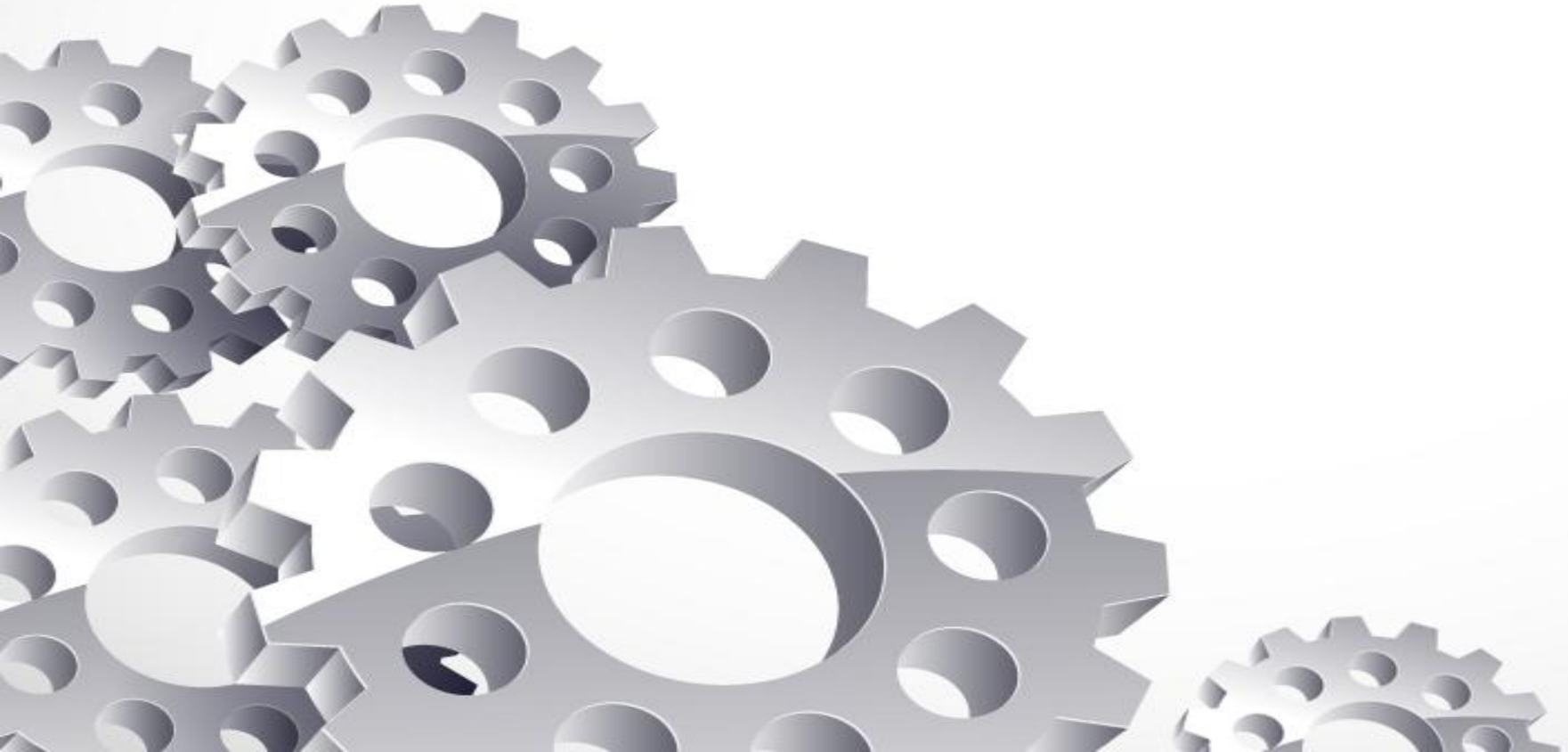
Presentation on credit EDA Case study



Presented by:-
Mayank Jain

Problem Statement

Analyze the risk involved of the Banks/NBFC for loan

- As if they will give the loan there are chances of payment default which may lead to loss for the banks/NBFC
- If they don’t give the loan it will lead to their revenue loss which could also lead to closure of business.

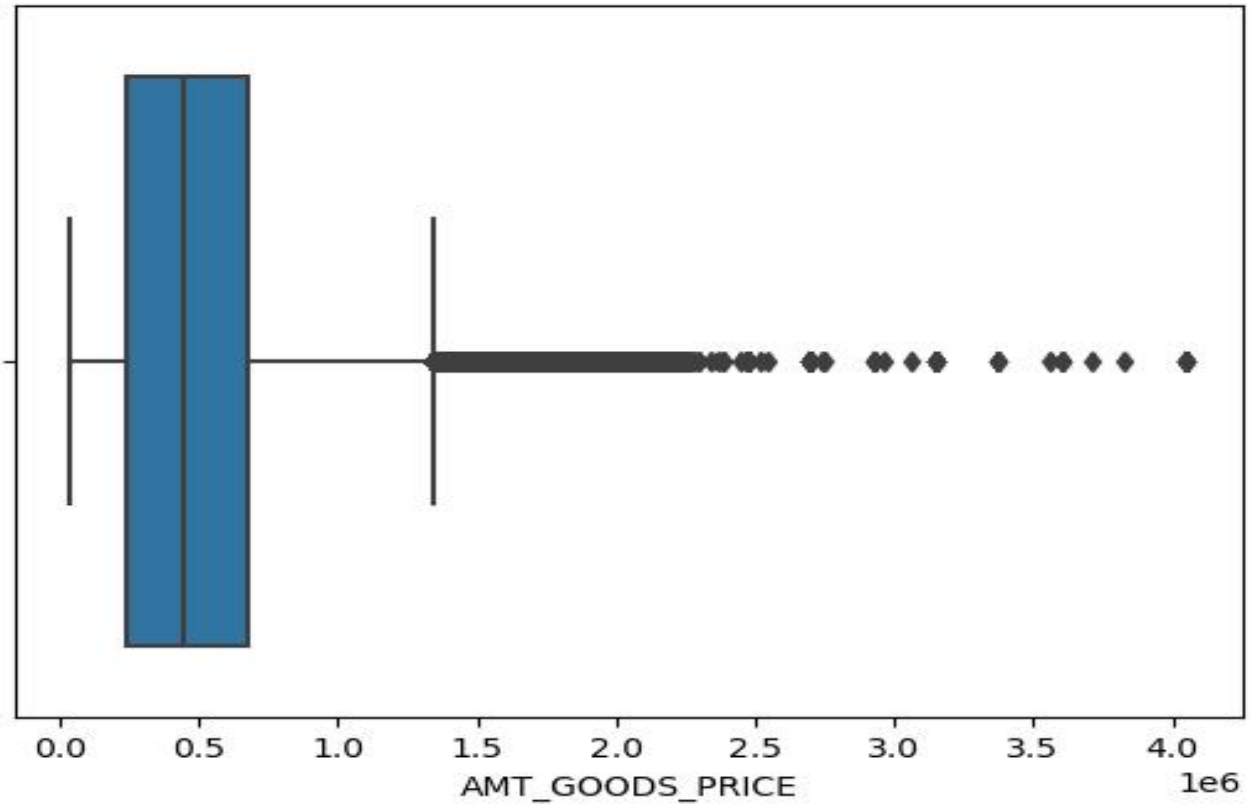
Steps involed in analysis

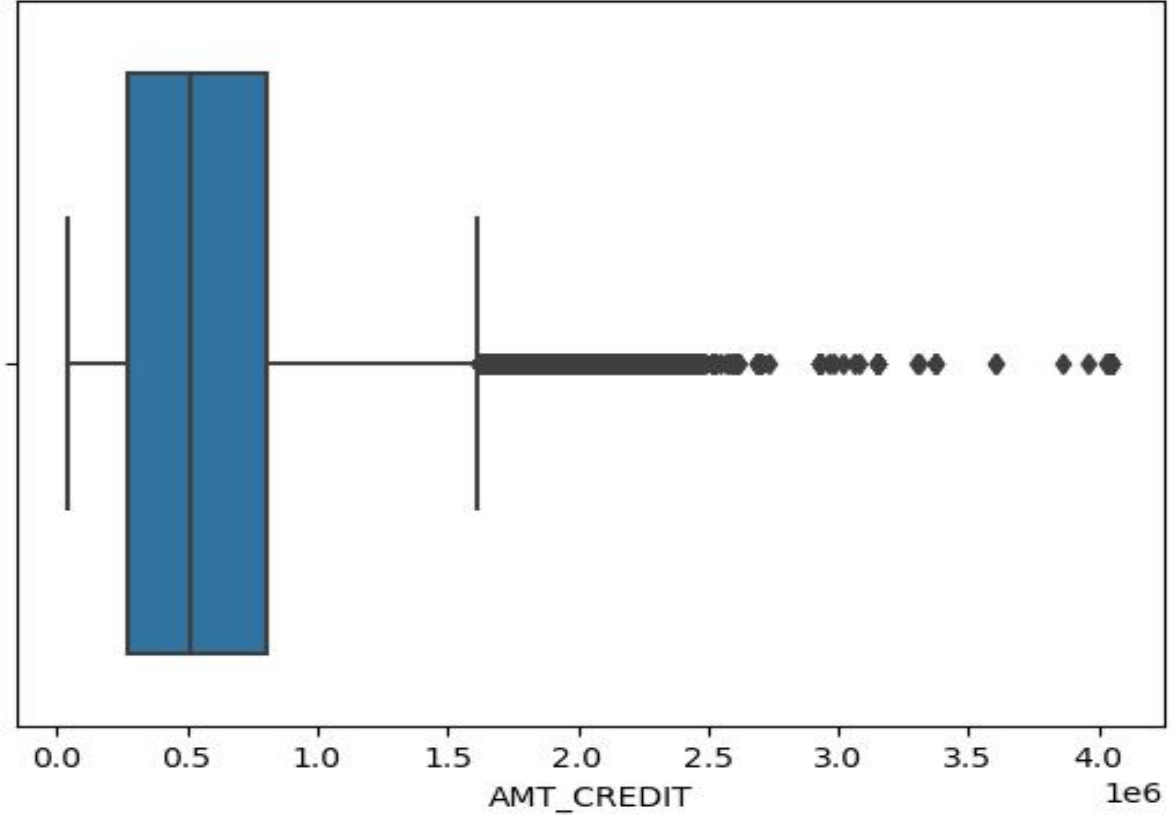
1. Import the Application Data
2. Check the structure of the data.
3. Missing Value Check
4. Dropping the column with more than 40% of missing values
5. Handling missing values
6. Divide the data into Numerical, Categorical and Extra list
7. Dropping the extra column from the data set.
8. Check for outliers (Not taking action on the outliers but the outliers can be dropped or edited as per data requirement)
9. Analyzing the dataset Unilateral, Bilateral and Multilateral analyses
10. Repeating all the above steps with Previous application dataset
11. Comparision analysis between Numerical and Categorical list of both the dataset with each other.

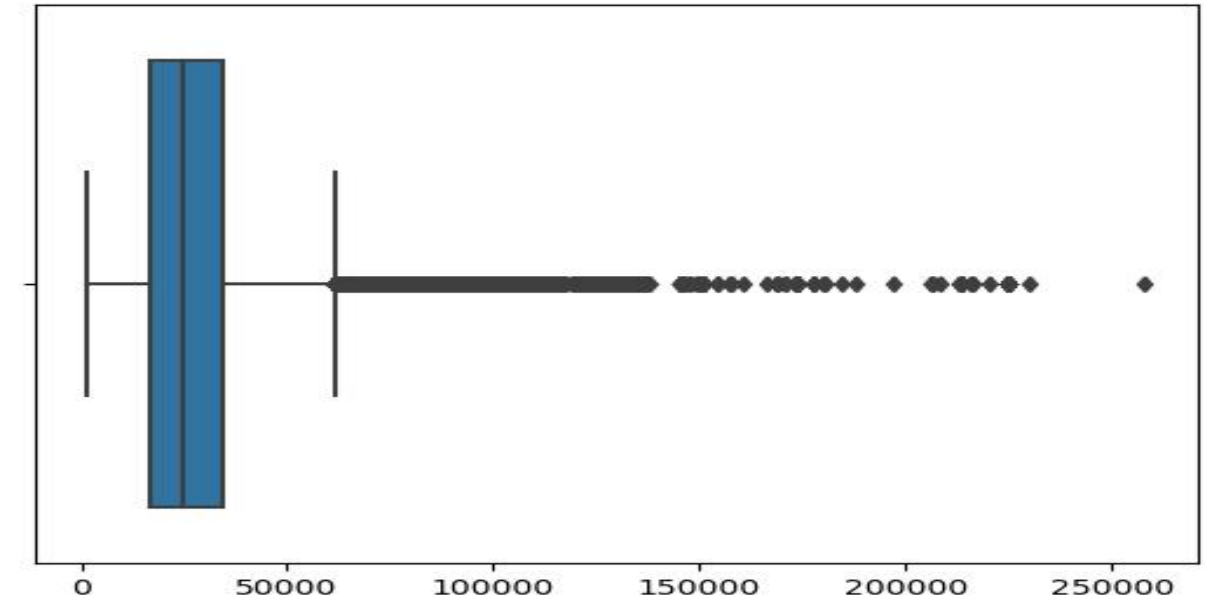
Assumptions

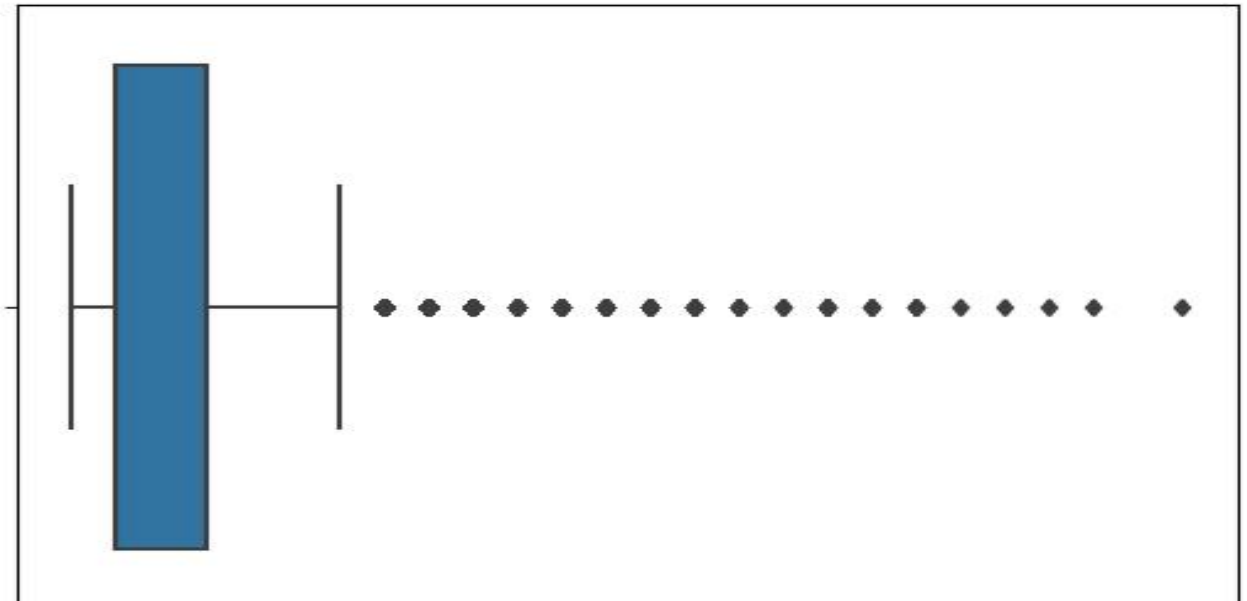
1. Assumed that the columns with more than 40% of missing values will affect our analysis. So dropped the columns.
2. No action on outliers as it will affect our analysis
3. Dropped the columns considering that they are not in use for analyzing the data

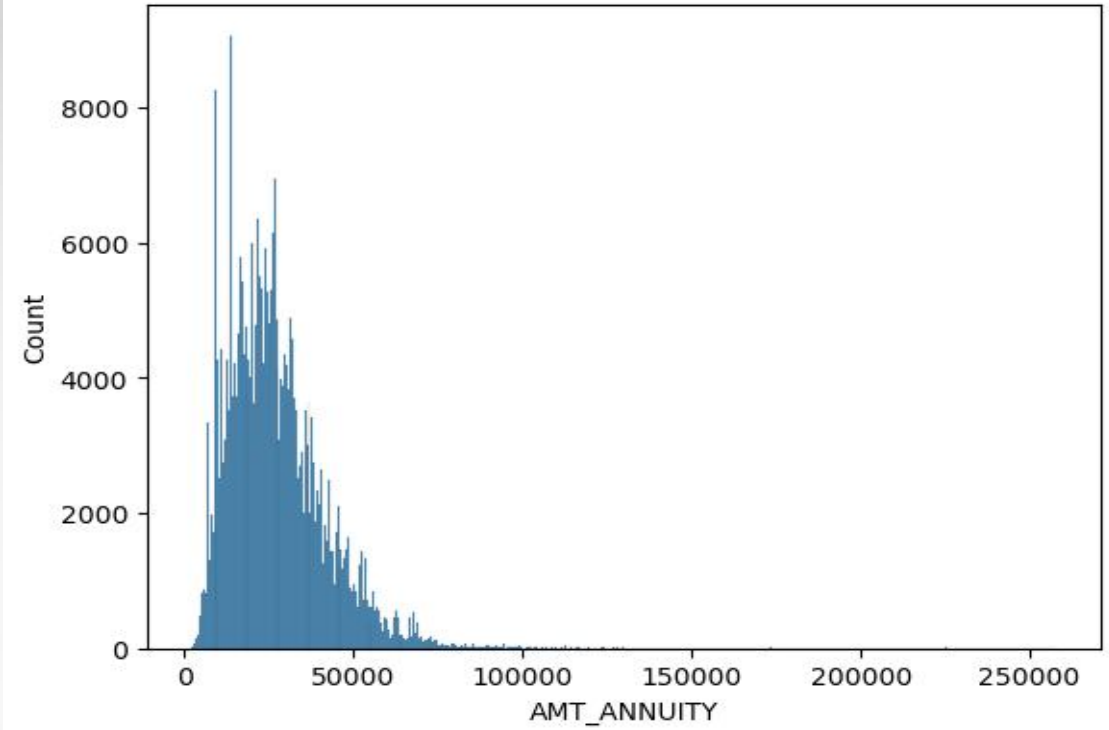
Graphical Analysis For Application Data

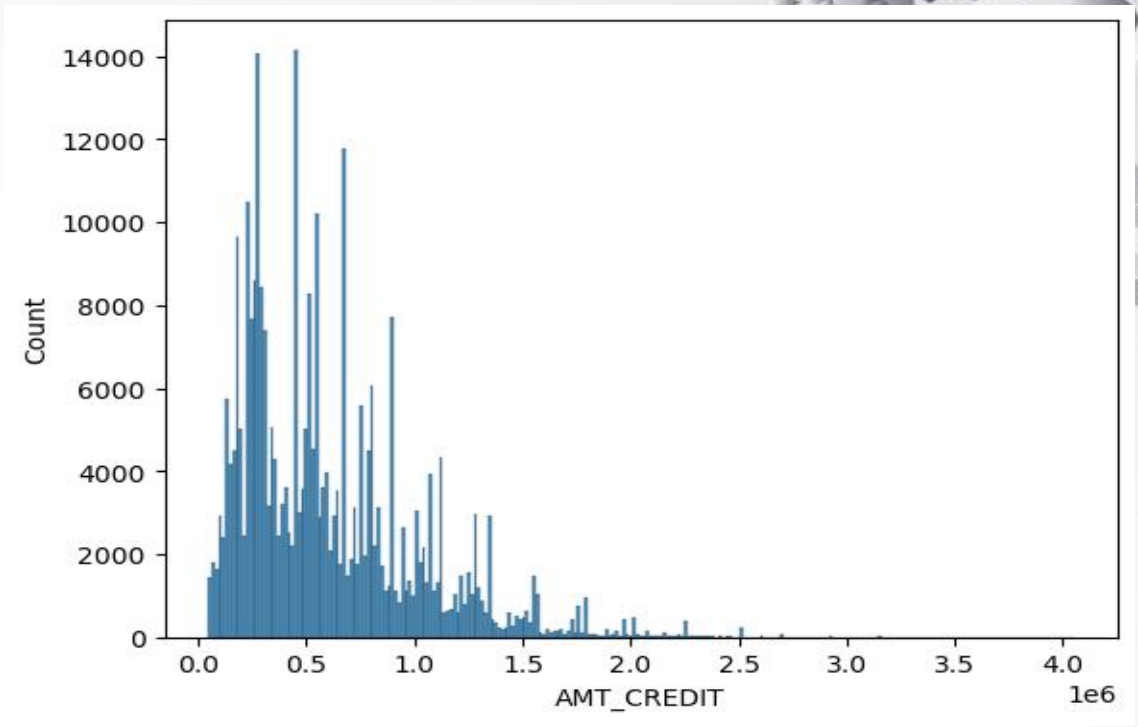

0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
AMT_GOODS_PRICE
1e6

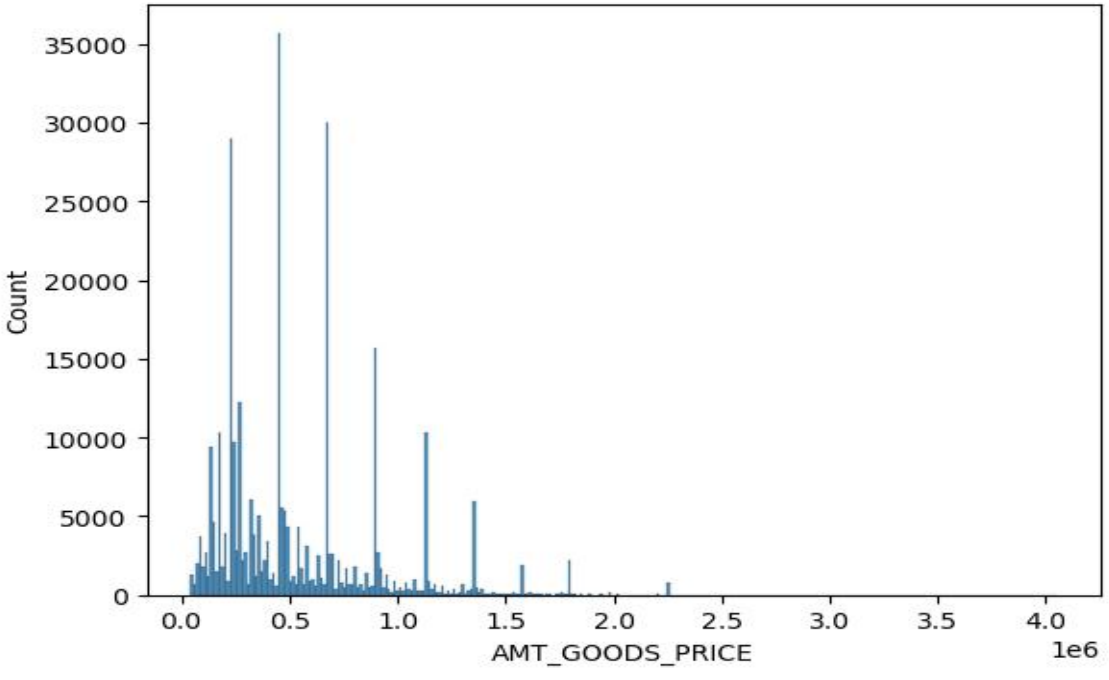

0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
AMT_CREDIT
1e6

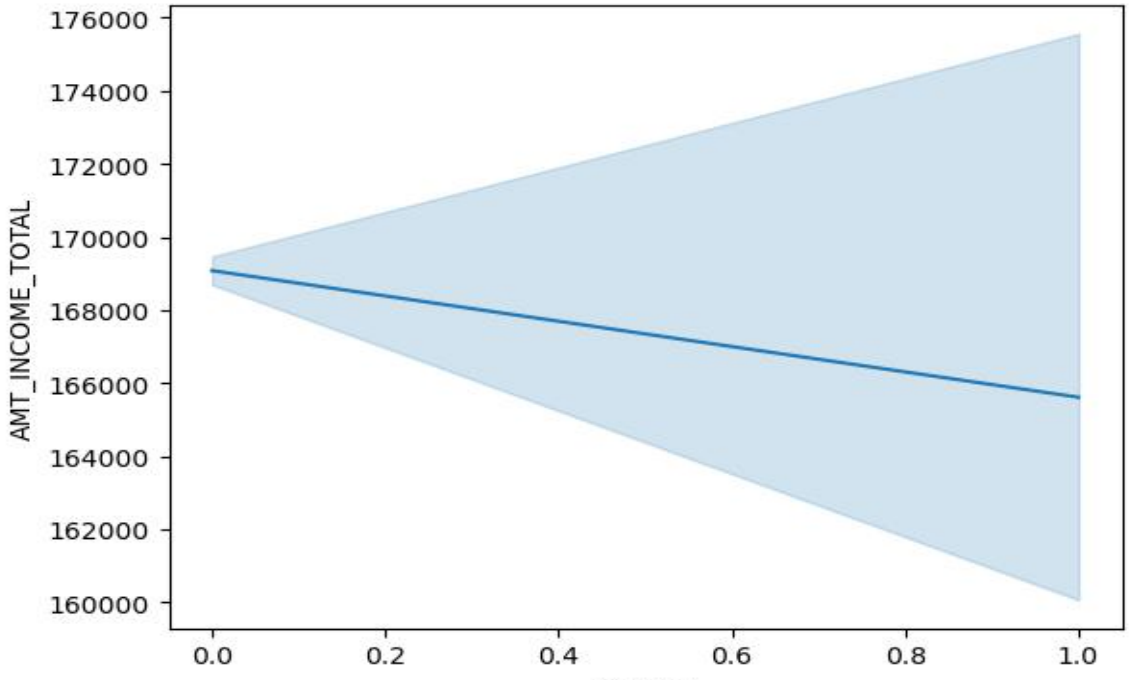

0
50000
100000
150000
200000
250000

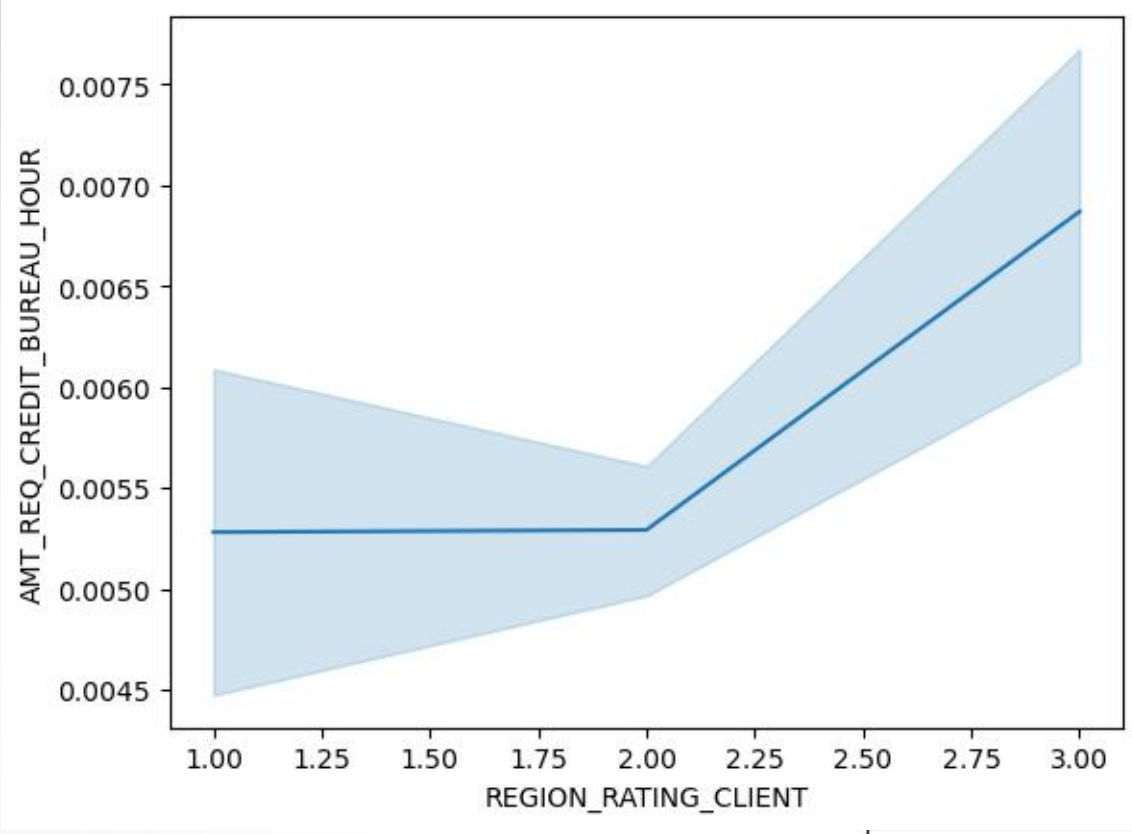


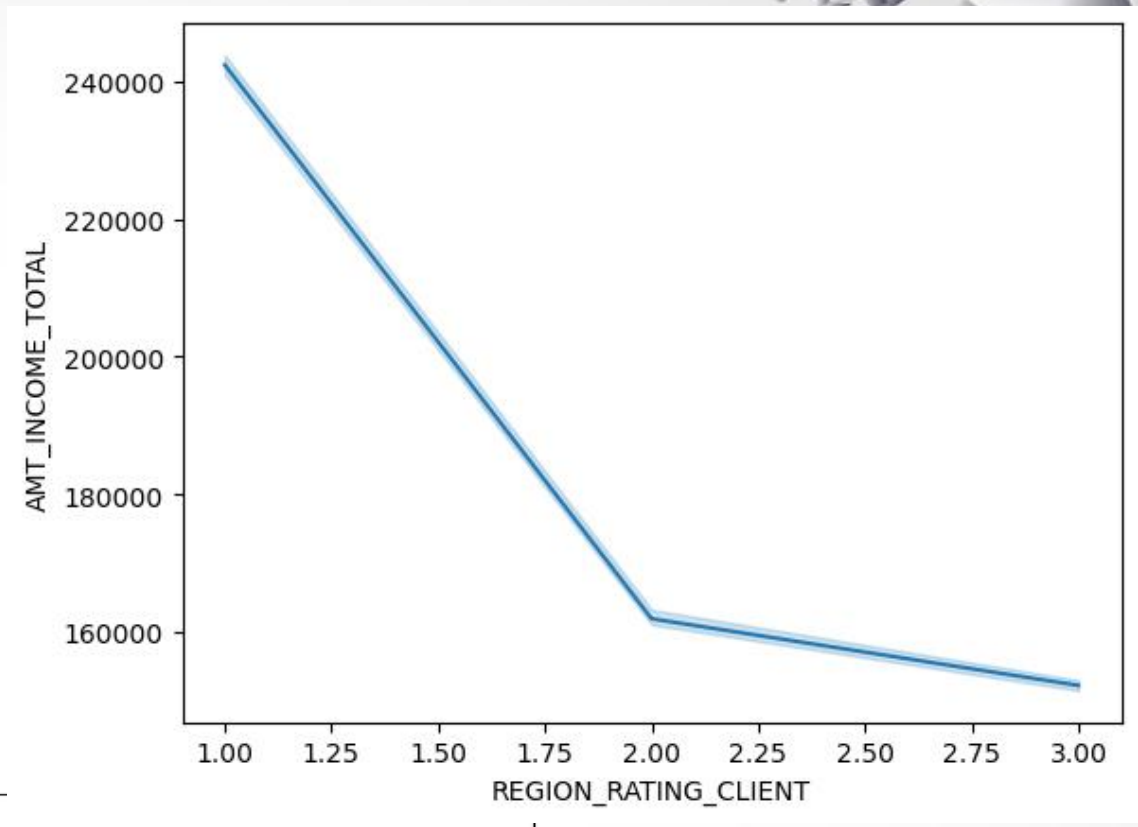

Count
8000
6000
4000
2000
0
0
50000
100000
150000
200000
250000
AMT_ANNUITY

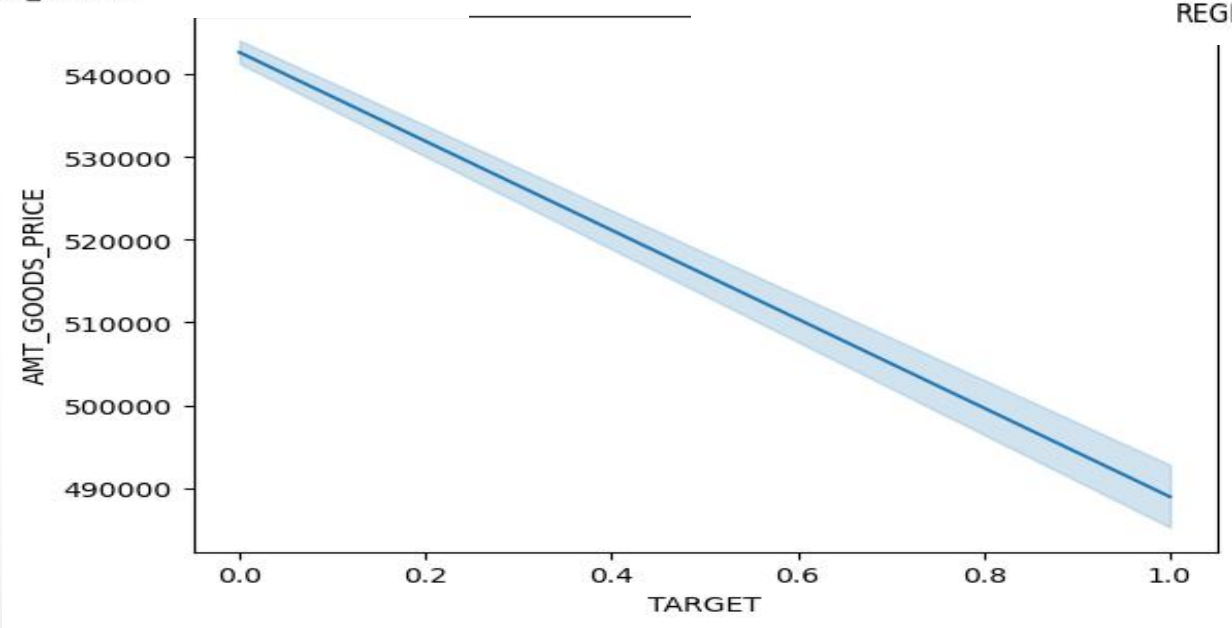

Count
14000
12000
10000
8000
6000
4000
2000
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1e6
AMT_CREDIT


Count
35000
30000
25000
20000
15000
10000
5000
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1e6
AMT_GOODS_PRICE

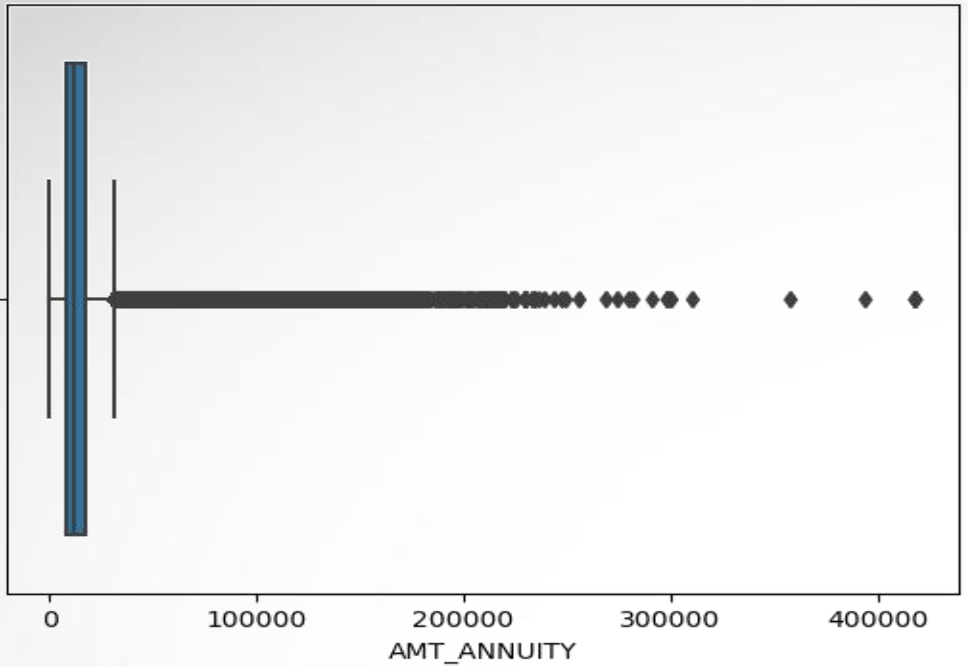

AMT_INCOME_TOTAL
176000
174000
172000
170000
168000
166000
164000
162000
160000
0.0
0.2
0.4
0.6
0.8
1.0

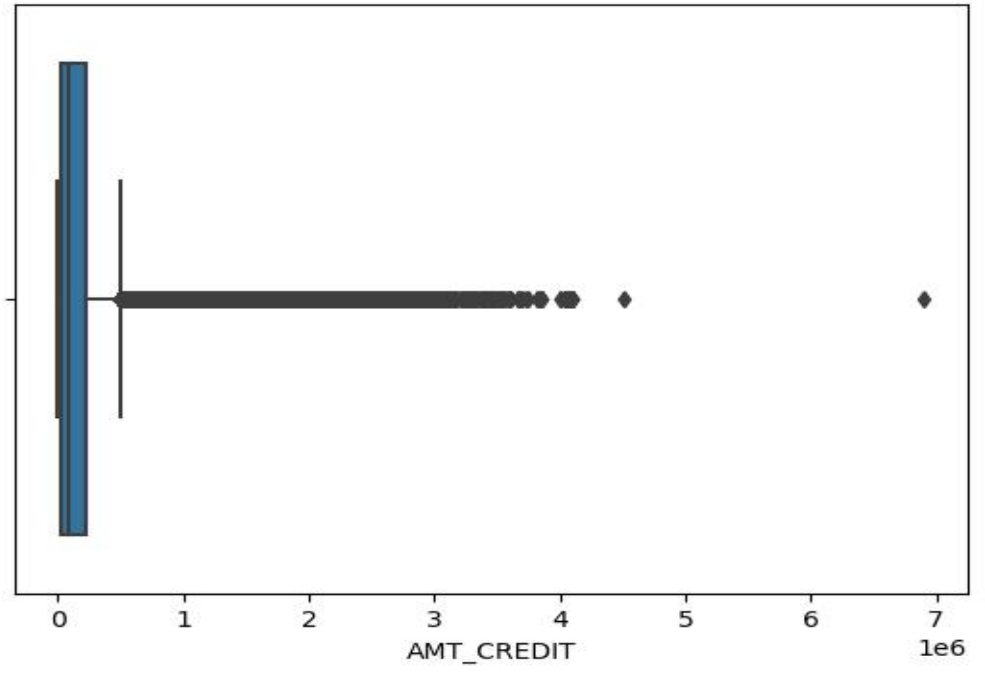

AMT_REQ_CREDIT_BUREAU_HOUR
0.0075
0.0070
0.0065
0.0060
0.0055
0.0050
0.0045
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
REGION_RATING_CLIENT

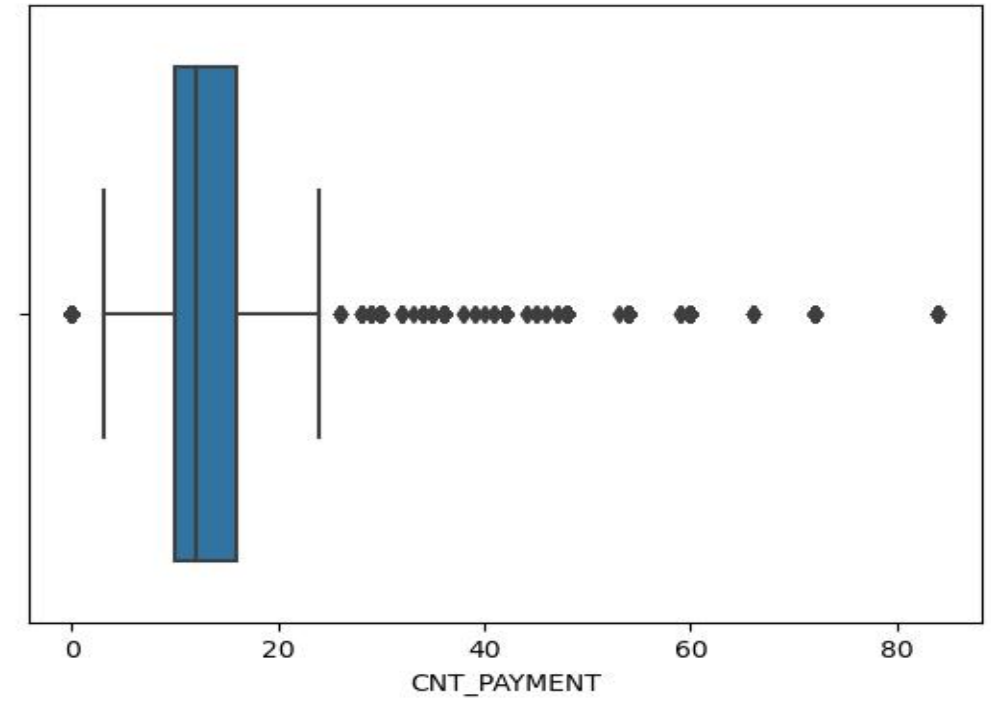

AMT_INCOME_TOTAL
240000
220000
200000
180000
160000
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
REGION_RATING_CLIENT

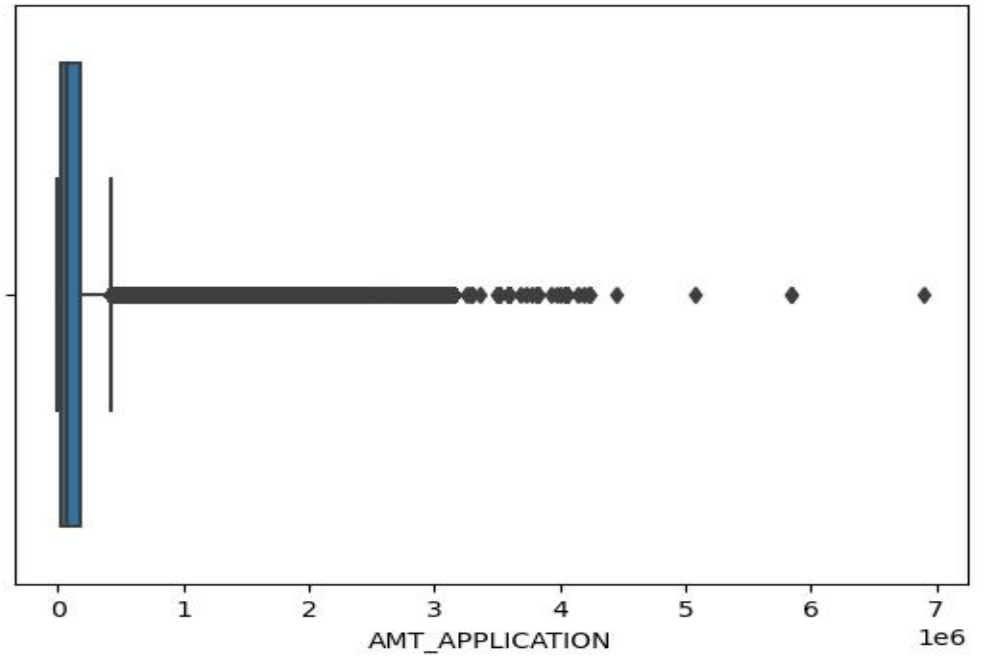

AMT_GOODS_PRICE
540000
530000
520000
510000
500000
490000
0.0
0.2
0.4
0.6
0.8
1.0
TARGET

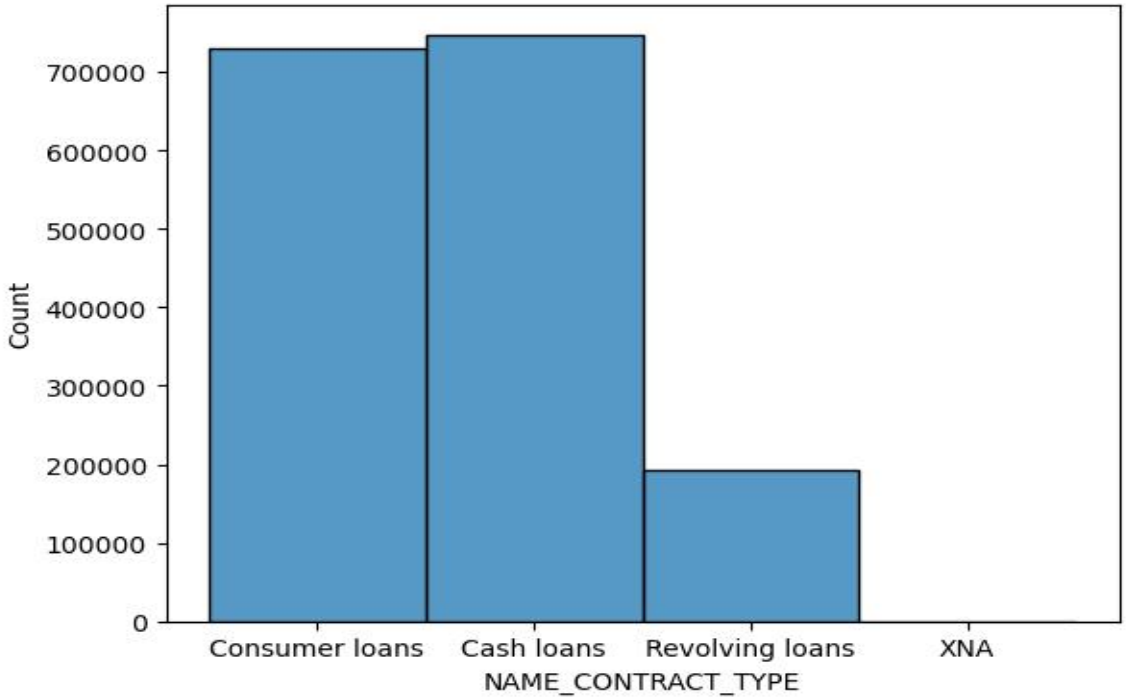
Graphical Analysis For Previous Application Data

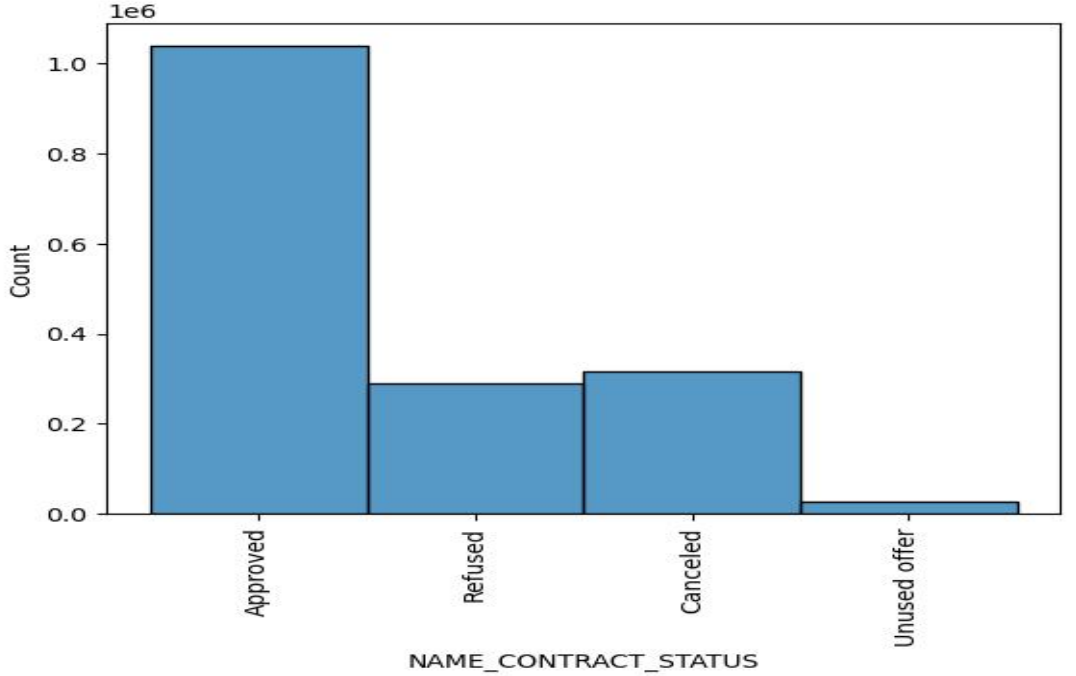

0
100000
200000
300000
400000
AMT_ANNUITY

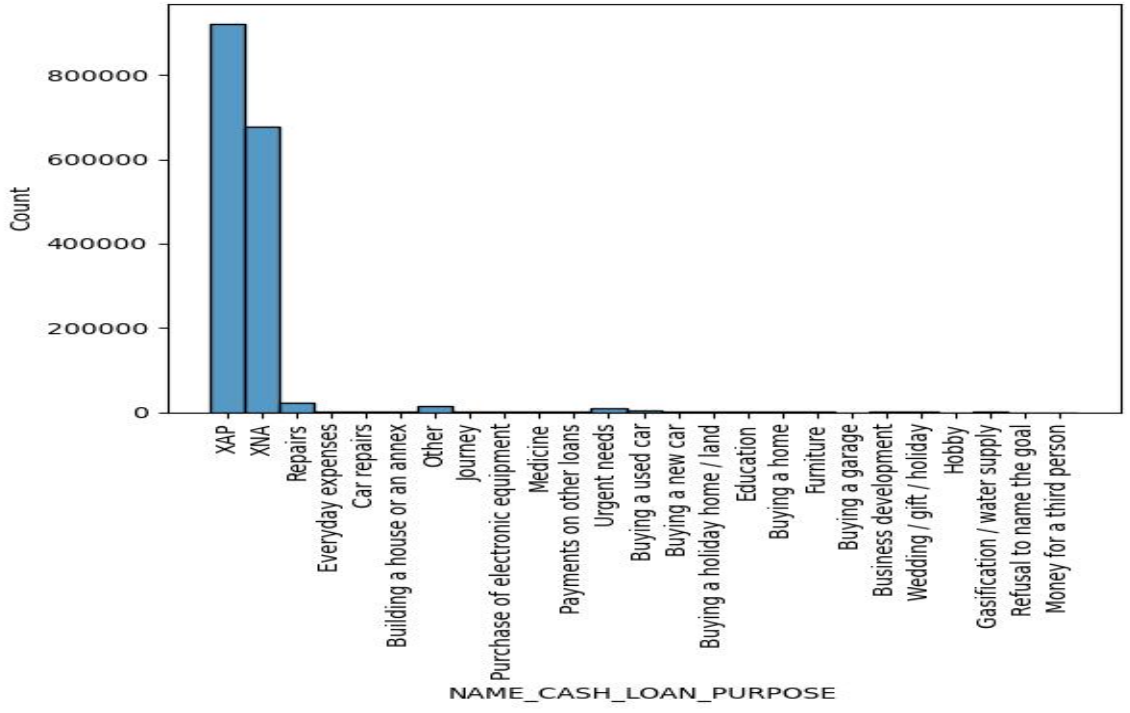

0
1
2
3
4
5
6
7
1e6
AMT_CREDIT

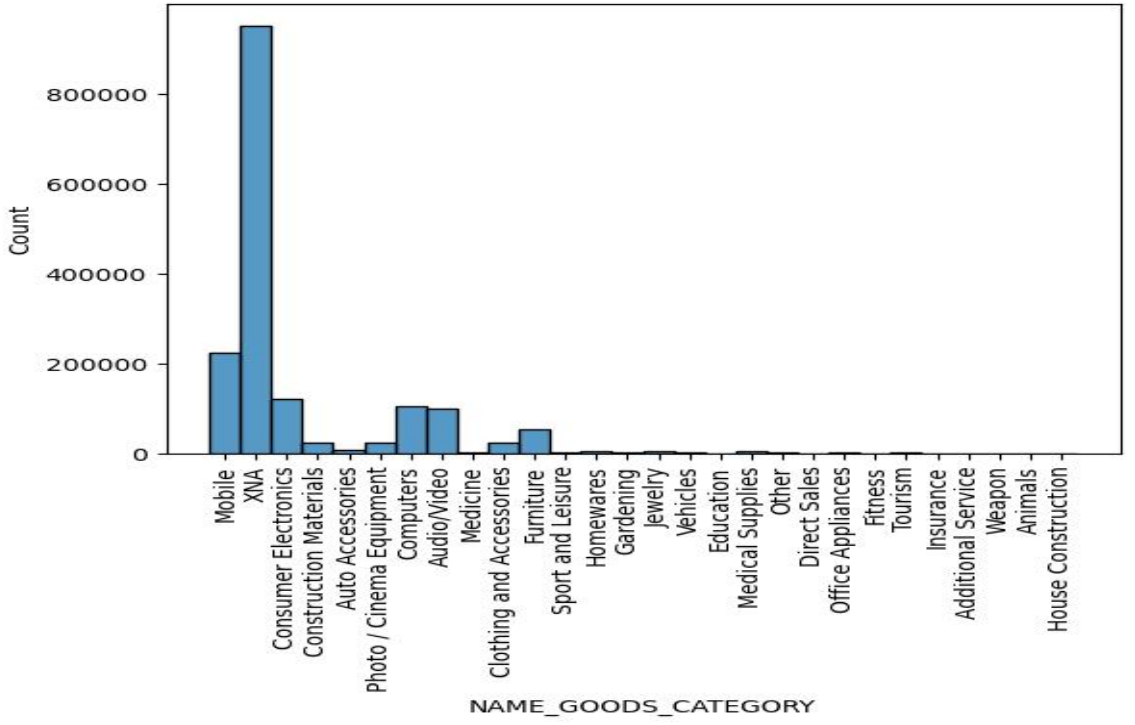

0
20
40
60
80
CNT_PAYMENT

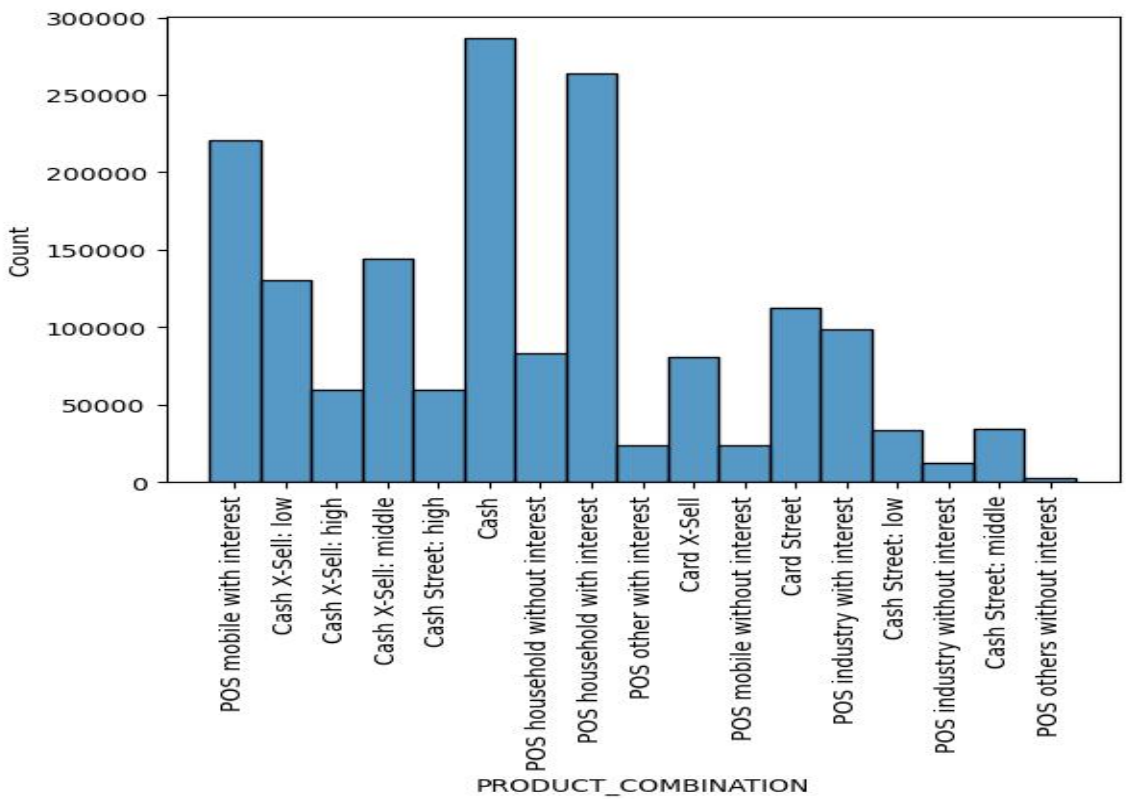

0
1
2
3
4
5
6
7
1e6
AMT_APPLICATION

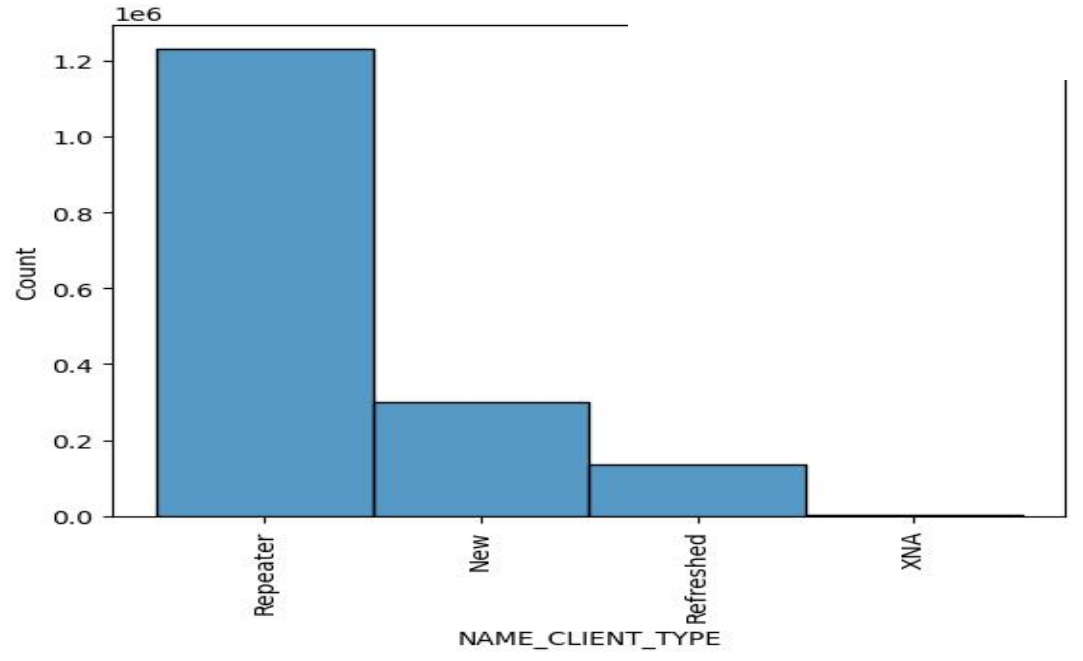

Count
700000
600000
500000
400000
300000
200000
100000
0
Consumer loans
Cash loans
Revolving loans
XNA
NAME_CONTRACT_TYPE

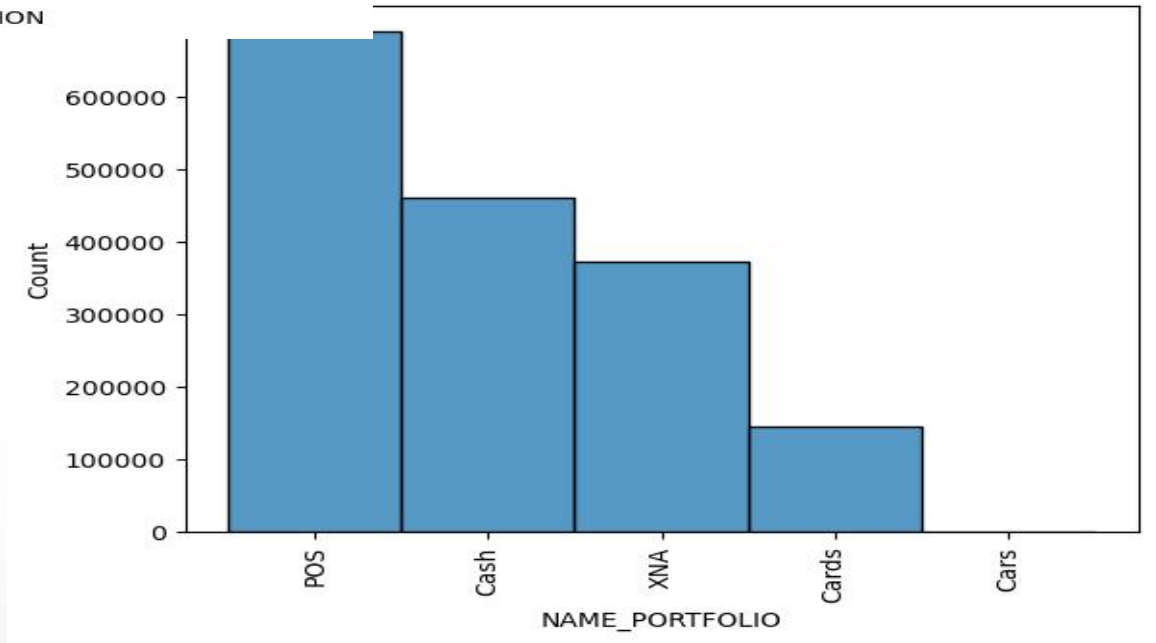

1e6
Count
1.0
0.8
0.6
0.4
0.2
0.0
Approved
Refused
Canceled
Unused offer
NAME_CONTRACT_STATUS


Count
800000
600000
400000
200000
0
XAP
XNA
Repairs
Everyday expenses
Car repairs
Building a house or an annex
Other
Journey
Purchase of electronic equipment
Medicine
Payments on other loans
Urgent needs
Buying a used car
Buying a new car
Buying a holiday home / land
Education
Buying a home
Furniture
Buying a garage
Business development
Wedding / gift / holiday
Hobby
Gasification / water supply
Refusal to name the goal
Money for a third person
NAME_CASH_LOAN_PURPOSE


Count
800000
600000
400000
200000
0
Mobile
XNA
Consumer Electronics
Construction Materials
Auto Accessories
Photo / Cinema Equipment
Computers
Audio/Video
Medicine
Clothing and Accessories
Furniture
Sport and Leisure
Homewares
Gardening
Jewelry
Vehicles
Education
Medical Supplies
Other
Direct Sales
Office Appliances
Fitness
Tourism
Insurance
Additional Service
Weapon
Animals
House Construction
NAME_GOODS_CATEGORY


Count
300000
250000
200000
150000
100000
50000
0
POS mobile with interest
Cash X-Sell: low
Cash X-Sell: high
Cash X-Sell: middle
Cash Street: high
Cash
POS household without interest
POS household with interest
POS other with interest
Card X-Sell
POS mobile without interest
Card Street
POS industry with interest
Cash Street: low
POS industry without interest
Cash Street: middle
POS others without interest
PRODUCT_COMBINATION


1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
Repeater
New
Refreshed
XNA
NAME_CLIENT_TYPE


600000
500000
400000
300000
200000
100000
0
Count
POS
Cash
XNA
Cards
Cars
NAME_PORTFOLIO

Conclusion

- From the analysis it is clear that the maximum number of loan disbursed by the banks/NBFC are cash loans which has high risk potential
- People from low income goup can also be threat.
- The maximum amount in previous application used is for purchasing mobiles.
- Maximum number of loans were distributed from POS.
- More than 60% of the previous applications were from existing customers
- More than 70% of the applications are in 1-10 lac of loan amount range