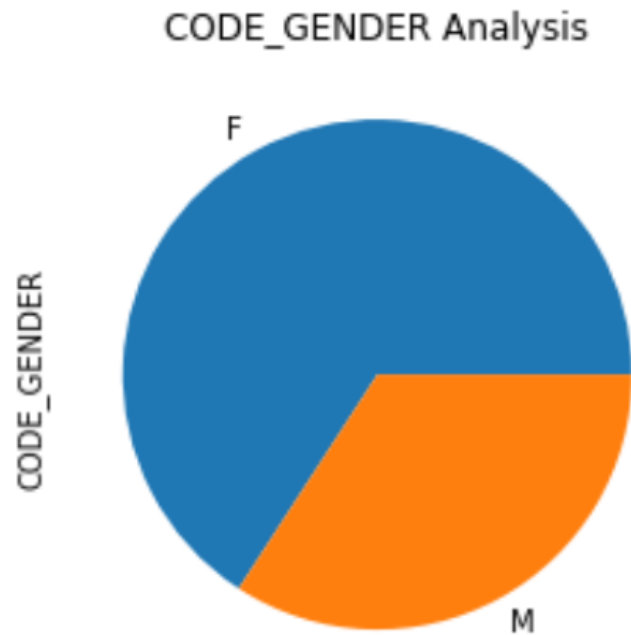


Credit Card EDA Case study

- ▶ Rohit Patnaik
 - ▶ rohitpatnaik27@gmail.com
 - ▶ +91-8895-291-619
- ▶ Yoshitha Palavali
 - ▶ yoshi.palavali@gmail.com
 - ▶ +91-8096-579-756

Introduction

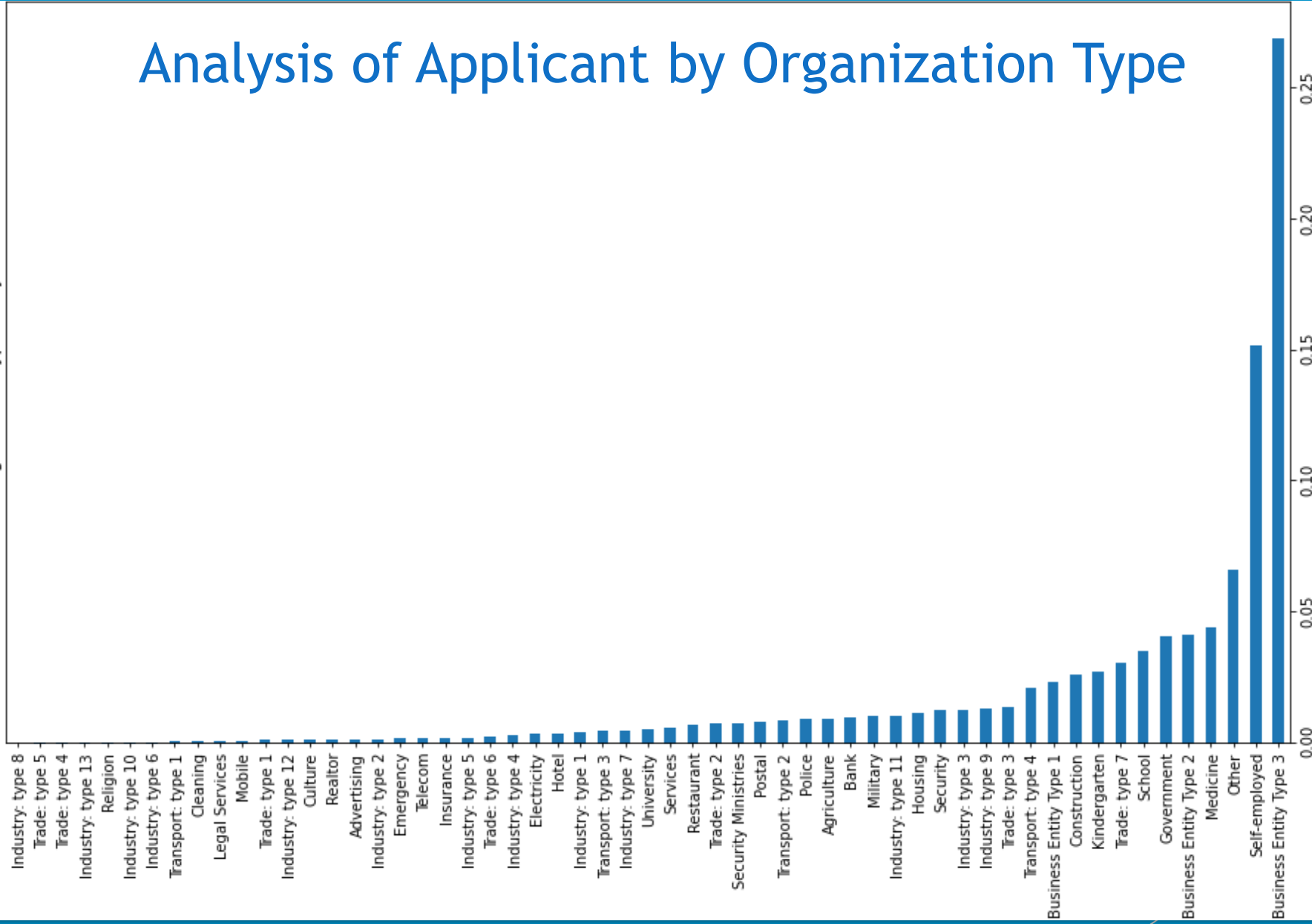
- The Case Study is about a loan providing companies which have difficulty in approving loans. We have used EDA to analyse the patterns present in the data. This will ensure that the applicants capable of repaying the loan are not rejected.
- When the company receives a loan application, the company must decide for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:
 - If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
 - If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.



Code Gender Analysis – It indicates that the loan applicants are majorly female than compared to male.

Analysis of Applicant by Organization Type

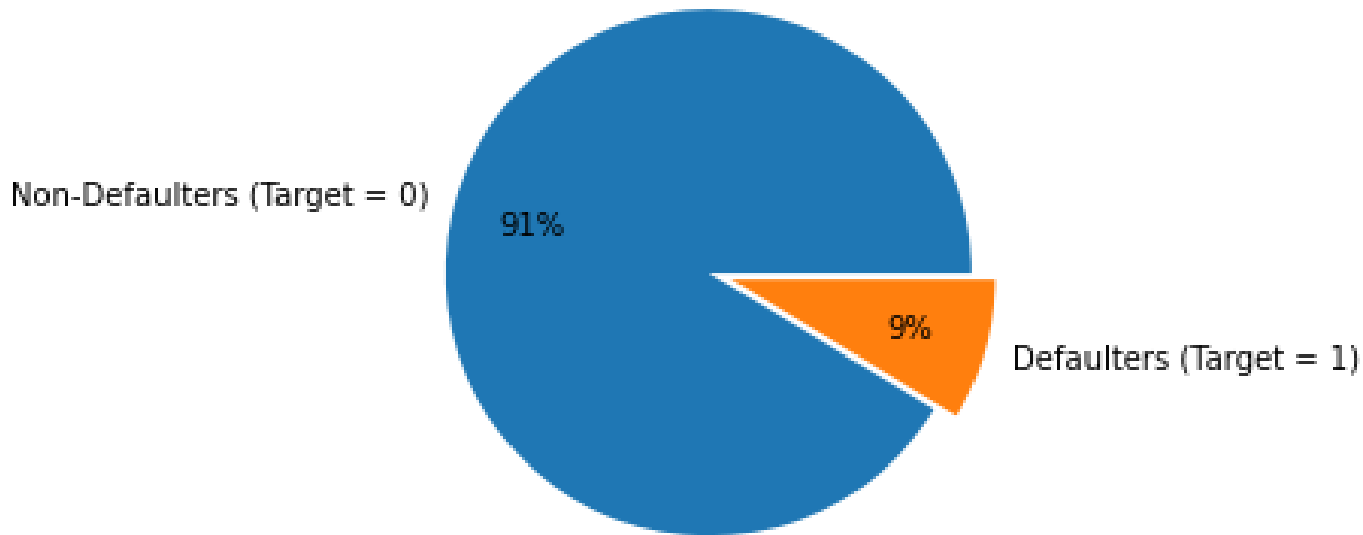
Organization Type Analysis



The graph indicates the loan intake of people by type. Business Entity and Self Employed are high in comparison to other occupation.

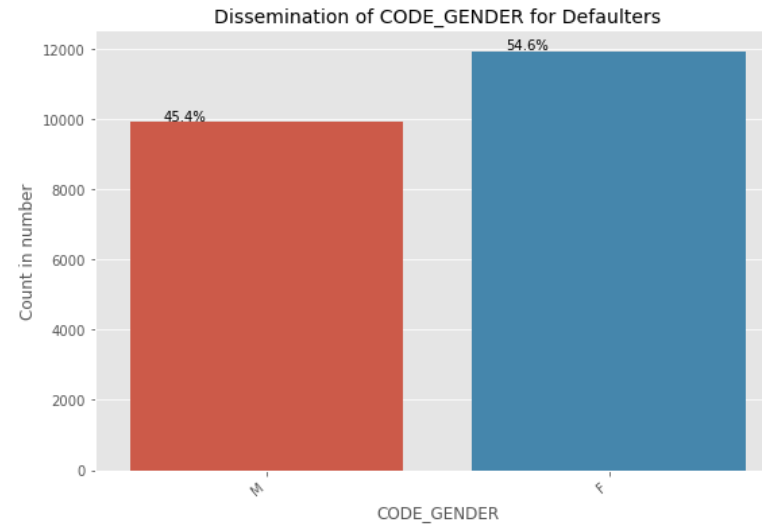
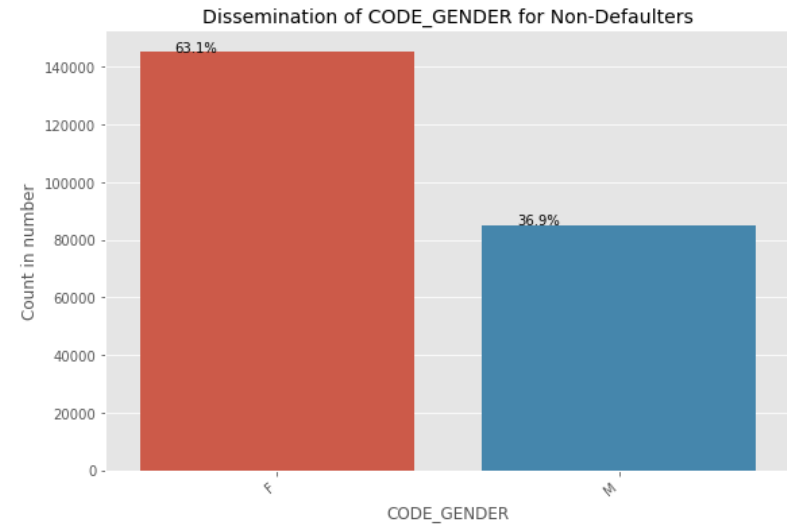
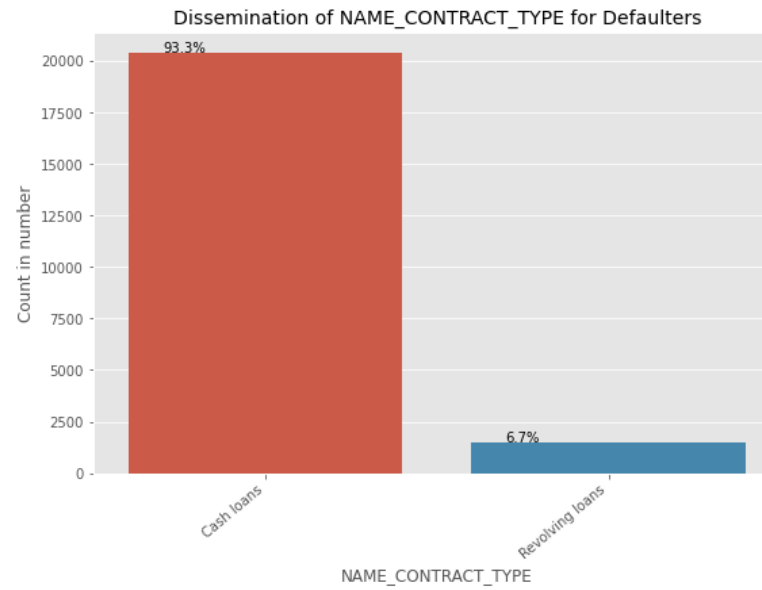
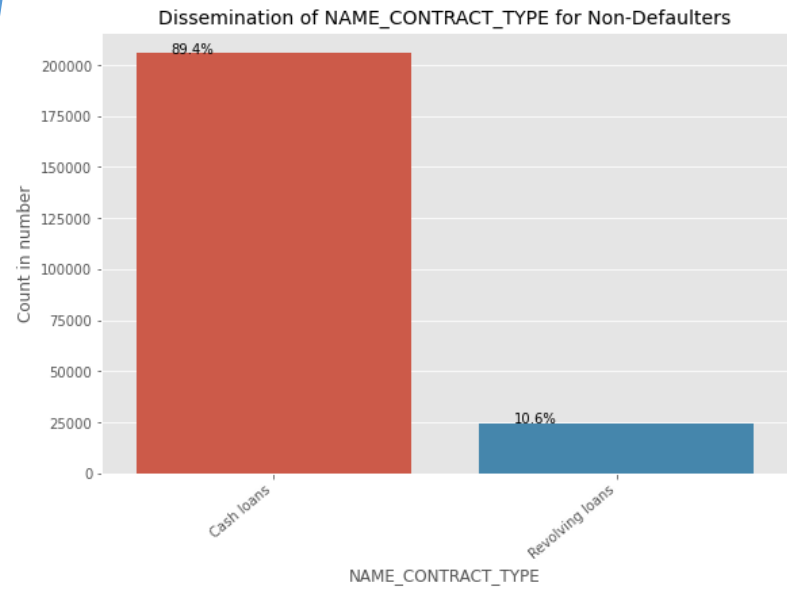
Analysis of Applicant by Target Variable (Payment)

TARGET Variable; Defaulter Vs Non-Defaulter



- ▶ Around 91% of client aren't having difficulties in payment i.e., non-defaulters list and rest 9% of clients are facing difficulties in payment.

Categorical Analysis based on Gender and Loan type



Inference : Categorical Analysis based on Gender and Loan type

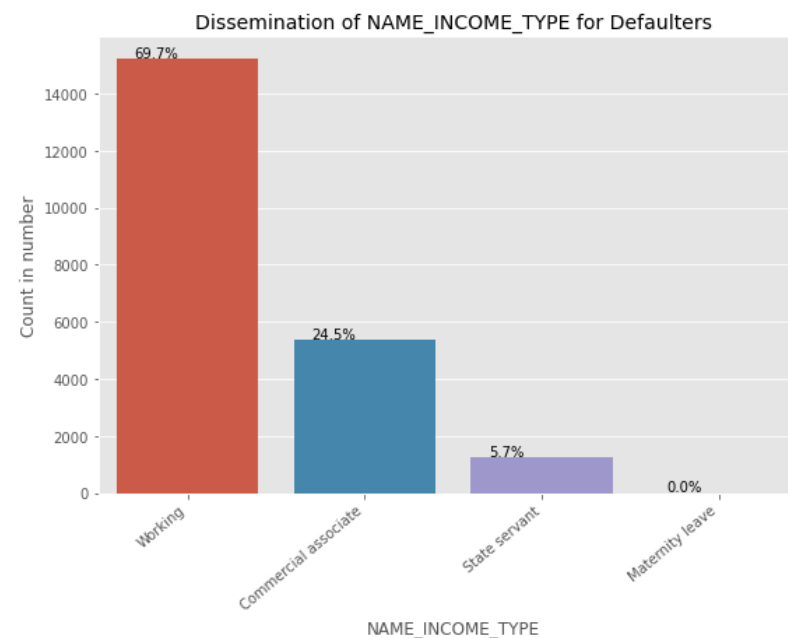
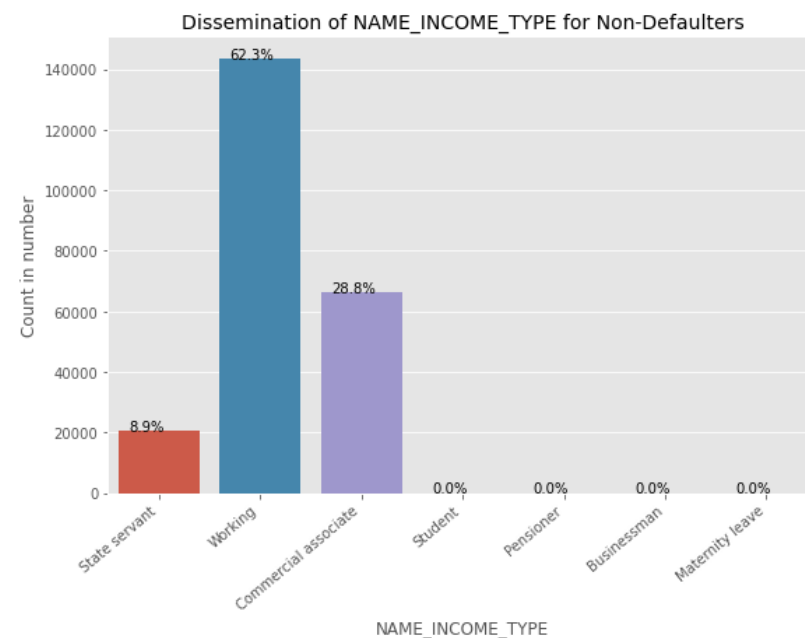
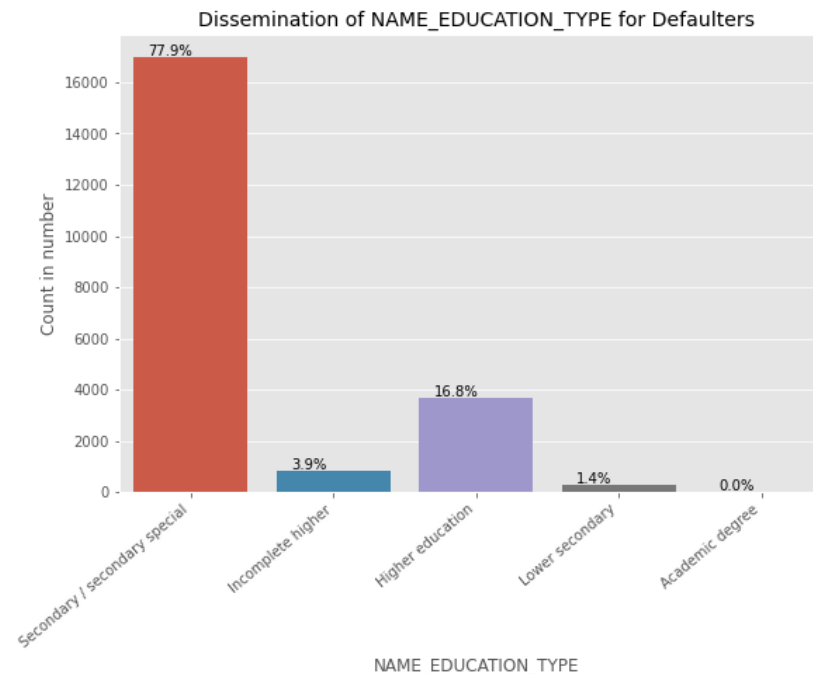
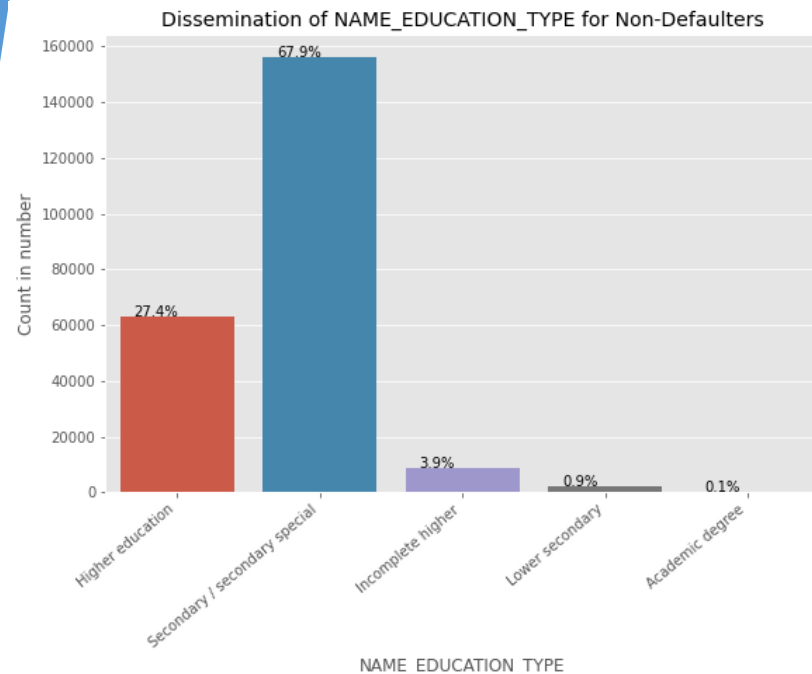
▶ LOAN_TYPE(CONTRACT) DEFAULTER & NON-DEFAULTER ANALYSIS:

- ▶ Most number of defaulters are having cash loans rather than revolving loans.

▶ CODE_GENDER DEFAULTER & NON-DEFAULTER ANALYSIS:

- ▶ Female contributes to 63% to the non-defaulters whereas Male contributes to 37% of the non-defaulters.
- ▶ To the defaulters list, Female contributes more than Male.
- ▶ Female applying for loans are more than the number of male applying for loan. Hence, we may infer that number of female defaulters are more.

Categorical Analysis based on Education Type & Income



Inference: Categorical Analysis based on Education Type and Income

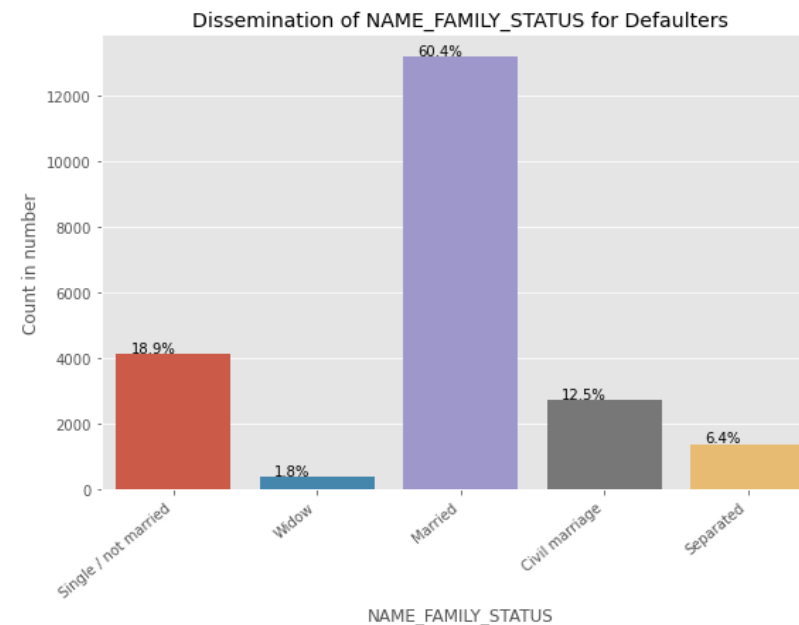
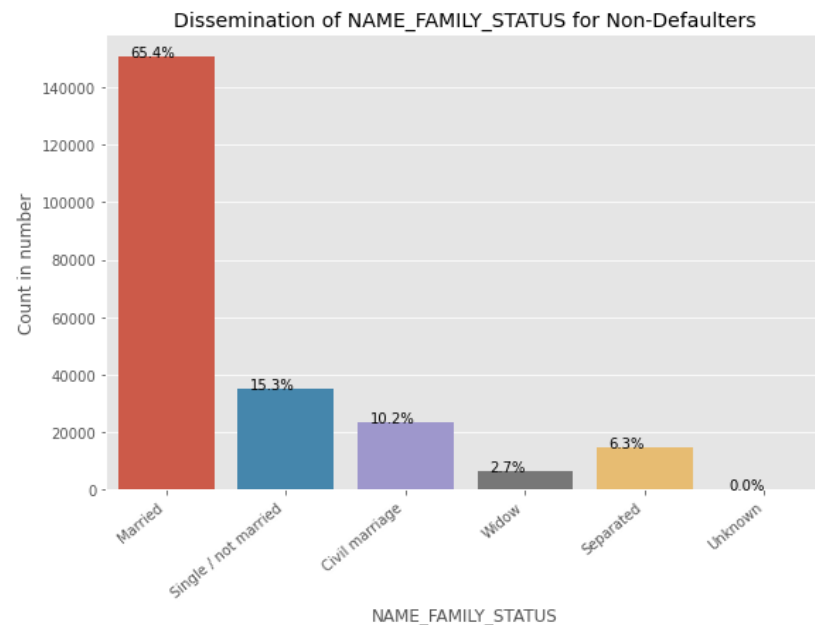
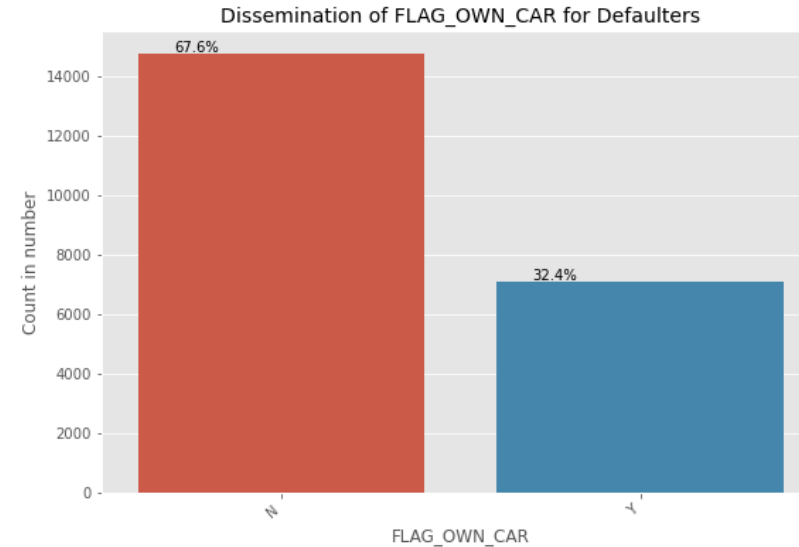
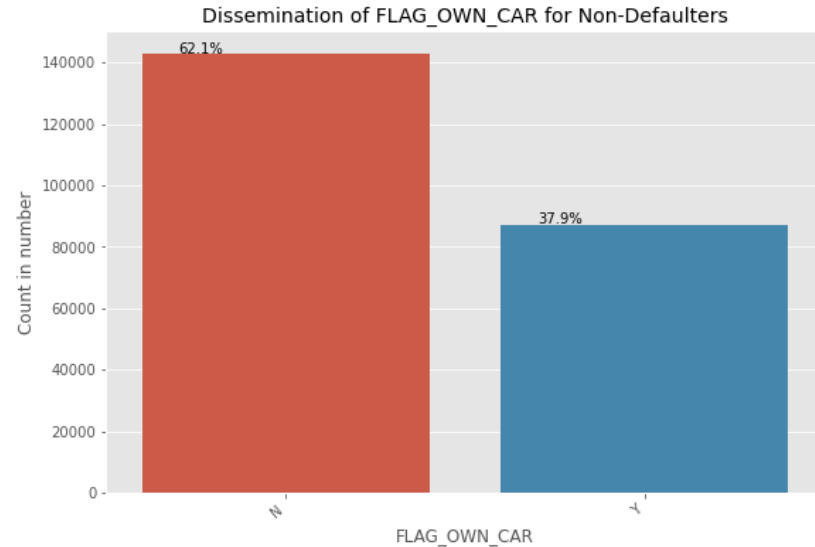
▶ EDUCATION TYPE DEFAULTERS & NON_DEFAULTERS ANALYSIS:

- ▶ Mostly loans are distributed to "Secondary/Secondary Special Education" group.
- ▶ As the number of loans distributed to Secondary/Secondary Special Education group is high, secondary/secondary special group contribution to defaulters list is also high.

▶ INCOME TYPE DEFAULTERS & NON_DEFAULTERS ANALYSIS:

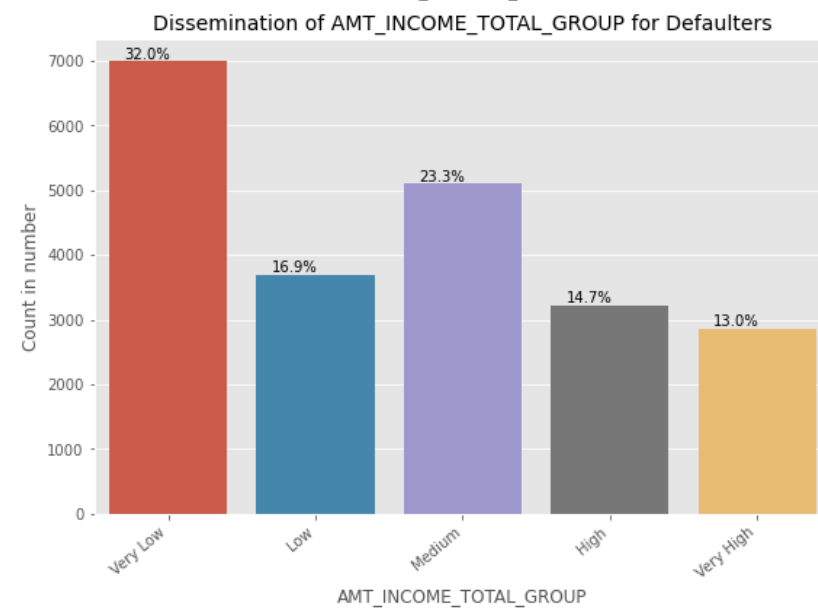
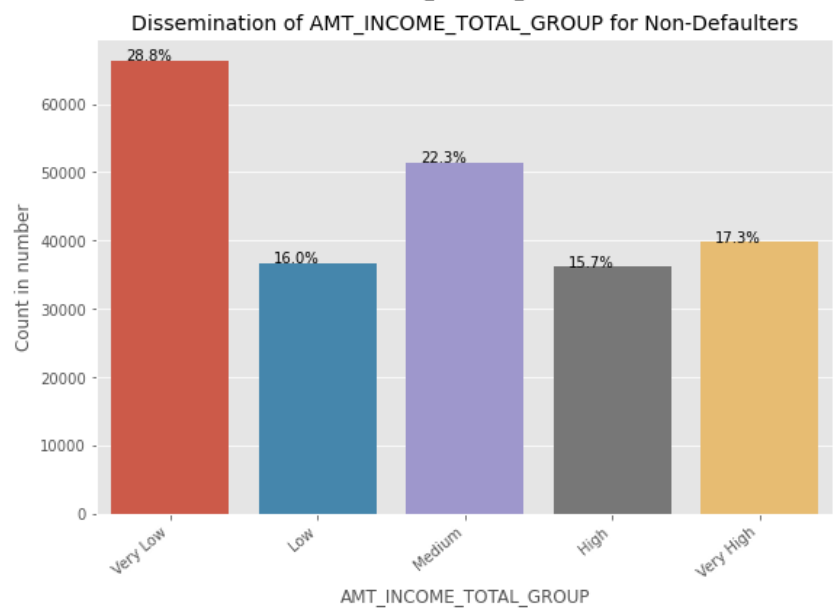
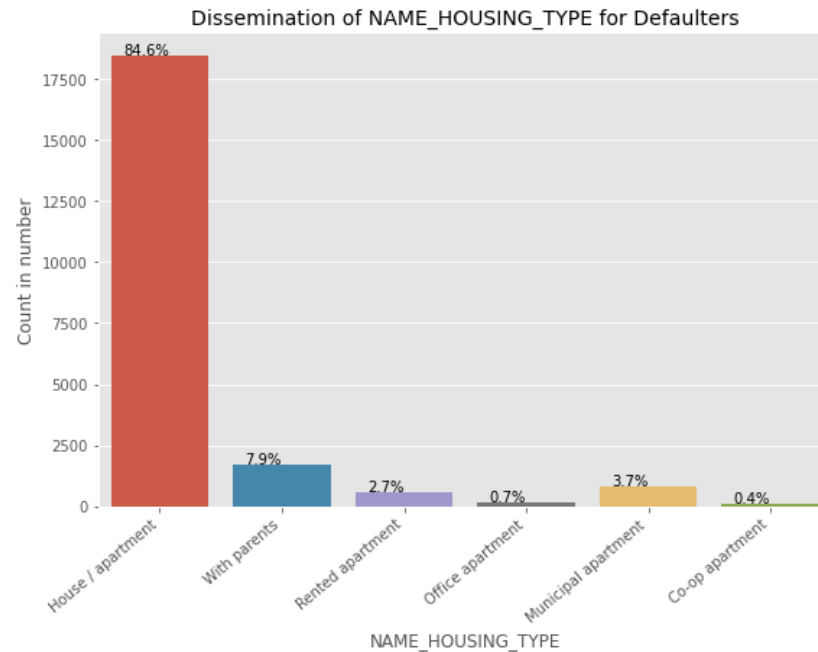
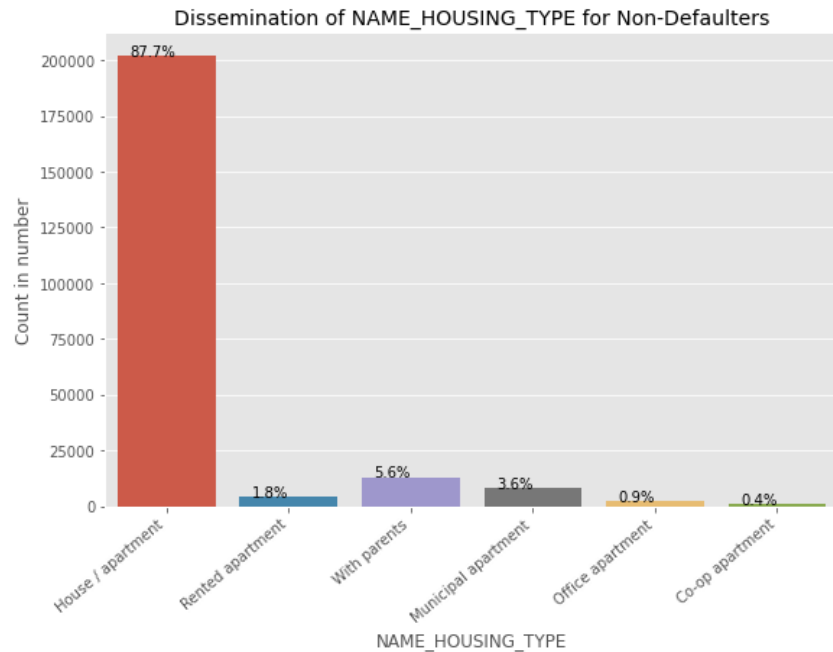
- ▶ Student, Pensioner, Businessman and women with maternity leaves are not likely to be in the defaulter list.
- ▶ It can be inferred that most of the loans are distributed to working class, commercial associate and state servant.
- ▶ We can also observe that, working class persons contributes 62% to non-defaulters list whereas around 70% to the defaulters list. It shows that the working-class clients are more in the defaulters list.

Categorical Analysis - Owns car and Family Status



Inference: Categorical Analysis - Owns car and Family Status

- ▶ **FLAG OWN CAR DEFAULTER & NON-DEFAULTER ANALYSIS:**
 - ▶ We observe that, People having their own car contributes to 38% of the non-defaulters & 32% of the defaulters.
 - ▶ The person who owns a car makes into default list more than people without car.
 - ▶ The rate of defaulters of people having a car is low compared to rate of people who don't own a car.
- ▶ **FAMILY STATUS DEFAULTER & NON-DEFAULTER ANALYSIS:**
 - ▶ Married People tend to apply loan more than any other group.
 - ▶ We observe that single/ unmarried, civil marriages and separated persons contribute more number to the defaulters list. So, it might be a risk to give loans to these categories.

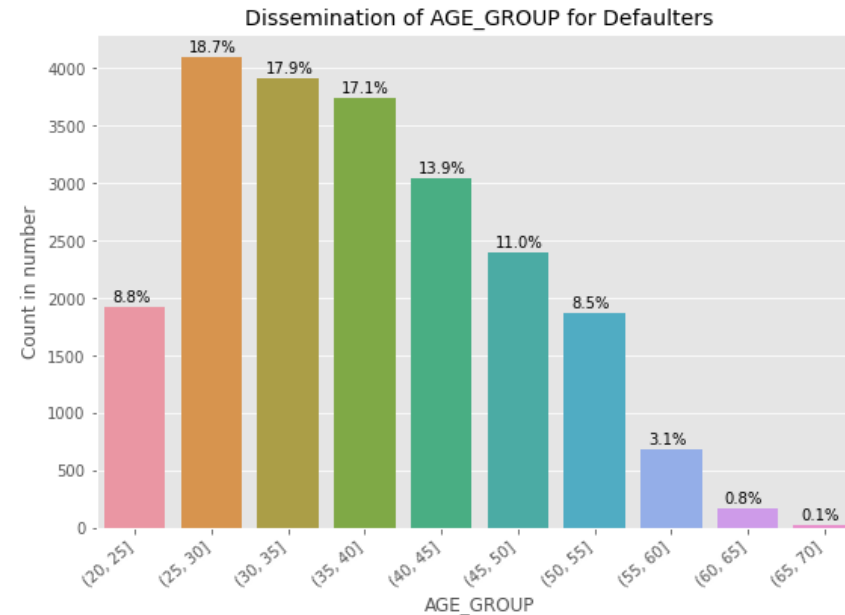
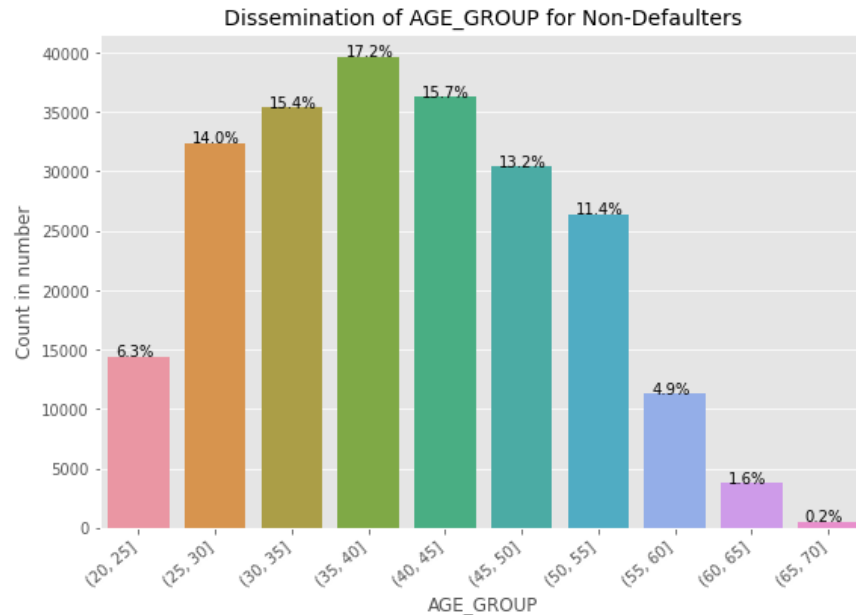


Categorical Analysis - Housing Type & Total Income Group

Inference: Categorical Analysis - Housing Type & Total Income Group

- ▶ HOUSING TYPE DEFAULTERS & NON-DEFAULTERS ANALYSIS:
 - ▶ Persons with own house/apartment apply more for loans rather than any other group.
 - ▶ Persons living with parents contribute more to defaulters list rather than any other group. It is quite possible that as parents are staying near by, so living expenses are higher and so contributed to defaulters list.
- ▶ TOTAL INCOME GROUP DEFAULTERS & NON-DEFAULTERS ANALYSIS:
 - ▶ The very high-income group tend to default less than the other class of income group.
 - ▶ Very low-income group tend to have more defaulters.

Categorical Analysis by Age Group

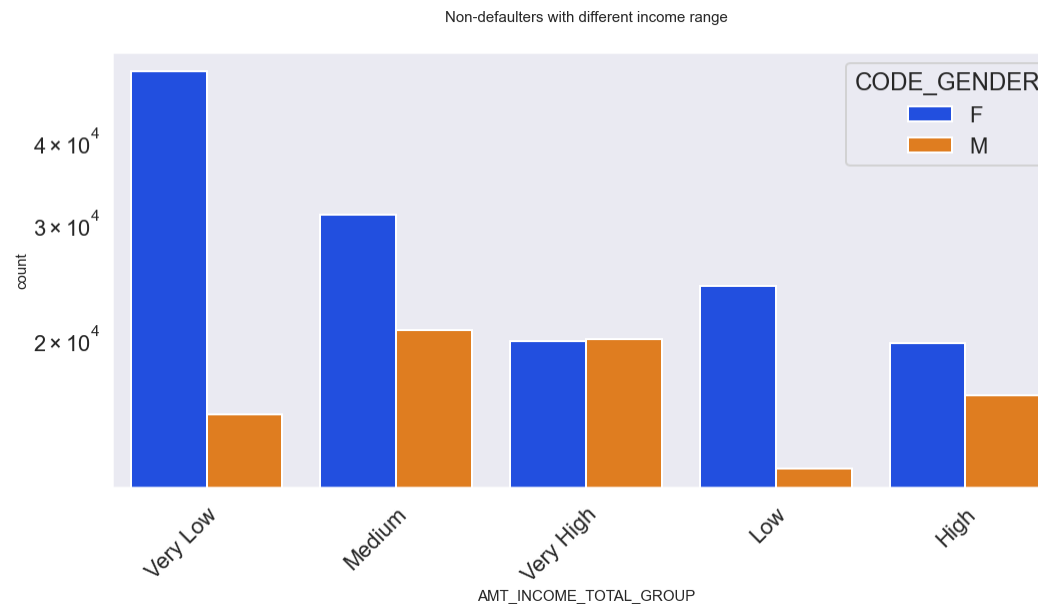


- ▶ It is observed that persons between the age of (25-35), tend to default more. So, it is risky to approve loan for this age group.
- ▶ With increase in age, persons tend to default less. One of the reason could be they get employed during this age and their salary also increases. So, they repay the loan.

Analysis by Total Income Group and Gender

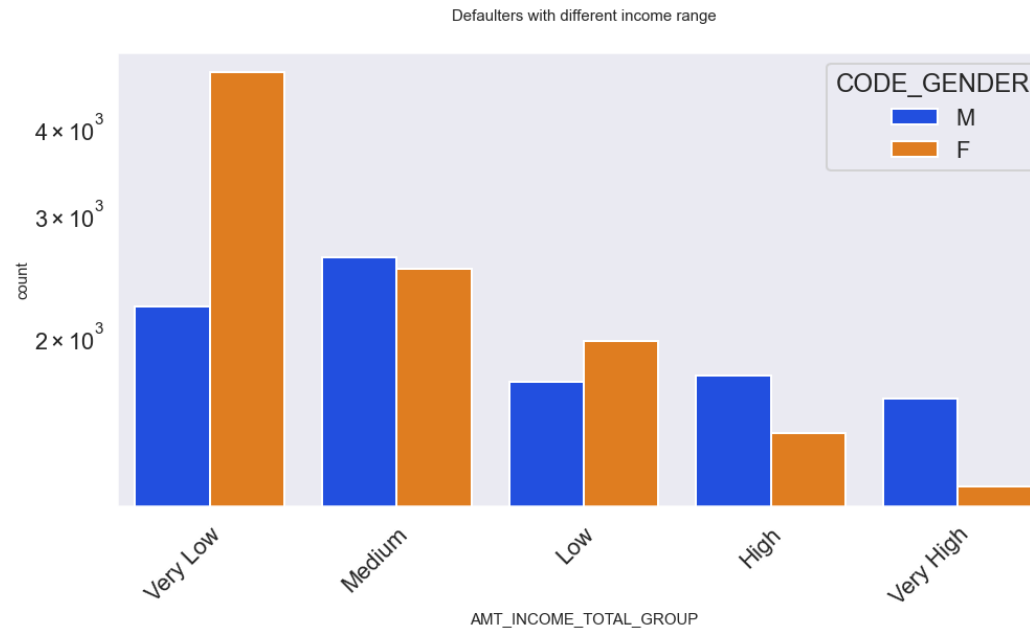
► NON-DEFAULTERS:

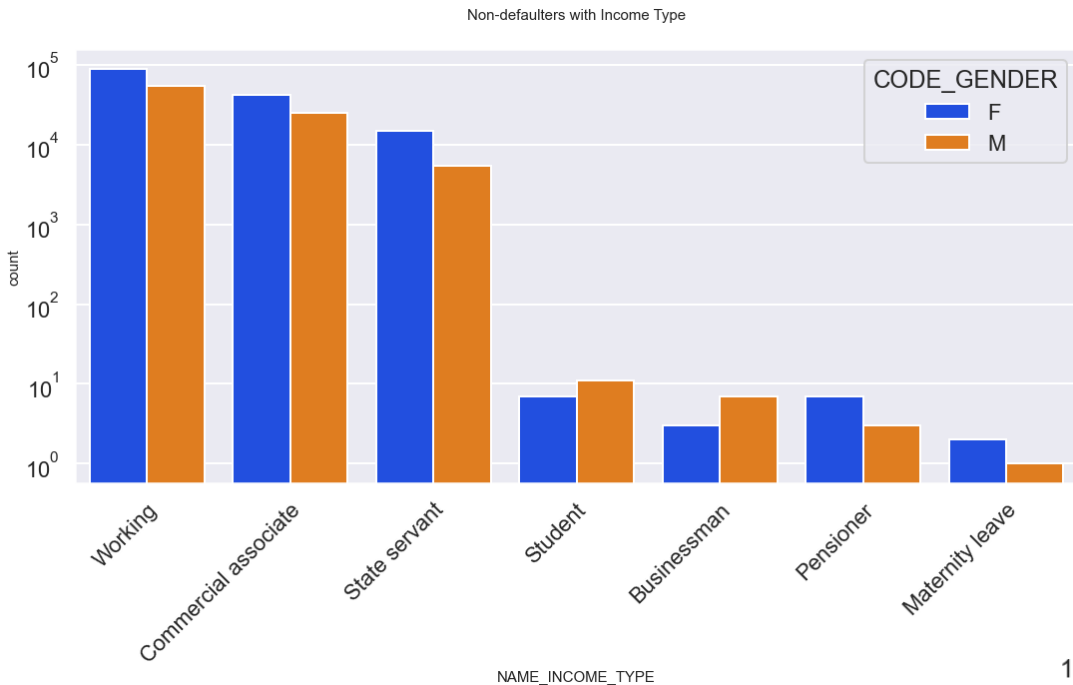
- Count of female is more than that of count of male in the non-defaulters category.



► DEFAULTERS:

- In the defaulters list, most likely Male in the Medium, High and Very High-income range tend to be defaulters.



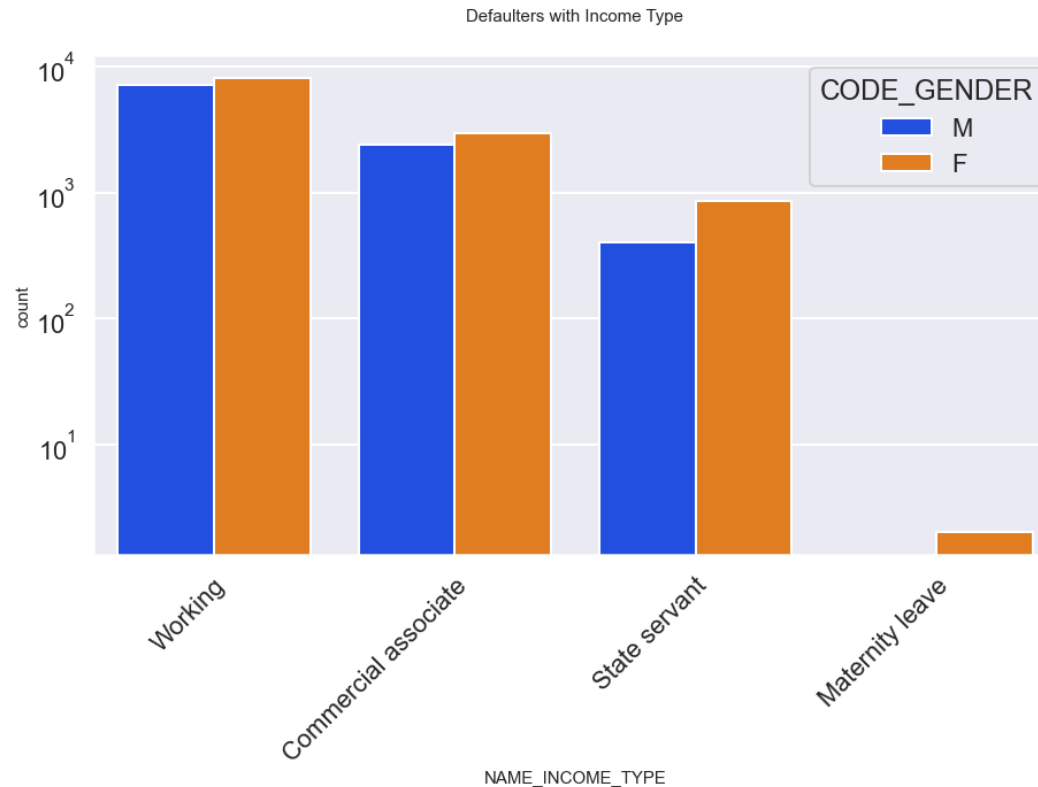


► NON-DEFAULTERS:

- According to working group, Female are having more credit than Male
- High credit goes for working, commercial associate, state servant
- Low credit goes for student, pensioner, businessman and maternity leave.

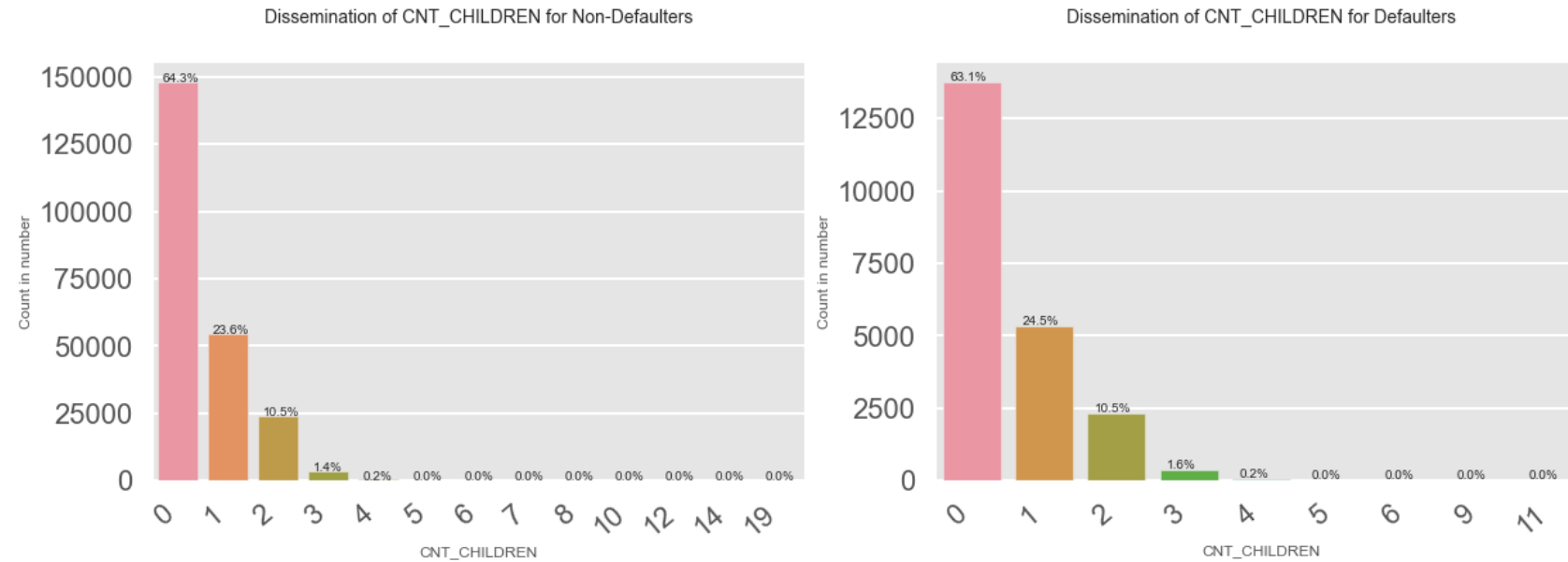
► DEFAULTERS:

- Mostly Female who are working, commercial associate and state servant tend to be more defaulters than any other categories.



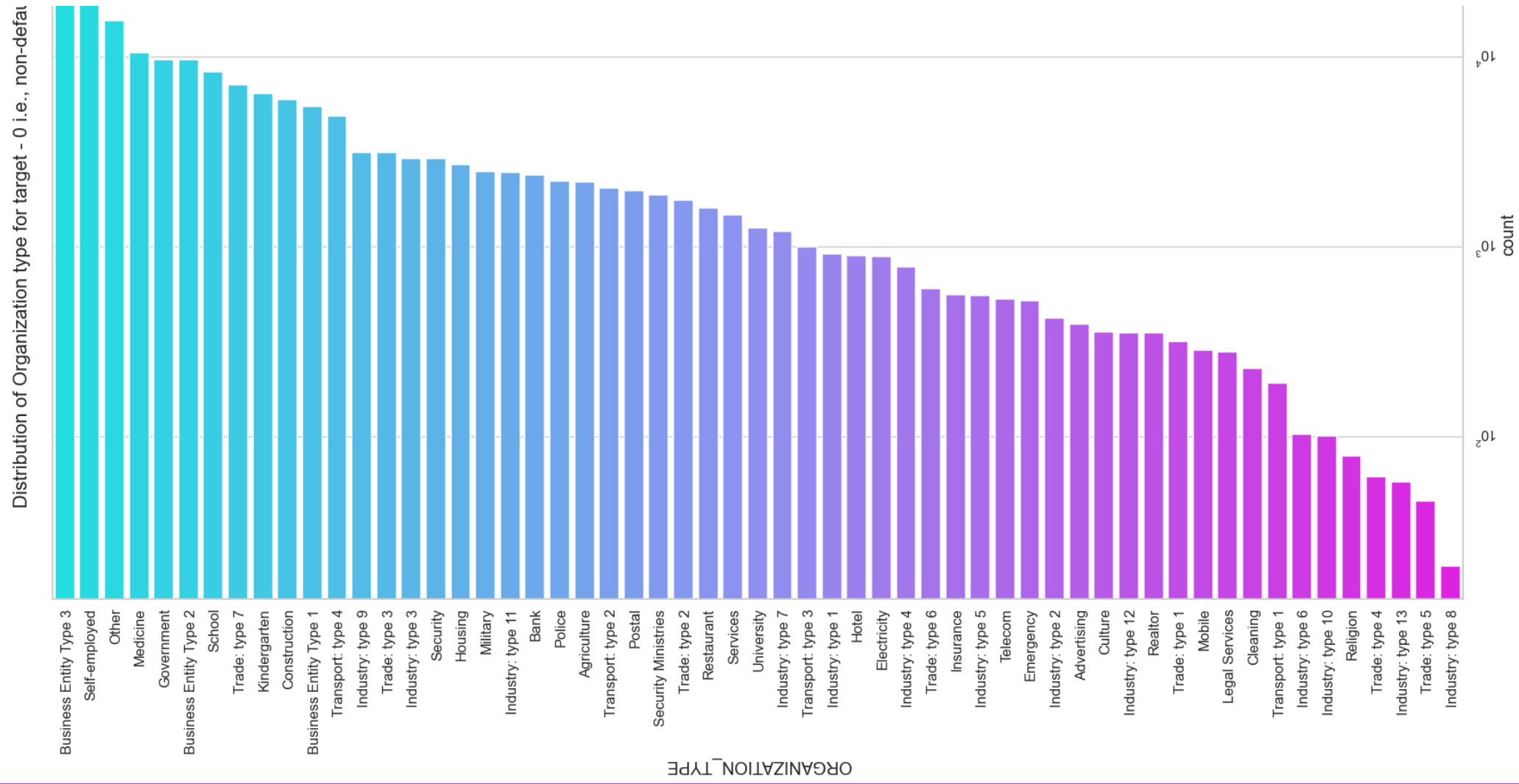
Analysis
by Income
Type and
Gender

Analysis by Count of Children for the Applicants



- The above chart is unpredictable as low number of children contribute to both defaulters and non-defaulters.

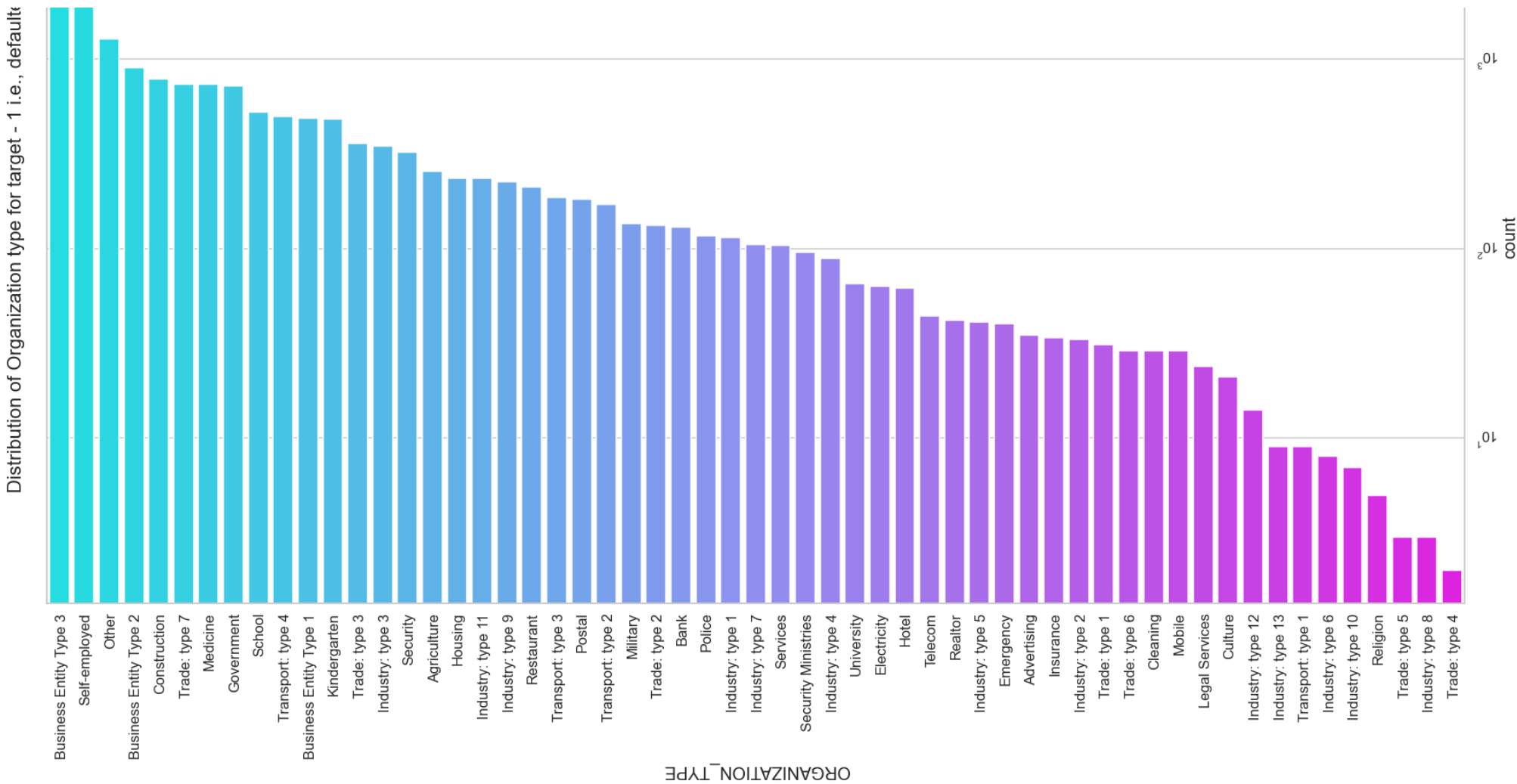
Distribution of Organization type for target - 0 i.e., non-defaulters



Inference : Distribution of Organization type for target - 0 i.e., non-defaulters

- ▶ Business Entity, Self-employed, Other, Medicine, Government organization type are the top 5 organizations applying for credit.
- ▶ Industry, Trade, Religion organization type are the bottom 5 organizations to apply for credit. We might infer that bank might not provide credit to these type organizations as risk is involved in repayment.

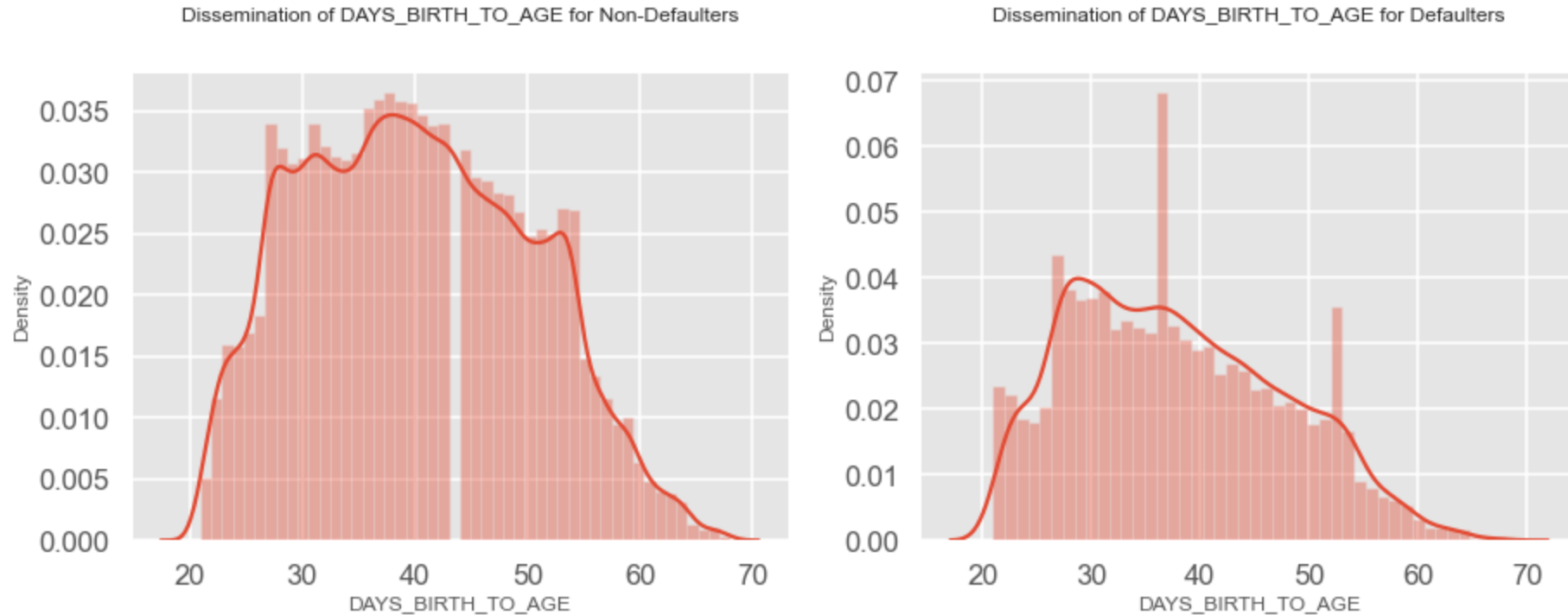
Distribution of Organization type for target - 1 i.e., Defaulters



Inference : Distribution of Organization type for target - 1 i.e., Defaulters

- ▶ Business Entity, self employed, other, trade and constructions are most likely to get on defaulters list. We may infer that as these type group persons does not get a continuous flow of income as salary, so they are the one's who are likely to get in the defaulters list.
- ▶ Trade, Industry, Religion are bottom 5 organizations who are getting into the defaulters list.

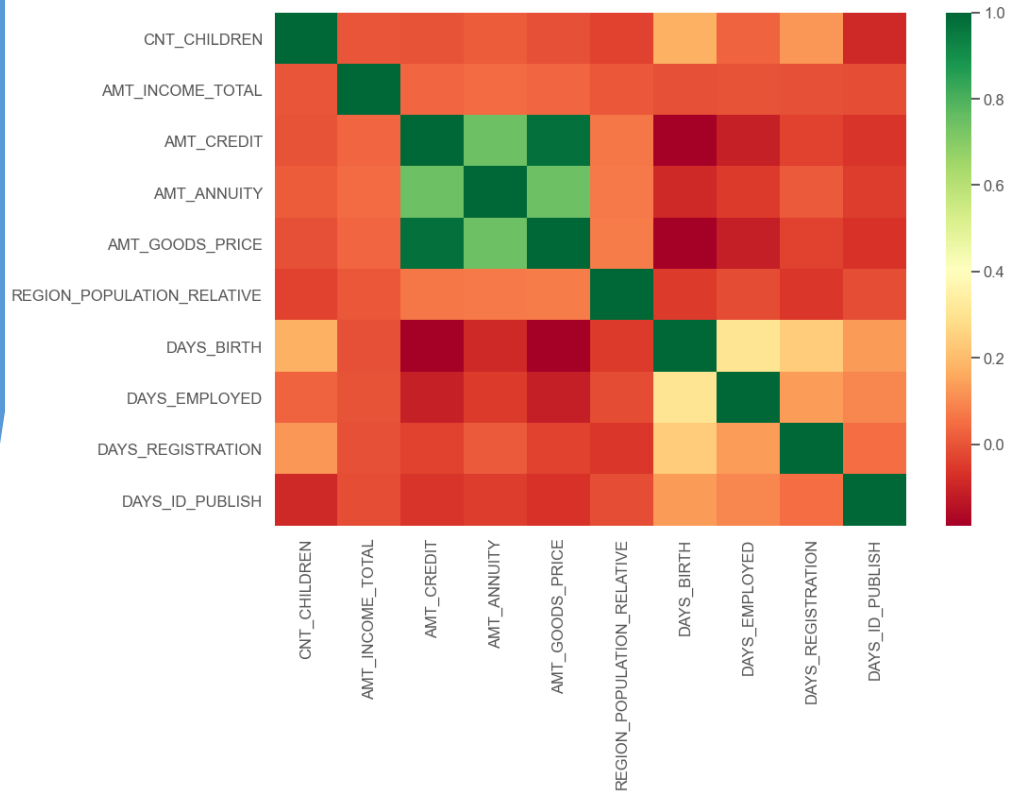
Analysis by Days Of Birth



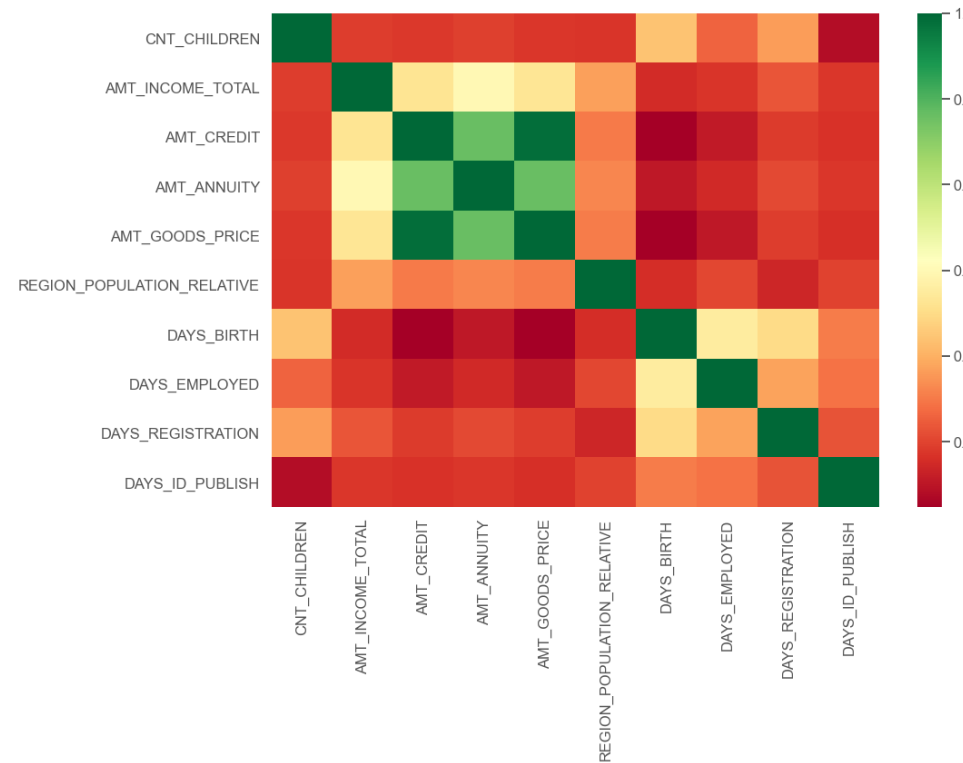
- Persons applying for credit are on the age group of (25-55) whereas mostly persons from age group (35-40) and (50-55) are likely to be defaulters.

Correlation Analysis

Correlation for defaulters



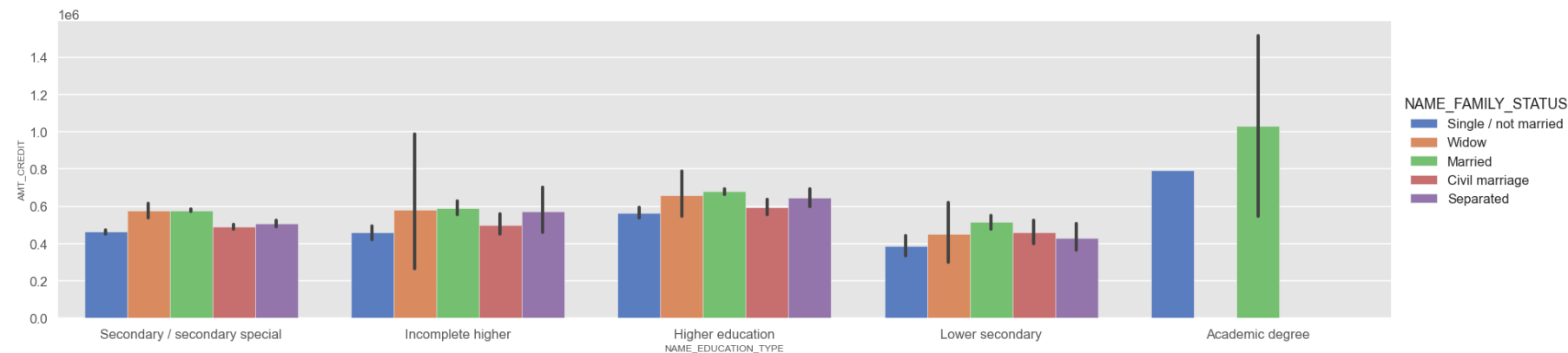
Correlation for non-defaulters



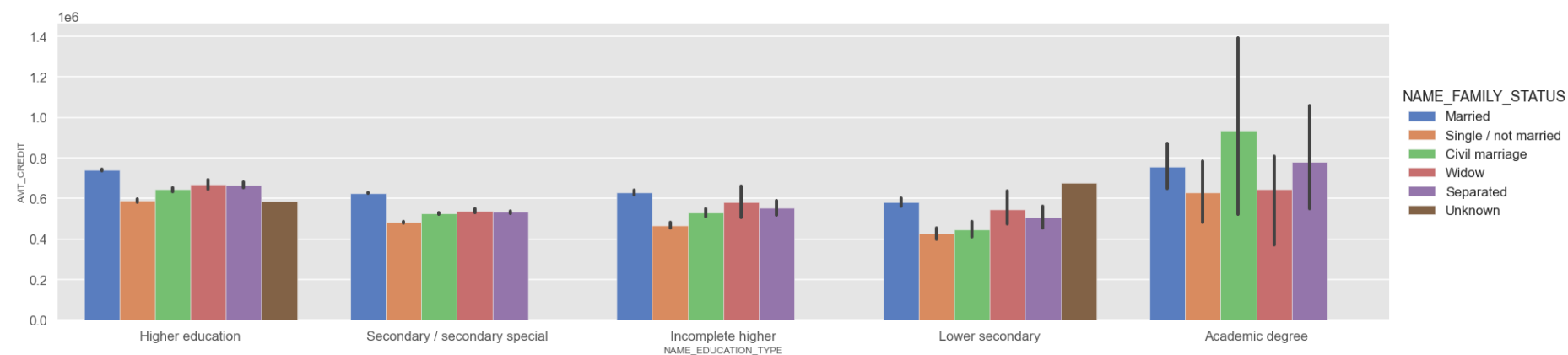
- The following columns are having high correlation :
 - AMT_CREDIT and AMT_GOODS_PRICE
 - AMT_CREDIT and AMT_ANNUITY
 - AMT_ANNUITY and AMT_GOODS_PRICE

Credit Amount vs Education Status Analysis

Credit amount vs Education Status for defaulters



Credit amount vs Education Status for non-defaulters



Inference: Credit Amount vs Education Status Analysis

► DEFAULTER:

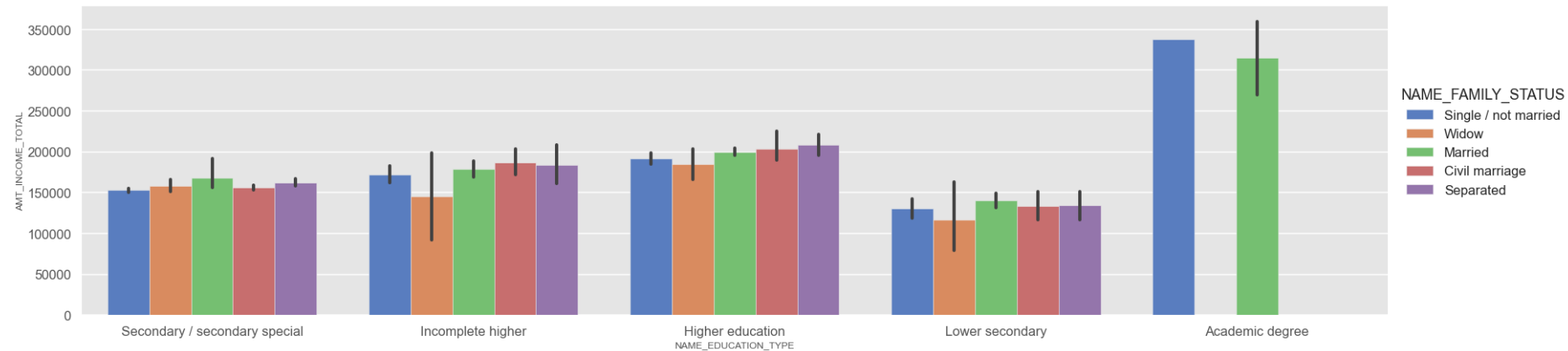
- Persons falling under married, single and having academic degree are most likely to get on defaulters list.
- Among all the education categories, married persons are most likely to be on the defaulters list.
- Persons with lower educational background will tend to have low credit amount.

► NON-DEFAULTER:

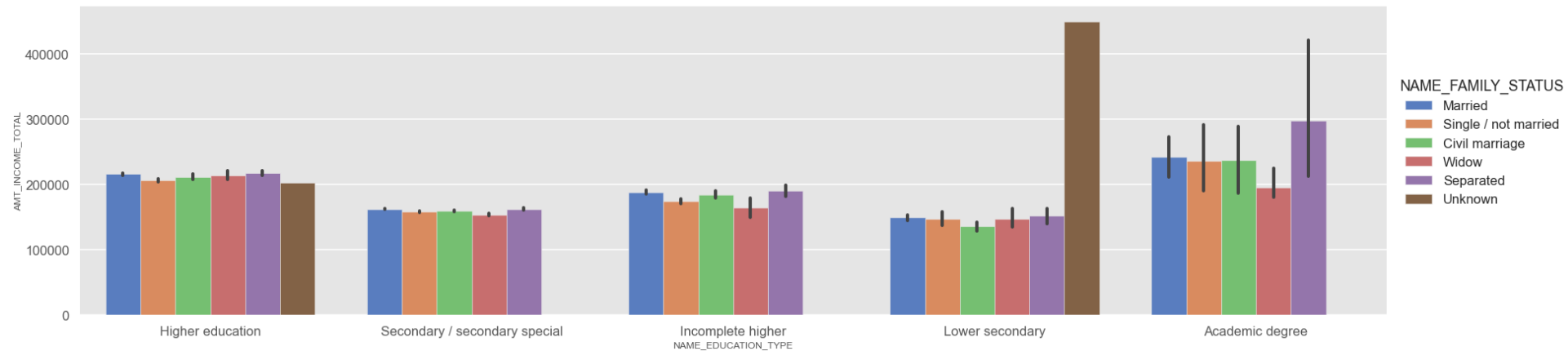
- Among all the categories, Lower secondary single/not married persons are having lower credit amount.
- Person with Education status as Academic Degree relatively have higher credit amount with civil marriage segment considered the best among all.
- Married persons are having greater credit amount among all education categories other than academic degree status.

Total Income Vs Education Status

Total income vs Education Status for defaulters



Total income vs Education Status for non-defaulters



Inference: Total Income Vs Education Status

► DEFAULTERS:

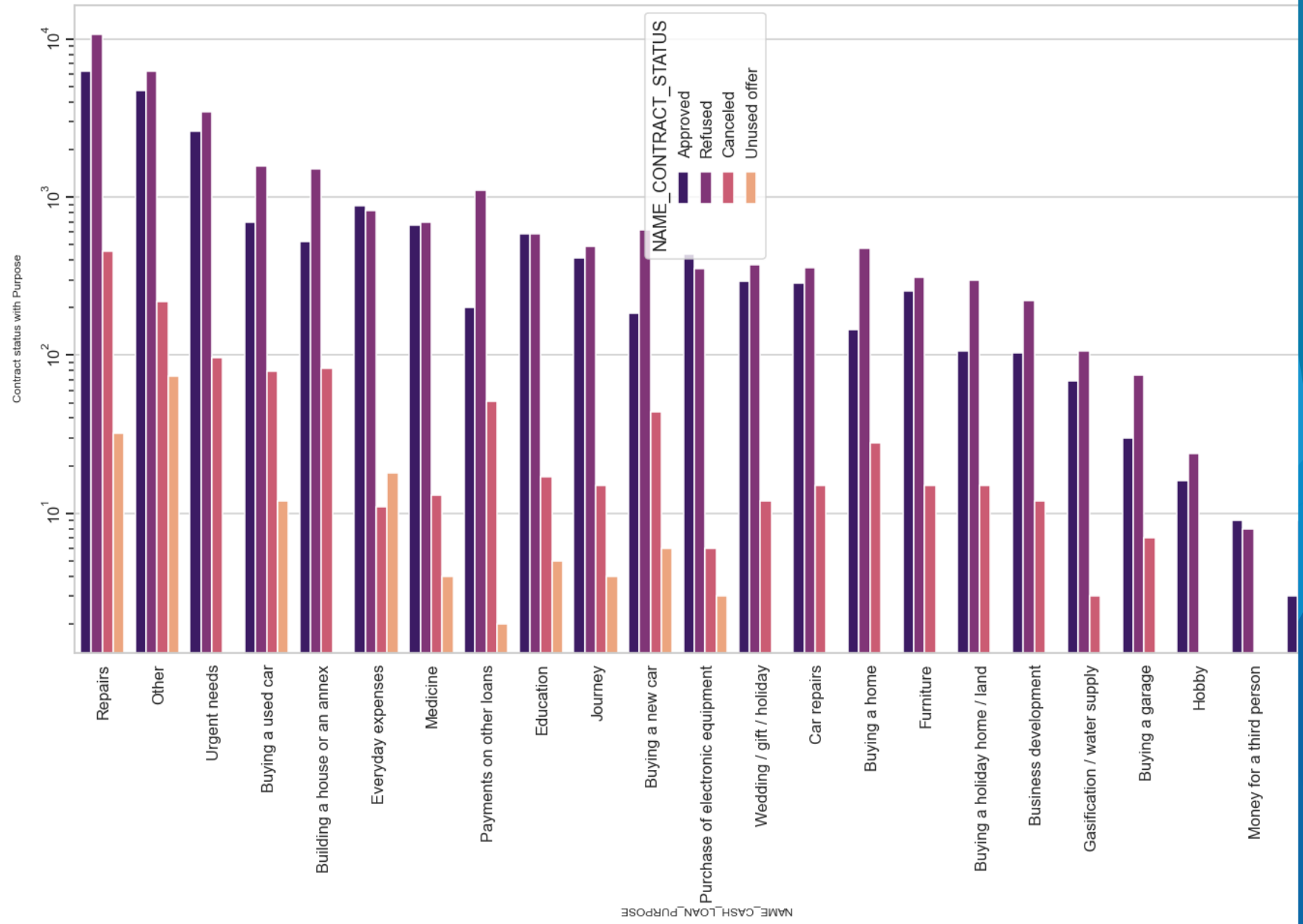
- Under academic degree, we just have two segments i.e., single and married who are likely to be defaulters.
- Person with lower secondary education are less likely to be defaulters. We may infer that as they are getting less income, so defaulters' segment is also less.
- Married persons among all education group are more likely to be defaulters but single person from the academic degree is the highest among all.

► NON-DEFAULTERS:

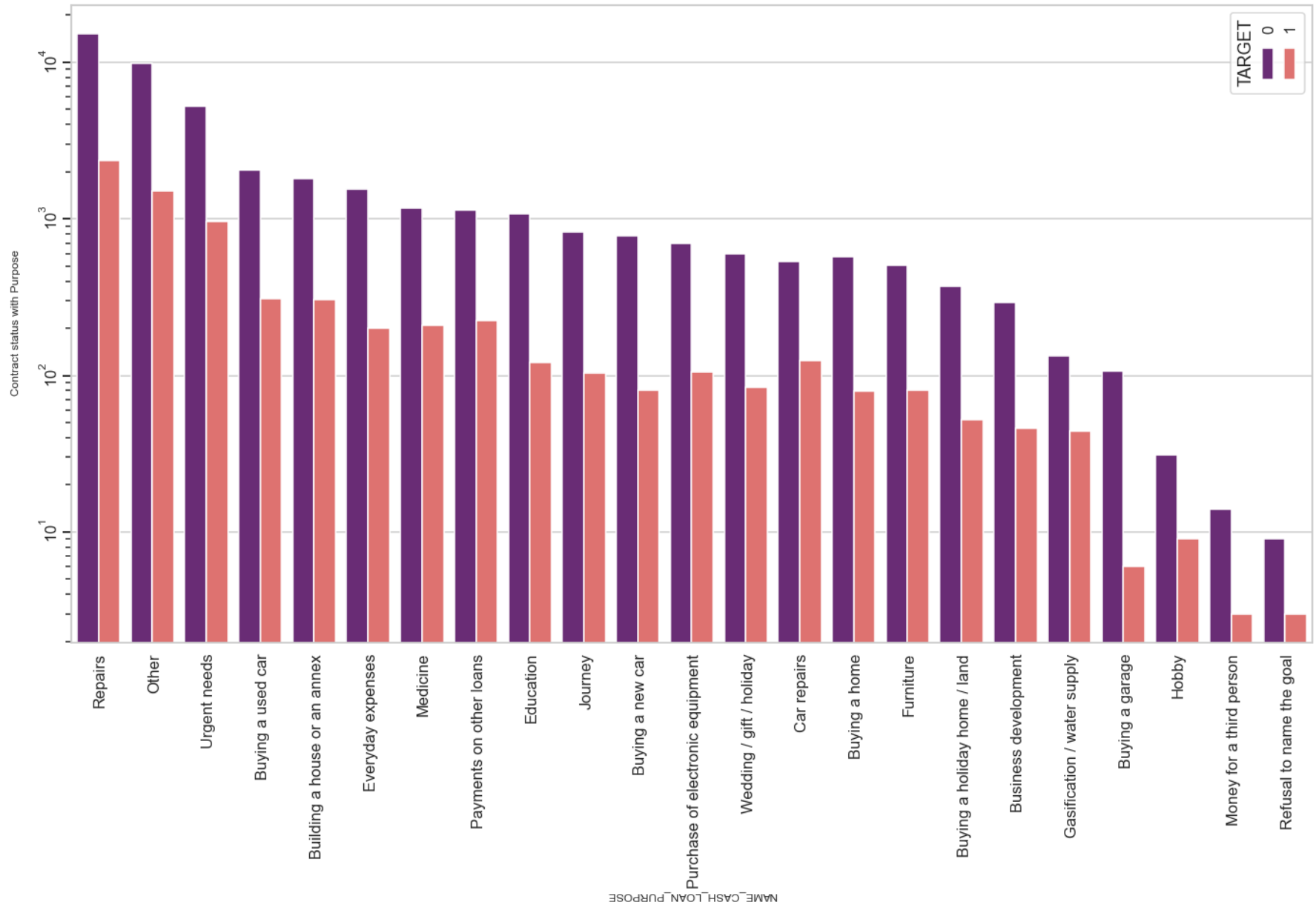
- We can observe that persons with academic degree are having more income than any other segment but unknown family status with lower education is having the highest income among all.
- Lower secondary education are comparatively less income than any other sources.
- Usually separated persons are having more total income than any other status.

Loan Status Vs Loan Purpose

- ▶ The greatest number of rejection and approval of loan is for repair purpose.
- ▶ Those person who refused to name the goal have least number of approval of loan and a greater number of rejections associated.
- ▶ In many cases, rejection is more than approved case except for everyday expense loan, education loan, purchase of electronic equipment loan, and money for third person

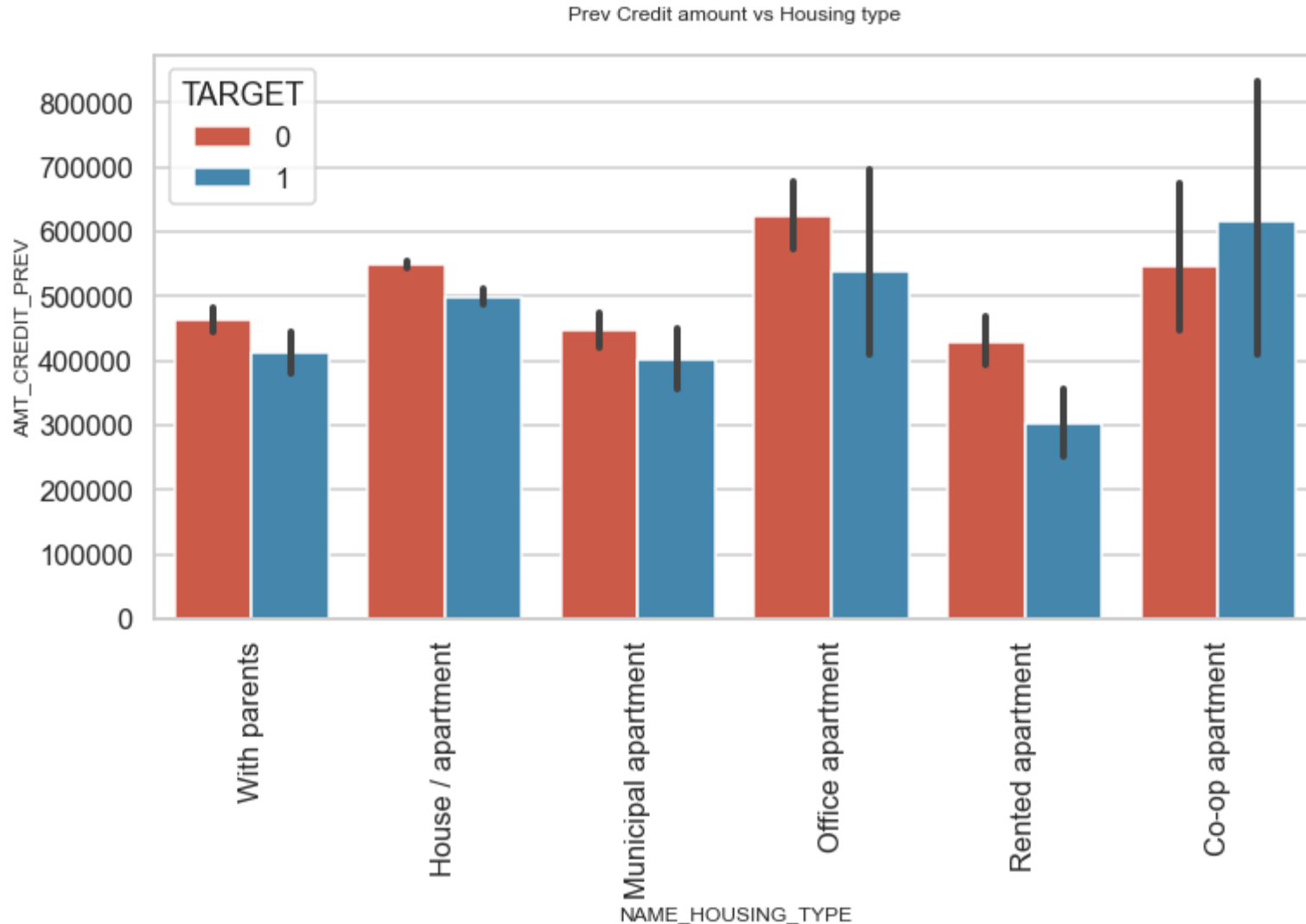


Target Vs Loan Purpose



- ▶ Usually person taking loan for repair work, other works, urgent needs are most likely to have on defaulter list.
- ▶ Persons taking loan for "Refusal to name the goal", "money for third person", "Hobby" are less likely to be on defaulter list. It may be inferred that these persons are getting less approval for loan so less likely to get on defaulter list.

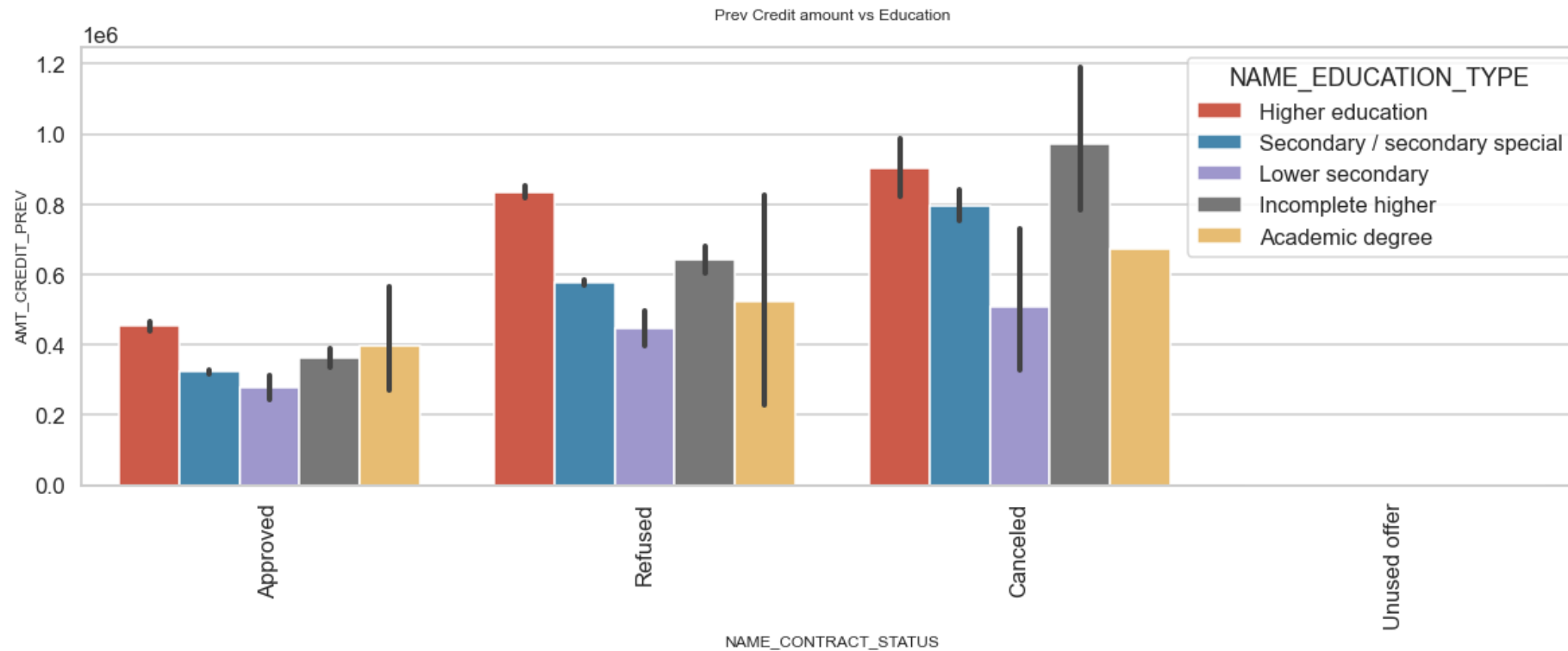
Target Vs Previous Credit Amount



- ▶ Persons staying in Office apartment are having highest credit. It may be inferred that, persons in office apartment might be getting salary in time are most likely to be non-defaulters than any other group.
- ▶ Bank should avoid giving loans to co-op apartment as they are the persons who are most likely to get on to defaulters list.
- ▶ Bank can focus on other type of housing like:
 - ▶ Staying with parents
 - ▶ House/apartment
 - ▶ Municipal apartment
 - ▶ Rented apartment

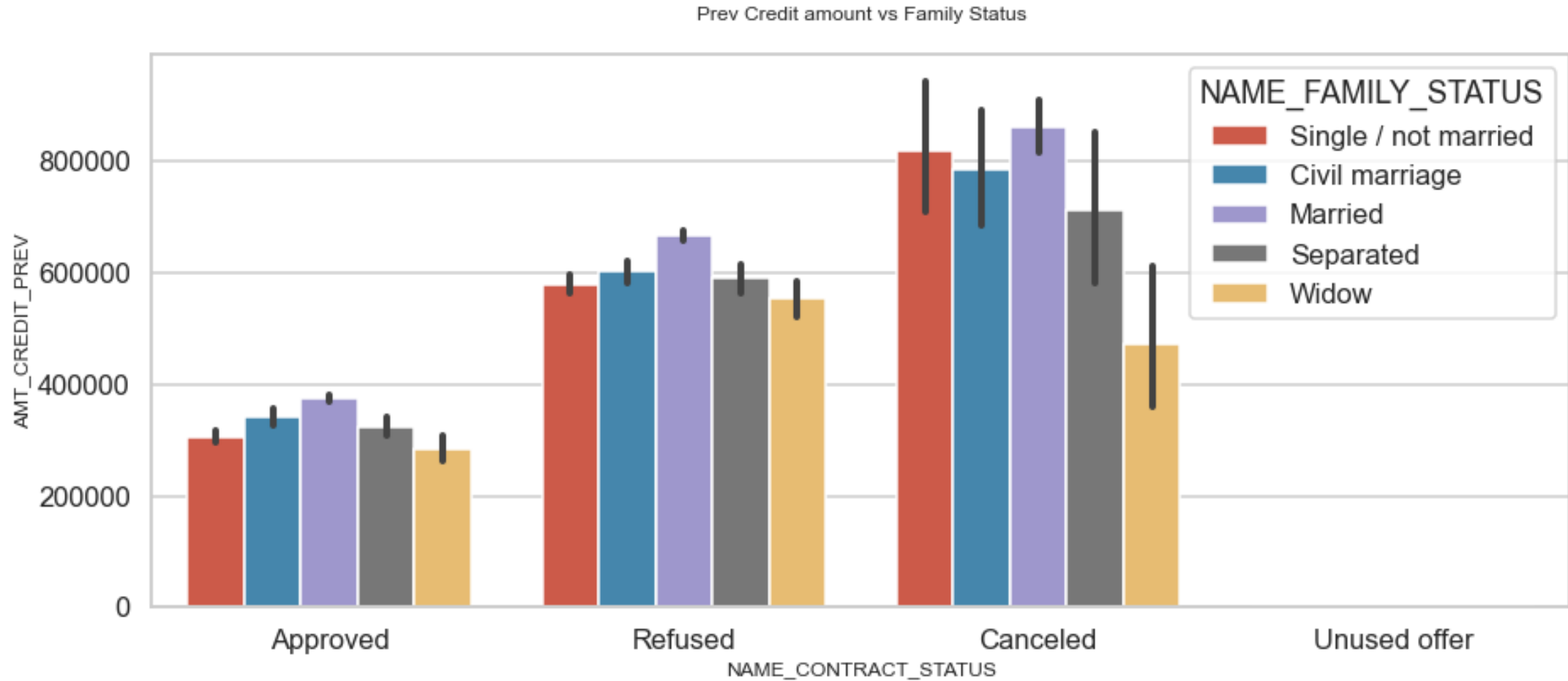
in order to get successful payment

Education Vs Previous Credit Amount



- ▶ Mostly client with higher educations are applying for loan and it gets approved. It may be inferred that as they are applying for more loan, more number of refusals and cancellations of loan is there.
- ▶ Clients with Academic degree are less likely to get refused or cancelled.

Family Status Vs Previous Credit Amount



- ▶ Mostly married clients are safe to credit loan.

CONCLUSION

- Client who has higher education are safe to credit loans if the previous application was approved.
- Bank can focus more on 'Higher education' to provide loan and avoid lower secondary education as they face with payment difficulties.
- Bank should also focus on the clients from age group 40-70 as they will have financial stability and will not be on defaulters list.
- Bank should focus on the following group to credit loan as they are less likely to be defaulters:
 - Old people of any income group
 - Client whose previous loan was approved
 - Client with high income
 - Staying with parents, house/apartments, municipal apartments
 - Clients who are taking loan for repair work and his/her previous loan was approved
- Bank should avoid the following group to credit loan as they are more likely to be defaulters:
 - Lower secondary educated client
 - Client whose previous loan status was refused
 - Clients staying in Co-op apartment



Thank you