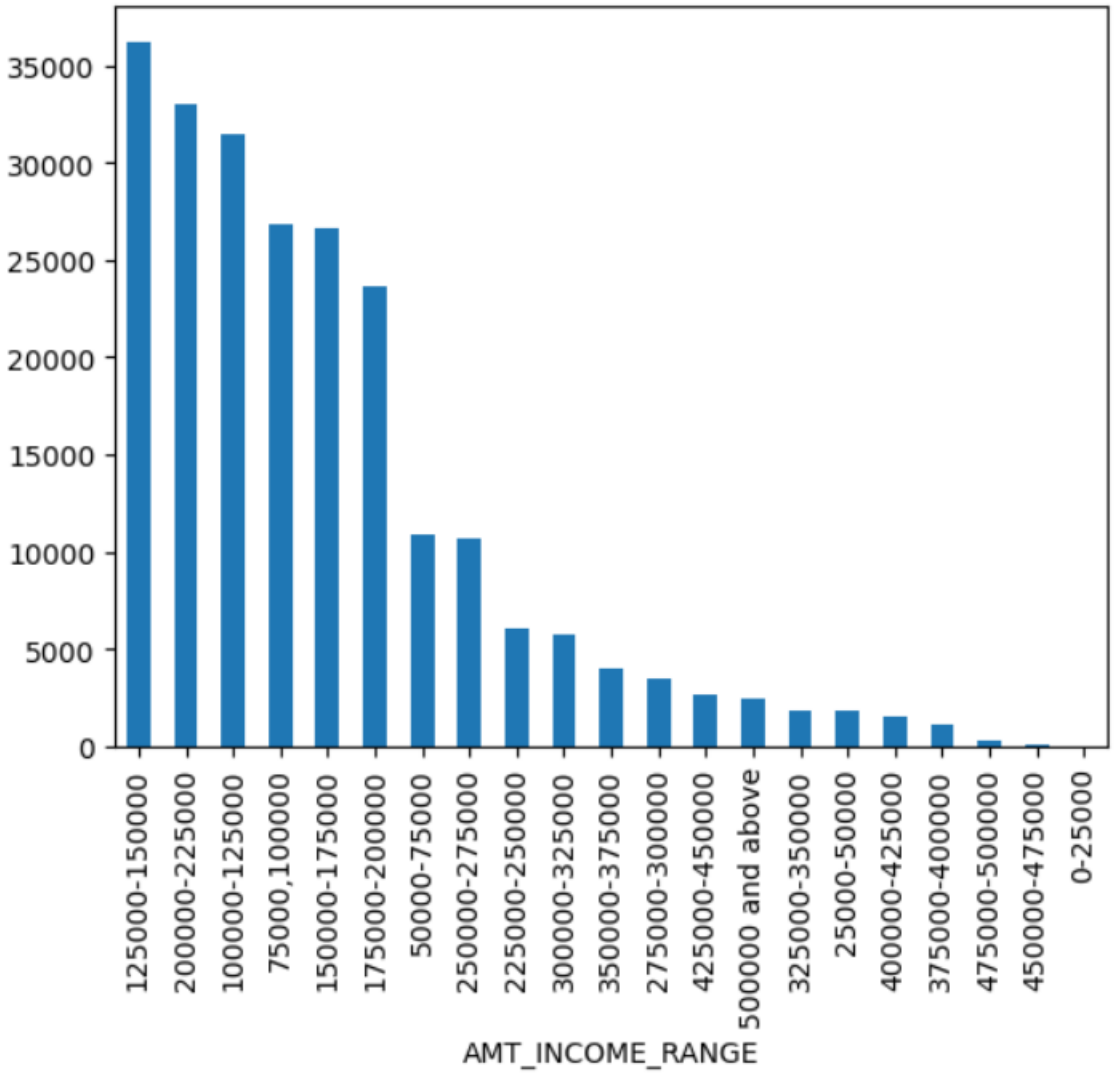
Assignment:

CREDIT EDA

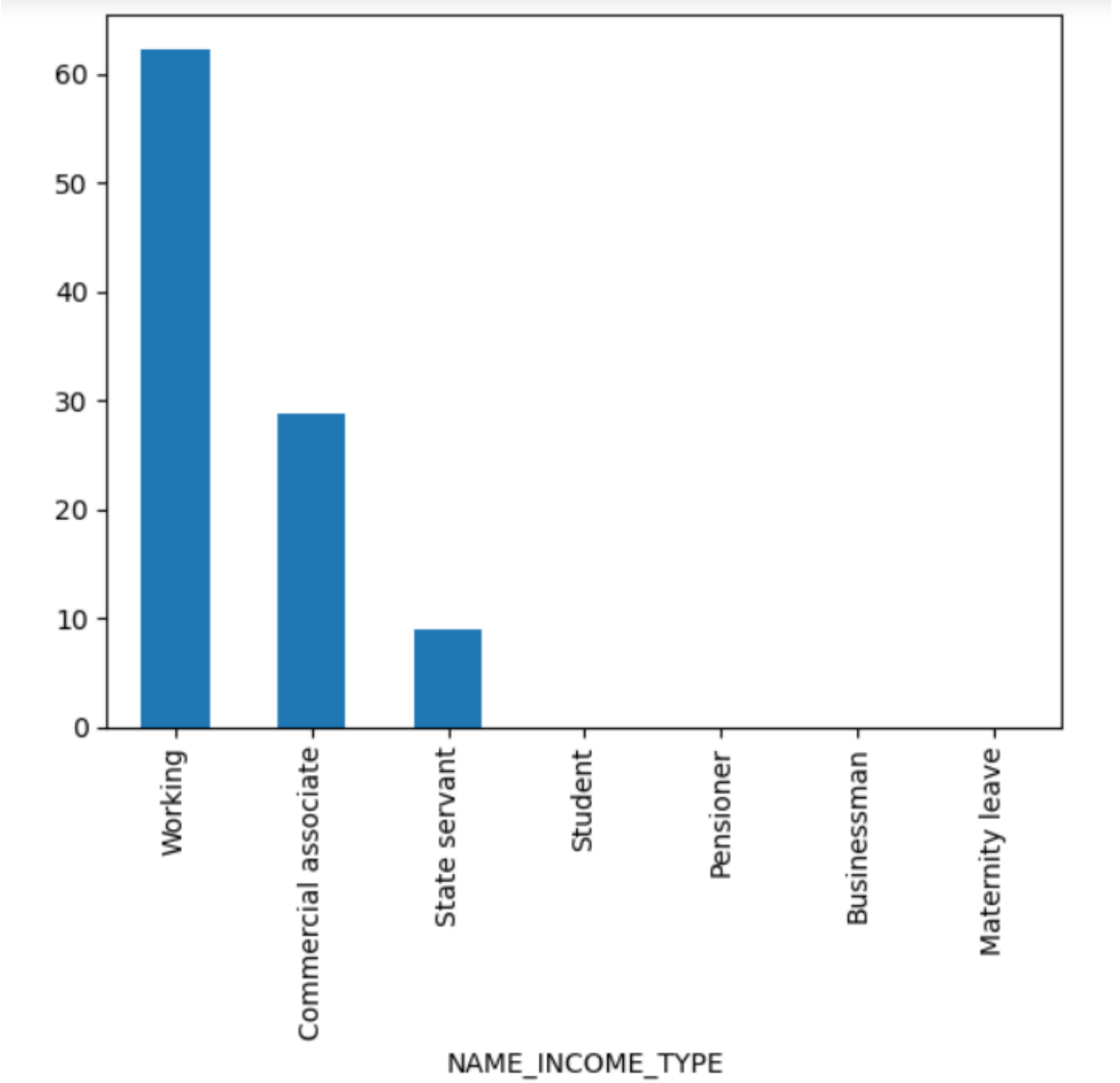
Categorical Univariate analysis for target 0

1. as we can see there are more number of persons in the income range "125000-200000"
2. the remaining income range has less number of persons in their respected income range



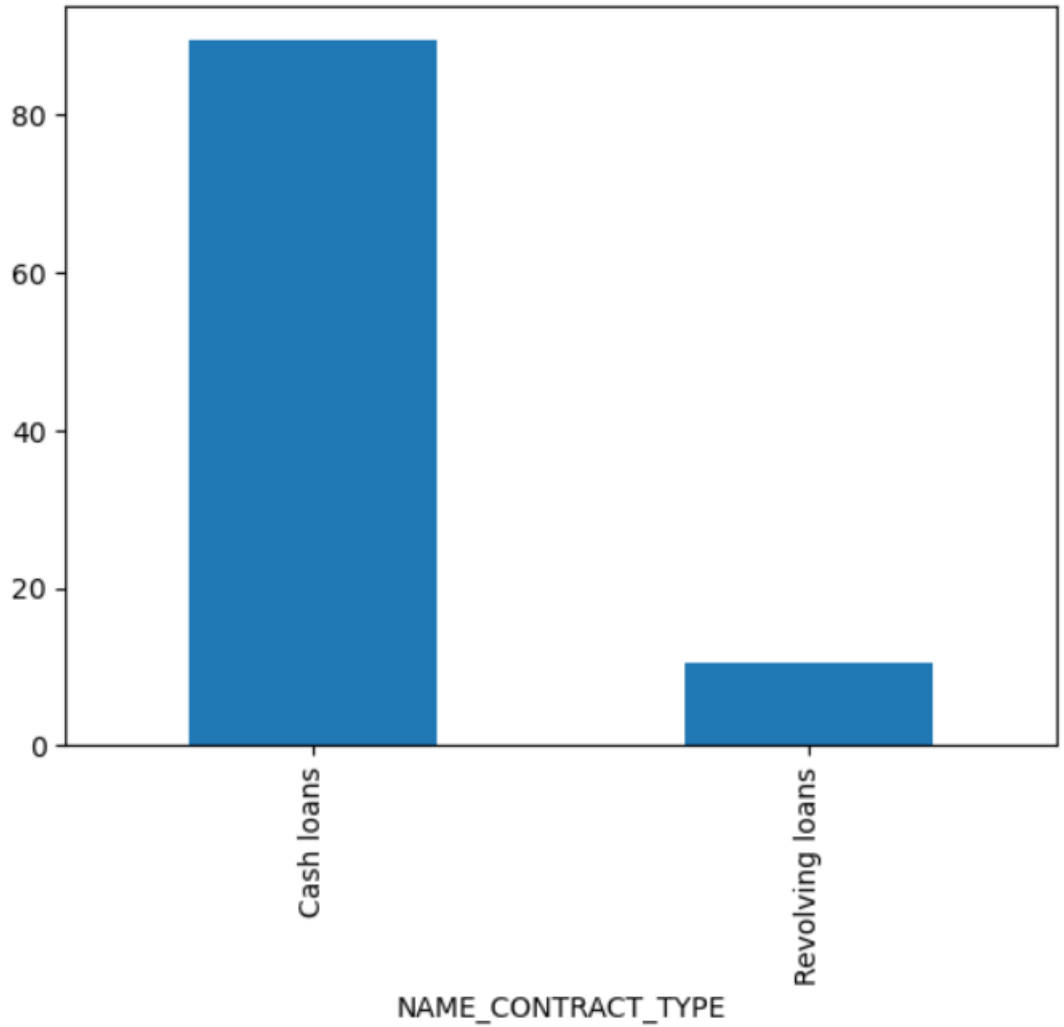
Based on income type

1. For income type ‘working’, ’commercial associate’, and ‘State Servant’ the number of credits are higher than others.
2. Less number of credits or no credits for income type ‘student’ ,’pensioner’, ‘Businessman’ and ‘Maternity leave’.



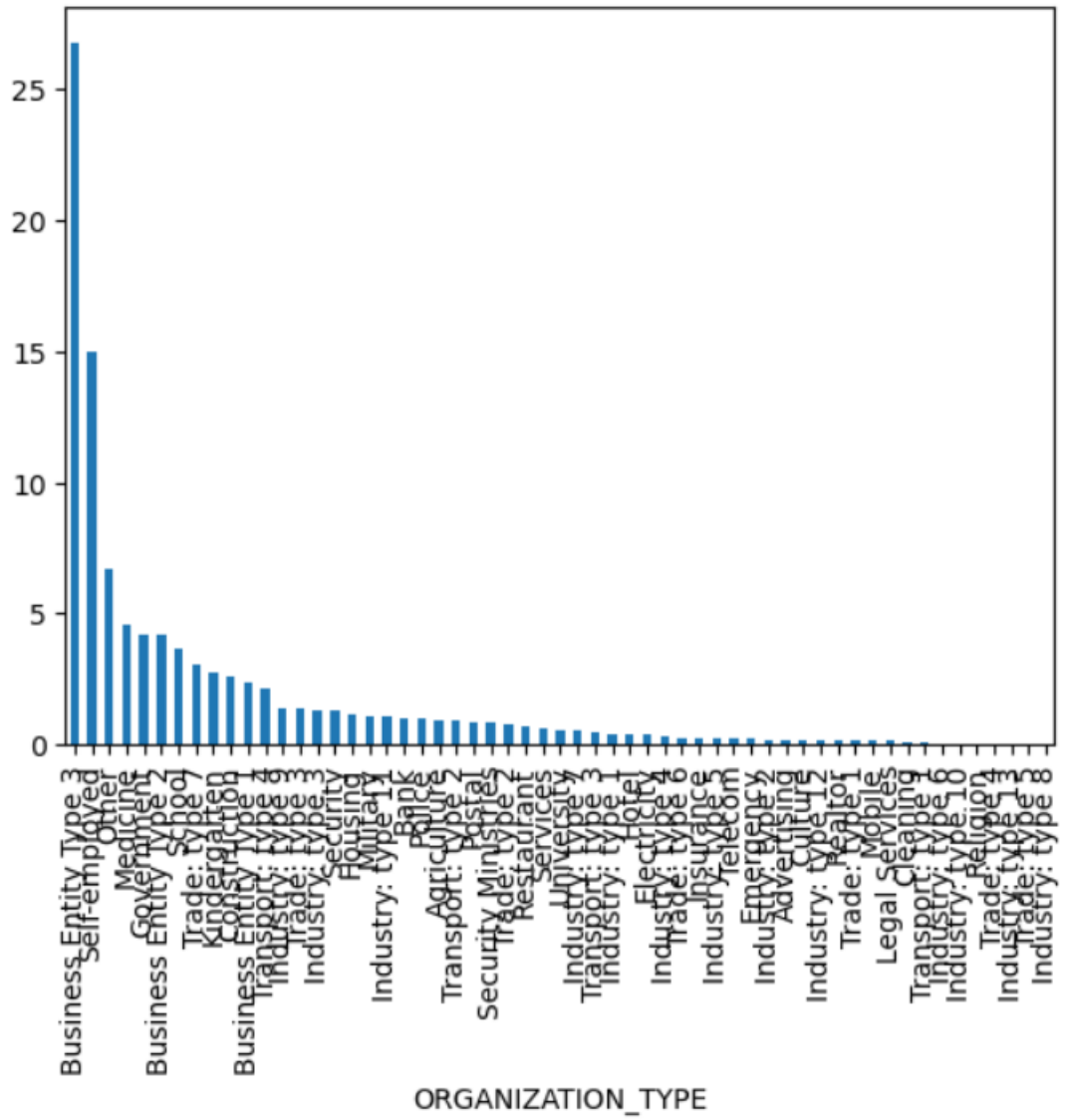
Based on contract type

1. cash loans=loans that is paid in the form of emi's every month.
2. revolving loans= loans that's paid whenever they want,e.g., credit card loan repayment.
3. cash loans are taken more in this intrest variable.



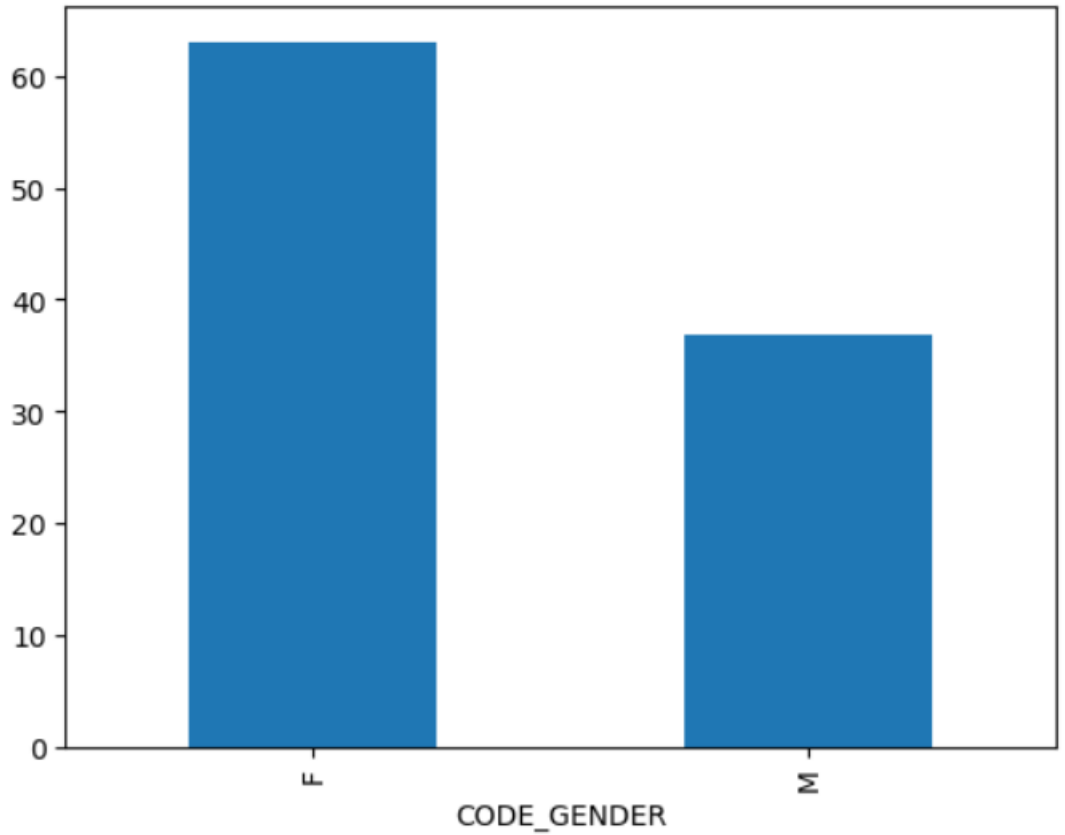
Based on organisation type

1. The clients from organization_types such as: business entity type3,self employed are most likely to take credit
2. The clients from organization_types such as: industry type 8,10,13 and trade type5,religion have no intent to take credit at all
3. Also it's reasonably good till income range upto 200000
4. This group act as a target variable for huge number of people
5. Range from 2.5l to 5l act as a target variable with less number of clients with high income



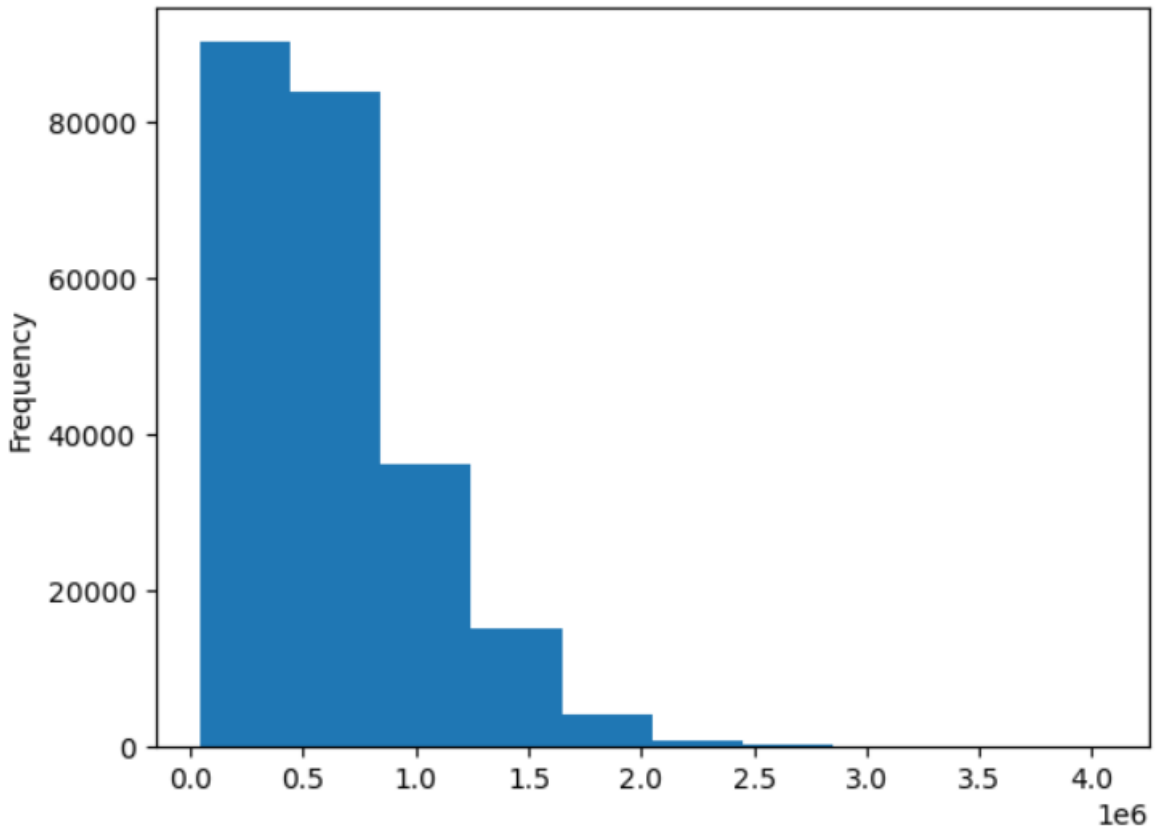
Code_gender

- The female has more chances of taking CREDIT than Male



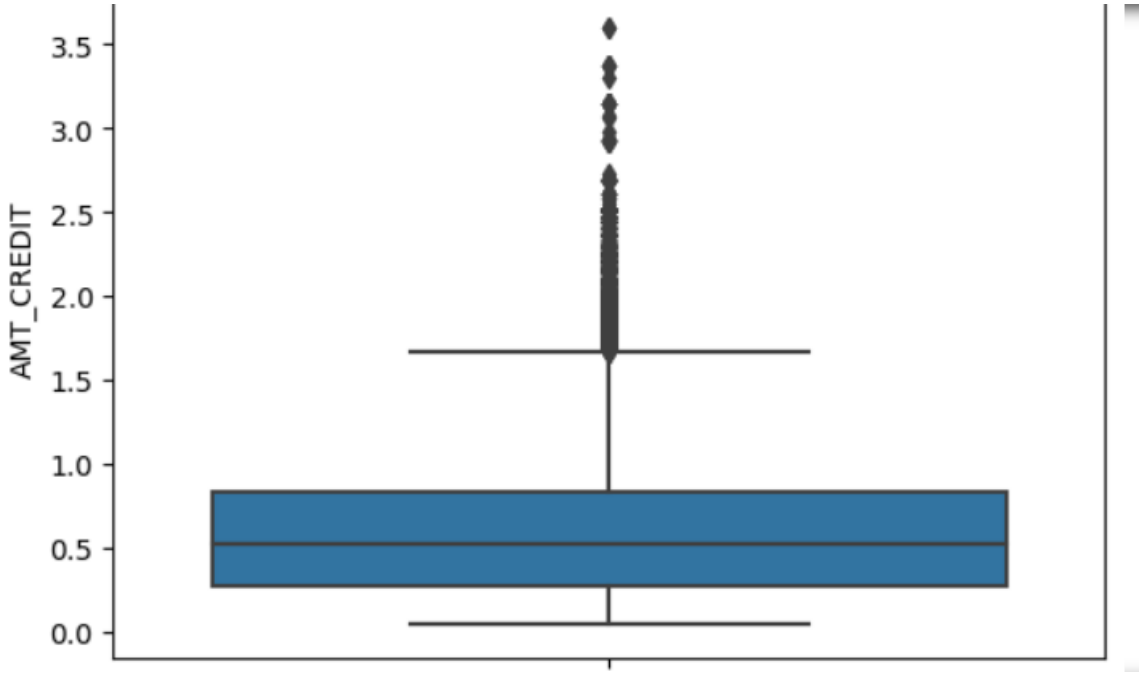
Univariate Numerical analysis for Credit_Amt

- The client having existing credit from 10k to 1L are more likely to take credit, and are likely to repay aswell
- upto 1.8L out of the entire data



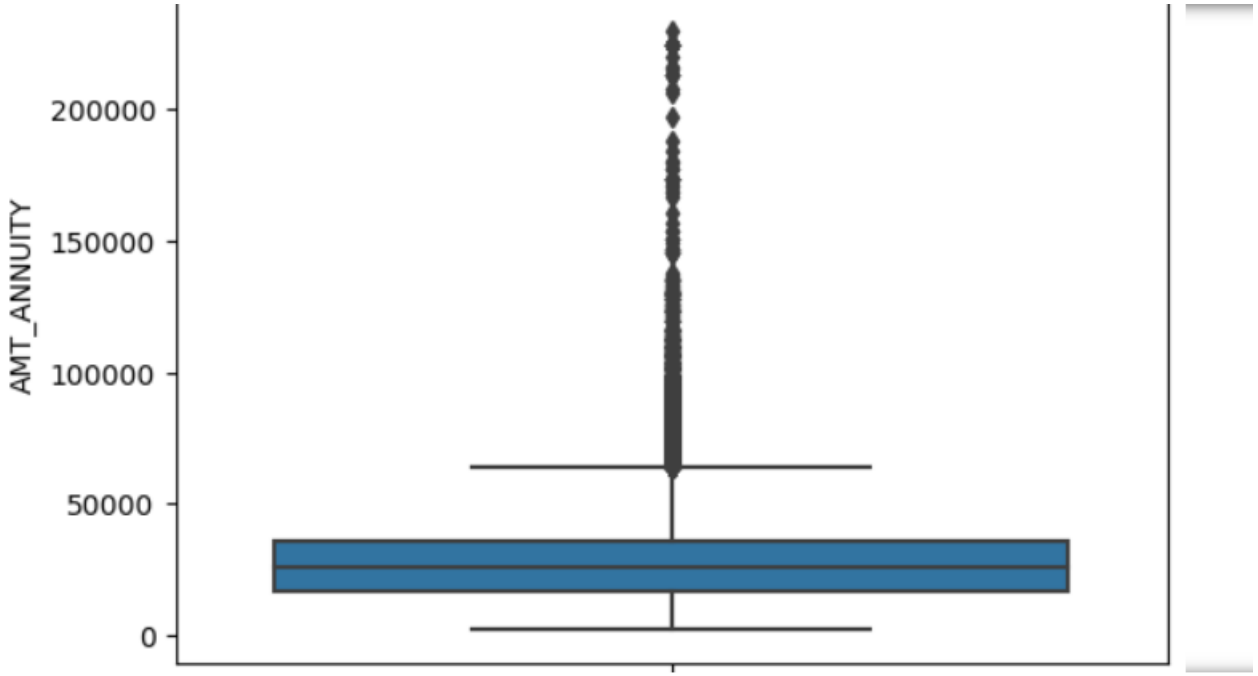
Boxplot of credit_amt

- OUTLIERS ARE NOTICED IN CREDIT AMOUNT
- THIRD QUARTILE IS BIGGER THAN FIRST QUARTILE=>MOST OF THE CREDITS OF CLIENTS ARE PRESENT IN THIRD QUARTILE



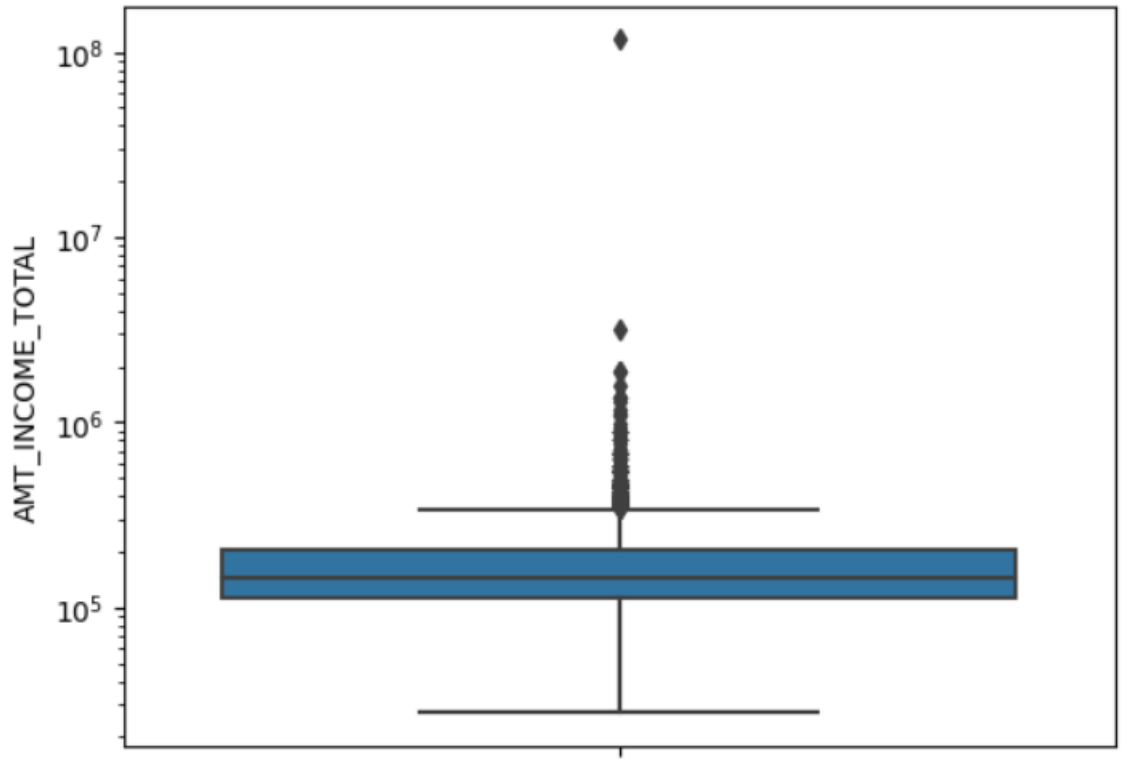
Amount annuity

1. OUTLIERS ARE NOTICED IN ANNUITY AMOUNT
2. 3RD QUARTILE BIGGER THAN FIRST QUARTILE FOR ANNUITY AMOUNT WHICH MEANS MOST OF THE ANNUITY CLIENTS ARE FROM FIRST QUARTILE



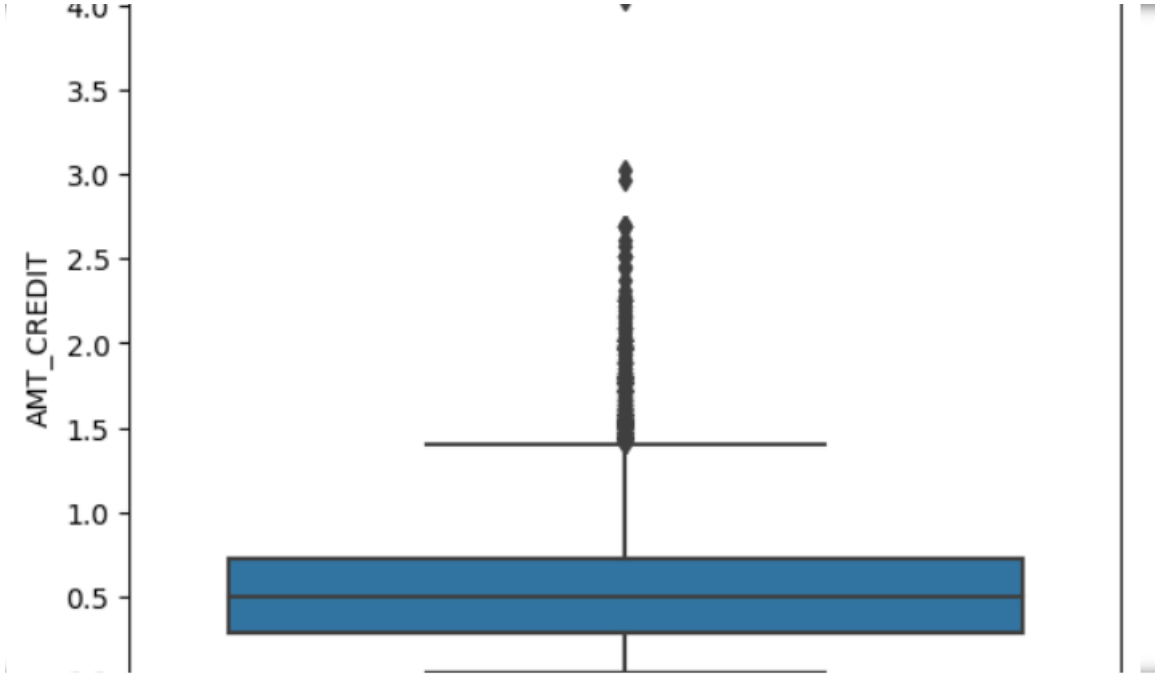
Univariate analysis for Target 1

- INCOME TOTAL
- SOME OUTLIERS ARE NOTICED IN INCOME AMOUNT



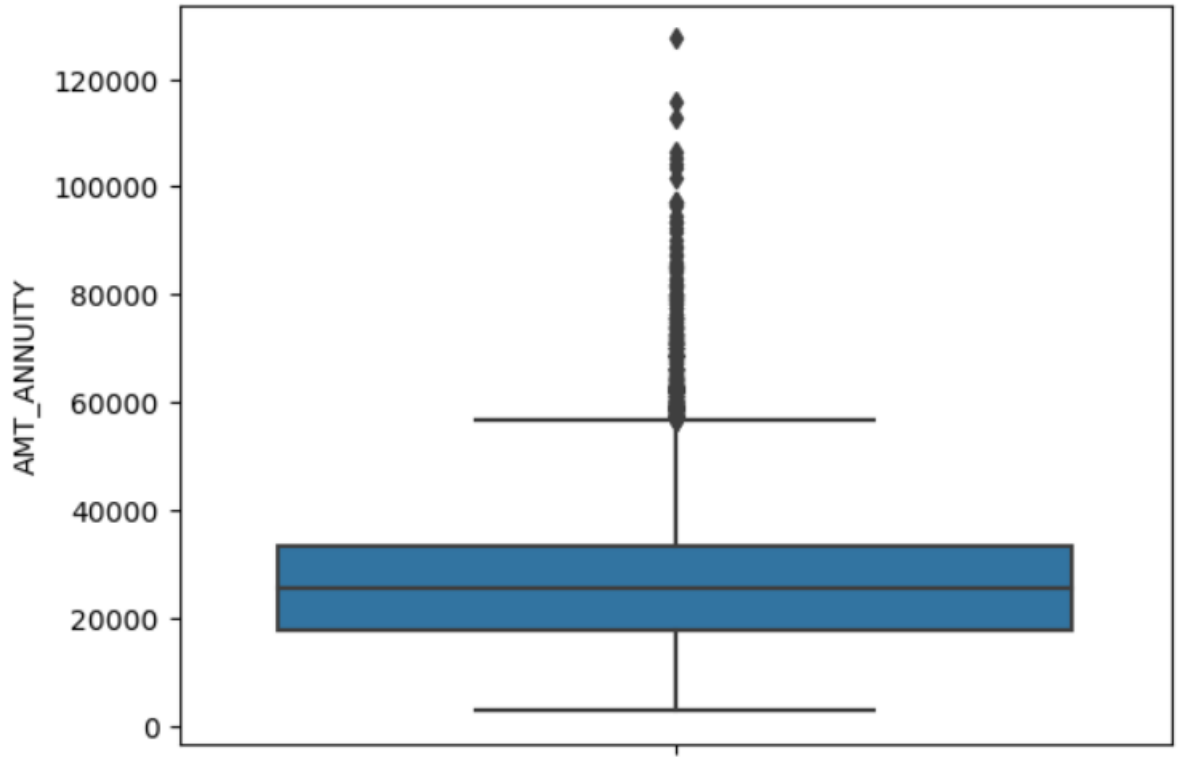
CREDIT AMOUNT

1. OUTLIERS CAN BE SEEN
2. FIRST QUARTILE IS VERY SLIM FOR CREDIT AMOUNT
3. MOST OF THE CREDITS OF CLIENTS ARE PRESENT IN THE 3RD QUARTILE



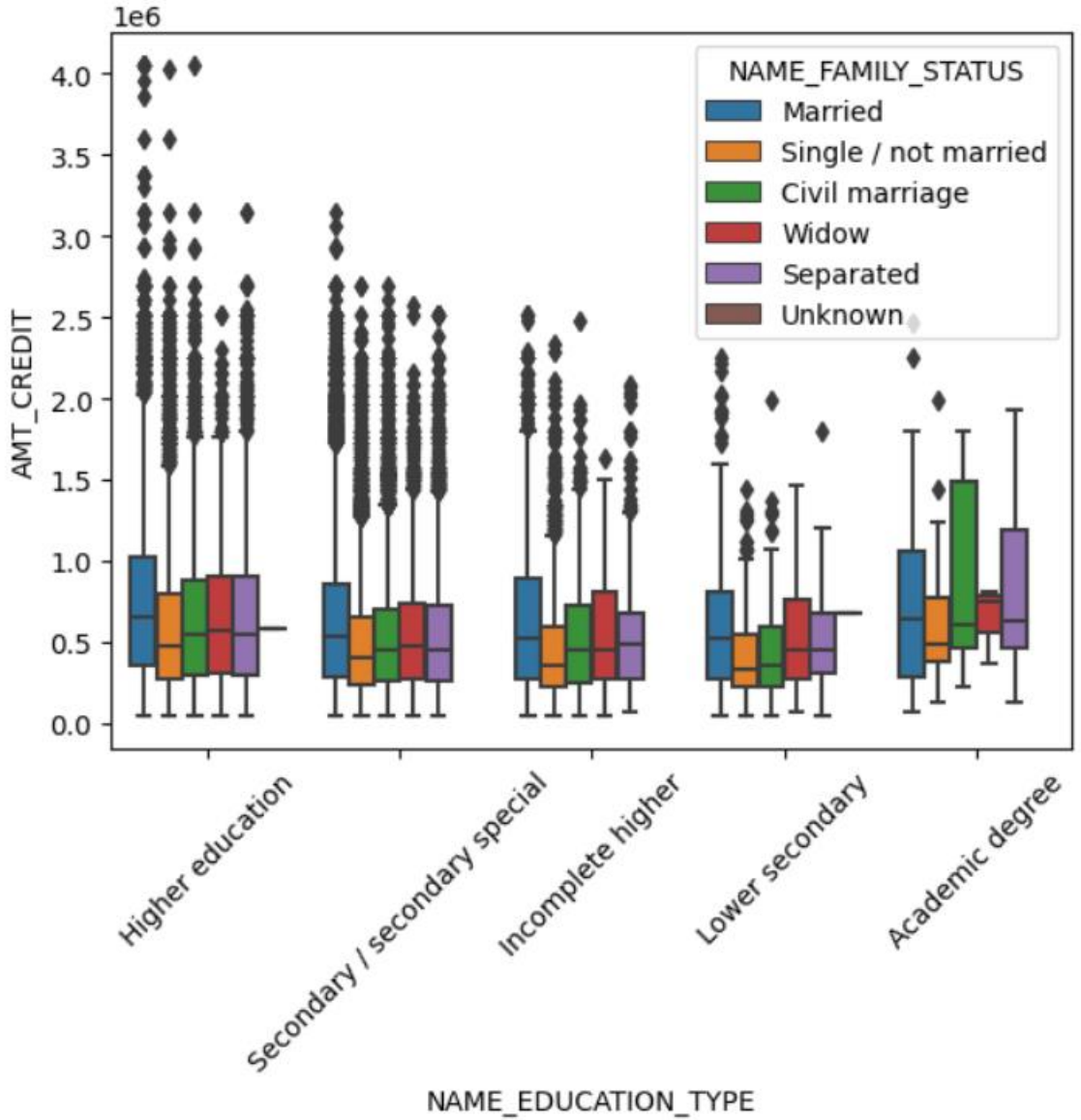
ANNUITY AMOUNT

- OUTLIERS ARE PRESENT
- 3RD QUARTILES HAS MORE CLIENTS THAT HAVE AMOUNT ANNUITY
- 35000-60000/month debts paying clients are more in number



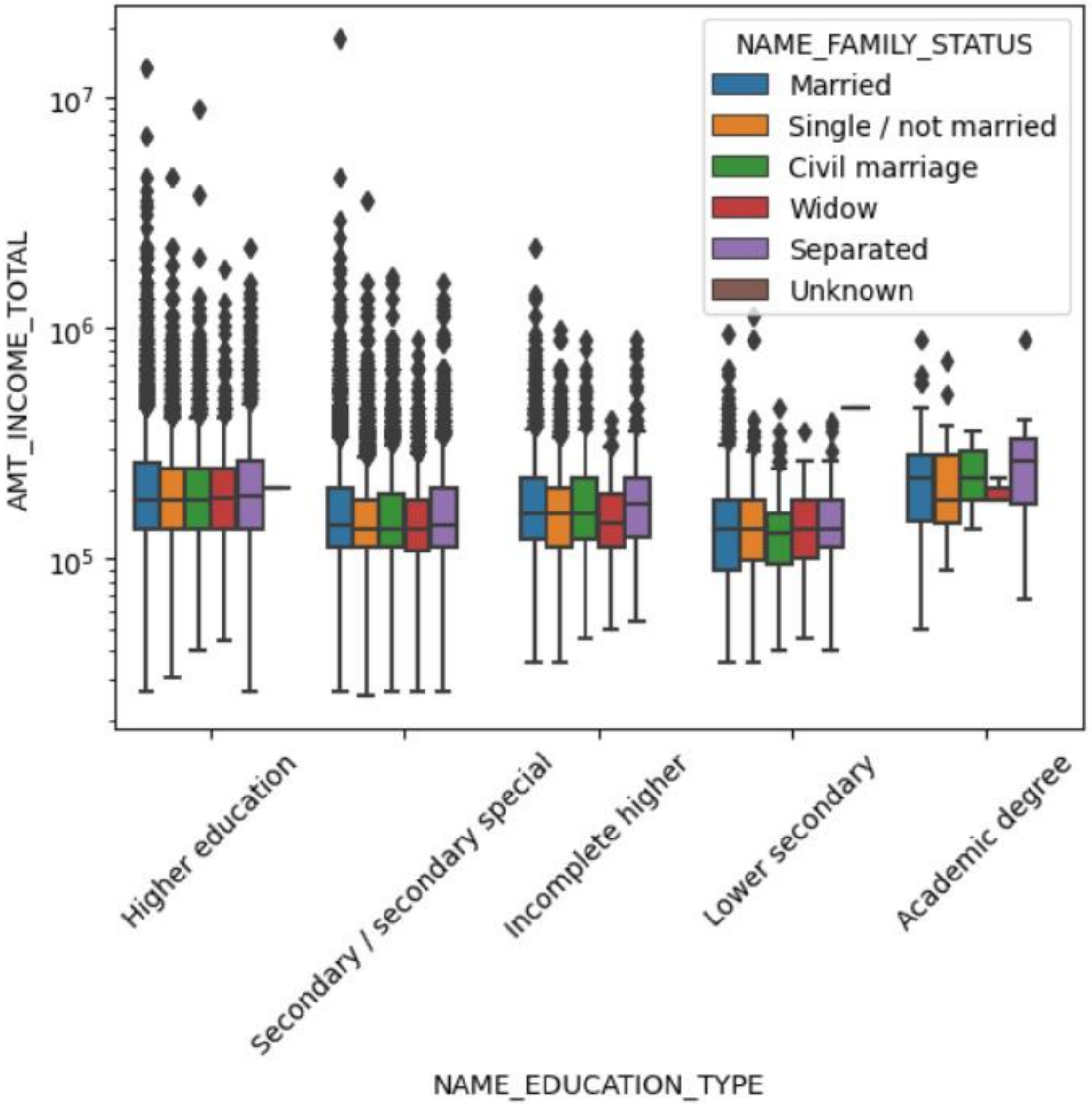
BIVARIATE ANALYSIS FOR NUMERICAL VARIABLES

1. From the above box plot we can conclude that Family status of 'civil marriage', 'marriage' and 'separated' of Academic degree education are having higher number of credits than others.
2. Also, higher education of family status of 'marriage', 'single' and 'civil marriage' are having more outliers.
3. Civil marriage for Academic degree is having most of the credits in the third quartile.



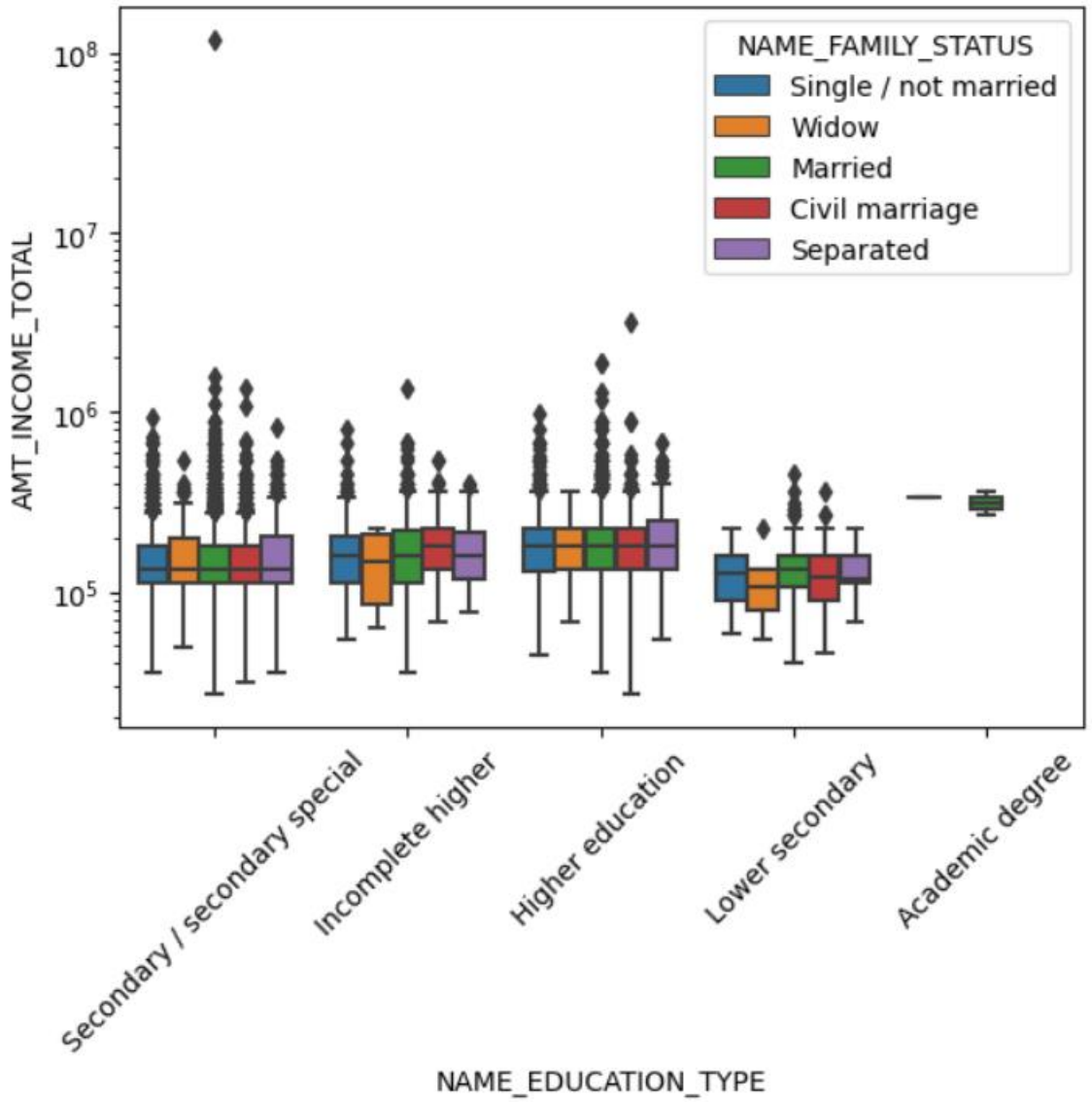
Income amount vs education status

1. HIGHER EDUCATION CLIENTS WI HAVE EQUAL INCOME.
2. LESS OUTLIERS IN ACADEMIC DE
3. INCOME AMOUNT OF ACADEMI HIGHER EDUCATION
4. Lower secondary of civil marriag amount than others.



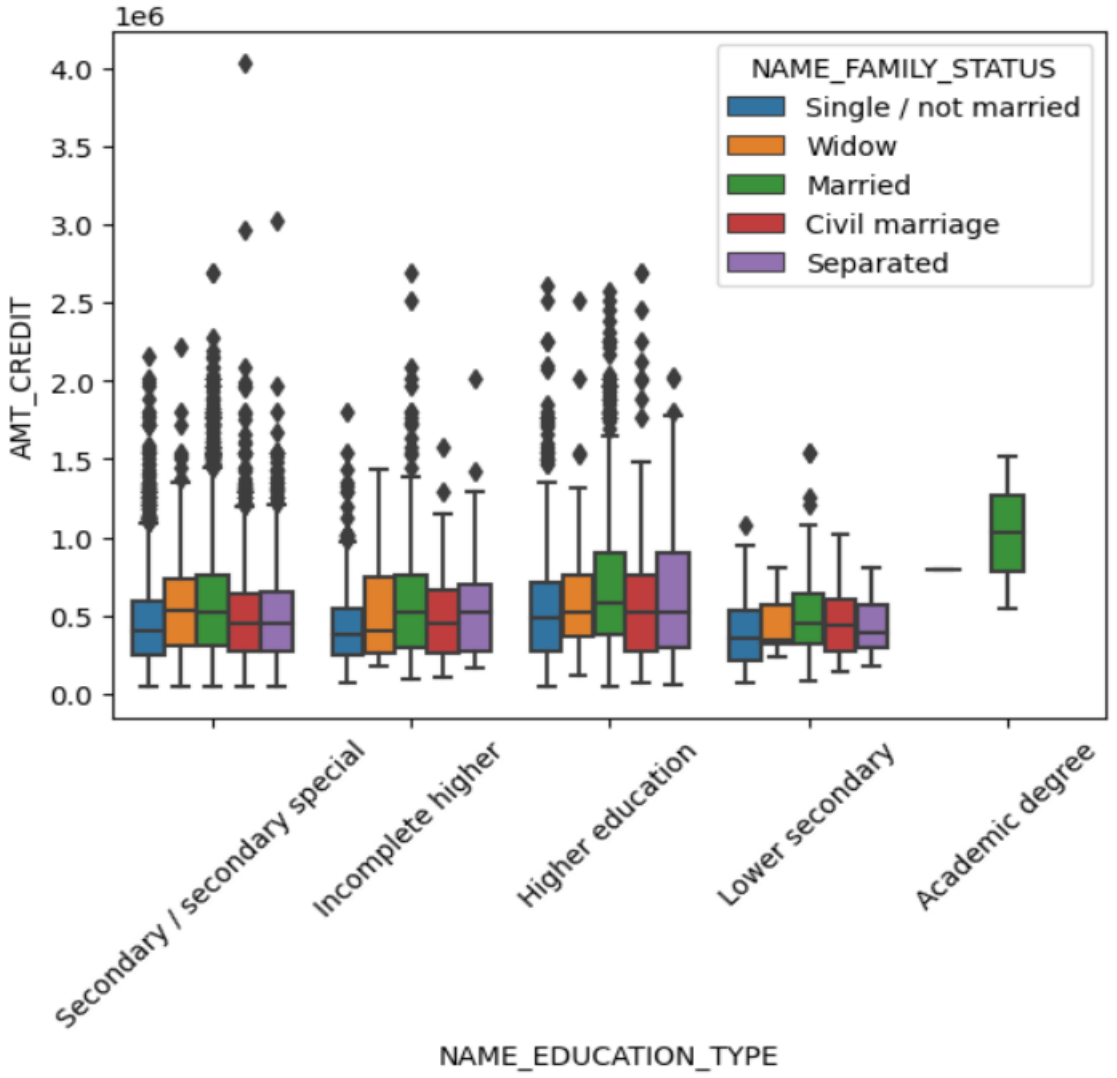
BIVARIATE ANALYSIS FOR NUMERICAL VARIABLES(for target1)

1. Have some similarity with Target0.
2. From above boxplot for Education type 'Higher education' the 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.
3. Lower secondary are have less income amount than others.

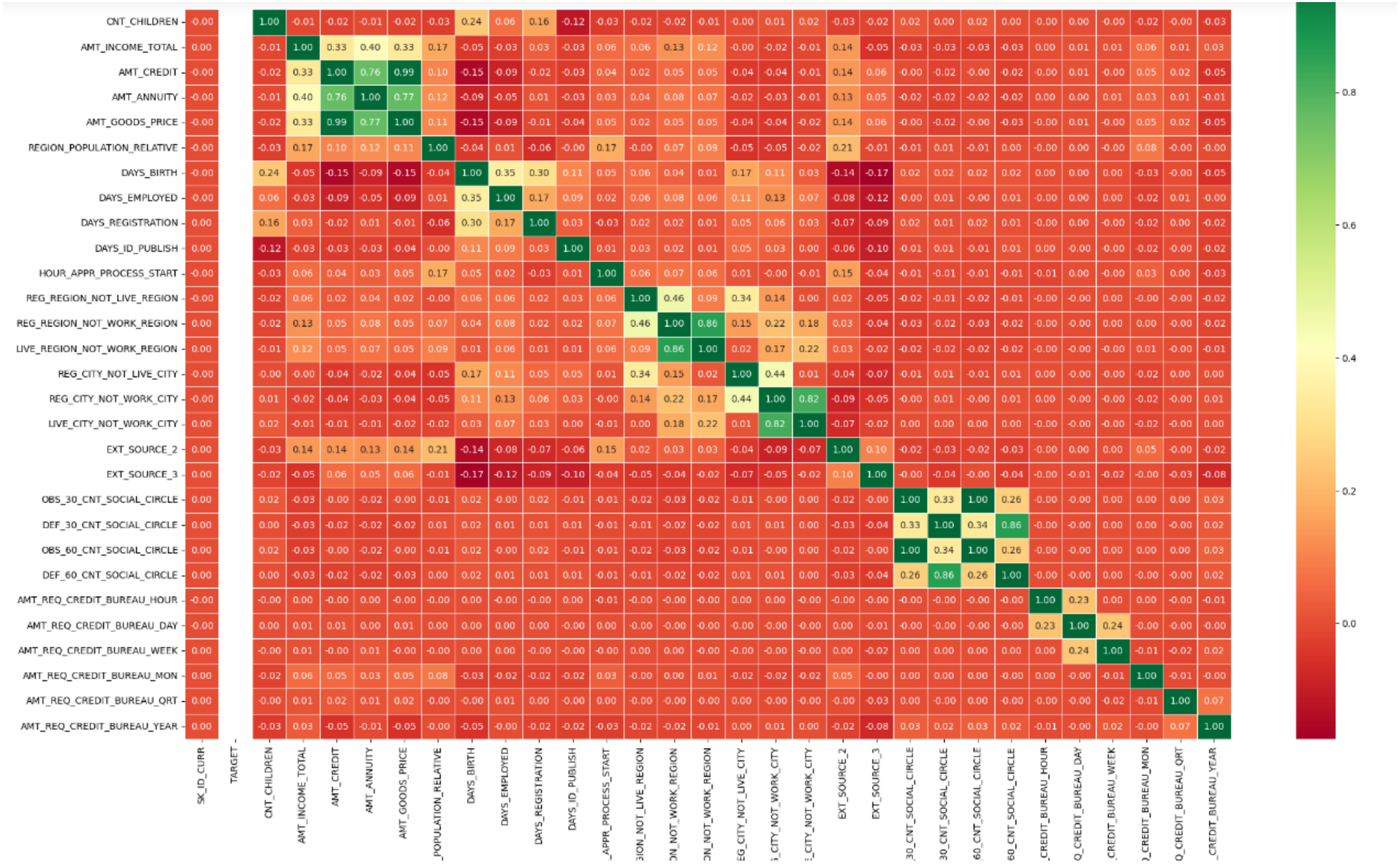


Credit amount vs Education status

1. FROM ACADEMIC BACKGROUND "ACADEMIC DEGREE"WE HAVE ONLY CLIENTS FROM FAMILY BACKGROUND " MARRIED"
2. CLIENTS HAVE MORE CREDIT UNDER FAMILY STATUS"MARRIED" IN ALL EDUCATION TYPE



Multivariate analysis-**for target0**

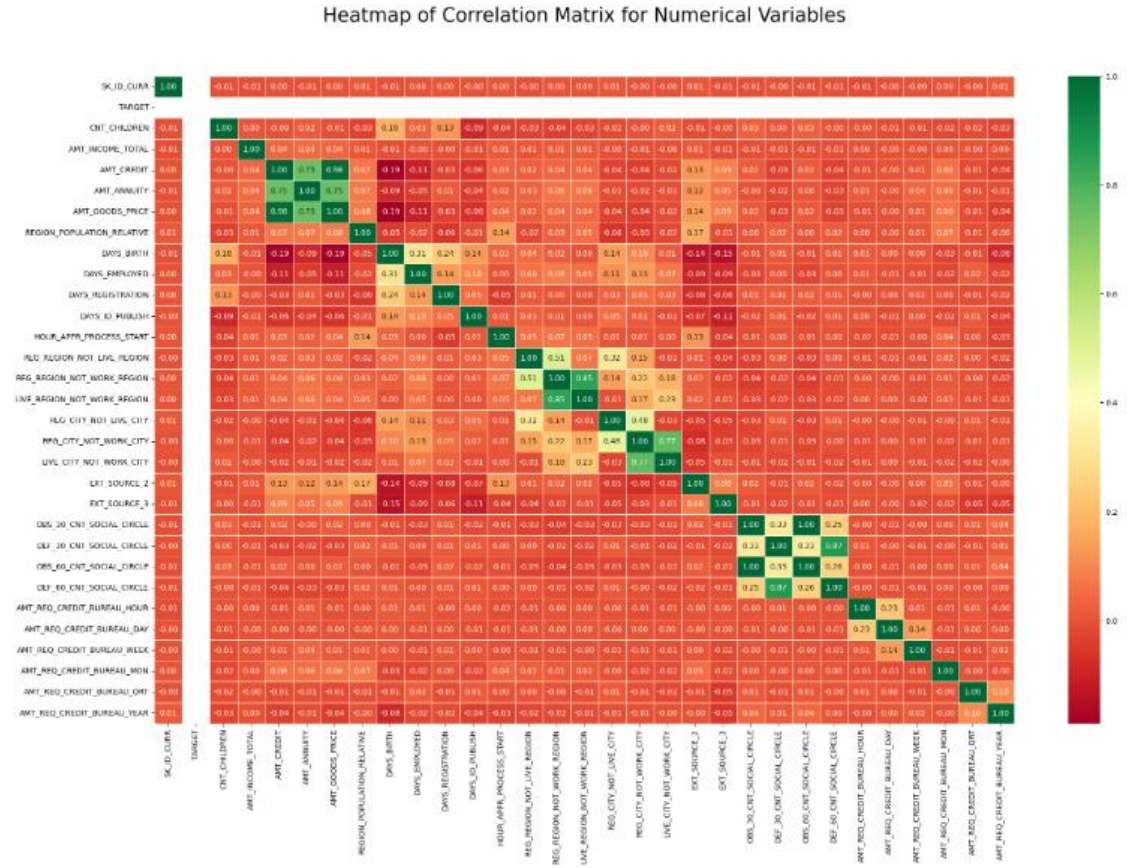


Cont.d

- 1. Credit amount is inversely proportional to the date of birth, which means Credit amount is higher for low age and vice-versa.
- 2. 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.
- 3. Income amount is inversely proportional to the number of children client have, means more income for less children client have and vice-versa.
- 4. less children client have in densely populated area.
- 5. Credit amount is higher to densely populated area.
- 6. The income is also higher in densely populated area.

Multivariate analysis for target1

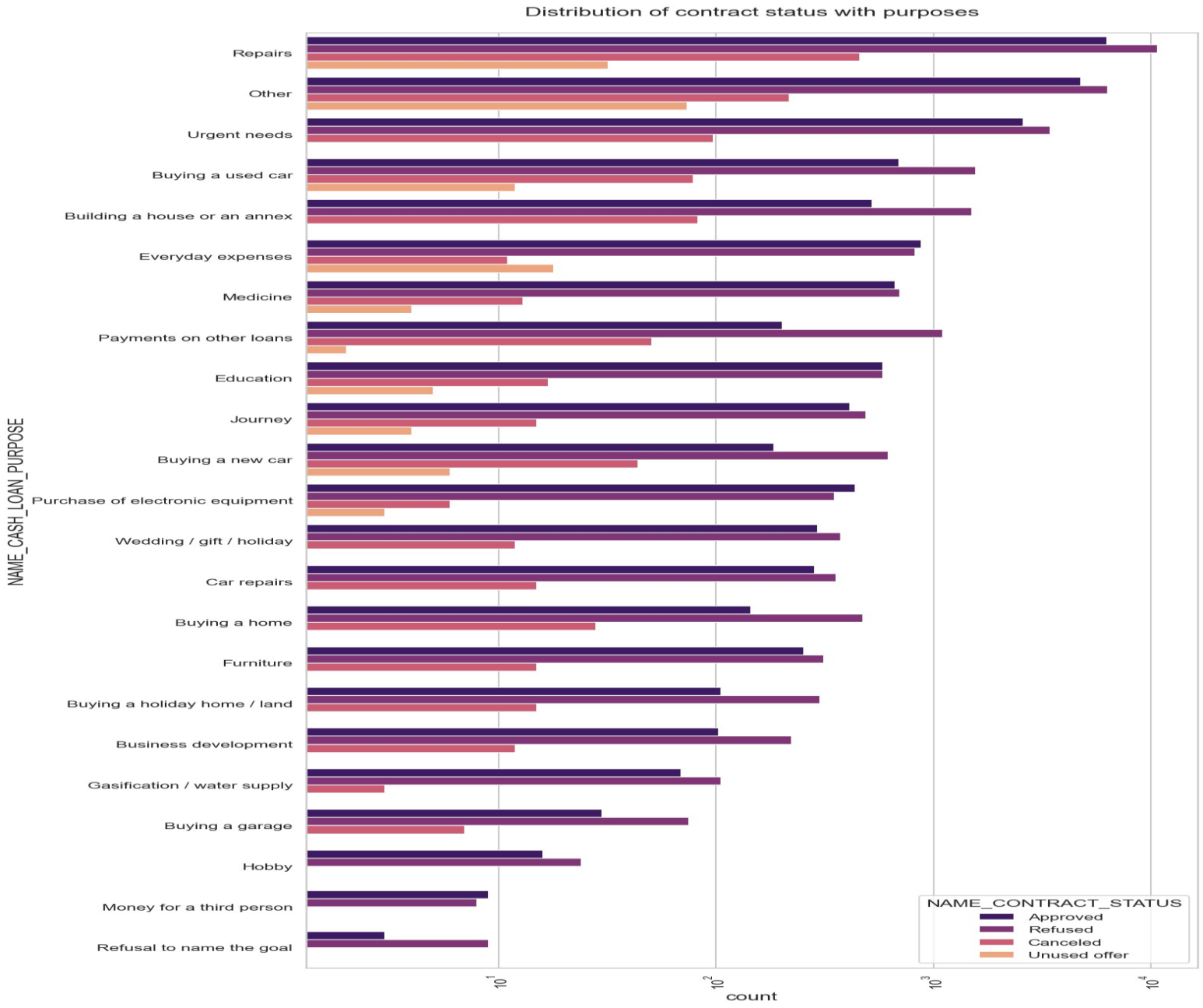
- almost same analysis as per target0
- for categorical variables of target1



Previous application

Previous_application Univariate analysis

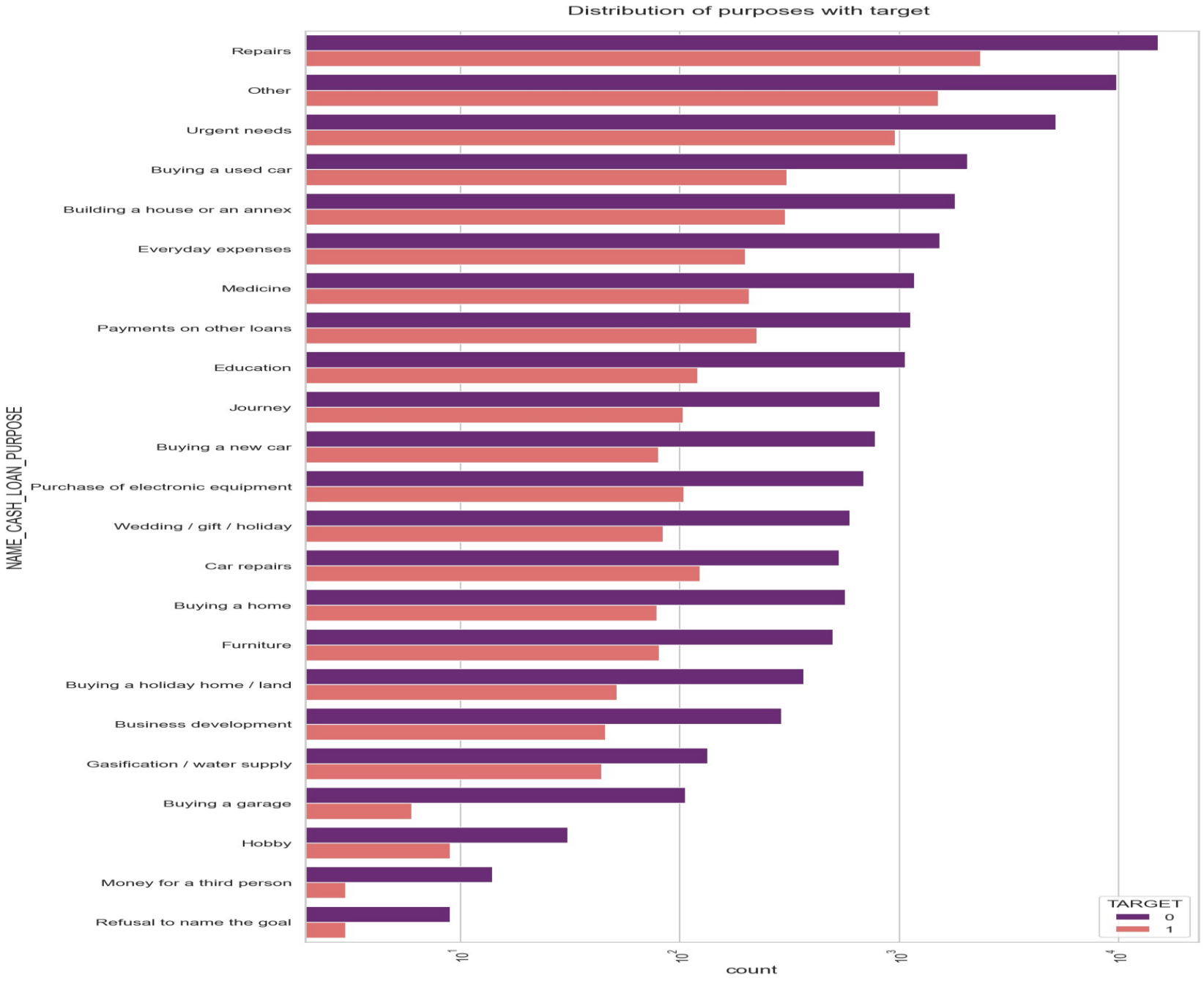
1. Most rejection of loans came from purpose 'repairs'.
2. For education purposes we have equal number of approves and rejection
3. Payign other loans and buying a new car is having significant higher rejection than approves.



Dist.of purposes with target

- Few points we can conclude from above plot:

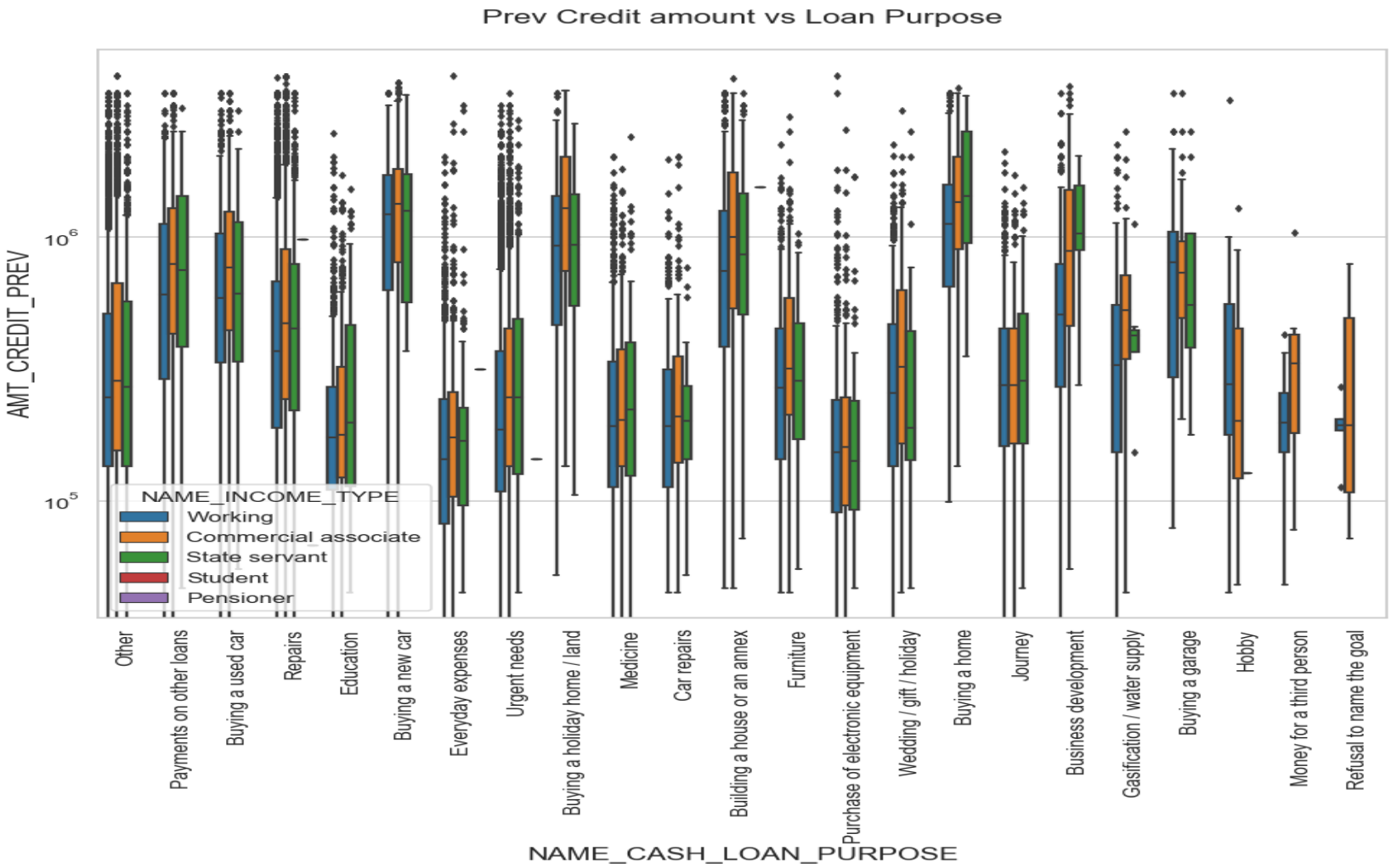
1. Loan purposes with 'Repairs' are facing more difficulites in payment on time.
2. There are few places where loan payment is significant higher than facing difficulties.
3. They are 'Buying a garage', 'Business developemt', 'Buying land','Buying a new car' and 'Education'
4. Hence we can focus on these purposes for which the client is having for minimal payment difficulties.



Prev credit amount vs Loan purpose

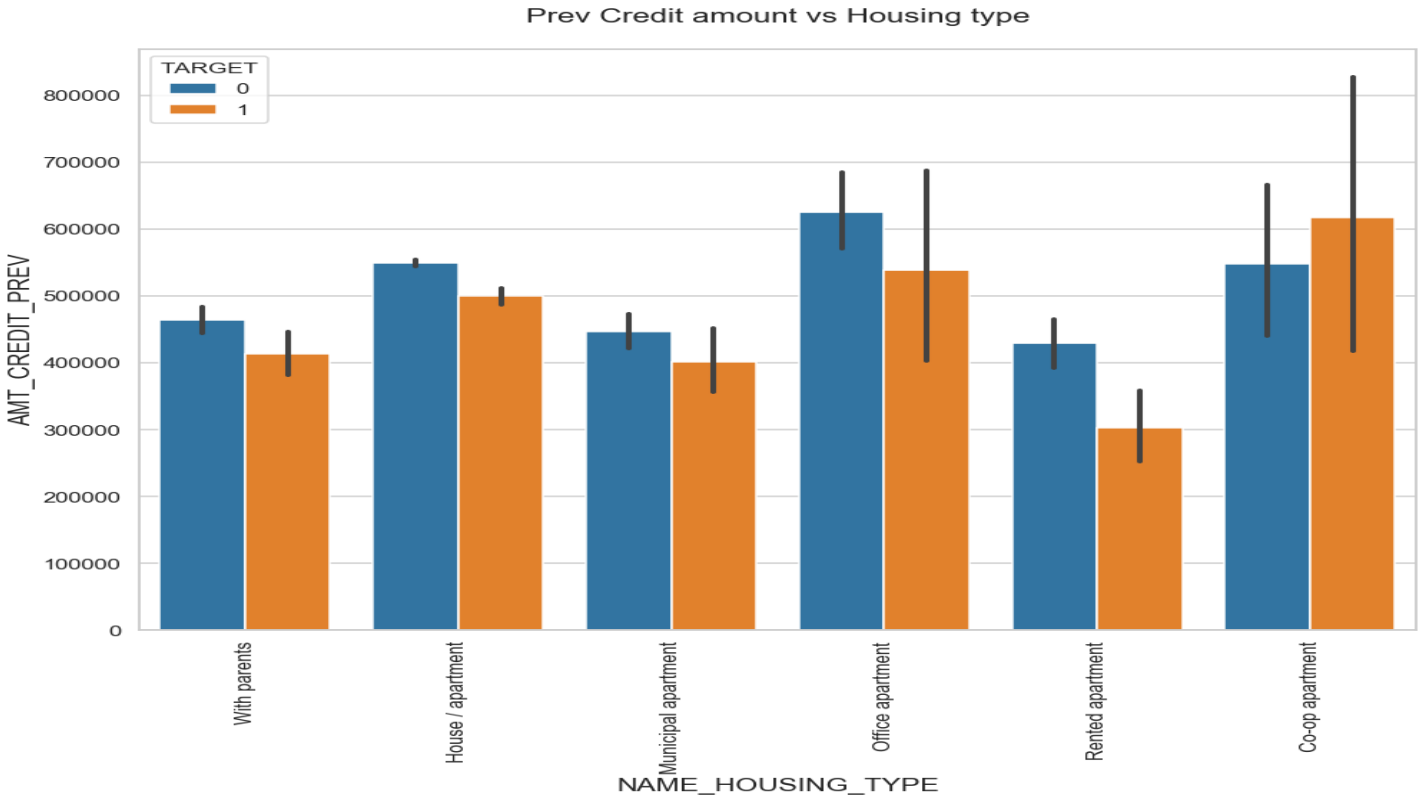
From the above we can conclude some points-

1. The credit amount of Loan purposes like 'Buying a home','Buying a land','Buying a new car' and'Building a house' is higher.
2. Income type of state servants have a significant amount of credit applied
3. Money for third person or a Hobby is having less credits applied for.



Prev vs housing type

- Here for Housing type, office appartment is having higher credit of target 0 and co-op apartment is having higher credit of target 1.
- So, we can conclude that 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 or House\appartment or miuncipal appartment for successful payments.



Conclusion

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

2. Banks should focus less on income type ‘Working’ as they are having most number of unsuccessful payments.

3. Also with loan purpose ‘Repair’ is having higher number of unsuccessful payments on time.

4. Get as much as clients from housing type ‘With parents’ as they are having least number of unsuccessful payments.