

# EDA CREDIT ASSIGNMENT

# BUSINESS OBJECTIVE

- IDENTIFY THE LOAN DEFAULTING PATTERNS
- IDENTIFY DRIVING FACTORS BEHIND DEFAULT
- DEFAULTERS ARE THE APPLICANTS WHO HAD LATE PAYMENTS MORE THAN X DAYS ON AT LEAST ONE OF THE Y INSTALLMENTS OF THE LOAN. WE NEED TO FIND THE PATTERNS TO AVOID DEFAULTERS

# EDA OF APPLICATION\_DATA.CSV DATA SET

# DATA CLEANING AND ANALYSIS

- DATA CLEANING
  - TREATMENT OF OUTLIERS
  - STANDARDISING THE DATA VALUES
  - TREATMENT OF NULL VALUES
  - CHECKING THE DATA TYPES AND STANDARDISE THE COLUMNS
- ANALYSIS
  - UNIVARIATE ANALYSIS
  - BIVARIATE ANALYSIS
  - MULTIVARIATE ANALYSIS

# TREATMENT OF NULL VALUES

- REMOVE THE COLUMNS WITH MISSING VALUES GREATER THAN 40%
- THE REMAINING COLUMNS HAVING NULL VALUES LESS THAN 40% REPLACE THE VALUES WITH MEAN FOR NUMERICAL COLUMNS THAT DOESN'T HAVE OUTLIERS AND MODE FOR COLUMNS WITH OUTLIERS.
- FOR ALL THE CATEGORICAL COLUMNS REPLACE WITH MODE VALUES(ONLY IF THE COUNT OF MISSING VALUES IS NOT GREATER THAN THE COUNT OF MODE VALUE OF THE COLUMN)
- OCCUPATION\_TYPE COLUMN IS LEFT UNTREATED BECAUSE THE NULL VALUES COUNT IS GREATER THAN THE MODE VALUE COUNT

# REMOVE THE UNWANTED COLUMNS FOR ANALYSIS

- LOOKING AT THE COLUMN NAME DESCRIPTION REMOVE THE UNWANTED COLUMNS THAT ARE NOT NEEDED FOR ANALYSIS
- WE CAN DROP THE COLUMNS IF THE MAJORITY OF THE DATA IS ZERO
- WE CAN DROP THE HIGHLY SKEWED COLUMNS-COLUMNS THAT HAS CONSTANT VALUES ,MEANING MAJORITY OF ITS ENTRIES ARE SAME
- WE CAN DROP THE COLUMNS HAVING INCONSISTENT DATA

# TREATMENT OF OUTLIERS

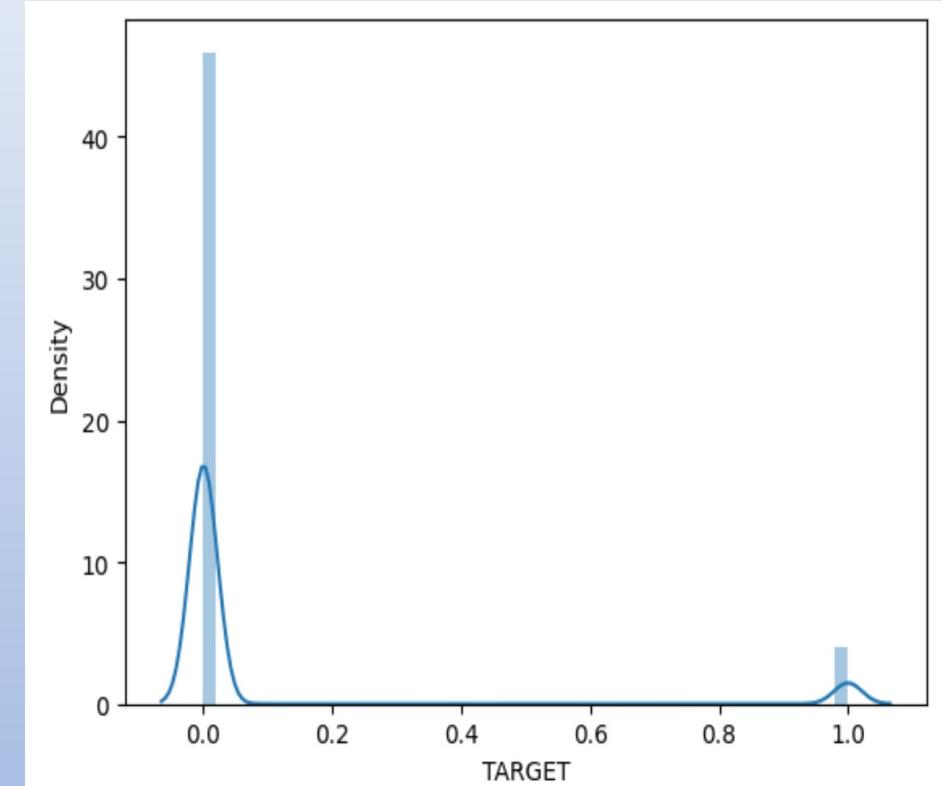
- IDENTIFY THE OUTLIERS FOR NUMERICAL COLUMNS USING BOXPLOT
- IF REPLACING THE OUTLIER WITH UPPERBOUND AND LOWER BOUND VALUES DOESN'T IMPACT THE ANALYSIS CAP THOSE VALUES ACCORDINGLY ELSE LEAVE THE COLUMN UNTREATED AND MAKE SURE YOU DON'T USE MEAN,SD OR VARIANCE FOR ANALYSIS OF THOSE COLUMNS
- WE CAN BUCKET THE VALUES TO IMPUTE OUTLIERS(EX.AGE COLUMN)

# CHECKING THE DATA TYPES AND STANDARDISE THE COLUMNS

- THE DATA TYPES OF THE NUMERICAL COLUMNS SHOULD BE INT OR FLOAT AND THE DATA TYPES OF THE CATEGORICAL COLUMN SHOULD BE OBJECT/FLOAT/INT. IF NOT THEN CONVERT THE DATA TYPES ACCORDINGLY
- CHECK FOR THE VALUES IN EACH COLUMN AND STANDARDISE UNITS/SCALE VALUES/STANDARDISE PRECISION ACCORDINGLY
- CONVERT DAYS\_BIRTH TO AGE FOR EASY ANALYSIS
- CONVERT DAYS\_EMPLOYED, DAYS\_REGISTRATION, DAYS\_ID\_PUBLISH, DAYS\_LAST\_PHONE\_CHANGE IN TO ABSOLUTE VALUES

# DATA IMBALANCE

- DATA IS HEAVILY SKEWED TOWARDS ONE OF THE TARGET VALUE(TARGET=1)
- IMBALANCE RATIO = 11.39
- DUE TO DATA IMBALANCE, WE WILL SPLIT THE DATA IN TO TWO GROUPS – NON-DEFAULT-TARGET=0 AND DEFAULT-TARGET=1 FOR FURTHER ANALYSIS

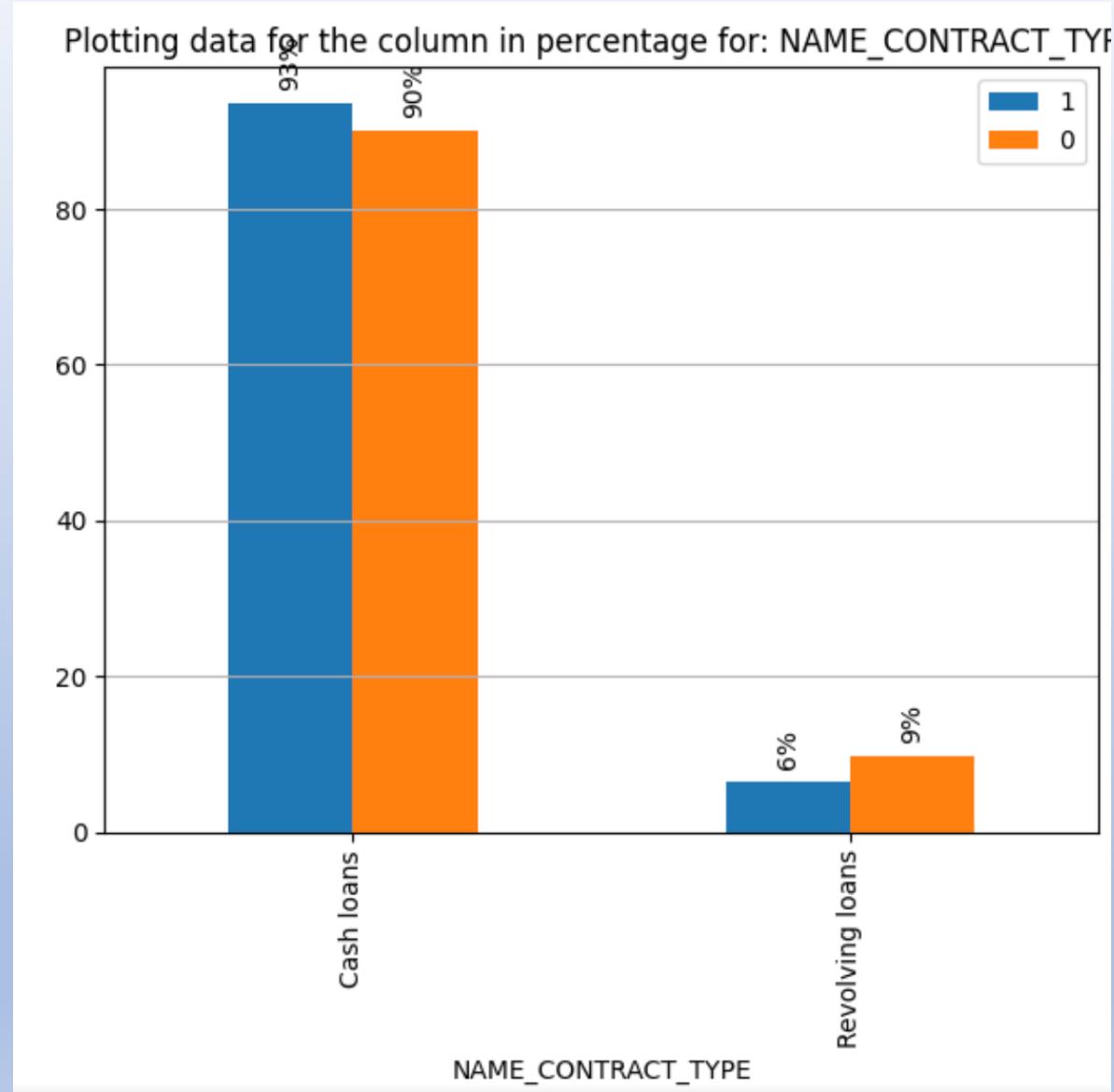


# UNIVARIATE ANALYSIS DEFAULT(TARGET=1) AND NON-DEFAULT (TARGET=0)

# CATEGORICAL UNIVARIATE ANALYSIS (COUNT PLOT AND PIE CHART)

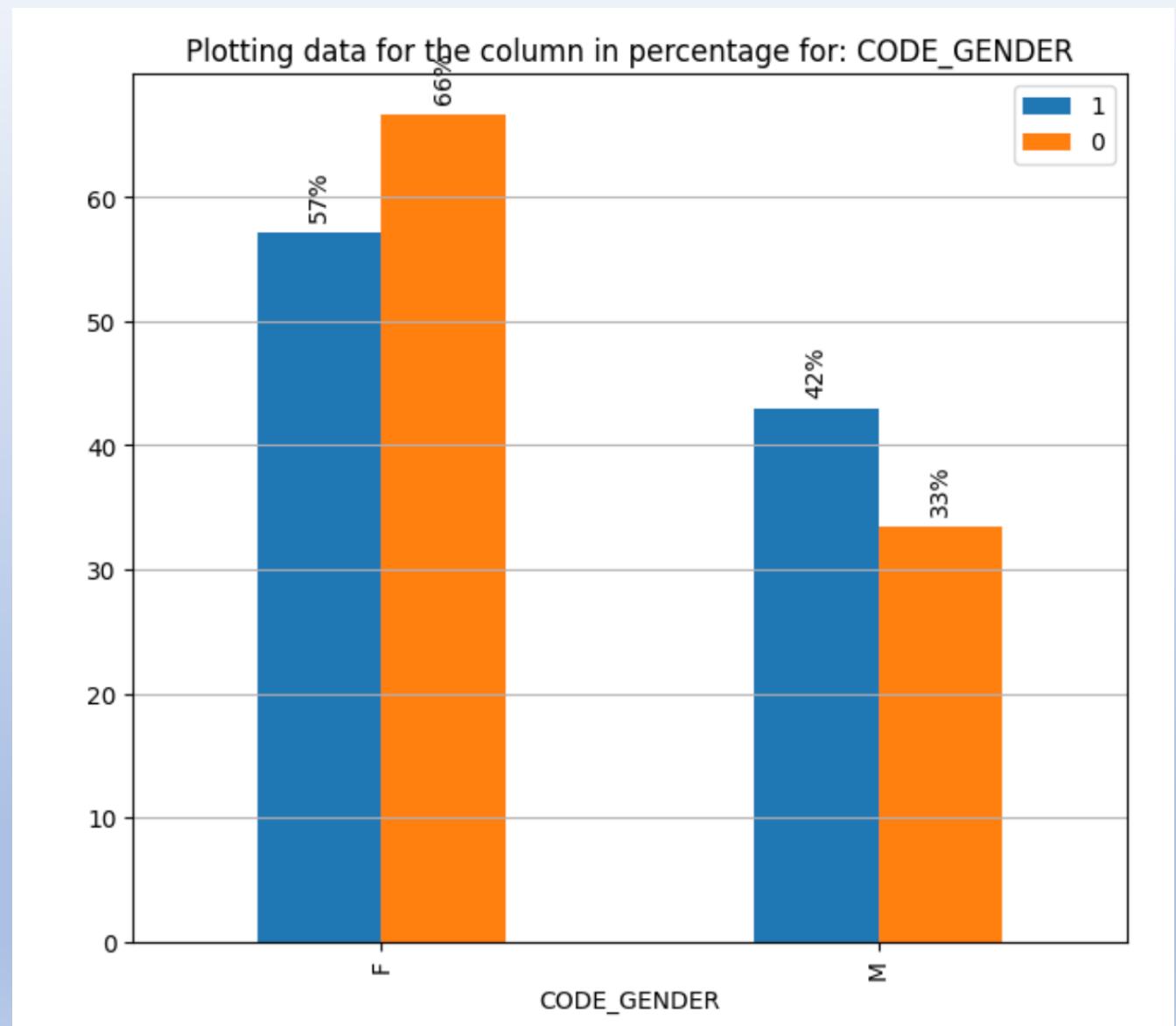
NAME\_CONTRACT\_TYPE(IF THE LOAN TYPE IS CASH OR REVOLVING)

- BOTH THE DEFAULTERS AND NON-DEFAULTERS HAVE GOT MORE CASH LOANS COMPARED TO REVOLVING LOANS
- PERCENTAGE OF DEFAULTERS WHO HAVE OPTED FOR CASH



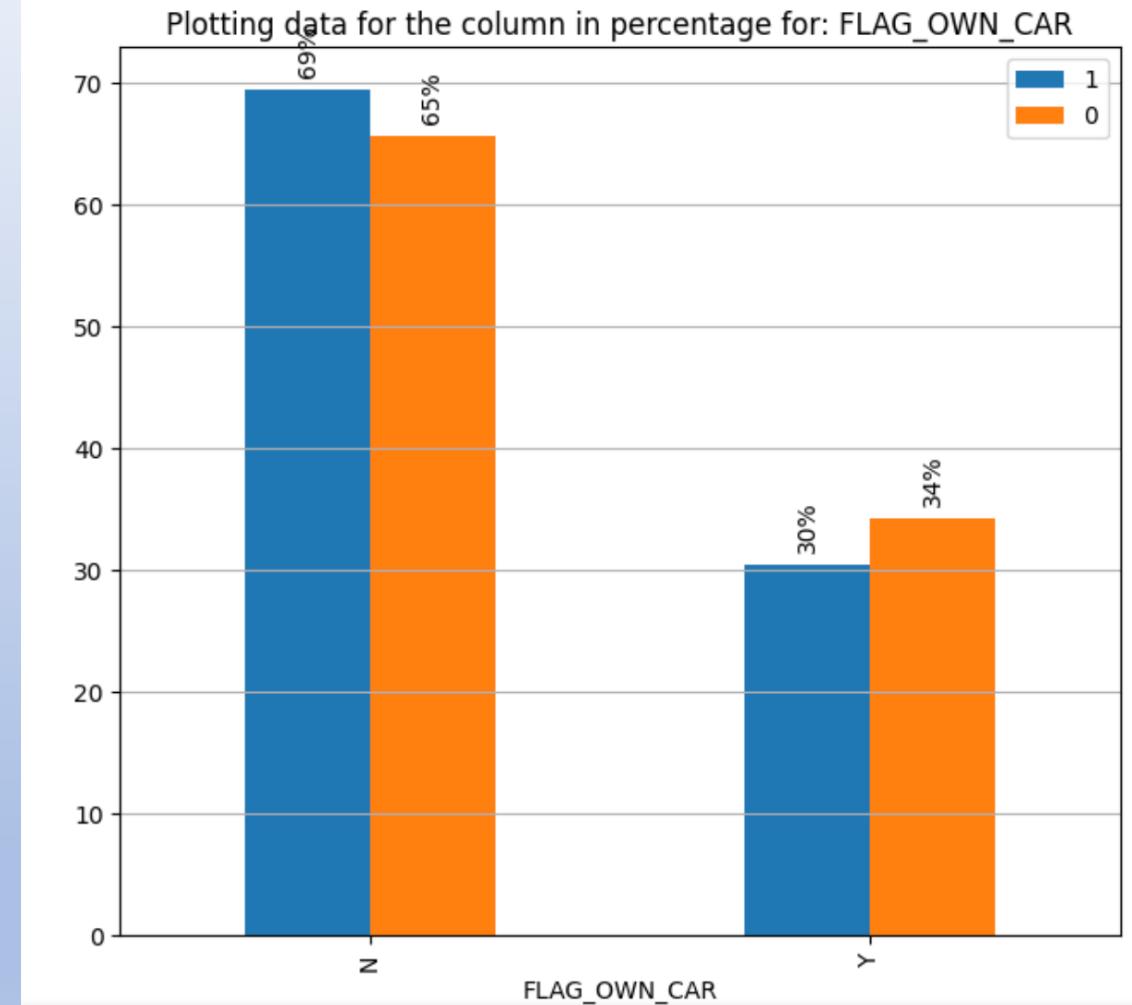
## CODE\_GENDER(MALE OR FEMALE)

- BOTH THE GROUPS FEMALE APPLICANTS ARE HIGHER THAN THE MALE APPLICANTS
- DEFAULTERS PERCENTAGE IS HIGH IN MALE GROUP



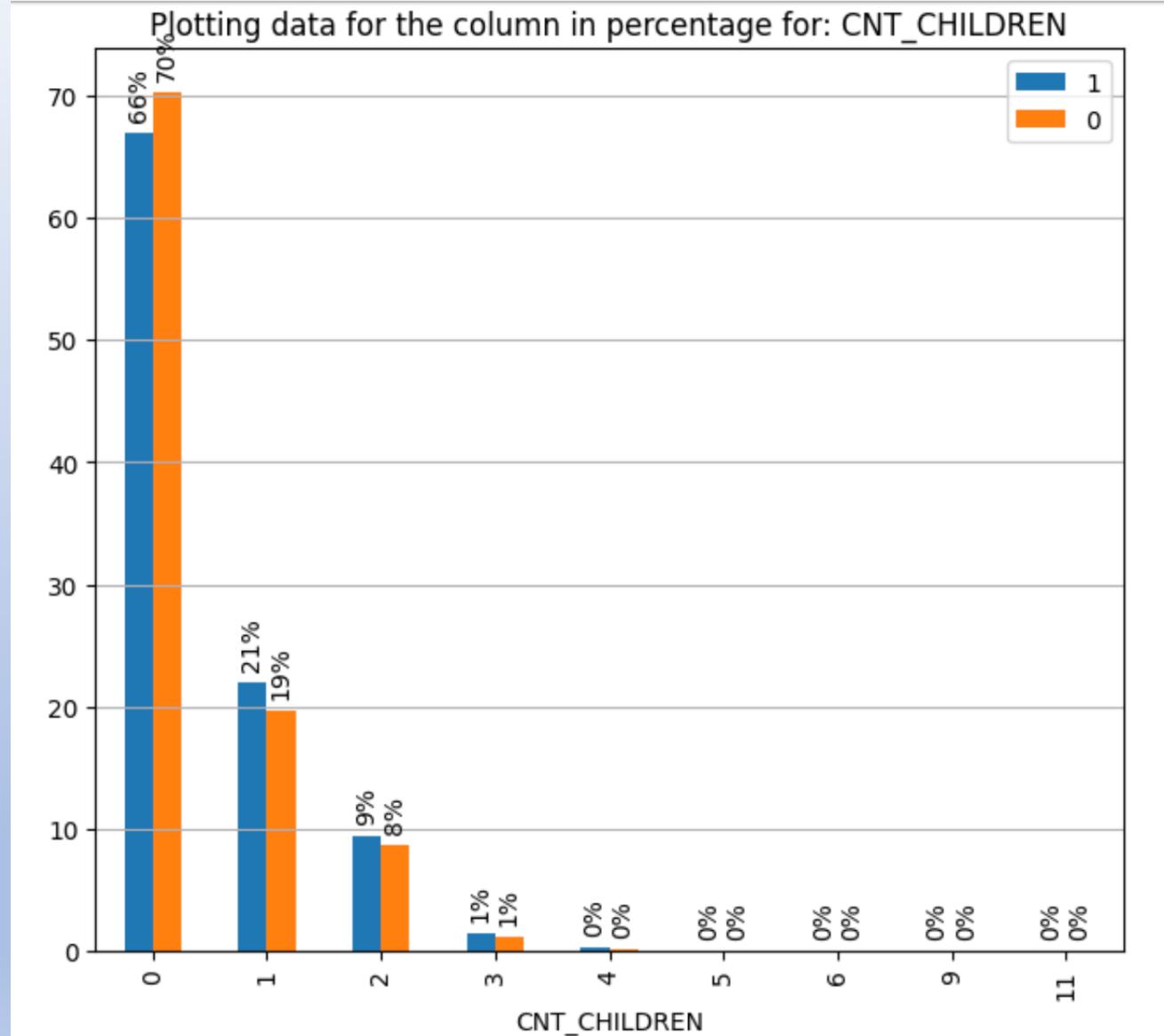
## FLAG\_OWN\_CAR(IF THE CLIENT OWNS A CAR OR NOT)

- IN BOTH THE GROUPS,CLIENTS WITHOUT CAR IS HIGH
- CLIENTS WITHOUT CAR DEFAULT MORE



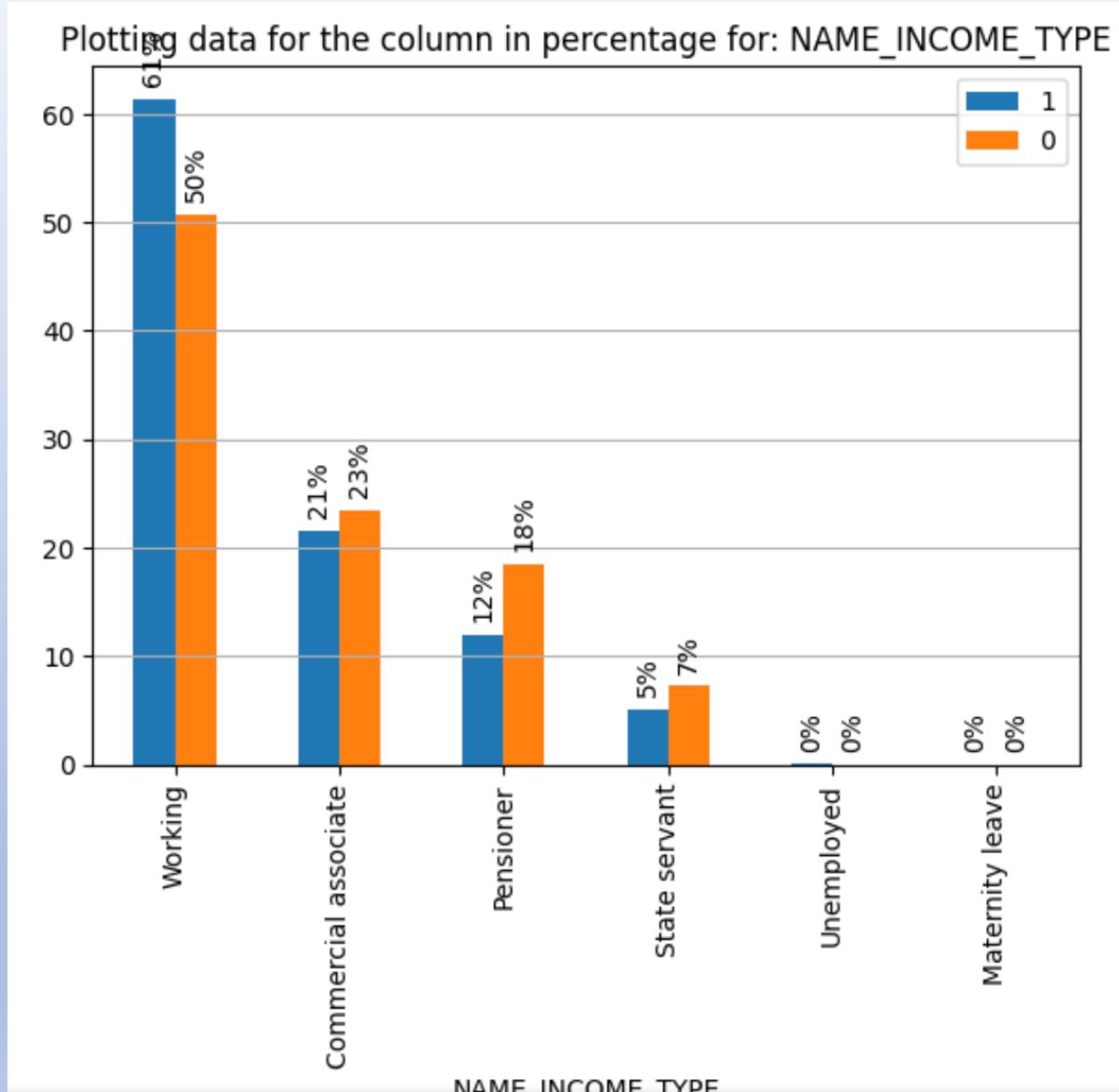
## CNT\_CHILDREN(NUMBER OF CHILDREN THE CLIENT HAS)

- MOST OF THE CLIENTS DOESN'T HAVE ANY CHILDREN
- CLIENTS WITH ONE CHILDREN TEND TO DEFAULT MORE



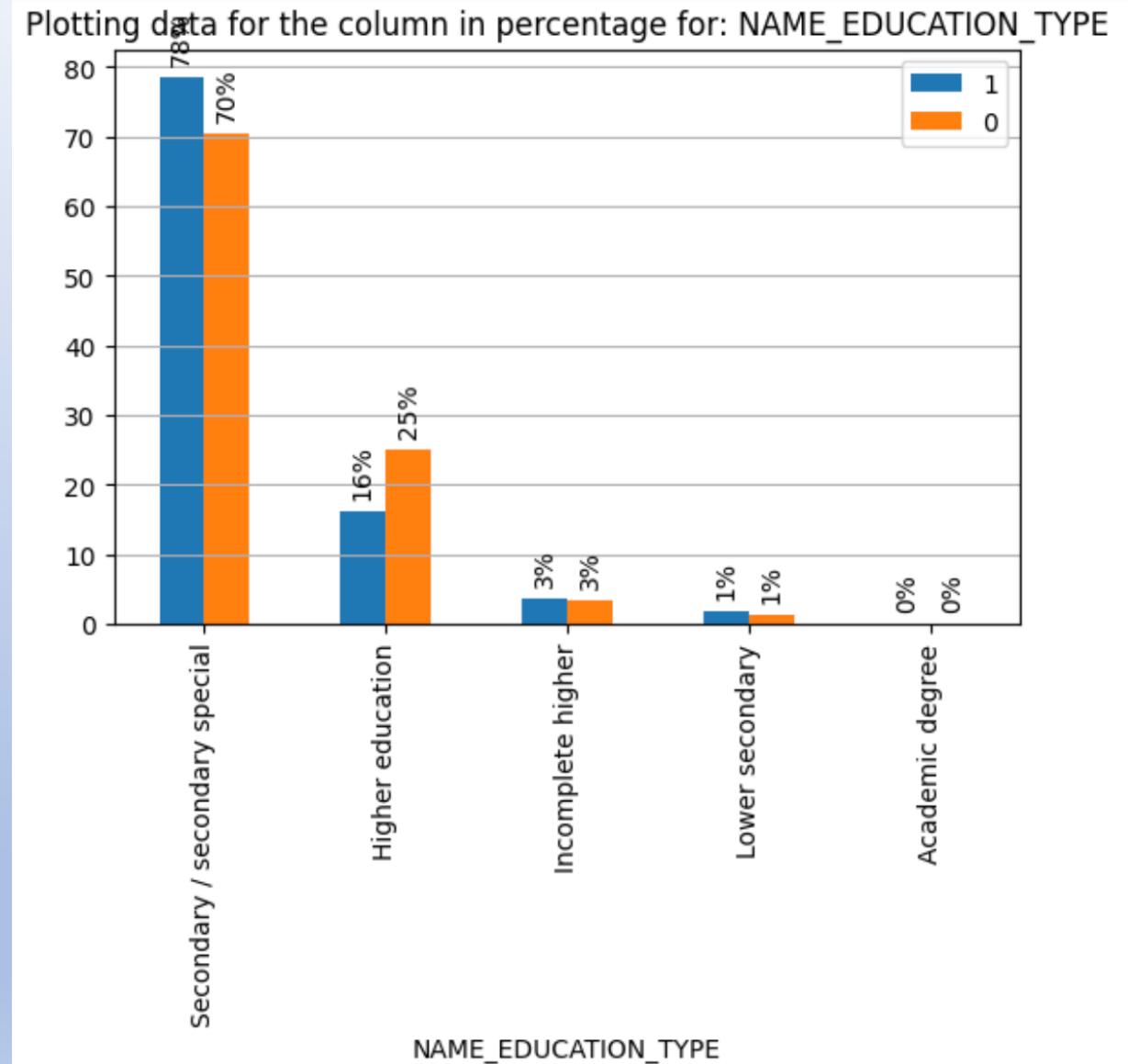
## NAME\_INCOME\_TYPE(TYPE OF THE INCOME)

- CLIENTS WITH WORKING CATEGORY TEND TO DEFAULT MORE



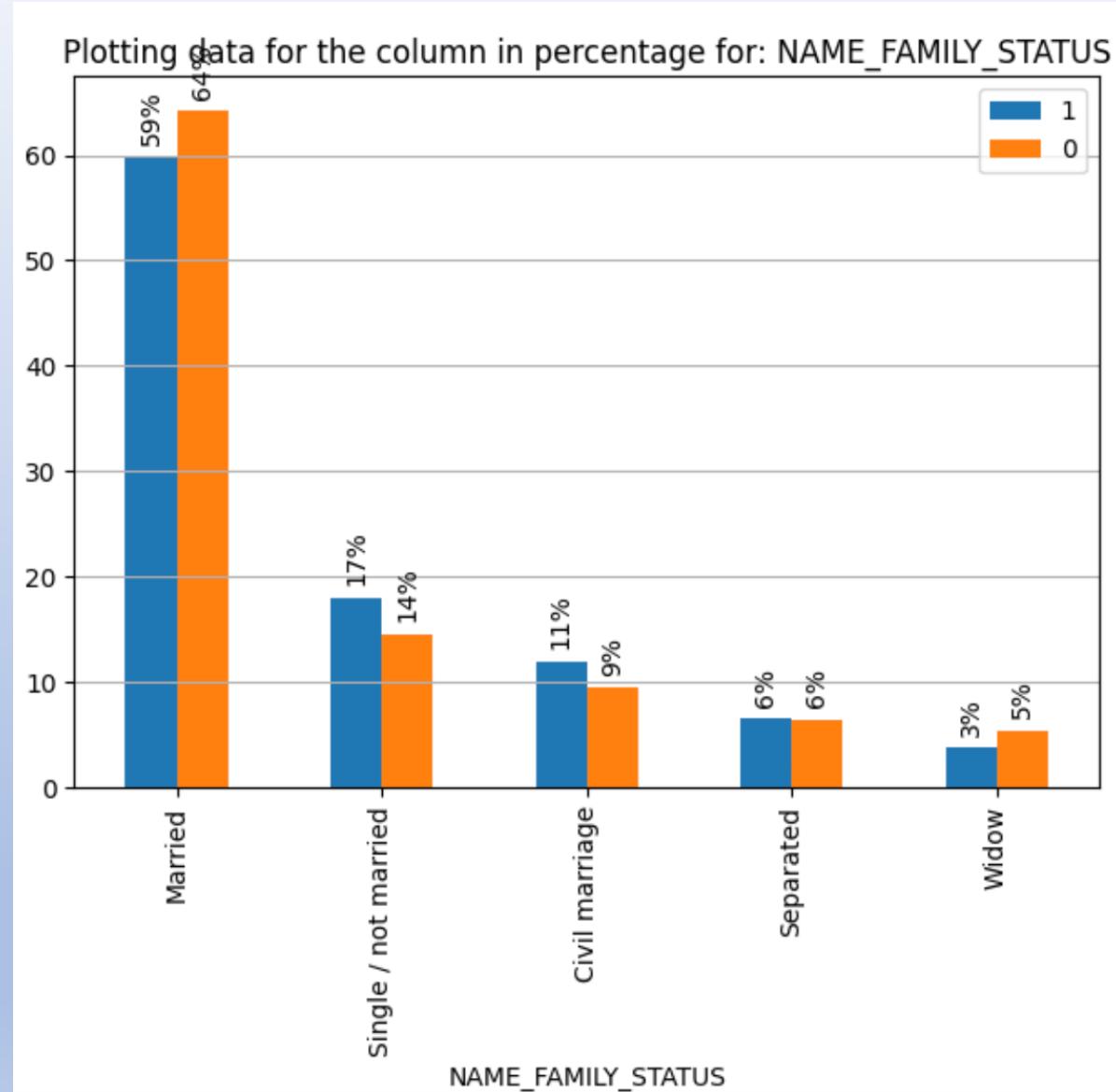
## NAME\_EDUCATION\_TYPE(LEVEL OF HIGHEST EDUCATION)

- CLIENTS WITH SECONDARY/SECONDARY SPECIAL TEND TO DEFAULT MORE



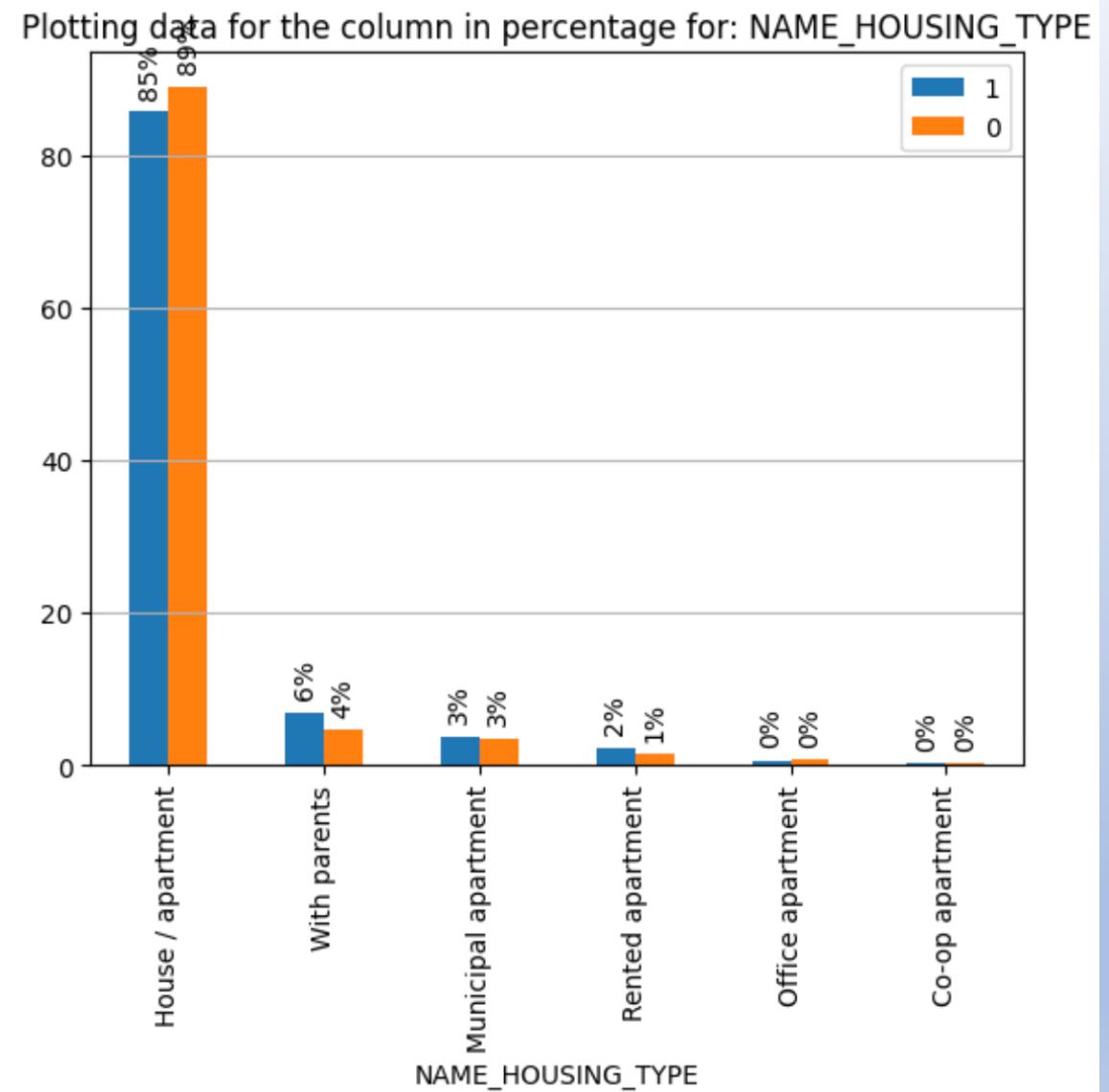
NAME\_FAMILY\_STATUS(FAMILY STATUS OF THE CLIENT)

- CLIENTS WHO ARE SINGLE/NOT MARRIED AND CIVIL MARRIAGE TEND TO DEFAULT



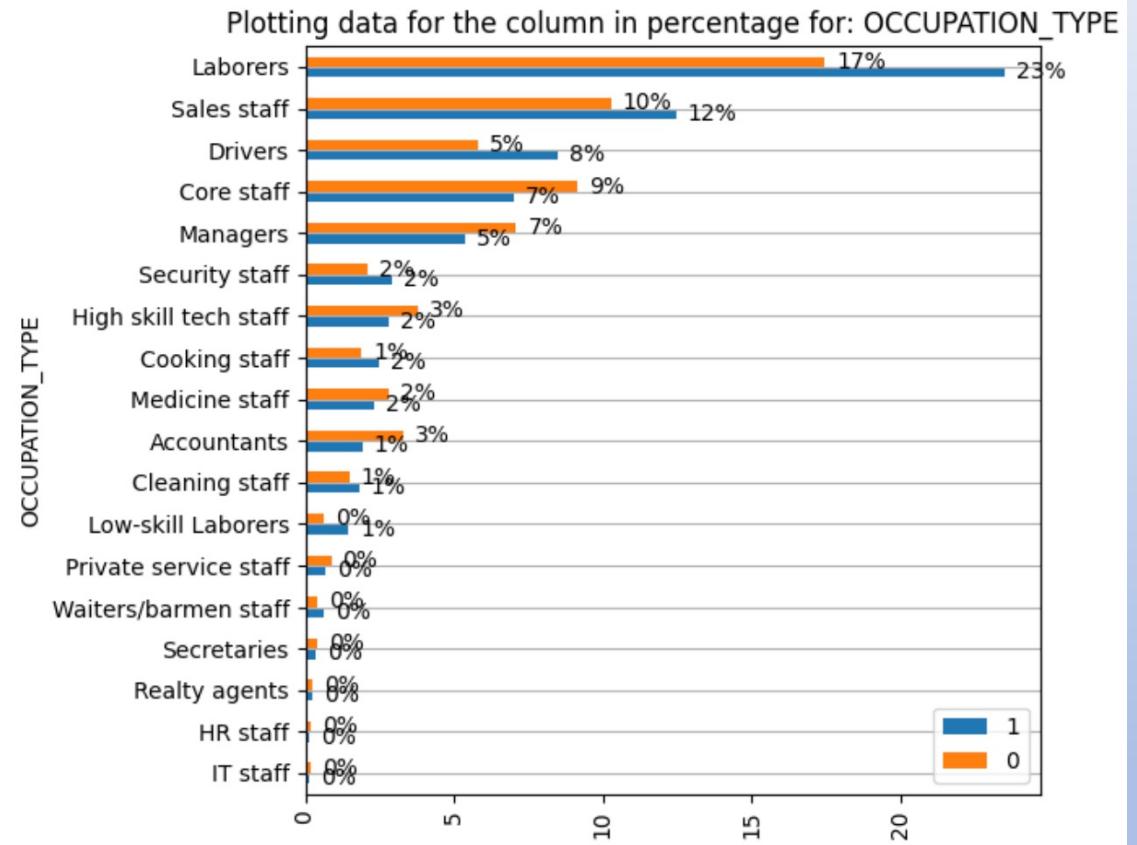
NAME\_HOUSING\_TYPE(HOUSE  
SITUATION LIKE RENTING/LIVING  
WITH PARENTS)

- MAJORITY OF CLIENTS WERE LIVING IN HOUSE/APARTMENT
- CLIENTS LIVING WITH PARENTS TEND TO DEFAULT



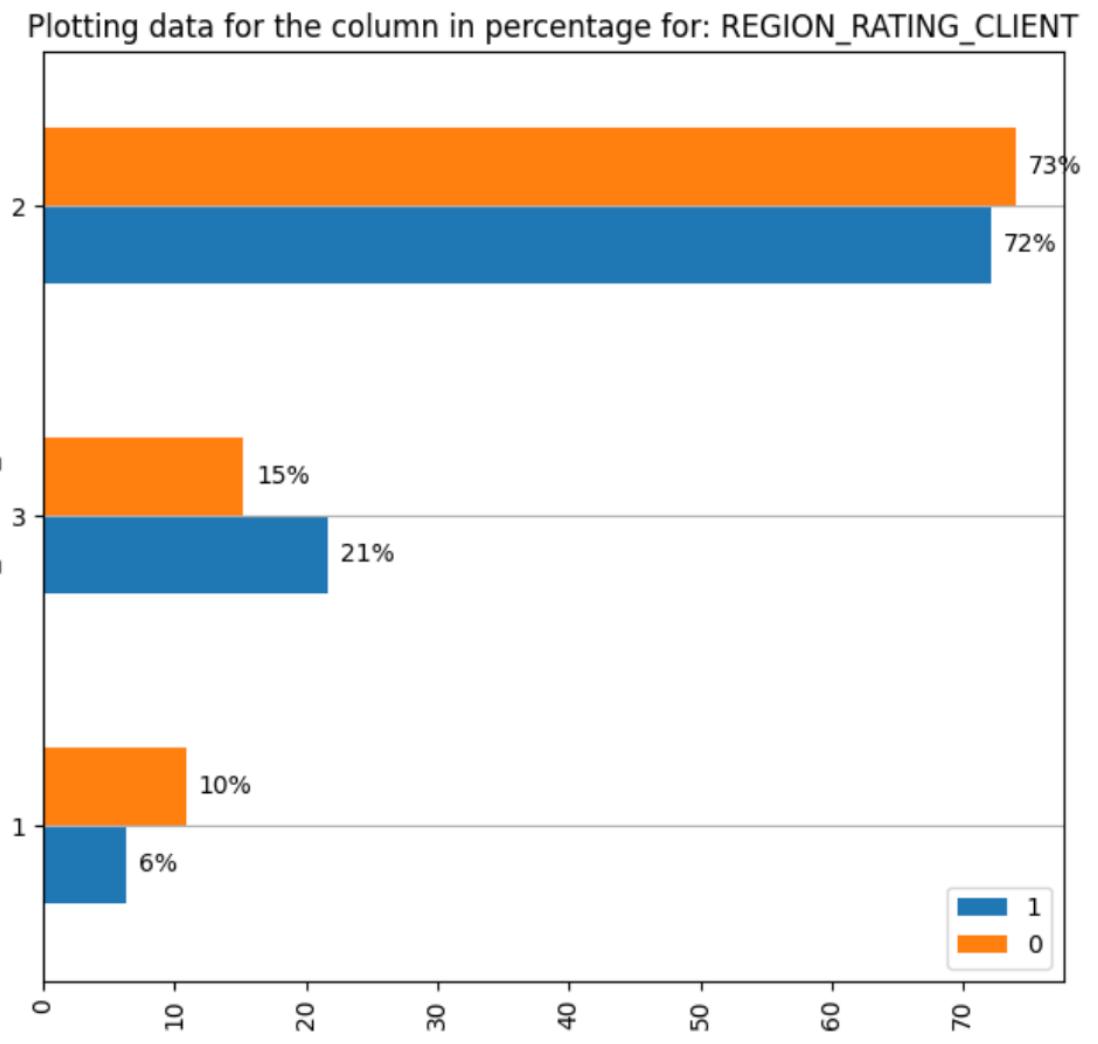
# OCCUPATION\_TYPE(TYPE OF OCCUPATION)

- CLIENTS HAVING OCCUPATION TYPE-LABORERS,SALES STAFF,DRIVERS,COOKING STAFF AND LOW SKILLED LABORERS TEND TO DEFAULT



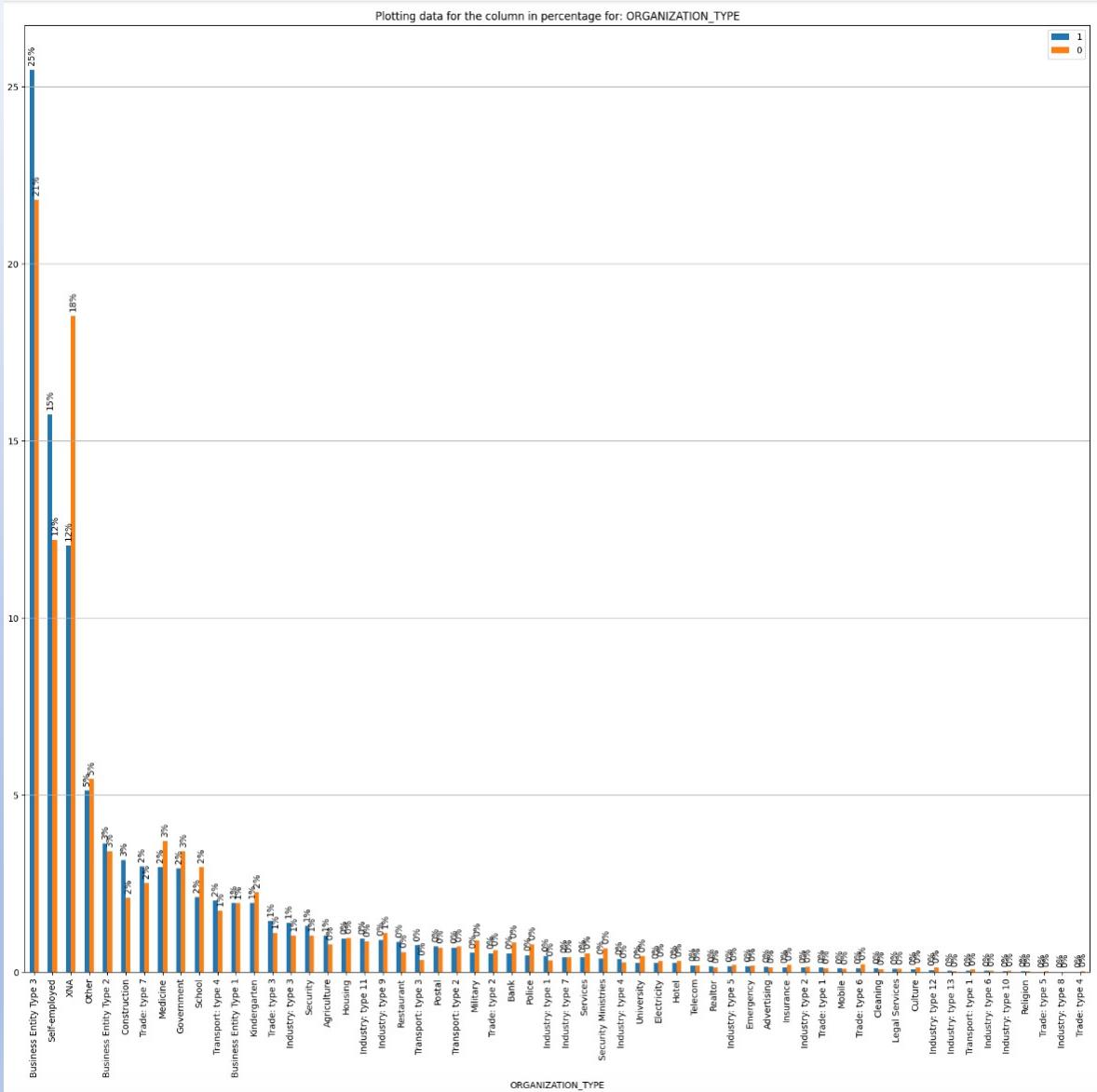
## REGION\_RATING\_CLIENT(RATING OF THE LIVING REGION)

- CLIENTS LIVING IN THE REGION RATING=3 TEND TO DEFAULT



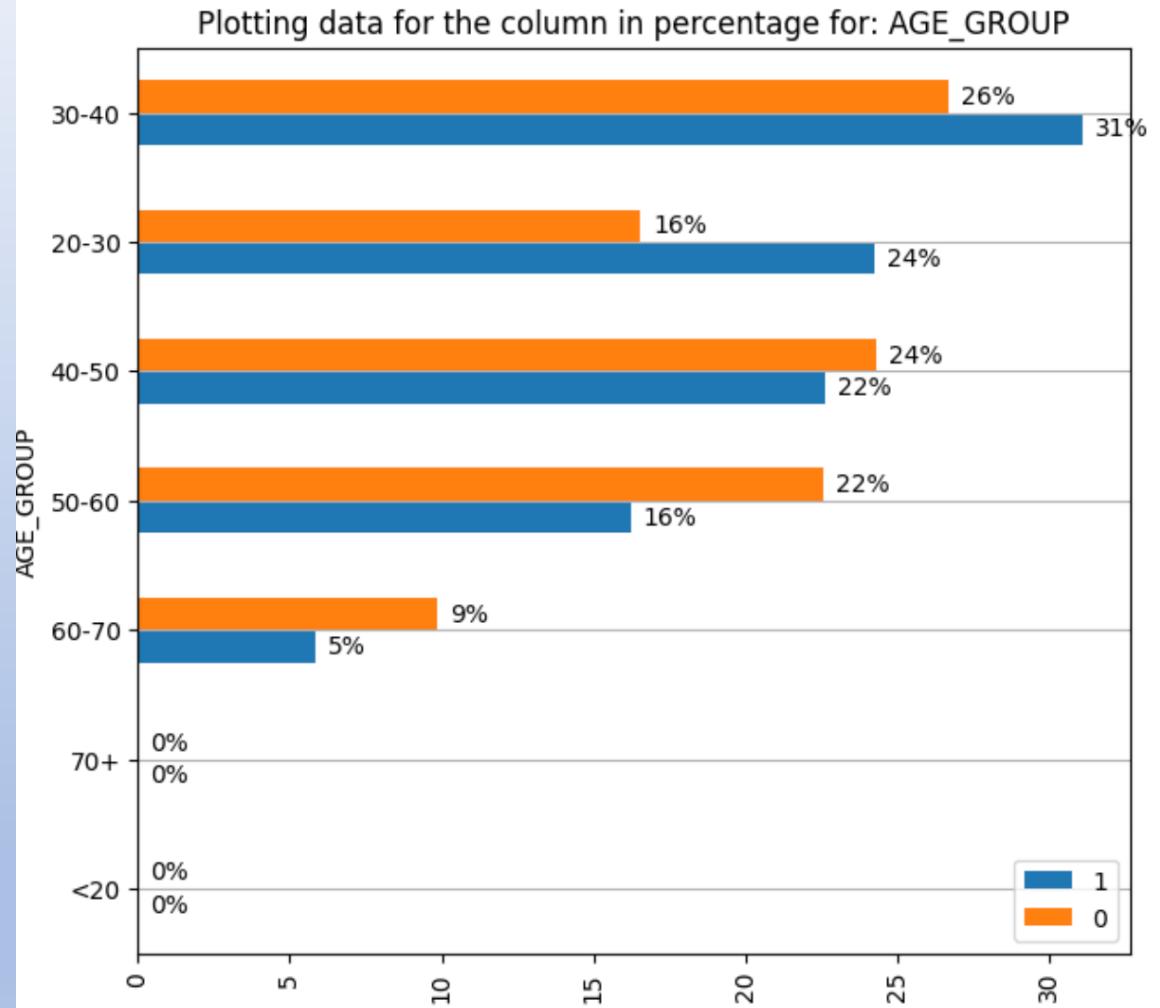
## ORGANIZATION\_TYPE(TYPE OF WORKING ORGANIZATION)

- BUSINESS ENTITY TYPE3,SELF EMPLOYED AND CONSTRUCTION BASED CLIENTS TEND TO DEFRAUD



# AGE\_GROUP(AGE GROUP OF THE CLIENTS)

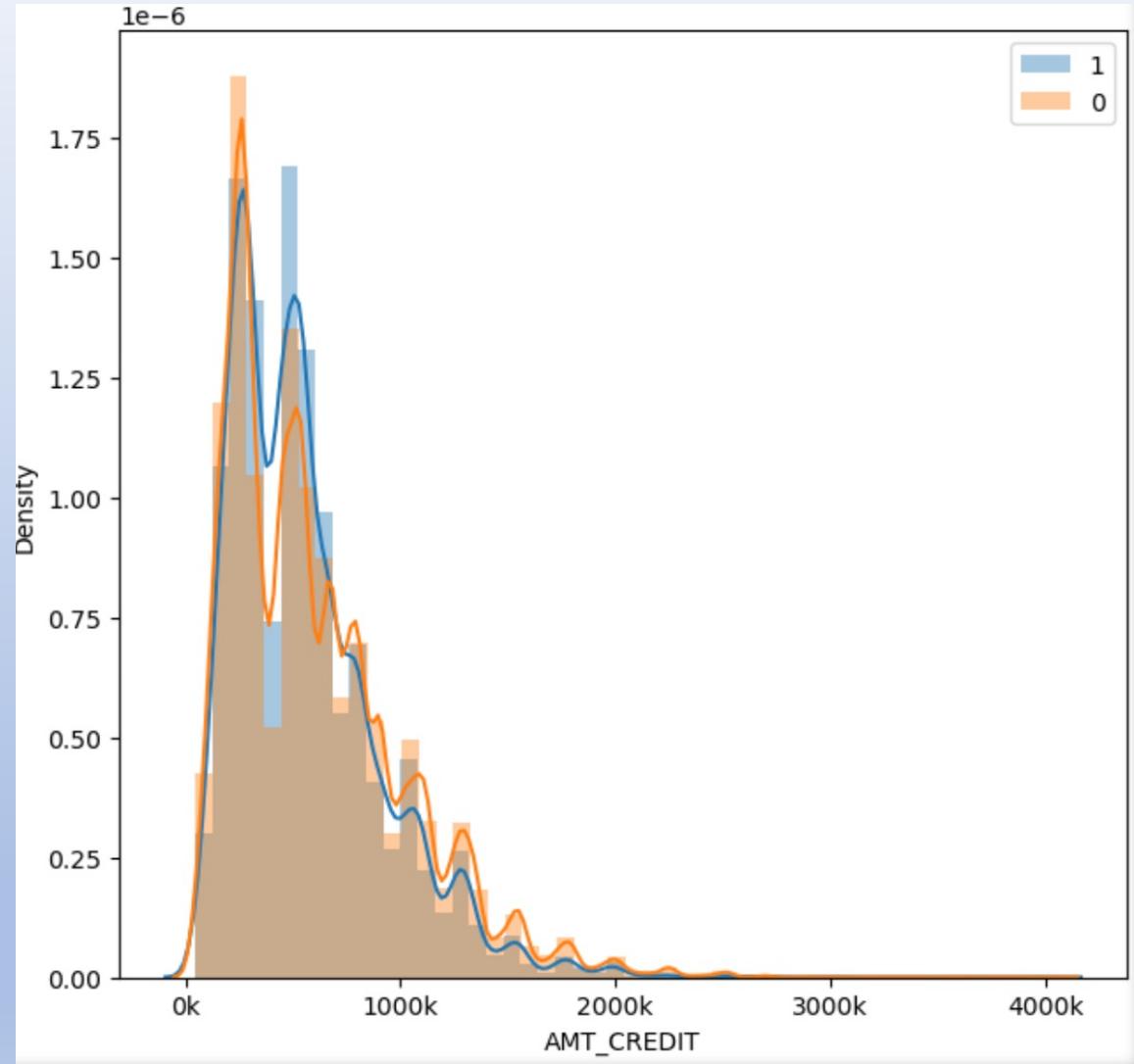
- CLIENTS IN THE AGE GROUP 20-40 TEND TO DEFAULT



# NUMERICAL UNIVARIATE ANALYSIS(HISTOGRAM AND DISTPLOT AND BOX PLOT)

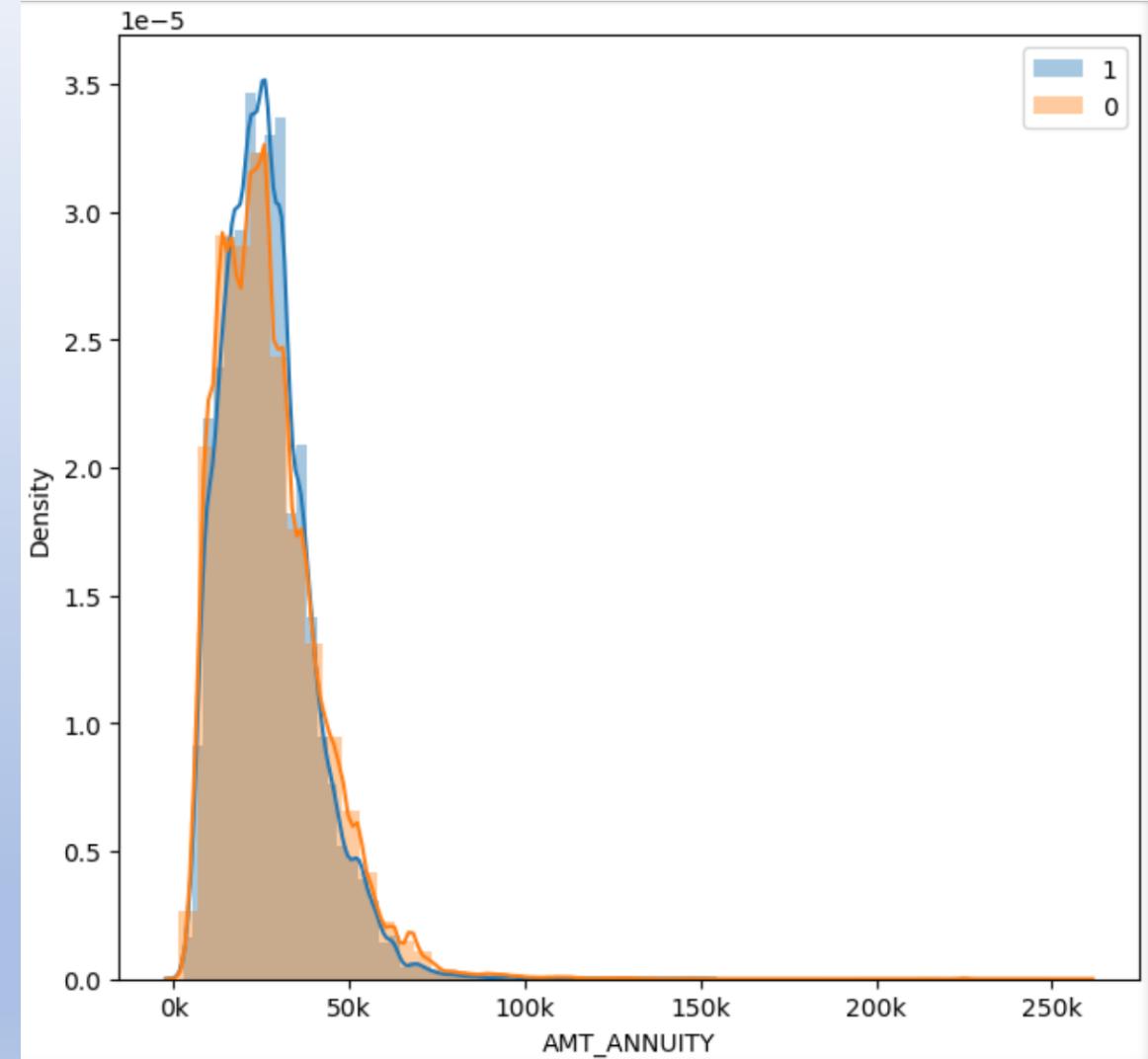
# AMT\_CREDIT(FINAL CREDIT AMOUNT IN THE APPLICATION)

- CLIENTS WHO TEND TO DEFAULT HAVE CREDIT AMOUNT IN THE RANGE OF 50K TO 2000K



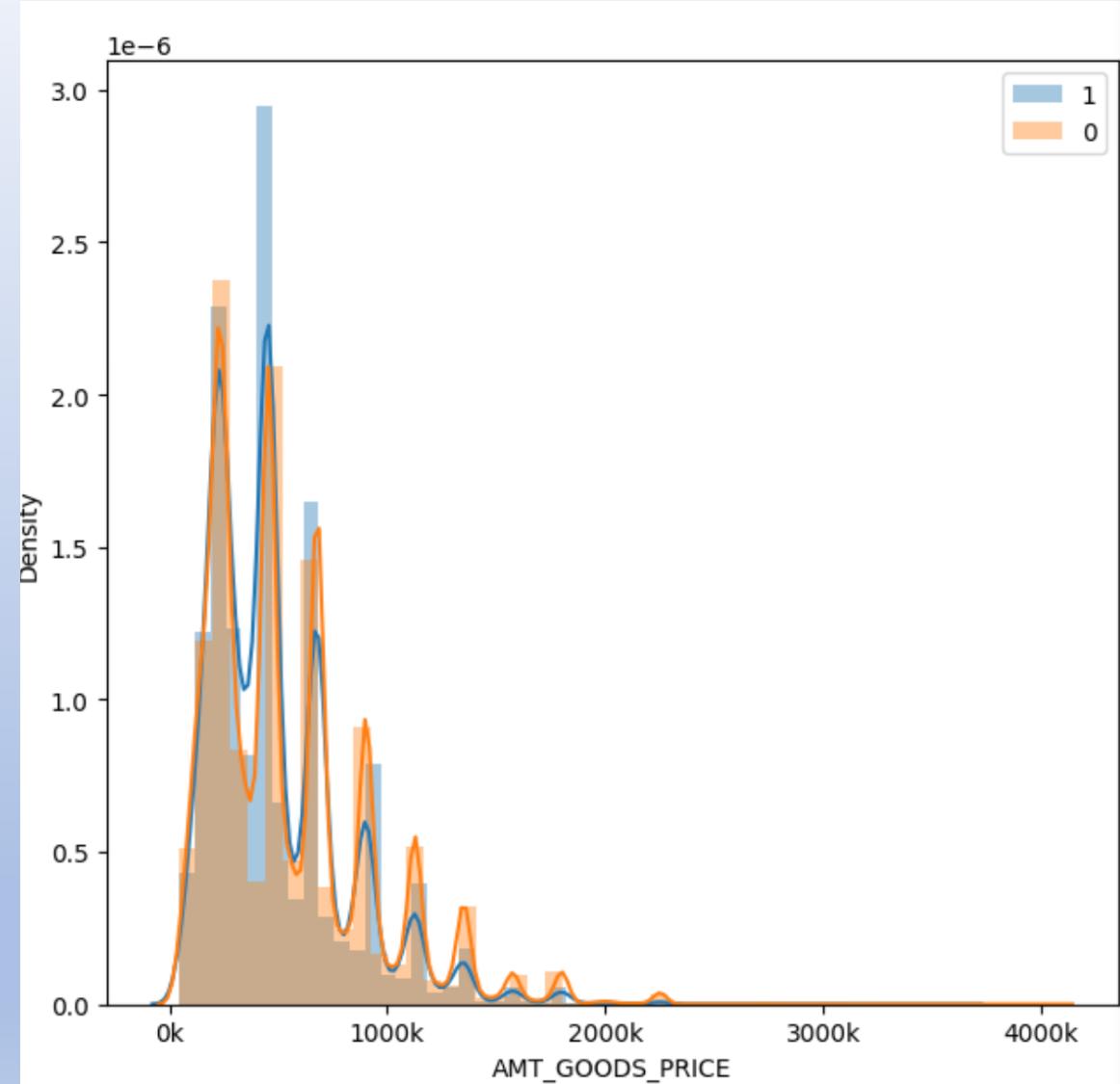
# AMT\_ANNUITY(LOAN ANNUITY-THE FIXED PAYMENY LOAN)

- CLIENT WHO TEND TO DEFAULT HAVE ANNUITY AMOUNT IN THE RANGE OF 2K-100K



## AMT\_GOODS\_PRICE(PRICE OF GOODS FOR WHICH THE LOAN IS GIVEN)

- CLIENTS WHO TEND TO DEFAULT HAVE  
GOODS PRICE IN THE RANGE OF 45K-  
2000K

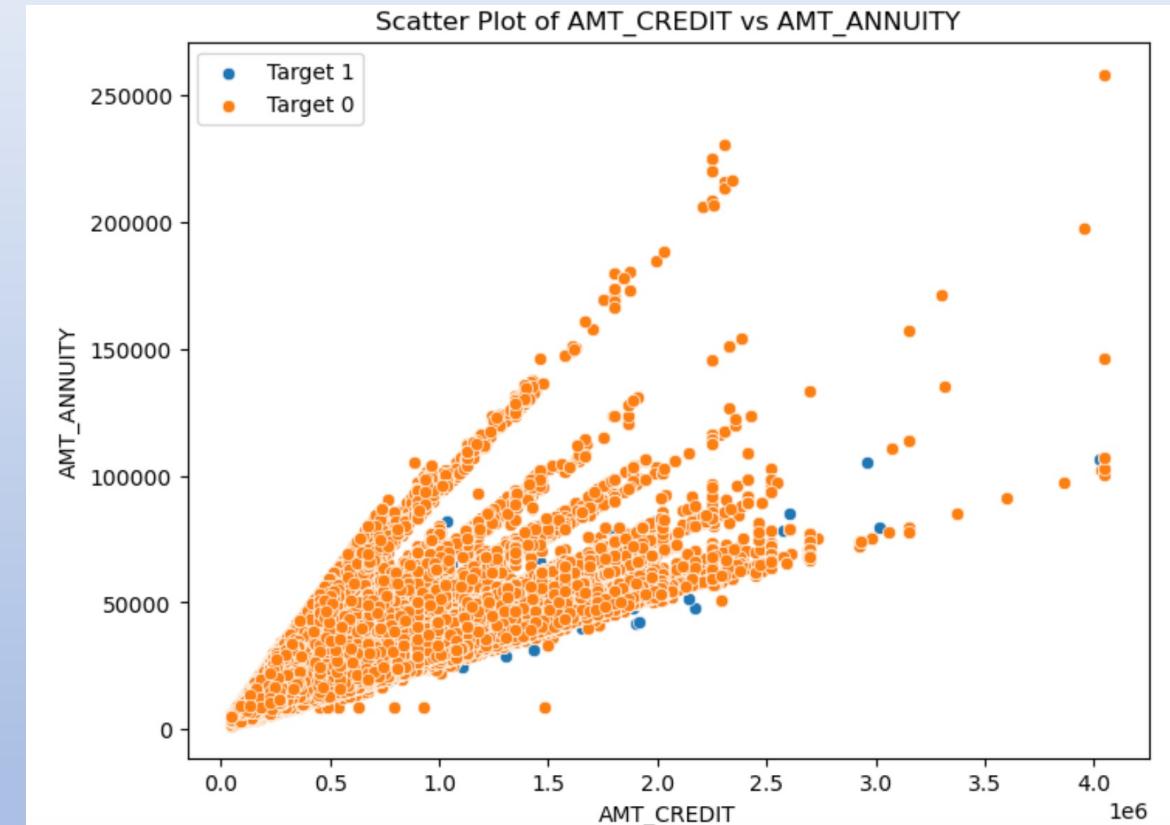


# BIVARIATE ANALYSIS

# NUMERICAL VS NUMERICAL (SCATTER PLOT)

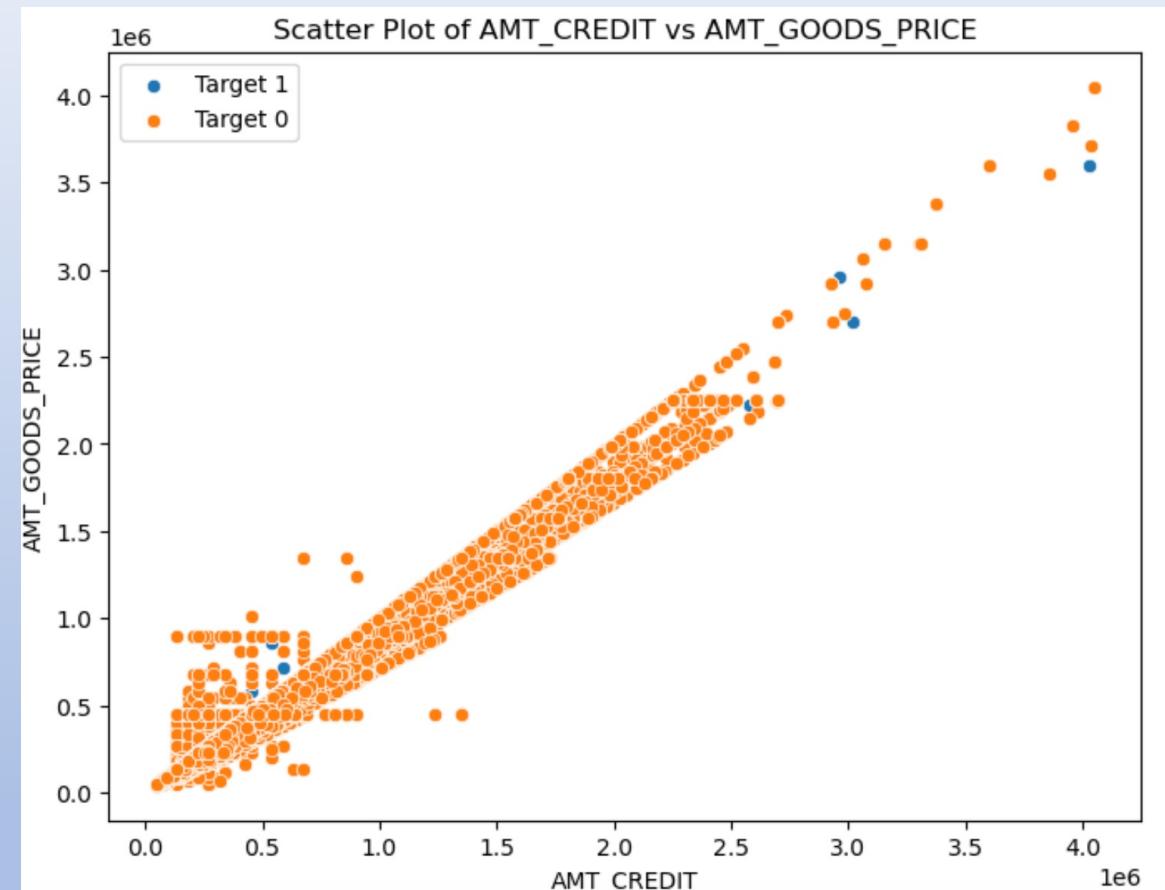
# AMT\_CREDIT VS AMT\_ANNUITY

- BOTH THE VARIABLES HAVE POSITIVE CORRELATION(IF ONE INCREASE, THE OTHER ALSO INCREASES)



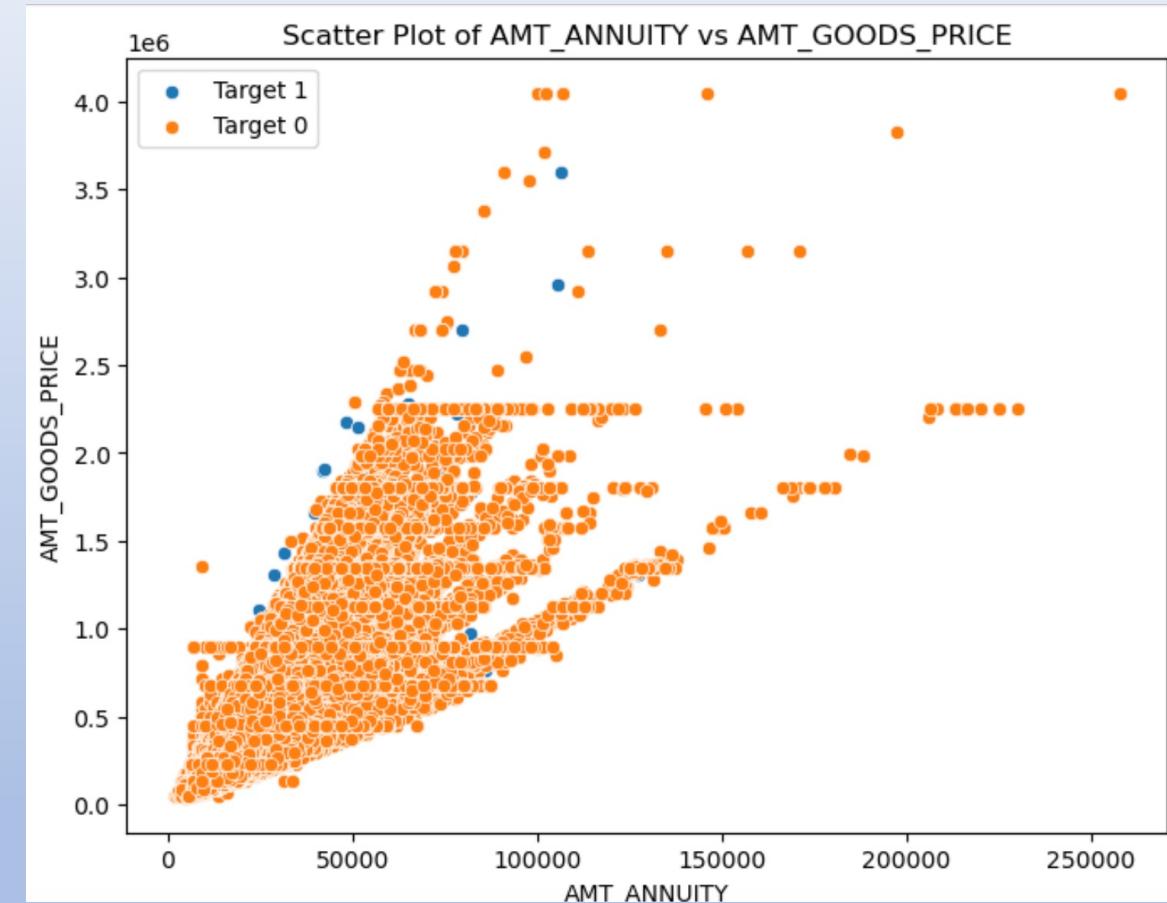
# AMT\_CREDIT VS AMT\_GOODS\_PRICE

- BOTH THE VARIABLES HAVE POSITIVE CORRELATION



# AMT\_ANNUITY VS AMT\_GOODS\_PRICE

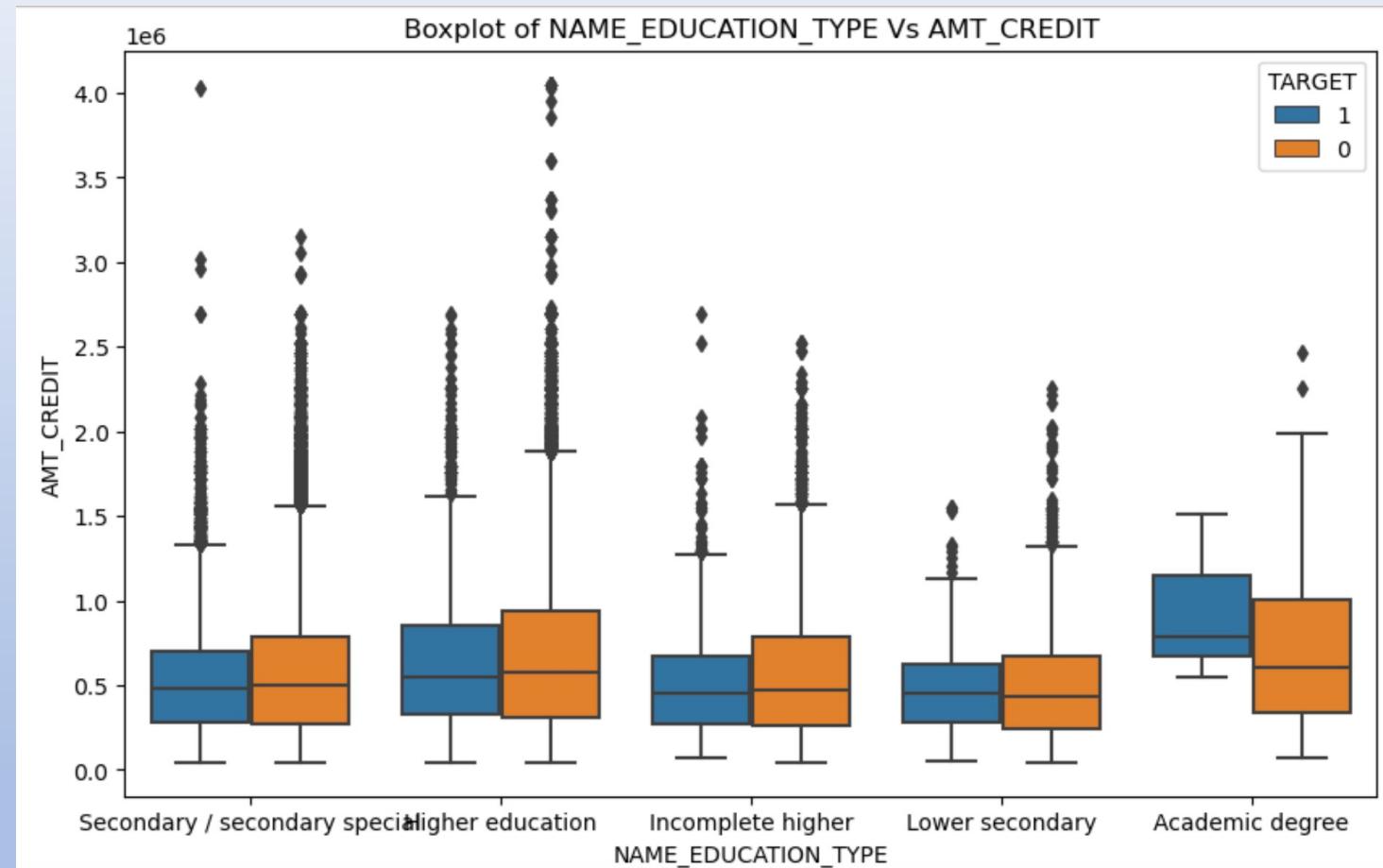
- BOTH THE VARIABLES HAVE POSITIVE CORRELATION



# NUMERICAL VS CATEGORICAL (BOX PLOT)

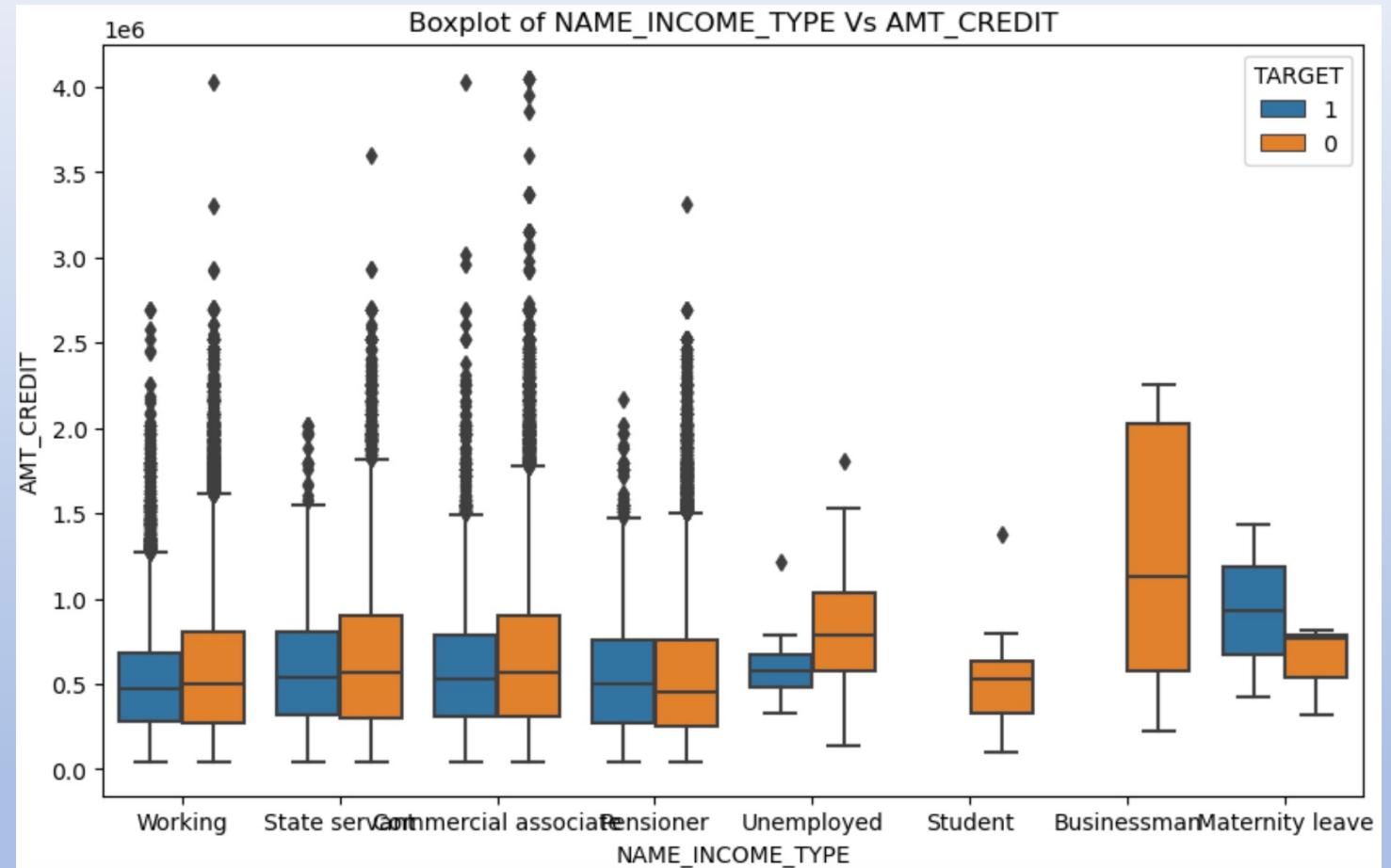
## NAME\_EDUCATION\_TYPE VS AMT\_CREDIT

- CREDIT AMOUNT INCREASES AS THE EDUCATION LEVEL INCREASES
- SECONDARY AND HIGHER EDUCATION CATEGORY HAVE BIGGER OUTLIER WHICH LOOKS GENUINE



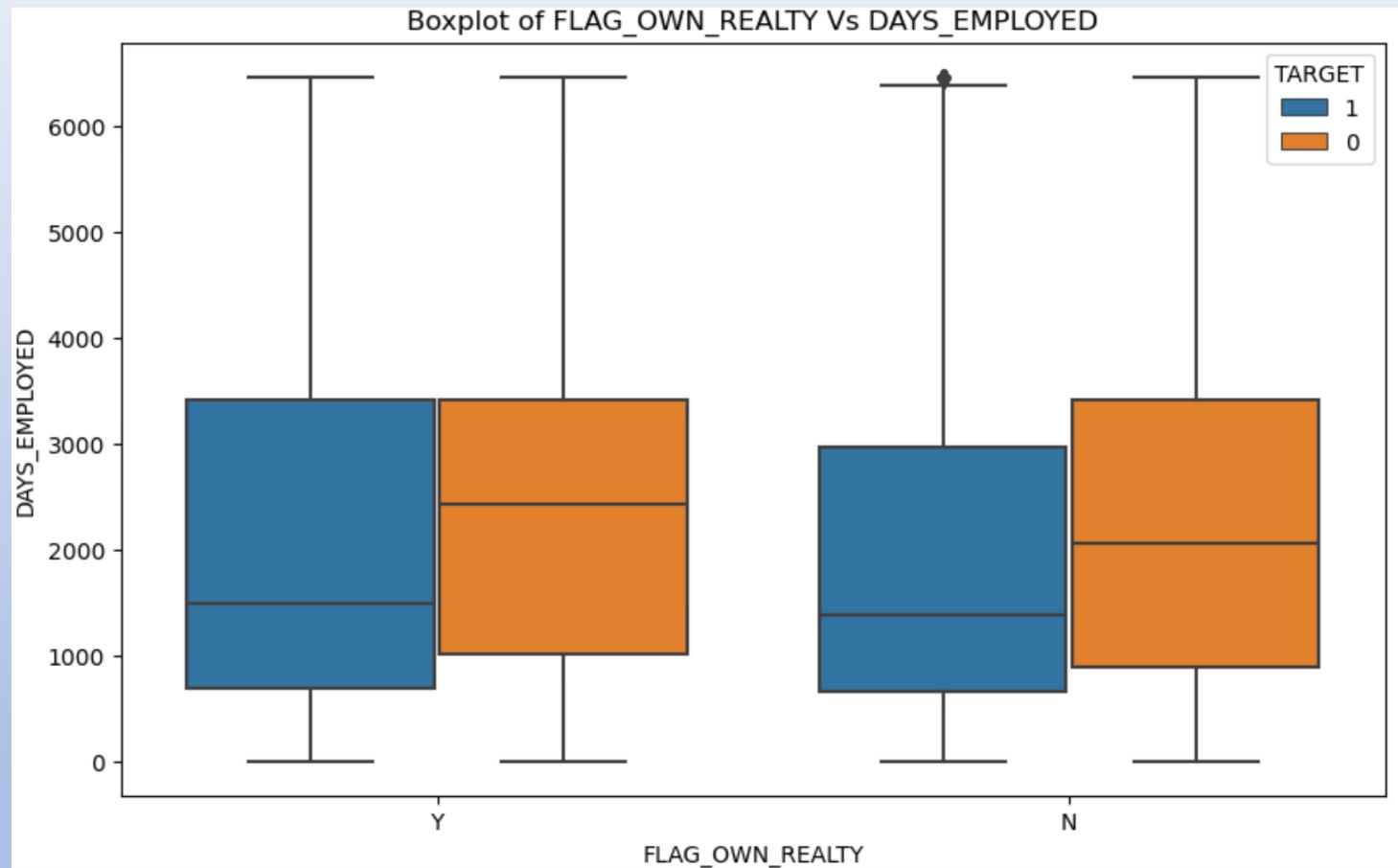
## NAME\_INCOME\_TYPE VS AMT\_CREDIT

- BUSINESSMAN CATEGORY HAS HIGH CREDIT AMOUNT
- THERE ARE NO DEFAULTERS IN STUDENT AND BUSINESSMAN CATEGORY



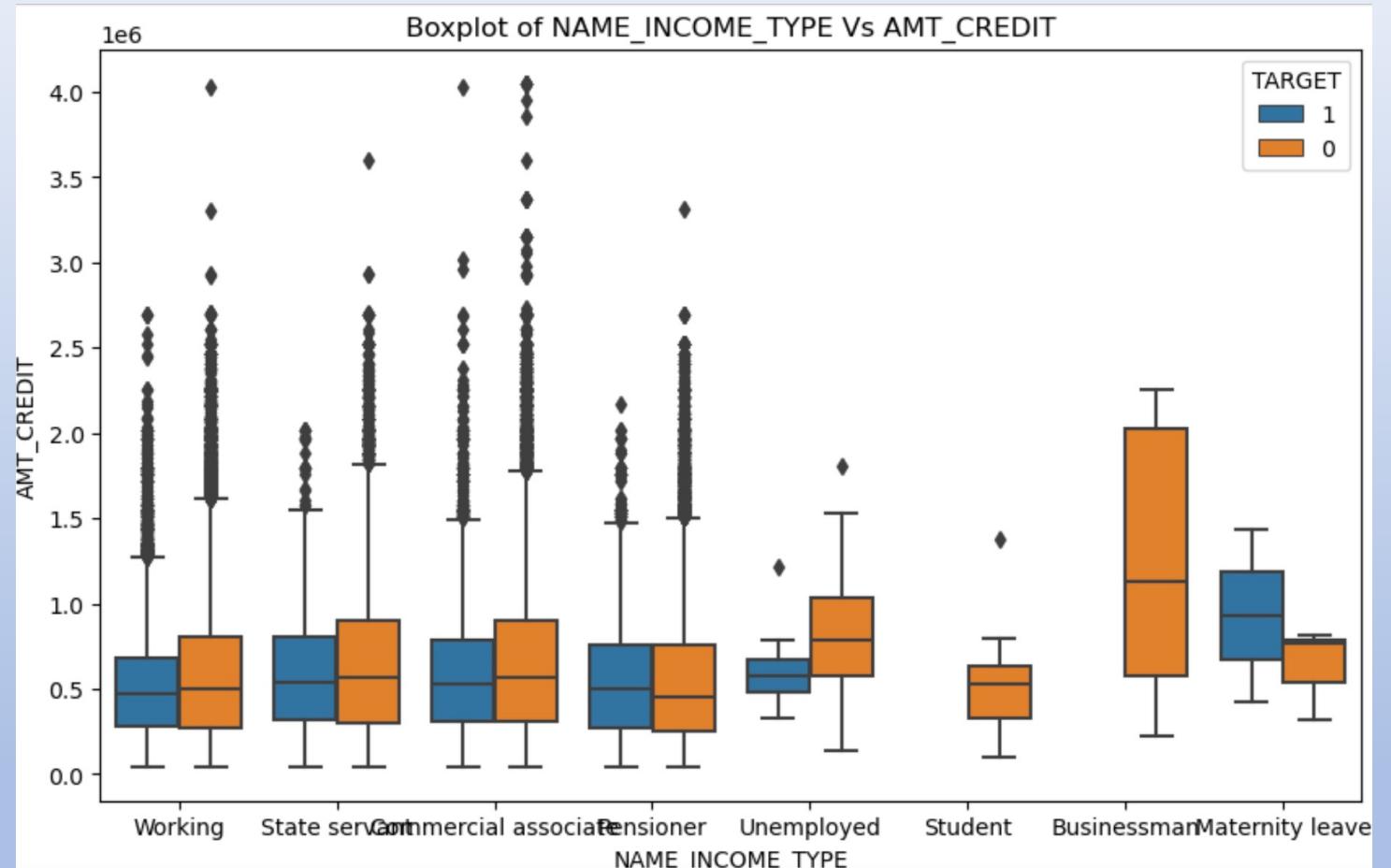
## DAYS\_EMPLOYED VS FLAG\_own\_REALTY

- THE MEDIAN VALUE OF NUMBER OF DAYS EMPLOYED IS LOW FOR DEFAULTERS



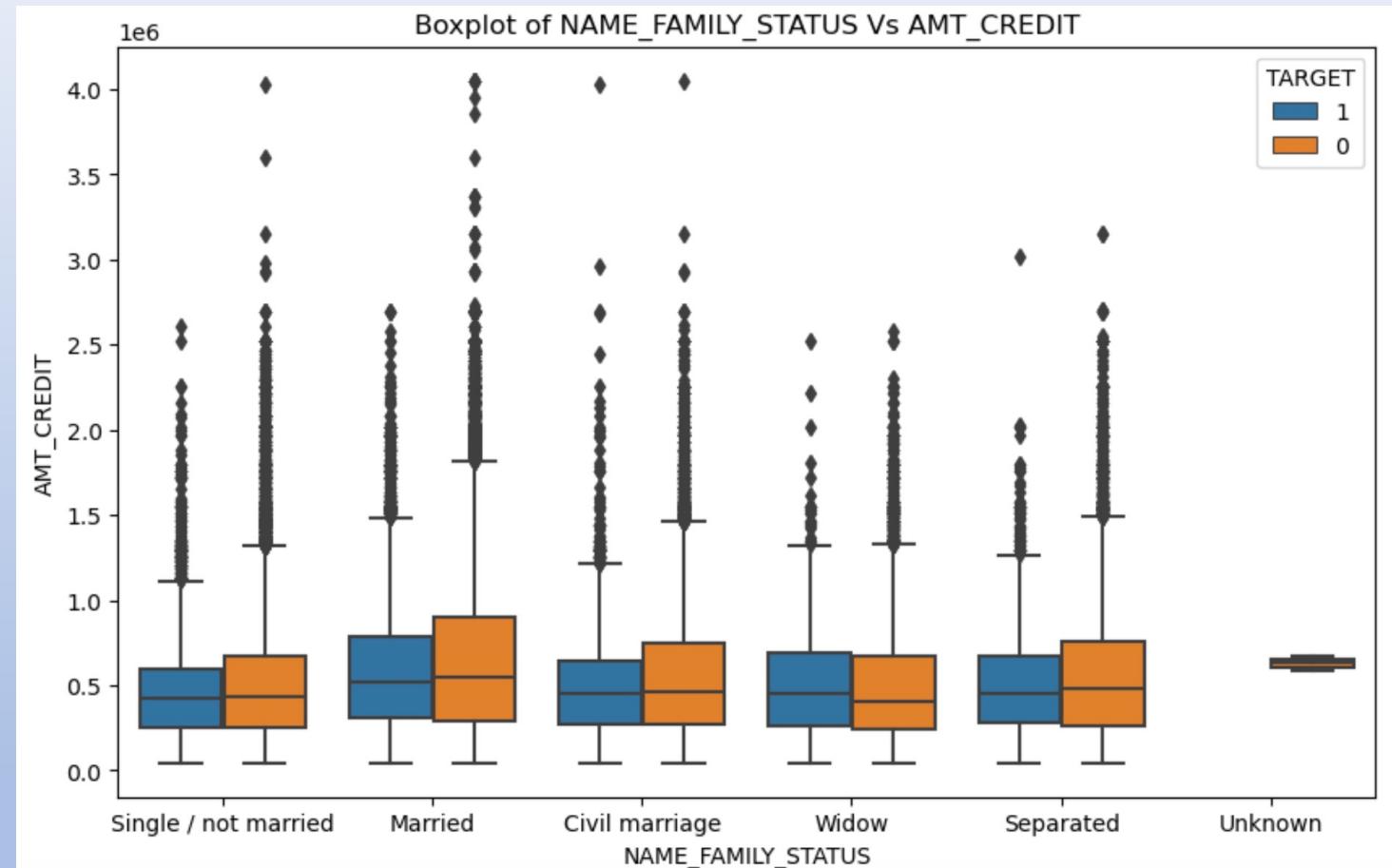
## NAME\_INCOME\_TYPE VS AMT\_CREDIT

- CREDIT AMOUNT FOR DEFAULTERS WHO ARE ON MATERNITY LEAVE IS HIGH



## AMT\_CREDIT VS NAME\_FAMILY\_STATUS

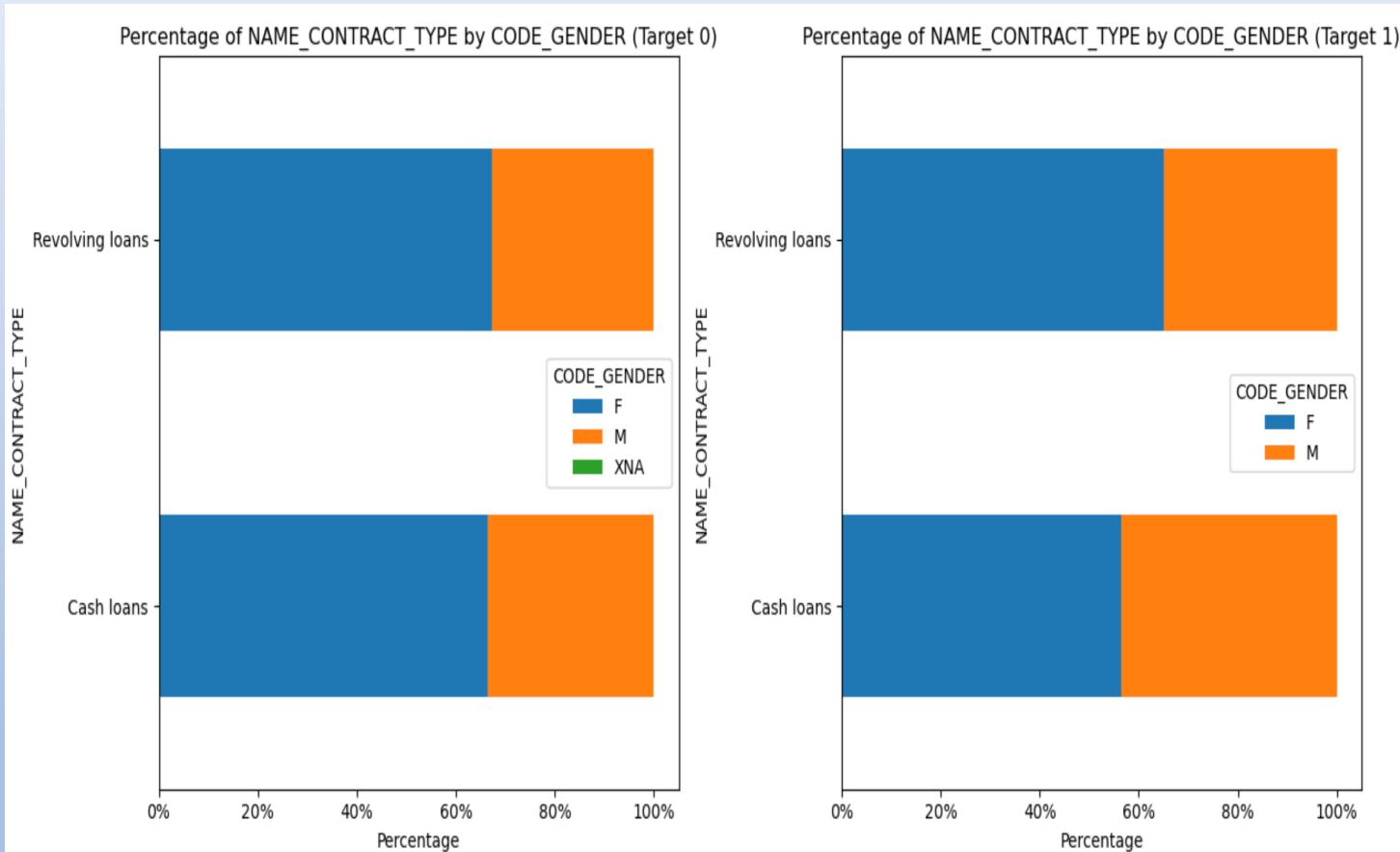
- MARRIED PEOPLE HAVE HIGH AVERAGE CREDIT



CATEGORICAL VS CATEGORICAL

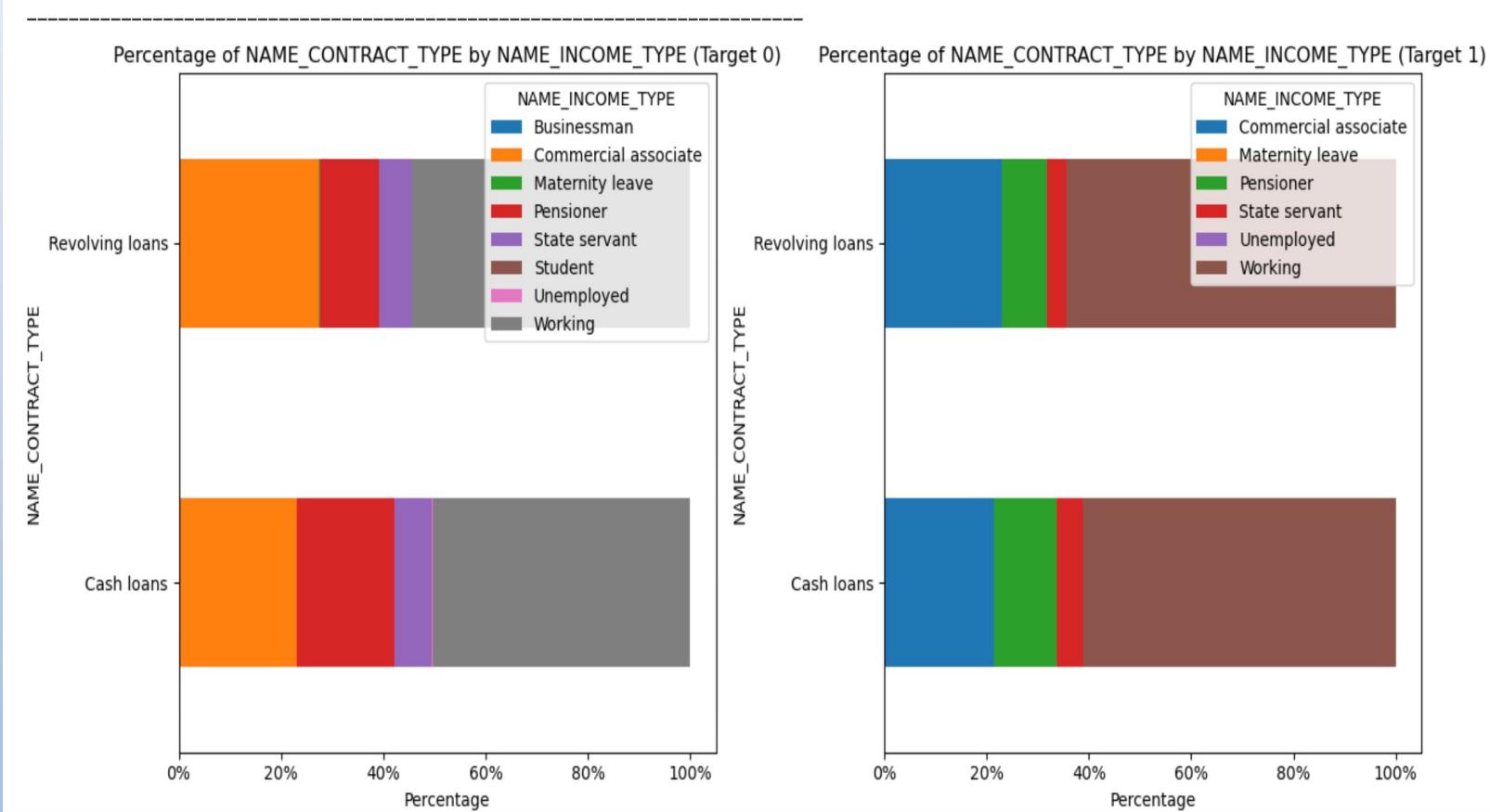
# CONTRACT\_TYPE VS CODE\_GENDER

- BOTH THE CASH LOANS AND REVOLVING LOANS IN BOTH DEFAULTERS AND NON-DEFAULTERS, FEMALE NUMBER IS HIGH



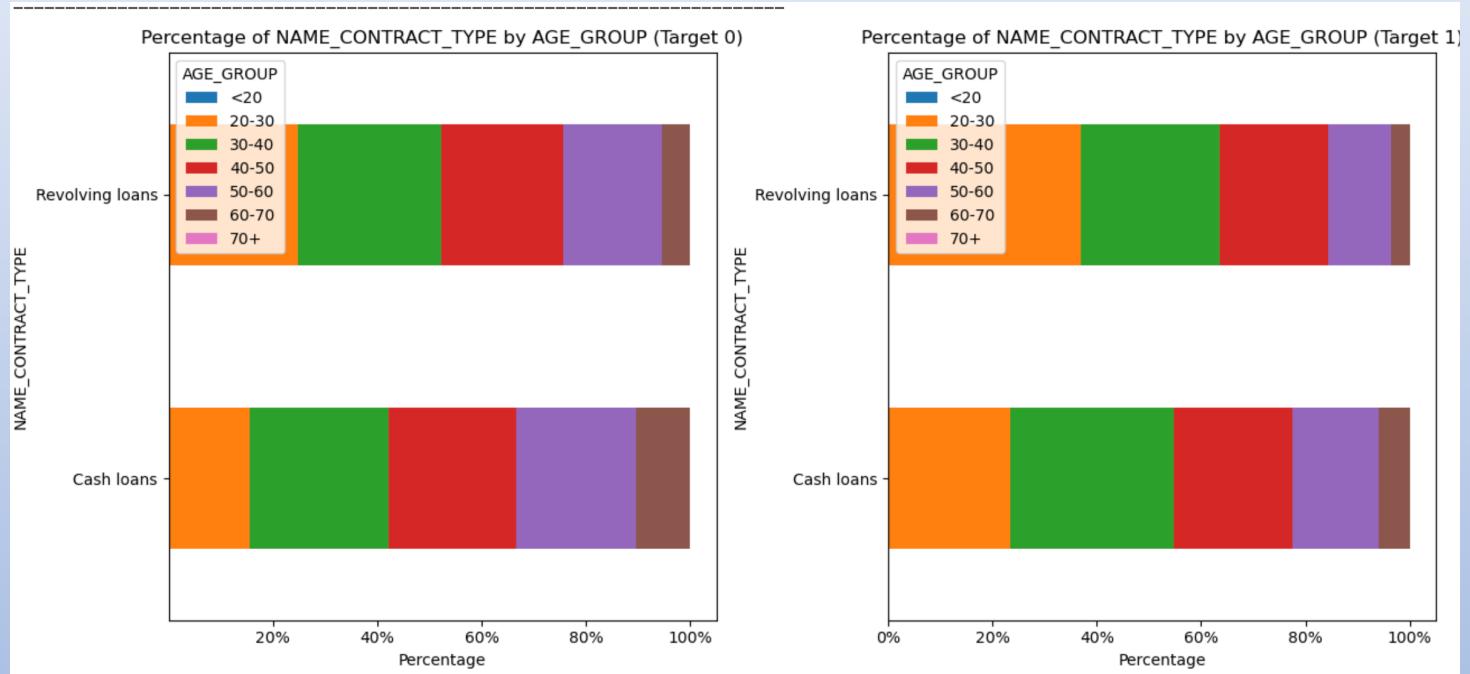
# CONTRACT\_TYPE VS NAME\_INCOME\_TYPE

- COMMERCIAL ASSOCIATES ARE HIGH IN BOTH THE CATEGORIES



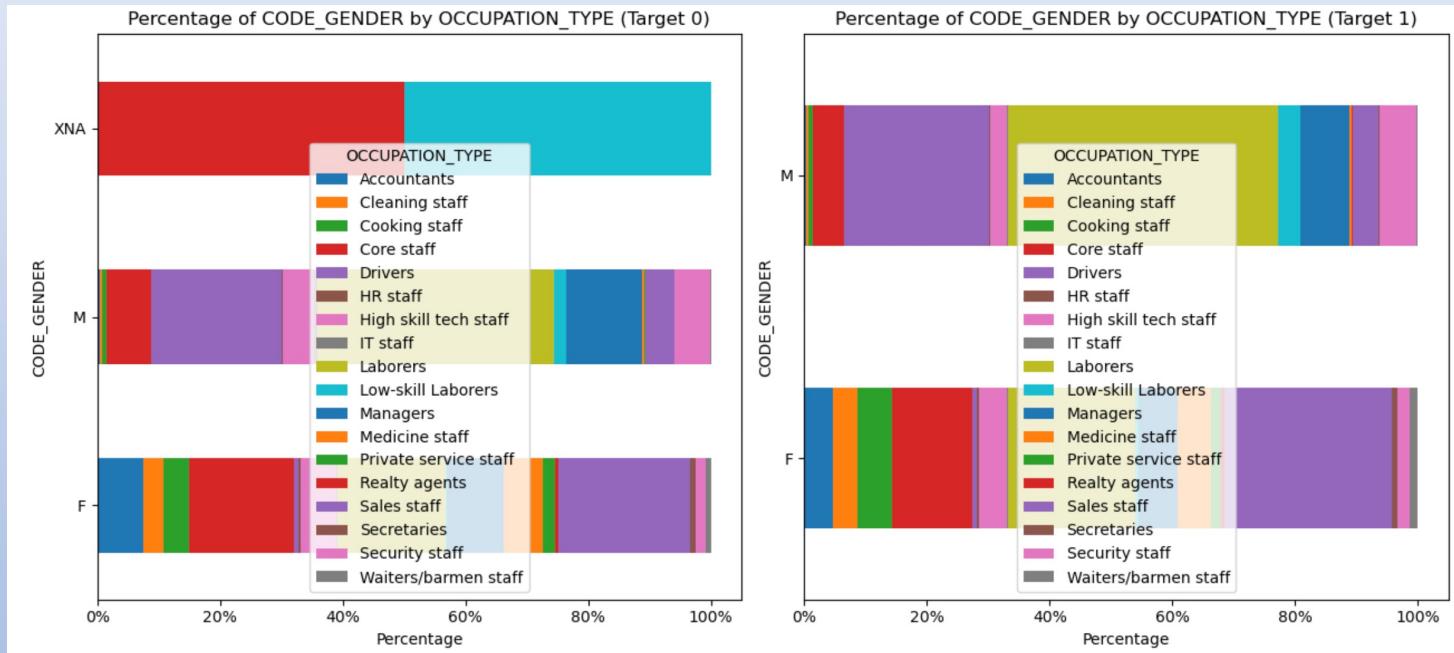
# CONTRACT\_TYPE VS AGE\_GROUP

- CLIENTS IN THE REVOLVING LOAN AND IN THE AGE GROUP 20-30 TEND TO DEFAULT MORE



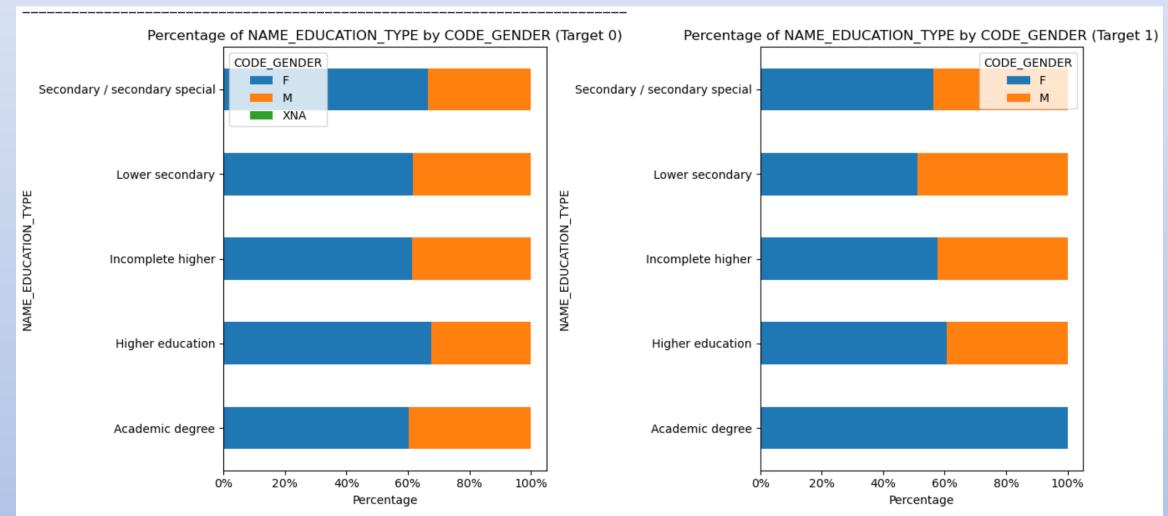
# CODE\_GENDER VS OCCUPATION\_TYPE

- IN DEFAULTERS' MALE AND FEMALE, LABORERS AND SALES STAFF TEND TO DEAFAULT MORE



# NAME\_EDUCATION\_TYPE VS CODE\_GENDER

- FEMALES WHO HAVE ASSOCIATE DEGREE TEND TO DEFAULT MORE



## AGE\_GROUP VS NAME\_CONTRACT\_TYPE

- CLIENTS IN THE AGE GROUP <20 OPTED ONLY FOR CASH LOANS AND THERE ARE NO DEFAULTERS IN THAT AGE GROUP

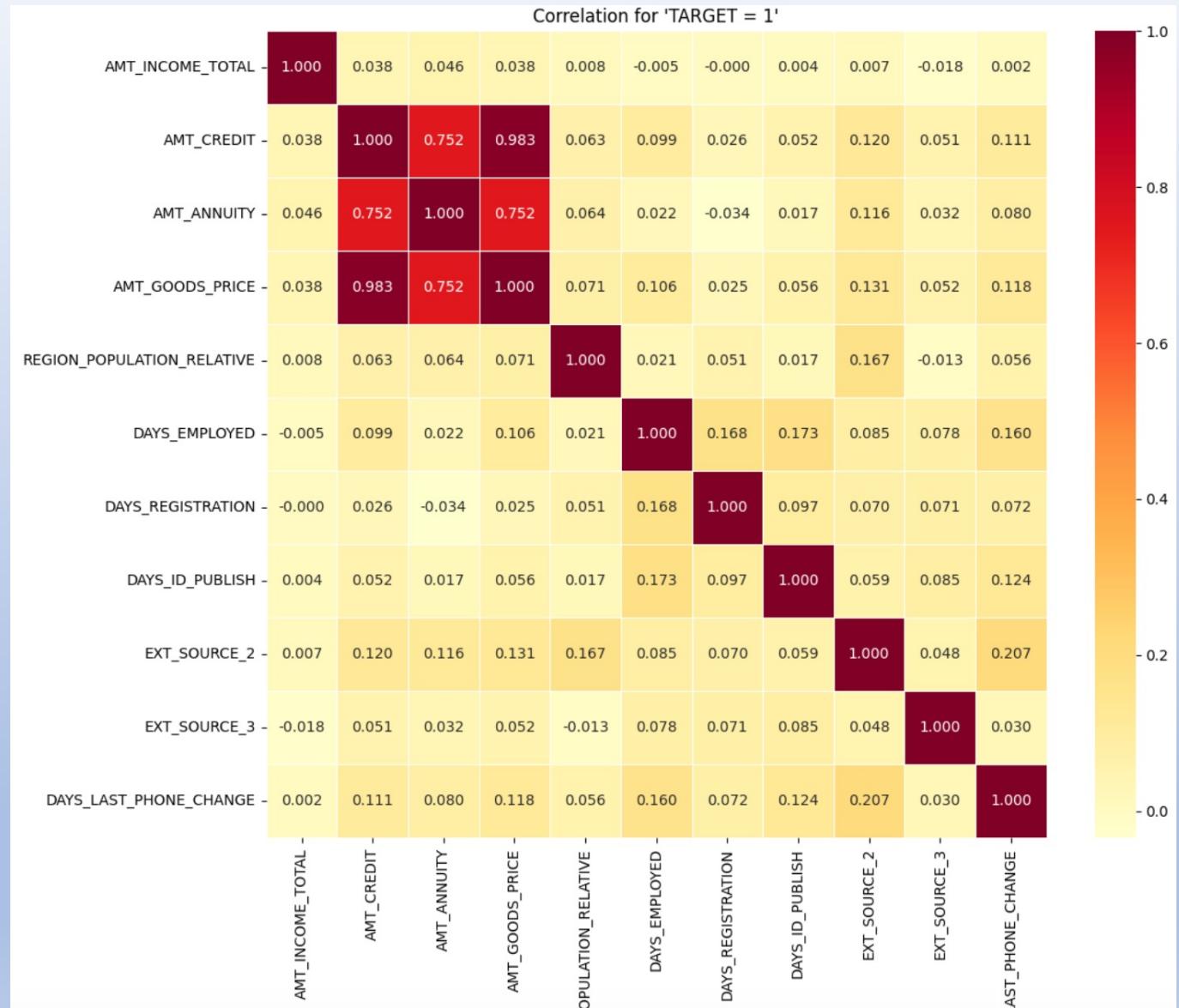


# MULTIVARIATE ANALYSIS (HEAT MAP)

# CORRELATION FOR TARGET=1

---

- THE BRIGHTER RED INDICATES HIGH POSITIVE CORRELATION BETWEEN THE VARIABLES

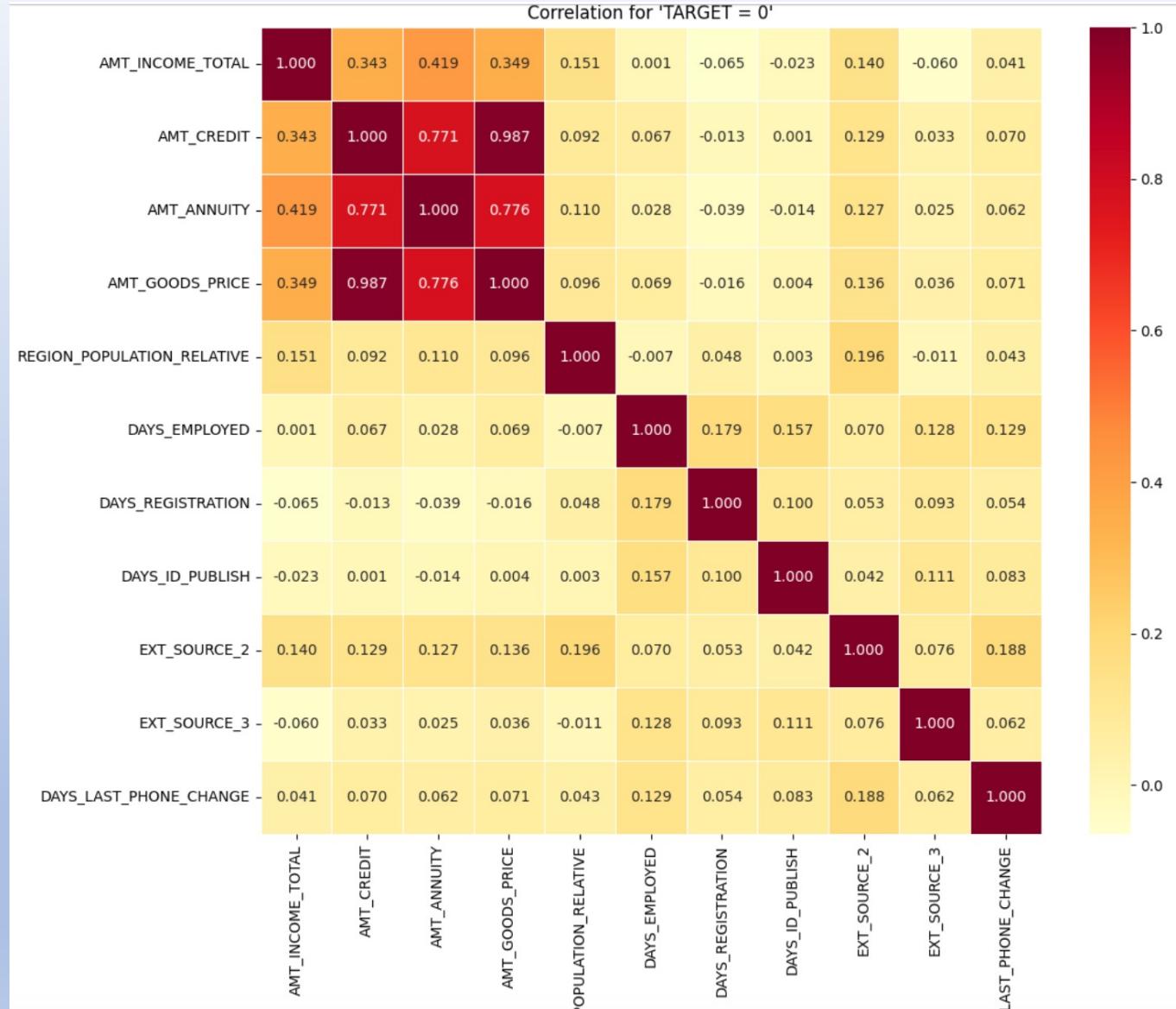


# TOP 3 CORRELATION FOR TARGET=1

- AMT\_GOODS\_PRICE AMT\_CREDIT 0.982783
- AMT\_ANNUITY AMT\_GOODS\_PRICE 0.752295
- AMT\_CREDIT AMT\_ANNUITY 0.752195

# CORRELATION FOR TARGET=0

- THE BRIGHTER RED INDICATES HIGH POSITIVE CORRELATION BETWEEN THE VARIABLES



# TOP 3 CORRELATION FOR TARGET=0

- AMT\_CREDIT AMT\_GOODS\_PRICE 0.987022
- AMT\_ANNUITY AMT\_GOODS\_PRICE 0.776400
- AMT\_CREDIT AMT\_ANNUITY 0.771276

EDA OF MERGED DATA OF  
PREVIOUS\_APPLICATION DATA SET AND  
APPLICATION\_DATA SET

# DATA CLEANING AND ANALYSIS

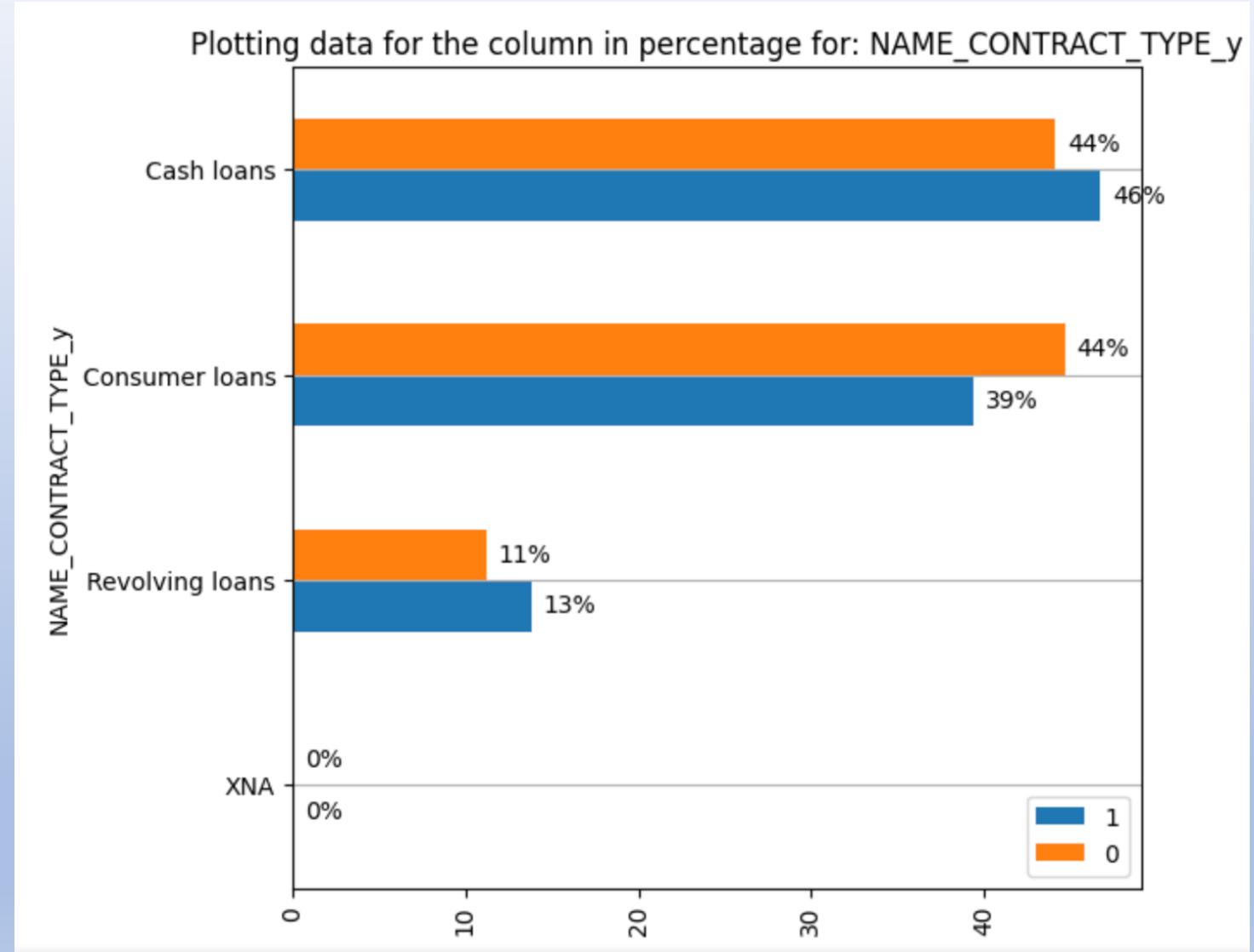
- DATA CLEANING
  - TREATMENT OF OUTLIERS
  - STANDARDISING THE DATA VALUES
  - TREATMENT OF NULL VALUES
  - CHECKING THE DATA TYPES AND STANDARDISE THE
- MERGE BOTH THE DATA
- ANALYSIS
  - UNIVARIATE ANALYSIS
  - BIVARIATE ANALYSIS
  - MULTIVARIATE ANALYSIS

# UNIVARIATE ANALYSIS (PREVIOUS\_APPLICATION)

# CATEGORICAL UNIVARIATE ANALYSIS (COUNT PLOT AND PIE CHART)

## NAME\_CONTRACT\_TYP E\_y(Contract type of previous application)

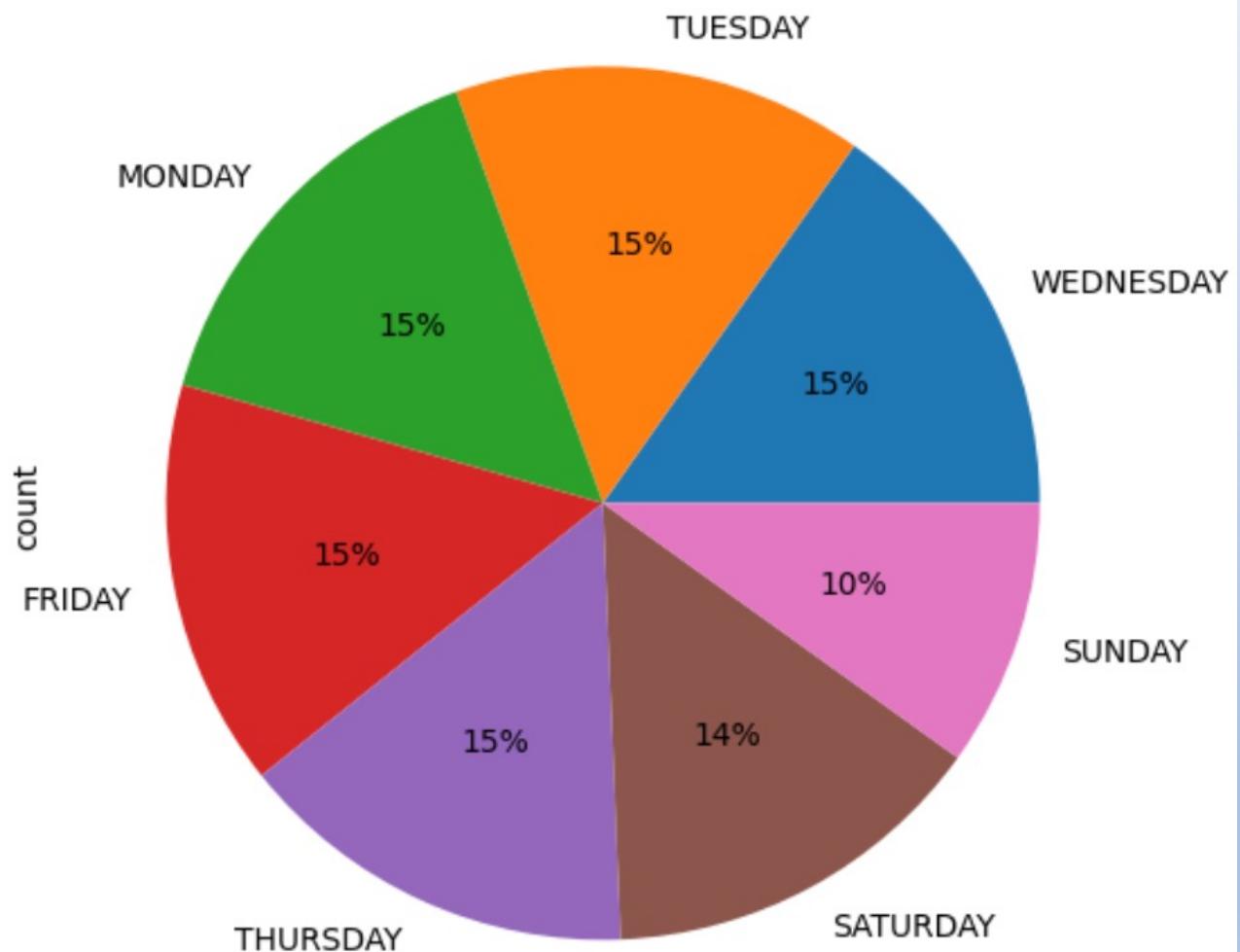
- PERCENTAGE OF REVOLVING LOAN TYPE IS LOW
- CLIENTS IN CASH LOANS TEND TO DEFAULT MORE



WEEKDAY\_APPR\_PROCESS\_S  
TART(ON WHICH DAY OF  
THE WEEK APPLICATION  
PROCESS STARTED)

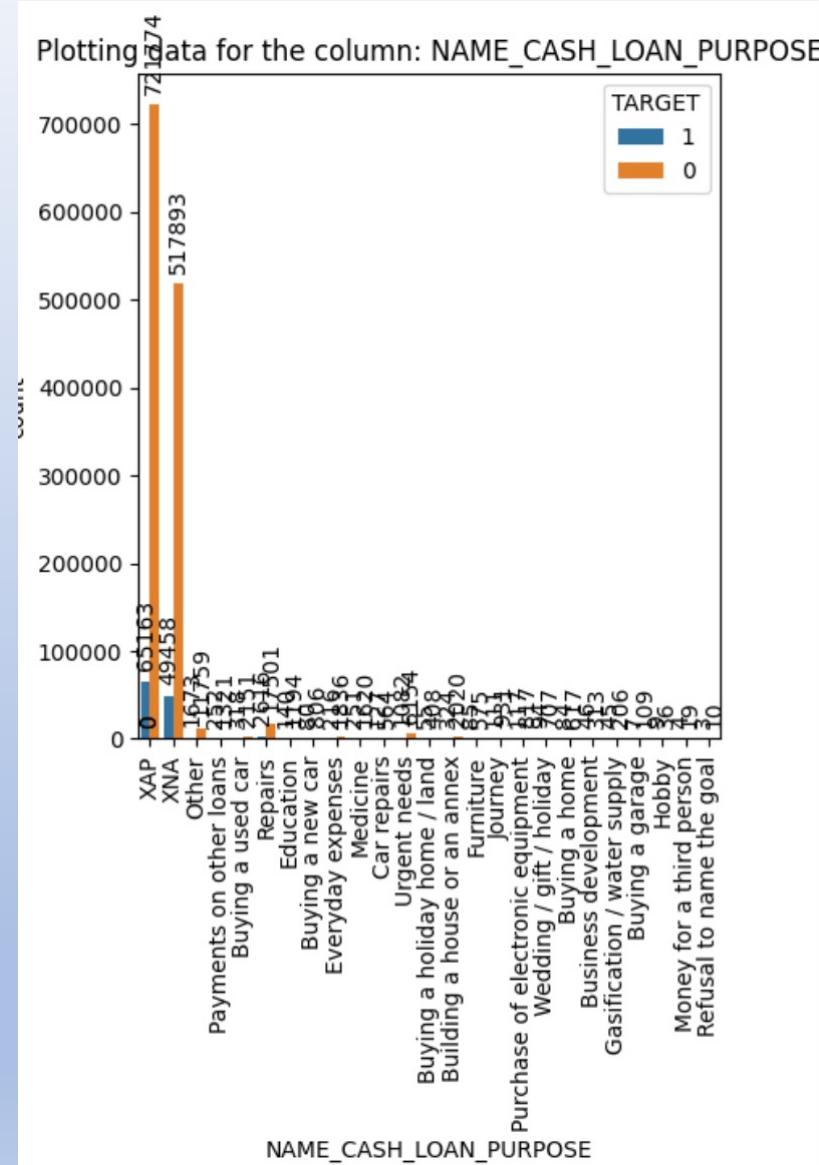
- PERCENTAGE OF  
APPLICATION ON ALL  
WEEKDAYS ARE ALMOST  
SAME
- PERCENTAGE OF  
APPLICATION ON SUNDAY  
LOOKS LOWER AND  
THAT'S EXPECTED SINCE  
SUNDAY IS A HOLIDAY

Plotting data for the column: WEEKDAY\_APPR\_PROCESS\_START



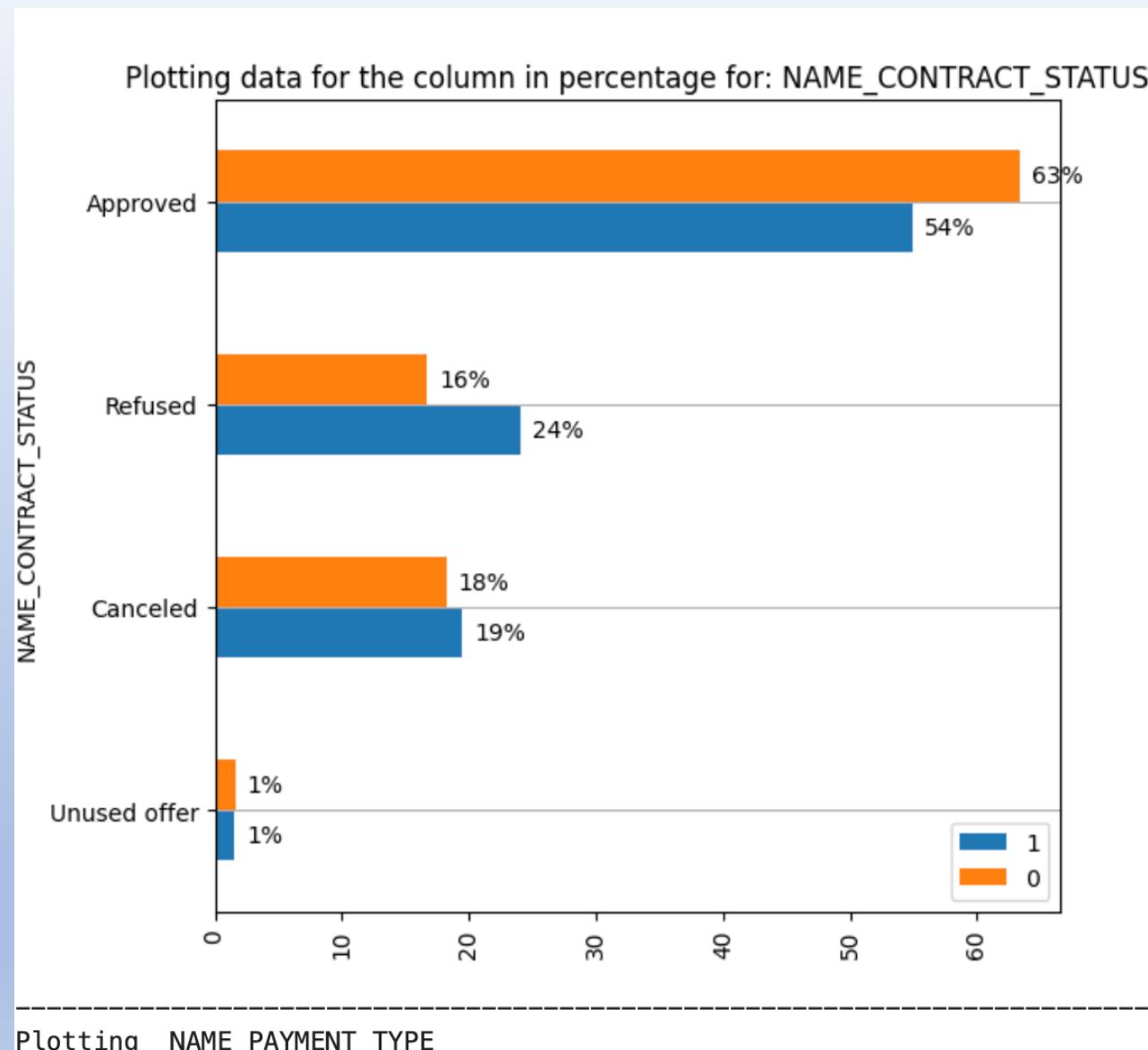
## NAME\_CASH\_LOAN\_PURPOSE

- THE PURPOSE OF CASH LOAN IN THE CATEGORY XAP AND XNA ARE HIGHER



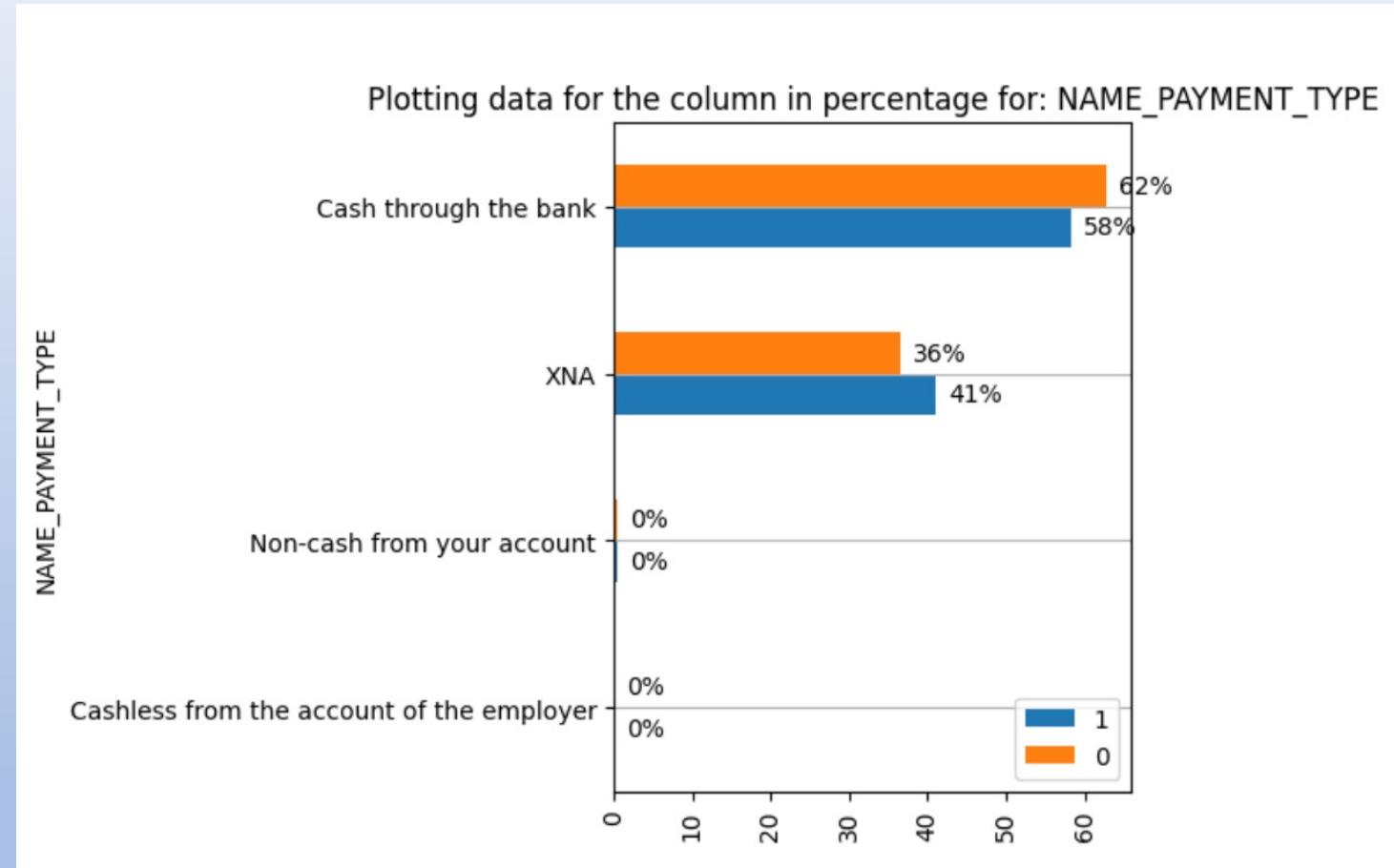
NAME\_CONTRACT\_STATUS  
US(APPROVED,CANCELLED,REFUSED OR  
UNUSED)

- MORE THAN 50% OF THE APPLICATIONS WERE APPROVED



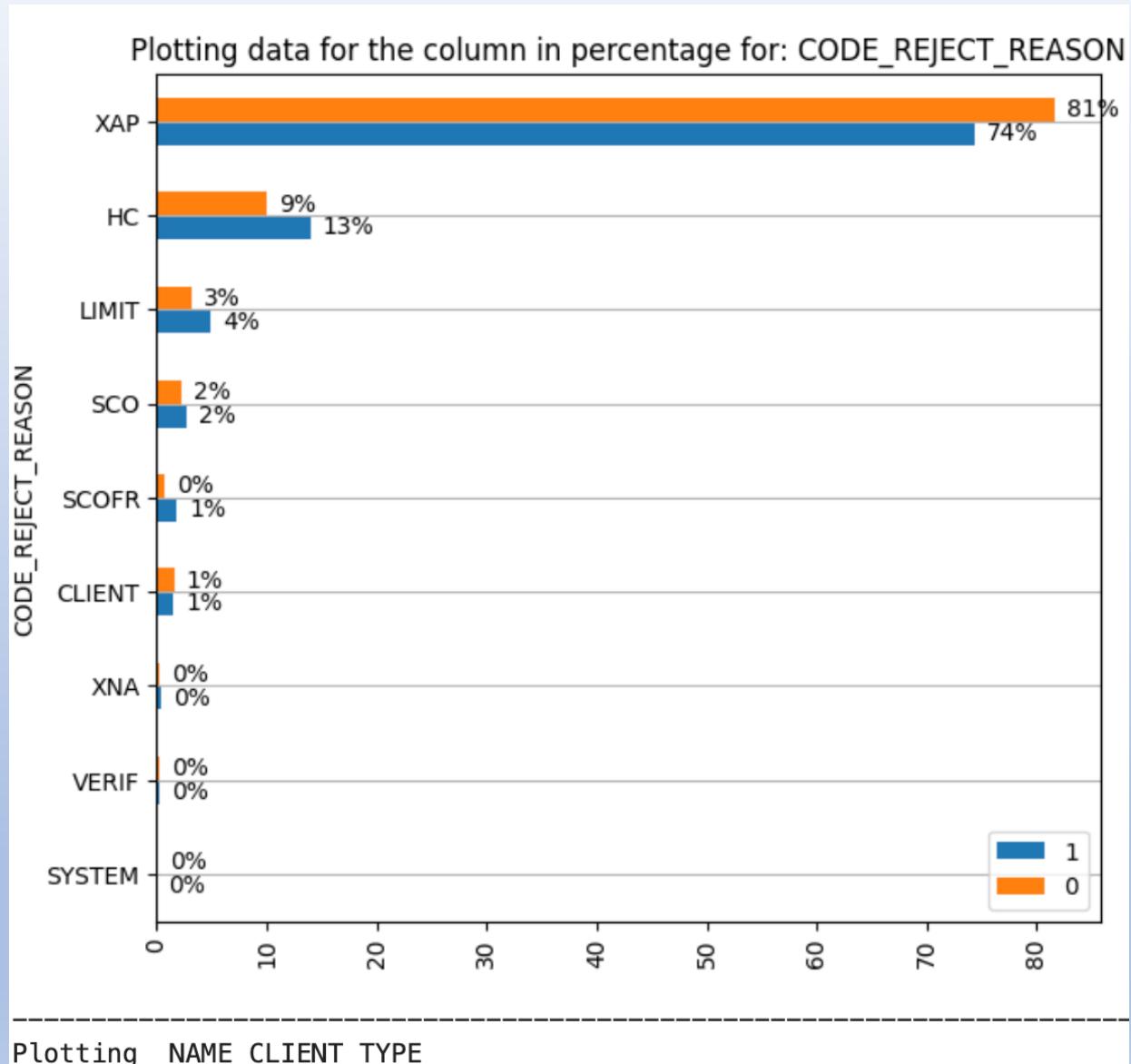
## NAME\_PAYMENT\_TYPE (REPAYMENT METHOD OF CLIENT)

- MOST OF THE CLIENTS CHOSE TO PAY CASH THROUGH THE BANK
- CLIENTS WHO CHOSE TO PAY THROUGH XNA MORE TEND TO DEFAULT



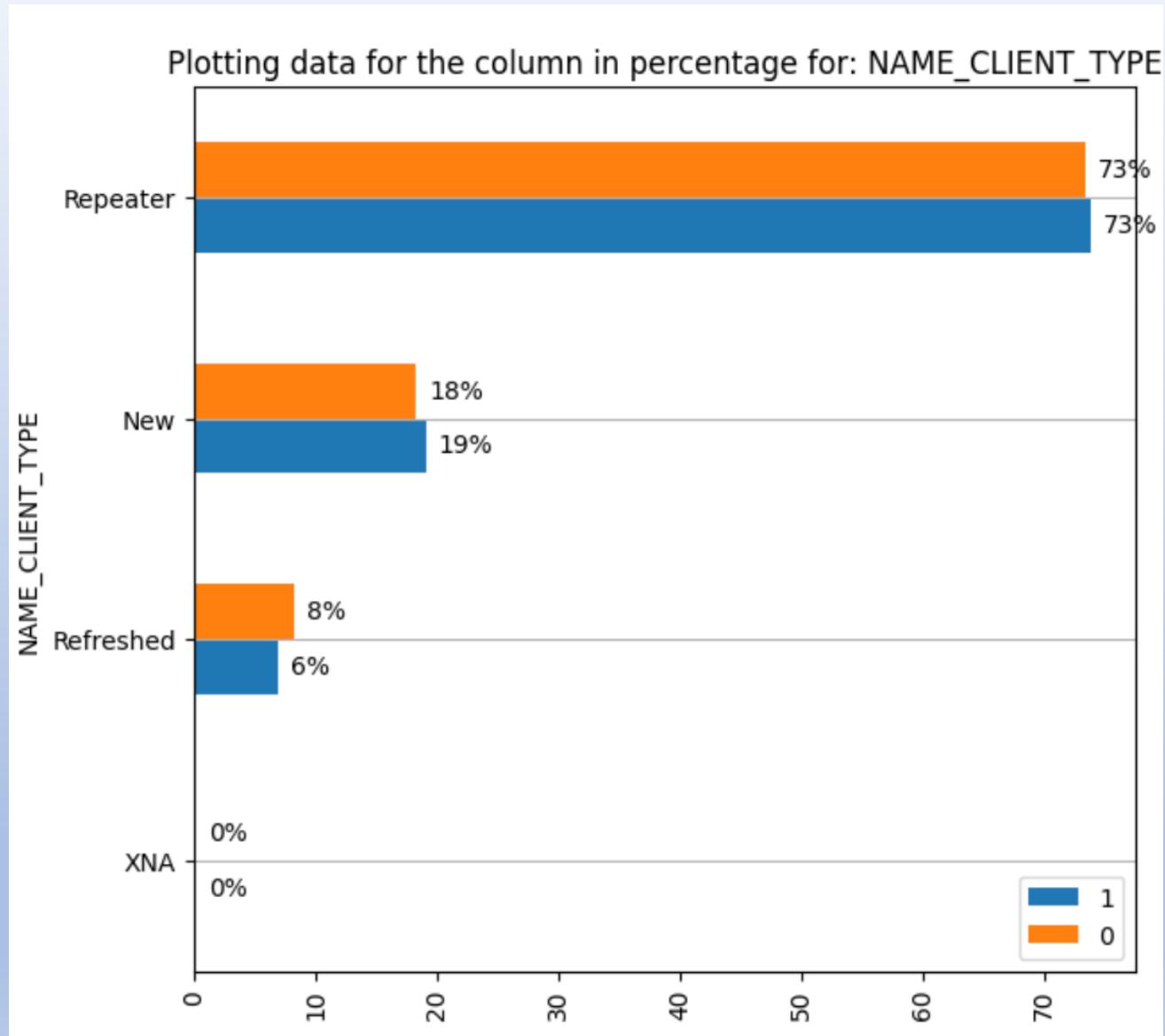
CODE\_REJECT\_REASON(WHY THE LOAN WAS REJECTED)

- MOST OF THE CLIENTS HAD REASON XAP FOR REJECTION
- CLIENTS WHO WERE REJECTED WITH REASON HC TEND TO DEFAULT MORE



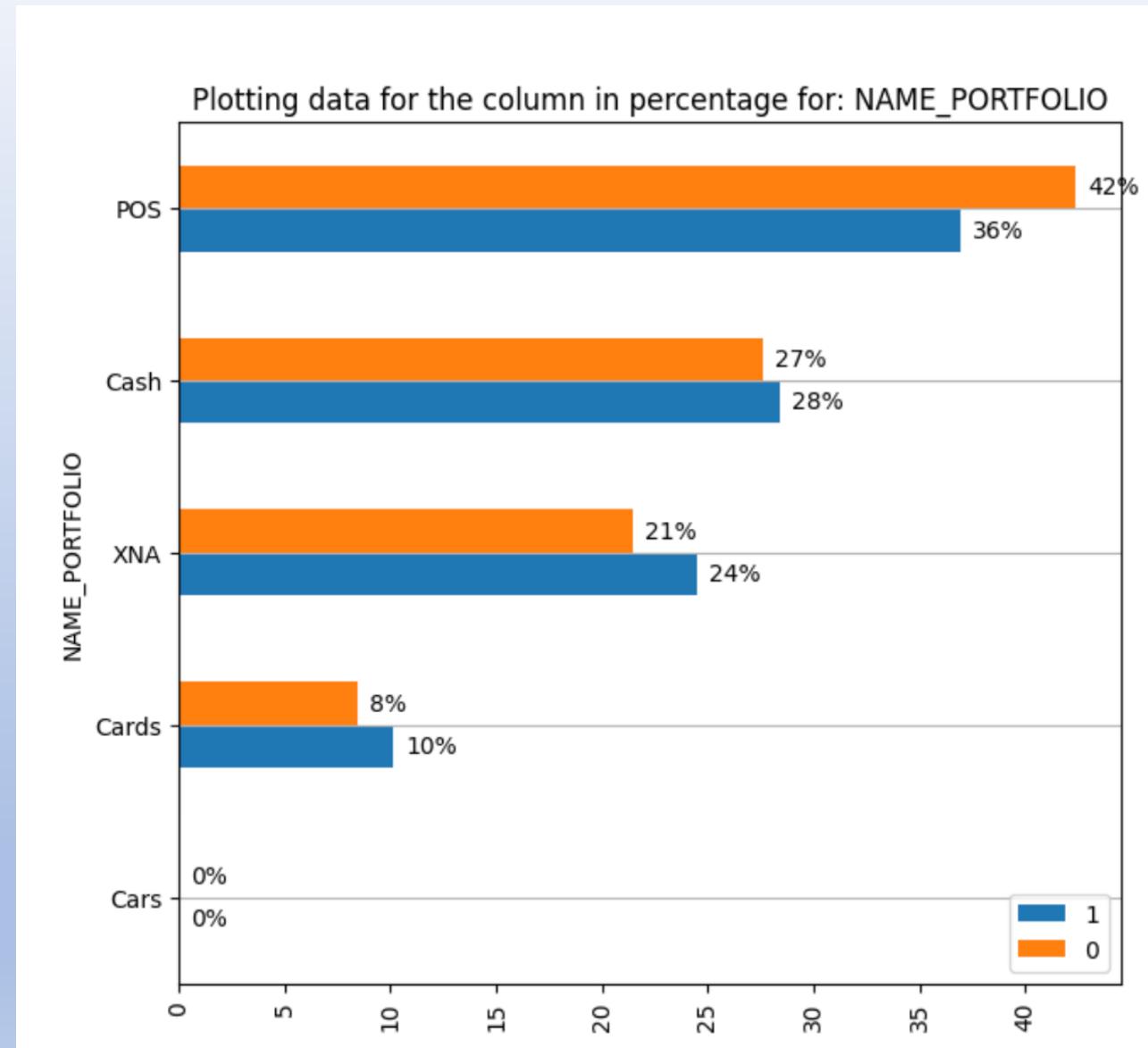
NAME\_CLIENT\_TYPE  
(REPEATER, NEW, REFRESHER)

- REPEATER TYPE CLIENTS ARE MORE



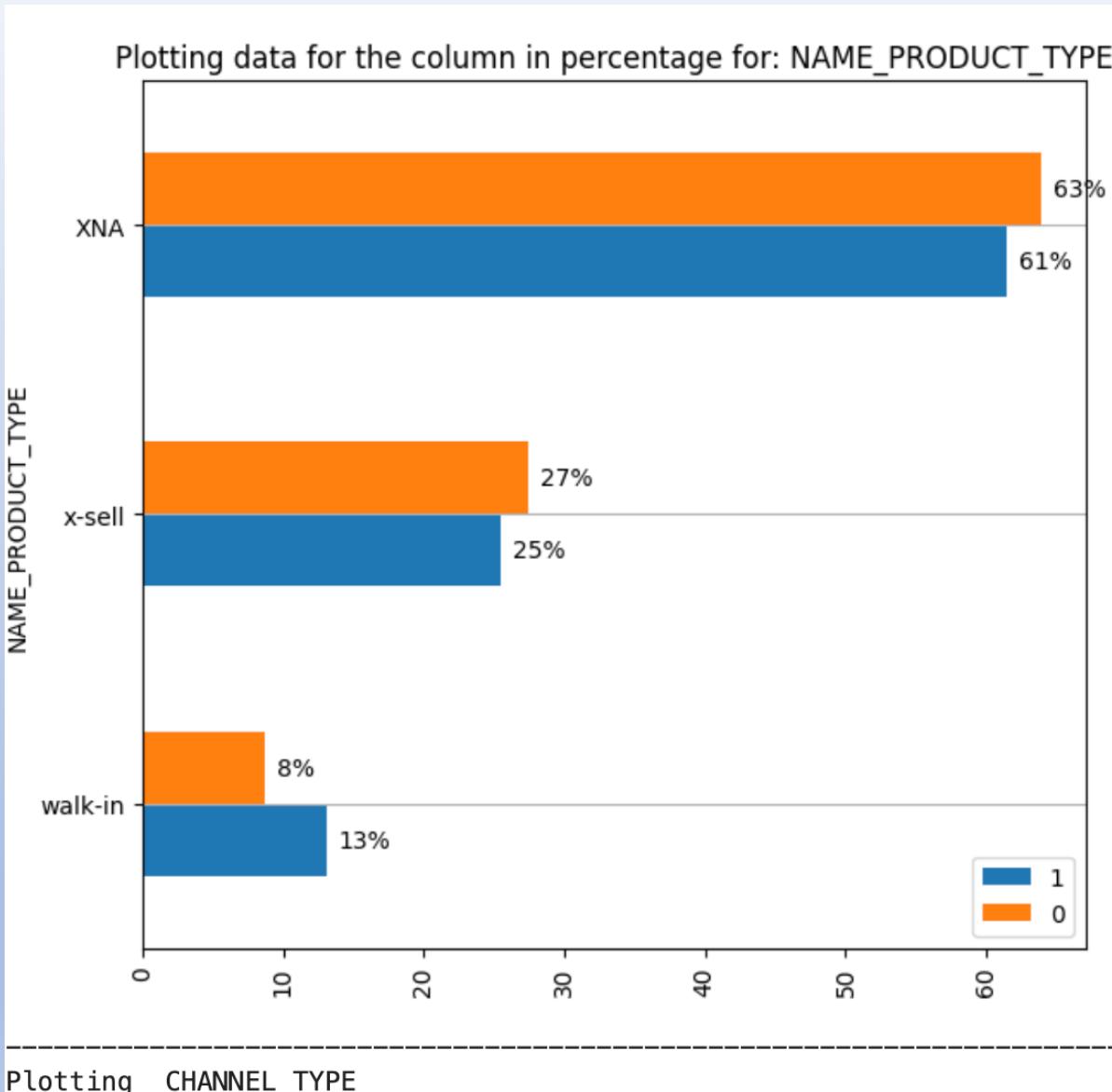
# NAME\_PORTFO LIO(CASH,POS,C AR,...)

- LOAN TYPE FOR POS(POINT OF SALE) CATEGORY IS HIGHER
- CLIENTS WHO HAVE NAME\_PORTFOLIO AS CARDS TEND TO DEFAULT MORE



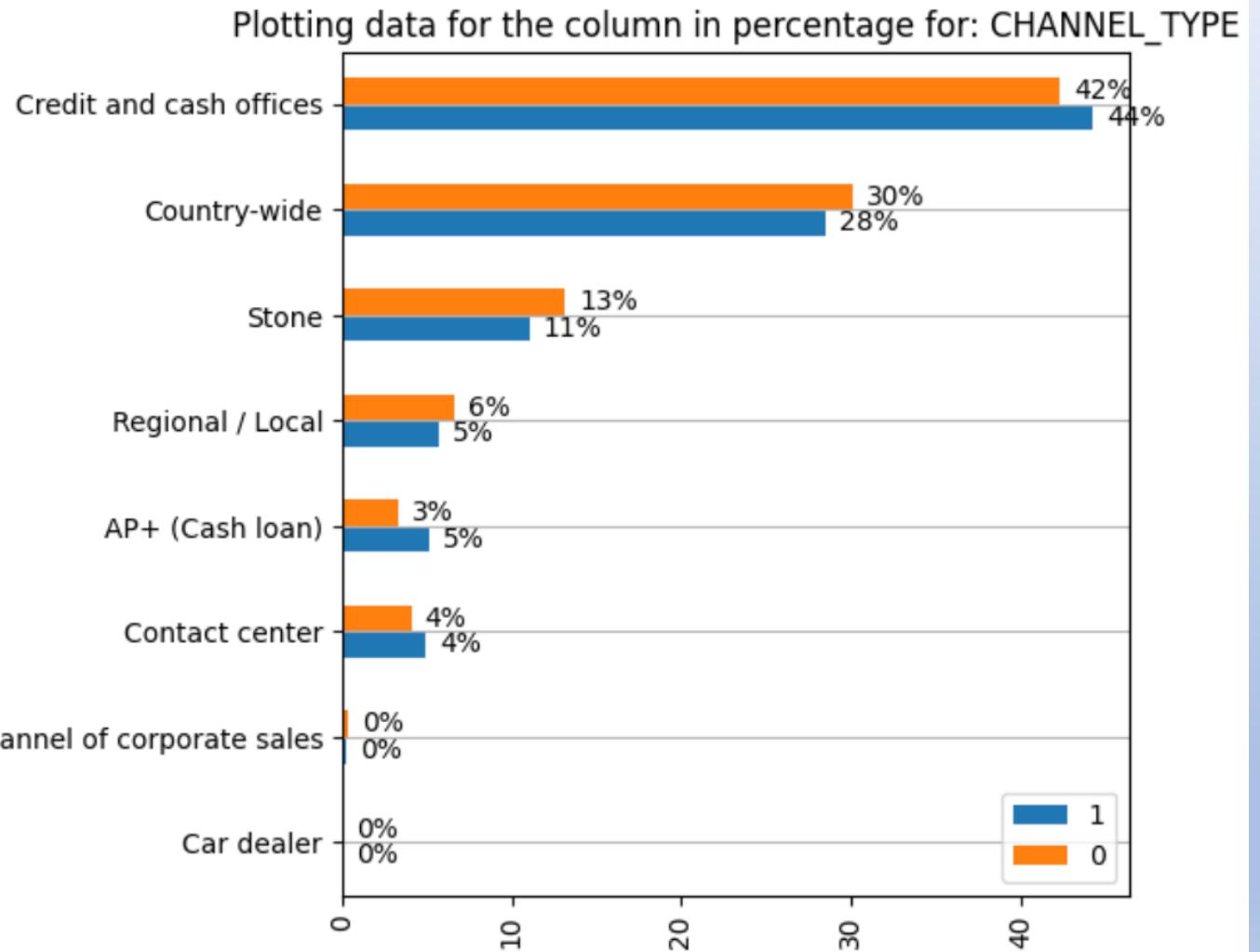
## NAME\_PRODUCT\_TYPE(X-SELL OR WALKIN)

- X-SELL (PRODUCT OR SERVICE OFFERED TO A CUSTOMER AS A PART OF CROSS SELLING STRATEGY) CATEGORY IS MORE COMPARED TO WALKIN
- WALKIN CATEGORY CLIENTS TEND TO DEFAULT MORE



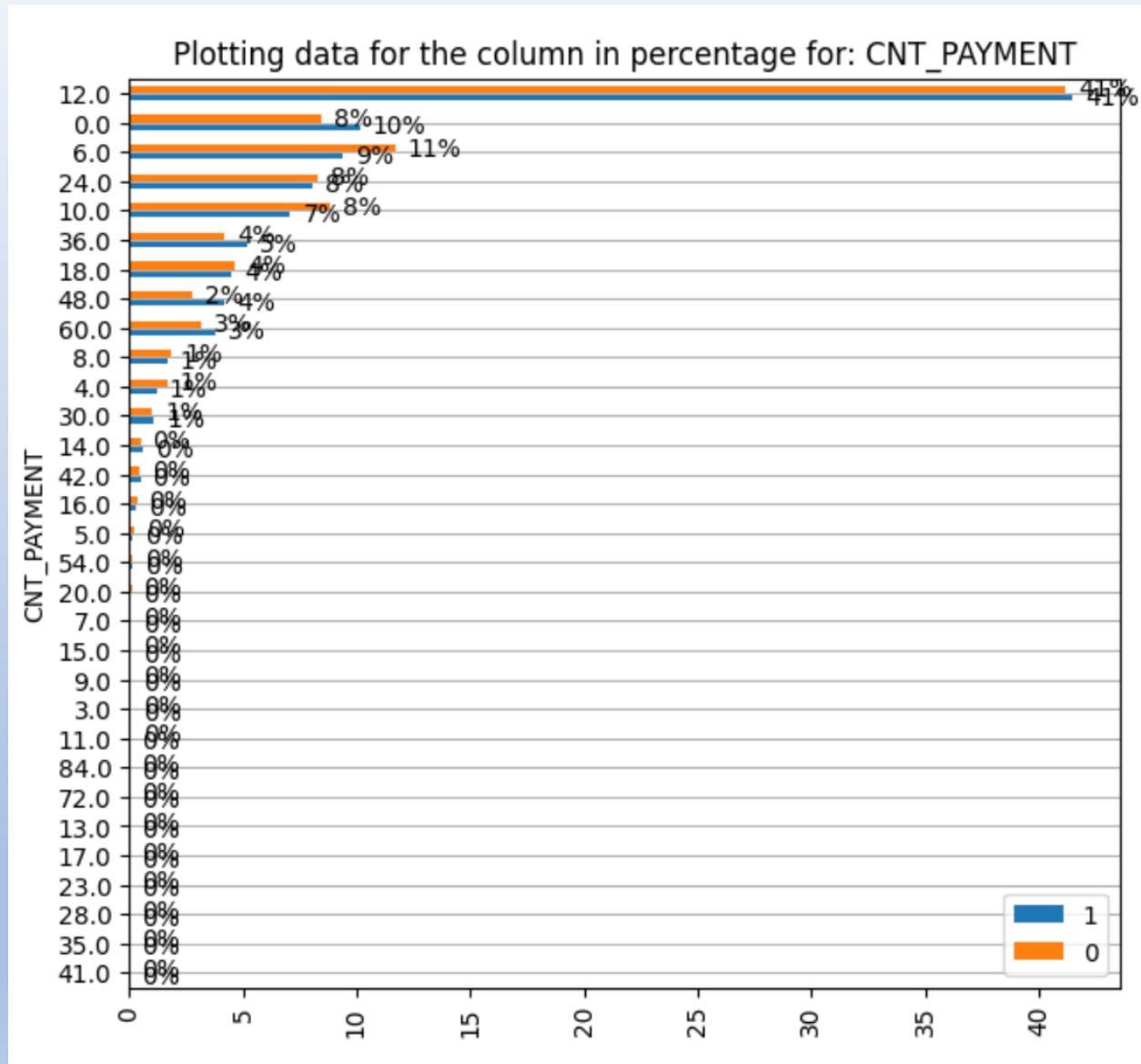
## CHANNEL\_TYPE(THROUGH WHICH CHANNEL CLIENT IS ACQUIRED)

- CLIENTS ACQUIRED THROUGH CREDIT AND CASH OFFICES, AP+(AFFINITY PARTNERS) TEND TO DEFAULT MORE



## CNT\_PAYMENT(TERM OF THE LOAN)

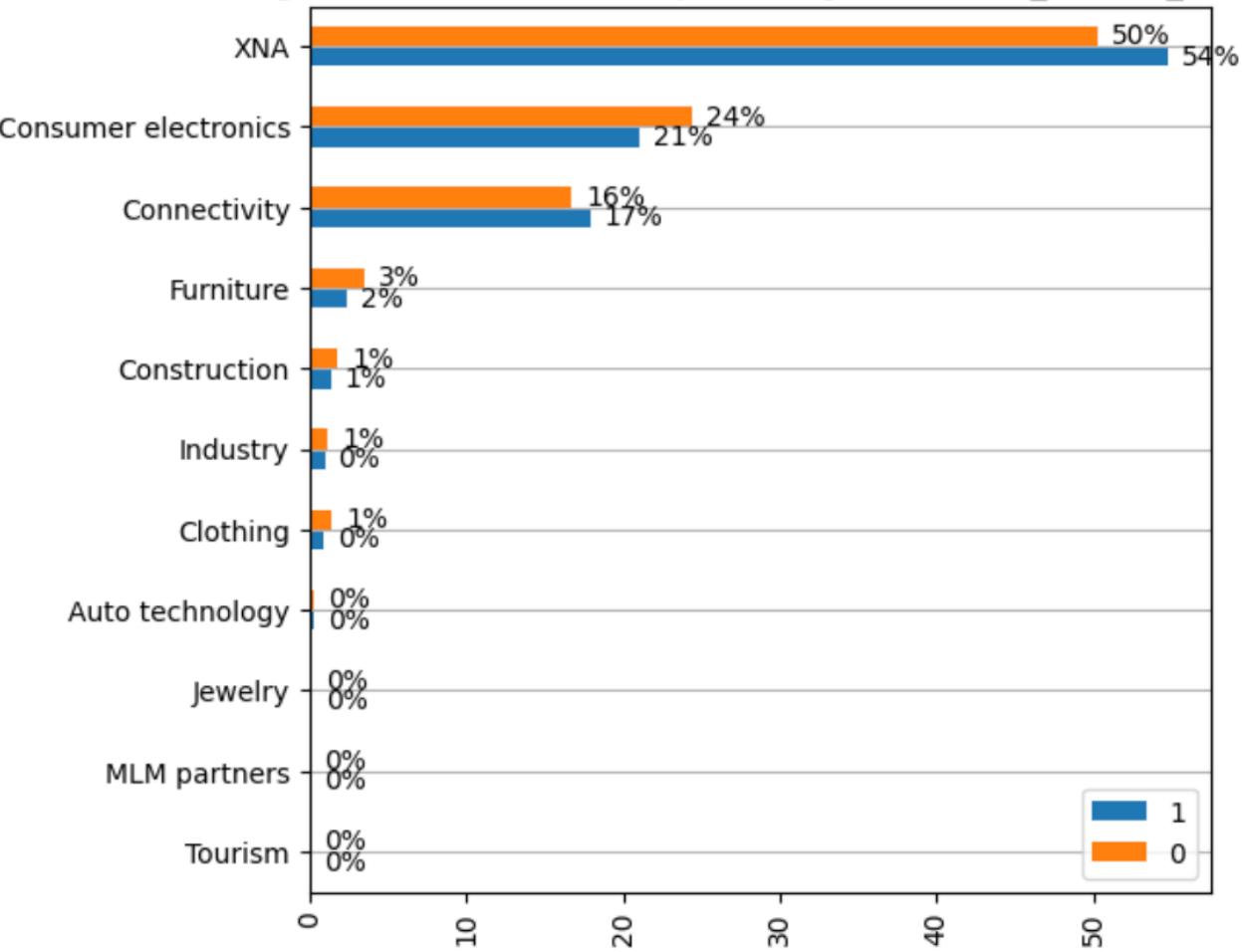
- CLENTS WITH SINGLE PAYMENT(CNT\_PAYMENT =0) TEND TO DEFFAULT MORE
- CATEGORY OF CLIENTS WITH 12 INSTALLMENTS ARE HIGH



# NAME\_SELLER\_IN DUSTRY(INDUSTRY OF THE SELLER)

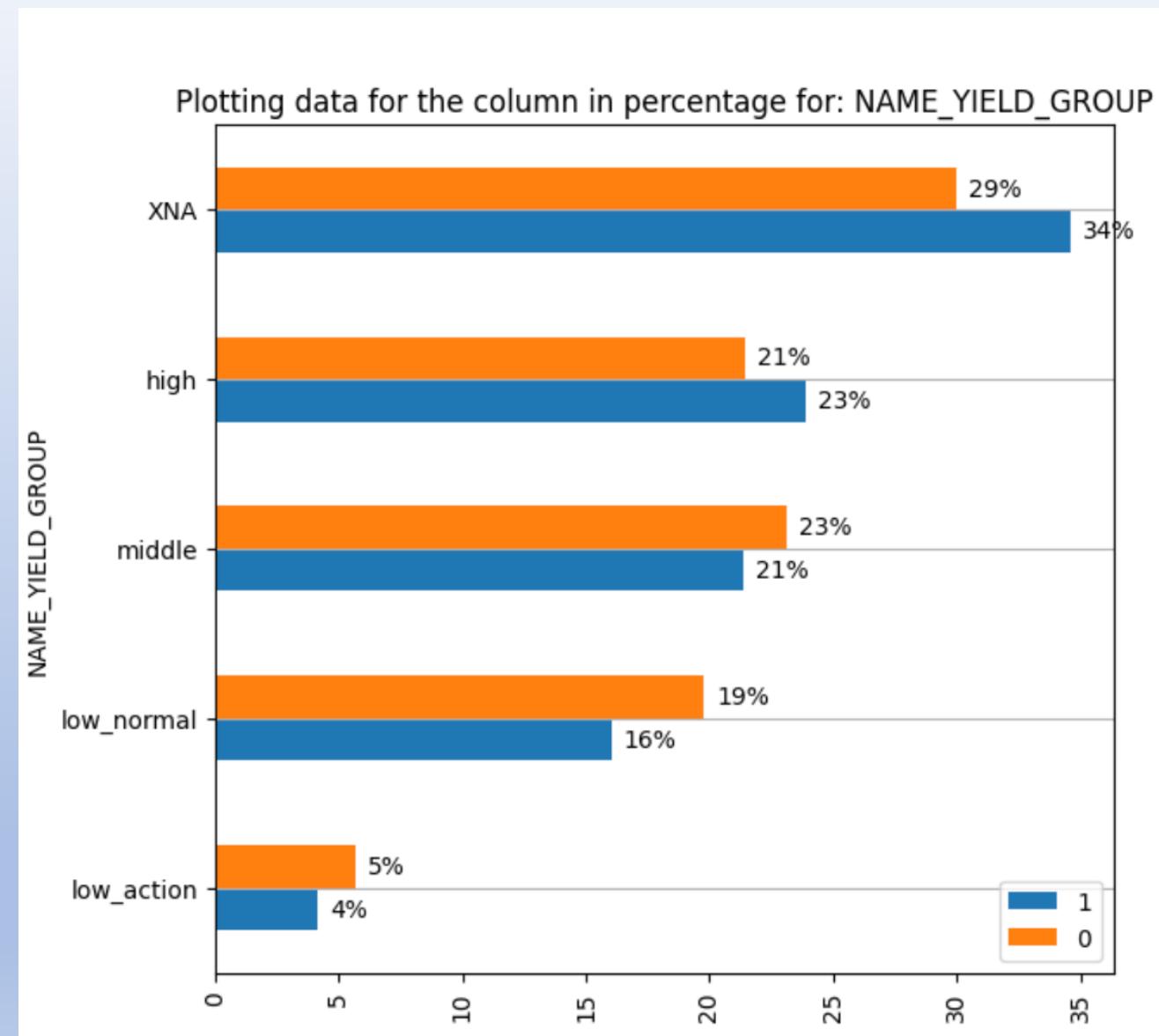
- CONSUMER ELECTRONICS CATEGORY IS MORE
- CONNECTIVITY CATEGORY TEND TO DEFAULT MORE

Plotting data for the column in percentage for: NAME\_SELLER\_INDUSTRY



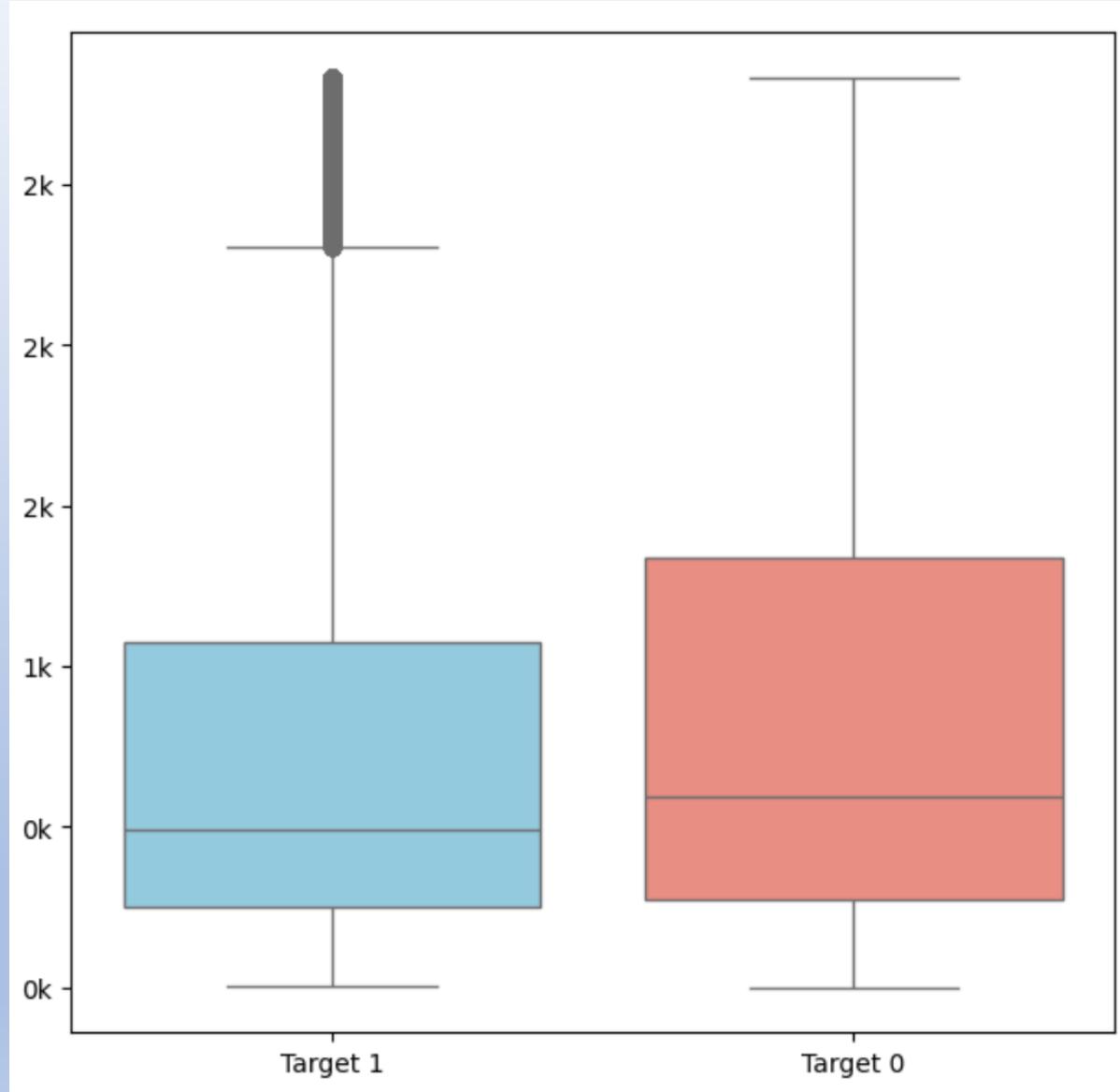
NAME\_YIELD\_GROUP(CATEGORY  
OF INTEREST RATE –  
HIGH,MIDDLE,LOW)

- CLIENTS WITH HIGH  
INTEREST RATE TEND TO  
DEFAULT MORE
- CLIENTS IN THE  
LOW\_NORMAL CATEGORY  
HAVE LESS DIFFICULTY IN  
REPAYING THE LOAN



## DAYS\_DECISION(No OF DAYS TO TAKE DECISION ON THE APPLICATION)

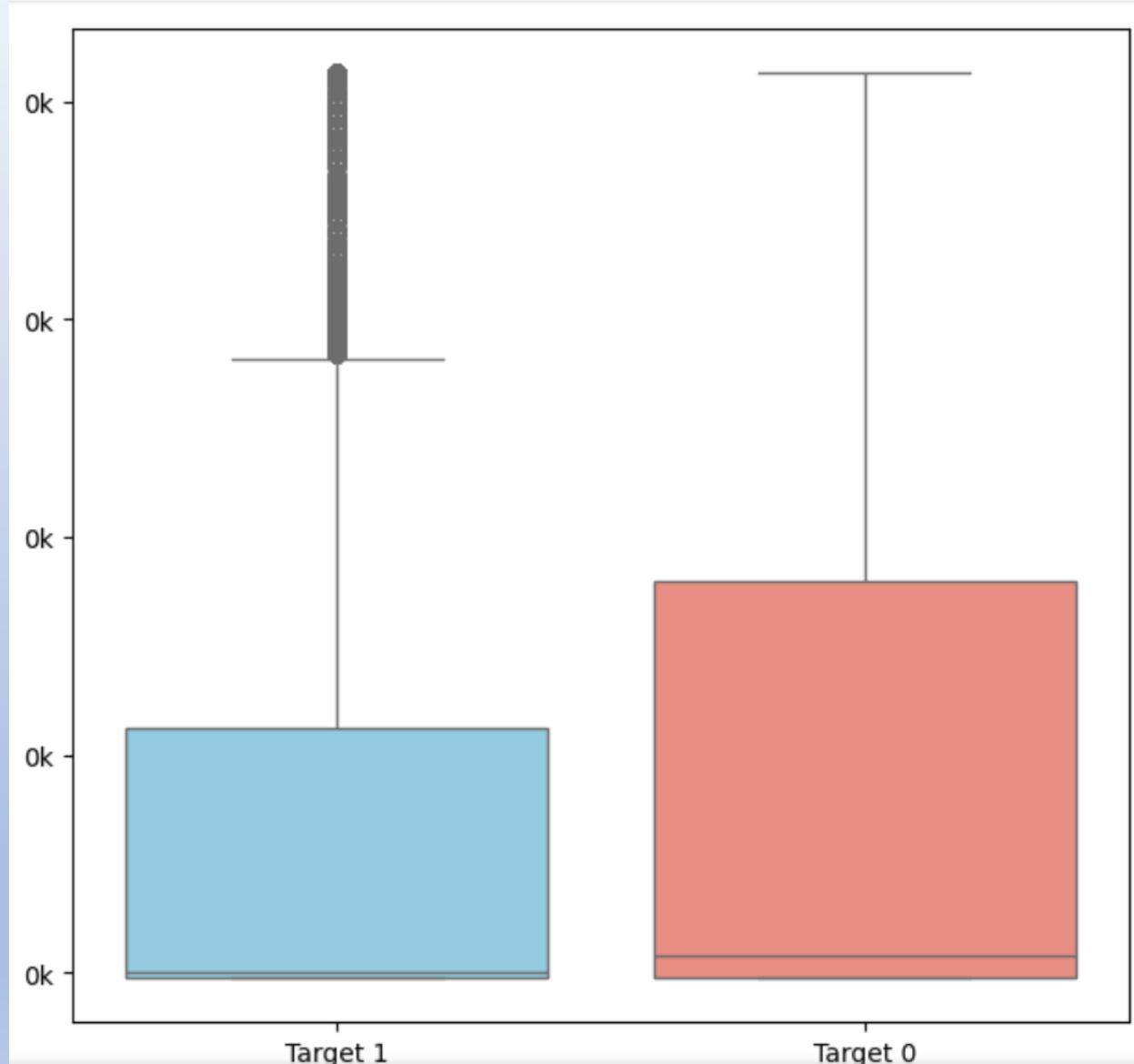
- NUMBER OF DAYS TAKEN FOR NON-DEFAULT CATEGORY IS HIGHER COMPARED TO DEFAULT CATEGORY



# NUMERICAL UNIVARIATE ANALYSIS(HISTOGRAM AND DISTPLOT AND BOX PLOT)

## SELLERPLACE\_AREA (SELLING AREA OF A SELLER)

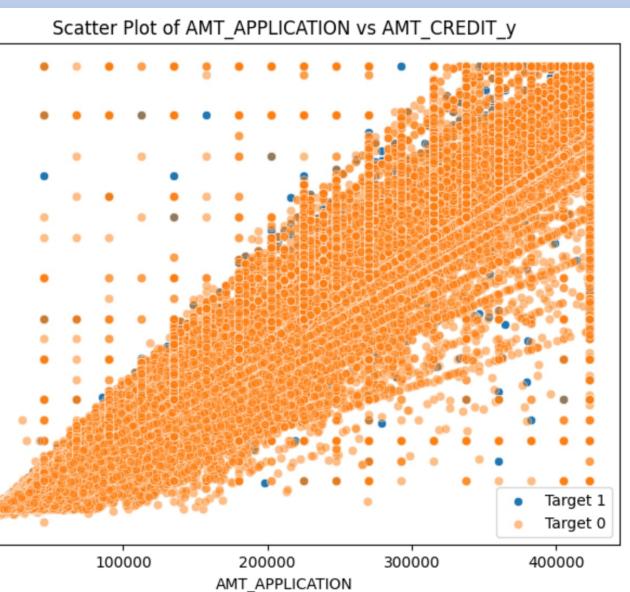
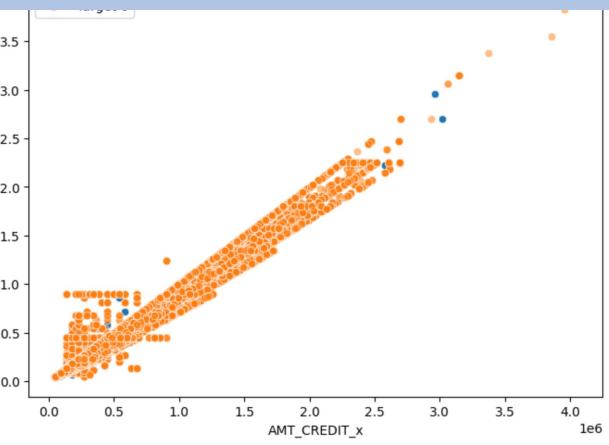
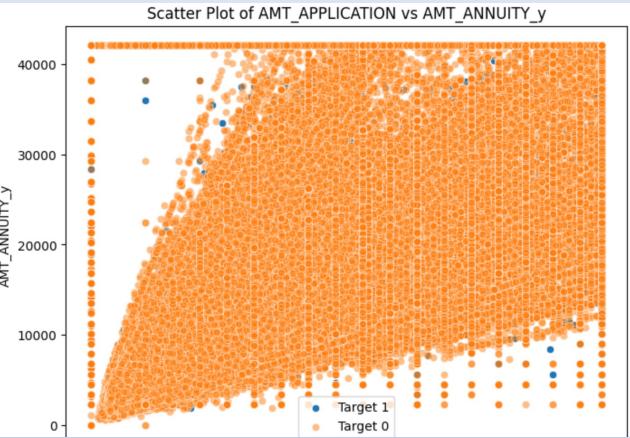
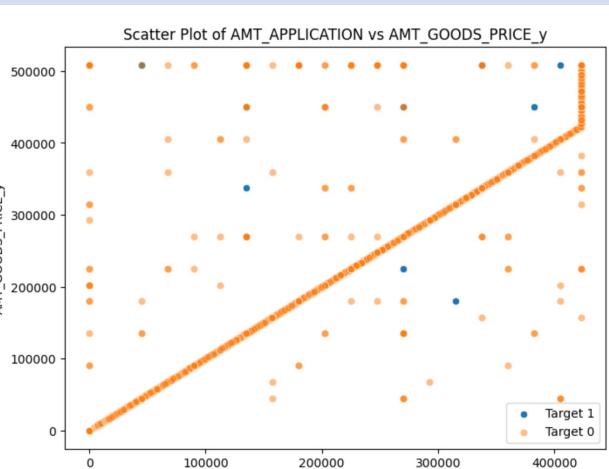
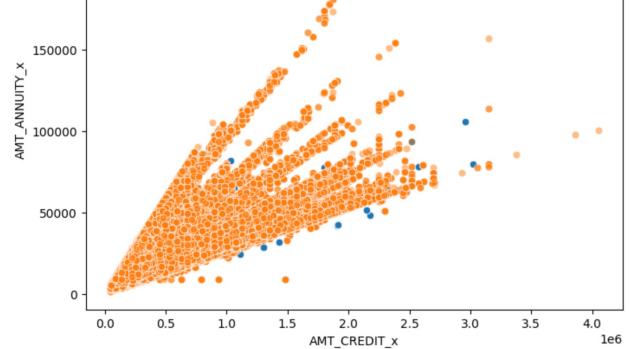
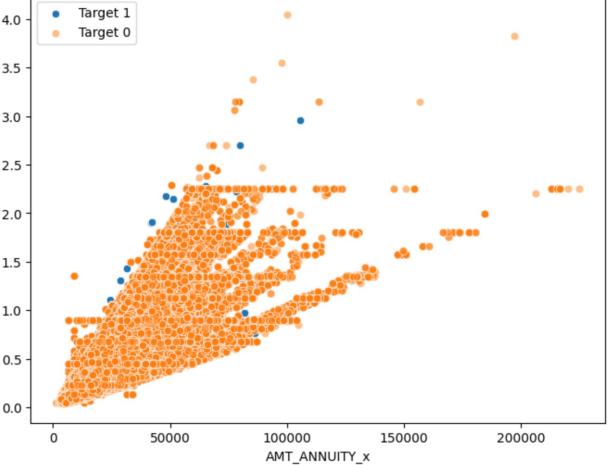
- CLIENTS IN THE HIGH SELLERPLACE\_AREA DOESN'T HAVE MUCH DIFFICULTY IN REPAYING THE LOAN



# BIVARIATE ANALYSIS

# AMT\_APPLICATION VS AMT\_CREDIT\_X VS AMT\_GOODS\_PRICE\_X VS AMT\_ANNUITY\_X

- ALL THESE FOUR VARIABLES HAVE POSITIVE CORRELATION(IF THE VALUE OF ONE INCREASES THE OTHER INCREASES)

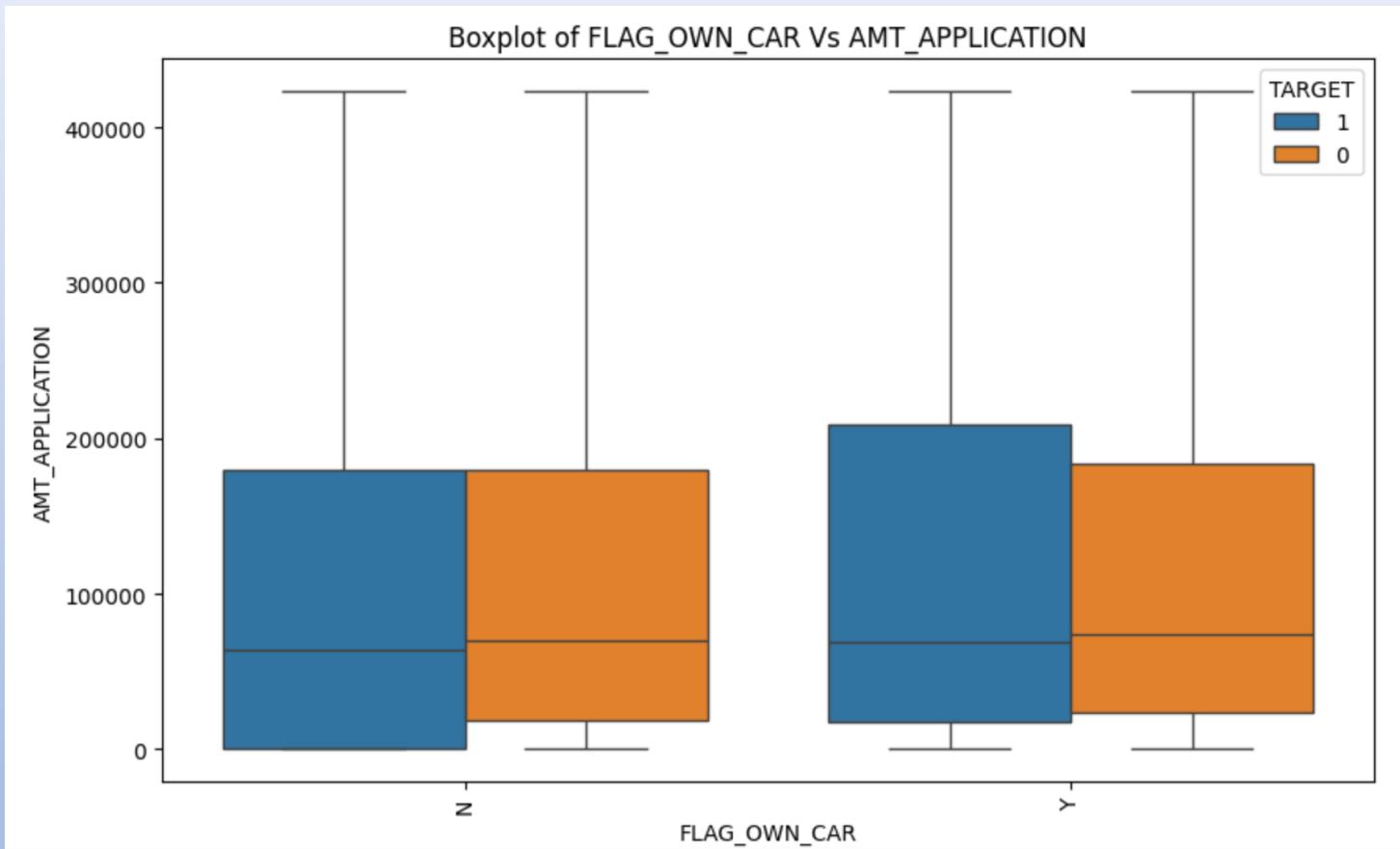


# NUMERICAL VS NUMERICAL (SCATTER PLOT)

# NUMERICAL VS CATEGORICAL (BOX PLOT)

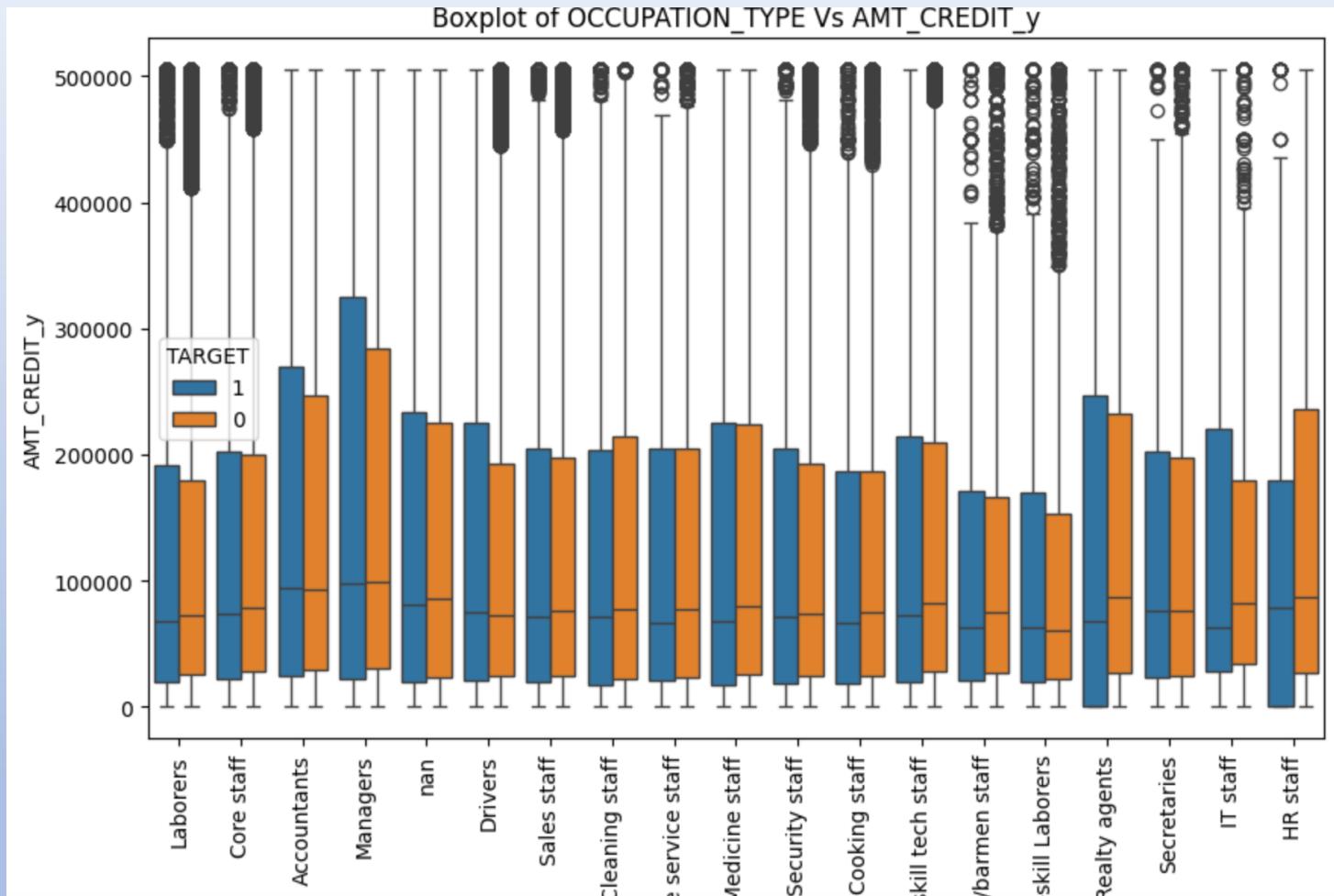
# AMT\_APPLICATION VS FLAG\_own\_CAR

- THE CLIENTS OWNING THE CAR HAVE HIGHER AMT\_APPLICATION



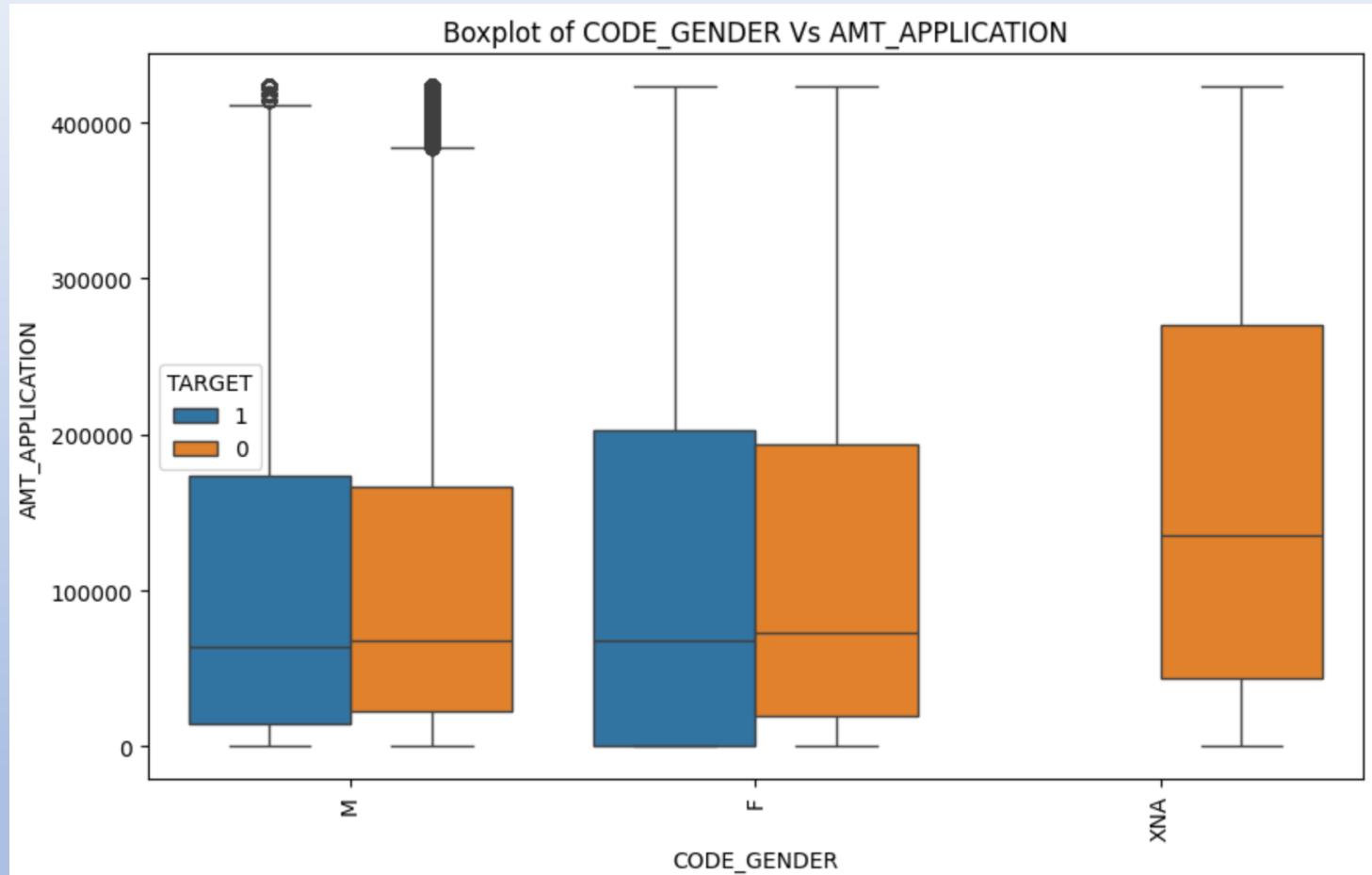
# OCCUPATION\_TYPE VS AMT\_CREDIT\_y

- THE AMT\_CREDIT\_Y FOR MANAGERS VARY MORE IN THE UPPER QUARTILE REGION



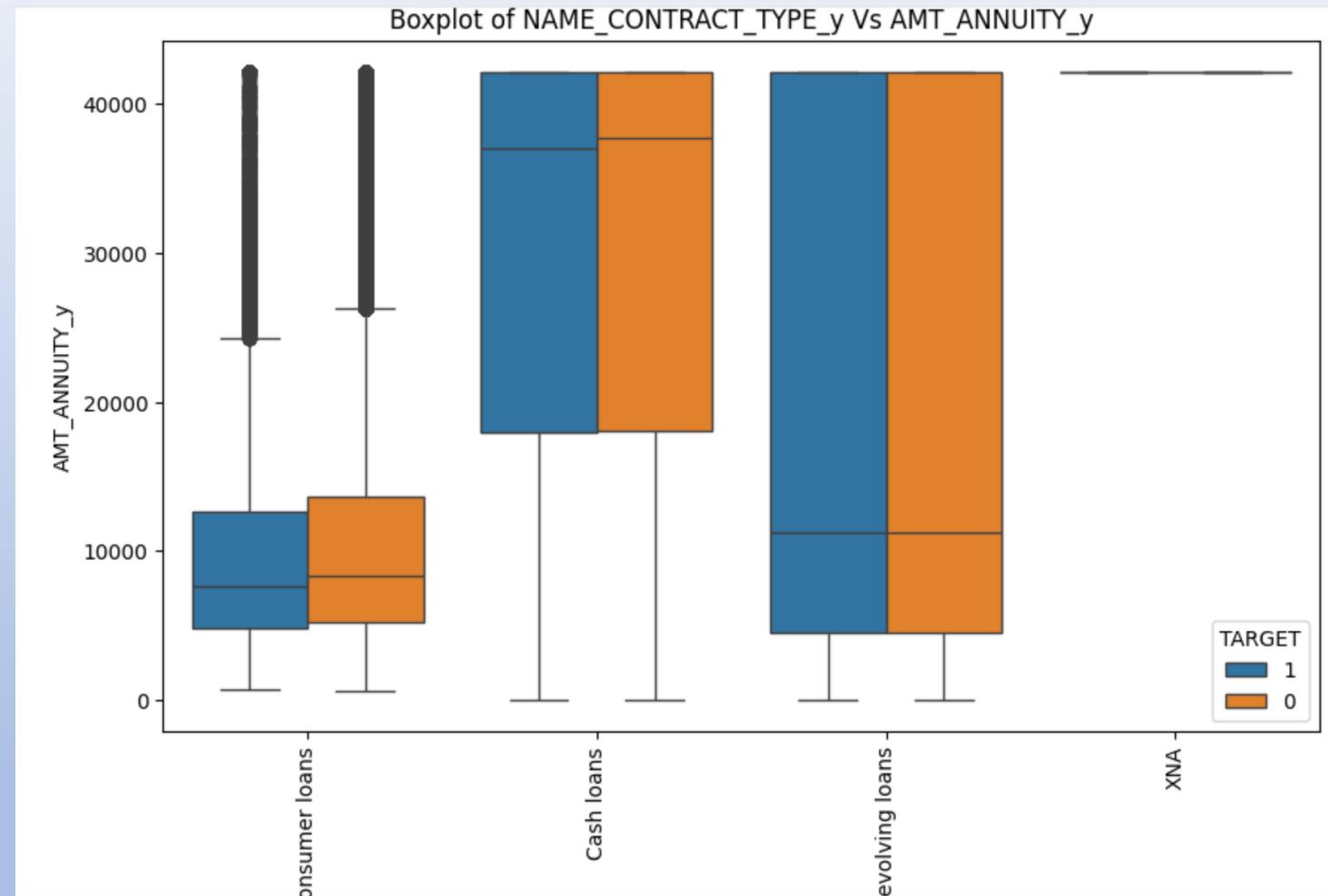
# CODE\_GENDER VS AMT\_APPLICATION

- AMT\_APPLICATION IS HIGHER IN THE FEMALE CATEGORY AND FEMALE CATEGORY TEND TO DEFAULT MORE



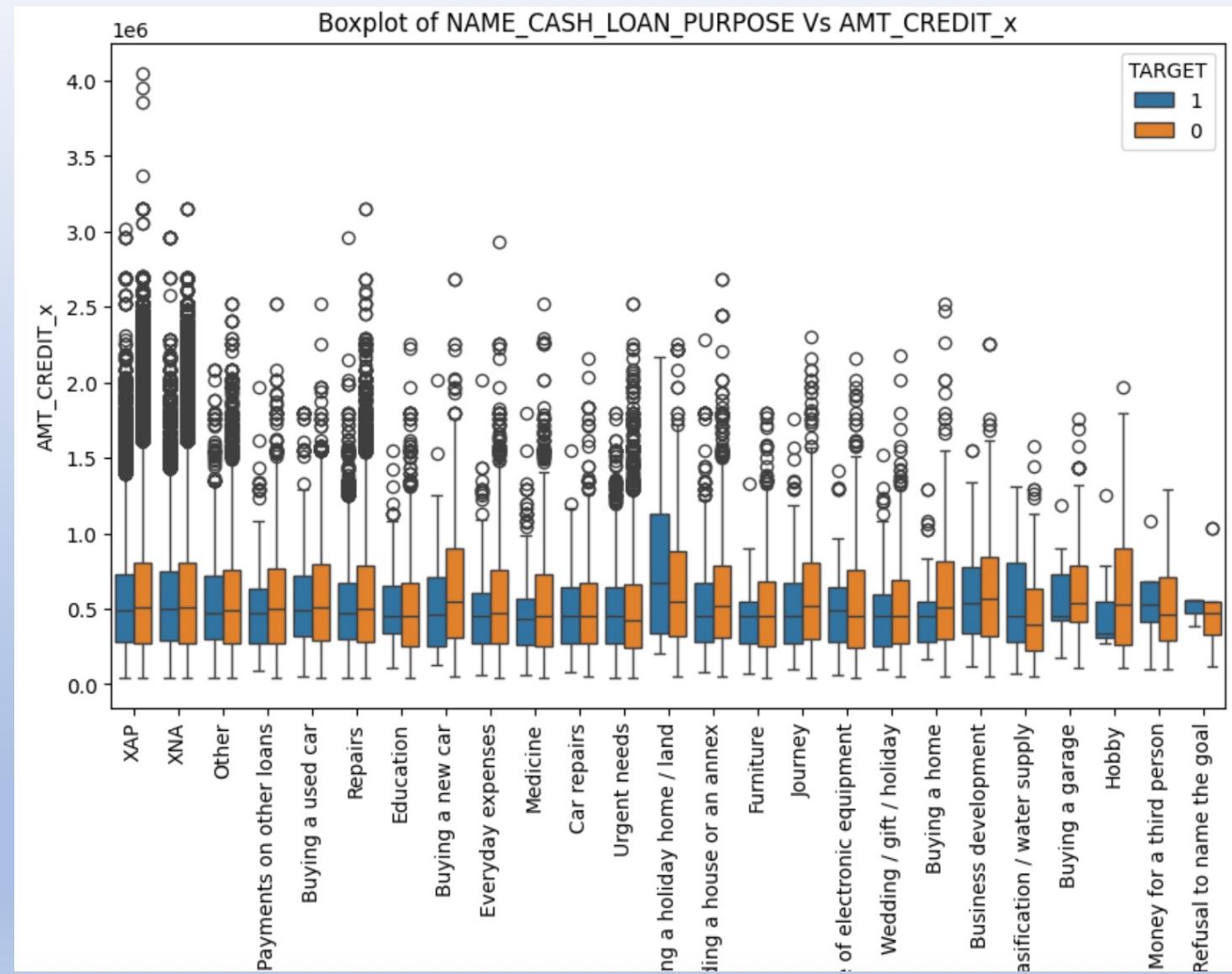
## NAME\_CONTRACT\_TYPE VS AMT\_ANNUITY\_y

- THE MEDIAN  
AMT\_ANNUITY IS HIGH  
FOR CASH LOANS



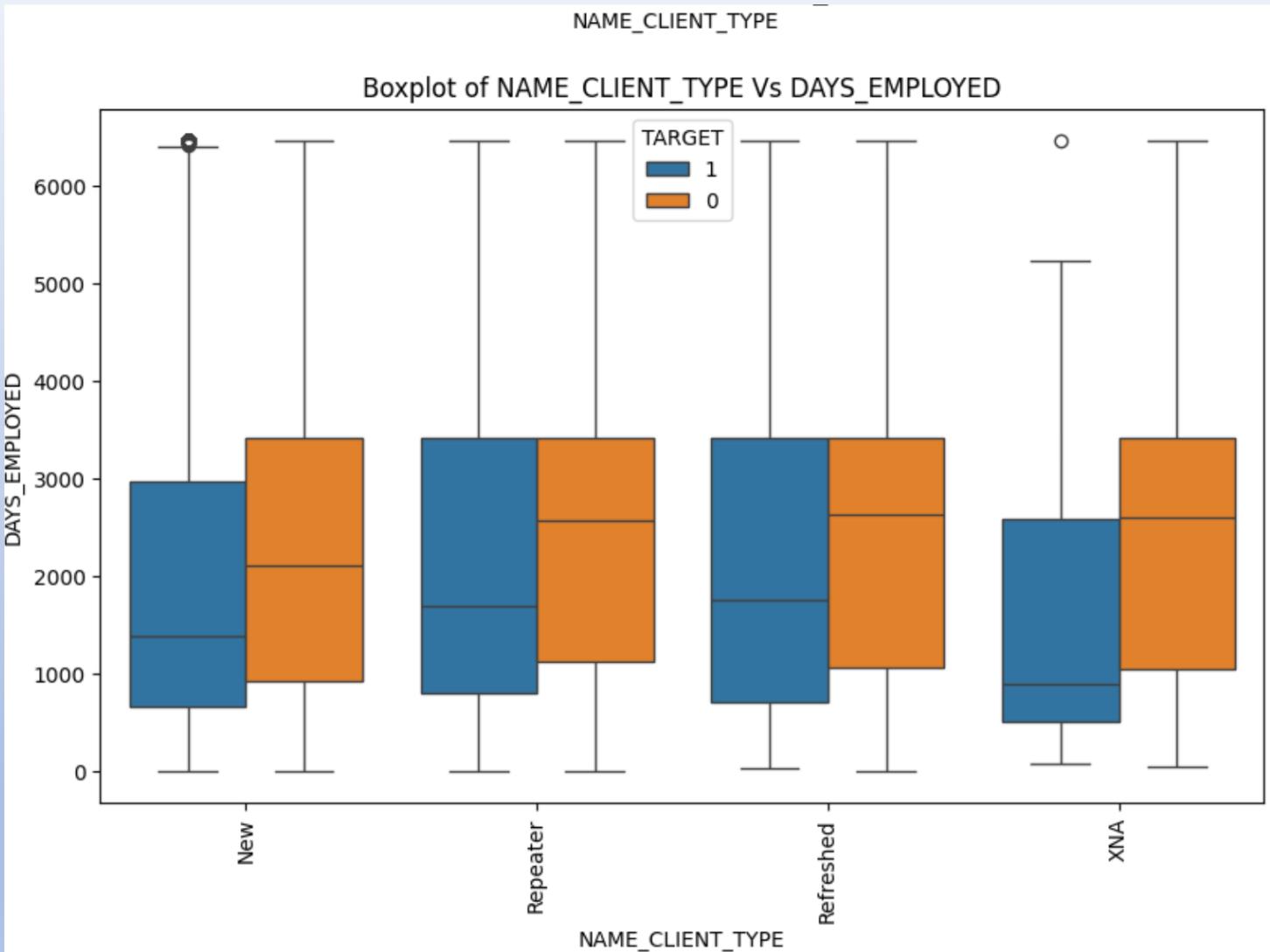
## NAME\_CASH\_LOAN\_PURPOSE VS AMT\_CREDIT\_y

- THE CLIENTS WHO ARE BUYING A HOLIDAY HOME/LAND AND TEND TO DEFAULT HAVE HIGHER MEDIAN VALUE FOR AMT\_CREDIT\_Y



## NAME\_CLIENT\_TYPE VS DAYS\_EMPLOYED

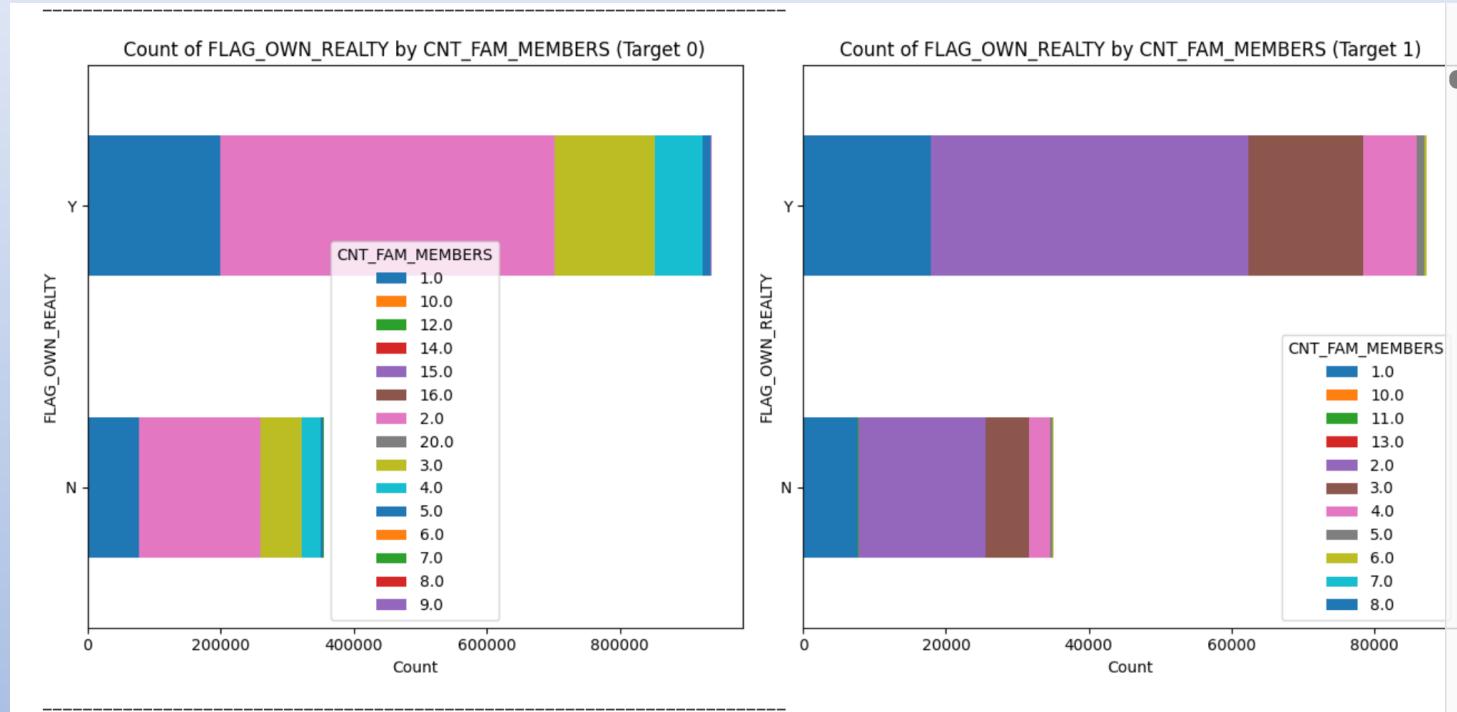
- FOR ALL THE CLIENT TYPES, THE MEDIAN VALUE OF DAYS\_EMPLOYED IS LESS FOR DEFAULTERS



CATEGORICAL VS CATEGORICAL

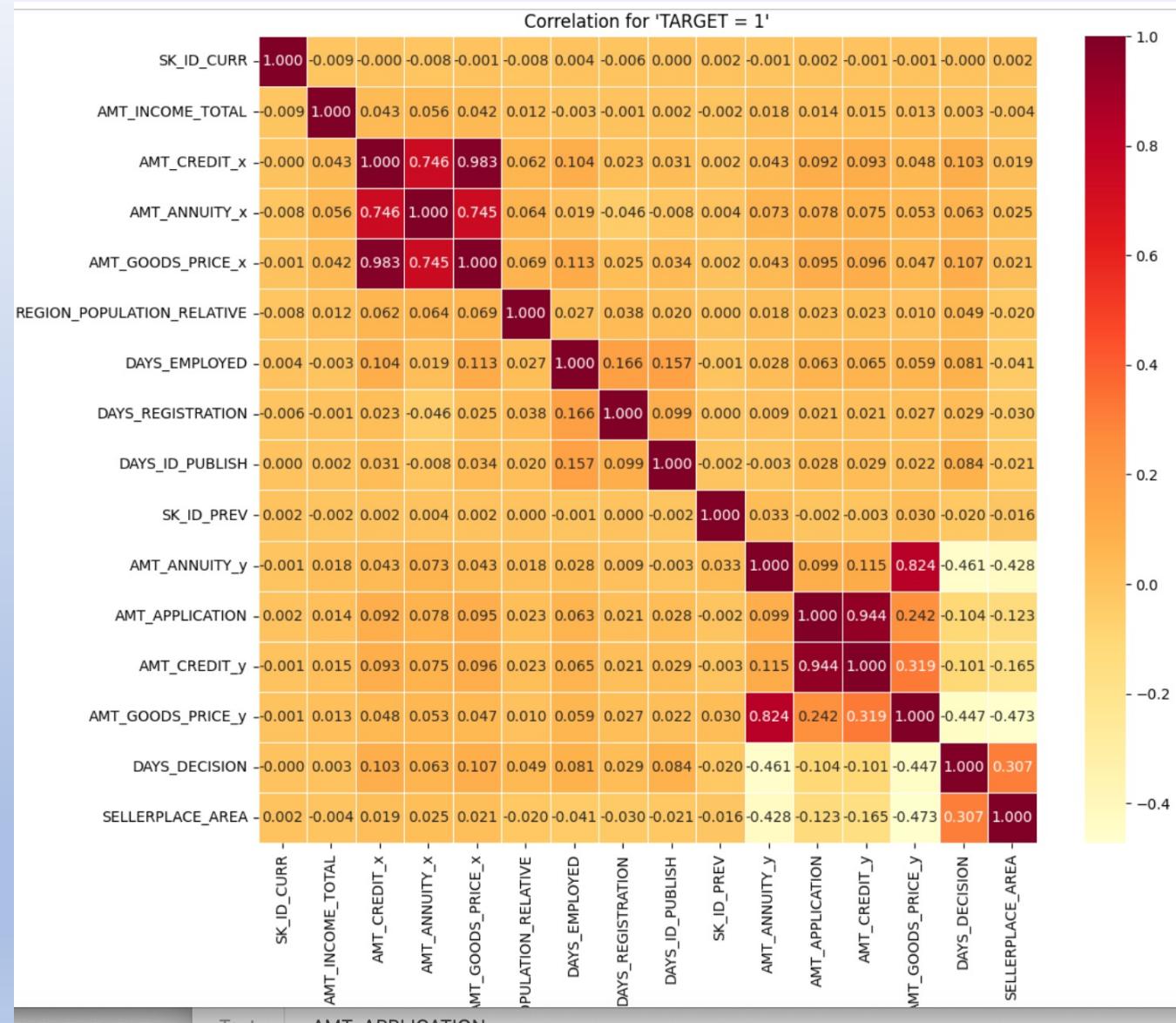
## FLAG\_OWN\_REALTY VS CNT\_FAM\_MEMBERS

- MOST OF THE CLIENTS WHO OWN HOME HAVE COUNT OF FAMILY MEMBERS AS 2.



# MULTIVARIATE ANALYSIS

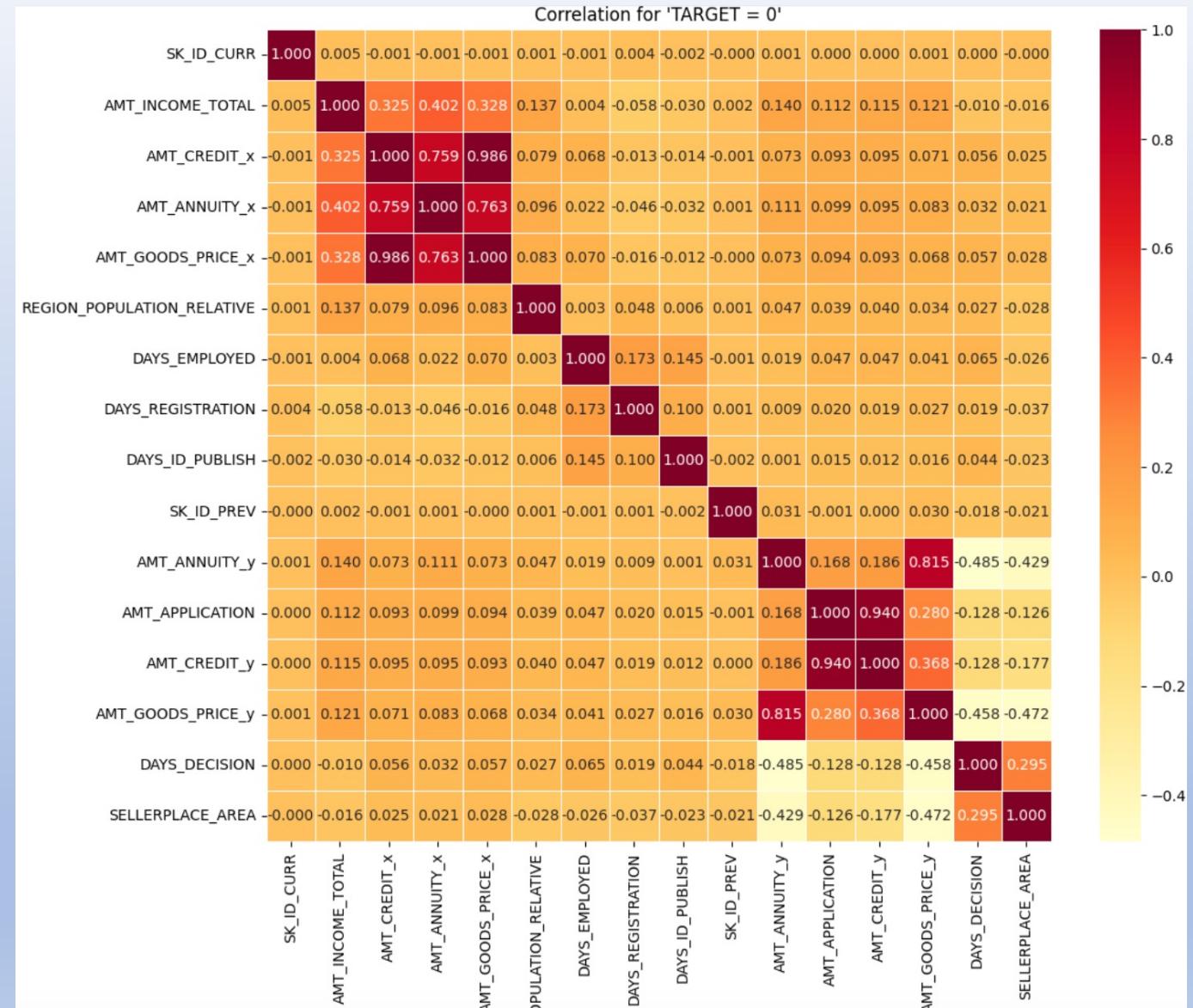
# CORRELATION FOR TARGET=1



# TOP 5 CORRELATIONS FOR TARGET=1

- AMT\_GOODS\_PRICE\_x AMT\_CREDIT\_x 0.982525
- AMT\_CREDIT\_y AMT\_APPLICATION 0.943839
- AMT\_GOODS\_PRICE\_y AMT\_ANNUITY\_y 0.823913
- AMT\_ANNUITY\_x AMT\_CREDIT\_x 0.745679
- AMT\_ANNUITY\_x AMT\_GOODS\_PRICE\_x 0.745083

# CORRELATION FOR TARGET=0



## TOP 5 CORRELATIONS FRO TARGET=0

- AMT\_GOODS\_PRICE\_x AMT\_CREDIT\_x 0.986402
- AMT\_CREDIT\_y AMT\_APPLICATION 0.939569
- AMT\_GOODS\_PRICE\_y AMT\_ANNUITY\_y 0.814557
- AMT\_ANNUITY\_x AMT\_GOODS\_PRICE\_x 0.763397
- AMT\_CREDIT\_x AMT\_ANNUITY\_x 0.759380

# CONCLUSION

- AMT\_INCOME\_TOTAL HAVE OUTLIERS INDICATING THERE ARE HIGH INCOME HOLDERS
- SHOULD CONCENTERATE ON FEMALE CLIENTS COMPARED TO MALE CLIENTS AS MALE CLIENTS FIND MORE DIFFICULTY IN REPAYING
- CLIENTS WITH WORKING CATEGORY FIND MORE DIFFICULTY IN REPAYING THE LOAN
- CLIENTS WITH SECONDARY EDUCATION FIND DIFFICULTY IN REPAYING THE LOAN
- SHOULD CONCENTERATE MORE ON CLIENTS WITH HIGHER EDUCATION
- SHOULD CONCENTERATE MORE ON CLIENTS WHO ARE MARRIED

# CONCLUSION

- CLIENTS WHO ARE LIVING WITH PARENTS FIND DIFFICULTY IN REPAYING THE LOAN
- CLIENTS HAVING OCCUPATION TYPE-LABORERS,SALES STAFF,DRIVERS,COOKING STAFF AND LOW SKILLED LABORERS TEND TO DEFAULT
- CLIENTS LIVING IN THE REGION RATING=3 TEND TO DEFAULT
- BUSINESS ENTITY TYPE3,SELF EMPLOYED AND CONSTRUCTION BASED CLIENTS TEND TO DEFFAULT
- CLIENTS IN THE AGE GROUP 20-40 TEND TO DEFAULT
- THERE ARE NO DEFALTERS IN STUDENT AND BUSINESSMAN CATEGORY

# CONCLUSION

- THE MEDIAN VALUE OF NUMBER OF DAYS EMPLOYED IS LOW FOR DEFAULTERS.SO CONCENTERATE ON CLIENTS WHO HAVE MORE NUMBER OF DAYS EMPLOYED
- CLIENTS IN THE AGE GROUP <20 OPTED ONLY FOR CASH LOANS AND THERE ARE NO DEFAULTERS IN THAT AGE GROUP
- CLIENTS WHO HAVE NAME\_PORTFOLIO AS CARDS TEND TO DEFAULT MORE
- WALKIN CATEGORY CLIENTS TEND TO DEFAULT MORE
- CLIENTS ACQUIRED THROUGH CREDIT AND CASH OFFICES, AP+(AFFINITY PARTNERS) TEND TO DEFAULT MORE
- CLENTS WITH SINGLE PAYMENT(CNT\_PAYMENT=0) TEND TO DEFFAULT MORE
- CONNECTIVITY CATEGORY TEND TO DEFAULT MORE

# CONCLUSION

- CLIENTS IN THE LOW\_NORMAL CATEGORY(NAME\_YIELD\_GROUP) HAVE LESS DIFFICULTY IN REPAYING THE LOAN
- NUMBER OF DAYS TAKEN FOR NON-DEFAULT CATEGORY IS HIGHER COMPARED TO DEFAULT CATEGORY
- CLIENTS IN THE HIGH SELLERPLACE\_AREA DOESN'T HAVE MUCH DIFFICULTY IN REPAYING THE LOAN
- THE CLIENTS WHO ARE BUYING A HOLIDAY HOME/LAND AND TEND TO DEFAULT HAVE HIGHER MEDIAN VALUE FOR AMT\_CREDIT\_Y