

Credit Analysis

Submitted By

Hitha M V(APFE19a51034)

Anandu (APFE19802793)

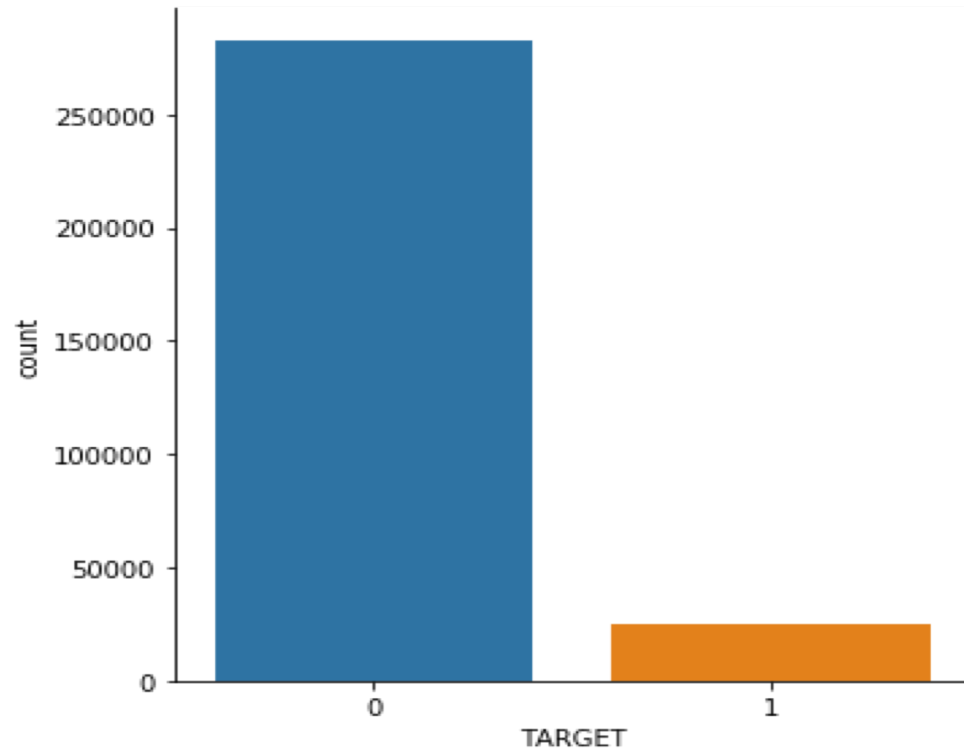
Objective

Analysis of the loan attributes which provides the strong indication to default

Date Cleaning

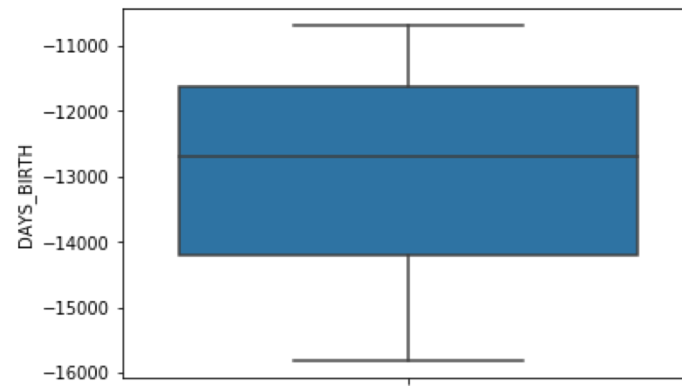
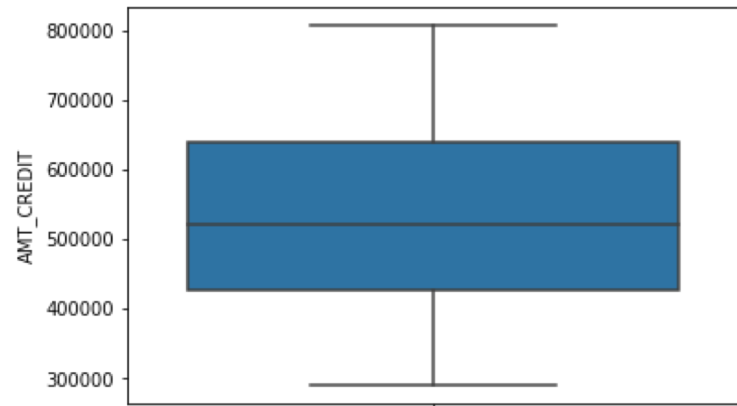
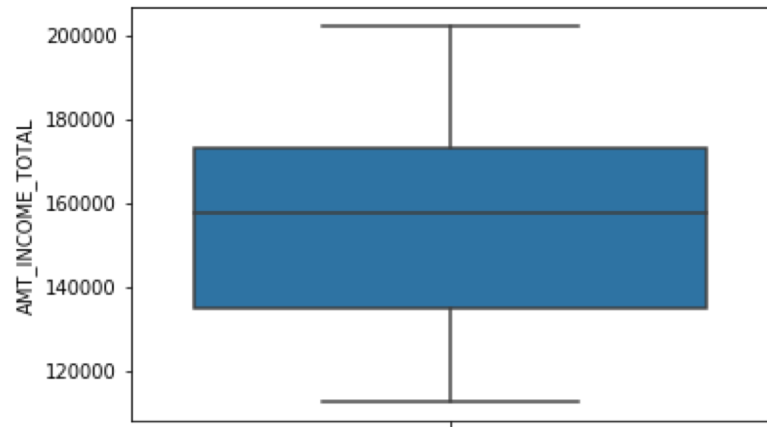
- Remove row/column with all NA
- Removing the columns with lot of Missing values
- Replace XNA/XAP with NA
- Binning of continuous variable
- Derive Columns

Data Imbalance



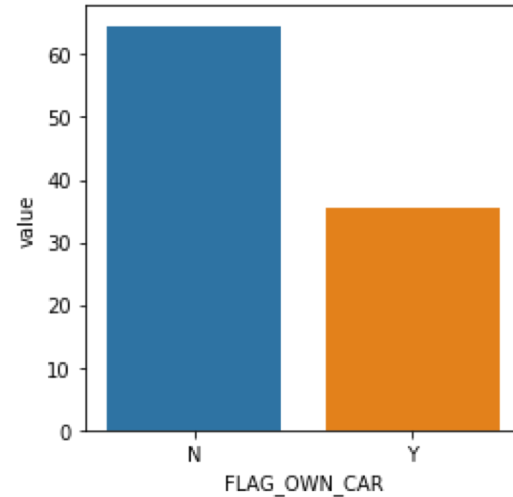
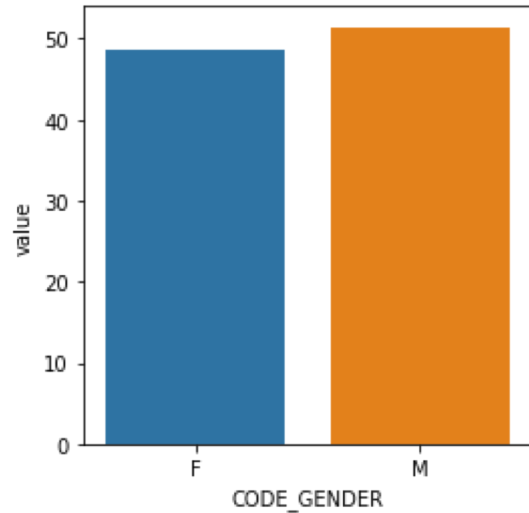
- As the data is imbalanced, using the percentage to the target, 8 % before removal, 13% after removal

Numeric Attribute Analysis

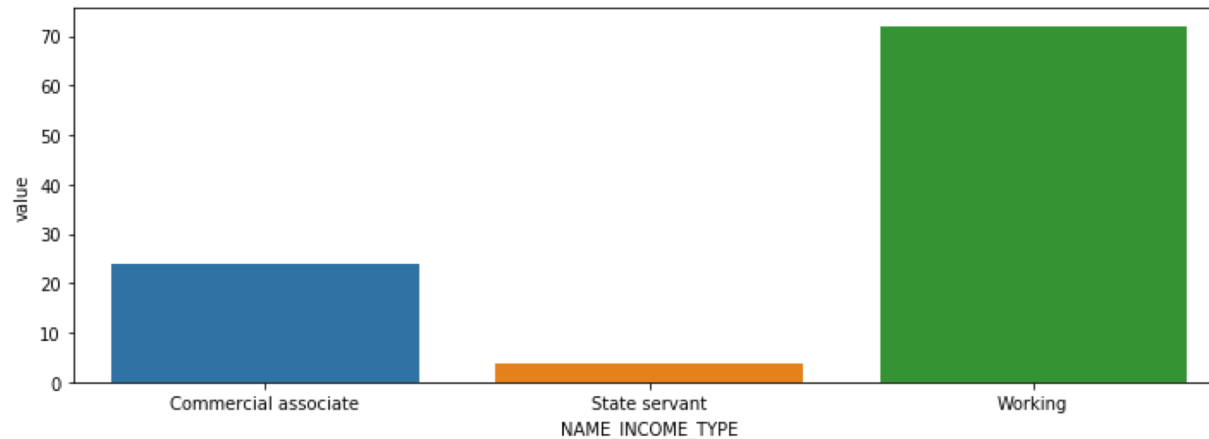


- Outliers are removed by taking the values between 0.25 and 0.75 quantile
- Outliers are removed for AMT_INCOME_TOTAL, AMT_CREDIT, AMT_GOODS_PRICE, DAYS_EMPLOYED, DAYS_REGISTRATION and DAYS_BIRTH

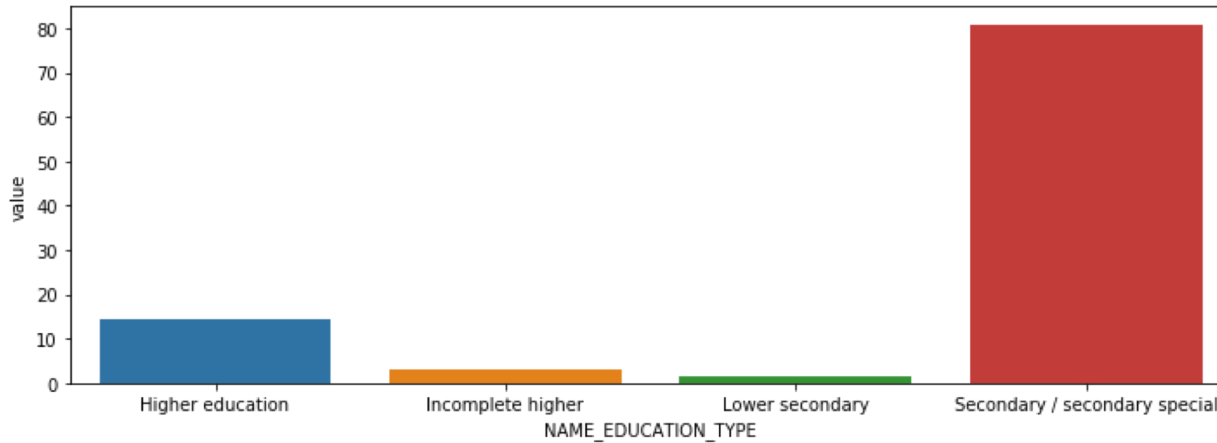
Categorical Attribute Analysis



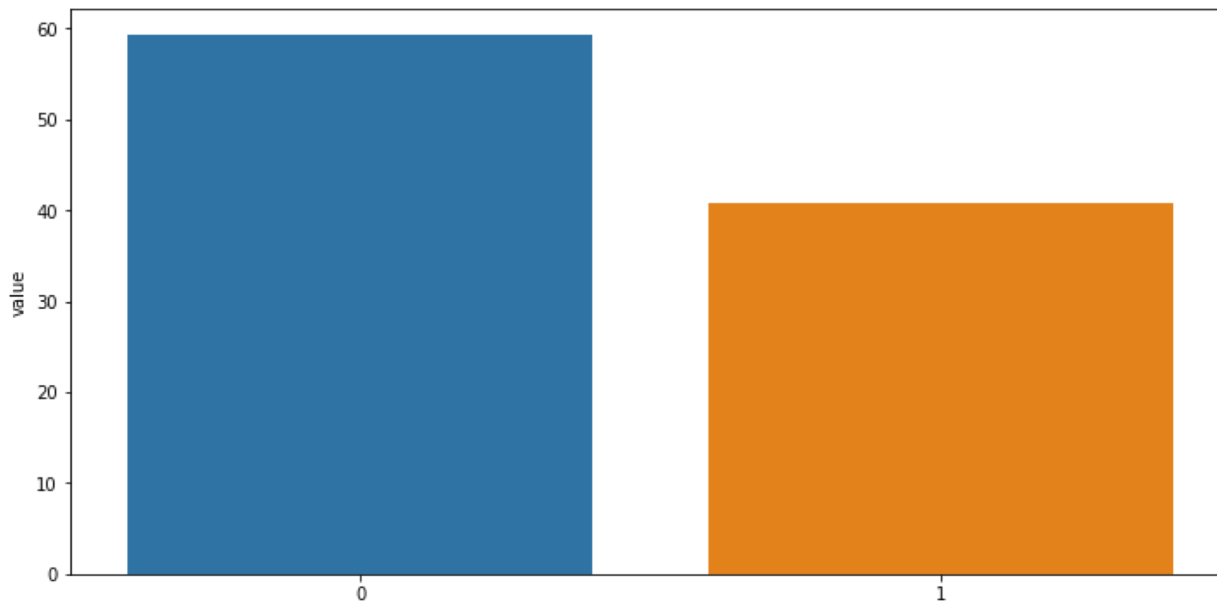
- Males are defaulting more than females
- Clients who don't have cars are more defaulting
- Working clients are defaulting more



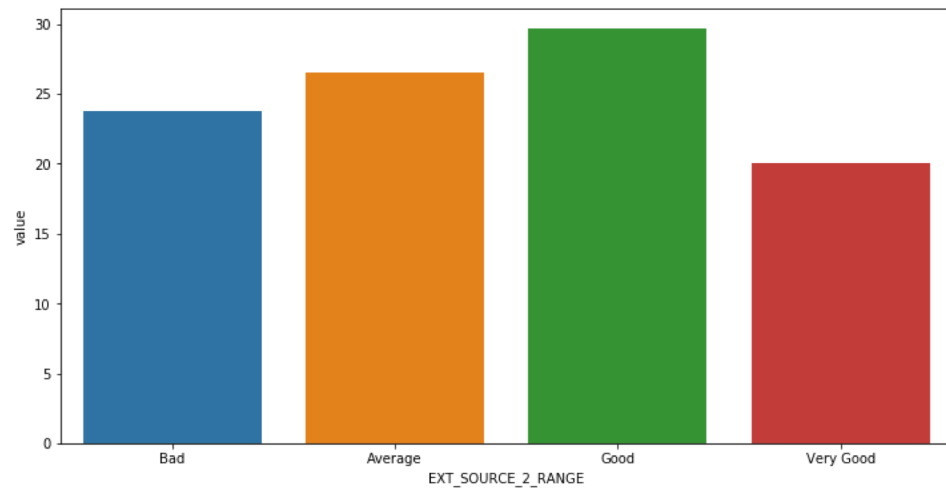
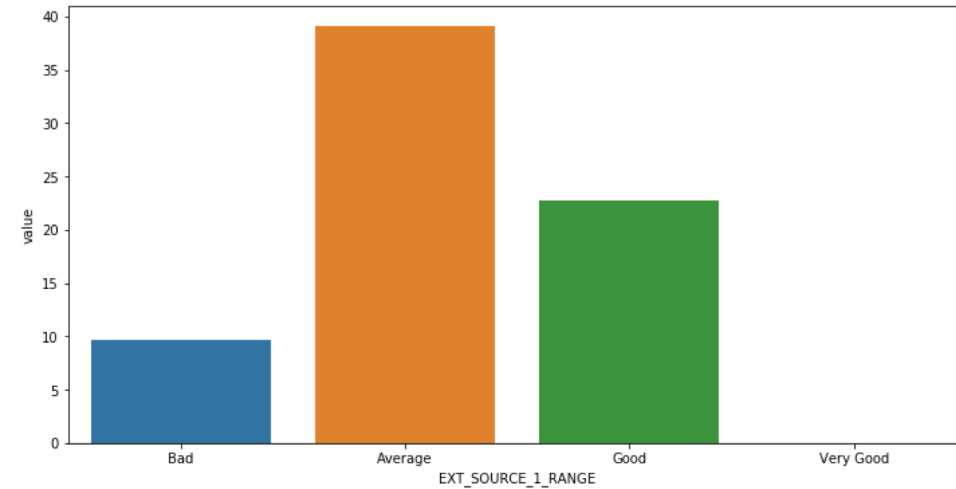
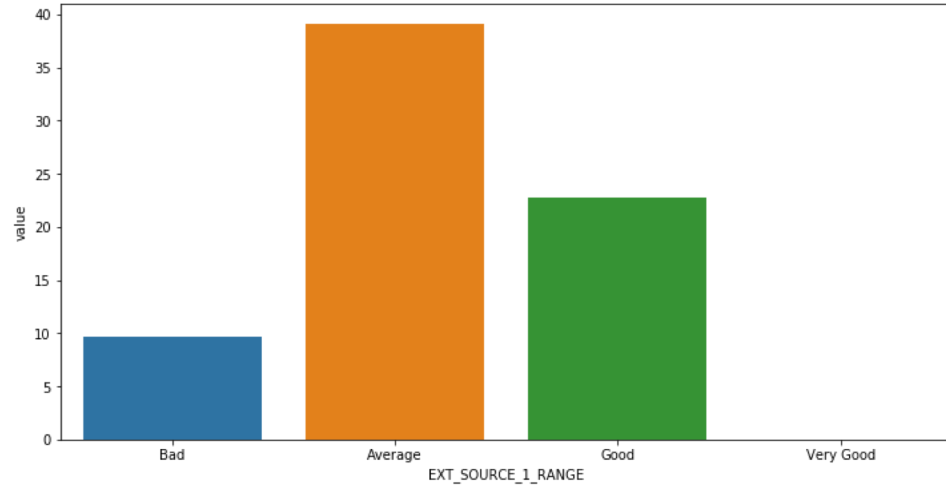
Categorical Attribute Analysis



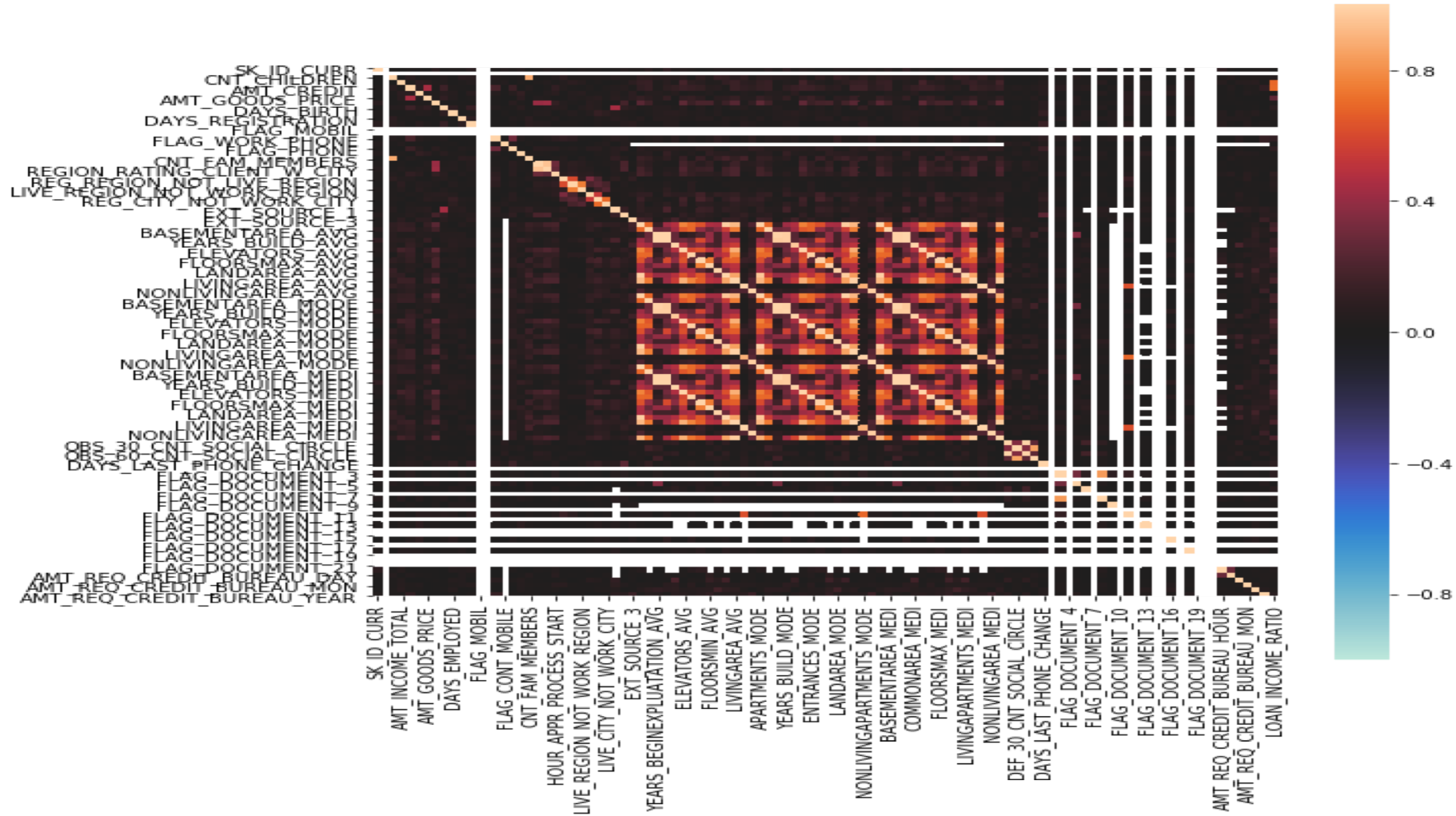
- Clients with Secondary education are defaulting more
- if clients permanent address does not match contact address defaults more



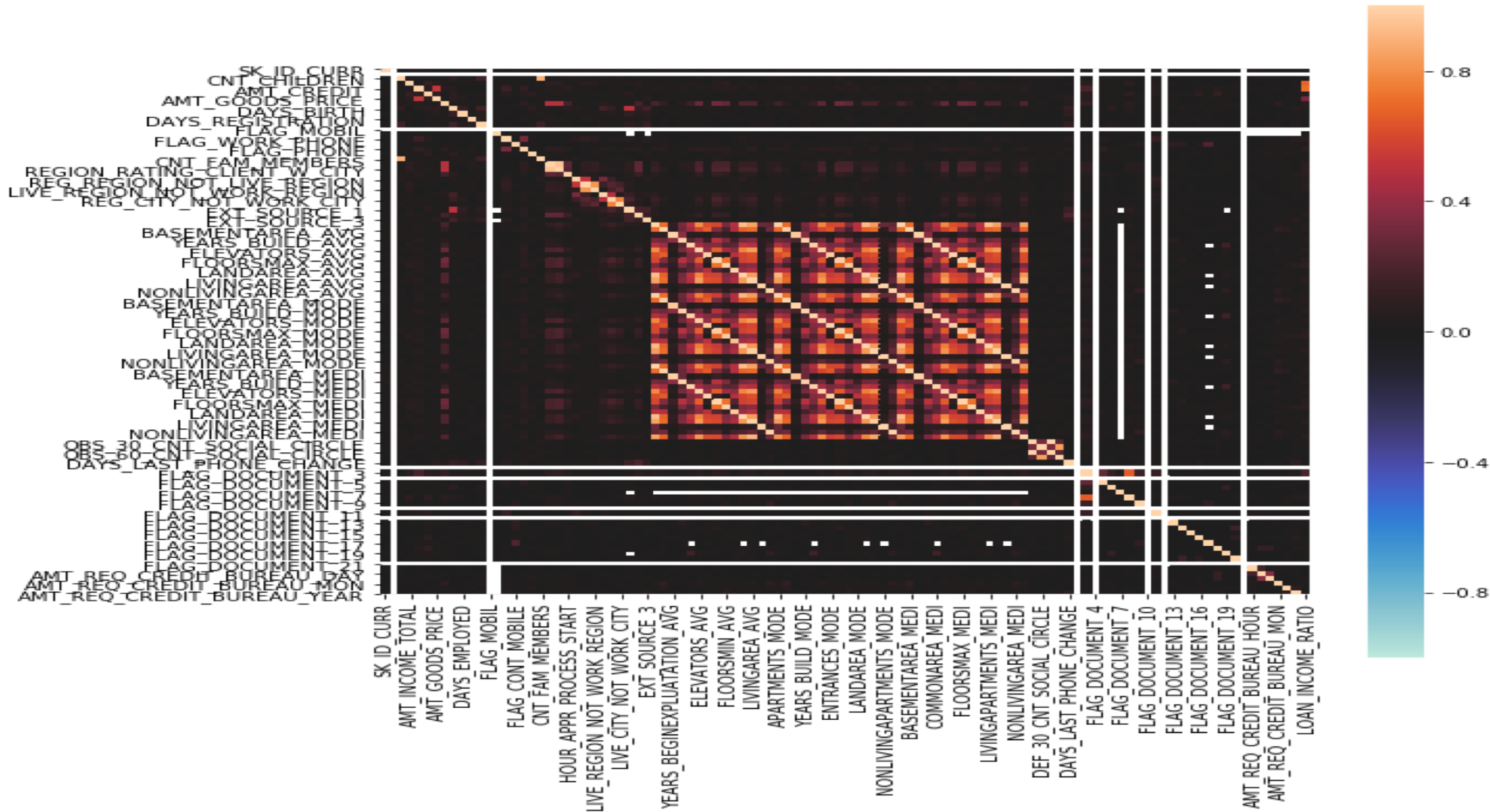
External Source Analysis after binning



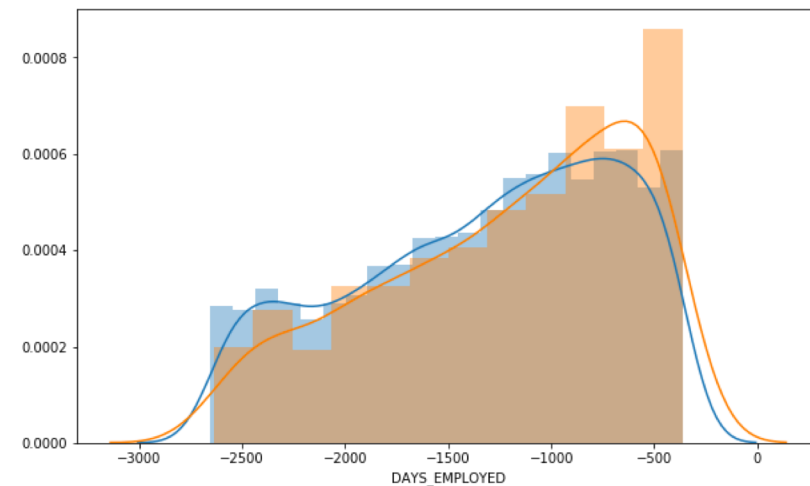
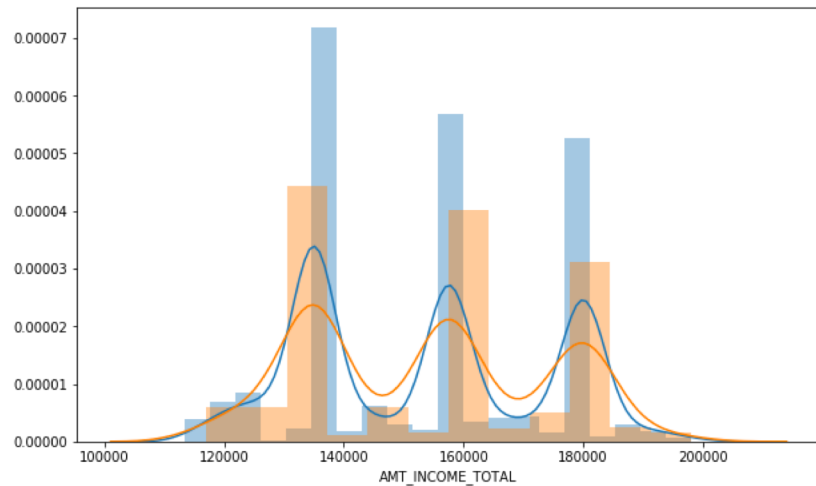
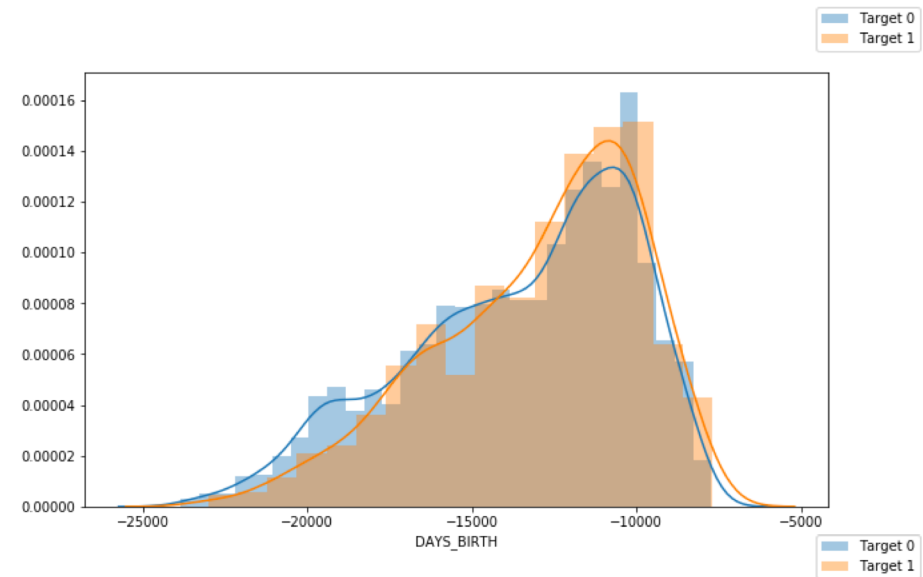
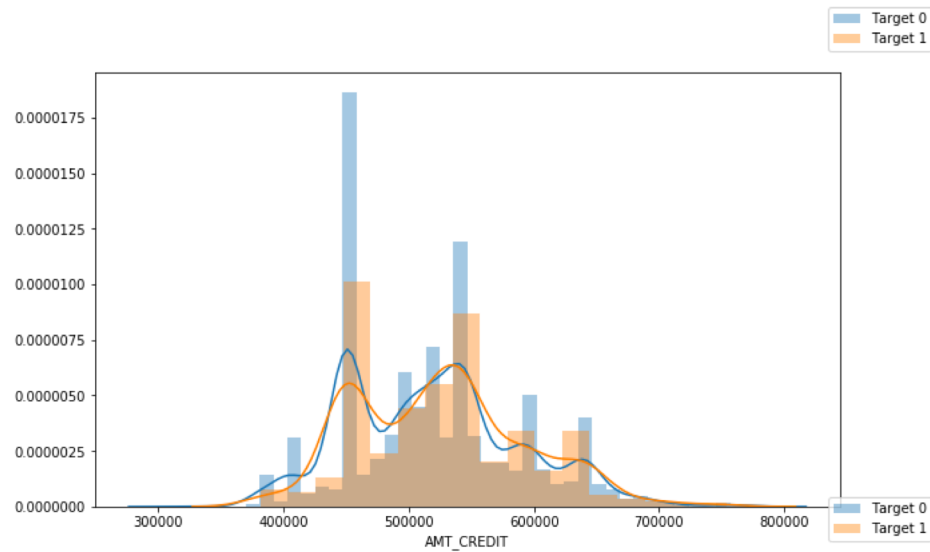
Default Correlation



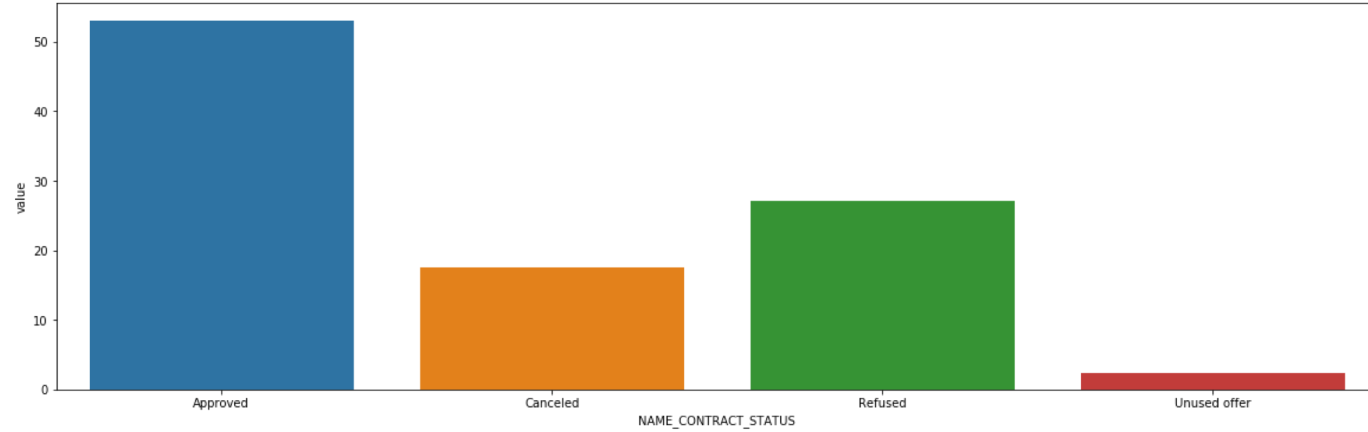
Non Default Correlation



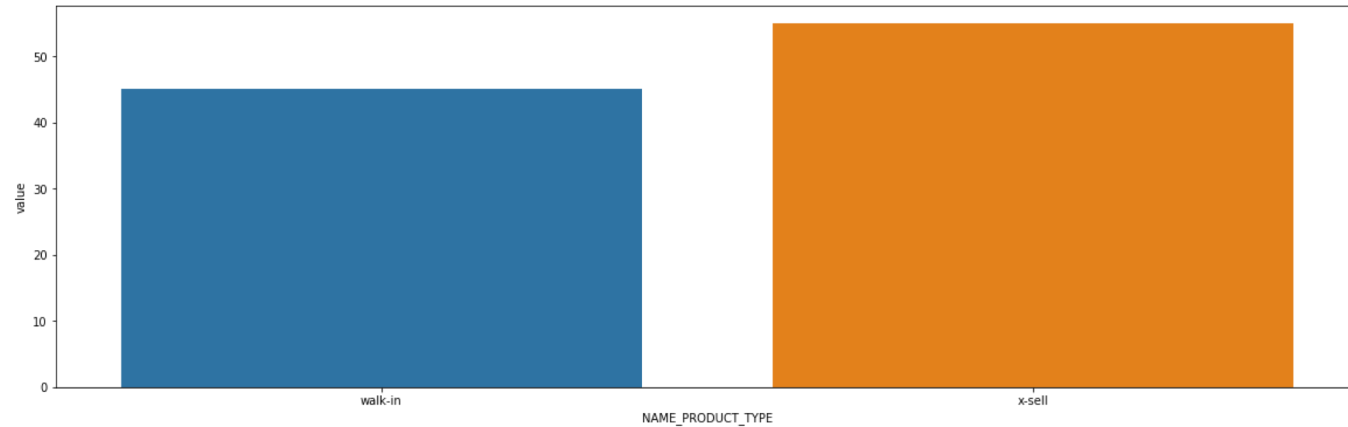
Numeric Attribute Distribution



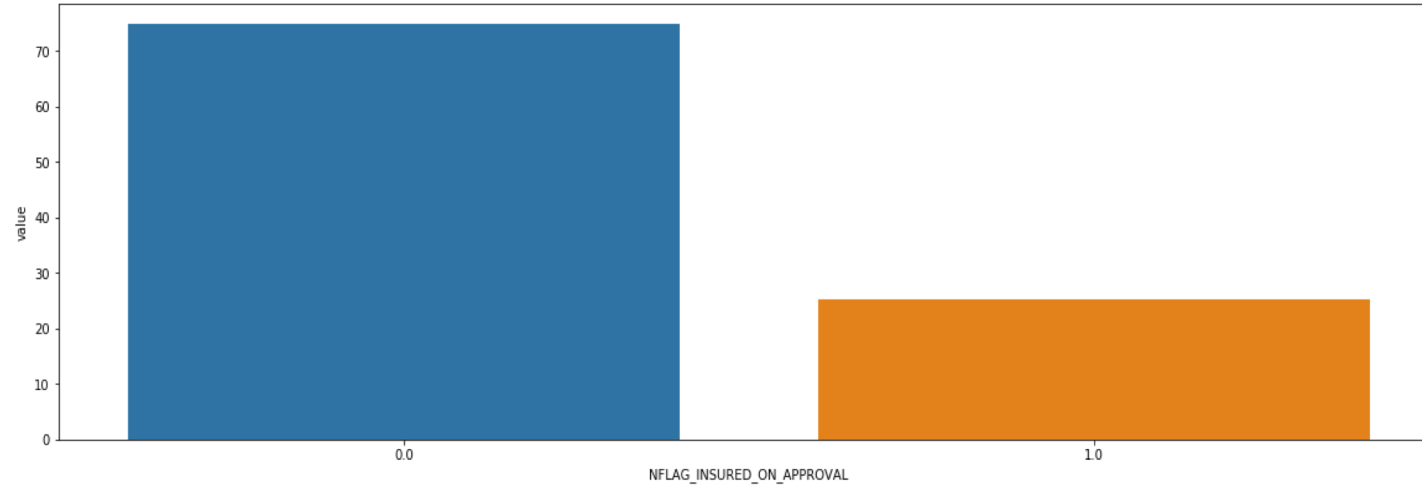
Category Analysis



- Walk in clients are more likely to default
- Previous Refused clients are likely to default



Insured or Not



- If client not requested insurance, will default

Summary

- Working Clients with Secondary Education , whose previous application is rejected is more likely to default
- Client who didn't request insurance, and walked in clients are more like to default