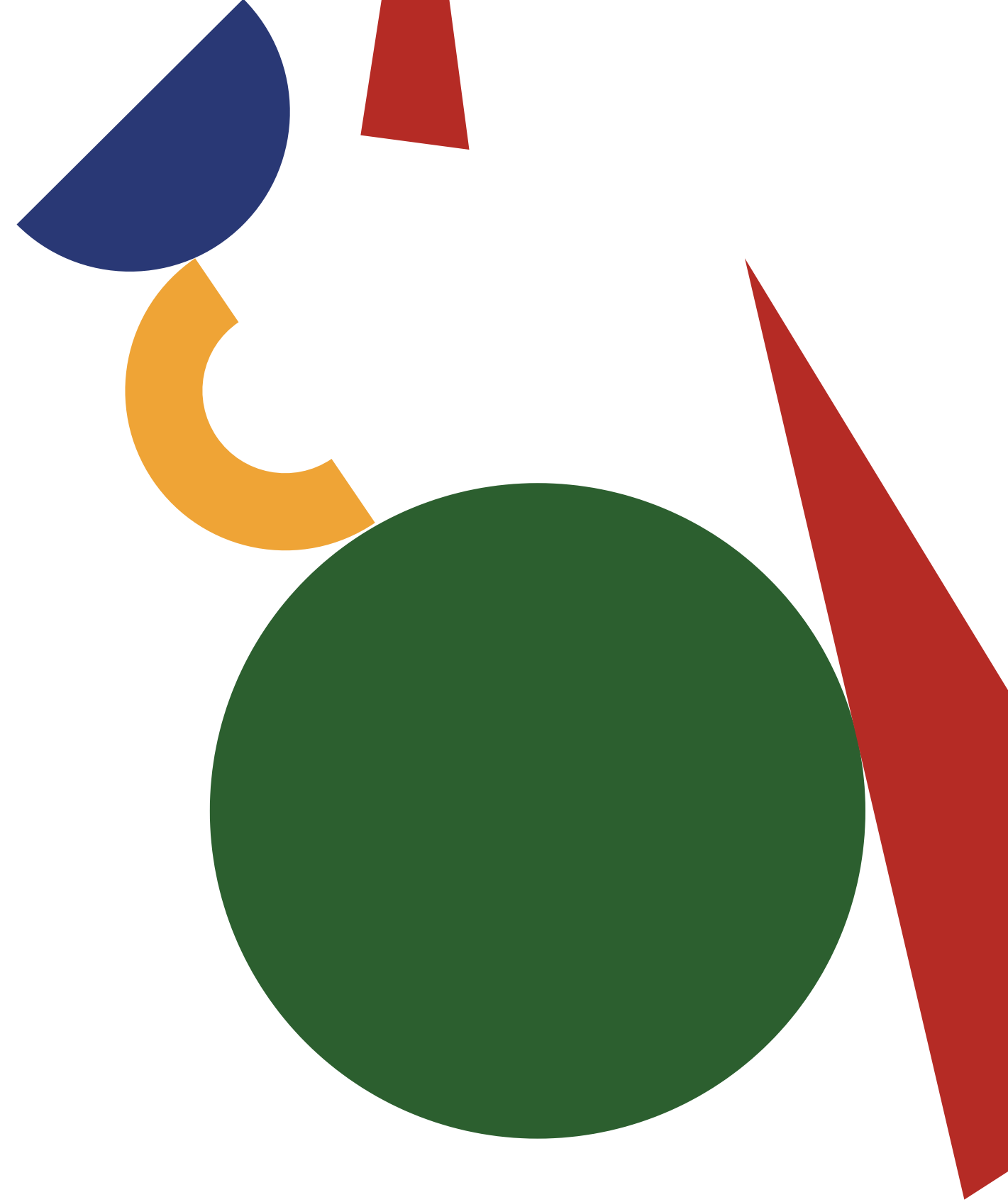
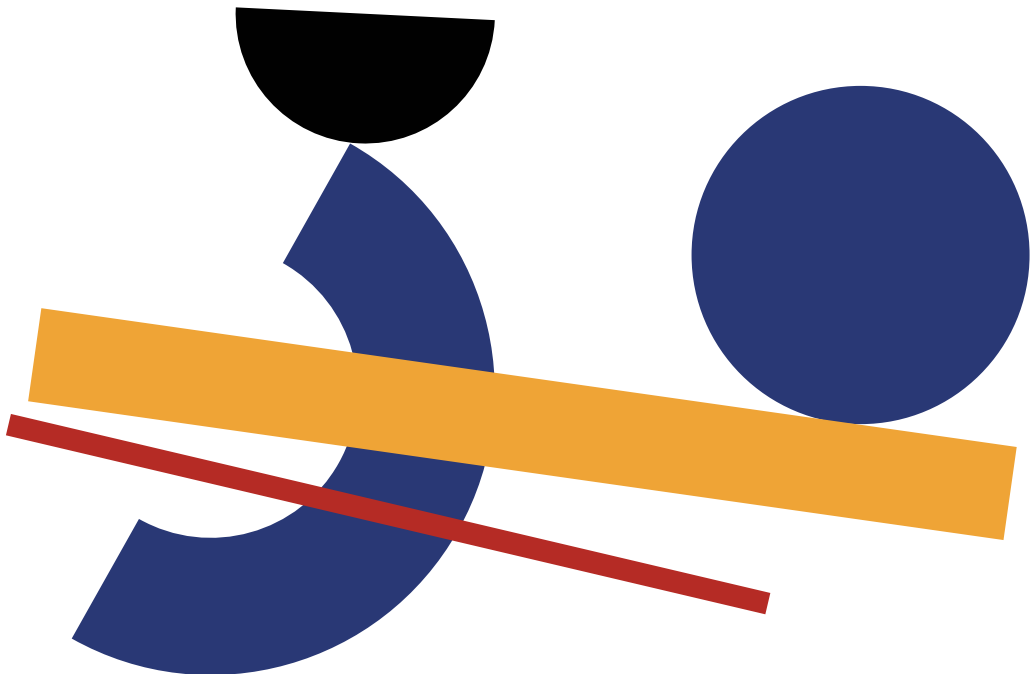




CREDIT EDA CASE STUDY

SAITEJA GUNDETI





Steps for Analysis



01

Data Cleaning of 'application_data' Dataset

02

Visualization of 'application_data' Dataset

03

Binning & Imbalance Percentage

04

Univariate Analysis & Bivariate Analysis

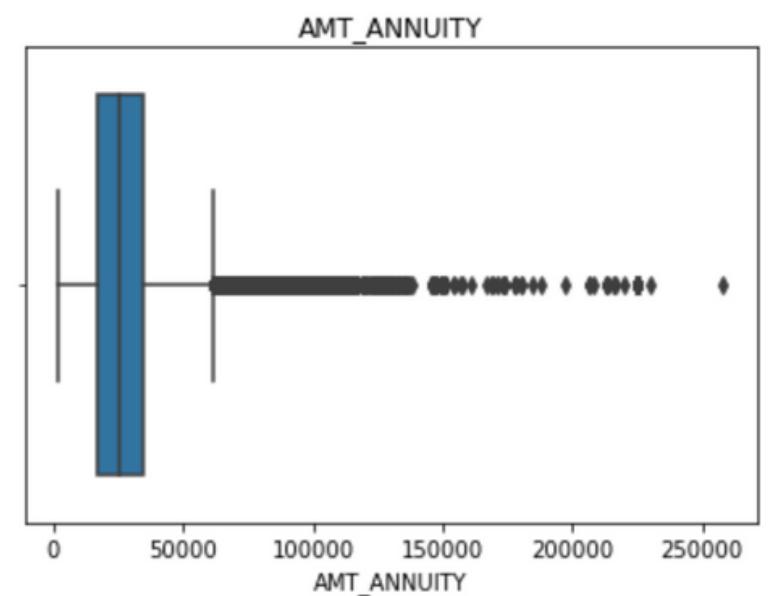
05

Data Cleaning of 'previous_data' Dataset

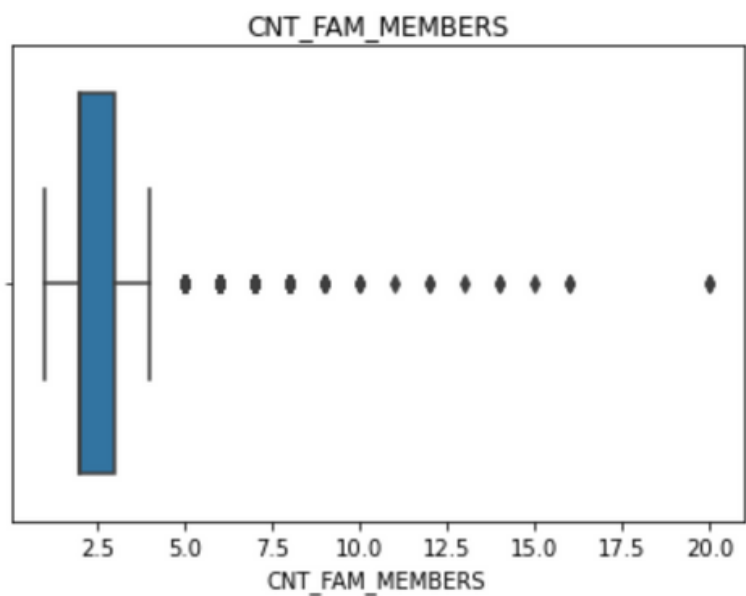
06

Univariate Analysis

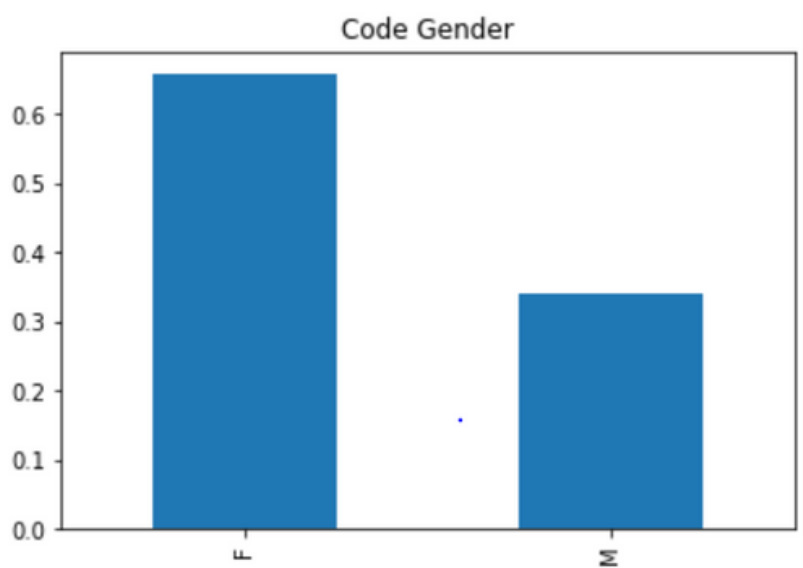
Visualization of 'application_data' Dataset



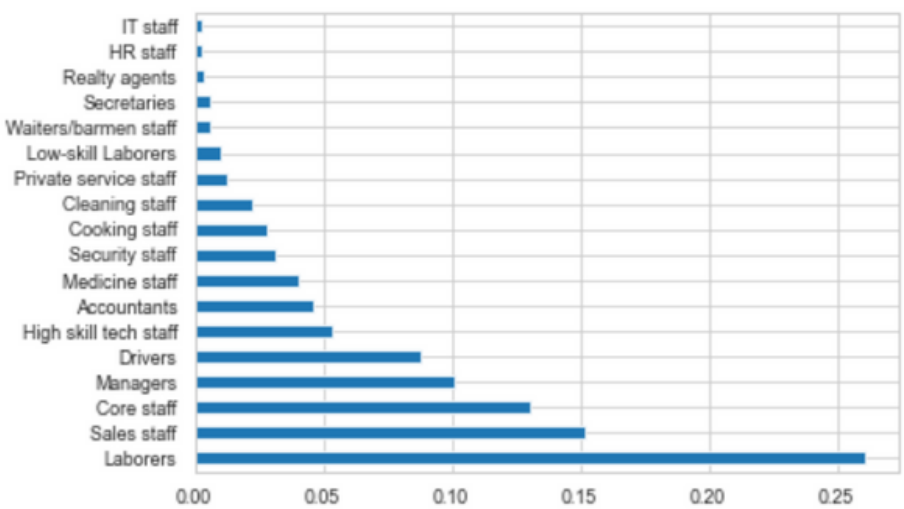
In AMT_ANNUIITY, we observed that the median, 25th and 75th Percentile lies between 0 and 50000.



In CNT_FAM_MEMBERS, we observe that the majority of the family members lie in the range of 2.5

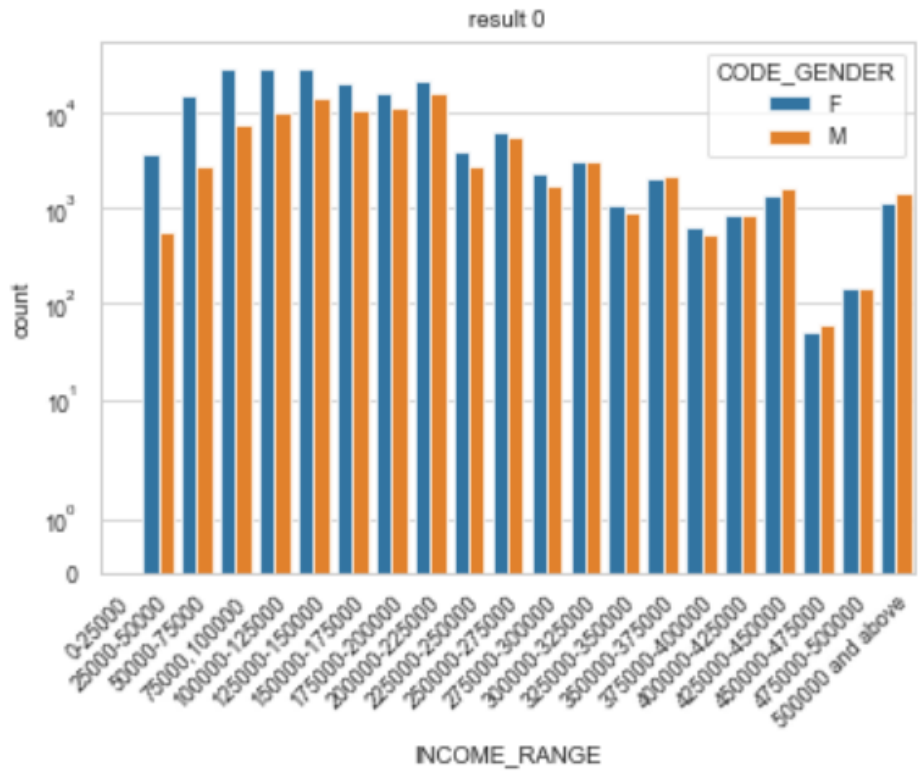


In CODE_GENDER, we observe that the majority of the clients are Female

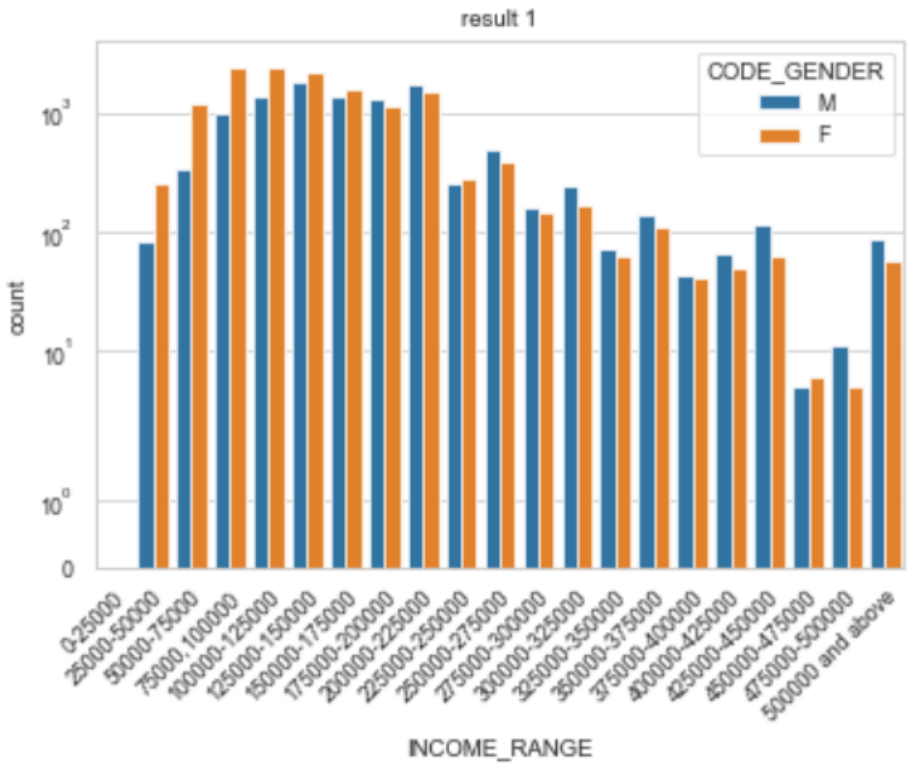


In OCCUPATION_TYPE, we observe that majority of clients are labourers

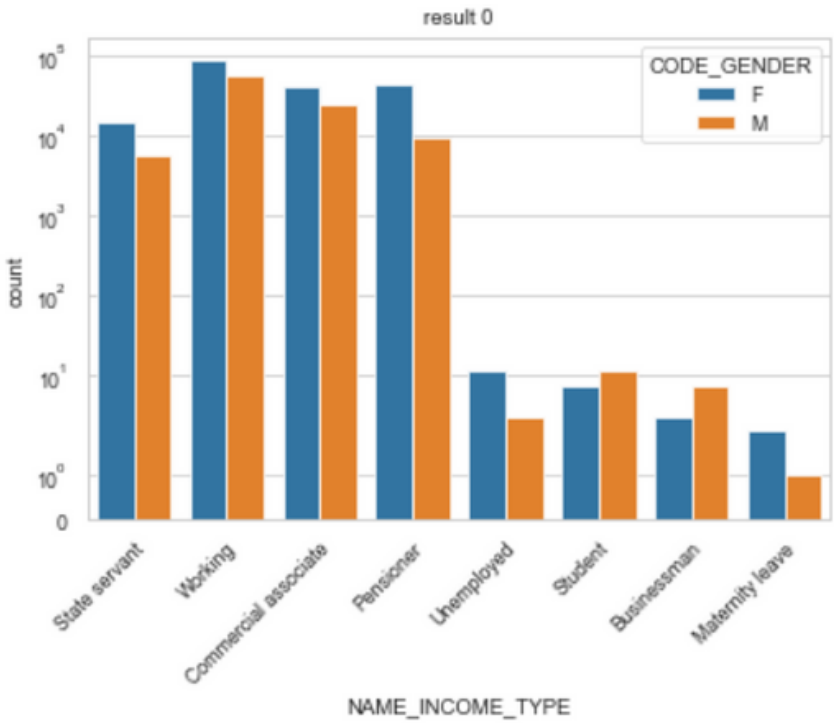
Univariate Analysis of 'application_data' Dataset



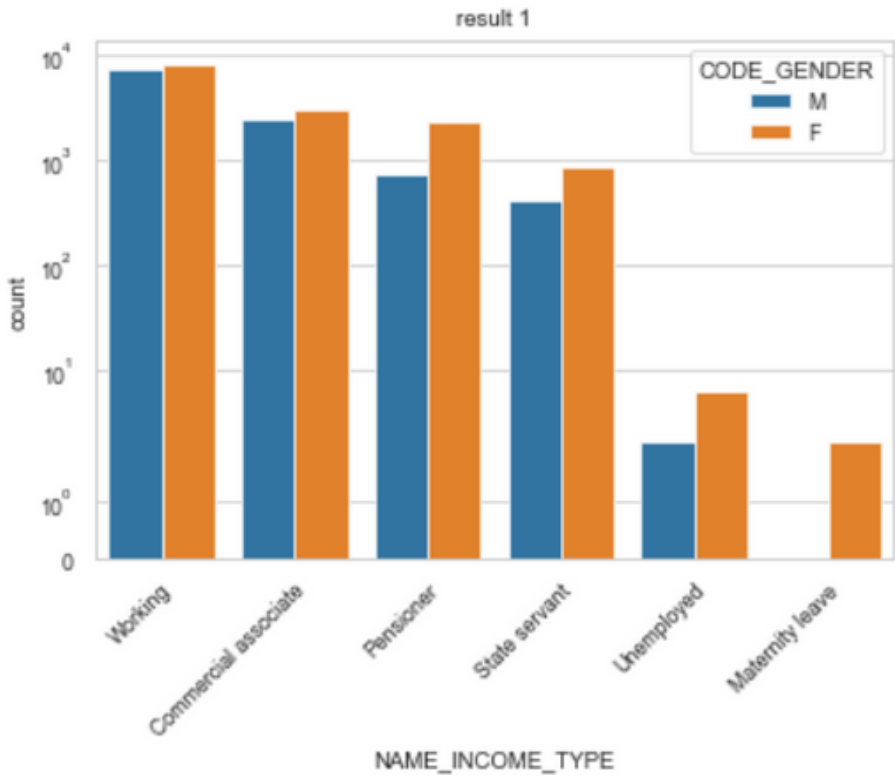
The clients who are not defaulters have their maximum income range between 100000-150000 for Females & 100000-225000 for Males



The clients who are defaulters have their maximum income range between 100000-150000 for Females & 100000-225000 for Males

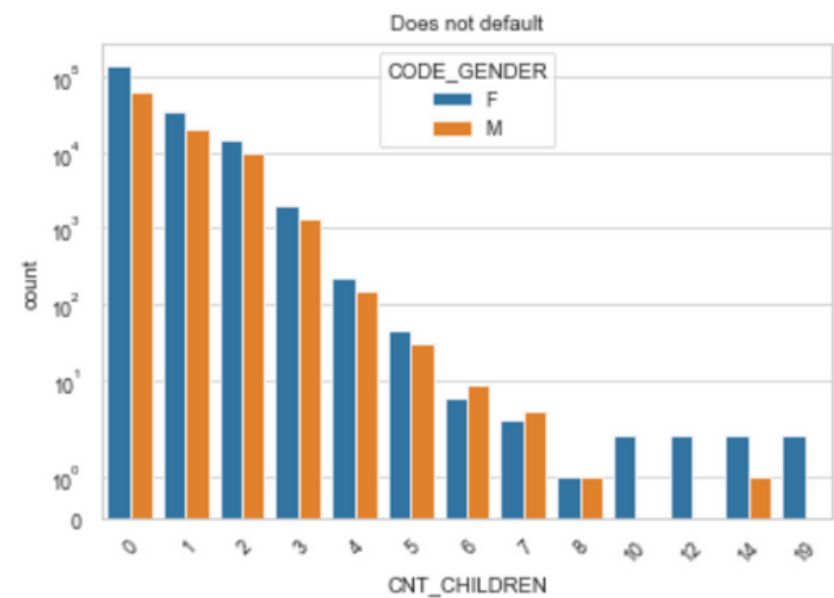


The income type of maximum clients who are not defaulters for both Male & Female are State servant, Commercial associate, Pensioner & Working

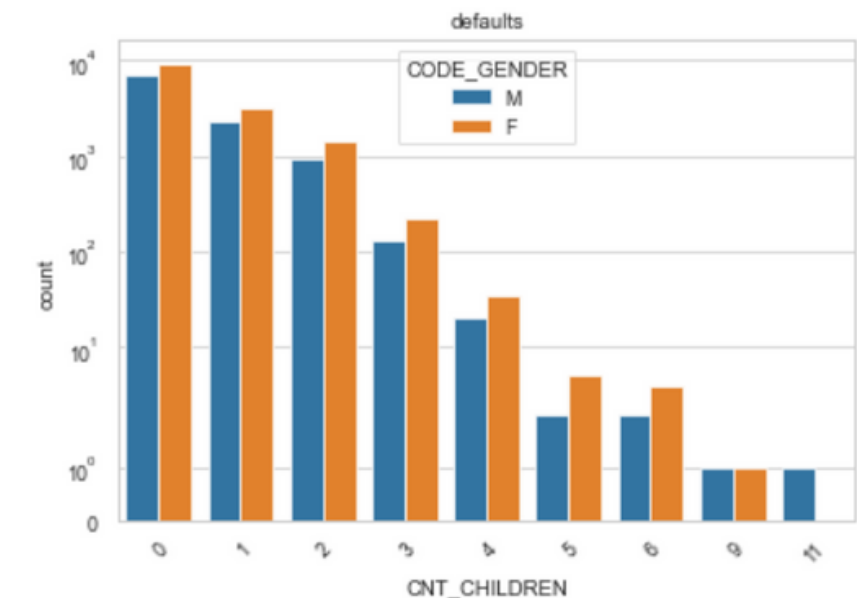


The income type of maximum clients who are defaulters for both Male & Female are Working, Commercial associate & Pensioner

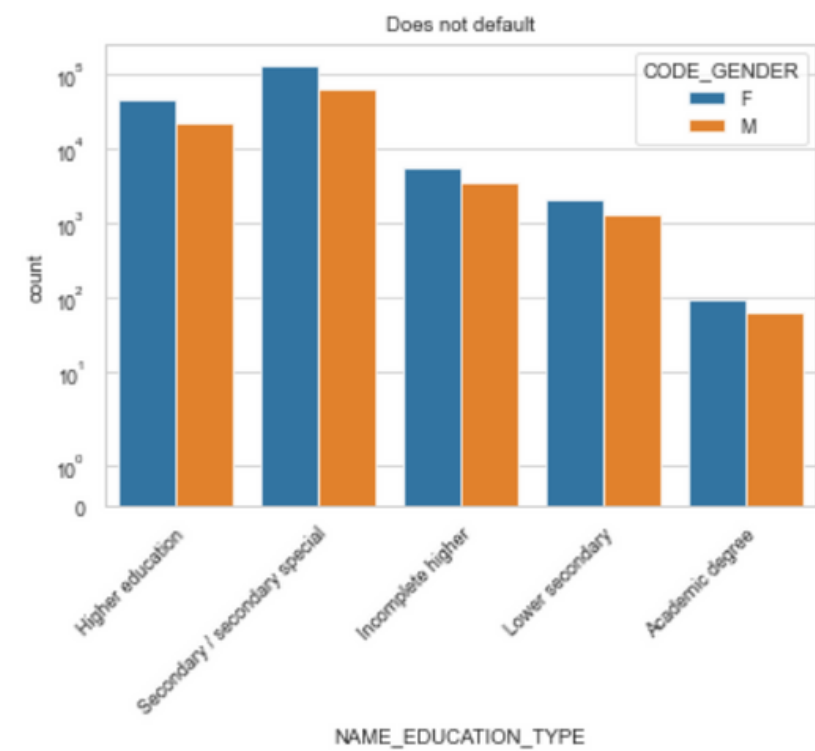
Univariate Analysis of 'application_data' Dataset



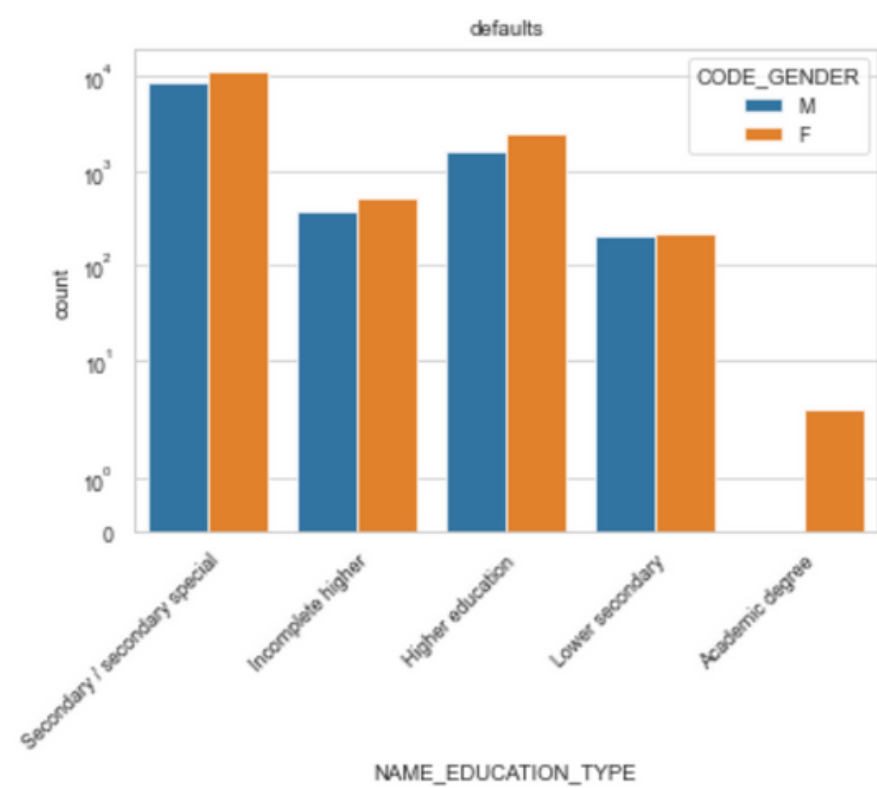
The clients who don't default have maximum children between 0-1 for both the Male & Female clients



The clients who default have maximum children between 0-1 for both the Male & Female clients

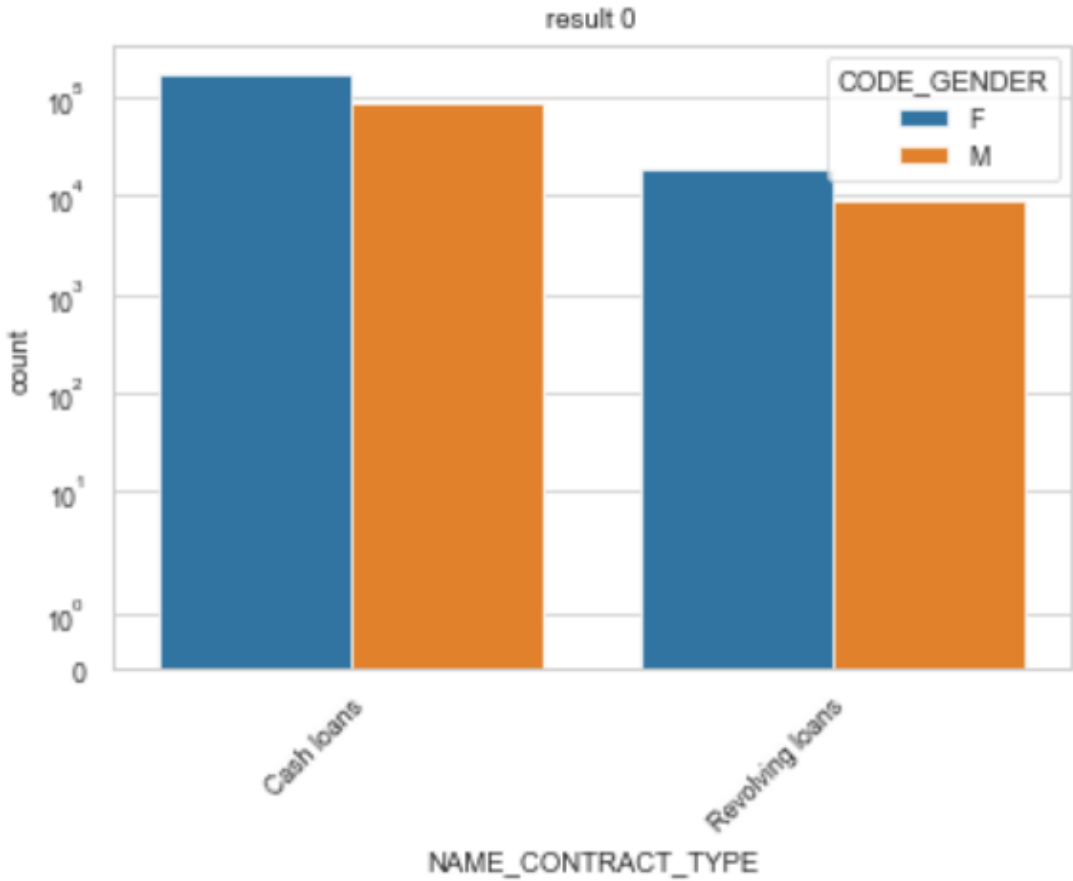


The clients who don't default have maximum education of Secondary education



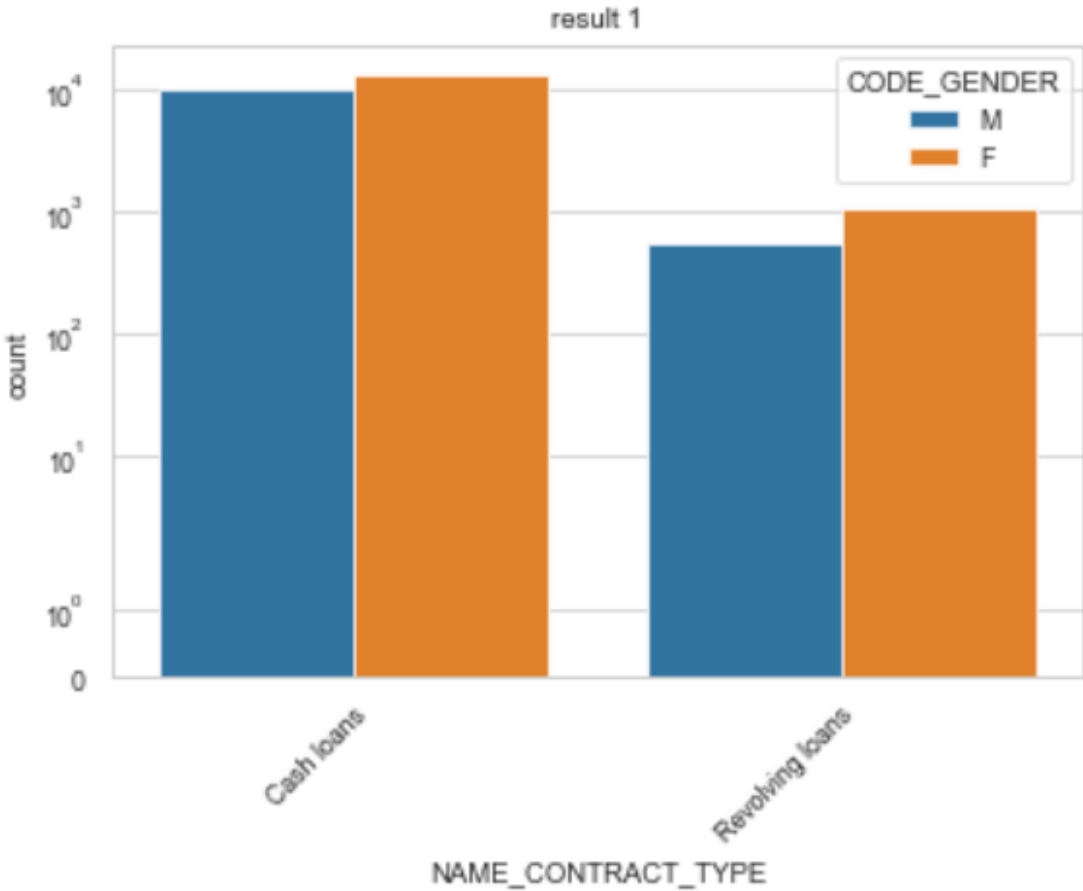
The clients who default have maximum education of Secondary education

Univariate Analysis of 'application_data' Dataset

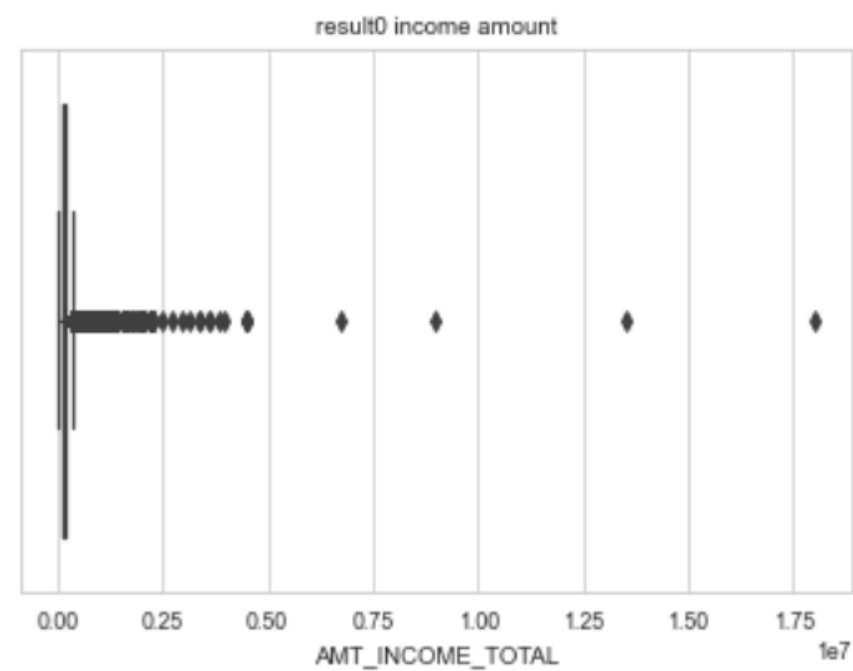


The maximum Non defaulters have contract type of cash loans

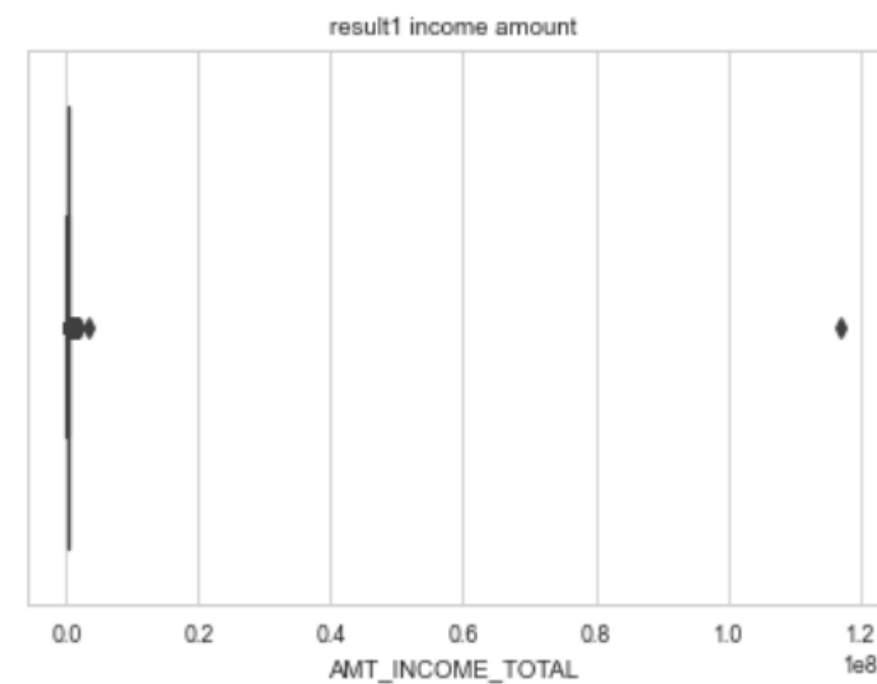
The maximum defaulters have contract type of cash loans



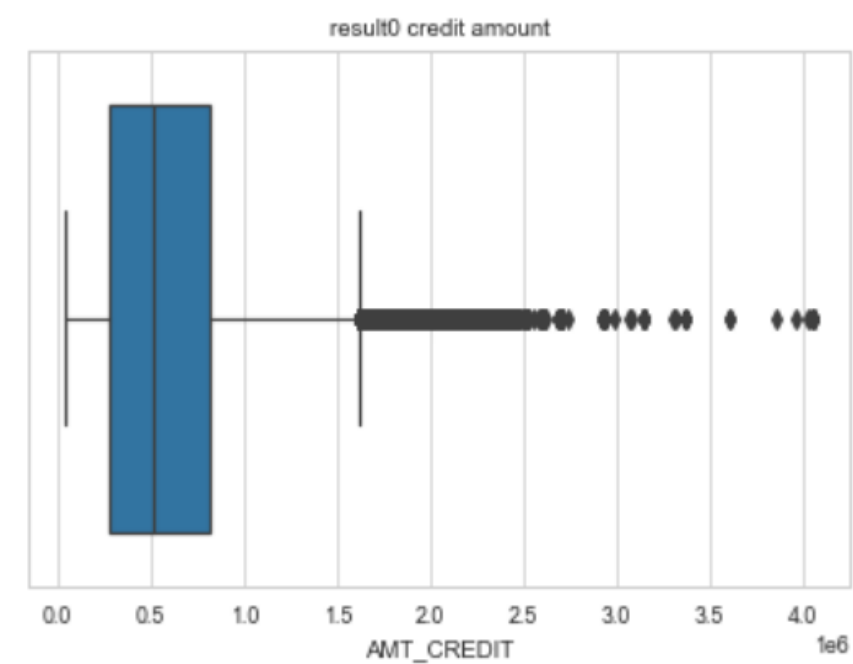
Univariate Analysis of 'application_data' Dataset



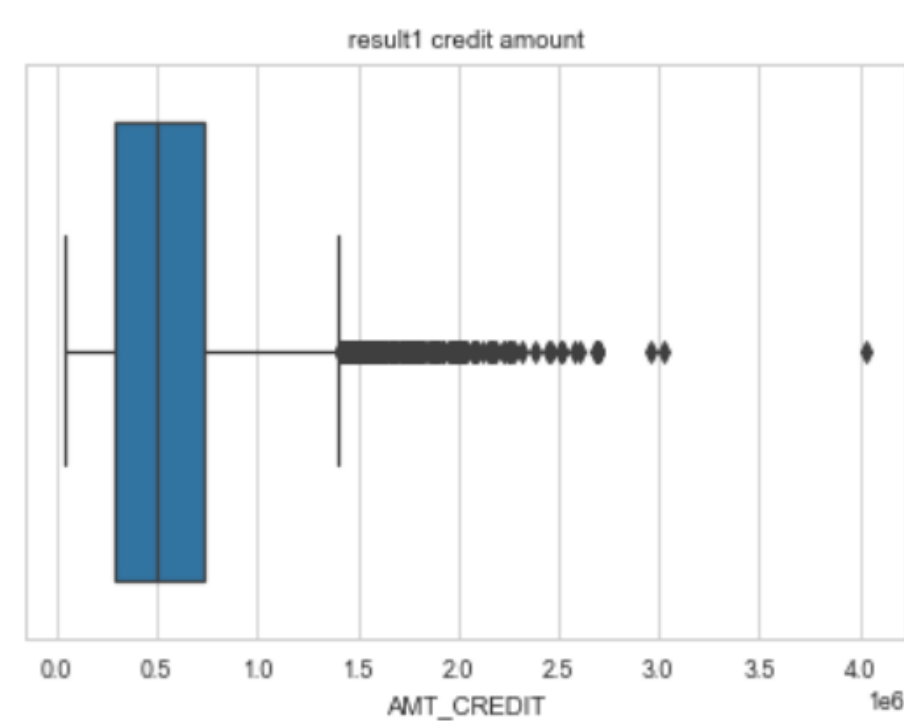
The Non Defaulters have their income in the 1st Quartile and there are Outliers too



The Defaulters have their income in the 1st Quartile and there are Outliers too

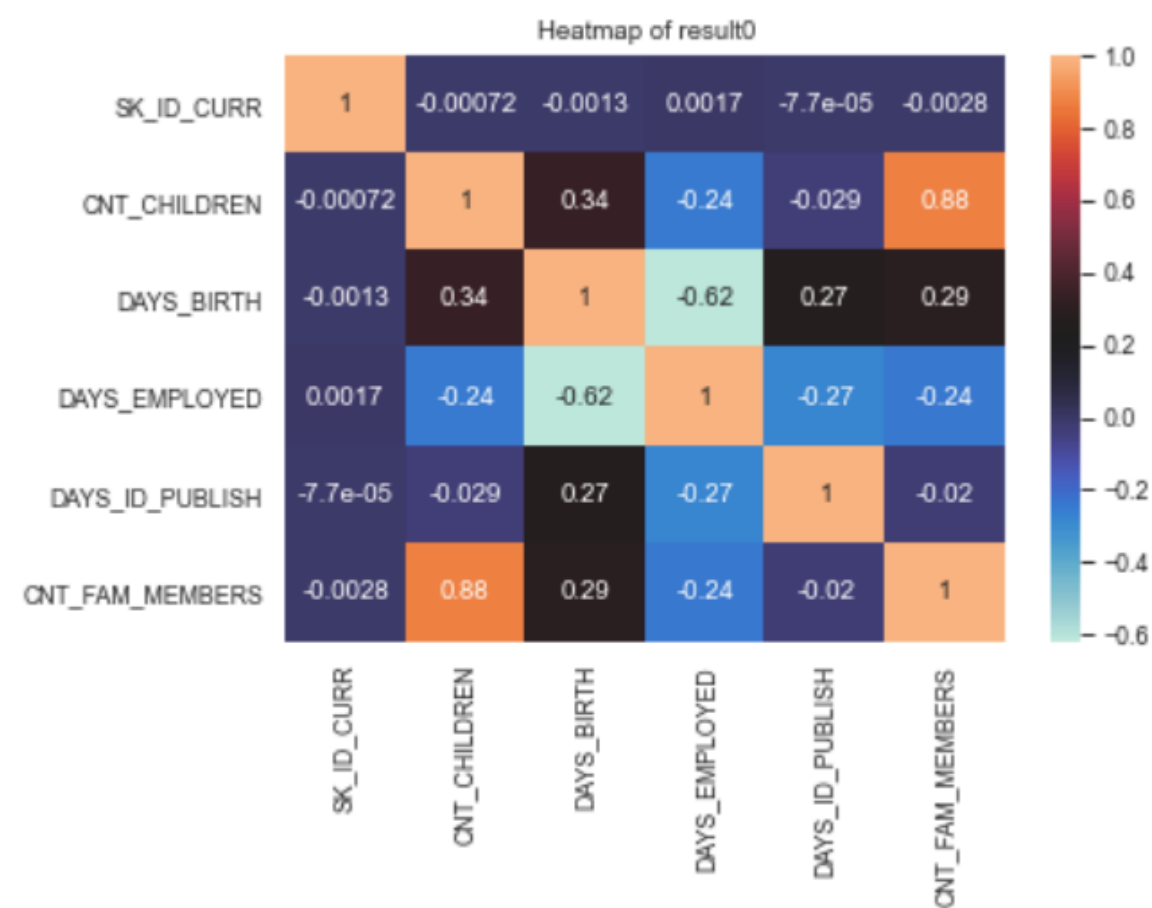


The Non Defaulters have their credit amount in the 1st Quartile and there are Outliers too



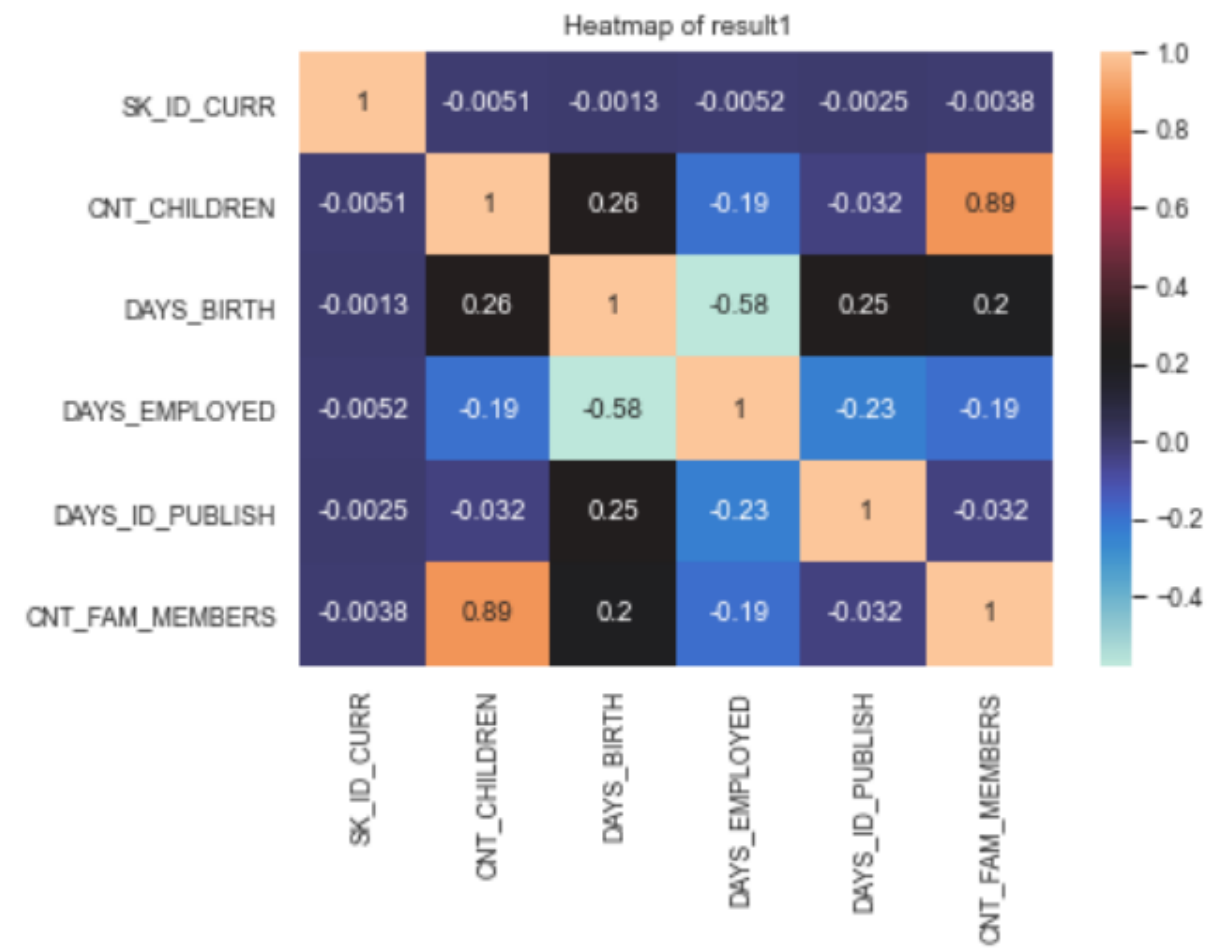
The Defaulters have their credit amount in the 1st Quartile and there are Outliers too

Bivariate Analysis of 'application_data' Dataset

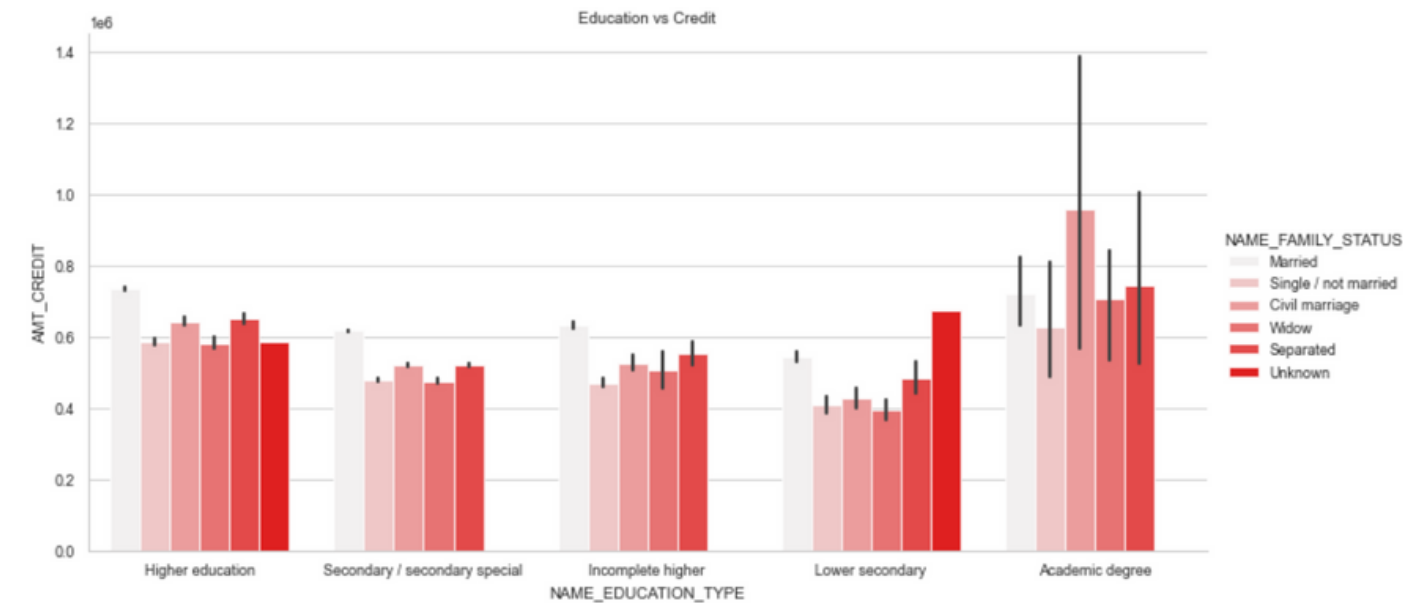


The Non Defaulters have High Correlation Values in
CNT_CHILDREN & CNT_FAM_MEMBERS
CNT_CHILDREN & DAYS_BIRTH

The Defaulters have High Correlation Values in
CNT_CHILDREN & CNT_FAM_MEMBERS



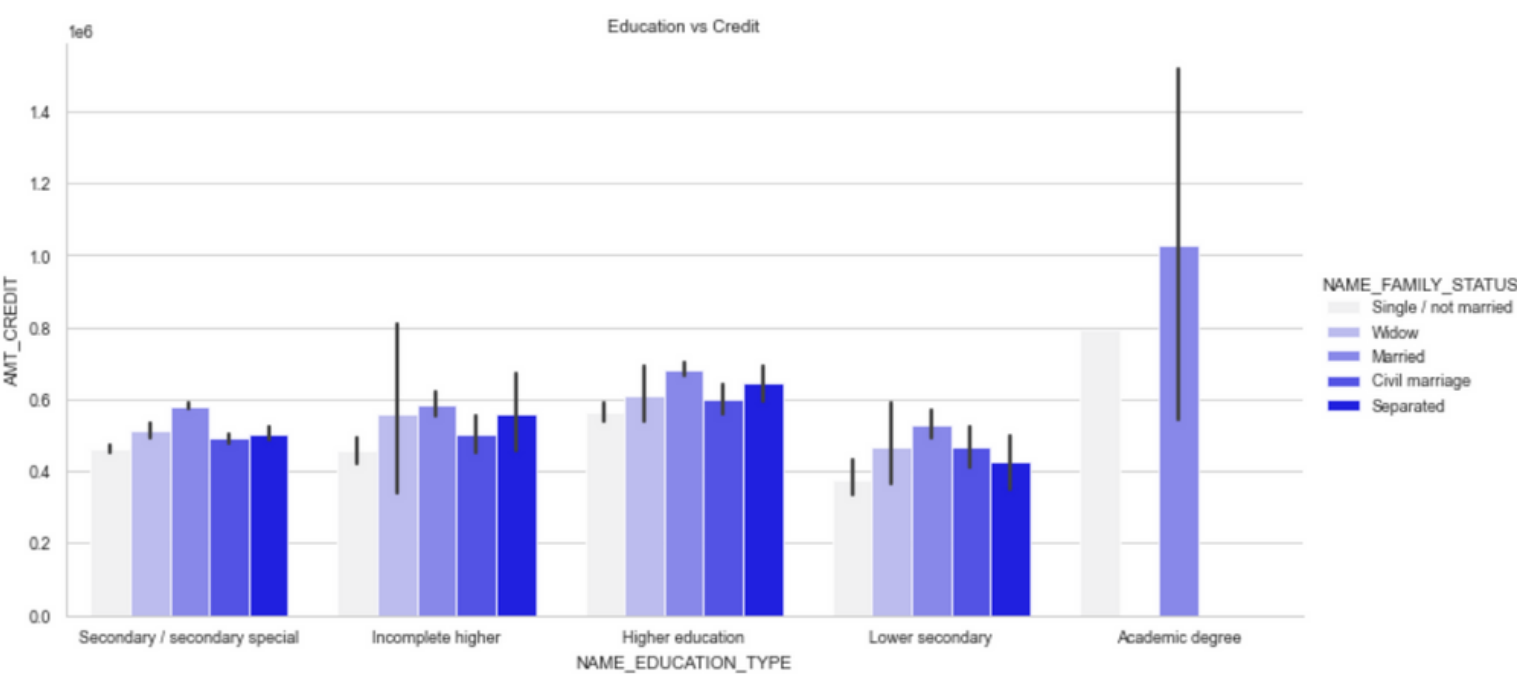
Bivariate Analysis of 'application_Data' Dataset



The Non Defaulters who have Academic Degree are the ones having maximum credit amount in which 'civil marriage' is the highest & 'single' is the lowest

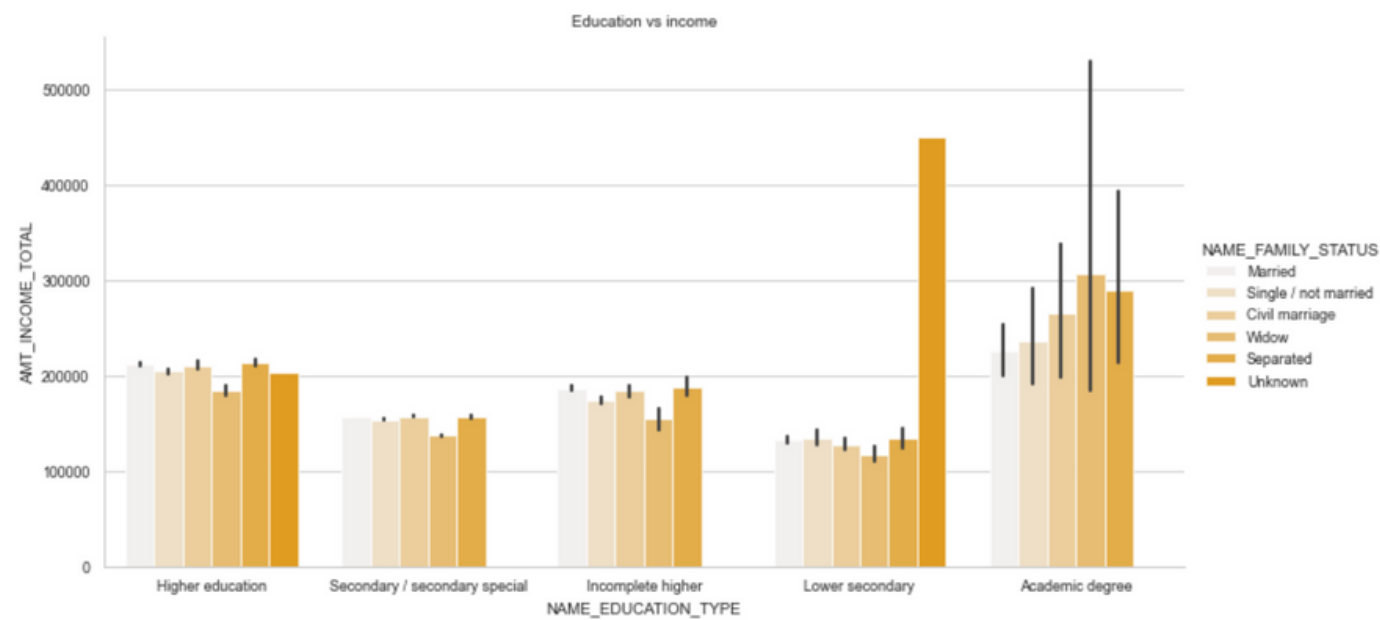
result 0

The Defaulters who have Academic Degree are the ones having maximum credit amount in which 'married' is the highest & 'single' is the lowest



result 1

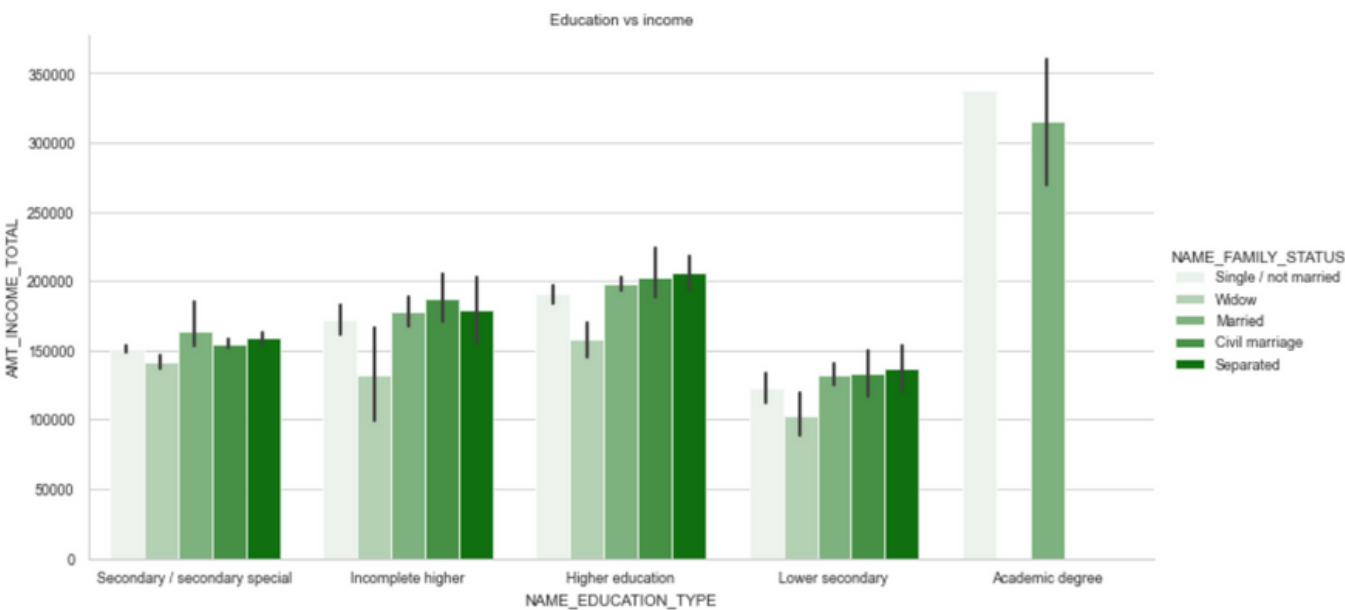
Bivariate Analysis of 'application_data' Dataset



The Non Defaulters who have Academic Degree have the highest income in which 'widow' earn the highest & 'married' earn the lowest

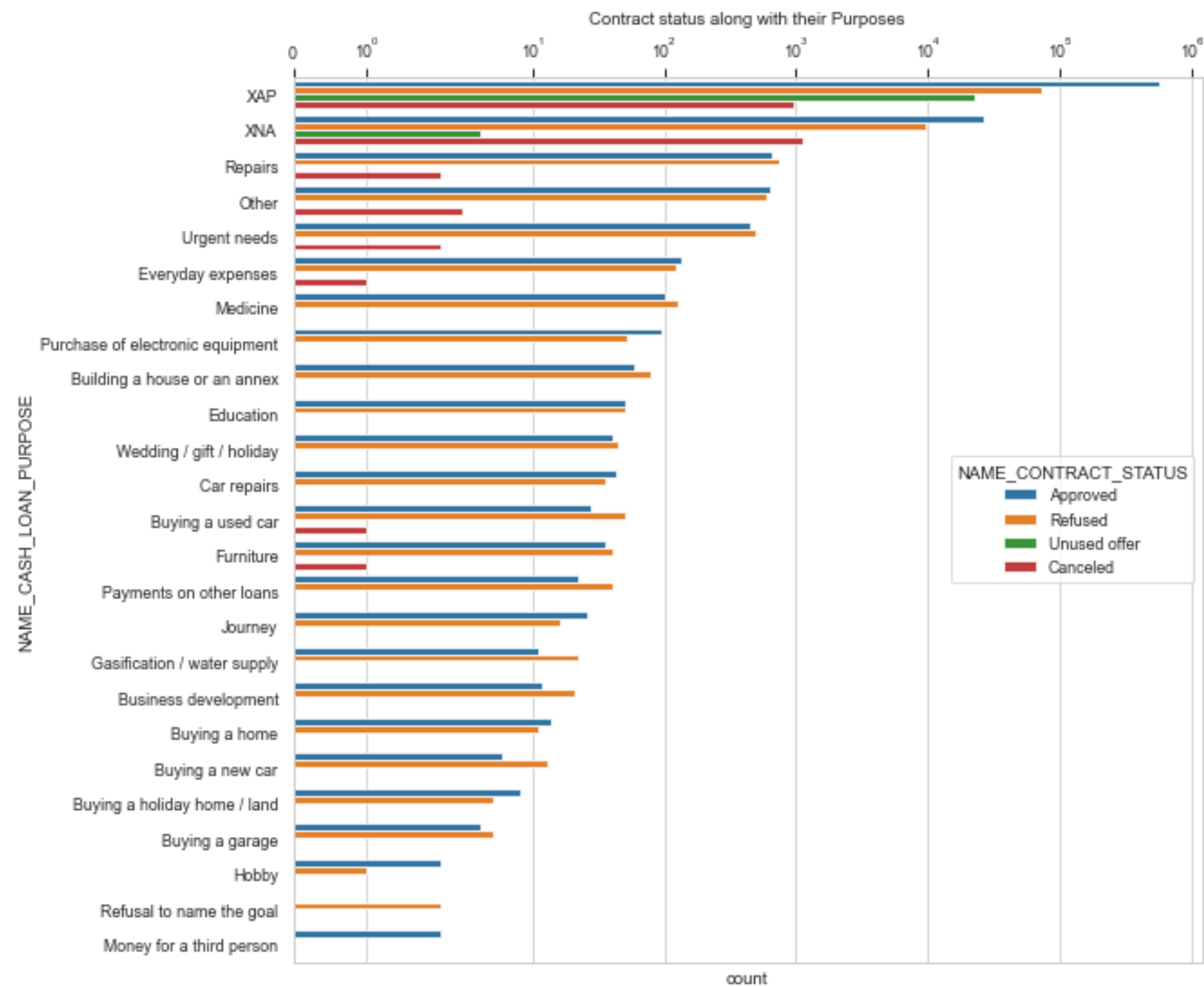
result 0

The Defaulters who have Academic degree have the highest income in which 'single' earn the highest & 'married' earn the lowest



result 1

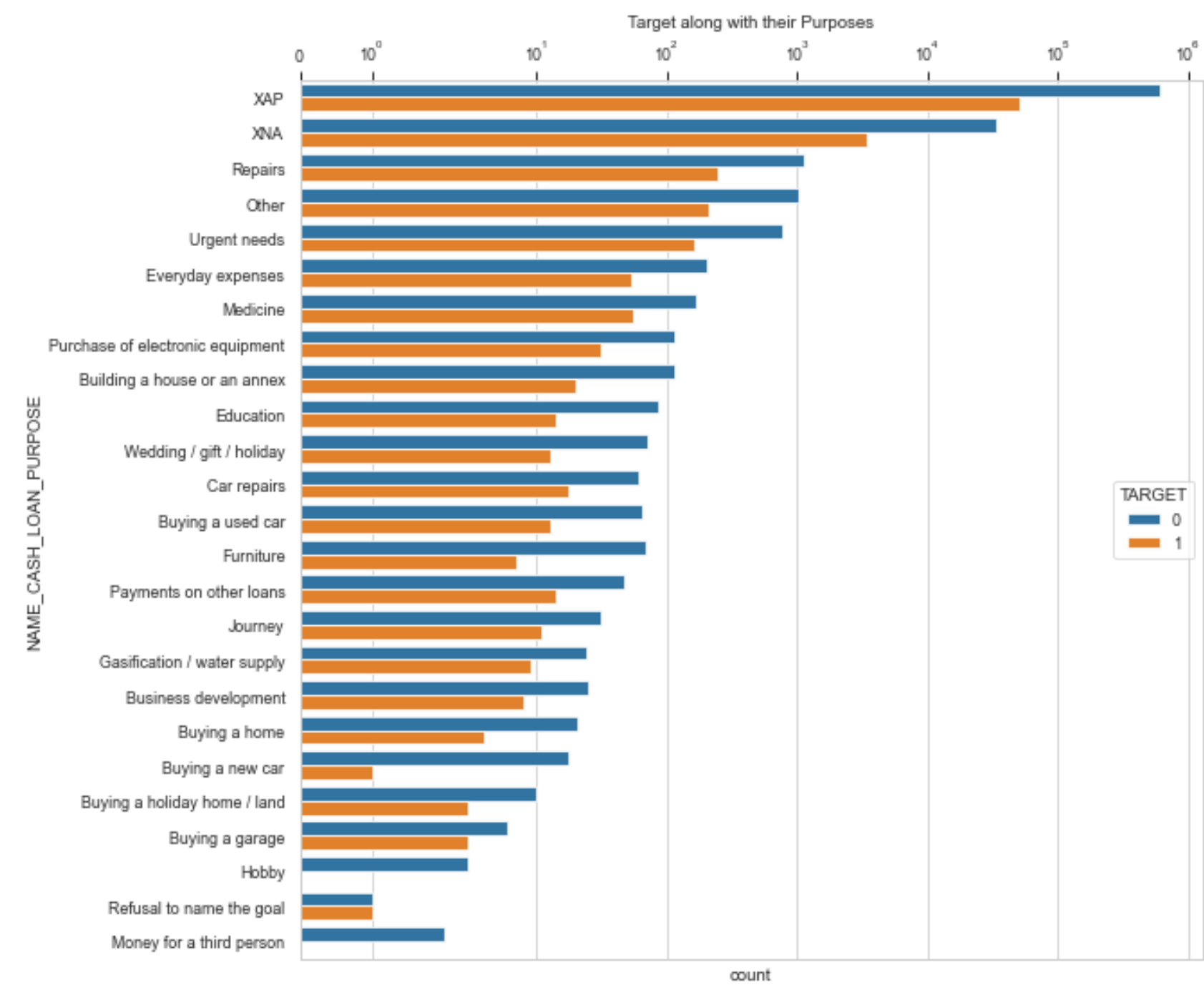
Univariate Analysis of 'previous_application' Dataset



Contract Status along with their Purposes

From the above graph, we can analyse that the Purpose 'Repairs' have the highest rejection of loans & Purpose 'Education' have equal no of Accepted & Rejected loans.

Univariate Analysis of 'previous_application' Dataset



Target along with their Purposes

From the above graph, we can analyse that Purpose 'Repairs' have maximum no of Non Defaulters & Defaulters & Purpose 'Education' have equal no of Non Defaulters & Defaulters

01

For Non Defaulter Clients, Females are more than Males whereas for Defaulter Clients, Males are more than Females.

02

For Non Defaulters & Defaulters, the No of Children does not affect so thus we cannot conclude anything.

03

The Clients with 'Secondary Education' have Defaulted the most.

04

'Cash loans' have higher no of credit for both Non Defaulters & Defaulters.

05

For Non Defaulters, Customers having 'Academic Degree' have high amount of credit in which 'Civil marriage' are the highest & 'lower educated' clients have less amount of credit in which 'widows' are the lowest.

06

For Defaulters, Customers having 'Academic Degree' have high amount of credit in which 'Married' are the highest defaulters & 'single' are the lowest among them.

07

From the analysis, it is concluded that the clients with Purpose 'repairs' have more no of rejection of loans & clients with Purpose 'Education' have equal amount of approved & rejected loans.

08

It is concluded that the clients with Purpose 'repairs' are the ones who have defaulted the most.

Conclusion