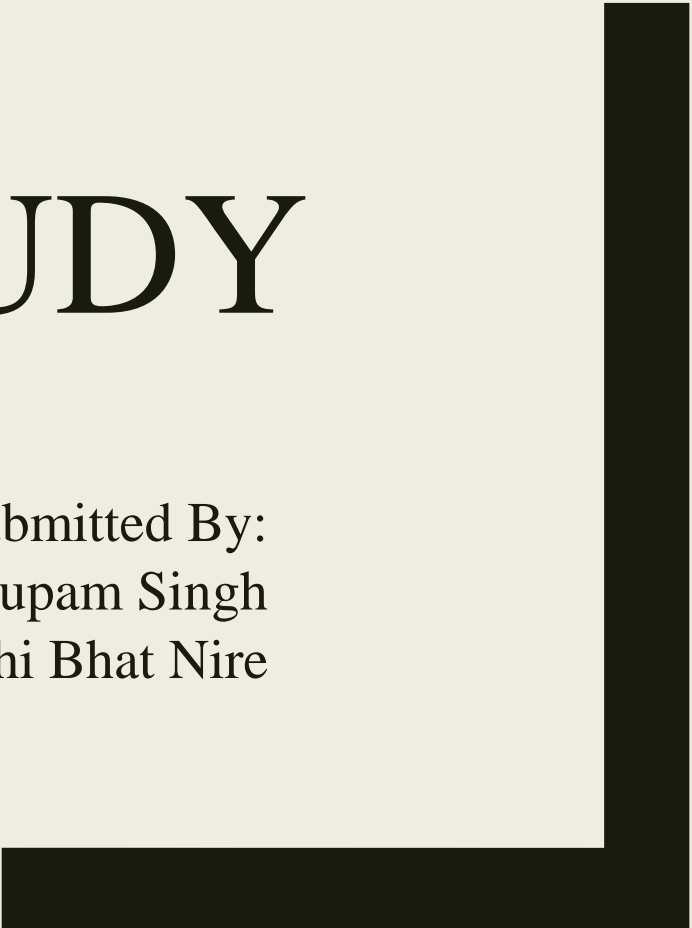




EDA CASE STUDY

Submitted By:
Anupam Singh
Srinidhi Bhat Nire



Problem Statement:

The loan providing companies want to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.

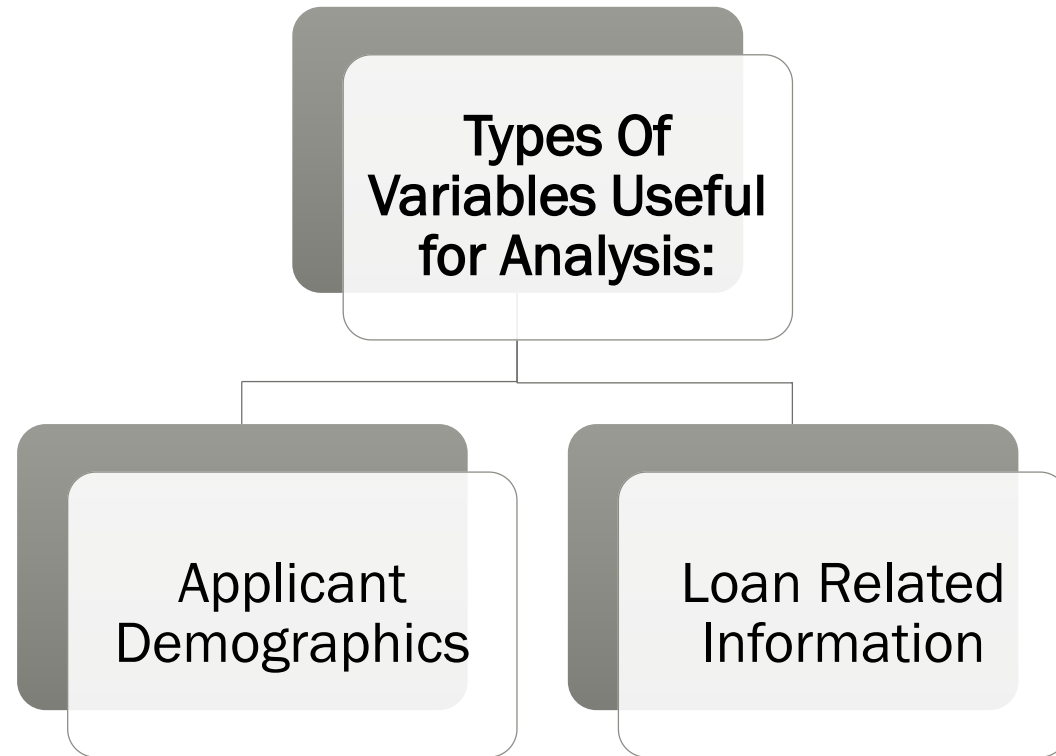
Two types of risks that are associated with the bank's decision:

- If the applicant is likely to repay the loan, then not approving the loan results is 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.

Statement of Purpose:

Primary Objective of this analysis is to identify variables which are strong indicators or drivers to find out clients potential for loan credit and customers who aren't and to use these insights in approval / rejection decision making.

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

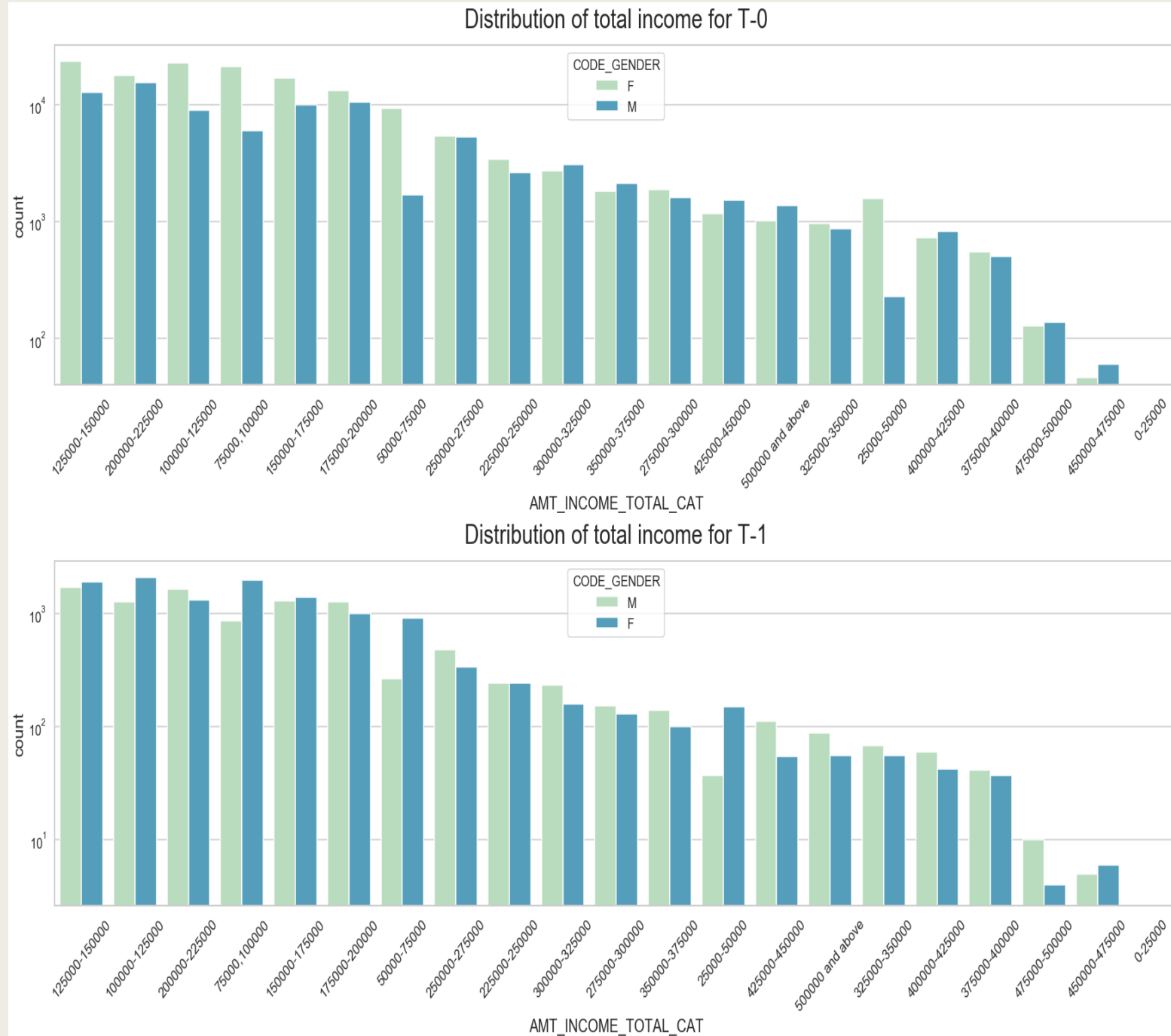
Total Income Distribution for Defaulters and Non-defaulters:

1. For T-0 type customers, Female counts are higher than male.
2. Income range from 75000 to 125000 is having more number of credits.
3. Highest income group of males and females are 200000-225000 and 125000-150000 respectively.

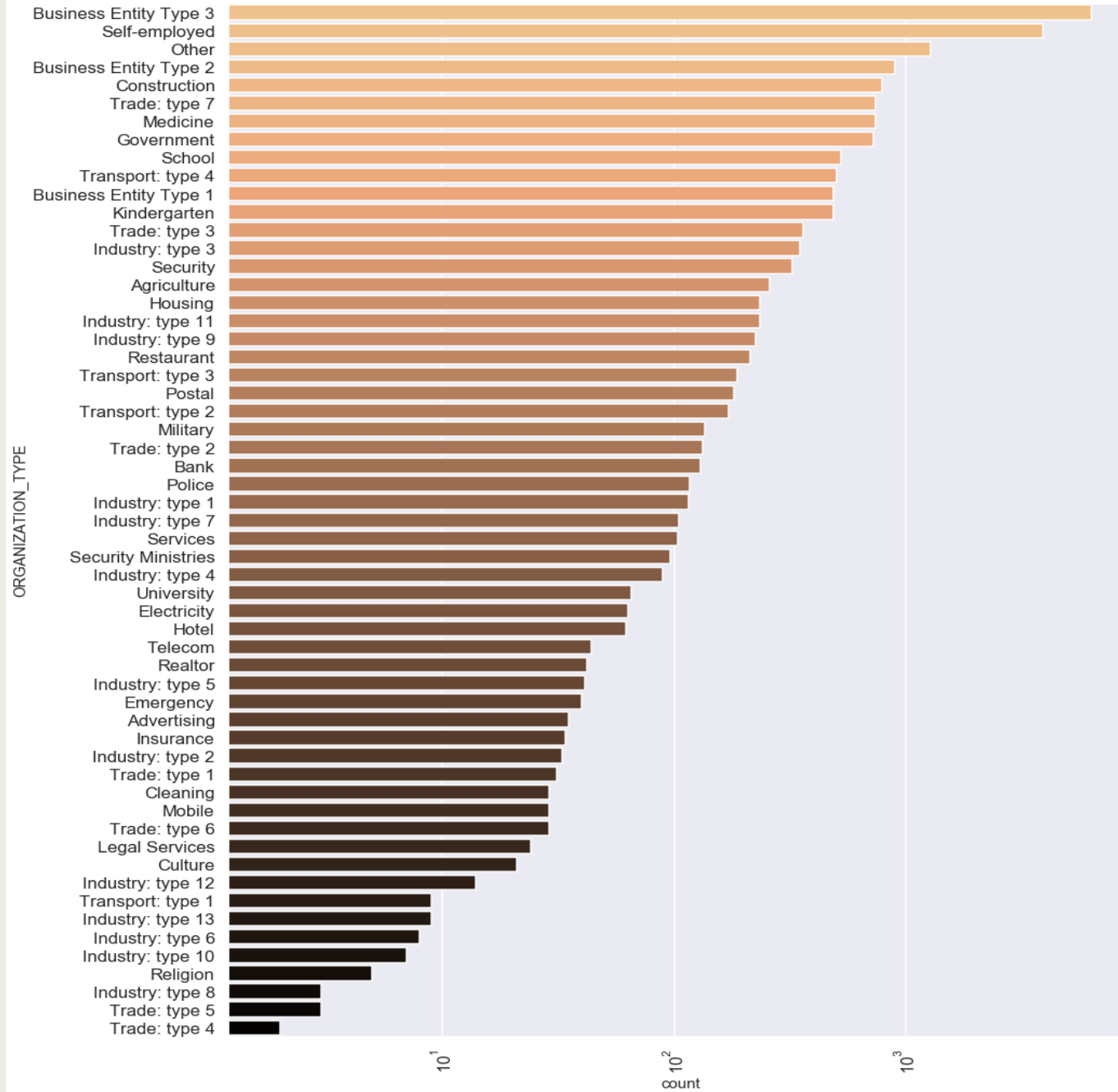
1.For T-1 type customers, Female counts are nearly equal for most of the categories as compared to male.

2.Income range from 75000 to 125000 is having more number of credits.

3.Lowest income ranges for males and females are 450000-475000 and 475000-500000 respectively.

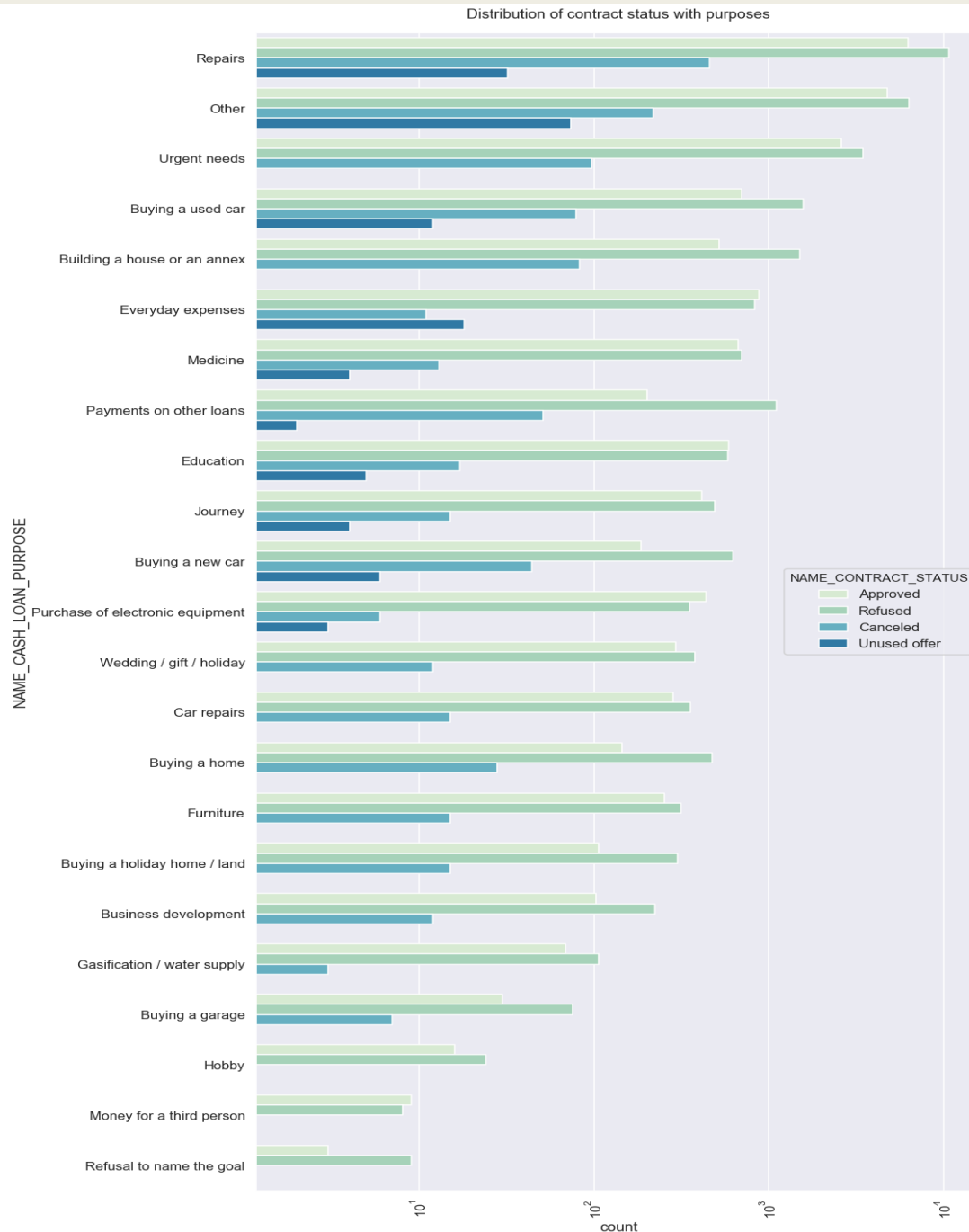


Distribution of organization type for T-1



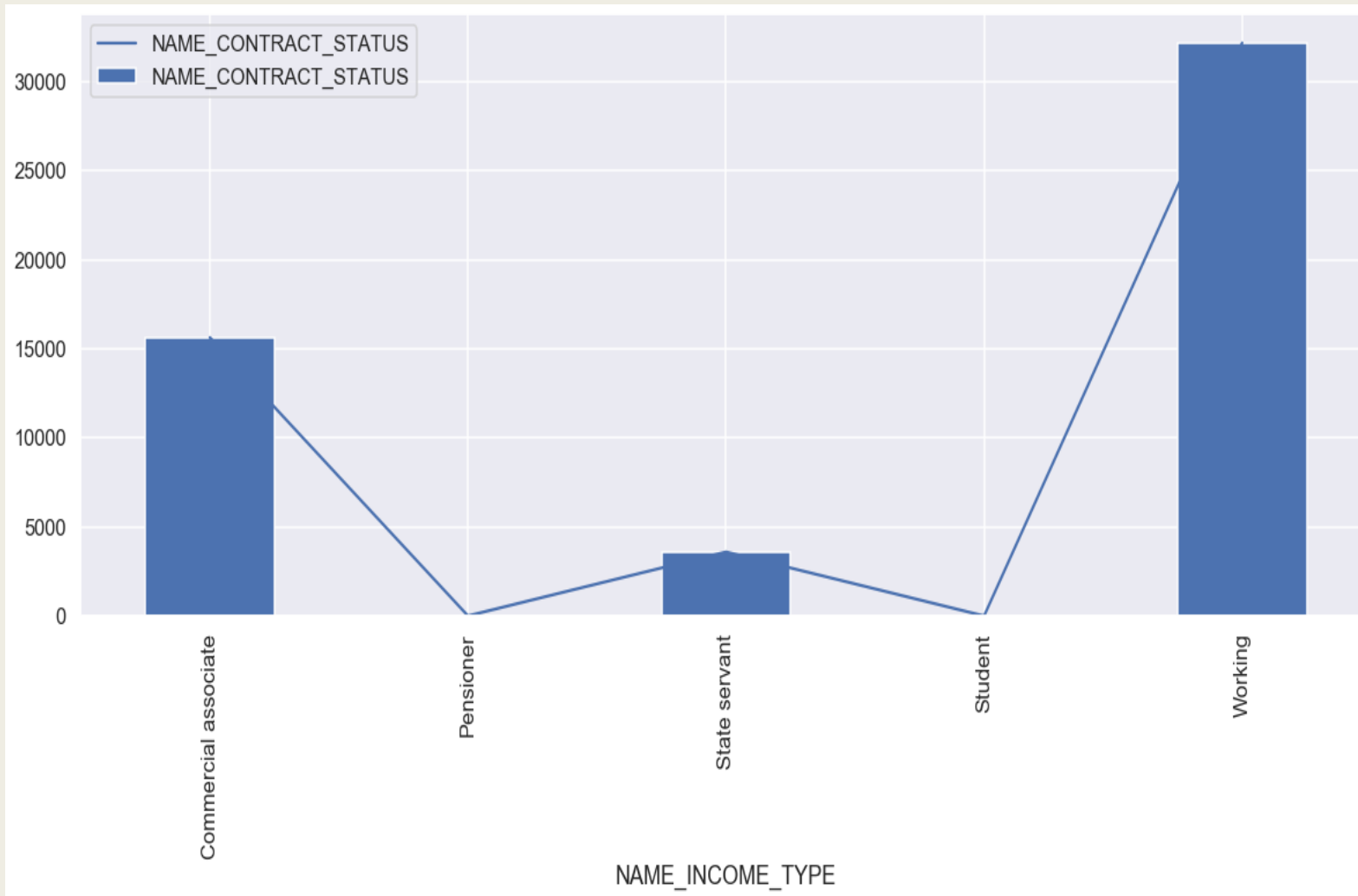
Top-5 Organizations that have payment difficulties:

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:

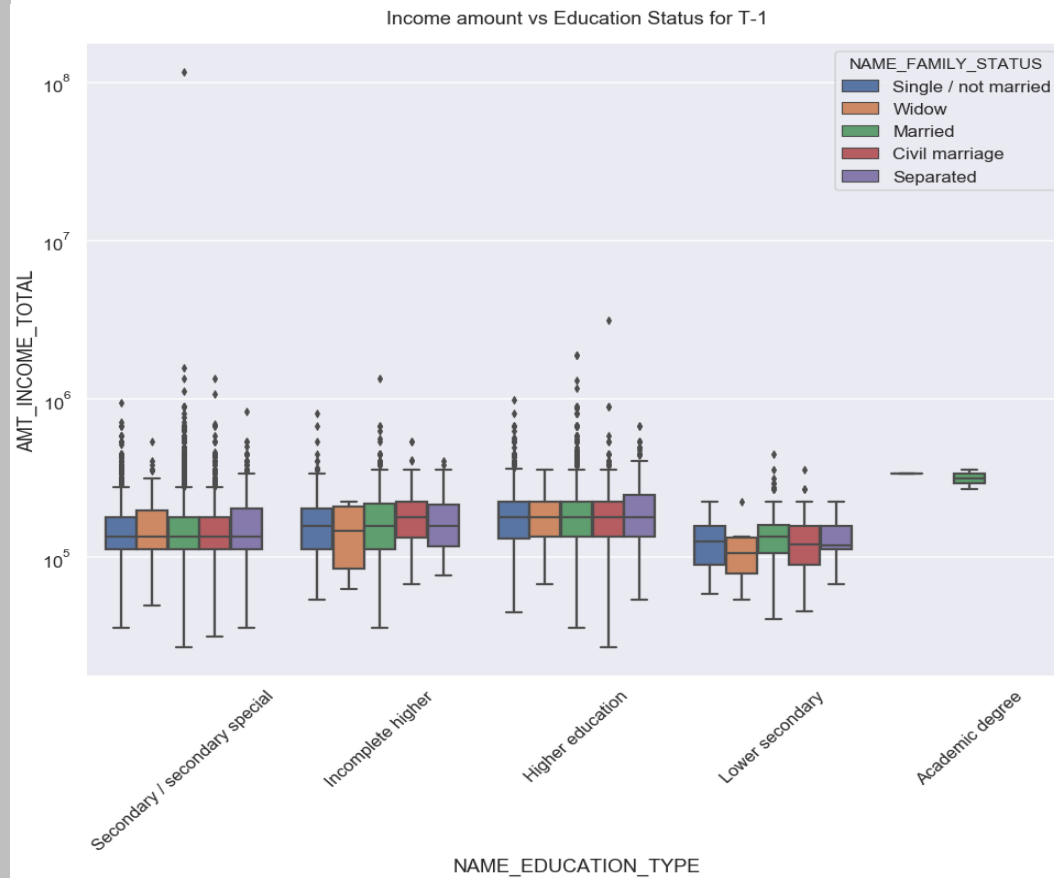
- Approximately 10000 request for “Repair” type loans and 1000 request for “Payment on other loans” have been refused.
- 100+ requests for “Buying a home” and “Buying a home / land” have been approved.
- Highest number of rejections are faced for “Repair” and “Payment of other loans” type.
- “Education” loan requests have same number of approvals as the number of rejections.



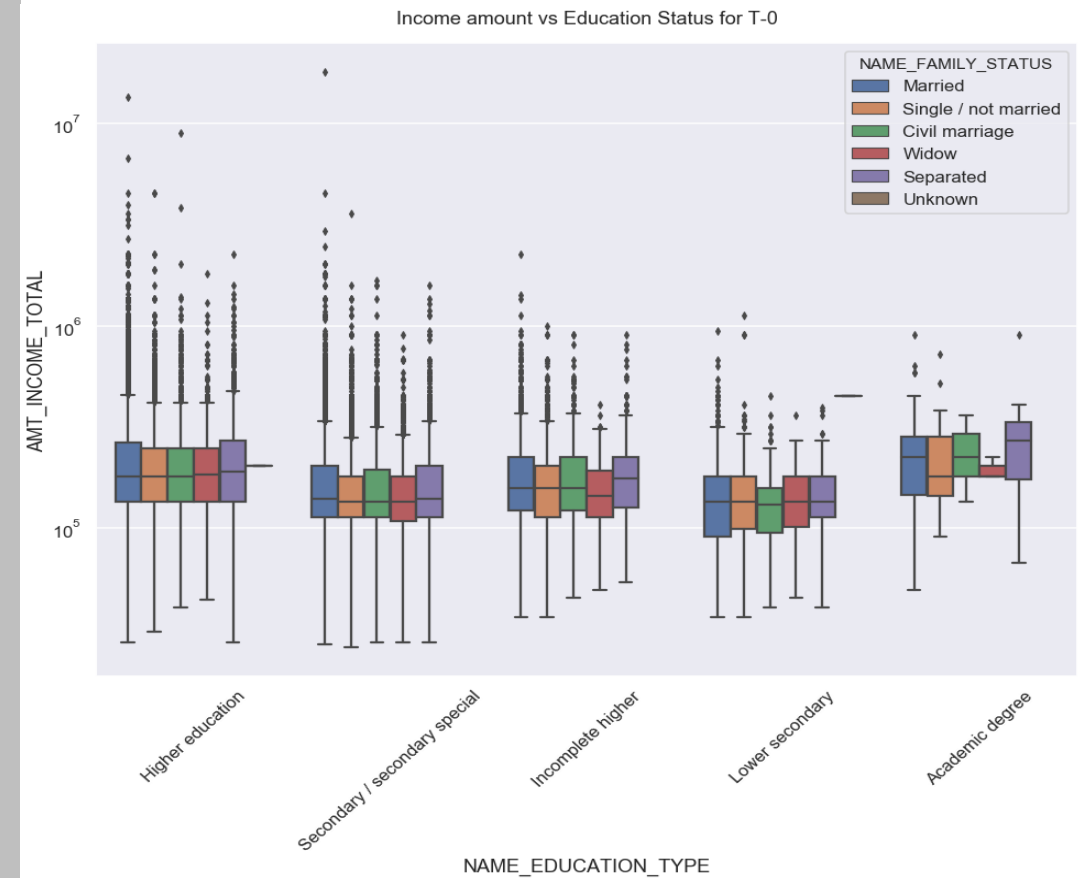
Income Based Loan Approvals:

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

Dependency of Income vs Education Status for Loan Approvals:

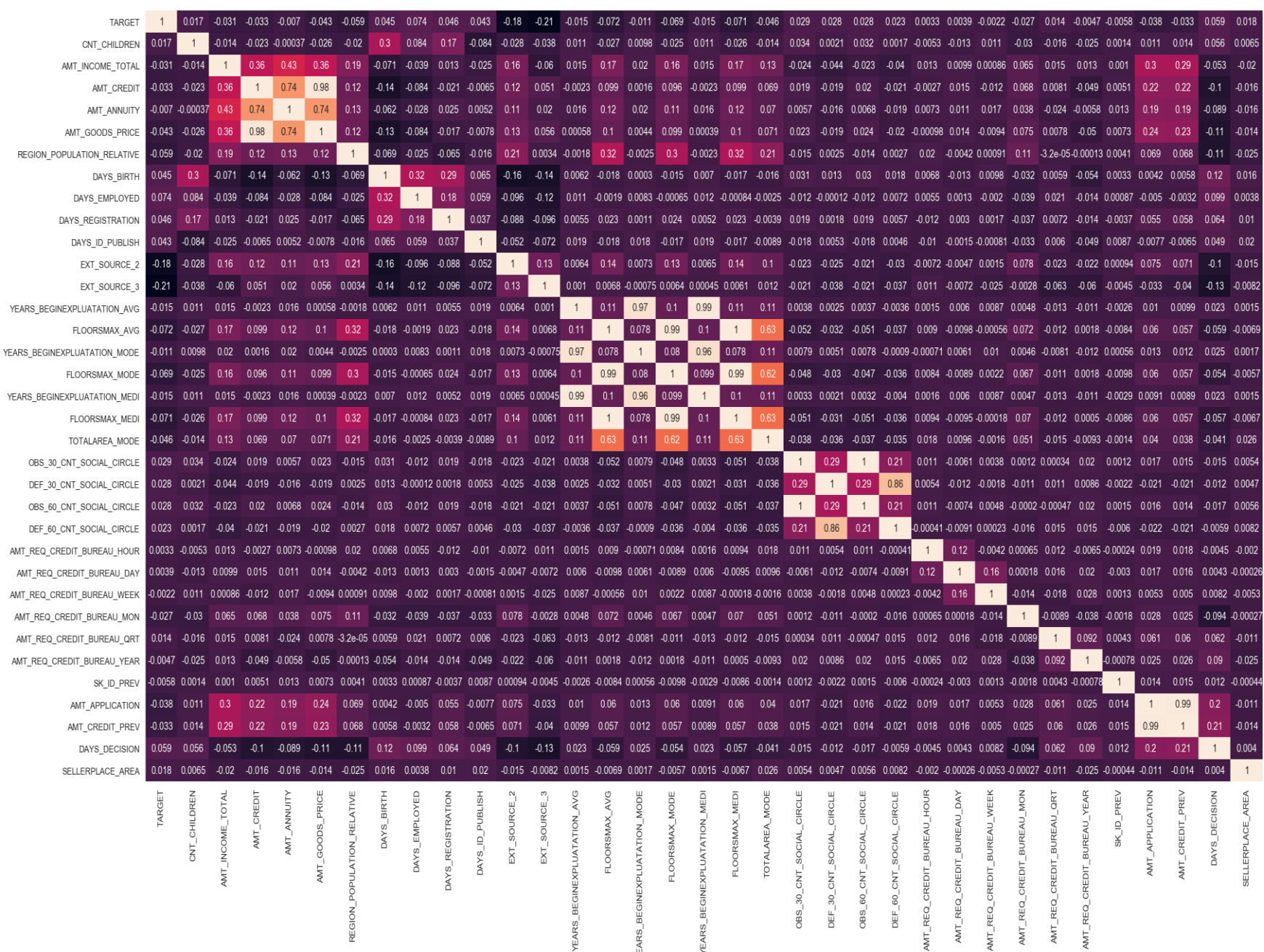


- T-1 plot has some similarity with T-0. For the 'Higher education' type, the income amount is mostly equal with family status.
- Less numbers of outliers are present for Academic degree and the income amount is higher than that of Higher education.
- Lower secondary are have less income amount than others.
- High number of outliers are present for Married status people who are secondary/ secondary special qualified.

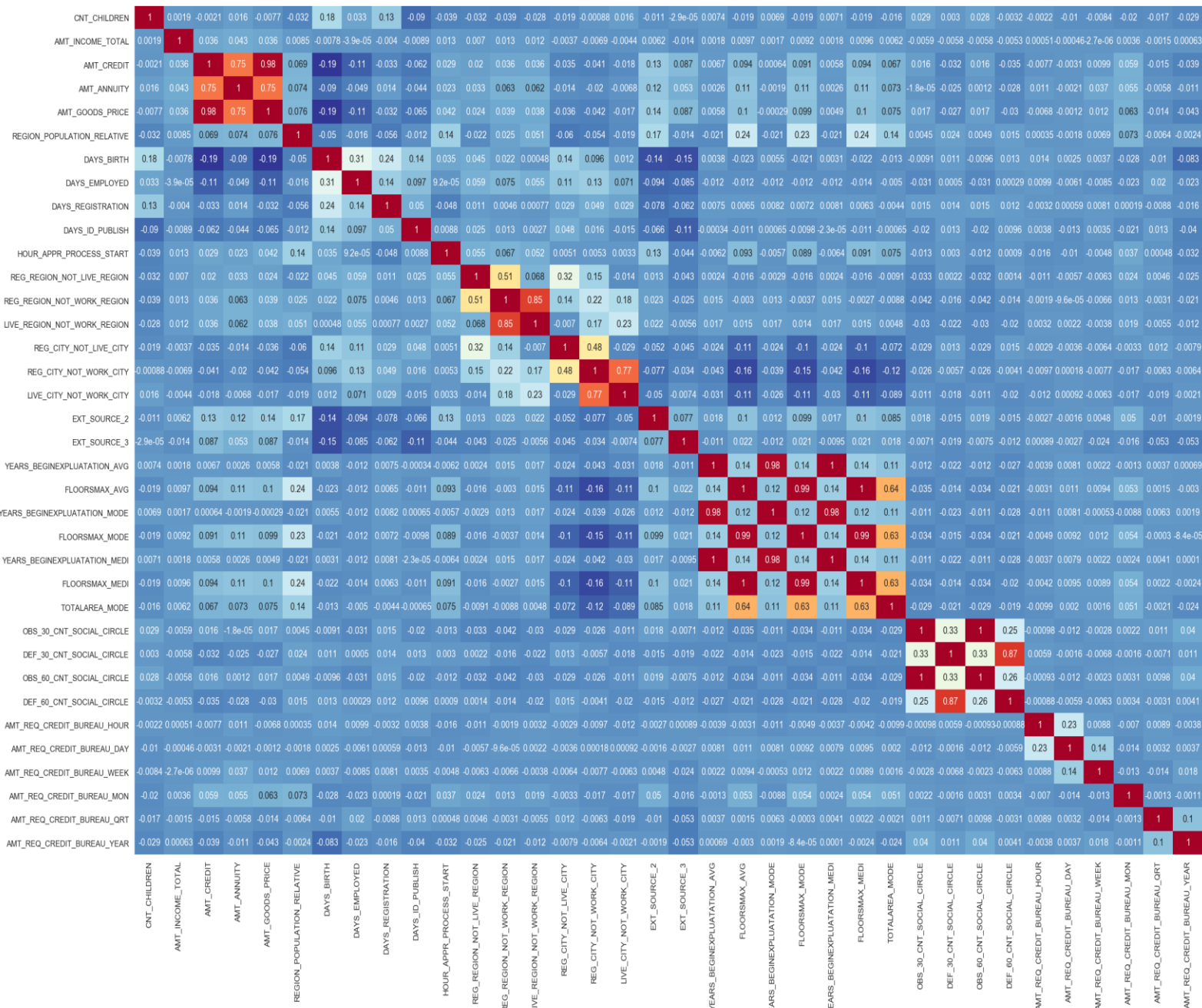


- For Education type 'Higher education' the income amount is mostly equal with family status and has many outliers.
- Less number of outliers are present for Academic degree and the income amount is higher than that of Higher education.
- Lower secondary of civil marriage status have less income amount than others.
- Lowest number of outliers are present for Academic Degree type people.

Correlation between driver and non driver variables for Loan Approvals:



Correlation for target 1



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 : 75000-150000 and for men who are of following income range categories : 175000-225000
- Business Entity Type 3 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 T-0 and co-op apartment is having higher credit of T-1. 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.