

Credit EDA Assignment

By Rohan Pandey

Missing Values

- Columns with more than 40% missing values were dropped
- Rows with majority missing values were dropped
- Imputed missing values of AMT_ANNUIITY variable using the median of the field since there are outliers present

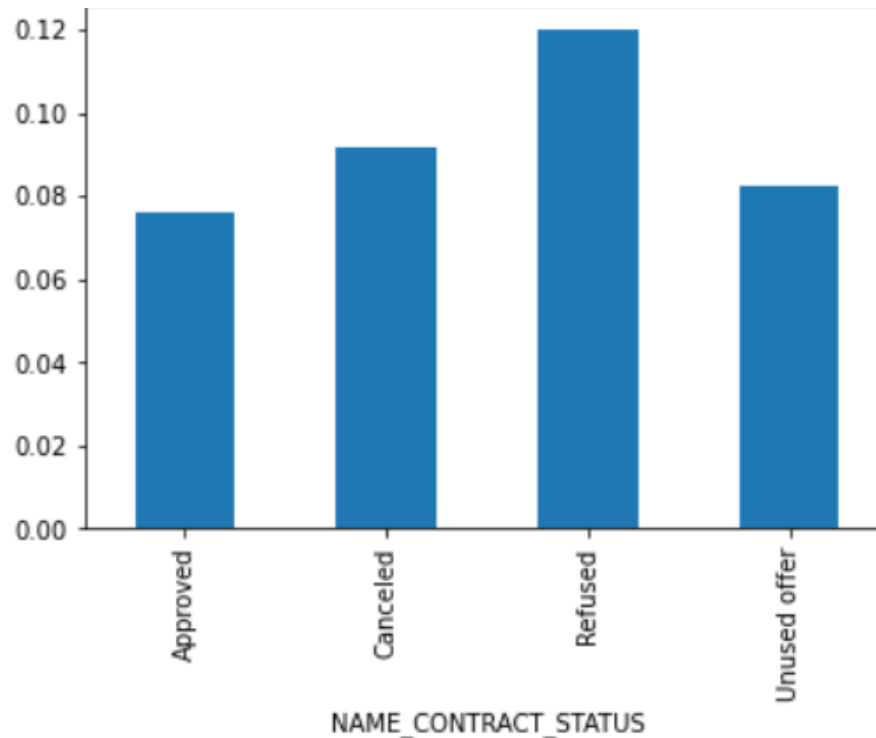
Outliers Analysis

- Outliers lie in the following fields:
 - One outlier lying at 11.7 crore in AMT_INCOME_TOTAL
 - Many above 15 lacs in AMT_CREDIT
 - Many above 60,000 in AMT_ANNUITY
 - Many above 13 lacs in AMT_GOODS_PRICE
 - A few around 1000 years in EMPLOYED_YEARS which is not possible so can be removed
- These outliers should be handled separately or removed if it is an incorrect value

List of Defaults and Detail Analysis of the following

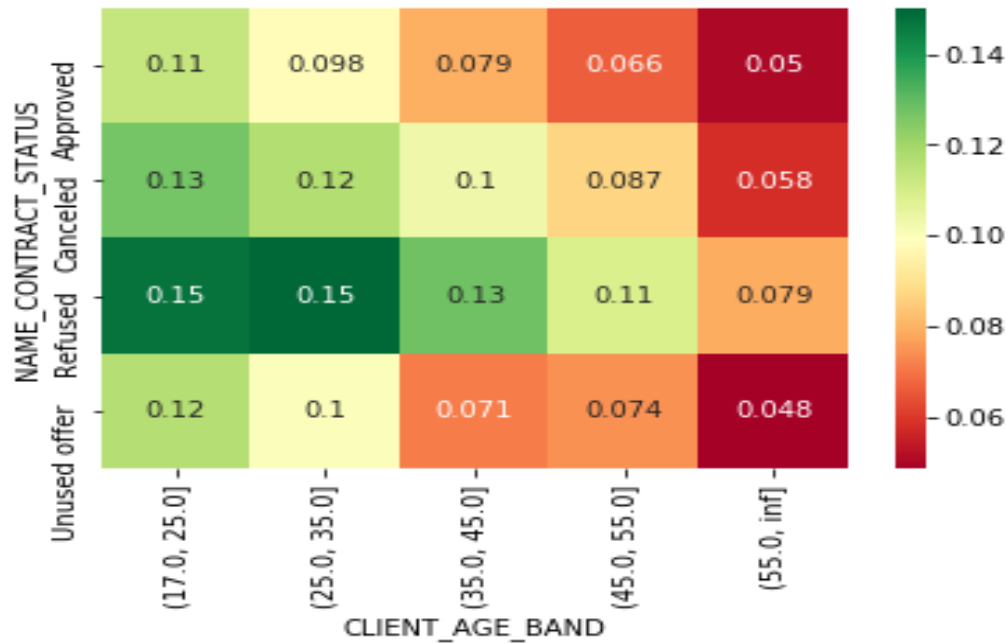
- Previously refused loans
- Client's age
- Client's family status
- Contract type
- Client's gender
- Client's education level
- Loan amount

Previously Refused Loans



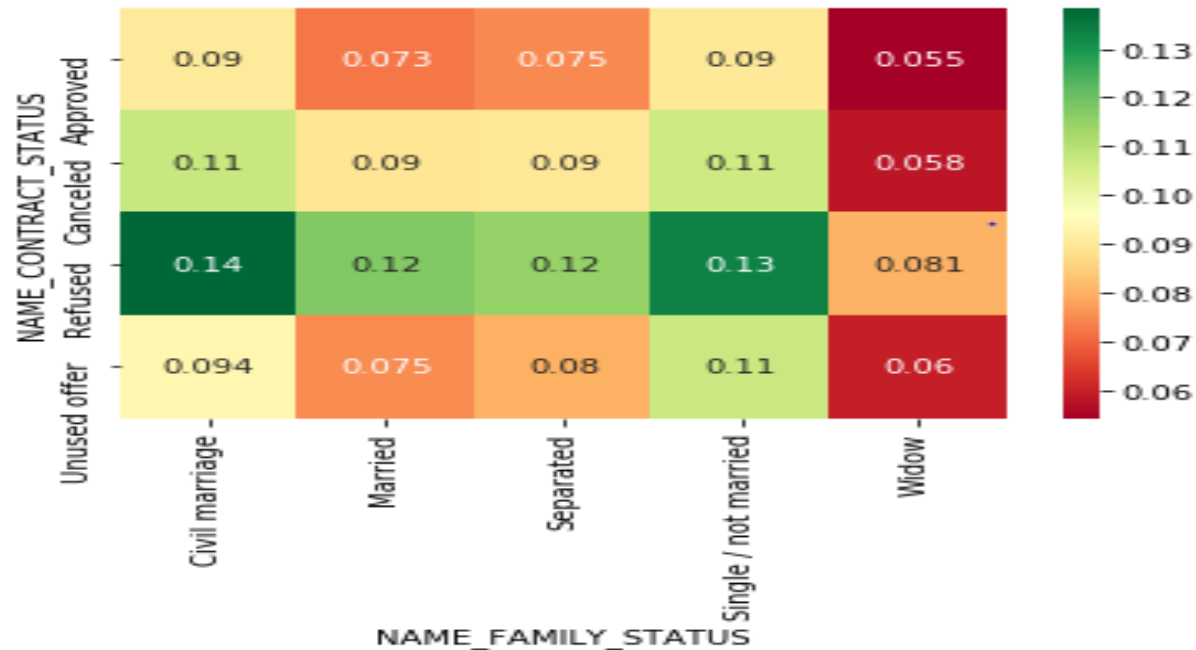
- The clients whose loans were previously refused, larger proportion of them have payment difficulties currently more than other cases
- The clients whose loans were previously approved, least proportion of them have payment difficulties currently
- As seen above, it is almost 12% for refused loans while 75% for the approved loans

Age



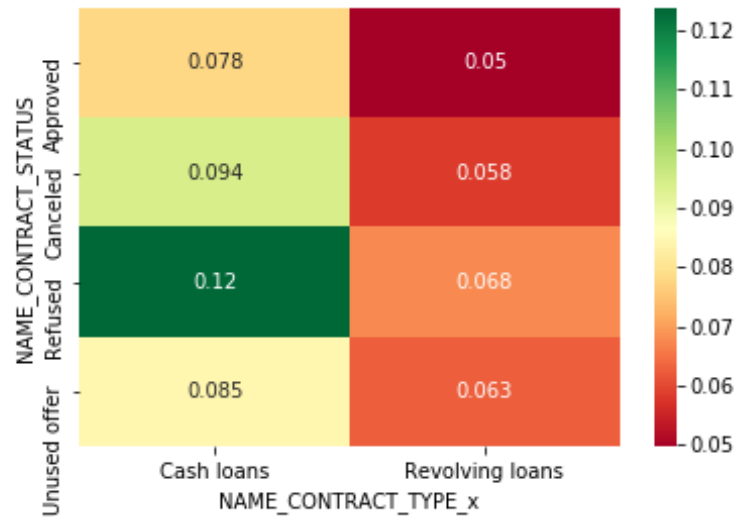
- Age group 17-35 show the highest proportions of clients who have payment difficulties so the riskiest for the bank
- Age group 55 and above shows the lowest proportions of clients who have payment difficulties and so are the least risky for the bank

Family Status



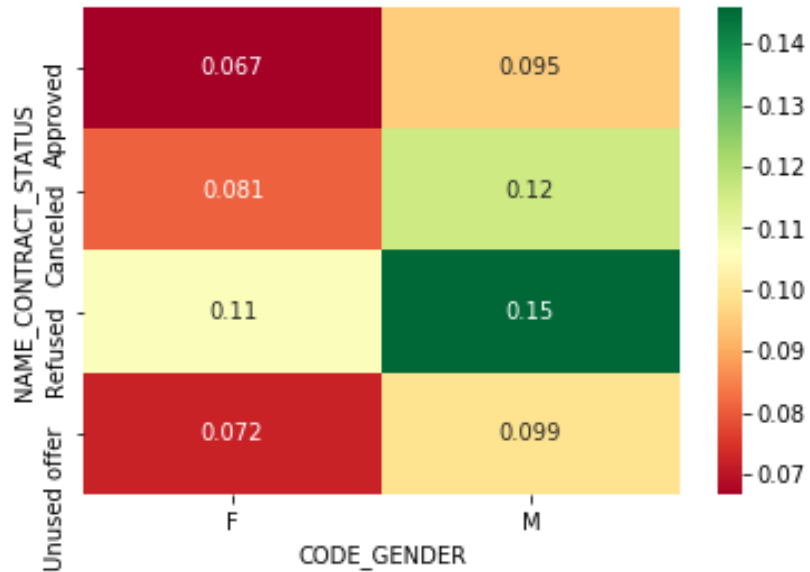
- Widows show the lowest proportion of clients who have payment difficulties so are the least risky for the bank
- Unmarried and civil married clients show the highest proportion who have payment difficulties and so are the riskiest for the bank

Contract Type



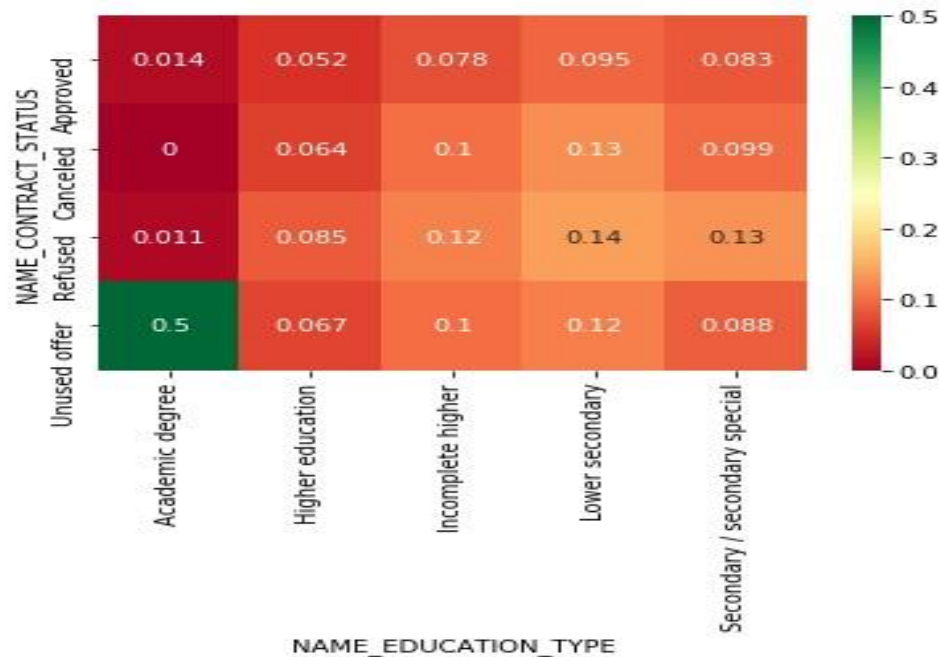
- Cash loans have a significantly higher proportion of clients who have payment difficulties compared to revolving loans
- On an average, proportion of clients with payment difficulties is ~95% or cash loans while only ~0% for revolving loans

Gender



- Female clients have a lower proportion of clients who have payment difficulties compared to males and so are less riskier for the bank
- On an average, the proportion of clients who have payment difficulties is ~11.6% for Males while only ~8.2% for Females

Education Level



- Clients obtained education only till lower secondary have a higher proportion of ~12% compared to rest
- Clients who have an academic degree and previously did not use the offer have a very high proportion of ~50% of having payment difficulties

Loan Amount



- Loan credit amount in the range of 4lacs to 6lacs has the highest proportion of ~11.5% of clients having payment difficulties

Recommendations

- Avoid sanctioning loans to clients whose applications were previously rejected.
- Prefer sanctioning loans to widows and avoid sanction of loans to unmarried and civil married clients.
- Female clients should be preferred over males since that is low at risk.
- Avoid sanctioning loans to the age group of 17 to 35 and to sanction more loans to an age group of 55 and more.
- Clients who have completed education till lower secondary only should be avoided because they are prone to be defaulters.
- Loan ranging from 4-6 lacs should be avoided since that is at the most risk.
- Sanctioning revolving loans is a better option than cash loan since lower risk for the bank and hence, it could also be used to hedge the risk of cash loans.