Credit based ETA

Problem statement

Loan companies face challenges in lending to people without sufficient credit history, as some may exploit this by defaulting on loans.

As an analyst at a consumer finance company, my task is to conduct Exploratory Data Analysis (EDA) on loan applicant data to identify patterns.

The company's decision to approve a loan involves assessing two risks:

Approving loans for applicants who can repay minimizes the risk of losing potential business.

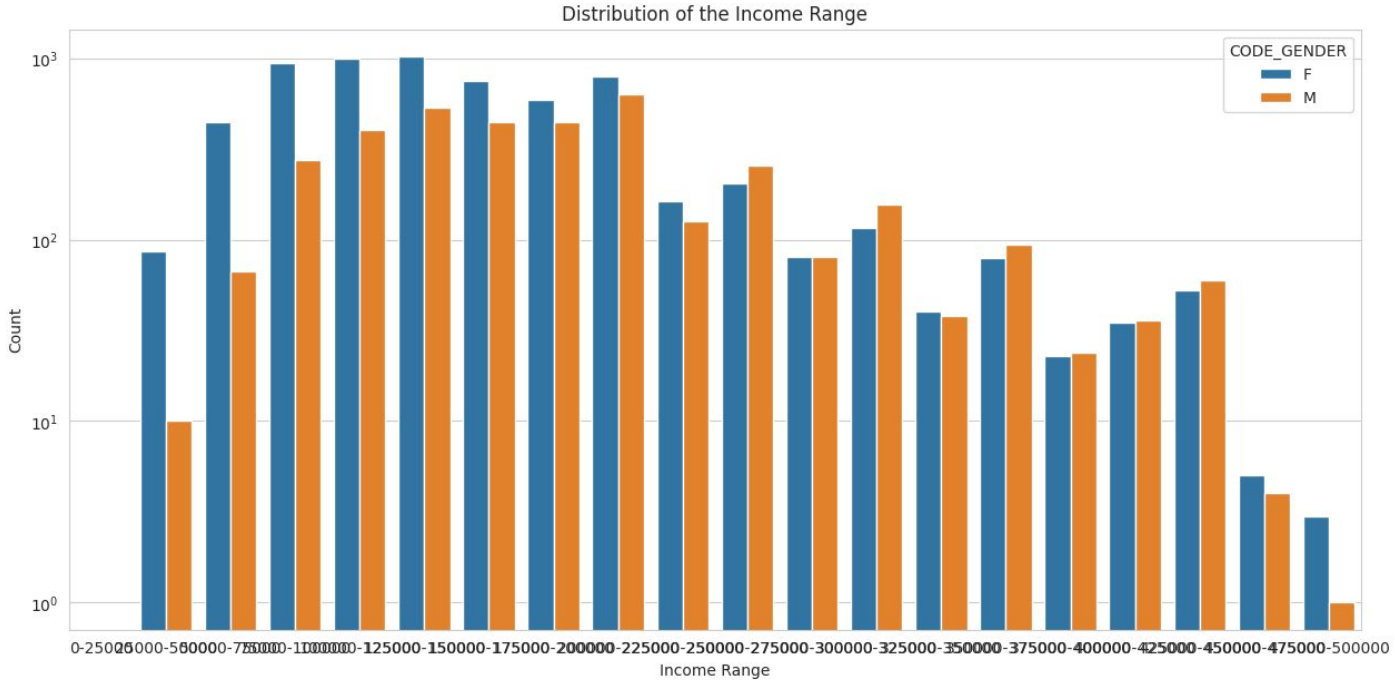
However, approving loans for applicants likely to default can result in financial losses for the company.

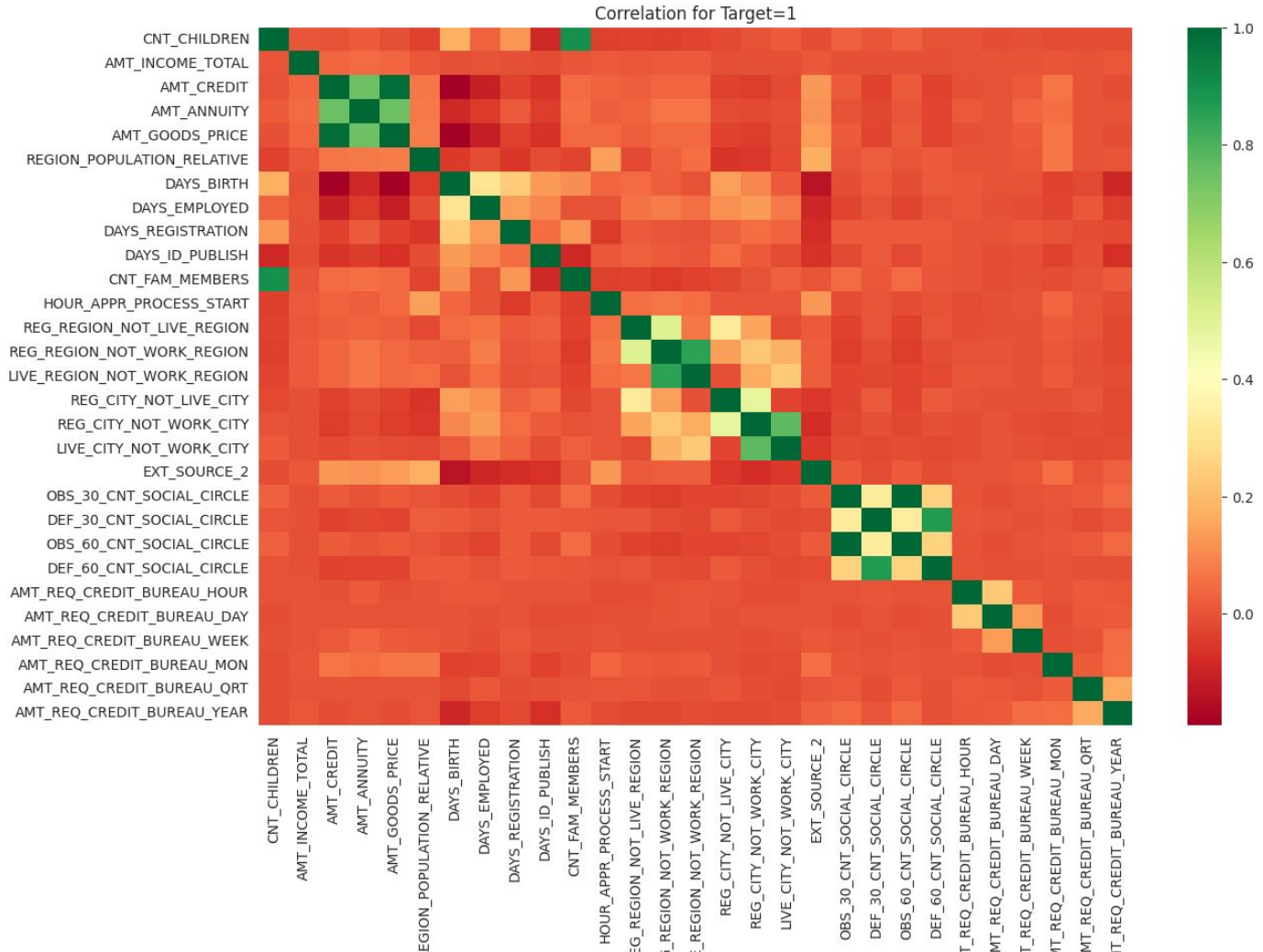
Approach

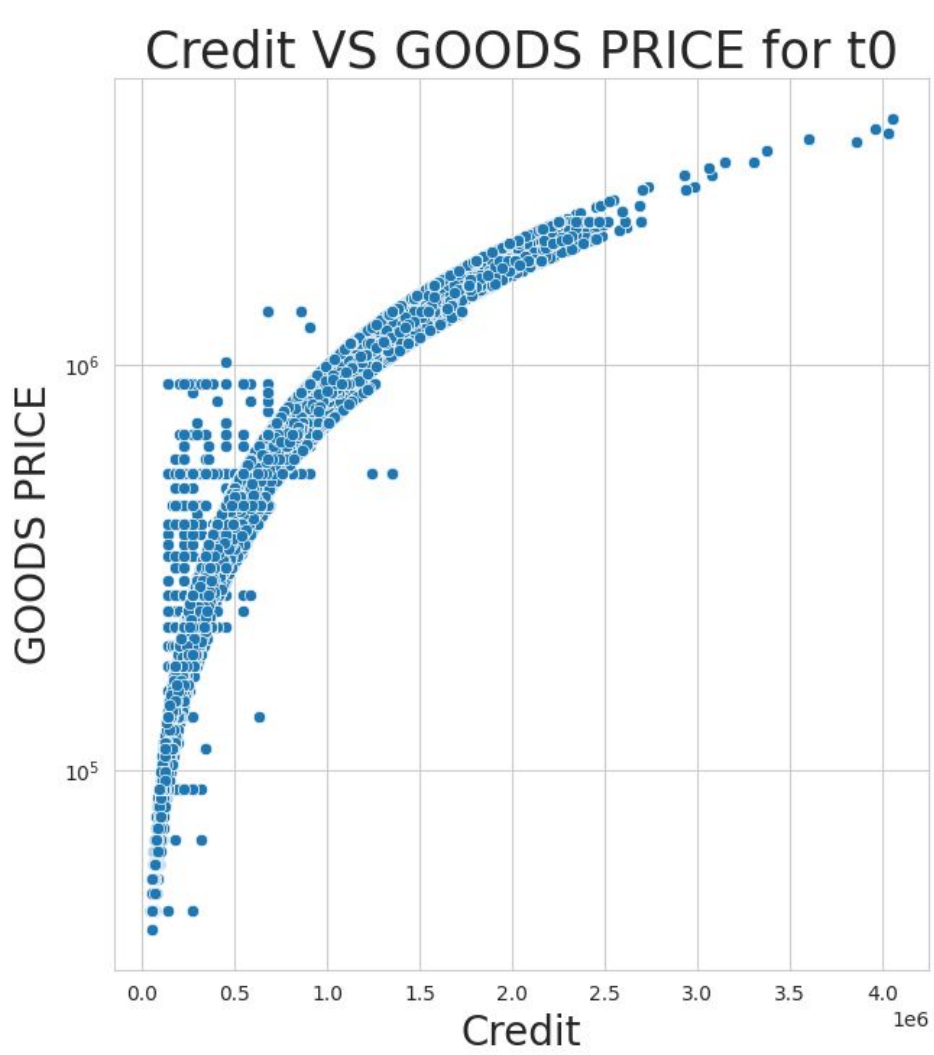
- Importing the required libraries: Started by importing necessary libraries like pandas, numpy, and matplotlib to facilitate data manipulation, analysis, and visualization.
- Data Cleaning on application dataset: Conducted data cleaning procedures such as handling missing values, removing unwanted columns, and checking for errors in gender and organization columns.
- Imputing the missing values: Employed imputation techniques to handle missing values in both the application dataset and the previous application dataset.
- Creating bins for numerical columns: Segmented numerical columns like AMT_INCOME_TOTAL and AMT_CREDIT into bins to simplify analysis and interpretation.
- Checking the imbalance ratio: Assessed the class imbalance ratio in the target variable to determine if any class is significantly underrepresented.

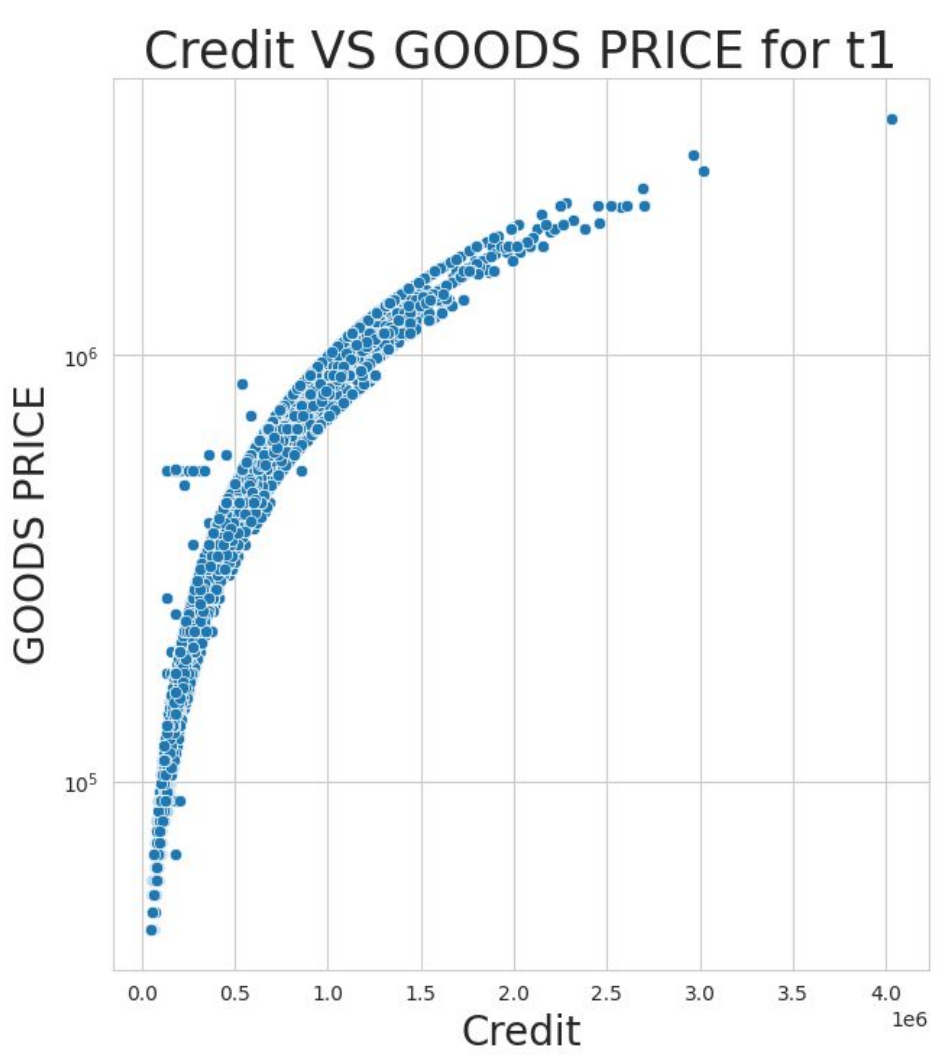
- Univariate Analysis: Performed univariate analysis by plotting charts separately for customers who did not face repayment issues (Target = 0) and those who did (Target = 1).
- Defining correlation: Calculated correlation coefficients to identify relationships between variables and the target variable.
- Finding outliers: Detected outliers in numerical columns using statistical methods or visualization techniques to ensure data quality.
- Multivariate Analysis: Conducted multivariate analysis to explore relationships between multiple variables and the target variable, considering interactions and dependencies.
- Work on previous_application dataset: Extended the analysis to the previous application dataset by merging it with the application dataset and conducting both univariate and bivariate analyses to gain insights into historical loan applications.

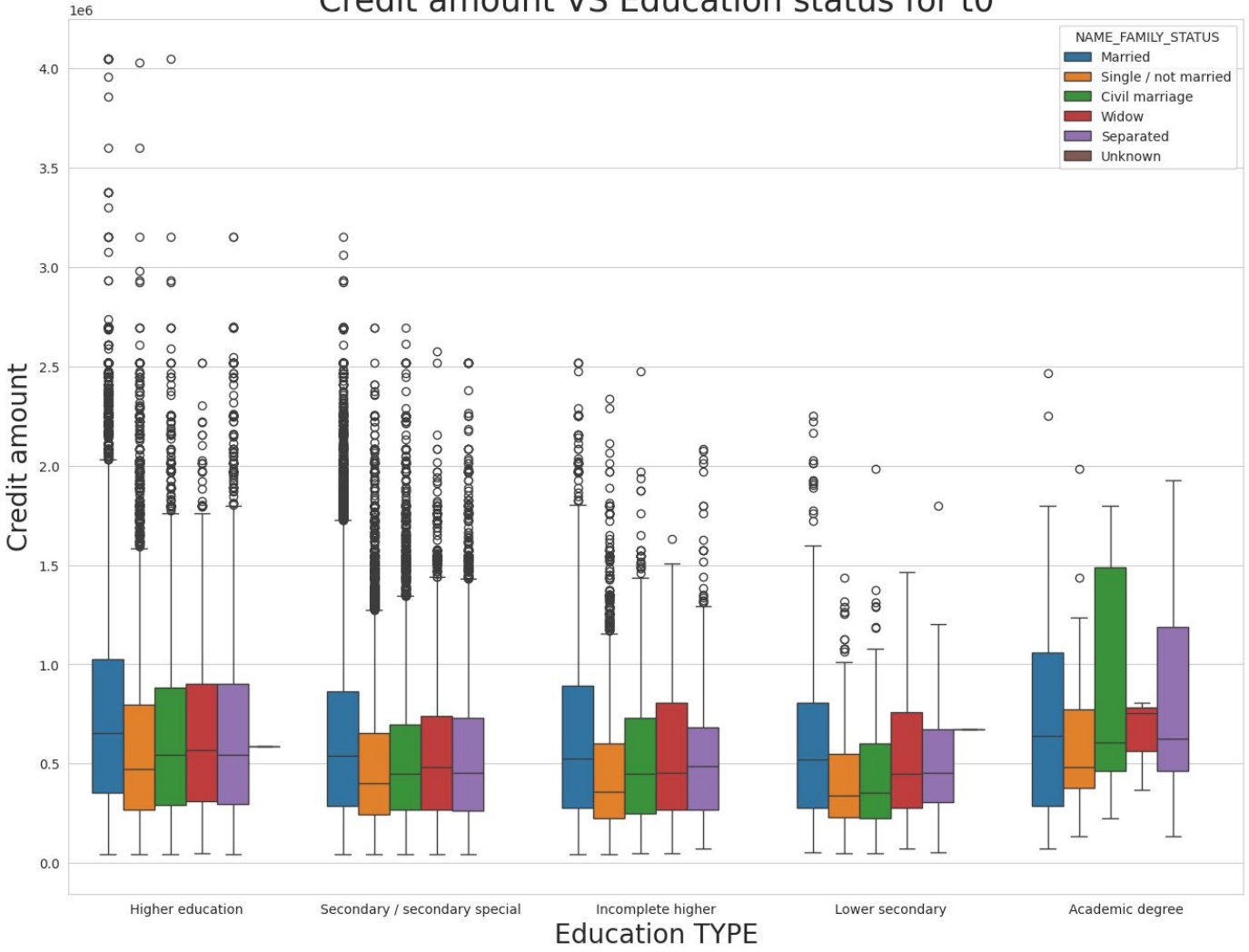
Graphs used in the representation

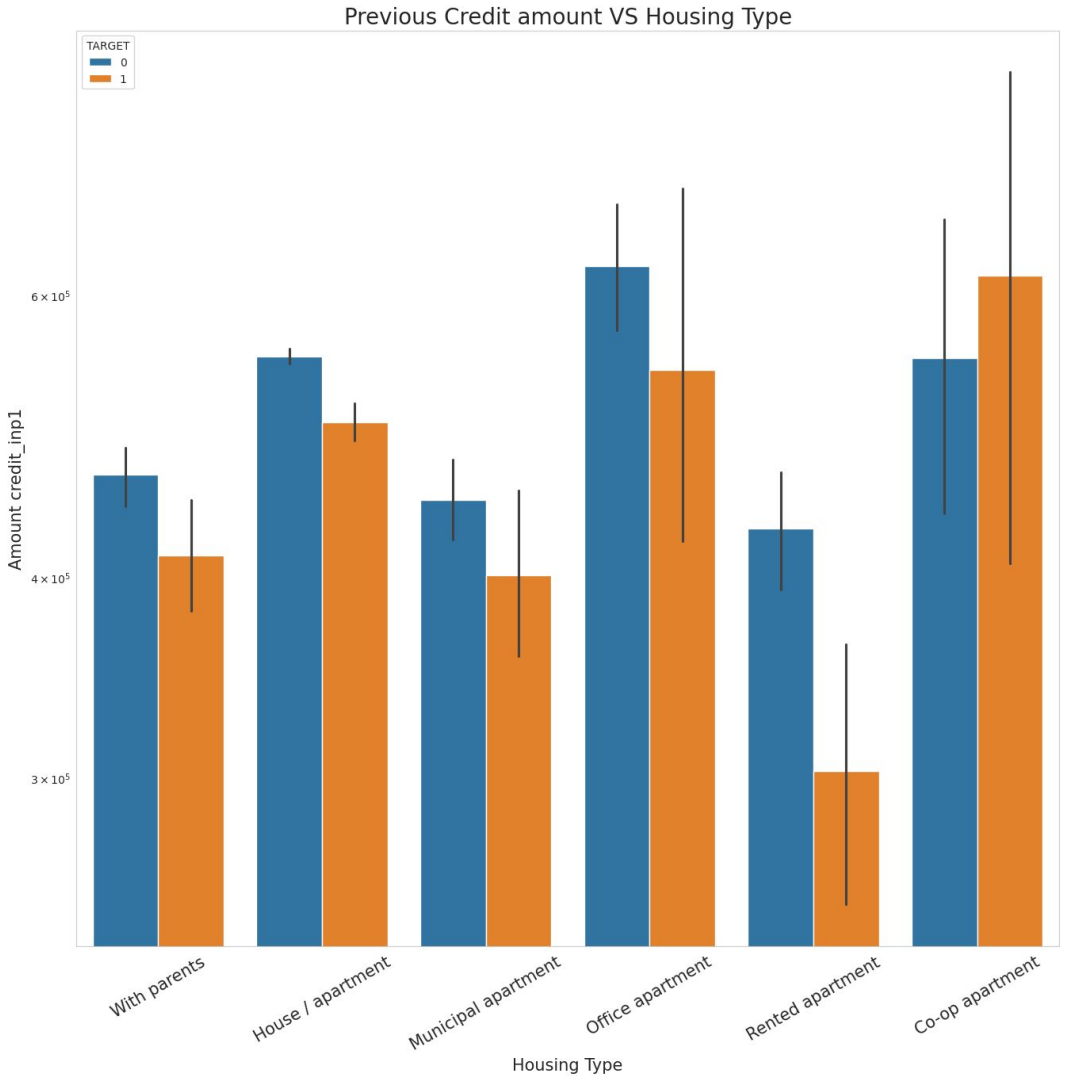



Correlation for Target=1
CNT_CHILDREN
AMT_INCOME_TOTAL
AMT_CREDIT
AMT_ANNUITY
AMT_GOODS_PRICE
REGION_POPULATION_RELATIVE
DAYS_BIRTH
DAYS_EMPLOYED
DAYS_REGISTRATION
DAYS_ID_PUBLISH
CNT_FAM_MEMBERS
HOUR_APPR_PROCESS_START
REG_REGION_NOT_LIVE_REGION
REG_REGION_NOT_WORK_REGION
LIVE_REGION_NOT_WORK_REGION
REG_CITY_NOT_LIVE_CITY
REG_CITY_NOT_WORK_CITY
LIVE_CITY_NOT_WORK_CITY
EXT_SOURCE_2
OBS_30_CNT_SOCIAL_CIRCLE
DEF_30_CNT_SOCIAL_CIRCLE
OBS_60_CNT_SOCIAL_CIRCLE
DEF_60_CNT_SOCIAL_CIRCLE
AMT_REQ_CREDIT_BUREAU_HOUR
AMT_REQ_CREDIT_BUREAU_DAY
AMT_REQ_CREDIT_BUREAU_WEEK
AMT_REQ_CREDIT_BUREAU_MON
AMT_REQ_CREDIT_BUREAU_QRT
AMT_REQ_CREDIT_BUREAU_YEAR
1.0
0.8
0.6
0.4
0.2
0.0


Credit VS GOODS PRICE for t0
GOODS PRICE
10^6
10^5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1e6
Credit


Credit VS GOODS PRICE for t1
GOODS PRICE
10^6
10^5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1e6
Credit


Credit amount VS Education status for t0
1e6
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Credit amount
NAME_FAMILY_STATUS
Married
Single / not married
Civil marriage
Widow
Separated
Unknown
Higher education
Secondary / secondary special
Incomplete higher
Lower secondary
Academic degree
Education TYPE


Previous Credit amount VS Housing Type
TARGET
0
1
Amount credit_inp1
6 × 10^5
4 × 10^5
3 × 10^5
With parents
House / apartment
Municipal apartment
Office apartment
Rented apartment
Co-op apartment
Housing Type

Conclusion

- Banks ought to give office apartment loans more approvals. cooperative apartment building so there are fewer financial challenges.
- Lending should be offered by banks for "Repairs" and "Others" purposes.
- 'Business Entity Type 3' and 'Self Employed' individuals should be eligible for loans from banks.
- *Those who are "working," particularly women, are the ideal candidates for loans.