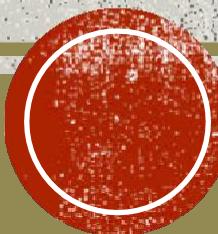


CREDIT EDA ASSIGNMENT



By Devanthi Bagh

CATEGORICAL UNIVARIATE ANALYSIS FOR TARGET 0



DISTRIBUTION OF INCOME RANGE

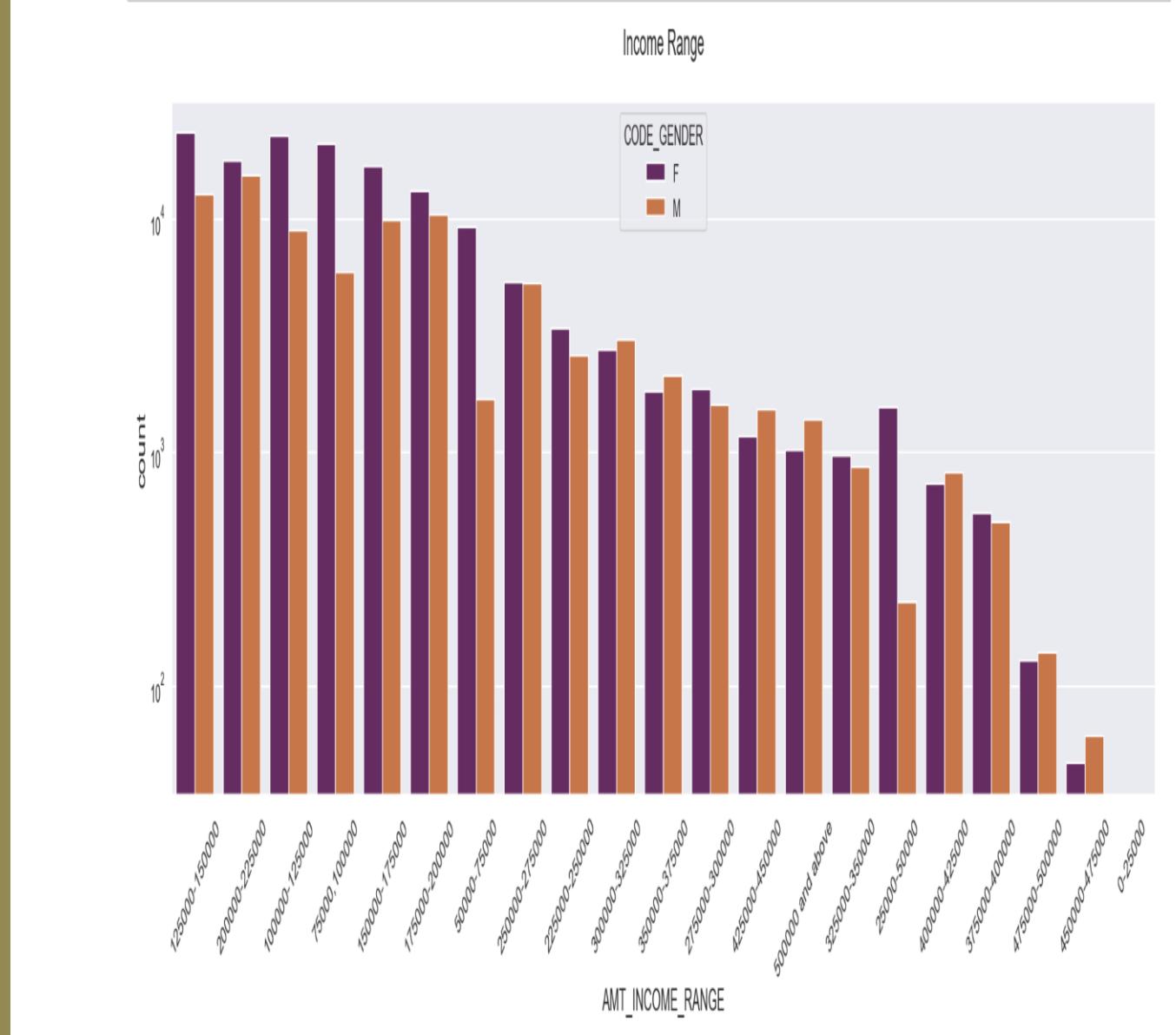
Observations:

=>The range of income between 100k and 200k shows a higher impact of credited instances.

=> Female counts surpass those of Males, In 10 sections, although there exist instances where males have higher counts, with one section having equal counts for both genders.

=>Within this income range, females predominantly receive credits.

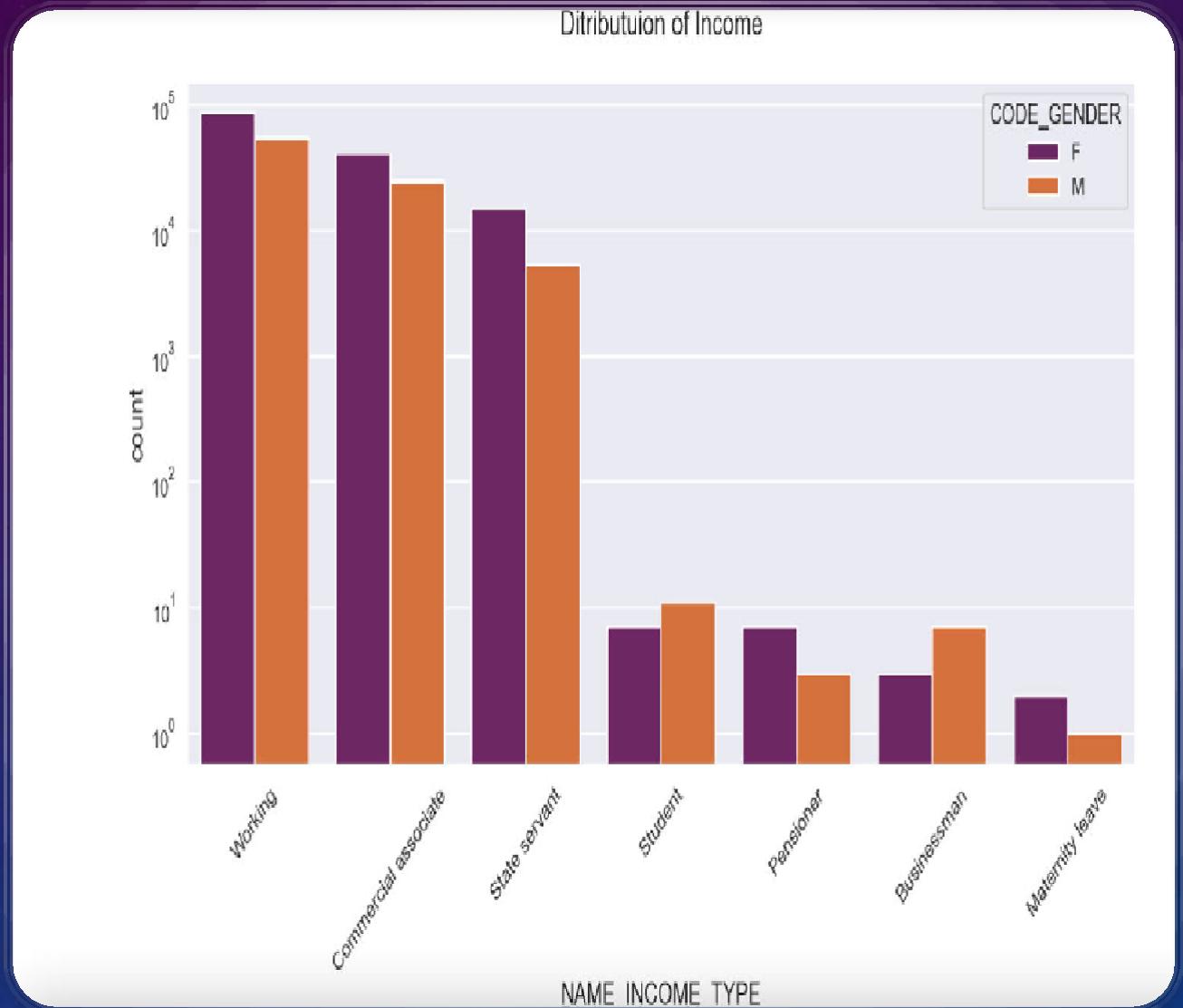
=>Instances of income exceeding 400k must be noted.



DISTRIBUTION OF INCOME TYPE

Observations from the Graph:

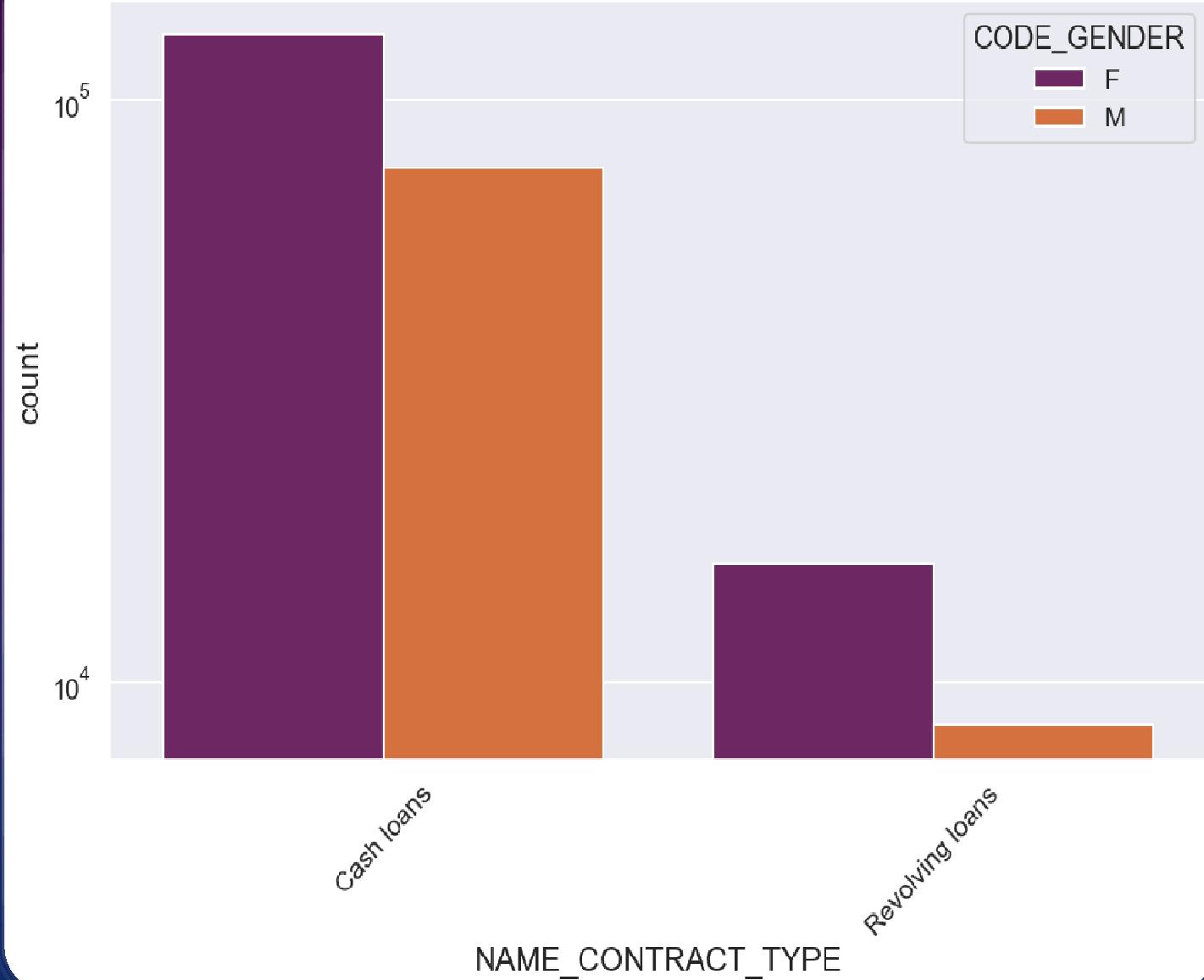
- Notably, income types like 'student', 'pensioner', and 'Businessman' do not have any instances of late payments, indicating a lack of occurrences in this context
- Among income types such as 'working', 'commercial associate', and 'State Servant', the count of credits is higher and compared to 'Maternity leave'.
- Additionally, females have a greater count of credits than males in these income types.
- 'Maternity leave' shows a lower number of credits.



DISTRIBUTION FOR CONTRACT TYPE

- The number of credits is greater for the 'cash loans' contract type compared to the 'Revolving loans' contract type.
- Similarly, in this context, female applicants are demonstrate a higher tendency to apply for credits.

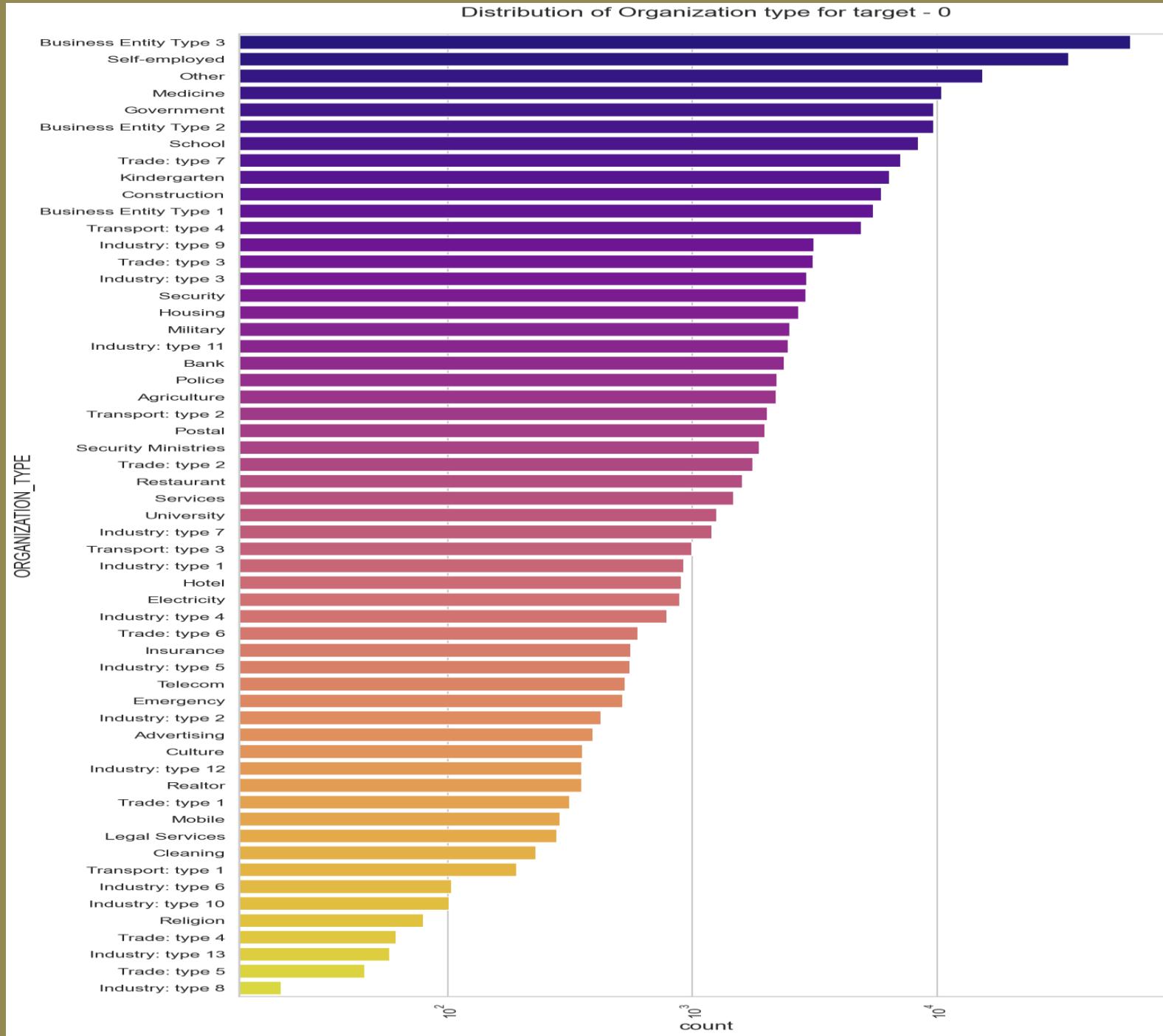
Distribution of contract type



DISTRIBUTION OF ORGANIZATION TYPE

Main observations from the graph:

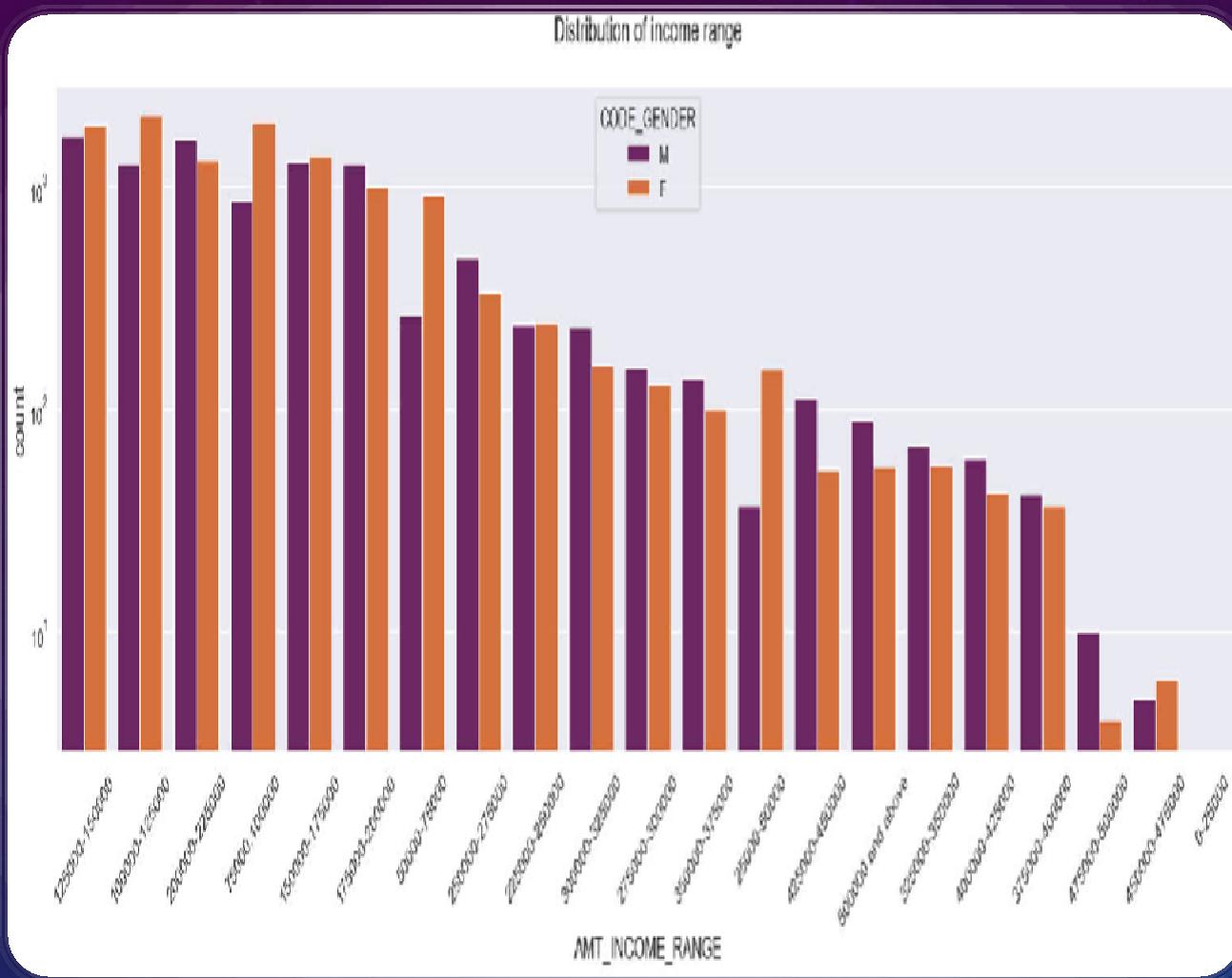
- Credit applicants primarily belong to diverse organization types, including 'Business entity Type 3', 'Self employed', 'Other', 'Medicine', and 'Government'.
- There is a lower representation of clients originating from the Industry types 8, 6, and 10, as well as specific categories such as religion and trade type 5 and 4.



CATEGORICAL UNIVARIATE ANALYSIS TARGET 1



DISTRIBUTION OF INCOME RANGE



observations from graph:

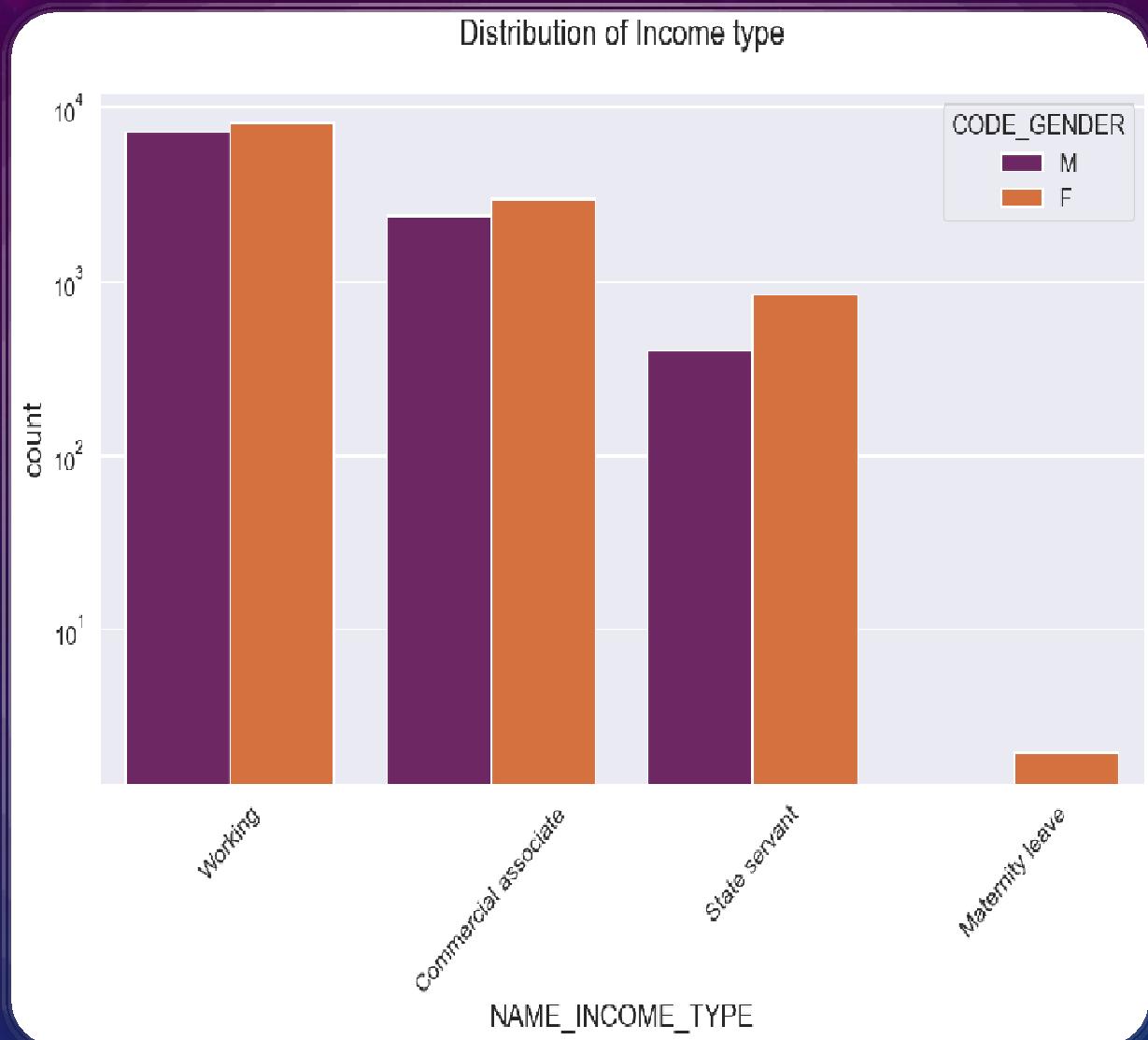
- The credit count is notably higher within the income range of 100000 to 200000.
- Male population surpasses the count of females.
- This visualization underscores the predominance of males in securing credits within this income bracket.
- There is a minimal occurrence of instances within the income range of 400000 and above.



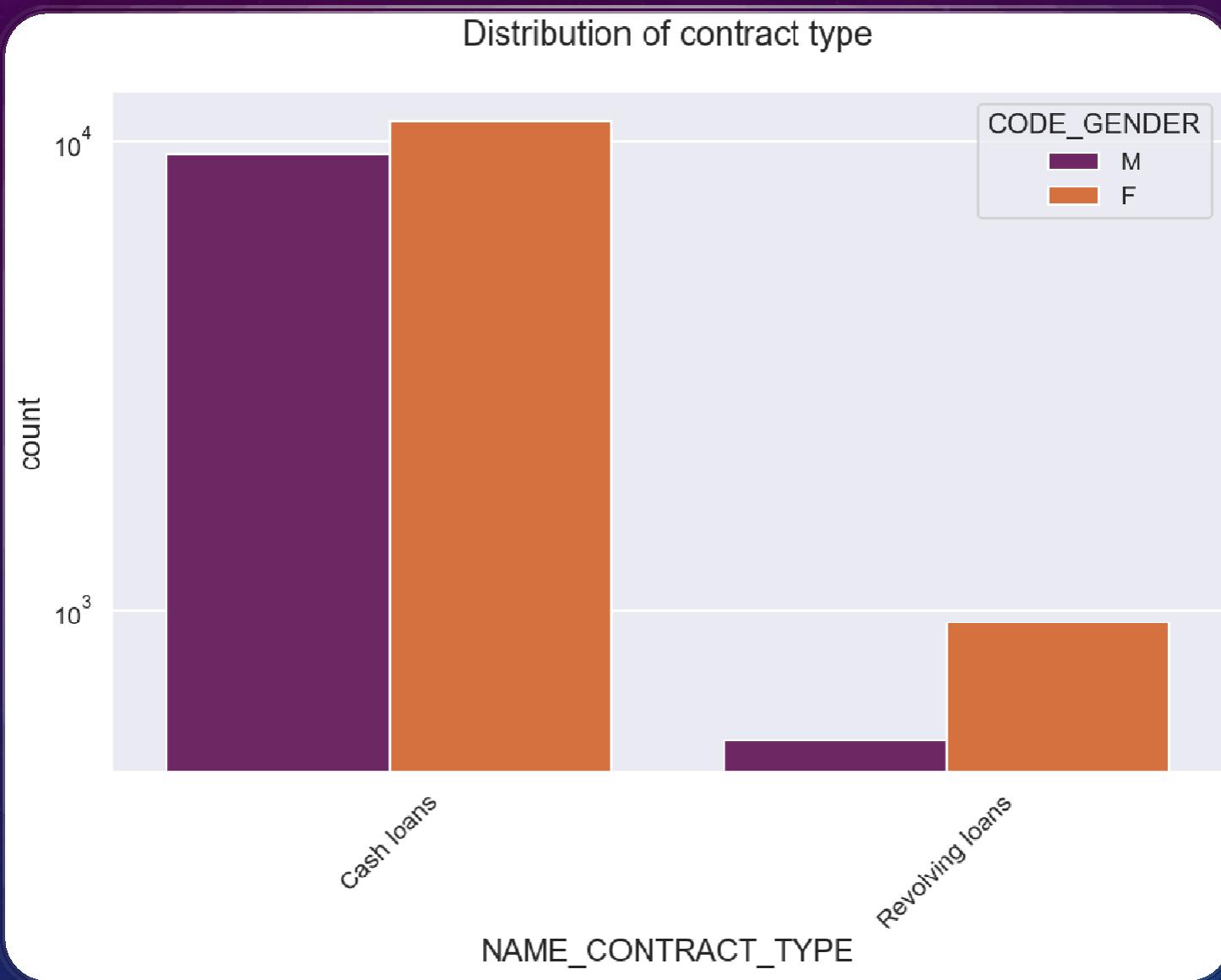
DISTRIBUTION OF INCOME TYPE

Key insights drawn from the graph:

- The prevalence of credits is notably higher for income types labeled as 'working', 'commercial associate', and 'State Servant' compared to other categories like 'Maternity leave'.
- Within this context, females exhibit a higher credit count compared to males.
- A limited number of credits are associated with the 'Maternity leave' income type.
- Notably, for income type '1', there are no instances of 'student', 'pensioner', or 'Businessman', implying their absence in delayed payments.

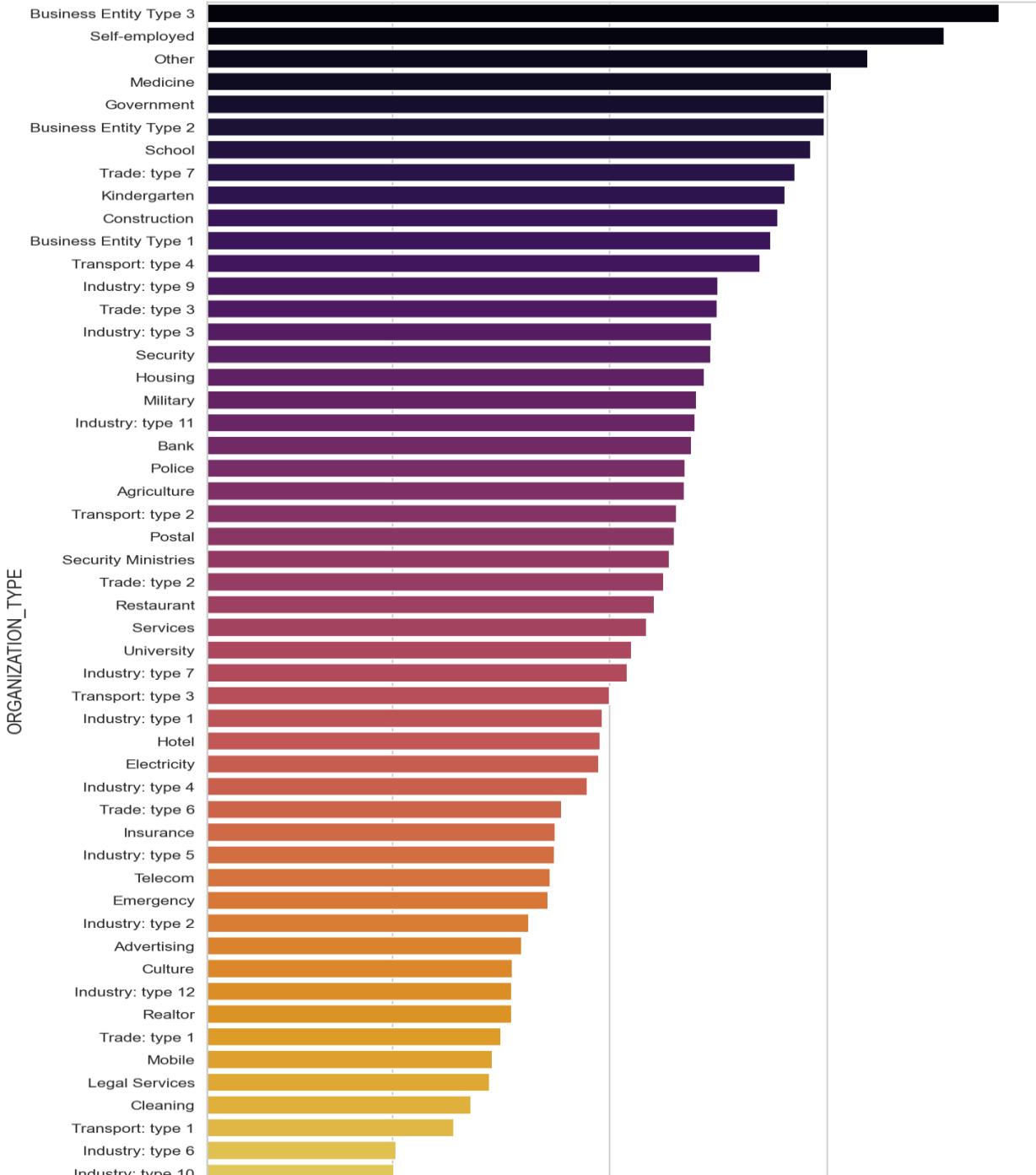


DISTRIBUTION FOR CONTRACT TYPE



- Points are concluded from the above graph.
- 1. For contract type ‘cash loan’ is having more number of credits than ‘Revolving loans’ contract type.
- 2. For this also Female is leading for applying credits.
- 3. For type 1 :there is only Female Revolving loans.

Distribution of Organization type for target - 1



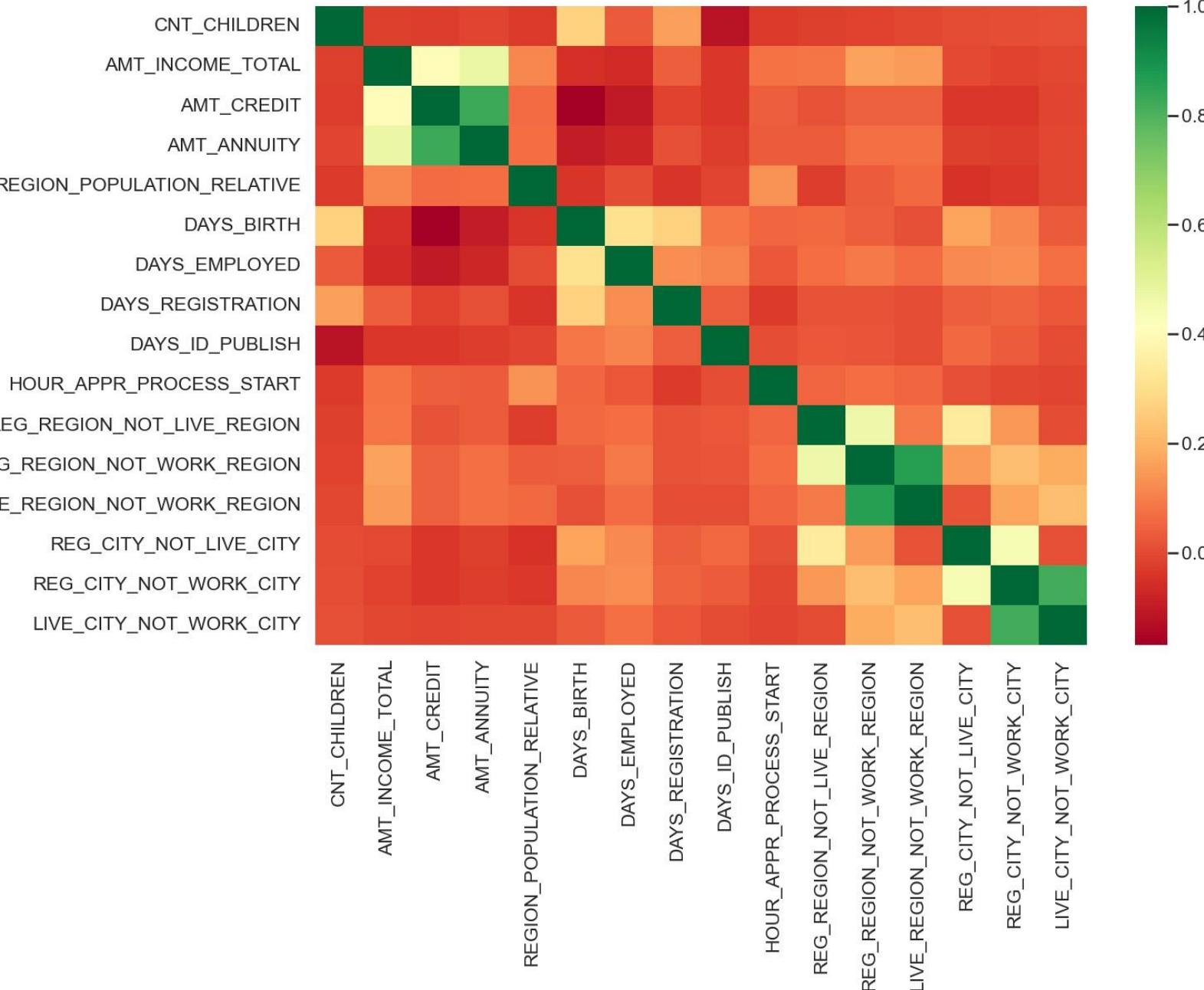
DISTRIBUTION OF ORGANIZATION TYPE

Key observations extracted from the provided graph:

- Credit applicants predominantly originate from a diverse set of organization types, notably including 'Business entity Type 3', 'Self employed', 'Other', 'Medicine', and 'Government'.
- In contrast, there is a lower representation of clients hailing from Industry types 8, 6, and 10, as well as specific categories such as religion and trade type 5 and 4.
- Similar patterns in the distribution of organization types are discernible for type 0.
-



Correlation for target 0



CORRELATION OF TARGET
0



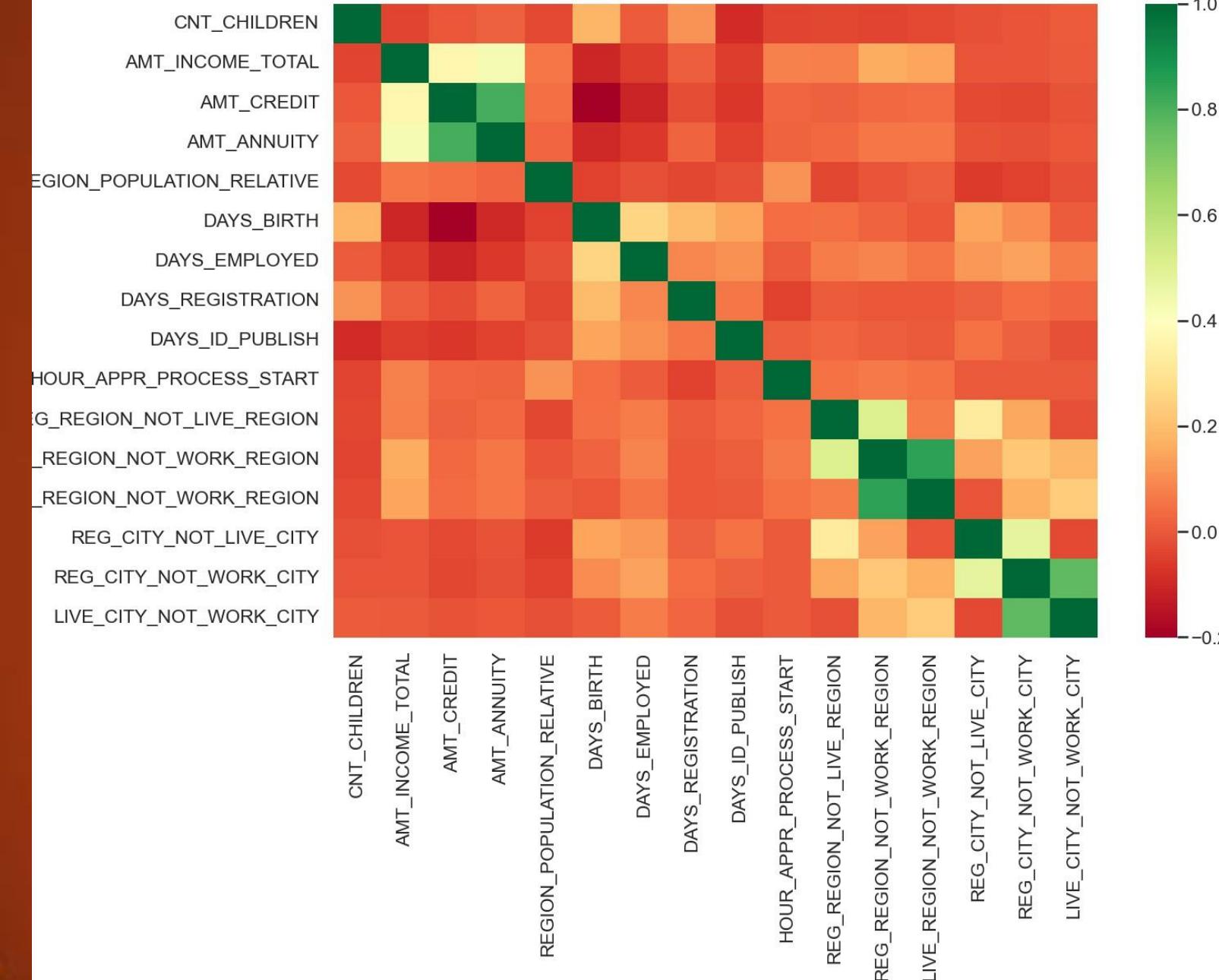
CORRELATION FOR TARGET 0

Analyzing the correlation heatmap above, we can deduce several noteworthy observations:

- There exists an inverse relationship between credit amount and the client's age, implying that lower age is associated with higher credit amounts and vice versa.
- Credit amount demonstrates an inverse correlation with the number of children a client has, suggesting that clients with fewer children tend to have higher credit amounts, and vice versa.
- The income amount showcases an inverse correlation with the number of children a client has; clients with fewer children tend to have higher income, and vice versa.
- Densely populated areas seem to have fewer clients with children.
- Higher credit amounts are linked with densely populated areas.
- Densely populated areas exhibit higher income levels as well.



Correlation for target 1



CORRELATION OF TARGET

1

CORRELATION FOR TYPE 1

The correlation heatmap for Target 1 reveals similar patterns to Target 0, but with a few distinct observations:

- Clients whose permanent address does not match the contact address tend to have fewer children, and the reverse is also true.
- Clients whose permanent address does not match the work address also tend to have fewer children, and the converse holds as well.

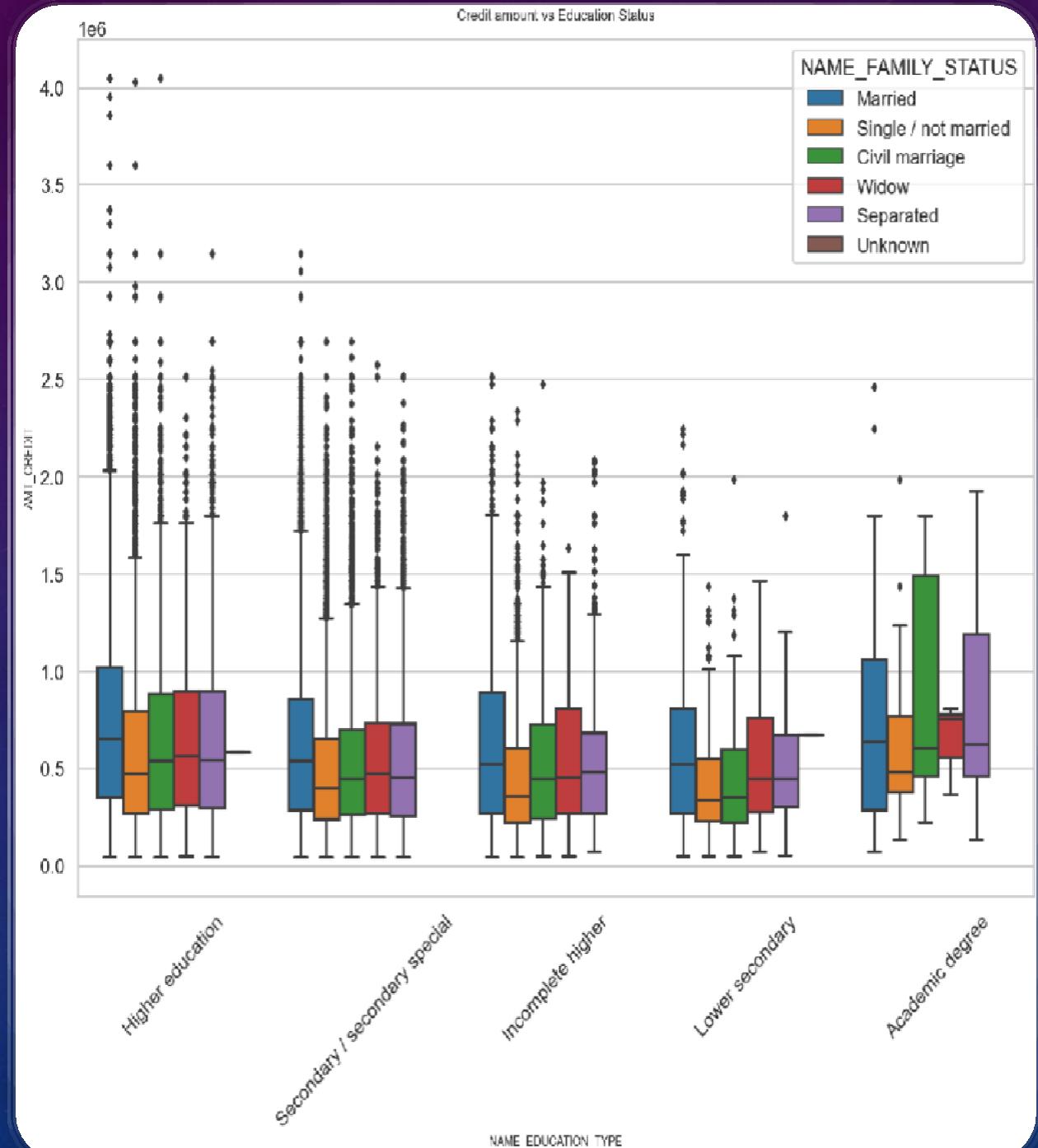


BIVARIATE ANALYSIS FOR TYPE 0



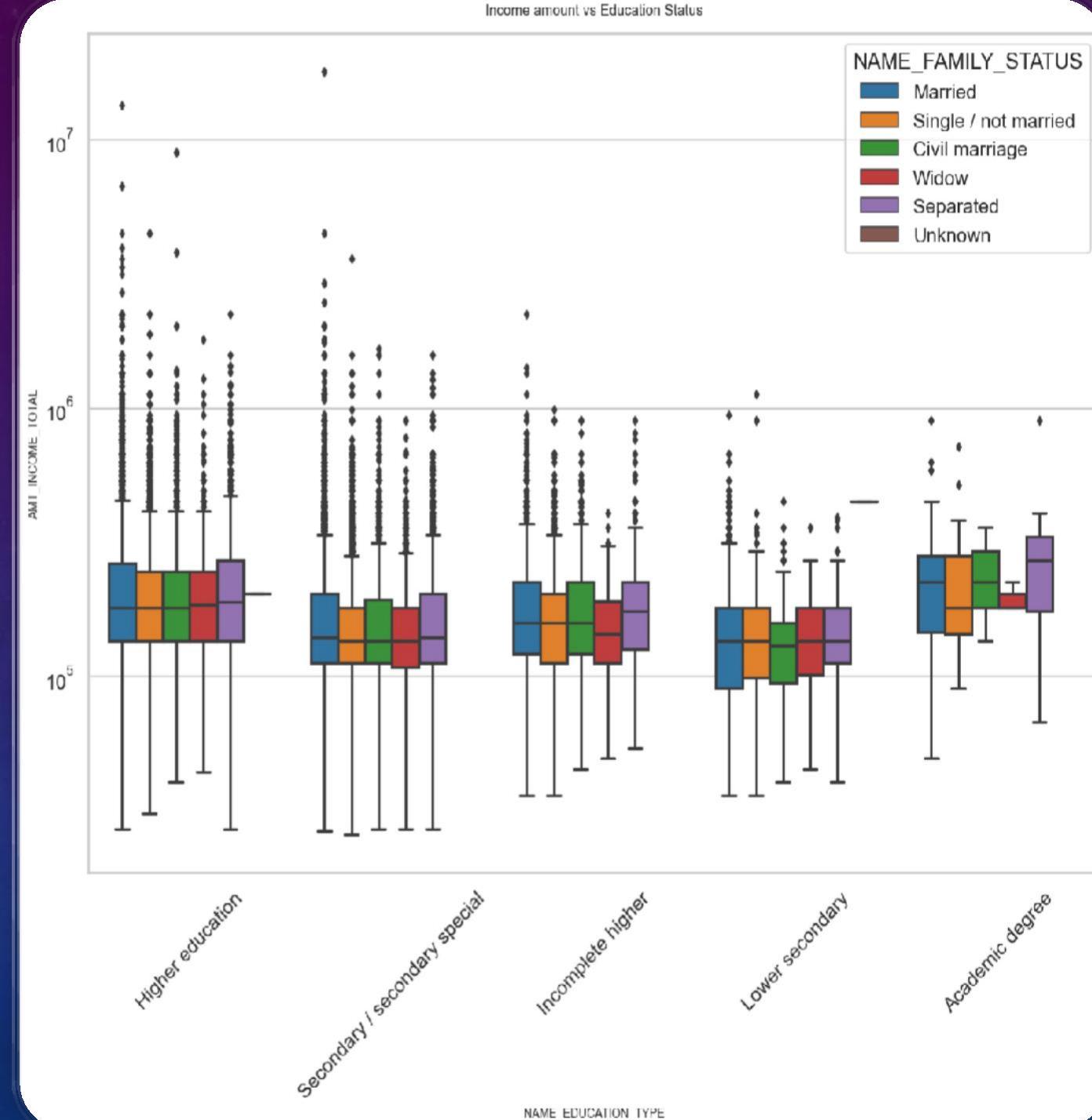
CREDIT AMOUNT VS EDUCATION STATUS

- Among the various family statuses, individuals with 'civil marriage', 'marriage', and 'separated' tend to have a higher frequency of credits when possessing an Academic degree education.
- Notably, the 'marriage', 'single', and 'civil marriage' family statuses with higher education levels exhibit a greater number of outliers.
- In the case of 'Academic degree' education, 'civil marriage' emerges as the dominant category, predominantly situated within the third quartile in terms of credit distribution.



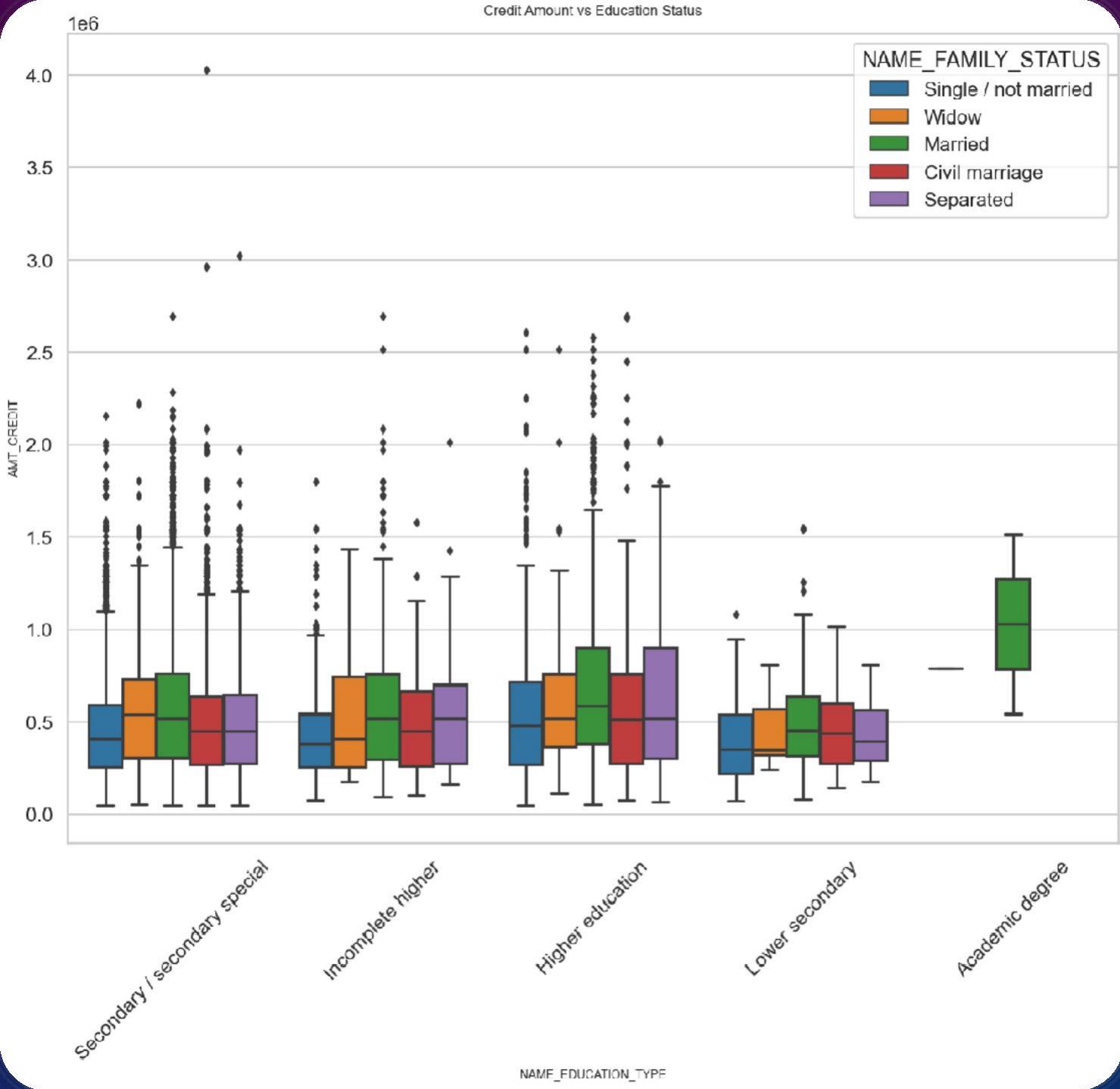
INCOME AMOUNT VS EDUCATION STATUS

- => there Within the context of 'civil marriage' family status and 'Lower secondary' education, the income amount tends to be lower compared to other situations.
- =>Analyzing the depicted boxplot for 'Higher education', it's evident that income amounts are predominantly comparable across various family statuses. While 'Academic degree' shows fewer outliers, its income levels are slightly higher than those for 'Higher education'.



BIVARIATE ANALYSIS FOR TYPE 1





CREDIT AMOUNT VS EDUCATION STATUS

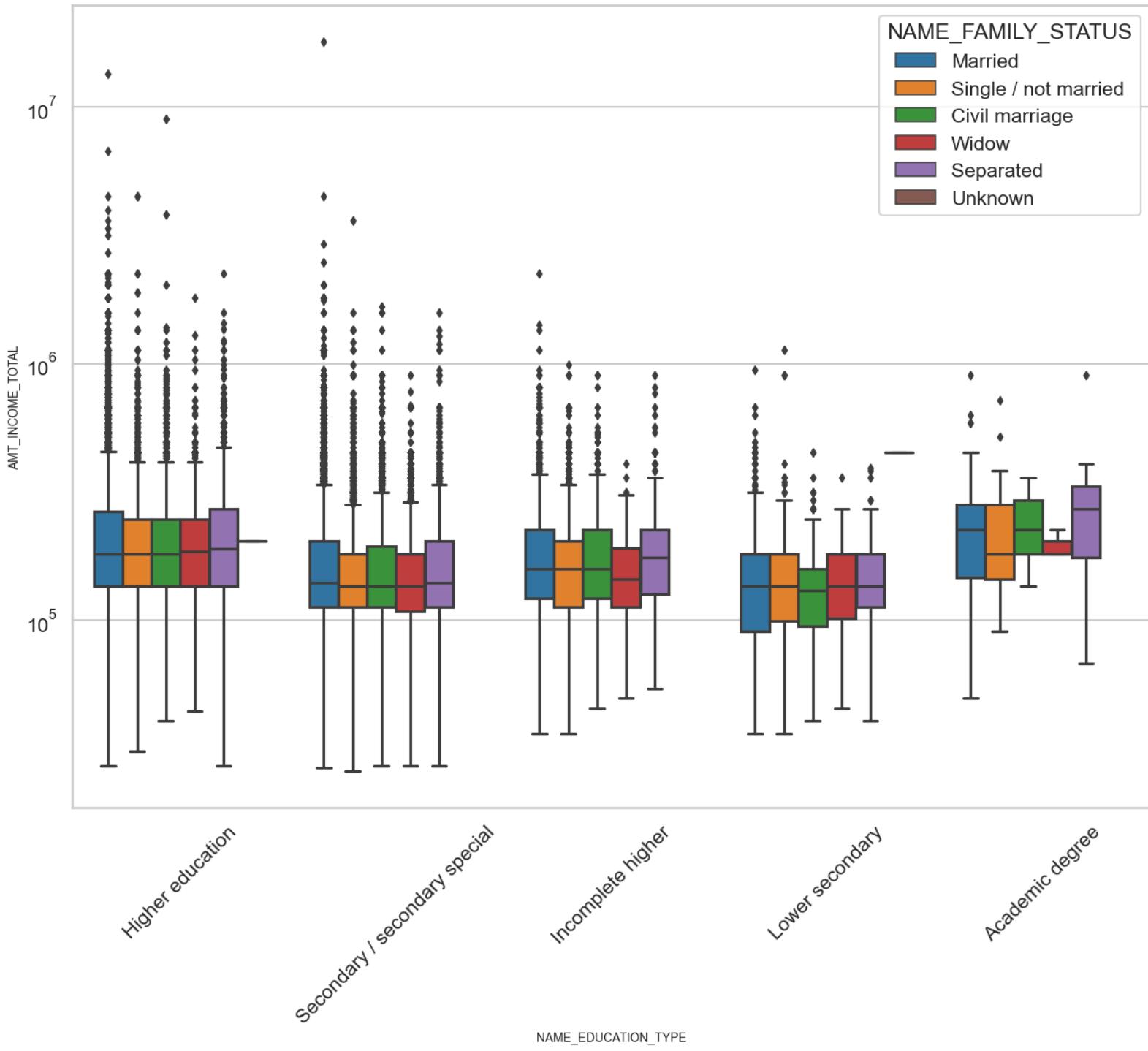
observations:

- People who are low educated are more likely to default whereas people with higher education and academic degree are less likely to default.
- Most of the loans are given to married people and default rate is approximately 8%. Single and Civil marriage folks have a default rate of 10.5% approximately. So comparatively Married people are less likely to default.
- Family statuses such as 'civil marriage', 'marriage', and 'separated', coupled with an Academic degree education, exhibit a greater propensity for higher credit counts compared to other cases.
- A substantial portion of outliers is apparent within 'Higher education' and 'Secondary' education categories.
- For 'Academic degree' education and 'civil marriage' family status, the third quartile predominantly encompasses the bulk of credit instances.



INCOME AMOUNT VS. EDUCATION STATUS

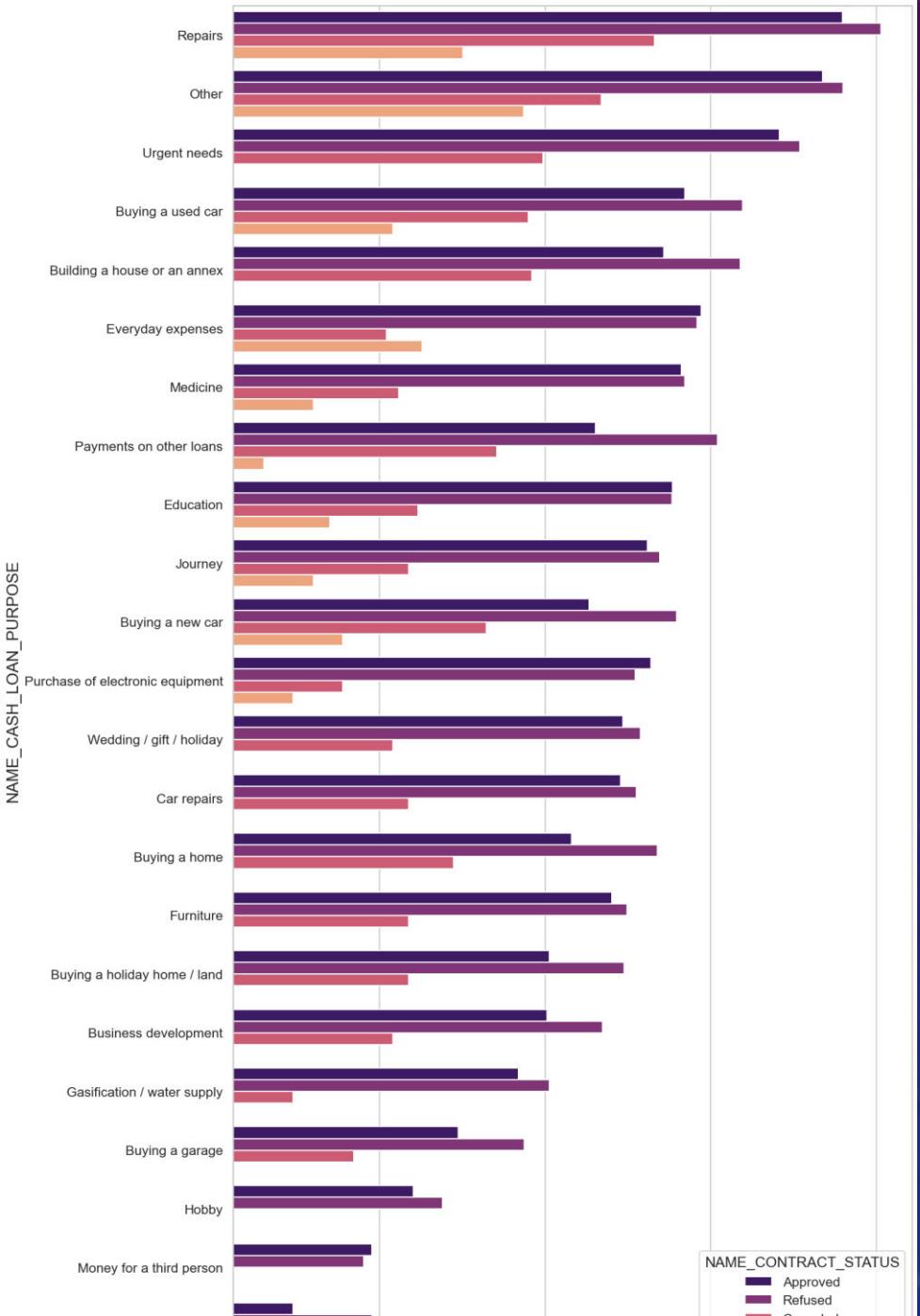
In the provided box plot, when considering 'Higher education', the income amounts often align closely with various family statuses. 'Academic degree', while exhibiting fewer outliers, displays slightly elevated income levels compared to 'Higher education'.



UNIVARIATE ANALYSIS AFTER MERGING THE DATA



Distribution of contract status with purposes

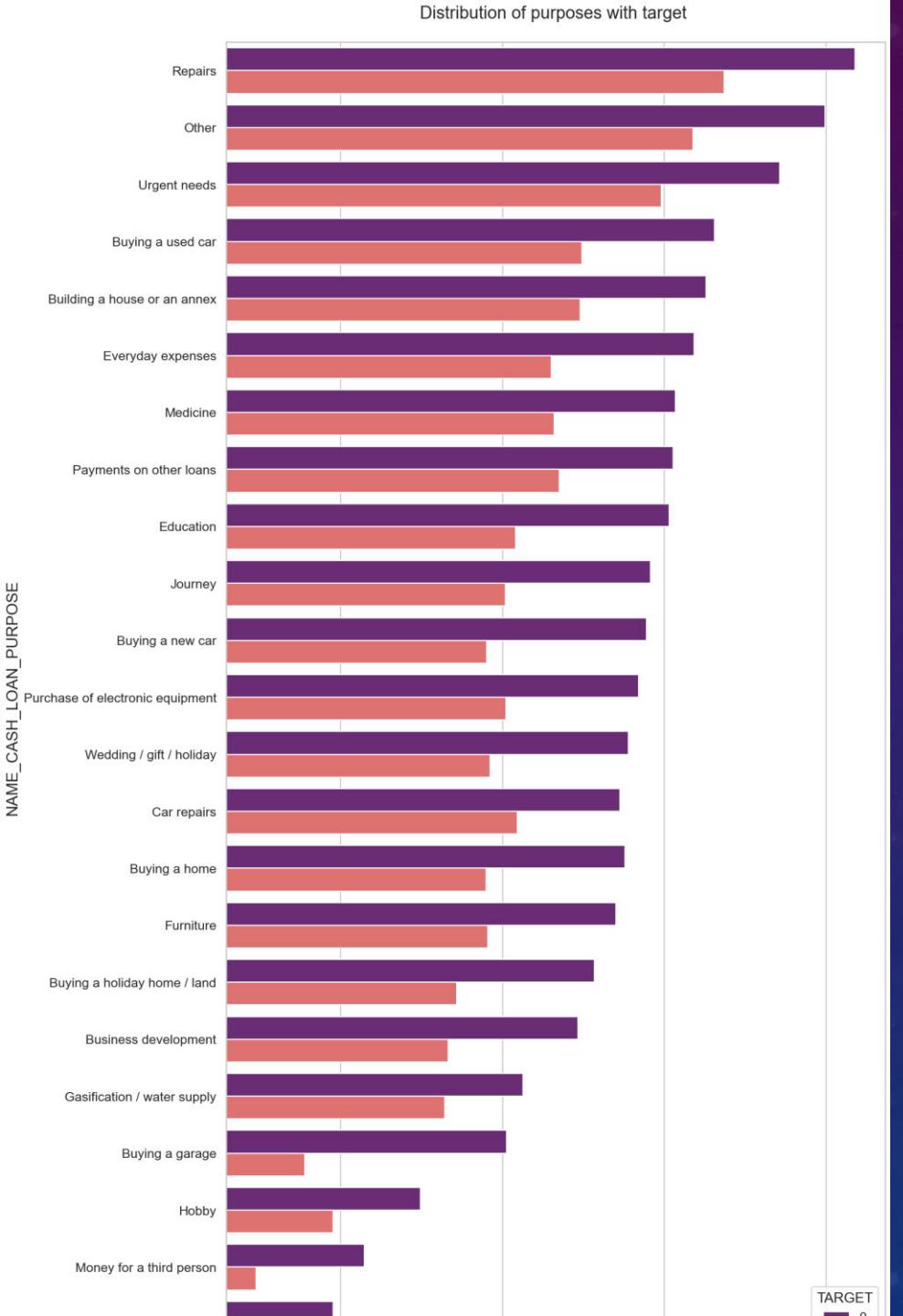


DISTRIBUTION OF CONTRACT STATUS WITH PURPOSES

Key observations derived from the presented plot:

- 'Repairs' emerge as the primary reason for loan rejections.
- Educational purposes yield an equilibrium between loan approvals and rejections.
- Notably, loan rejections significantly outweigh approvals for purposes involving 'Paying other loans' and 'Buying a new car'.





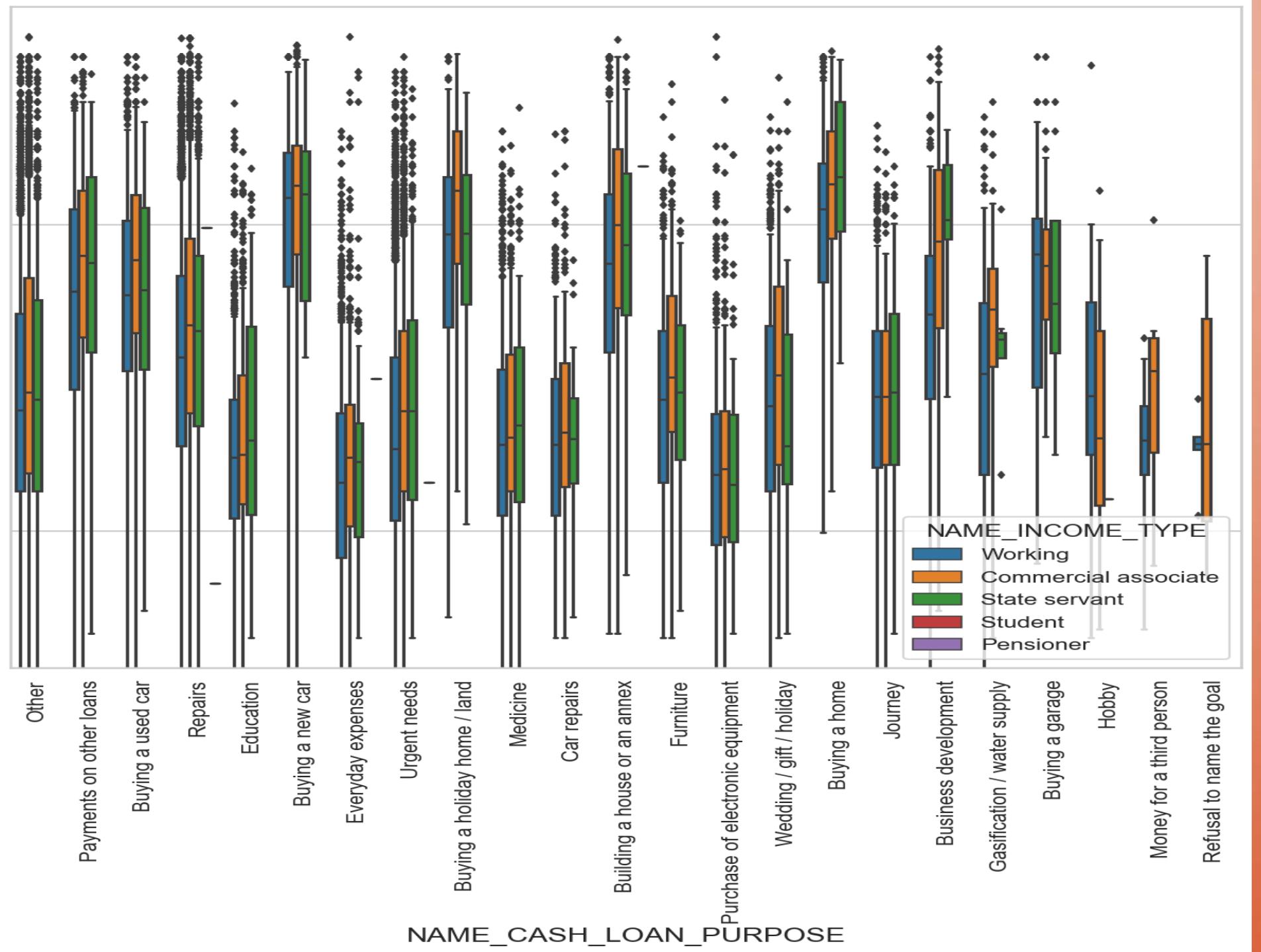
DISTRIBUTION OF PURPOSES WITH TARGET

Observations

- Borrowers who utilize loans for 'Repairs' encounter greater challenges in making timely payments.
- Notably, specific loan purposes exhibit a notably higher propensity for successful repayments. These purposes include 'Buying a garage,' 'Business development,' 'Buying land,' 'Buying a new car,' and 'Education.' Thus, these purposes merit focused attention due to their association with minimal payment difficulties.

BIVARIATE ANALYSIS AFTER MERGING THE DATA



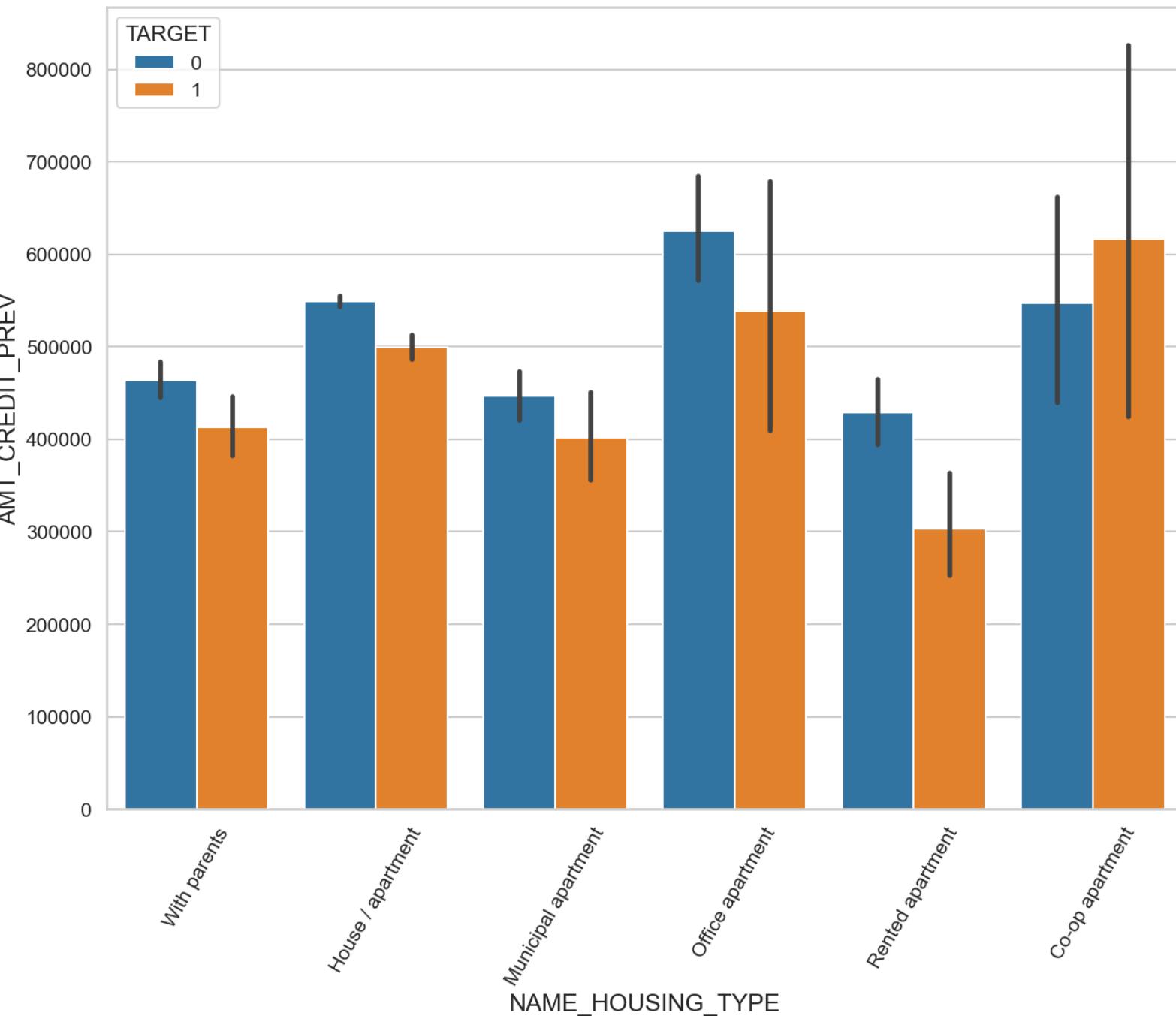


PREVIOUS CREDIT AMOUNT VS. PURPOSE OF CASH LOAN

- Loan purposes such as 'Buying a home,' 'Buying land,' 'Buying a new car,' and 'Building a house' are associated with higher credit amounts.
- Applicants classified as state servants exhibit a noteworthy demand for credit.
- There is a relatively lower propensity for credit applications related to third-party funding or hobbies.



Previous Credit amount vs Housing type



PREVIOUS CREDIT AMOUNT VS. HOUSING TYPE

- 'Office apartment' housing type shows a occurrence for credit target 0, when 'Co-op apartment' housing type is related with the credit target 1.
- Loan purposes such as 'Buying a home,' 'Buying land,' 'Buying a new car,' and 'Building a house' are associated with higher credit amounts.



FINAL CONCLUSION

- Married individuals default less than single or civil-married counterparts and parental statuses are also impacting default rates.
- Higher-educated clients shows better repayment behavior.
- Certain occupations, like low-skilled laborers and drivers, are prone to higher default rates.
- Higher incomes correlate with lower defaults, but extreme high incomes again has higher default risk.
- loans intended for the purpose of 'Repair' show a higher frequency of untimely payment issues.
- Targeting clients from the 'With parents' housing type could prove advantageous,
- Advise for banks to prioritize individuals with contract types 'Student,' 'Pensioner,' and 'Businessman' who possess housing types other than 'Co-op apartment' for high favorable payment returns.



THANK YOU.

