PRESENTATION

ON CREDIT EDA CASE STUDY

NEHA SAHU

WHY WE NEED TO ANALYSE

Credit risk analysis helps company to get a idea about which kind of customers should have provide loan based on their profile . Because in some cases customer proves defaulter they may not able to pay loan on time & company gets loss in business so it's necessary to do some extra work before providing loan to them .

Which steps needs to follow after getting data

data sourcing & understanding the data

after understanding,data cleaning , checking data qualities ,dealing with outliers

data imbalance , according to data imbalance needs to check correlation

univariate , bivariate analysis

afte analyse first data going to other data & merging both the data

univariate , bivariate analysis

based on the analysis recommendations

coclusion

Approaches & methodology

- **data sourcing & understanding**

- First go through the basic insights about data, what kind of information we have about clients.

- **data cleaning**

- This step is very important in analysis because after getting the necessary informations only we can get more accurtate analysis.

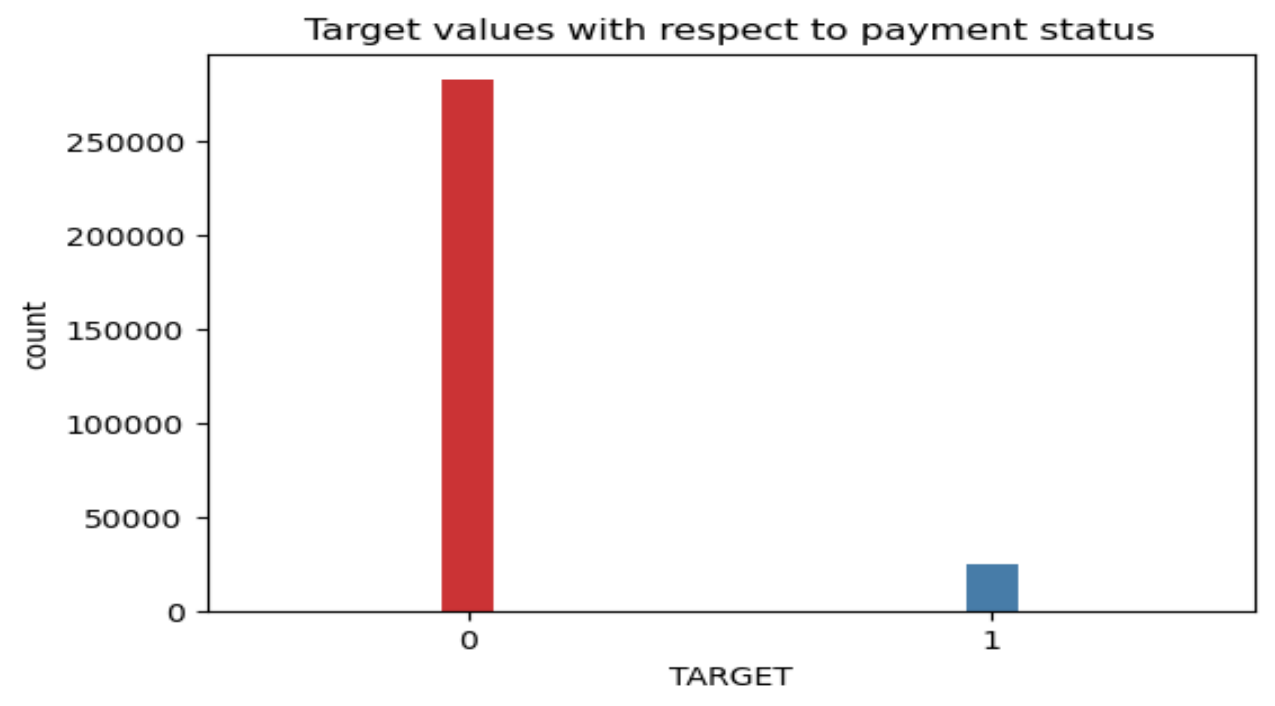
- First checked how much % of missing values in data
- If it's more than 40 then it's ok to drop them & if less than 40 then impute those null values with mean median or mode.
- If data is normal distributed then we can go for mean otherwise median or mode will be useful.

Data imbalance

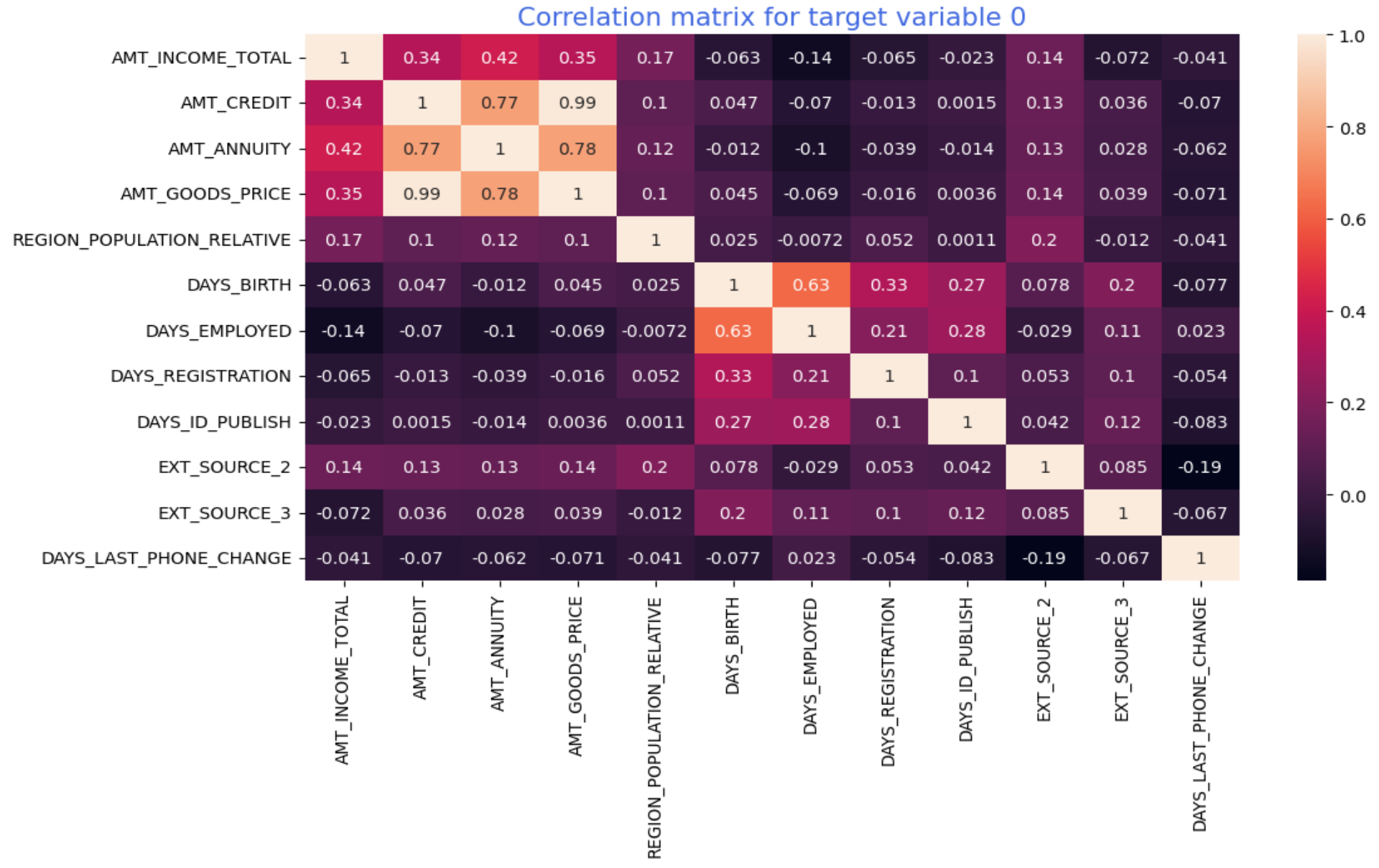
We would have top most column in every analysis which is called target column so based on target column just devide the datset into 2 datasets & let's check how data distributed

target =0 , indicates customer haven't payment difficulties
target=1 , , indicates customer have payment difficulties

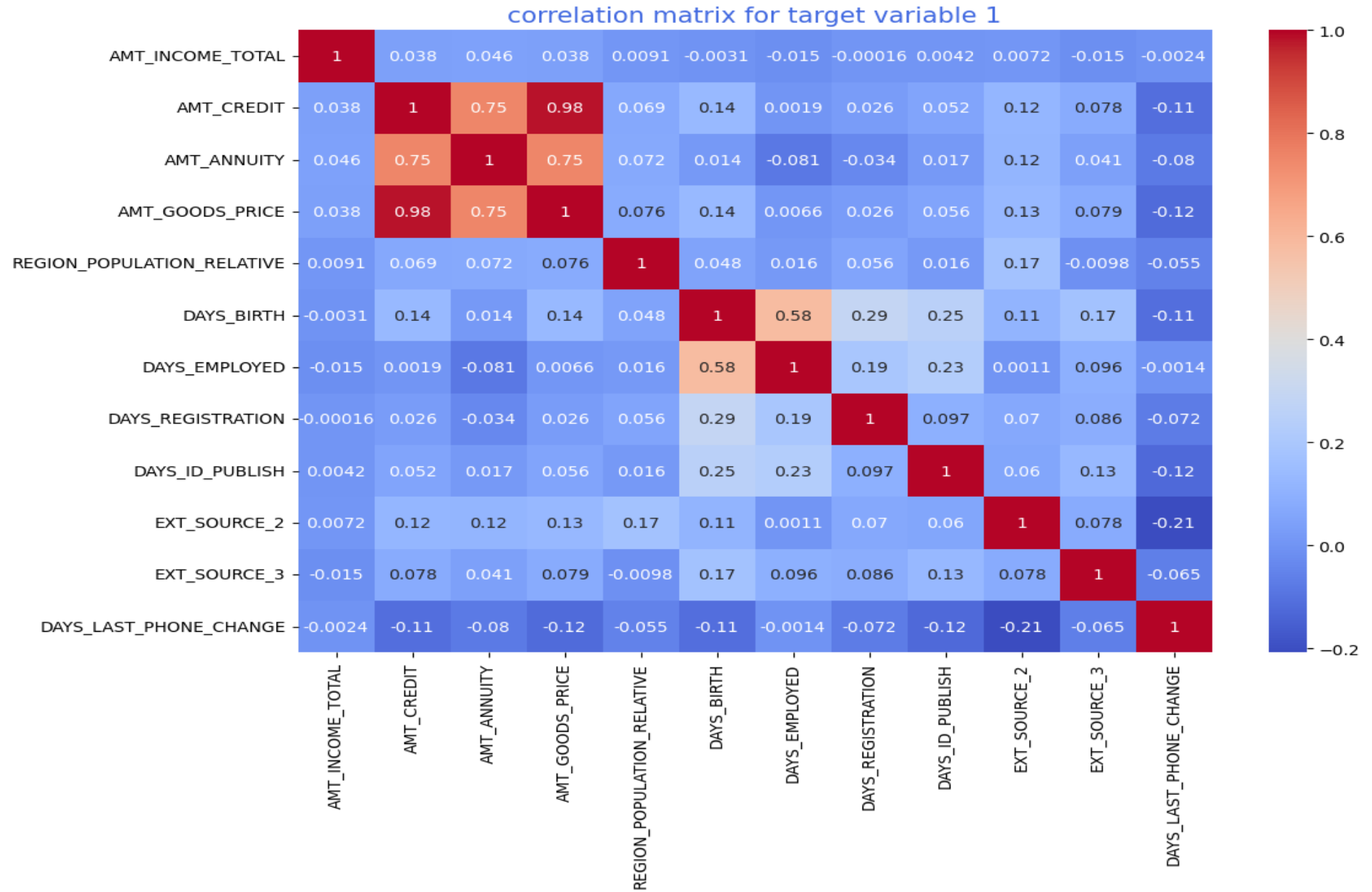
ratio of data imbalance = 11%



Correlation matrix for target 0

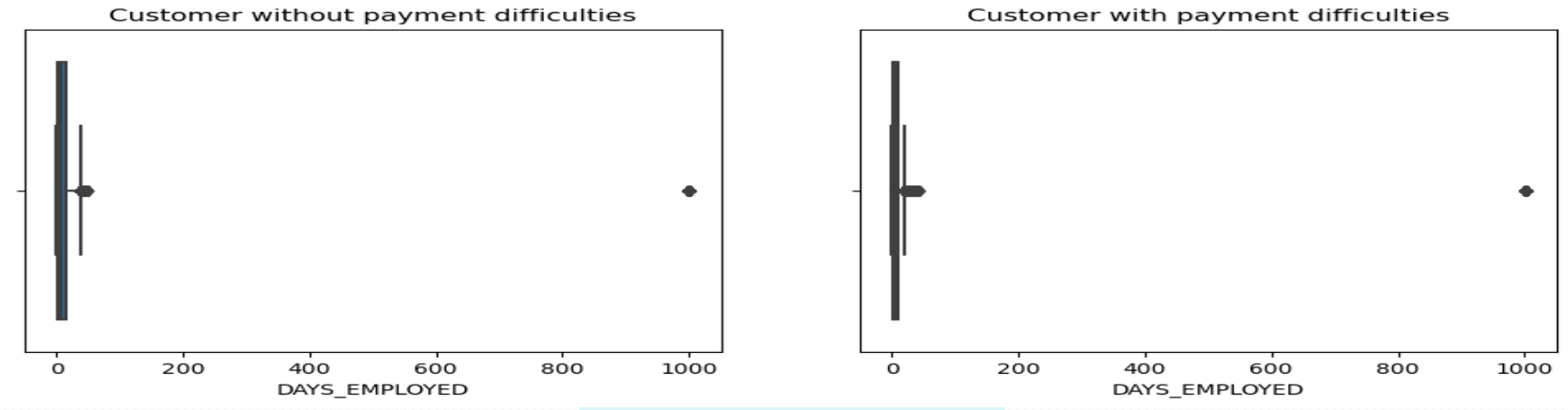


Correlation matrix for target 1

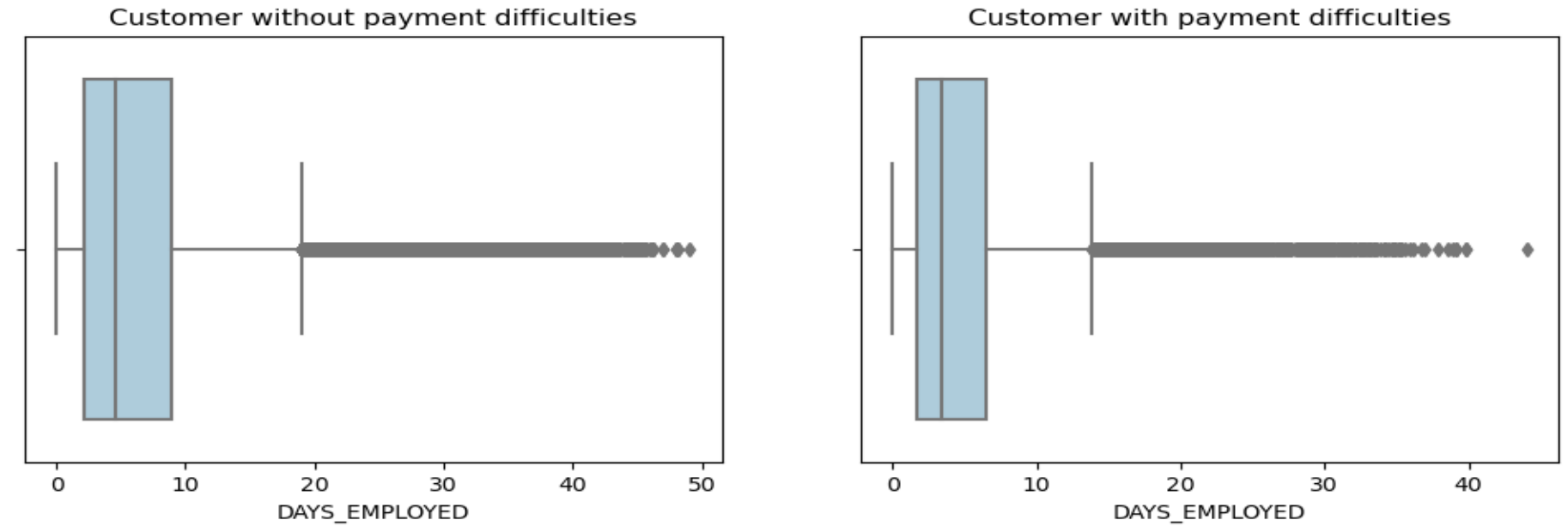


univariate numerical analysis (Dealing with outliers)

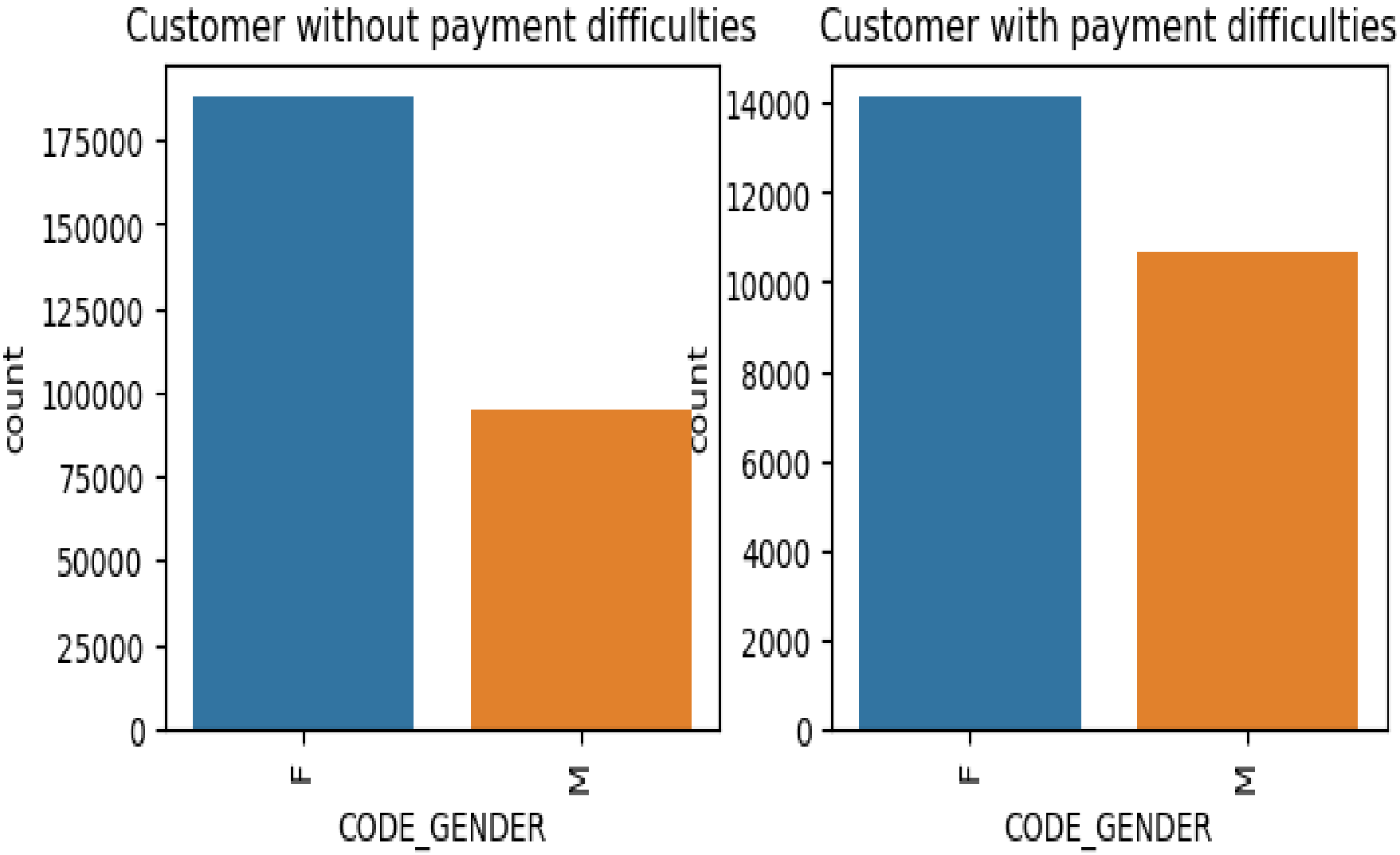
before removing the outlier

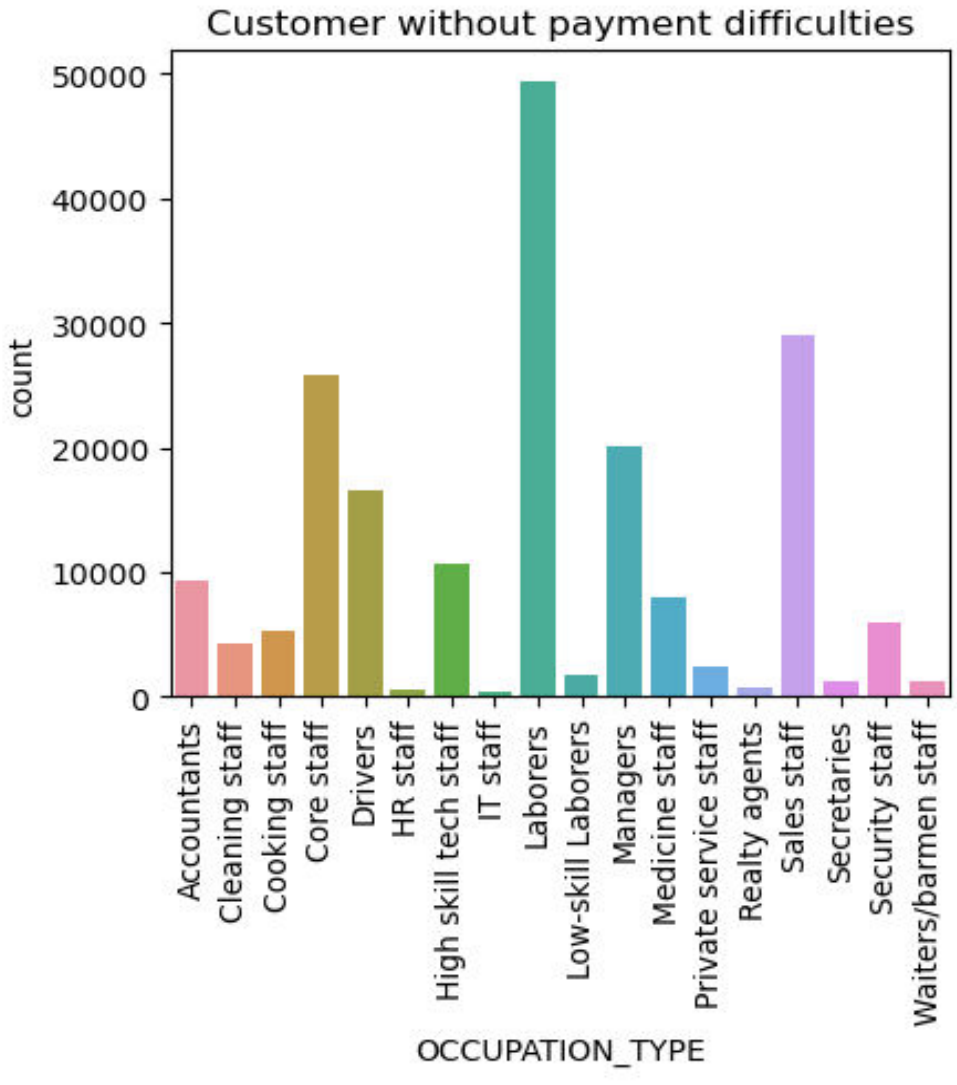


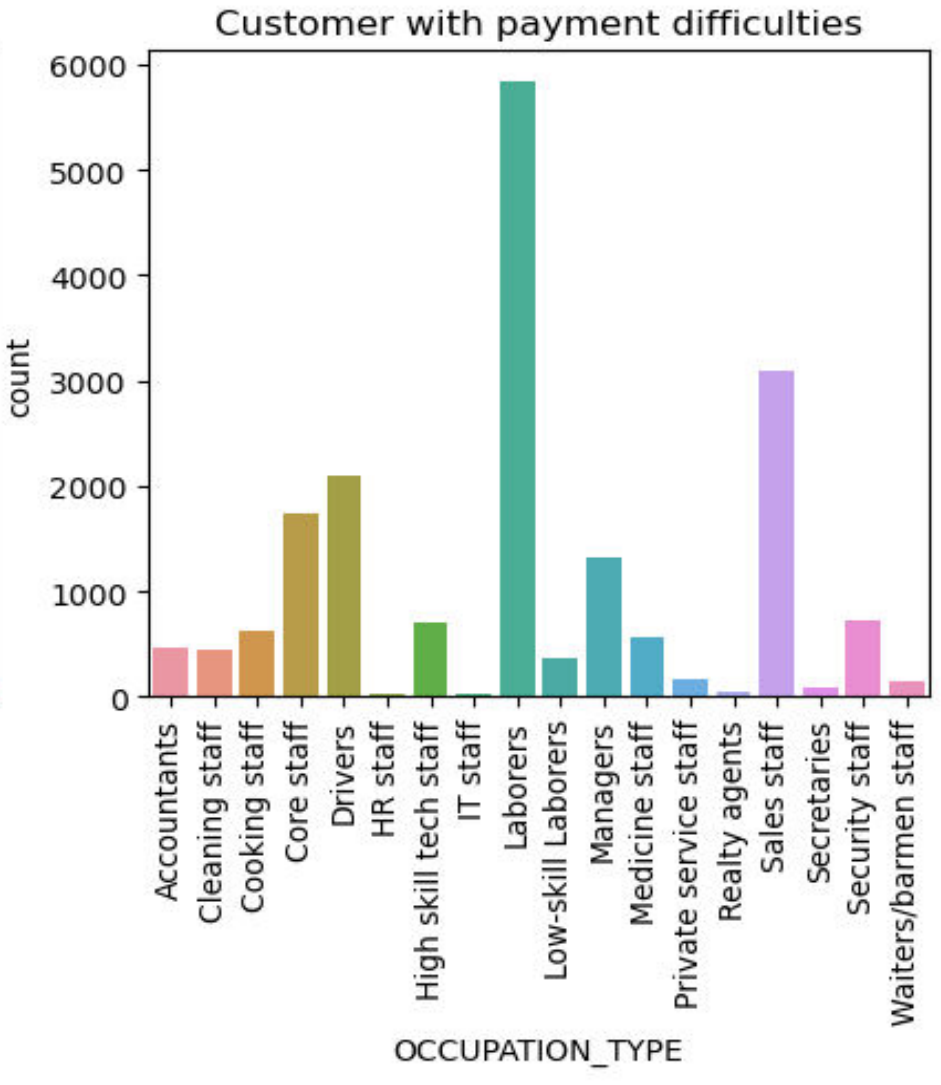
after removing the outlier



Unvariate categorical analysis

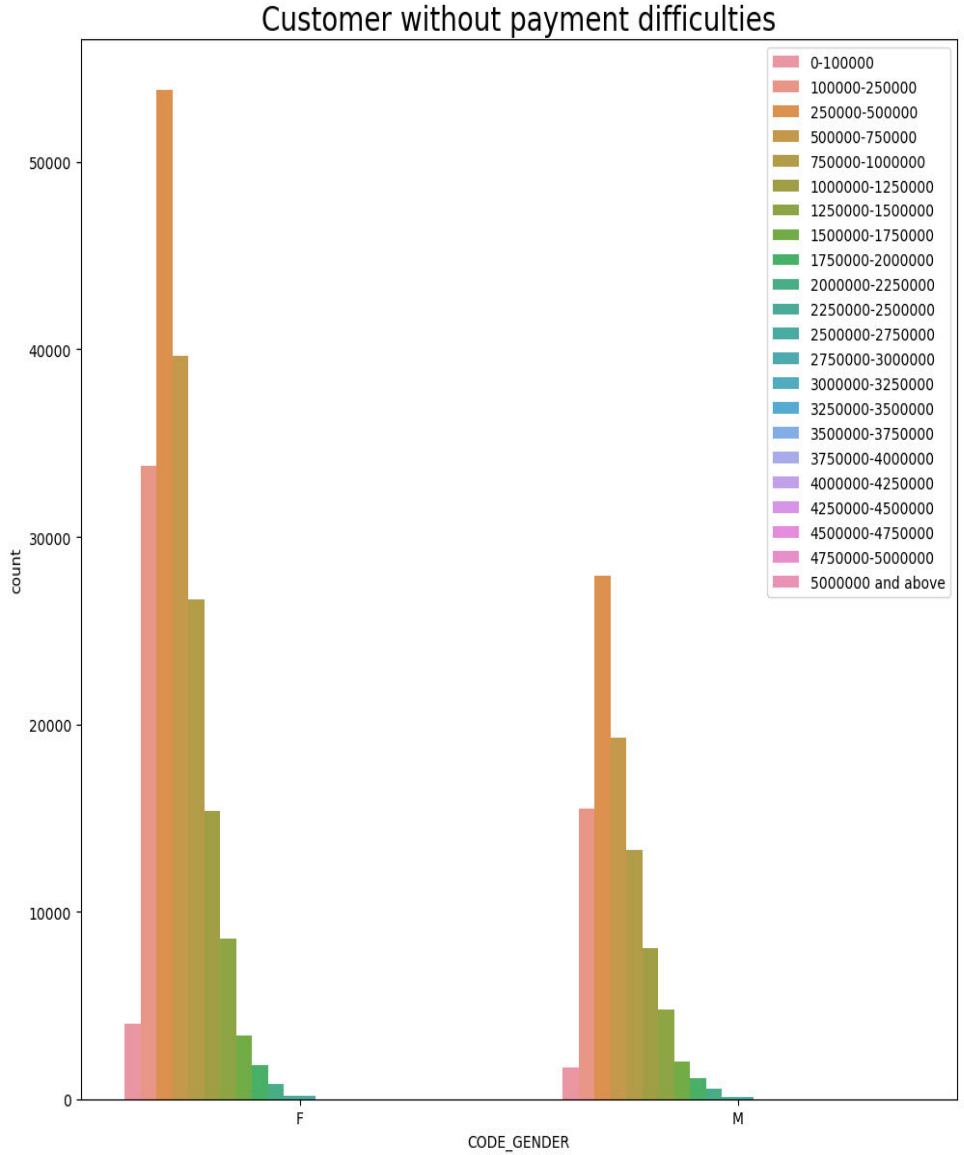


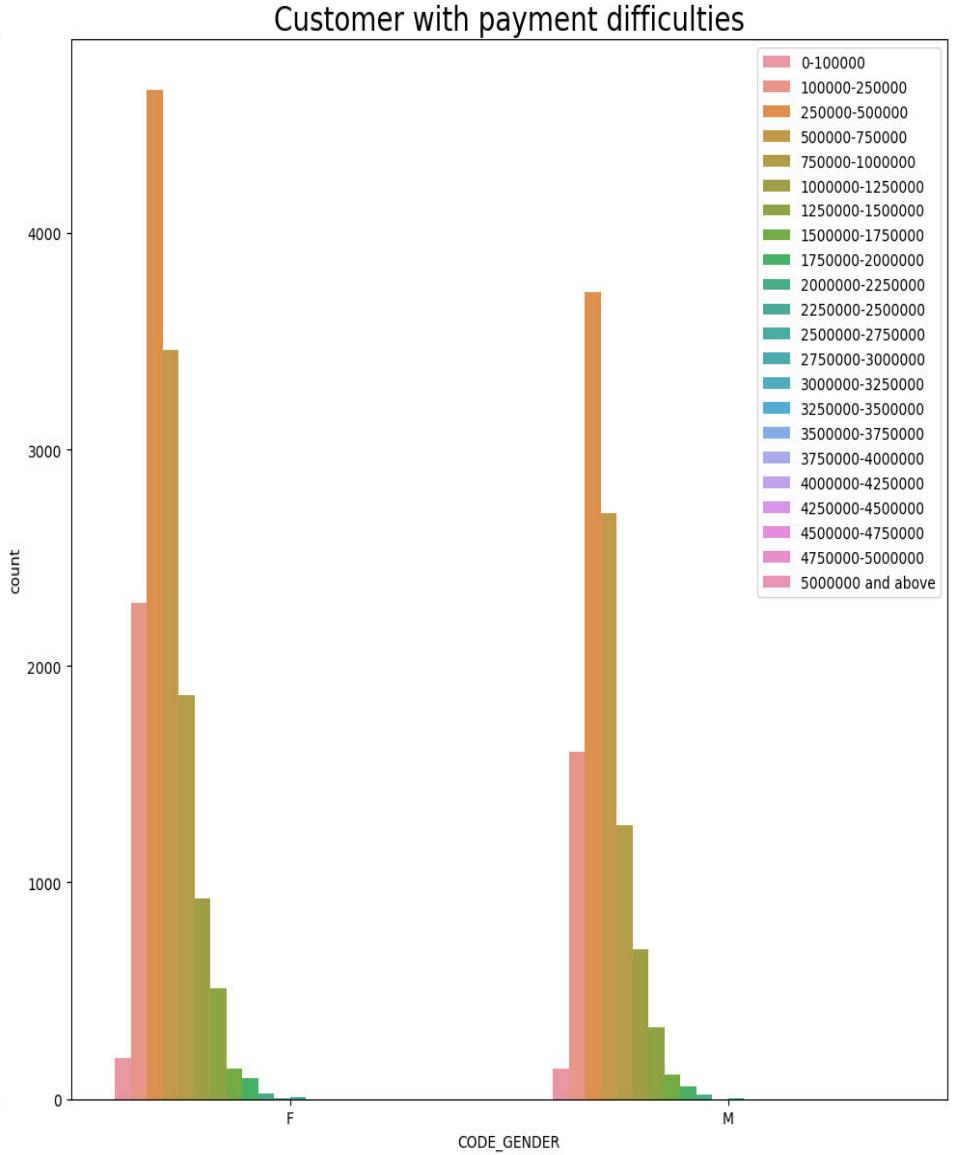

Customer without payment difficulties
count
50000
40000
30000
20000
10000
0
Accountants
Cleaning staff
Cooking staff
Core staff
Drivers
HR staff
High skill tech staff
IT staff
Laborers
Low-skill Laborers
Managers
Medicine staff
Private service staff
Realty agents
Sales staff
Secretaries
Security staff
Waiters/barmen staff
OCCUPATION_TYPE


Customer with payment difficulties
count
6000
5000
4000
3000
2000
1000
0
Accountants
Cleaning staff
Cooking staff
Core staff
Drivers
HR staff
High skill tech staff
IT staff
Laborers
Low-skill Laborers
Managers
Medicine staff
Private service staff
Realty agents
Sales staff
Secretaries
Security staff
Waiters/barmen staff
OCCUPATION_TYPE

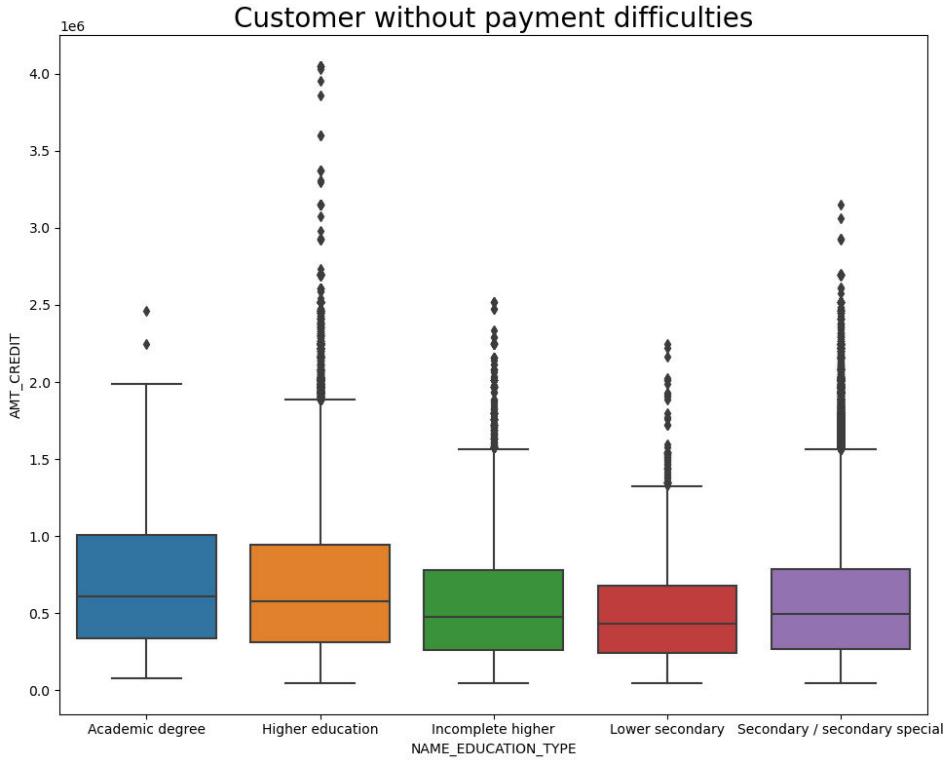
bivariate analysis(categorical-categorical)

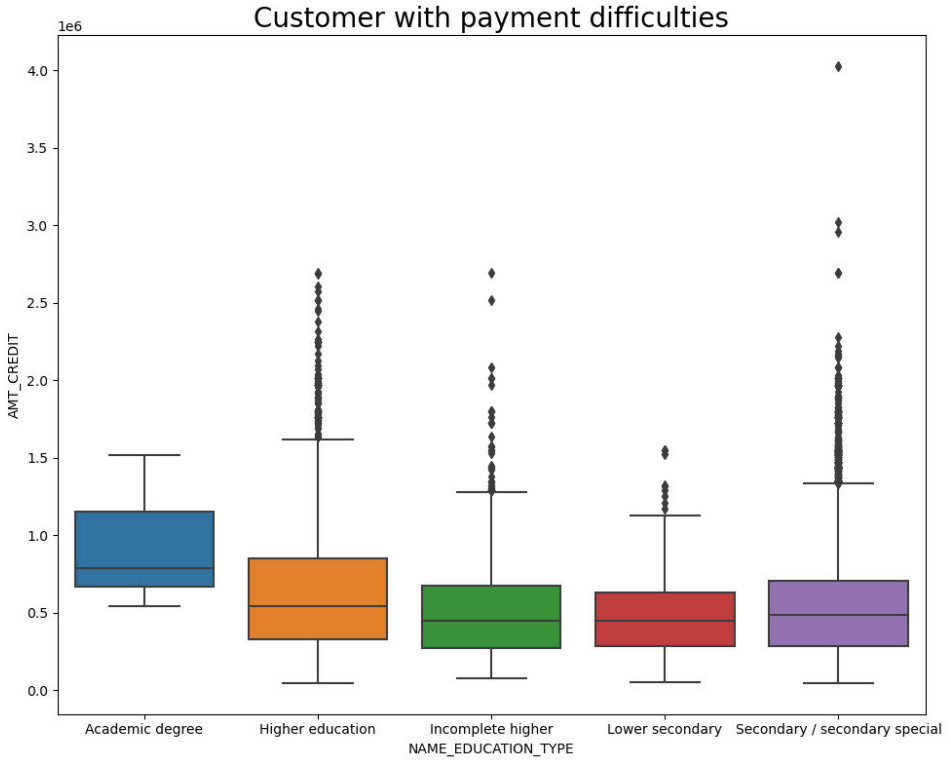
bivariate analysis on CODE_GENDER and Amt_Credit_Range



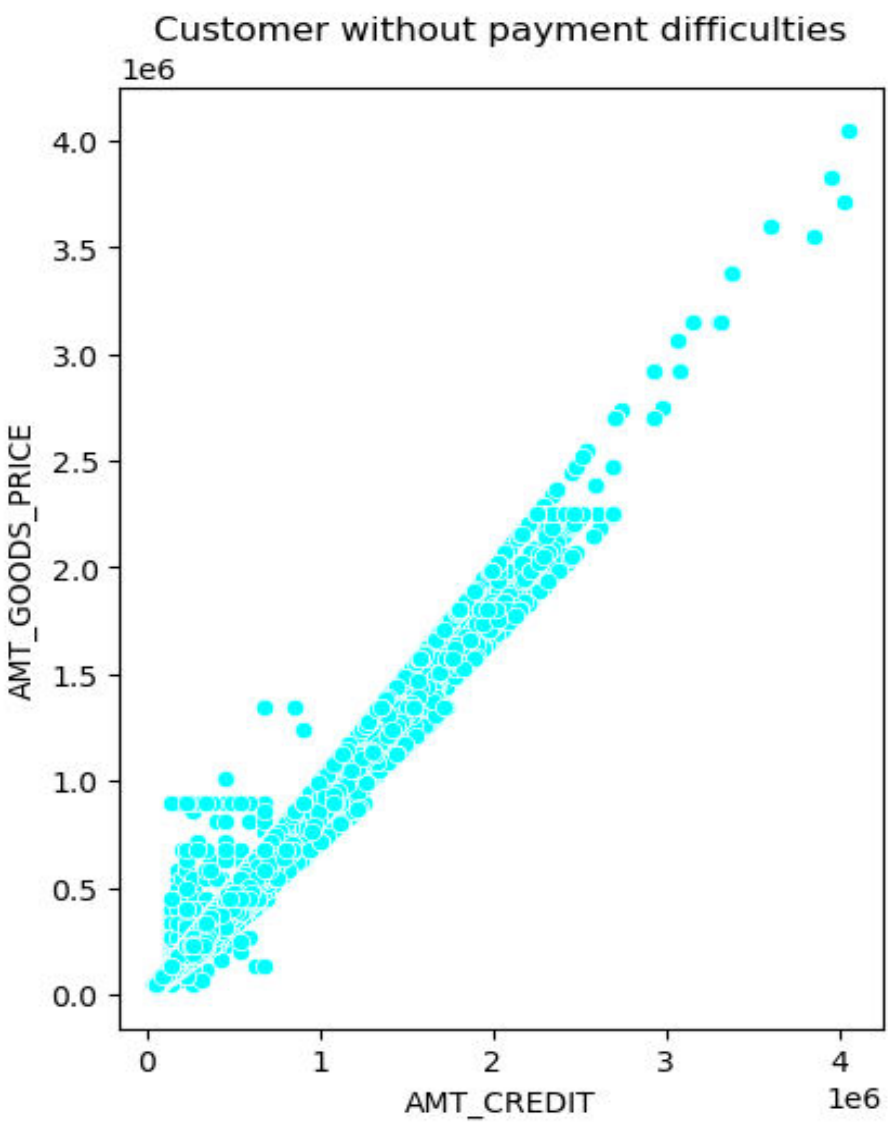


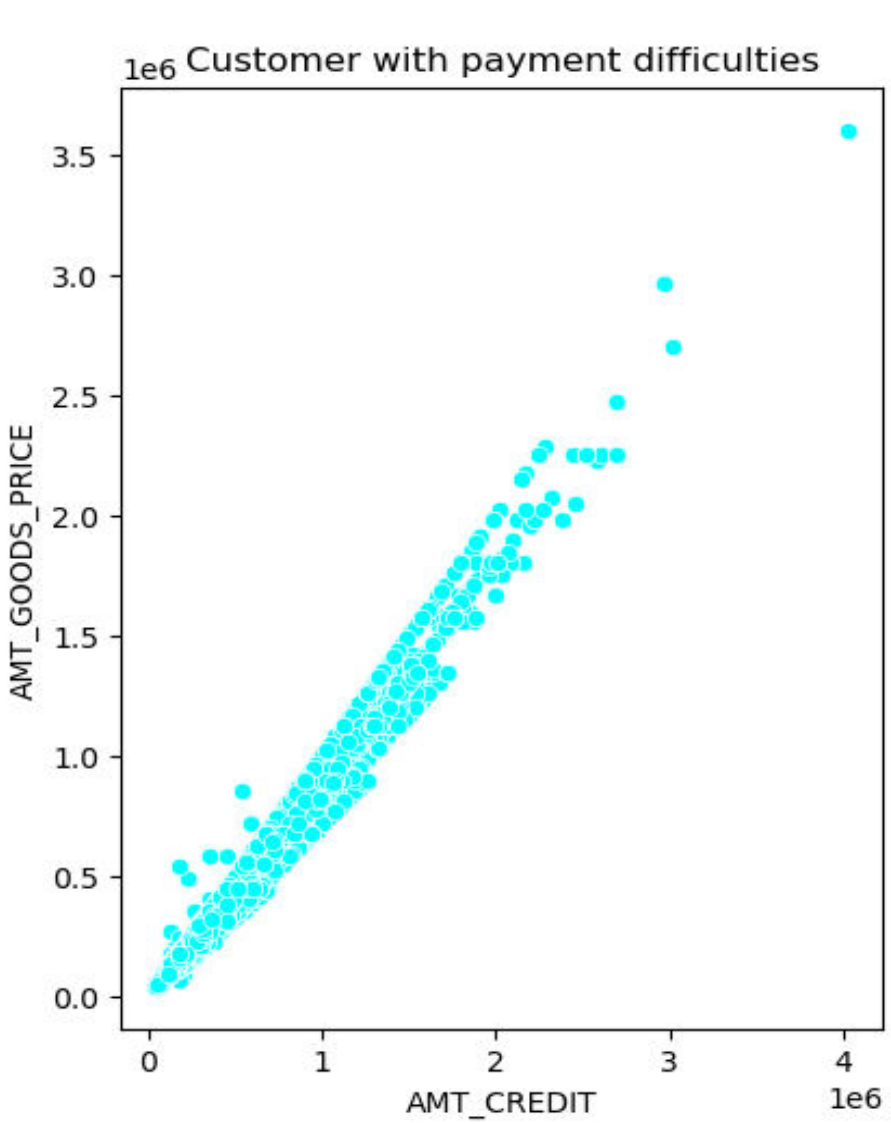
(categorical-numerical bivariate analysis)



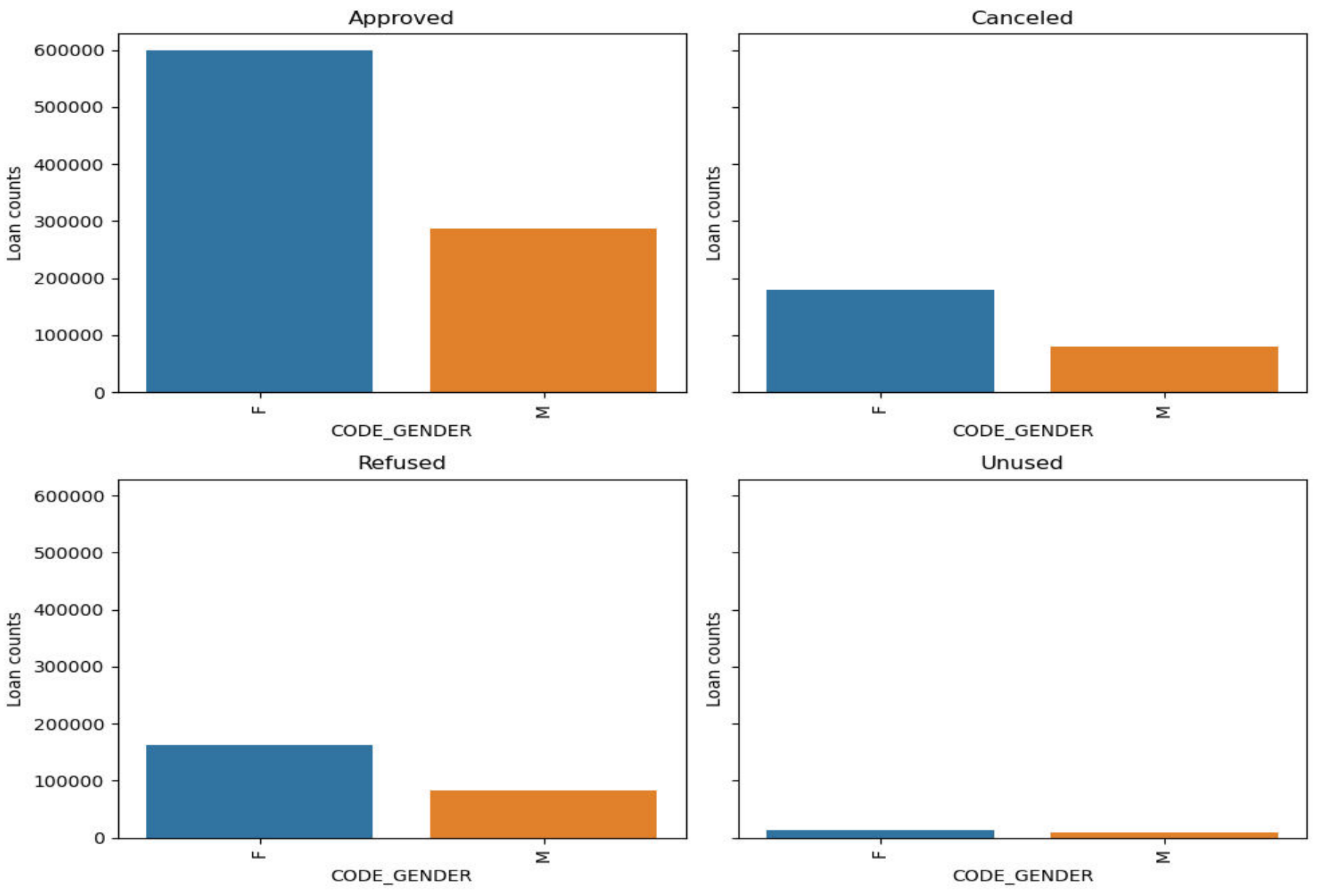


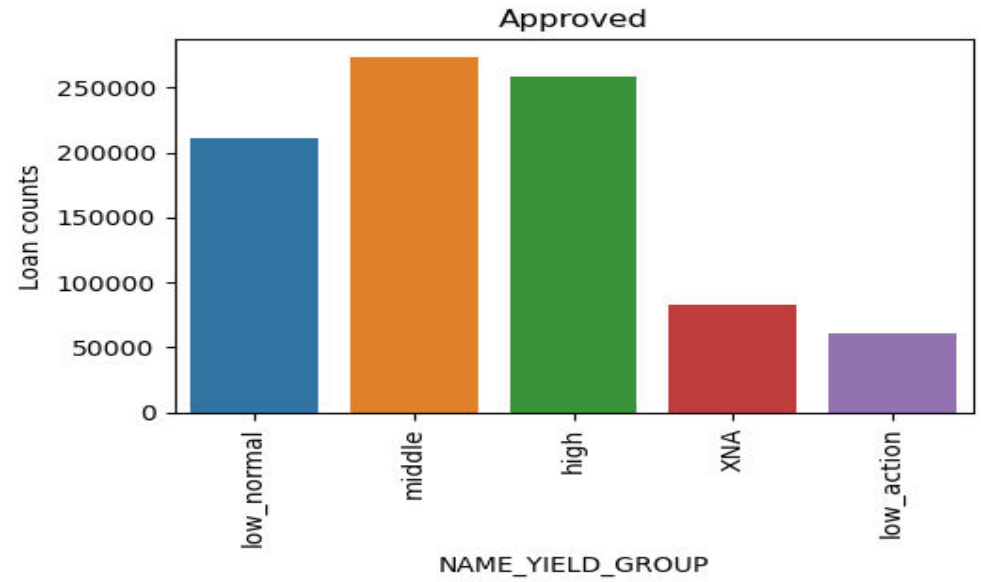
(numerical-numerical bivariate analysis)

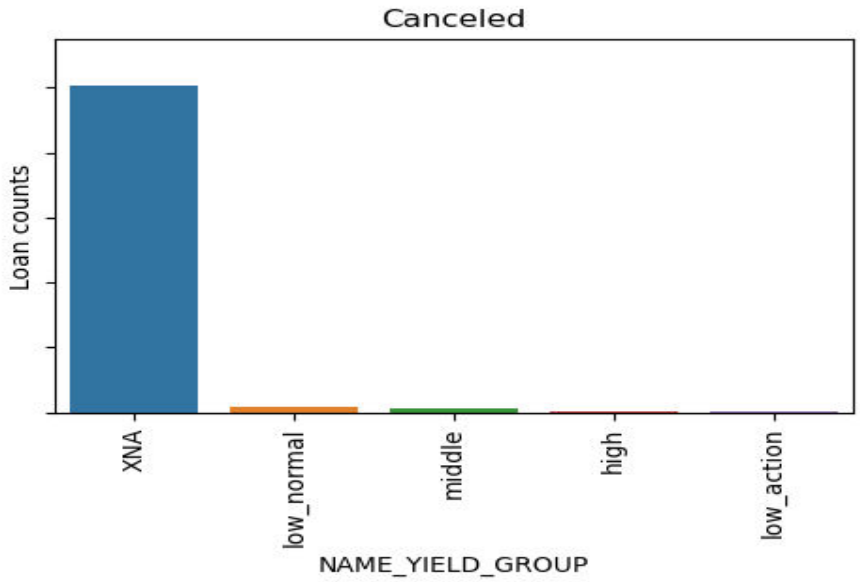


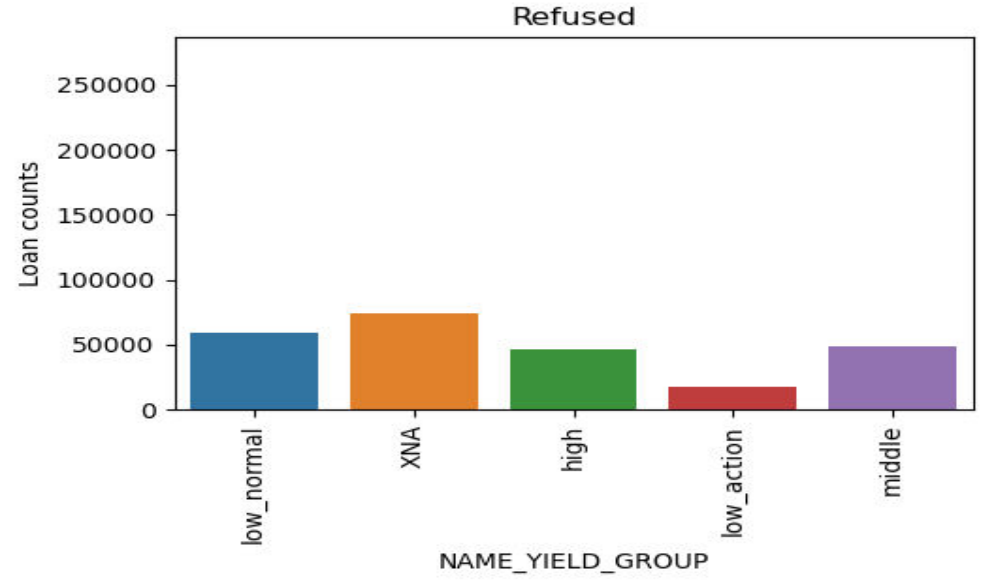


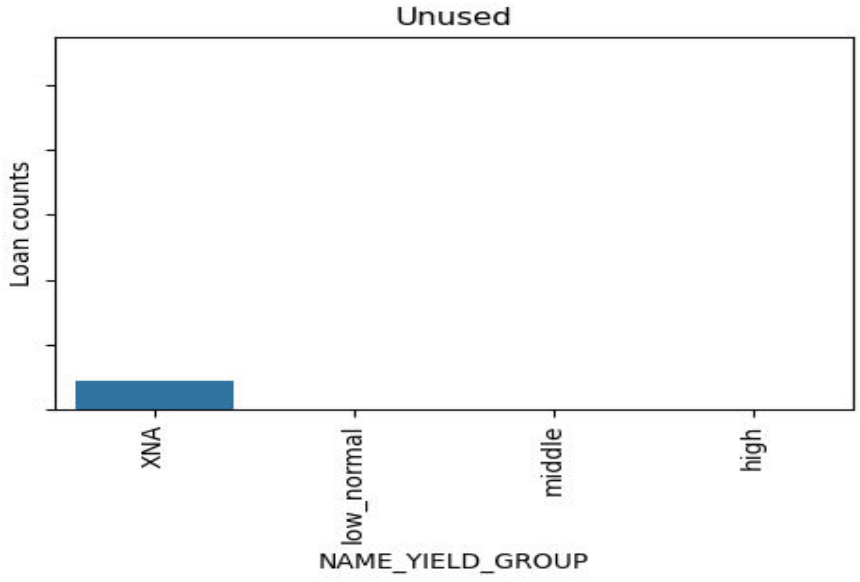
After merging both the dataset final outcomes regarding loan approval

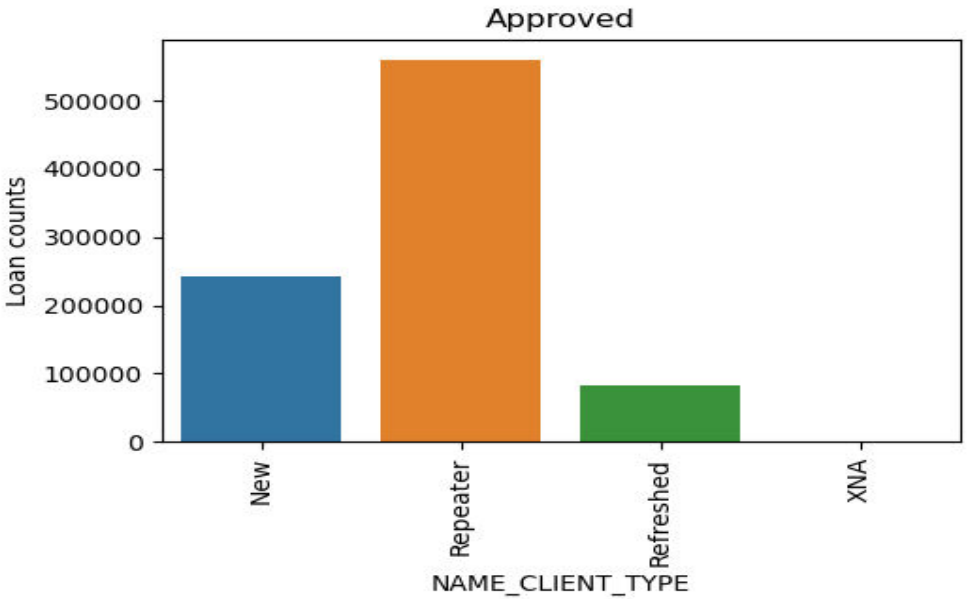


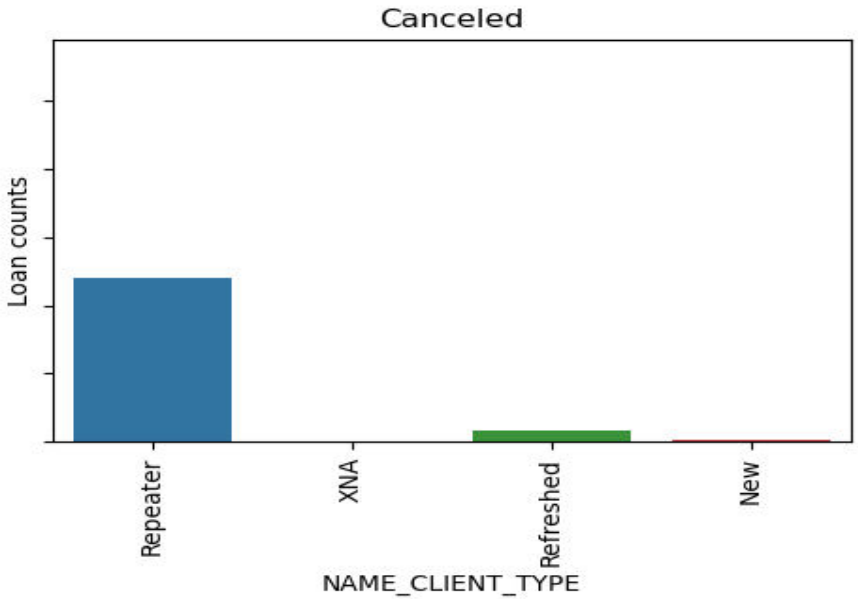

Approved
Loan counts
250000
200000
150000
100000
50000
0
low_normal
middle
high
XNA
low_action
NAME_YIELD_GROUP

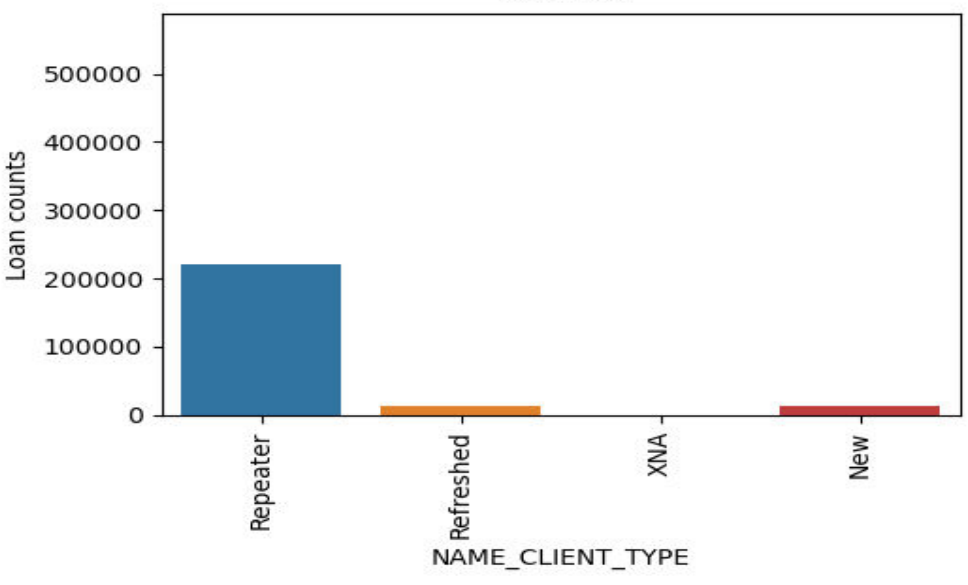

Canceled
Loan counts
XNA
low_normal
middle
high
low_action
NAME_YIELD_GROUP

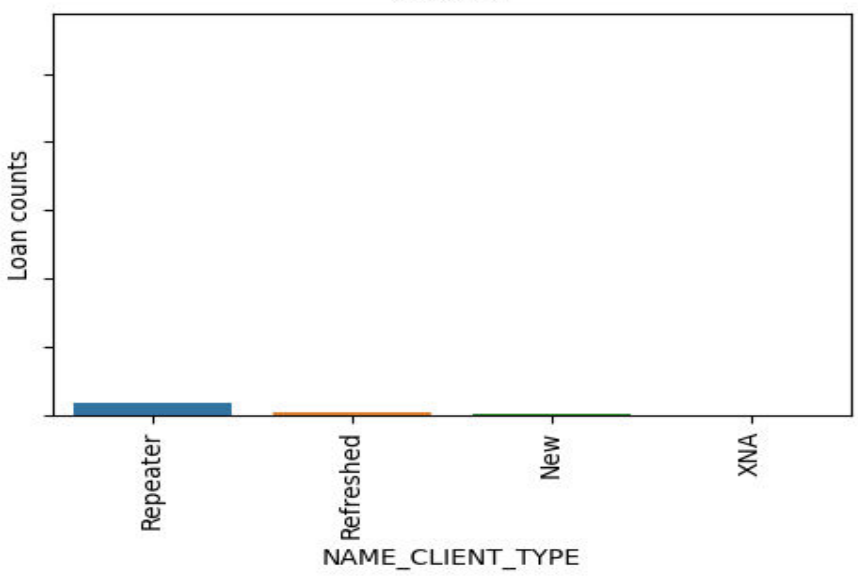

Refused
Loan counts
250000
200000
150000
100000
50000
0
low_normal
XNA
high
low_action
middle
NAME_YIELD_GROUP


Unused
Loan counts
XNA
low_normal
middle
high
NAME_YIELD_GROUP


Approved
Loan counts
500000
400000
300000
200000
100000
0
New
Repeater
Refreshed
XNA
NAME_CLIENT_TYPE


Canceled
Loan counts
Repeater
XNA
Refreshed
New
NAME_CLIENT_TYPE


Refused
Loan counts
500000
400000
300000
200000
100000
0
Repeater
Refreshed
XNA
New
NAME_CLIENT_TYPE


Unused
Loan counts
Repeater
Refreshed
New
XNA
NAME_CLIENT_TYPE

Recommendation

- Female customer can provide as they are less likely to be defaulter.
- Working persons (male & female).
- Married ones can provide.
- Who have their own houses/apartments can get loan.
- Age group of middle , high & normal can provide loan.
- Ones who alrrady taken loan before from the company they also can get loan.
- Refreshed & widow who have unused before approved loan.

Thankyou