# CREDIT EDA ASSIGNMENT

KARTIK REDDIWAR

Problem Statement

- A Consumer Finance company which specializes in lending various types of loans to urban customers.

- When the company receives a loan application, the company has to decide for loan approval based on the applicant's profile.

- Two types of risks are associated with the bank's decision:

     A. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company

     B. If the applicant is not likely to repay the loan, then approving the loan may lead to a financial loss for the company.

- The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.
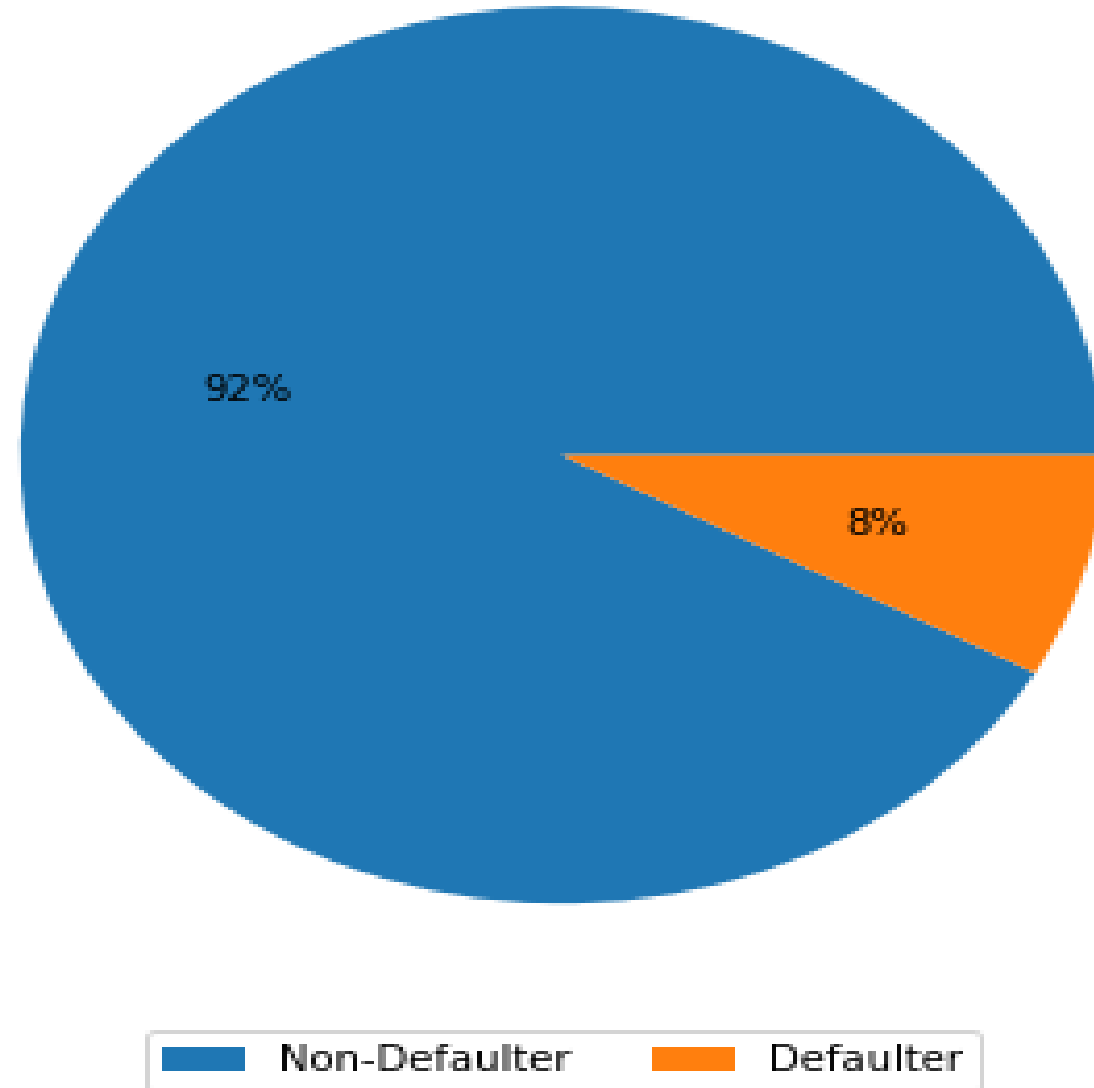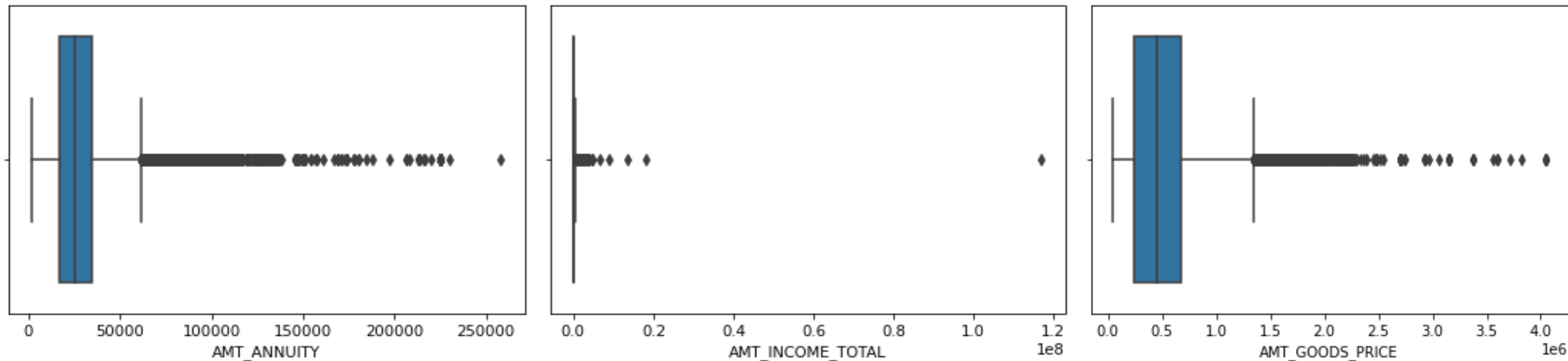
Analysis Approach.
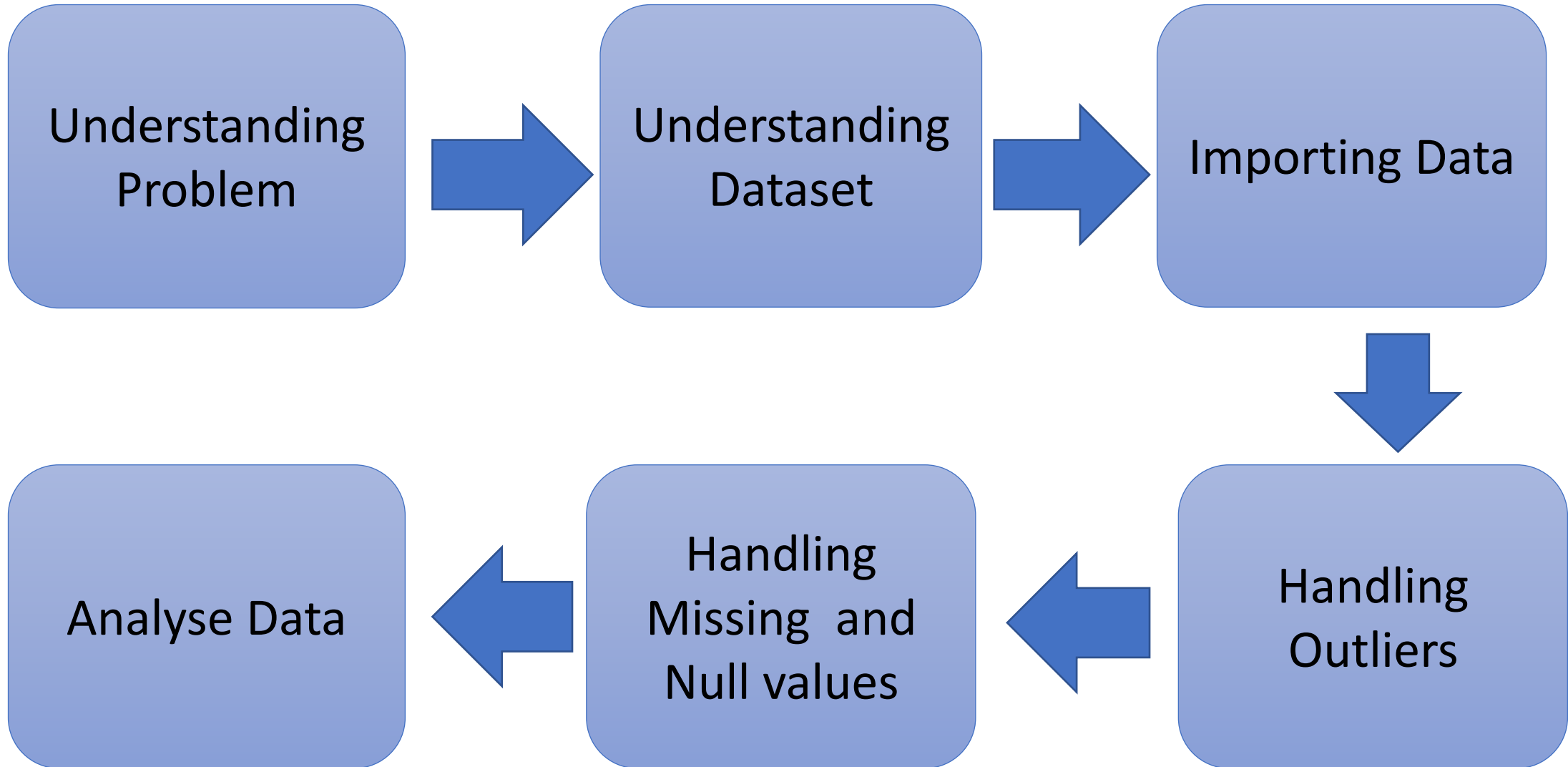
Data Imbalance

The Data-Imbalance of Target column

- The Percentage of Data Imbalance is

  Non-Defaulter= 92%

  Defaulter = 8%

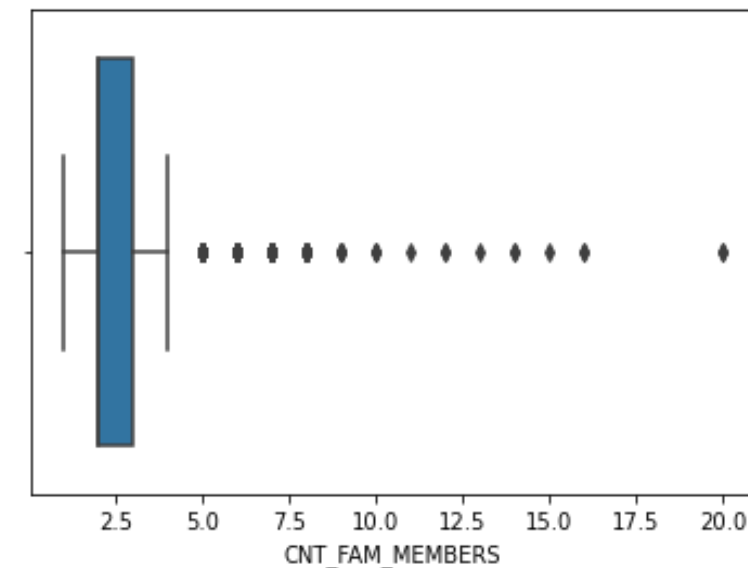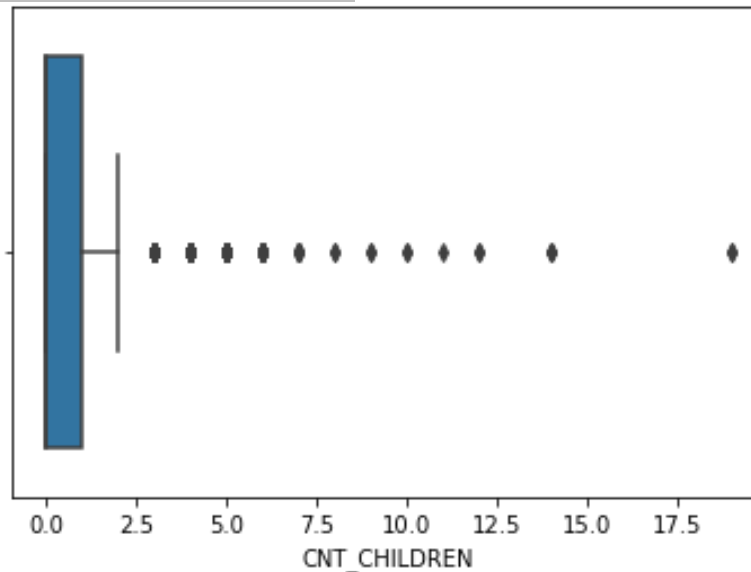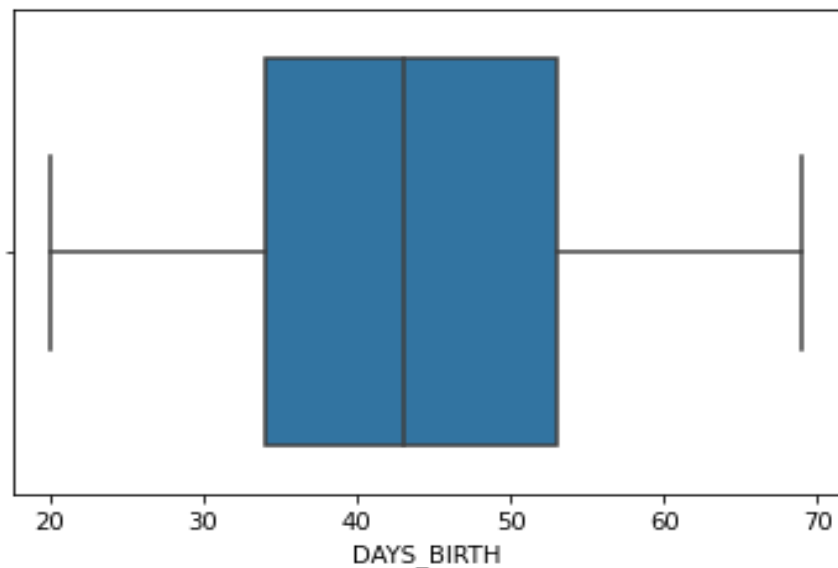- The Ratio of Data Imbalance is 11:39:1 (approx.)

Non-Defaulter 92%   Defaulter 8%

Outliers Analysis

❑ INSIGHTS.

- AMT_ANNUITY , AMT_INCOME_TOTAL, AMT_GOODS_PRICE have many values beyond IQR range hence they are outliers

Outliers Analysis
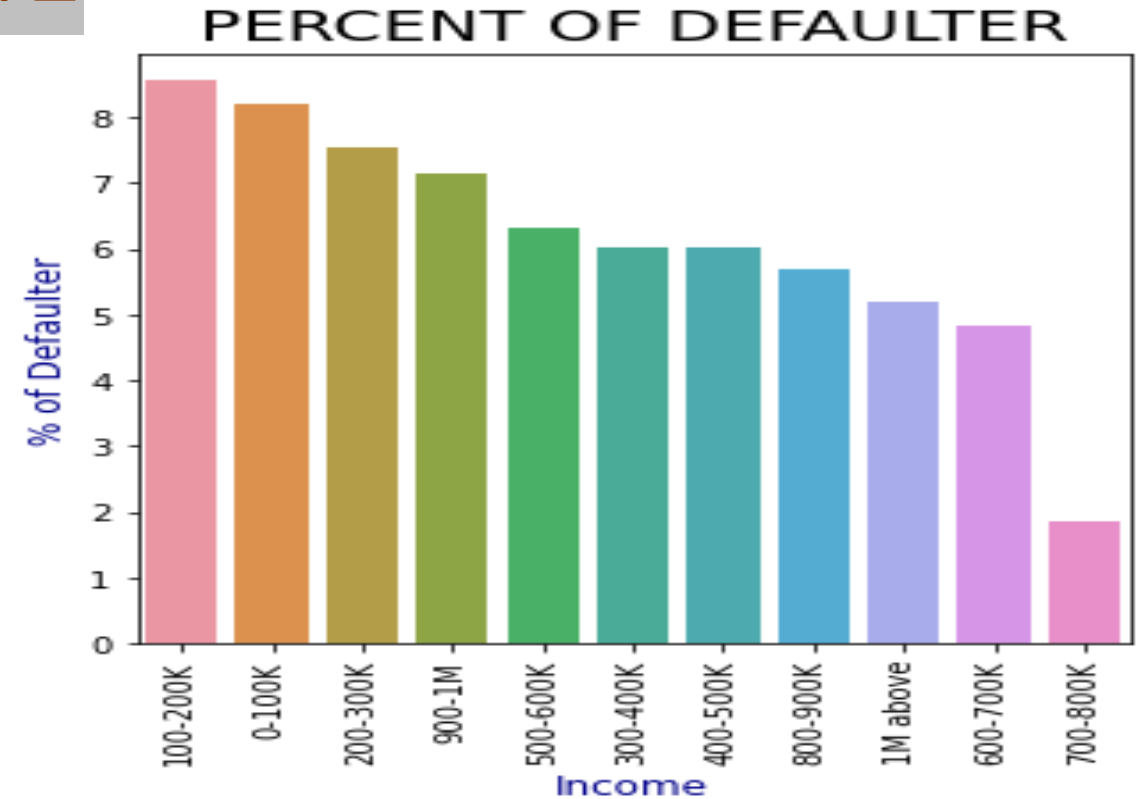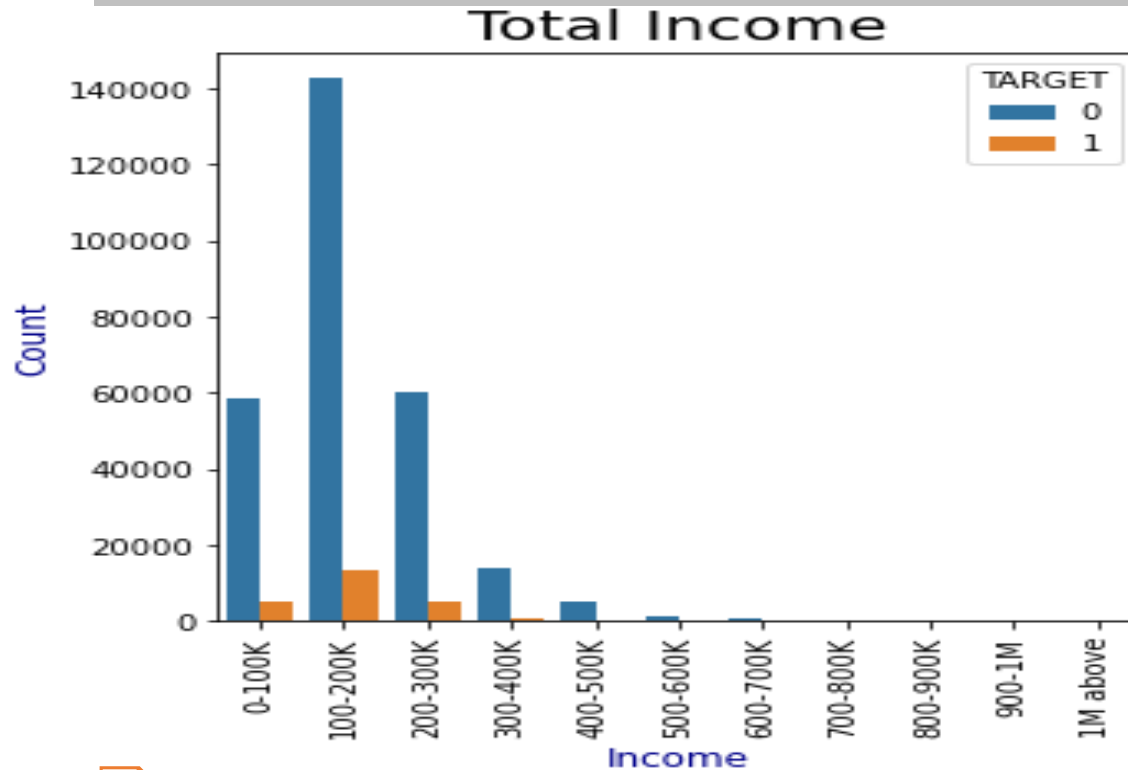

❑INSIGHTS.

- DAYS_BIRTH have no outlier present hence given data is genuine.
- CNT_CHIDREN , CNT_FAM_MEMBER have very less outliers .

UNIVARIATE ANALYSIS

AMOUNT OF INCOME

❑INSIGHTS.

- Most of the clients have salary in range of 100K-200K.

- Clients having salary range 700K-800K have very less default rate means they are likely to replay loan .

- Client having salary less than 300K have high default rate so approval of application should be done at higher interest rate.
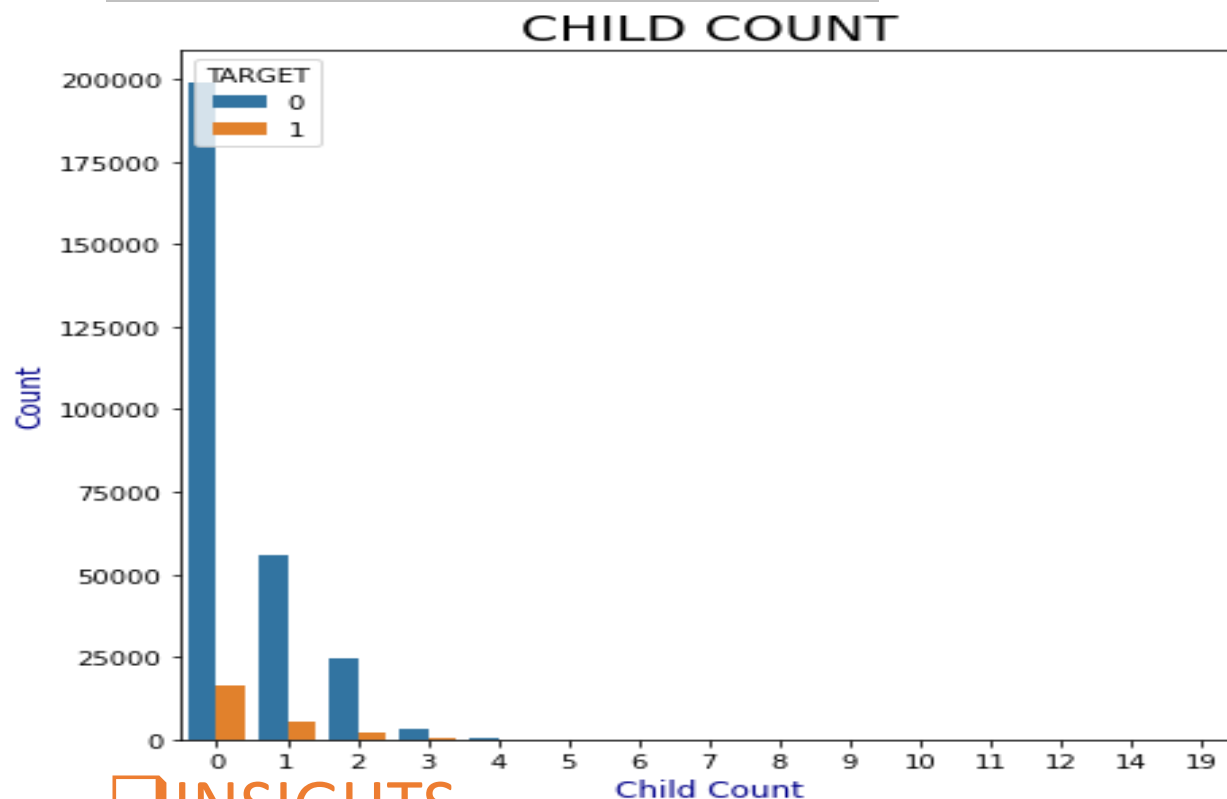
GENDER

CODE_GENDER


PERCENT OF DEFAULTER
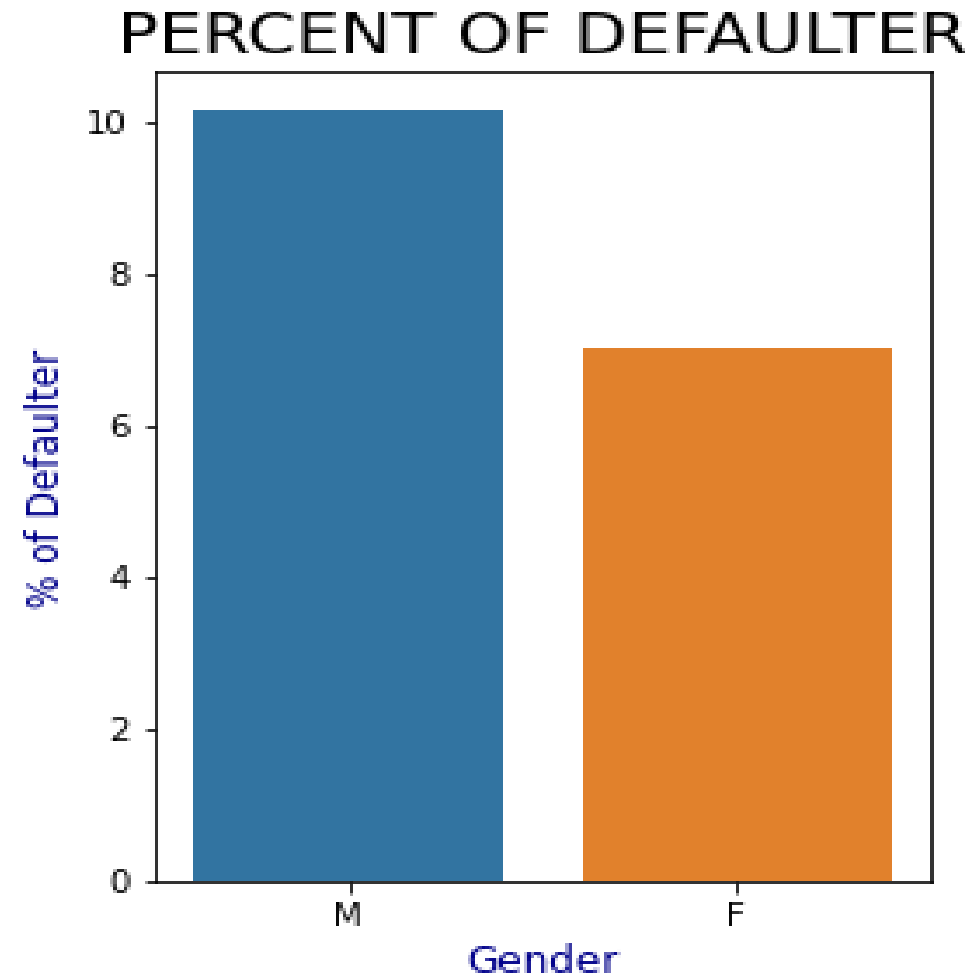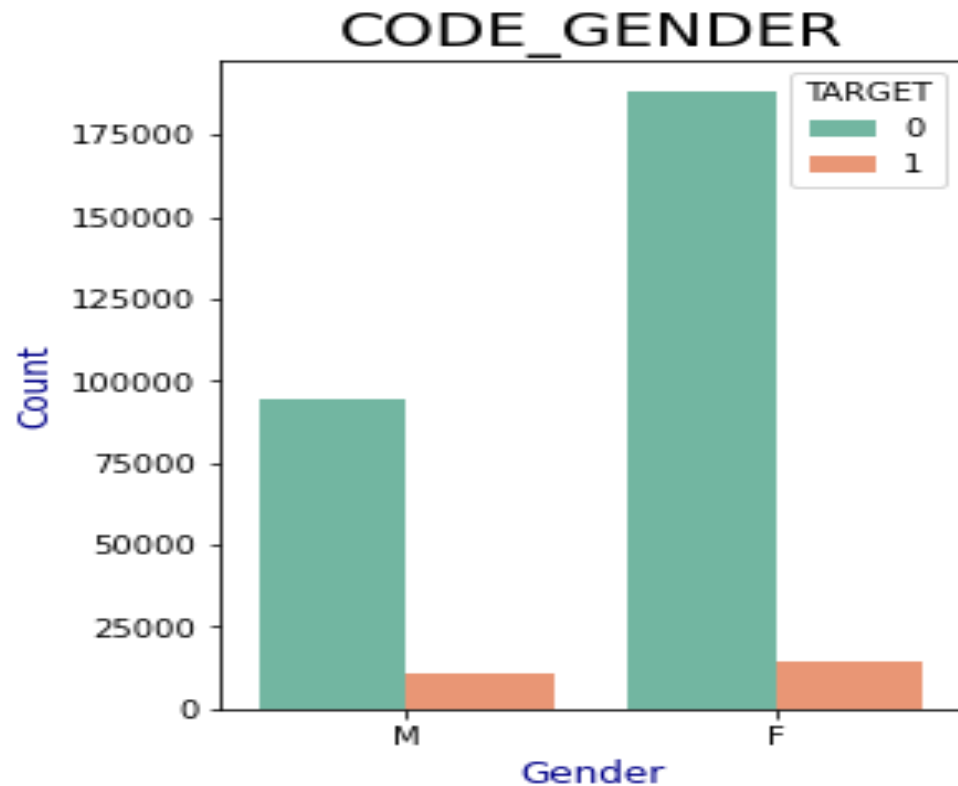
❑INSIGHTS.

- The counts of Female is almost double than Male hence female are more willing to take loans.
- Male Clients have high default rate so chances of they to replay is very less than female clients.
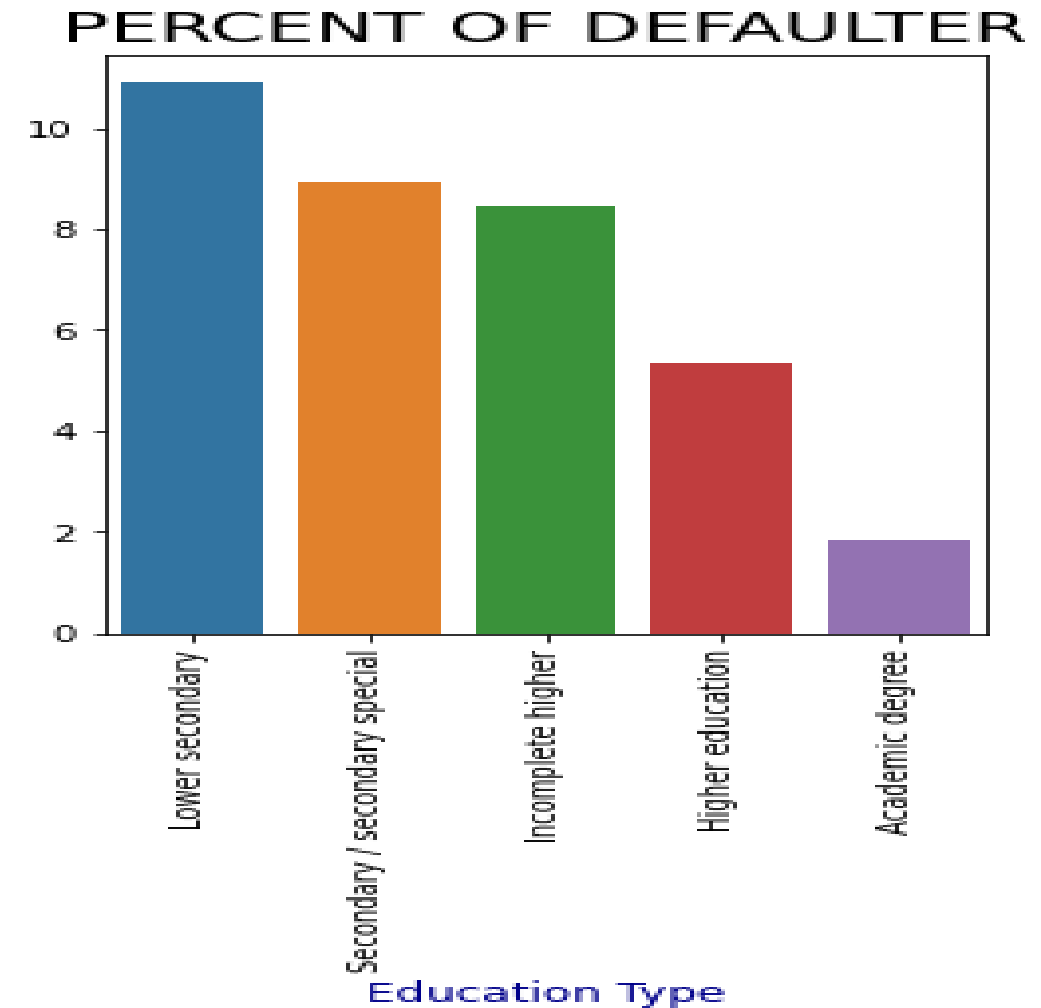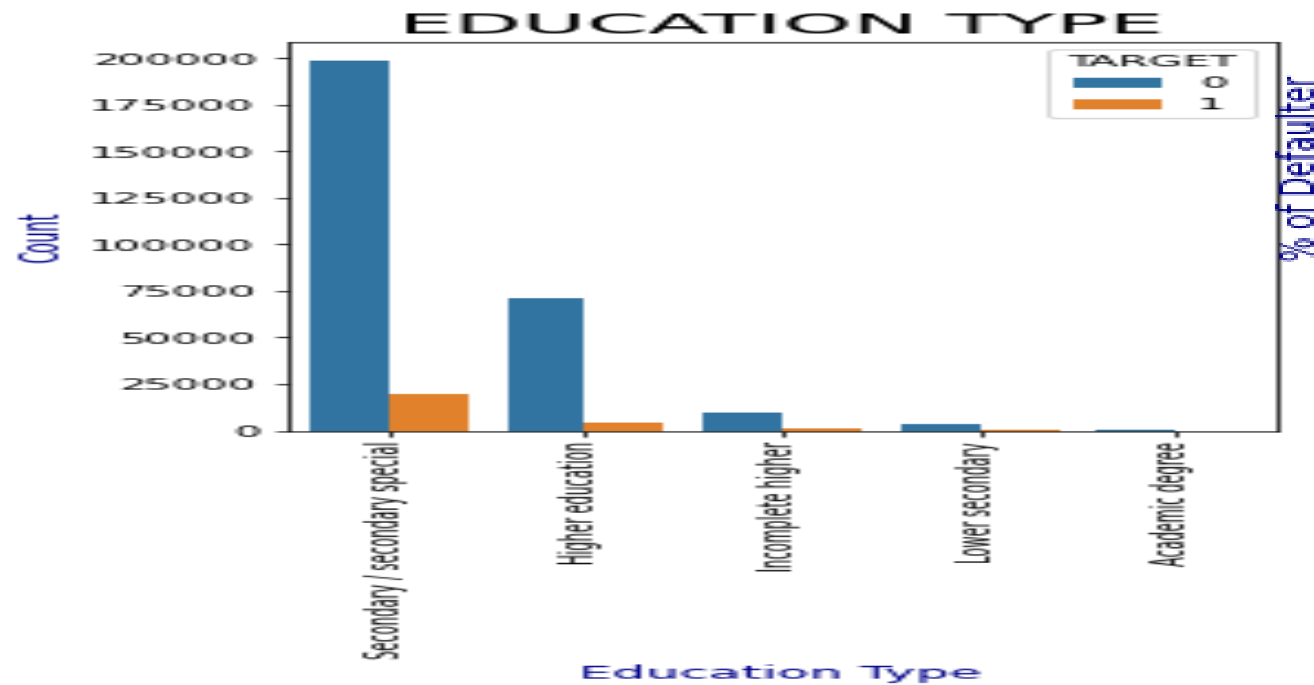
CHILDCOUNT

CHILD COUNT

PERCENT OF DEFAULTER

❑INSIGHTS.

- Most of the applicant have no child.

- The applicant who have child count about 11 and 9 have 100% default rate so the application of such applicant should be refused

- The applicant who have child count 1,2 or 0 would likely to repay loan.

EDUCATION TYPE

EDUCATION TYPE

PERCENT OF DEFAULTER
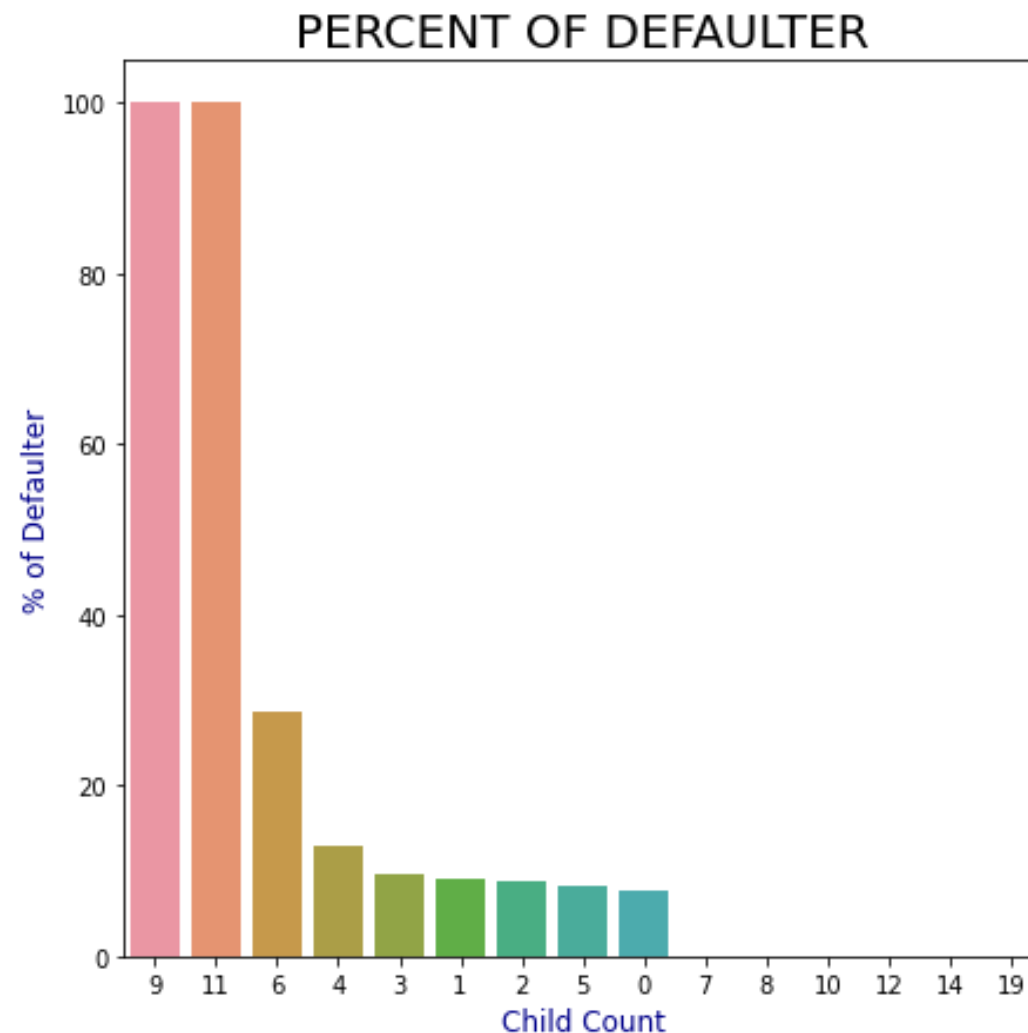
☐INSIGHTS.

- Many Applicant have Education qualification of secondary class.
- Applicant with lower secondary qualification have higher default rate
- Applicant with Academic degree have very less default rate and applicant count is very less.
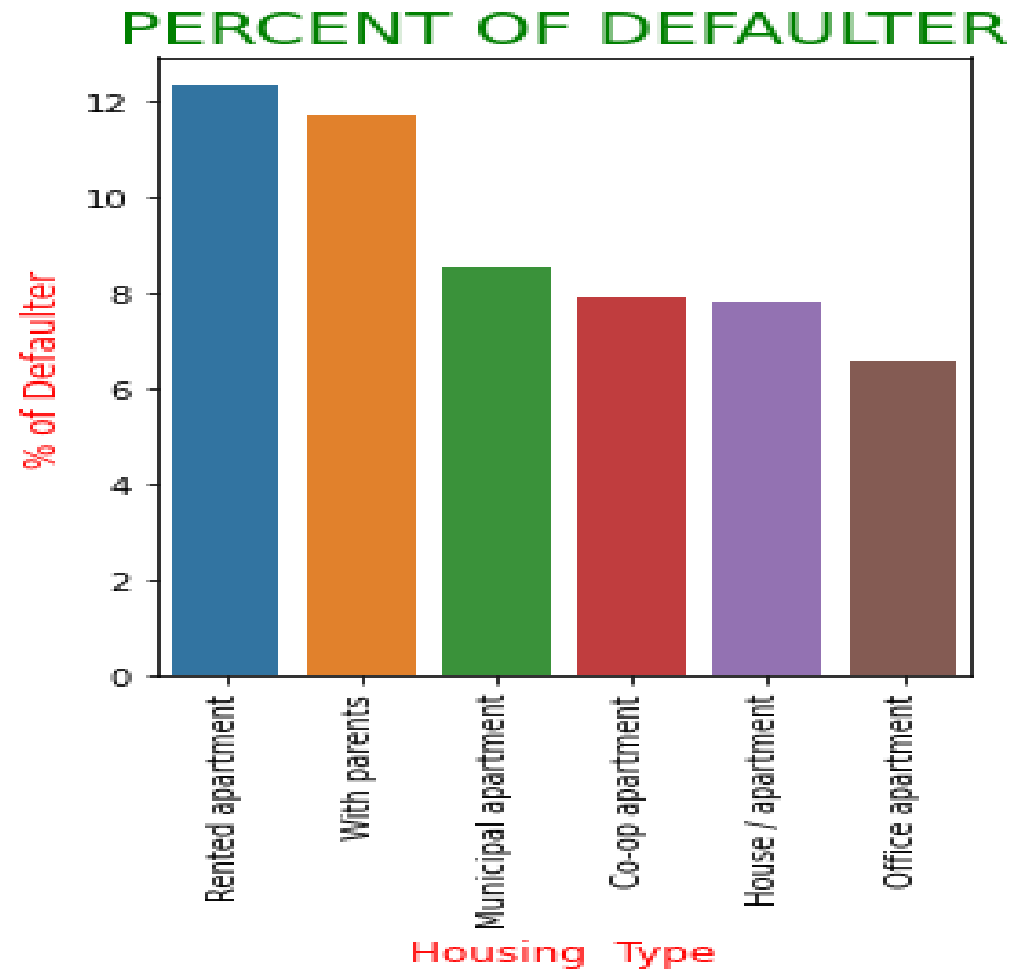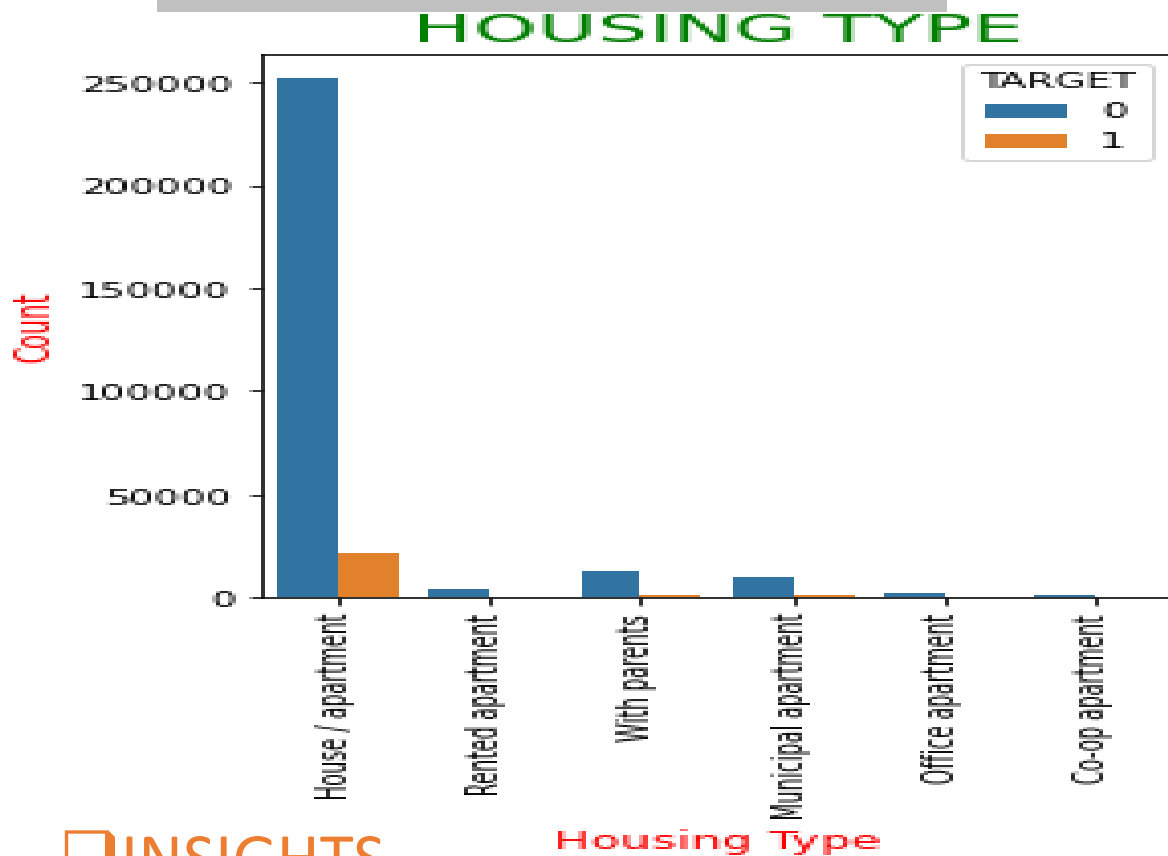
FAMILY STATUS

FAMILY STATUS

PERCENT OF DEFAULTER
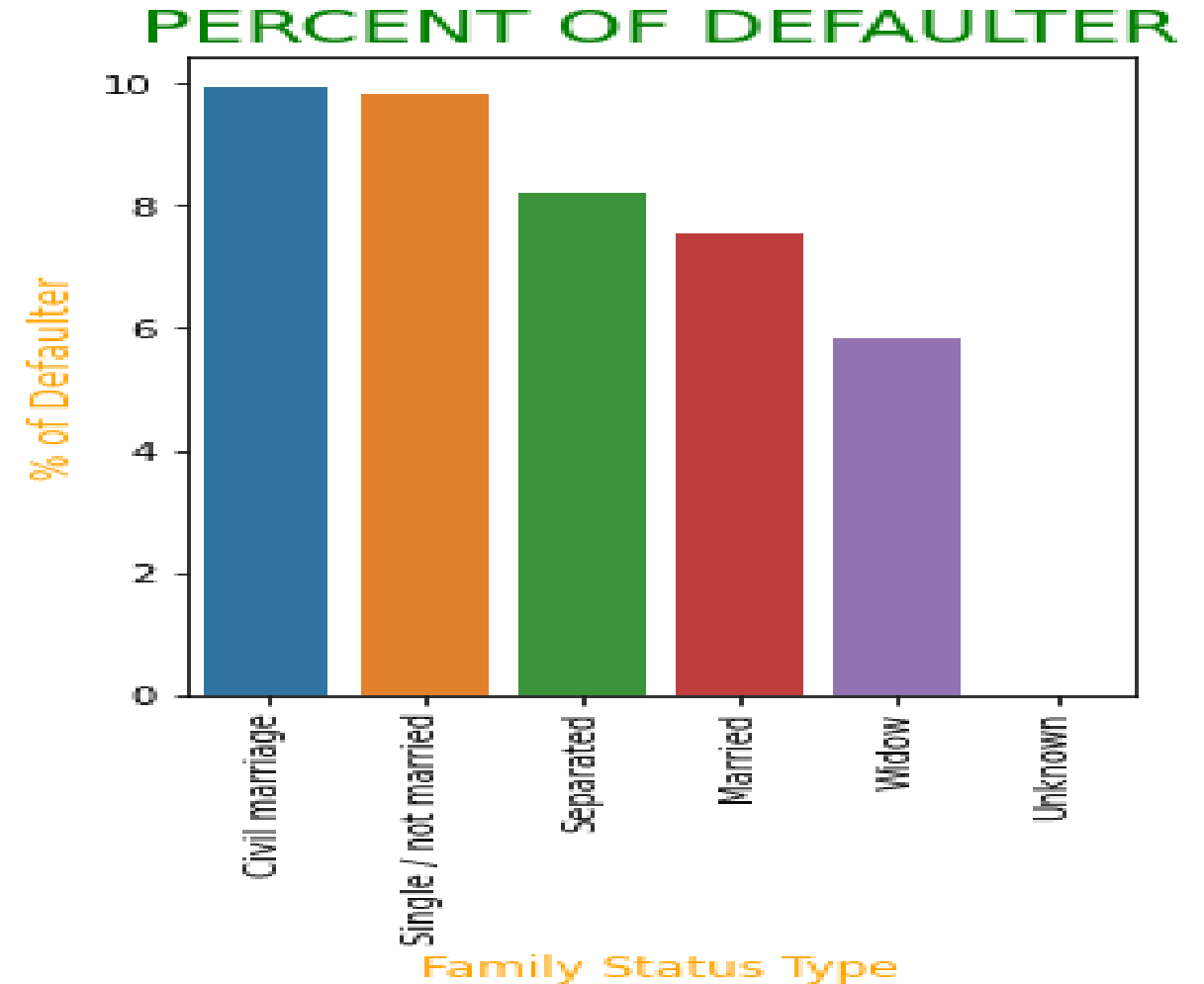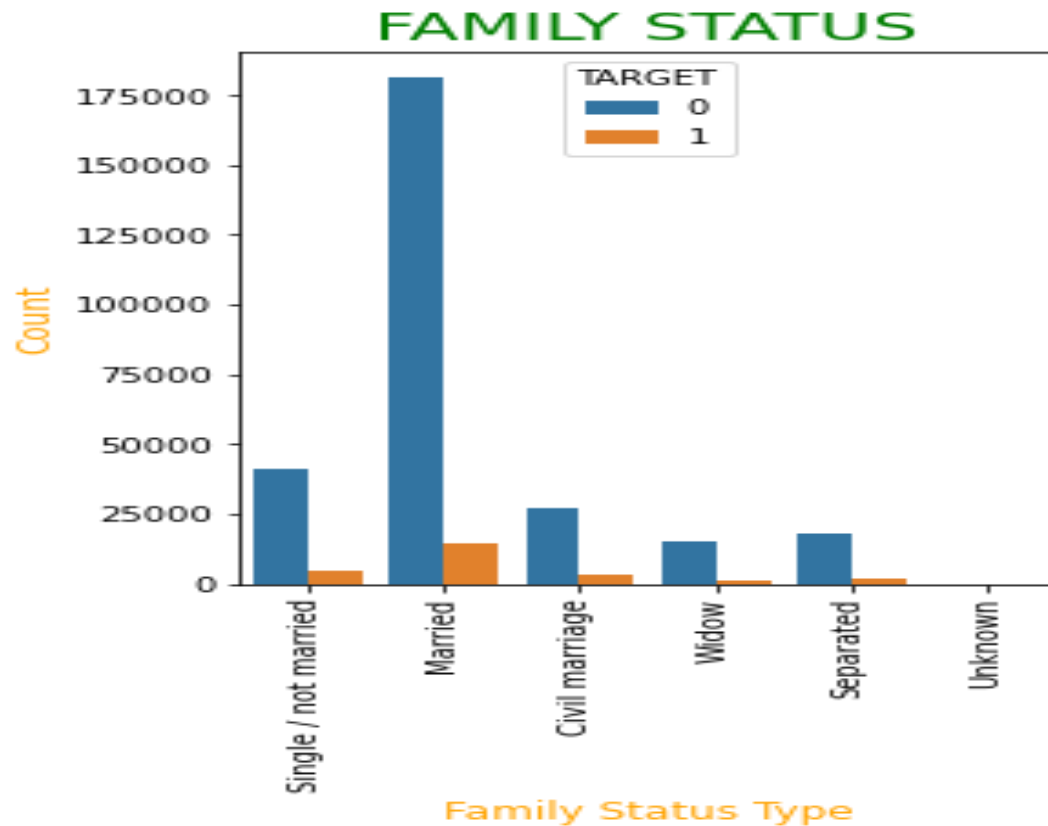
❑INSIGHTS.

- Most no of applicant applied for loan are married .
- Applicant who are widow have very less interest in loans.
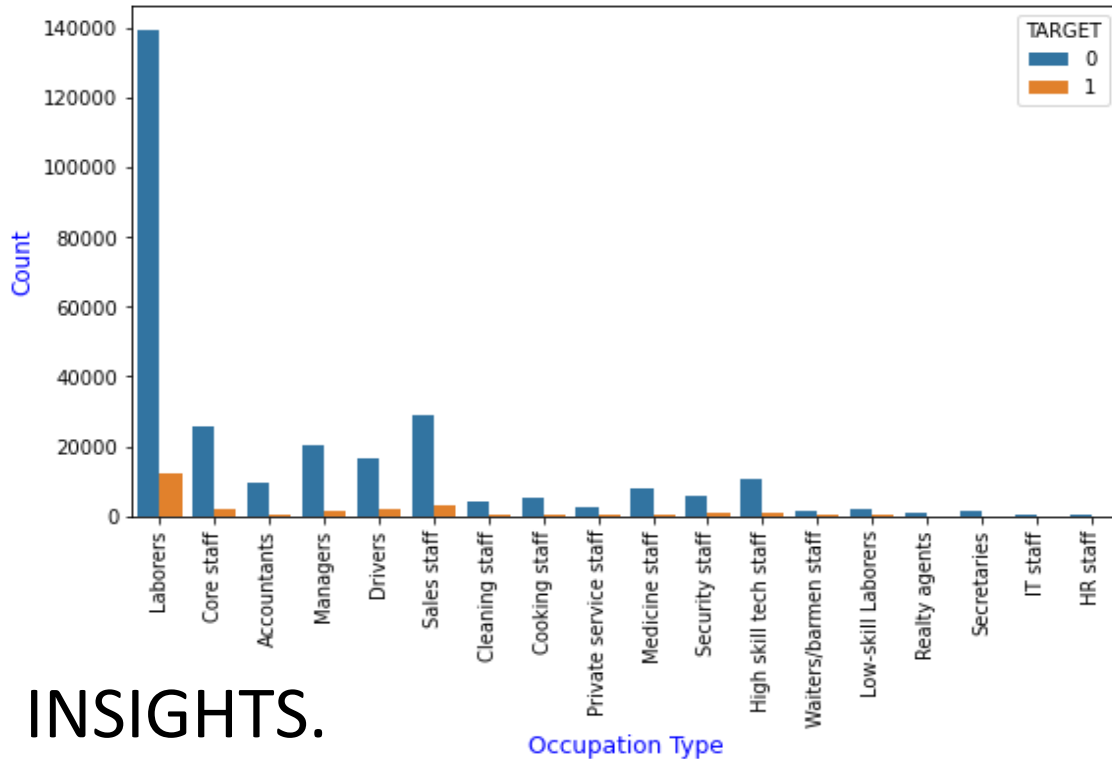- Single and Civil Marriage applicant have high default rate.
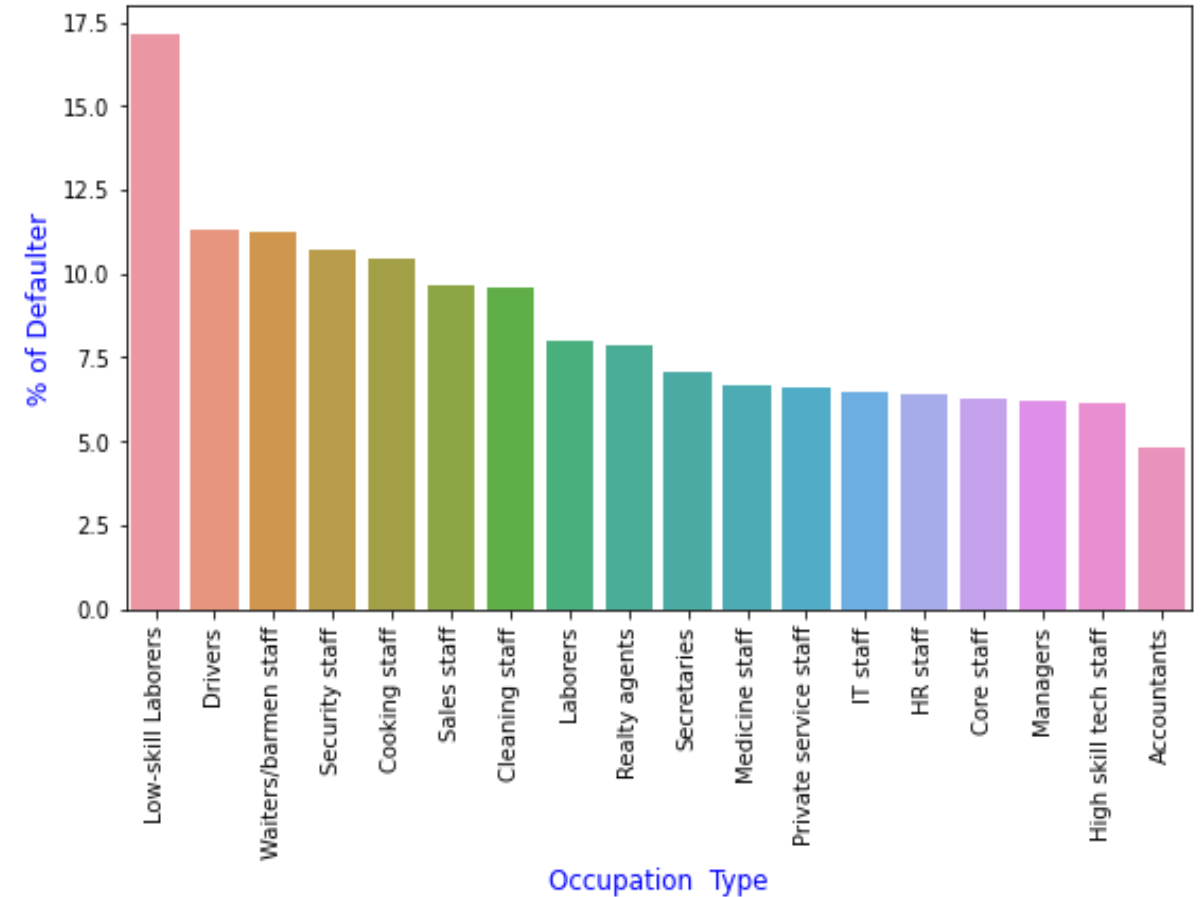
HOUSING TYPE

□ **INSIGHTS**.

- Most of the application are of applicant live in House/Apartment .

- Applicant live in office apartment are more likely to repay the loan.

- Applicant live with parents and live in rented apartment are more likely to be defaulter as default rate is high.
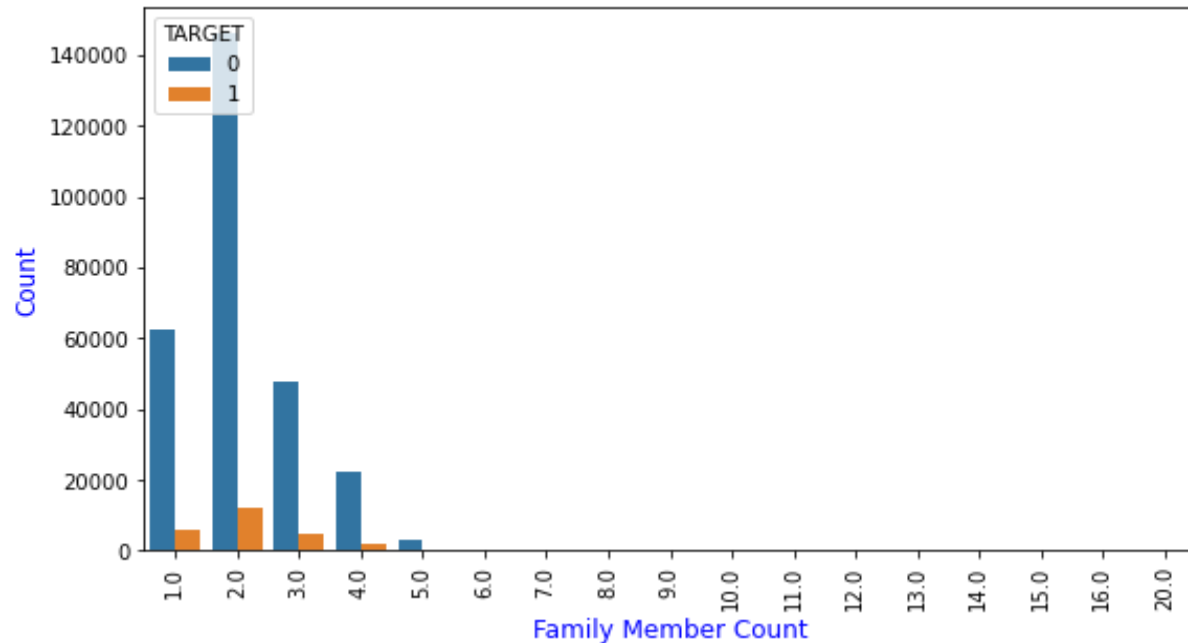
OCCUPATION

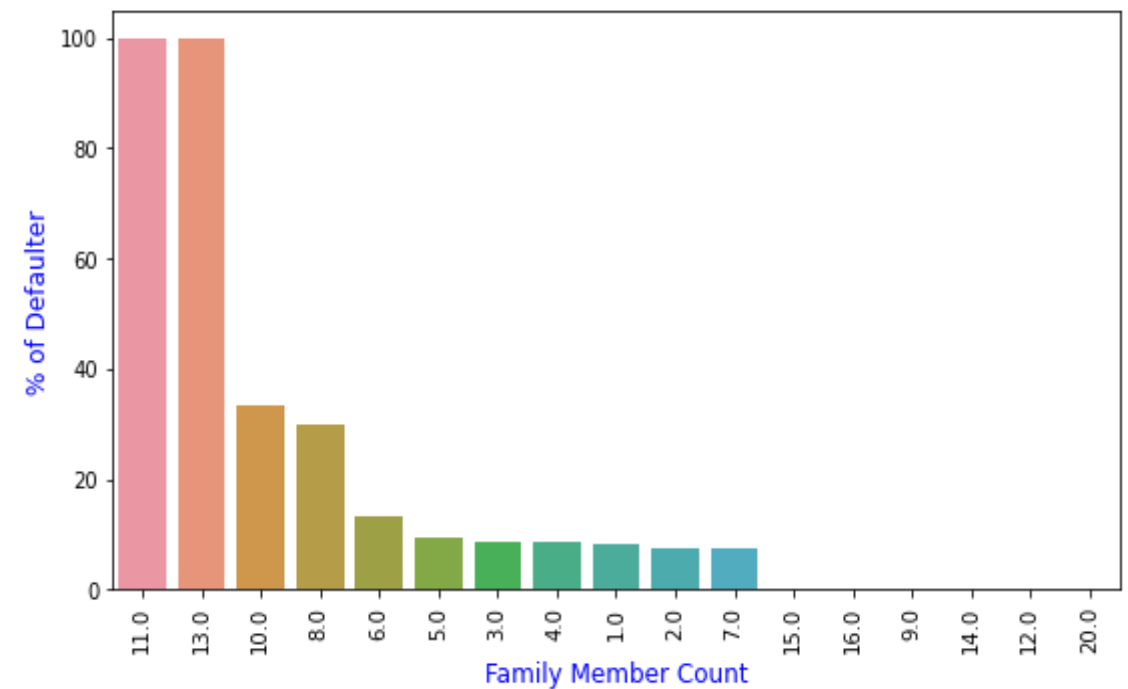## OCCUPATION TYPE

PERCENT OF DEFAULTER

- INSIGHTS.
- IT & HR staff applicant have very less interest in loans.
- Most no of application are of applicant working as a laborers.
- Clients working as low-skilled labour have very high default rate .
- Accountants are good repayer .

FAMILY MEMBER COUNT
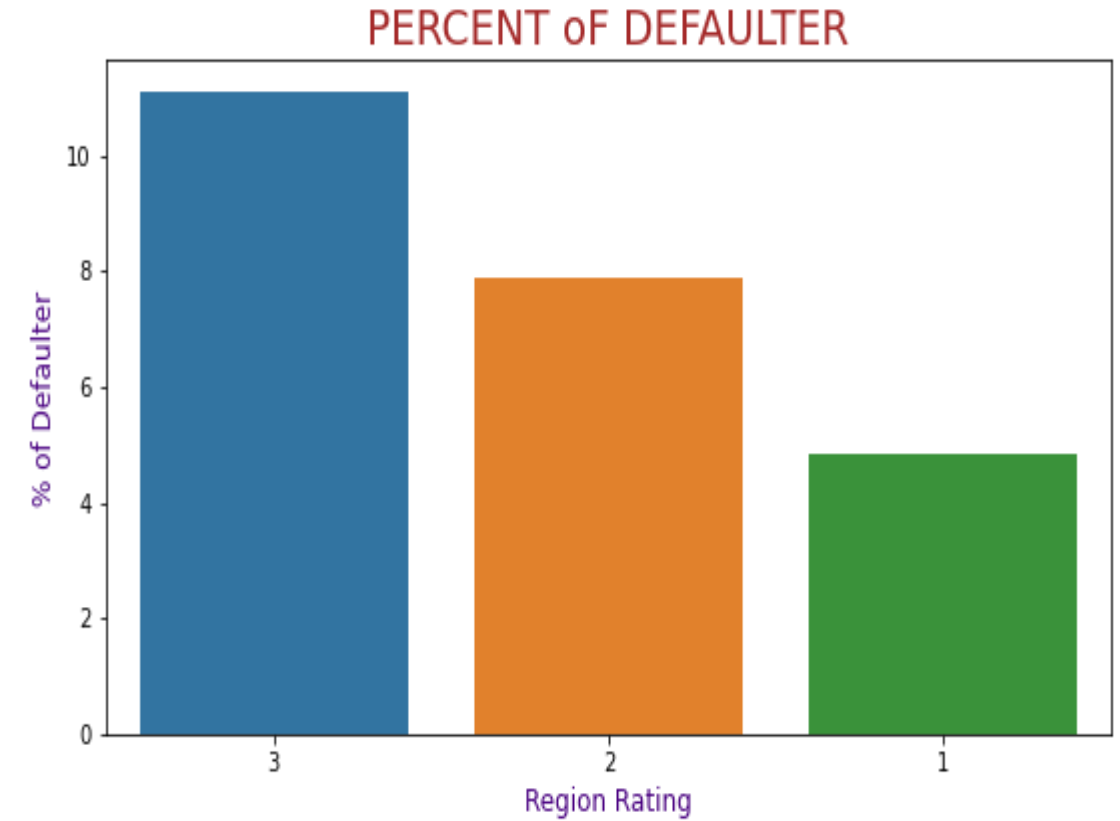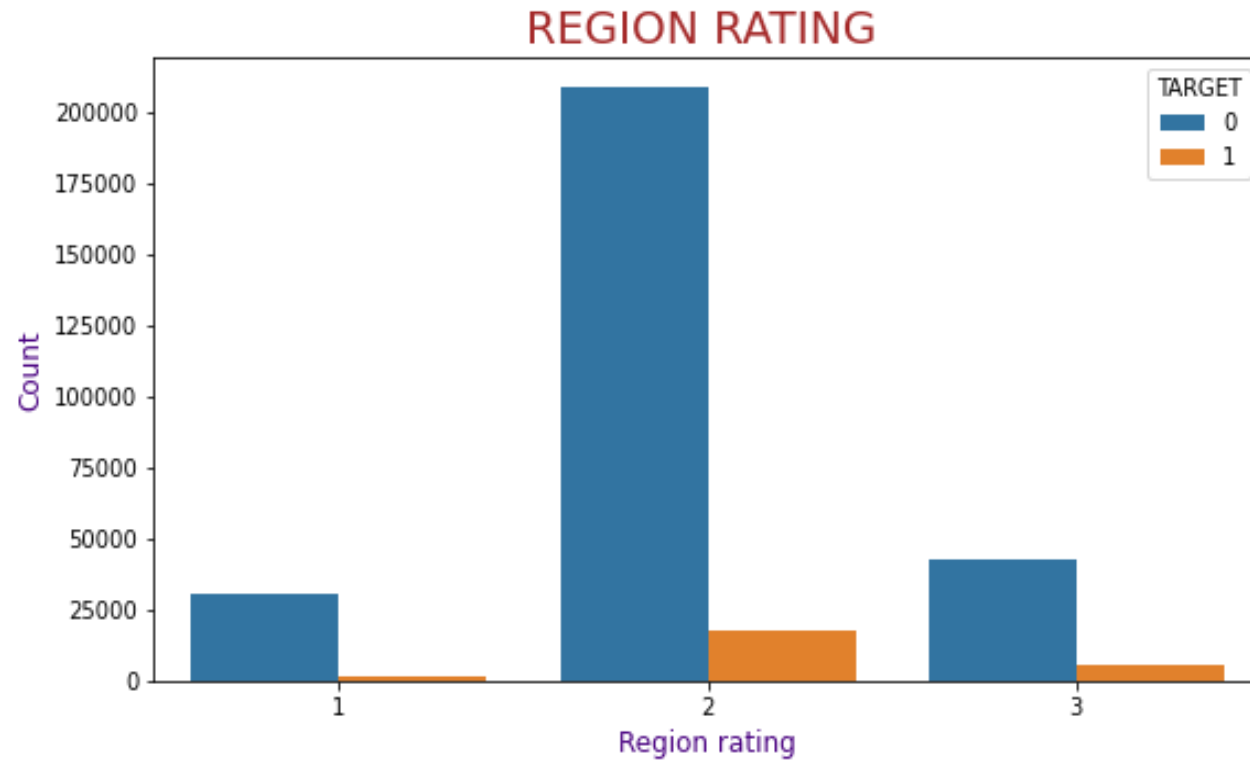

INSIGHT.

- Applicant have family member 2 are the most.

- Applicant having family member 11 and 13 are defaulter as default rate is 100% and company have to refused the application of such applicant.
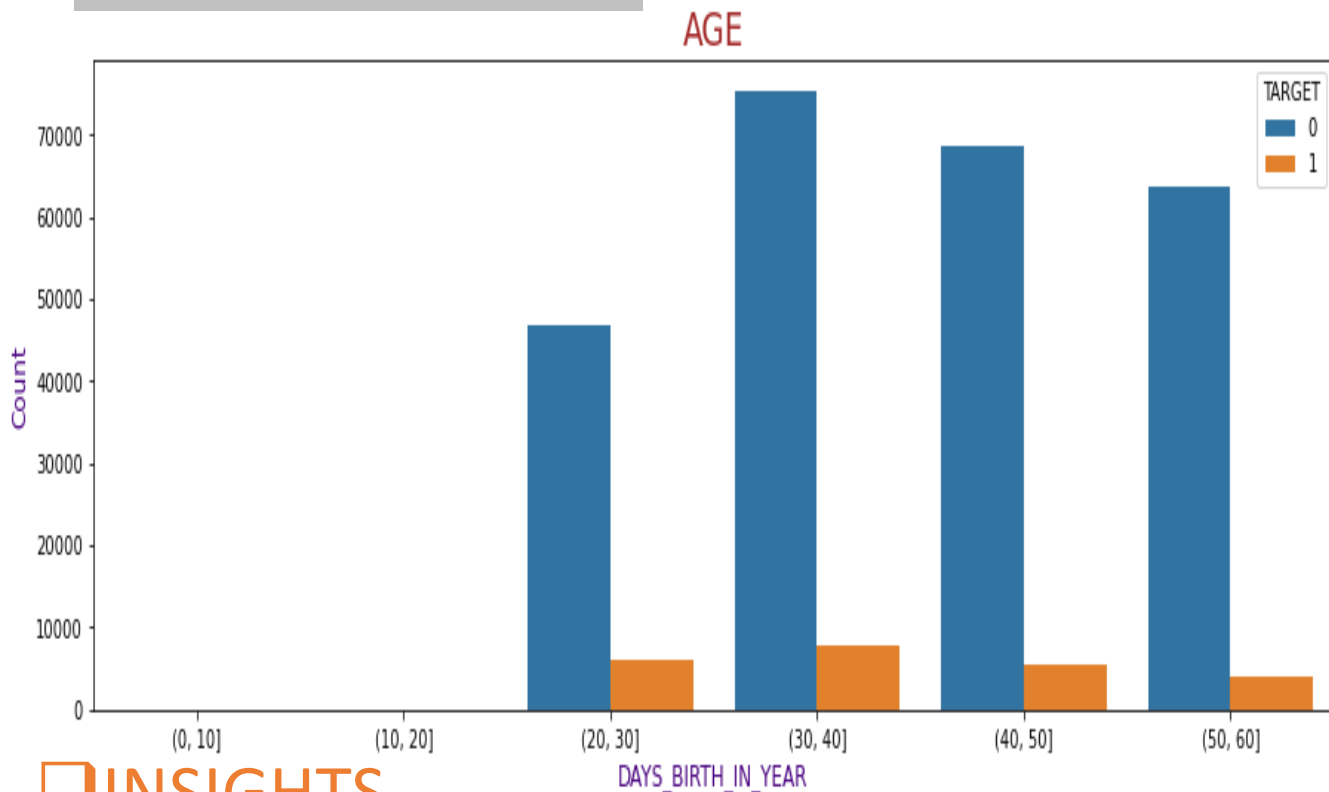
REGION RATING

❑INSIGHTS.

- Most of the applicant leave in area of Region-Rating 2

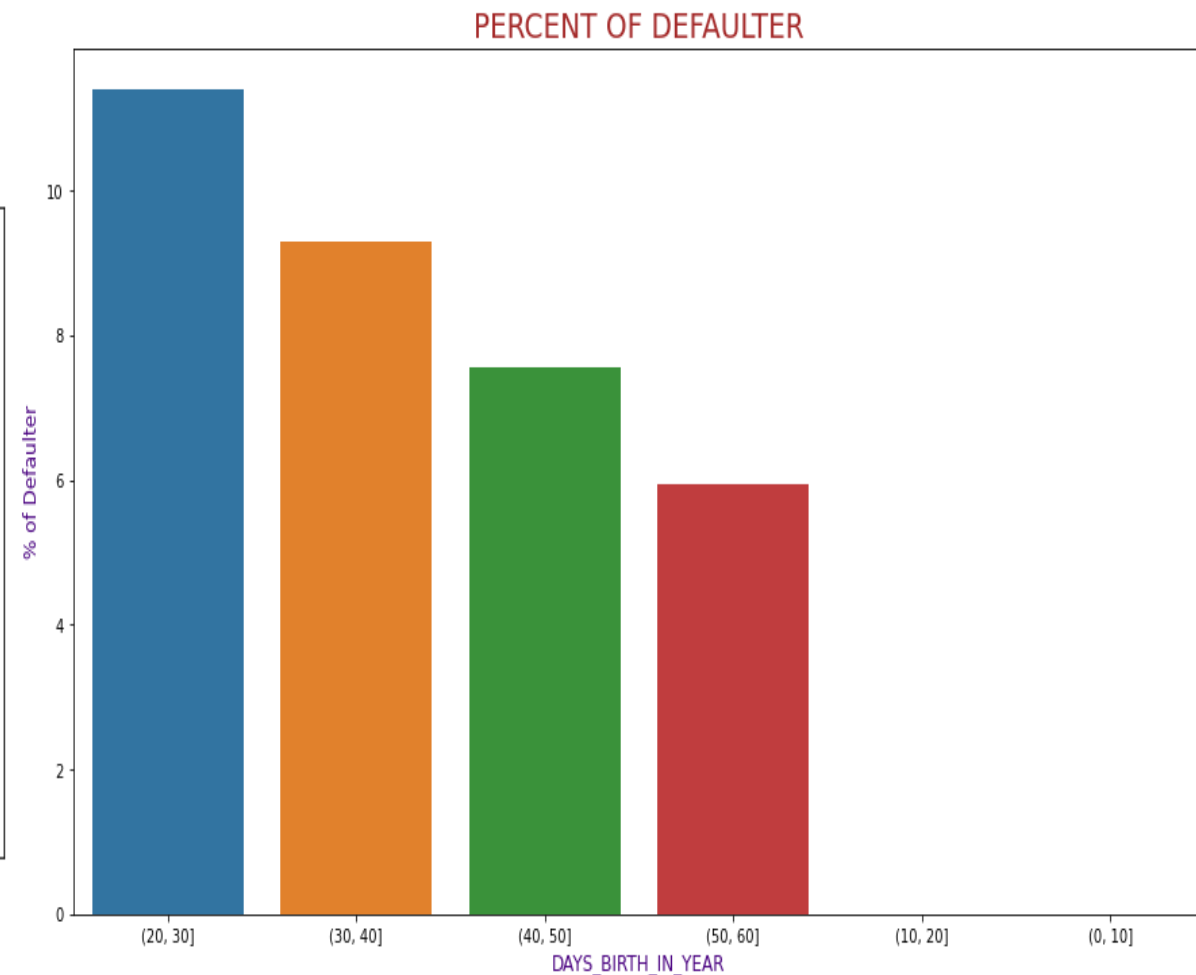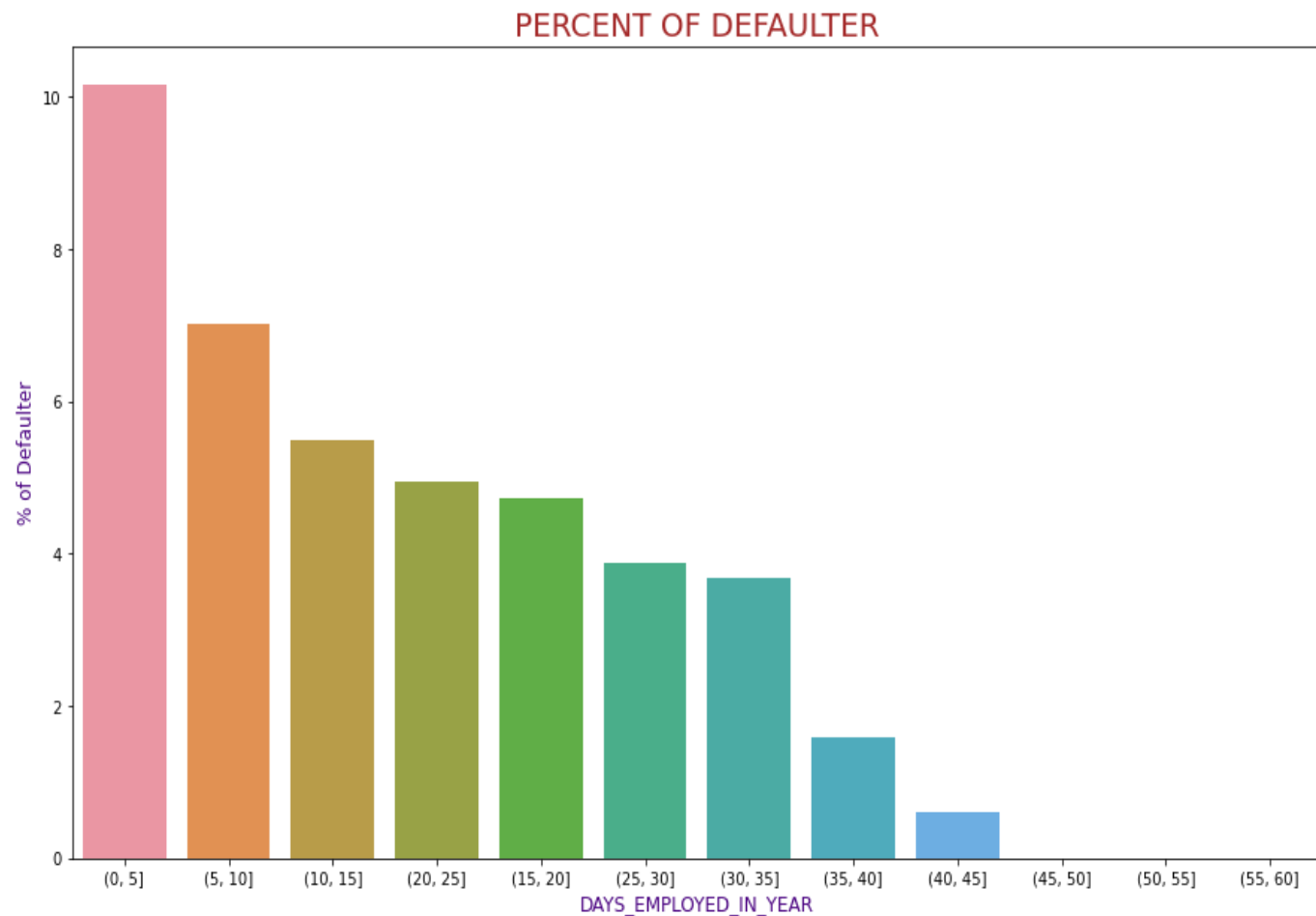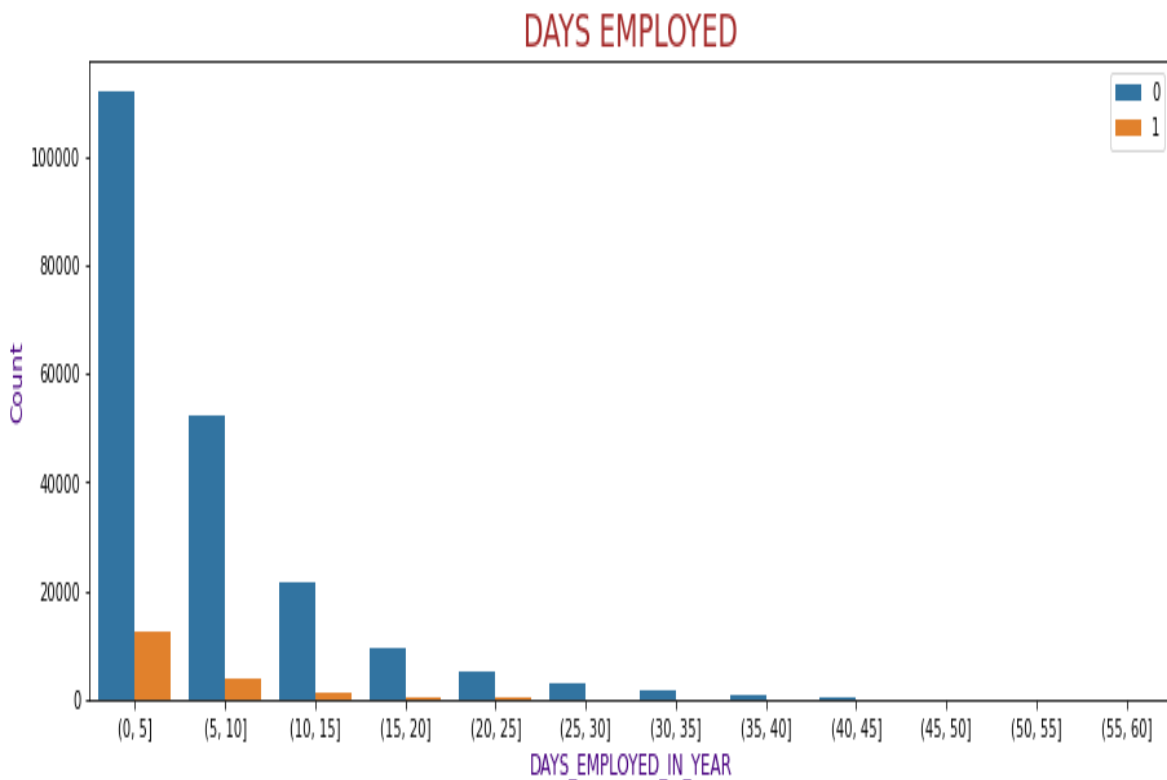- Applicant lives in Region-Rating 1 are safe to provide loans.

AGE GROUP

- Clients of age group 30-40 and 40-50 are most likely to take loan.
- Applicant of age above 50 are less chance of defaulting
- Applicant in age group 20-30 have high chances of defaulting.

DAYS EMPLOYED

DAYS EMPLOYED

PERCENT OF DEFAULTER
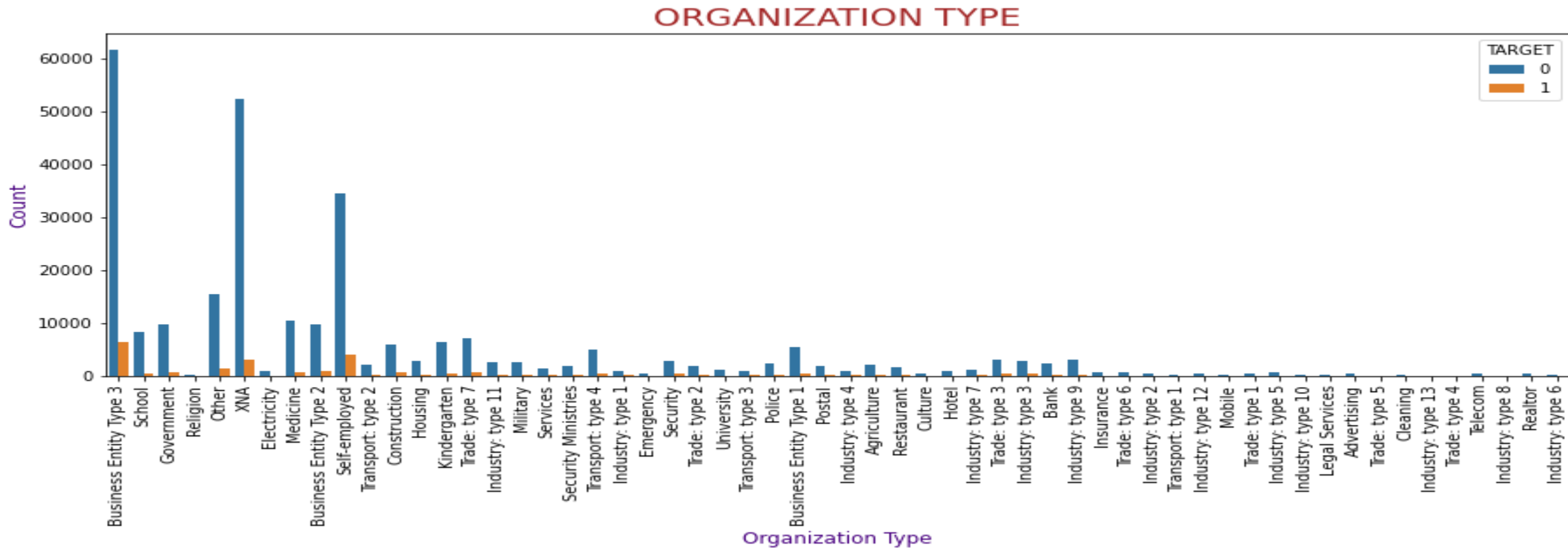

☐INSIGHTS

- Clients having less than 5 years of experience are more likely to take loan.
- Clients have 40 years are experience are less chance of defaulting .
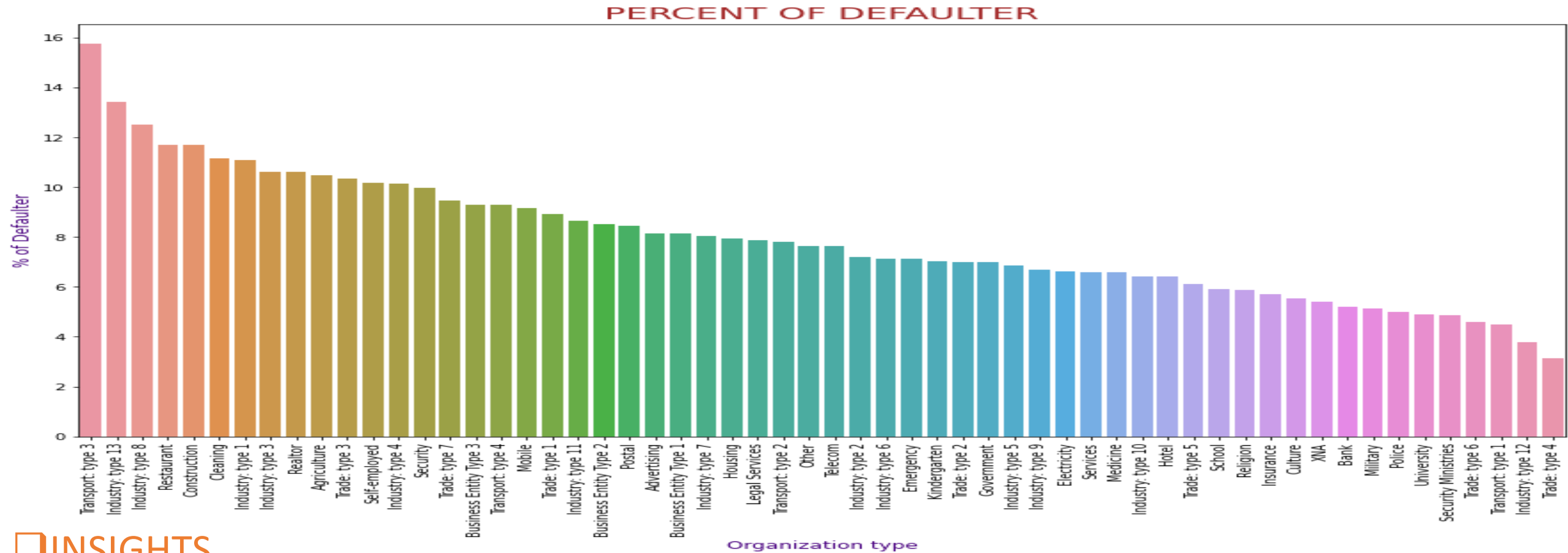- Clients having less than 5 years of experience have more chances of defaulting.

ORGANIZATION TYPE

☐ INSIGHTS.

- Most of the Applicant are from Business Entity Type 3.

- XNA and Self- Employed Clients also have high interest in loans.
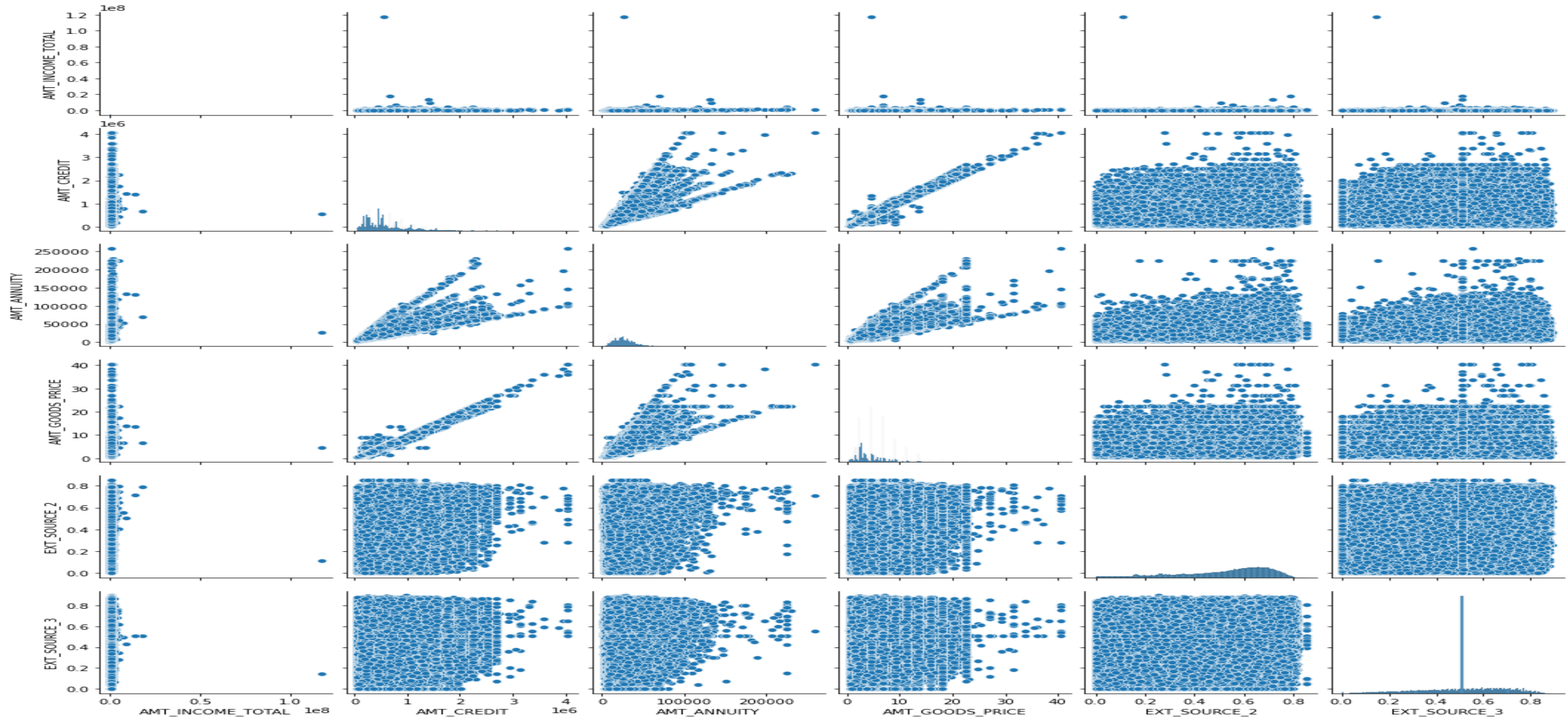
ORGANIZATION TYPE

PERCENT OF DEFAULTER

☐INSIGHTS.

- Transport type3 , Industry type 13 and Industry type 8 clients have high chances of not repaying load as defaulting percent is high

- Trade type4 clients have very low chance of defaulting.

BIVARIATE AND MULTIVARIATE ANALYSIS

ANALYSIS OF AMT_INCOME_TOTAL, AMT_CREDIT, AMT_GOODS_PRICE, EXT_SOURCE_2, AMT_ANNUITY

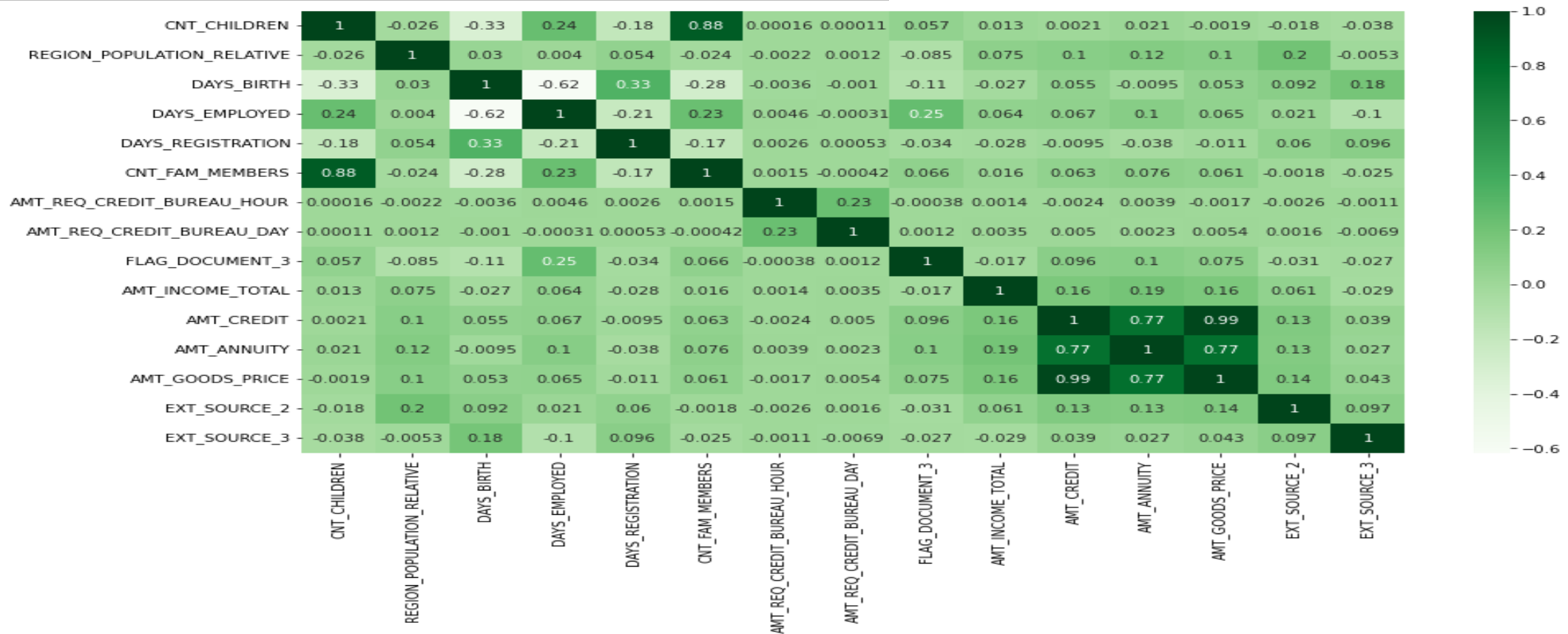# ANALYSIS OF AMT_INCOME_TOTAL, AMT_CREDIT, AMT_GOODS_PRICE, EXT_SOURCE_2, AMT_ANNUITY

• AMT_CREDIT and AMT_ANNUITY
have a very high positive correlation
coefficient. Clients who have for
high goods price will get a high loan
Amount.
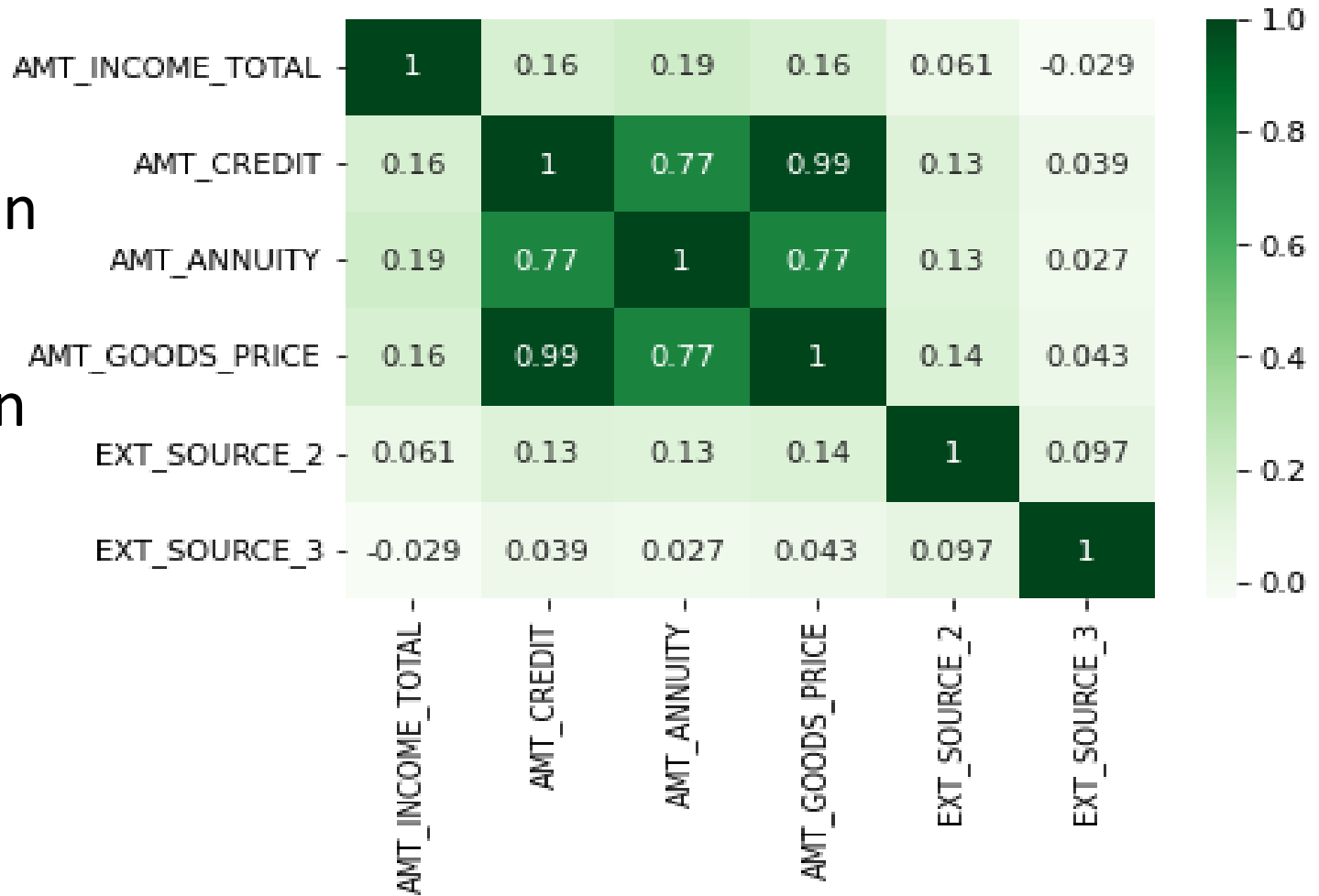
MULTIVARIATE ANALYSIS

- CNT_FAM_MEMBER has a good positive correlation coefficient with CNT_CHILDERN column.
- AMT_CREDIT has a good positive correlation with AMT_ANNUITY and AMT_GOODS.
- DAYS_BIRTH has a string negative correlation with DAYS_EMPLOYED.
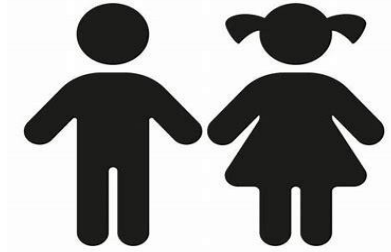
SUMMARY

FACTOR DERIVE FOR SAFE BORROWER (REPAYER)

Income should be more than
700K have less default rate

Female Applicant
Should be preferred

Client Child Count
Less than 2

Applicant with Academic Degree
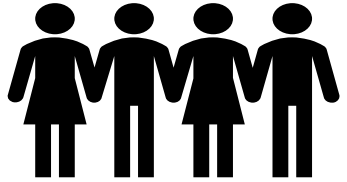has very less default rate .

Applicant live in own House /
Apartment are likely to repay loan

Accountant , high tech
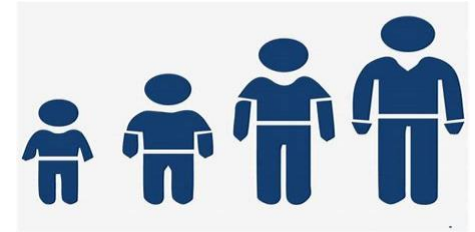Staff Employee are less
defaulter

FACTOR DERIVE FOR SAFE BORROWER (REPAYER)

Applicant having family Member count less than 4 have less default rate.

Applicant Live in area of Region Rating 1 are more likely to repay loan

Applicant with age 50+ have less chance of defaulting

Applicant having 40+ of experience are good repayer.

Applicant with Trade type 4 and 5 and industry type 12 have less default rate
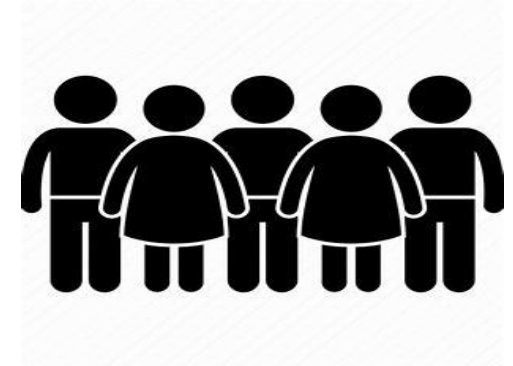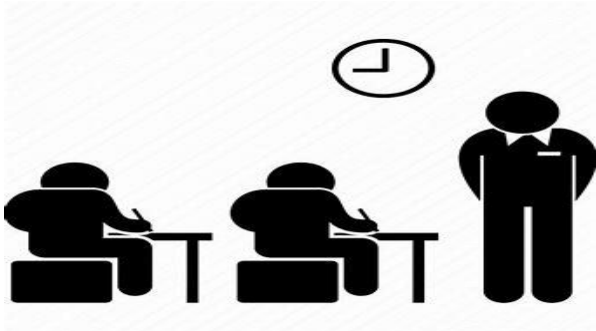
FACTOR DERIVE FOR IMPLICIT DEFAULTER

- Applicant salary less than
- 300K have high default rate

Male Applicant have default rate of 10%

Applicant child count more than 4 have high default rate

Applicant with lower/ secondary Education have more defaulter rate

Applicant live in rented apartment or with parents have high default rate of 12%

Civil Marriage/ Single applicant should provide at higher interest rate.
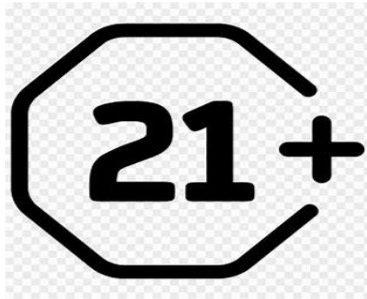
FACTOR DERIVE FOR IMPLICIT DEFAULTER

Applicant work as
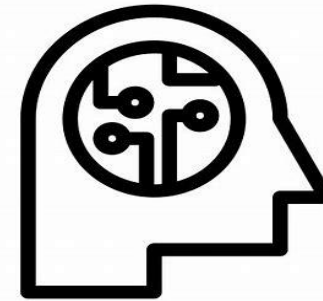a labour/driver have high
default rate

Applicant family member
more than 8 have default rate
100% so reject such application

Applicant live in area
of Region Rating 3
have high default rate

20+ age applicant have high
Default rate so such applicant
Should provide loan at higher rate.

Applicant having experience
less than 5 years have high
default rate.

THANKYOU !