CREDIT EDA ASSIGNMENT

SUBMITTED BY –

Kamna Sethi

DSC-56

Data Science

Business Objective

This case study aims to identify patterns which indicate if a client has difficulty paying their instalments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

Problem Statement

When the company receives a loan application, the company has to decide for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

1. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.

2. 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.

Approach

•Steps used in EDA
1. Importing data
2. Meta data check
3. Data cleaning
>- Imputing missing values
>-Detection of outlier
>- Binning
4. Analysis
>- Univariate
>- Bivariate
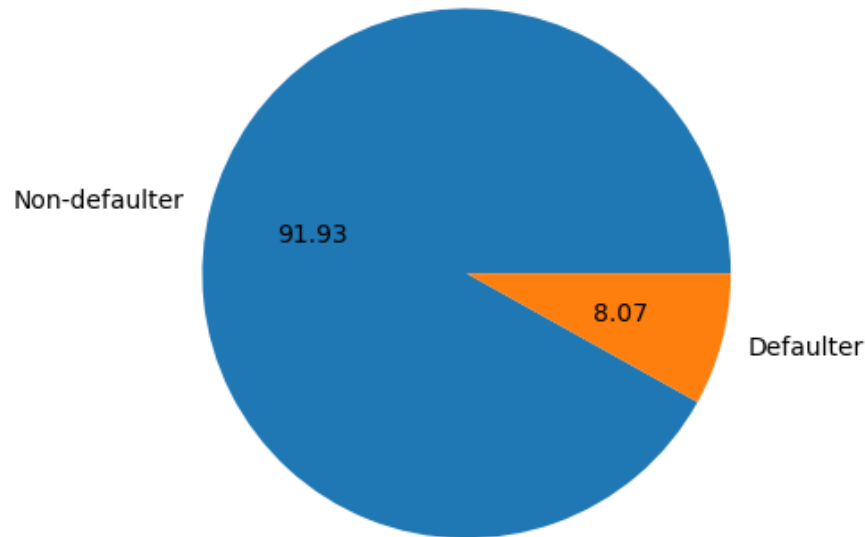>-Multivariate
5. Merging data
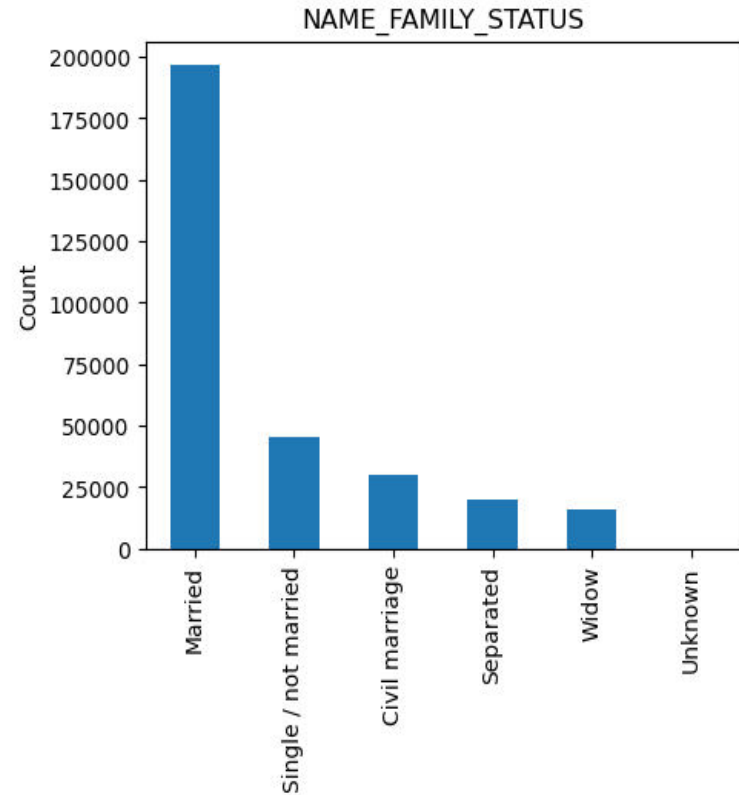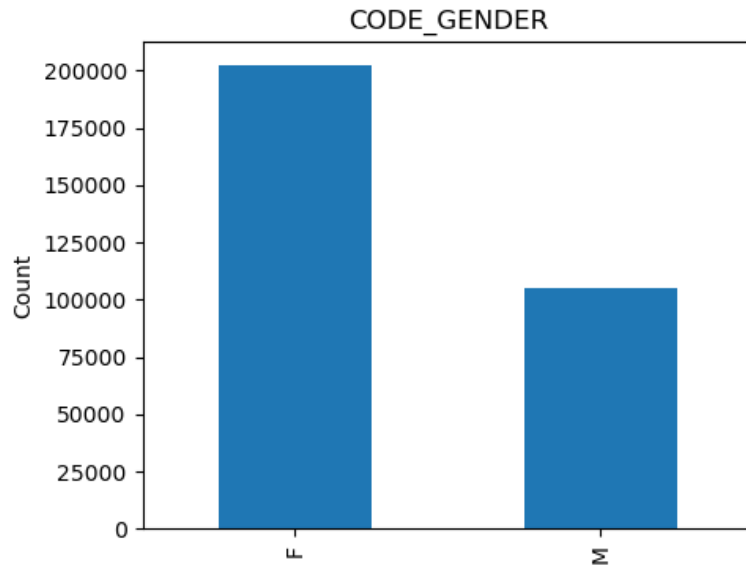6. Conclusion

Data Imbalance

Pie chart for target column

1. Percentage of Non-defaulters are 91.93% and of Defaulter are 8.07%.
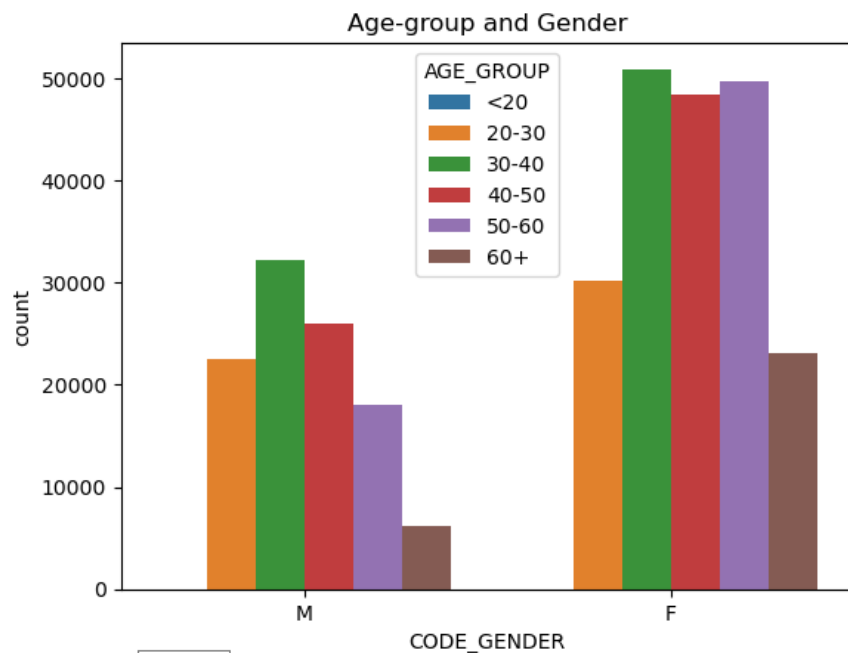
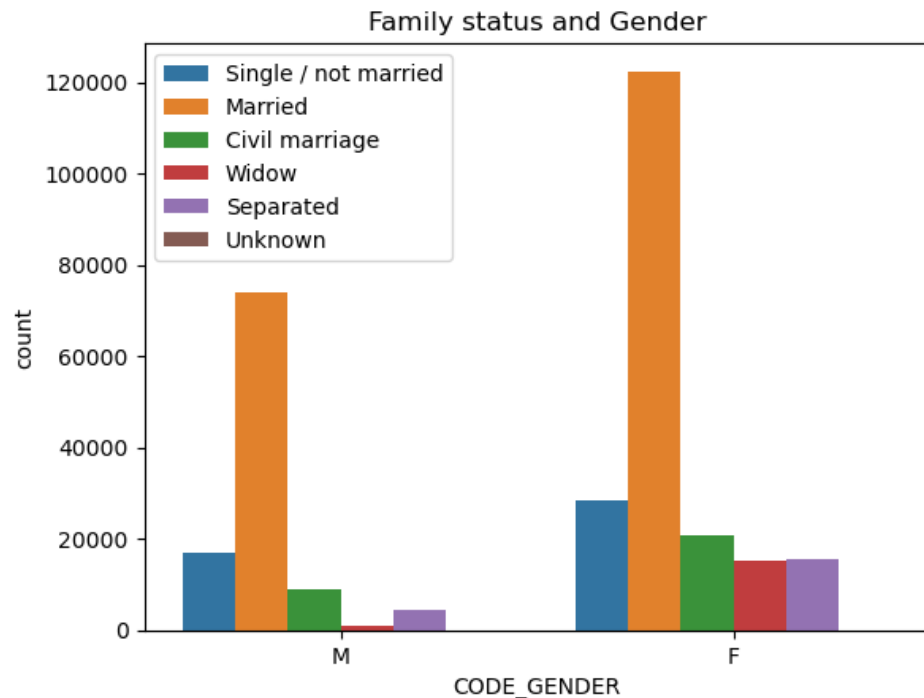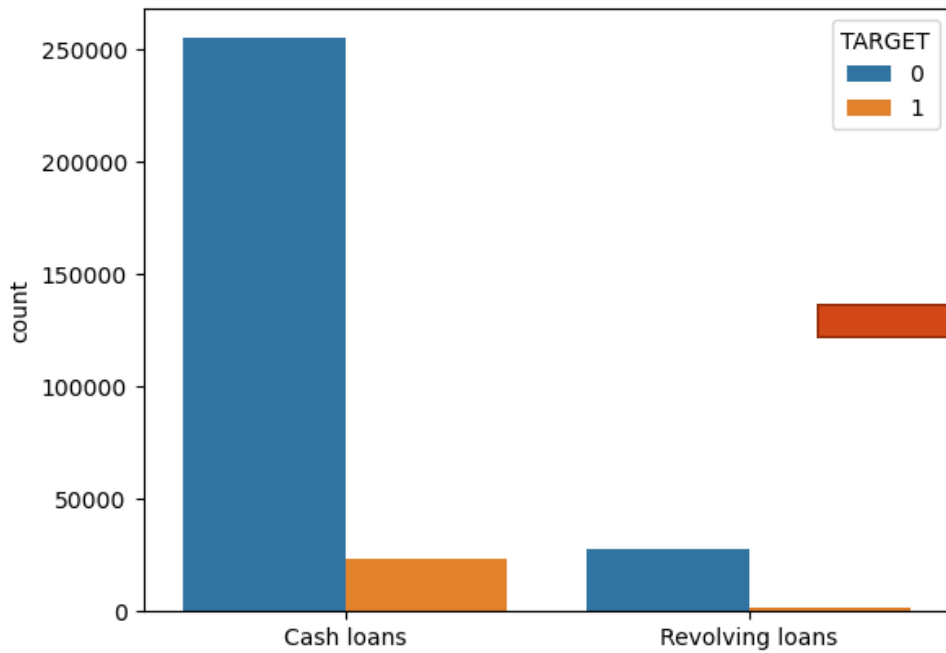**2. Imbalance ratio is 11.39.**

Univariate Analysis

CONCLUSION
More Females have applied for loan than Males.
Mostly Married people have applied for loan.

Age-group and Gender

Family status and Gender

Most Females of 30-40 age group have applied for loan followed by 50-60 age group.
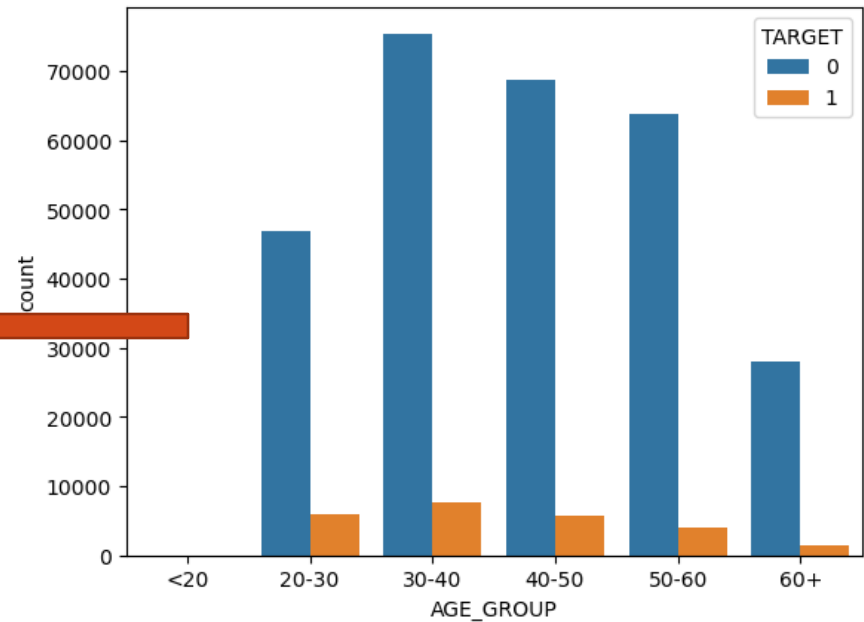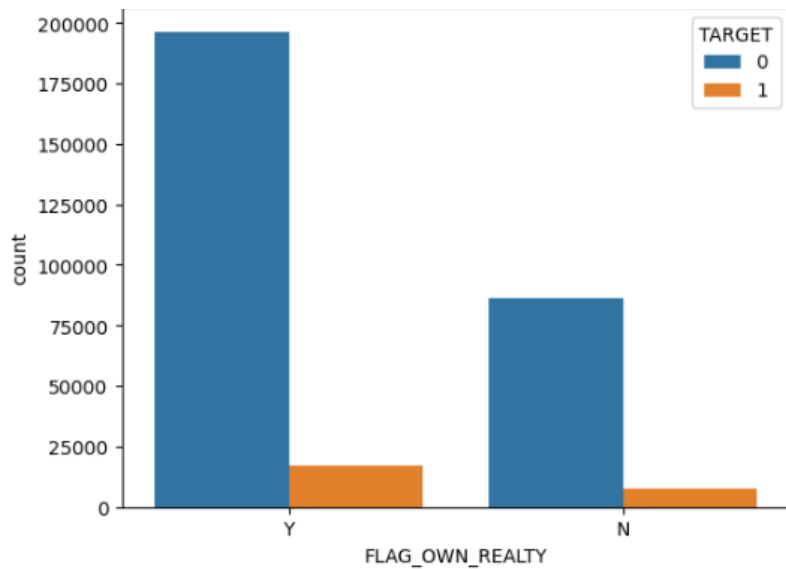
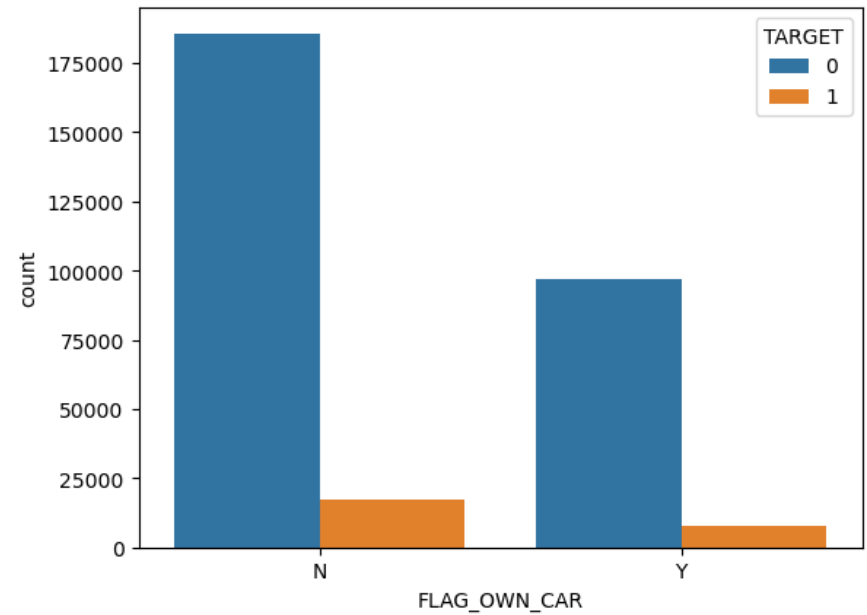**Mostly Married Females are taking more loan.**

More chances of default on cash loans.

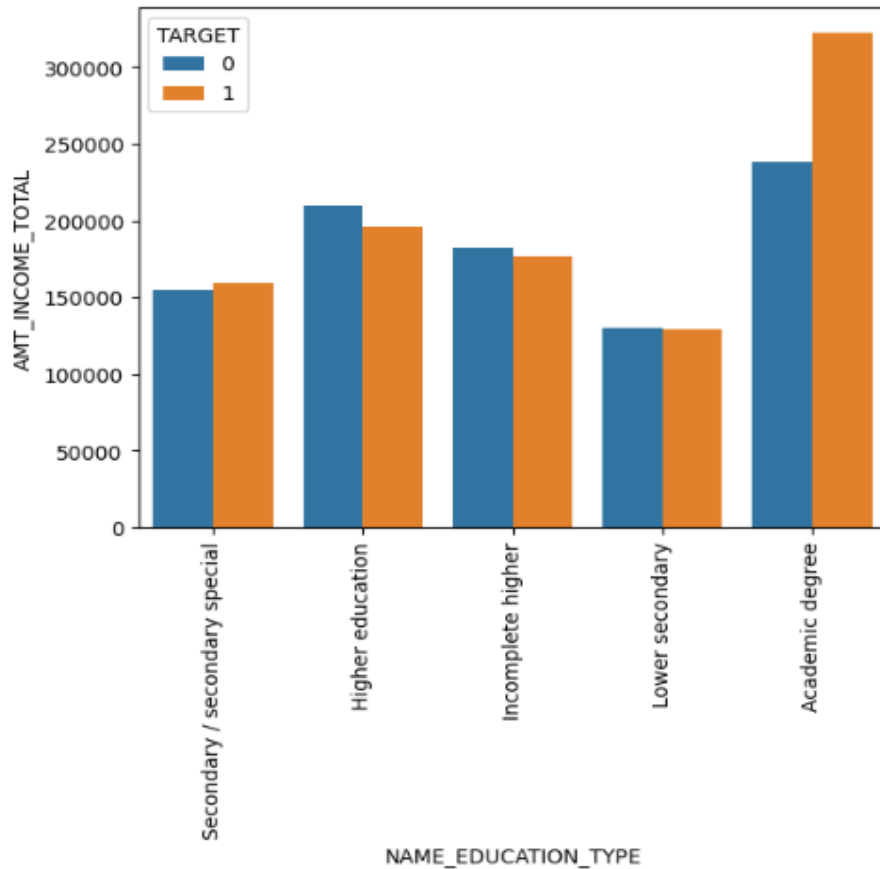**People less than the age group of 30 are more likely to default.**

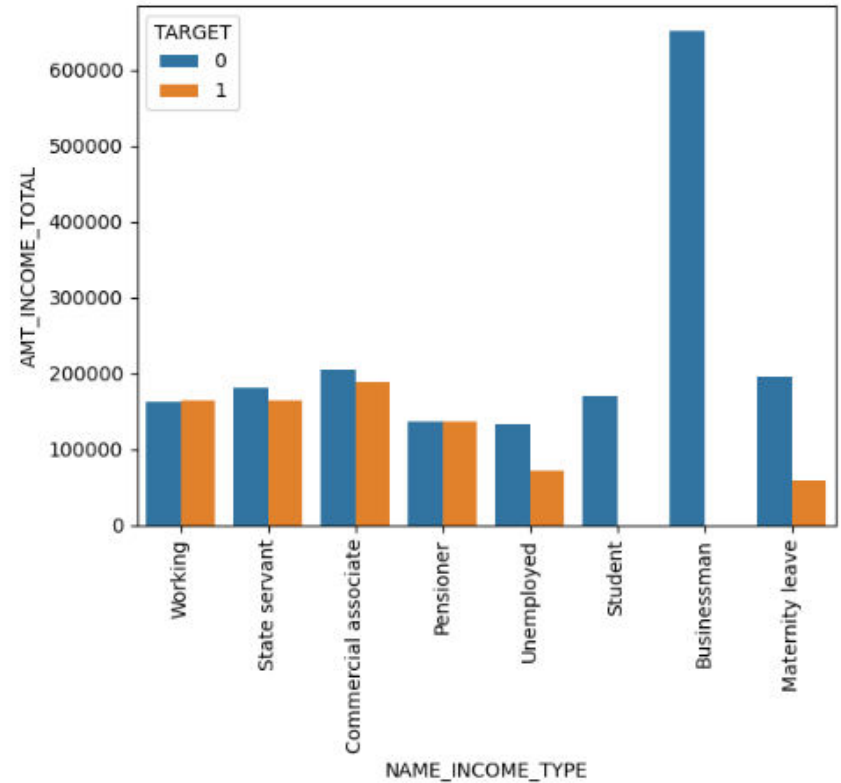**More chances of default on People who don't own Realty**

People who own a car are more likely to default
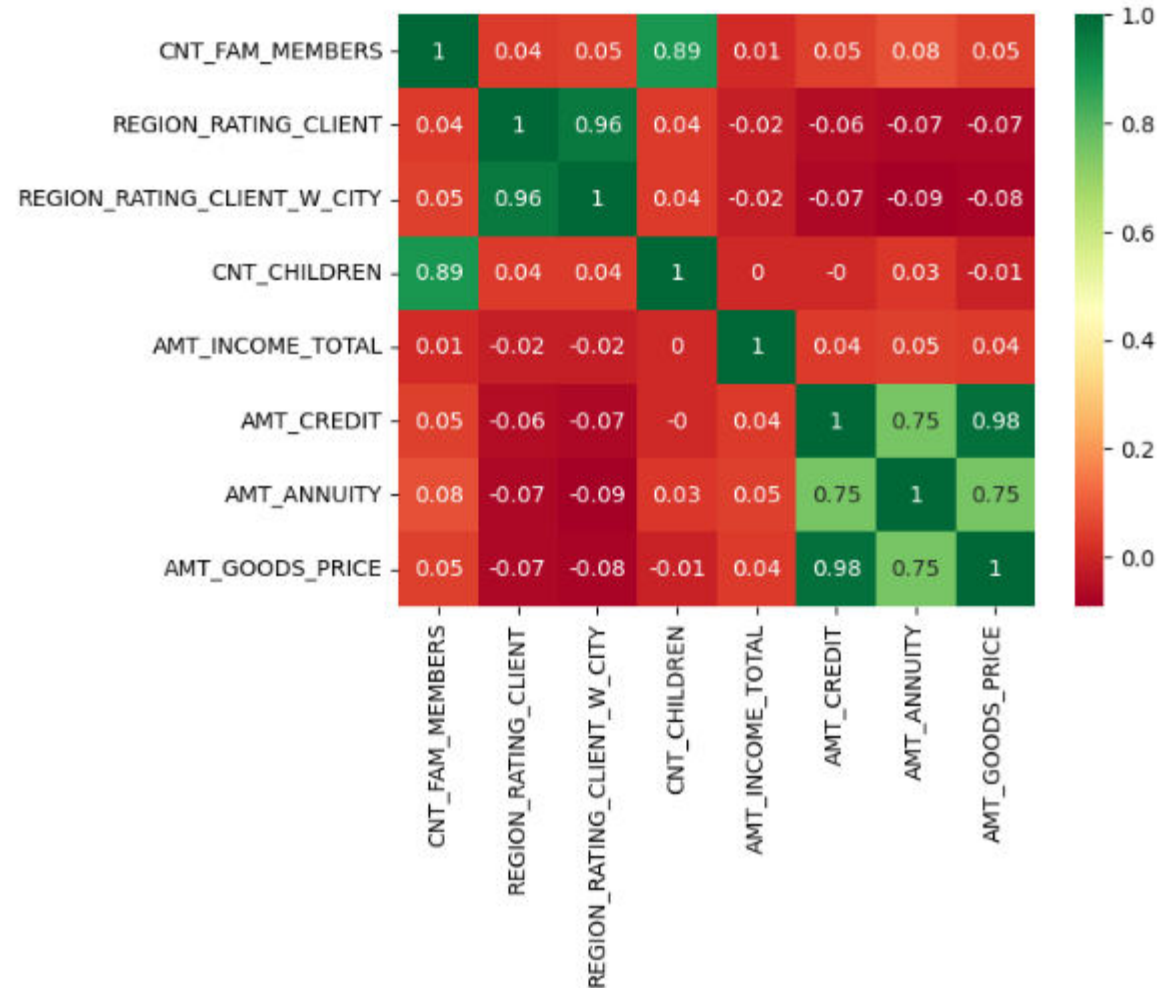
Bivariate Analysis

More number of defaulters in academic degree and very less number of defaulters are in higher education .

No defaulters are in business class and students category.
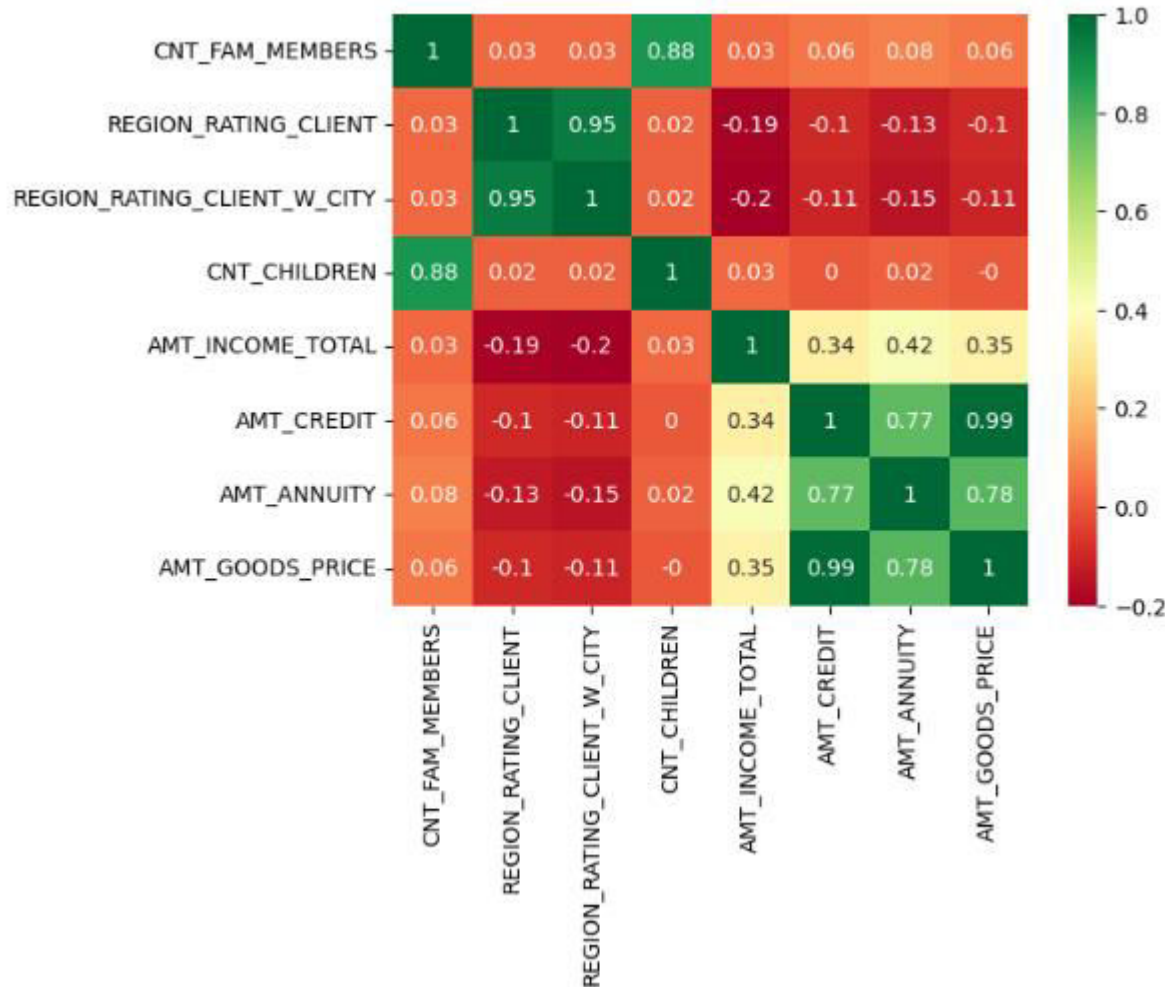
HEAT MAP FOR DEFAULTER

Top correlations

CNT_CHILDREN and CNT_FAM_MEMBERS=0.89

REGION_RATING_CLIENT and REGION_RATING_CLIENT_ W_CITY =0.96

AMT_GOODS_PRICE and AMT_CREDIT=0.98
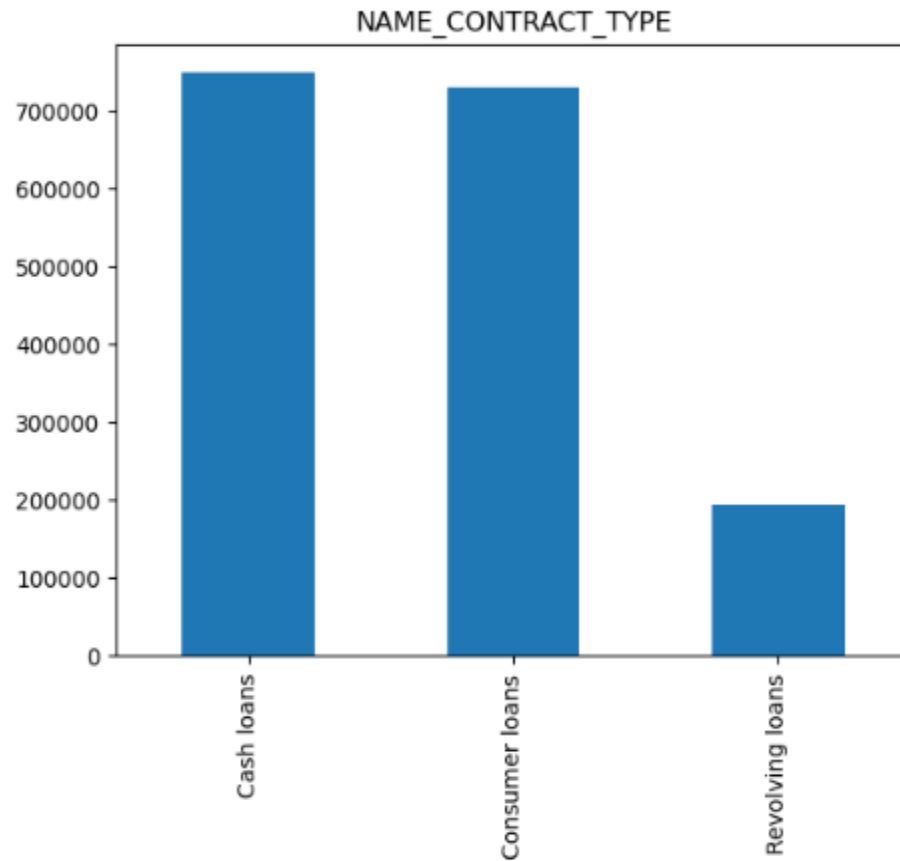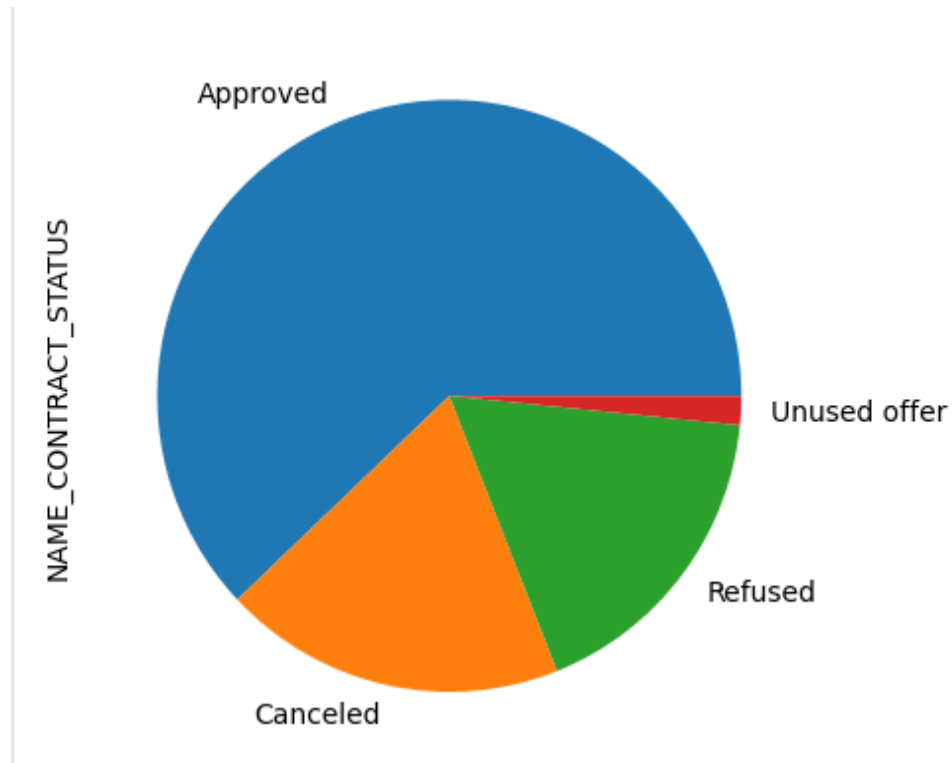
HEAT MAP FOR NON- DEFAULTER

Top correlations

AMT_GOODS_PRICE and AMT_CREDIT=0.99

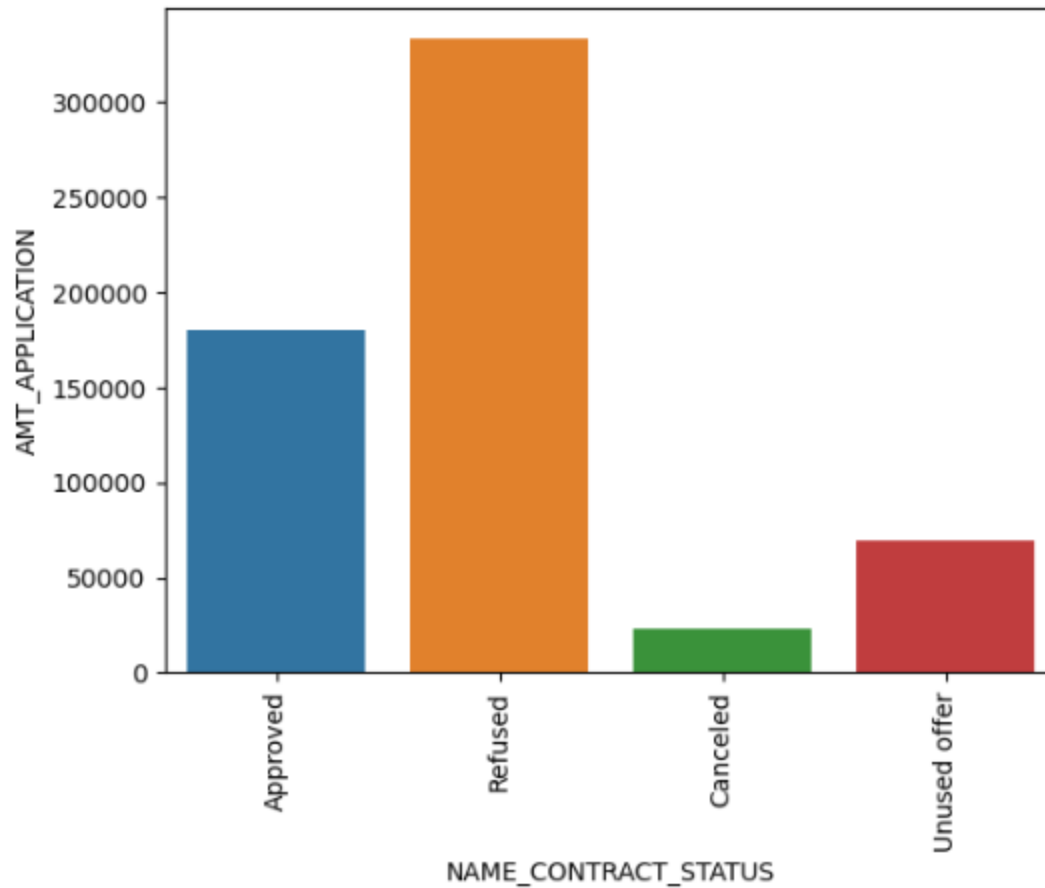**CNT_CHILDREN and CNT_FAM_MEMBERS= 0.88**

REGION_RATING_CLIEN T and REGION_RATING_CLI ENT_W_CITY =0.96
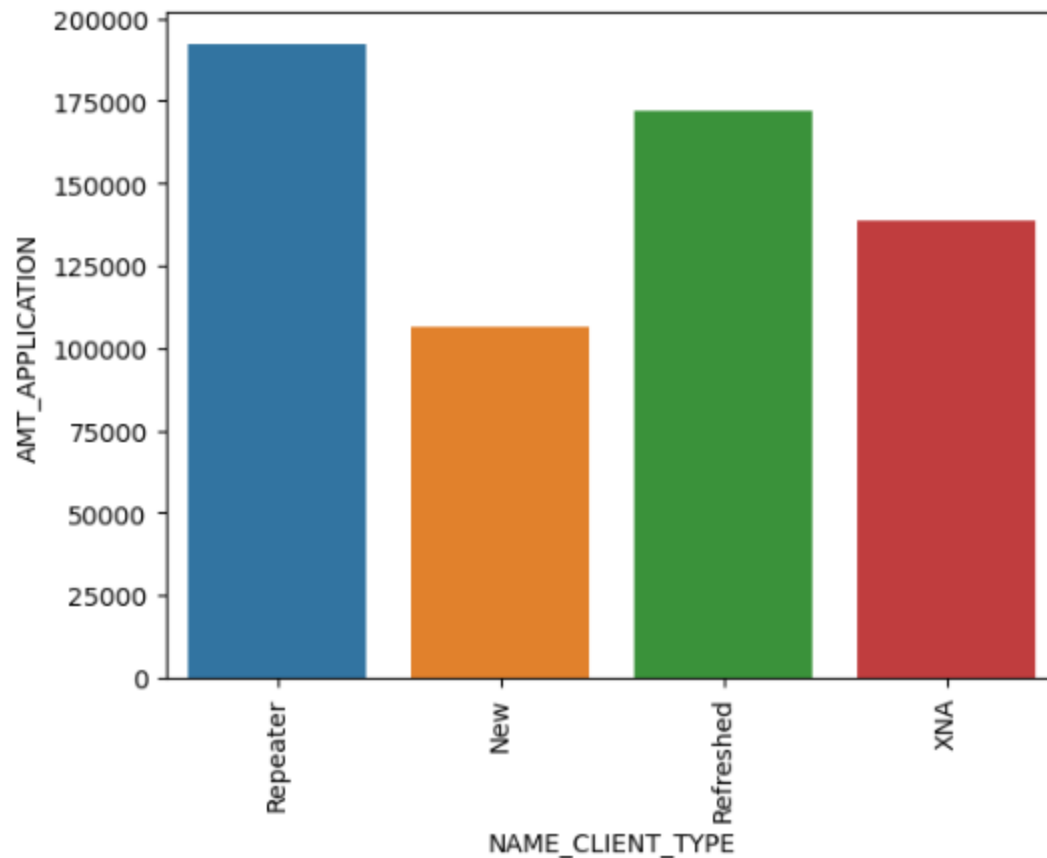
NAME_CONTRACT_TYPE

There are more number of cash and consumer loans provided bank in previous loan. But in current loan bank is only providing cash and revolving loans.

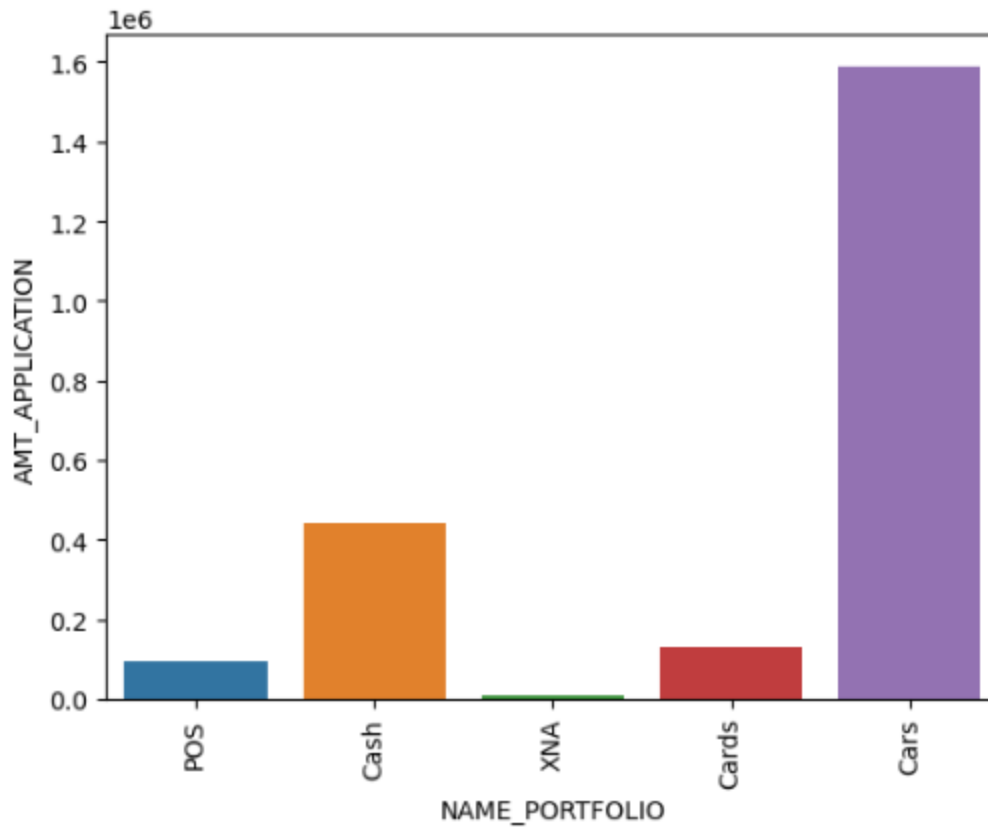More number of approved loans can be seen, very less number of unused offers are there.

A large number of amount application has been Refused and very less are canceled.

More number of repeaters can be seen as compared to new clients. So, Bank should focus on making more new clients in order to generate more profit.

A large number amount application are on cars, bank should also focus on POS.
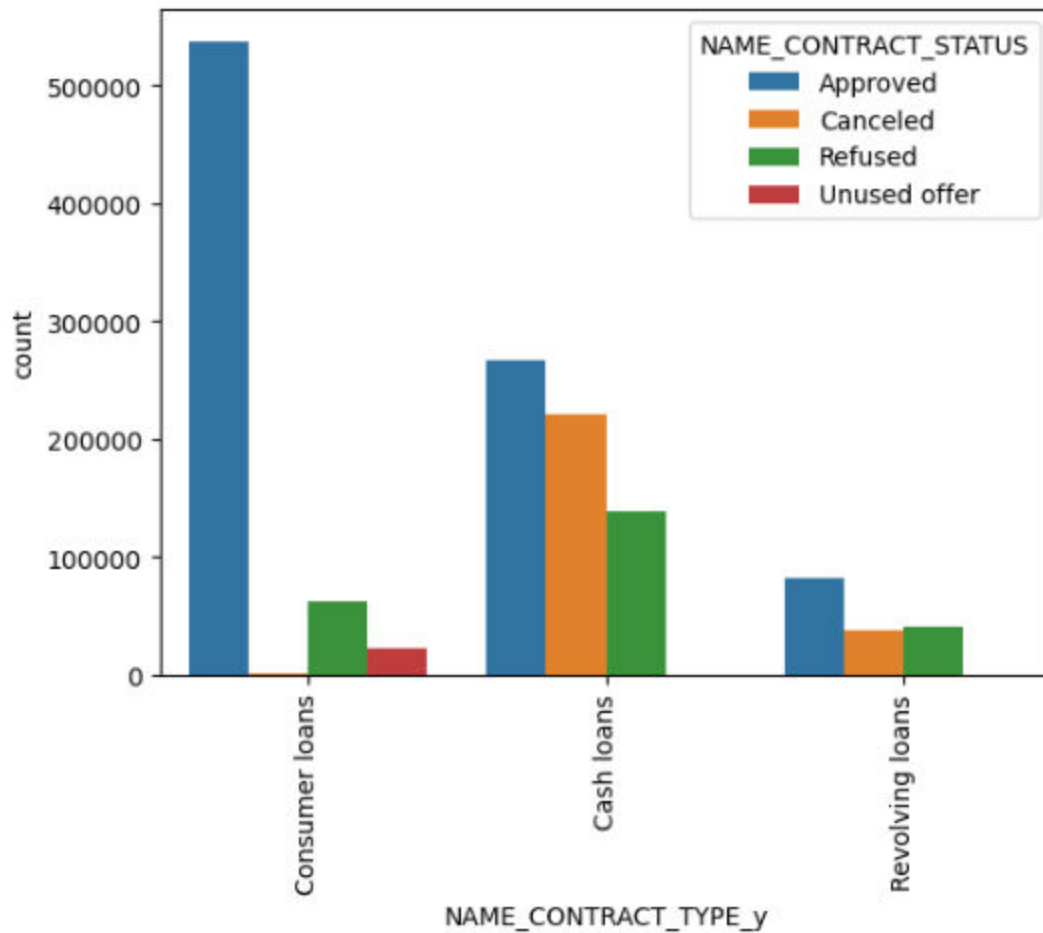
There are more number of consumer loans approved and very less are canceled.
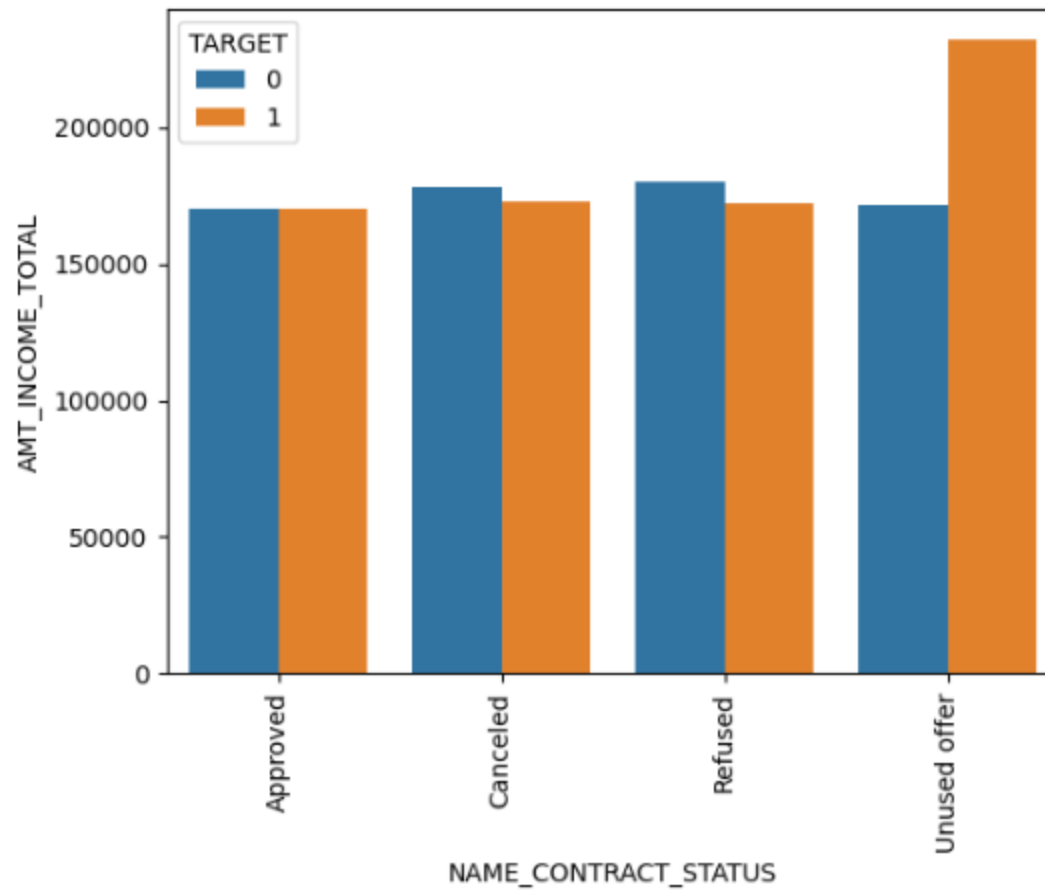
A large number of unused offers are expected to default despite of high income.

DRIVING FACTORS

Banks should focus on the following -

- Income
- Gender
- Family status Housing type
- Age
- Occupation Education
- Contract type
- Income type

CONCLUSION

• Females are more likely to repay loan than males.

• People with high education, academic degree will be less likely to default.

• Married people are likely to repay loan than those who are single, separated or in civil marriage.

• People of age less than 30 are more chances to default than higher age group people.

• People having no children have less chances to default than people having more than two children.

• People with low income and lower secondary education type have more chances to default.

• High income people like Business , Managers, IT staff have more chances to repay.

•People who own a car are having more chances to be a defaulter.
•Although women have high chances to repay loans but womens on maternity leave have less chances of repaying loan because of low income

•Pensioner people and unemployed people have higher chances of default.

•People whose previous application was refused have more chances of default.

•Males who are single and less than 30 years in age are more likely to be in defaulter list.

THANK YOU