

CREDIT EDA CASE STUDY

Presented By

Vishakha Kude

Viraj Malgi

Lending Risk Management and Decision Support System

Purpose:

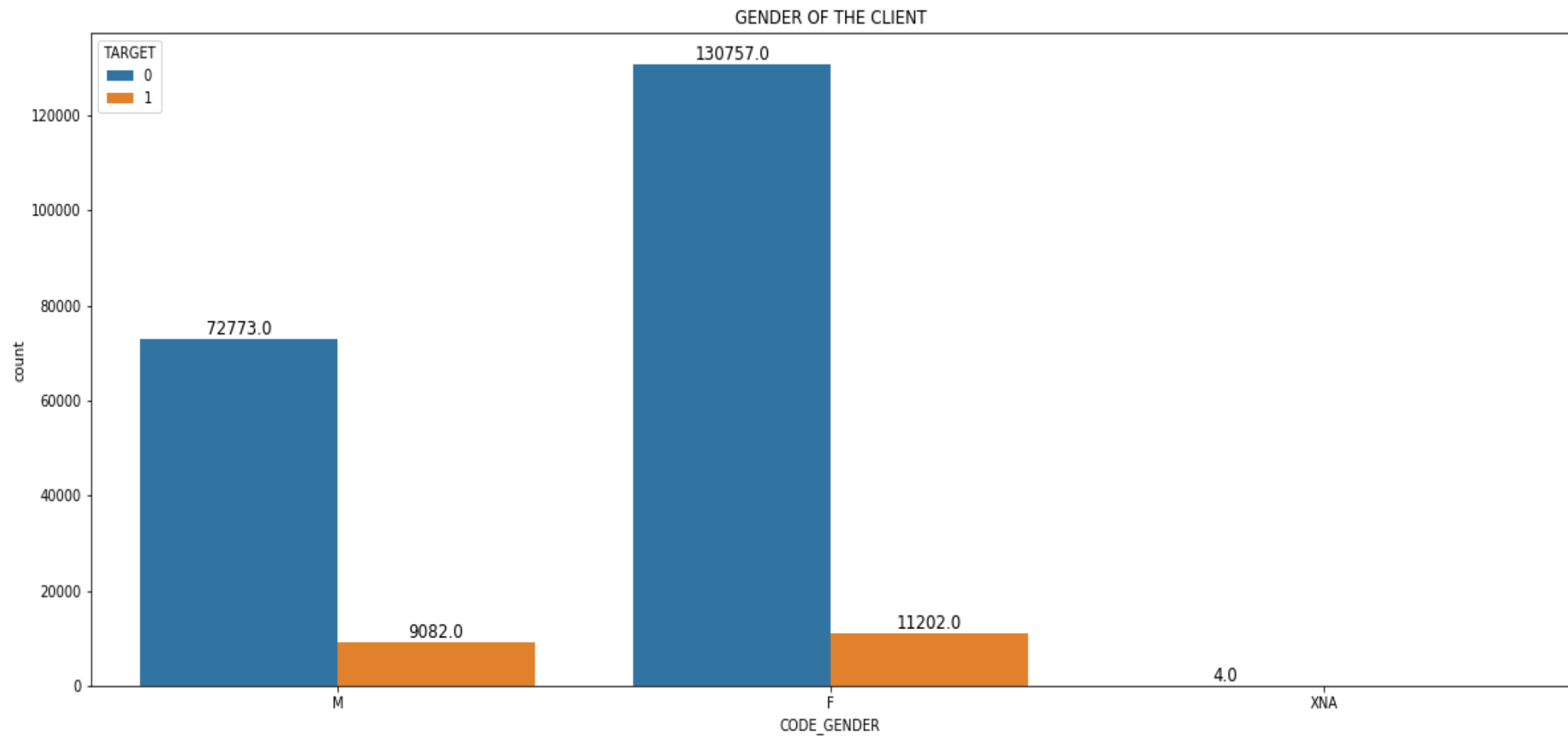
To minimize risk of lending.

Risks Involved:

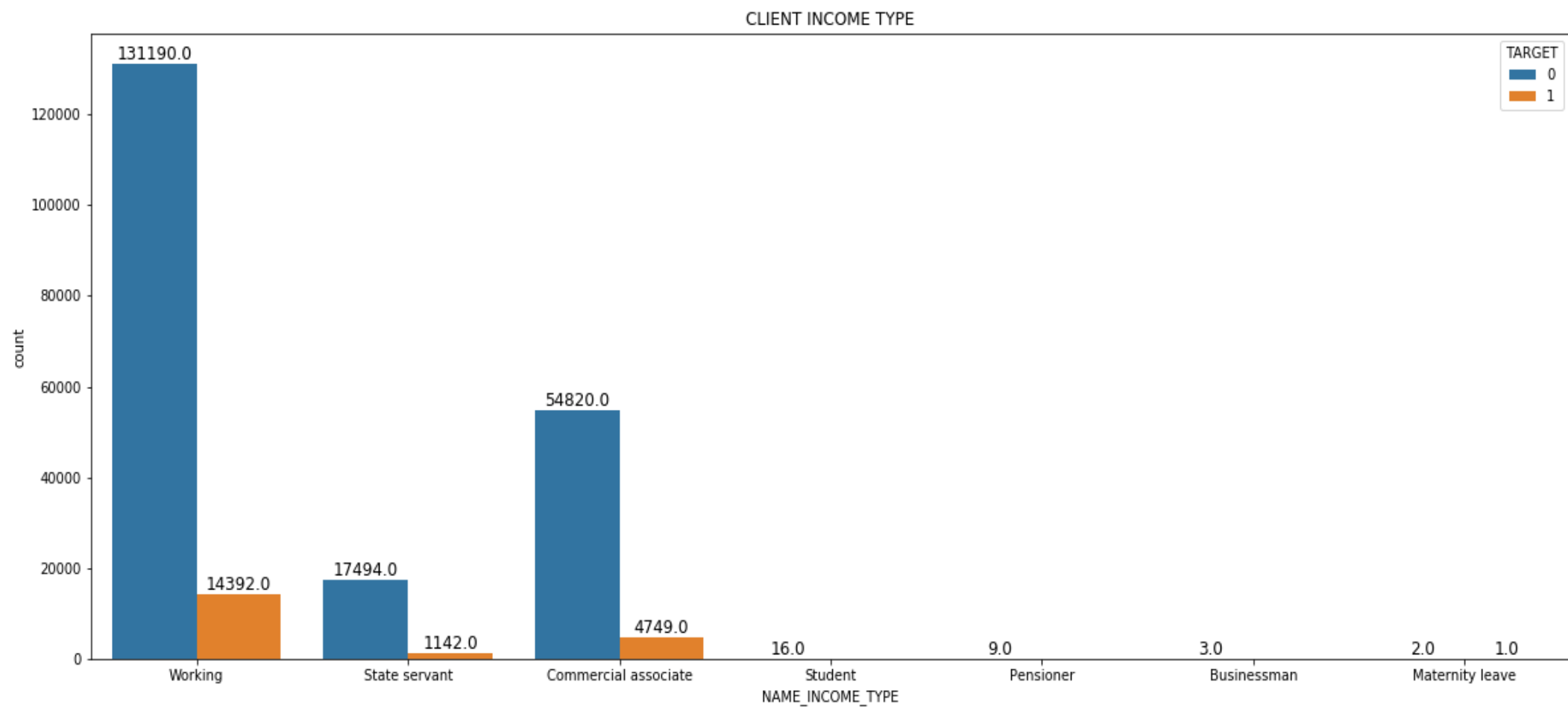
- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

Business Objectives:

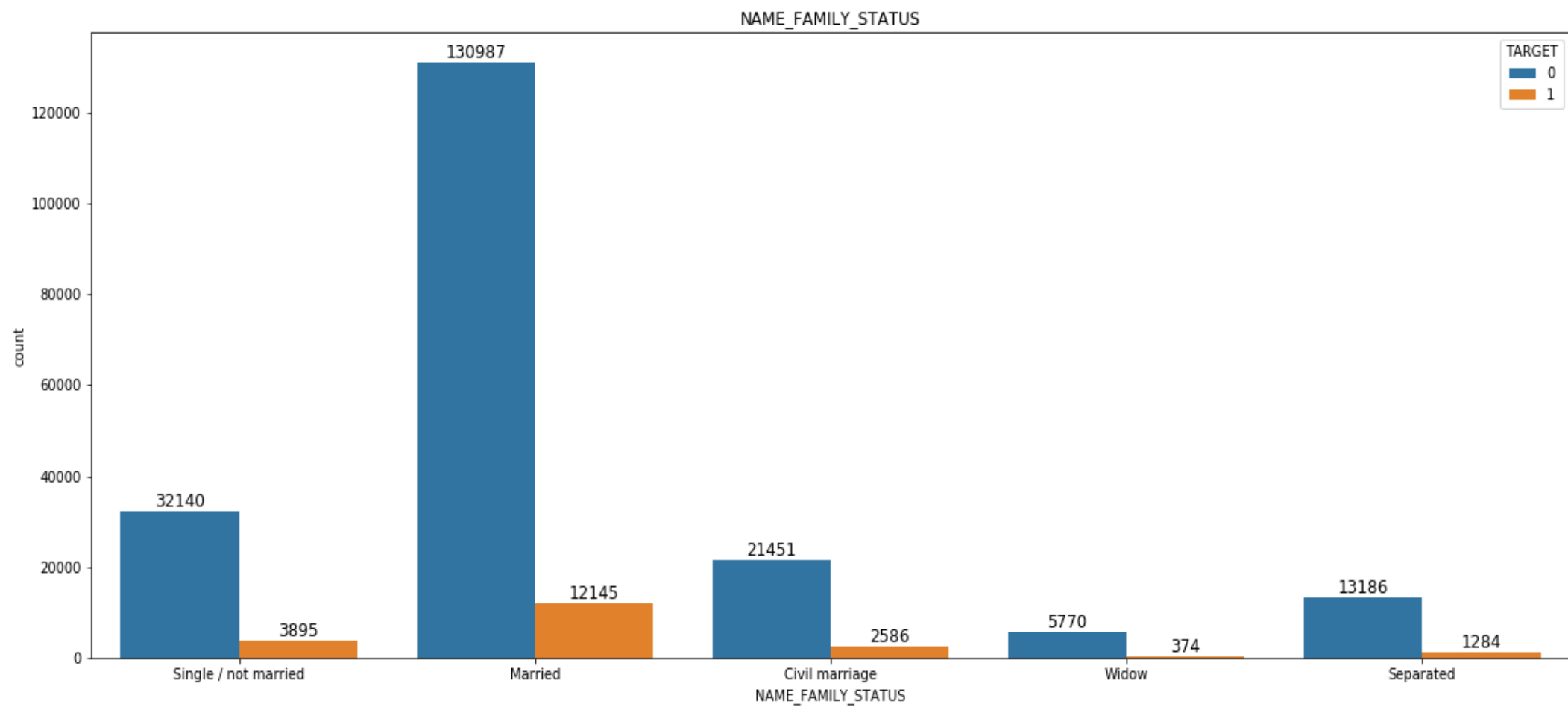
- Identify patterns which indicate if a client has difficulty paying their instalments.
- Understand the driving factors (or driver variables) behind loan default.
- Utilisation of this knowledge for one's portfolio and risk assessment.



Female application are more compared to Male because of low interest rates for females

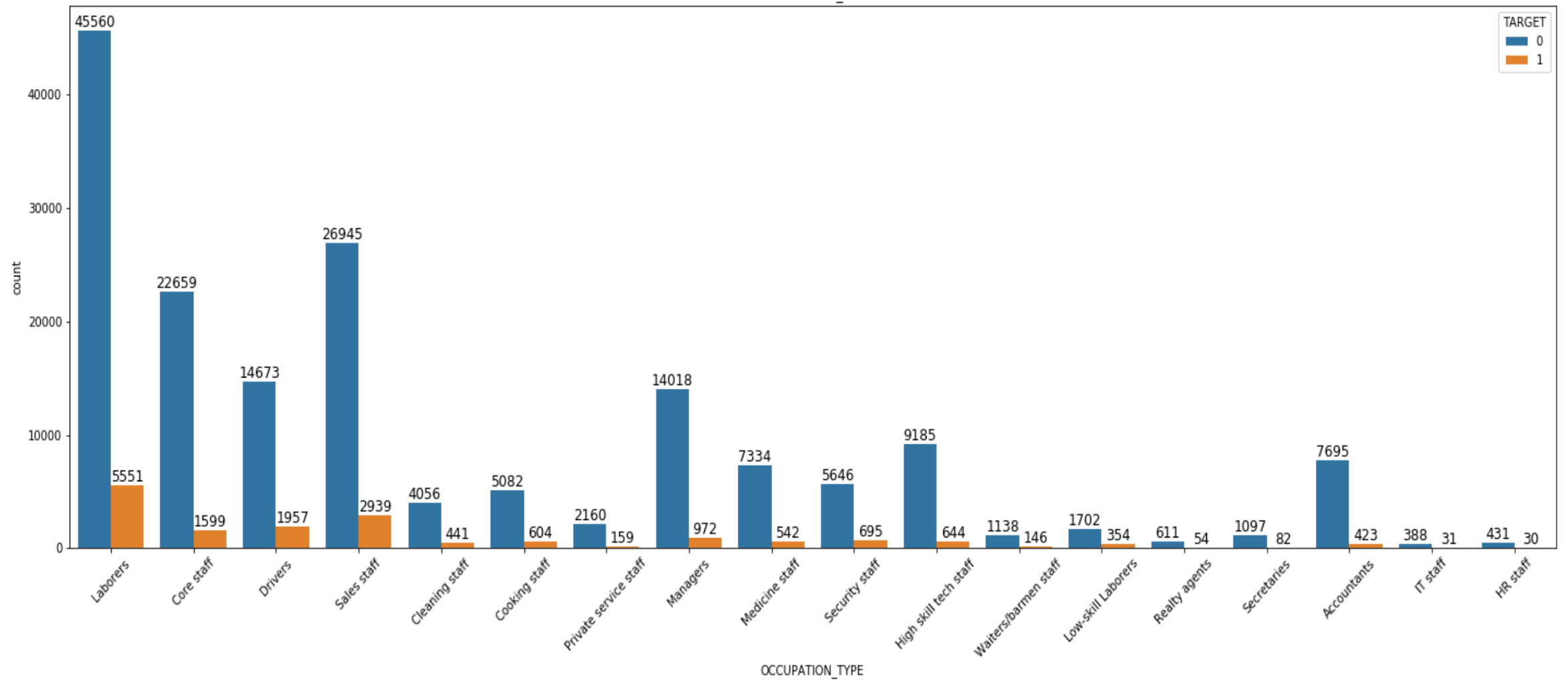


Most of the working people apply for the loan application

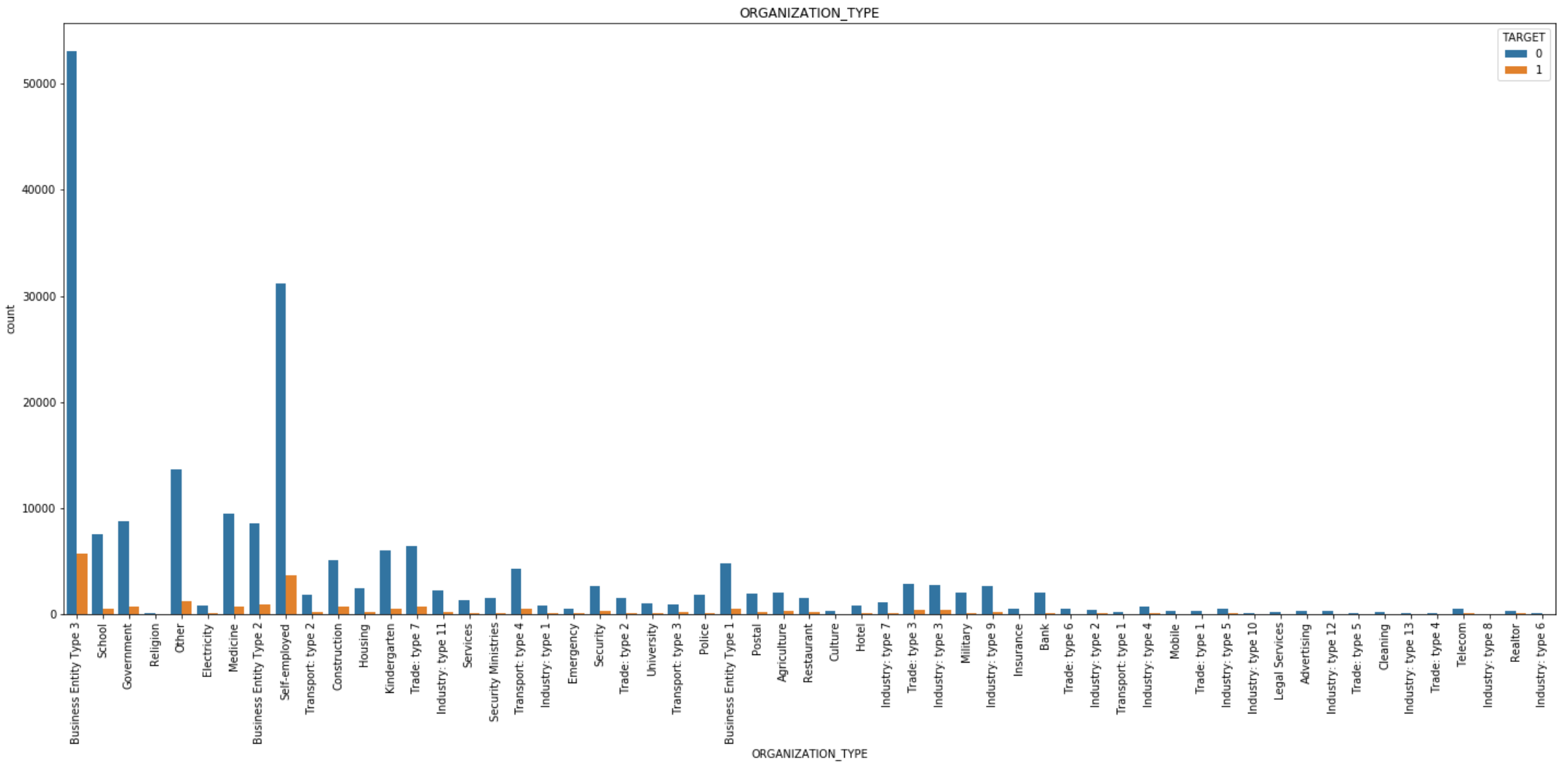


Married clients tend to have less chance of becoming defaulters

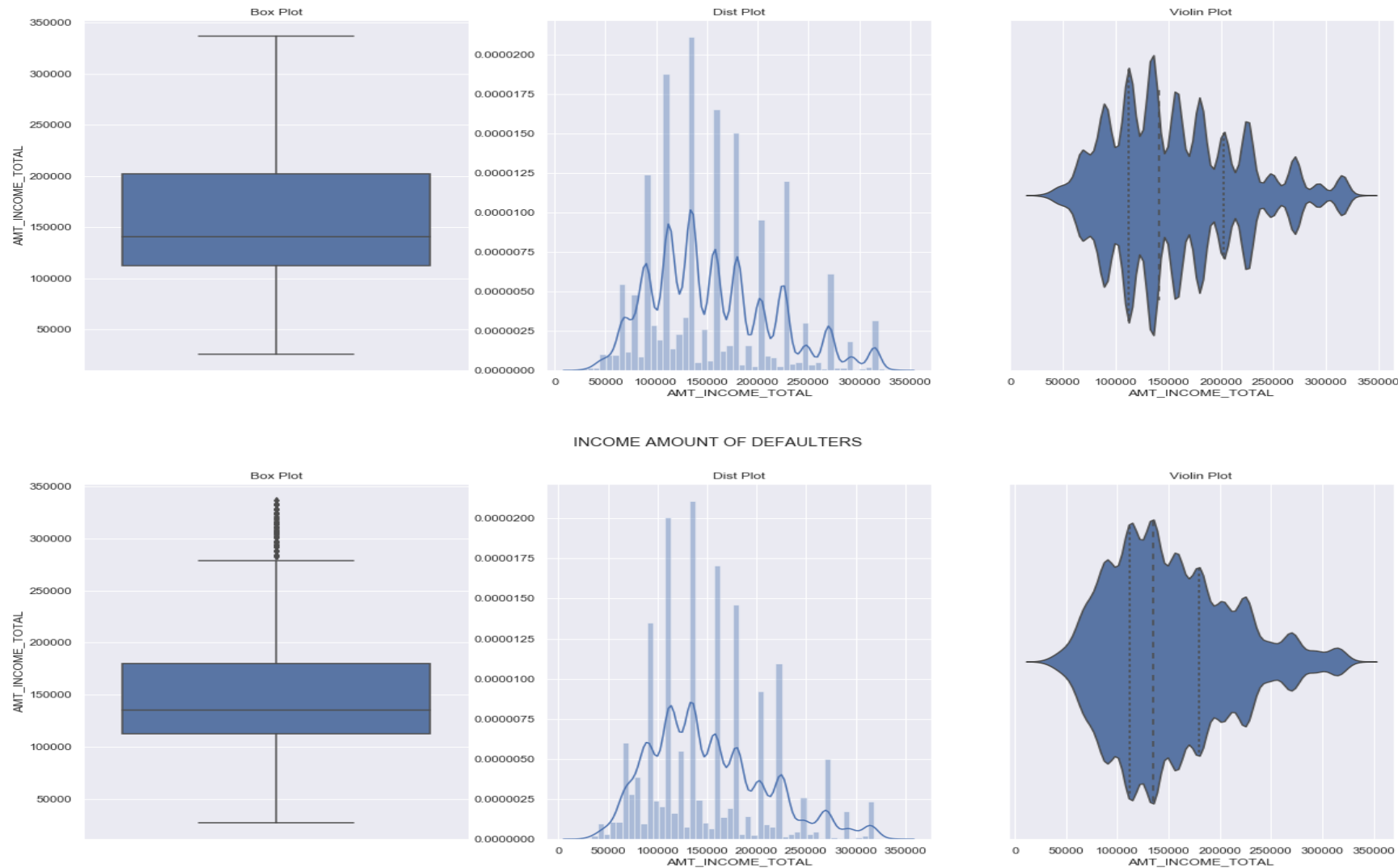
OCCUPATION_TYPE



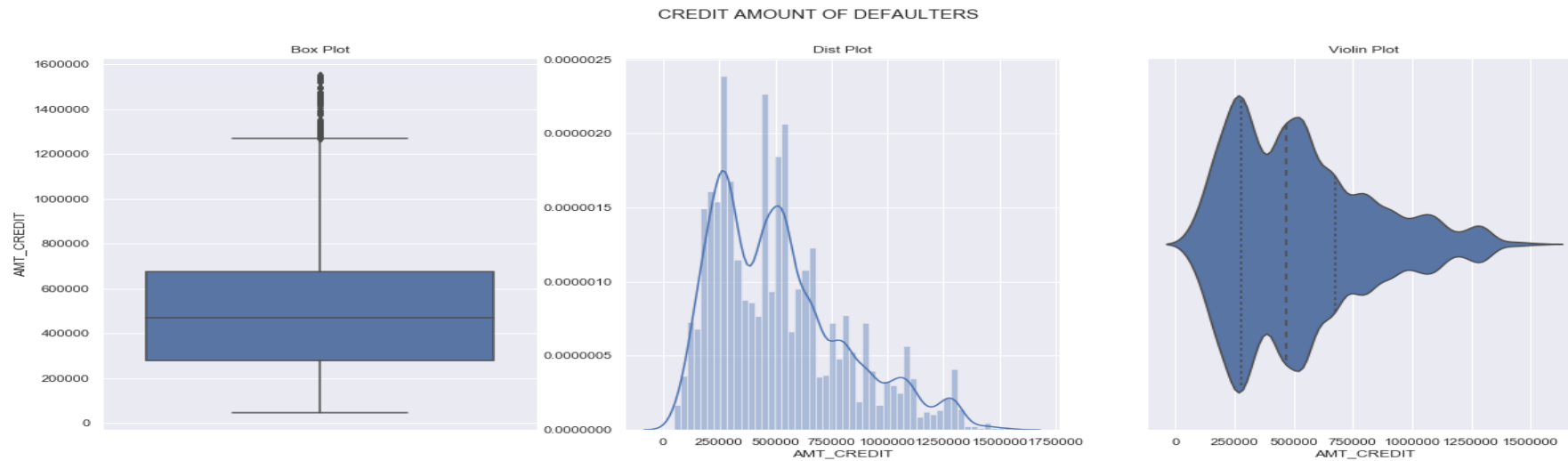
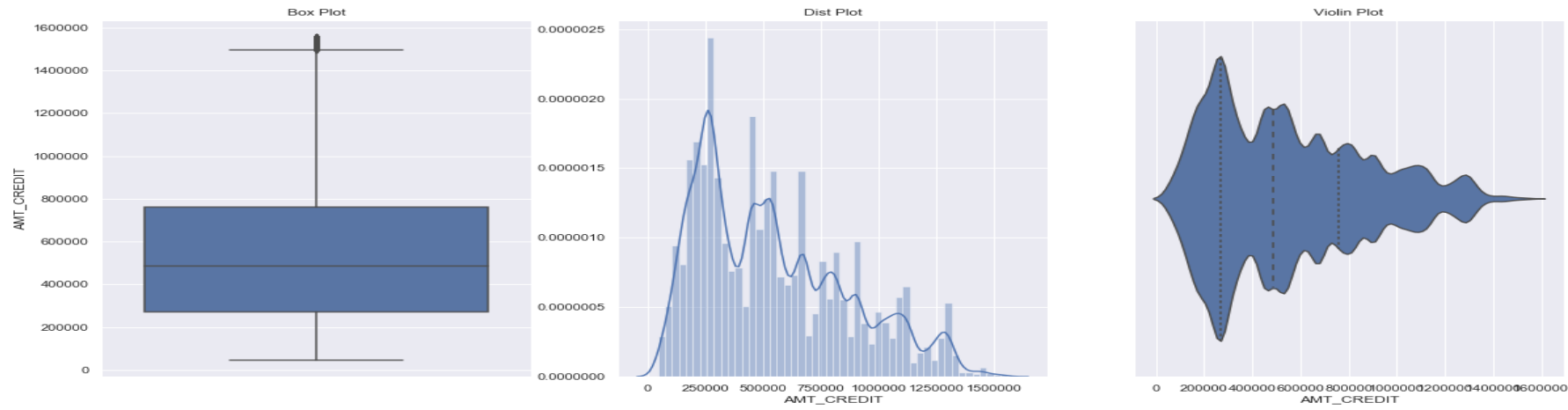
Labourers, core staff, drivers and sales staff require loan mostly and these occupation people have less chance of becoming defaulters



Most of the business people and self-employed are more likely to pay every instalment in scheduled period

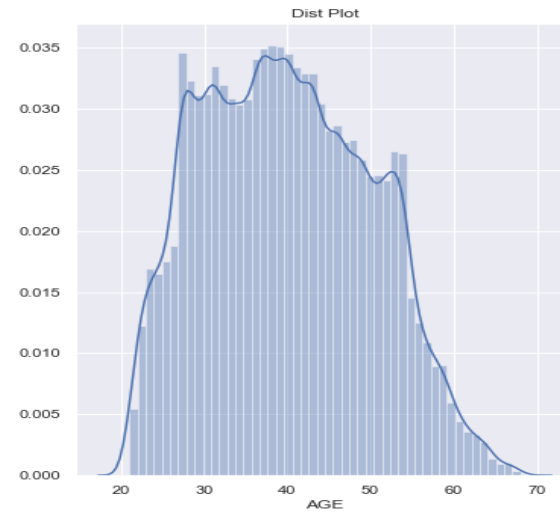
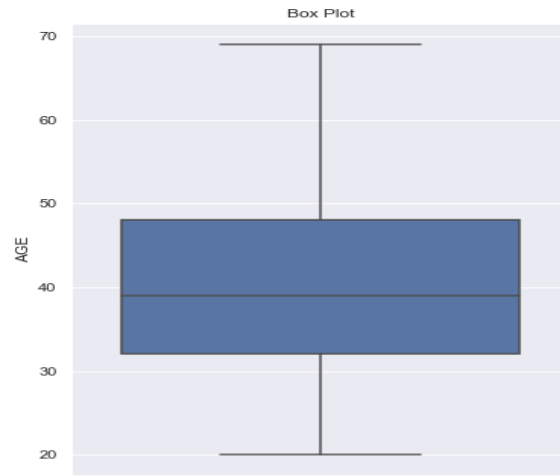


- Clients income is spread equally in both the target cases but from the box plots, one difference is that most of the non-defaulters clients have income more than the median income.
- This tells us that clients with income between 150000-200000 may be beneficial for getting non-defaulters clients
- Clients with income between 100000-150000 are most likely to become defaulters

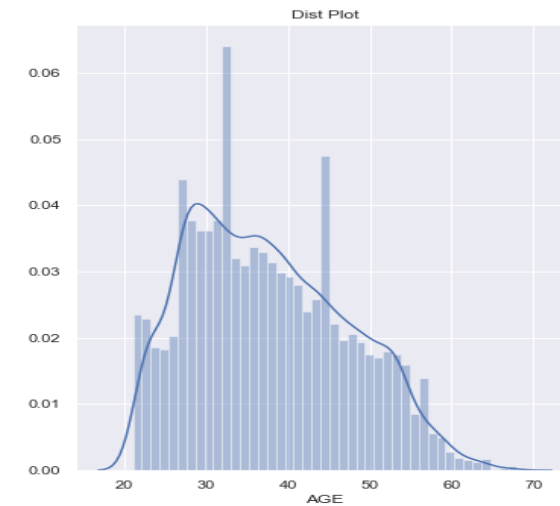
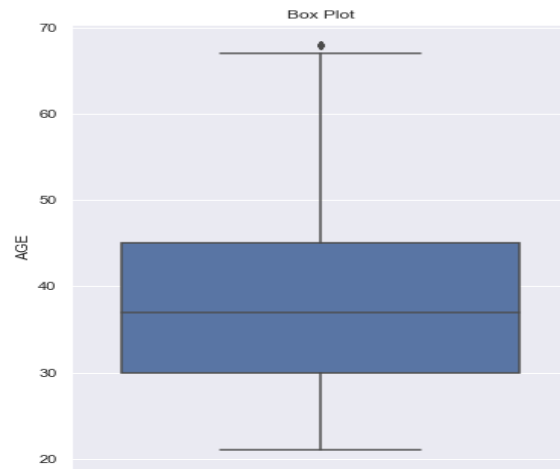


- Most clients gets 250000 loan amount approved in both the cases
- Clients with loan credit between 250000-500000 have more chances to become defaulters

CLIENT'S AGE OF NON_DEFAULTERS

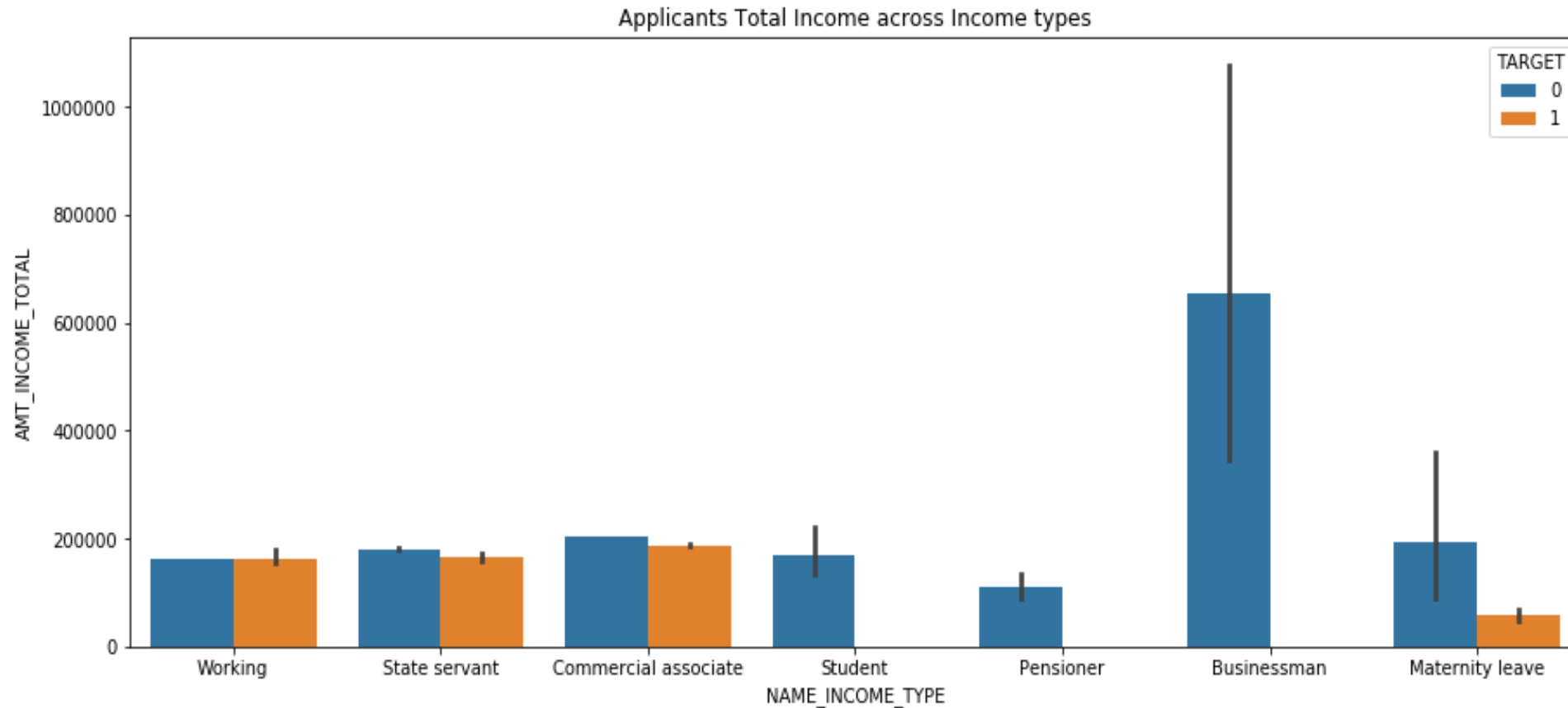


CLIENT'S AGE OF DEFAULTERS



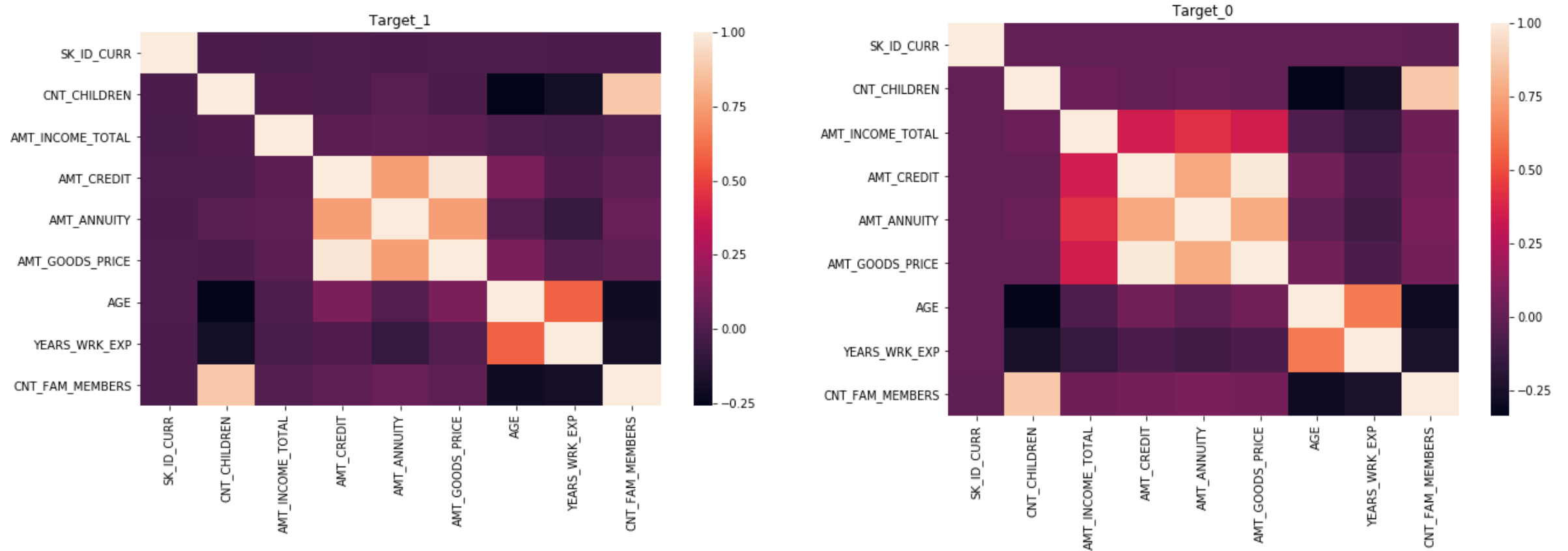
- Many clients with age 32 and 45 years are defaulters
- Client's age show a normal trend such that clients with every age category requires loan
- But clients with age between 30-40 can pay the loan without any difficulties

Income for various category and its distribution across defaulters and non defaulters group :



- Students/Pensioners/Businessman are absolute non defaulters.
- Females who are on maternity leaves are less likely to default.
- For rest of the income type categories, bank should verify as chances of default solely depends on situation as most of the defaulters are from working, state servant or commercial associates.
- Income Type seems influential factor in deciding if client default or not.

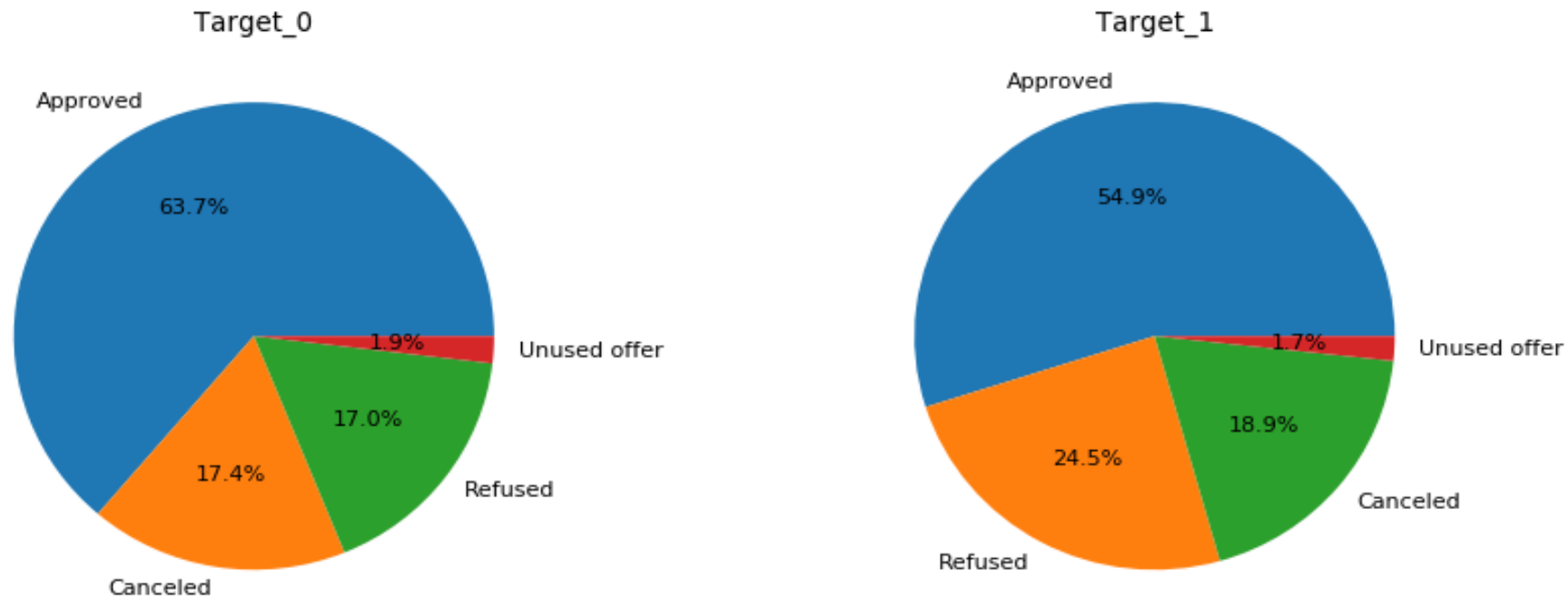
Correlation between numerical fields :



1. We can clearly see correlation difference for AMT_INCOME_TOTAL against columns AMT_CREDIT, AMT_ANNUITY, AMT_GOODS_PRICE
2. For Target 1, clients total income have low correlation with credit amount, goods price and annuity. However, for target 0 its high.
3. This clearly indicates, clients total income value is influencing factor to decide client will default or not.

Overall % contribution of each previous application's status for rows associated with Target category 1 or 0 from current application dataset

Percentage Distribution of Contract Status of Previous applications



From above graph : % of refused previous application are more for Target = 1 (24.0%) compared to that of Target 0 (16.7%) It looks like applicants with status of previous applications = Refused are more likely to have difficulties in payments.

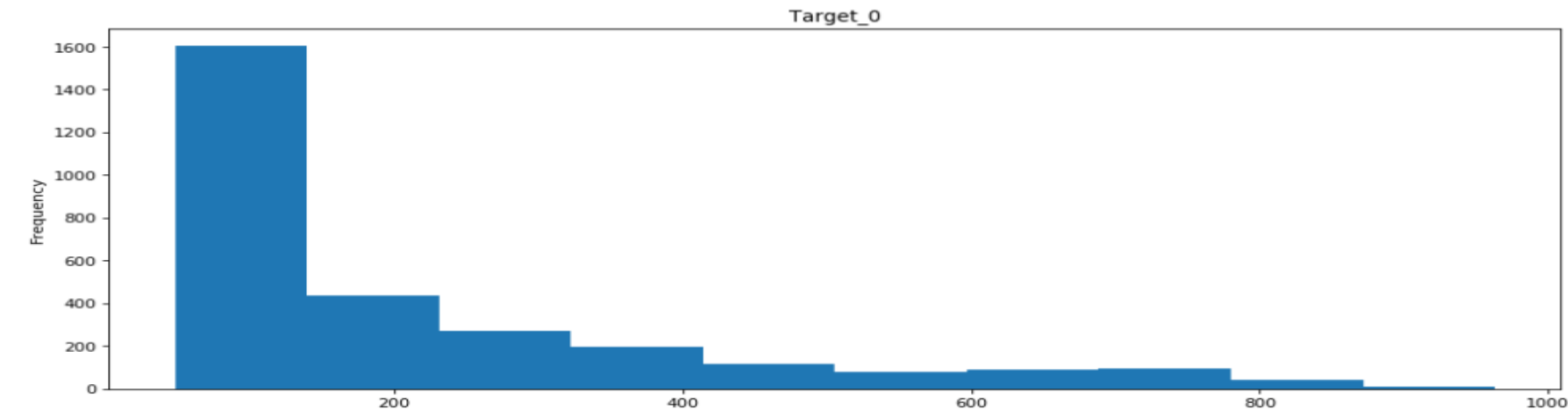
Distribution of overall Code Reject Reason of previous application where previous applications with contract status as Refused :



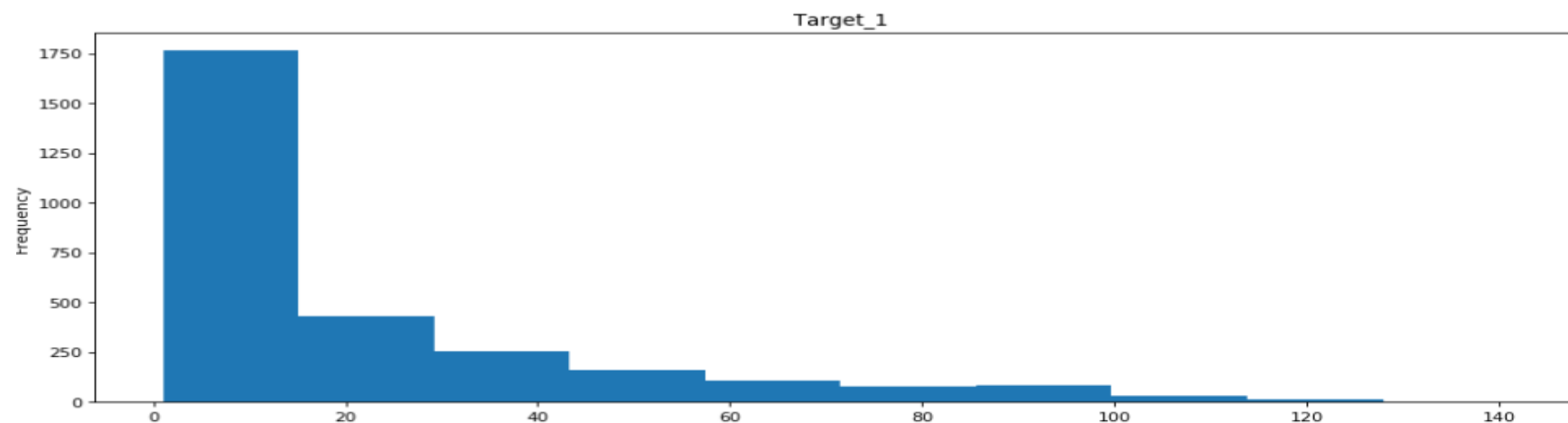
It looks like previous application with rejection reason with SCOFR are more likely to have payment difficulties. Rest of the reasons distribution is not that significant across Targets.

Distribution of number of days taken for decision of previous applications:

Distribution of number of days taken for decision of previous application

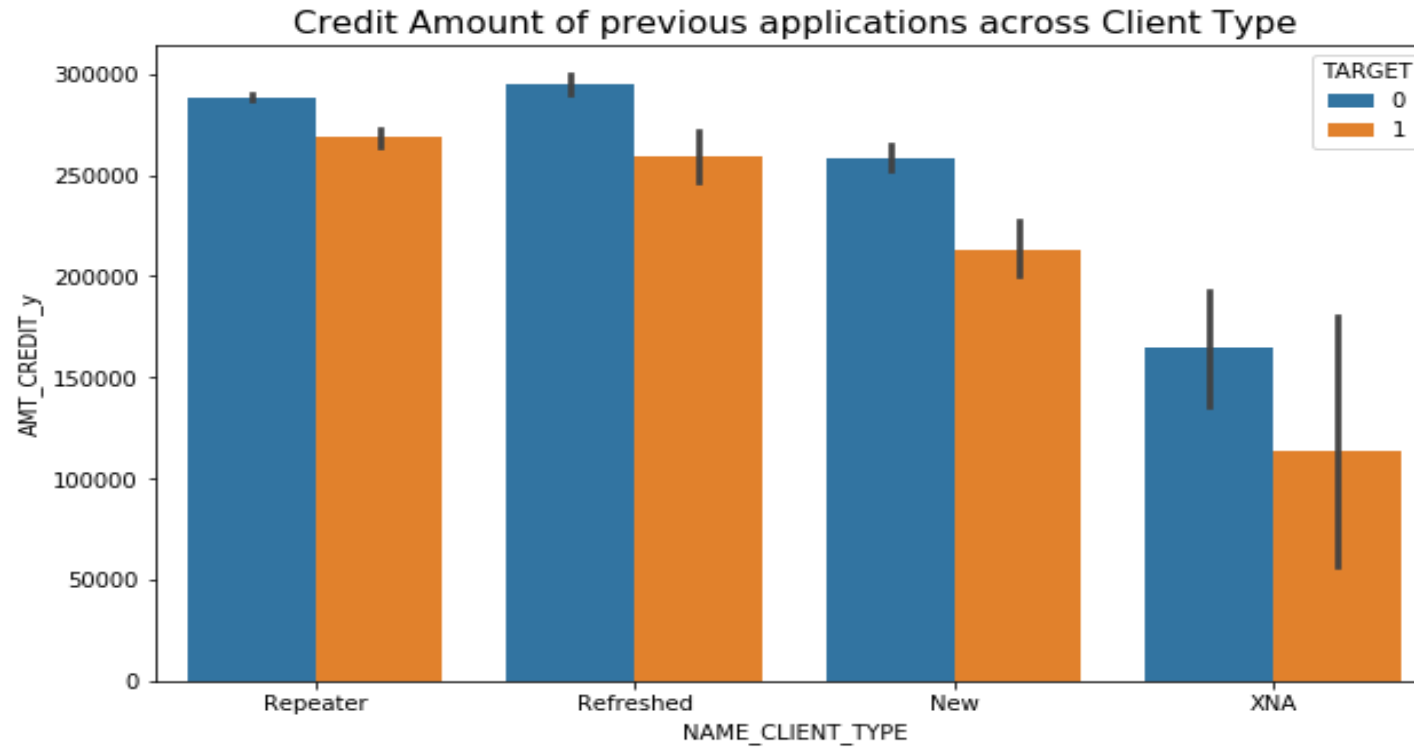


1. From the graphs above it can be seen that, Relative to current application when was the decision about previous application made is influencing factor.



2. Less the days for decision about previous application made, more likely to have payment difficulties

Total income for client across contract status for previous applications :



- 1.Credit amount for Revolving loans and Cash loans are checked against client type for previous applications.
- 2.There is no significant difference seen for Repeater/Refreshed/New client types.
- 3.For client type 'New' who are likely to default have average credit amount < 250000

Key points :

The most influencing factors from the both current and previous application data sets are as below :

- 1.Family Status (Married clients tend to have less chance of becoming defaulters.)
- 2.Occupation type and Organization type
(Both the points give strong evidence for not to default if fall under certain common types)
- 3.Clients Total Income (less the income more likely to default).Also the respective Amount Credit/Goods Price/Annuity should be verified. For defaulters its inversely correlated.
- 4.Income type (Students/Pensioners/Businessman are absolute non defaulters.)
- 5.Age of Client (Trend from current as well as previous application shows age group from 32 to 45 yrs are likely to default)
- 6.Previous application contract status (24.5% previous application in refused status contribute in defaulters list)
- 7.Rejection reason for previous application should be checked to verify compliance check/blacklisted customer etc)
- 8.Number of days taken for decision of previous applications (Less the number of days , more likely to default)
- 9.Client type against credit amount should be verified.