

# EDA ASSIGNMENT

NAME : AAGAM SHAH





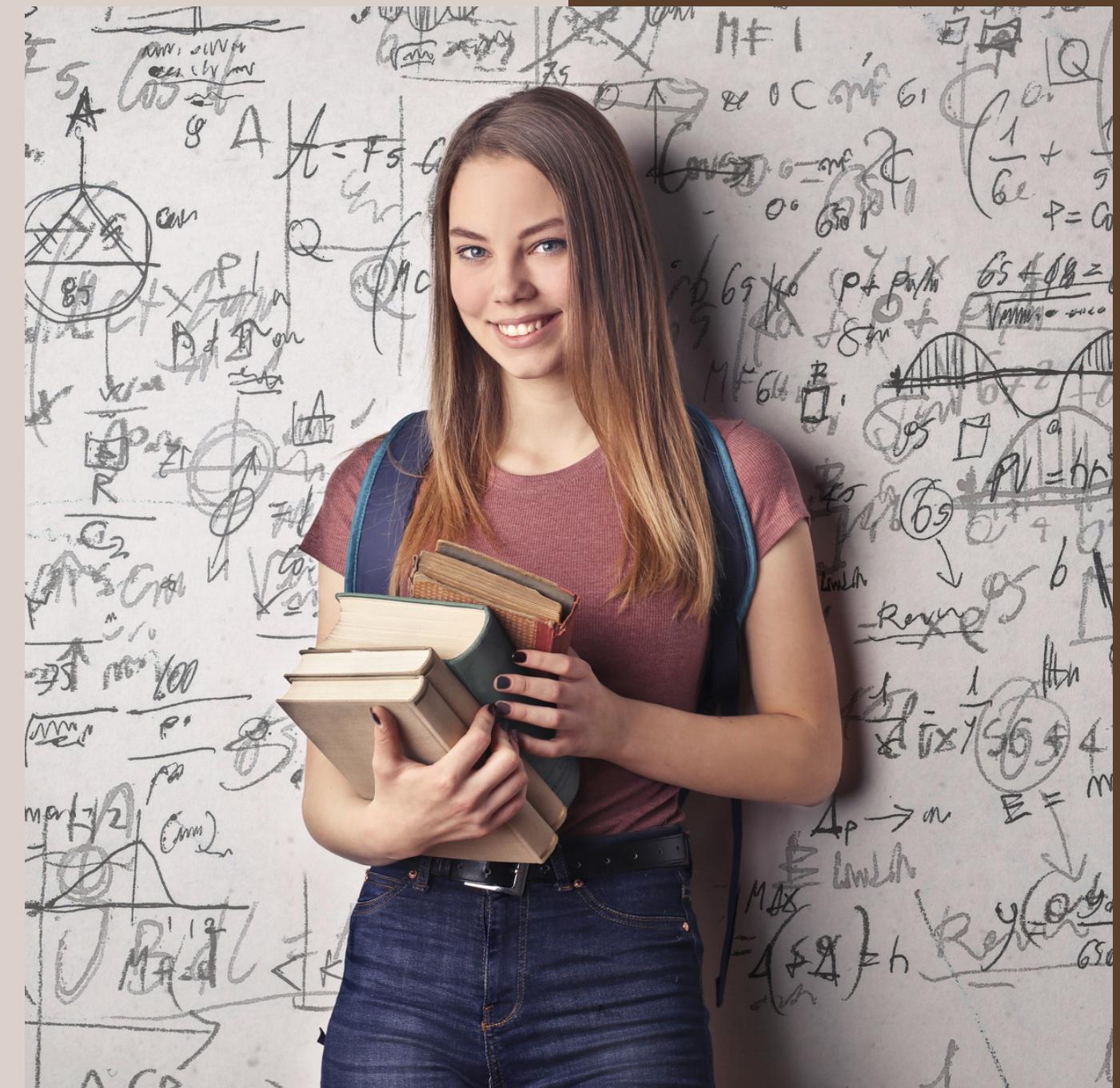
# OUTLIER

1. Objective
2. Approach / Steps
3. Graph and Insight
4. Conclusion

# OBJECTIVES

When a loan application is received, the business must evaluate whether to approve the loan based on the applicant's profile. The bank's choice is subject to two different kinds of risks:

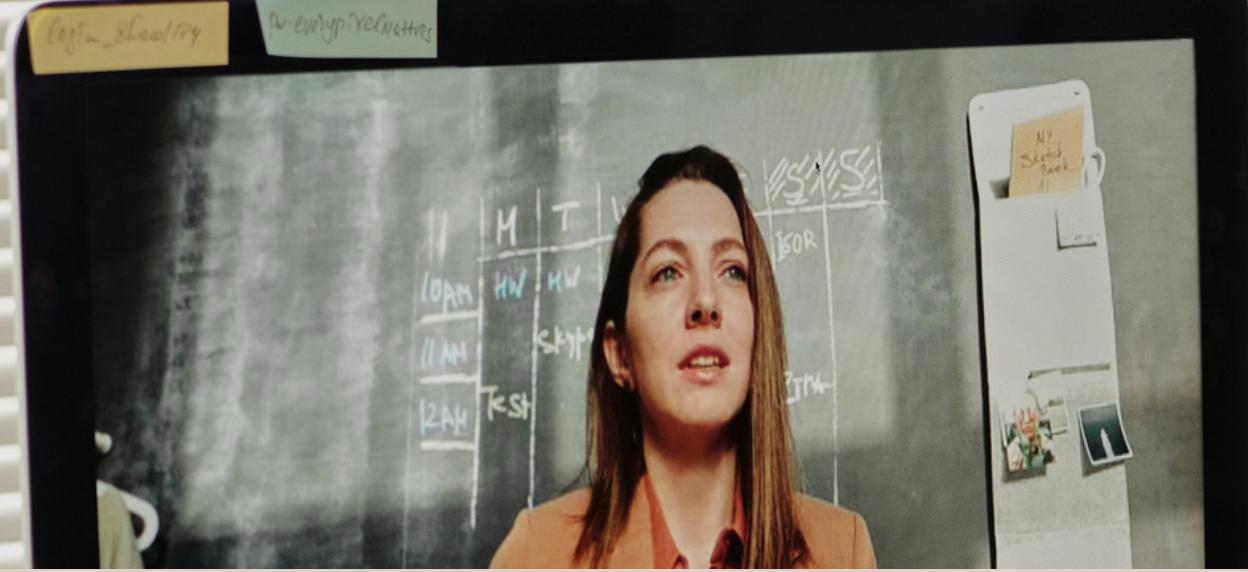
- If the borrower is likely to repay the loan, refusing to provide it results in the company losing business.
- If the applicant is unlikely to repay the loan, i.e., if he or she will default, accepting the loan may result in a financial loss for the firm.
- Identification of driving factors/indicators behind loan default for portfolio and risk assessment.



# S T E P S



- 1. Loading data from CSV file.  
(application\_data.csv)**
- 2. Getting basic info of the data.**
- 3. Analyzing the data.**
- 4. Checking the data imbalance percentage.**
- 5. Univariate analysis for Categorical and Numerical columns.**
- 6. Bivariate Analysis : Numerical - Categorical**
- 7. Bivariate Analysis : Numerical - Numerical**
- 8. Finding the correlation**
- 9. Loading the data from CSV file.  
(previous\_application.CSV)**
- 10. Getting basic info and analyzing it.**
- 11. Final Conclusion.**



# GETTING BASIC INFO

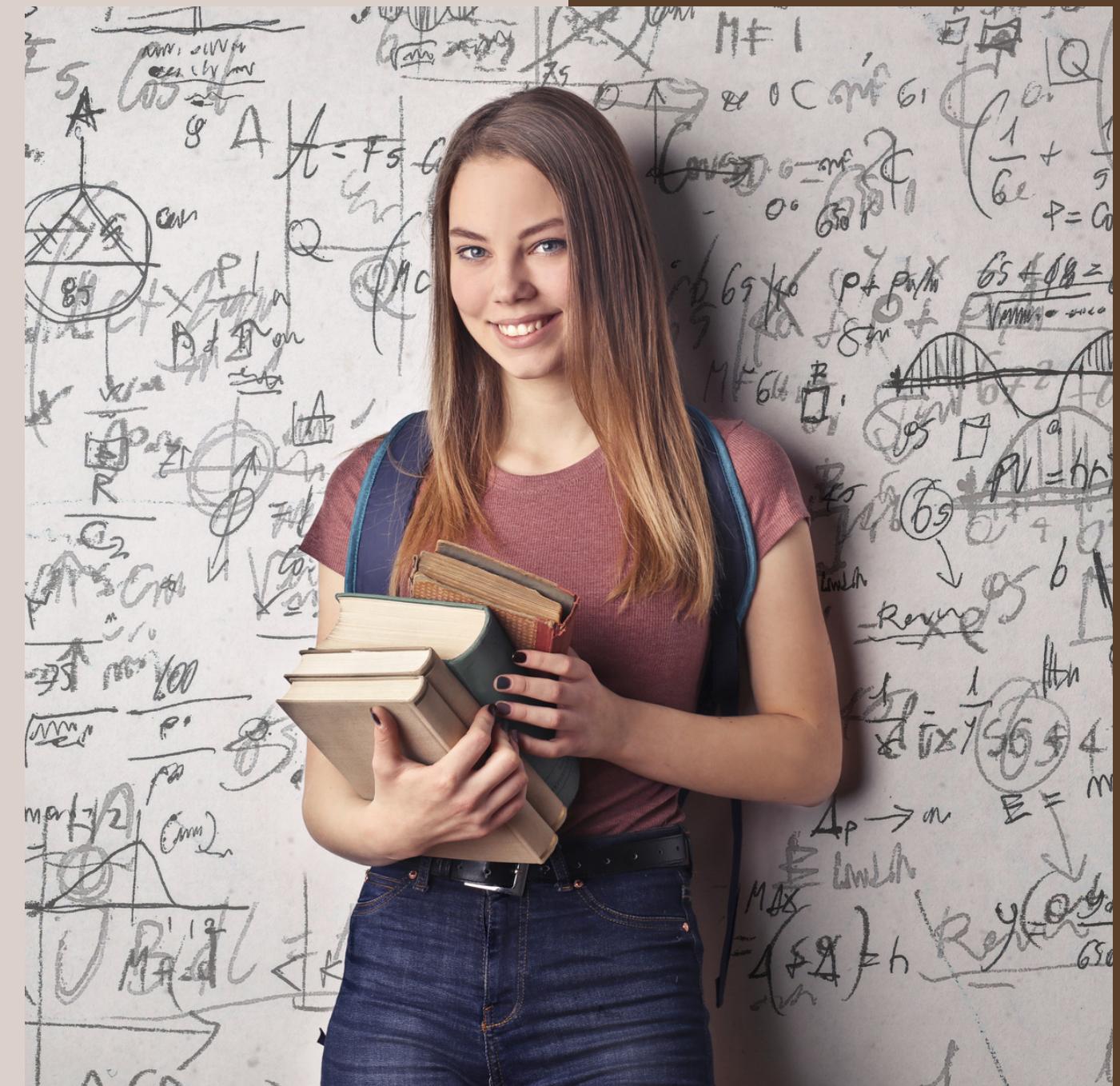
- Inspecting and understanding Data
  - Checking a few Dataset records such as `.shape`, `.info()`, and `.describe()`
- Data Cleaning
  - In descending order, check the percentage of null values in the data frame.
  - Examining the amount of null columns
  - Columns with null values more than 35% were removed.
  - Columns with empty values were imputed via Mode values for numerical columns, except for recurring numeric columns, which were imputed with the Median value.
- Handling Errors in Data types and Data
  - We discovered columns having negative or mixed values in the data frame: '`DAYS_BIRTH`', '`DAYS_EMPLOYED`', '`DAYS_REGISTRATION`', '`DAYS_ID_PUBLISH`', and '`DAYS_LAST_PHONE_CHANGE`', therefore we imputed them as absolute values for the purpose of our analysis.

- Then, for ease of analysis, we modified the values of columns FLAG\_OWN\_CAR and FLAG\_OWN\_REALTY from 'Y' to 1 and 0 accordingly.
- We discovered the value 'XNA' in the field CODE\_GENDER, which basically indicates Not Available. As a result, we imputed those values with the most common value (median value), i.e., F.
- Then we saw that the column ORGANIZATION\_TYPE has 18% 'XNA' entries as well. So we first determined if the values are Not present Completely at Random (MCAR), Missing at Randomised (MAR), or Missing Not at Random (MNAR). After comparing the contents of ORGANIZATION\_TYPE with the values of column NAME\_INCOME\_TYPE, we discovered that clients who are pensioners had XNA values in ORGANIZATION\_TYPE. So we changed "XNA" to "Pensioner".
- Similarly, we discovered the similar relationship between OCCUPATION\_TYPE and NAME\_INCOME\_TYPE, thus we substituted "Pensioner" for null values.
- We developed binning depending on the the quantiles for the columns listed below for ease of analysis:
- AMT\_CREDIT, AMT\_INCOME\_TOTAL, AND AGE\_DAYS



# ANALYSIS THE DATA

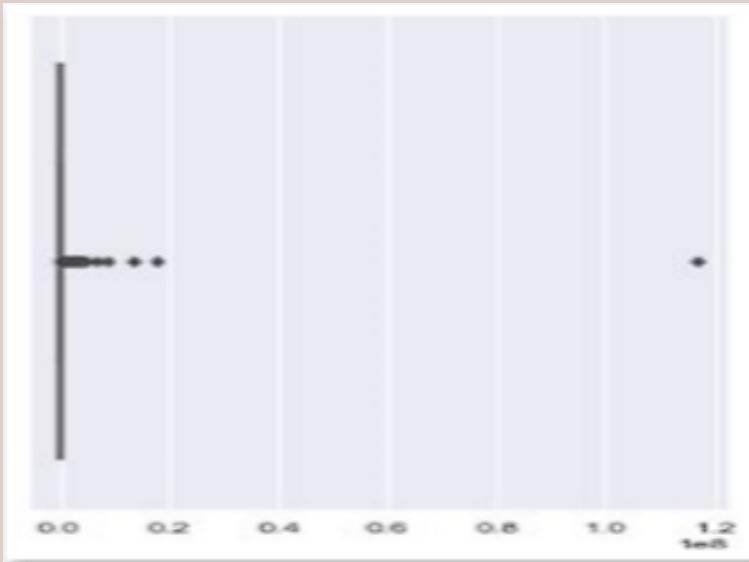
We noticed several columns in the data frame with the d-type "object," which can be transformed to the d-type "category," which will be more suitable for our analysis and will save memory. The data frame has a large number of columns. We'll eliminate any columns that aren't required for further investigation.



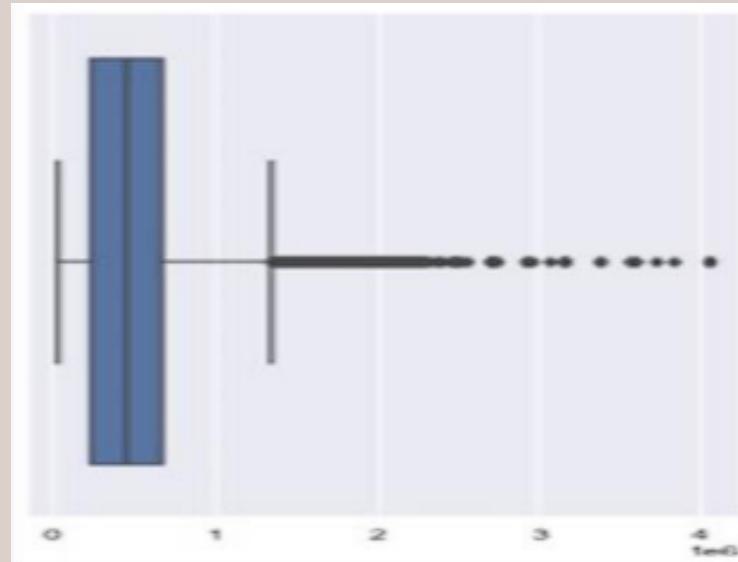
**WE MADE LIST OF ALL NUMERICAL COLUMNS .AND THEN PLOT BOXPLOTS FOR EACH NUMERICAL COLUMNS.**

**(VISUAL REPRESENTATION BY BOXPLOT IS SHOWN)**

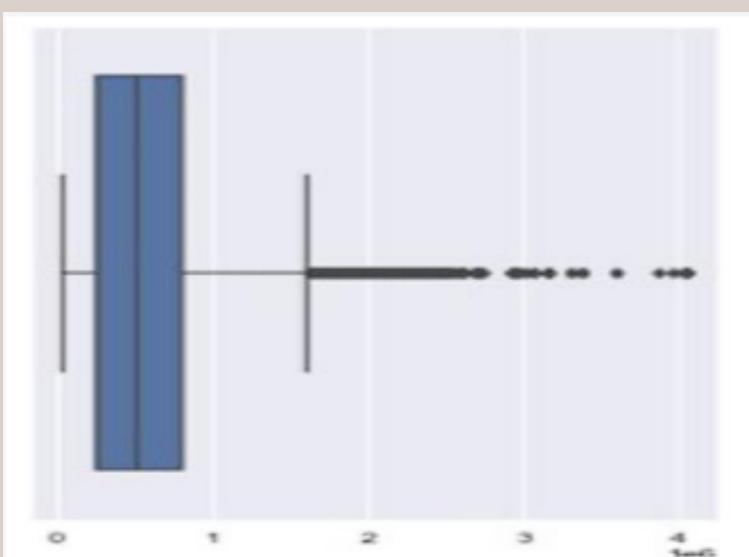
- The IQR for AMT\_INCOME\_TOTAL is unusually narrow, with several outliers. It has a maximum value of 117000000, which is a significant difference from the 75th percentile.
- The third quarter of AMT\_GOODS\_PRICE, AMT\_CREDIT is bigger than the first quartile, indicating that the majority of the Credit amount of the loan of consumers is included in the third quartile. The largest value here is 4050000, which deviates significantly from the 75th percentile.
- DAYS\_EMPLOYED has a maximum value of 375000, which varies greatly from the mean and the 75th percentile.



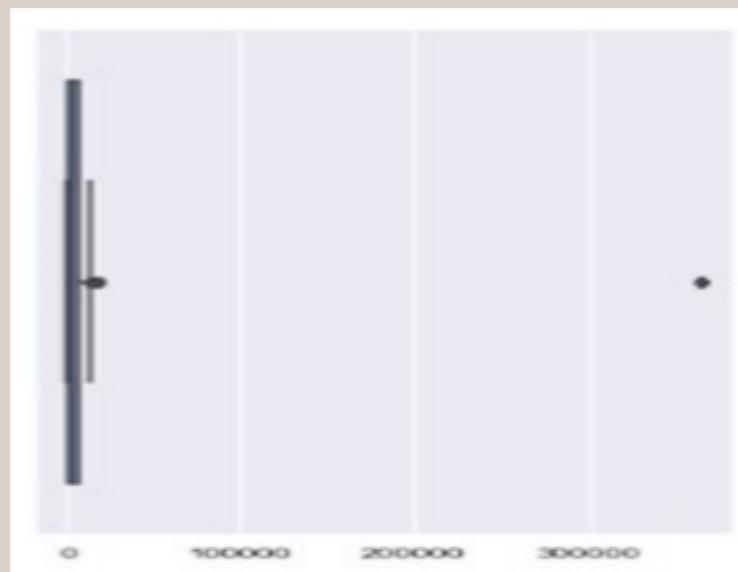
**AMT\_INCOME\_TOTAL**



**AMT\_GOODS\_PRICE**



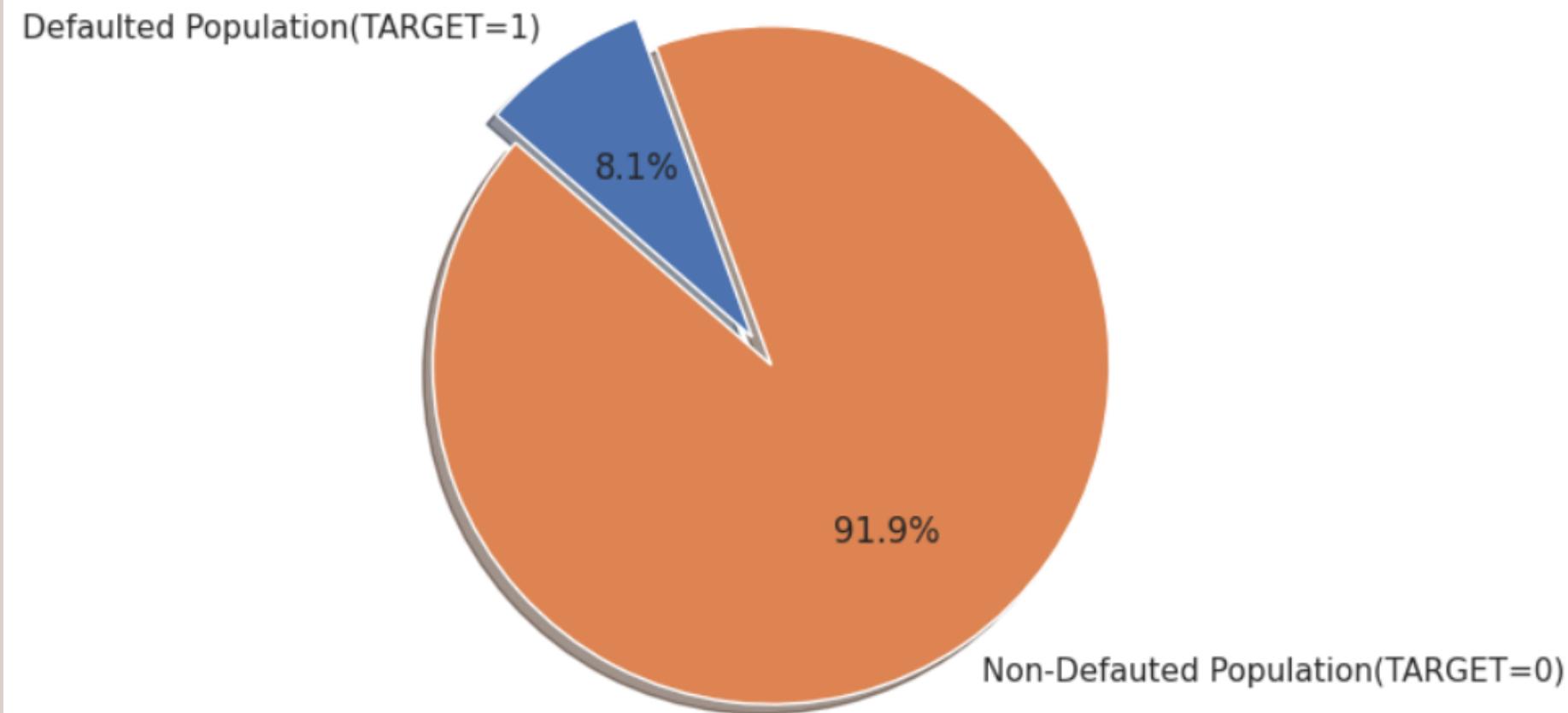
**AMT\_CREDIT**



**DAYS\_EMPLOYED**

## CHECKING DATA IMBALANCE FOR TARGET VARIABLE

### Data imbalance



- Because there is a significant imbalance between the TARGET variables 0 and 1, it makes more sense to partition the data frame into two sub datasets before continuing our research.
- We have splits data frame as follows:
  - Target0 : (Non-Defaulted Population) Clients without Payment Difficulties.
  - Target1 : (Defaulted Population) Clients with Payment Difficulties.

## CONCLUSION

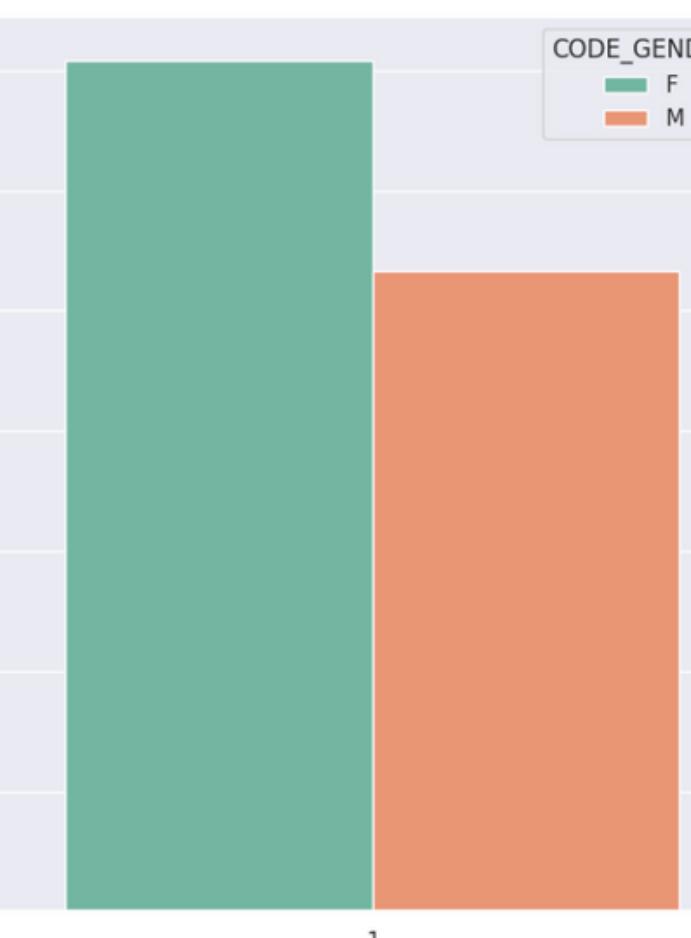
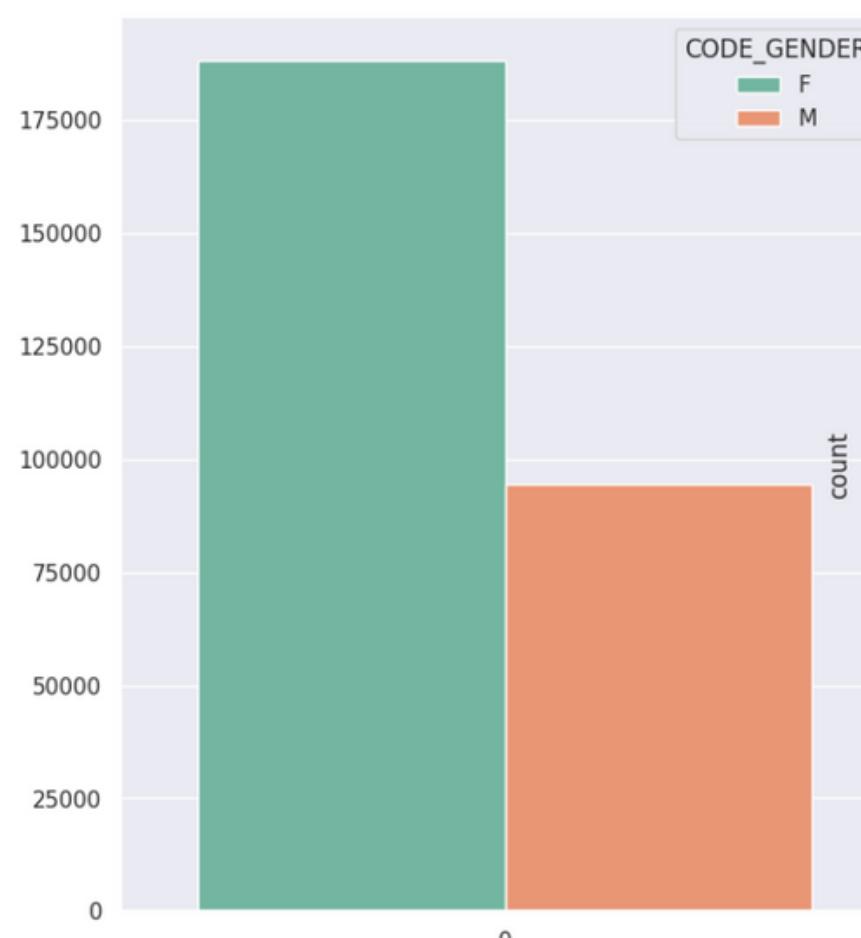
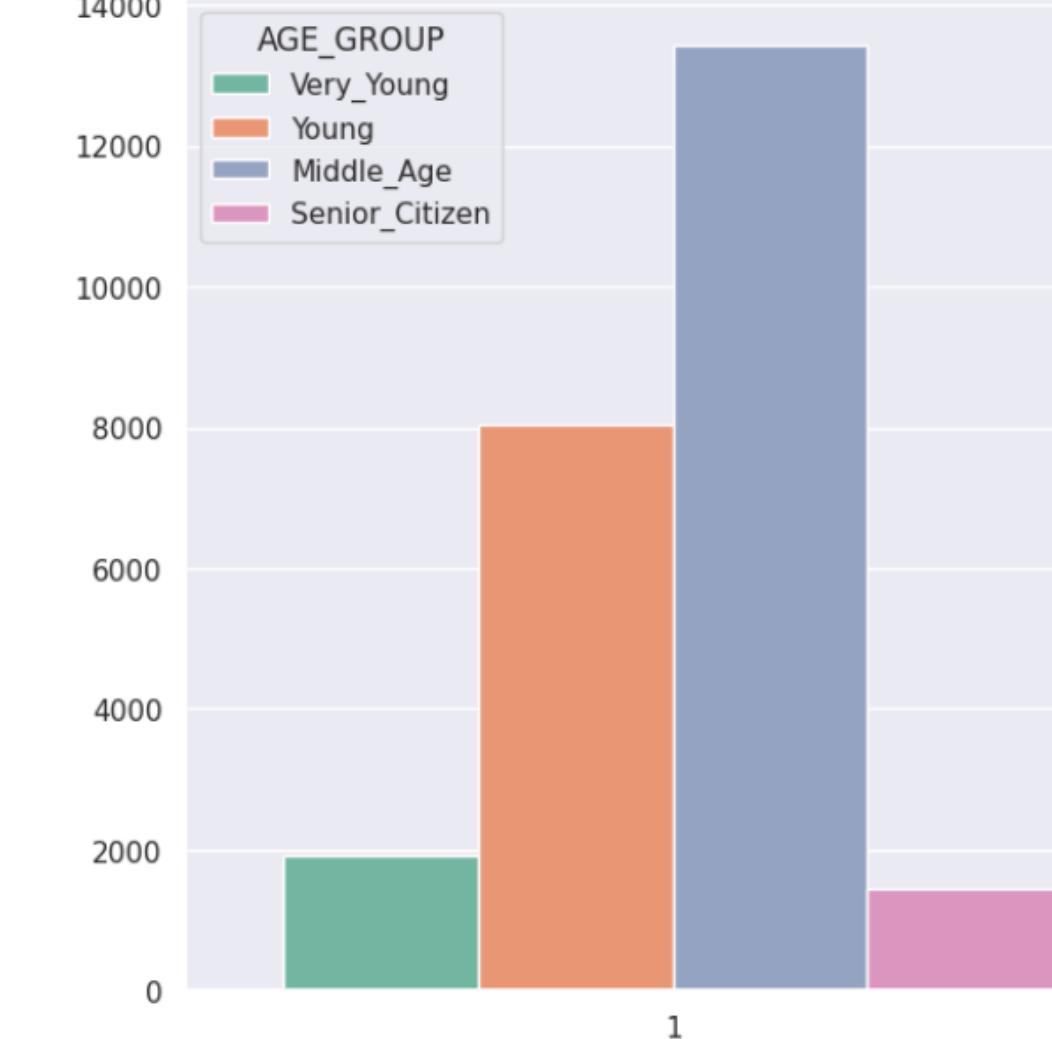
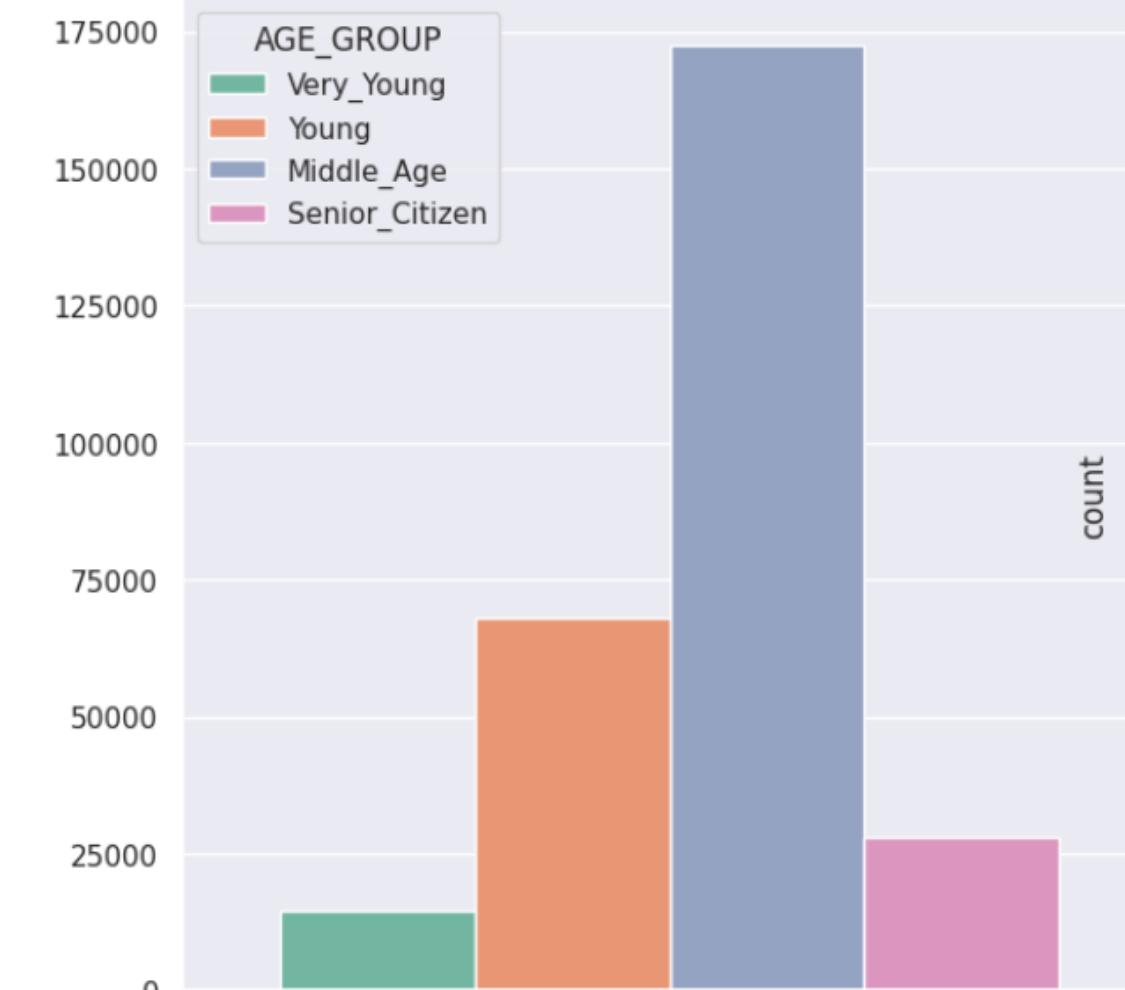
- Ratio of Data Imbalance is 11.3.
- In order to analyze the imbalance and various aspects of data we will perform various types of analysis such as: Univariate analysis , Bivariate analysis , Multivariate analysis

## AGE GROUP WITH RESPECT TO TARGET VARIABLES

- MIDDLE-AGED PEOPLE (35-60) HAVE APPLIED THE MOST AND HAVE THE MOST PAYMENT ISSUES OF ANY AGE GROUP.**
- WHILE SENIOR CITIZENS (60 - 100) AND THE VERY YOUNG (19 - 25) FACE LESS FINANCIAL ISSUES THAN OTHER AGE GROUPS,**

Gender Distribution in Target0

Gender Distribution in Target1



## GENDER WITH RESPECT TO TARGET VARIABLES

- FEMALE CLIENTS APPEAR TO HAVE APPLIED FOR LOANS IN GREATER NUMBERS THAN MALE CONSUMERS.**
- PAYMENT TROUBLES AFFECT 66.6% OF FEMALE CLIENTS AND 33.4% OF MALE CONSUMERS.**
- 57% OF FEMALE CUSTOMERS AND 42% OF MALE CLIENTS ARE HAVING FINANCIAL ISSUES.**



## SIMILARLY INSIGHT WHERE DRAWN FROM OTHER COLUMN.

(DUE TO LAST MOMENT ERROR IN CODE, COULD NOT ABLE TO PASTE PHOTO OF GRAPH)

### 1. NAME\_CONTRACT\_TYPE

- MOST OF THE CLIENT HAVE APPLIED FOR CASH LOAN WHILE VERY SMALL PROPORTION HAVE APPLIED FOR REVOLVING LOAN FOR BOTH.

### 2. NAME\_TYPE\_SUITE

- MOST OF THE ACCOMPANIED WHILE APPLYING FOR THE LOAN AND WITH FEW CLIENT A FAMILY MEMBER WAS ACCOMPANYING FOR BOTH.
- BUT FOR THE ONE WHO WAS ACCCOMPANYING CLIENT WHILE APPLYING FOR THE LOAN DOESN'T IMPACT ON THE DEFAULT. ALSO BOTH HAVE SAME PROPORTION.

### 3. NAME\_INCOME\_TYPE

- CLIENTS SUCH AS WORKING, COMMERCIAL ASSOCIATE AND PENSIONER ARE MORE LIKELY TO APPLY FOR LOAN.
- BUSINESSMAN, STUDENTS AND UNEMPLOYED ARE LESS LIKELY TO APPLY FOR LOAN .
- WORKING HAVE MORE RISK TO DEFAULT.
- STATE SERVANT IS AT MINIMAL RISK TO DEFAULT.

### 4. NAME\_EDUCATION\_TYPE

- CLIENTS HAVING EDUCATION SECONDARY OR SECONDARY SPECIAL ARE MORE LIKELY TO APPLY FOR THE LOAN.
- CLIENTS HAVING EDUCATION SECONDARY OR SECONDARY SPECIAL ARE HAVING MORE RISK TO THE DEFAULT AS COMPARE TO OTHER.

## **5. NAME\_FAMILY\_STATUS**

- MARRIED CLIENTS APPEAR TO APPLY FOR THE LOAN FOR DEFAULT AND NON-DEFAULTER OF THEM THE MOST.
- CLIENT BEING SINGLE RELATIONSHIP ARE LESS RISKY AND WIDOWER SHOWS MINIMAL RISK. (DEFaulTER)

## **6. NAME\_HOUSING\_TYPE**

- THE CHART MAKES IT EVIDENT THAT MOST CLIENTS, BOTH DEFULTERS AND NON-DEFULTERS, OWN A HOME OR RESIDE IN A FLAT.

## **7. OCCUPATION\_TYPE**

- PENSIONERS ARE MOSTLY GOING TO APPLY FOR THE LOAN AND ARE HIGHLY RISKY.(DEFaulTER AND NON-DEFaulTER)

## **8. WEEKDAY\_APPR\_PROCESS\_START**

- THERE IS NO CONSIDERABLE DIFFERENCE.

## **9. AMT\_INCOME\_TYPE**

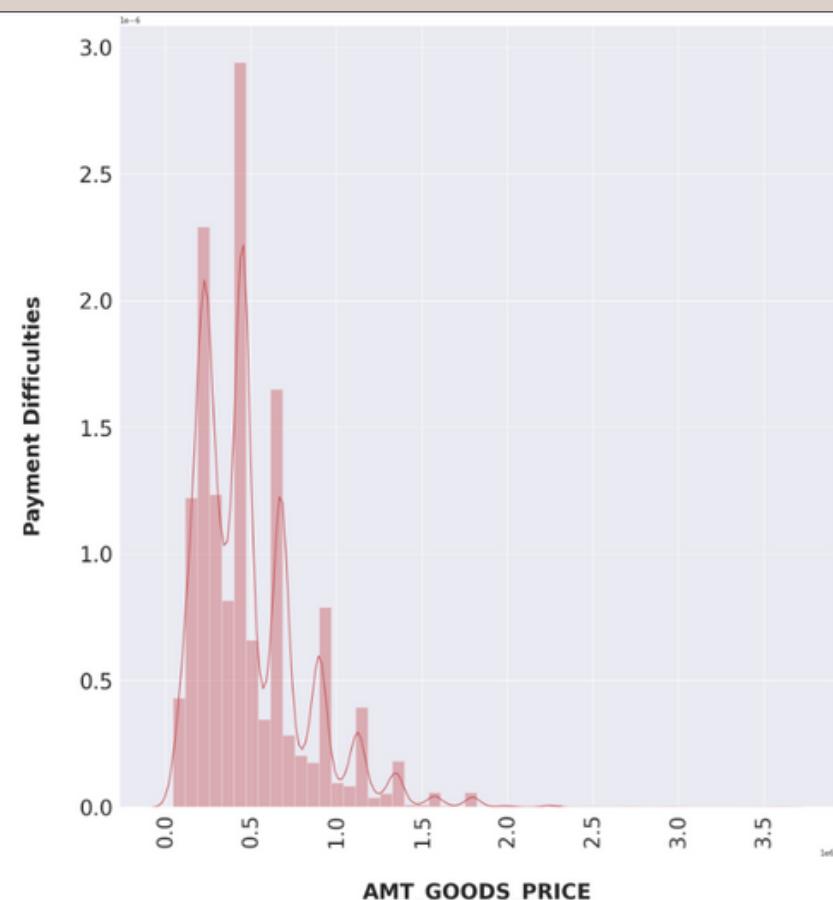
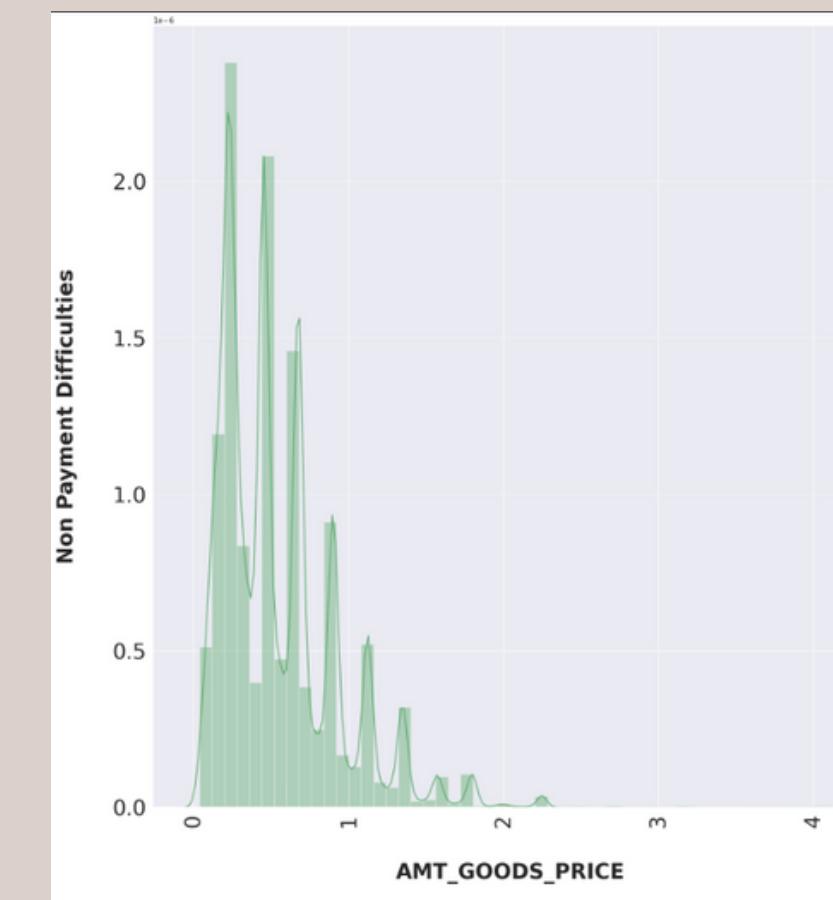
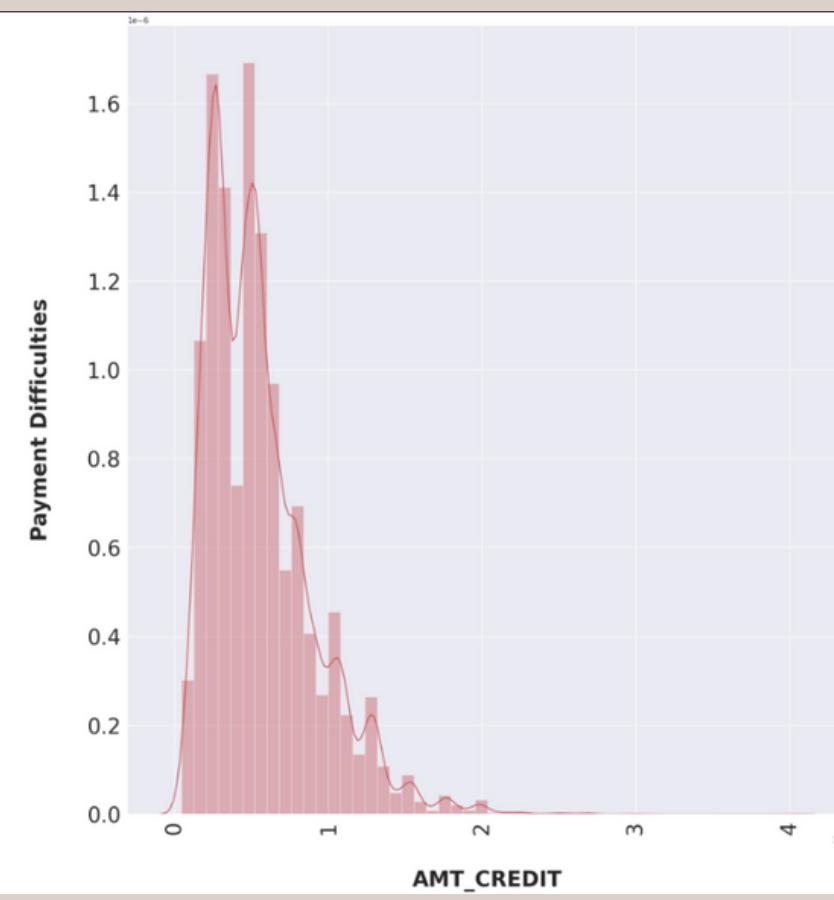
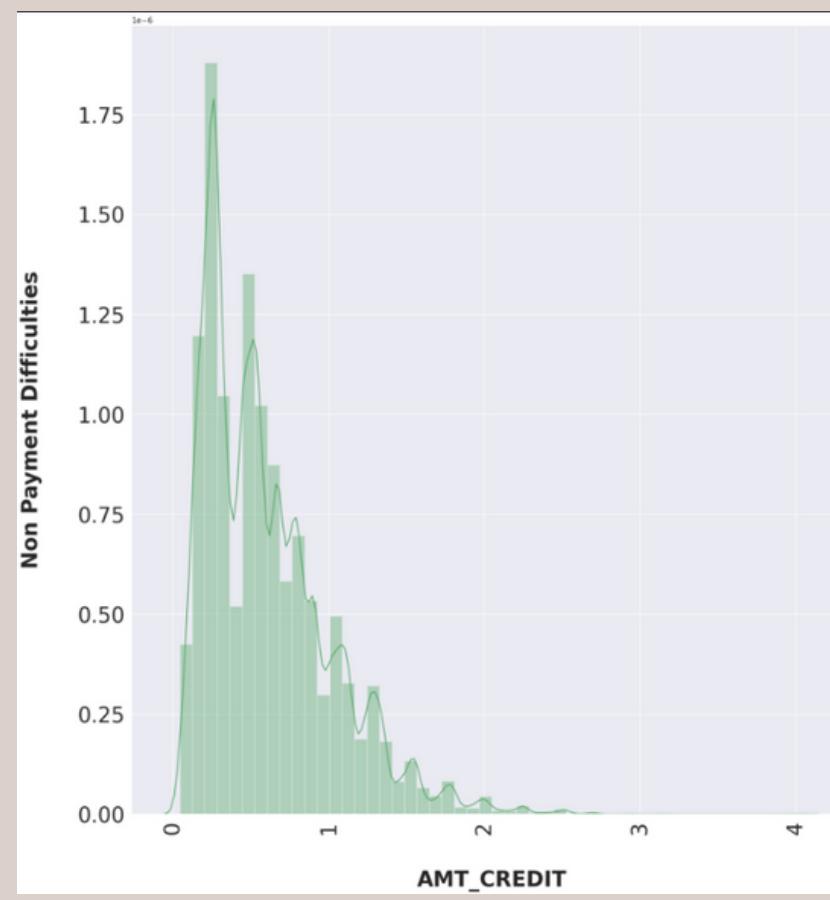
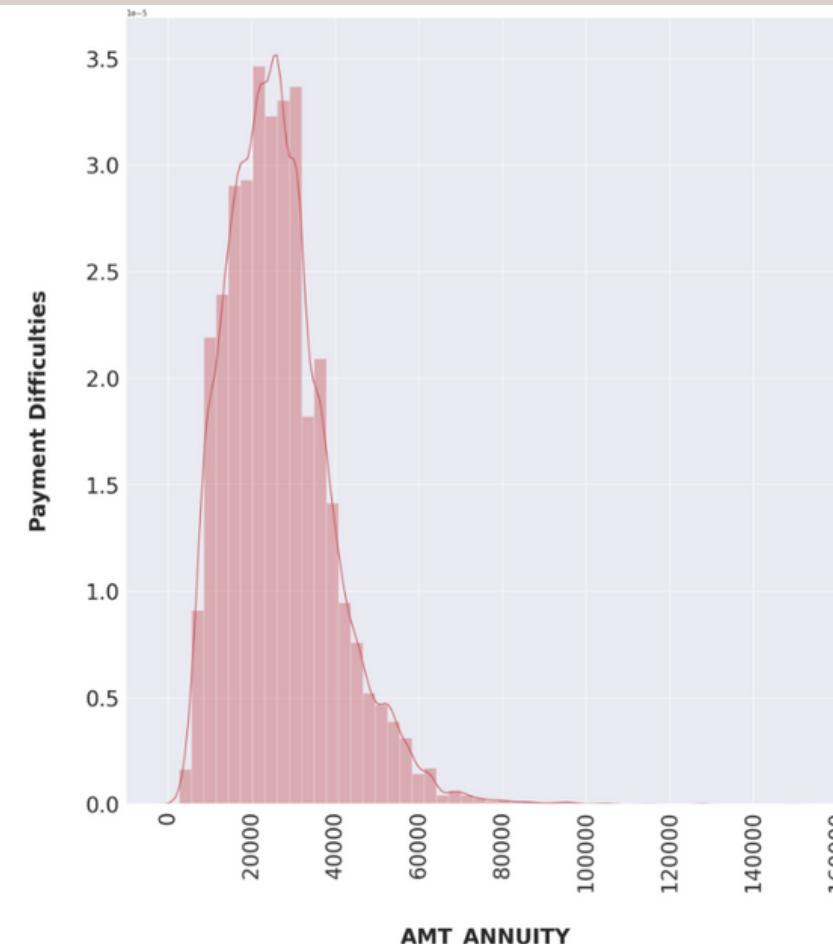
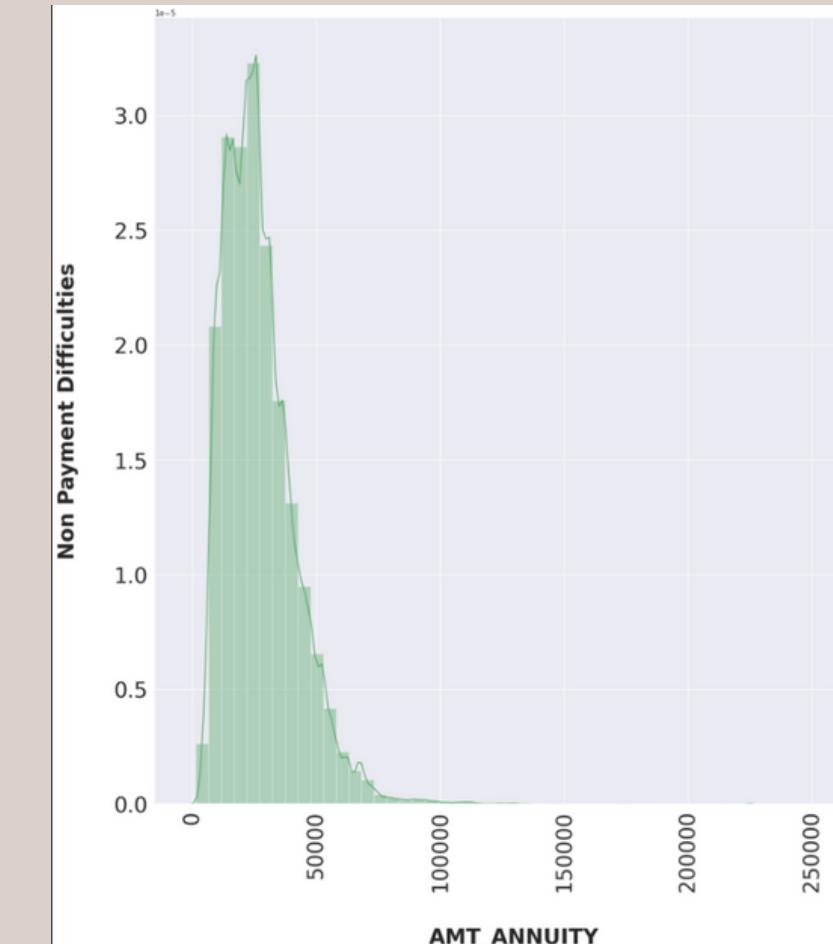
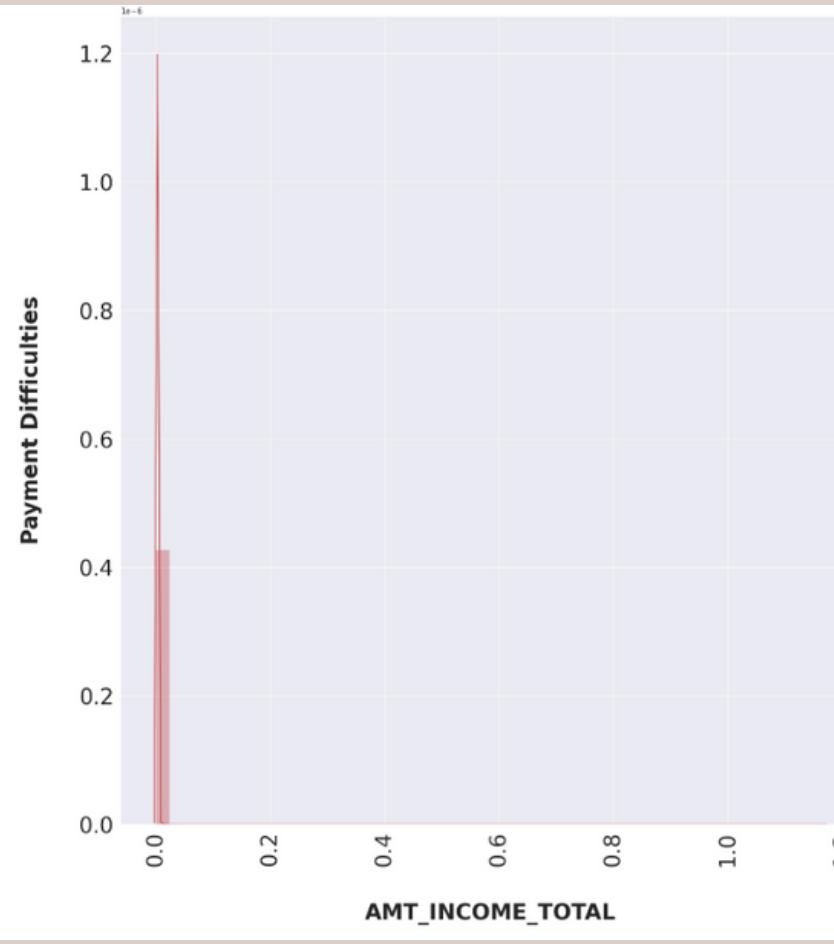
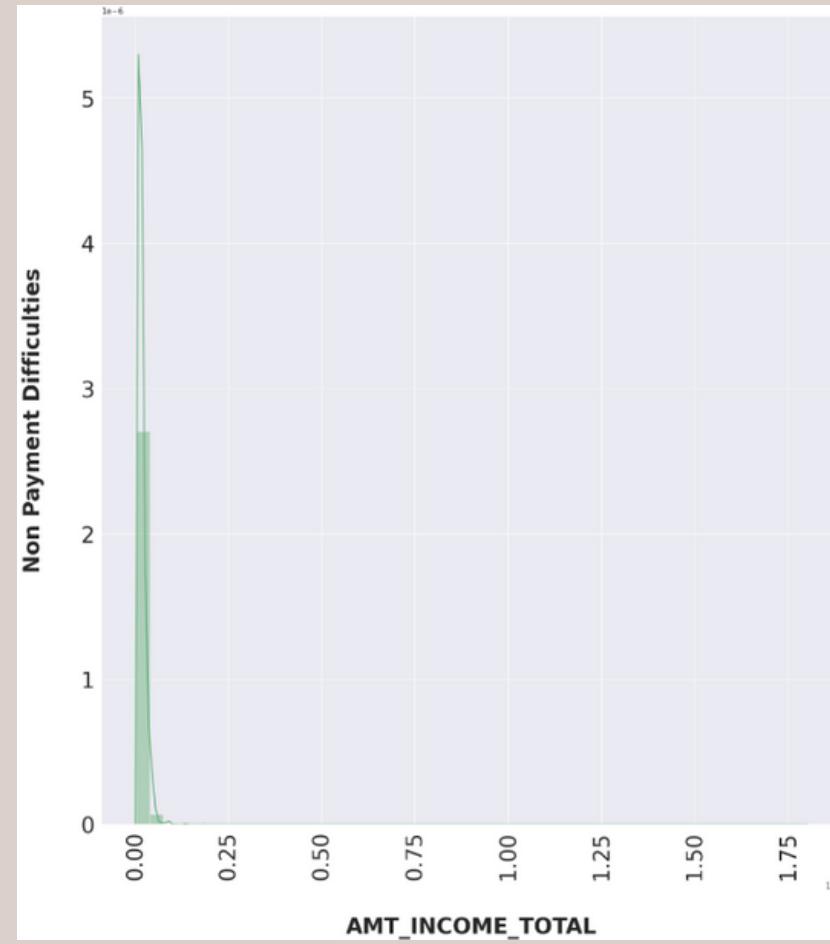
- CLIENT HAVING MEDIUM SALARY RANGE ARE MORE LIKELY TO APLY FOR THE LOAN.
- CLIENT WITH LOW AND MEDIUM SALARY ARE MORE RISKY.

## **10. AMT\_CREDIT\_TYPE**

- CLIENTS WHO APPLY FOR HIGH AND LOW CREDIT ARE AT THE HIGH RISK.
- MEDIUM CREDIT AMOUNT OF THE LOAN IS CHOSEN MORE BY CLIENT FOR BOTH DEFaulTER AND NOMN DEFaulTER.

# UNIVARIANTE ANALYSIS

## NUMERICAL



# CONCLUSION FROM GRAPH IN PREVIOUS SLIDE

01

Compared to target 0, those with target 1 have significantly more staggered income.

02

The distance plot clearly demonstrates that the shapes of the income total, annuity, credit, and good price for targets 0 and 1 are identical.

03

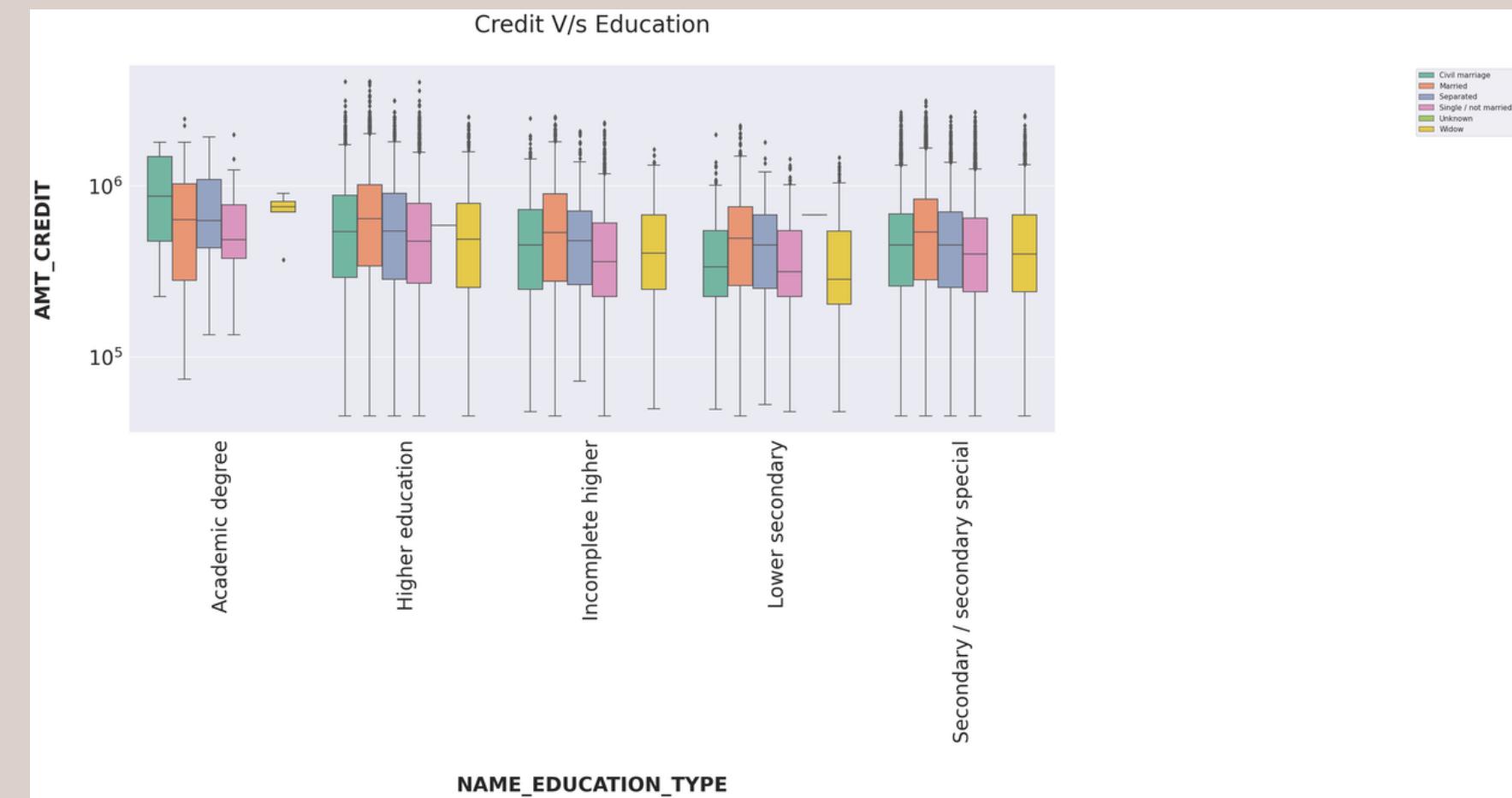
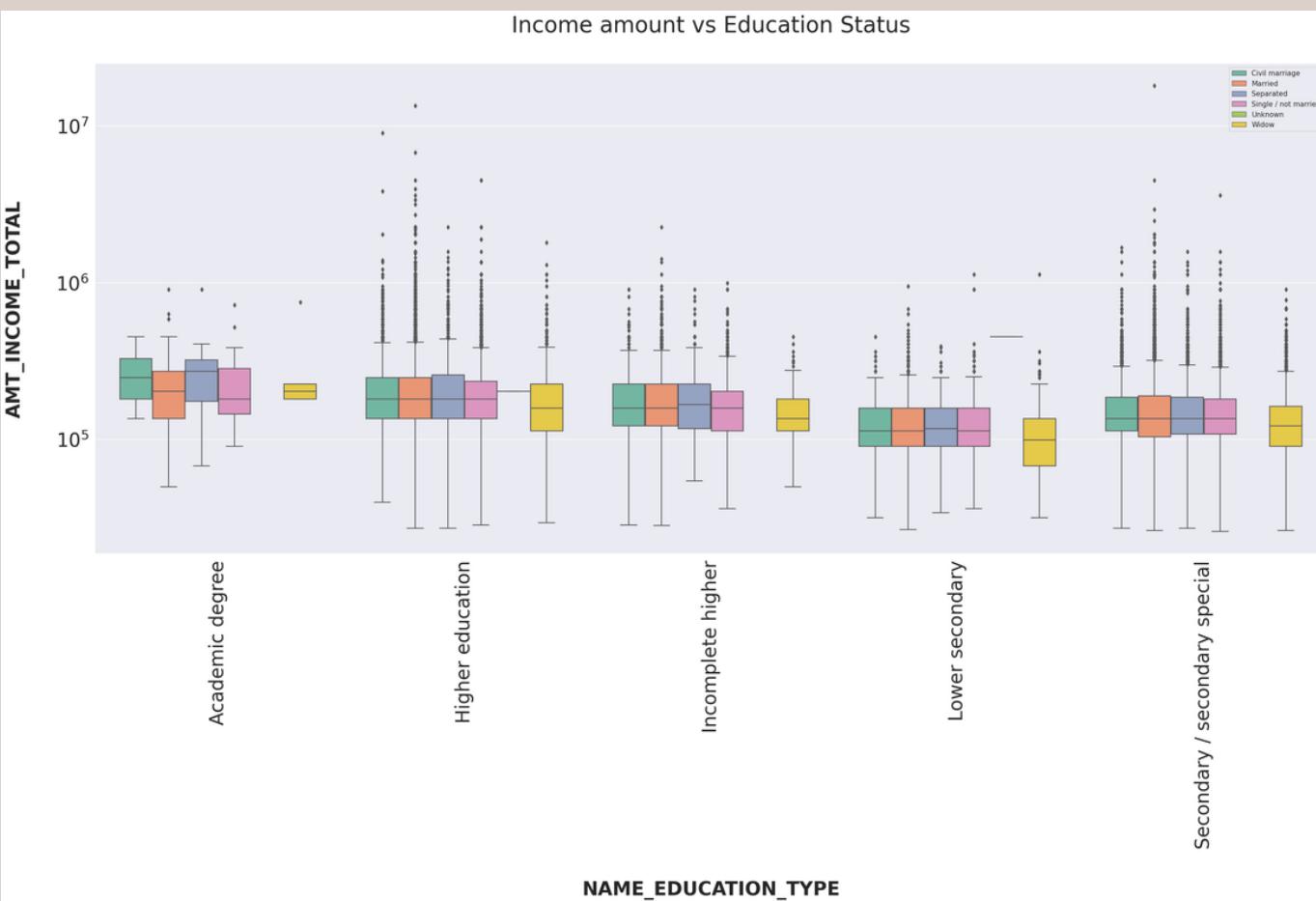
The plots also emphasise that those who struggle to repay loans due to their income, loan amount, cost of the commodities used to secure the loan, and Annuity.

04

The difference in the figure shows that Target 1 has a larger curve shape than Target 0, which has a smaller curve shape with clearly defined boundaries.

# BIVARIATE ANALYSIS NUMERICAL V/S CATEGORICAL

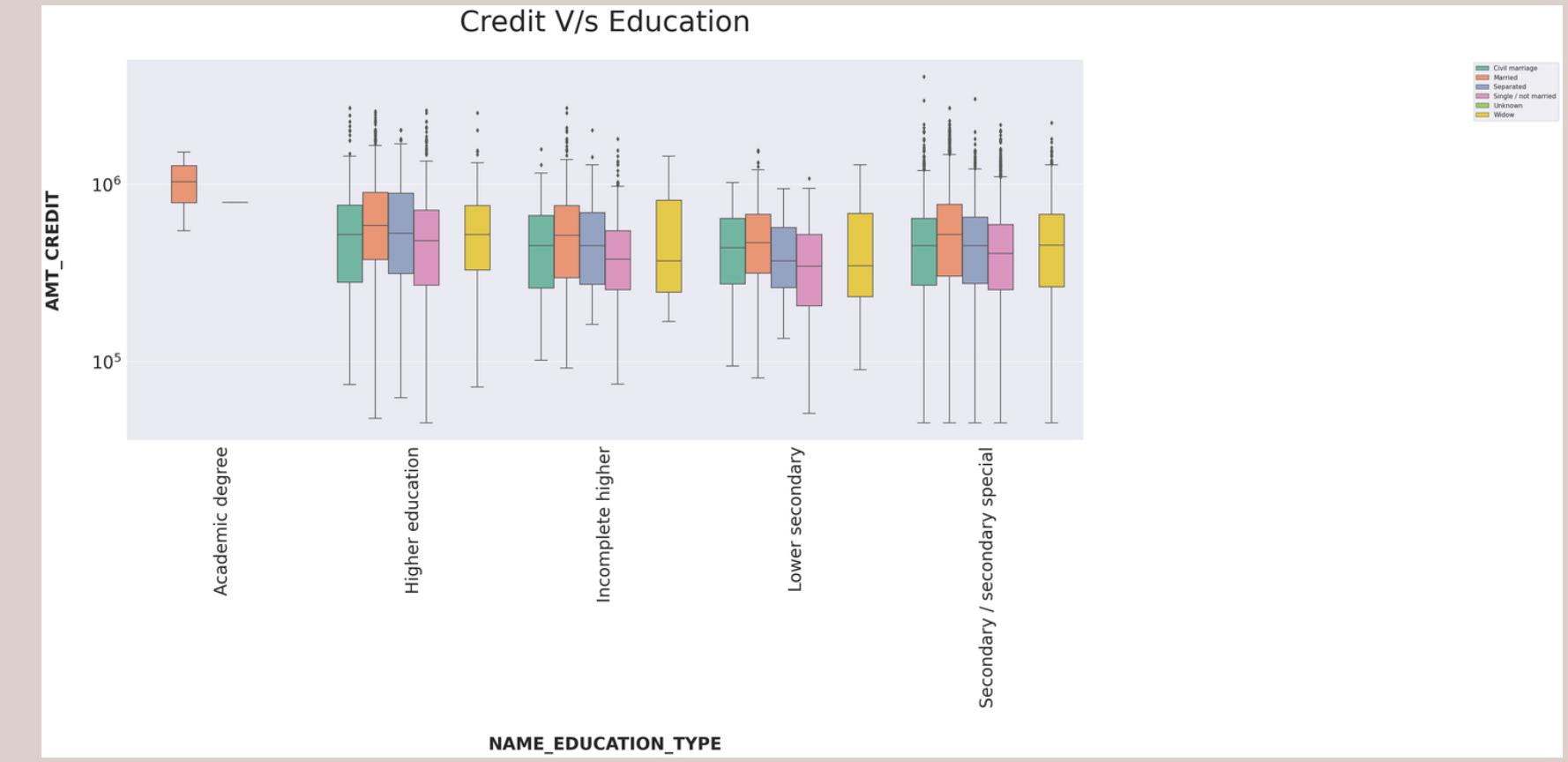
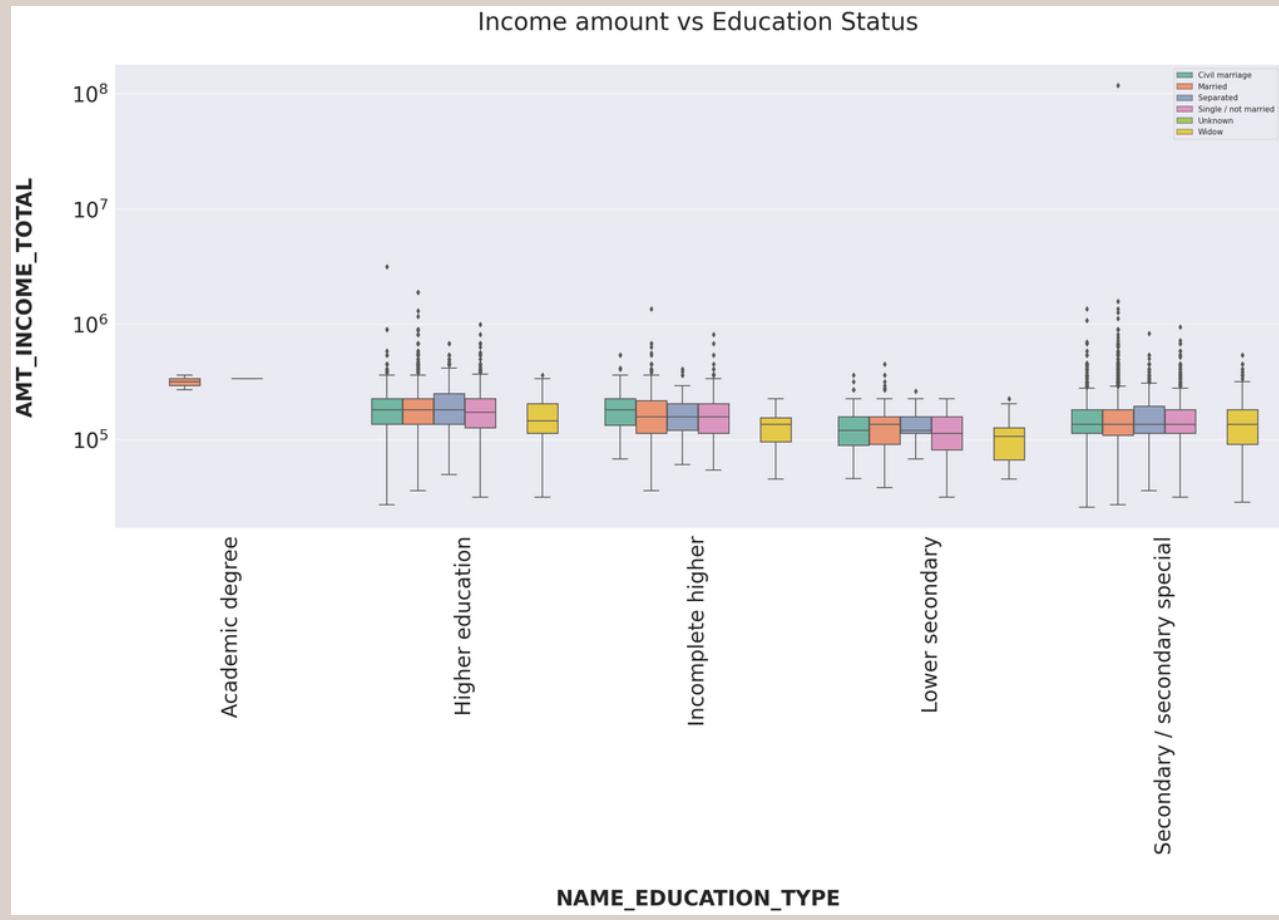
TARGET



- **WIDOW CLIENTS WITH ACADEMIC DEGREES DON'T FALL INTO THE FIRST OR THIRD QUARTILE AND ONLY HAVE A SMALL NUMBER OF OUTLIERS.** ADDITIONALLY, COMPARED TO CLIENTS WITH ACADEMIC DEGREE, AND ALL FAMILY STATUSES HAVE A RELATIVELY LOW LEVEL OF OUTLIERS.
- **CLIENTS' INCOMES, REGARDLESS OF FAMILY BACKGROUND OR KIND OF EDUCATION, ARE 25% OR LESS BELOW THE FIRST QUARTILE.**
- **HIGHER EDUCATION, INCOMPLETE EDUCATION IN HIGHER EDUCATION, LOWER SECONDARY EDUCATION, AND SECONDARY/SECONDARY SPECIAL EDUCATION CLIENTS HAD A LARGER PERCENTAGE OF OUTLIERS.**
- **ACCORDING TO THE ABOVE GRAPH, CLIENTS WITH HIGHER EDUCATION TEND TO HAVE HIGHER INCOMES THAN CLIENTS WITHOUT HIGHER EDUCATION.**
- **DESPITE THE FACT THAT CERTAIN CONSUMERS TEND TO HAVE BETTER INCOMES AND HAVE NOT COMPLETED THEIR HIGHER EDUCATION. SOME OF THE CLIENTS WHO HAVE COMPLETED SECONDARY OR SECONDARY SPECIAL EDUCATION TYPICALLY MAKE MORE MONEY.**

- **CLIENTS WITHOUT AN ACADEMIC DEGREE BUT WITH OTHER TYPES OF EDUCATION HAVE A HIGH PROPORTION OF OUTLIERS.**
- **THE MAJORITY OF CLIENTS' CREDIT BALANCES, ARE BELOW 25%.**
- **CLIENTS WITH ACADEMIC DEGREES AND WIDOWS FREQUENTLY TAKE OUT LOANS WITH GREATER CREDIT LIMITS.**
- **HIGHER EDUCATION, INCOMPLETE HIGHER EDUCATION, LOWER SECONDARY EDUCATION, AND SECONDARY/SECONDARY SPECIAL EDUCATION ARE AMONG THE CLIENTELE GROUPS THAT ARE MORE PRONE TO TAKE OUT LARGE CREDIT LOANS.**

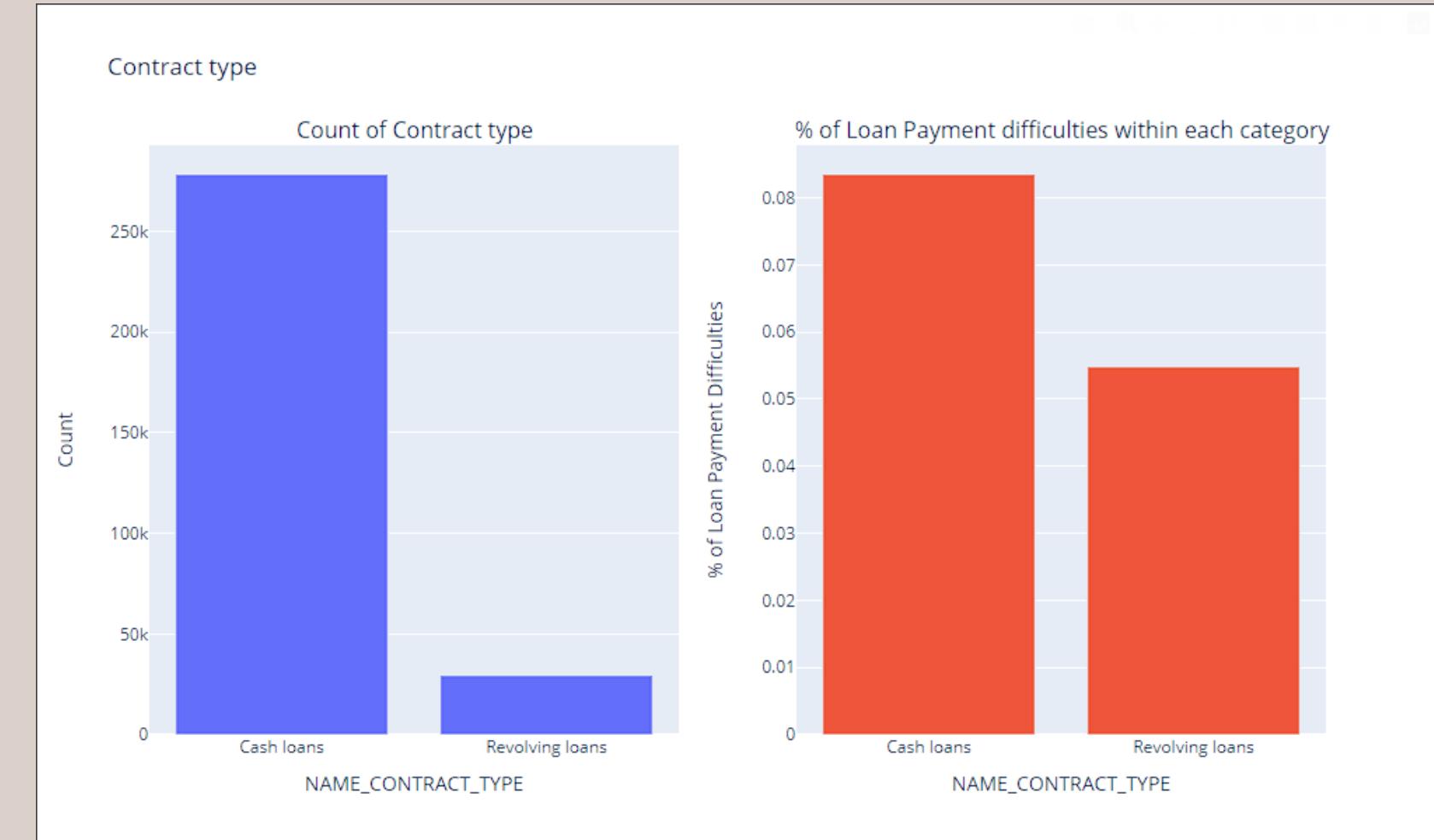
# TARGET 1



- MARRIED CLIENTS WITH ACADEMIC DEGREES EARN SIGNIFICANTLY LESS MONEY THAN OTHER CLIENTS.**
- DEFULTER CLIENTS MAKE A LOT LESS MONEY THAN NON-DEFULTERS DO.**

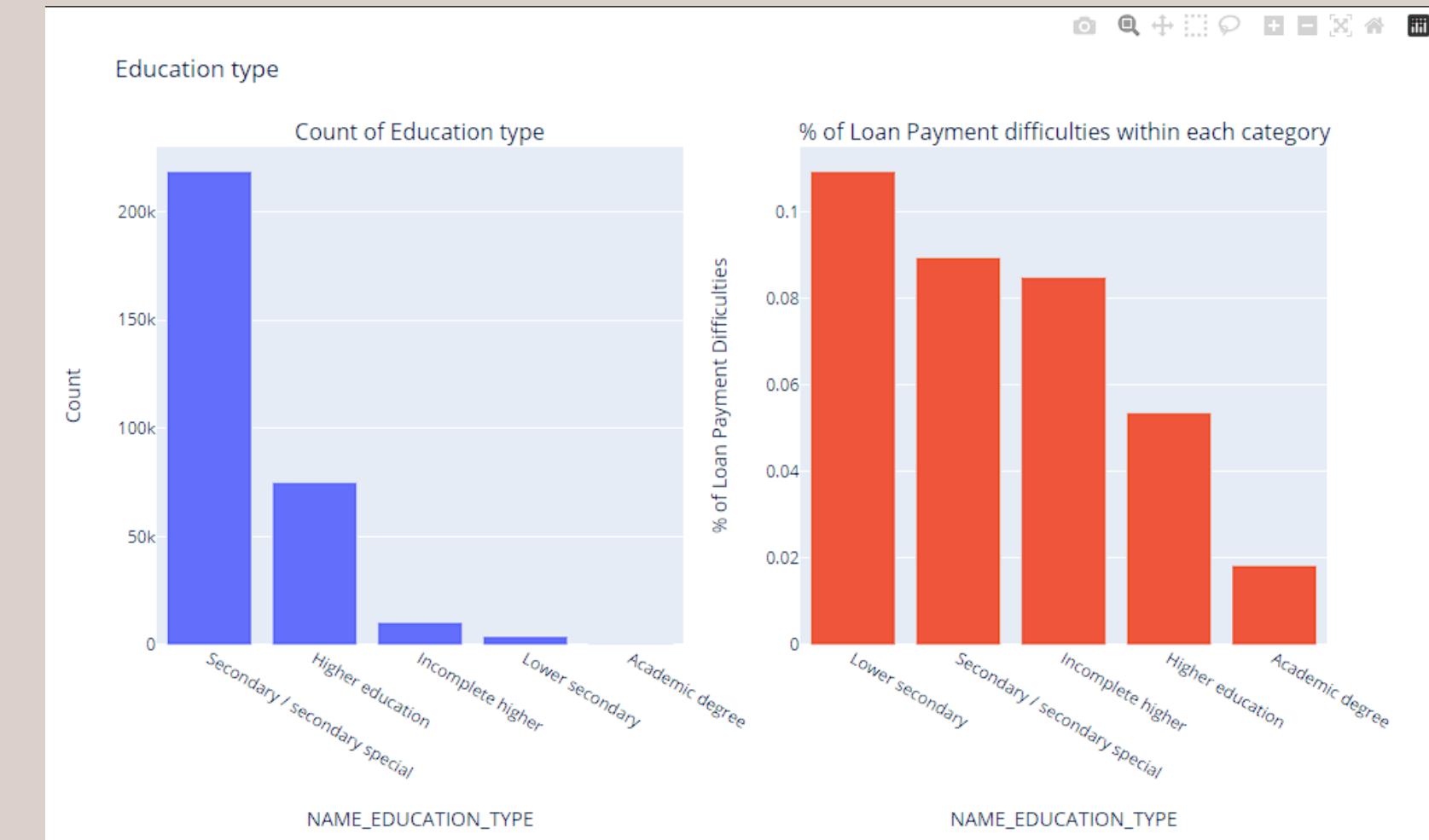
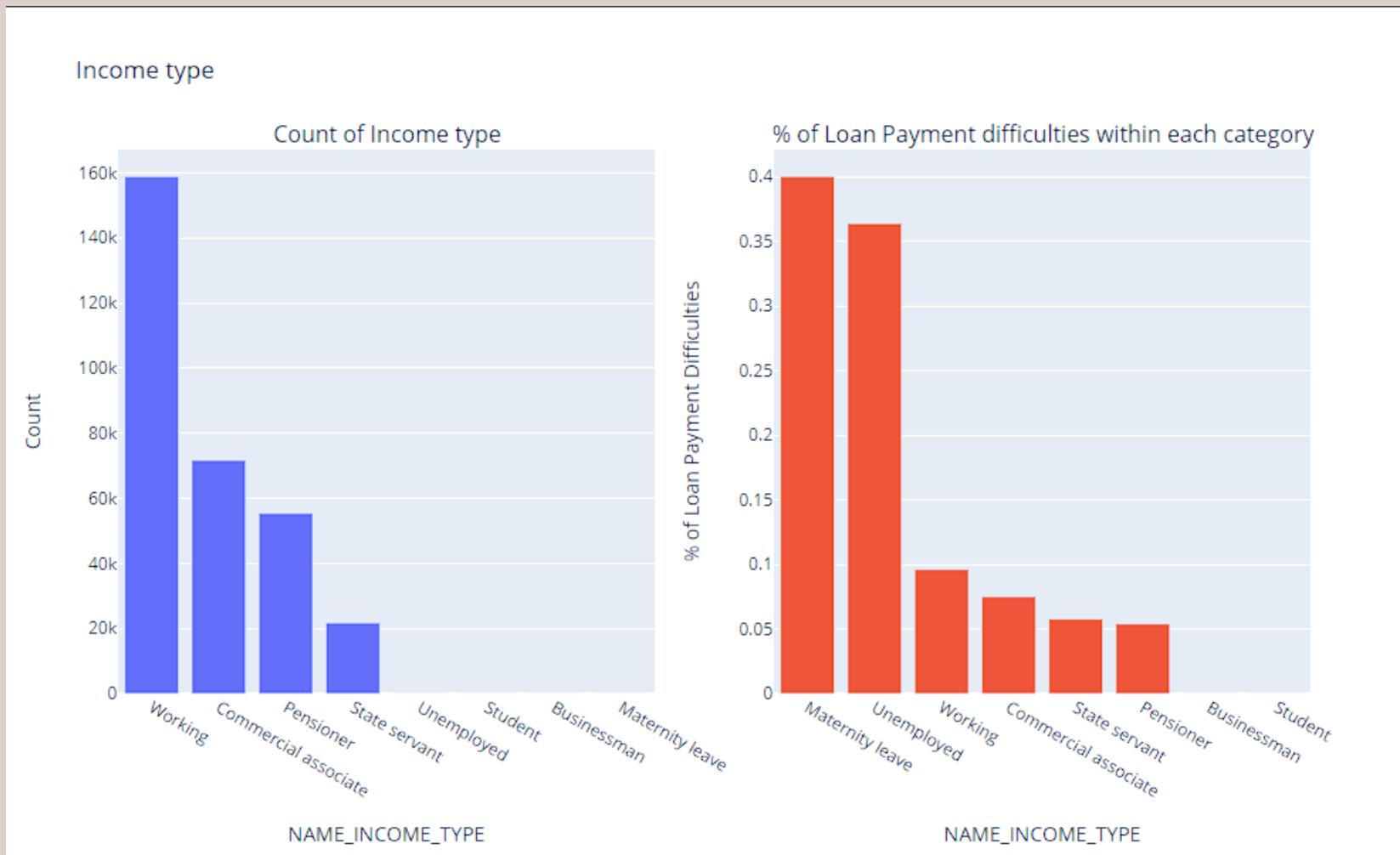
- MARRIED CLIENT WITH ACADEMIC SEEKING BETTER CREDIT LOAN SUBMITTED APPLICATION. LIKEWISE LACKS OUTLIERS. WITH NO OUTLIERS, THE BOXPLOT FOR SINGLE CLIENTS WITH ACADEMIC DEGREES IS RELATIVELY SMALL.**
- HIGHER EDUCATION, INCOMPLETE HIGHER EDUCATION, LOWER SECONDARY EDUCATION, AND SECONDARY/SECONDARY SPECIAL EDUCATION ARE AMONG THE CLIENTELE GROUPS THAT ARE MORE PRONE TO TAKE OUT LARGE CREDIT LOANS.**

# BIVARIATE ANALYSIS CATEGORICAL V/S CATEGORICAL



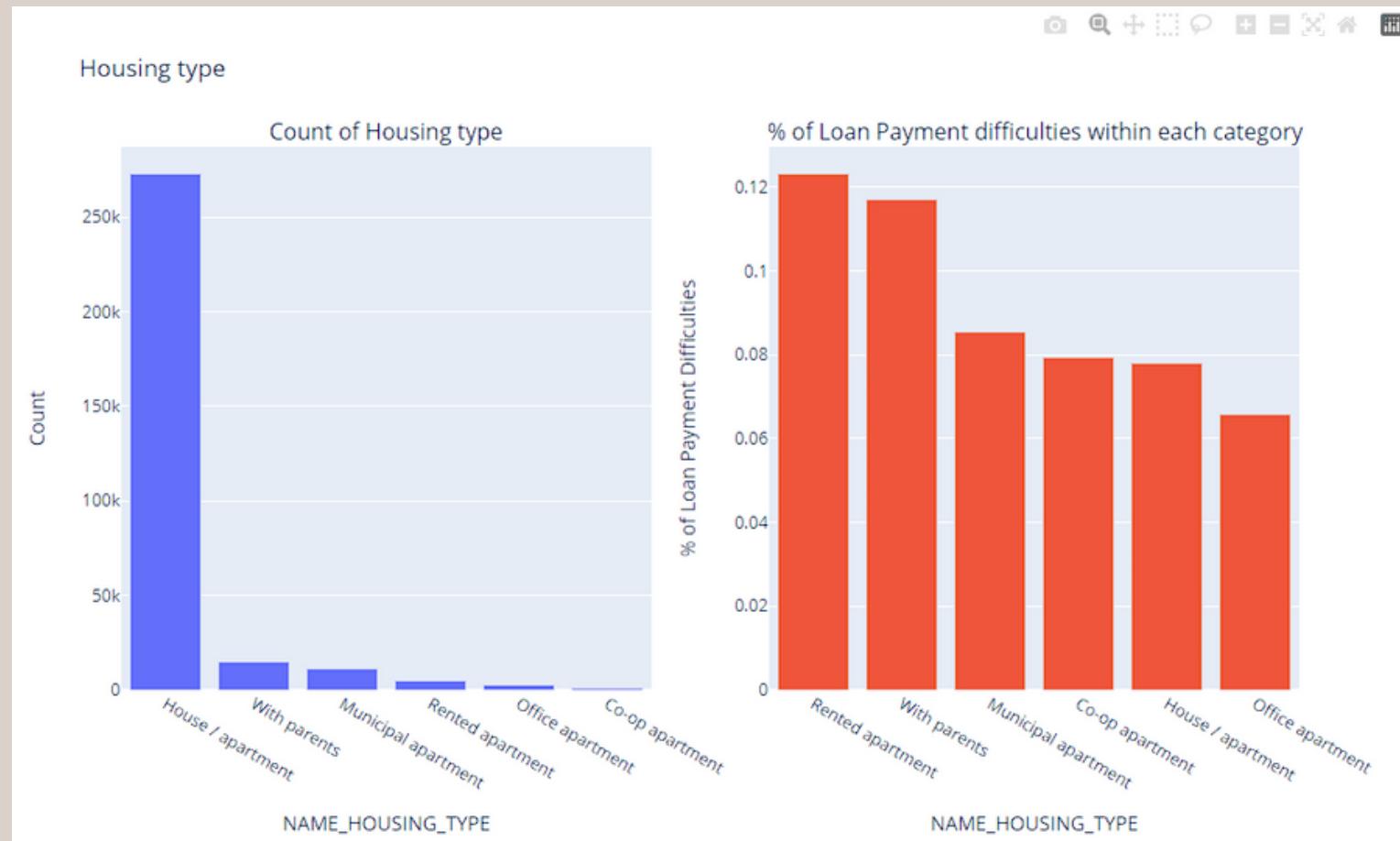
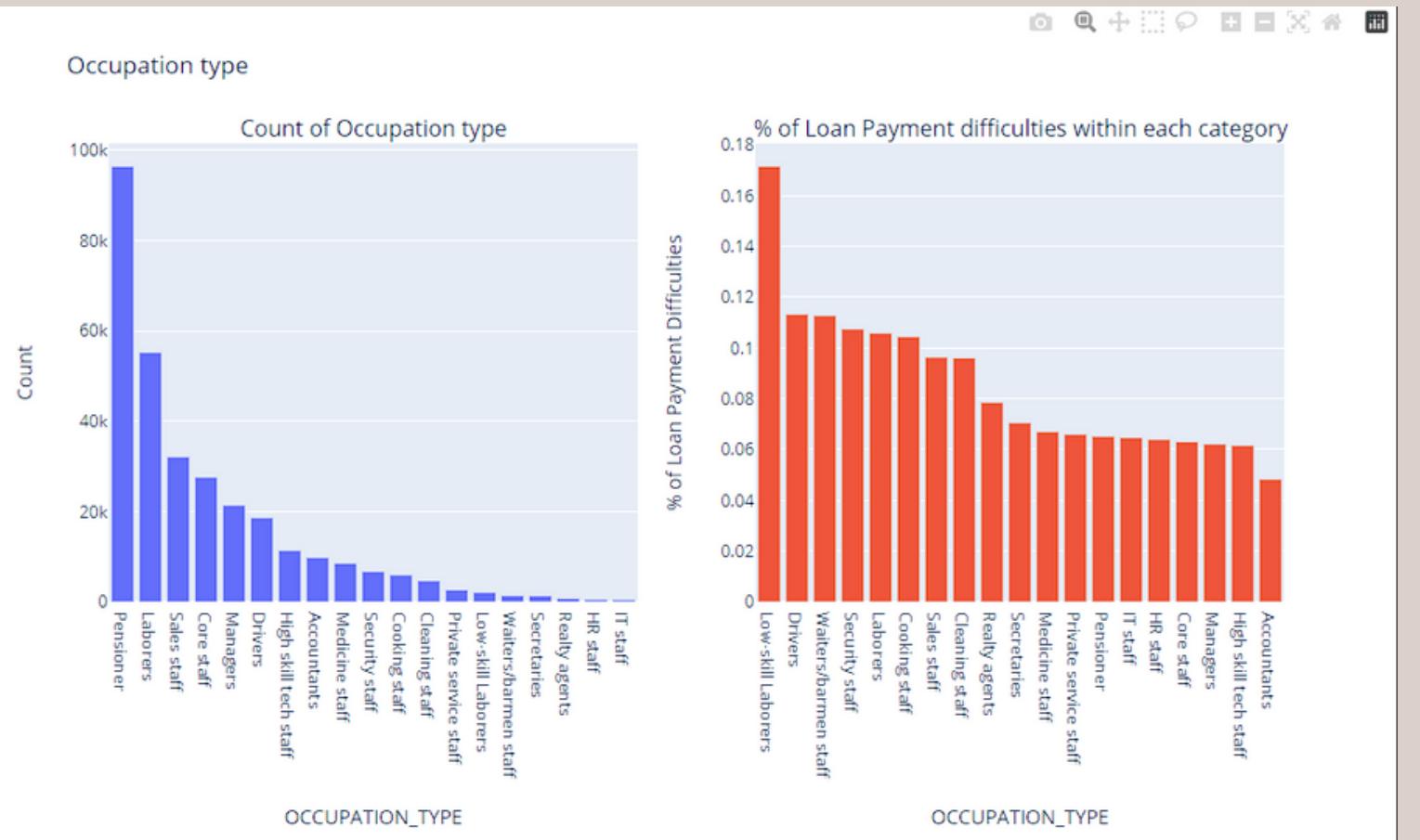
- FOLLOWING CLIENTS WITH MEDIUM AND EXTREMELY LOW INCOMES IN TERMS OF DEFAULT RISK ARE CLIENTS WITH LOW INCOMES.
- HIGH-PAYING CLIENTS ARE LEAST LIKELY TO GO INTO DEFAULT.

- CASH LOANS ARE MORE LIKELY TO DEFAULT THAN REVOLVING LOANS, WHICH HAVE A COMPARABLY LOWER RISK.



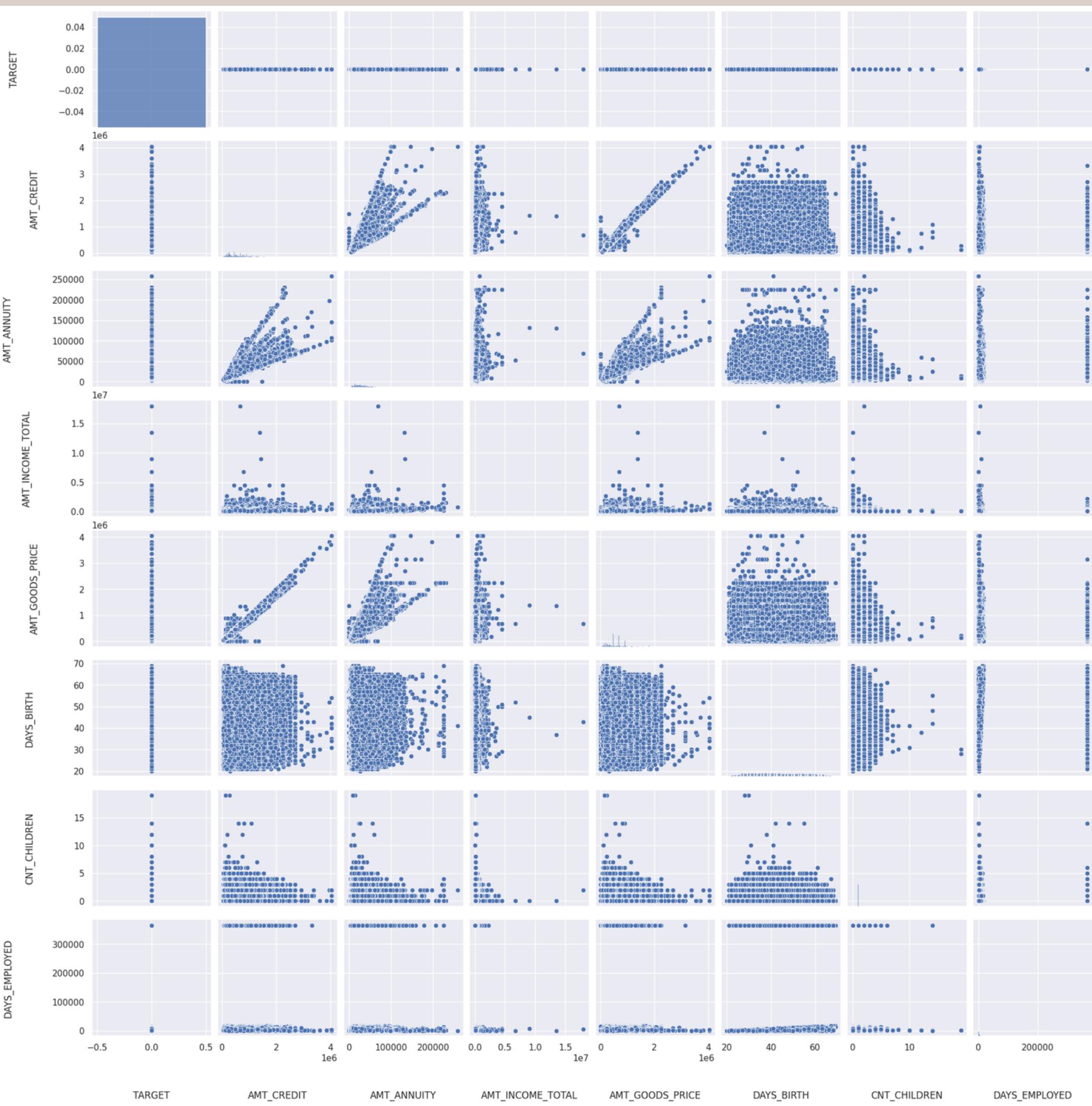
- **NUMBER OF CLIENTS BY INCOME CATEGORY EVEN THOUGH MATERNITY LEAVE IS JUST FIVE WEEKS LONG, THE PROBABILITY OF PAYMENT DEFAULT IS THE LOWEST OF ALL INCOME CATEGORIES.**
- **THE SAME SITUATION IS SEEN WITH UNEMPLOYED PEOPLE. DESPITE THE EXCEEDINGLY LOW COUNT, THERE IS LITTLE CHANCE OF PAYMENT DEFAULT.**
- **PENSIONERS, PUBLIC EMPLOYEES, AND BUSINESS PARTNERS ARE MORE LIKELY TO DEFAULT.**
- **ALTHOUGH THE NUMBER OF WORKING CLIENTS ASKING FOR LOANS IS QUITE HIGH, THE DANGER OF PAYMENT DEFAULT IS LOWER THAN IT IS FOR OTHER CLIENTS.**

- **HIGHER EDUCATED AND ACADEMICALLY QUALIFIED CLIENTS ARE LESS LIKELY TO DEFAULT.**
- **LOWER SECONDARY AND SECONDARY SPECIAL EDUCATION CLIENTS ARE AT A VERY HIGH RISK OF DEFAULT.**

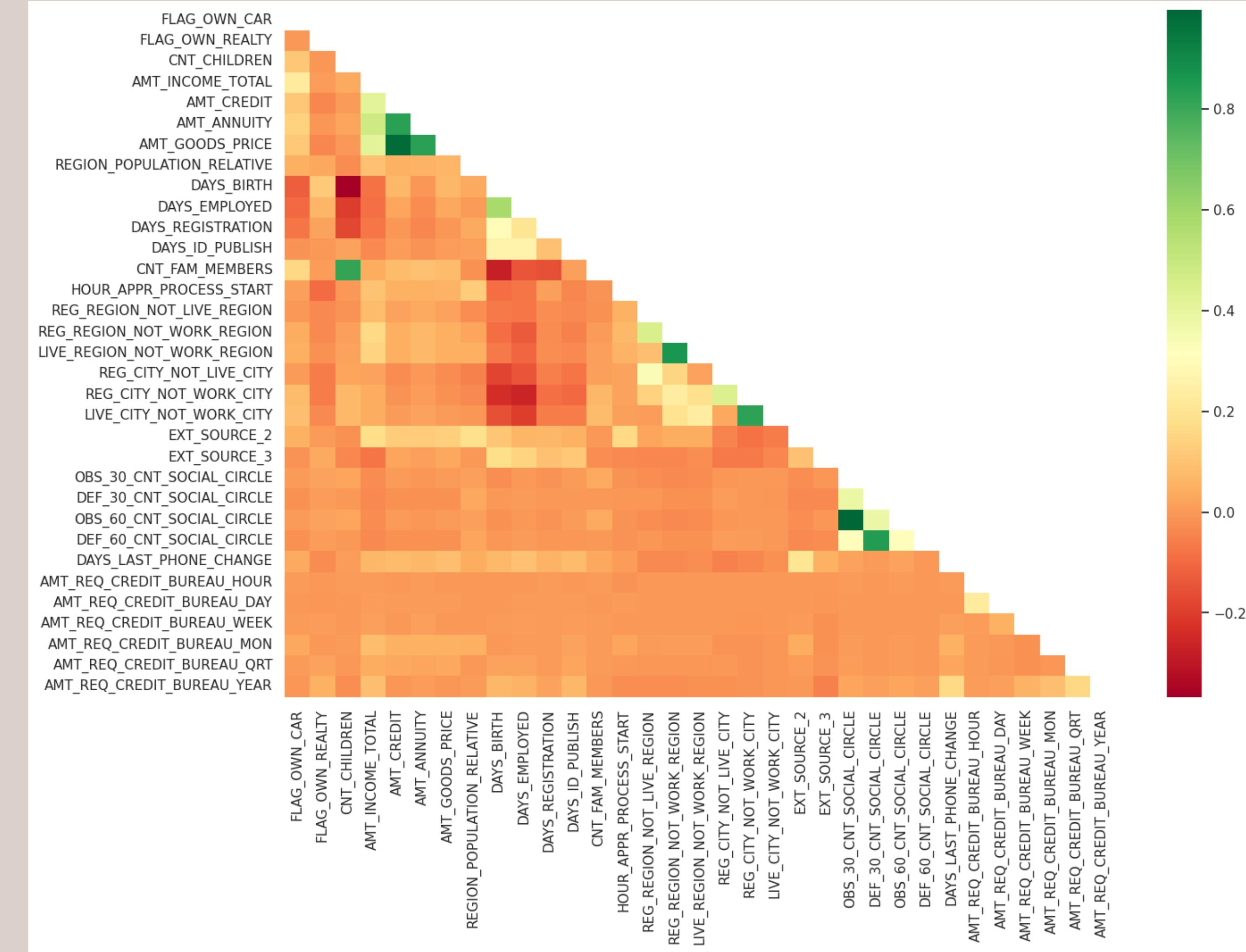


- **LOW-SKILLED WORKERS ARE MORE LIKELY TO DEFAULT.**
  - **MANAGERS, HIGHLY SKILLED TECHNICIANS, AND ACCOUNTANTS ARE LESS LIKELY TO DEFAULT.**

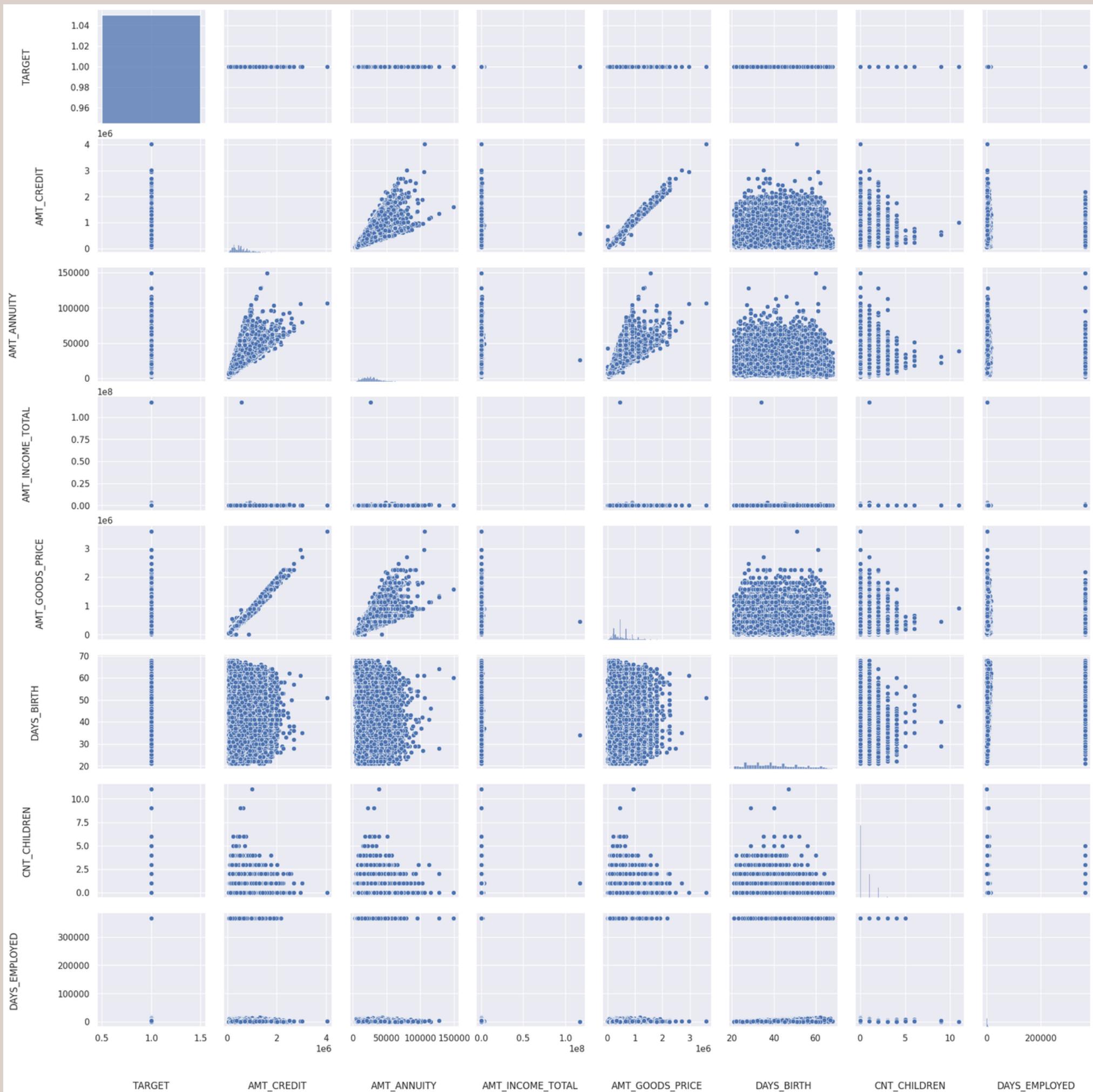
## CORRELATION FOR TARGET 0 WITH COLUMN 'AMT\_CREDIT','AMT\_ANNUITY','AMT\_INCOME\_T OTAL','AMT\_GOODS\_PRICE','DAYS\_BIRTH','CNT \_CHILDREN','DAYS\_EMPLOYED'



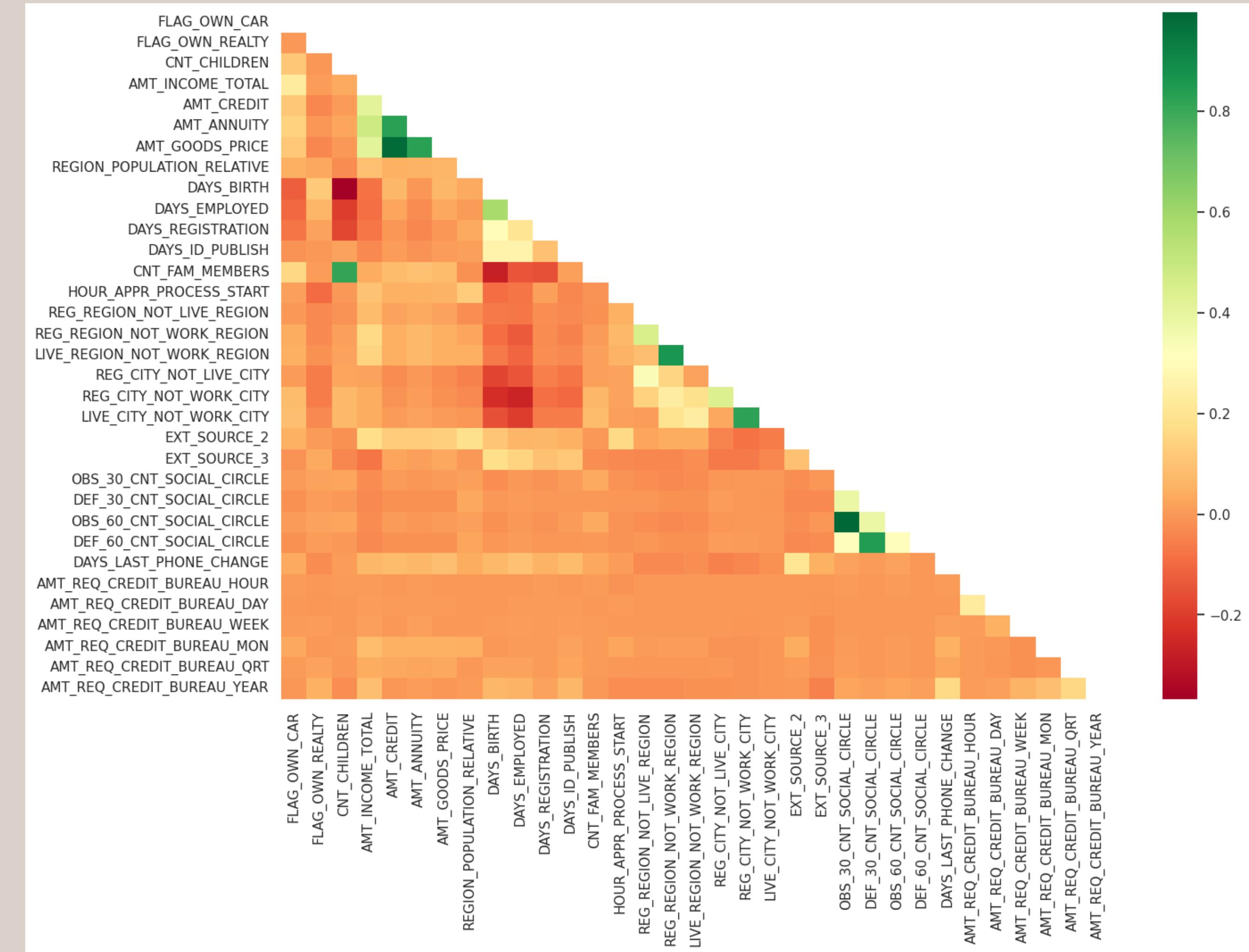
# T A R I G T



# CORRELATION FOR TARGET 1 WITH COLUMN 'AMT\_CREDIT','AMT\_ANNUITY','AMT\_INCOME\_TOT AL','AMT\_GOODS\_PRICE','DAYS\_BIRTH','CNT\_CHIL DREN','DAYS\_EMPLOYED'



# T A R G T



# CONCLUSION FOR CORRELATION

## FROM PAIR PLOT

- BOTH DEFAULTERS AND NON-DEFAULTERS HAVE A STRONG CORRELATION BETWEEN THE VARIABLES AMT\_CREDIT AND AMT\_GOODS\_PRICE. CONSEQUENTLY, IF THE COST OF A PROPERTY RISES, SO DOES THE LOAN BALANCE.
- BOTH DEFAULTERS AND NON-DEFAULTERS HAVE SUBSTANTIAL CORRELATIONS BETWEEN THE VARIABLES AMT\_CREDIT AND AMT\_ANNUITY (EMI). THEREFORE, IT MAKES SENSE THAT IF THE PRICE OF A PROPERTY RISES, THE EMI WOULD AS WELL.
- DUE TO THE SIGNIFICANT CORRELATION BETWEEN DEFAULTERS AND NON-DEFAULTERS FOR ALL THREE VARIABLES (AMT\_CREDIT, AMT\_GOODS\_PRICE, AND AMT\_ANNUITY), AN EFFECTIVE SIGNAL FOR DEFaulTER IDENTIFICATION MAY NOT BE POSSIBLE.

## FROM HEAT MAP

- LESS CHILDREN ARE BEING BORN AT THE CLIENT'S PERMANENT ADDRESS, WHICH DIFFERS FROM THE CONTACT ADDRESS.
- LESS CHILDREN ARE BEING BORN, AND THE CLIENT'S PERMANENT ADDRESS DIFFERS FROM THEIR WORKPLACE ADDRESS.

# TOP 10 CORRELATION

**FOR TARGET 0**

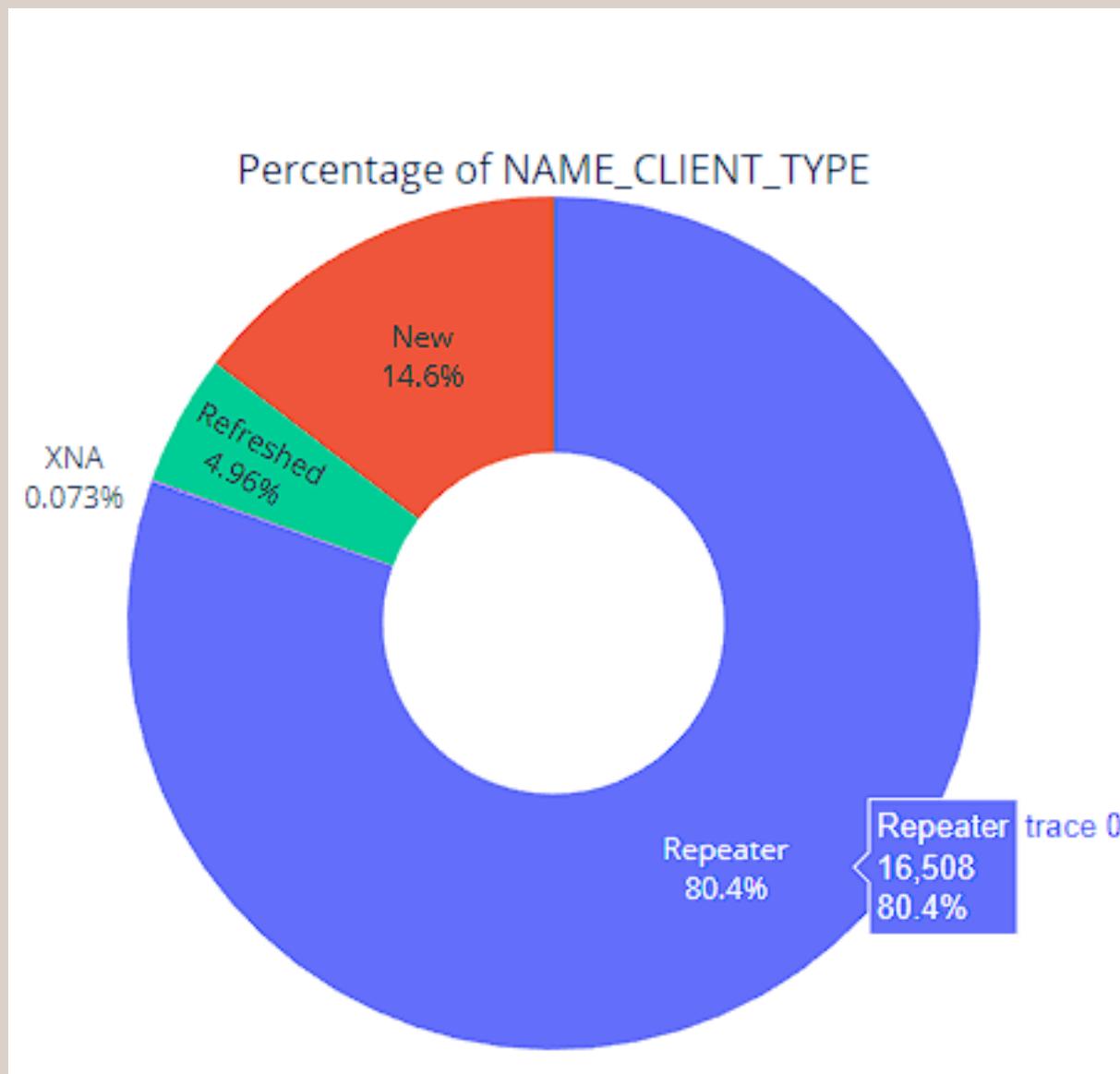
	VAR1	VAR2	CORRELATION	CORR_ABS
934	OBS_60_CNT_SOCIAL_CIRCLE	OBS_30_CNT_SOCIAL_CIRCLE	0.998510	0.998510
286	AMT_GOODS_PRICE	AMT_CREDIT	0.987250	0.987250
494	CNT_FAM_MEMBERS	CNT_CHILDREN	0.878571	0.878571
647	LIVE_REGION_NOT_WORK_REGION	REG_REGION_NOT_WORK_REGION	0.861861	0.861861
970	DEF_60_CNT_SOCIAL_CIRCLE	DEF_30_CNT_SOCIAL_CIRCLE	0.859371	0.859371
755	LIVE_CITY_NOT_WORK_CITY	REG_CITY_NOT_WORK_CITY	0.830381	0.830381
287	AMT_GOODS_PRICE	AMT_ANNUITY	0.776686	0.776686
251	AMT_ANNUITY	AMT_CREDIT	0.771309	0.771309
395	DAYS_EMPLOYED	DAYS_BIRTH	0.626028	0.626028
611	REG_REGION_NOT_WORK_REGION	REG_REGION_NOT_LIVE_REGION	0.446101	0.446101

**FOR TARGET 1**

	VAR1	VAR2	CORRELATION	CORR_ABS
934	OBS_60_CNT_SOCIAL_CIRCLE	OBS_30_CNT_SOCIAL_CIRCLE	0.998270	0.998510
286	AMT_GOODS_PRICE	AMT_CREDIT	0.983103	0.987250
494	CNT_FAM_MEMBERS	CNT_CHILDREN	0.885484	0.878571
647	LIVE_REGION_NOT_WORK_REGION	REG_REGION_NOT_WORK_REGION	0.847885	0.861861
970	DEF_60_CNT_SOCIAL_CIRCLE	DEF_30_CNT_SOCIAL_CIRCLE	0.869016	0.859371
755	LIVE_CITY_NOT_WORK_CITY	REG_CITY_NOT_WORK_CITY	0.778540	0.830381
287	AMT_GOODS_PRICE	AMT_ANNUITY	0.752699	0.776686
251	AMT_ANNUITY	AMT_CREDIT	0.752195	0.771309
395	DAYS_EMPLOYED	DAYS_BIRTH	0.582441	0.626028
611	REG_REGION_NOT_WORK_REGION	REG_REGION_NOT_LIVE_REGION	0.497937	0.446101

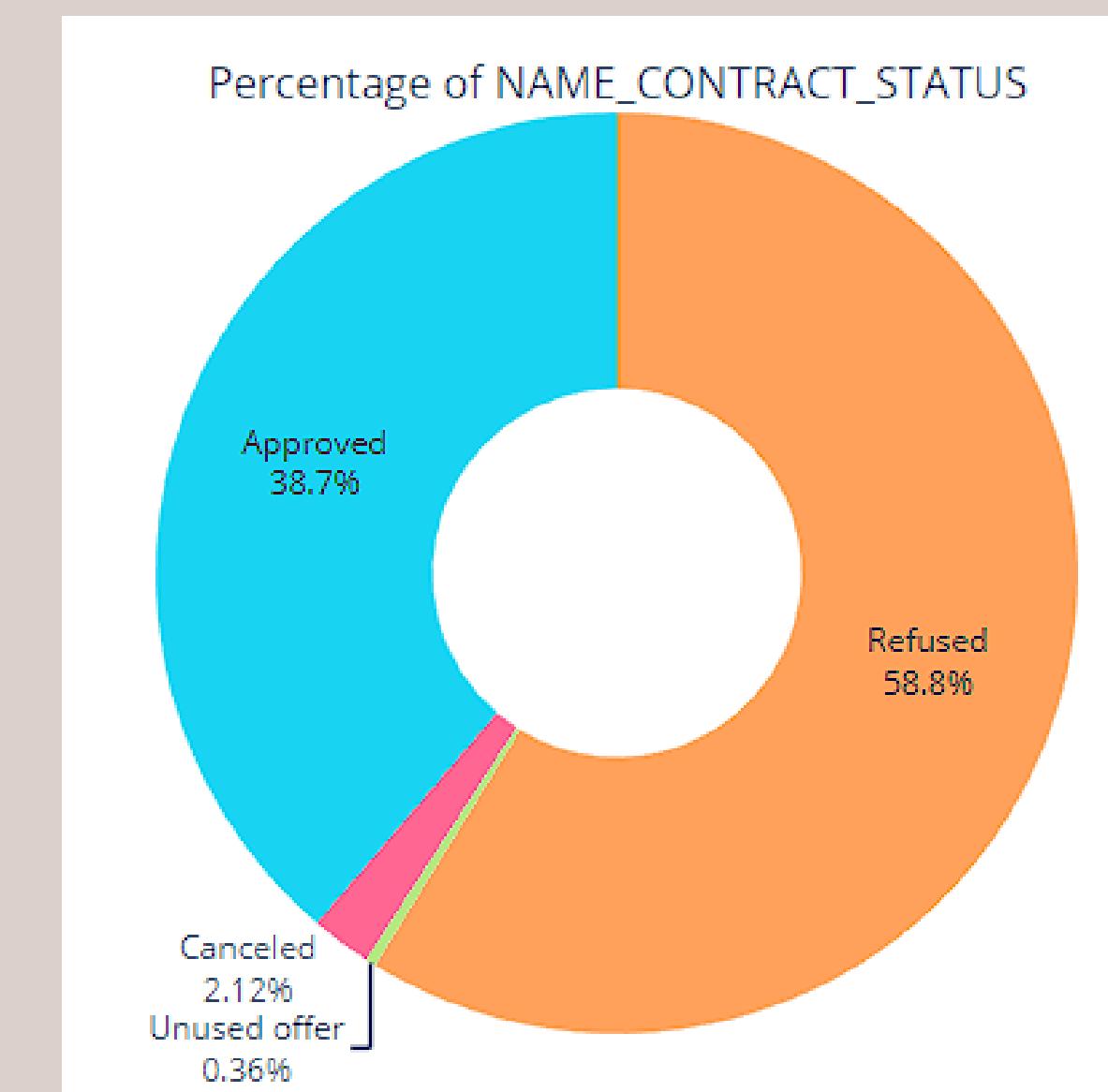
**TOP 10 CORRELATIONS IN THE POPULATIONS WITH DEFAULTS AND NON DEFULTER ARE ALMOST AT THE SAME LEVEL.**

# PERCENTAGE OF `NAME\_CONTRACT\_STATUS` AND `NAME\_CLIENT\_TYPE`

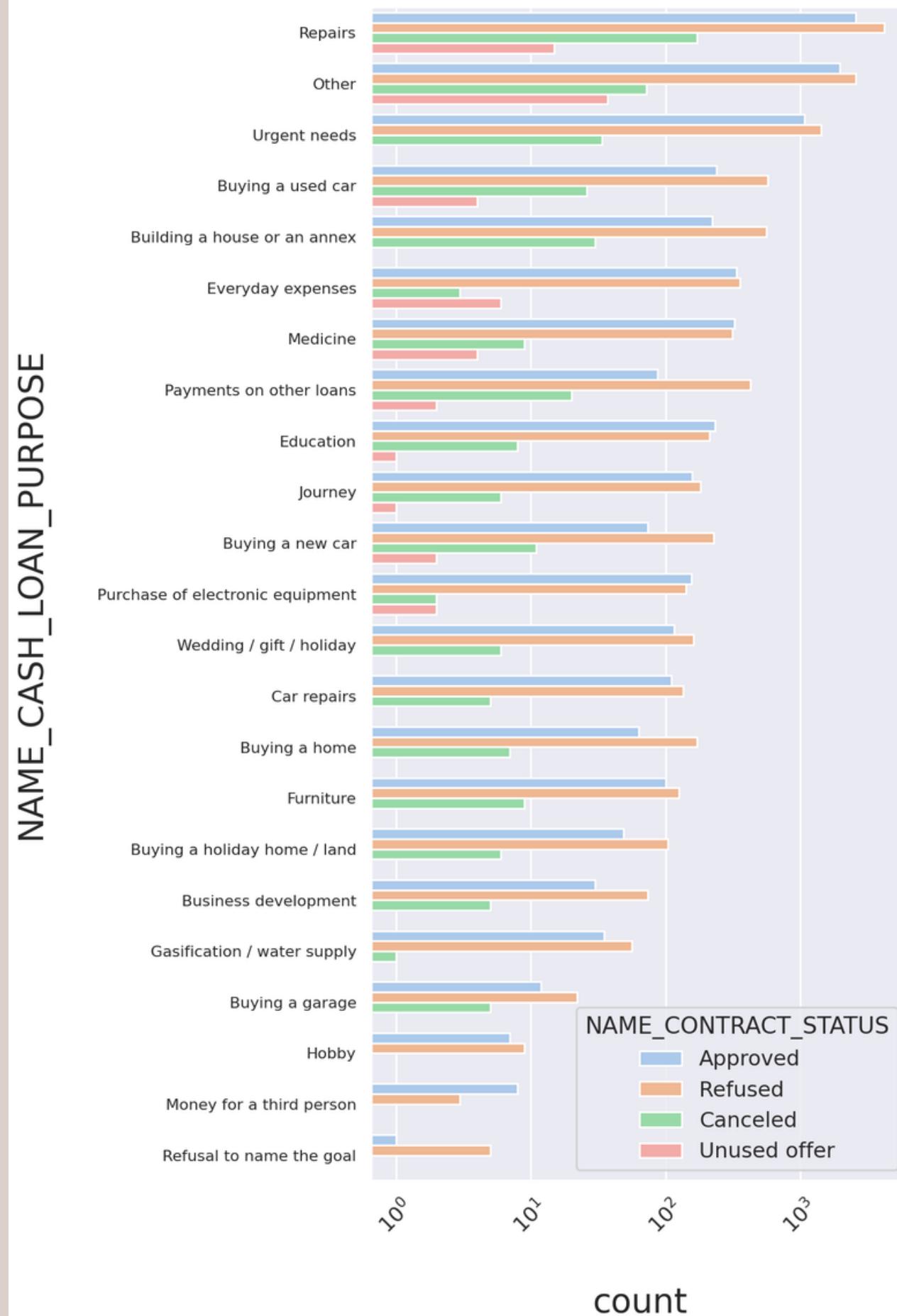


- ABOUT 80.7% OF THE CUSTOMERS WHO APPLIED FOR LOANS AGAIN.
- 14.5% OF THE LOAN APPLICANTS ARE NEW CUSTOMERS.

- APPROVED :- 38.8%
- REFUSED :- 58.5%
- CANCELED :- 2.3%
- UNUSED OFFER :- 0.31%



Distribution of contract status with loan purposes

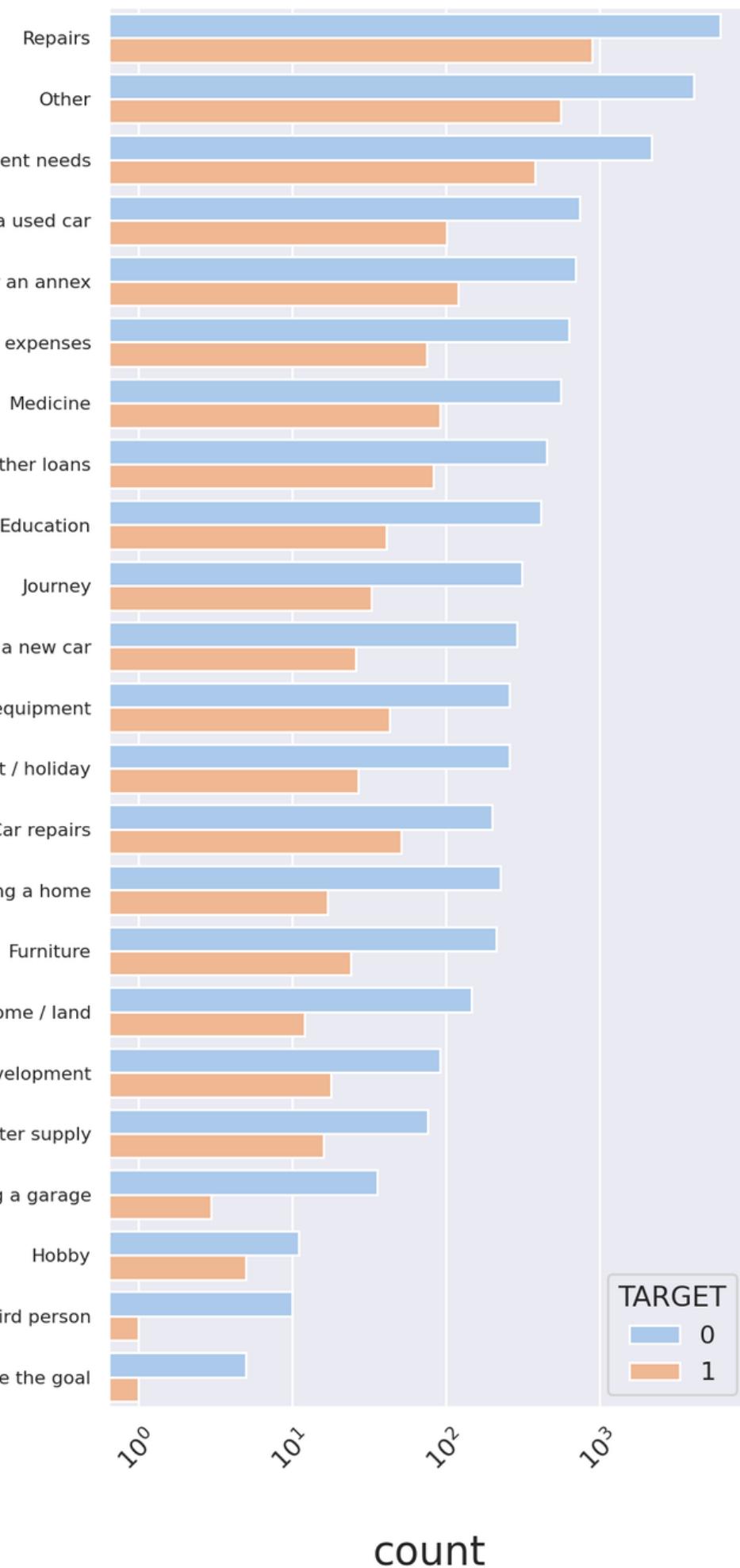


## CONTRACT\_STATUS V/S LOAN\_PURPOSE

- THE MAJORITY OF LOAN REJECTIONS WERE FOR "REPAIRS".
- WE HAVE AN EQUAL NUMBER OF APPROVALS AND REJECTIONS FOR "EDUCATION" AND "MEDICINE" OBJECTIVES.
- PAYING OFF OTHER DEBT AND PURCHASING A NEW AUTOMOBILE ARE FAR MORE LIKELY TO BE REJECTED THAN APPROVED.

Distribution of Target with Purposes of loan

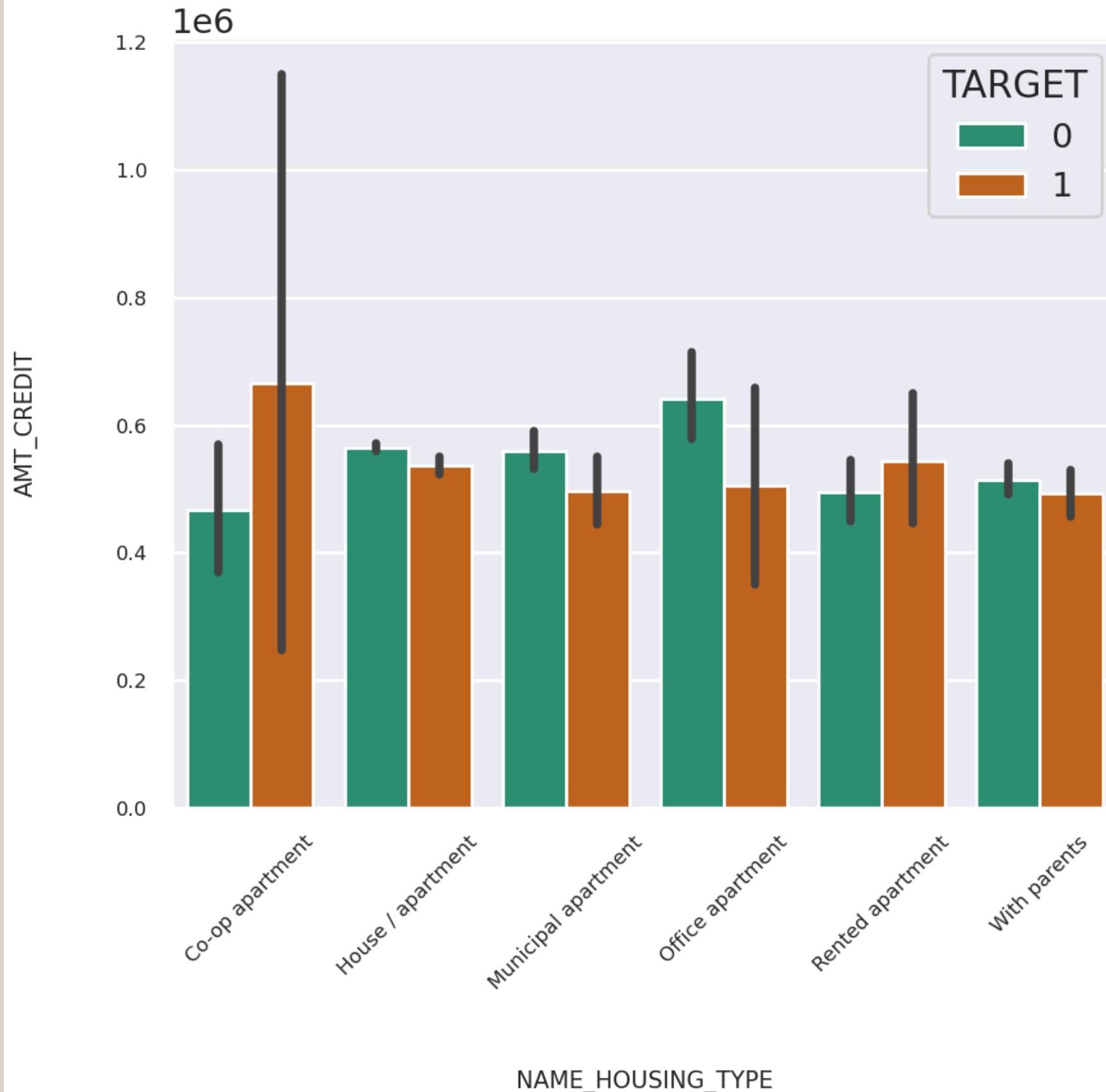
NAME\_CASH\_LOAN\_PURPOSE



## TARGET V/S LOAN\_PURPOSE

- FOR BOTH OF THE GOALS, "REPAIRS" ALSO EXHIBITS A CONSIDERABLE DEGREE OF VARIANCE.
- COMPARING "EDUCATION" & "MEDICINE" REASONS, A HIGHER PERCENTAGE OF PATIENTS ARE HAVING TROUBLE REPAYING THEIR LOANS IN THE MEDICAL FIELD THAN IN "EDUCATION."
- PURCHASING A SECONDHAND CAR" AND "BUILDING PURPOSE CUSTOMER EXPERIENCING FINANCIAL TROUBLES" HAD A SAME PERCENTAGE.

Distibution for columns AMT\_CREDIT & NAME\_HOUSING\_TYPE



## AMT\_CREDIT VS NAME\_HOUSING\_TYPE

- HERE, THE OFFICE APARTMENT HAS A HIGHER GOAL CREDIT OF 0 WHILE THE CO-OP APARTMENT HAS A HIGHER TARGET CREDIT OF 1. WE MAY THUS DRAW THE CONCLUSION THAT A BANK SHOULD REFRAIN FROM LENDING TO THE CO-OP FLAT HOUSING TYPE BECAUSE OF THEIR PAYMENT ISSUES. FOR SUCCESSFUL PAYMENTS, THE BANK CAN CONCENTRATE MOSTLY ON DWELLING TYPES SUCH AS THOSE WITH PARENTS, HOUSES, APARTMENTS, OR MUNICIPAL APARTMENTS.



# CONCLUSION

- WHEN THE PERCENTAGES OF PAYMENT DIFFICULTIES AND NON-PAYMENT DIFFICULTIES ARE COMPARED, WE SEE A DROP IN THE NUMBER OF RETIREES AND AN INCREASE IN THE PERCENTAGE OF WORKING PAYMENT DIFFICULTIES.
- WHEN COMPARED TO THE PROPORTIONS OF BOTH LOAN PAYMENT DIFFICULTIES AND LOAN NON-PAYMENT DIFFICULTIES, WE SEE A DROP IN THE PROPORTION OF MARRIED AND WIDOWED WITH LOAN PAYMENT DIFFICULTIES AND AN INCREASE IN THE PERCENTAGE OF SINGLE AND CIVIL MARRIED WITH LOAN PAYMENT DIFFICULTIES.
- WHEN COMPARING THE PERCENTAGES OF LOAN PAYMENT ISSUES AND LOAN NONPAYMENT DIFFICULTIES, WE SEE AN INCREASE IN THE PERCENTAGE OF LOAN PAYMENT DIFFICULTIES WITH SECONDARY/SECONDARY SPECIAL EDUCATION AND A REDUCTION IN THE PROPORTION OF LOAN PAYMENT CHALLENGES WITH HIGHER EDUCATION.
- BANKS SHOULD PAY GREATER ATTENTION TO CONTRACT TYPE. FOR SUCCESSFUL PAYMENTS, STUDENTS, PENSIONERS, AND BUSINESS OWNERS WITH HOUSING TYPES OTHER THAN CO-OP APARTMENTS AND OFFICE APARTMENTS SHOULD APPLY.

- BANKS NEED TO CONCENTRATE LESS ON EARNINGS TYPE WORKING SINCE THEY HAVE THE HIGHEST PROPORTION OF FAILED PAYMENTS.
- THE COUNT OF 'LOWER SECONDARY' IN 'NAME\_EDUCATION\_TYPE' IS RATHER LOW, AND IT ALSO HAS THE HIGHEST PERCENTAGE OF PAYMENT DIFFICULTIES- ABOUT 11%. AS A RESULT, CLIENTS WITH A 'LOWER SECONDARY' EDUCATION LEVEL ARE THE DRIVING CAUSES FOR LOAN DEFAULTERS.
- THE NUMBER OF 'LOW SKILLED LABOURERS' IN 'OCCUPATION\_TYPE' IS COMPARABLY SMALL, AND IT ALSO HAS THE HIGHEST PERCENTAGE OF PAYMENT DIFFICULTIES- ABOUT 17%. AS A RESULT, CLIENTS CLASSIFIED AS 'LOW SKILLED LABOURERS' ARE THE PRIMARY DRIVERS OF LOAN DEFAULTERS.
- GET AS MANY CLIENTS AS POSSIBLE BASED ON DWELLING TYPE. WITH PARENTS BECAUSE THEY HAVE THE FEWEST FAILED PAYMENTS
- FURTHERMORE, WITH LOAN PURPOSE REPAIR HAS A LARGER FREQUENCY OF FAILED PAYMENTS ON TIME.
- APPLICANTS RESIDING IN HOUSES/APARTMENTS HAVE THE MOST LOAN APPLICATIONS. WHILE WE CAN OBSERVE THAT RENTED APARTMENTS AND CANDIDATES LIVING WITH THEIR PARENTS HAVE A VERY HIGH DEFAULT RATE,

# CONTACT ME FOR ANY QUERY



+91 88505 75863



snaagam@gmail.com



# THANK YOU

