# CREDIT EDA CASE STUDY

BY SANTHOSH GUDUR

CATEGORICAL UNIVARIATE ANALYSIS FOR TARGET 0

# DISTRIBUTION OF INCOME RANGE

Observations from the graph:

• Female counts are higher than male.

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

• Females are more than male in having credits for the range 100000 to 200000

• Count is very less for income range 400000 and above.
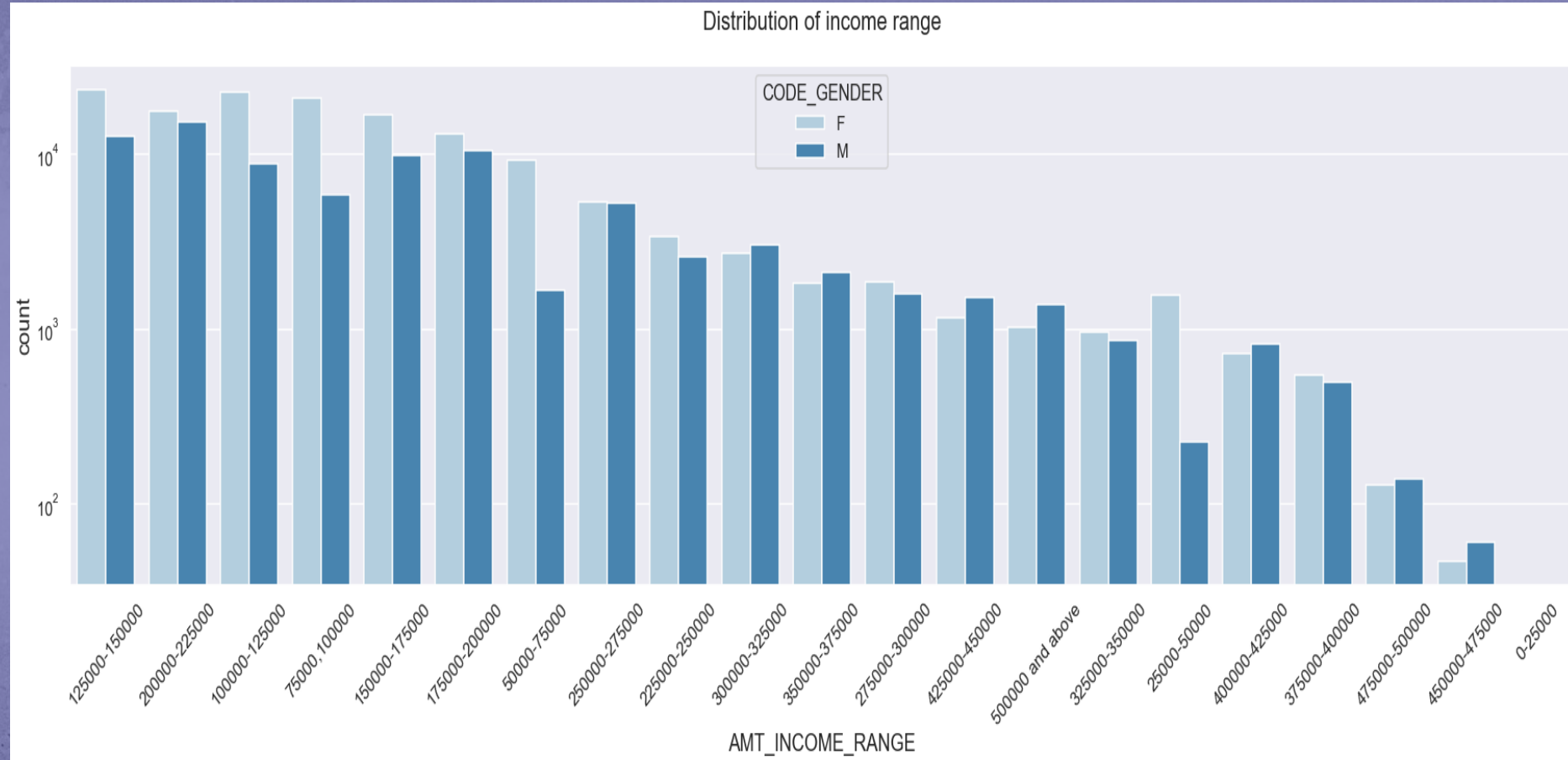
Distribution of income range

DISTRIBUTION OF INCOME TYPE

Observations from the graph:

- For income type 'working', 'commercial associate', and 'State Servant' the number of credits are higher than others.

- For income type 'working', 'commercial associate', and 'State Servant', Females are having more number of credits than male.

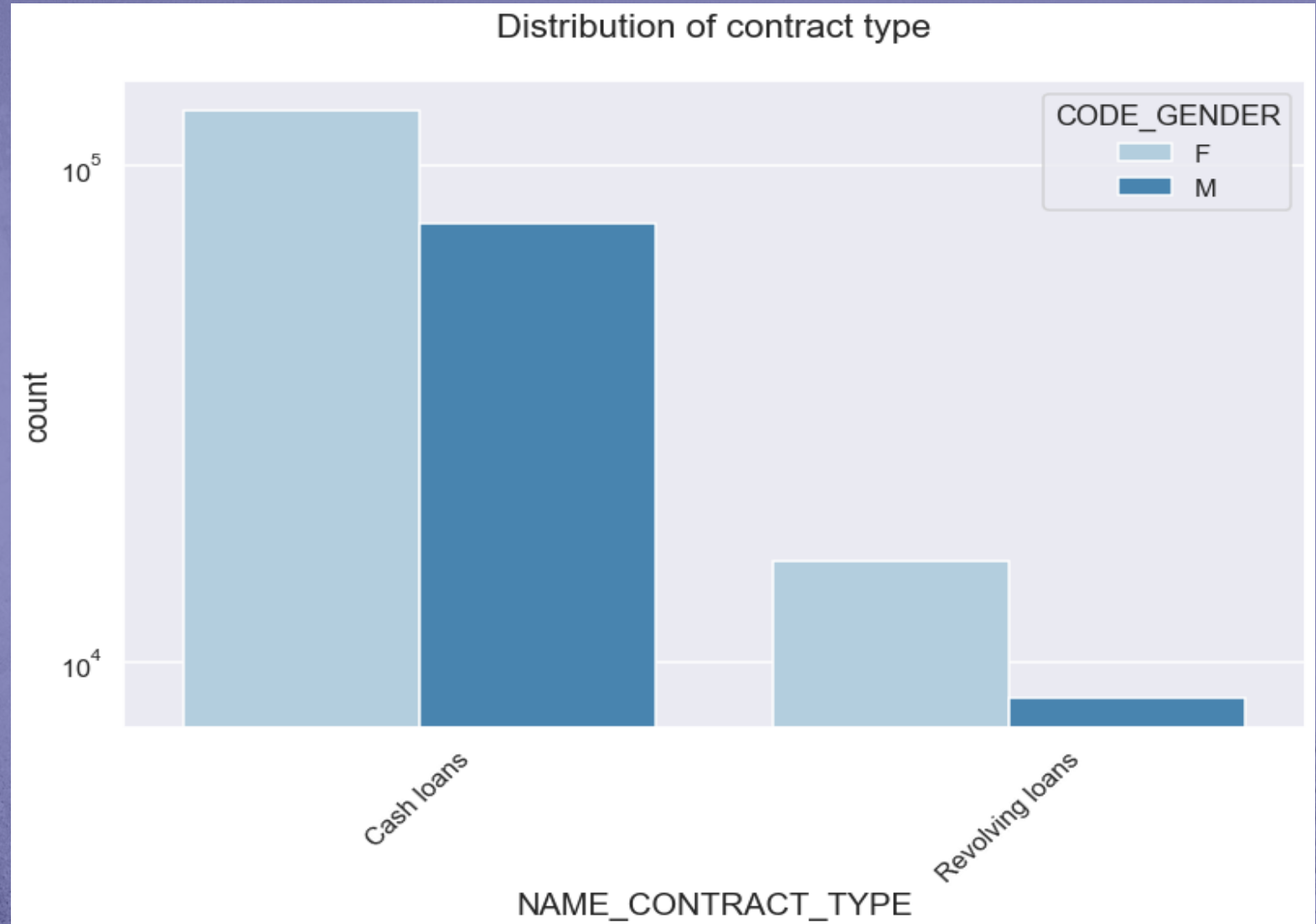- Credits are less for income type 'student' ,'pensioner', 'Businessman' and 'Maternity leave'.
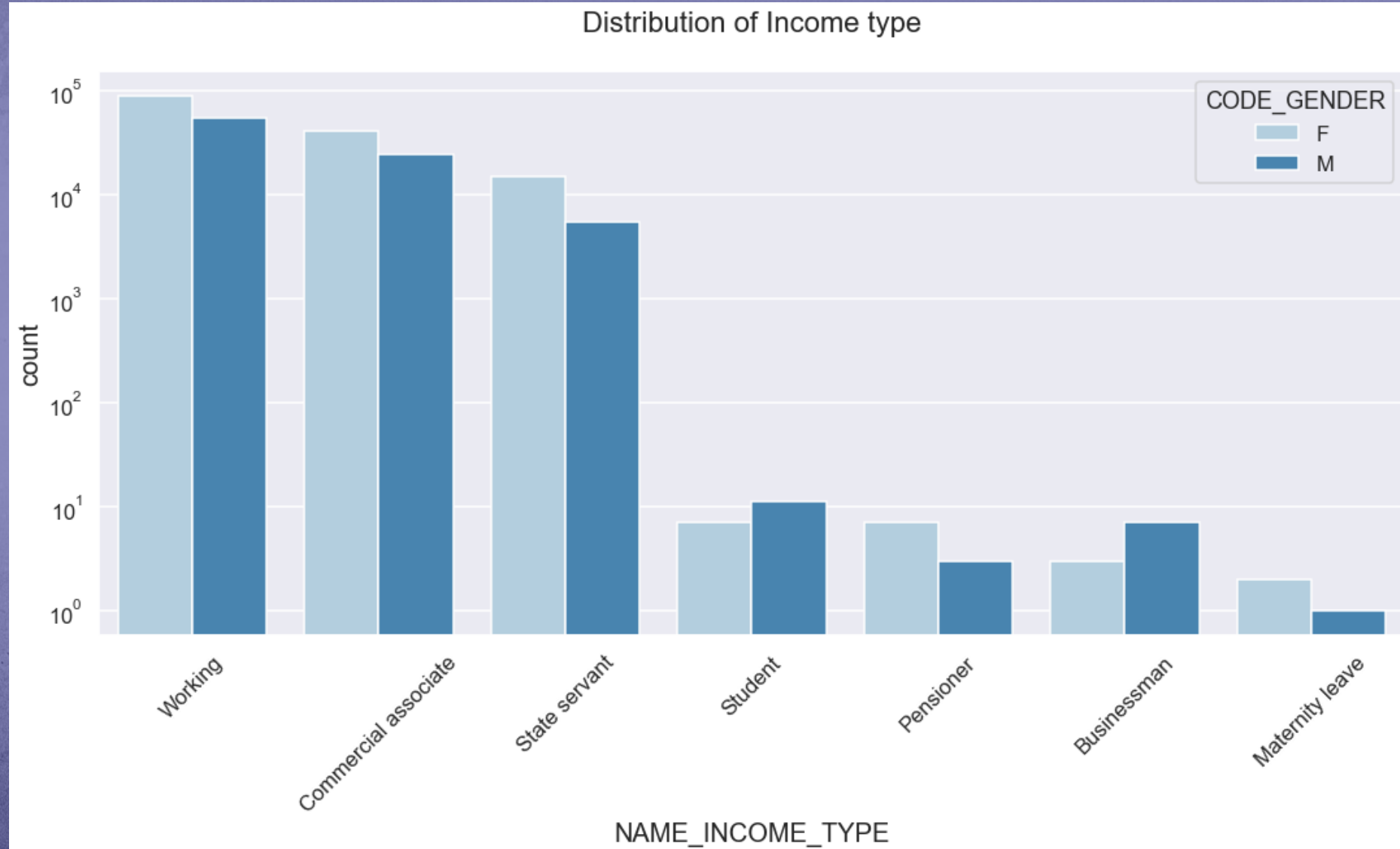
# DISTRIBUTION OF CONTRACT TYPE

Observations from the graph:

• For contract type 'cash loans' is having higher number of credits than 'Revolving loans' contract type.

• For this also Female is leading for applying credits.

DISTRIBUTION OF ORGANIZATION TYPE

Observations from the graph:

• Clients which have applied for credits are from most of the organization type 'Business entity Type 3' , 'Self employed', 'Other' , 'Medicine' and 'Government'.

• Less clients are from Industry type 8,type 6, type 10, type 13, religion and trade type 5, type 4.
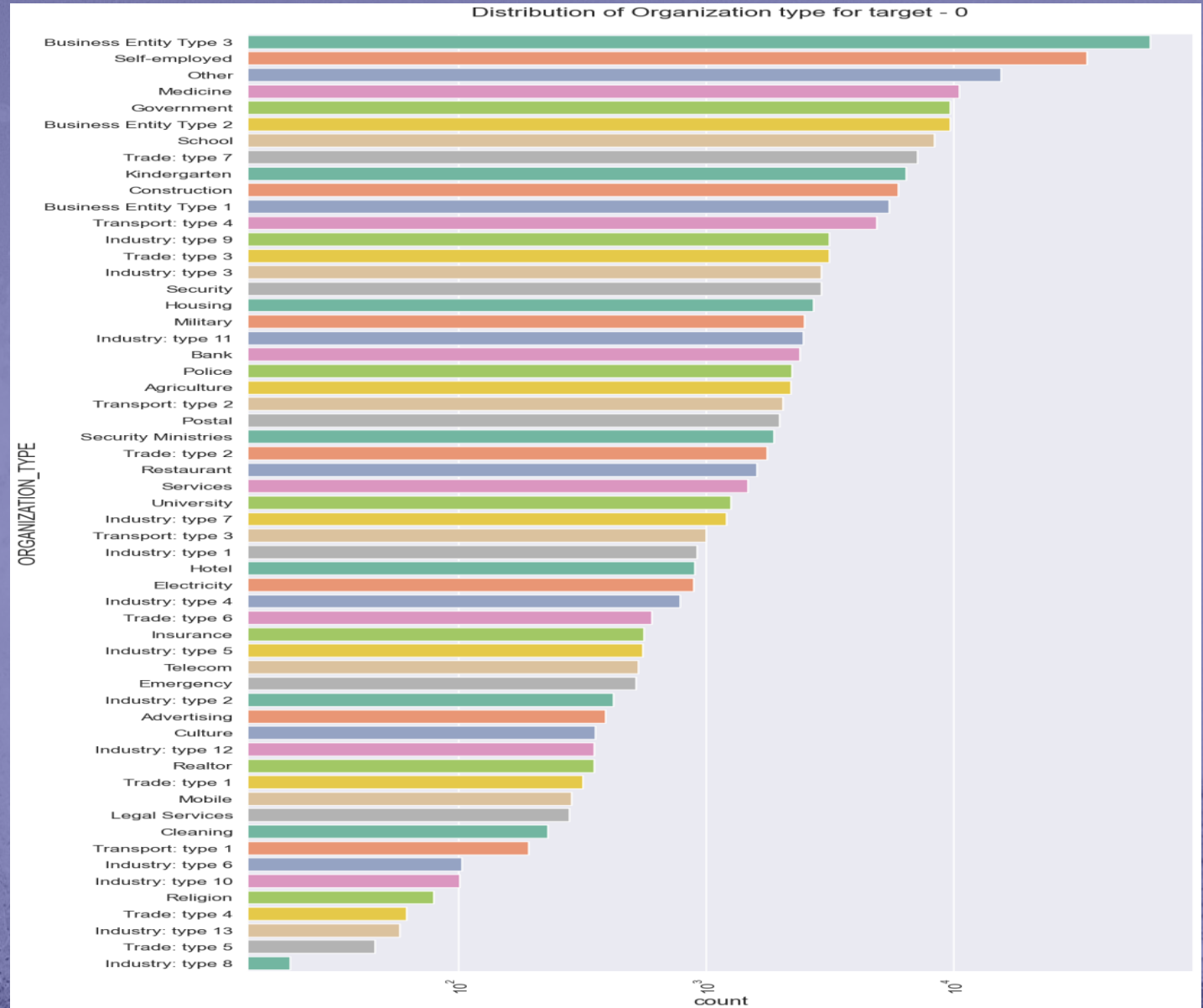
Distribution of Organization type for target - 0
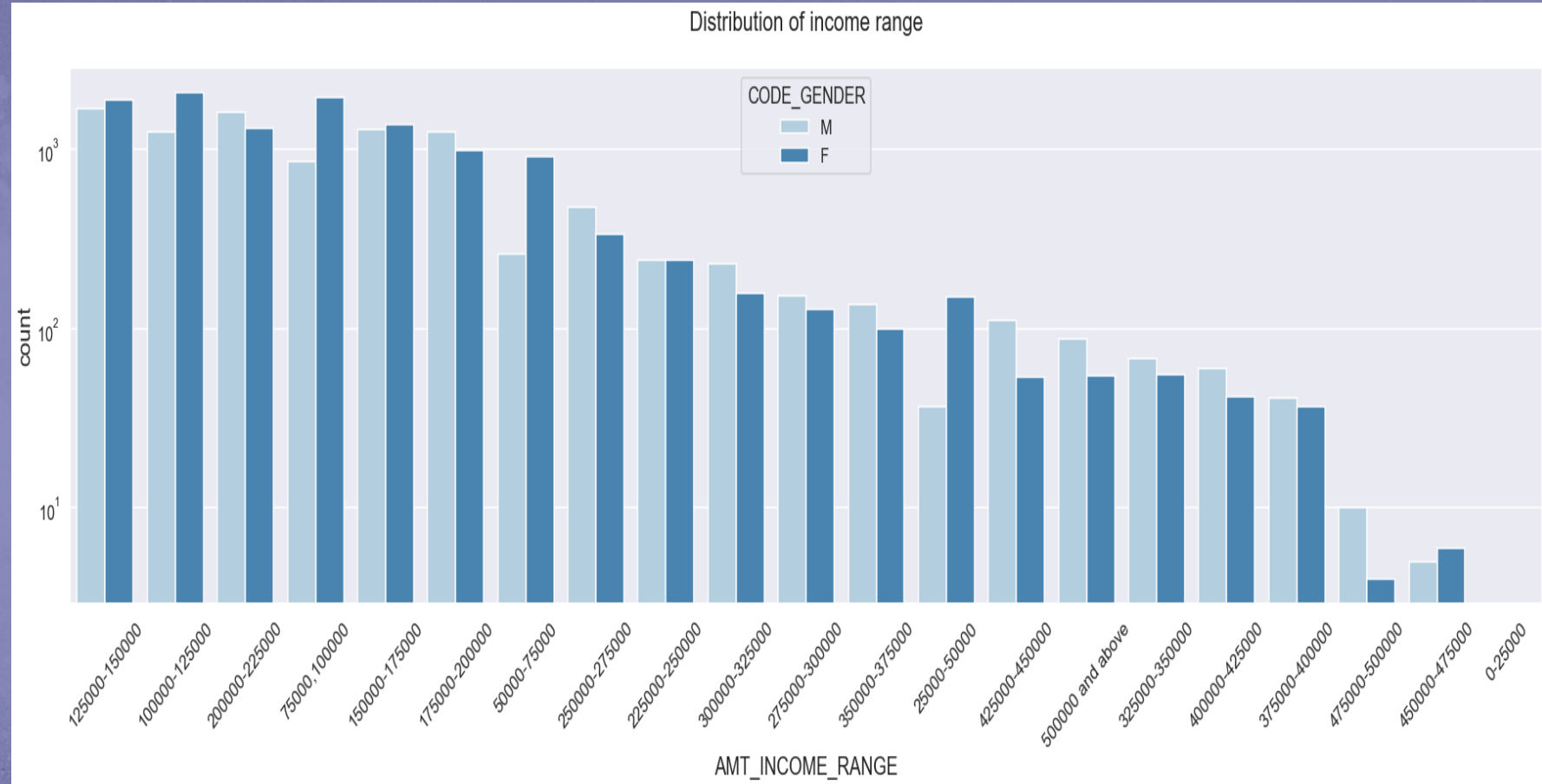
CATEGORICAL UNIVARIATE ANALYSIS FOR TARGET 1

DISTRIBUTION OF INCOME RANGE

Observations from the graph:

• Male counts are higher than female.

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

• This graph show that males are more than female in having credits for that range.

• Very less count for income range 400000 and above.


Distribution of income range

DISTRIBUTION OF INCOME TYPE

Observations from the graph:

• 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 are 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.
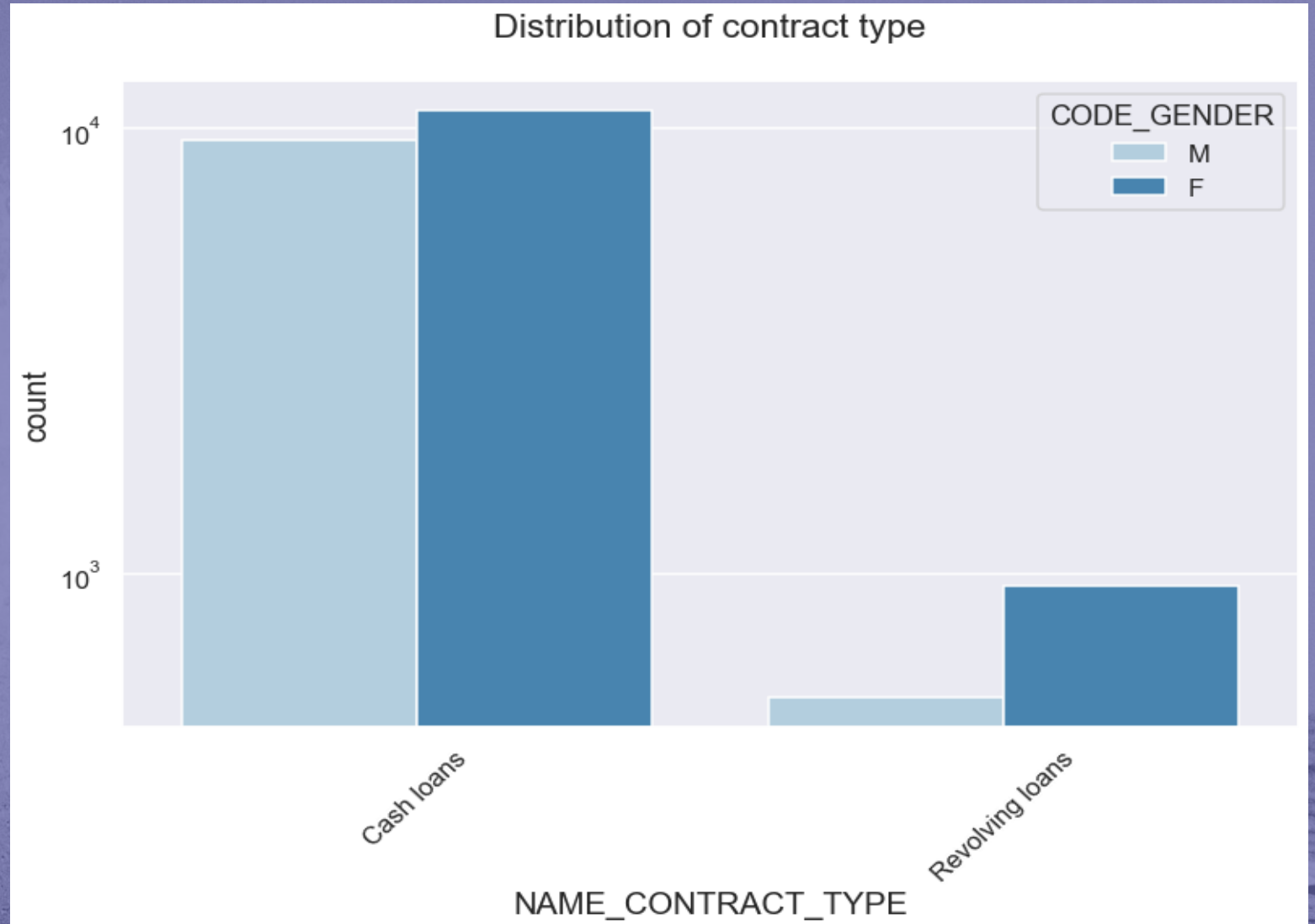

Distribution of Income type
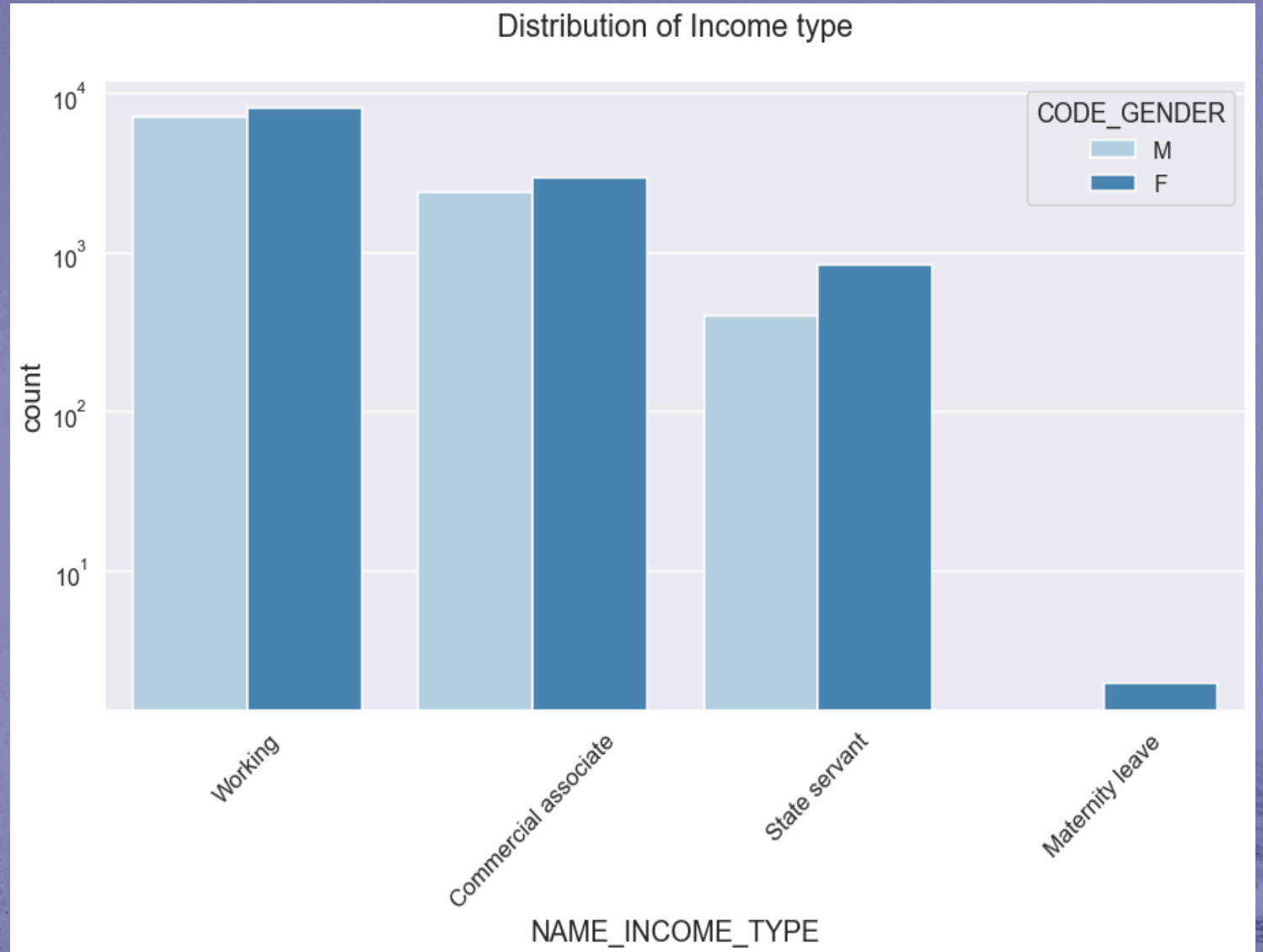
DISTRIBUTION OF CONTRACT TYPE

Observations from the above graph:

• For contract type, 'cash loans' is having higher number of credits than 'Revolving loans' contract type.

• Female is leading for applying credits.

• For type 1: there is only Female Revolving loans.

DISTRIBUTION OF ORGANIZATION TYPE

Observations from the graph:

• Clients which have applied for credits are from most of the organization type 'Business entity Type 3', 'Self employed', 'Other', and 'Business entity Type 2'.

• Less clients are from Industry type 8,type 6, type 10, religion and trade type 5, type 4.


Distribution of Organization type for target - 1
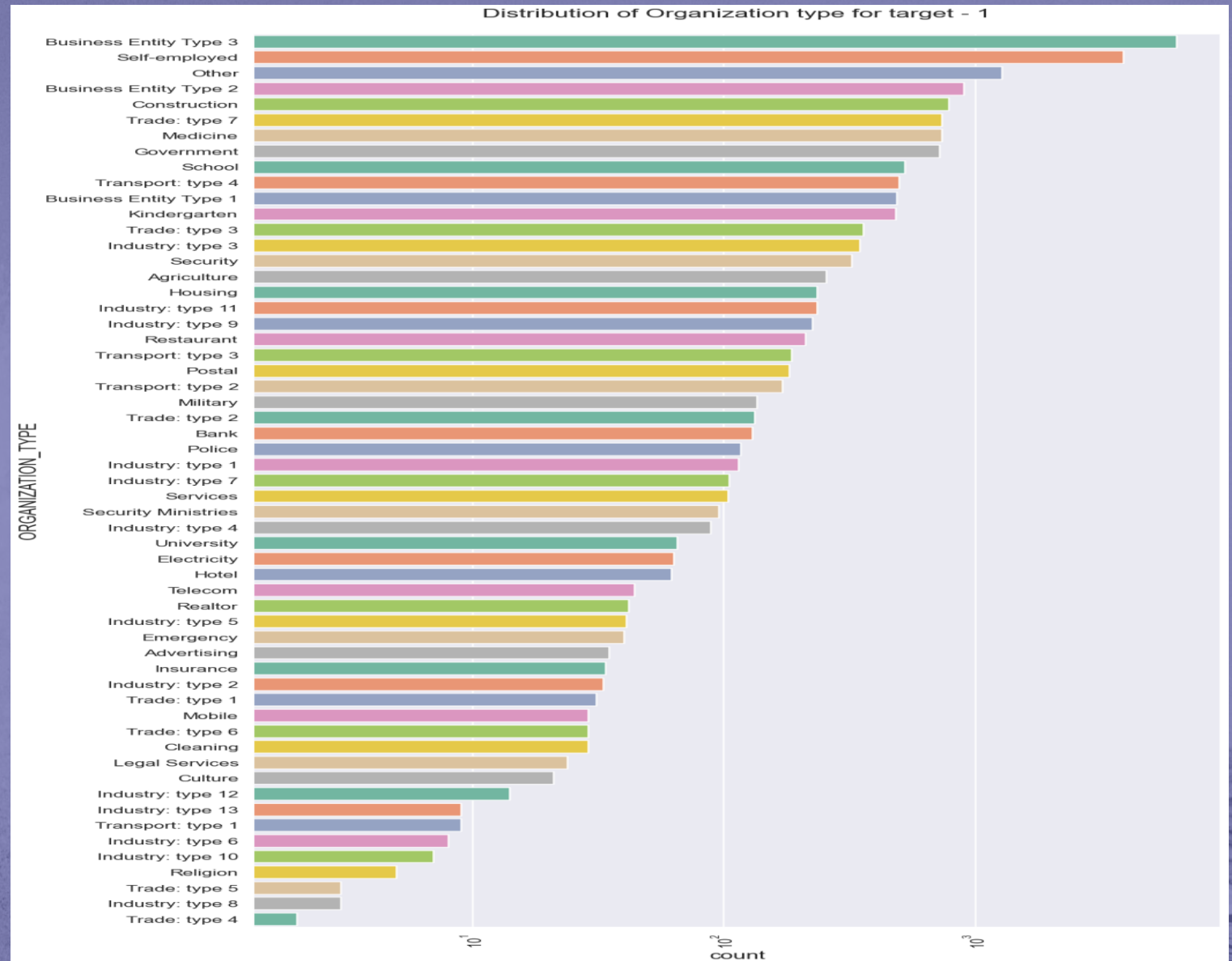
CORRELATION OF TARGET 0 AND TARGET 1

CORRELATION OF TARGET-0

Observations that can be derived from the correlation heatmap are:

• Credit amount is inversely proportional to the date of birth, which means Credit amount is higher for low age and vice-versa.

• Credit amount is inversely proportional to the number of children client have, means Credit amount is higher for less children count client have and vice-versa.

• Income amount is inversely proportional to the number of children client have, means more income for less children client have and vice-versa.

• Less children client are in densely populated area.

• Credit amount is higher in densely populated area.

• Income is also higher in densely populated area.
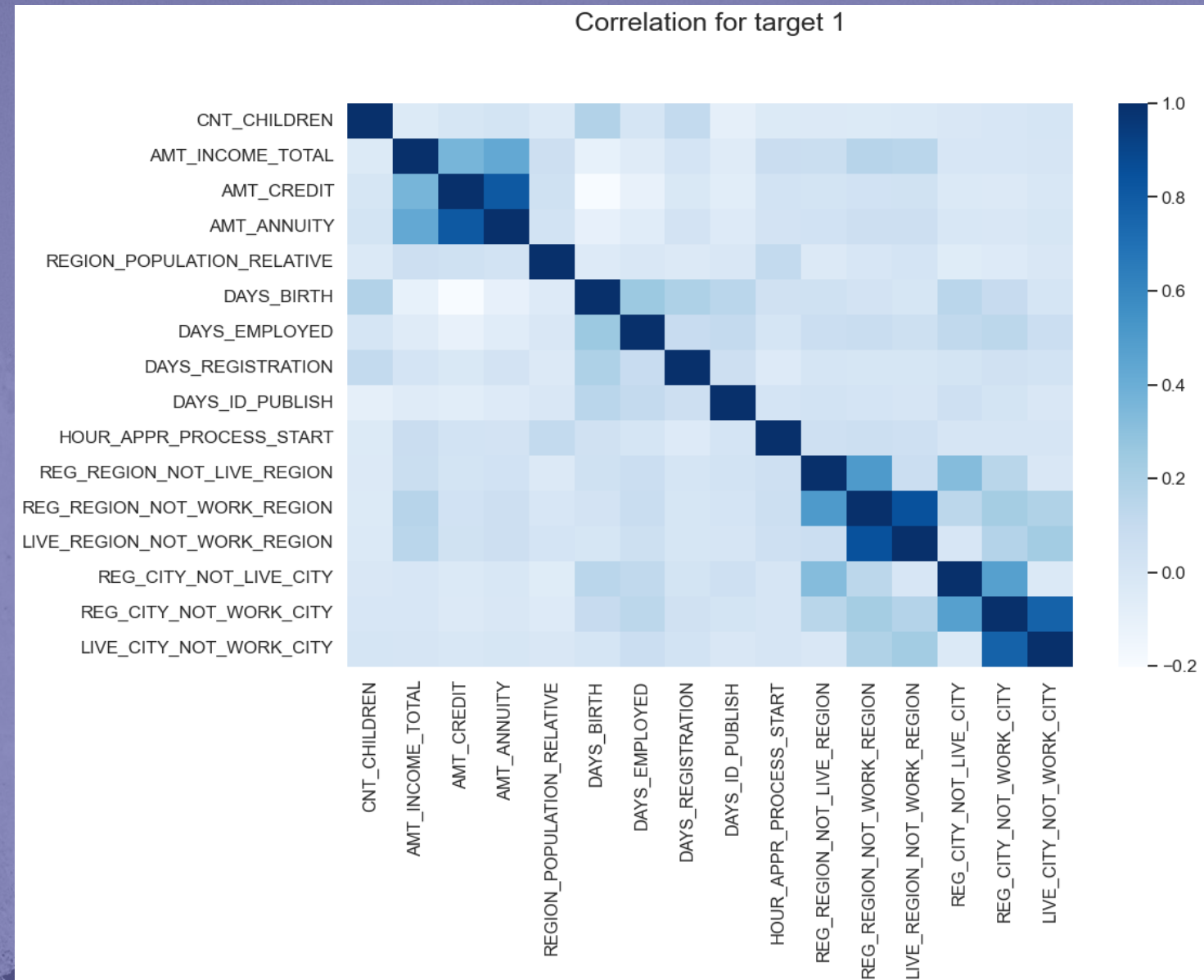

Correlation for target 0
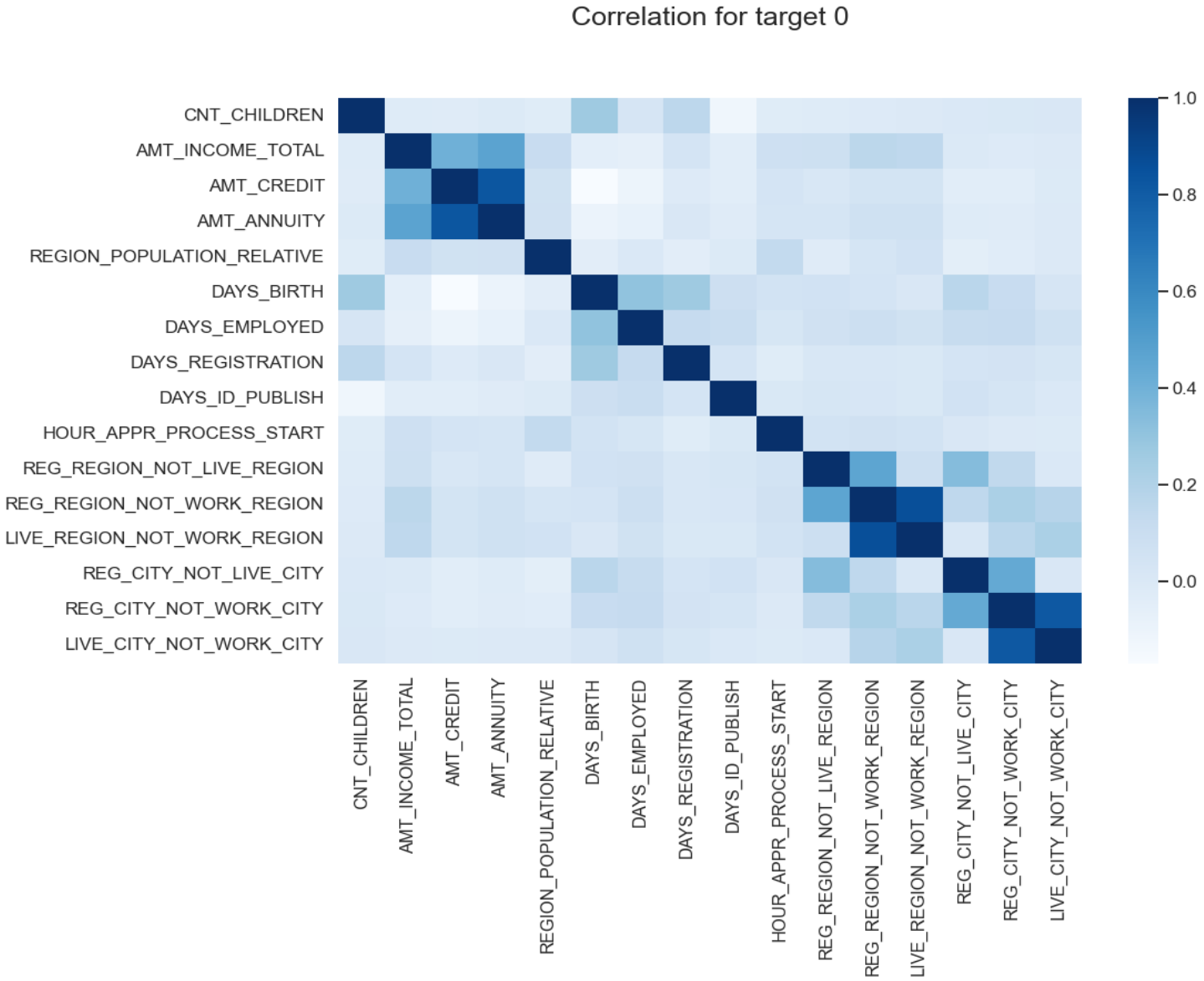
CORRELATION OF TARGET-1

Observations that can be derived from the correlation heatmap are as follows:

• The client's permanent address does not match contact address are having less children and vice-versa

• The client's permanent address does not match work address are having less children and vice-versa
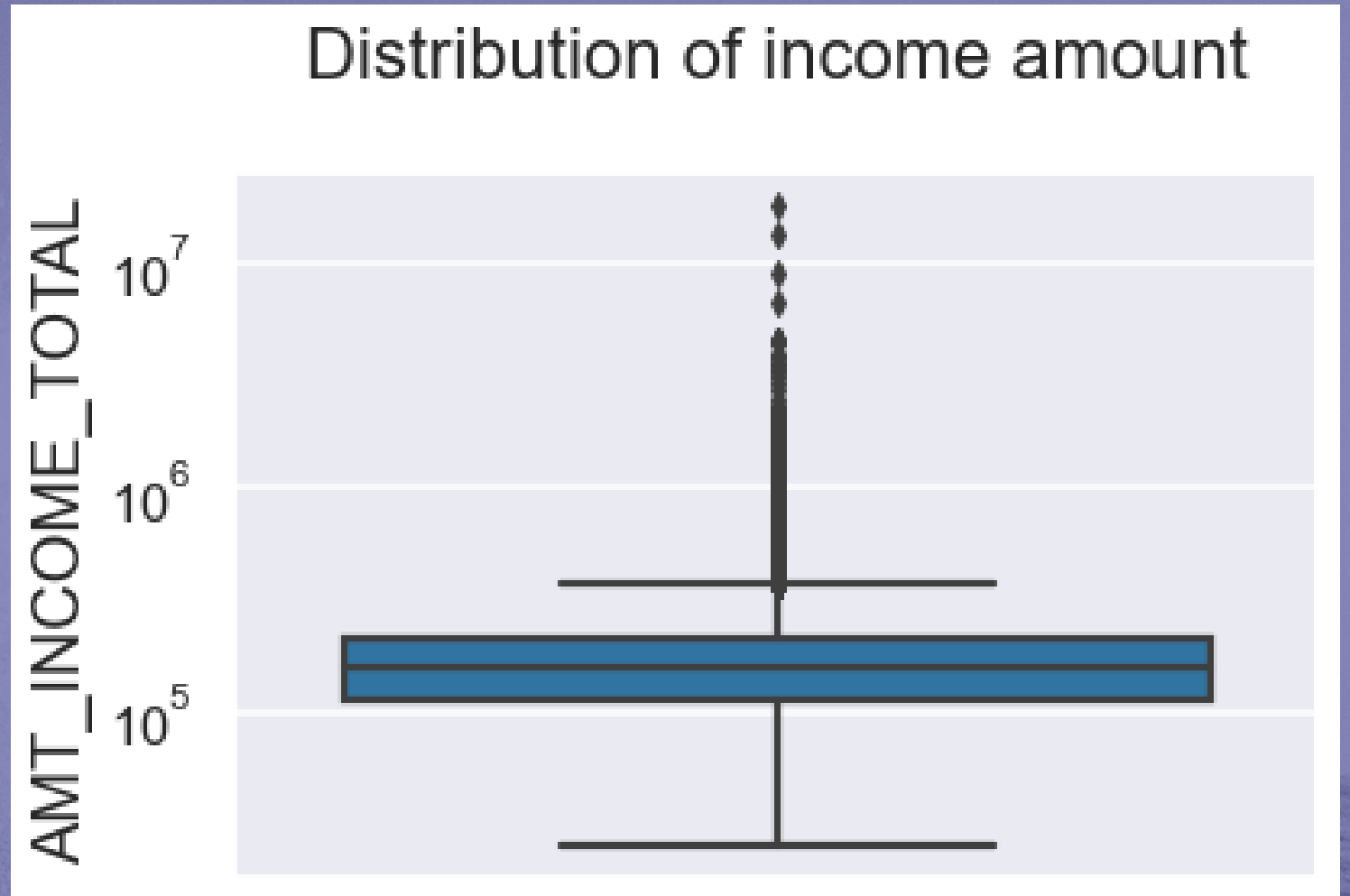

Correlation for target 1

CATEGORICAL UNIVARIATE ANALYSIS FOR VARIABLES TARGET 0

BOX PLOT FOR INCOME AMOUNT

Observation from the graph:

• Outliers are noticed in income amount.

• Third Quartile is very slim for income amount.
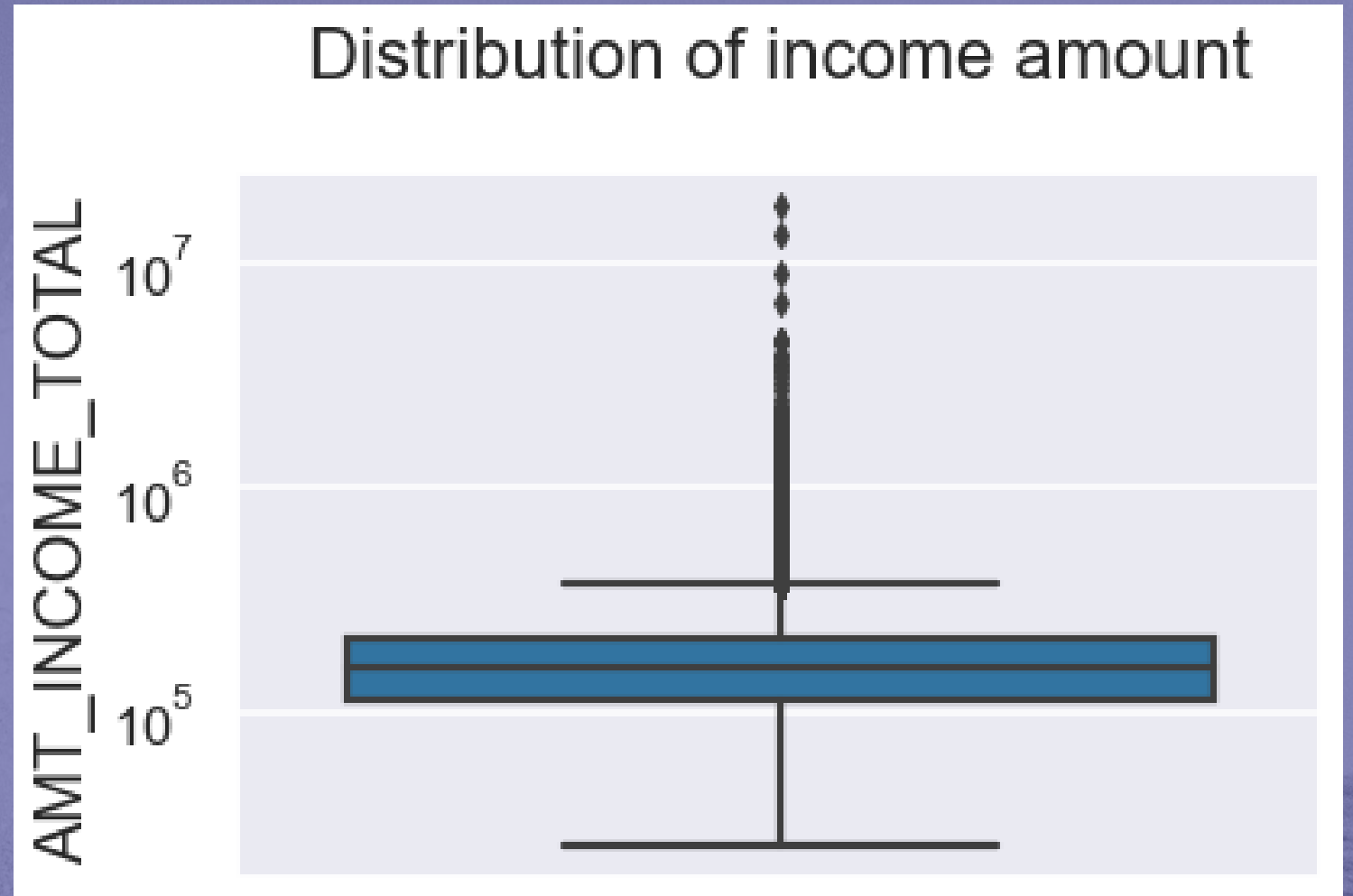

Distribution of income amount

BOX PLOT FOR CREDIT AMOUNT

Observation from the graph:

• Outliers are noticed in credit amount.

• First quartile is bigger than third quartile for credit amount which means most of the credits of clients are present in the first quartile.
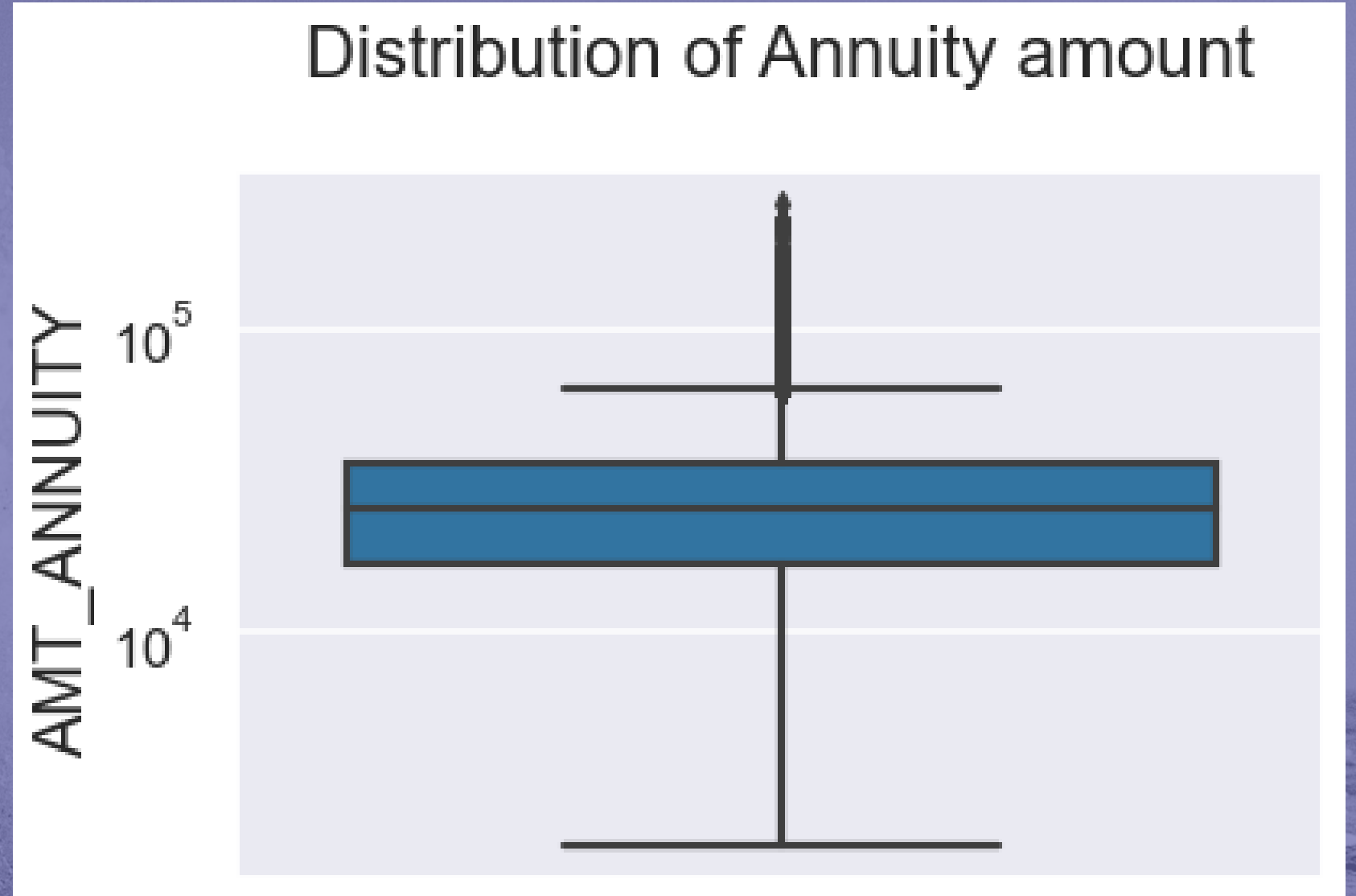

Distribution of income amount

BOX PLOT FOR ANNUITY AMOUNT

Observation from the graph:

• Outliers are noticed in annuity amount.

• First quartile is bigger than third quartile for annuity amount which means most of the annuity clients are from first quartile.

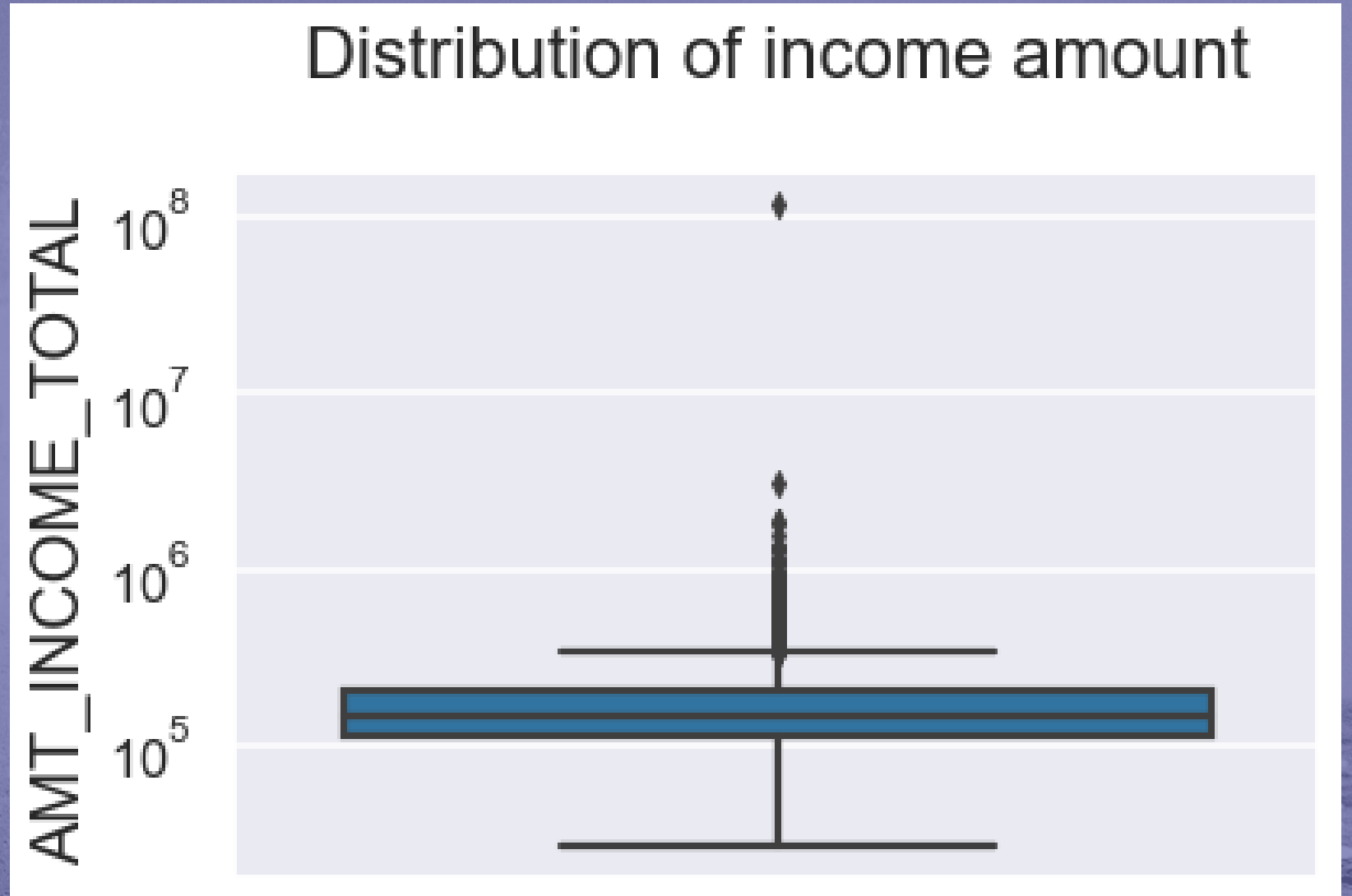Distribution of Annuity amount

CATEGORICAL UNIVARIATE ANALYSIS FOR VARIABLES TARGET 1

BOX PLOT FOR INCOME AMOUNT

Observation from the graph:

• Outliers are noticed in income amount.

• First quartiles is very slim for income amount.

• Most of the clients of income are present in third quartile.

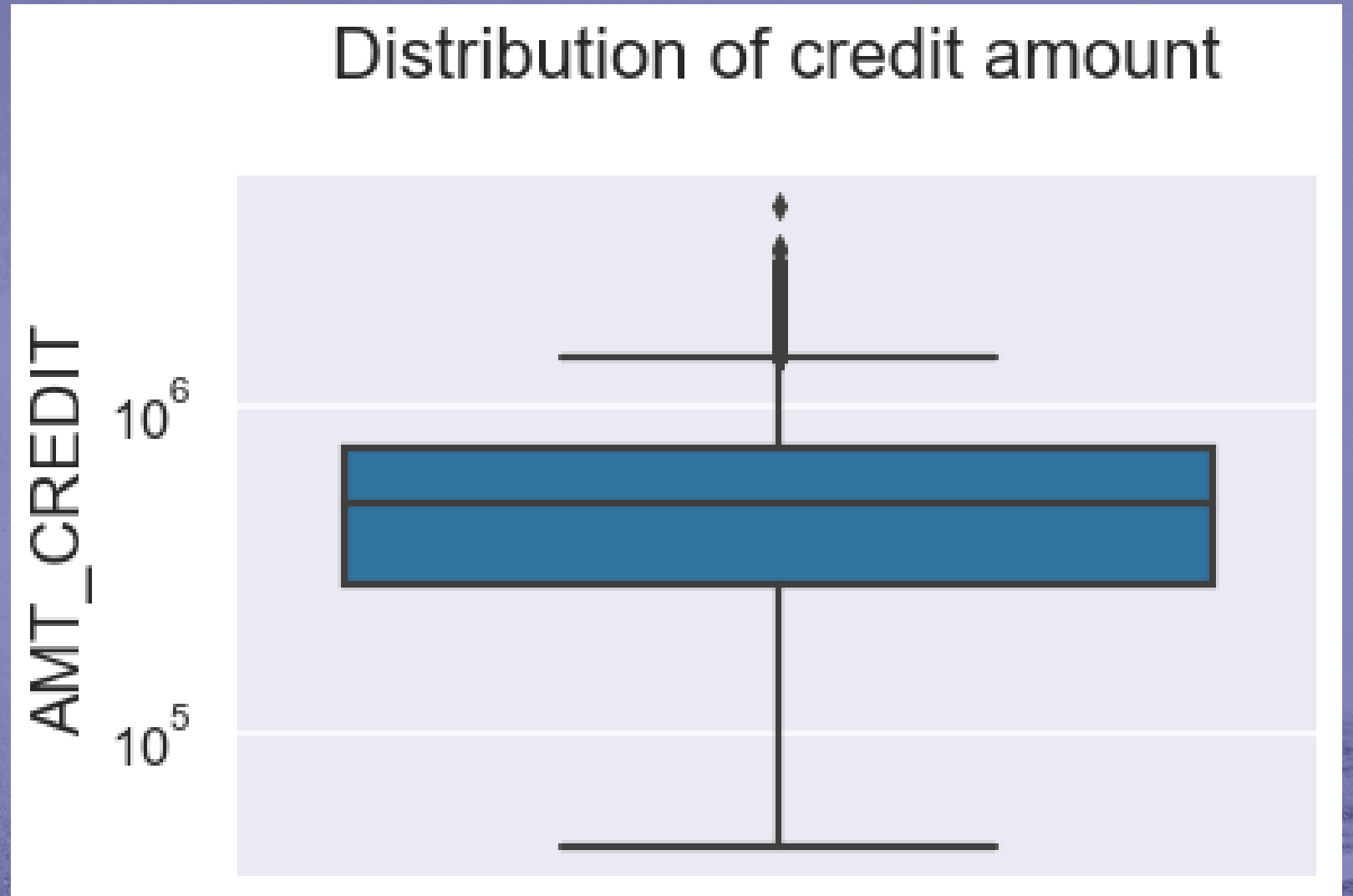Distribution of income amount

BOX PLOT FOR CREDIT AMOUNT

Observation from the graph:

• Outliers are noticed in credit amount.

• The first quartile is bigger than third quartile for credit amount which means most of the credits of clients are present in the first quartile.
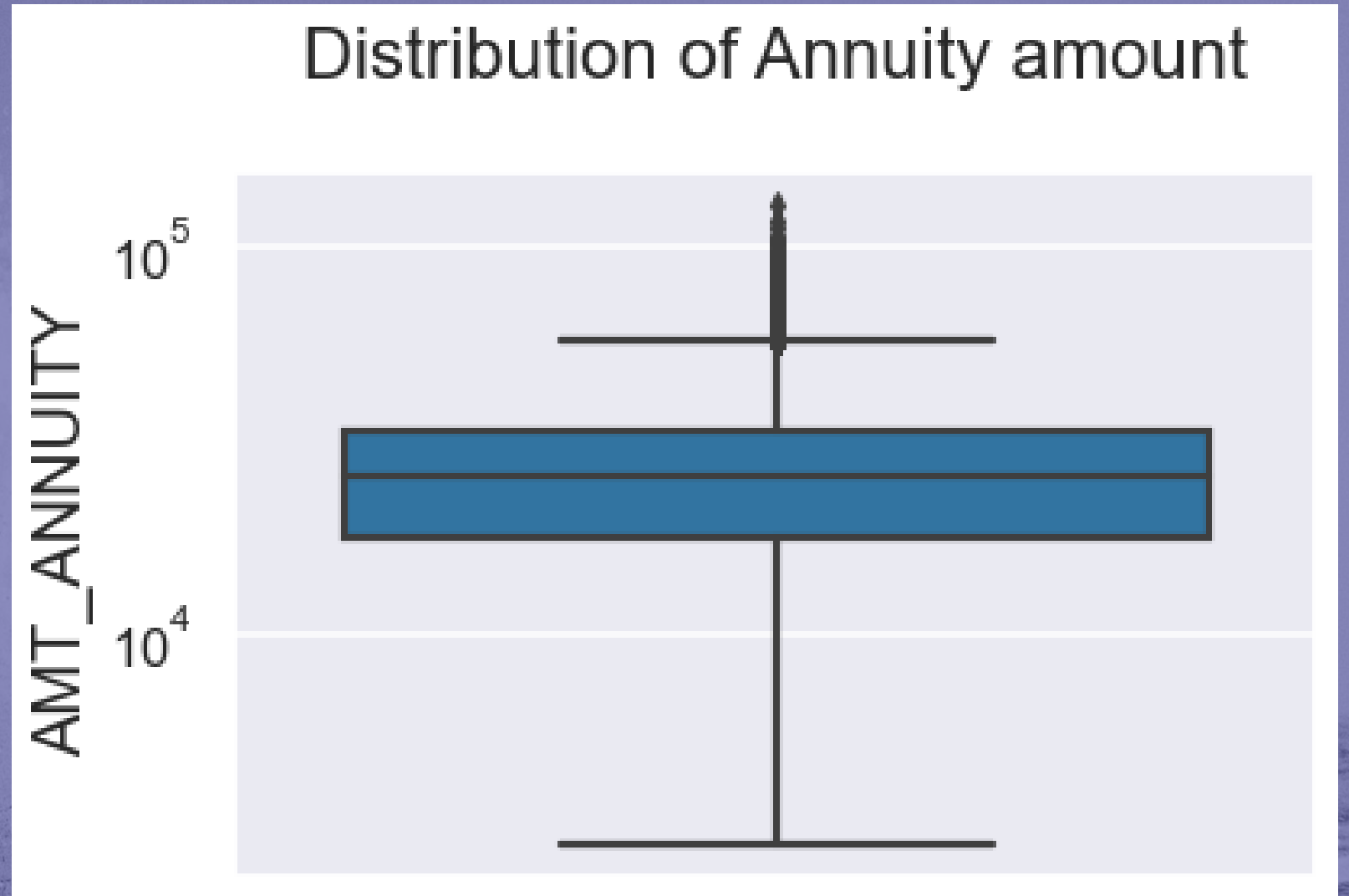

Distribution of credit amount

BOX PLOT FOR ANNUITY AMOUNT

Observation from the graph:

• Outliers are noticed in annuity amount.

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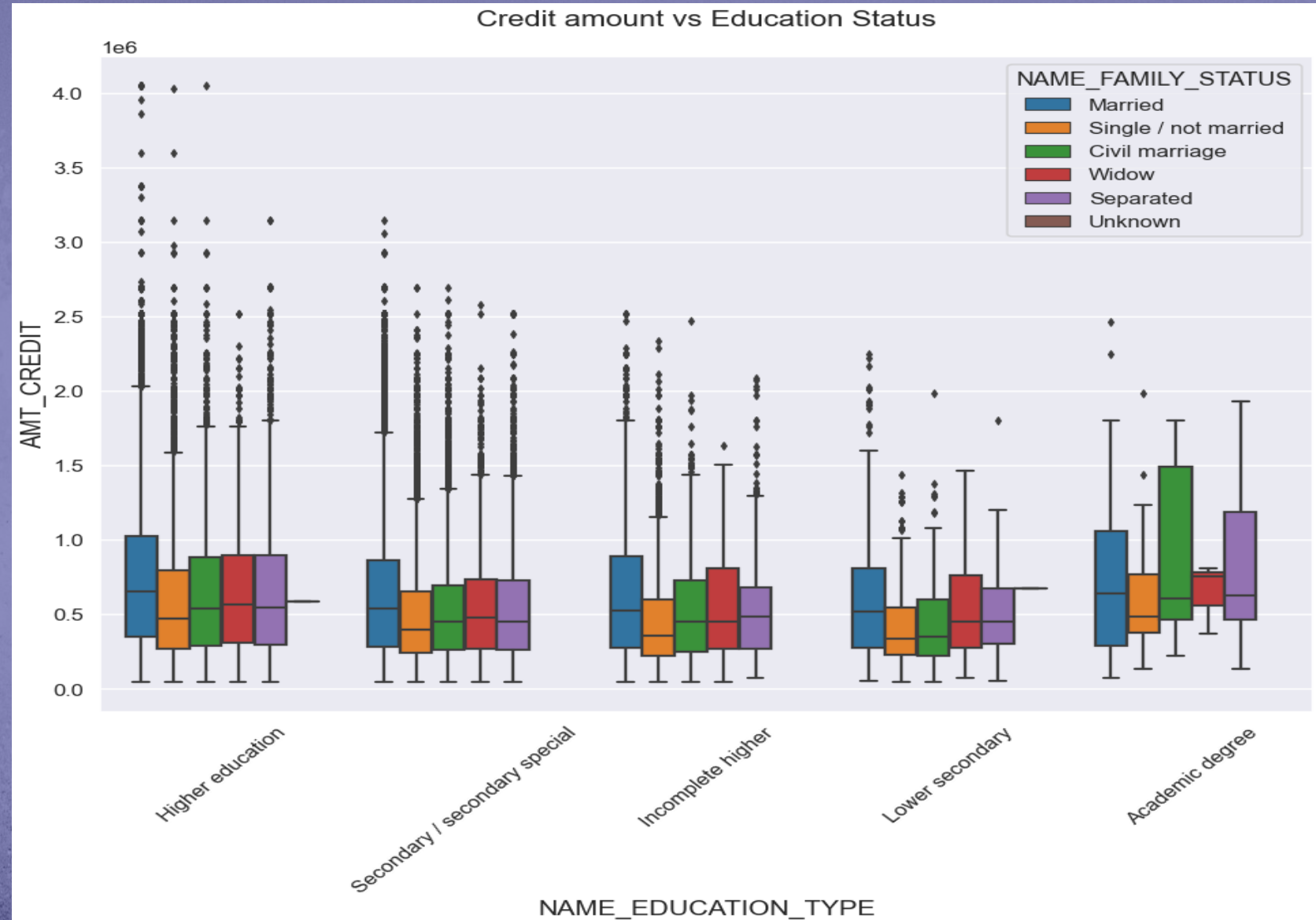

Distribution of Annuity amount

BIVARIATE ANALYSIS FOR TYPE 0

CREDIT AMOUNT VS. EDUCATION STATUS

Observation from the graph:

• Family status of 'civil marriage', 'marriage' and 'separated' of Academic degree education are having higher number of credits than others.

• Higher education of family status of 'marriage', 'single' and 'civil marriage' are having more outliers.

• Civil marriage for Academic degree is having most of the credits in the third quartile.
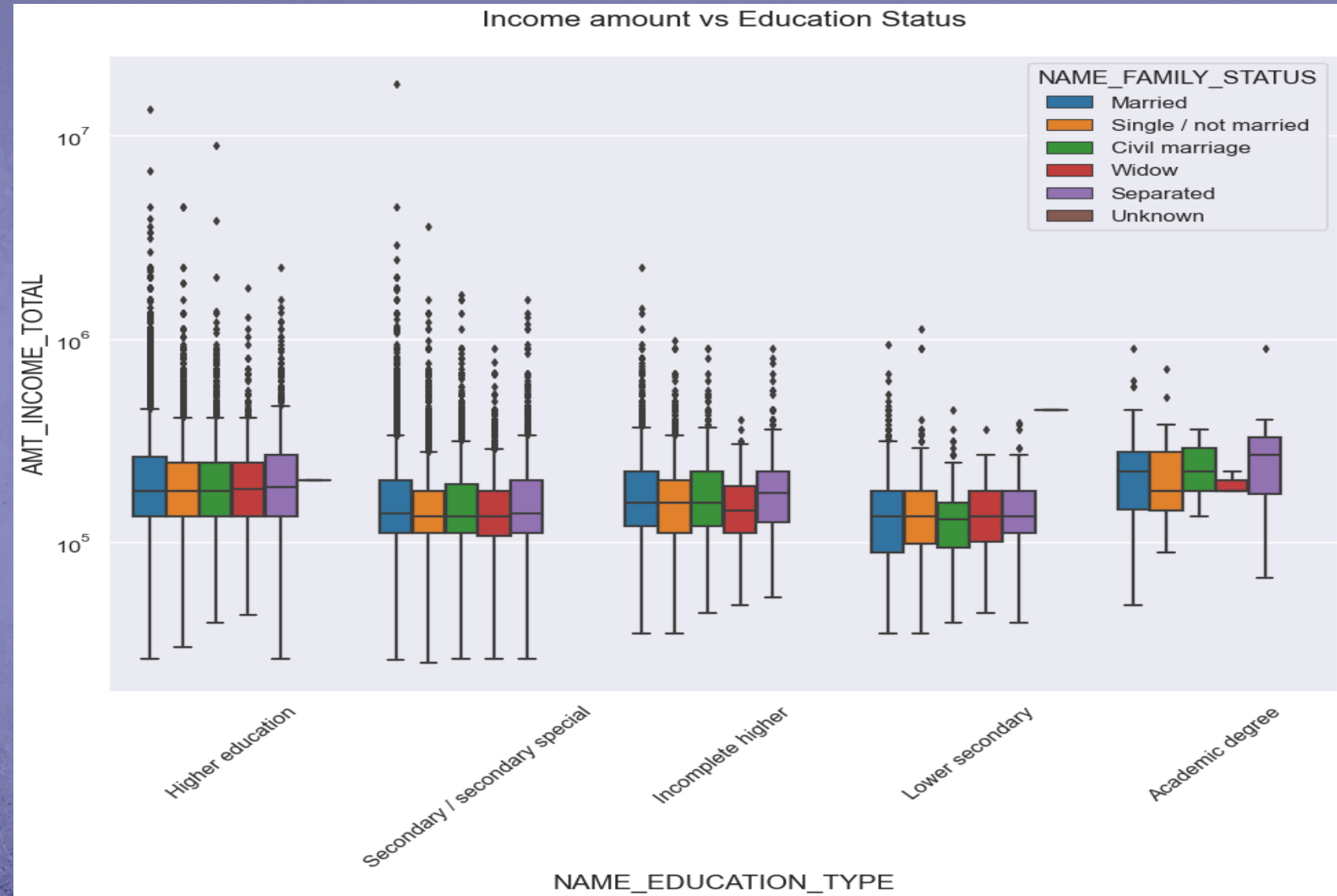

Credit amount vs Education Status

INCOME AMOUNT VS. EDUCATION STATUS

Observation from the graph:

• 'Higher education' in the income amount is mostly equal with family status. It does contain many outliers.

• Less outliers are having for Academic degree but there income amount is little higher that Higher education.

• Lower secondary of civil marriage family status are having less income amount than others.

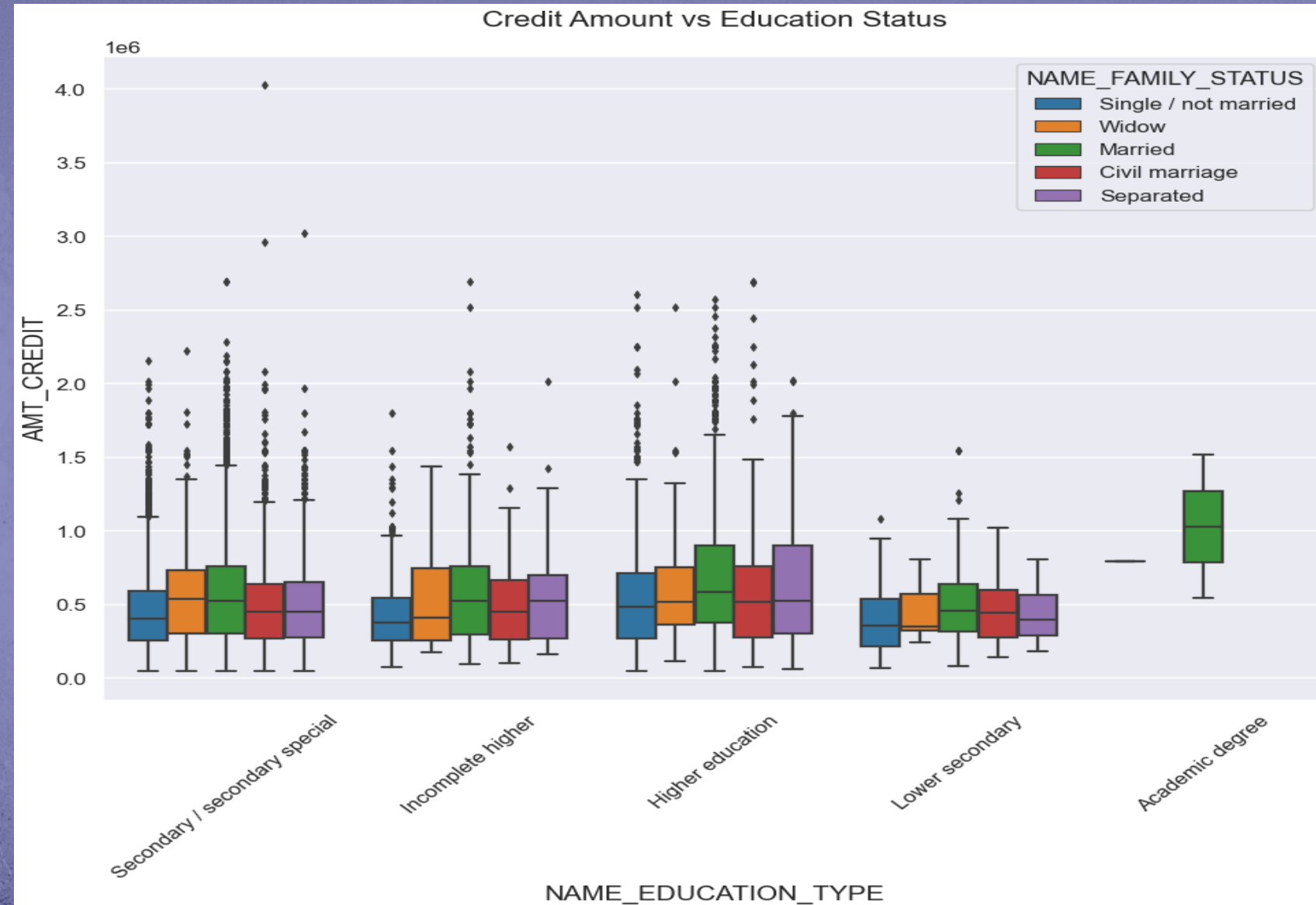
Income amount vs Education Status

BIVARIATE ANALYSIS FOR TYPE 1

CREDIT AMOUNT VS. EDUCATION STATUS

Observation from the graph:

• Most of the outliers are present in education type 'Secondary' and 'Higher education'.

• Family status of 'Married', 'Separated' and 'civil marriage' of 'Higher education' are having higher number of credits than others.

• 'Separated' for 'Higher education' is having most of the credits in the third quartile.
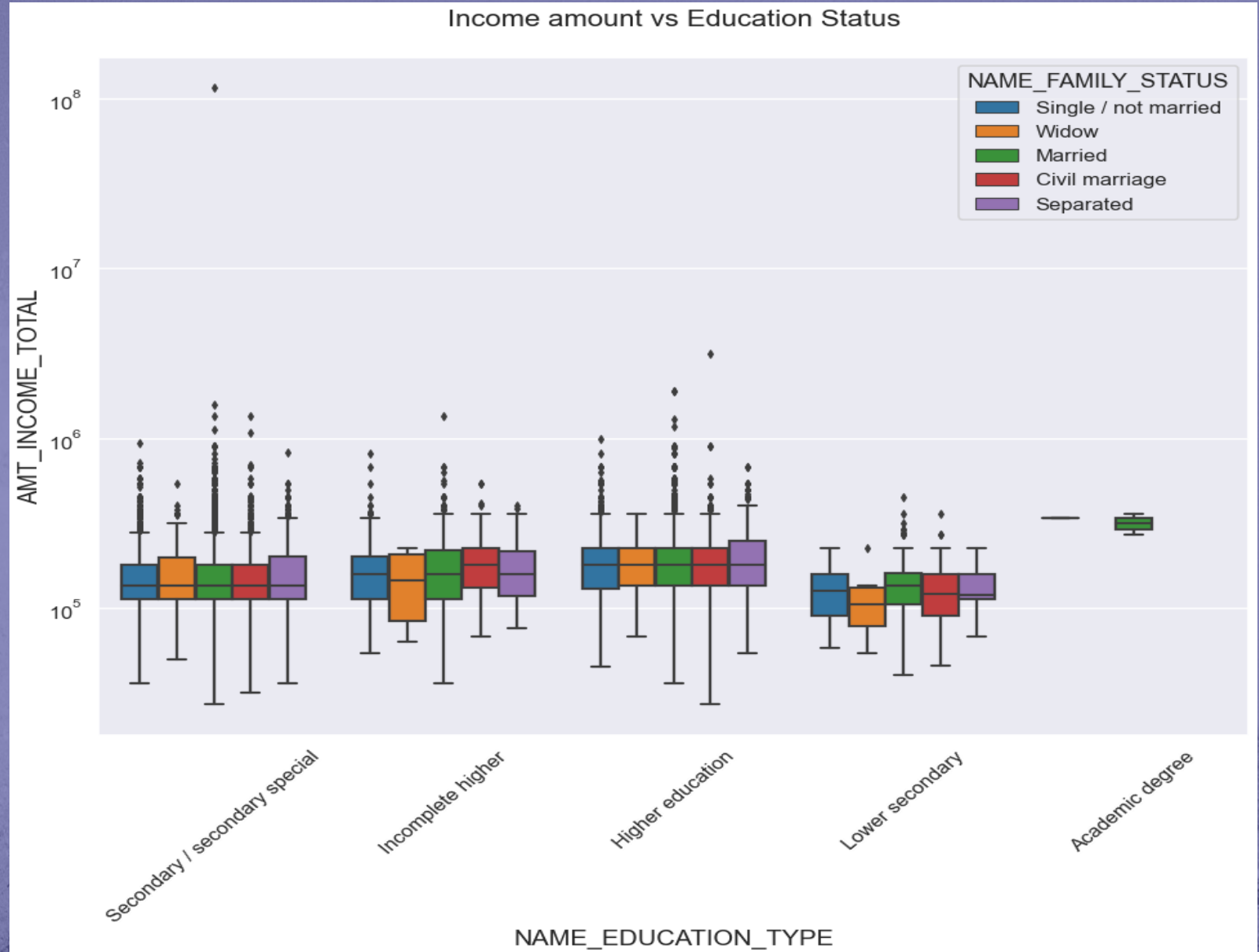
INCOME AMOUNT VS. EDUCATION STATUS

Observation from the graph:

• Income amount is mostly equal with family status.

• Less outlier are having for Academic degree but there income amount is little higher that Higher education.

• Lower secondary are have less income amount than others.

UNIVARIATE ANALYSIS AFTER MERGING PREVIOUS DATA

DISTRIBUTION OF CONTRACT STATUS WITH PURPOSES

Observation from the graph:

• Most rejection of loans came from purpose 'repairs'.

• For education purposes we have equal number of approvals and rejection

• 'Payment on other loans' and 'buying a new car' are having significant higher rejection than approvals

Distribution of contract status with purposes
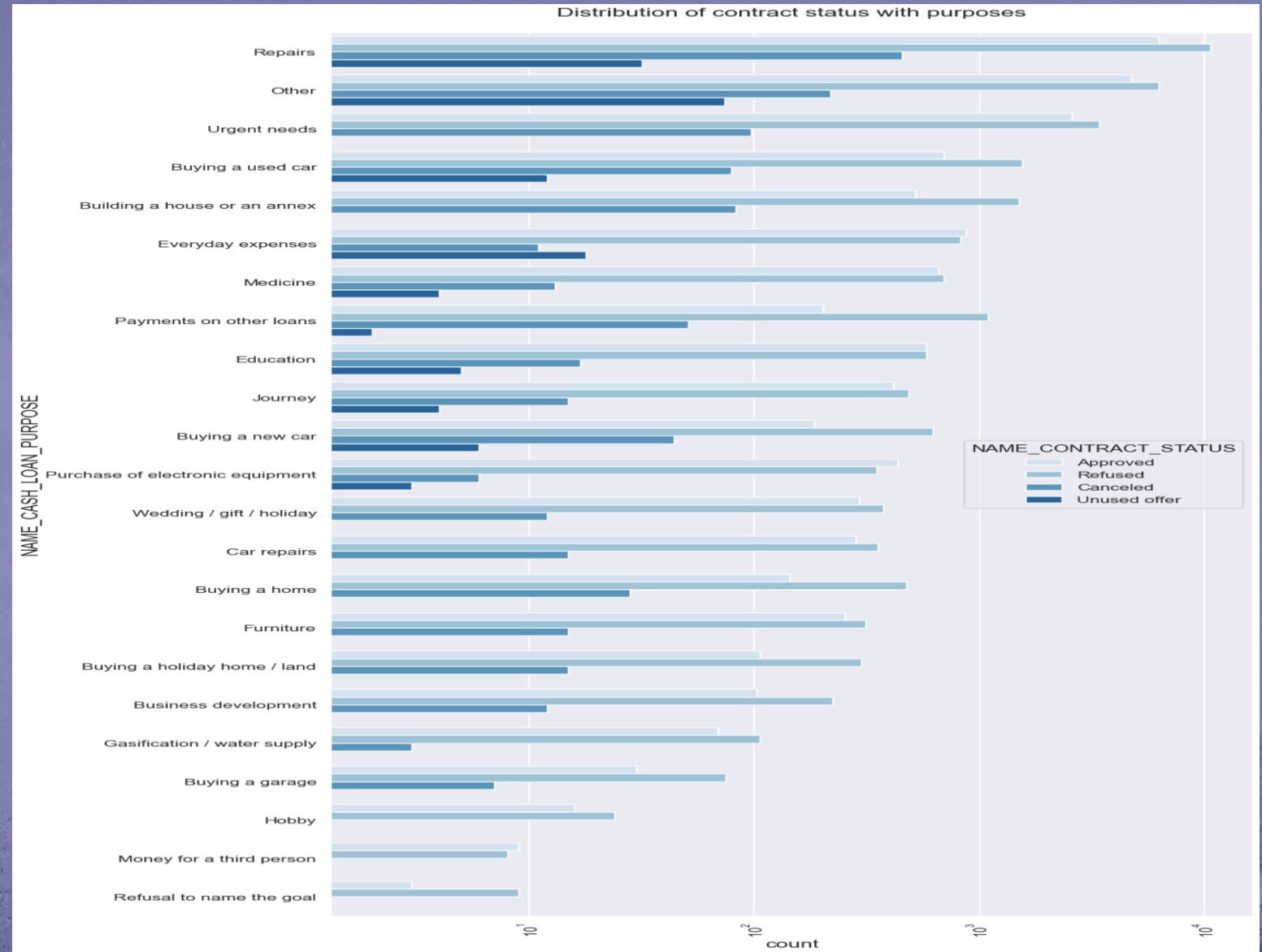
DISTRIBUTION OF PURPOSES WITH TARGET

Observation from the graph:

• Most rejection of loans came from purpose 'repairs'.

• For education purposes we have equal number of approvals and rejection

• 'Payment on other loans' and 'buying a new car' are having significant higher rejection than approvals
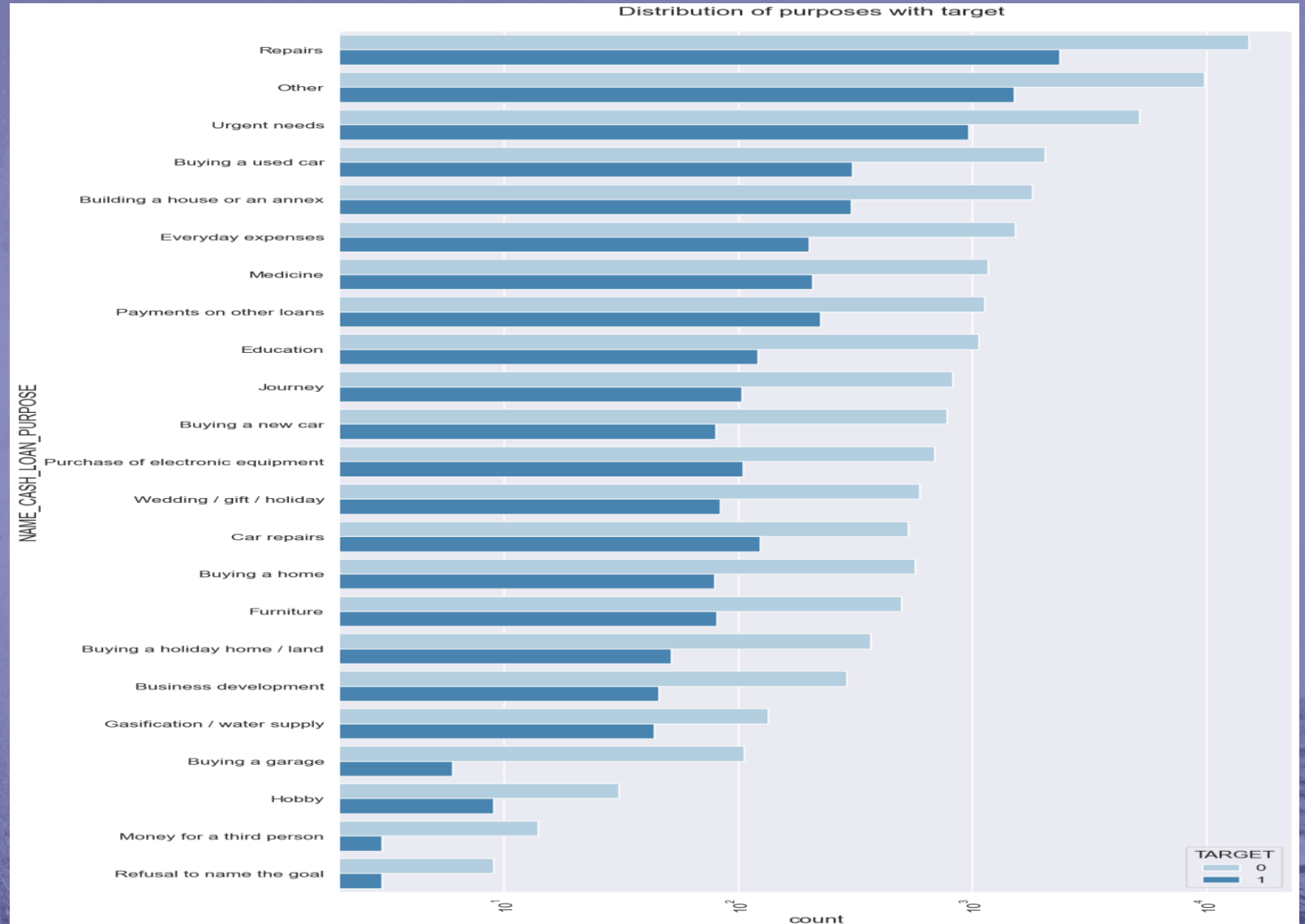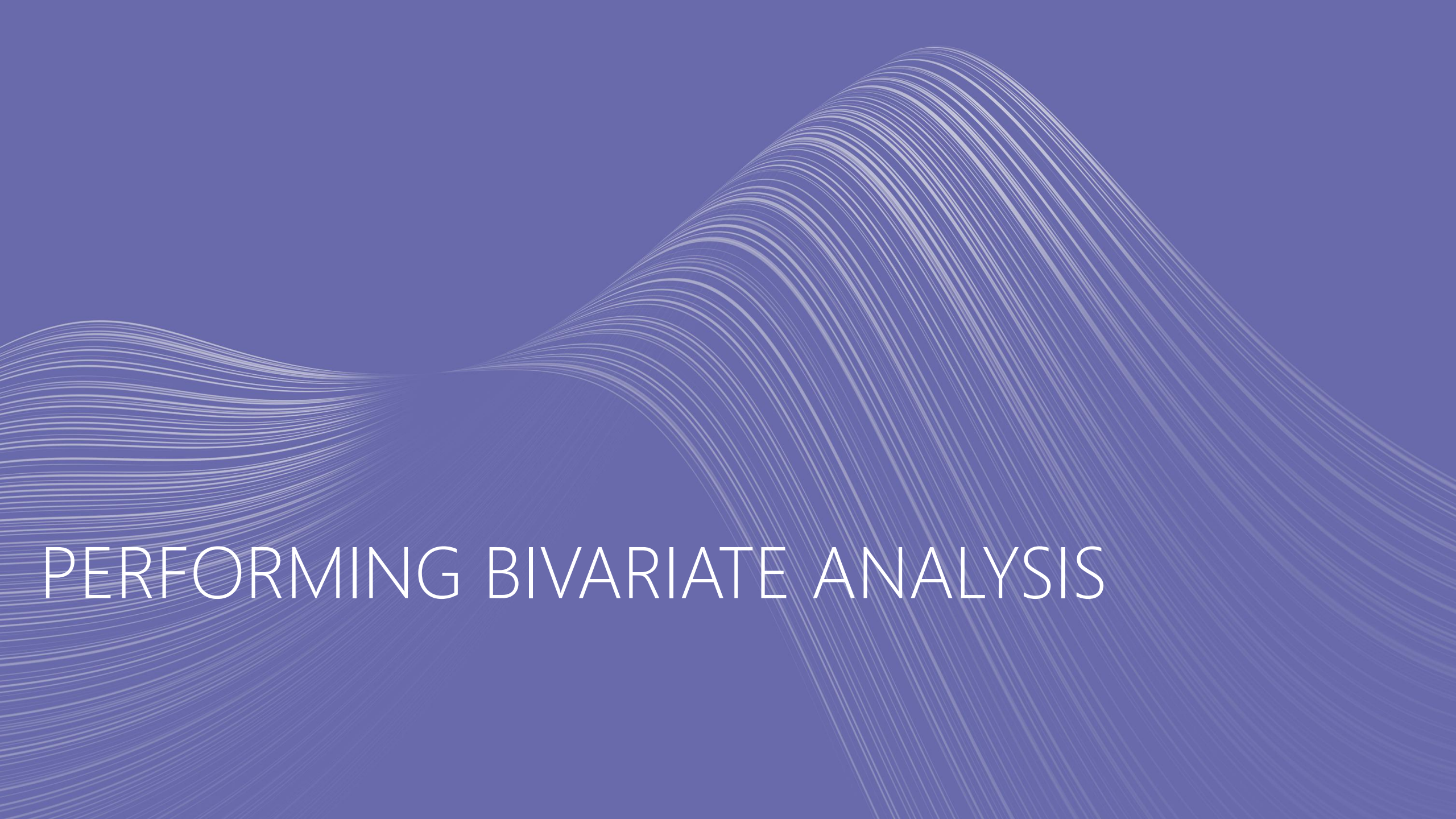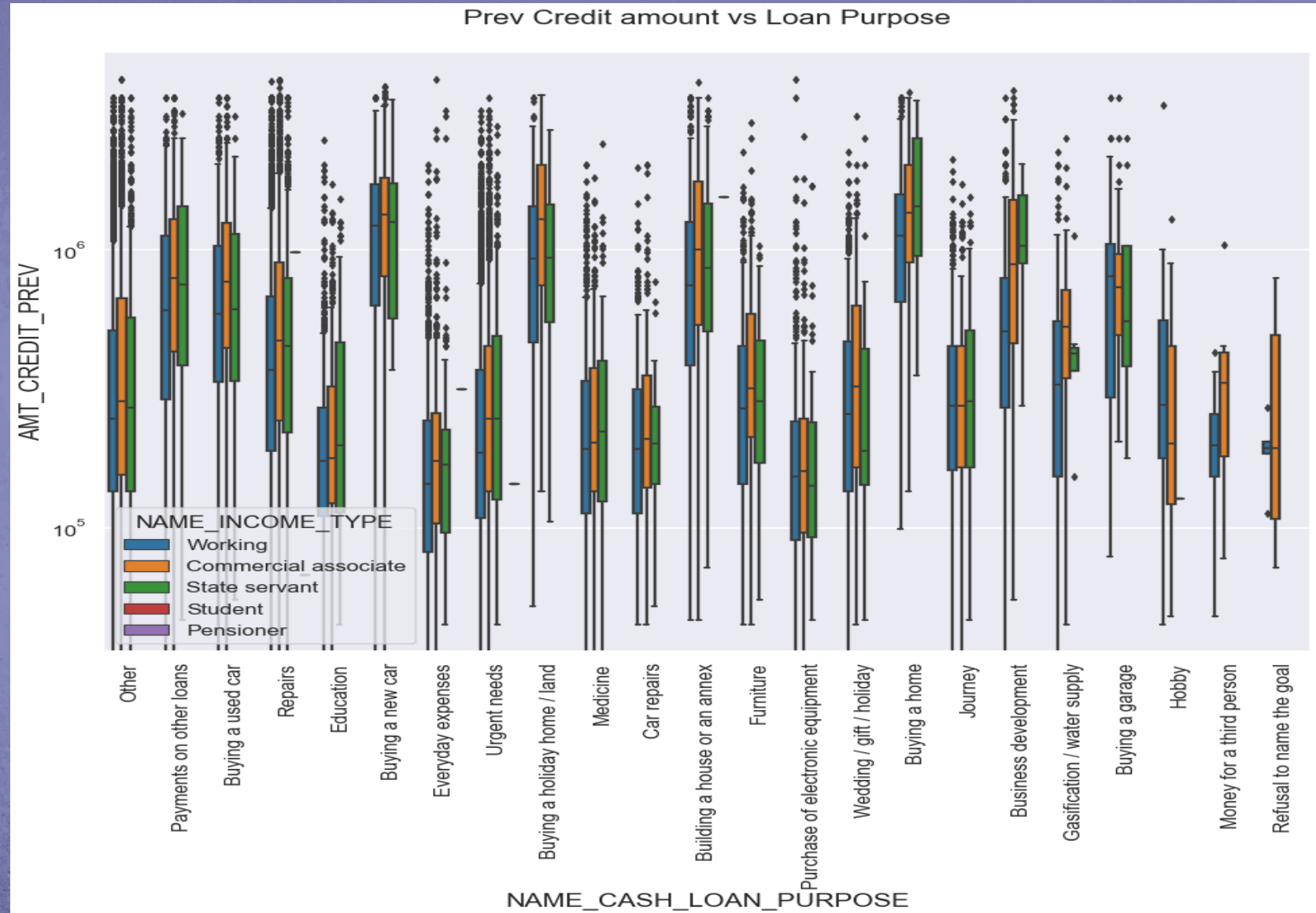
PERFORMING BIVARIATE ANALYSIS

PREVIOUS CREDIT AMOUNT VS. LOAN PURPOSE

Observation from the graph:

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

• Income type of state servants have a significant amount of credit applied

• 'Money for third person' and 'Hobby' are having less credits applied for.

PREVIOUS CREDIT AMOUNT VS. HOUSING TYPE

Observation from the graph:

• 'House / apartment', 'Office apartment' are having higher credit in target 0

• 'Co-op apartment' is having higher credit in target 1.

• Based on the above analysis, Bank should avoid giving loans to the housing type of 'co-op apartment' as they are having difficulties in payment.

• Bank can focus mostly on housing type 'With parents', 'House / apartment' or 'Municipal apartment' for successful payments.
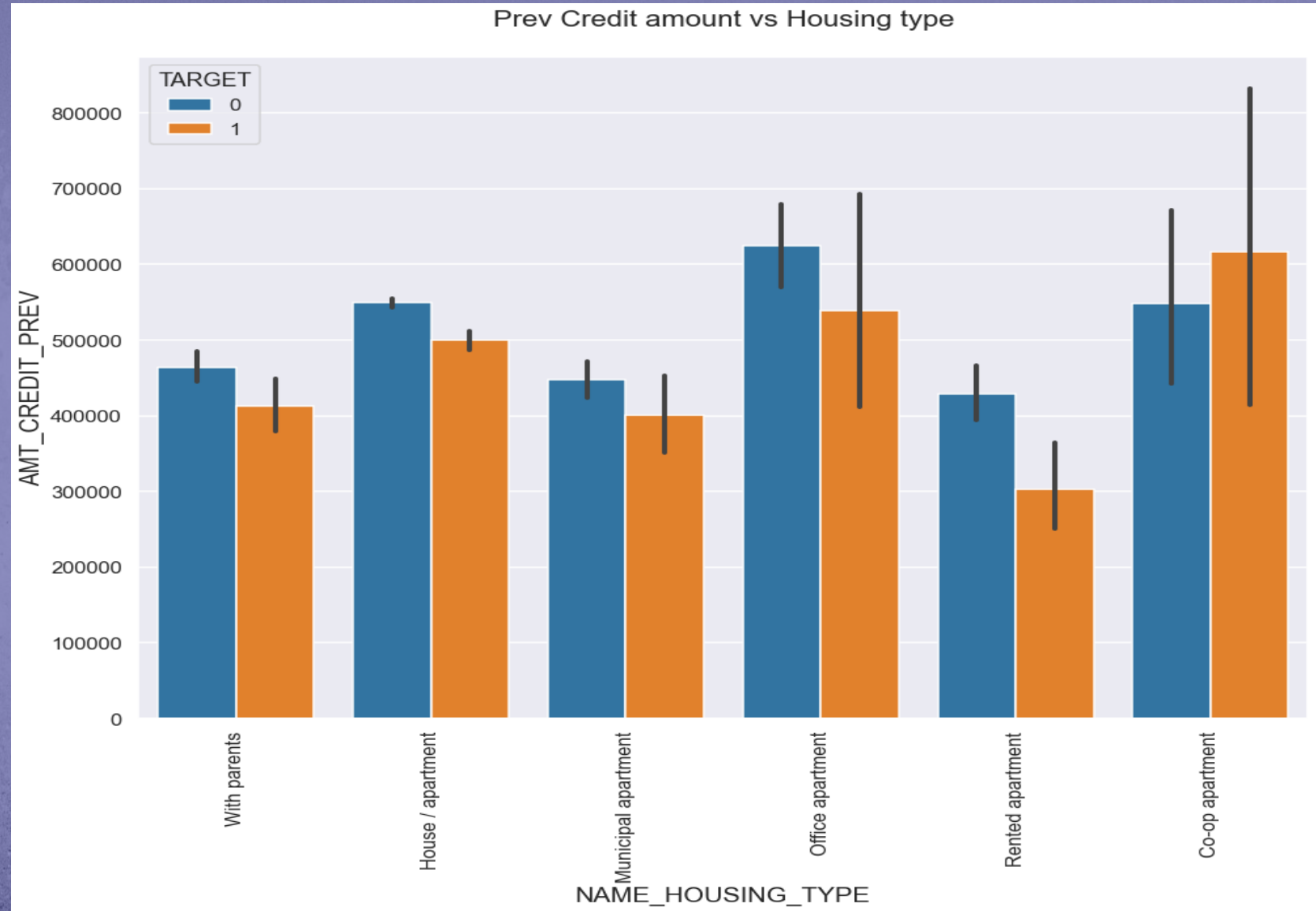
CONCLUSION

- High Focus Area for Banks:
  - 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 on clients from housing type 'With parents' as they are having least number of unsuccessful payments.

- Less Focus Area for Banks:
  - Banks should focus less on income type 'Working' as they are having most number of unsuccessful payments.
  - Loan purpose 'Repair' is having higher number of unsuccessful payments on time.

THANK YOU