

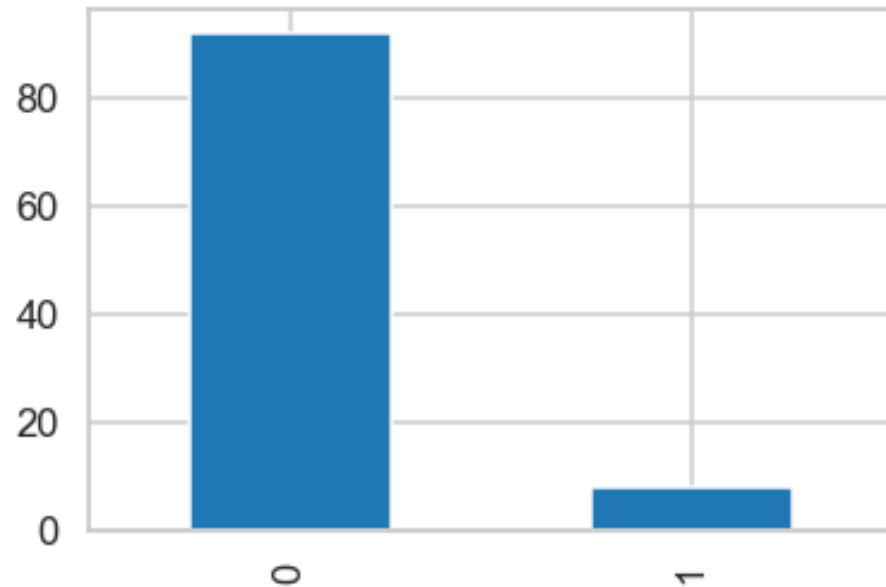
BANK DEFAULT RISK ANALYSIS

SHYAMALA RAJASEKAR

ARAVANAN R.



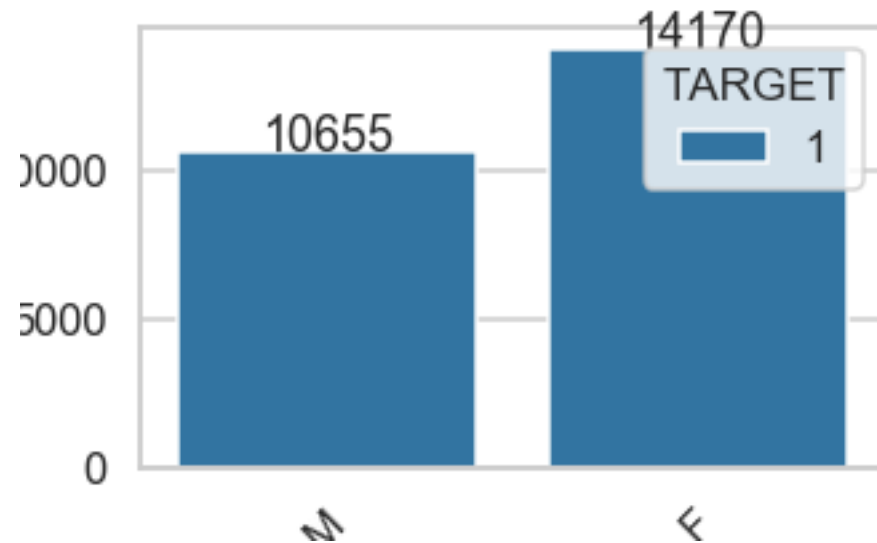
IMBALANCE PERCENTAGE



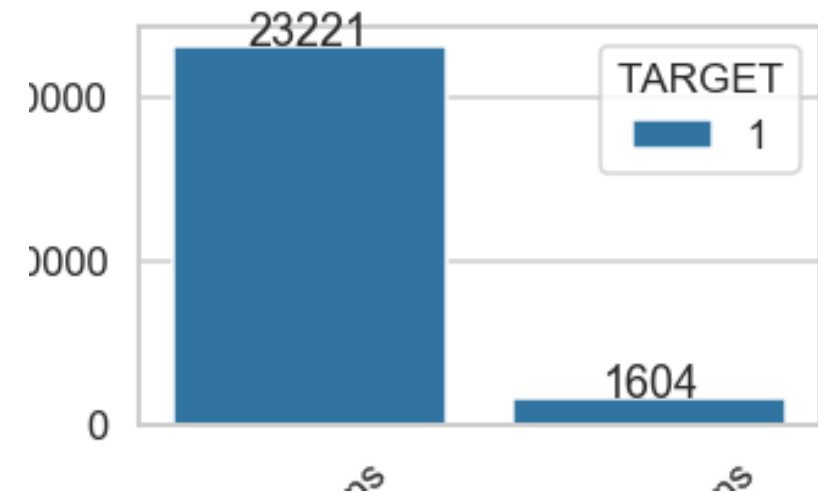
- Dataset is highly imbalanced.
- Default is denoted by 1
- Non-defaulters constitute 91.92%
- Defaulters constitute 8.07%

Univariate Analysis

- Categorical variables for Target 1(Defaulters)



57% of the females defaulted

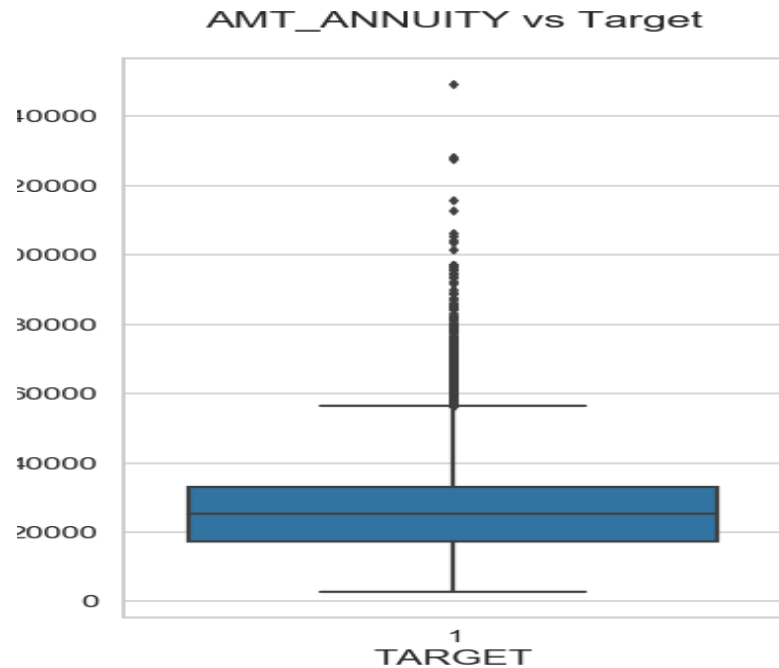


93% of defaulters are those with cash loans.

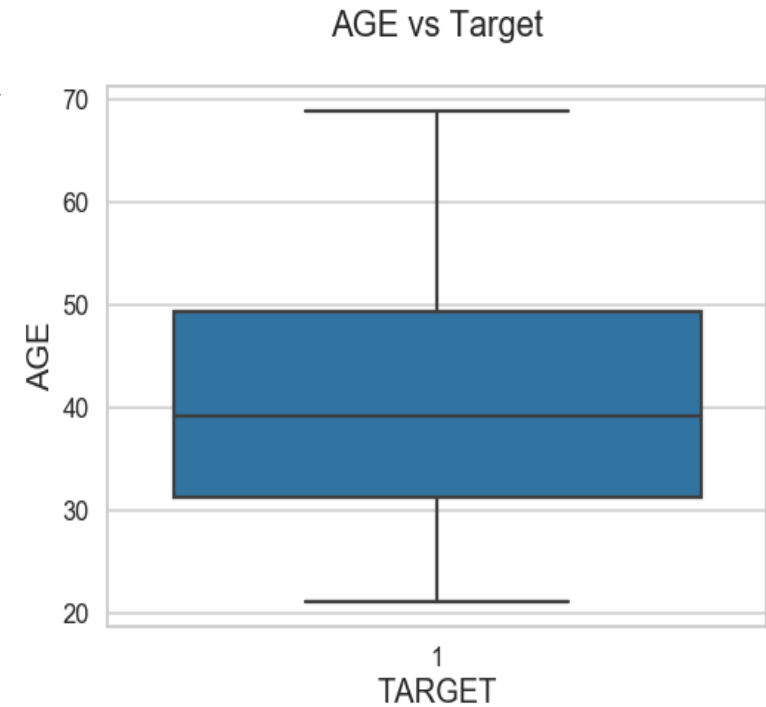
Other Conclusions from categorical variables in Target 1

- 37% of defaulters are females. 14% of men were defaulters.
- 93% of defaulters are those with cash loans.
- 69% of defaulters dont own a car.
- 68% of defaulters donot own a realty
- 78% of defaulters were secondary or secondary special level education.
- 6.9% defaulters lived with parents.
- 31% of defaulters were labourers.
- Around 14% of defaulters had high credit range of 900000 and above.

ANALYSIS OF NUMERICAL VARIABLES FOR TARGET 1



Median Amount annuity of non-defaulters is 25000. Median Amount annuity of defaulters is 30000.



Median value of defaulters is around 40 years.
No difference in median value of defaulters and non-defaulters.

- Median value for Amount goods price for defaulters is around 500000.
- Around 50% of those who defaulted belonged to the Amount income total of 1,35,000.

BIVARIATE ANALYSIS OF TARGET 1



Females in the income range of 25000 to 200000 tend be defaulters.

Correlation Analysis of Target 1 (defaulters) Dataframe

VAR1	VAR2	Correlation
AGE	DAYS_BIRTH	1.00
AMT_GOODS_PRICEx	AMT_APPLICATION	1.00
AMT_GOODS_PRICEx	AMT_CREDITx	0.99
AMT_GOODS_PRICE_	AMT_CREDIT_	0.98
AMT_CREDITx	AMT_APPLICATION	0.98
AMT_CREDITx	AMT_ANNUITYx	0.84
AMT_GOODS_PRICEx	AMT_ANNUITYx	0.84
AMT_APPLICATION	AMT_ANNUITYx	0.82
AMT_ANNUITY_	AMT_CREDIT_	0.75
AMT_GOODS_PRICE_	AMT_ANNUITY_	0.75

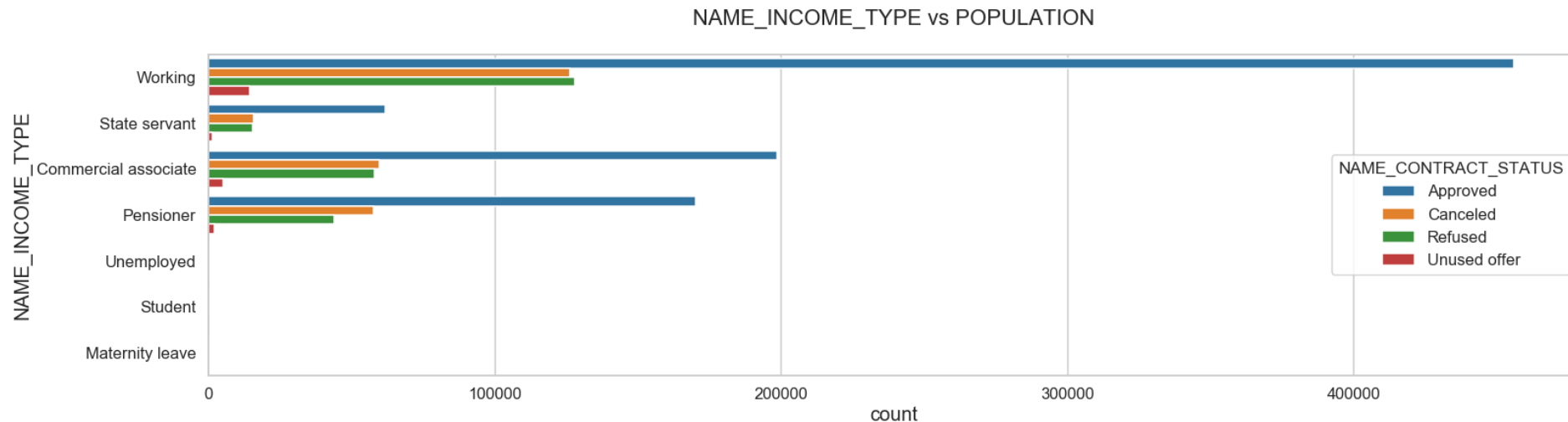
- Amount goods price and amount application of previous application are highly correlated.
- Amount goods price and amount credit of previous application are highly correlated.
- Amount Annuity and credit amount are positively correlated by an amount 0.75.

Correlation Analysis of Target 0 (non-defaulters) Dataframe

VAR1	VAR2	Correlation
AGE	DAYS_BIRTH	1.00
AMT_GOODS_PRICEx	AMT_APPLICATION	1.00
AMT_GOODS_PRICE_	AMT_CREDIT_	0.99
AMT_GOODS_PRICEx	AMT_CREDITx	0.99
AMT_CREDITx	AMT_APPLICATION	0.98
AMT_CREDITx	AMT_ANNUITYx	0.82
AMT_GOODS_PRICEx	AMT_ANNUITYx	0.82
AMT_APPLICATION	AMT_ANNUITYx	0.81
AMT_ANNUITY_	AMT_CREDIT_	0.76
AMT_GOODS_PRICE_	AMT_ANNUITY_	0.76

- Amount goods price of previous application and amount application are directly correlated.
- Amount goods price and amount credit are positively correlated.
- Amount credit of previous years and Amount of annuity are positively correlated by a factor 0.82.

UNIVARIATE ANALYSIS OF MERGED DATAFRAME

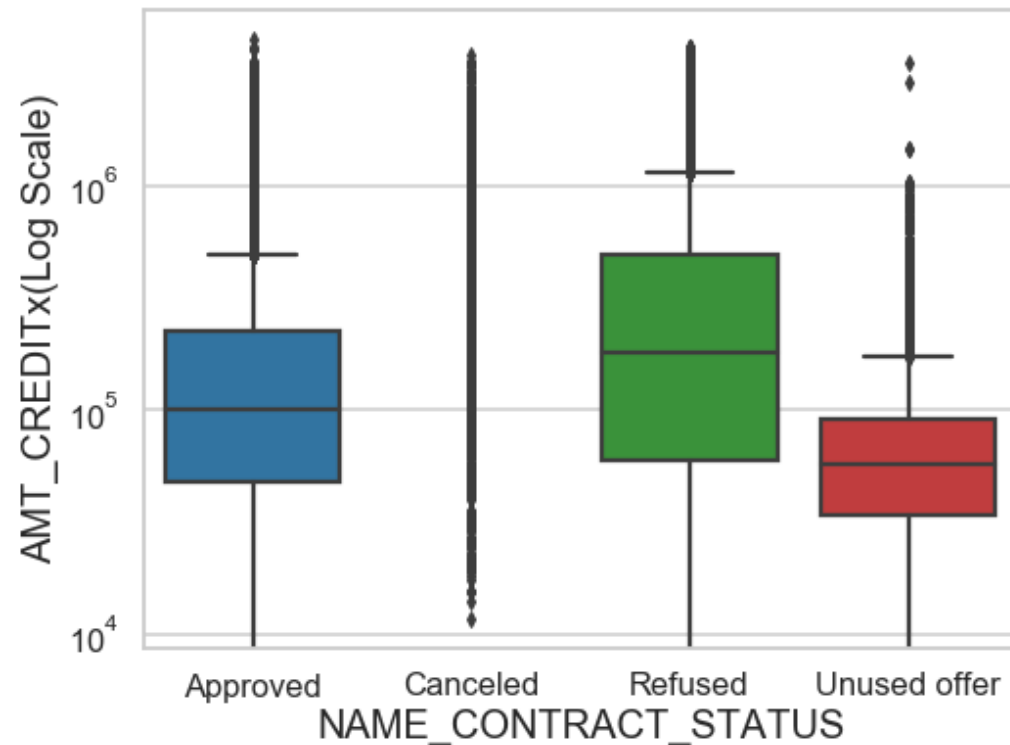


Most of the approved loans belonged to the working class. Followed by commercial associates and pensioners. More than one-fourth of commercial associates loans were rejected. Same is with pensioners.

It is safe to say that most of the approved loans belong to the working class

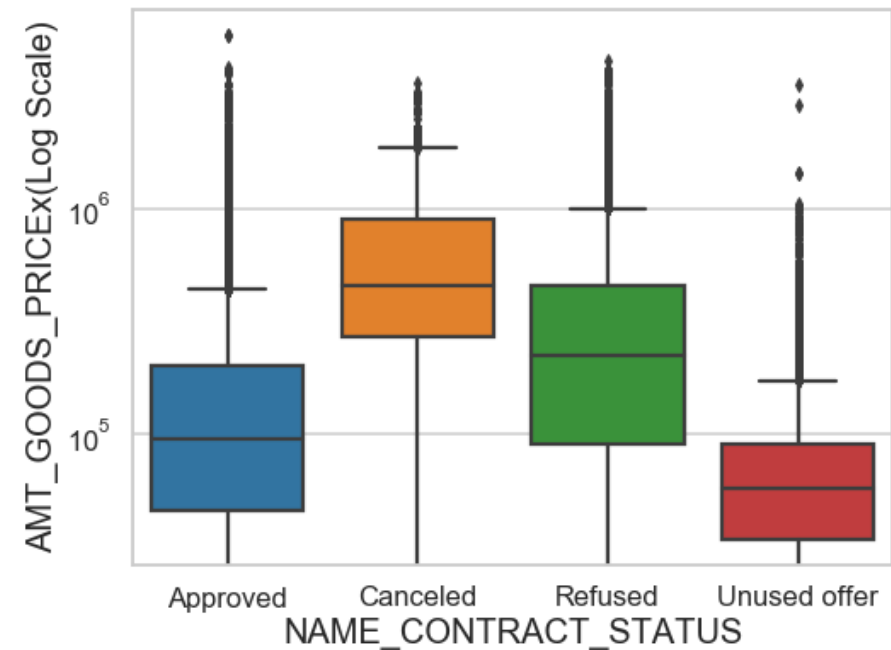
NUMERICAL DATA ANALYSIS OF MERGED DATAFRAME

AMT_CREDITx vs Contract Status



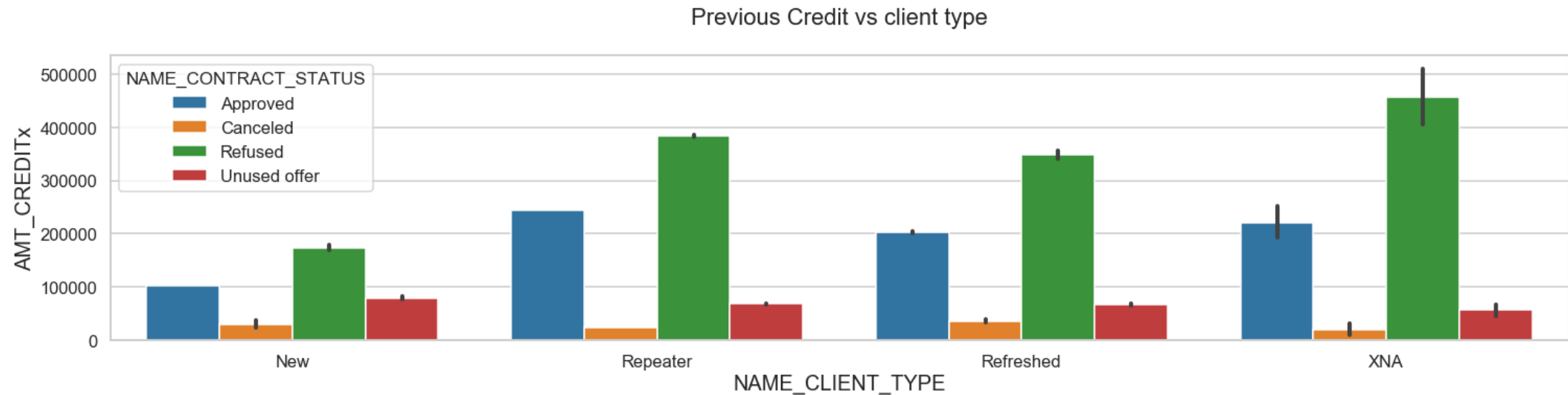
Those with higher previous credits got their loans refused.

AMT_GOODS_PRICEx vs Contract Status



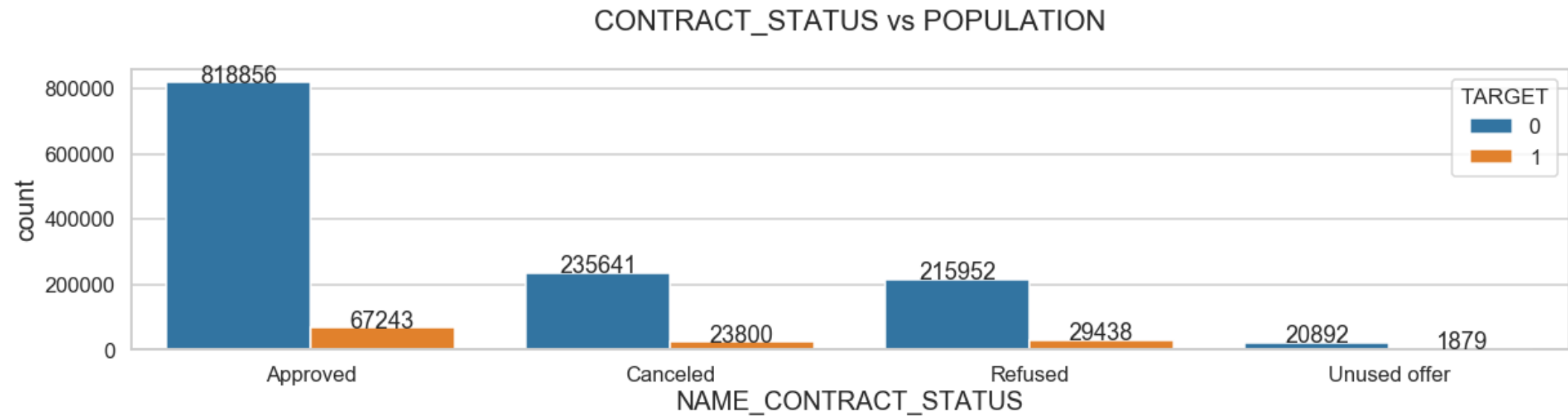
Those with higher goods price got their loans refused or cancelled.

BIVARIATE DATA ANALYSIS OF MERGED DATAFRAME



Repeater with high previous credit are likely to get their loans refused.

CONTRACT VS TARGET



7.5% of approved loans in the past were defaulted. 9.1% of those cancelled defaulted later. 12% of refused loans defaulted.

It is safe to say that refused and cancelled loan takers are likely to default.

CONCLUSIONS DRAWN

1. Credit range of 900000 or above has the maximum no. of approved loans.
2. Most of the approved loans are in the category of 50000 to 225000
3. Most of the approved loans are of the repeater category of clients.
4. POS portfolio of loans were approved the most followed by cash.
5. Most of the approved loans were of the secondary and higher educated folks.

RECOMMENDATION TO THE BANK.

1. Labourers and low skill labourers default often. Hence safe to leave them out.
2. Higher annuity amount for education purposes are likely to be rejected. Purchase of electronic equipment/wedding/gift, education, urgent needs loans with annuity around 40000 are most likely to be approved.
3. Repeater with high previous credit are likely to get their loans refused. Repeaters with high credit are likely to be defaulters.
4. Students with income of around 150000 are likely to get approved loans. A good Target.
5. Those who had high credit and lived in Municipal apartment, Office apartment, co-op apartment loans are likely to be refused. Those who live in municipal apartments and co-op apartments defaulted more. Rented and those who lived with parents defaulted less. Target those who rented or lived with parents.
6. Those with revolving loans and live with parents tend to default less. A Good category.
7. Those with academic degrees defaulted the Least. Target them
8. Single people defaulted often followed by separated people.
9. Those with cash Loans and secondary level education tend to default more. Safe to leave them out.