

C R E D I T E D A C A S E S T U D Y

EDA CASE STUDY



DATA TOOLKIT

BY
**SHUBHNEET AND
SAMTA**

Univariate Analysis

FOR CATEGORICAL
VARIABLES

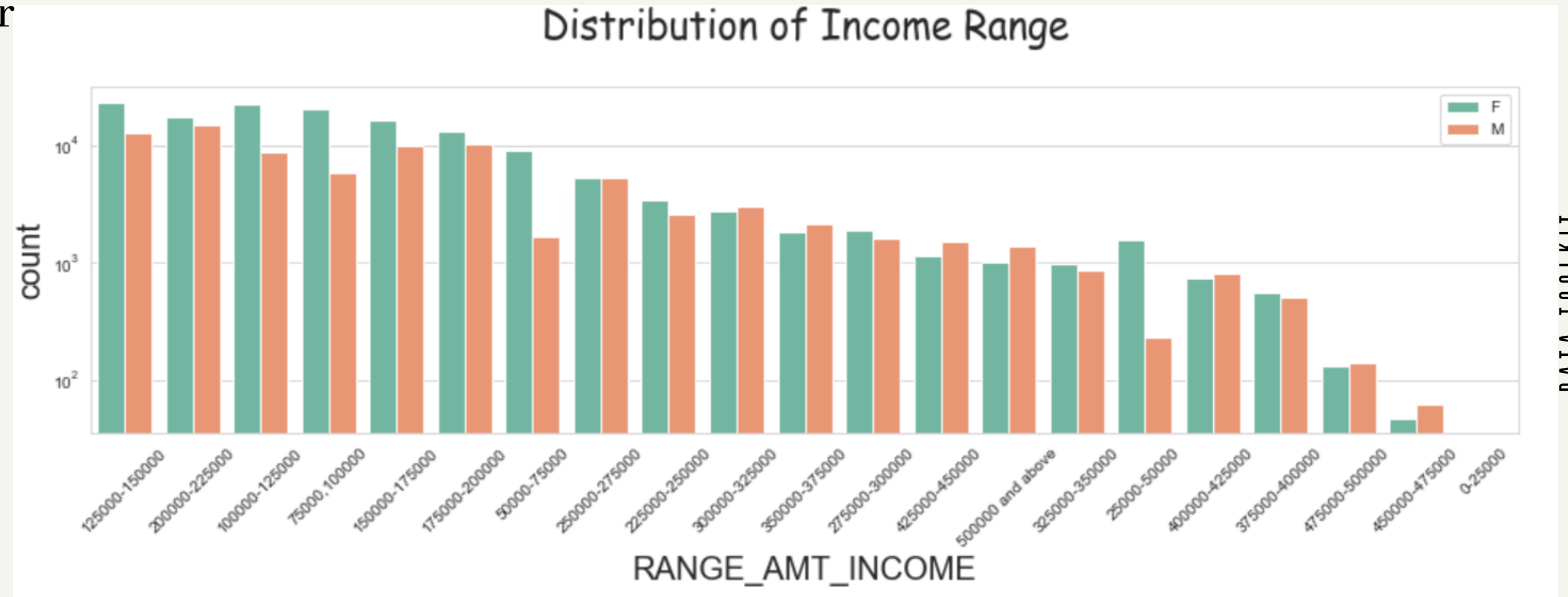
FOR NON-DEFAULT CLIENT
TARGET 0

INFERENCES

1. Female count is higher than male.

2. Income range from 100000 to 200000 is having more number of credits.

3. Income range below 400000 have least number of credits.

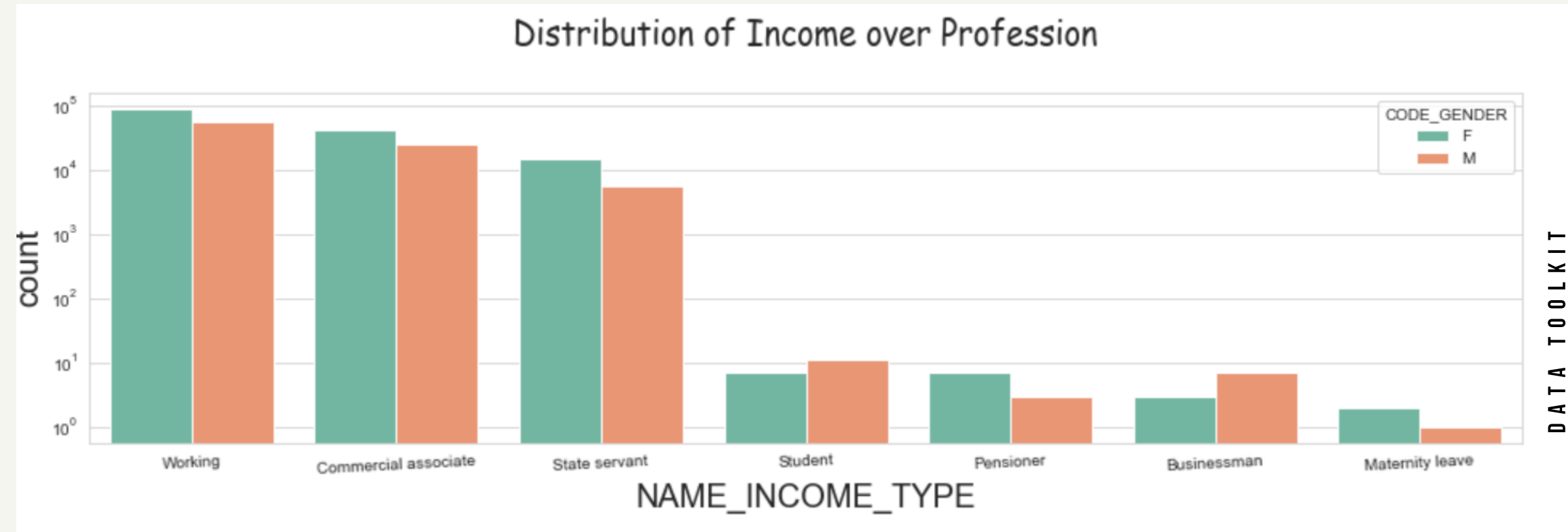


INFERENCES

1. 'working', 'commercial associate', and 'State Servant' professions have higher number of credits.

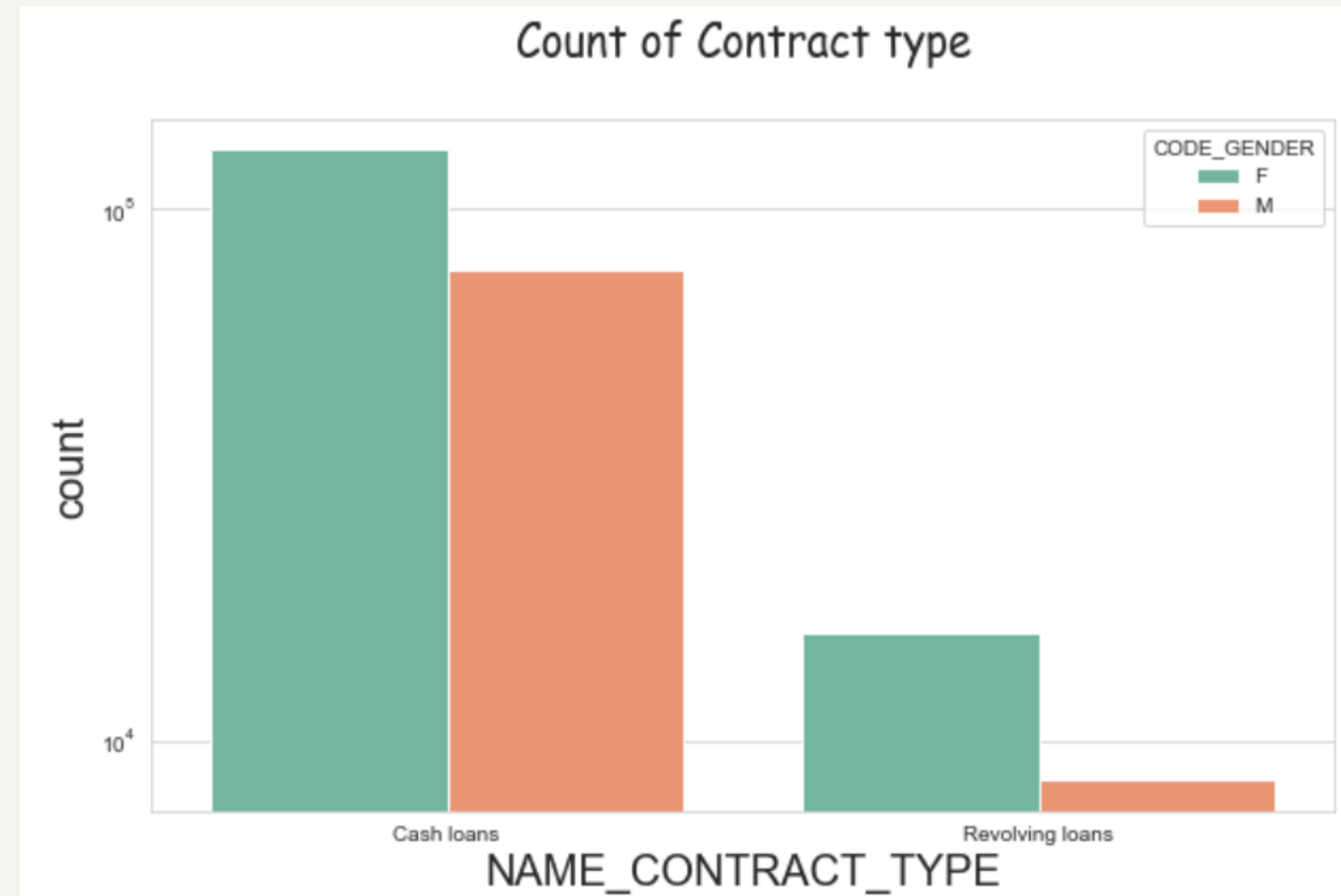
2. Females are having more number of credits than male.

3. 'student', 'pensioner', 'Businessman' and 'Maternity leave' have less credits.



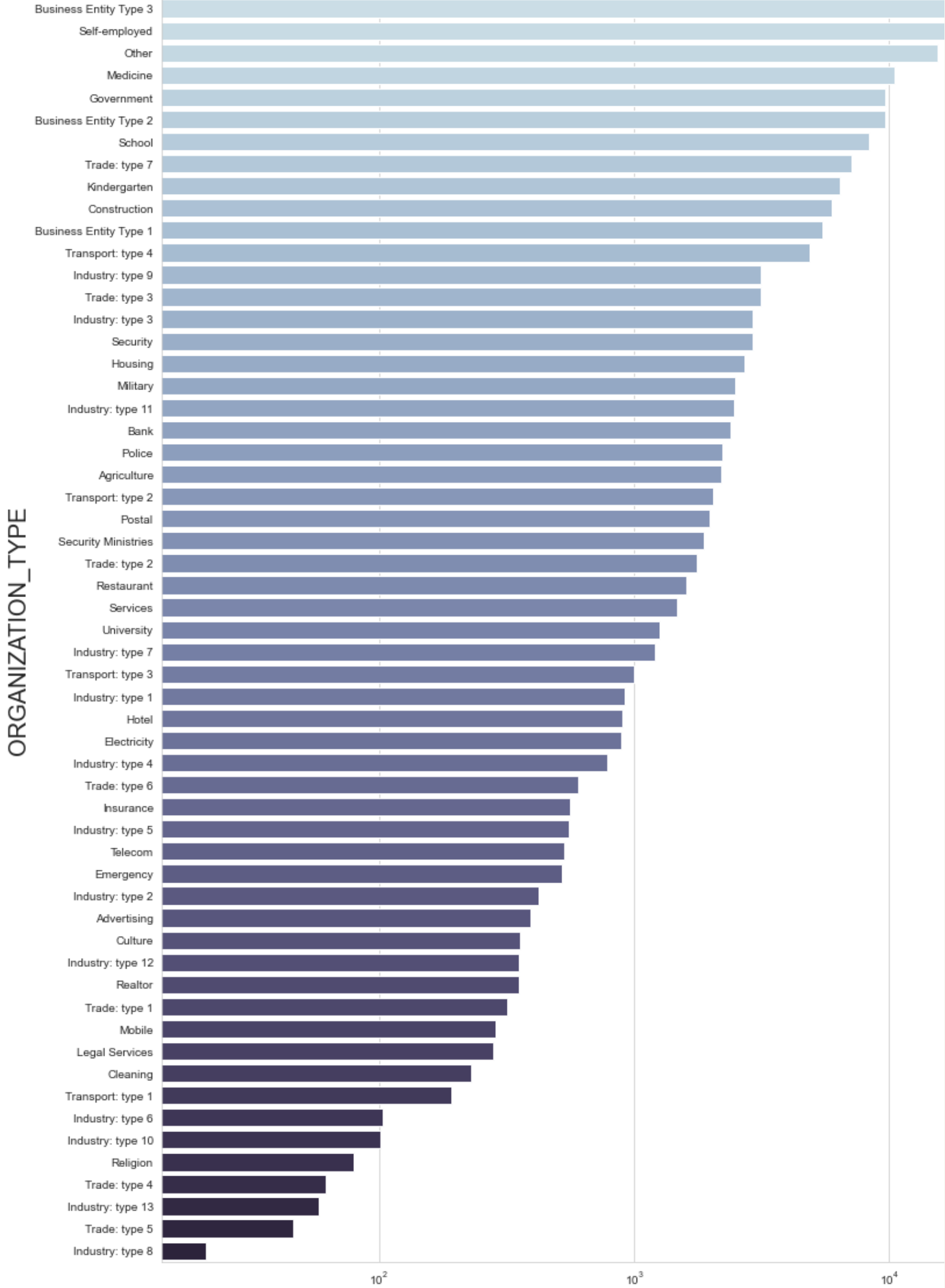
INFERENCES

1. Contract type 'cash loans' has higher number of credits than 'Revolving loans' contract type.
2. In both the contract types female count is higher.



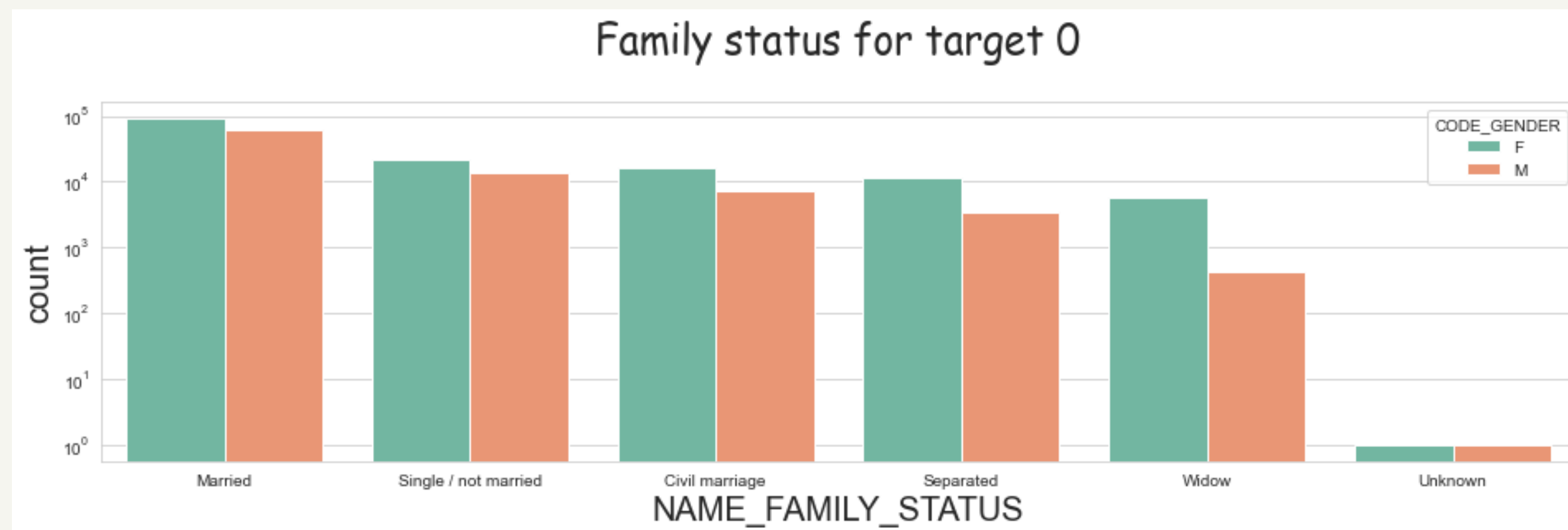
INFERENCES

1. Most of the clients who have applied for loan are mainly from the organization type ‘Business entity Type 3’ , ‘Self employed’, ‘Other’ , ‘Medicine’ and ‘Government’.
2. Very Few clients are from Industry type 8,type 6, type 10, religion and trade type 5, type 4.



INFERENCES

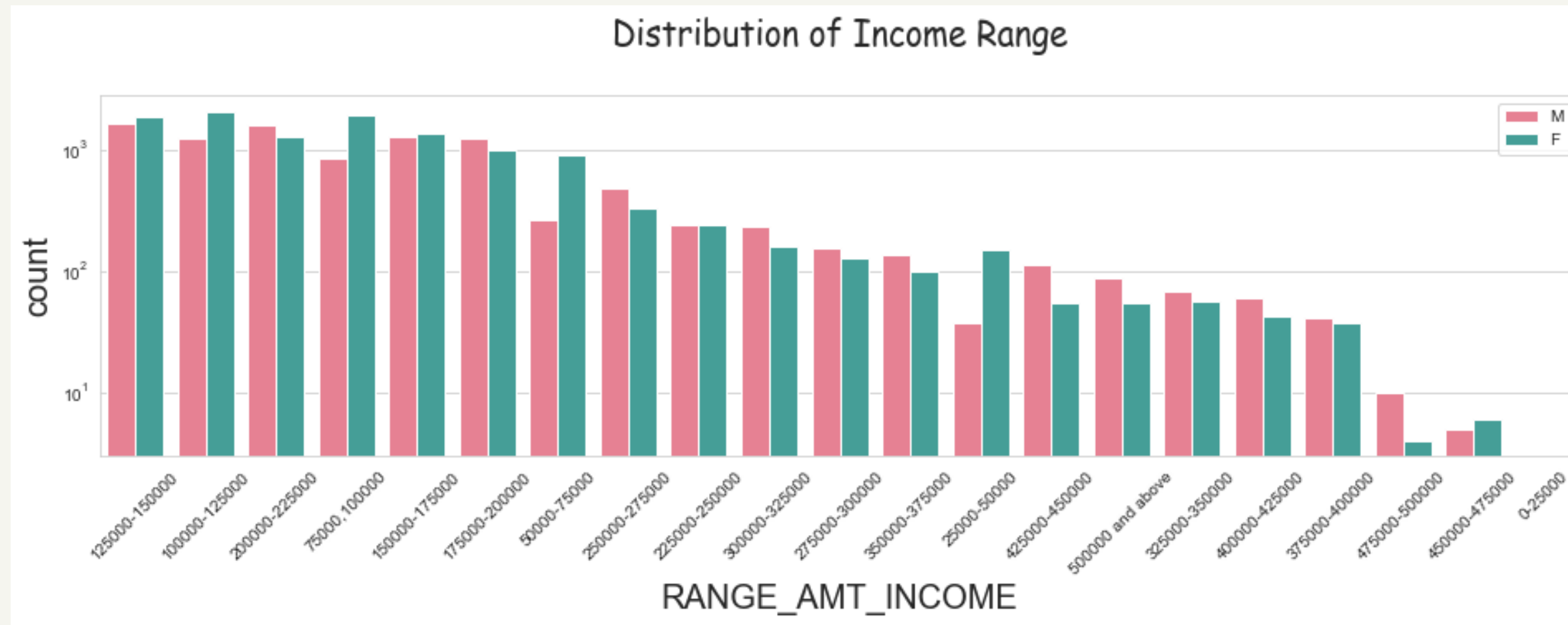
1. 'Married' customers have the maximum number of credits.
2. Females have more number of credits than male.
3. Less number of credits for status type 'widow' and 'unknowns'.



Univariate Analysis

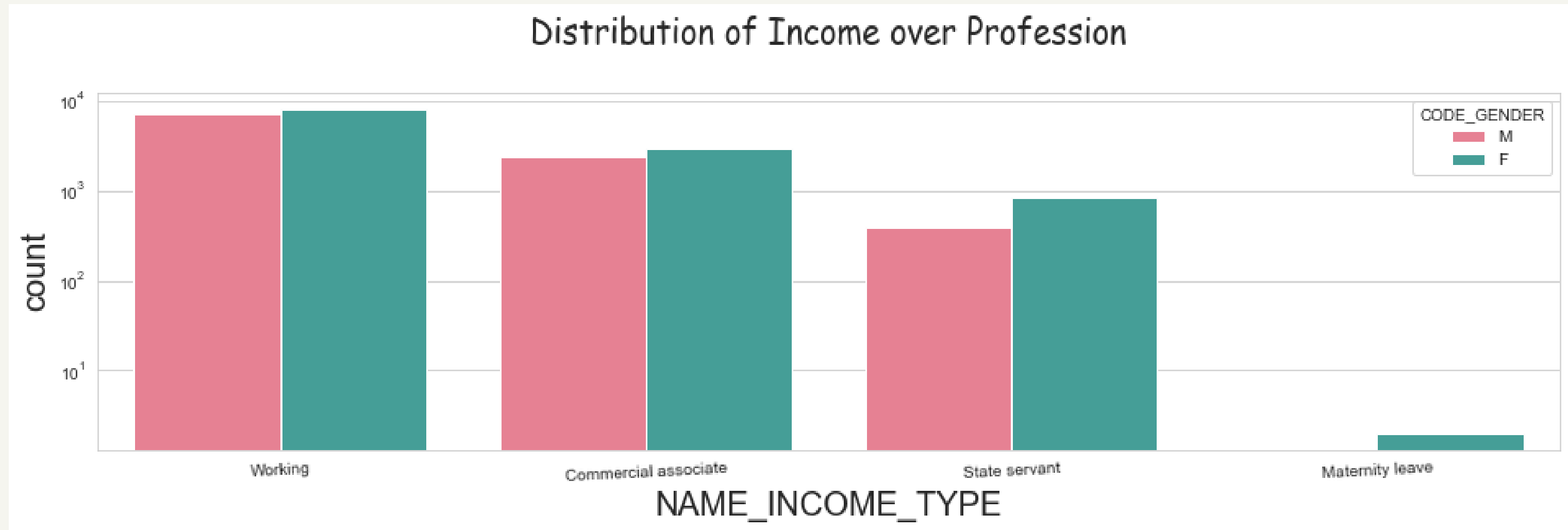
FOR CATEGORICAL
VARIABLES

FOR DEFAULT CLIENT
TARGET 1



INFERENCES

1. Female count is higher than male.
2. Income range from 100000 to 200000 is having more number of credits.
3. Very less count for income range 400000 and above.

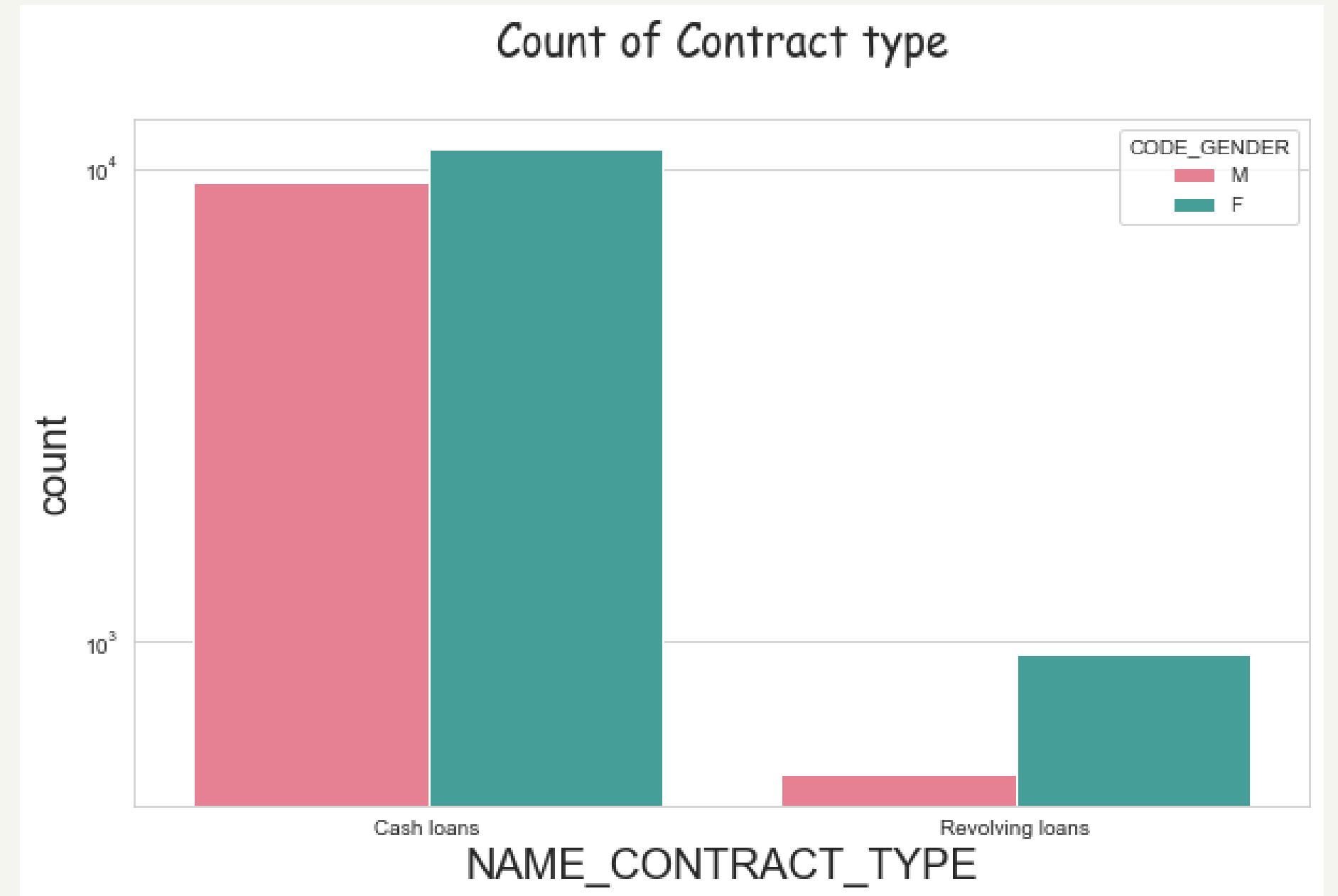


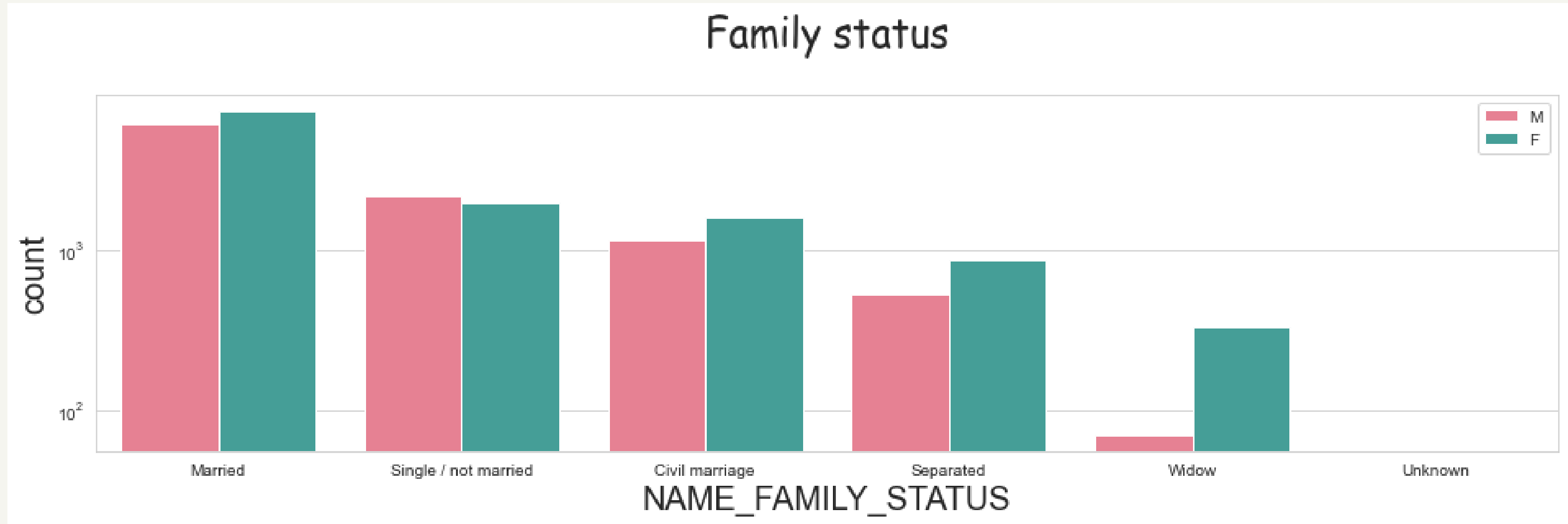
INFERENCES

1. 'working', 'commercial associate', and 'State Servant' professions have higher number of credits.
2. Females have high number of credits than male.
3. Less number of credits for income type 'Maternity leave'.

INFERENCES

1. Contract type 'cash loans' has higher number of credits than 'Revolving loans' contract type.
2. High number of female credits can be observed.



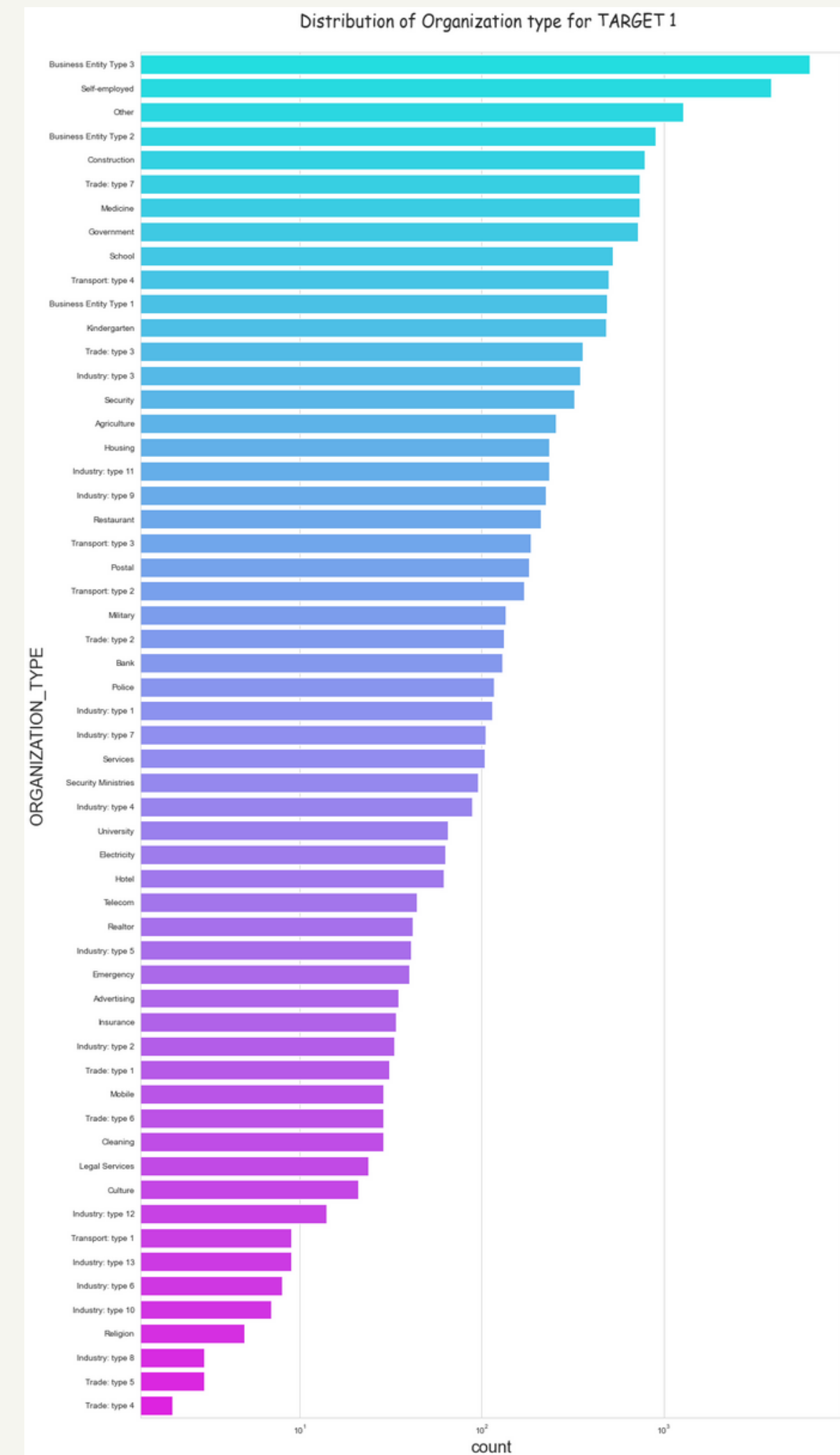


INFERENCES

1. Married clients have higher number of credits than others.
2. Less number of credits for status type 'widow' and 'unknowns'.

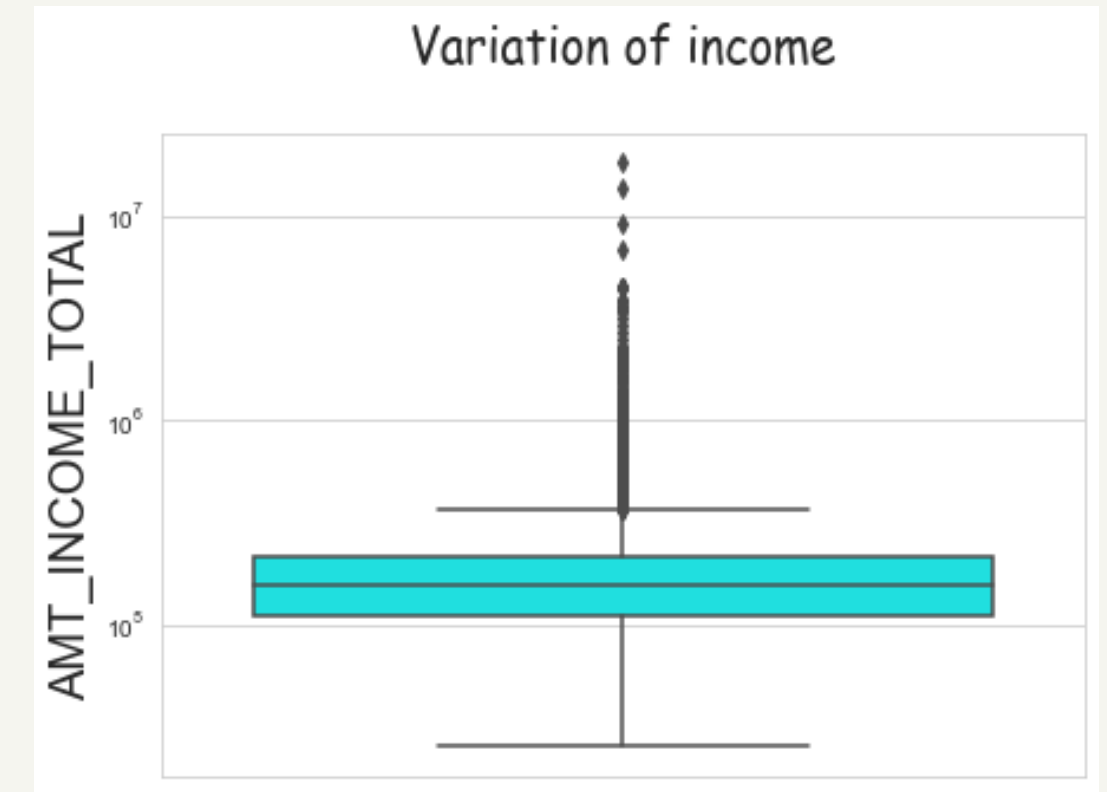
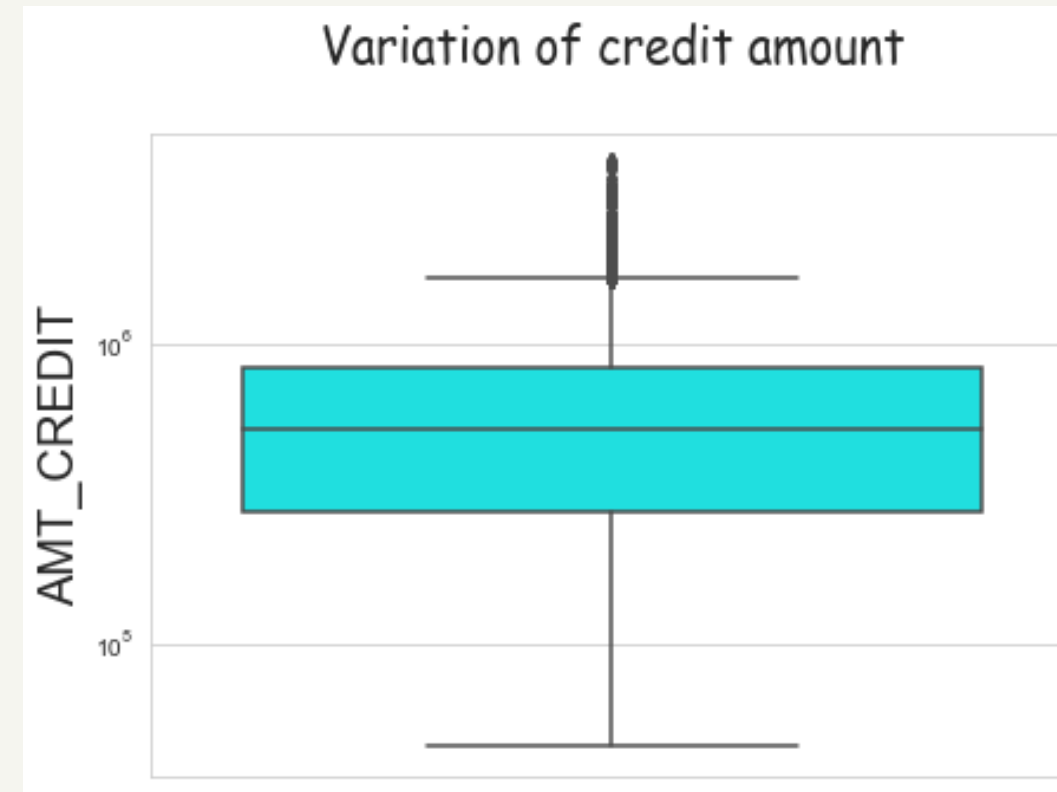
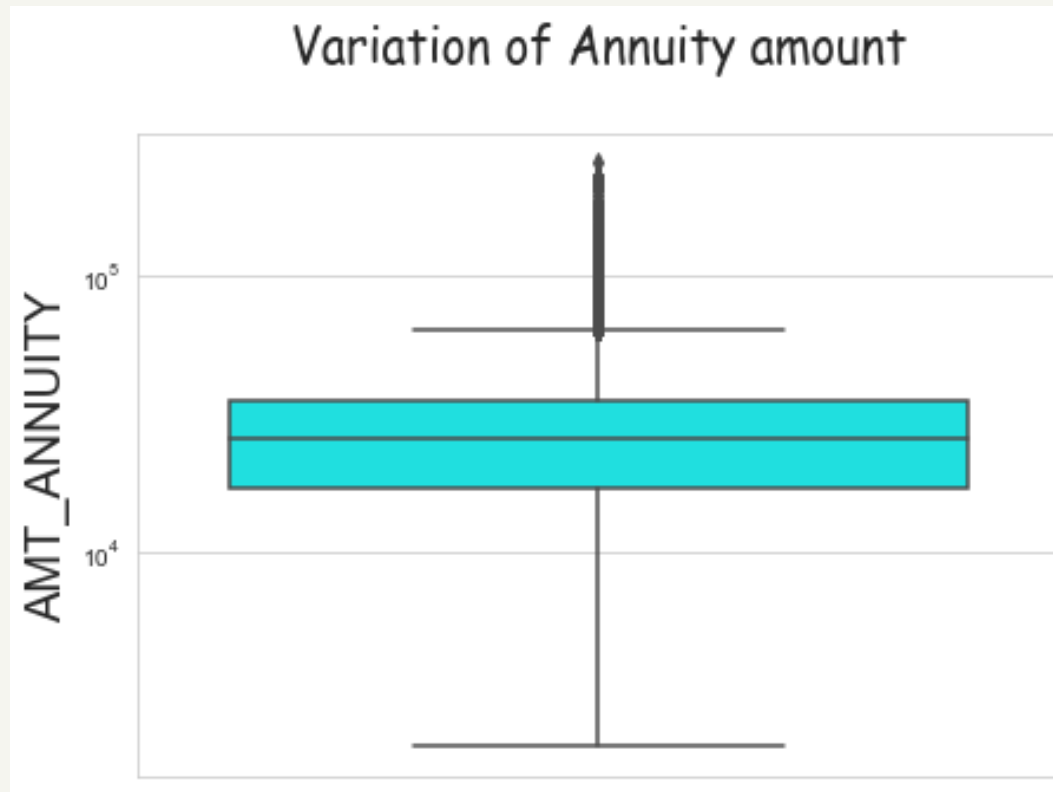
INFERENCES

1. Clients who have applied for credits are mostly from of the organization type 'Business entity Type 3', 'Self employed', 'Other', 'Medicine' and 'Government'.
2. Few clients are from Industry type 8,type 6, type 10, religion and trade type 5, type 4.
3. Same as target 0 in distribution of organization type.



Analysis for Outliers

FOR NON-DEFAULTING CLIENTS (TARGET 0)



INFERENCES

1. Outliers are present in annuity amount.

2. The first quartile is bigger than third quartile which means most of the annuity clients are from first quartile.

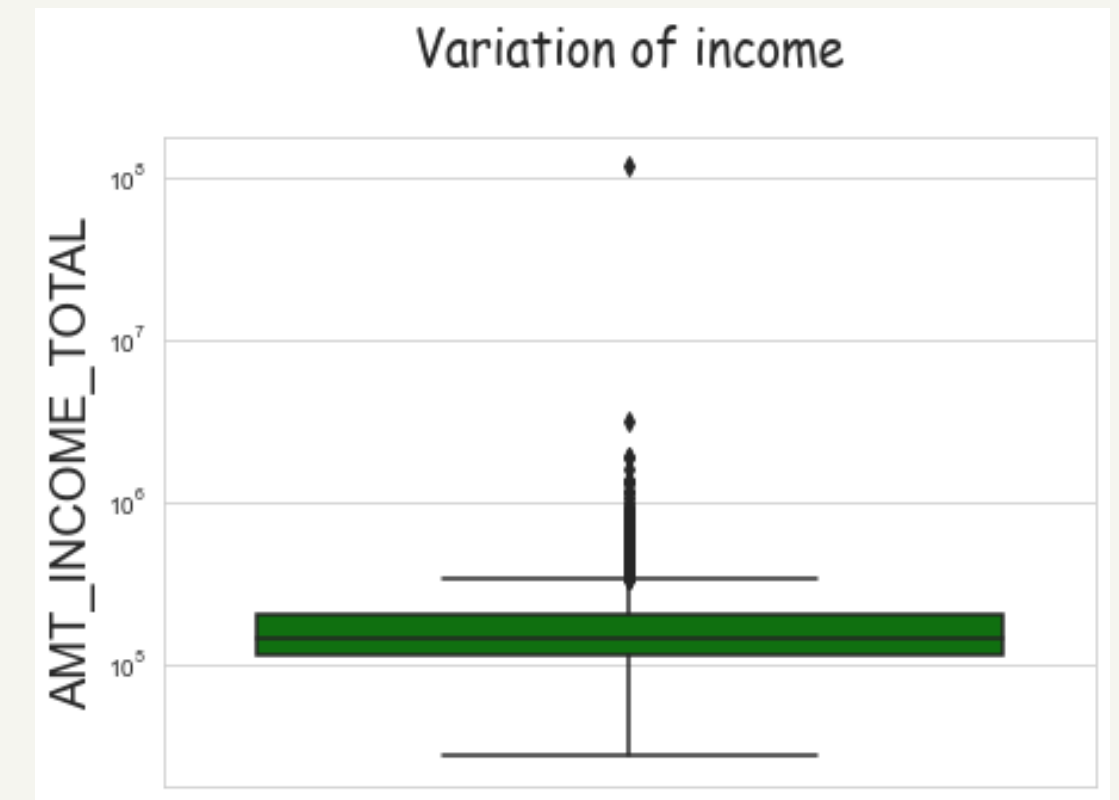
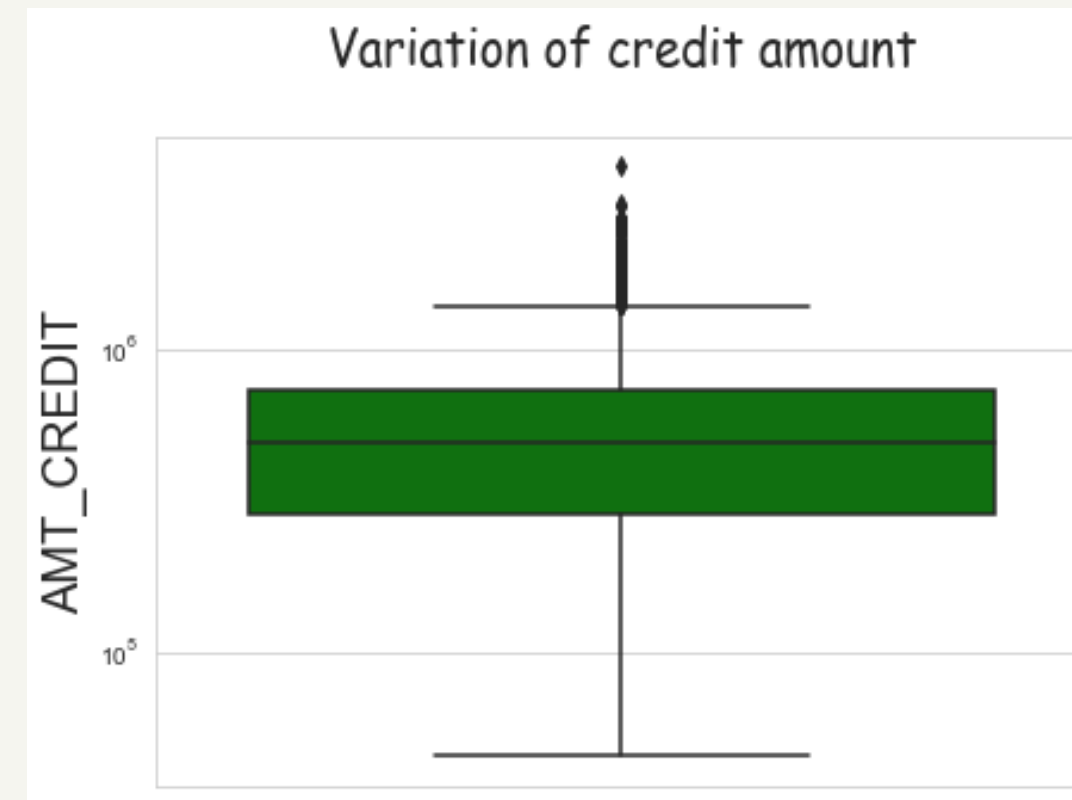
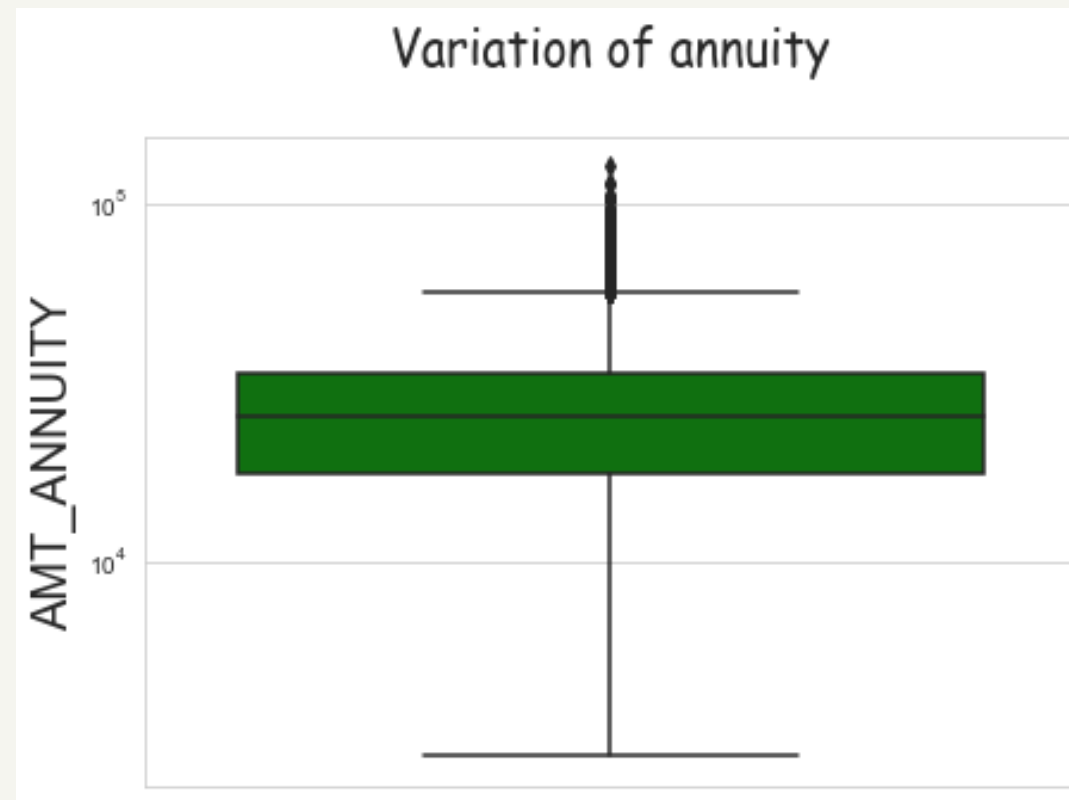
1. Outliers can be observed in credit amount.

2. The first quartile is bigger than third quartile which means most of the credits of clients are present in the first quartile

1. Significant outliers are noticed in income.

2. The third quartiles is very small for income amount.

FOR DEFAULTING CLIENTS (TARGET 1)



INFERENCES

1. Clustered outliers are noticed in annuity amount.

2. The first quartile is bigger than third quartile for which means most of the annuity clients are from first quartile.

1. Some outliers are noticed in credit amount.

2. The first quartile is bigger than third quartile which means most of the credits of clients are present in the first quartile.

1. The third quartiles is very small for income amount.

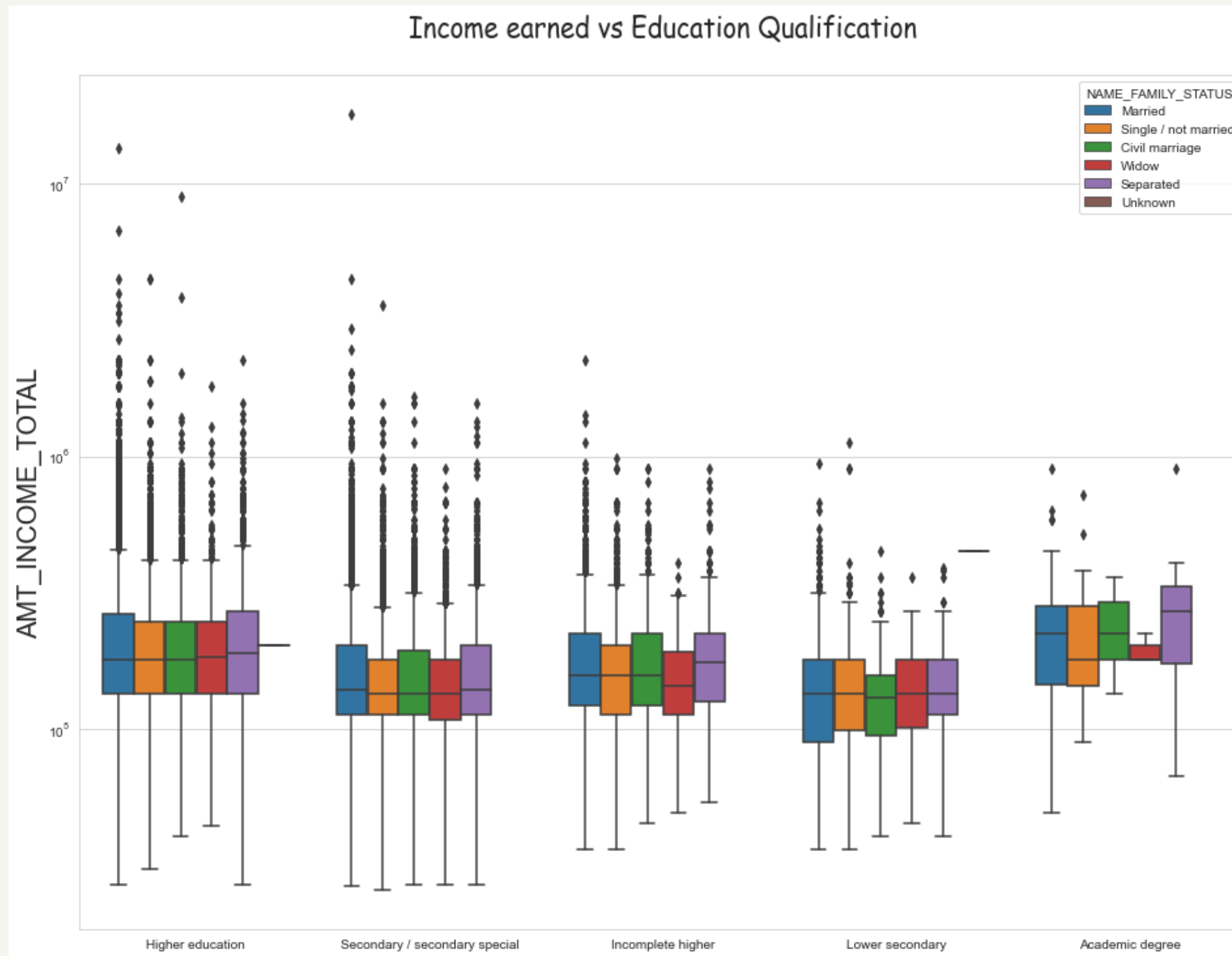
2. Far end outlier is observed in the box plot.

3. Most of the clients of income are present in first quartile.

Bivariate Analysis

FOR NON-DEFAULTING CLIENTS (TARGET 0)

INFERENCES

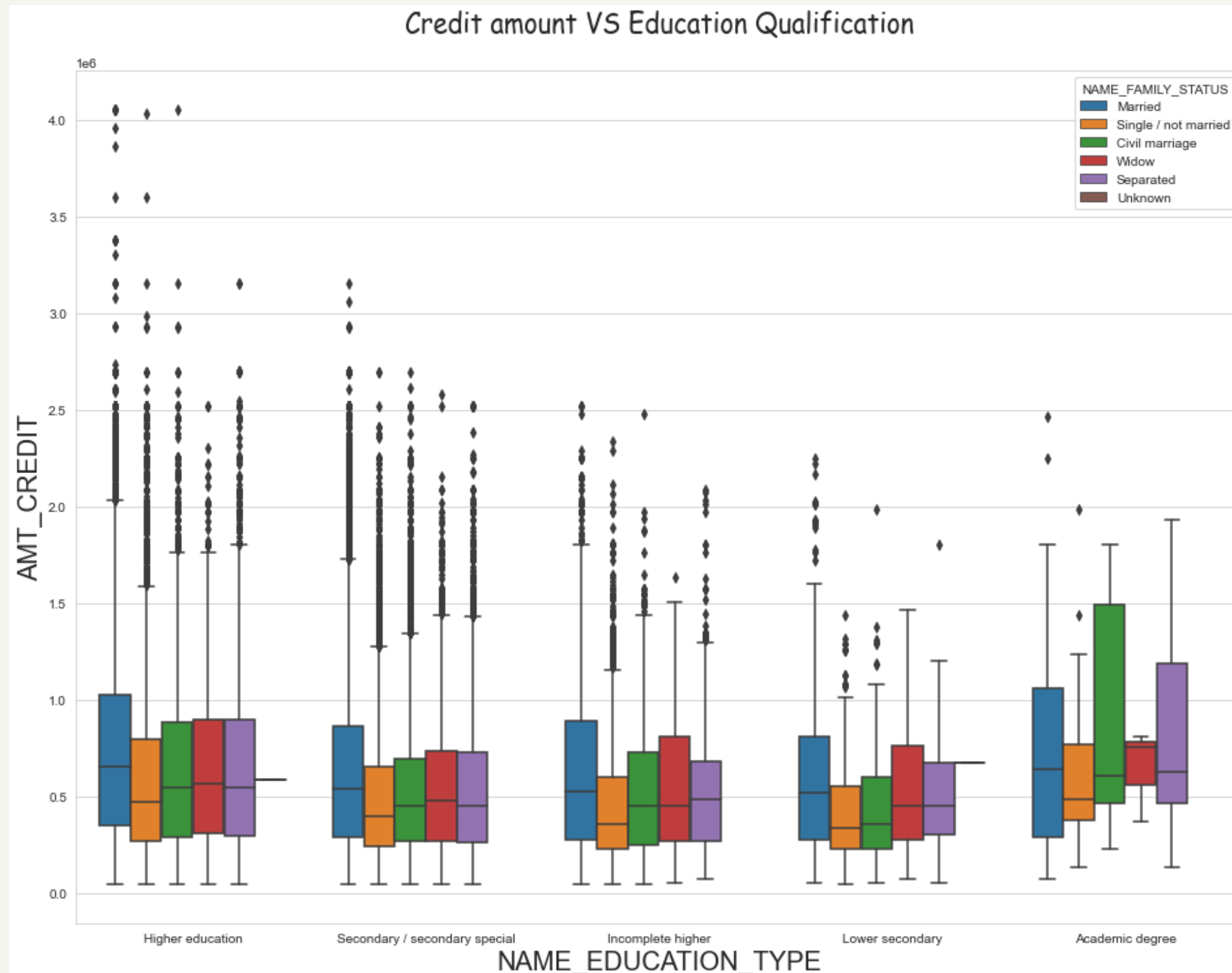


1. Education type 'Higher education' the income amount mean is mostly equal within all marital status. It contains many outliers.

2. Education type 'Academic degree', there are less outliers but they are having the income amount little higher than the Higher education.

3. Education type 'Lower secondary' of family status as 'civil marriage' have less income amount than the others.

FOR NON-DEFAULTING CLIENTS (TARGET 0) INFERENCES



1. Education type 'Higher education' the credit amount mean is different for the each marital status. It is lowest for single/unmarried people and highest for married people. It contains many outliers.

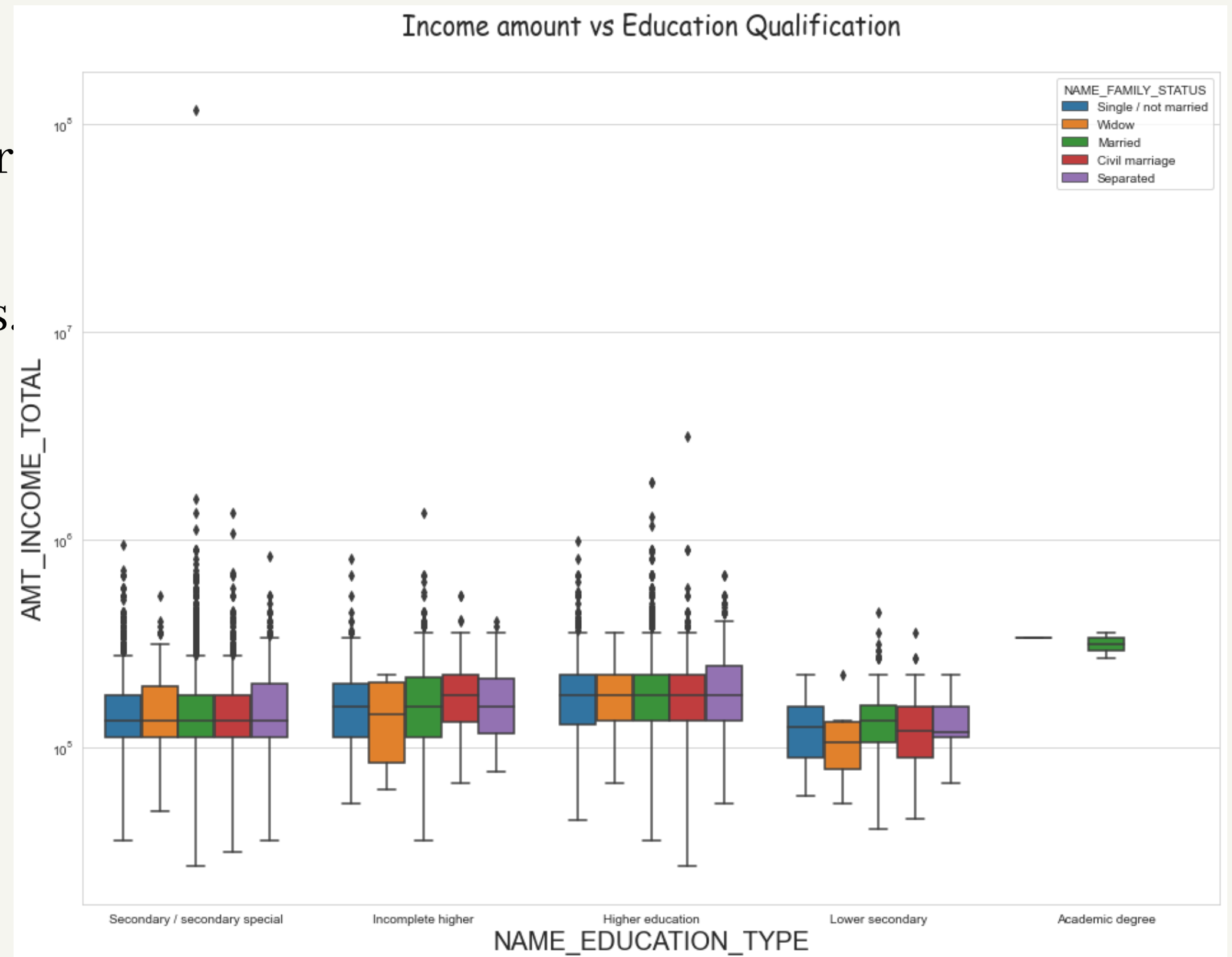
2. Clients with Education type 'Academic degree' have the credit amount little higher than the Higher education and there are less outliers.

3. Education type 'Lower secondary' of family status 'single/unmarried' are having less credit amount than the others.

FOR DEFAULTING CLIENTS (TARGET 1)

INFERENCES

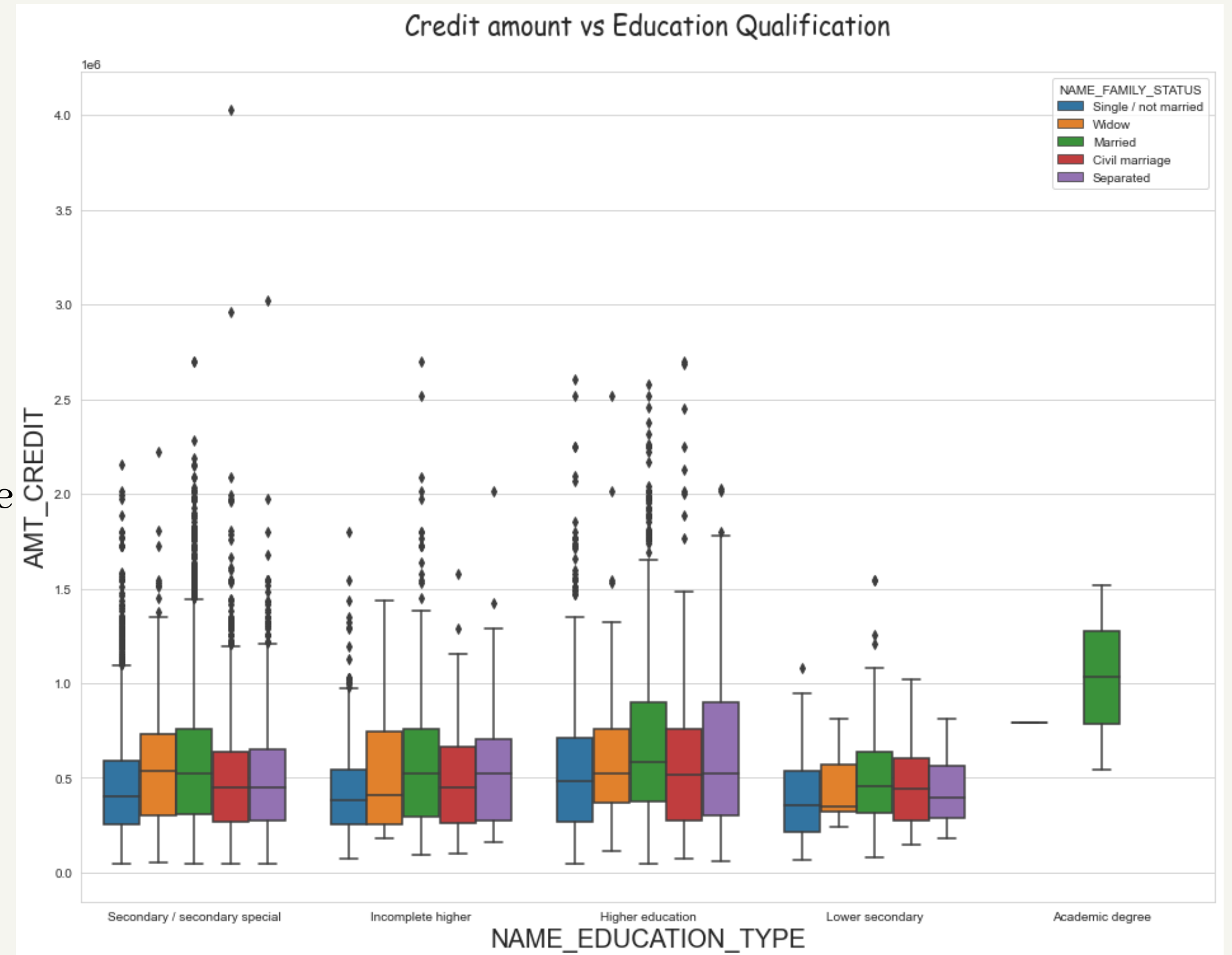
1. Clients with Education type 'Higher education' have similar income amount among all the family statuses.
2. Only few outliers can be observed for Academic Degree clients
3. Clients with Education type 'Lower secondary' have less income amount than the others.



FOR DEFAULTING CLIENTS (TARGET 1)

INFERENCE

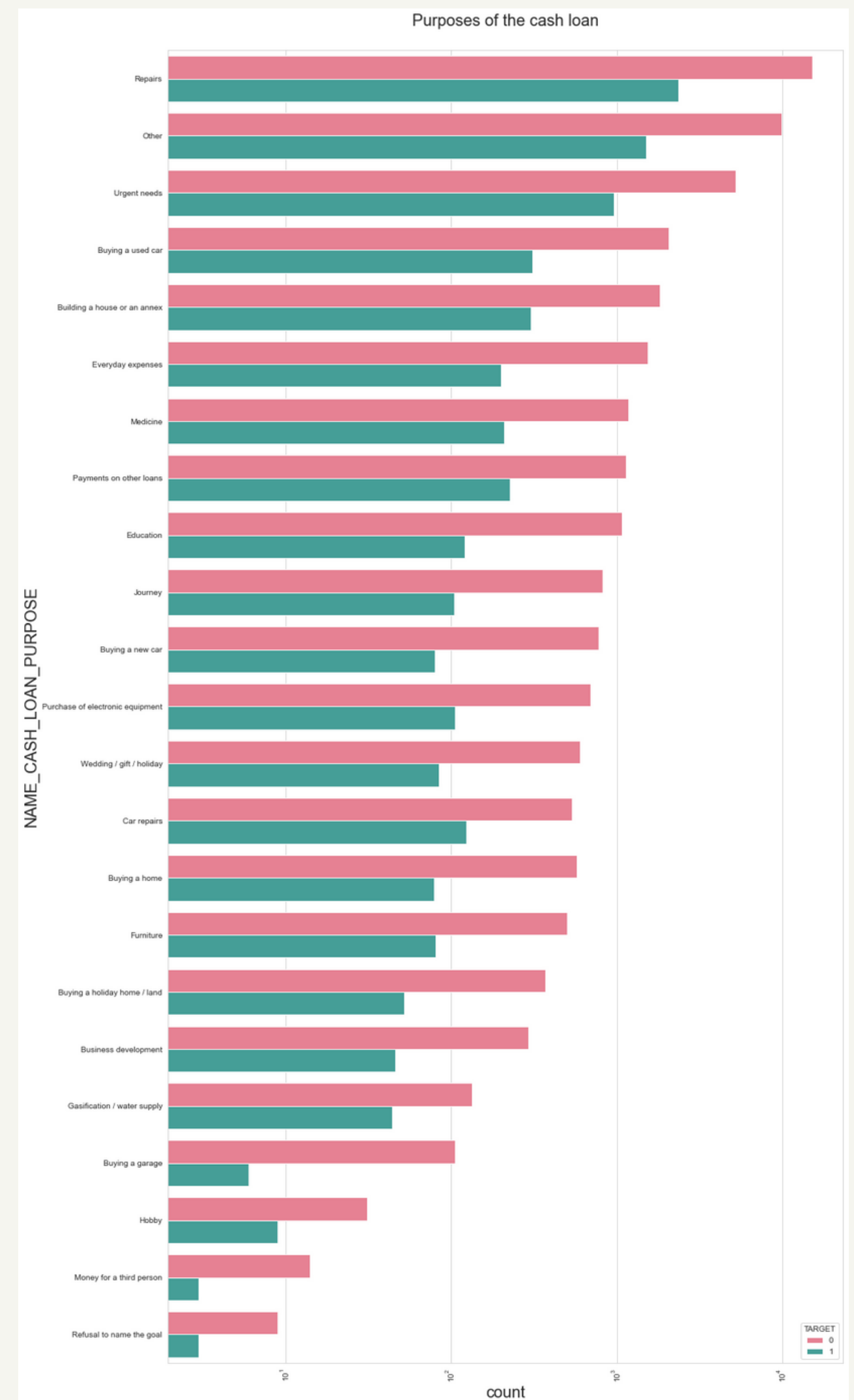
1. Education type 'Academic degree' with family status as 'married' people are having higher credit amount as compared to other family status people and also to the other education type people.
2. In the Education type 'Higher education', the family status of 'Married' have higher credit amount than the other family statuses.
3. Very few outliers present in the education type 'Lower Secondary'.
4. The family status 'Widow' and 'Separated' are having very few outliers for the education type 'Higher education'.



Analysis of Merged Dataset

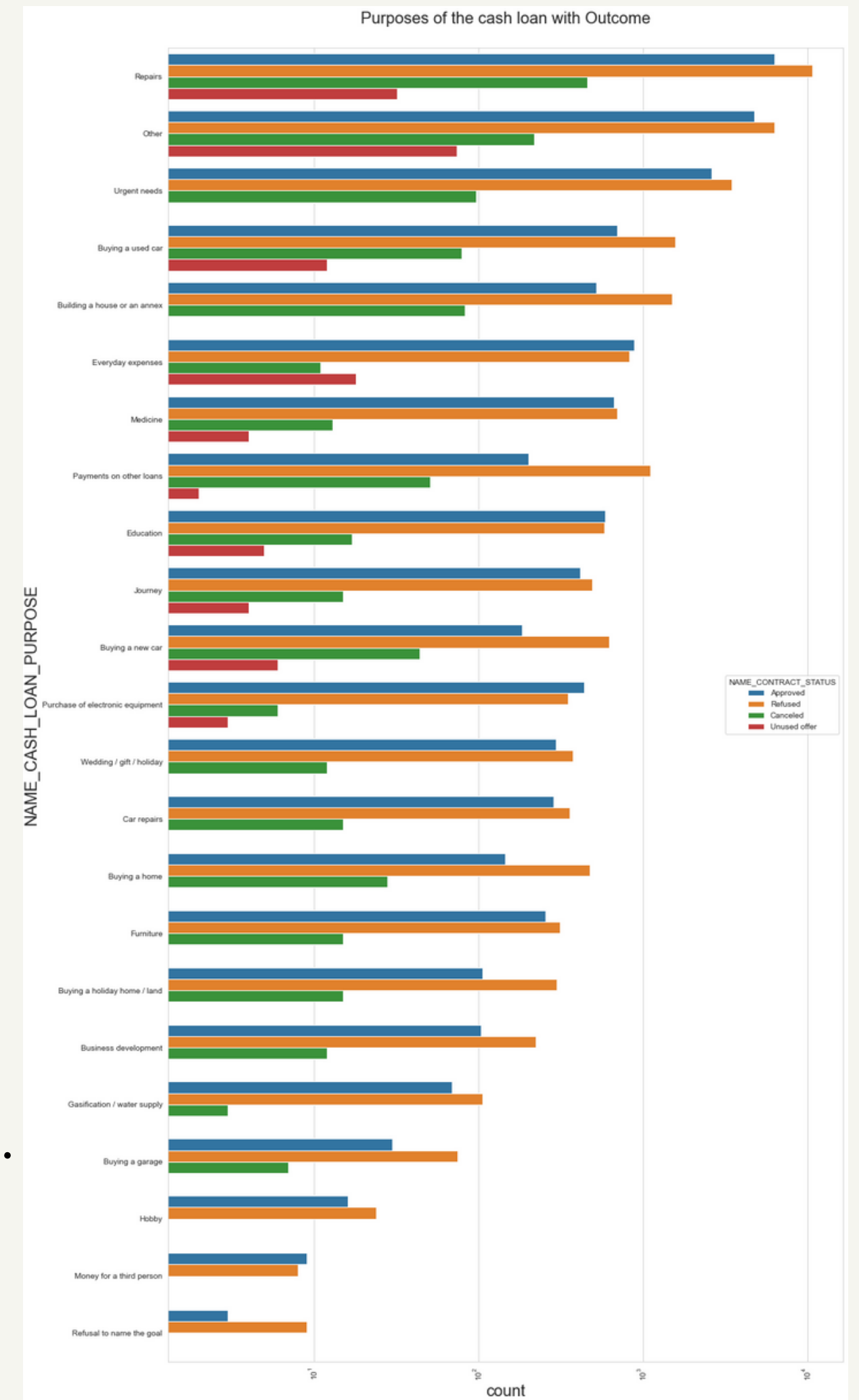
INFERENCES

1. In general, there are fewer non-default clients as compared to defaulting clients.
2. Most rejection of loans comes from 'repairs' requests.
3. For education purposes we see an equal number of approval and rejection.

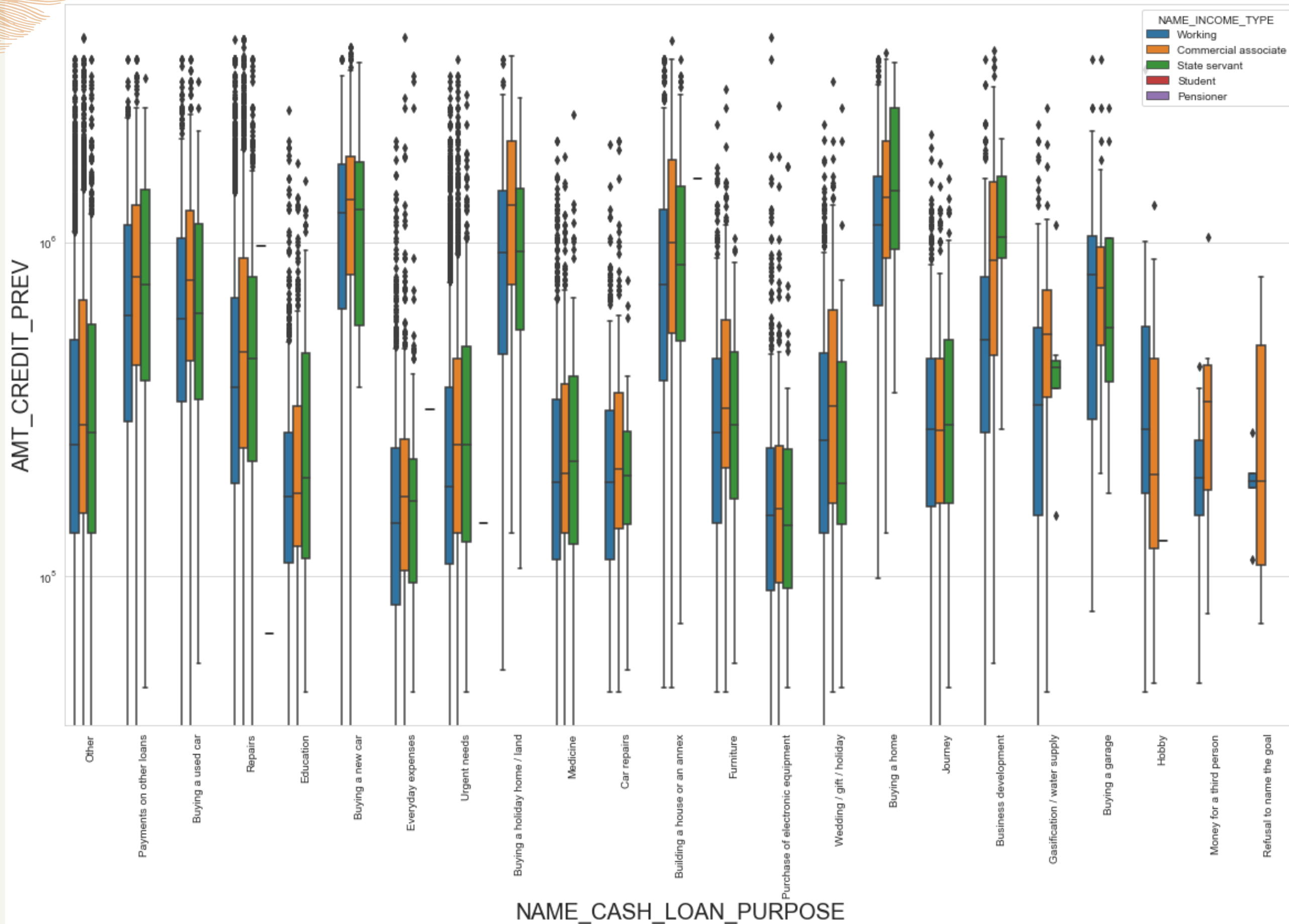


INFERENCES

1. Loan purposes with 'Repairs' are facing more difficulties in payment on time.
2. There are more number of rejections as compared to accepted loans.
3. They seems to be risk in repairs as it has highest number of refusals, cancelled and unused loan amounts.



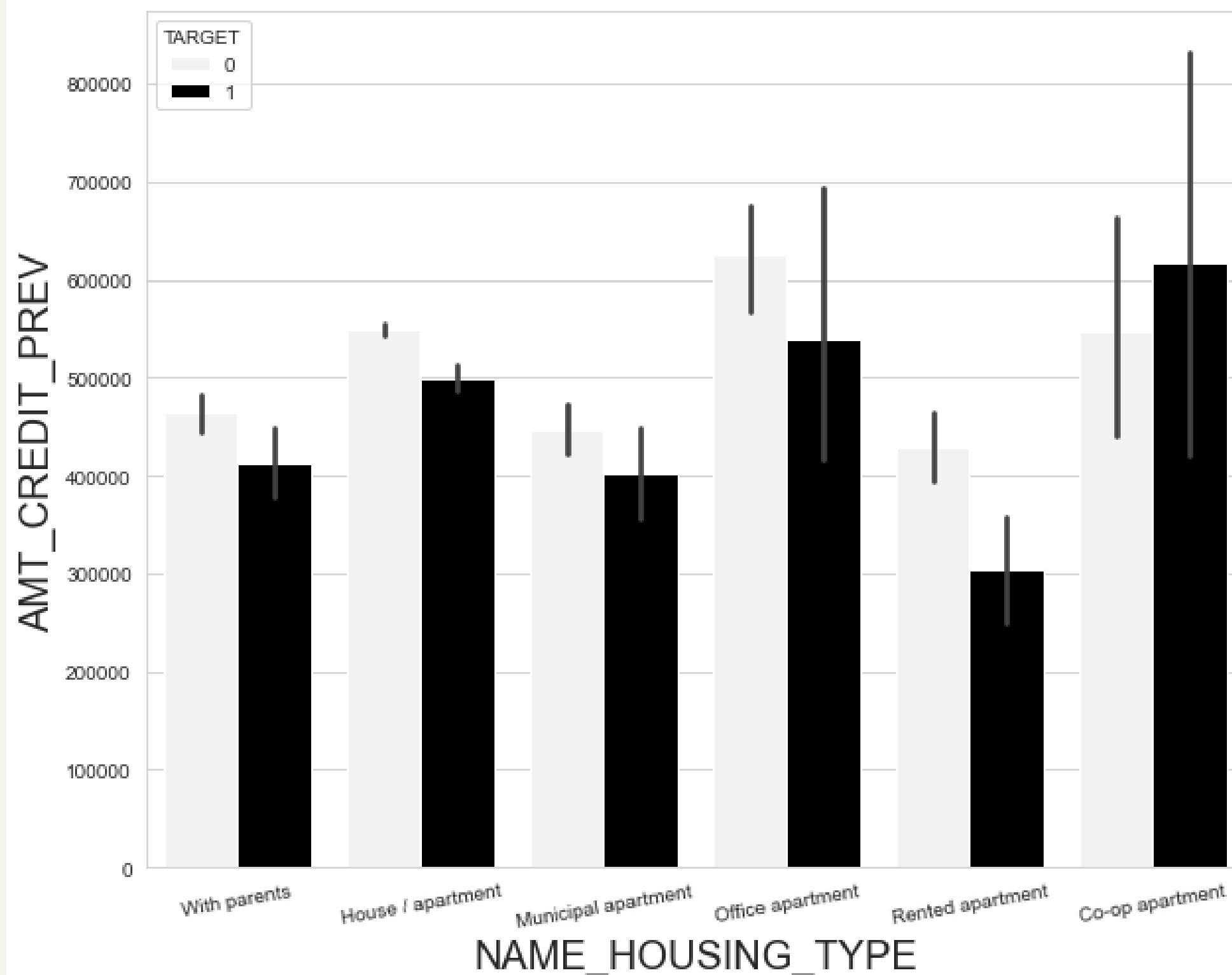
Prev Credit amount vs Loan Purpose



INFERENCES

1. Loan purposes like 'Buying a home', 'Buying a land', 'Buying a new car' and 'Building a house' have high number of credits.
2. Working profession have less amount of credit applied.
3. Less number of credits have been applied for the cash loan purposes personal everyday usage.

Previous Credit amount vs Housing type



INFERENCES

1. Office apartment is having higher credit for target 0.
2. Co-op apartment is having higher credit for target 1.
3. High variation can be observed in defaulting clients in office apartment and co-op apartment.

Recommendations

The company should be cautious of the following flags while providing loans to customers if:

- ~ Applicant's employment experience is less than 10 years.*
- ~ Applicant falls in the age group of 21-30.*
- ~ Applicant earns in the income bracket of 25K to 1.125K.*
- ~ Education type is Lower secondary or Secondary/Secondary .*
- ~ The applicant is a Low Skilled Labourer.*
- ~ Collaterals with value in between 40K to 4.5 Lakhs.*

Thank You!

