

EDA Bank Credit Case Study

Submitted by :

T.Karan Sai Reddy

Chitra Morwal

Problem Statement

Loan providing company wants to understand the driving factors behind loan default, i.e. the variables which are strong indicators of default.

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.

Analysis Approach

Analysis approach focuses on identifying key factors/variables that helps to find potential customers for loan & those insights can be used to move the application to “**Approved**”, “**Cancelled**”, “**Refused**”, “**Unused offer**”

Steps to carry out EDA Credit Case Study Analysis



Data
Understanding



Data Cleaning



Univariate &
Bivariate Analysis



Correlation
Analysis



Insight &
Conclusion

Data Understanding

Applicant Demographics

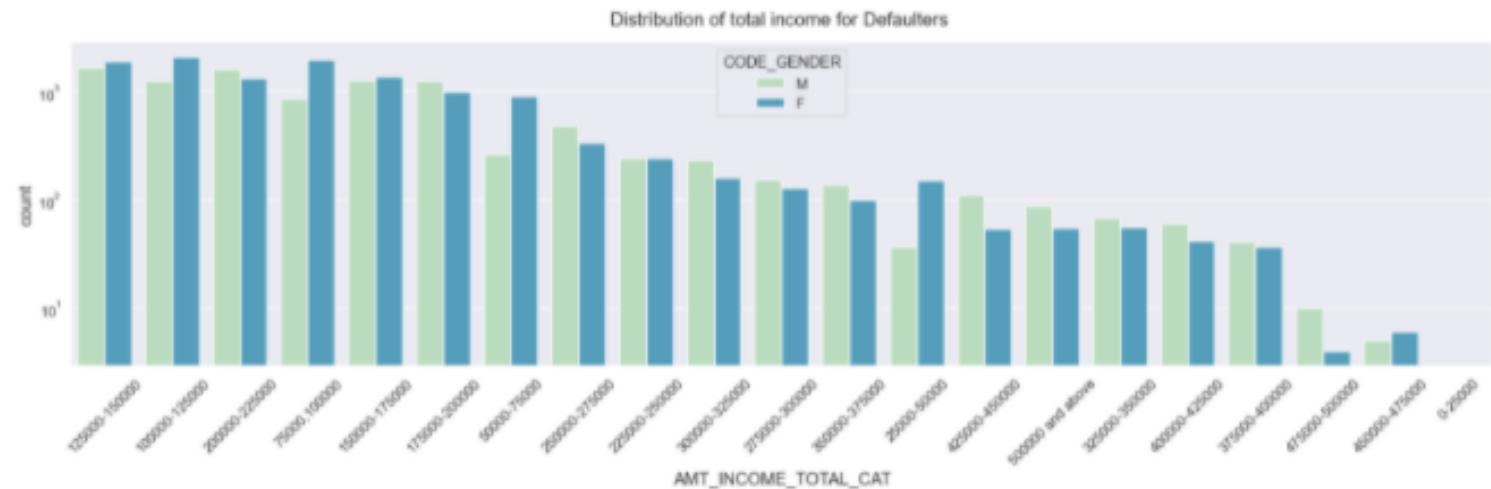
- AMT_INCOME_TOTAL
- CODE_GENDER
- NAME_FAMILY_STATUS
- NAME_EDUCATION_TYPE
- REG_CITY_NOT_LIVE_CITY
- ORGANIZATION_TYPE
- NAME_INCOME_TYPE
- OCCUPATION_TYPE

Loan related information

- AMT_CREDIT
- AMT_ANNUITY
- AMT_GOODS_PRICE
- NAME_CONTRACT_TYPE

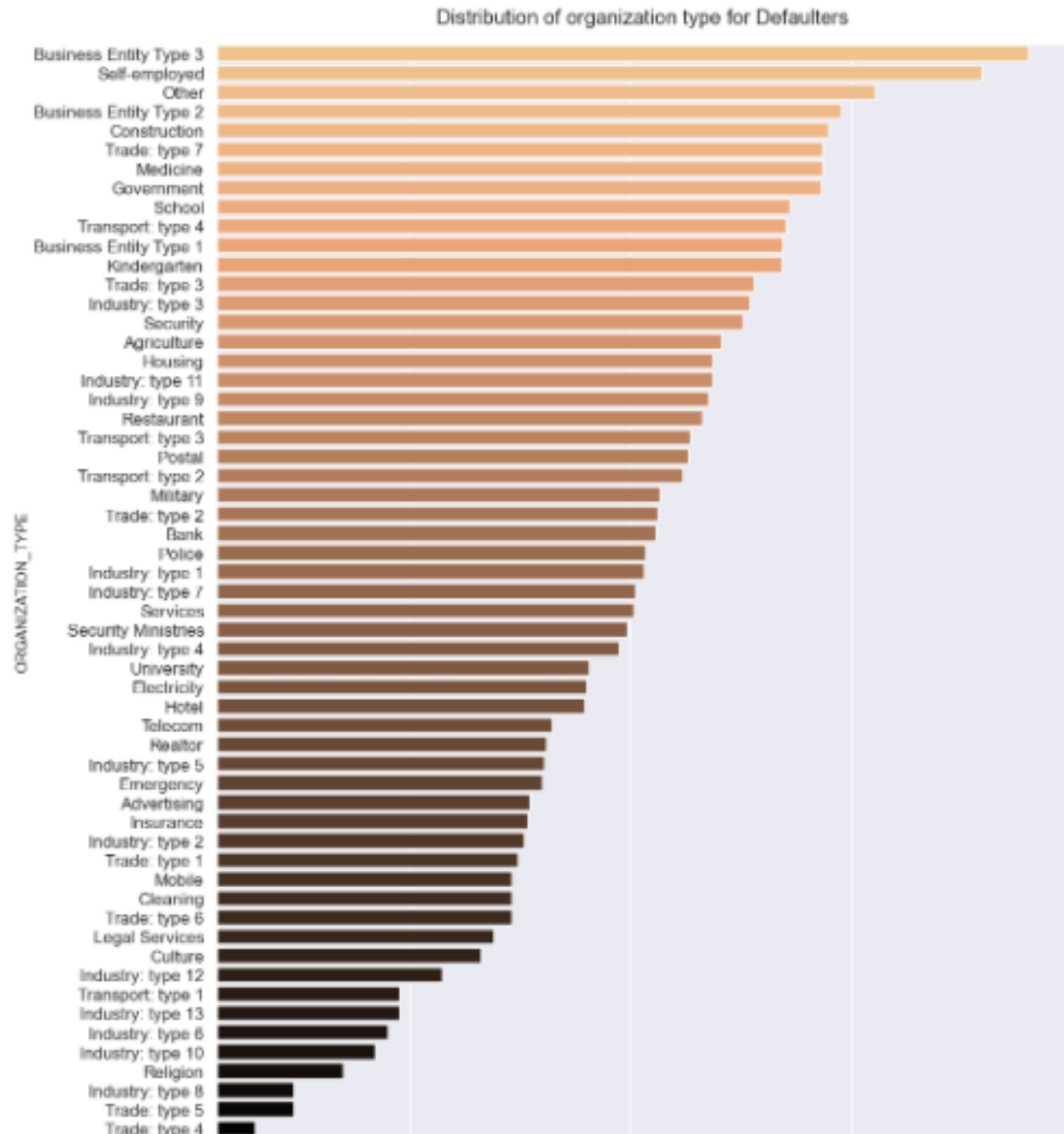
Income distribution for Defaulters & Non-Defaulters:

1. **For Non-Defaulters**, Female counts are higher than male.
2. Income range from 75000 to 225000 is having a greater number of credits.
3. Females have more credits for above range than males
4. Highest income group of males and females are 200000-225000 & 125000-150000 respectively
5. **For Defaulters**, Female counts are nearly equal for most of the categories as compared to male.
6. Income range from 75000 to 125000 is having a greater number of credits.
7. Lowest income ranges for males and females are 450000-475000 and 4750000-500000 respectively.



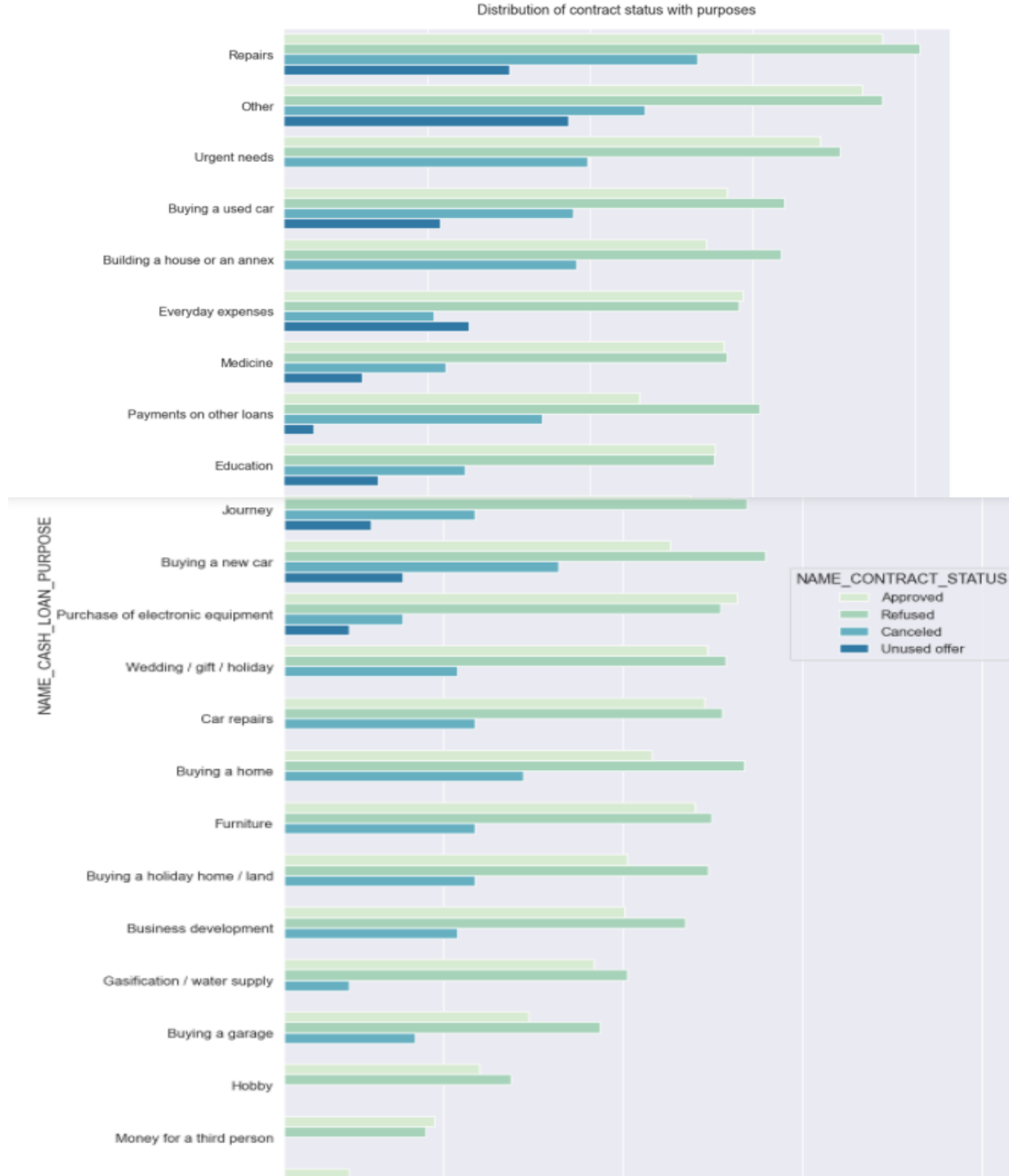
Top-5 Organizations that have payment difficulties:

Business Entity Type 3, Self-employed, Others ,business Entity Type 2,and Construction are the categories which have applied for the credits higher than any others and have payment difficulties.



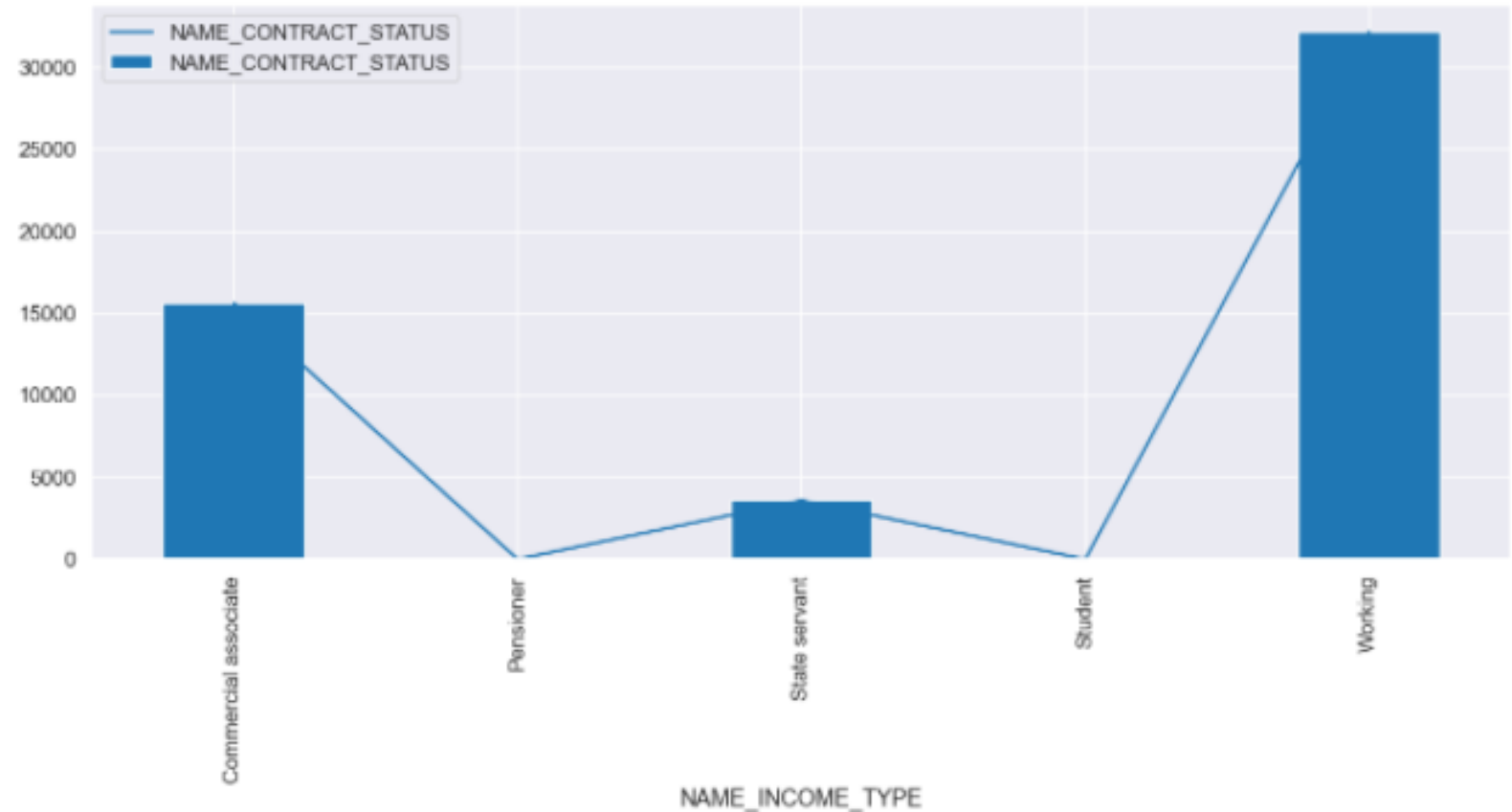
Loan Purpose Distribution:

1. Around 10000 requests for “Repair” type loans and “Payment on other loans” have been refused & holds highest rejections.
2. 100+ requests for “Buying a home” and “Buying a home / land” have been approved.
3. “Education” loan requests are balanced because they have equal number of approvals & rejections.



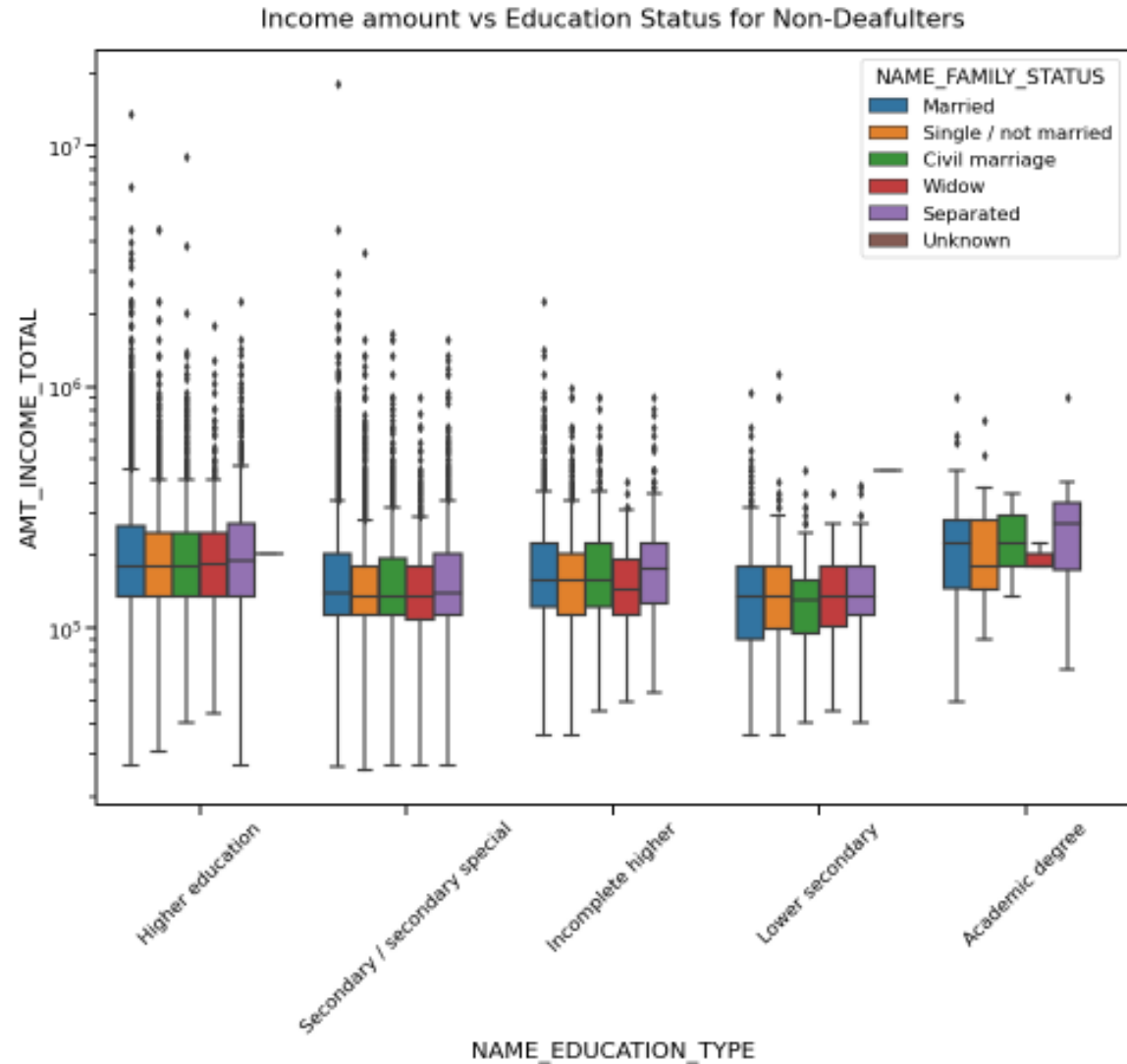
Income Based Loan Approvals:

1. Based on the income type , loans contract status shows that mostly status are approved for working, commercial associate and state servant
2. "Student" and "Pensioner" loans are less likely to be approved as there is a possibility of difficulty in repayment.



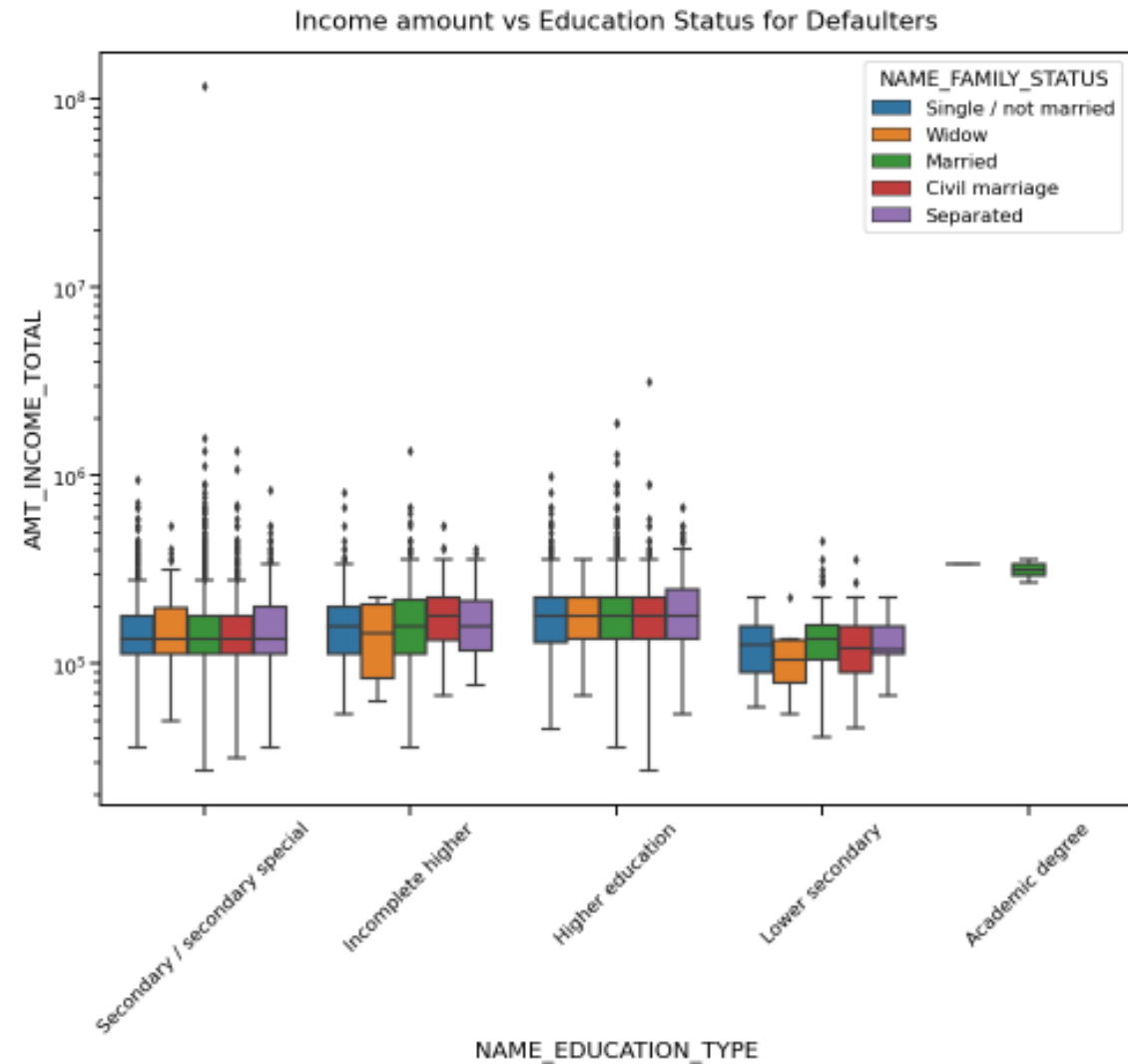
Income amount VS Education Status for Non-Defaulters

- The income amount of 'Higher education' is almost equal with family status and has many outliers.
- Least number of outliers are present for Academic degree & the income amount is higher than that of Higher education.
- When compared to others Lower secondary of civil marriage status have less income amount.

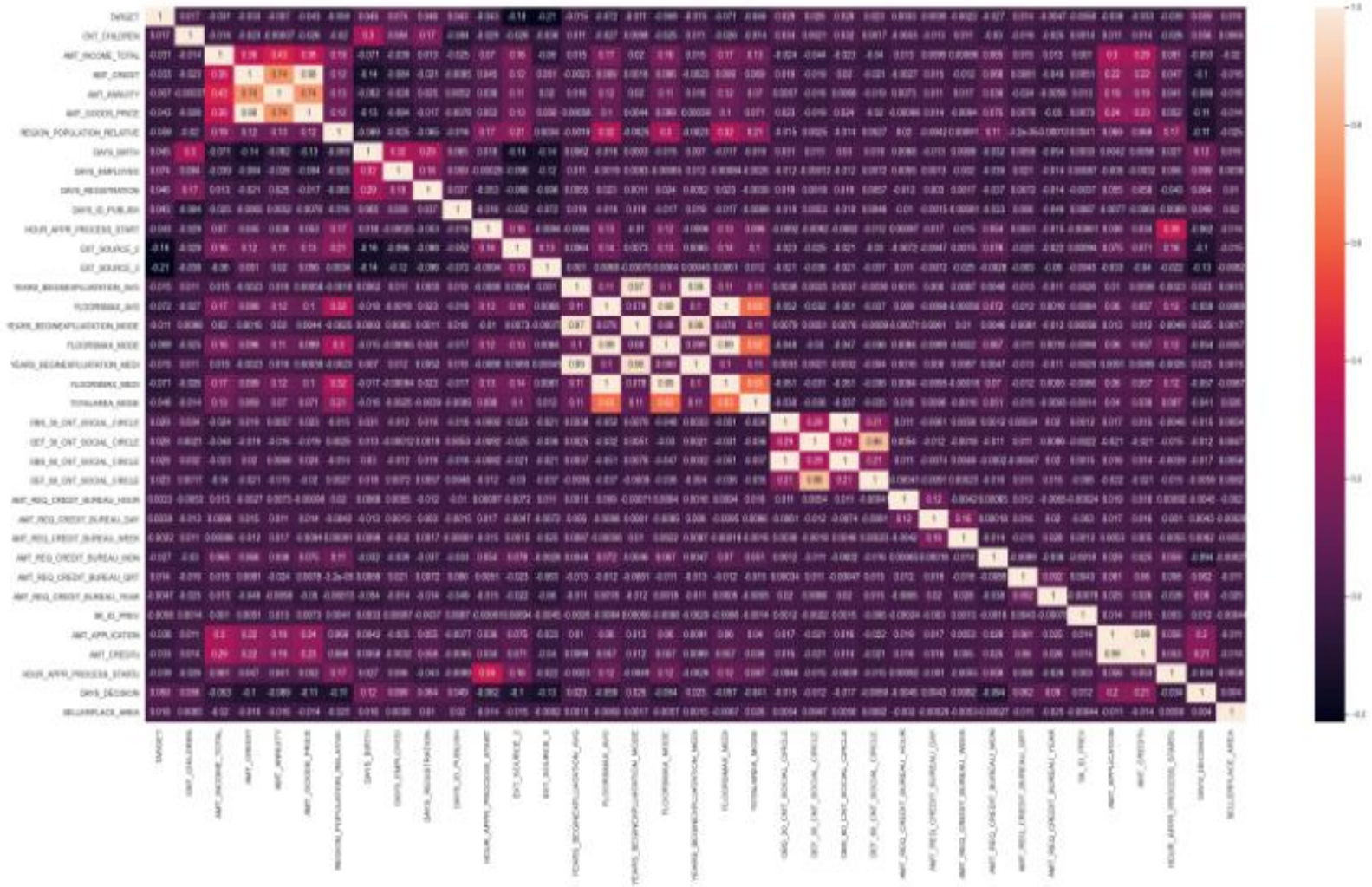


Income amount VS Education Status for Defaulters

- Defaulters plot has some similarity with Non-Defaulters.
- The income amount for 'Higher education' type & family status is almost equal.
- The income amount is higher for Academic degree than Higher education, but there are minimum number of outliers.
- Lower secondary have less income amount than others.
- More number of outliers are present for Married status people who are secondary/ secondary special qualified.



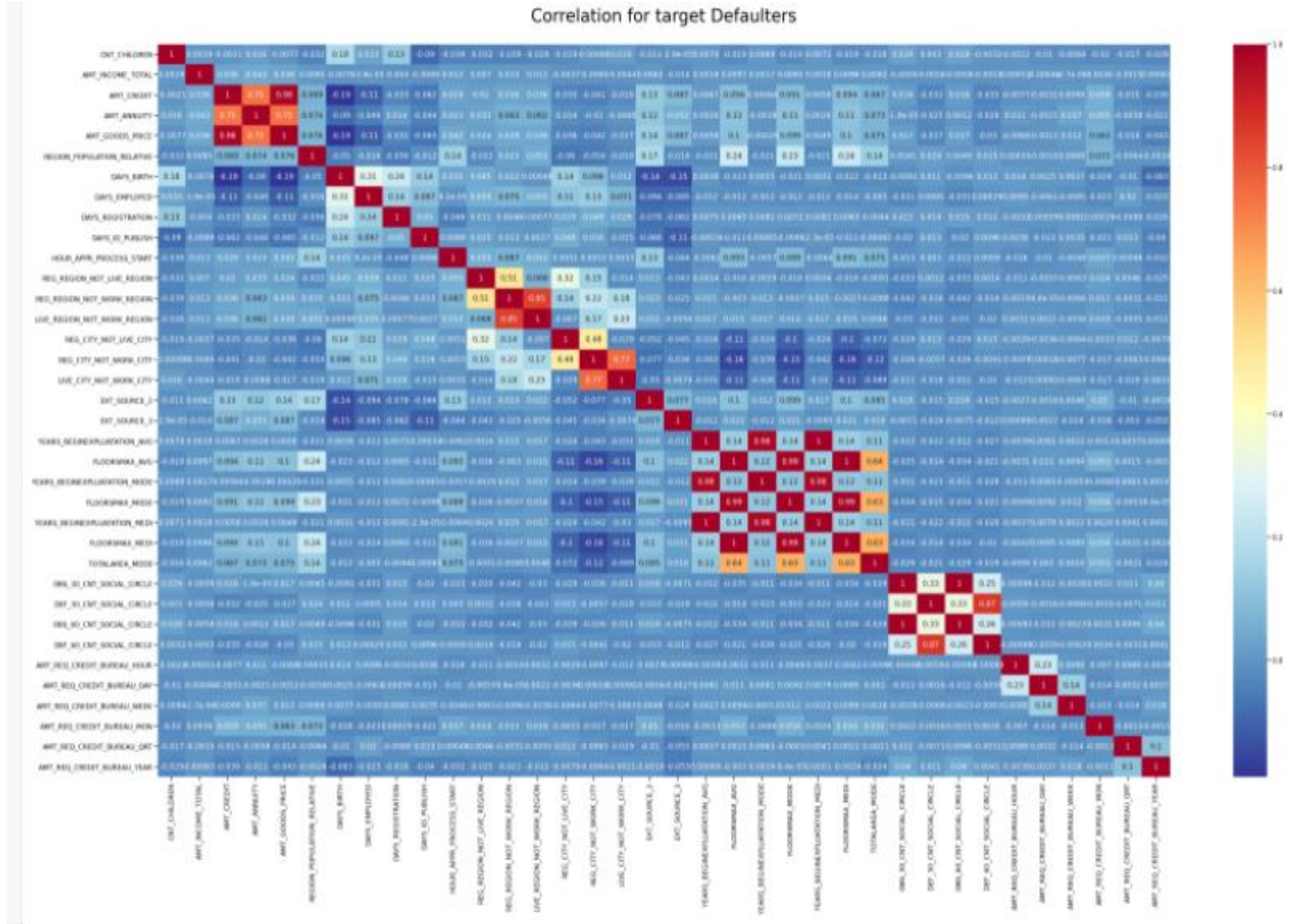
Correlation between variables leading to Loan Defaulters & Non-Defaulters:



1. From the above correlation we can say that Final credit amount on the previous application (AMT_CREDITo) has a high correlation with AMT_APPLICATION, for the amount which client initially applied.
2. For consumer loans, higher the price of the goods given in exchange for loan, higher is the loan credited given to the customer.
3. Higher the Age of the customer, less is the credit amount given for loan.
4. Amount credited is less if Number of enquiries to Credit Bureau about the client one day is very high.
5. Higher the count of children, less is the amount credited for loan.
6. Higher the count of children, less is the total income.

Loan Approvals dependency on client Demographics:

1. This heat map for Defaulters is also having quite a same observation just like Non-Defaulters. But for few points are different. They are listed below.
2. Suppose clients permanent address is not same as work address, he/she is having less children & vice-versa the amount credited is less.



Conclusion:

- As per our observation, loans repayment is most likely to be done without any difficulties for women who are of following income range categories: 75000-15000 and for men who are of following income range categories: 175000-225000
- Business Entity Type3 has equal chances of being defaulters and non-defaulters.
- Working officials and commercial associates are the potential customers a bank can give loans to.
- Housing type and office apartment are having higher credit of Non Defaulter and co-op apartment is having higher credit of Defaulter. Hence, 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/apartment or municipal apartment for successful payments.