

# CREDIT EDA CASE STUDY

---

BY DHARSHAN JAIN

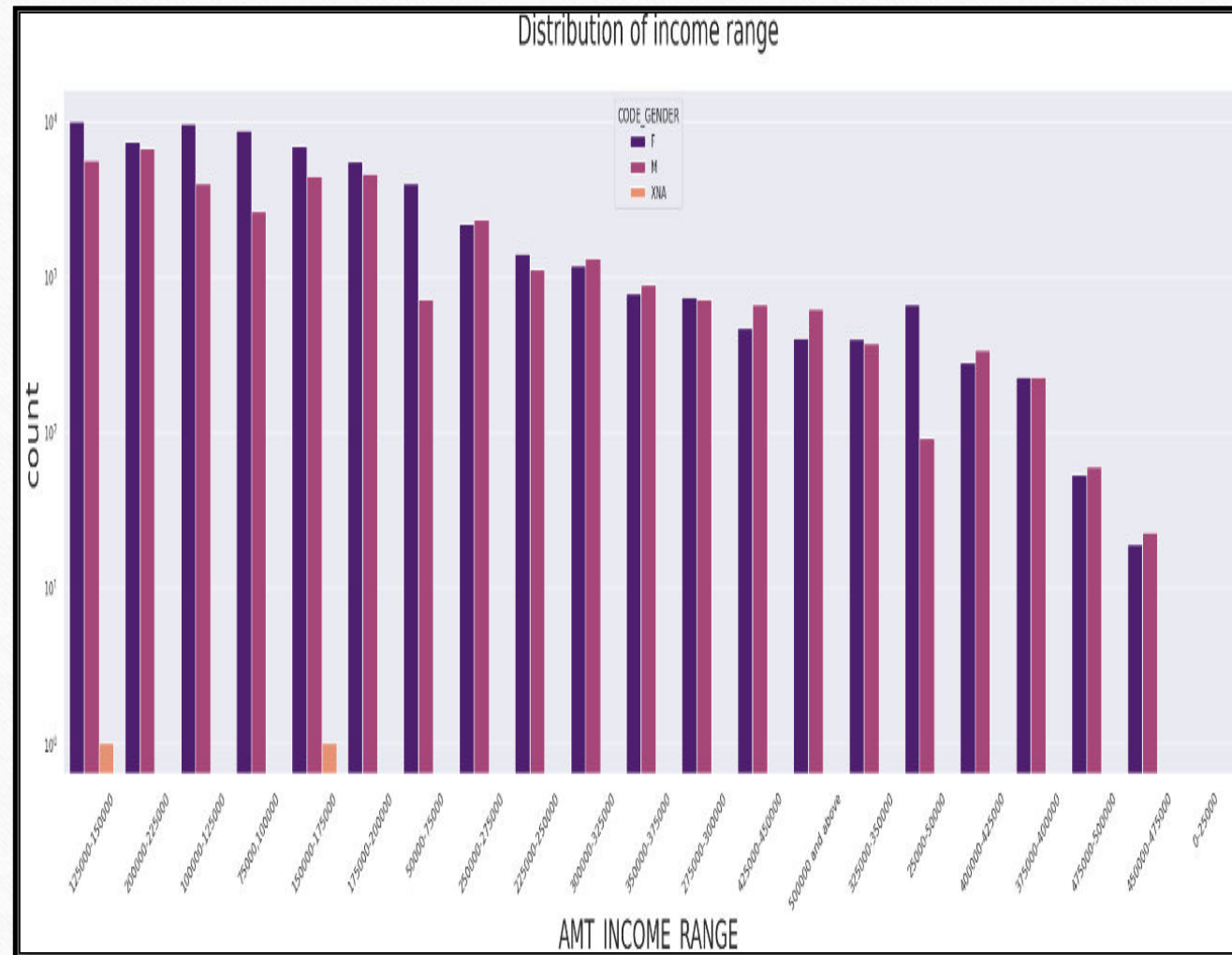
---

# CATEGORICAL UNIVARIATE ANALYSIS FOR TARGET 0



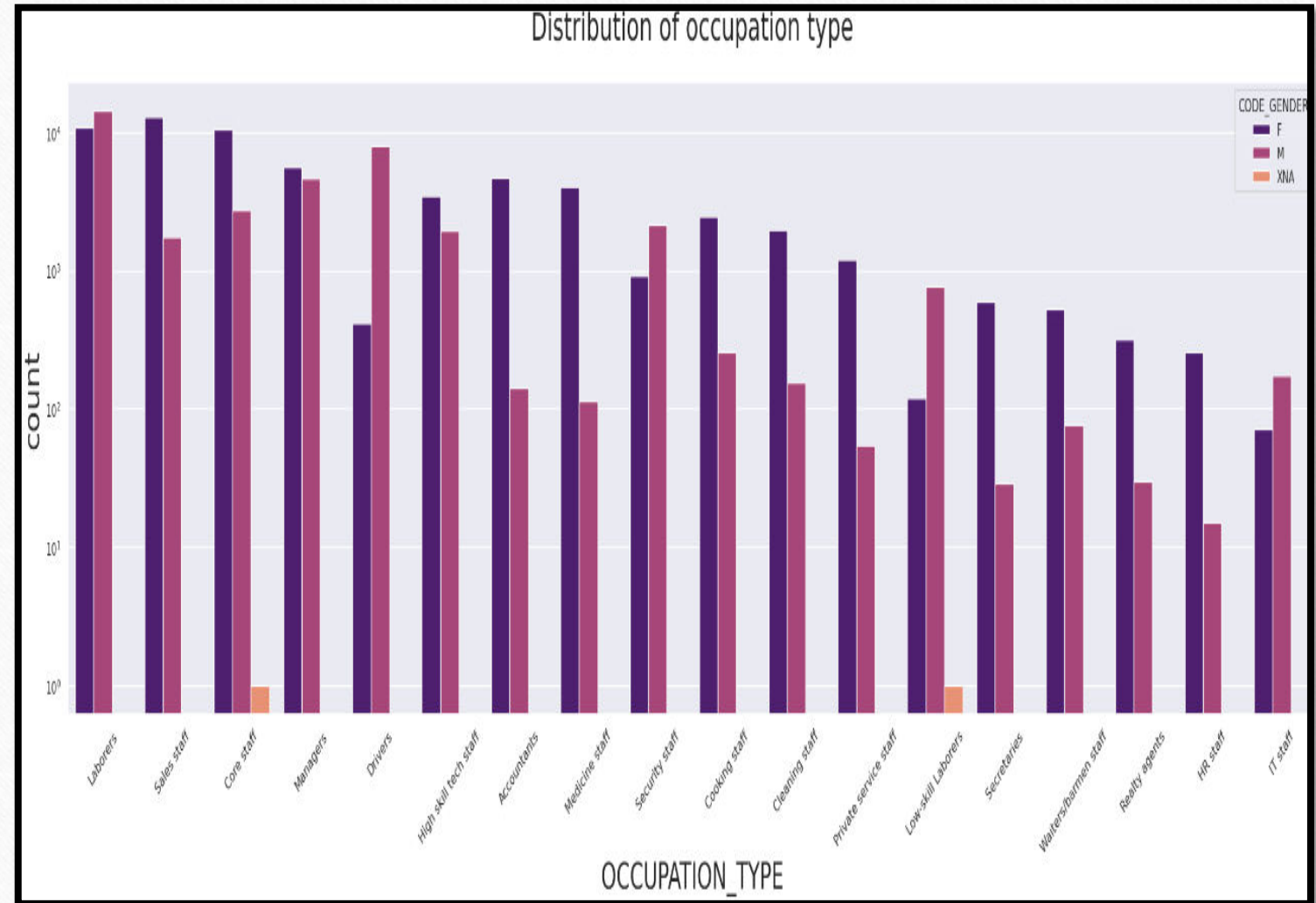
## Distribution of income range

1. Female counts are higher than male.
2. Income range from 200000 to 225000 is having more number of credits.
3. This graph shows that females are more than males in having credits for that range.
4. Very less count for income range 400000 and above.



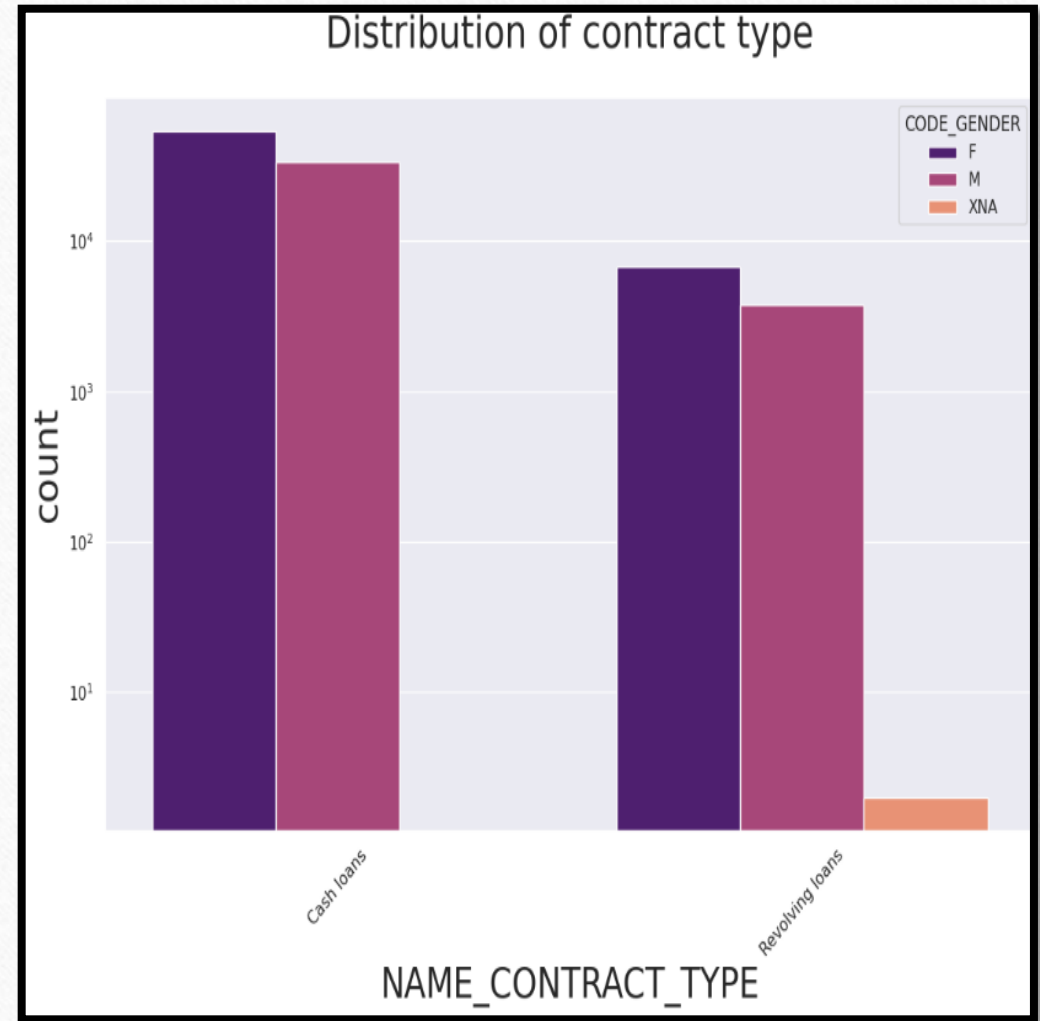
# Distribution of occupation type

1. Female counts are higher than male.
2. Laborers have more number of credits.
3. This graph shows that females are more likely than males to have credits for that range.
4. In IT Staff females have a very small number of credits.
5. In Medicine Staff Secretaries males have a very small number of credits.
6. In Realty agents males don't have credits.



## Distribution of contract type

1. For contract type 'cash loans' have a higher number of credits than 'Revolving loans' contract type.
2. For this also females are applying for more credits.



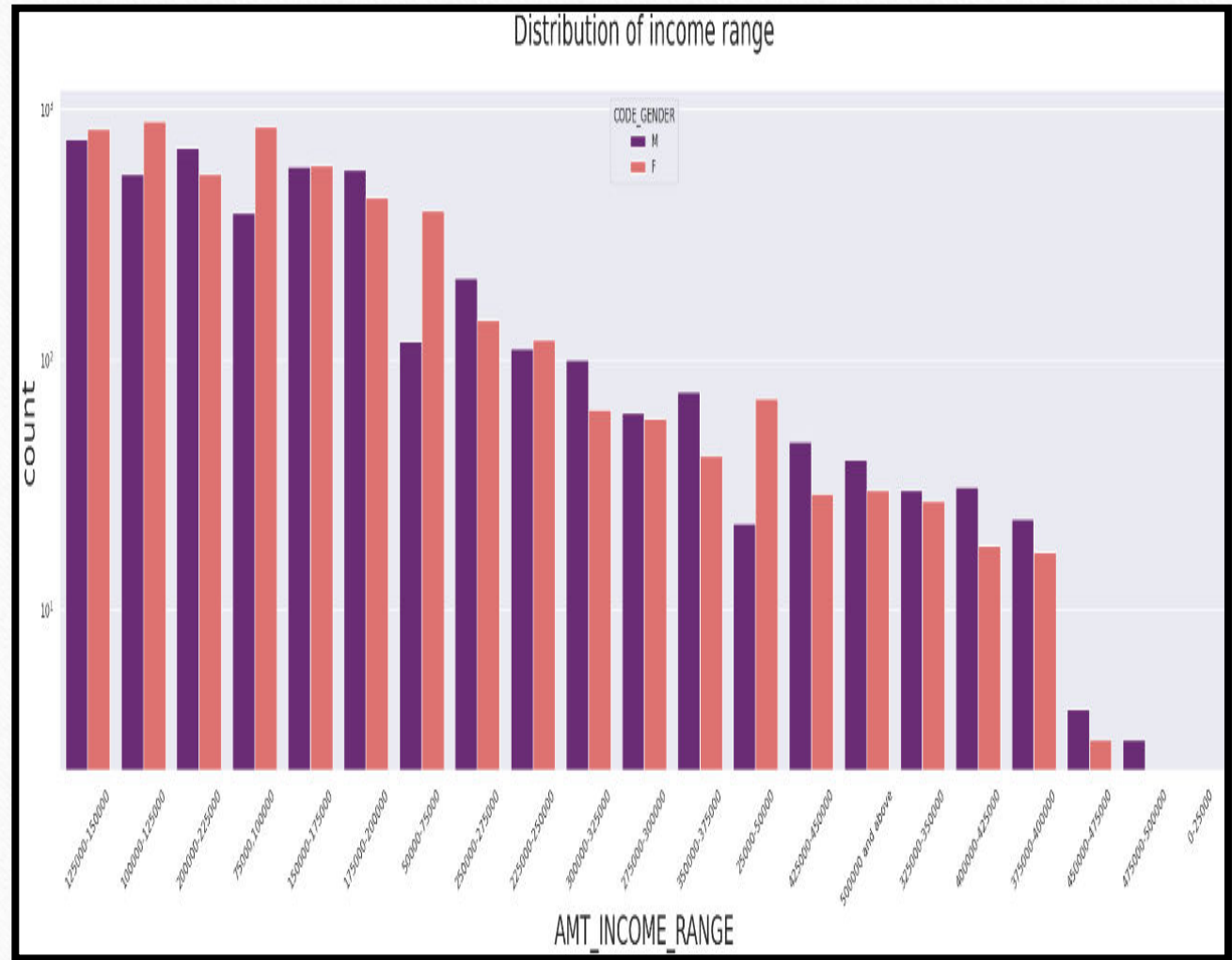


---

# CATEGORICAL UNIVARIATE ANALYSIS FOR TARGET 1

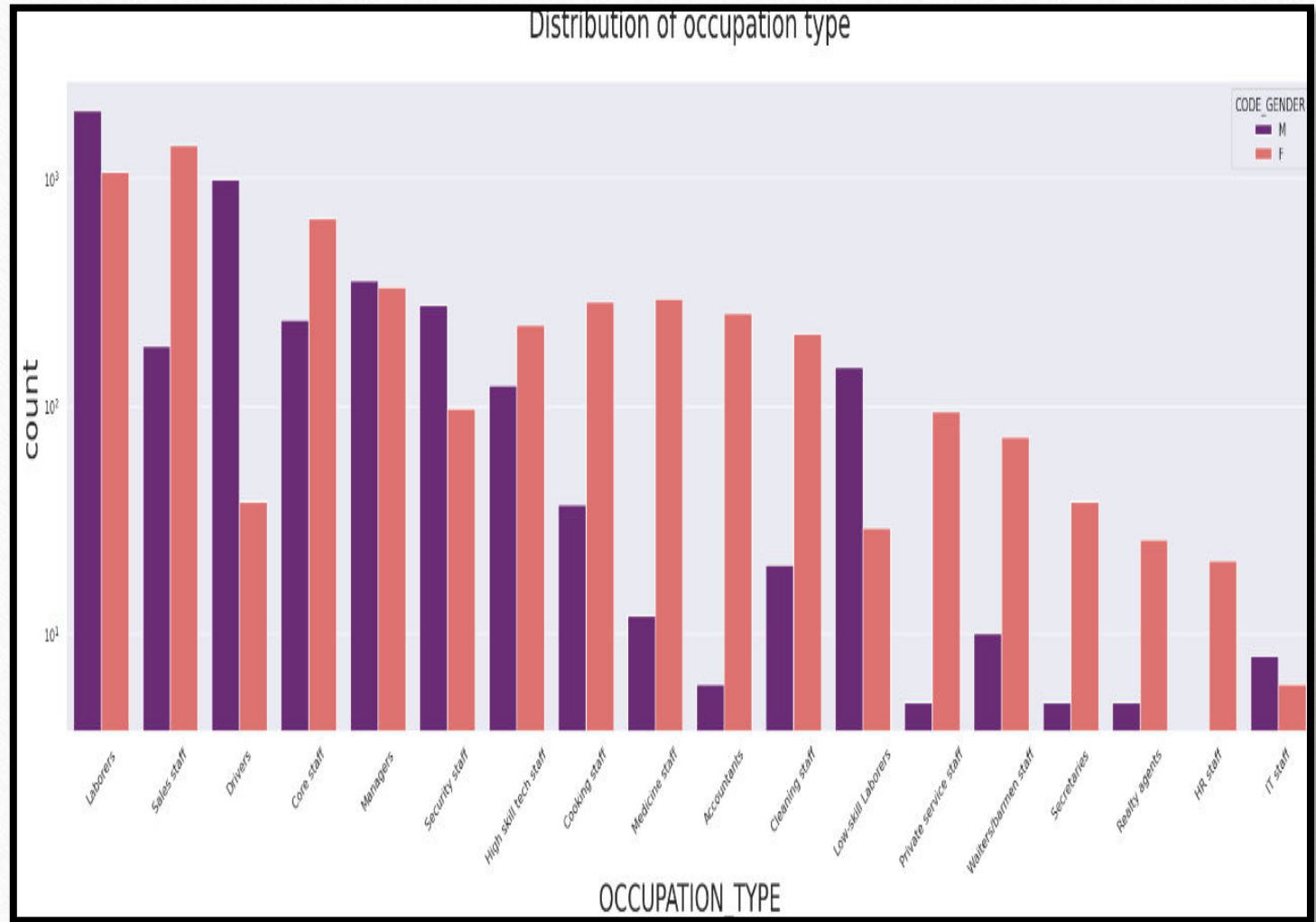
## Distribution of income range

1. Male counts are higher than female in making defaults.
2. Income ranges from 125000 to 150000 have more number of defaults.
3. This graph shows that males are more than female in having defaults for that range.
4. Very less count for income range 400000 and above.



## Distribution of occupation type

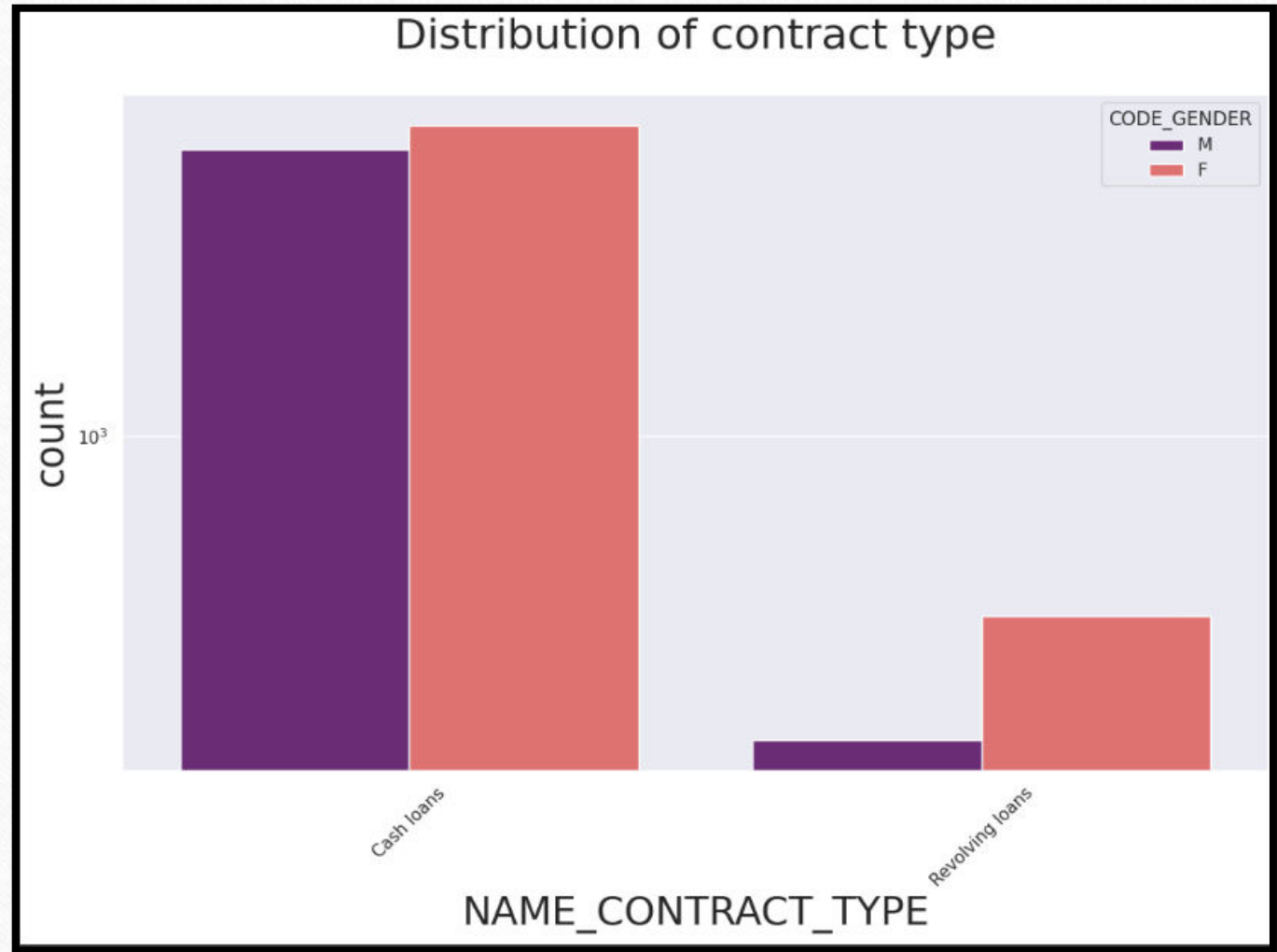
1. Male counts are higher than female.
2. Laborers having more number of defaults.





## Distribution of contract type

1. For contract type 'cash loans' is having higher number of defaults than 'Revolving loans' contract type.
2. For this also Female is leading for applying credits.
3. There is higher Female Revolving loans.



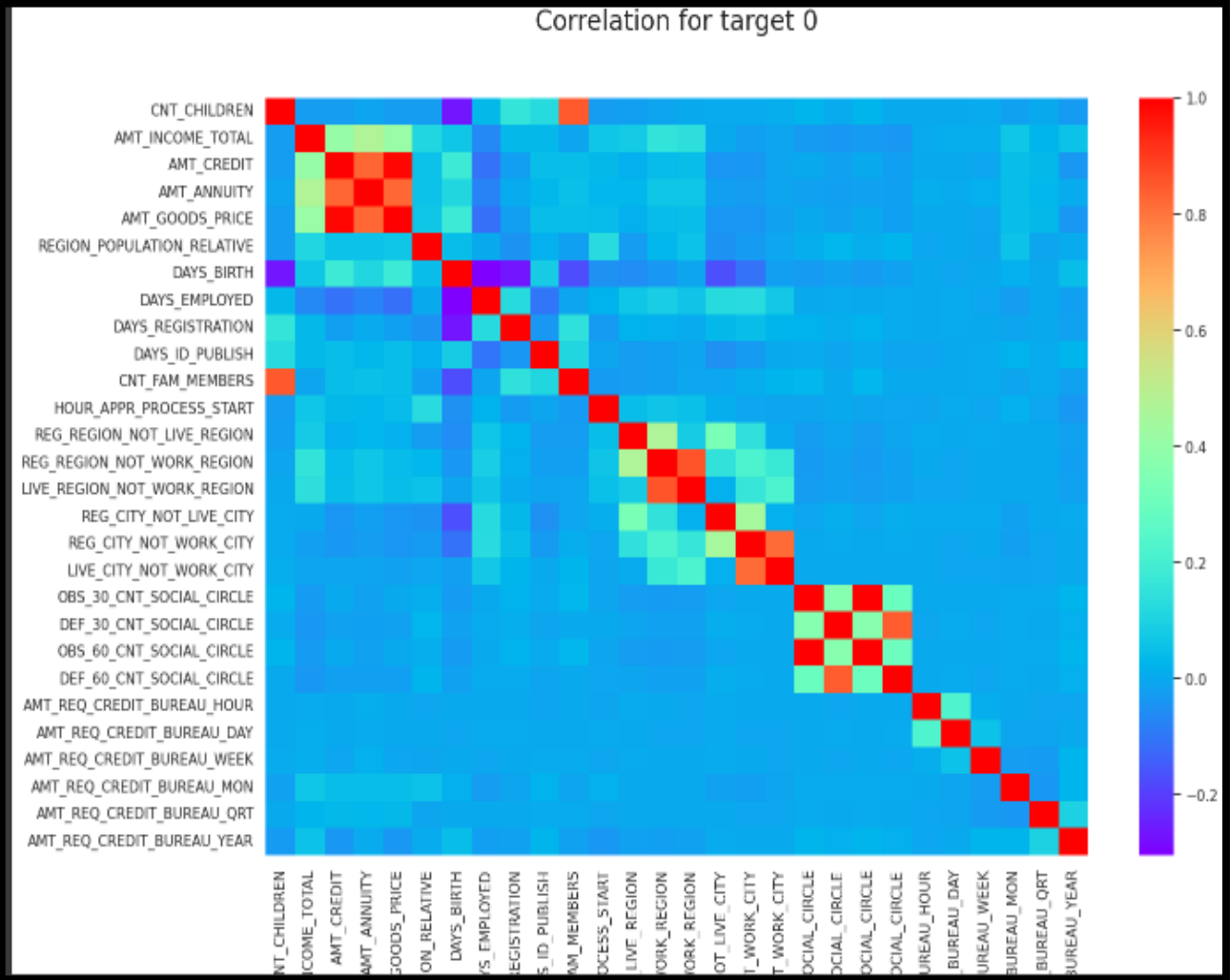


---

# CORRELATION OF TARGET 0

## Correlation for target 0

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.





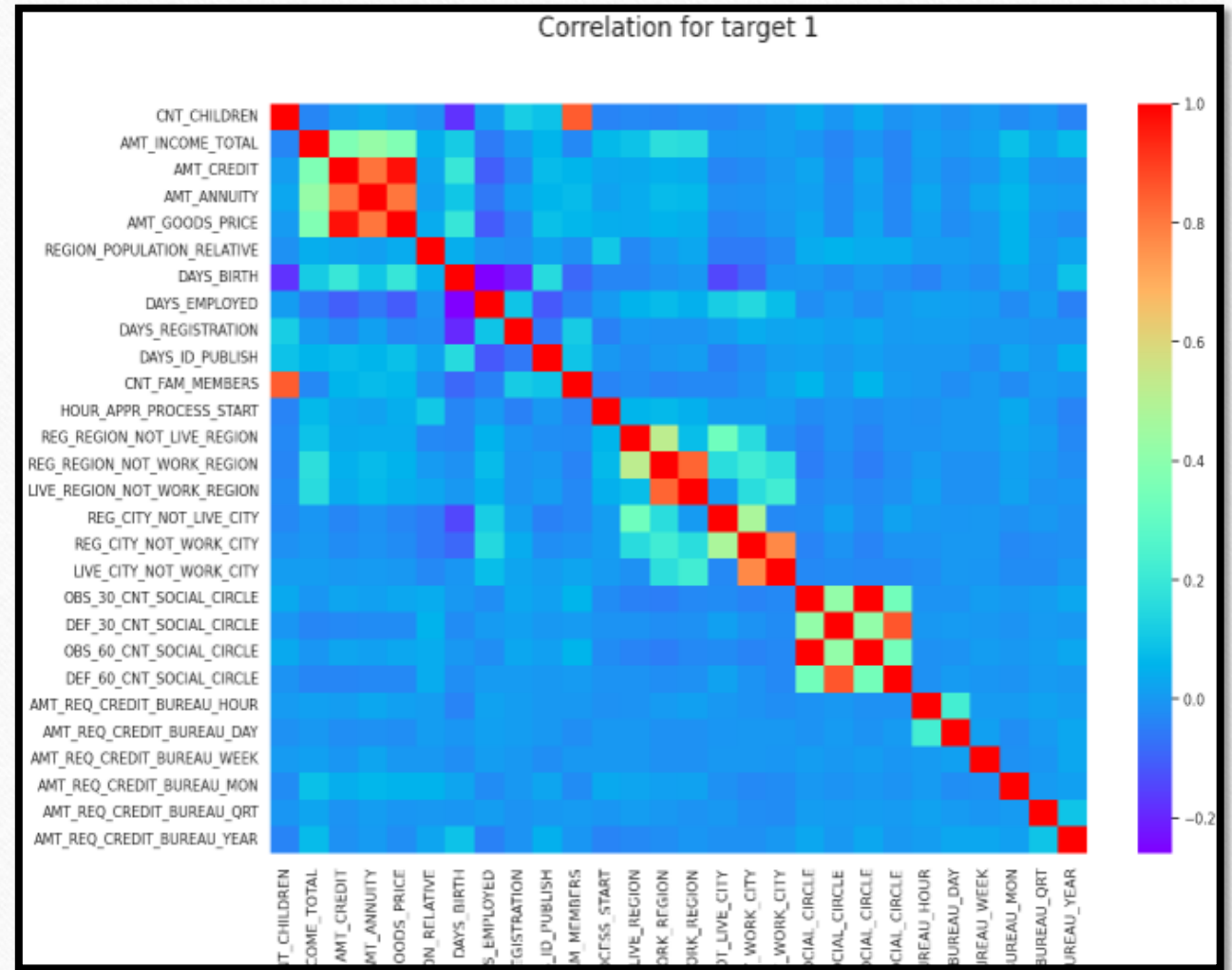


---

# CORRELATION OF TARGET 1

## Correlation for target 1

1. The client's permanent address does not match the contact address  
are having fewer children and vice-versa
2. The client's permanent address does not match the work address  
are having fewer children and vice-versa



---

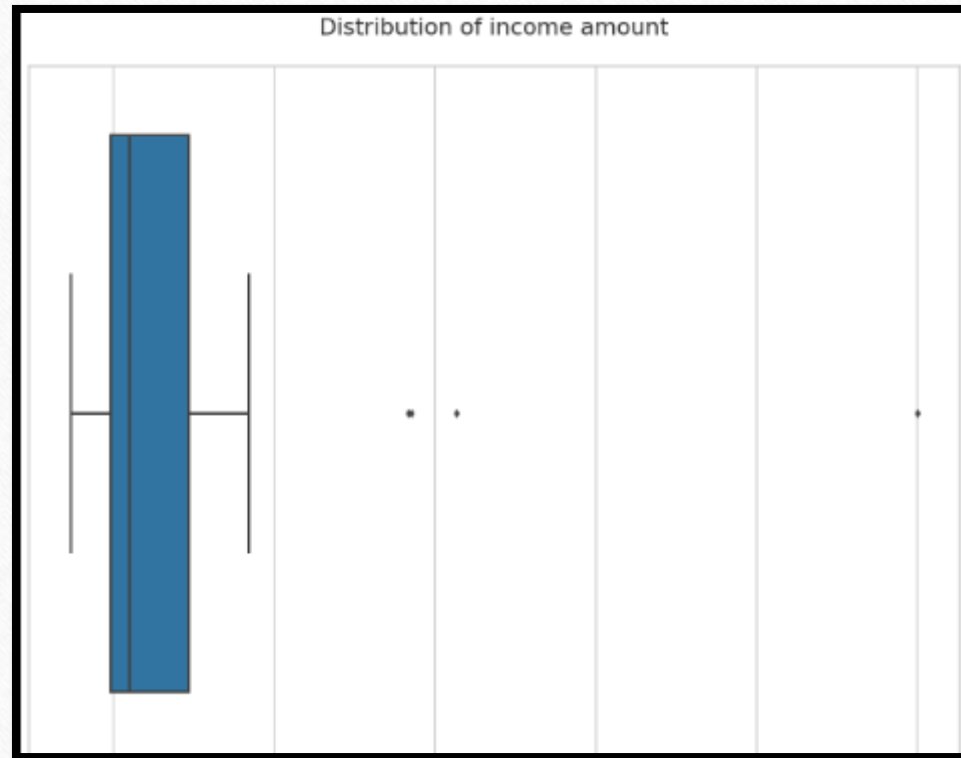
# CATEGORICAL UNIVARIATE ANALYSIS FOR VARIABLES TARGET 0



## Distribution of income amount

---

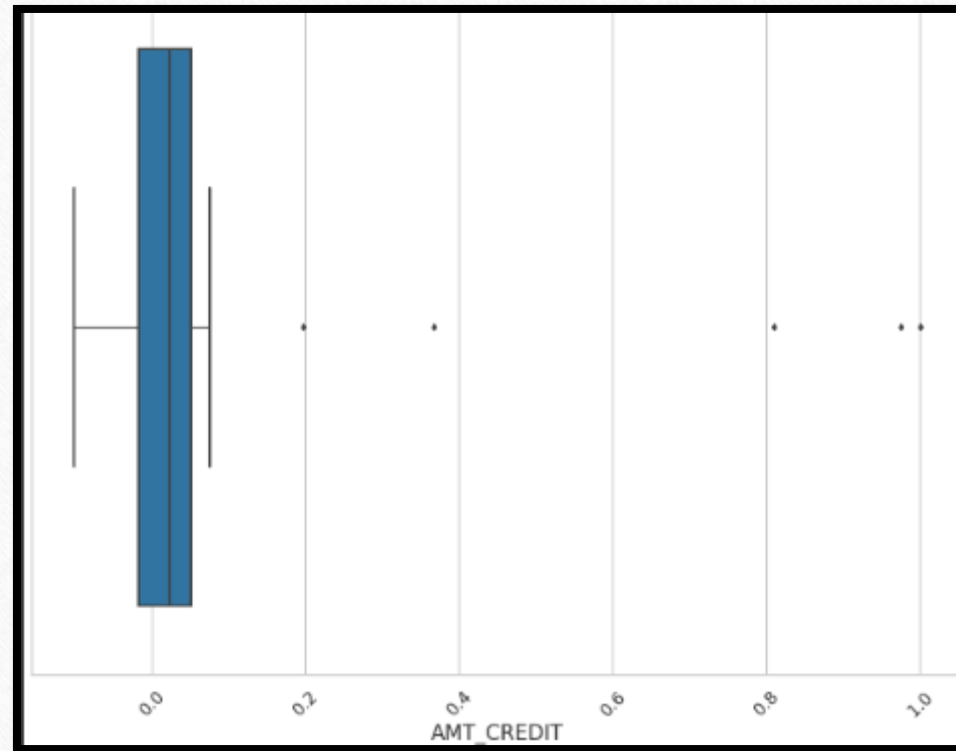
1. Some outliers are noticed in income amounts.
2. The third quartile is very slim for income amount.



## Distribution of credit amount

---

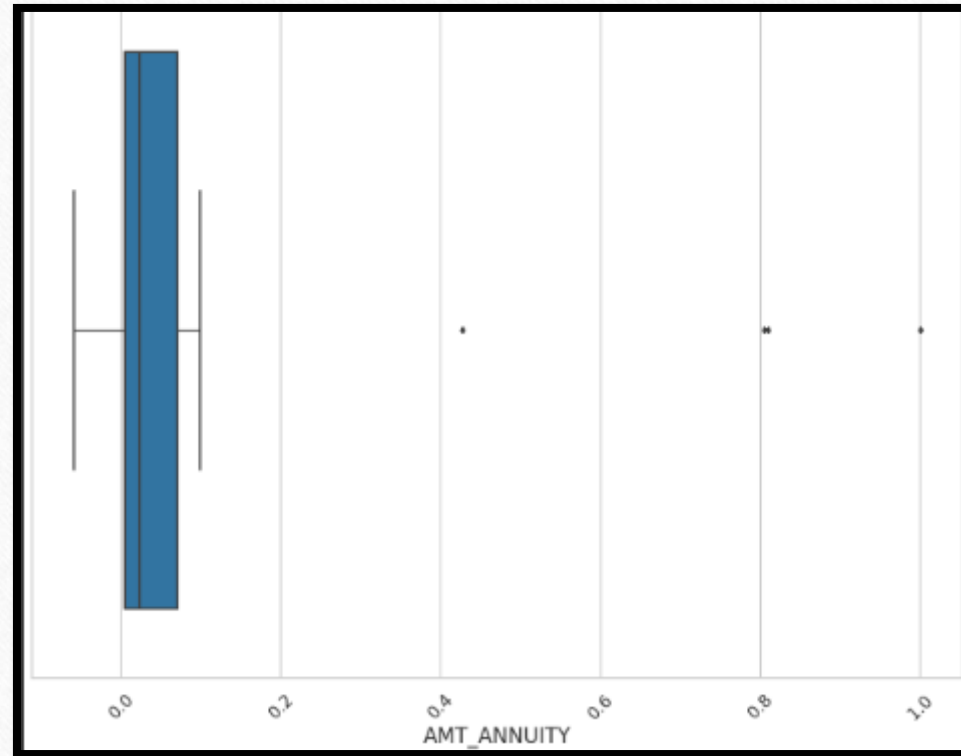
1. Some outliers are noticed in credit amounts.
2. The first quartile is bigger than the third quartile for credit amount which means most of the credits of clients are present in the first quartile.



## Distribution of Annuity amount

---

1. Some outliers are noticed in the amount of annuity.
2. The first quartile is bigger than the third quartile for the amount of annuity, which means most of the credits of clients are present in the first quartile.





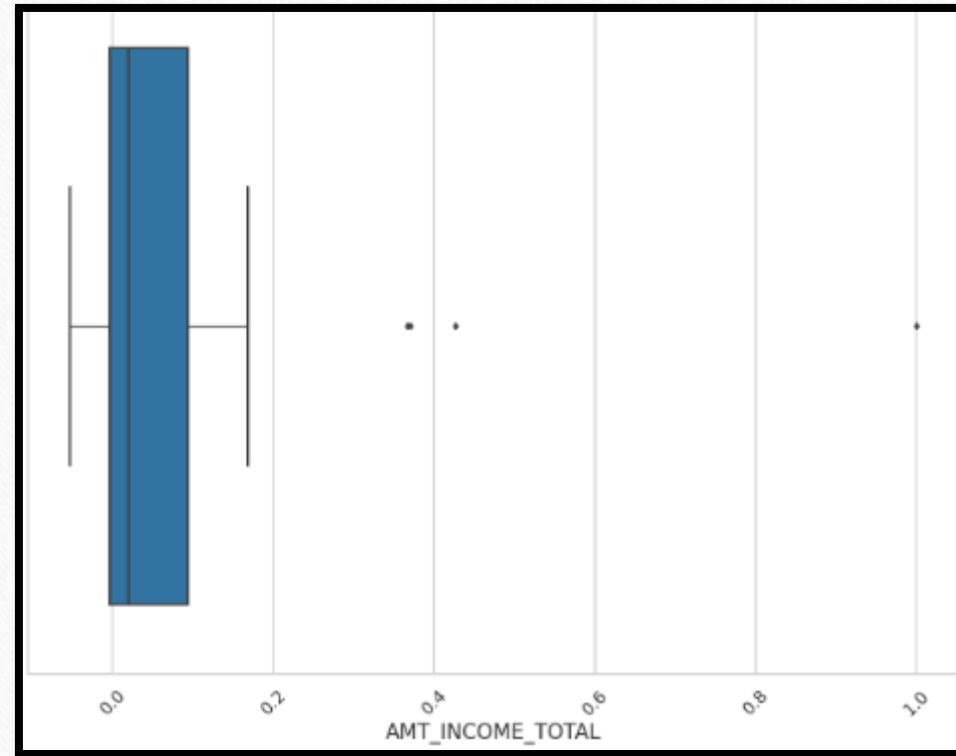
---

# CATEGORICAL UNIVARIATE ANALYSIS FOR VARIABLES TARGET 1

## Distribution of income amount

---

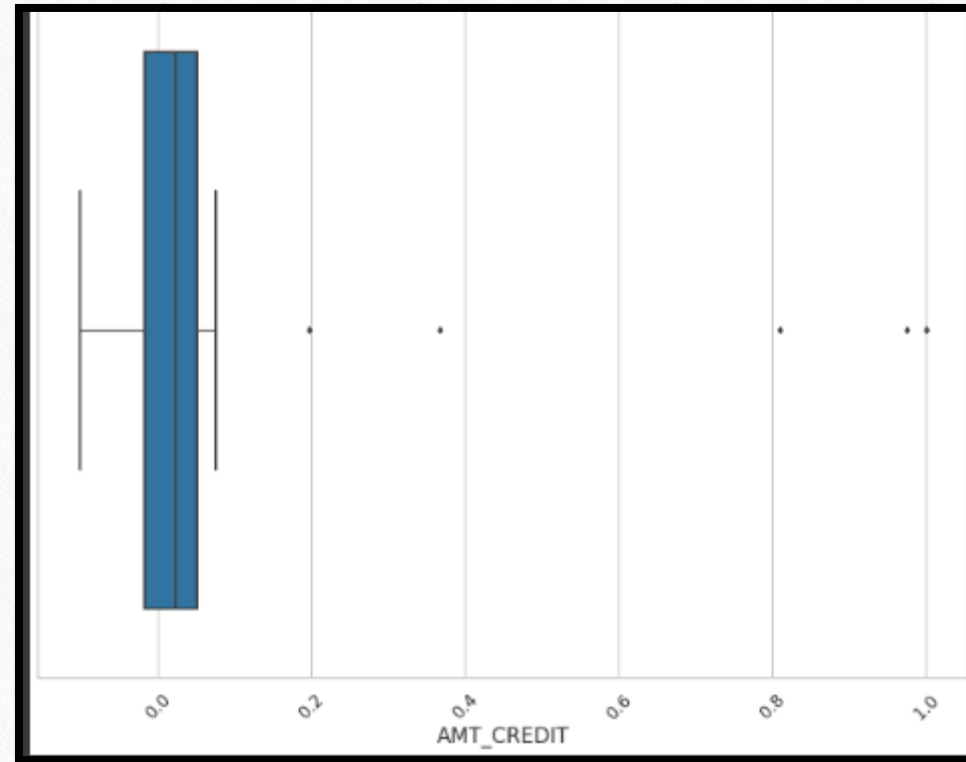
1. Some outliers are noticed in income amounts.
2. The third quartile is very slim for income amount.
3. Most of the clients of income are present in the first quartile.



## Distribution of credit amount

---

1. Some outliers are noticed in credit amounts.
2. The first quartile is bigger than the third quartile for credit amount which means most of the credits of clients are present in the first quartile.

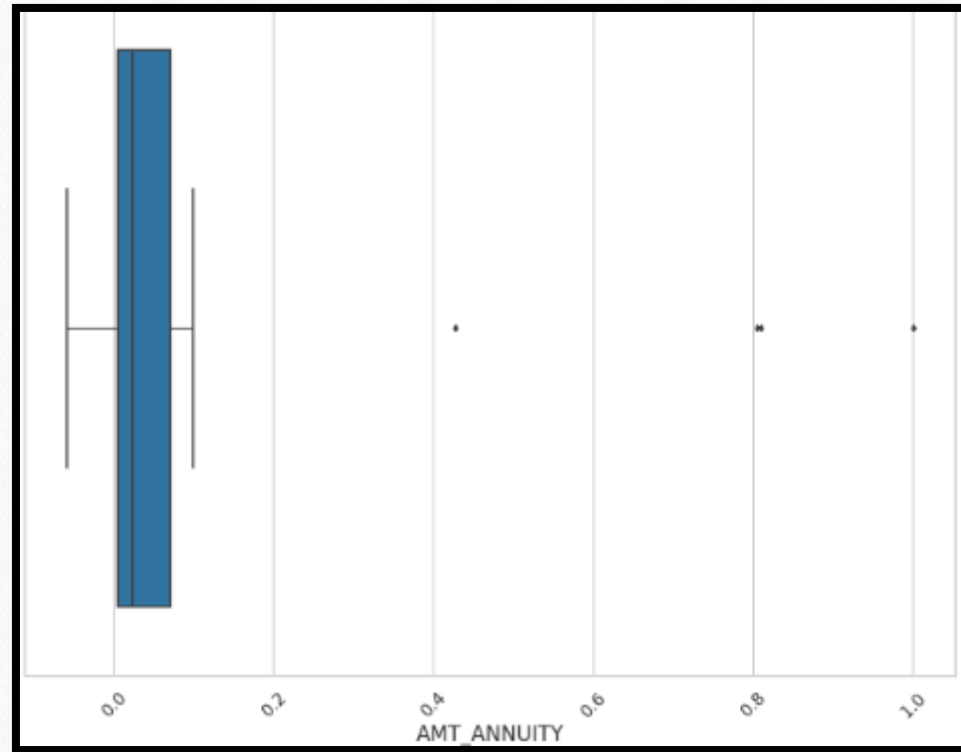




## Distribution of Annuity amount

---

1. Some outliers are noticed in the annuity amount.
2. The first quartile is bigger than the third quartile for the annuity amount which means most of the annuity clients are from the first quartile.



---

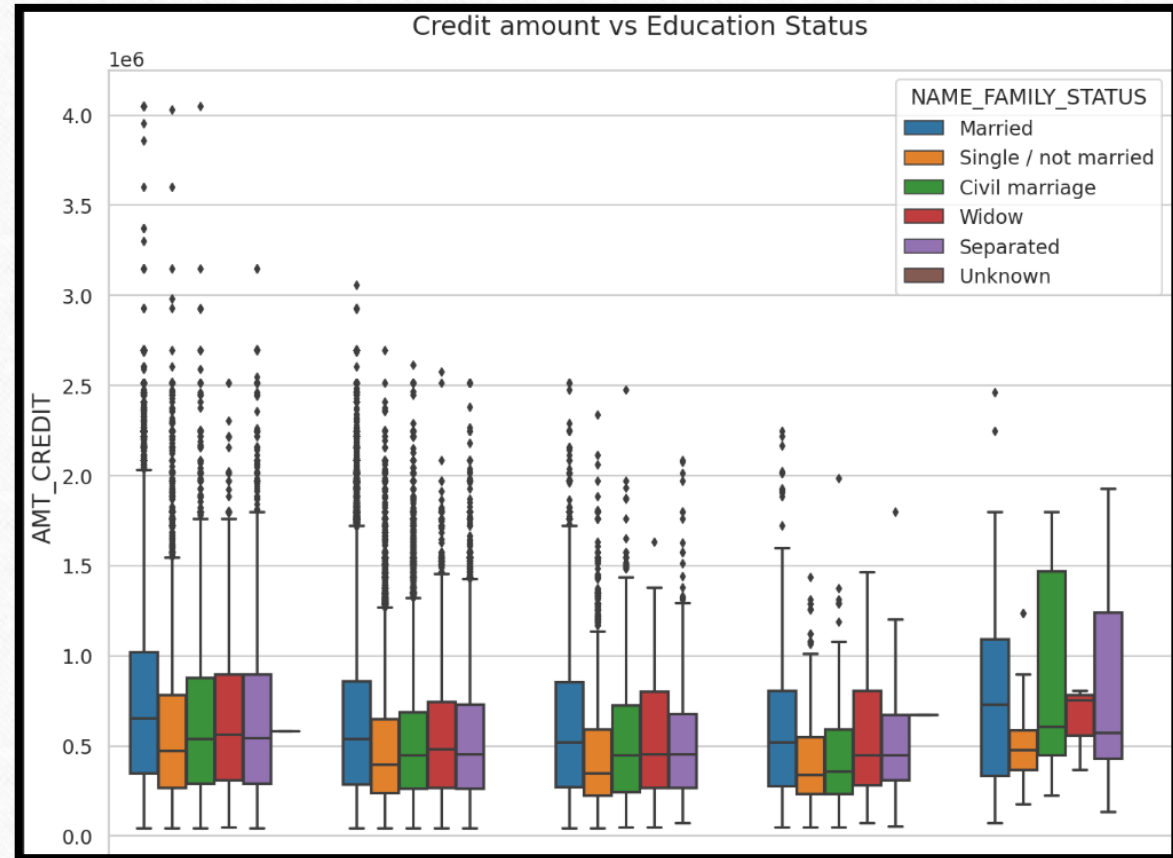
# BIVARIATE ANALYSIS FOR VARIABLES TARGET 0

## Credit amount vs Education Status

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.

Also, 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.

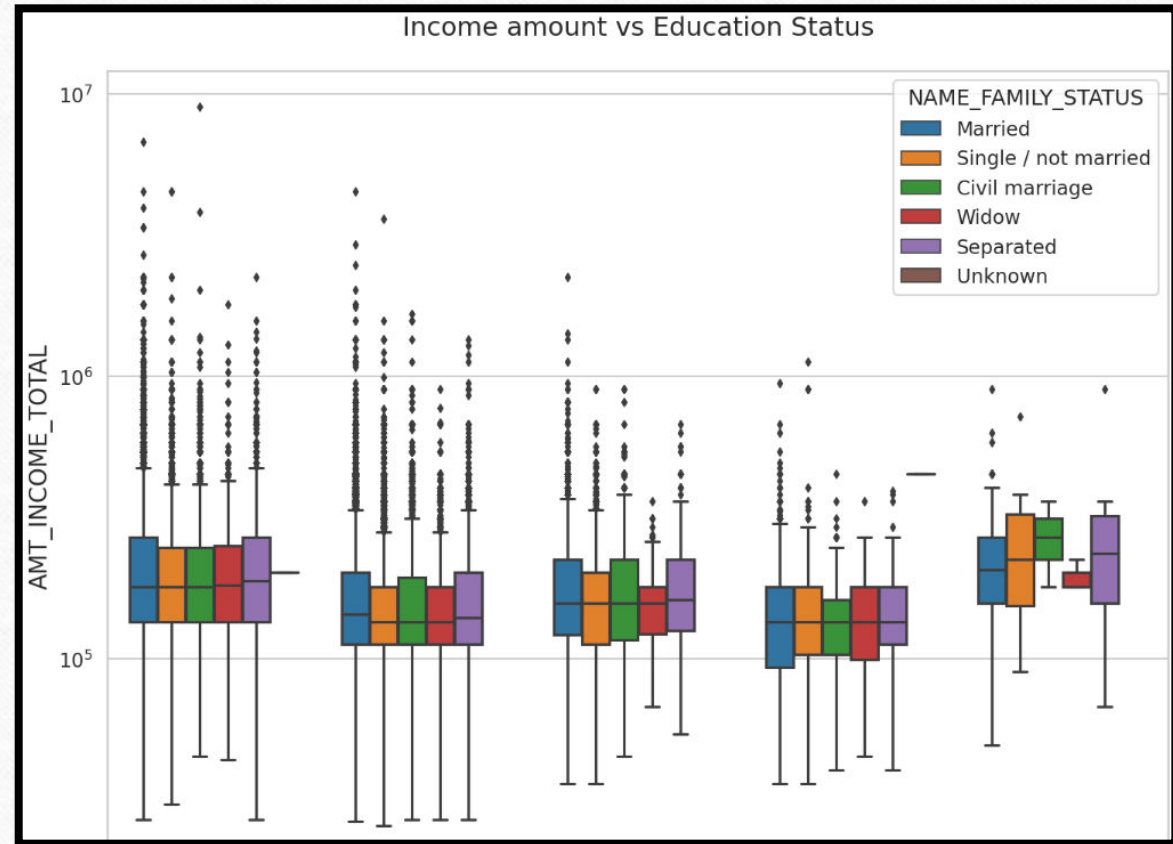




## Income amount vs Education Status

From the above boxplot for Education type 'Higher education' the income amount is mostly equal with family status. It does contain many outliers. Less outliers have an Academic degree but their income amount is a little higher than Higher education.

Lower secondary of civil marriage family status have less income amount than others.



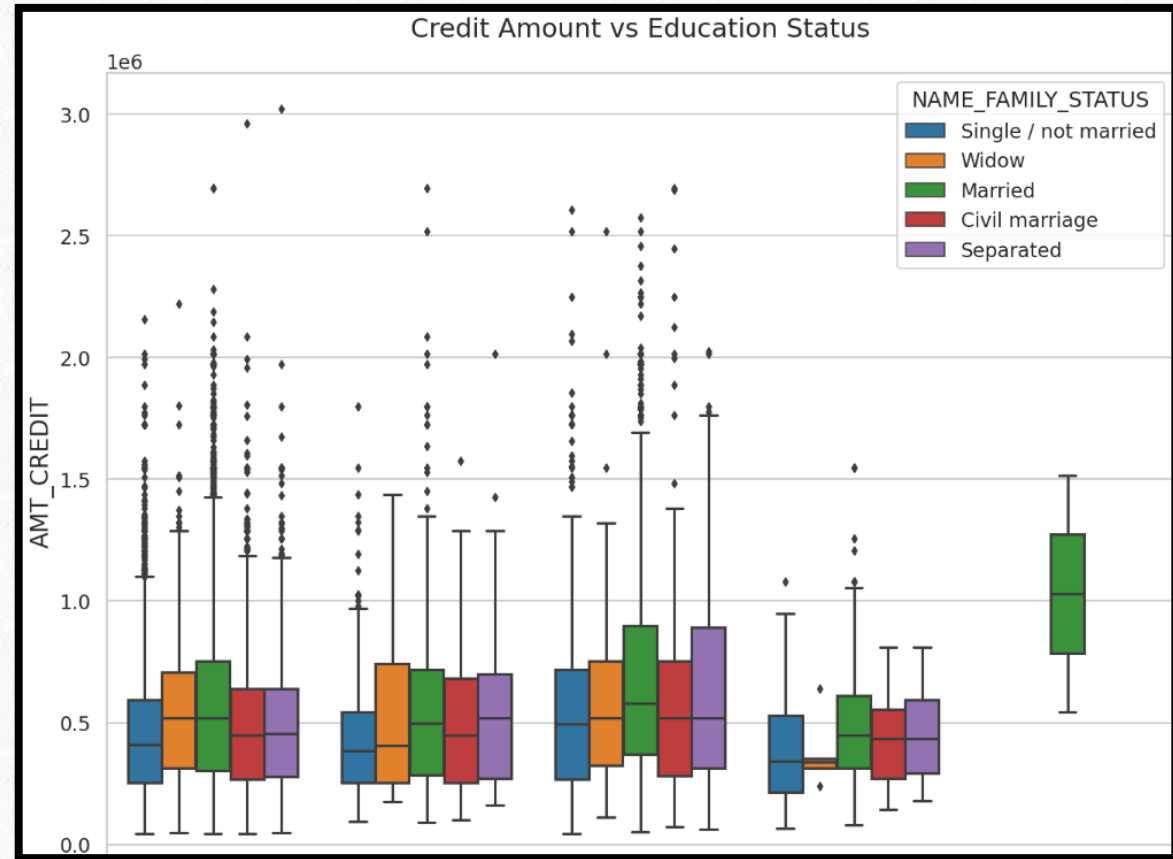
---

# BIVARIATE ANALYSIS FOR VARIABLES TARGET 1

## Credit Amount vs Education Status

From the above box plot, we can conclude that the Family status of 'Separated' whose higher education is incomplete, and those 'Civil marriage' who have done their higher education have a higher number of defaults than others.

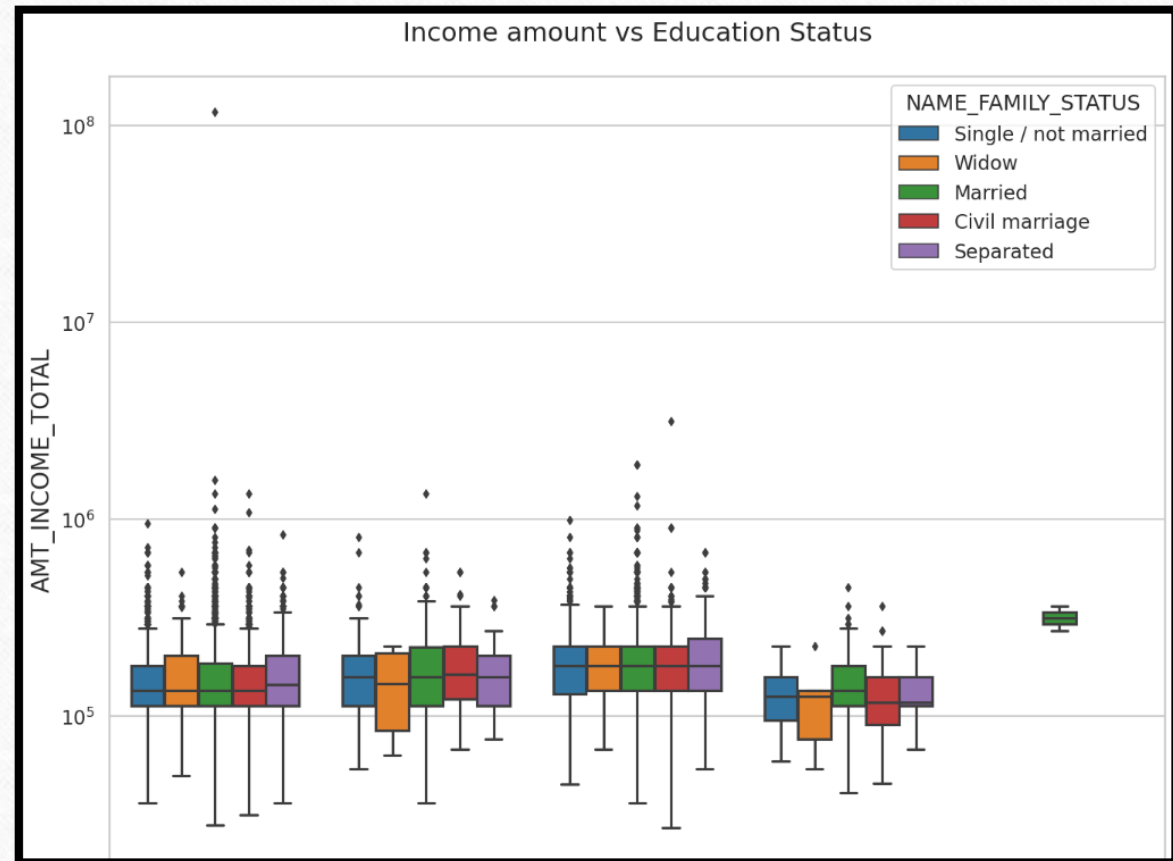
Also, higher education and Secondary on the family status of 'marriage', 'single', and 'civil marriage' have more outliers.





## Income amount vs Education Status

From the above boxplot for Education types 'Higher education' and 'Secondary' the income amount is mostly equal with family status. It does contain many outliers. Lower secondary of married and single-family status have less income amount than others.

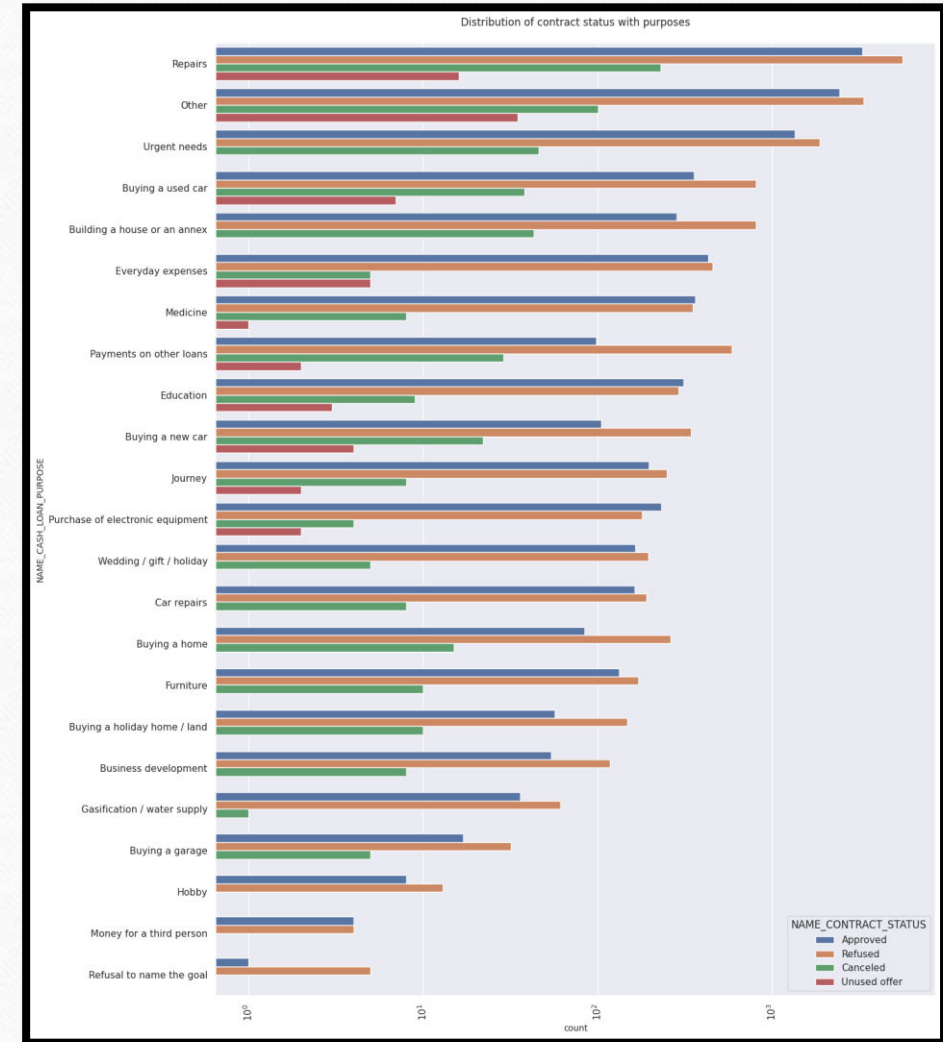


---

# UNIVARIATE ANALYSIS AFTER MERGING PREVIOUS DATA

## Distribution of contract status

1. Most rejections of loans came from purpose 'repairs'.
2. For education purposes we have an equal number of approves and rejections.
3. Paying other loans and buying a new car having significantly higher rejections than approves.



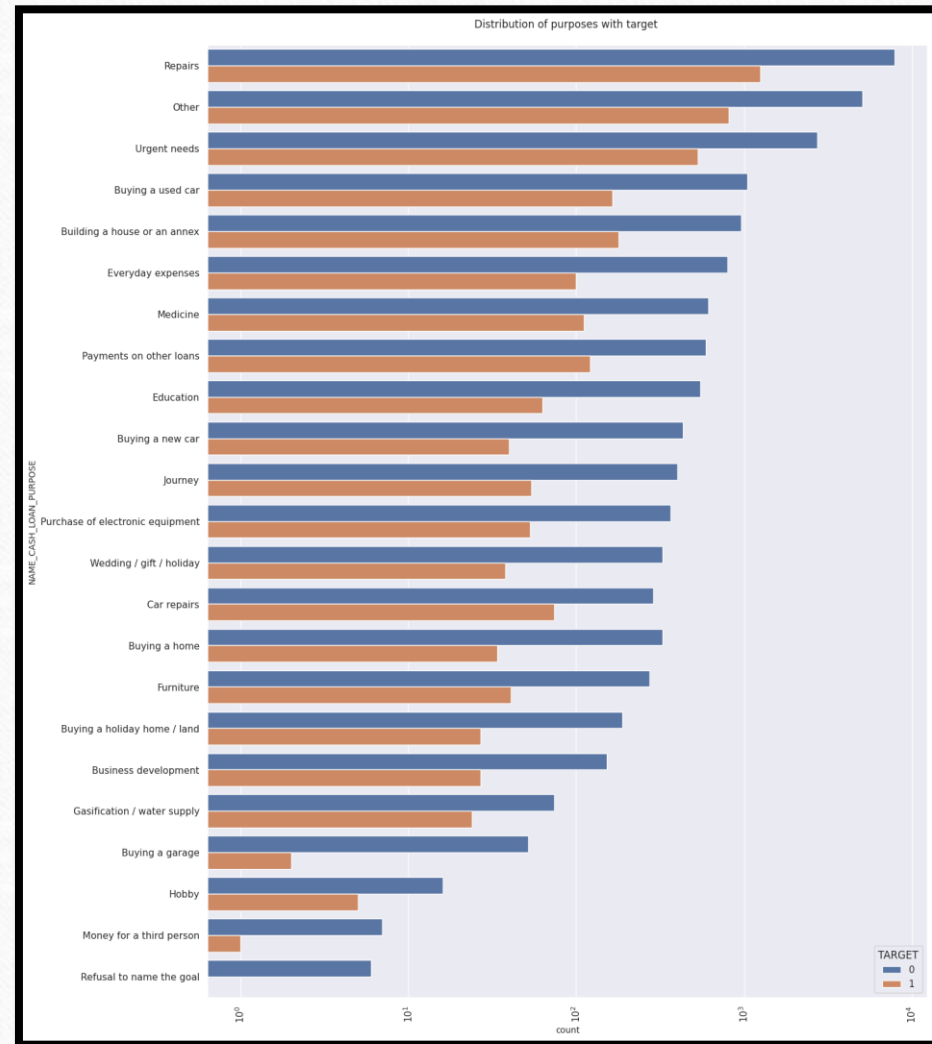


## Distribution of purposes with target

1. Loan purposes with 'Repairs' are facing more difficulties in payment on time.
2. There are few places where loan payment is significantly higher than facing difficulties.

They are 'Buying a garage', 'Business development', 'Buying land', 'Buying a new car', and 'Education'

Hence we can focus on these purposes for which the client is having minimal payment difficulties.

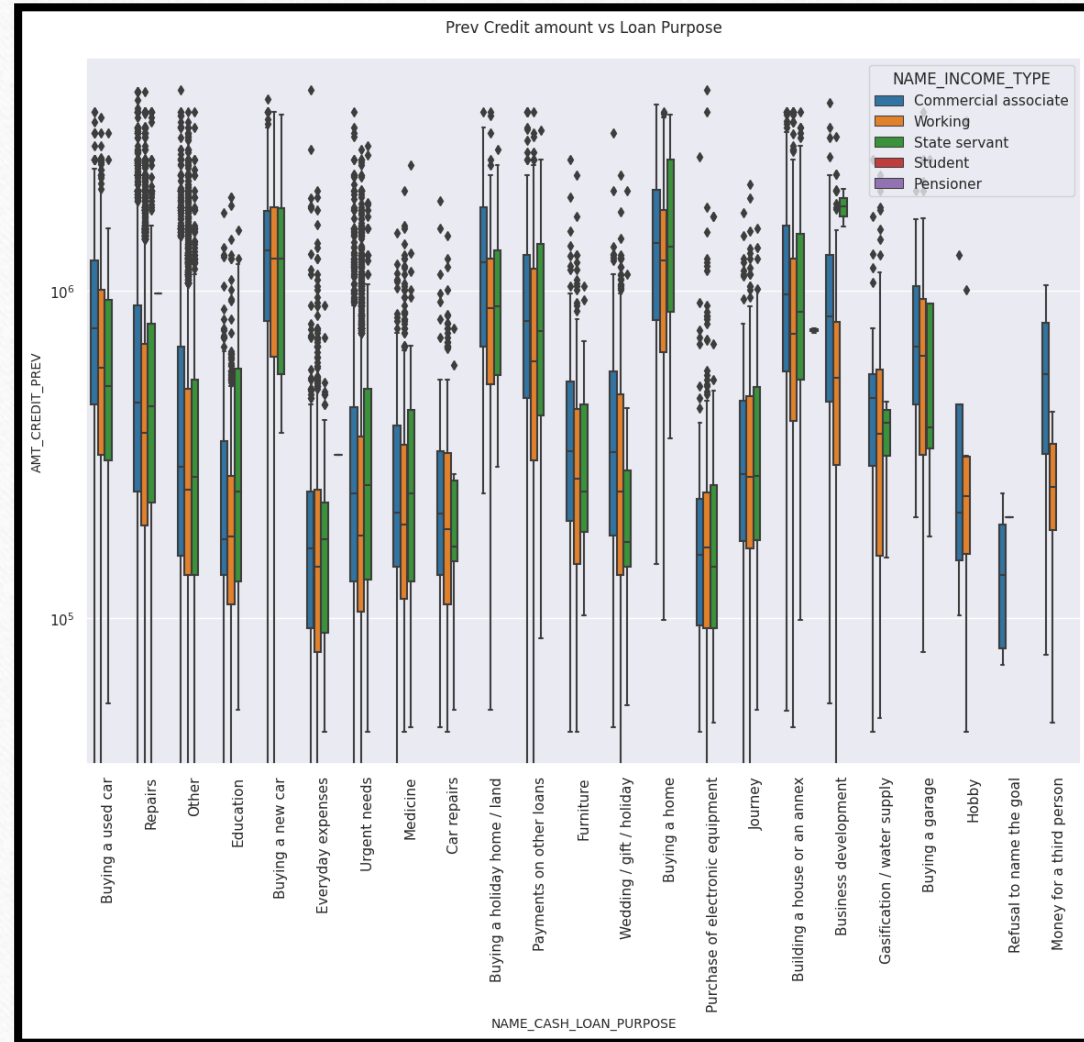


---

# BIVARIATE ANALYSIS AFTER MERGING PREVIOUS DATA

## Credit amount

1. The credit amount for Loan purposes like 'Buying a home', 'Buying 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 a third person or a Hobby is having fewer credits applied.



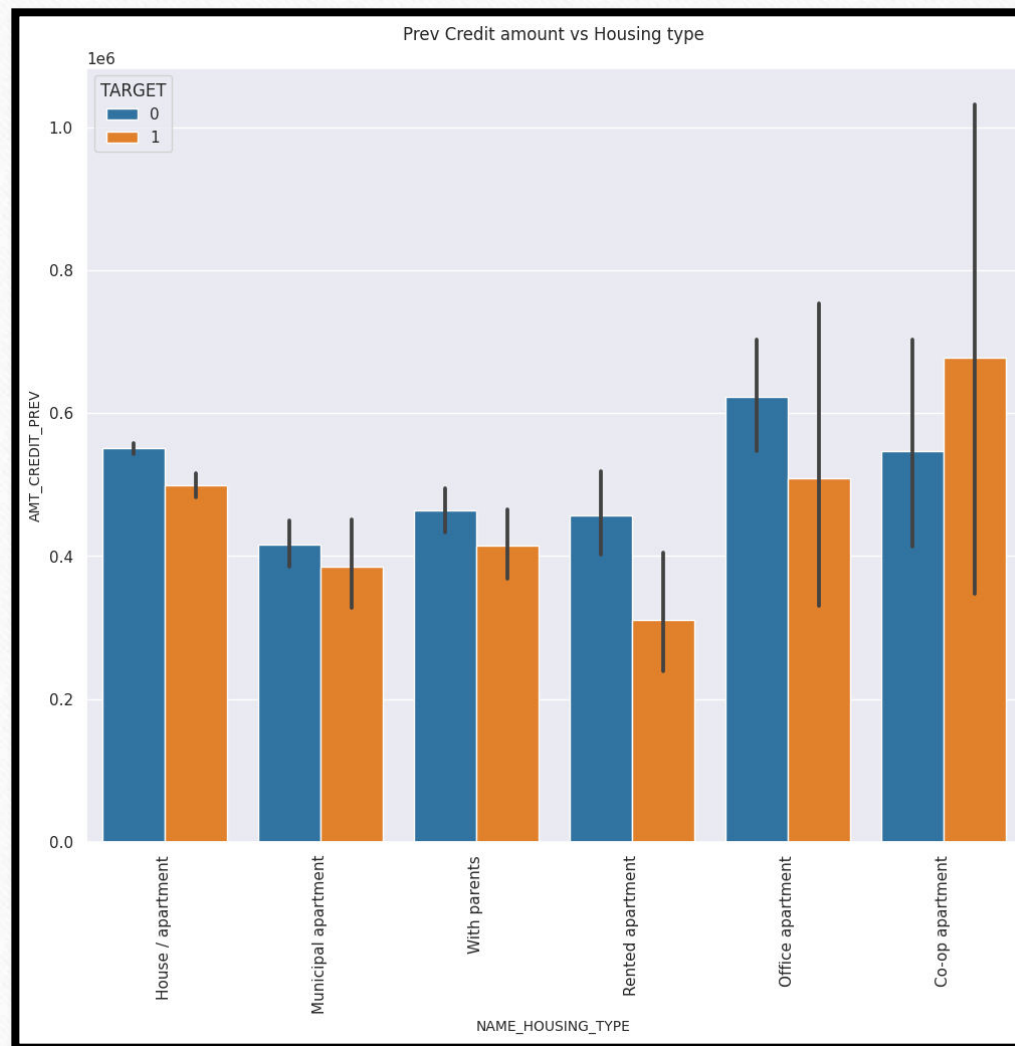


## Credit amount vs Housing type

Here for the Housing type, the office apartment has having higher credit of target 0, and the co-op apartment has having higher credit of target 1.

So, we can conclude that banks should avoid giving loans to the housing type of co-op apartments as they are having payment difficulties.

Bank can focus mostly on housing type with parents or House\apartment or municipal apartment for successful payments.



# CONCLUSION

---

1. For payments to be successful, banks should concentrate more on contract types such as "student," "pensioner," and "businessman" with housing types other than "co-op apartments."
2. Banks should pay less attention to "working" income types because they account for the majority of failure payments.
3. Additionally, "Repair" loans have a higher percentage of late or missed payments.
4. Acquire as many clients as possible from the dwelling type "With parents" since they have the fewest failed payments.



Thank You