

# CREDIT EDA CASE STUDY

By Nazibur Rahaman

# Problem Statement and Goal

The aim of this case study is to identify client behaviour-

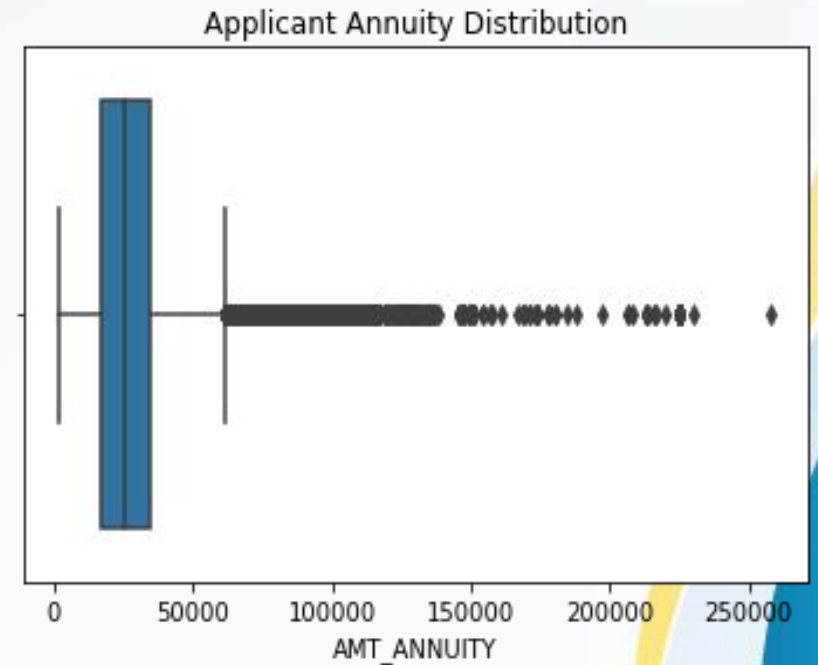
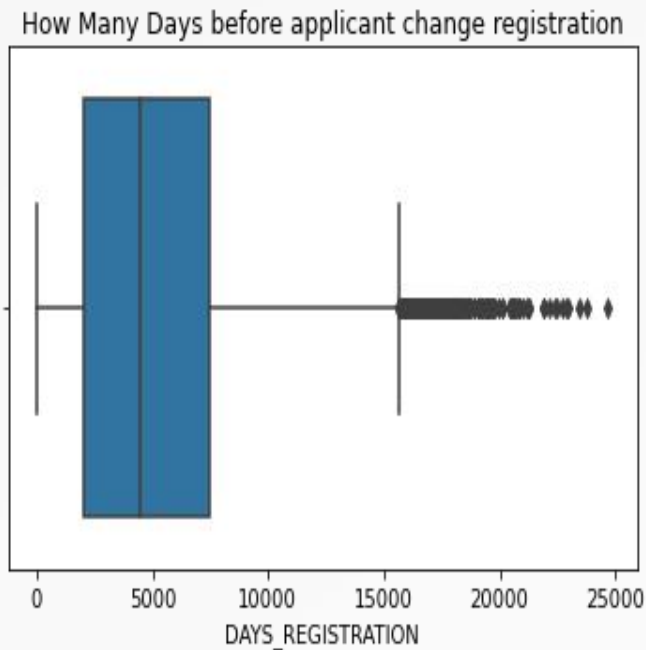
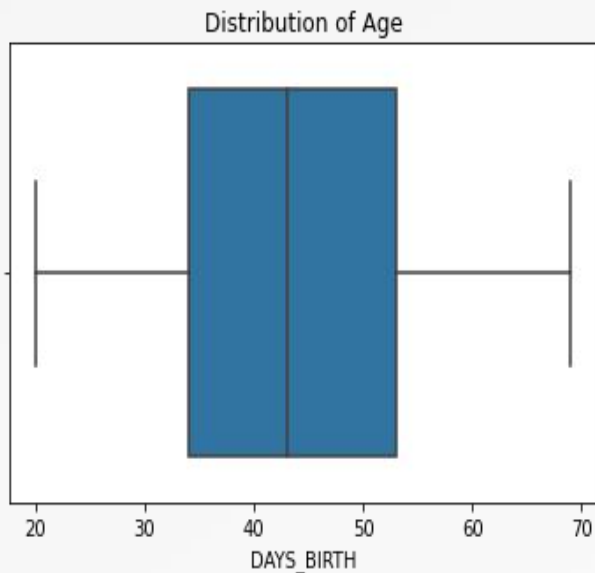
- 1. Who have problem with payment/installment, which will cause business loss.
- 2. And who is capable of repayment on time, should not miss this type client , as they will help to grow the business.

# Analysis Approach and procedure

- Import the data set, do check data quality, missing values, data type and data dictionary for better data understanding.
- Calculate the missing values percentage, i have drop the columns with more than 45% missing values we can impute the missing values but it can be biased.
- Less than 13% missing values , I have imputed with Best Metric, like median, mean and mode.

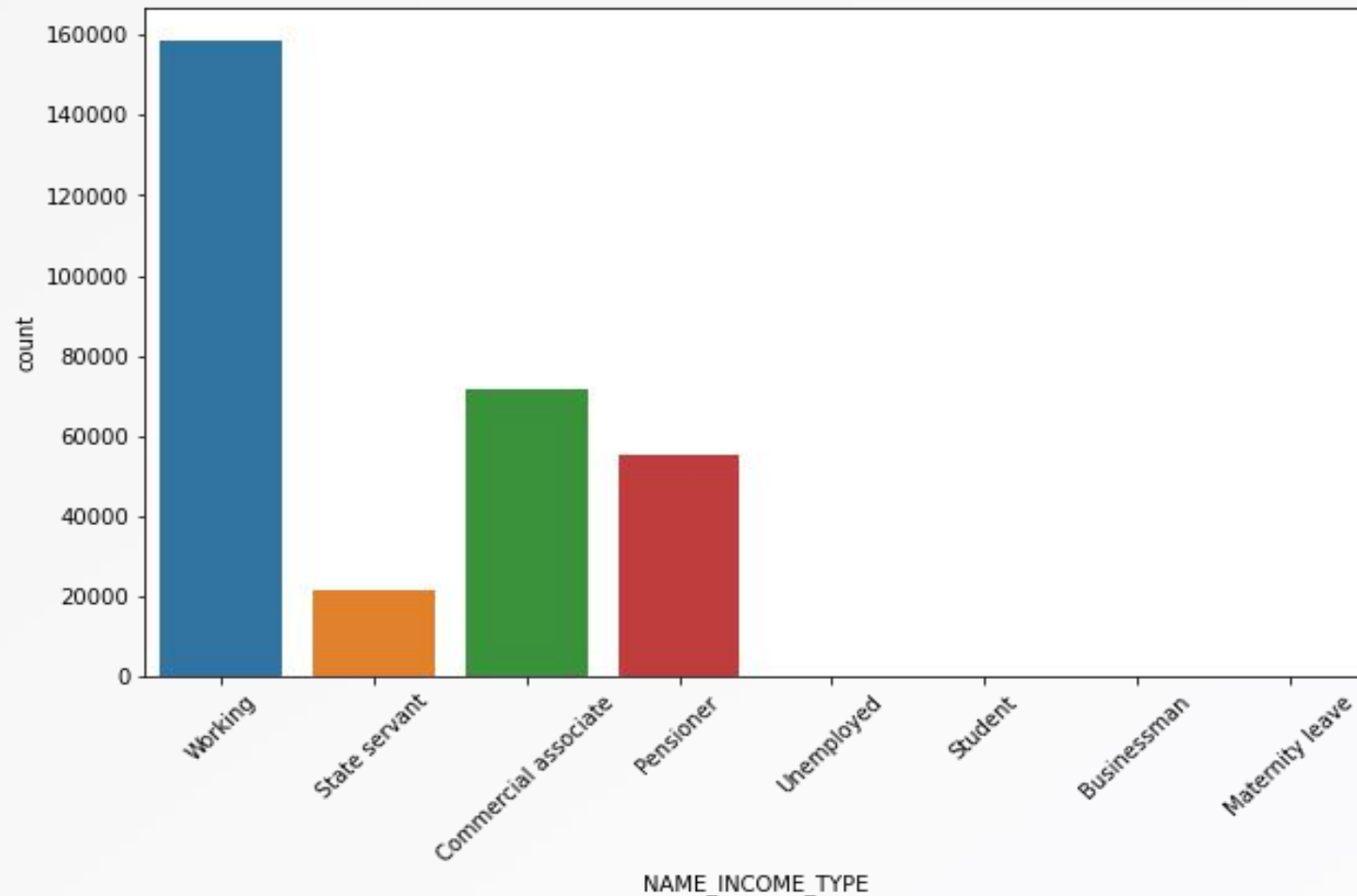
# Outliers checking for various columns-

1. Credit amount, Amount annuity, Days of Registration have outliers.
2. Age or Days Birth dont have any outliers.

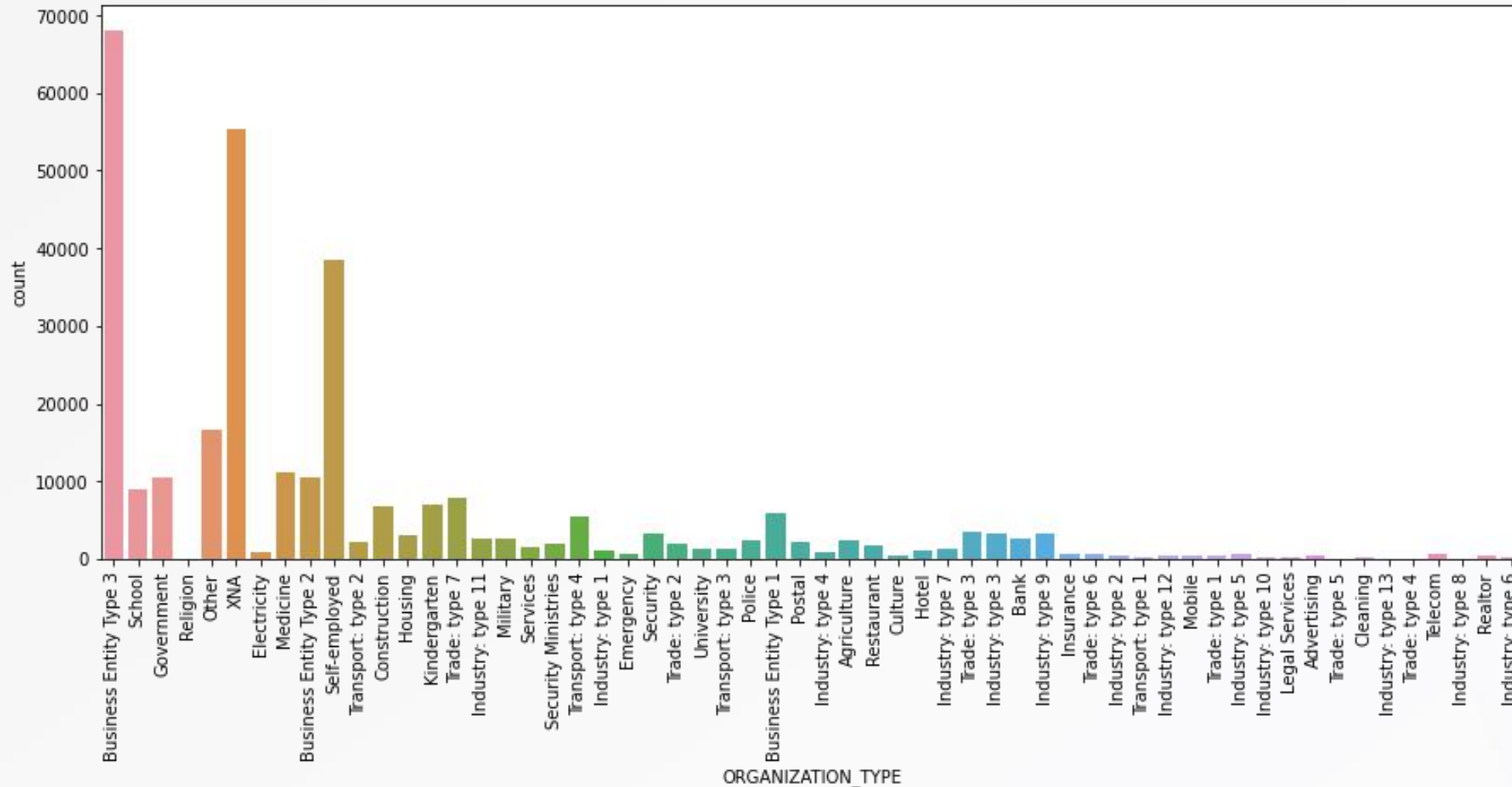


## Univariate Analysis-

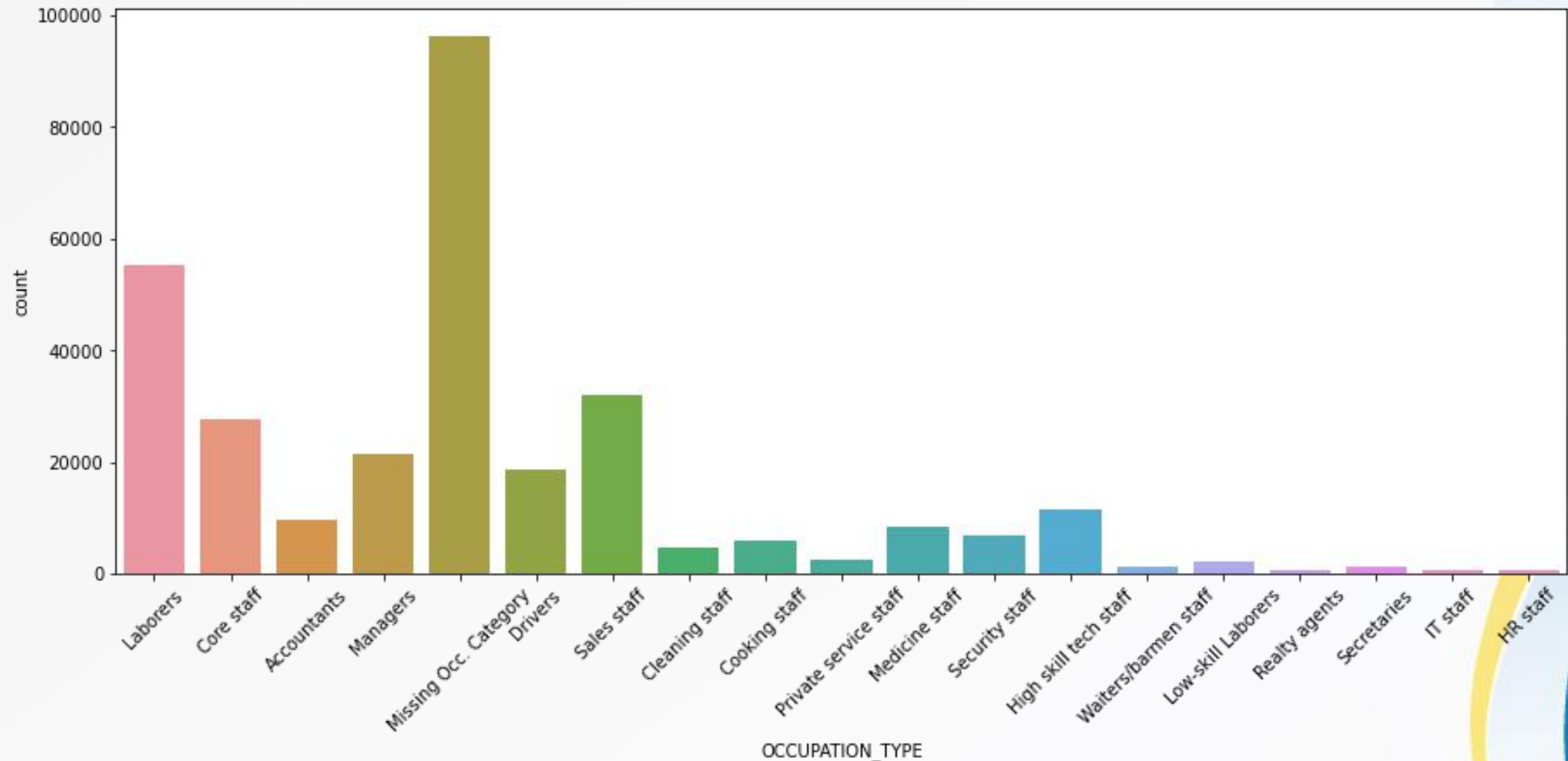
Around 51% are working, 23% commercial, 18% pensioner and 0.08% are state servant.



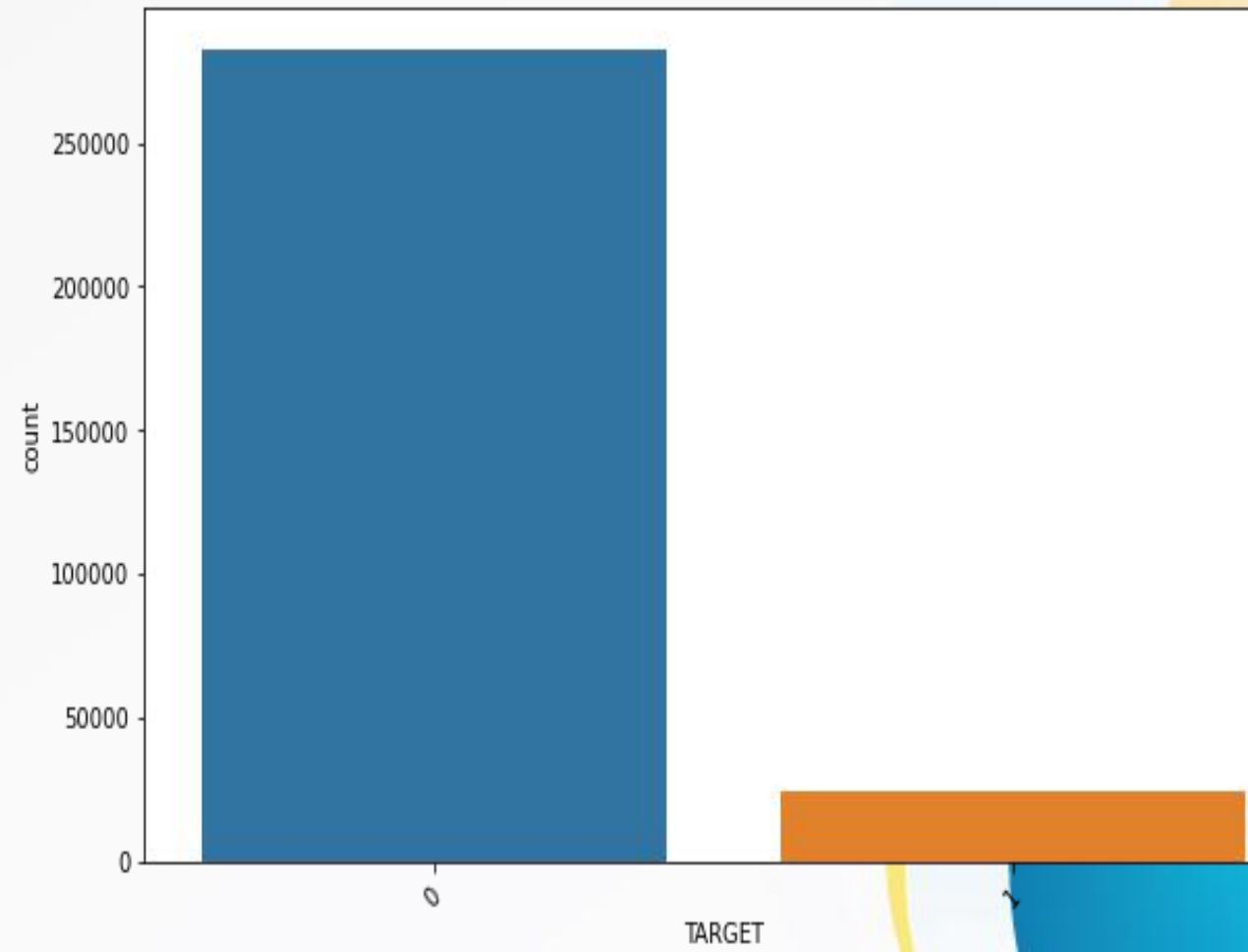
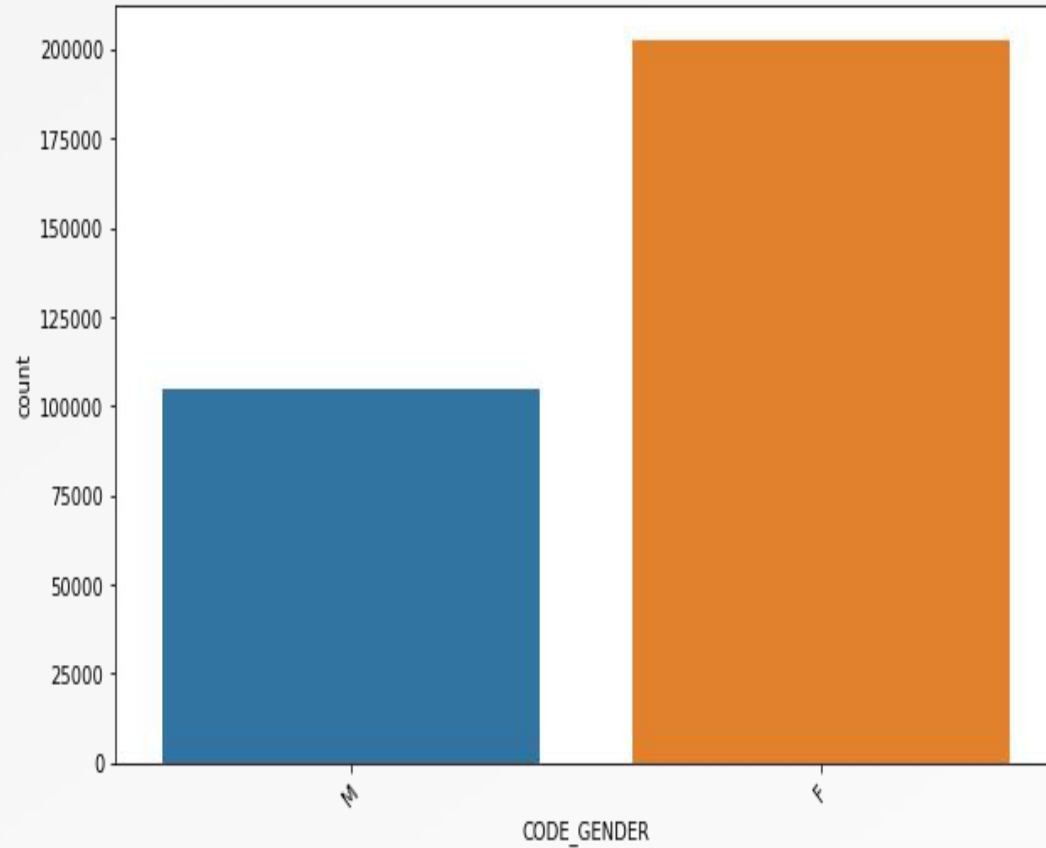
- 22% are from business entity type 3, 18% missing category and 12% self-employed so on.



- after 31% of missing category, around 17% are laborers, sales background 10% , core staff 8% and so on.

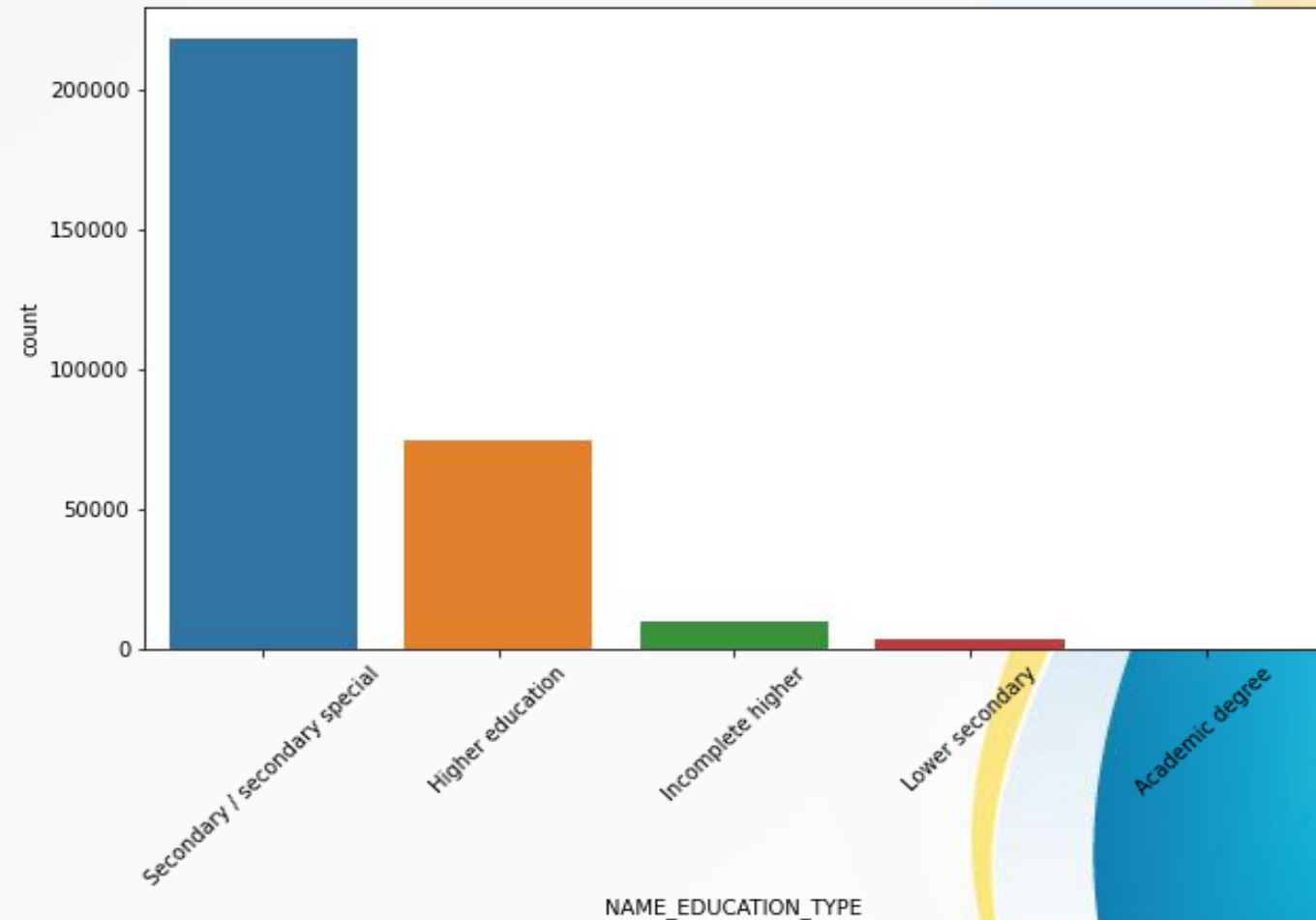
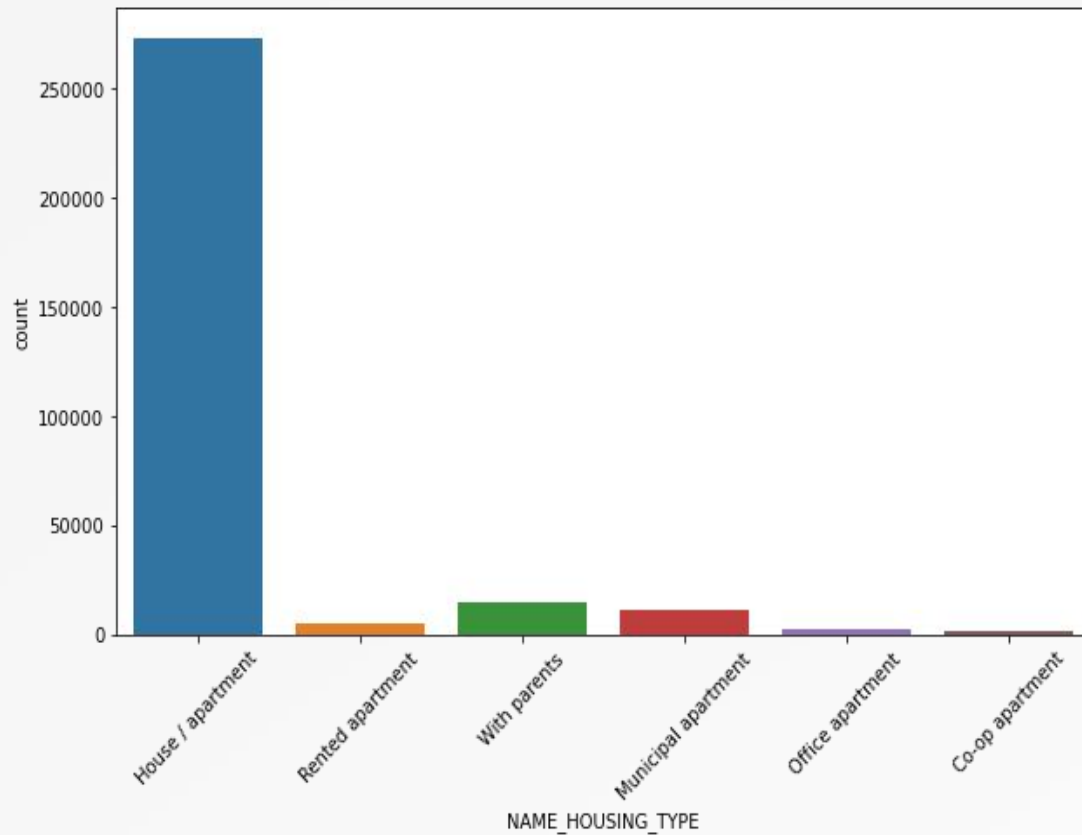


- 1. In Target column, 91.92 % have no payment difficulty , around 8% have payment problem
- 2. 65.83% are female and 34% are Male applicant.



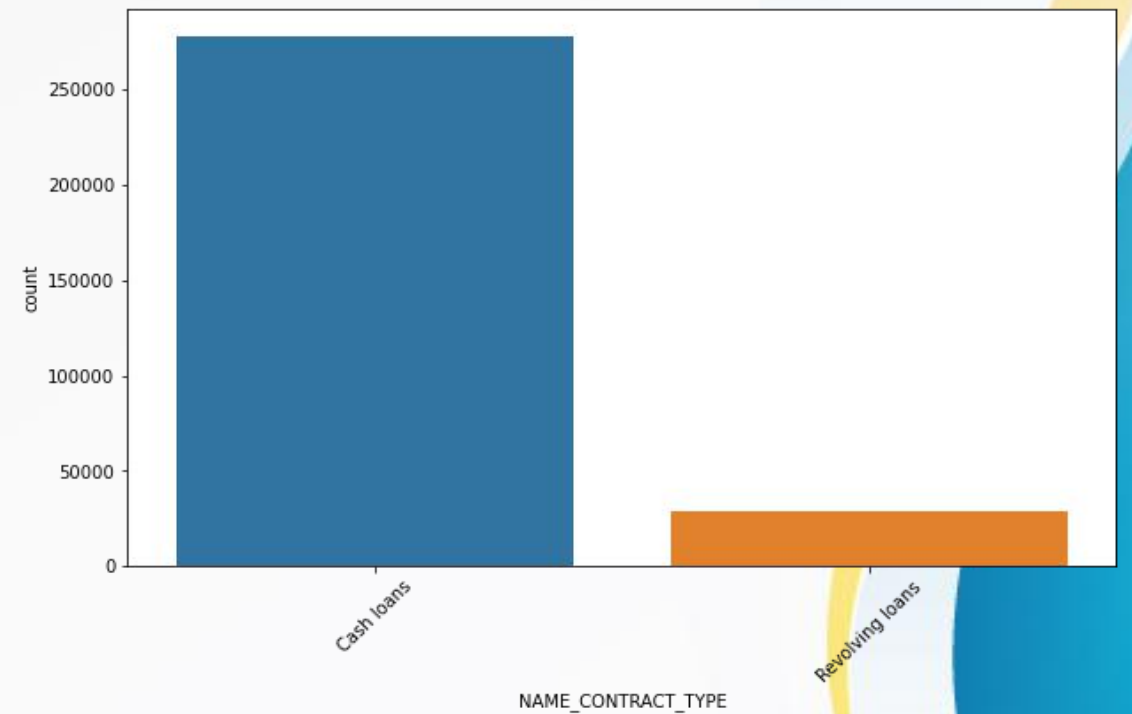
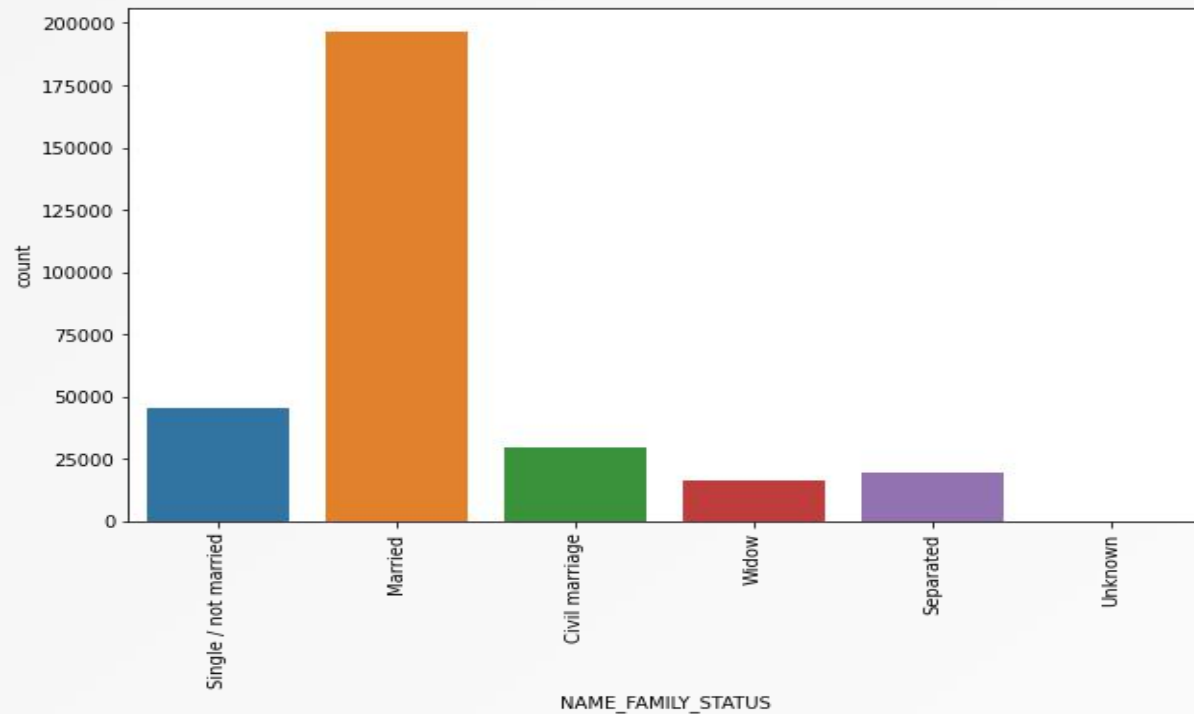


1.71% have secondary education and 24% have higher education.  
2. 88% have own apartment, 4% are with parents.



1.65% dont own any car, 34% are own car.

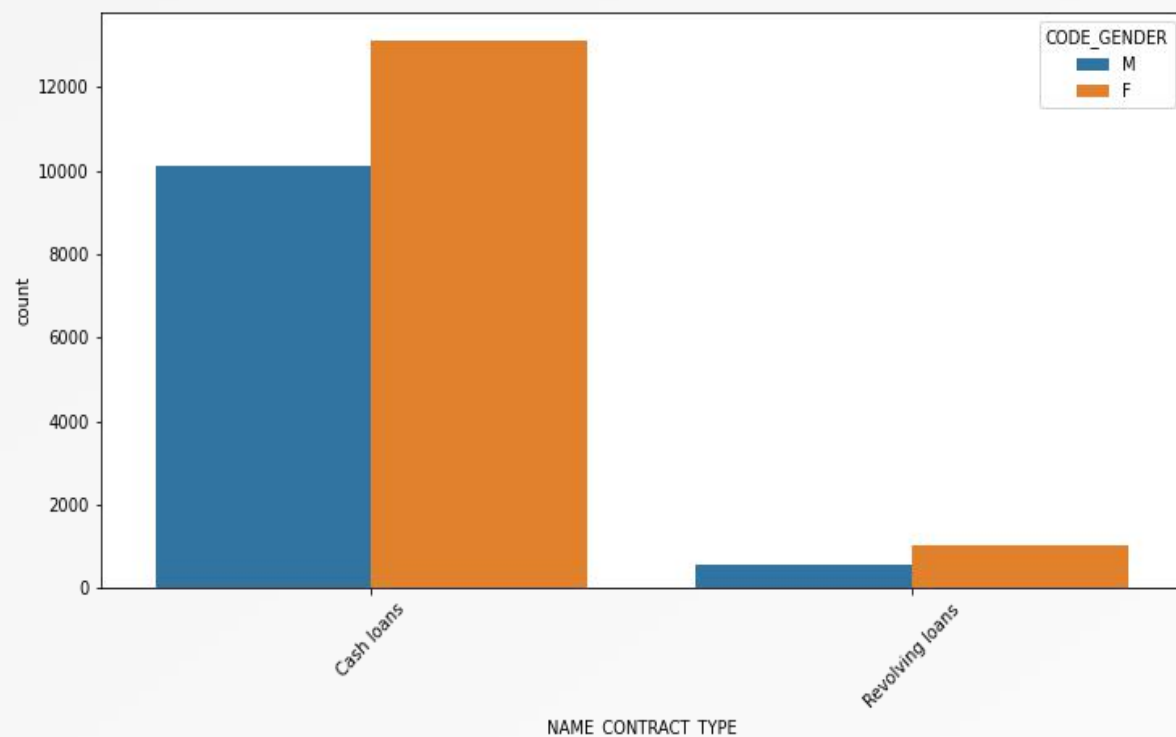
2. 81% are unaccompanied and 13% are married during loan application.



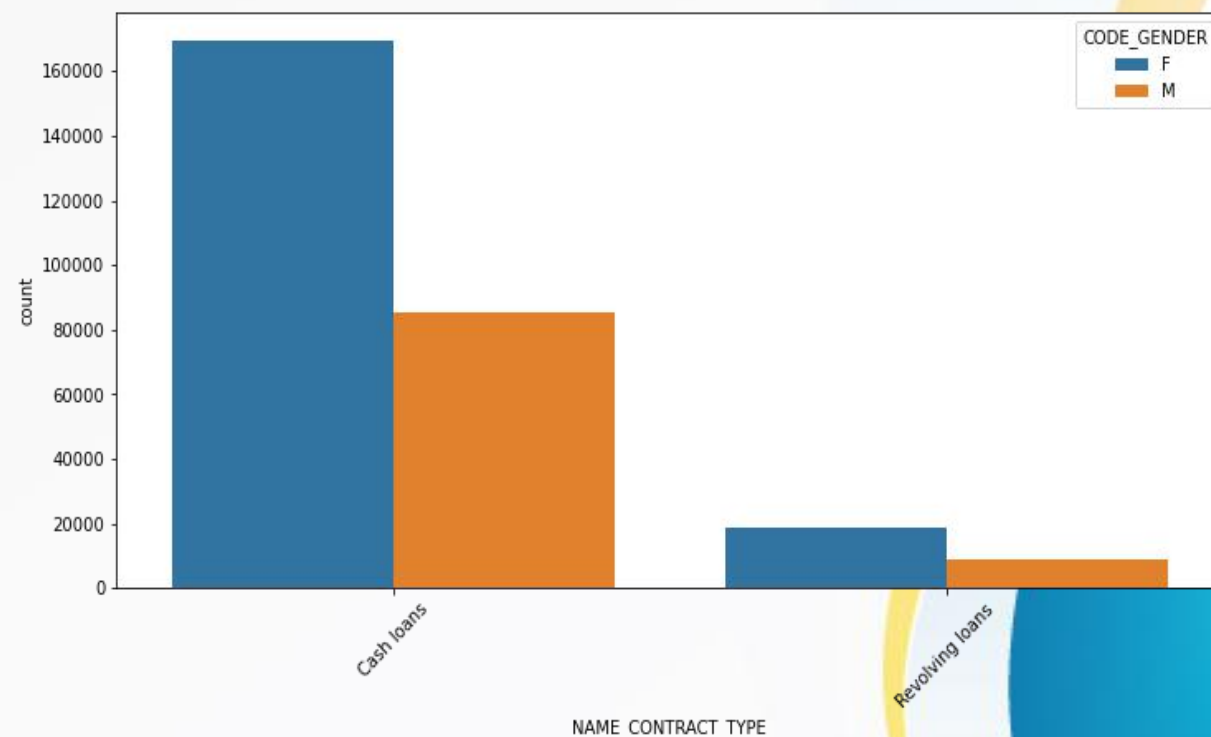
## Segmented Univariate Analysis

- For non defaulter Cash and revolving loans have applied by Female is more.
- Cash loan have much more than revolving loans
- Again defaulter same pattern for cash and revolving loans by female.
- Female Defulter are more than male in both loans types.

Target 1 contract distribution



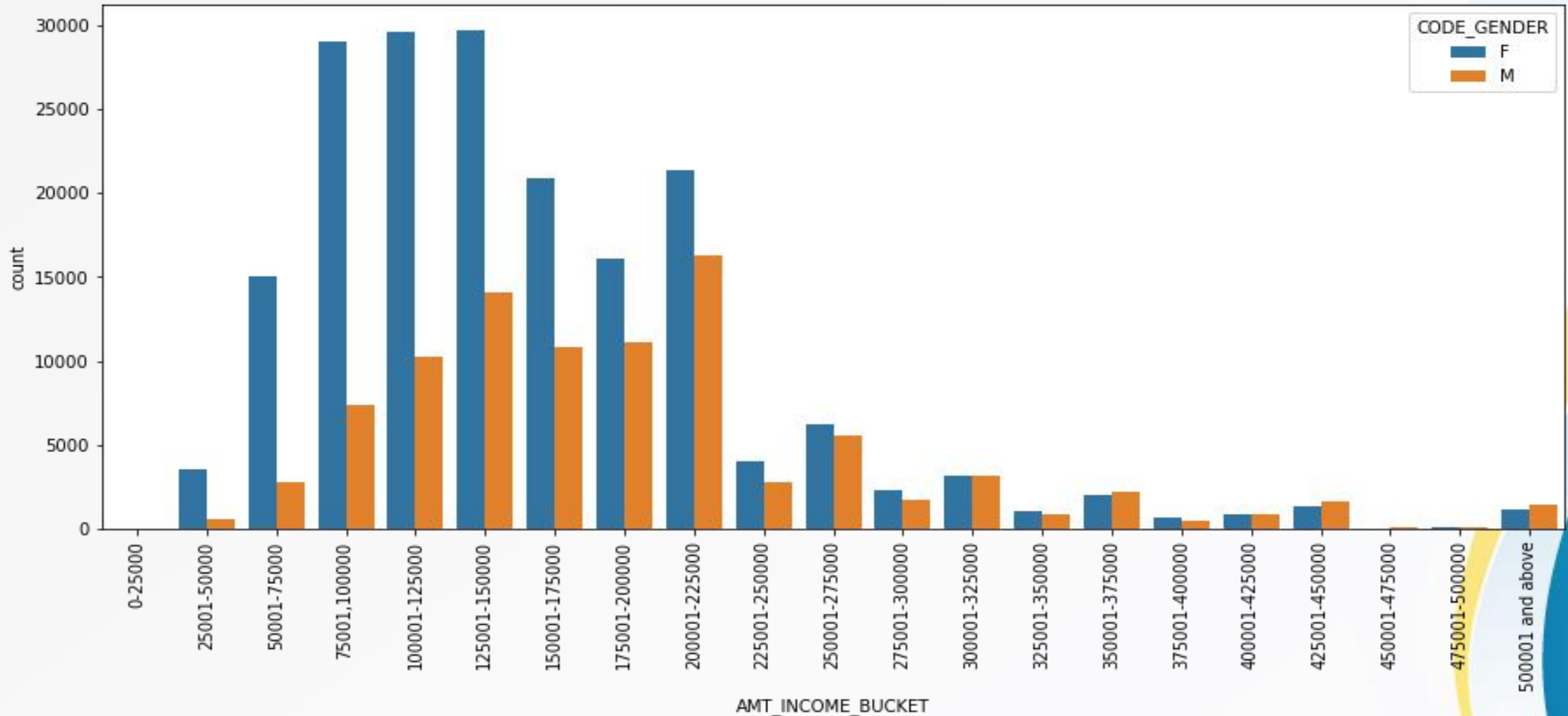
Target 0 contract distribution



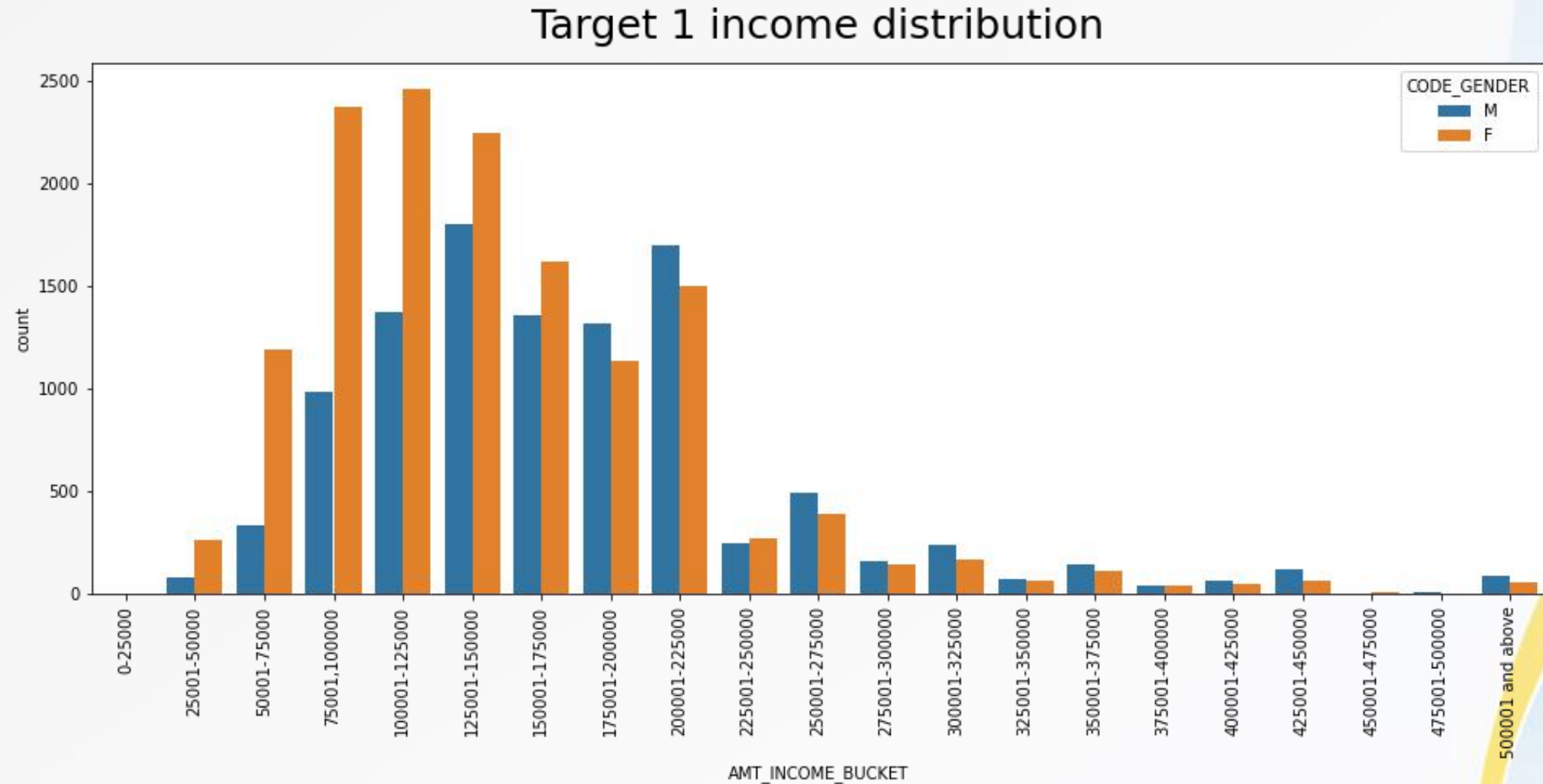
75001 to 150000 range of income group people have applied higher than others income group people.

High Income 225001 to above people have lesser applicant.

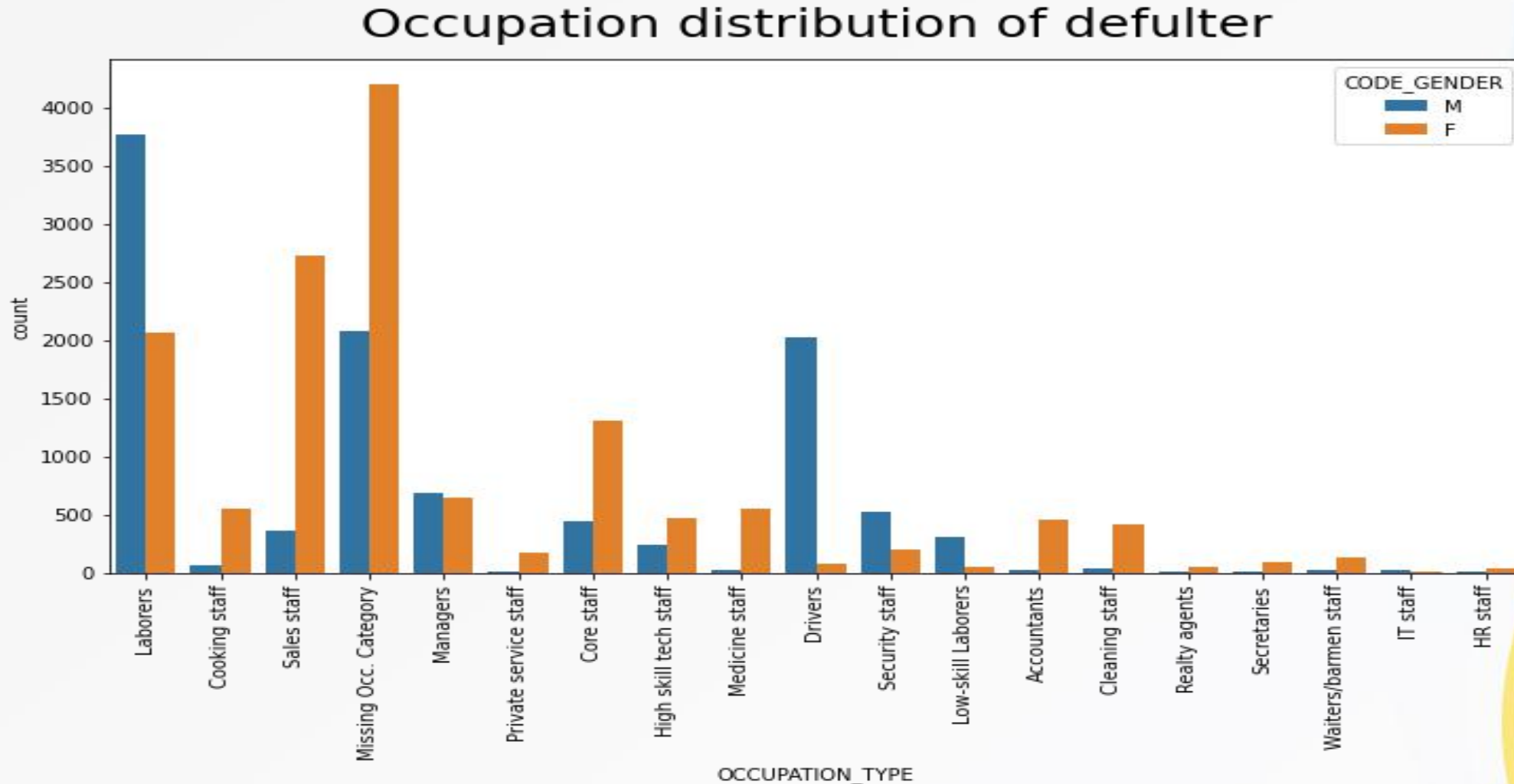
### Target 0 income distribution



75001 to 225000 income range female have applied for loans and have more defaulter.  
Female and male ratio difference is less for defaulter.

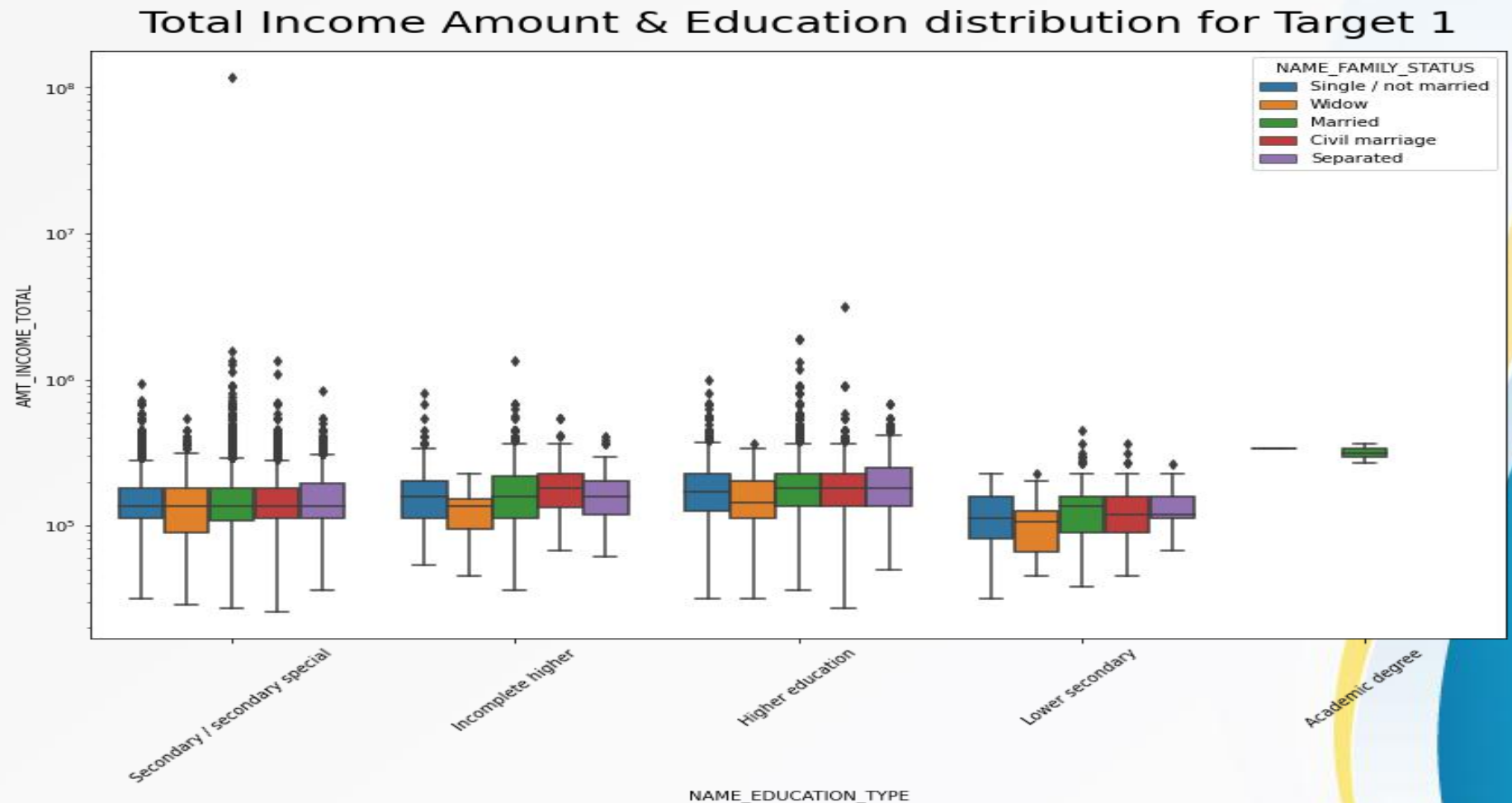


- 1.High number of male defaulter are from labours and driver occupation.
- 2.Sales and missing category have female defaulter are quite high

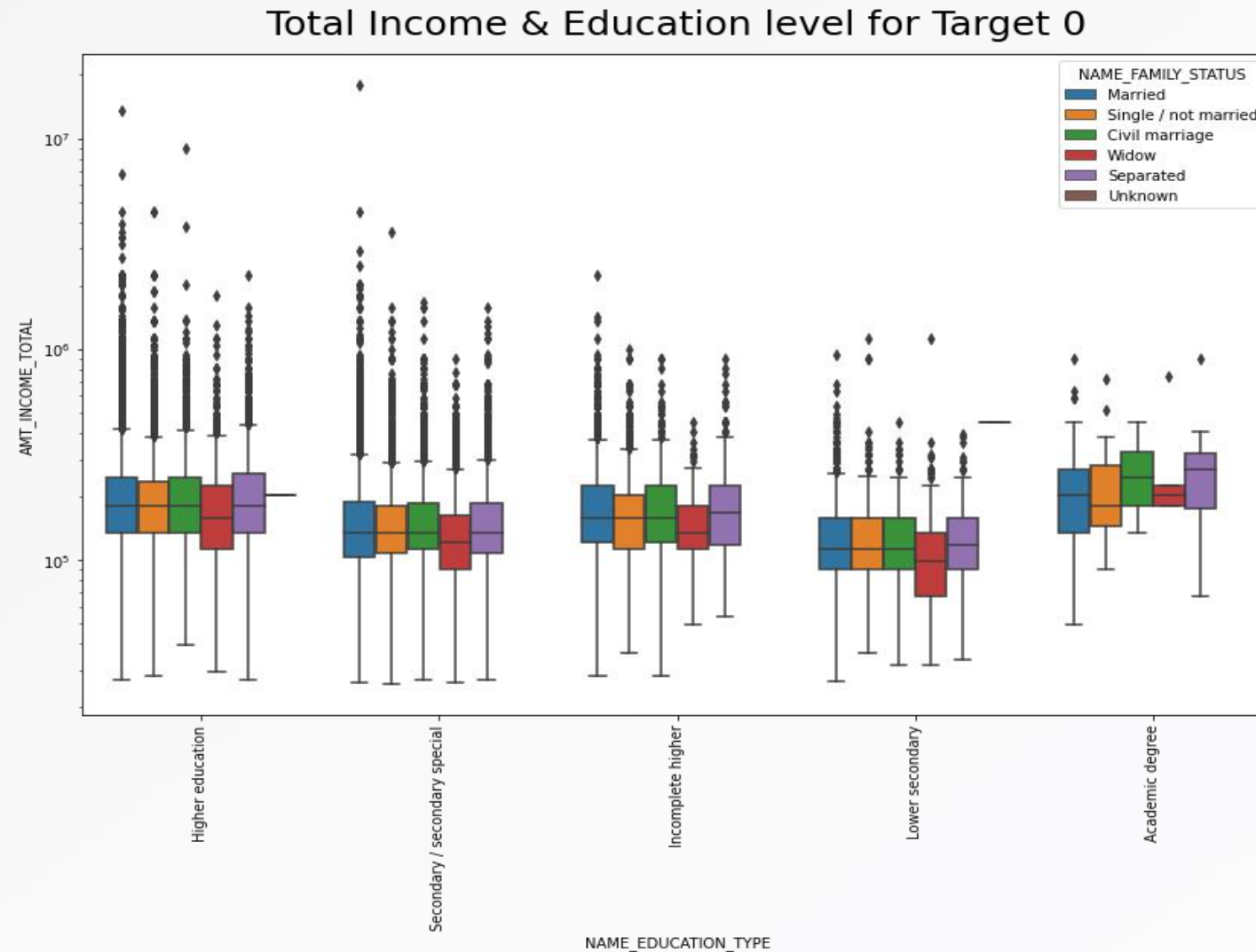


# Bivariate Analysis

1. For Secondary education and defaulter people have same median.
2. First quatile quite high.
3. For Academic degree and defaulter dont have any outliers which means limited income and they are married.

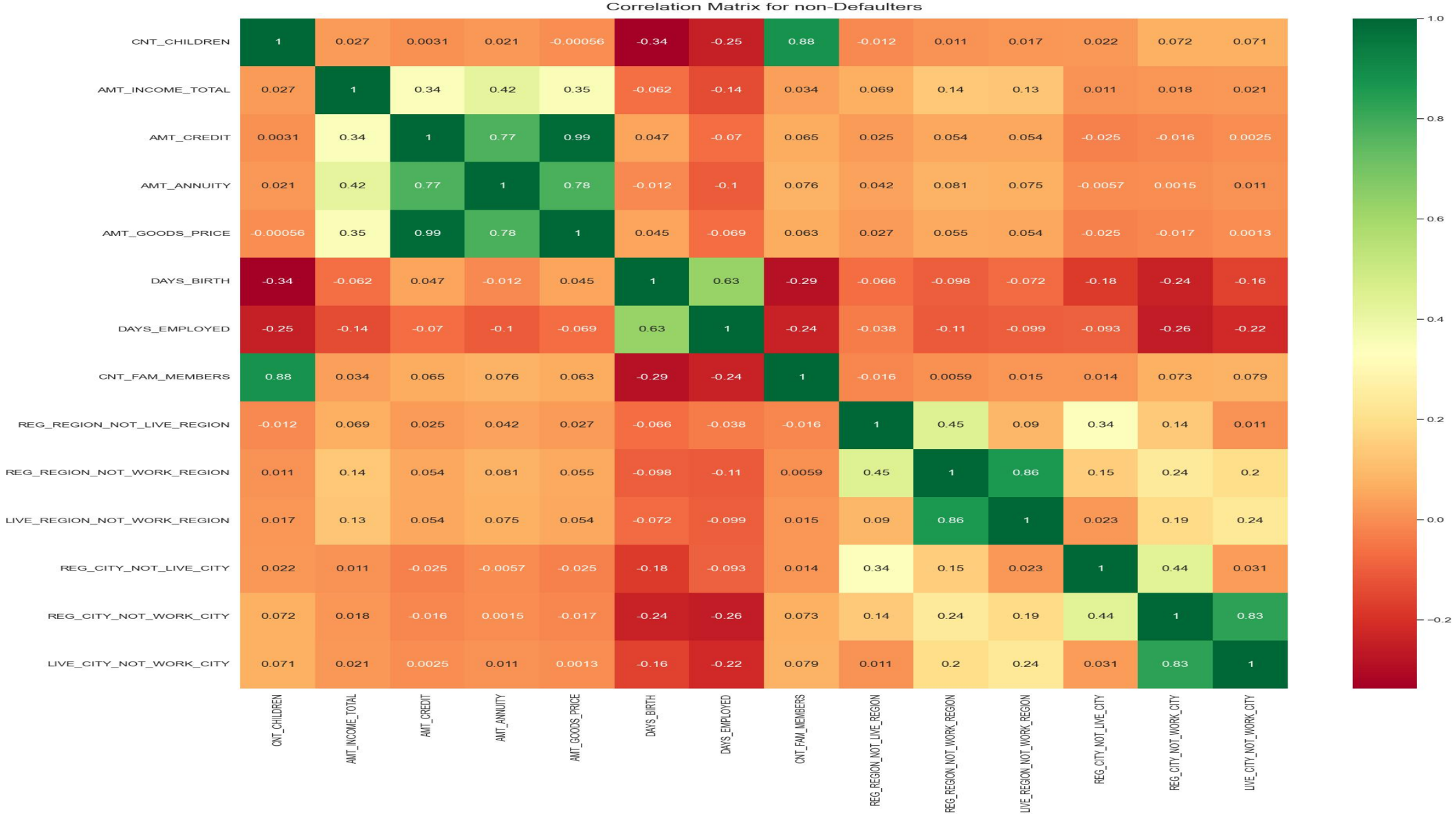


1. With Academic degree have high median values than rest of group which means total income is uniformly high.
2. Higher and secondary education almost same outliers and median values.



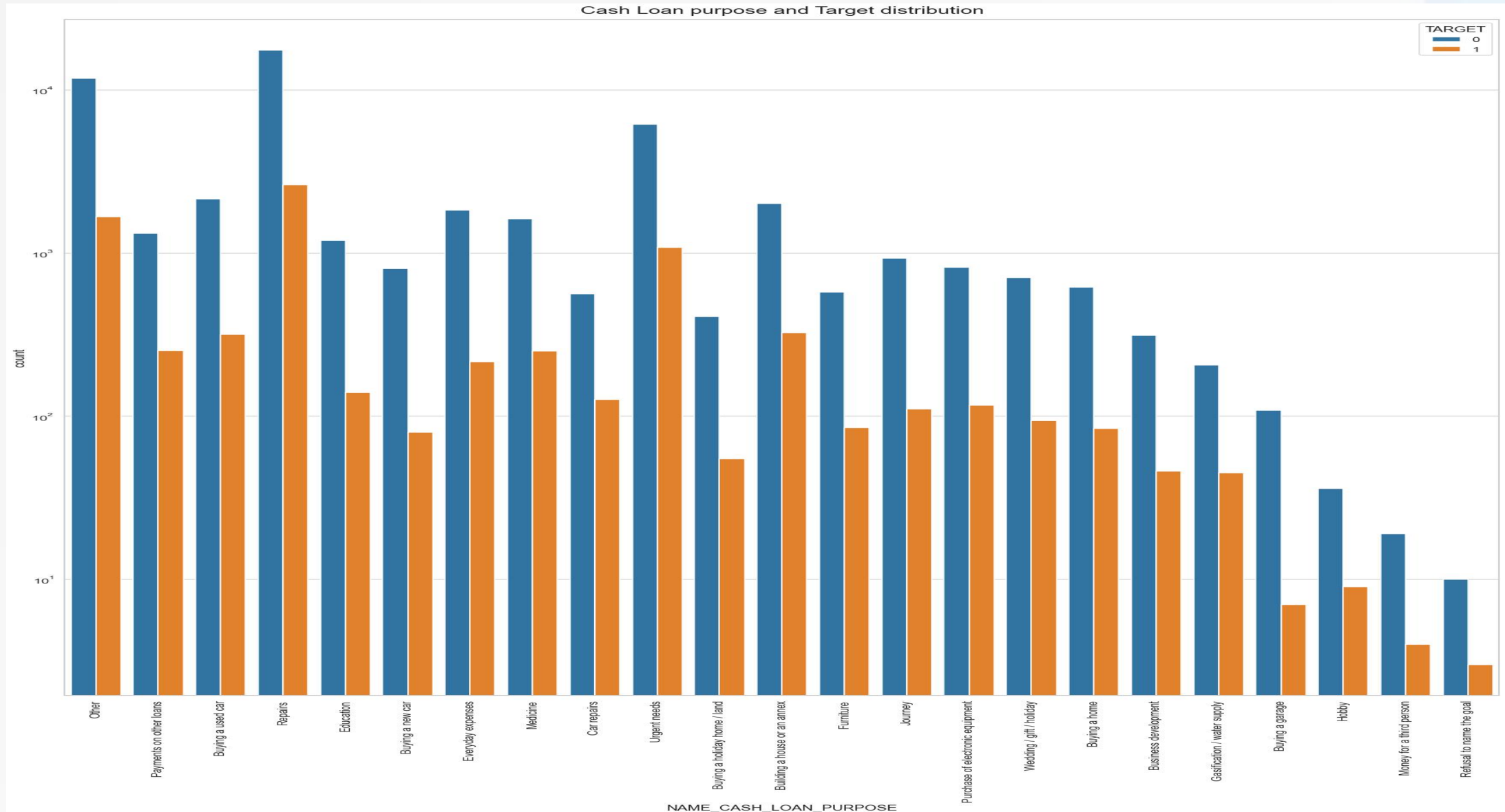


Top Ten Correlation of Variable for non defulter--Amount Credit and Goods Price have high corelation,Family Members and children number ,live city and registration city,live region and registration city,Days employed and age,Imcome and Annunity



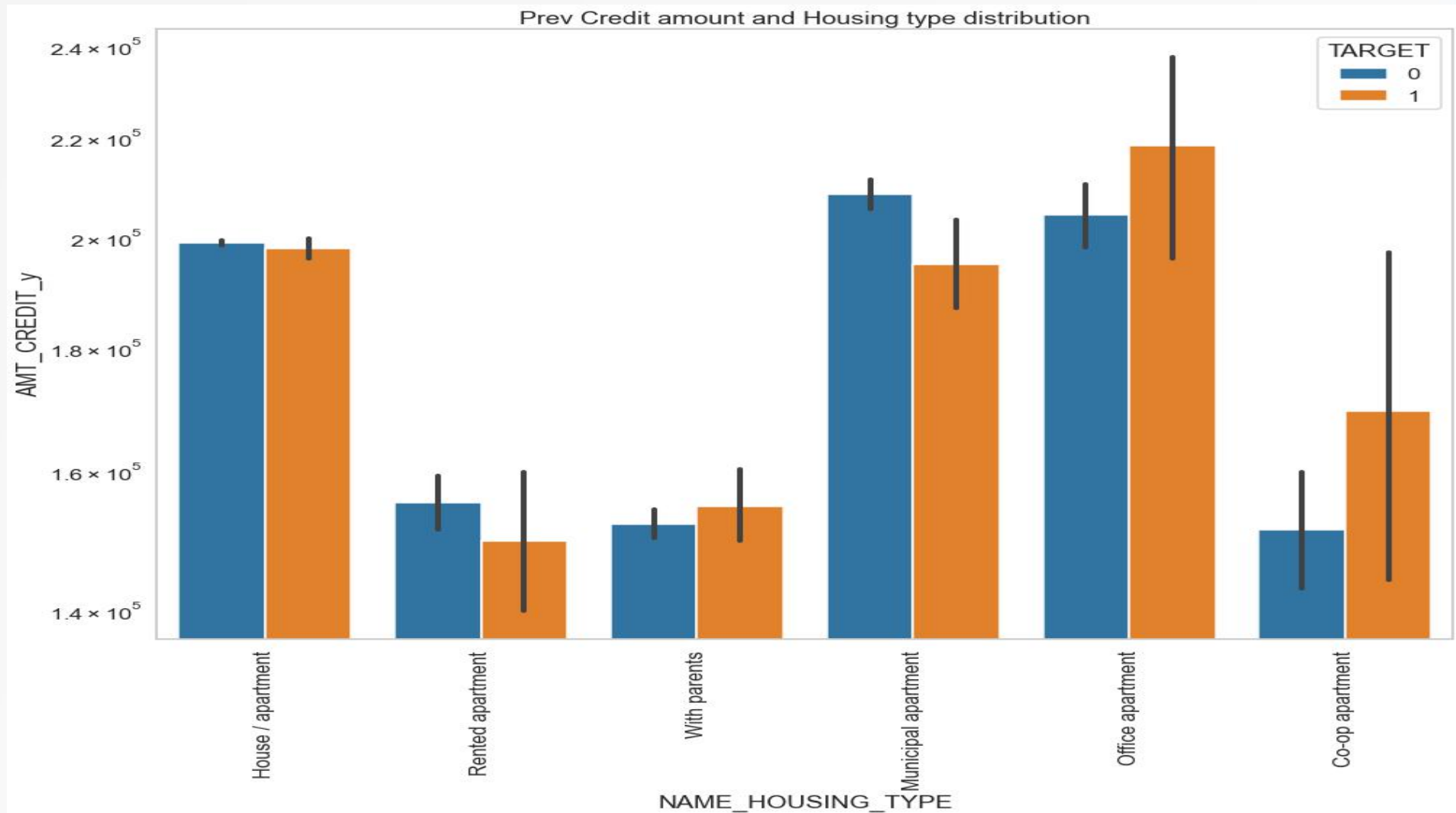
# Previous Application Analysis-

- 1.Repair category have highest number of refusal loan, than other and urgent needs.
- 2.On the other hand Education, Buying a new car, buying a home/land have high number of target.

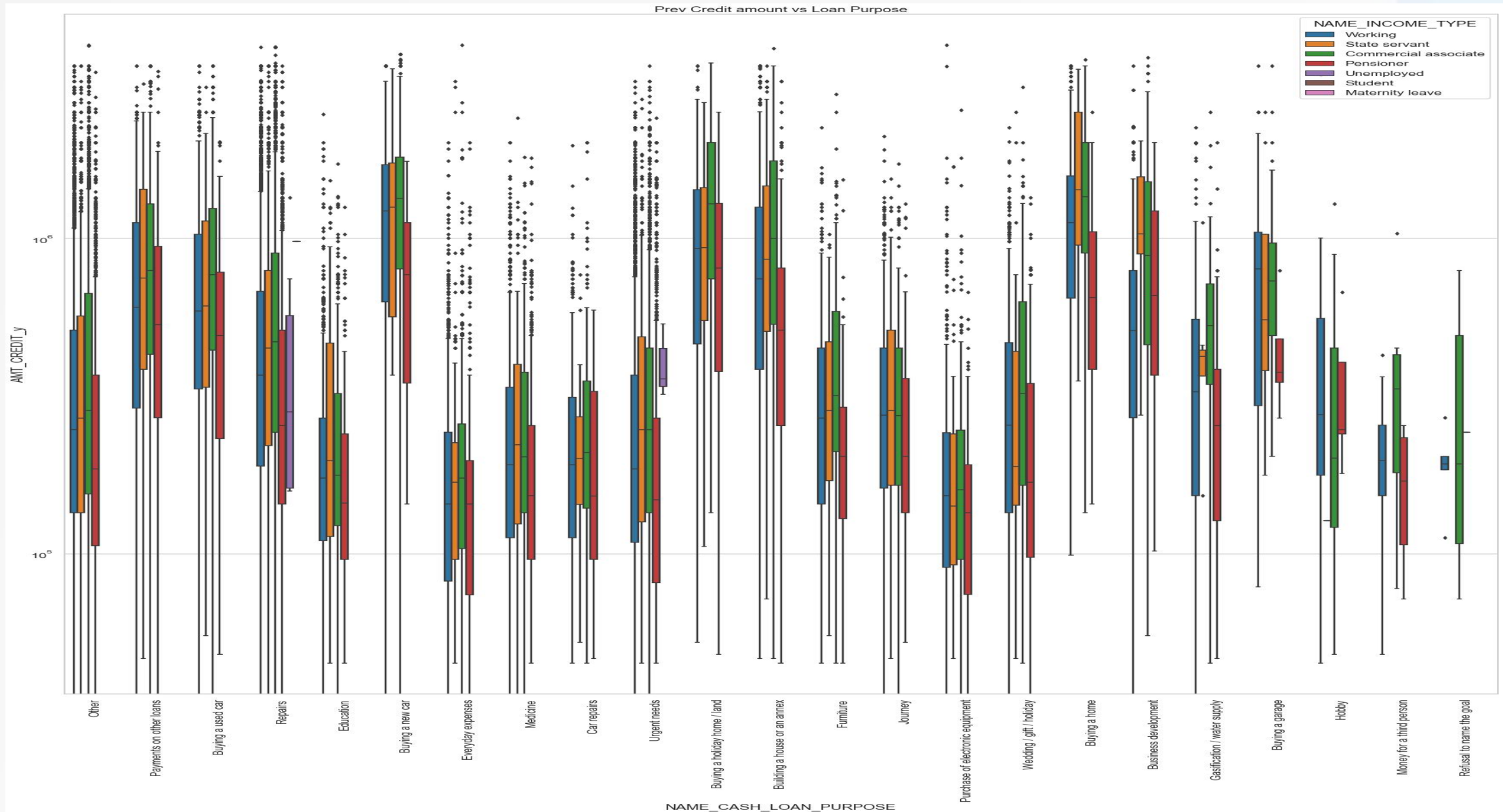


# Bivariate Analysis

- 1.Co-op and office apartment type people higher number defulter.
- 2.House/Apartment ,municipal apartment and rented type housing have good numbers of non-deulter.



1. Buying home, Buying a new car, and buying a holiday/home/land have higher loan amount credit.
2. Hobby, money for third person, refusal to name the goal have lesser number of loan credit.



# Conclusion

- 1. Bank Can try for with House/Apartment,Municipal apartment type customer as they have less payment problem or defaulter.
- 2. With loan purpose Repair Category ,have higher payment problem, so bank give lesser attention.
- 3. Very\_young Age group people have high defaulter.
- 4. Married and female have high number of defaulter.

THANK YOU

