Credit EDA Assignment

RIYA CHETWANI

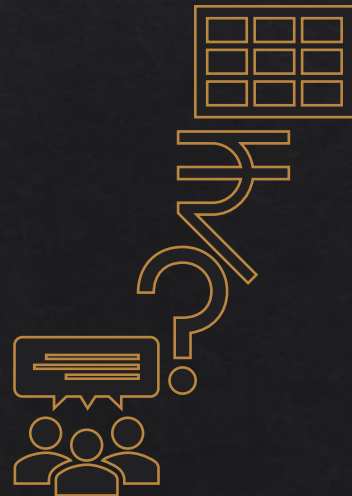# INDEX

INTRODUCTION

✓ The Case Study is related to Financial Institution especially of Lending Money.

✓ Risk for Loss is Greater in this Business Category.

✓ The Case Study aim applying EDA in real world problem

BUSINESS OBJECTIVE

➤ The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default through this case study.

➤ The company needs to understand the pattern of defaulter.

➤ & The company can utilize this knowledge for its portfolio and risk assessment.

EXPLORING DATA

➢ Data available are Previous Applicants, Column Description & Current Application.

➢ Previous Application : Customers who have already took loan.

➢ Column Description : This Data gives information of al columns present at other two data set.

➢ Current Application : This Data gives details of customer who have applied for loan currently & all other details related to them.

EXPLORING DATA

➢ We will start exploring the application of current Applicants then previous applicants.

➢ Data available for Current Application consist of :

➢ Entries : 307511*122 (Rows & Columns)

➢ Data available for Previous Aplliacation consist of :
➢ Entries :1048757*37 ((Rows & Columns)

DEALING WITH MISSING/NEGATIVE VALUE

- In the Data we have missing value.

- Sum of the Null Value in Current Application is 9152465.

- The Column which have more than **35%-40 % missing value** its advisable to drop that those columns.

- We make sure the columns which we are dropping doesn't effect our analysis.

- After dropping the Null Value Columns we have **75 Columns in the Data in current application & 26 in Previous Application Data**

- Missing Data with less than 40% is imputed with Mode, Median & Mean.

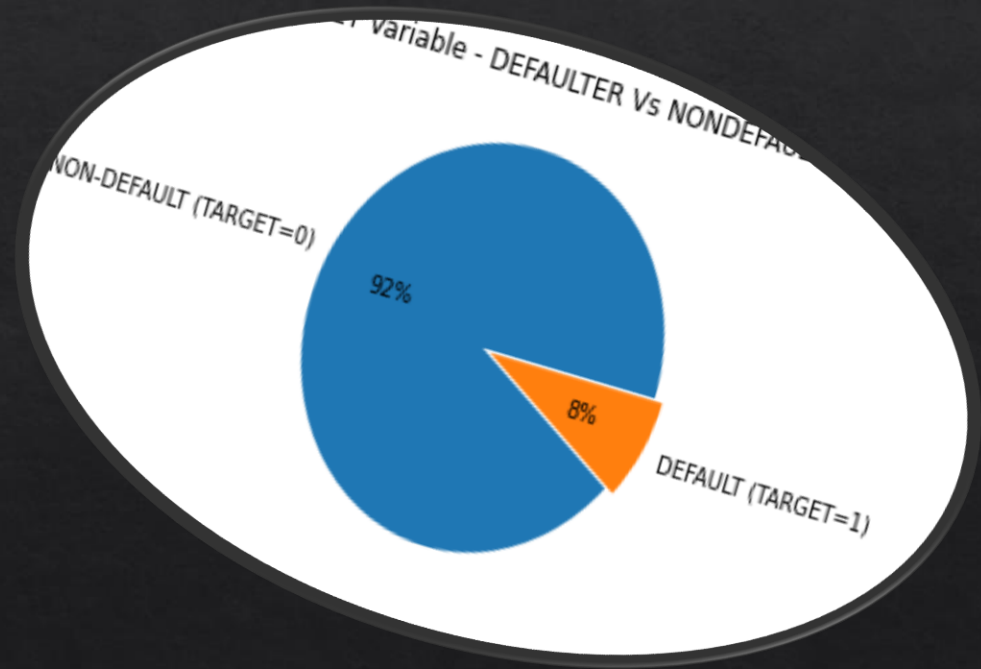- Negative will also neutralize with absolute number

ANALYSIS DATA

➢ During the process we have observe there are few columns which have the Outliner.

➢ So we deal with these with Binnig Methods & balance the Outliner with required quantile.
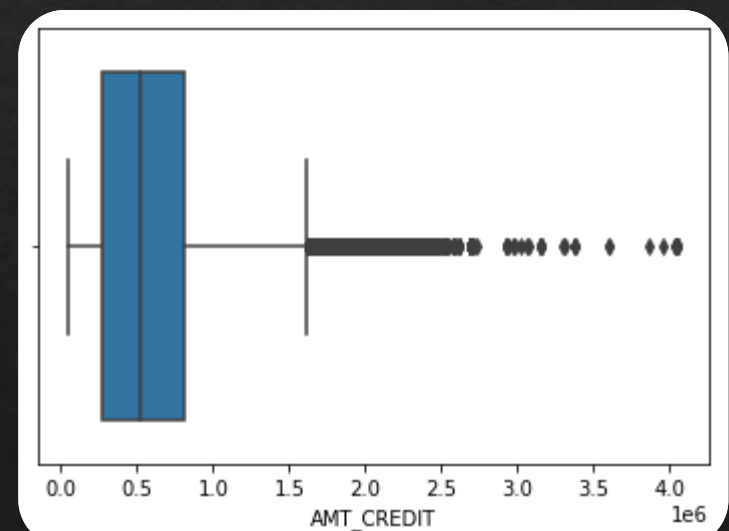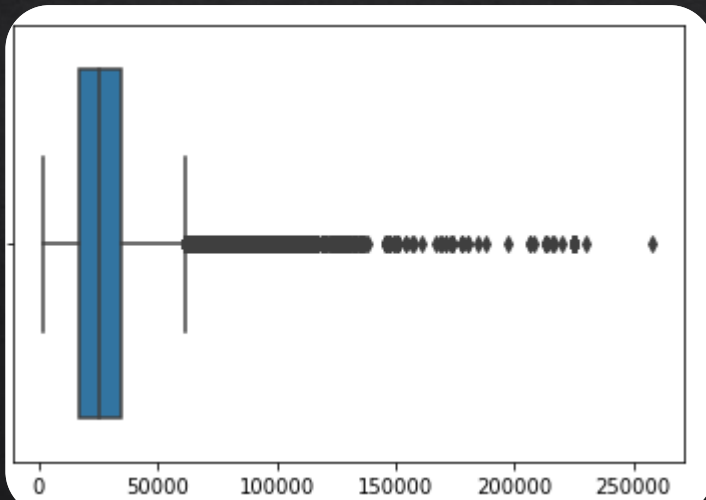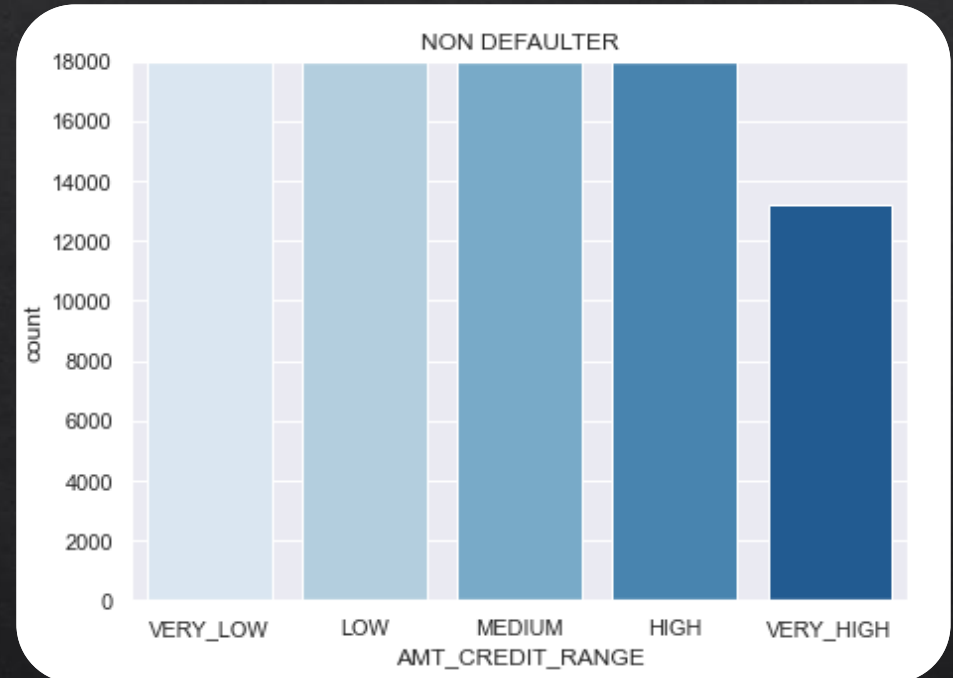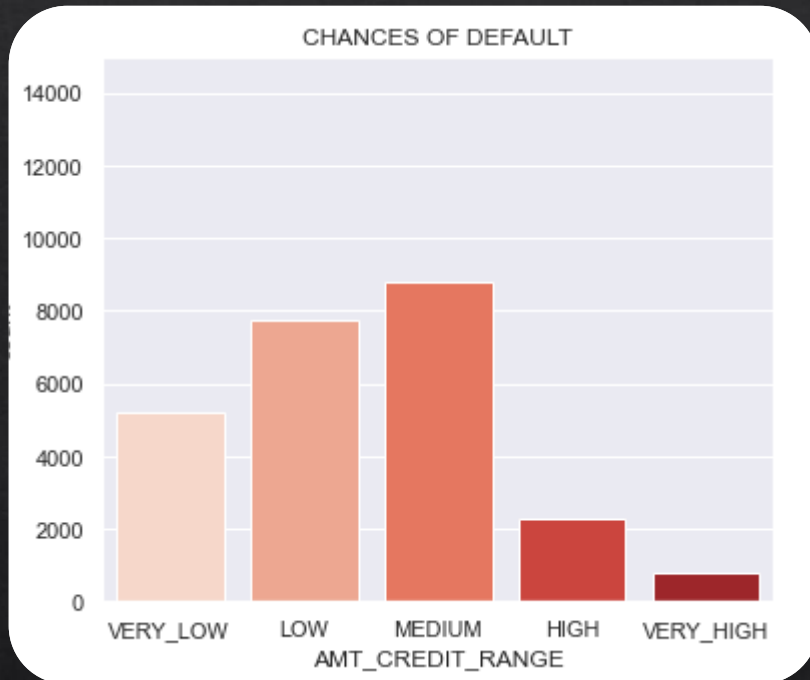
ANALYSIS DATA : TARGET

- ➤ *RATIO OF TARGET : 8% of APPLICANT difficulties in payments & 91.92% are having no difficulties*
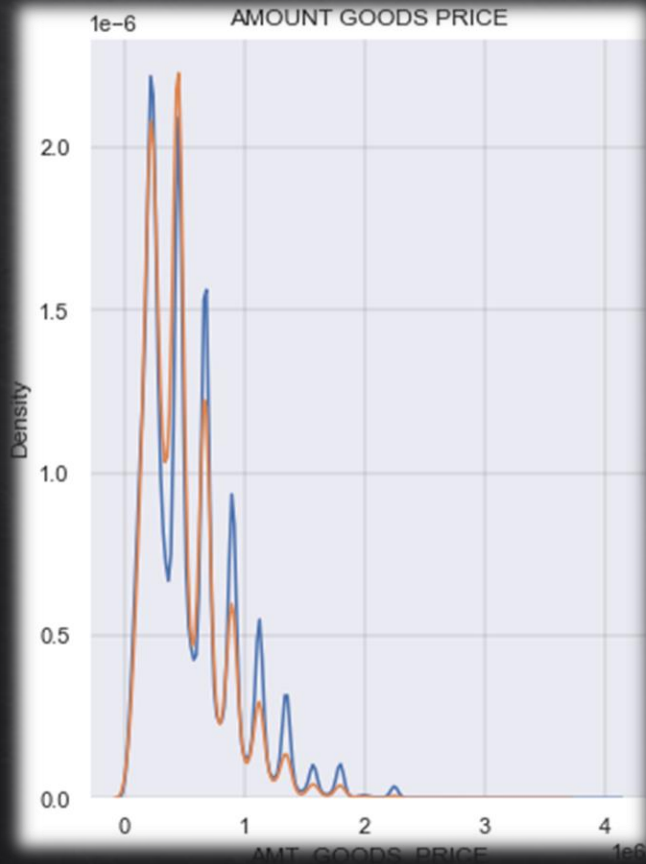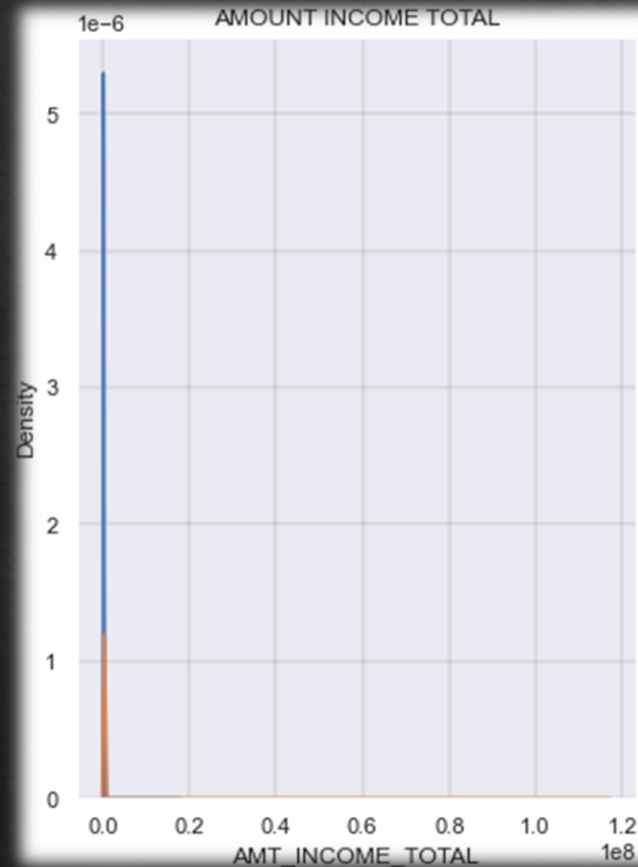
ANALYSIS DATA : TARGET

AMT_CREDIT_RANGE

➤ *APPLICANT WITH HIGH & VERY HIGH LIKELY TO PAY ON TIME*
➤ *APPLICANT WITH MEDIUM CREDIT ARE NOT LIKELY TO PAY ON TIME*
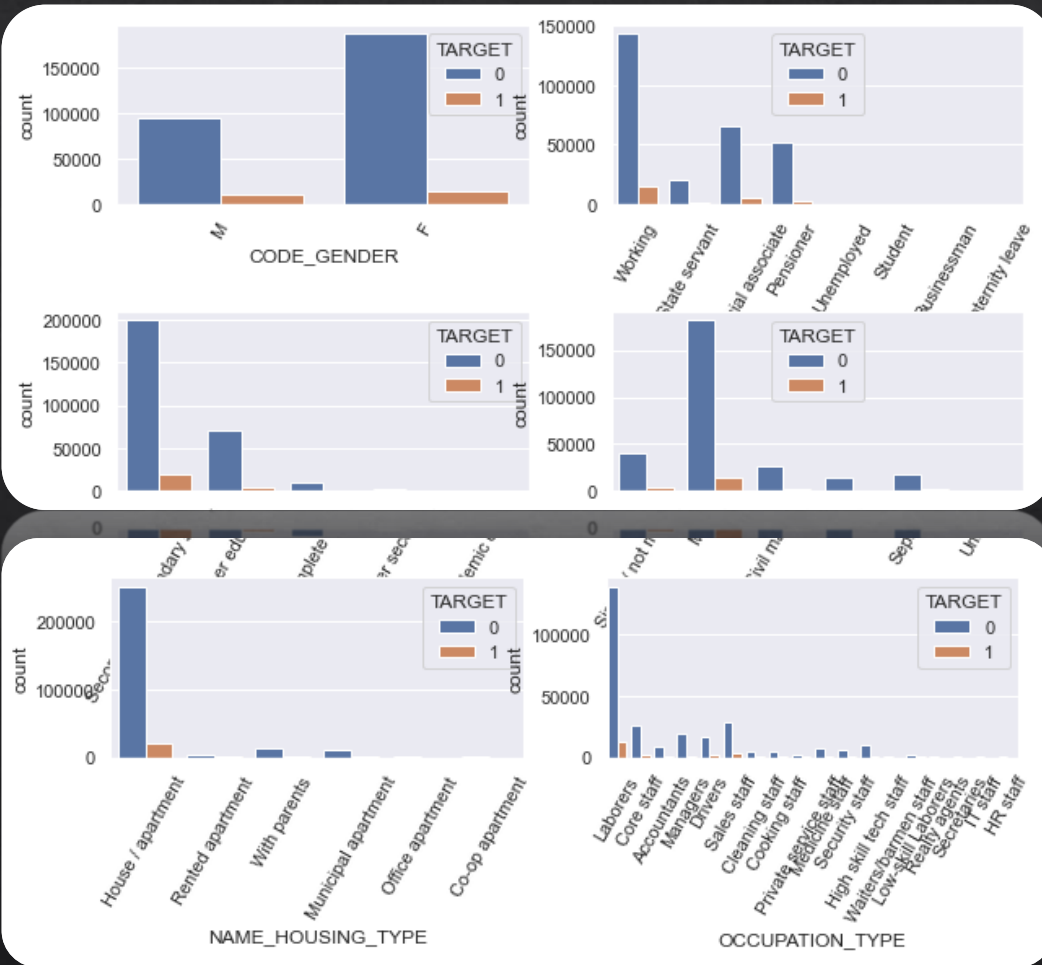
ANALYSIS DATA : TARGET

AMT_GOODS_PRICE

AMT_INCOME_TOTAL

➤ *APPLICANT WITH INCOME TOTAL MORE THAN 1.5 ISVERY HIGH LIKELY TO PAY ON TIME*
➤ *APPLICANT WITH INCOME TOTAL LESS THAN 1.0 ARE NOT LIKELY TO PAY ON TIME*

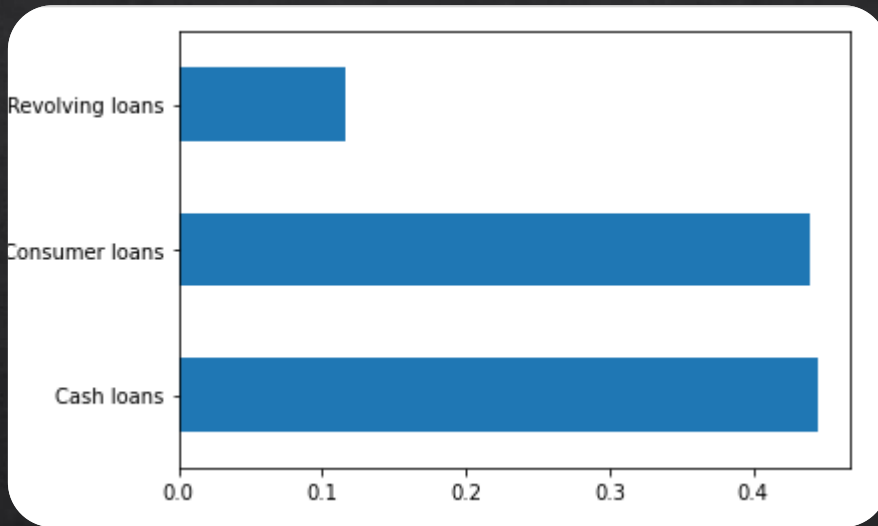# ANALYSIS DATA : TARGET

- *Applicant with Gender Female are more likely to pay on time*
- *Applicant with occupation type core staff, laborer are more likely to pay on time.*
- *Applicant with Secondary Education is more likely on time .*
- *Applicant as an working class are likely to pay on time.*

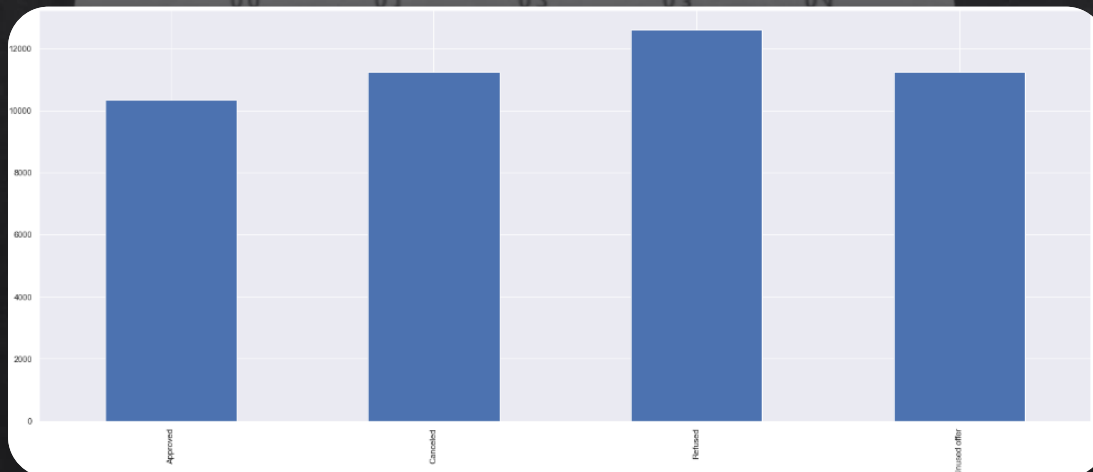- *Applicant with HR Profession is not likely to pay on time*
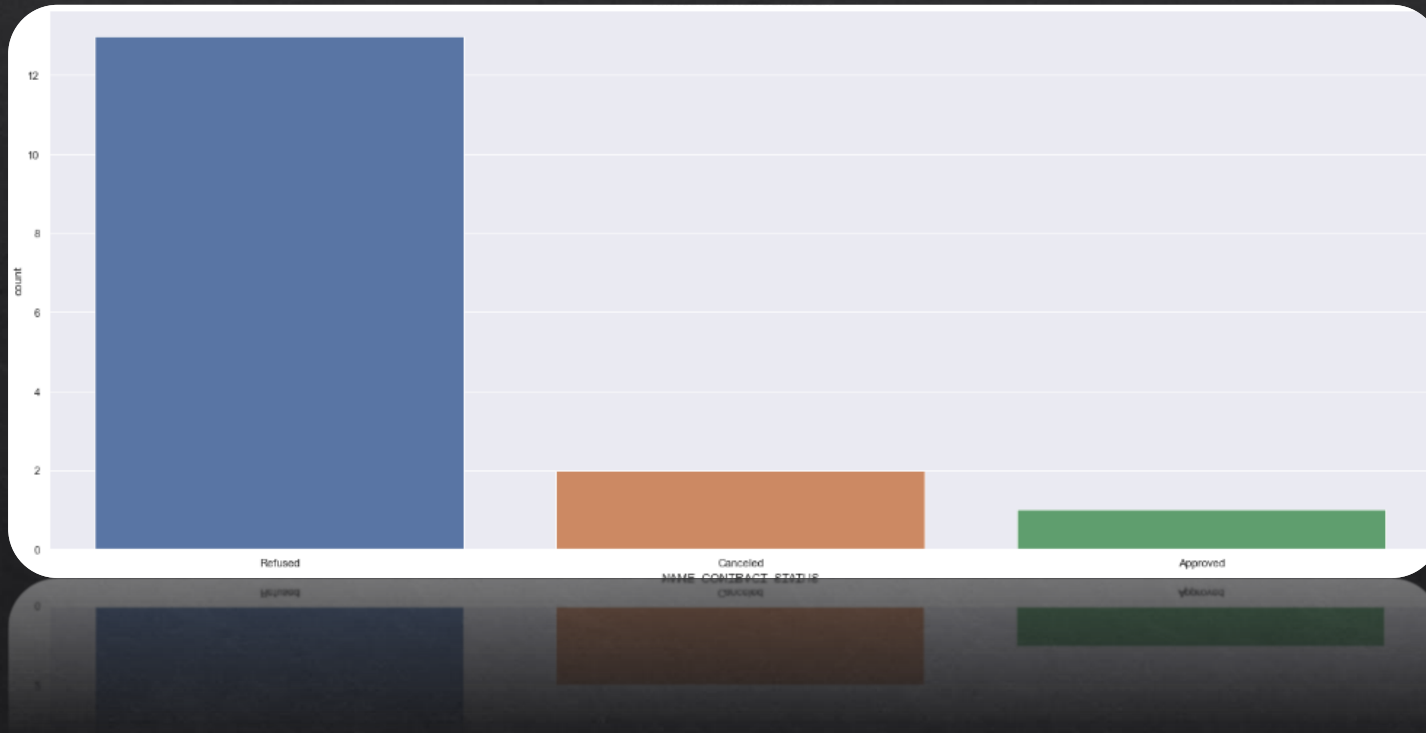
ANALYSIS DATA : CONTRACT_STATUS

PRODUCT TYPE

- ➤ *Applicants are more likely to take cash loan in Previous Data.*

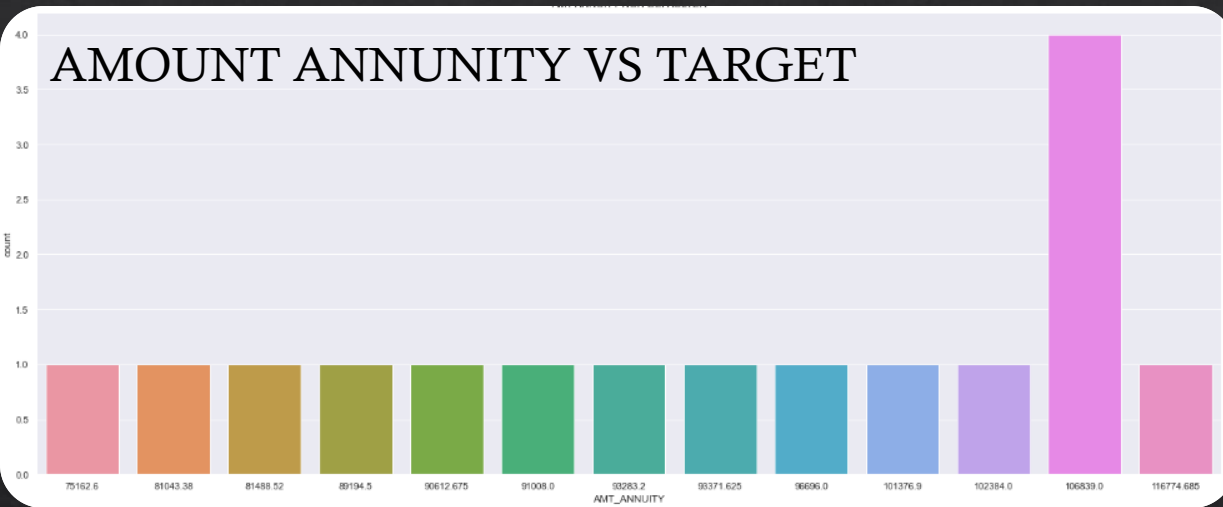- ➤ *The Rate of Rejected is high in Previous Data.*
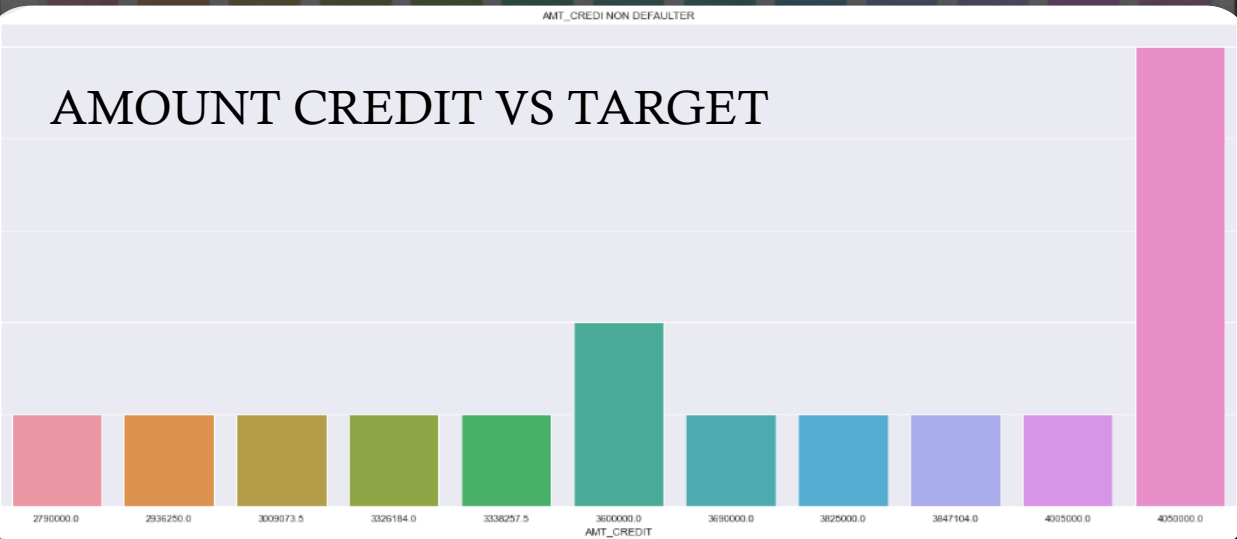

STATUS

ANALYSIS DATA AFTER MERGING THE DATA

CONTRACT STATUS VS TARGET

➢ Applicant which are more likely to pay on time are refused.
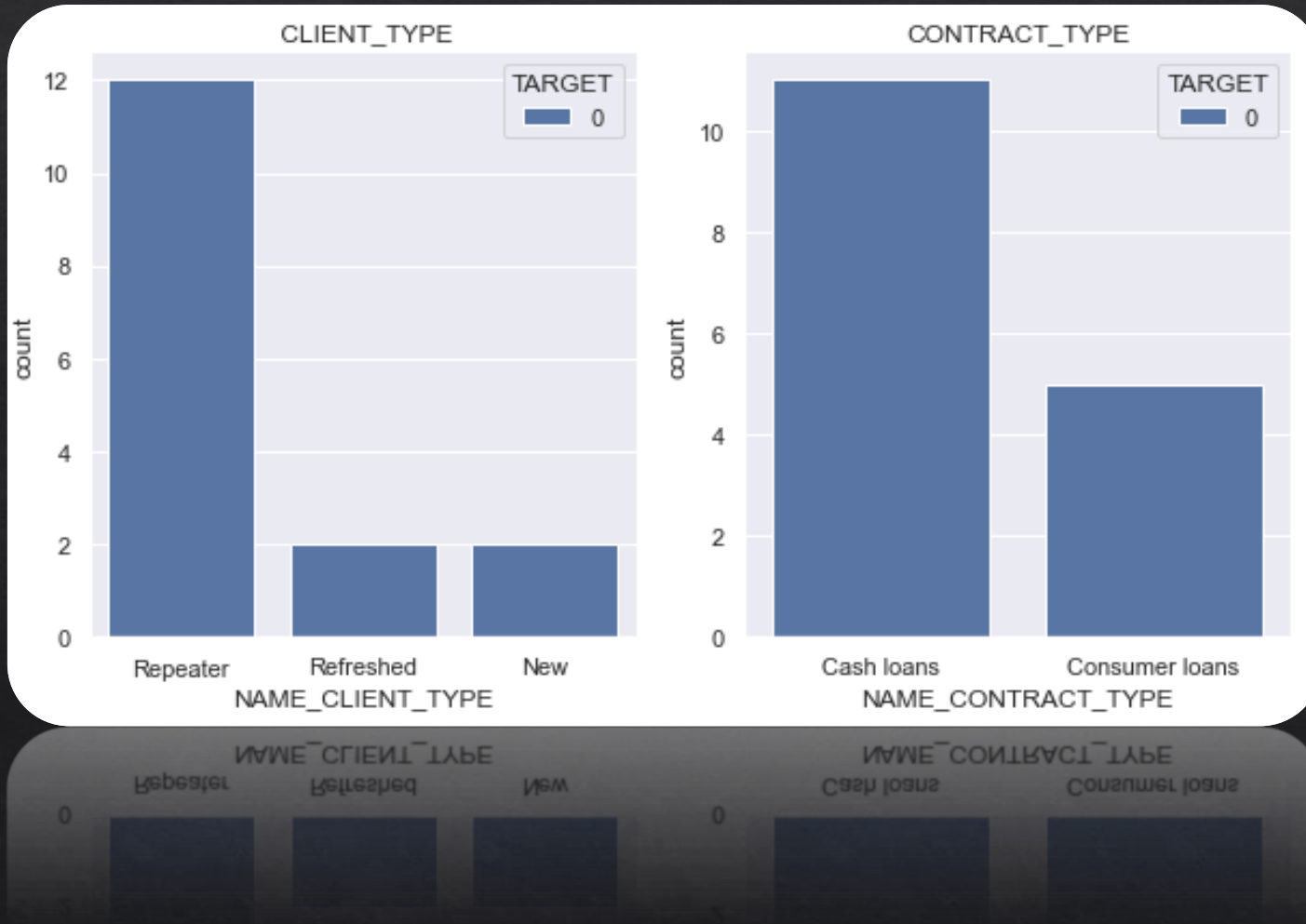
ANALYSIS DATA AFTER MERGING THE DATA

AMOUNT ANNUNITY VS TARGET

AMOUNT CREDIT VS TARGET

➢ *Applicants are high Annuity & credit are likely to pay on time.*

➢ *Applicant with more than 1,00,000 Credit & Annuity more thans 2 lacs.*

ANALYSIS DATA AFTER MERGING THE DATA

Applicant are more of Repeater & Loan type applied is most for Cash Loan.

CONCULSION

Female are more likely to pay on time.

Applicant with occupation Core Staff are more likely of Non-Defaulter Category.

Applicant with Married Status are likely to pay on time.

Applicant with Annuity more than 2 Lac is like to pay on time.

Applicant with Credit more than 1 Lac is likely pay on time.

Applicant of Working Class are likely pay on time

THANK YOU