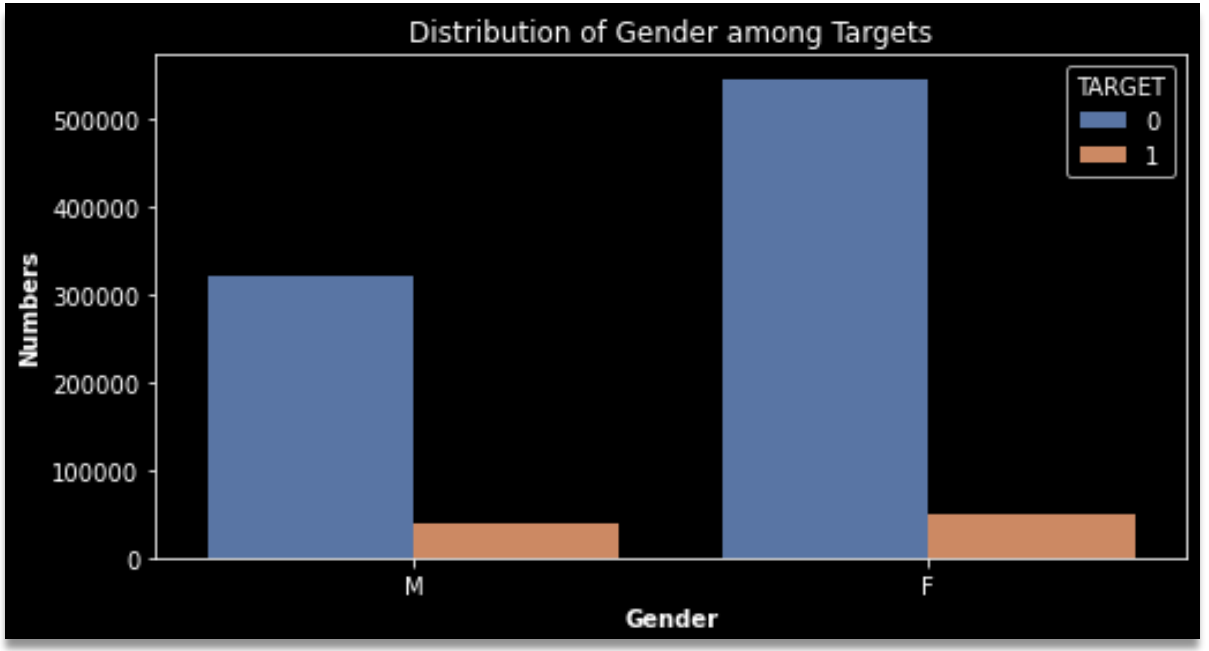
ANALYSIS OF TARGETS

- 91 % of the total clients are repayers. They are the clients who does not face payment difficulties. They are categorized as target 0.
- Remaining 9 % of the clients are defaulters. They are the clients who are facing difficulties in loan repayments. They are categorized as target 1.



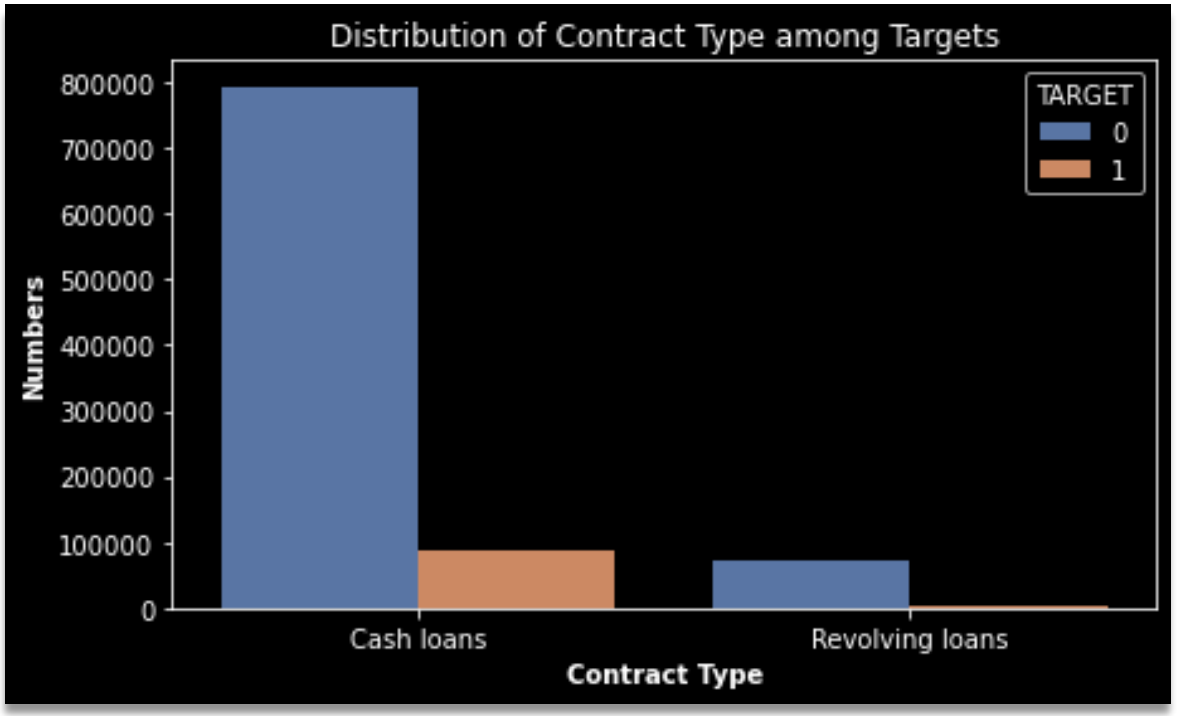
DISTRIBUTION OF GENDERS AMONG THE TARGETS

- Major part of both male and female clients does not have difficulty in payments
- Clients with payment difficulties are equally distributed among both genders



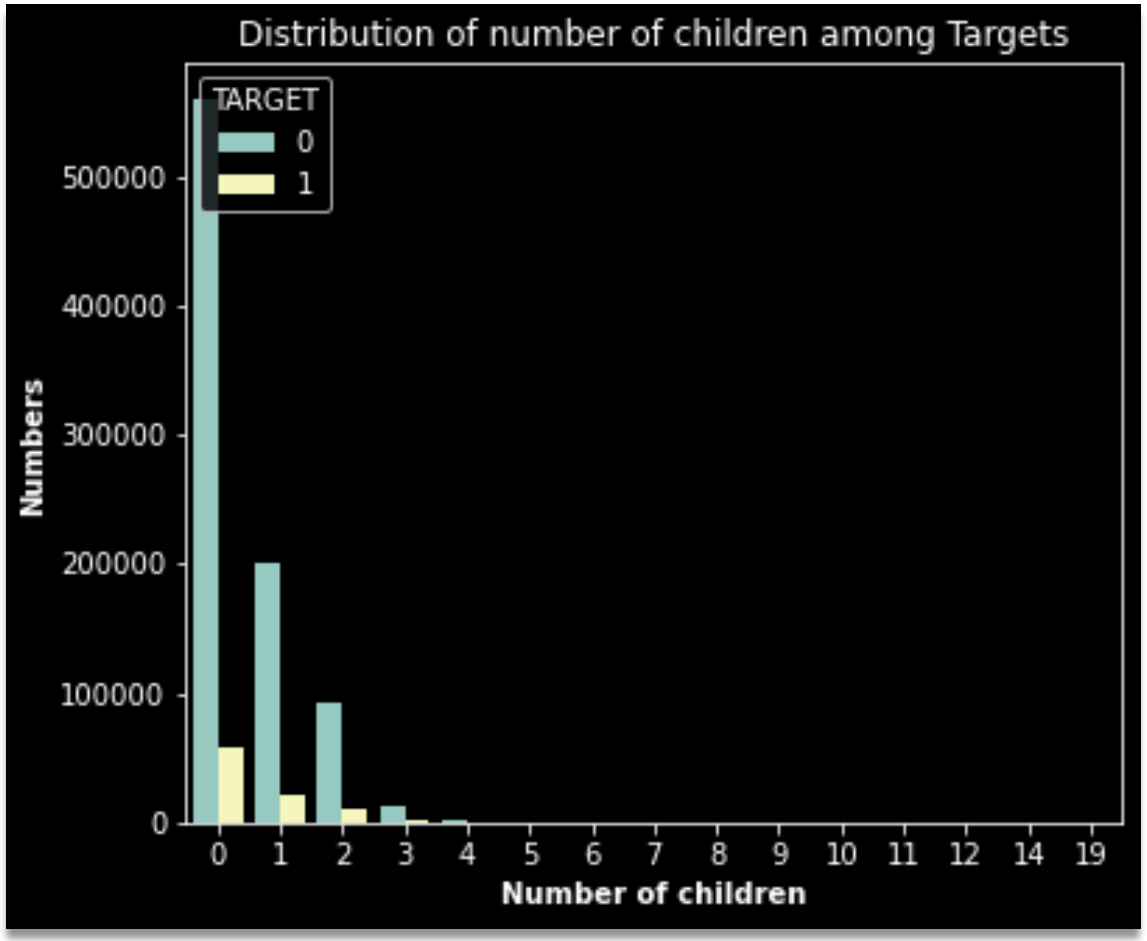
DISTRIBUTION OF CONTRACT TYPES AMONG TARGETS

- Both clients with payment difficulties and without payment difficulties are preferring cash loans than revolving loans
- Clients with payment difficulties are not at all preferring revolving loans



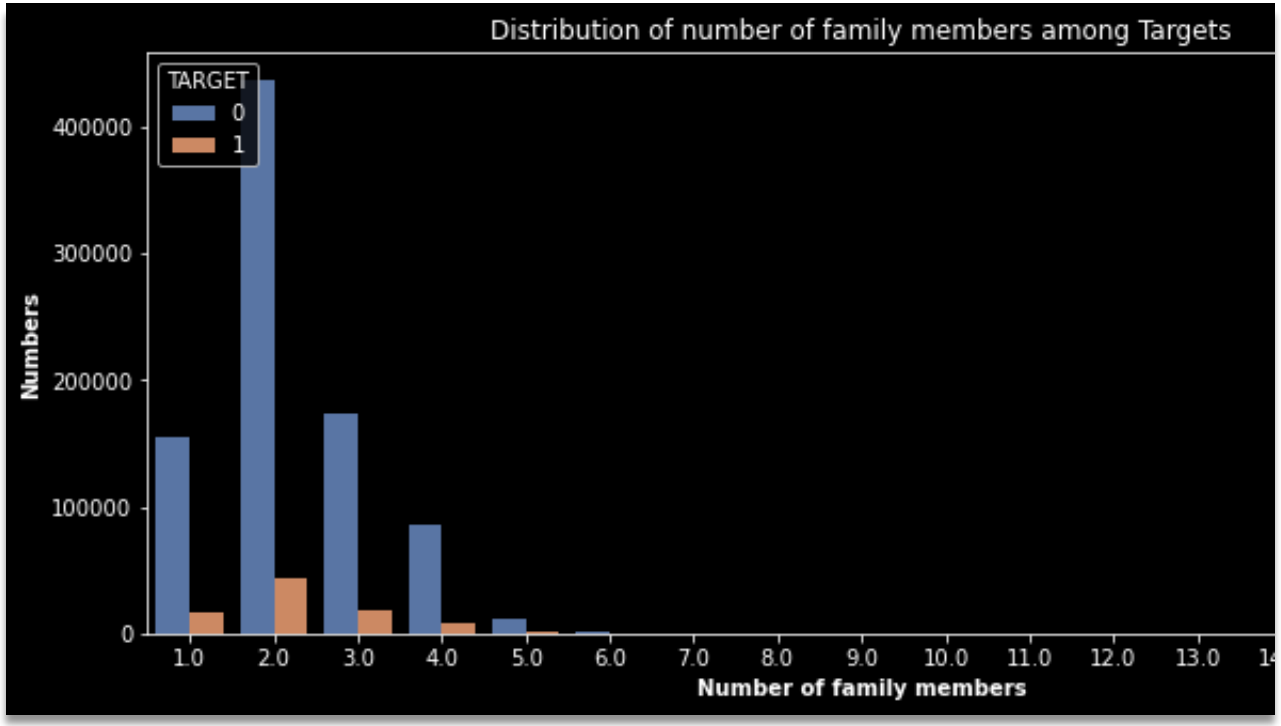
DISTRIBUTION OF NUMBER OF CHILDREN AMONG TARGETS

- Most of the clients with and without payment difficulties does not have children

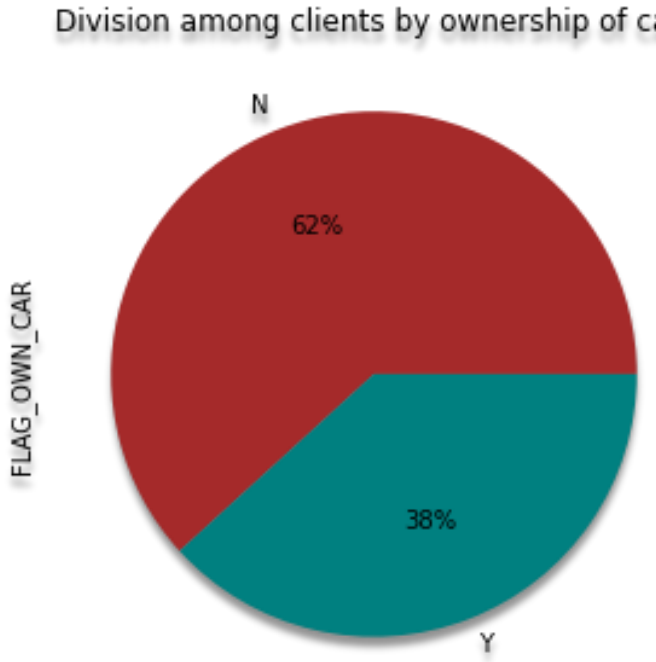


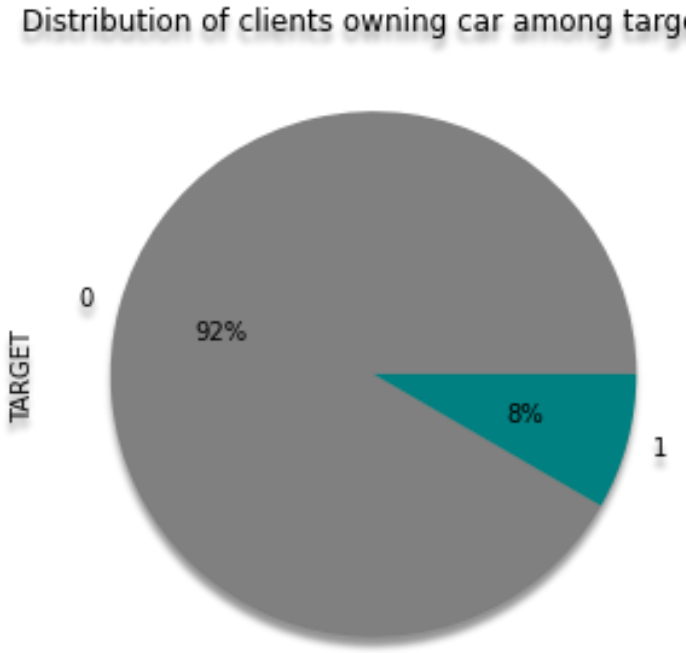
DISTRIBUTION OF NUMBER OF FAMILY MEMBERS AMONG TARGETS

- In both targets, clients with 2 family members are more in numbers
- Clients with more than 4 family members does not have difficulty in payment



DISTRIBUTION OF CLIENTS OWNING A CAR AMONG TARGETS

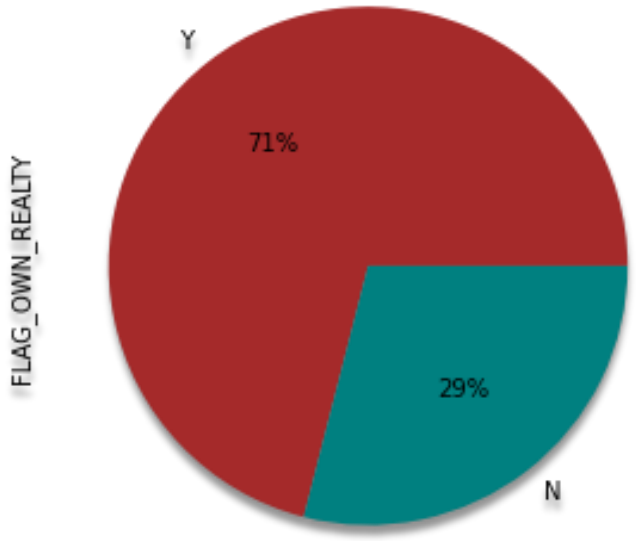




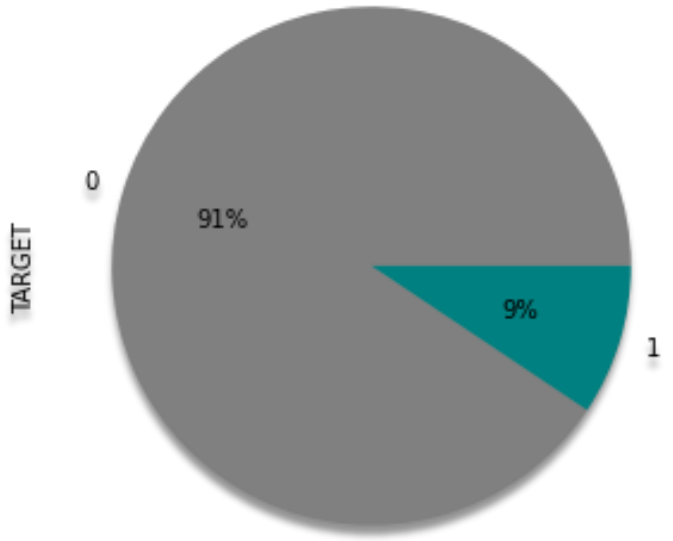
- 62% of the clients does not own a car
- 38% of the clients own a car
- 92% of the clients owning a car does not have difficulty in payment and 8% have difficulty in payment

DISTRIBUTION OF CLIENTS OWNING A FLAT/HOUSE AMONG TARGETS

Division among clients by ownership of flat/house

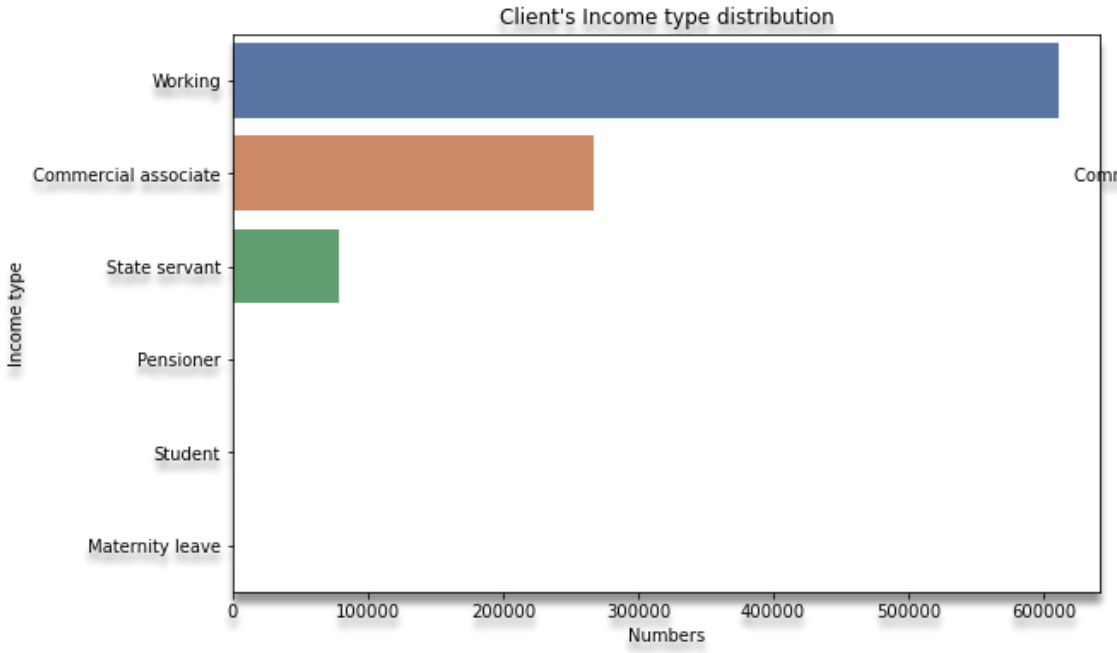


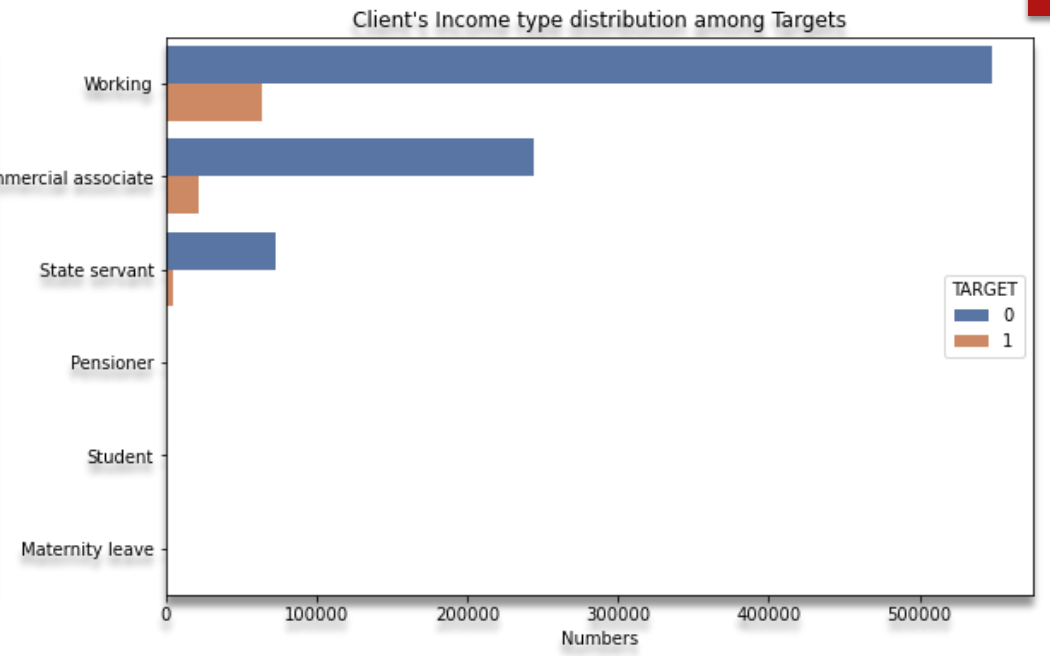
Distribution of clients owning flat/house among targets



- 71% of the clients own a flat/house and 29% do not.
- 91% of the clients owning a flat/house does not have difficulty in payment and 9% have difficulty in payment

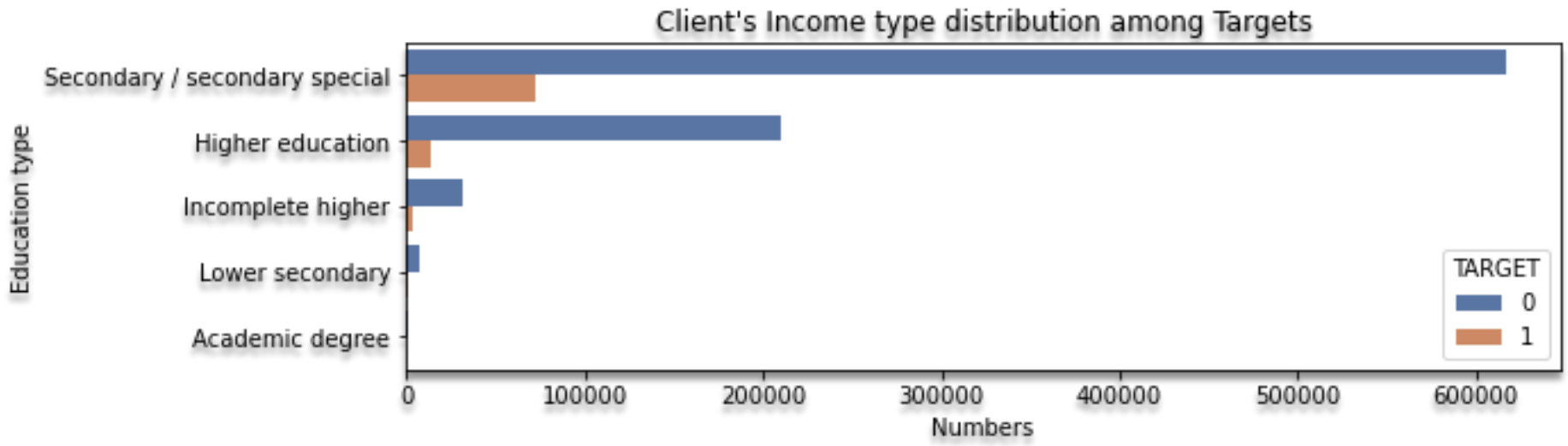
DISTRIBUTION OF CLIENTS INCOME TYPE AMONG TARGETS





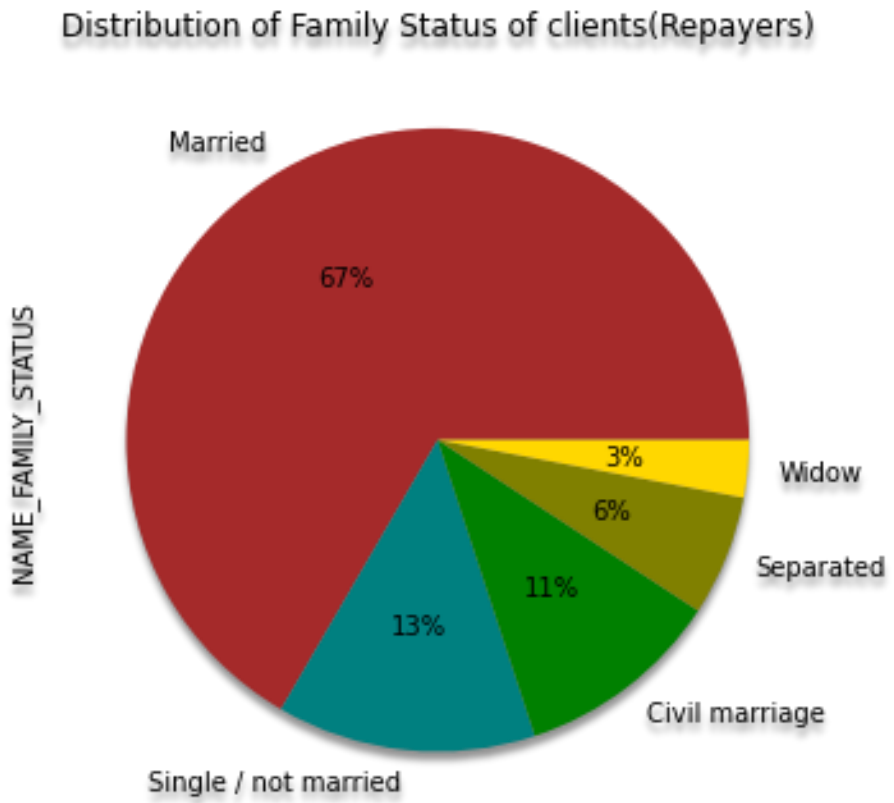
- Majority of the clients in both target types are working

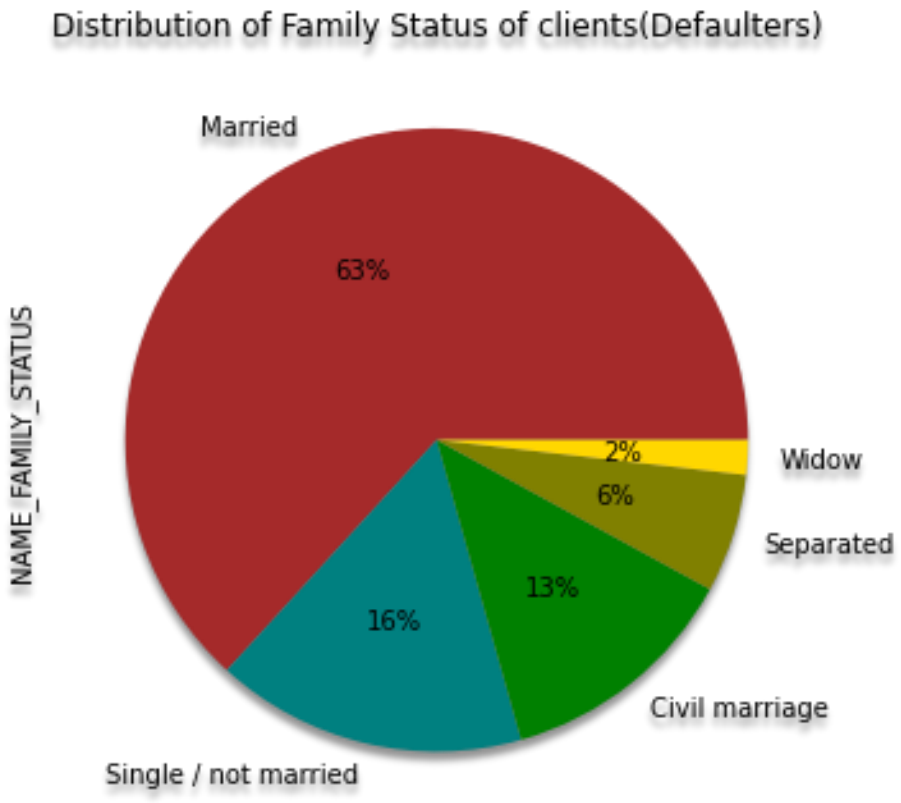
DISTRIBUTION OF EDUCATION LEVEL OF CLIENTS AMONG TARGETS



- Majority of the client's education level fall under secondary/secondary special category
- Both defaulters and re-payers are more in secondary/secondary special category
- There are no clients with academic degree
- All clients with lower secondary education are re-payers

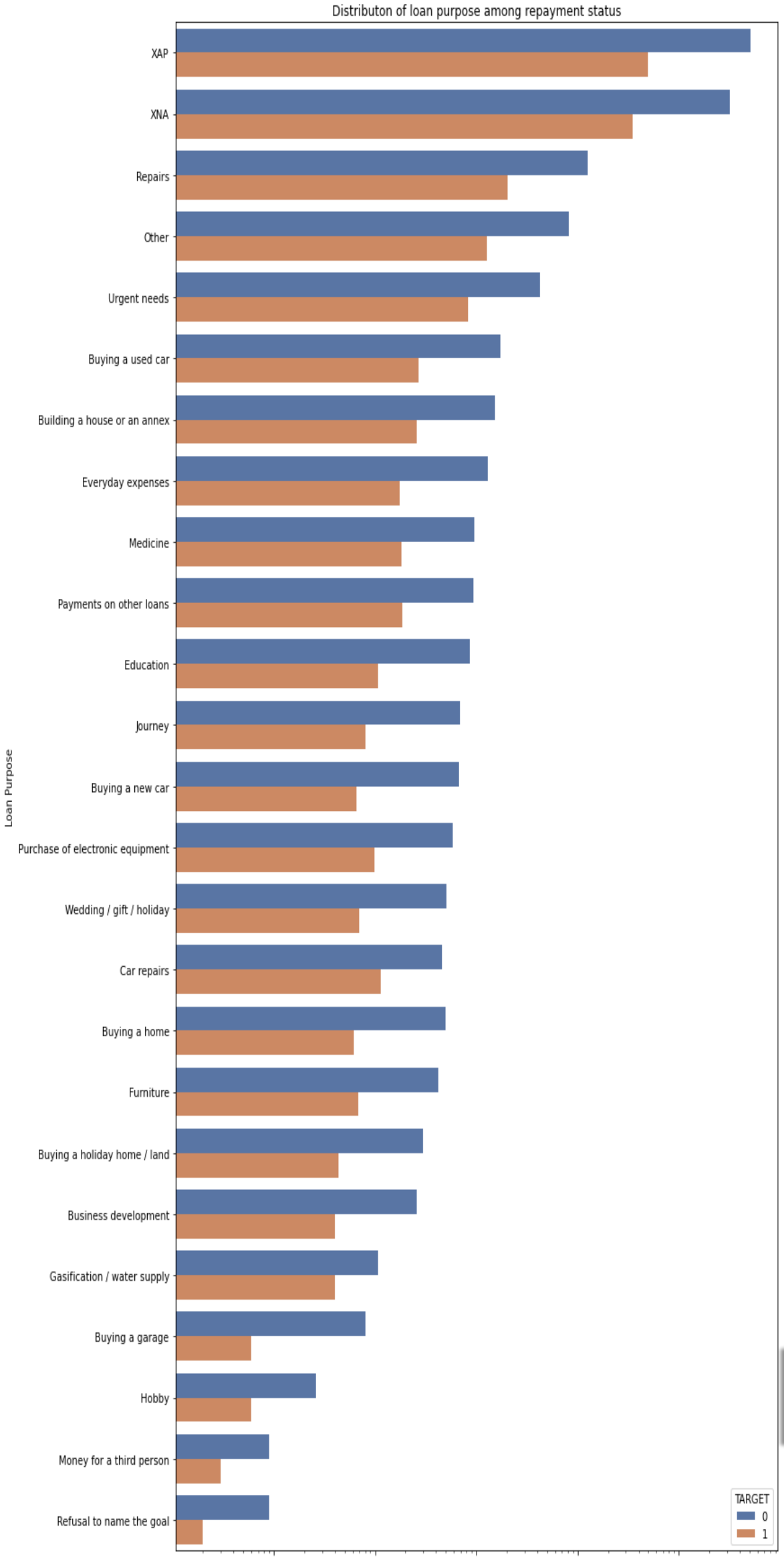
DISTRIBUTION OF FAMILY STATUS OF CLIENTS AMONG TARGETS





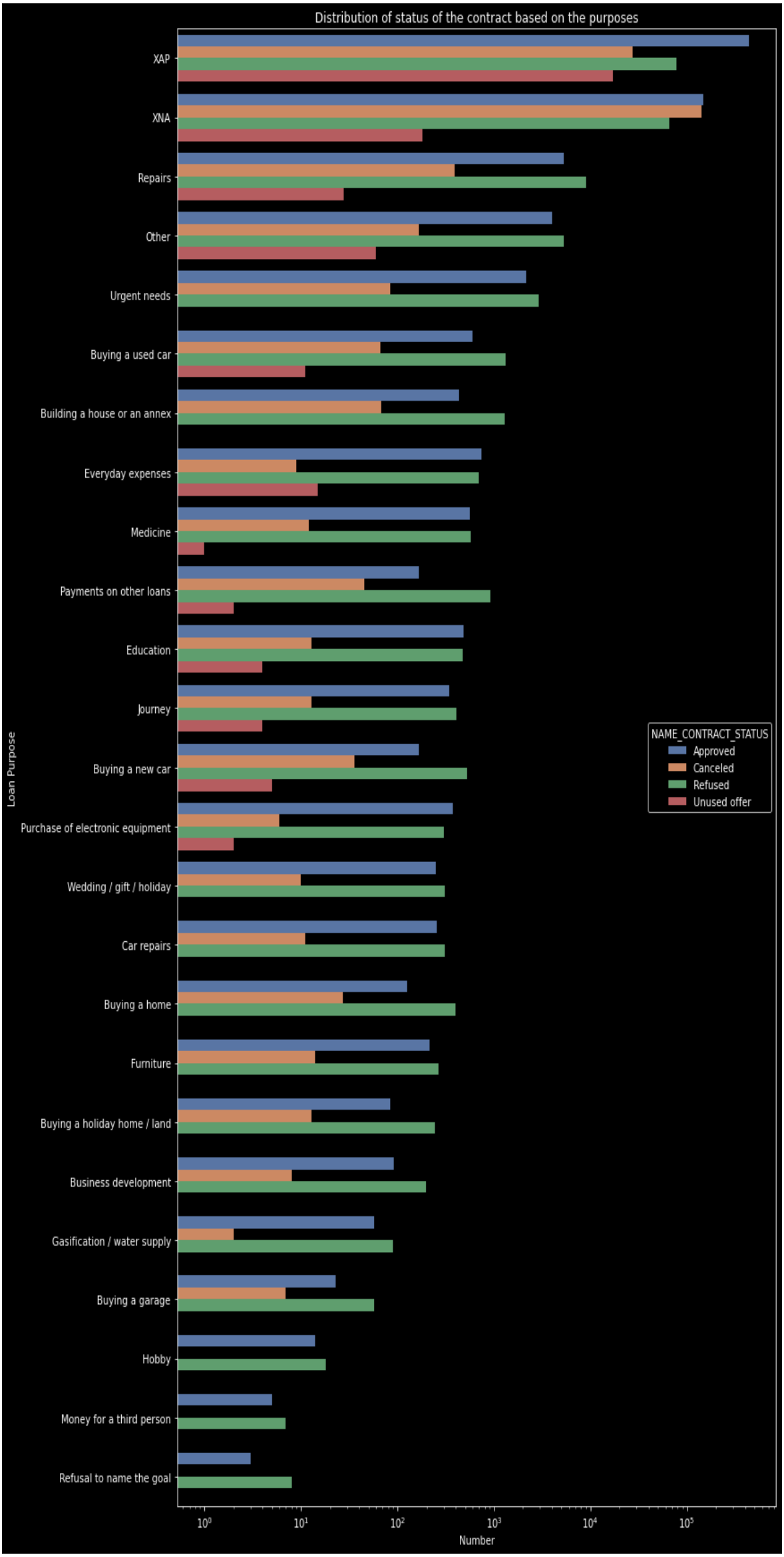
- Majority of the clients are married in both defaulters and re-payers
- Every category of client's family status is almost equally distributed between defaulters and re-payers

DISTRIBUTION OF PURPOSE OF LOAN AMONG TARGETS


Distributon of loan purpose among repayment status
XAP
XNA
Repairs
Other
Urgent needs
Buying a used car
Building a house or an annex
Everyday expenses
Medicine
Payments on other loans
Education
Journey
Buying a new car
Purchase of electronic equipment
Wedding / gift / holiday
Car repairs
Buying a home
Furniture
Buying a holiday home / land
Business development
Gasification / water supply
Buying a garage
Hobby
Money for a third person
Refusal to name the goal
Loan Purpose
TARGET
0
1

- The count of re-payers are more than defaulters in the list of clients who refuse to name the goal
- Majority of both re-payers and defaulters are applying for loan for repair purpose

DISTRIBUTION OF STATUS OF CONTRACTS AMONG PURPOSE OF LOAN


Distribution of status of the contract based on the purposes
XAP
XNA
Repairs
Other
Urgent needs
Buying a used car
Building a house or an annex
Everyday expenses
Medicine
Payments on other loans
Education
Journey
Buying a new car
Purchase of electronic equipment
Wedding / gift / holiday
Car repairs
Buying a home
Furniture
Buying a holiday home / land
Business development
Gasification / water supply
Buying a garage
Hobby
Money for a third person
Refusal to name the goal
Loan Purpose
NAME_CONTRACT_STATUS
Approved
Canceled
Refused
Unused offer
10^0
10^1
10^2
10^3
10^4
10^5
Number

- Most number of approvals and most number of rejections of loan are for repair purpose
- Most of the loans applied for payment of other loans are rejected
- There are equal number of approvals and rejections for education purpose

CONCLUSIONS

- Female clients should be targeted more. There more re-payers in female than male clients.
- Clients with no children should be targeted rather than those with children.
- There are no defaulters when contract type is revolving loan.
- Clients who owns a car are mostly re-payers.
- Clients who owns a flat/house are mostly re-payers.
- Clients who are having only 2 members in the family are has the maximum number of re-payers compared to those who are having more members.
- Clients who are working has highest number of re-payers and also defaulters. So there is risk in providing loans for them.
- Family status of the clients should not be considered for providing loans, as every categories has equal distribution of re-payers and defaulters.