

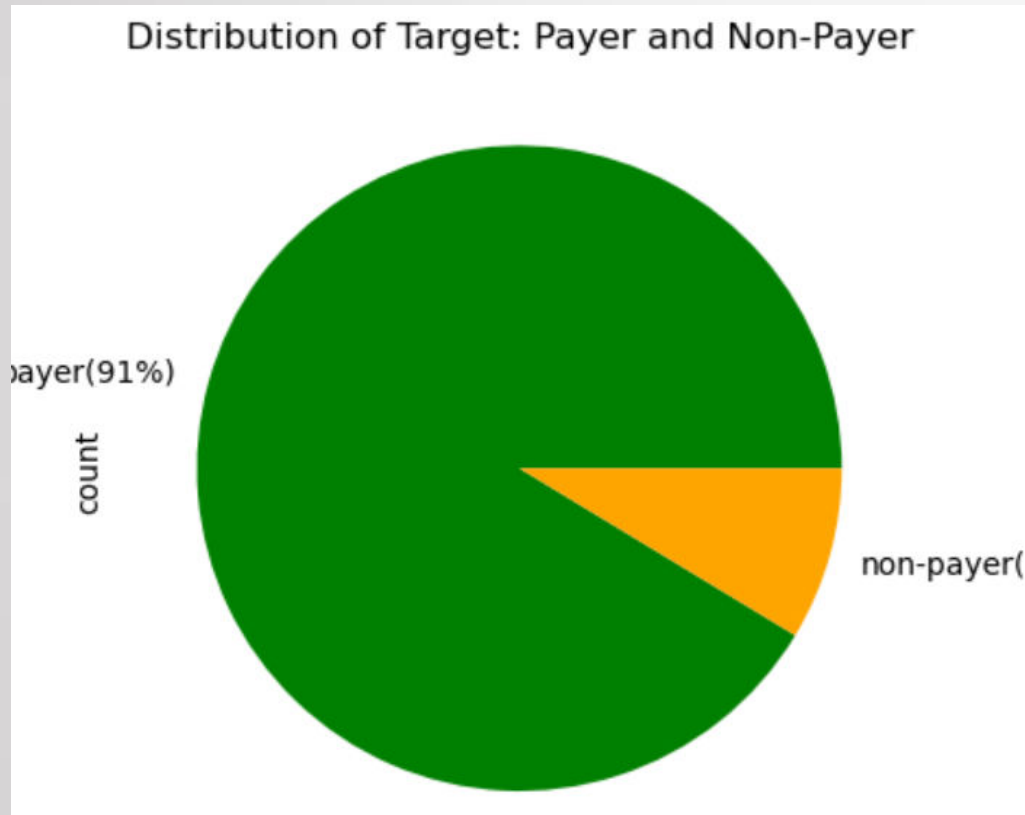
CASE STUDY

BY: AKSHITA GOEL

# CREDIT EDA ASSIGNMENT

PAYER AND NON-PAYER  
MALE AND FEMALE

# DATA IMBALANCE

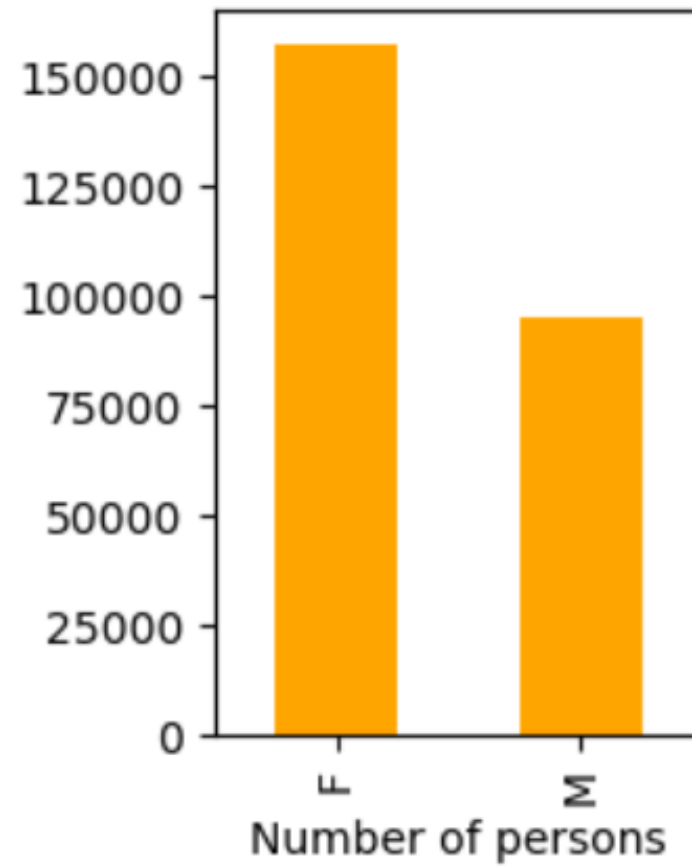


# TARGET DISTRIBUTION

- Target 0 are 91% approximately and
- Target 1 are 8%
- Target 0 are those who are non-defaulters to whom we can give loan easily
- Target 1 are people who are unable to give timely installments.



Females are 202448 in number and Males are 105059 in number


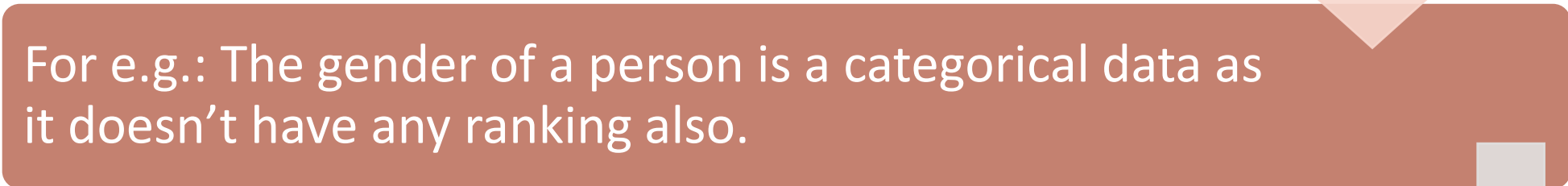


# UNIVARIATE CATEGORICAL ANALYSIS

Categorical univariate data – It consists of non-numerical observations which can be placed into categories.



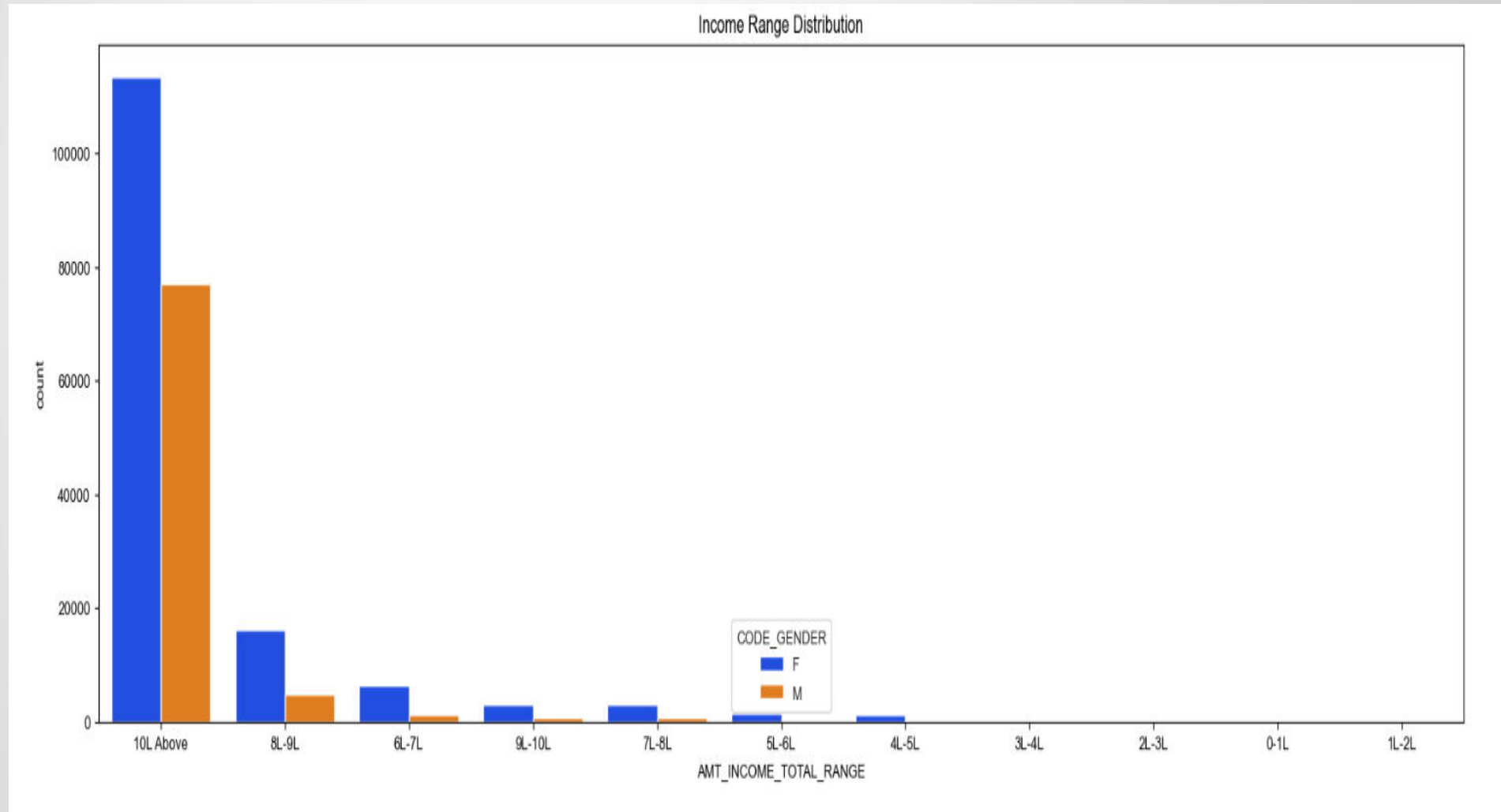
For e.g.: The gender of a person is a categorical data as it doesn't have any ranking also.



We cannot rank color of our skin, or gender.



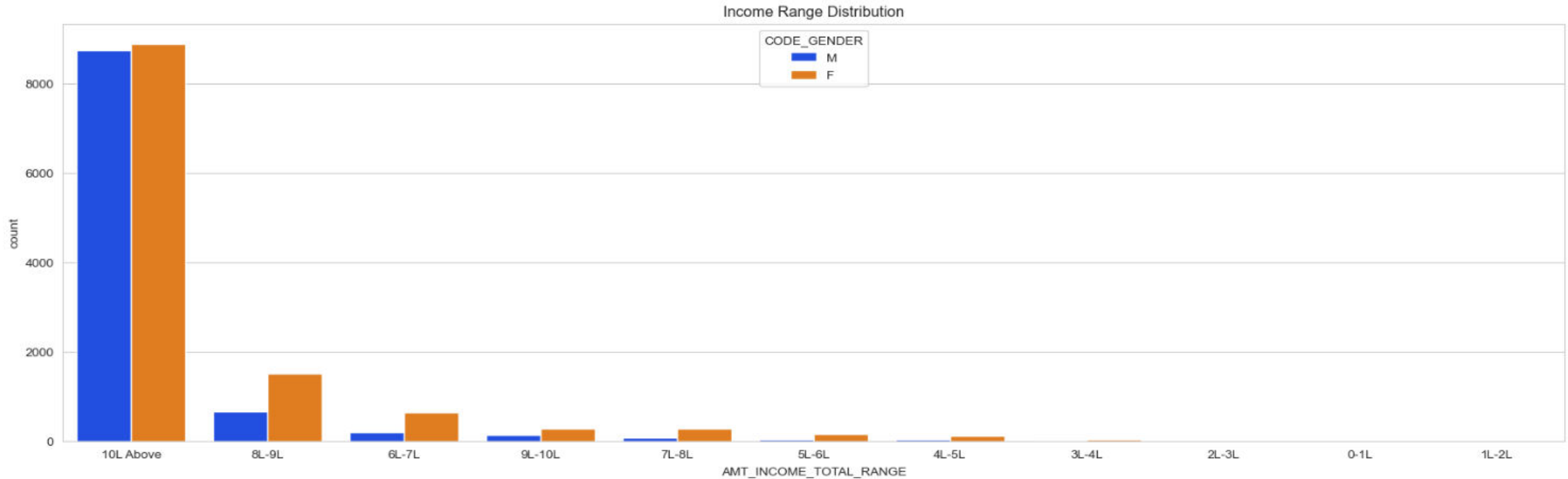
# Income range for target 0



## Observations For Target 0:

We can see that females count is higher in having credit compare to males  
Most of the people having credit are earning 10L and above

# Income range for Target 1



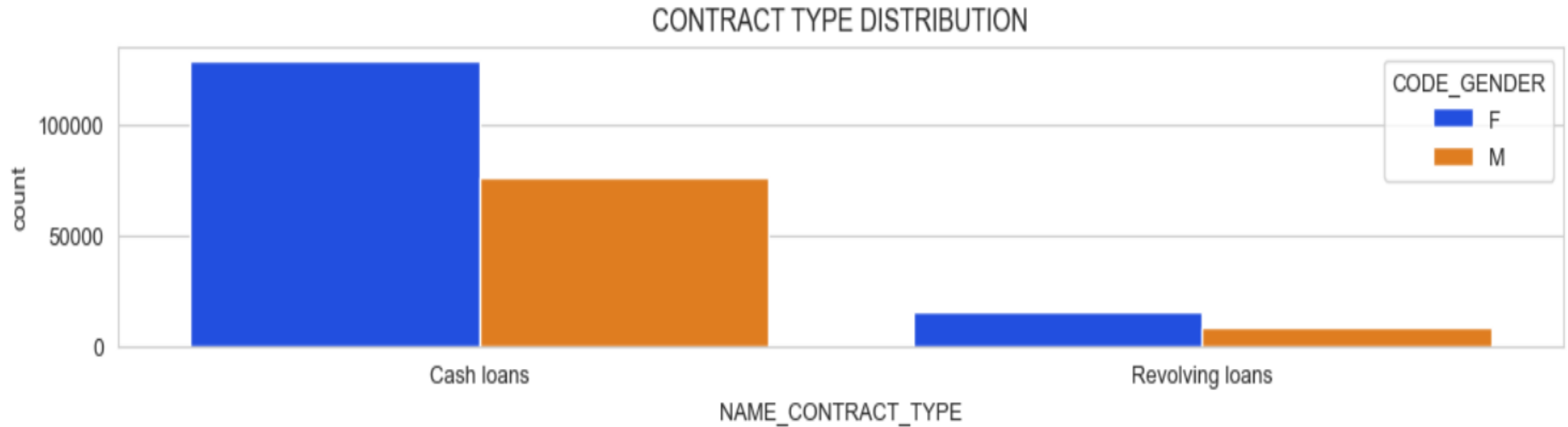
Observations For Target 1:

Females count is higher in having credit compare to males

Most of the people having credit with paying difficulty are earning 10L and above

Max count is very less compared to the max count for Target 0

# Contract Type For Target 0



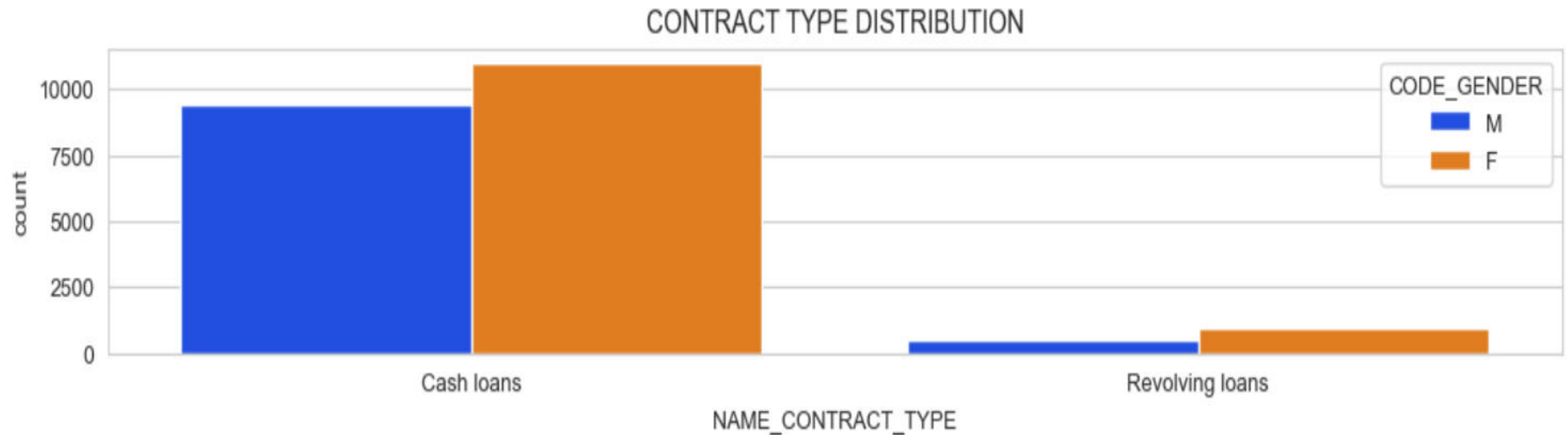
Observations For Target 0:

We can see that females count is higher in both type of loans

Most of the people prefer Cash Loans



# Contract Type for Target 1



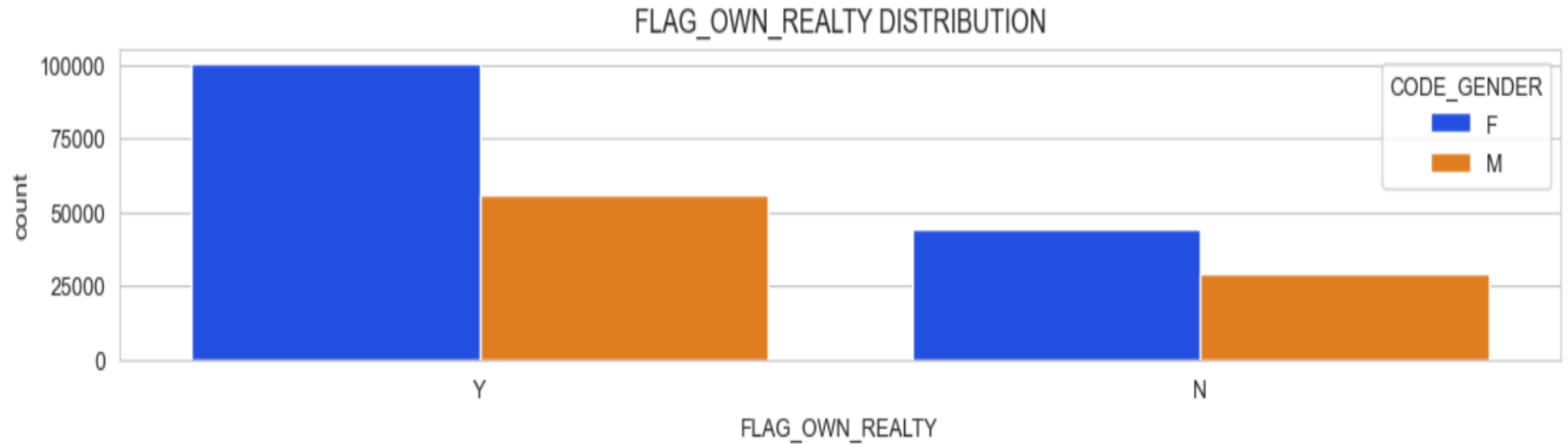
Observations For Target 1:

Females count is higher here as well

Most of the people having credit with paying difficulty are also taking cash loans

Revolving loan is taken by very few males

# FLAT OWN REALTY FOR TARGET 0

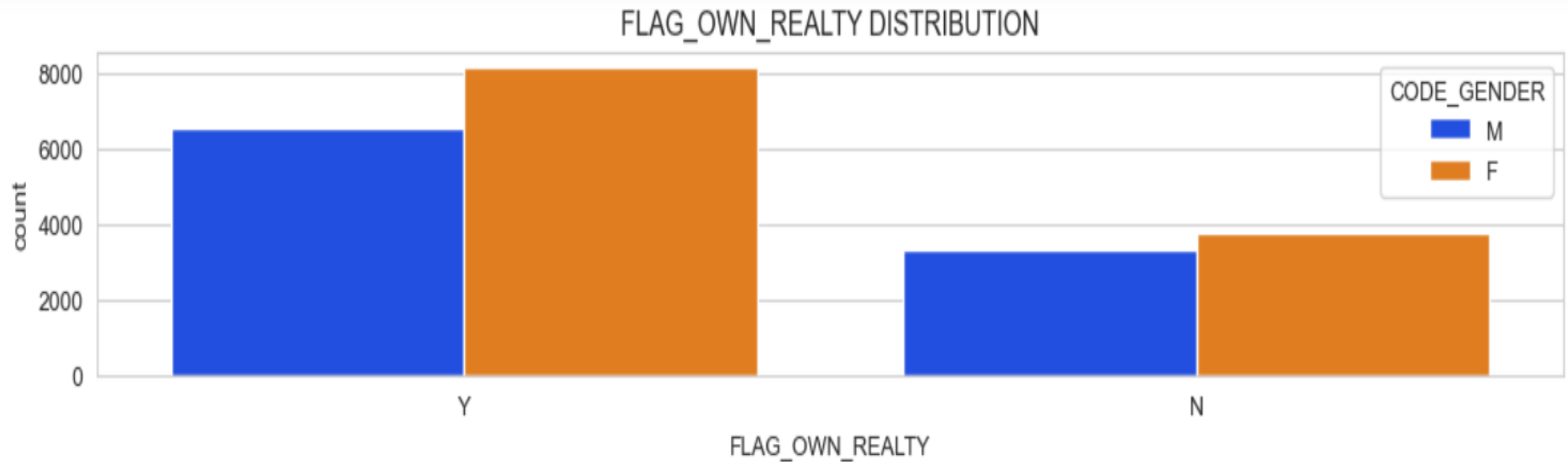


Observations For Target 0:

Most of the people don't have their house, still they are able to pay on time

Most of the people who have their own real\_estate are double than the people who don't have

# FLAT OWN REALTY FOR TARGET 1



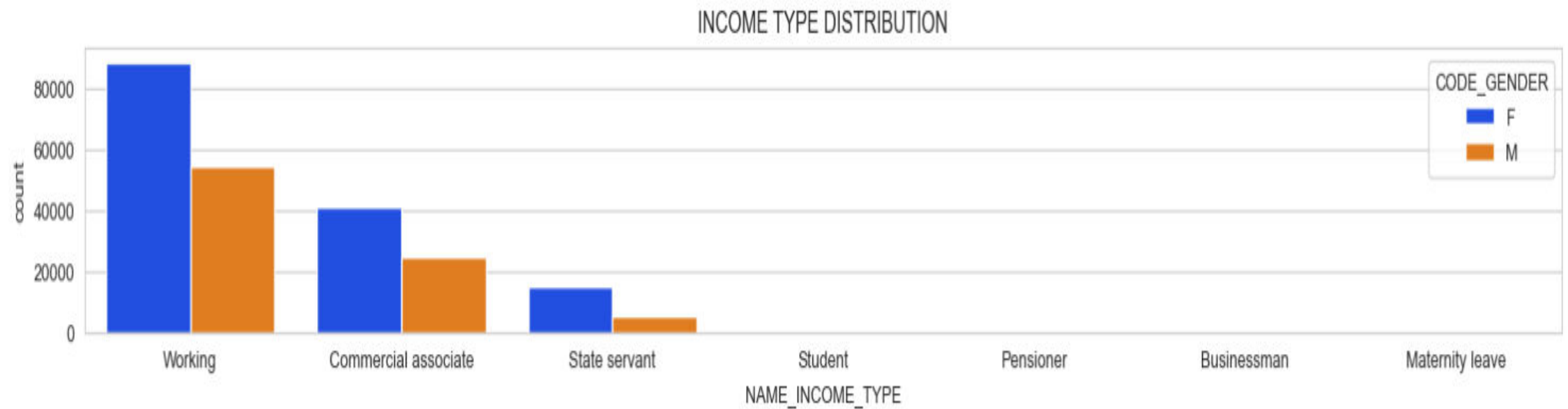
Observations For Target 1:

Females count is higher here as well

Most of the people who have their house, still they are defaulters

We can conclude that there is no correlation between owning a house and paying on time

# INCOME TYPE OF CLIENT FOR TARGET 0



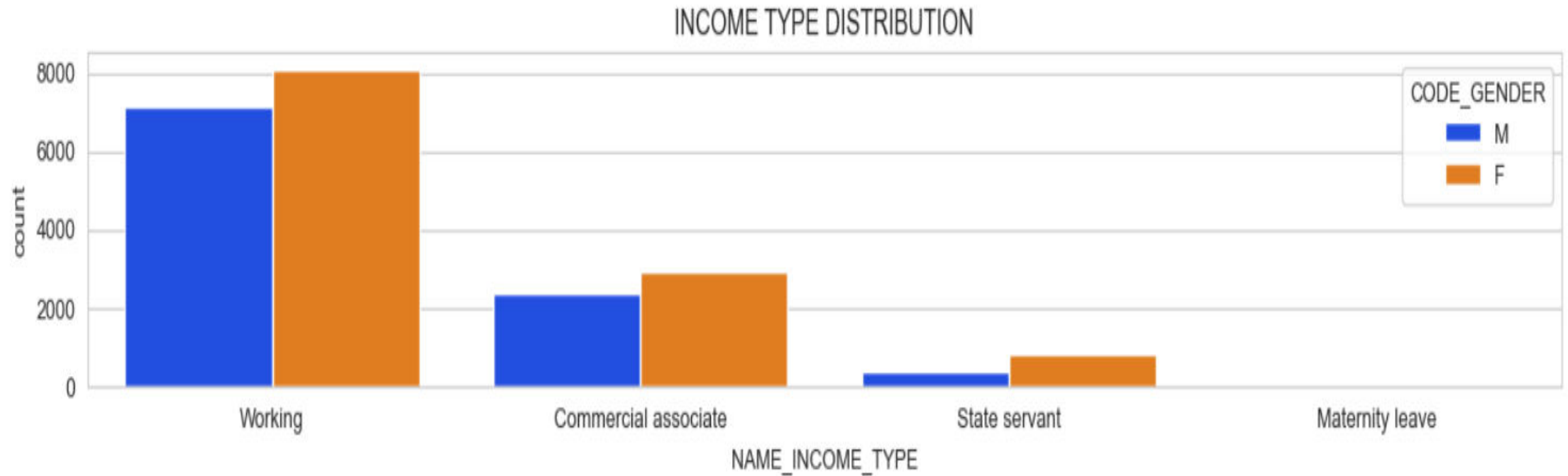
Observations For Target 0:

Most of the people come from working background followed by commercial associate

# INCOME TYPE OF CLIENT FOR TARGET 1

Observations for Target 1:

Most of the people come from working background followed by commercial associate



# UNIVARIATE NUMERICAL ANALYSIS

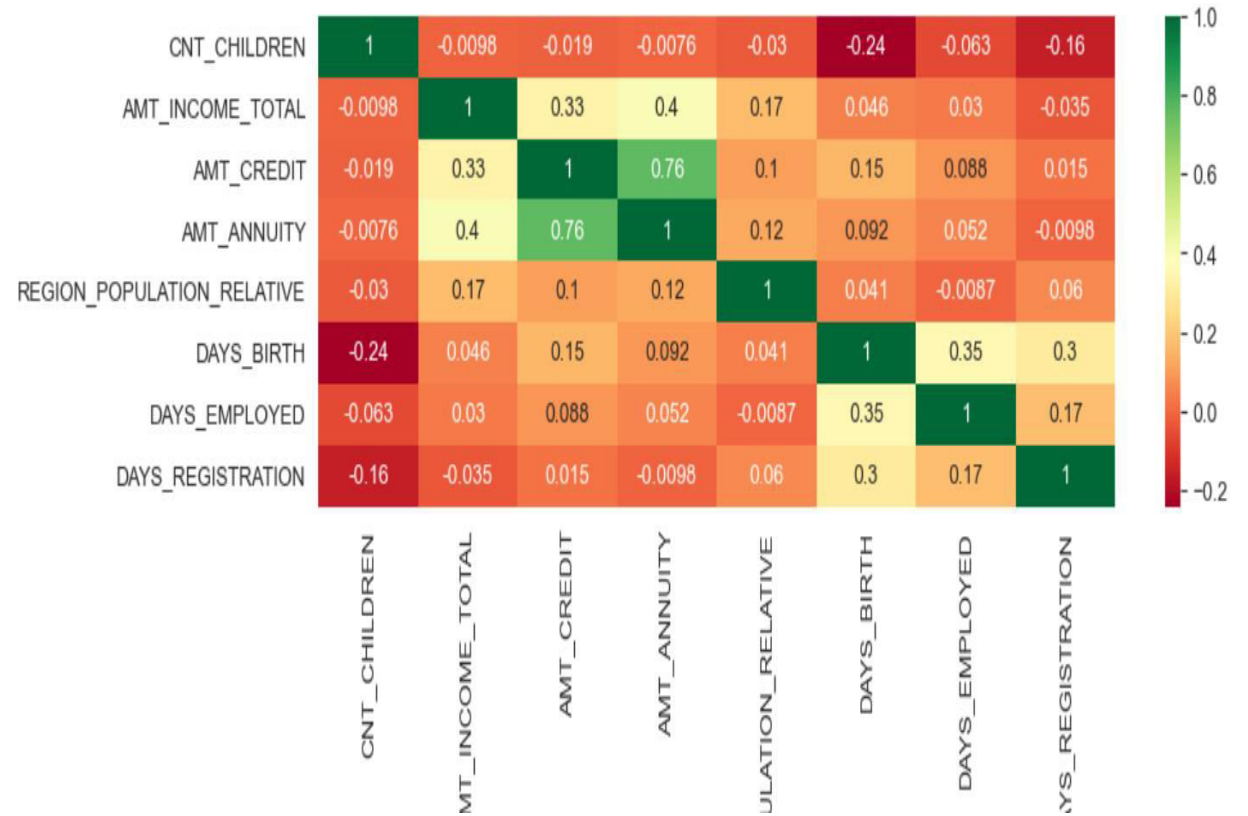


WE WILL ANALYSE ONE VARIABLE AT A TIME,  
THIS MEANS UNIVARIATE.



IF THAT VARIABLE IS NUMERICAL LIKE INCOME  
, ANNUITY ETC. THAT WILL COM UNDER  
NUMERICAL ANALYSIS

# CORRELATION OF NUMERICAL COLUMNS FOR TARGET 0



Observations:

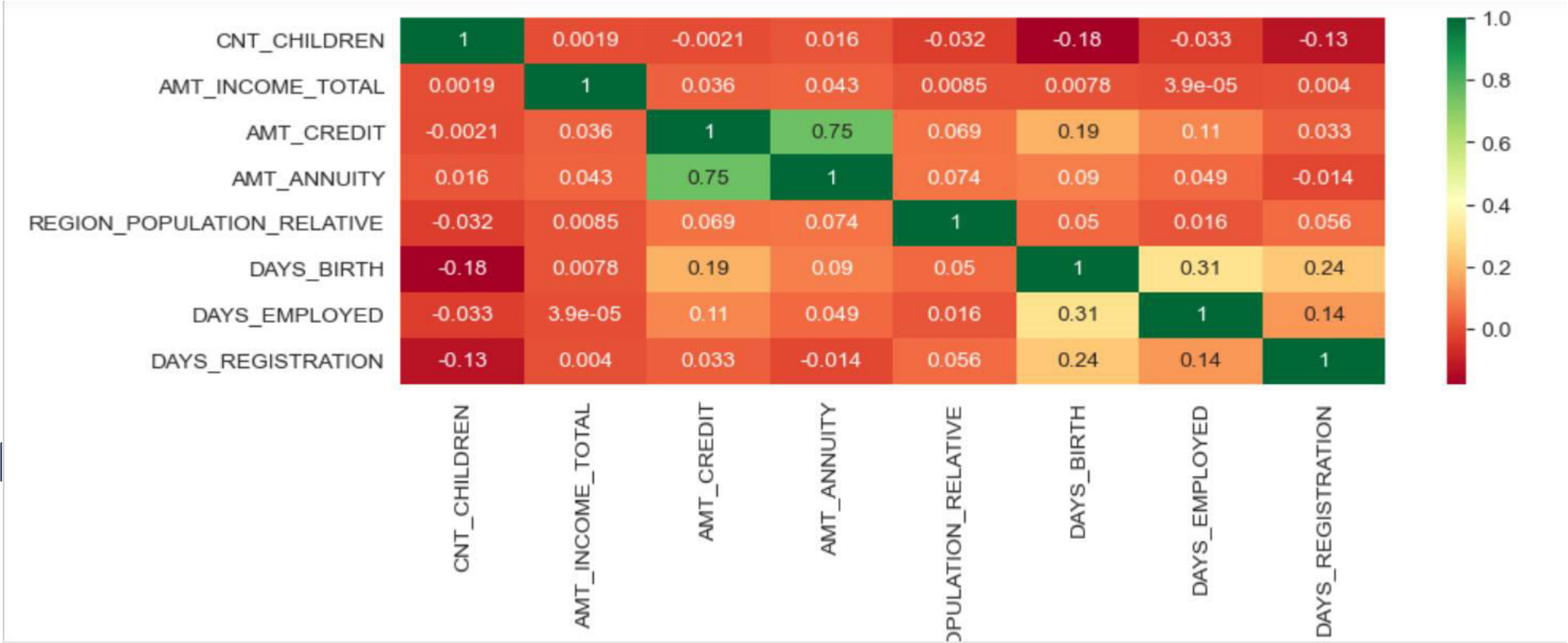
Amt\_Income\_Total is not that much connected to Days\_Employed, which means that no. of days employed is not much related to Income earned by that person

CNT\_Children is inversely proportional to DAYS\_BIRTH, which means old people have less children

Amt\_Annuity is related to Amt\_Credit, higher the credibility higher the annuity

Observations for Target 1:  
Most of the values are negatively correlated

CONF

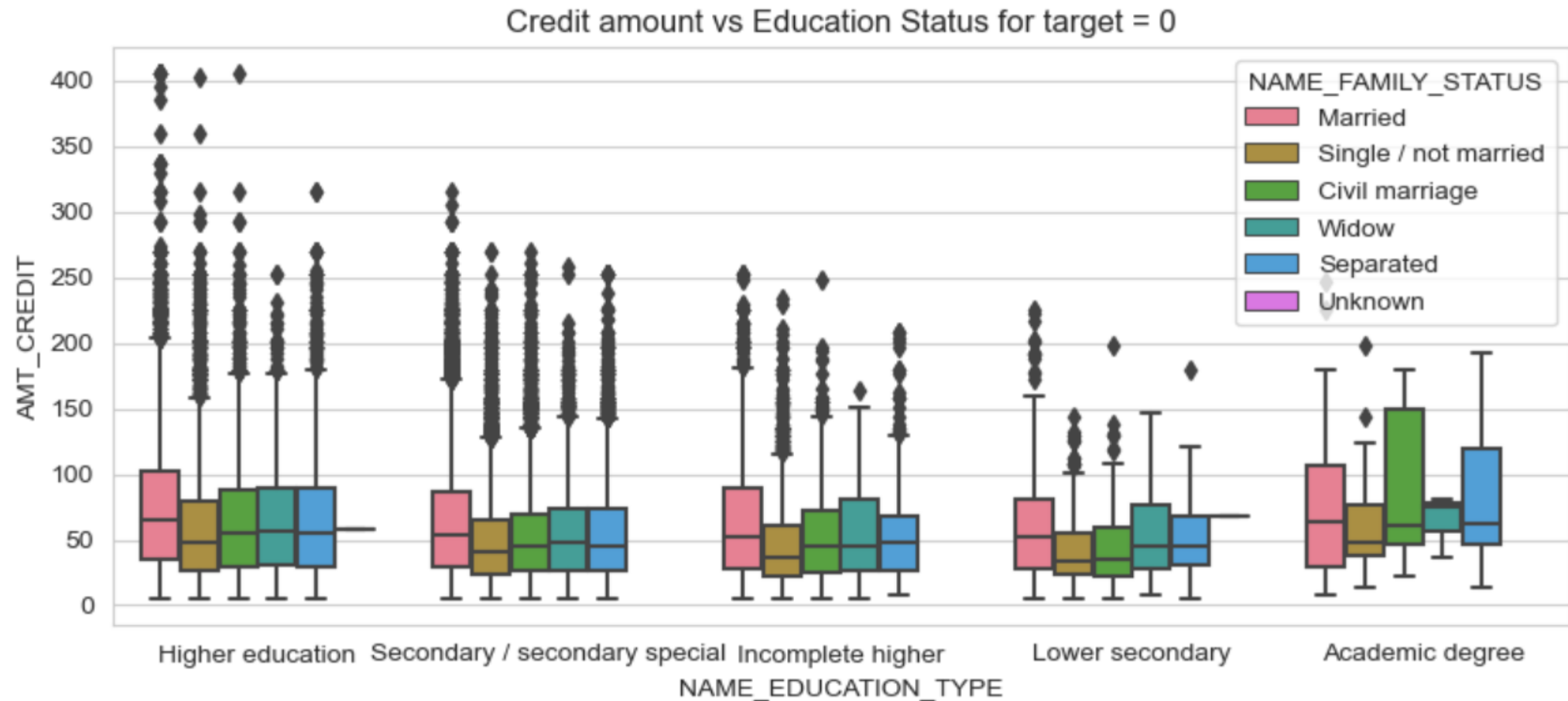




# BIVARIATE ANALYSIS

IT INVOLVES ANALYSIS OF 2 VARIABLES TO KNOW  
ABOUT THE RELATIONSHIP BETWEEN THEM

# CREDIT AMT VS EDUCATION STATUS FOR TARGET 0



# Analysis

Observations for Target 0:

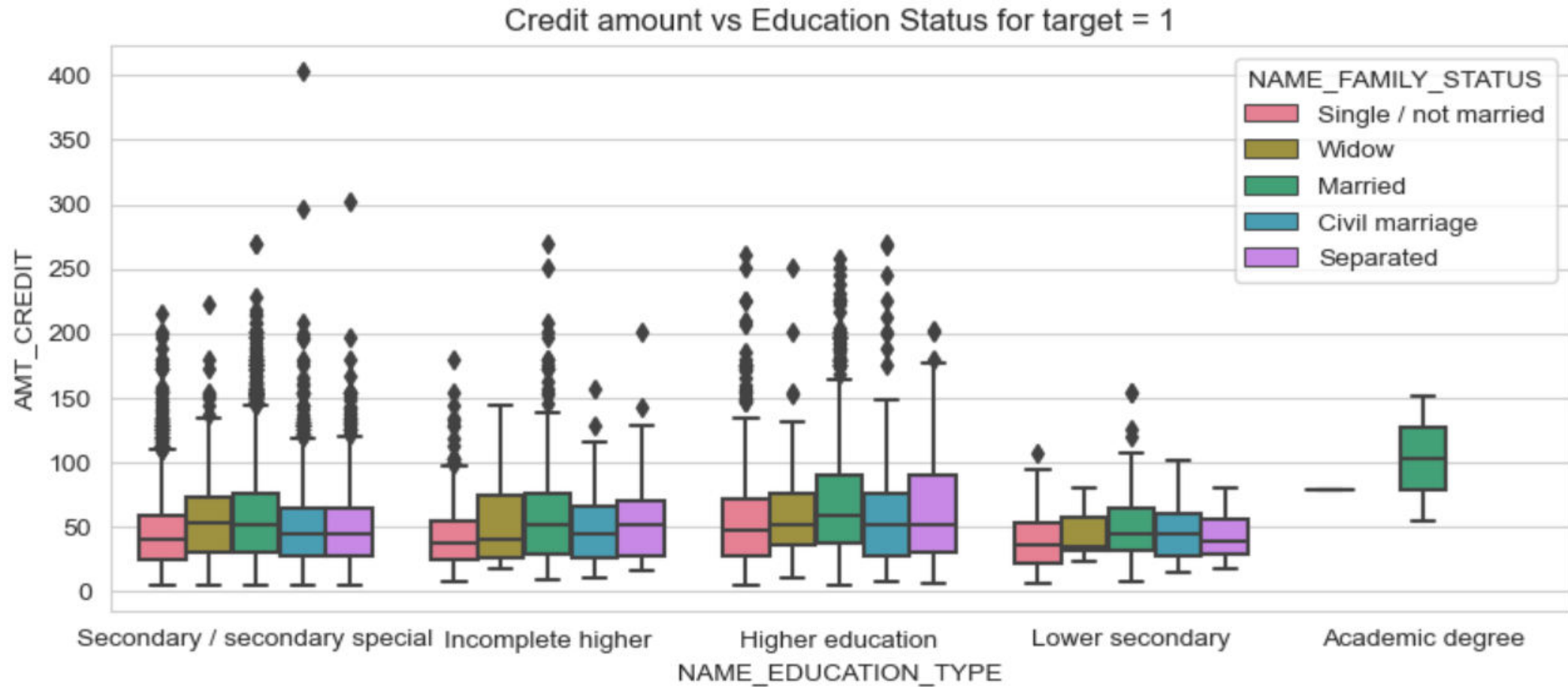
In most of the cases high amount is credited by the people who are married

People with status married and single are having high outliers.

People with higher education and who are also married has high probability of crediting

People with civil marriage are also crediting good amount

# CREDIT AMT VS EDUCATION STATUS FOR TARGET 1



# Analysis:

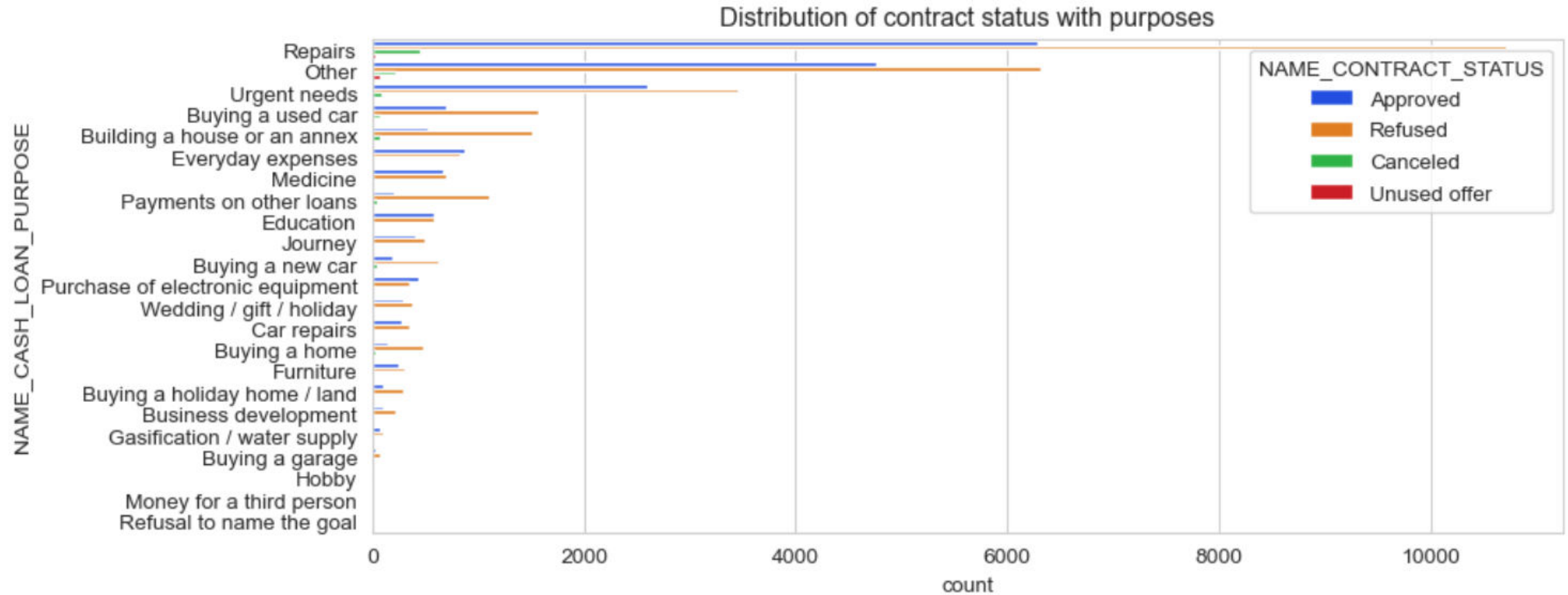
Observations for Target 1:

People with academic degree has least chances of being defaulter, condition- they should not be from the civil marriage status

People with family status marriage has high chances of being defaulter

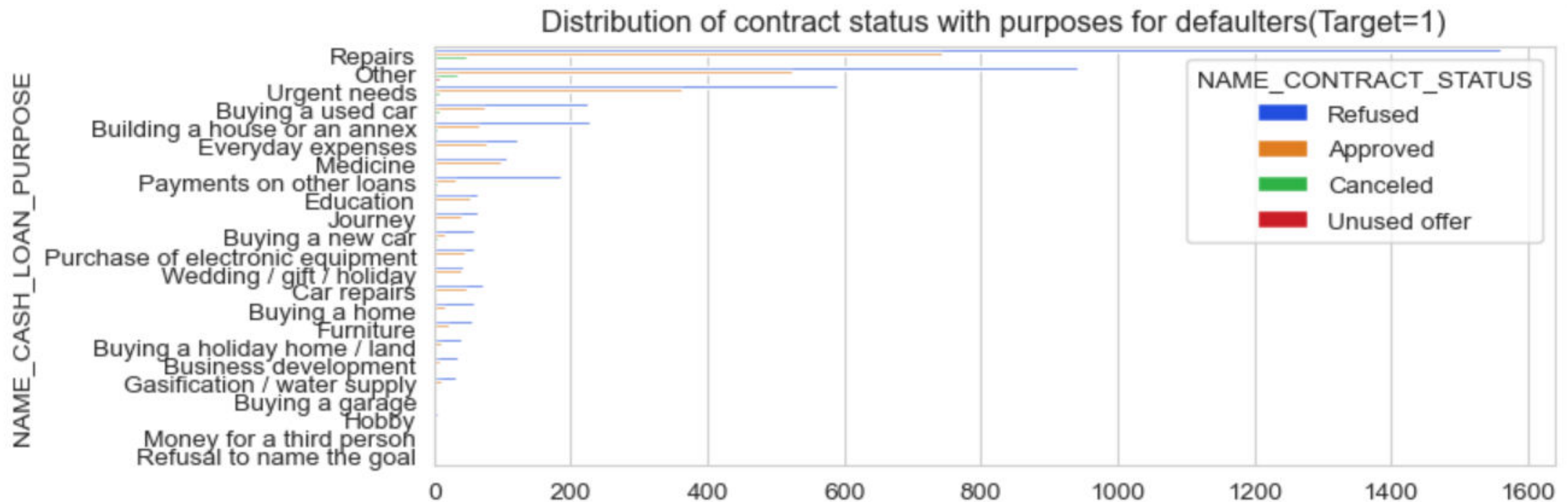
We can conclude that being married or not being married has nothing to do with amount crediting capacity

# MERGED DATA ANALYSIS



Maximum number of refusals and approvals are done for repairs.  
Very few people have taken loan for their hobbies.  
Unused offers are very less in count

# TARGET 1



We can see that refusals are maximum for each category, which means that in most of the cases analysers are analyzing correctly about the defaulters

Not the maximum number, but many loans are approved for defaulters

# CONCLUSION

- Banks should target high income people as high income means high AMT\_CREDIT and high annuity.
- Banks should see the Repairs department as maximum number of Approved loans which were defaulted were in Repairs.
- Banks should focus on married people as they have high probability of having high income
- People who are divorced and single and having academic degree are also a reliable source.