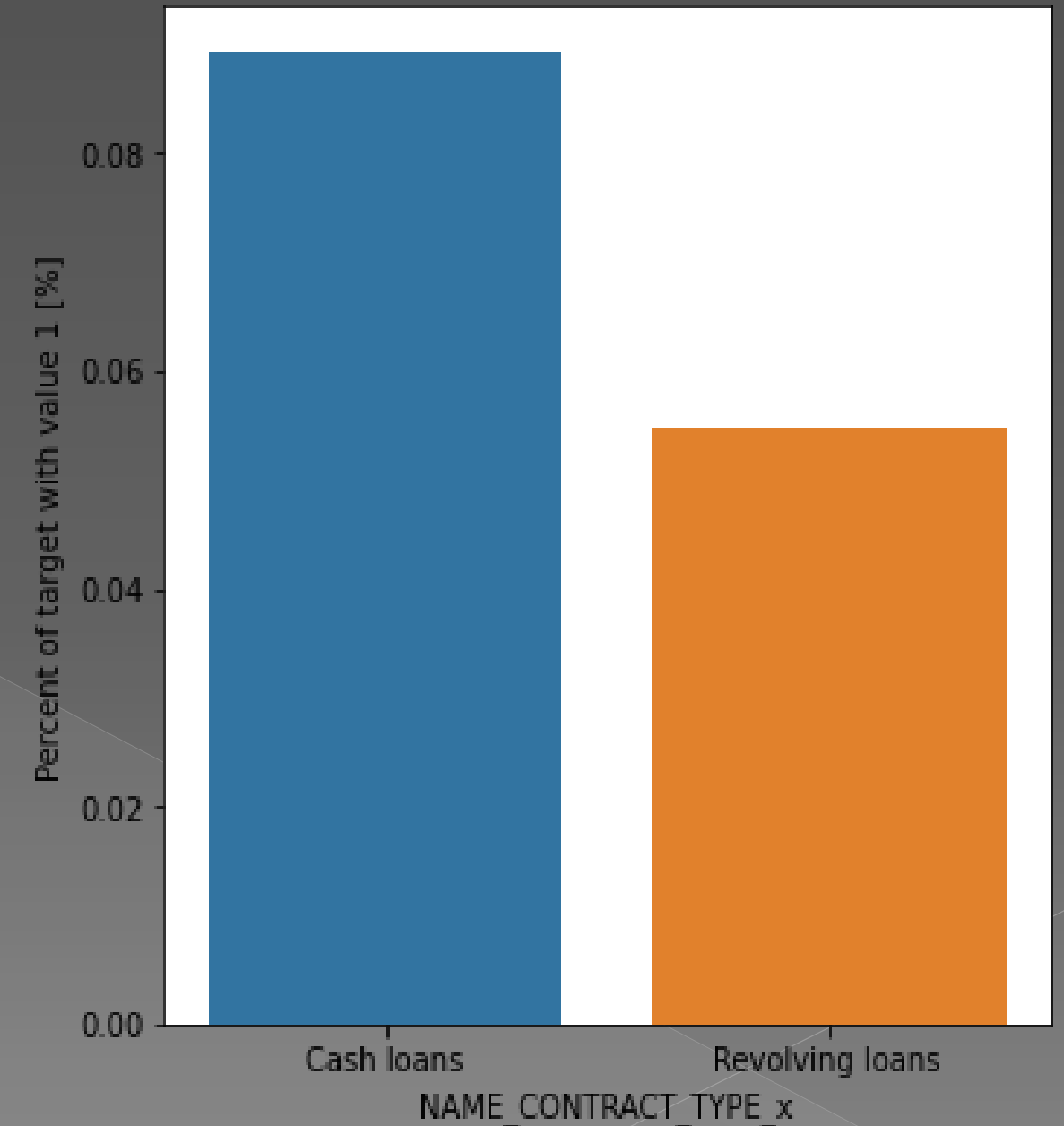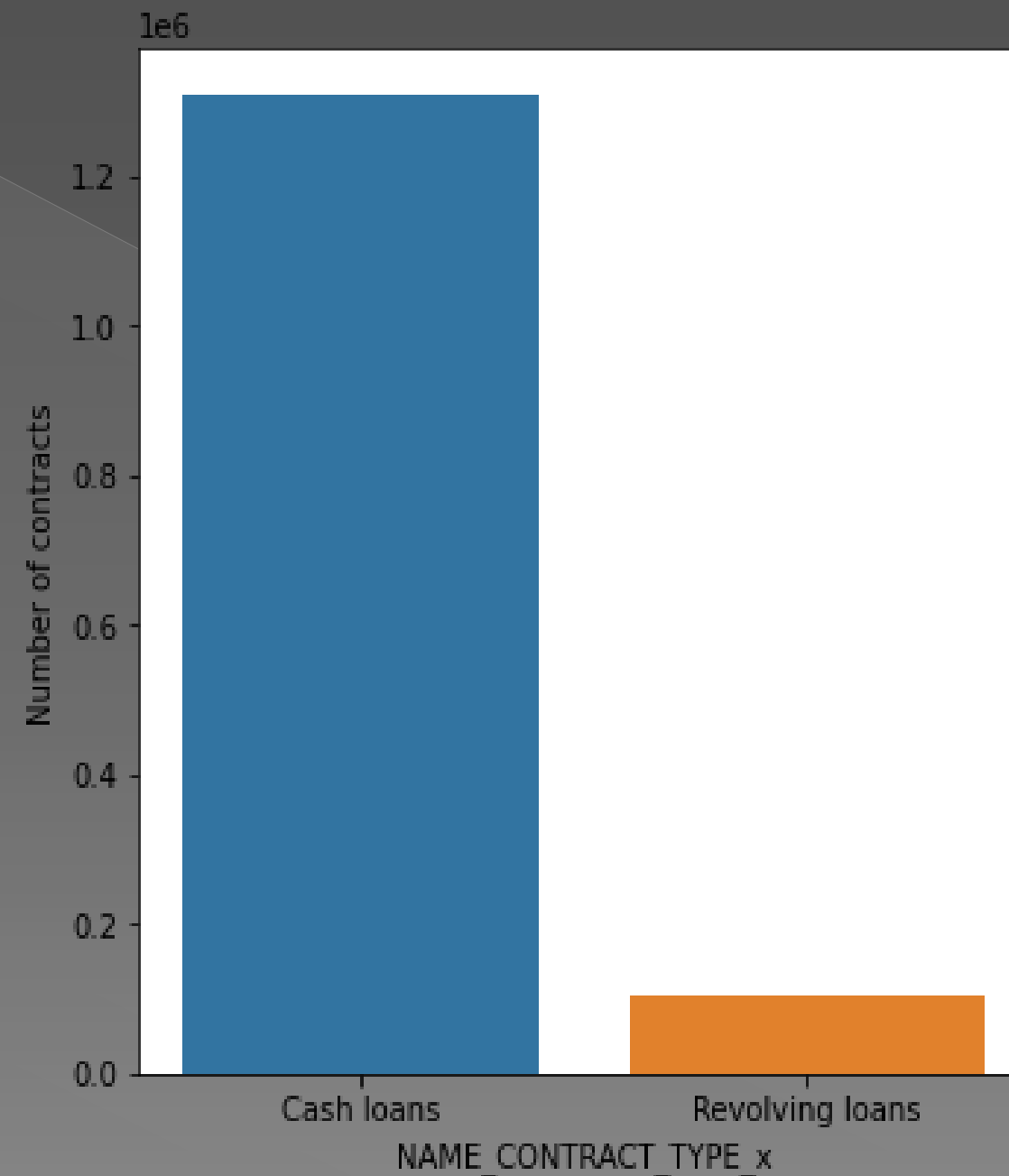# CREDIT EDA CASE STUDY

By Vijay Vinayak
M A Nadeem

Univariate analysis

Distribution for Contract Type

- We can see that the Revolving loans are small amount compared to Cash loans but the % of non payment for the revolving loans are comparatively high .
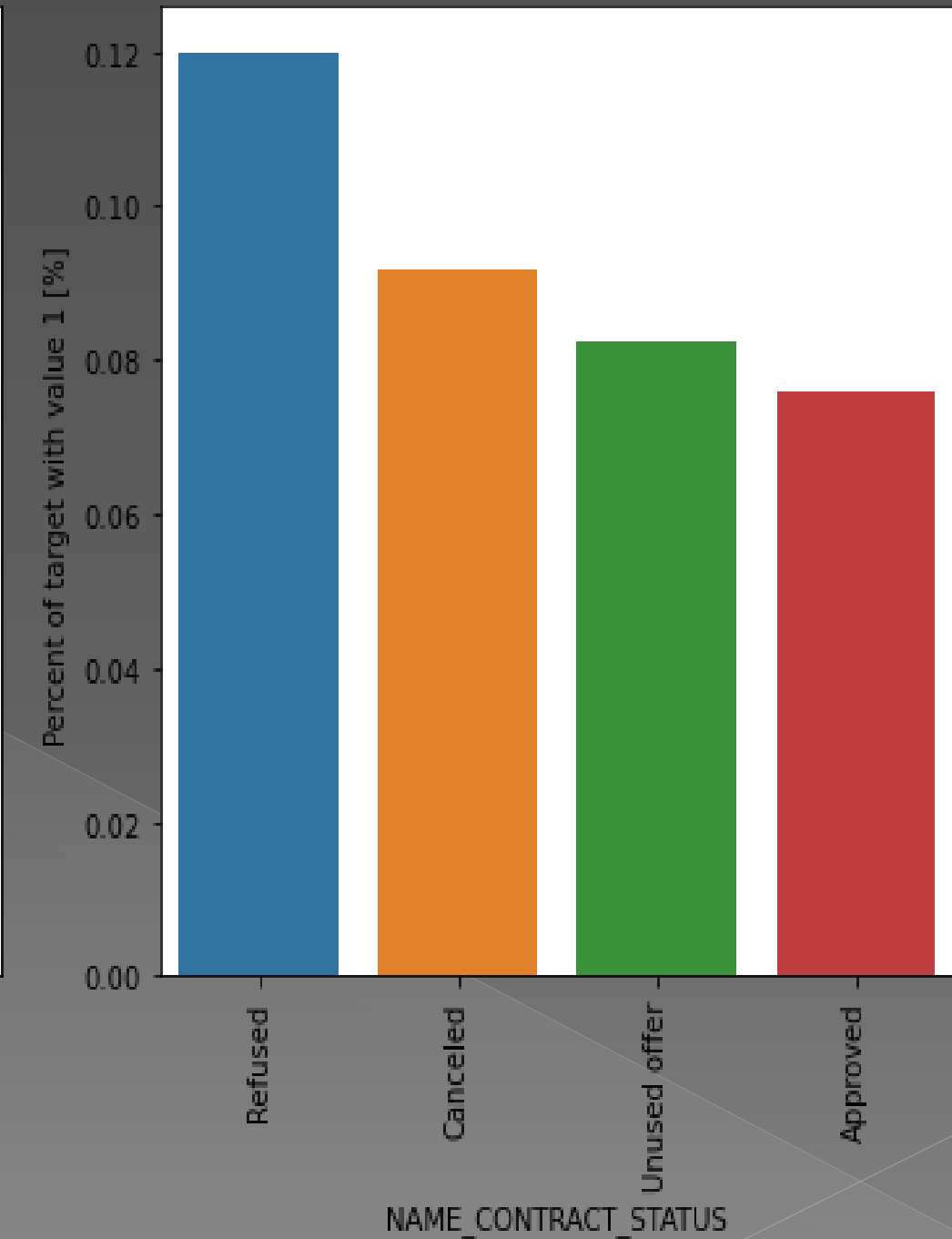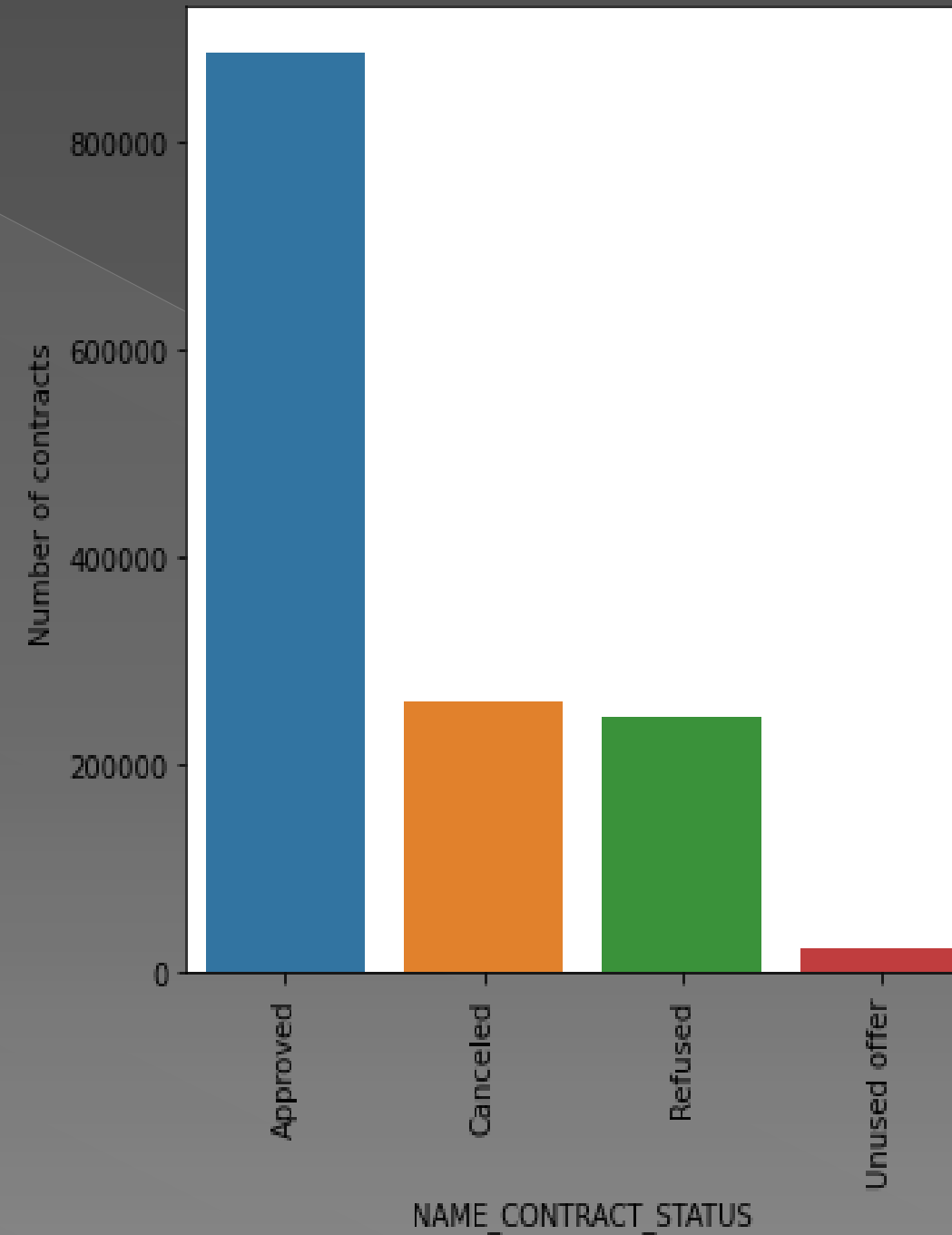
Purpose of Loan

- The credit amount of Loan purposes like 'Buying a home', 'Buying a land', 'Buying a new car' and 'Building a house' is higher.

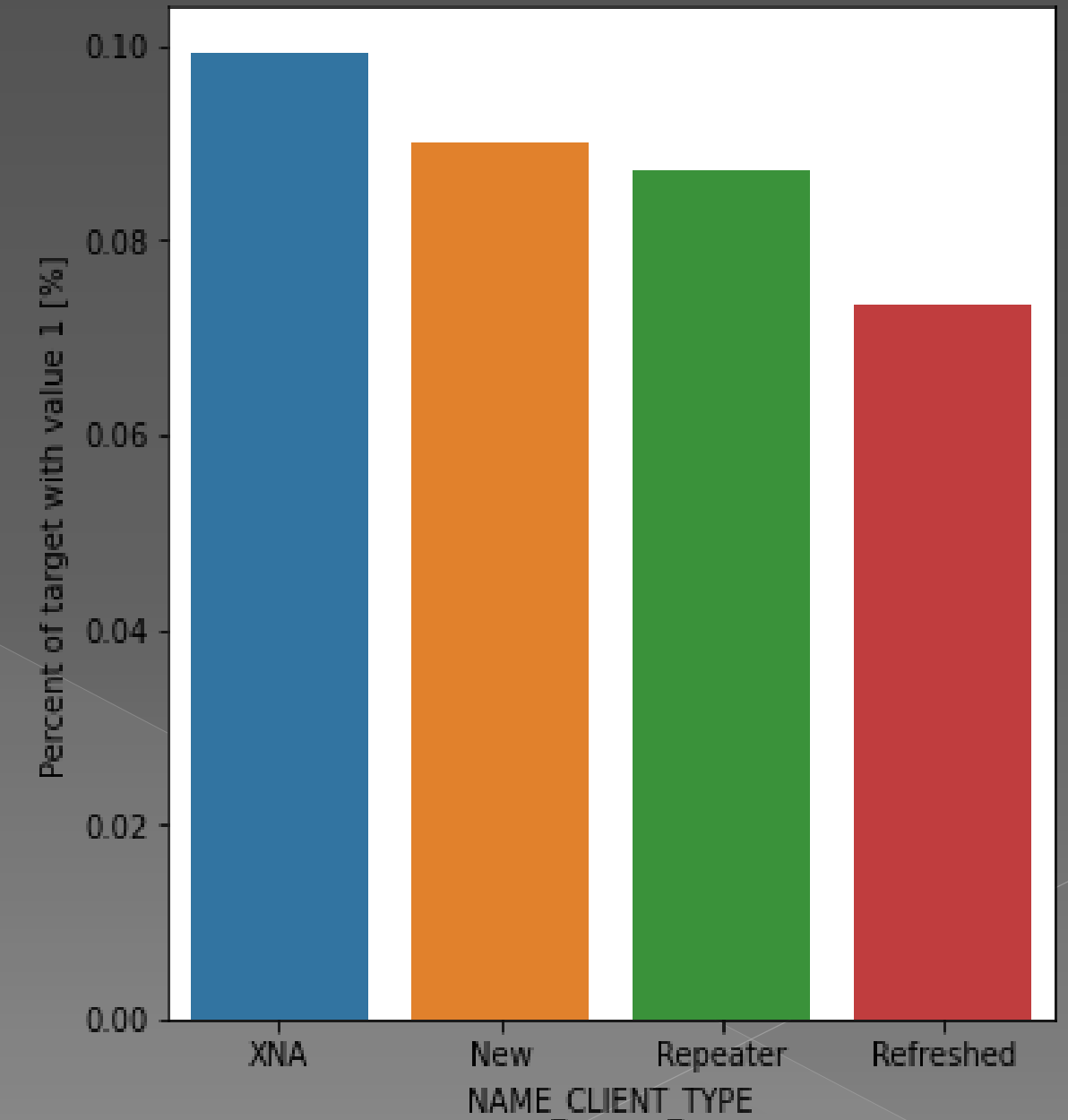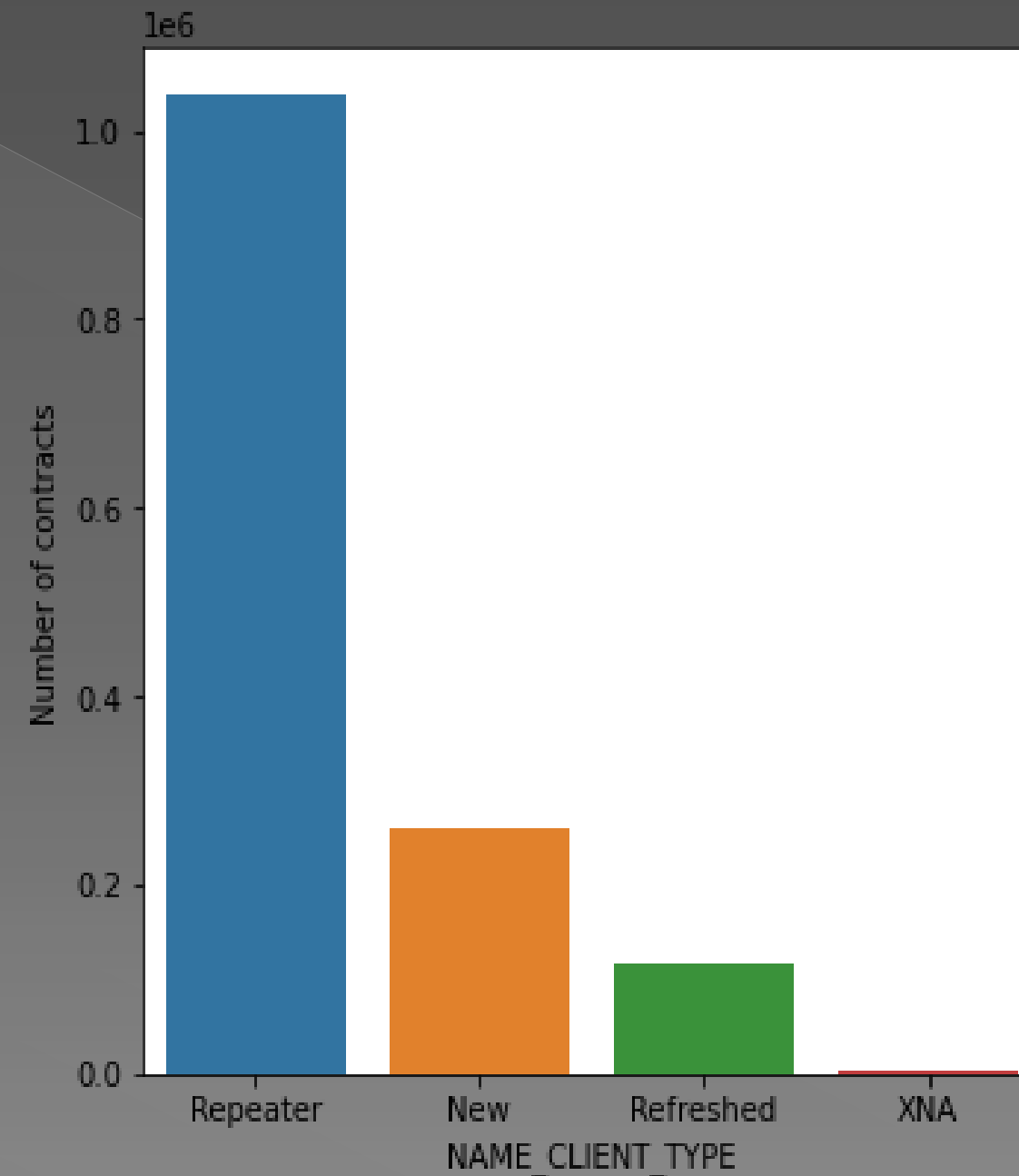- Loan purposes with 'Repairs' are facing more difficulties in payment on time.
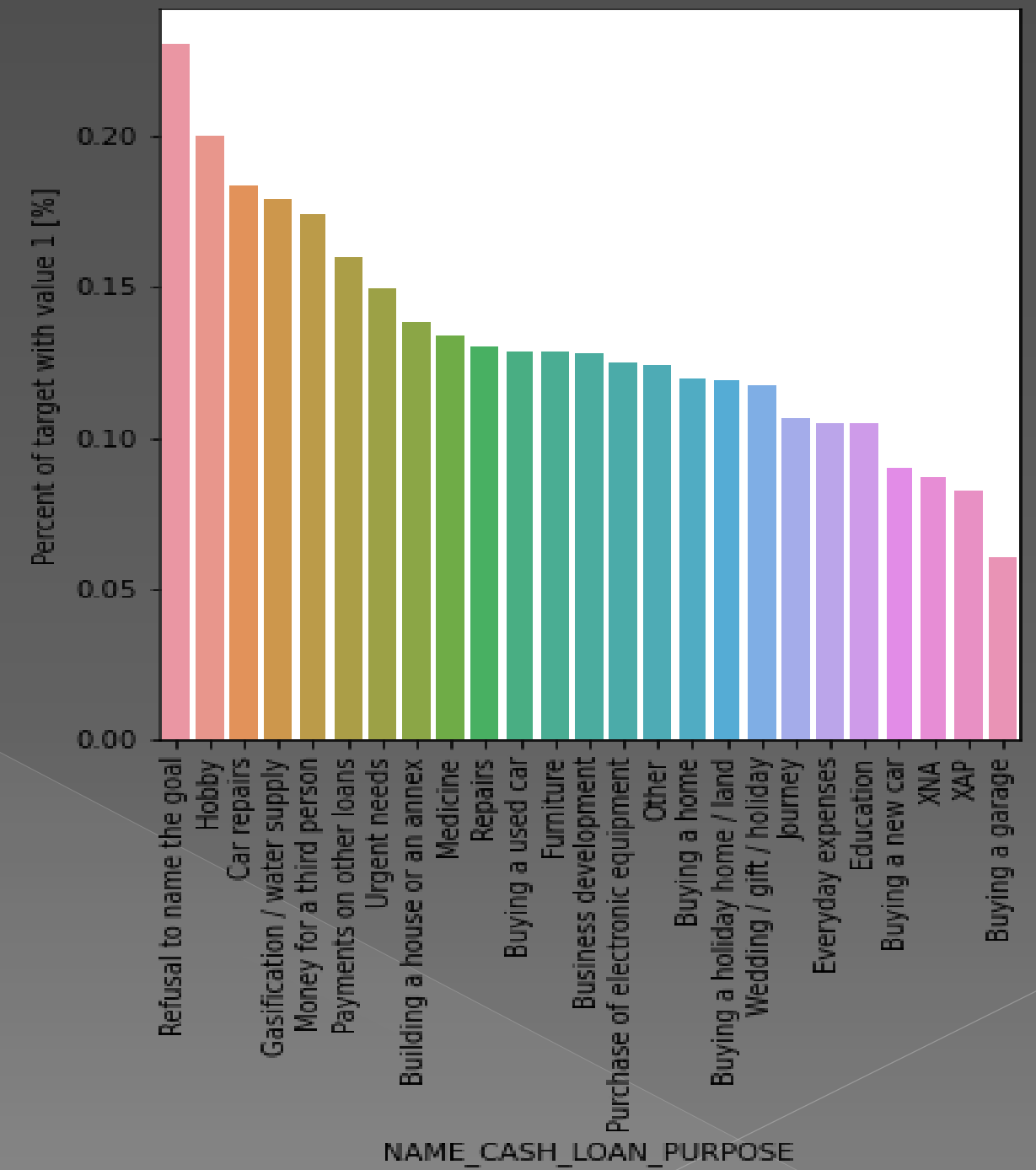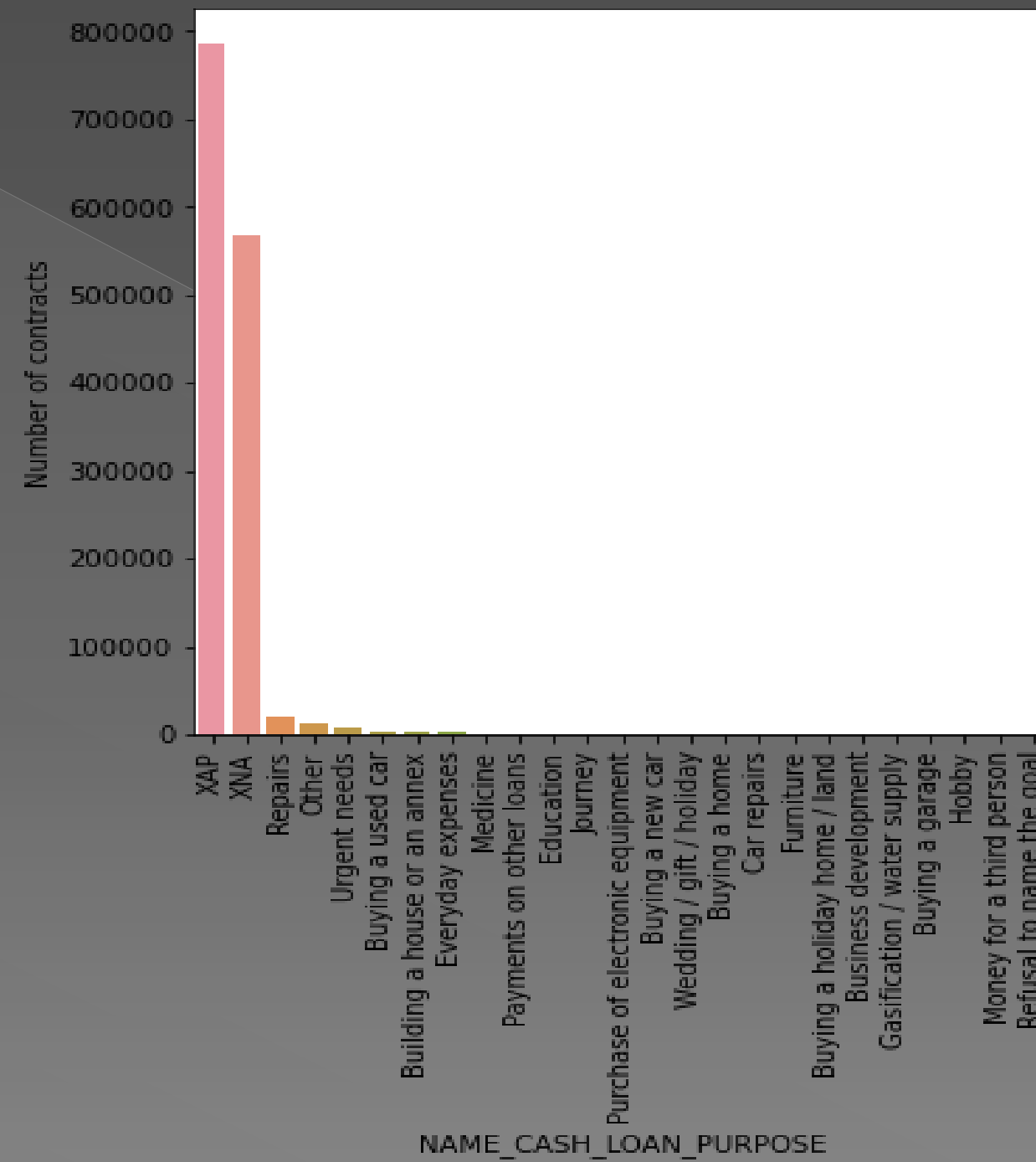
Contract Status

- As it has to be , the percentage of refused is more for the category 1 but still the people who didn't pay for after the approval is also more than 60% .
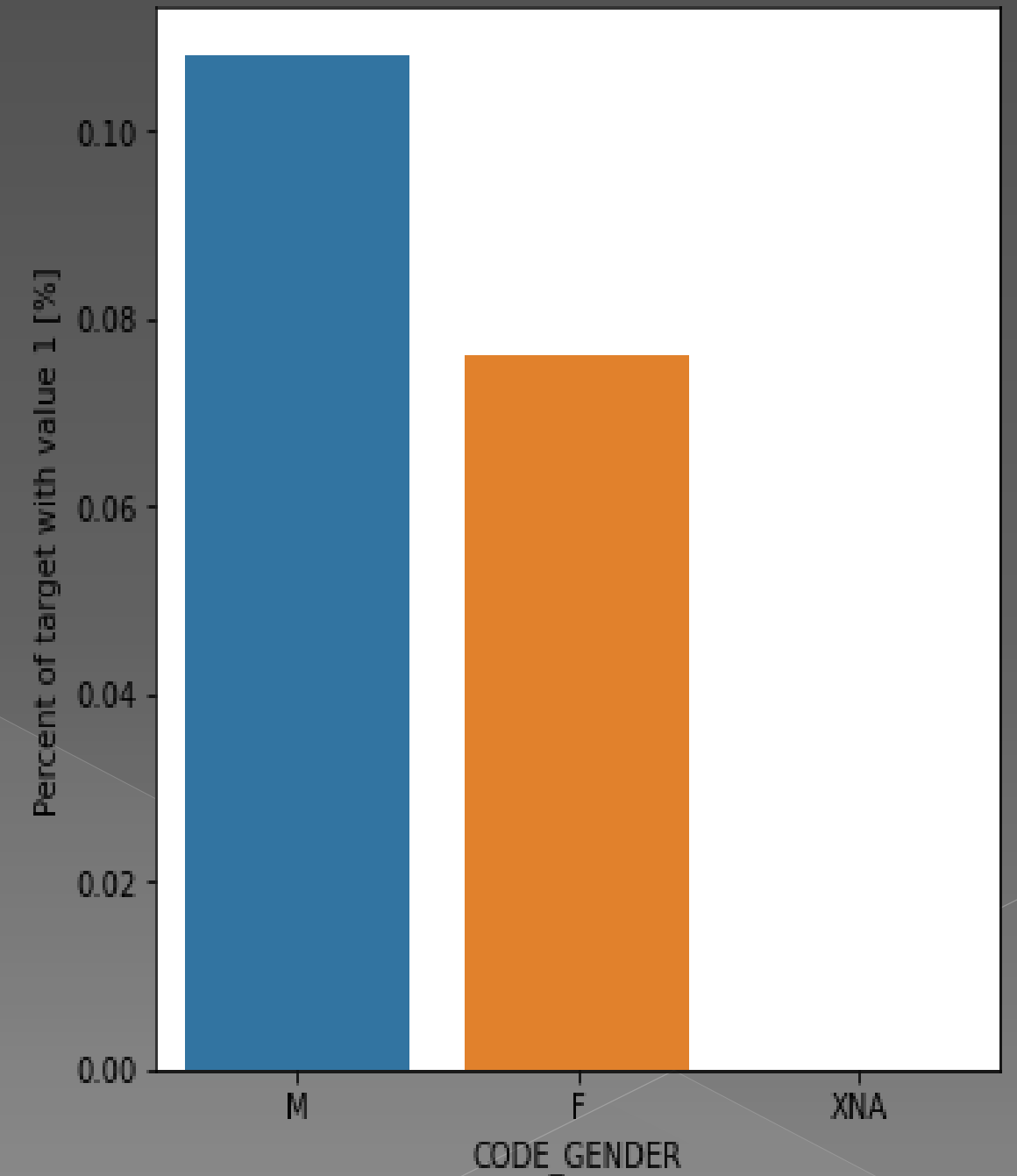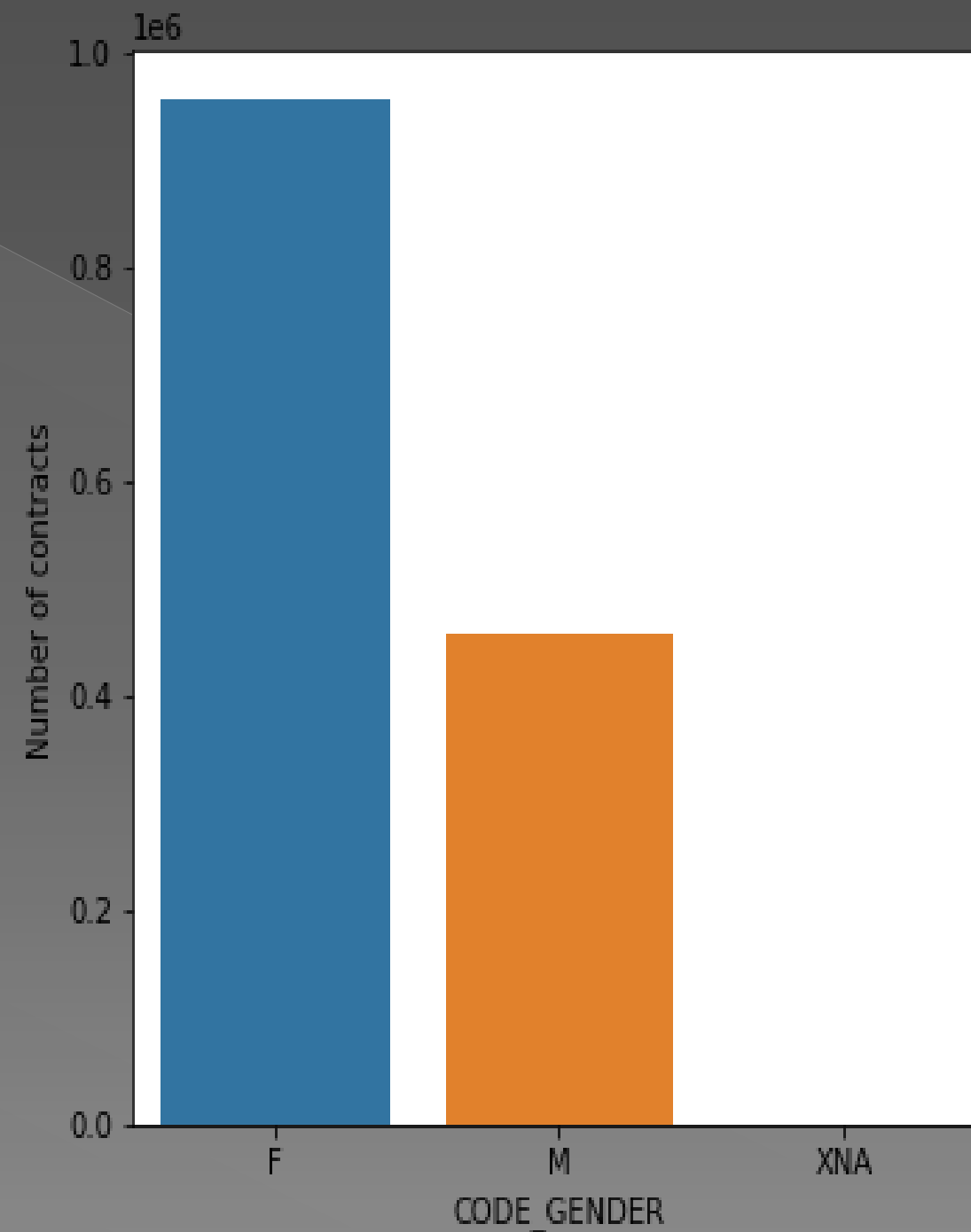
Type of Client

- We can see that repeated customers are also not repaying the loans
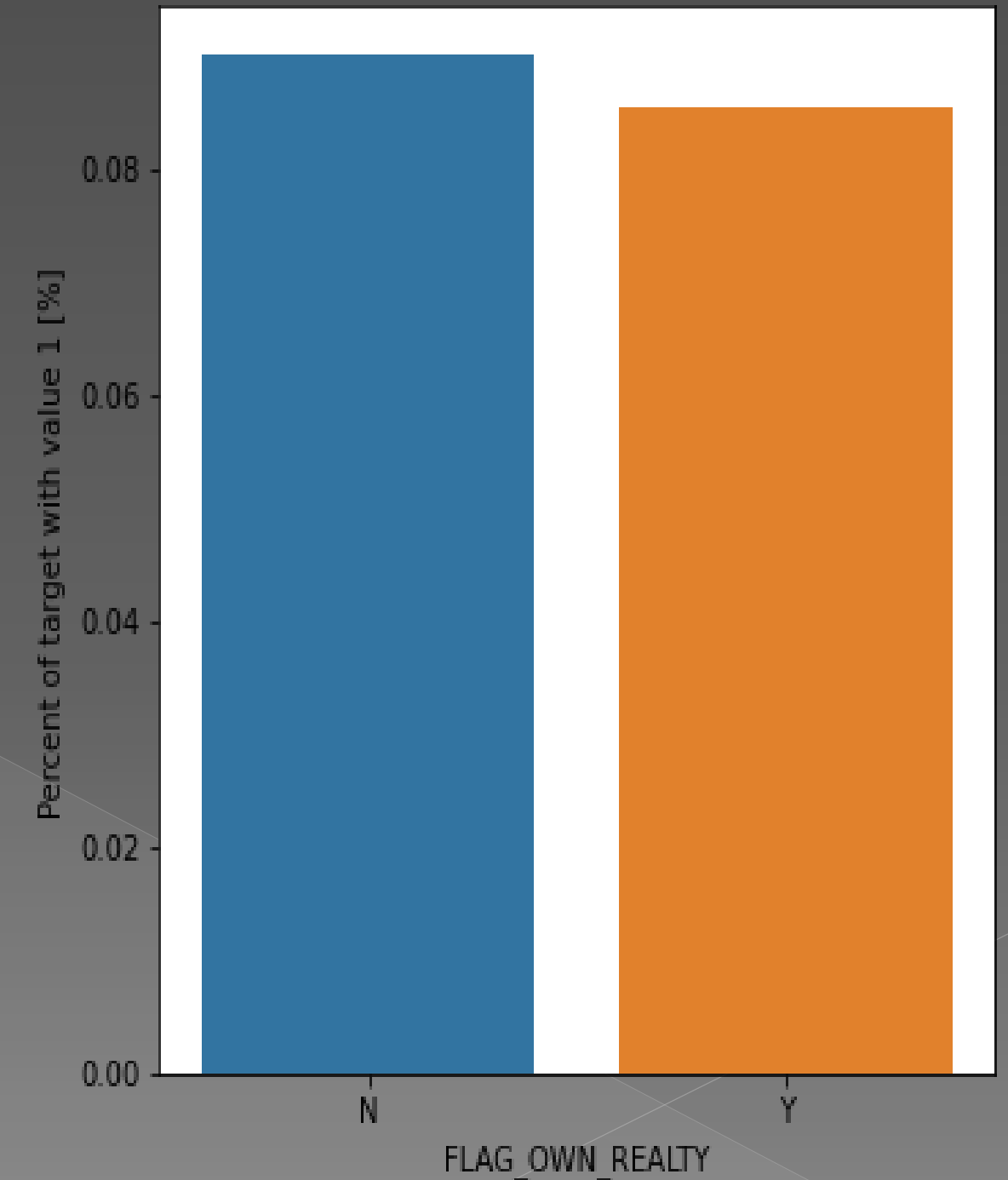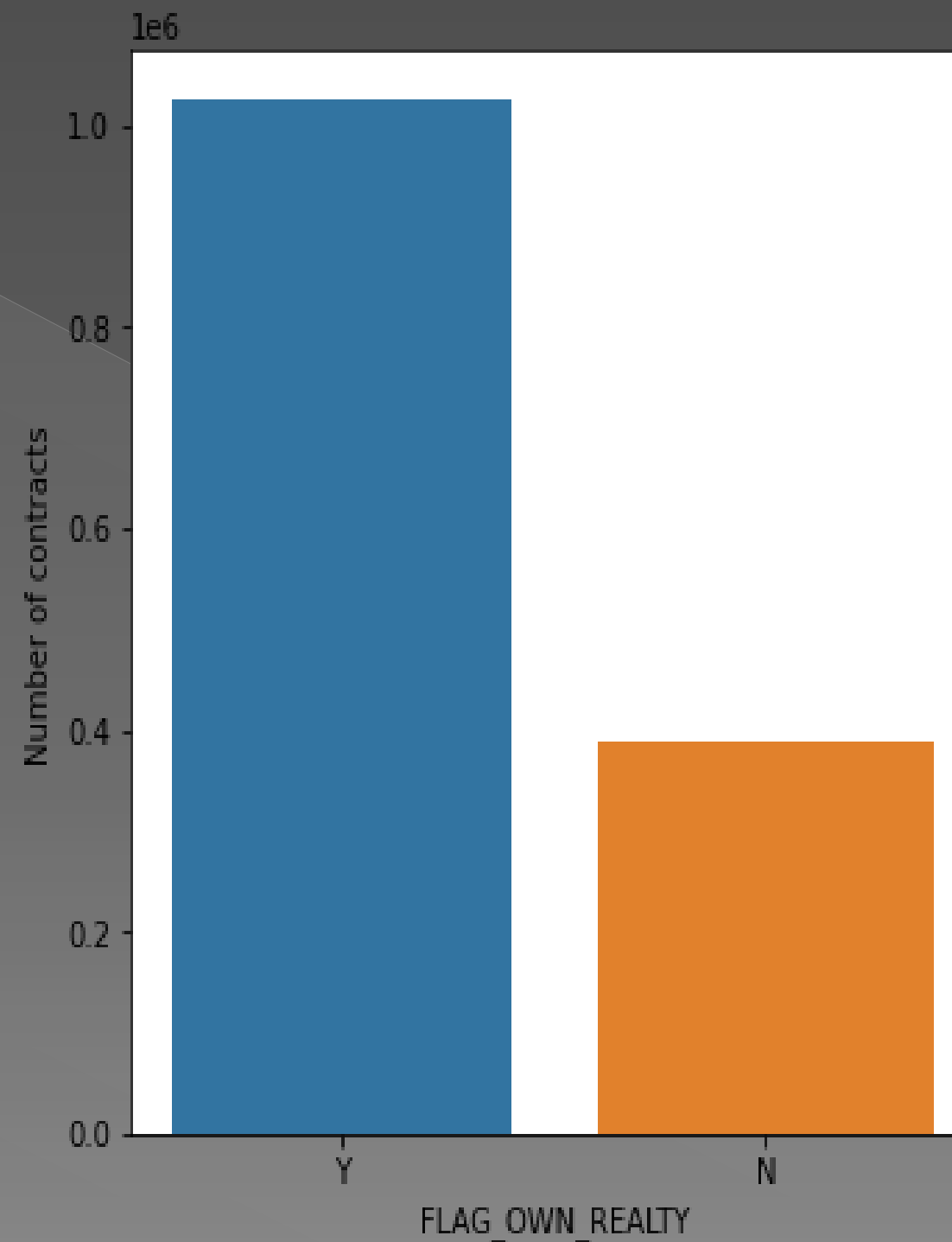
Gender

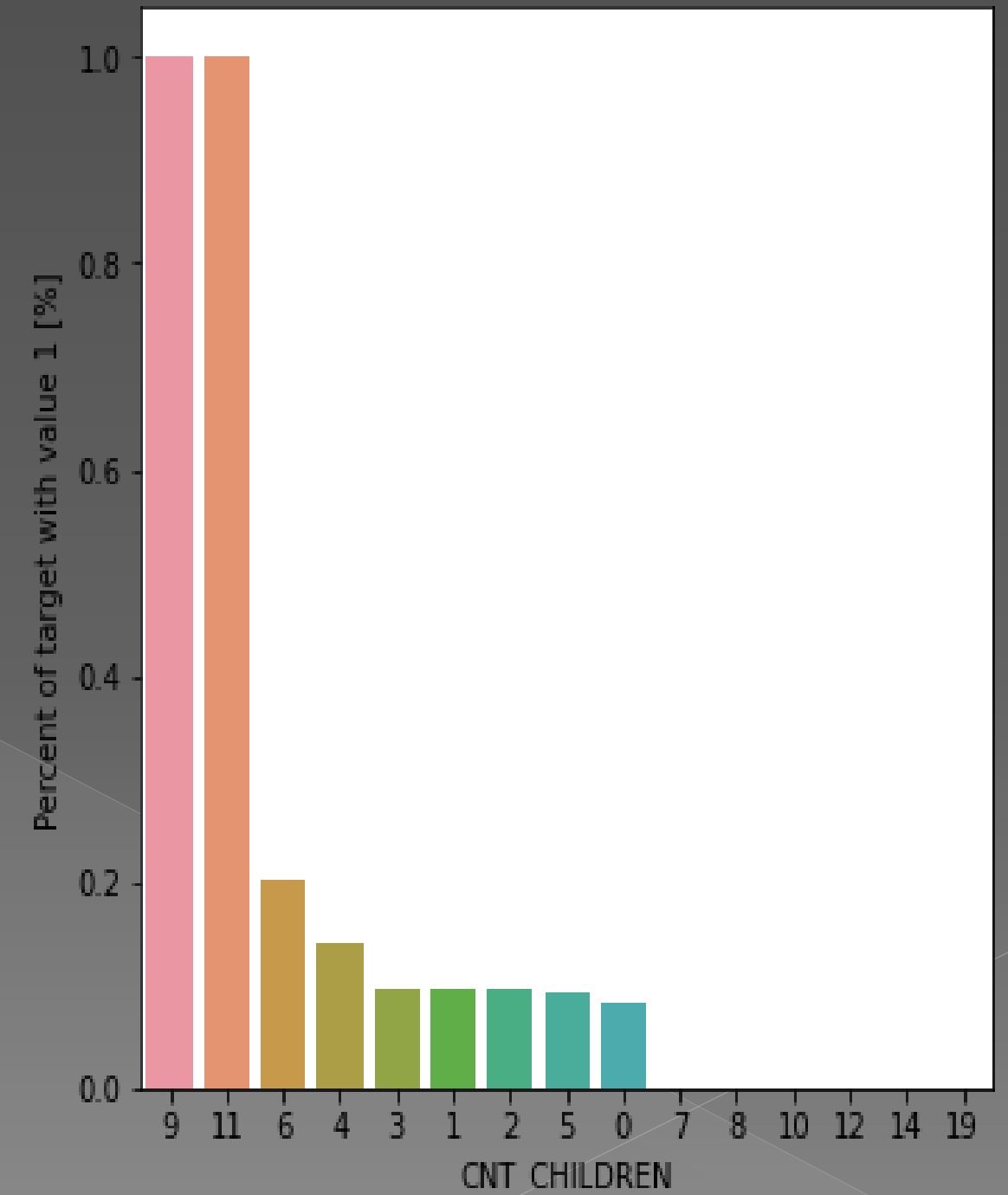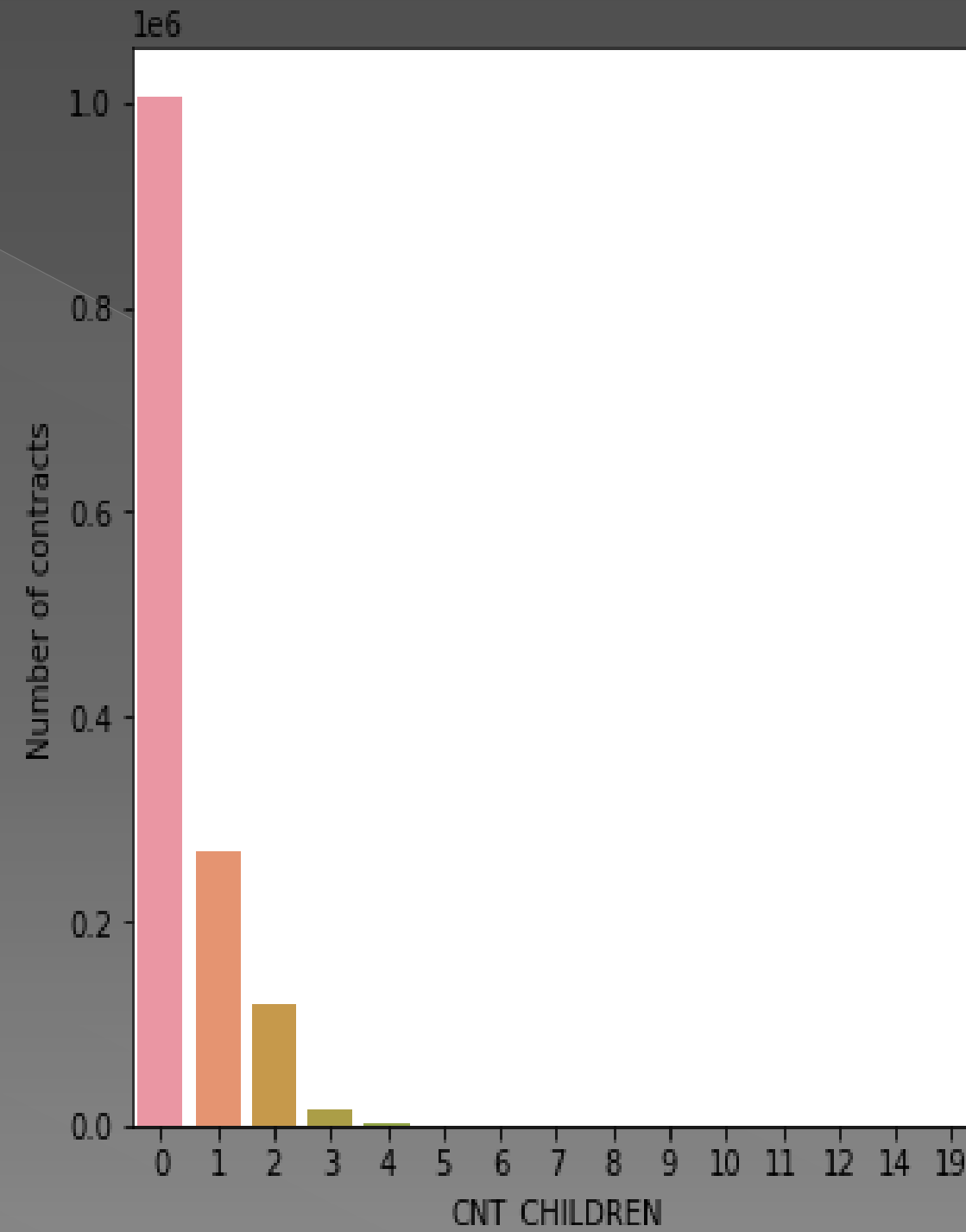- The percentage of defaulters are more in men than women.

FLAG_OWN_REALTY

- Both the cases above though there is own reality and own car , the % of non payment are equally the same so this cannot be a factor to be cosidered .

CNT_CHILDREN

- We can see that irrespective of 1,2,3,or 5 children the % of non payment is close but there is a difference for people with 6 and 4 kids

Type of Income

- For income type 'working', 'commercial associate', and 'State Servant' the number of credits are higher than other i.e. 'Maternity leave.
- For this Females are having more number of credits than male.
- Less number of credits for income type 'Maternity leave'.
- For type 1: There is no income type for 'student', 'pensioner' and 'Businessman' which means they don't do any late payments.
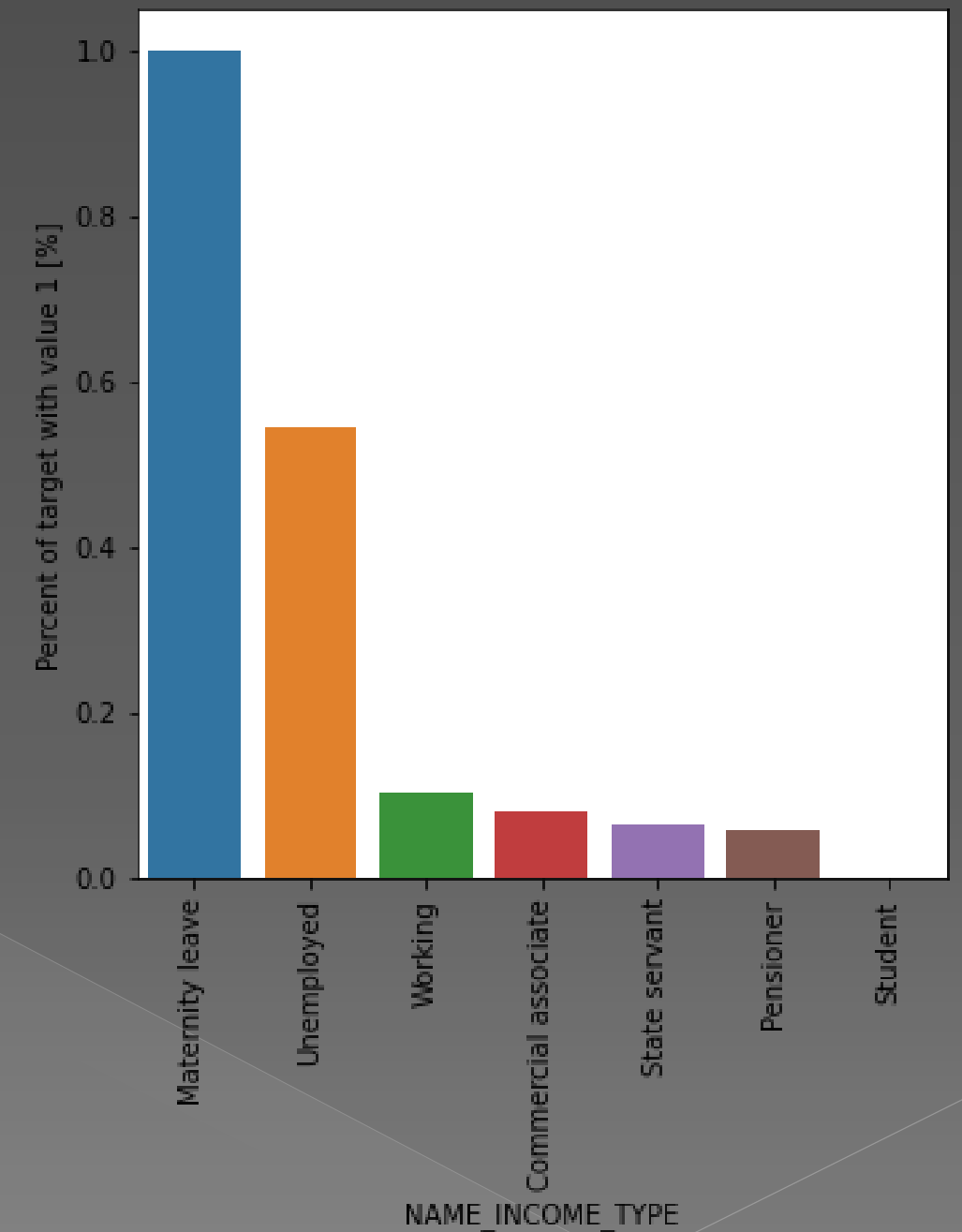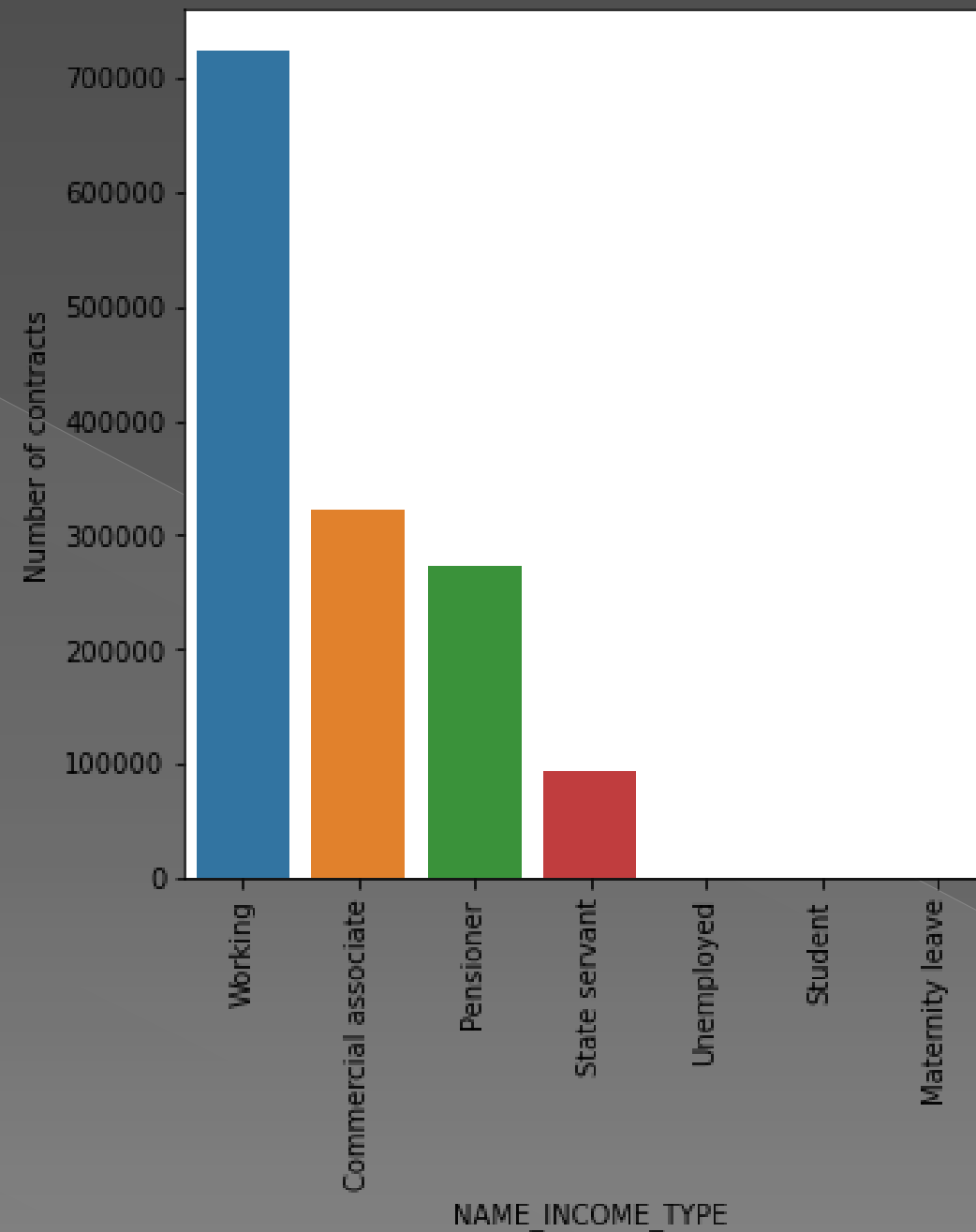
Type of Education

- Higher the education higher is the repayment rate
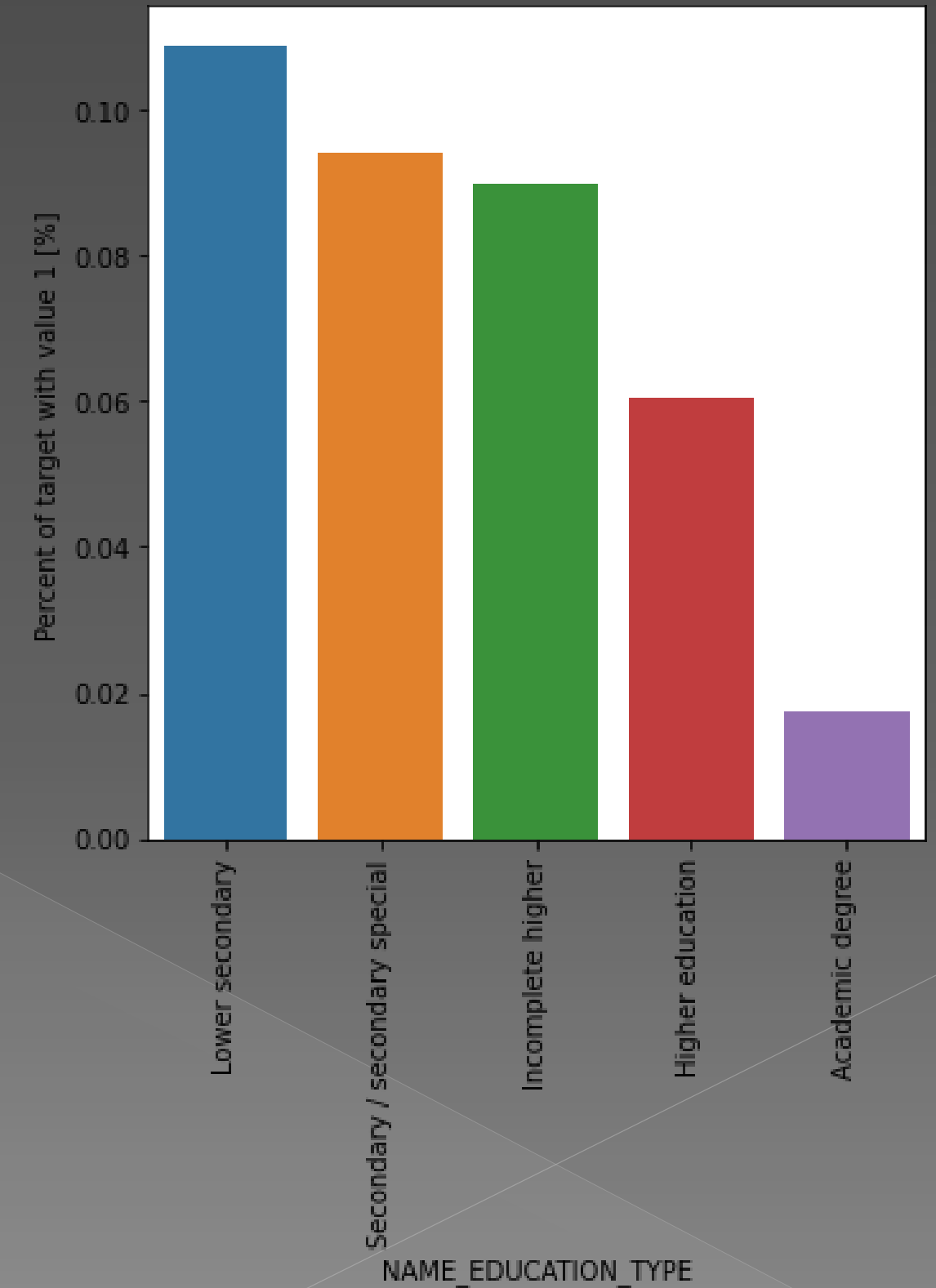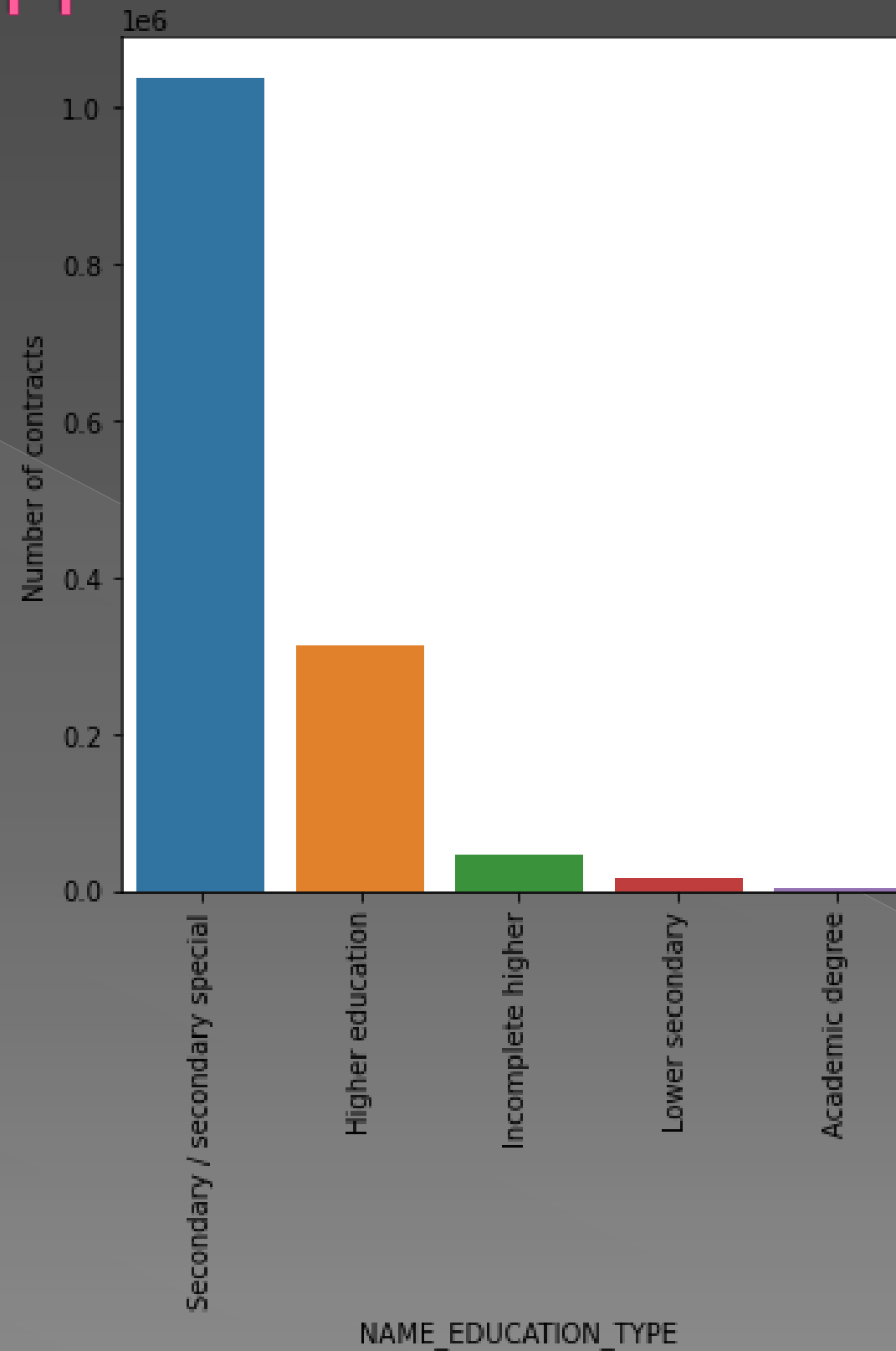
Family Status

- Civil Marriage and Single/not married family members have a higher percentage of non – repayment of loans.

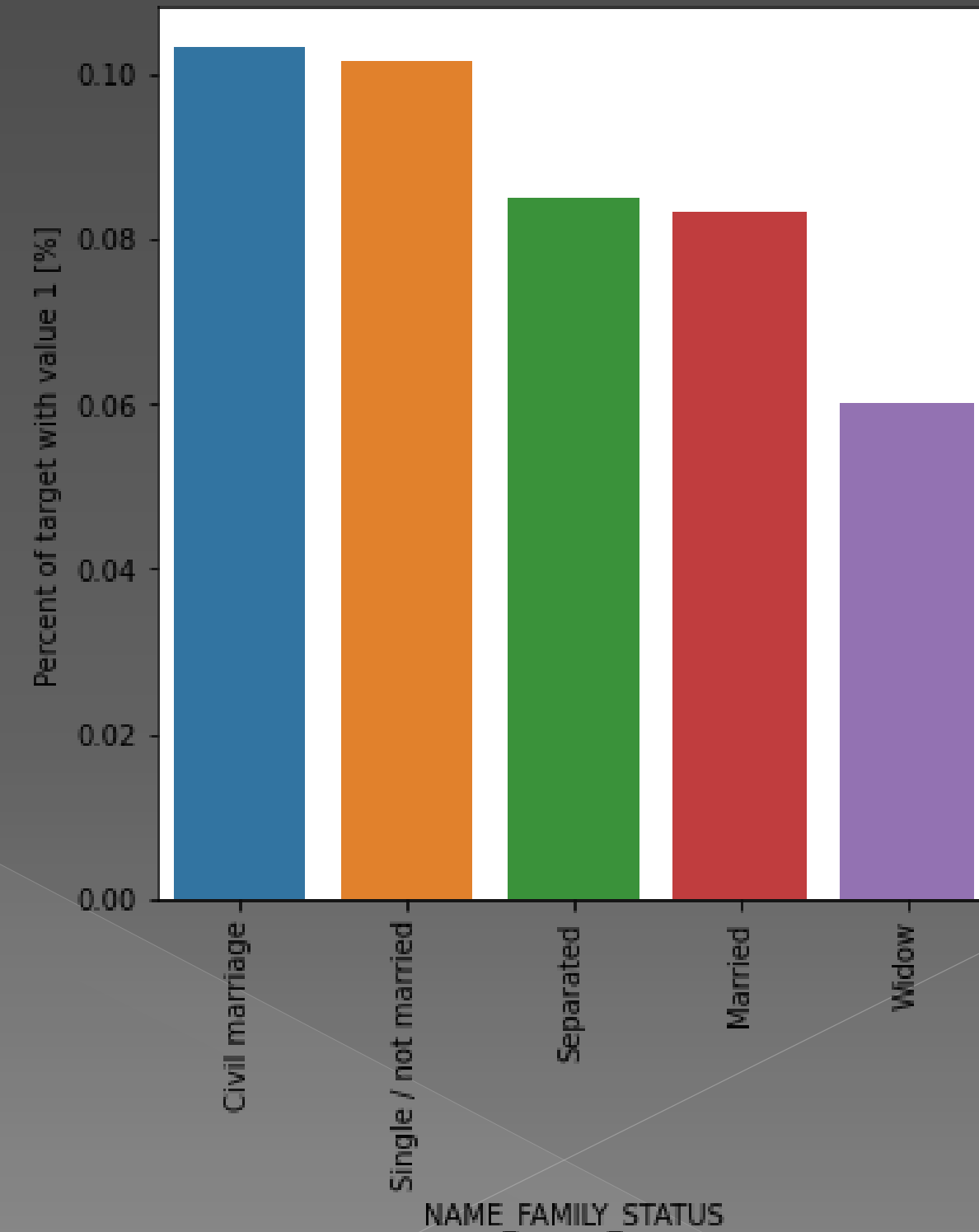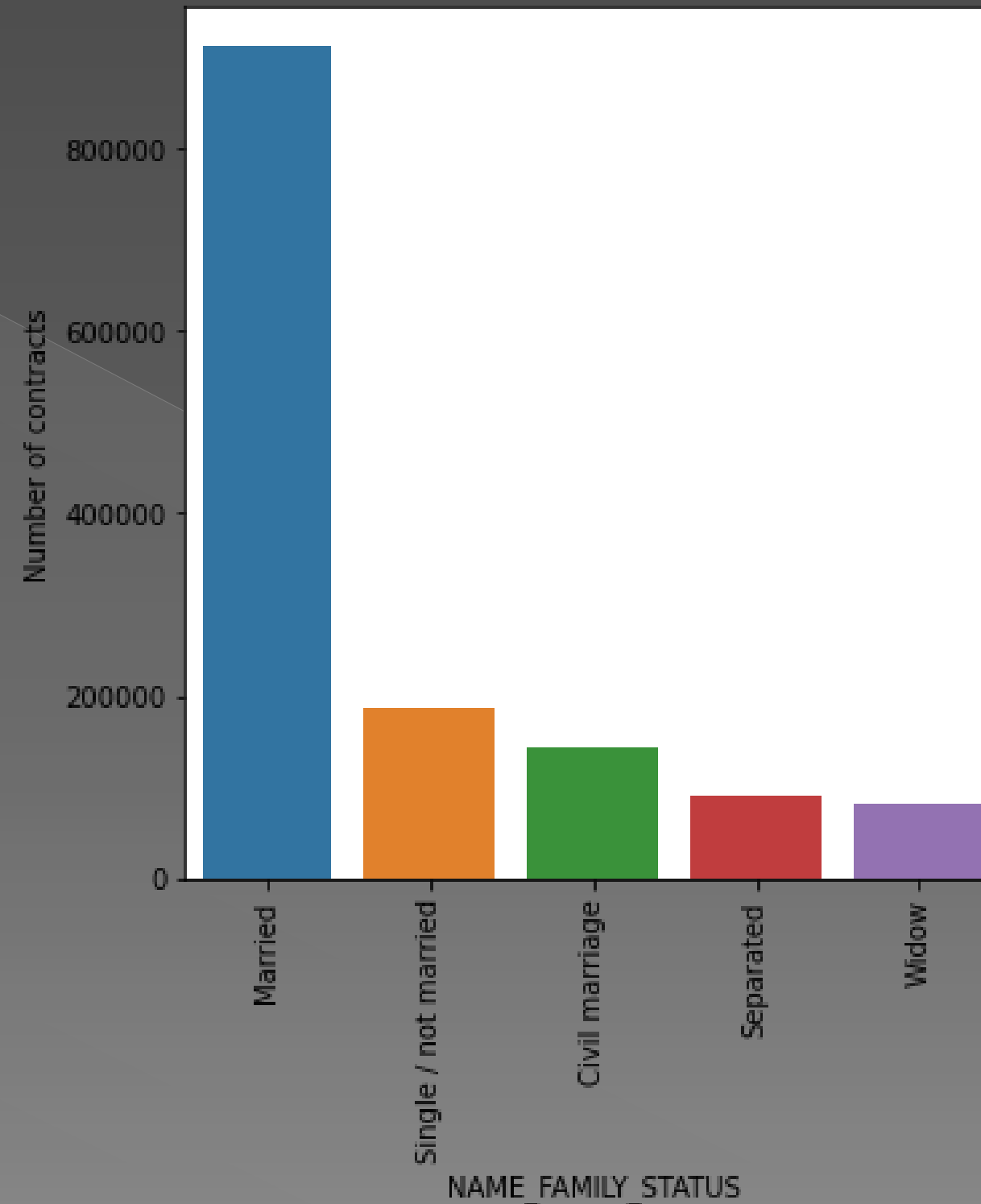- Married family members have a better repayment rate compared to single, separated and civil marriage members even though they contribute to the more than 60% of the total loan contracts.

Housing Type

- Bank can focus mostly on housing type with parents or House\ apartment or municipal apartment for successful payments.

- Bank should avoid giving loans to the housing type of co-op apartment as they are having difficulties in payment.
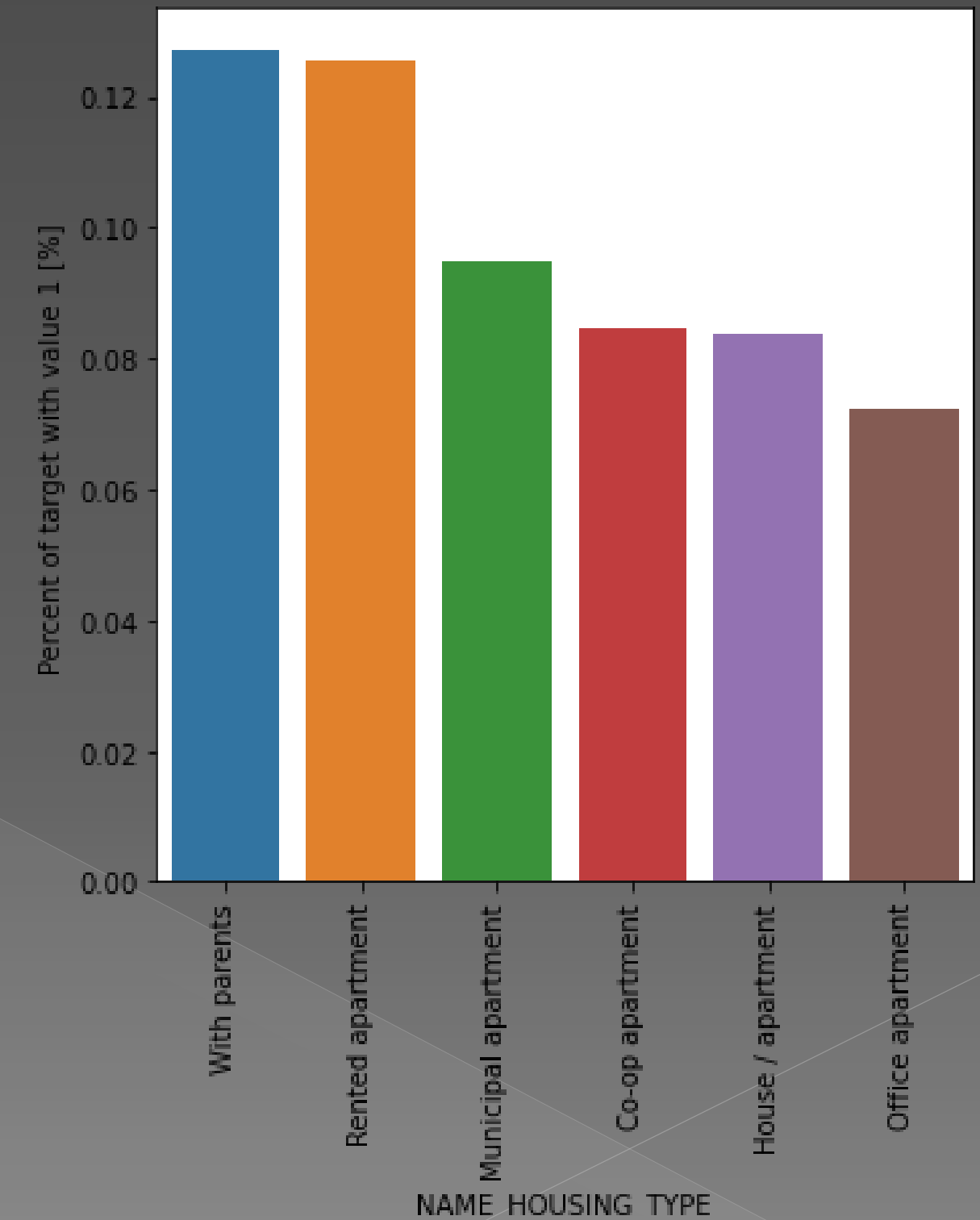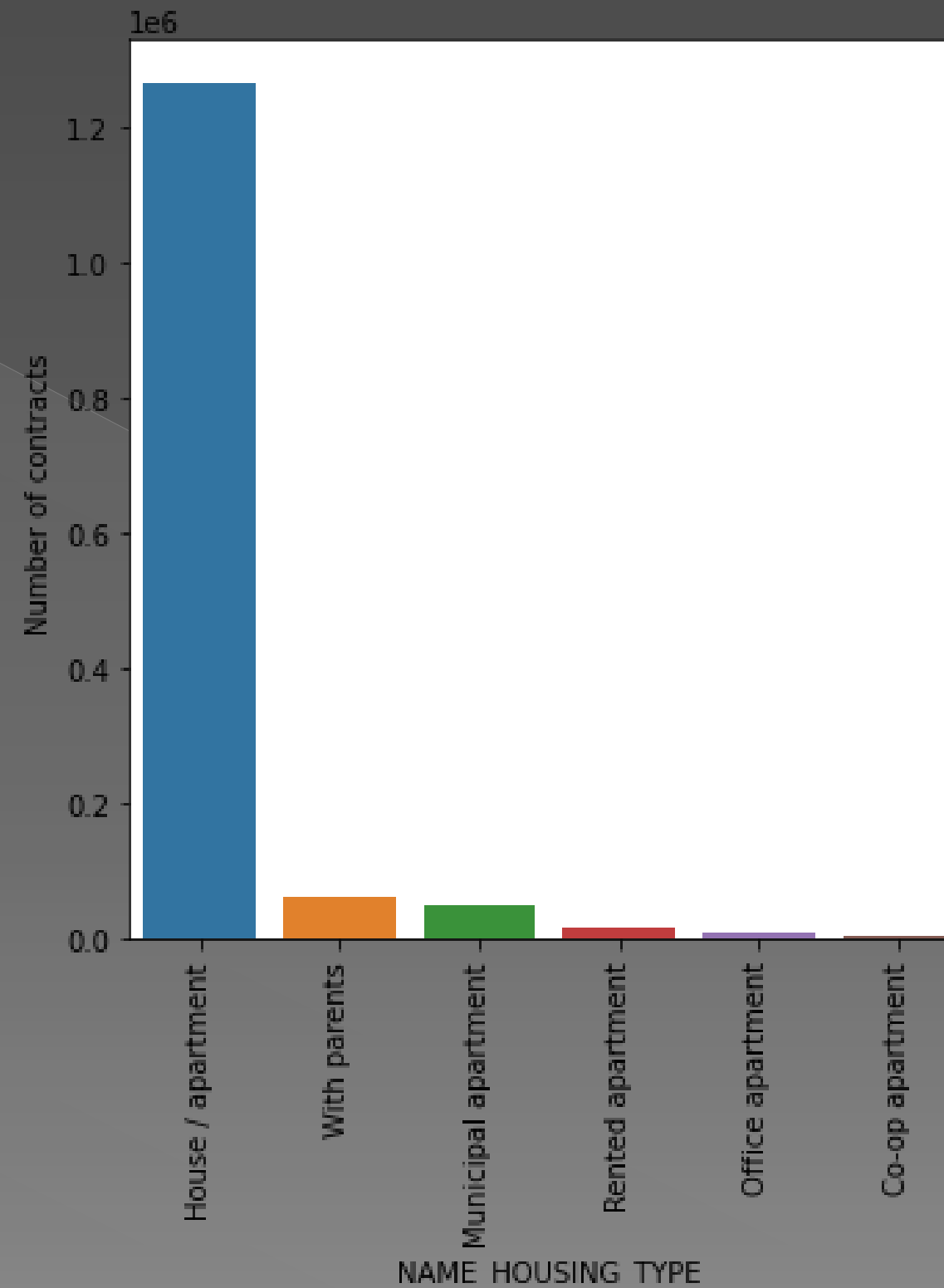
Count of Family Members

- Families with less than or equal to 3 people make over 60% of the overall contracts.
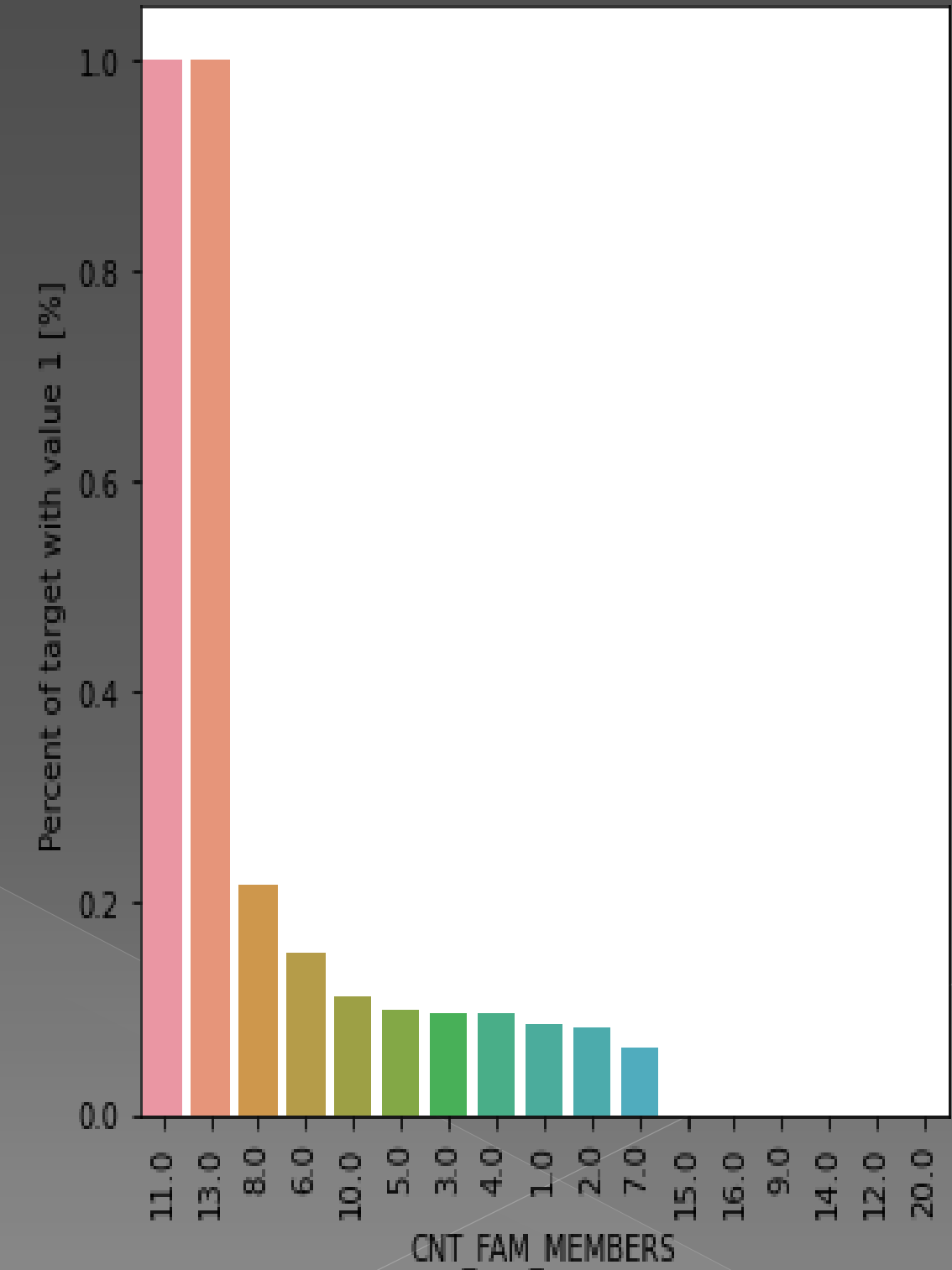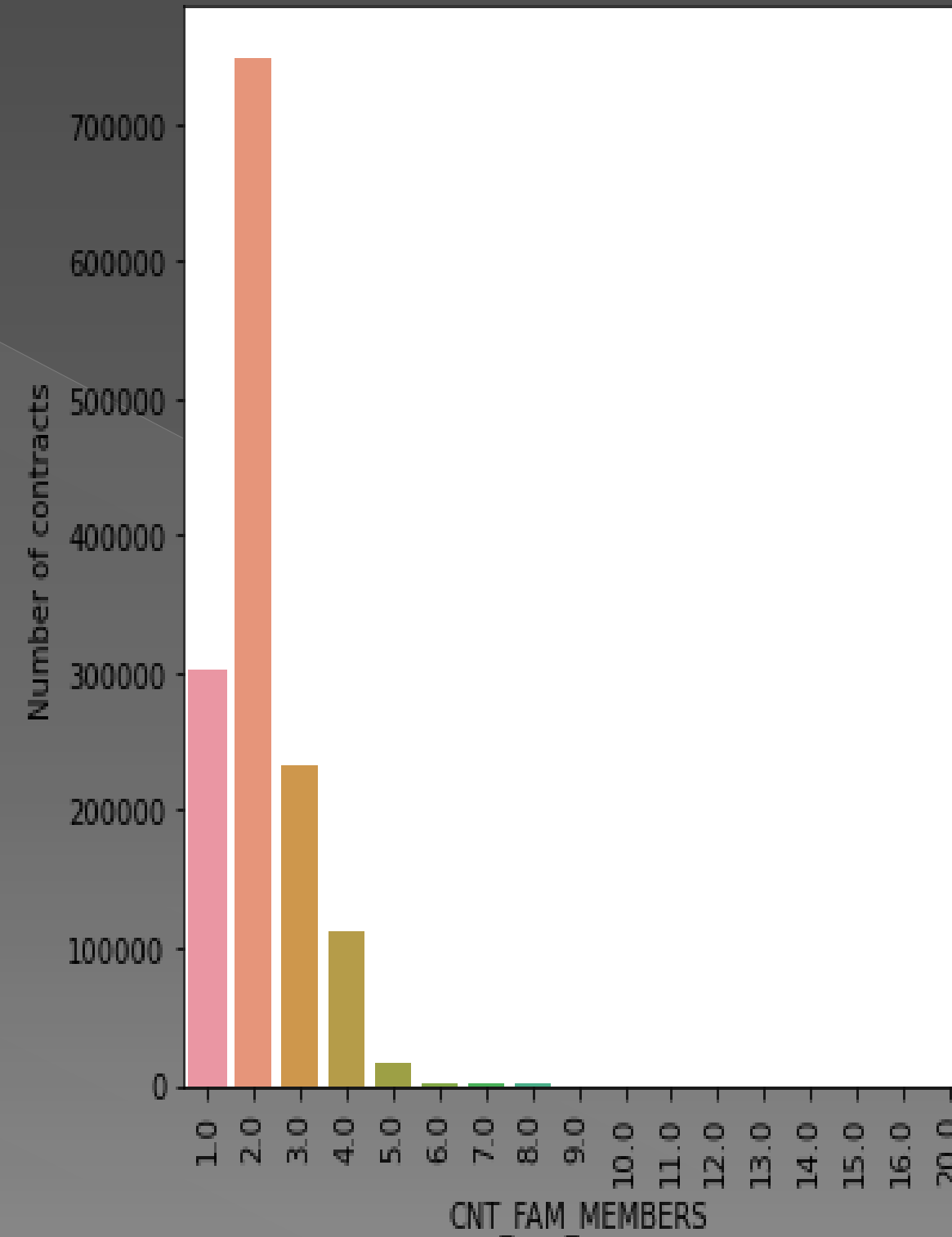
Loans Repaid

- Looking at the above graph we can see that the data is highly imbalanced .

- Class 0 = Loans repaid Class 1 = Loans not repaid

Heatmap

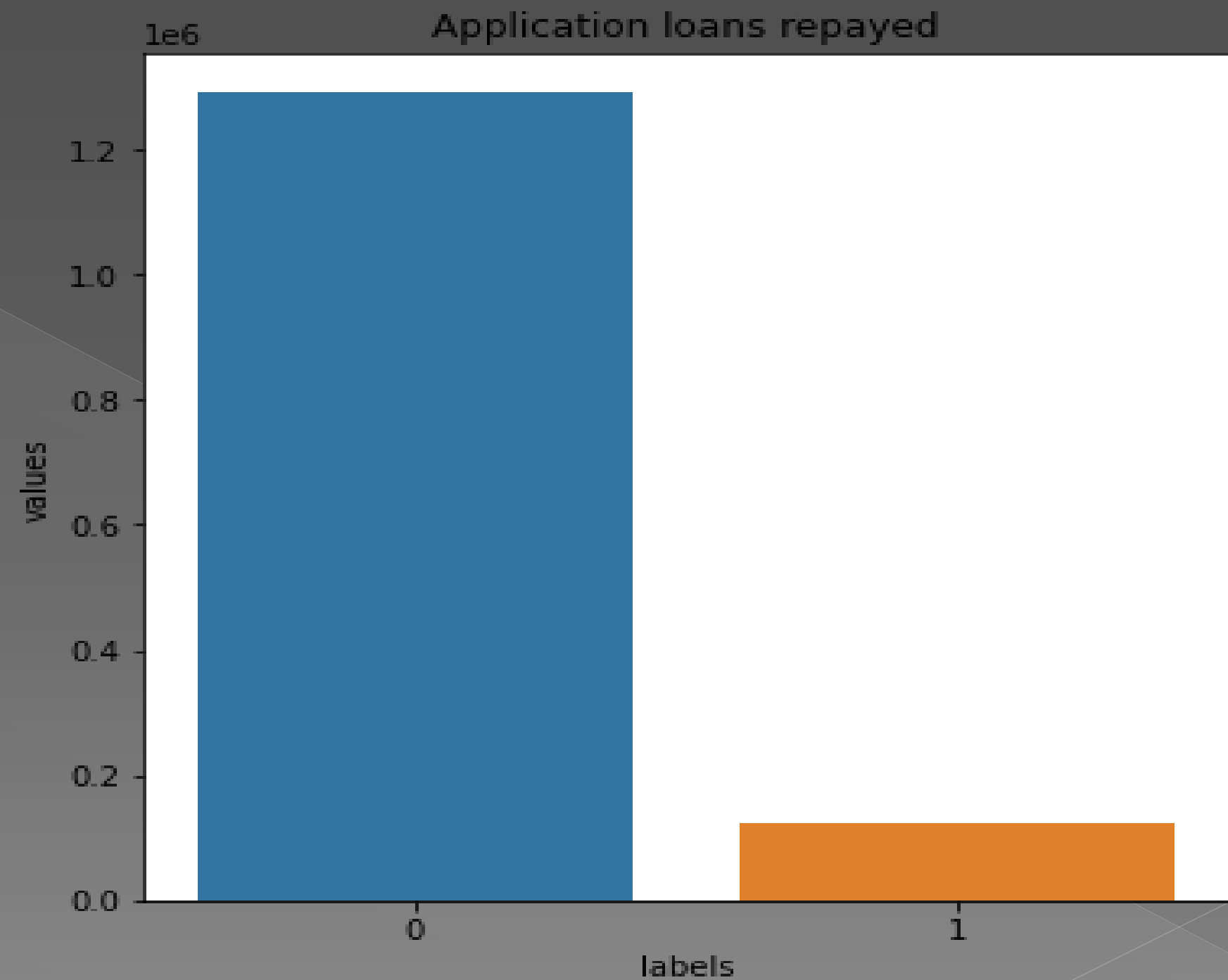|	CNT_CHILDREN	AMT_INCOME_TOTAL	DAYS_EMPLOYED	CNT_FAM_MEMBERS
CNT_CHILDREN	1.000000	0.011661	-0.247960	0.879224
AMT_INCOME_TOTAL	0.011661	1.000000	-0.067488	0.014119
DAYS_EMPLOYED	-0.247960	-0.067488	1.000000	-0.249285
CNT_FAM_MEMBERS	0.879224	0.014119	-0.249285	1.000000

Conclusion

- Banks should focus more on contract type 'Student' ,'pensioner' and 'Businessman' with housing 'type other than 'Co-op apartment' for successful payments.

- Banks should focus less on income type 'Working' as they are having most number of unsuccessful payments.

- Also with loan purpose 'Repair' is having higher number of unsuccessful payments on time.

- Get as much as clients from housing type 'With parents' as they are having least number of unsuccessful payments.

Thank You