CREDIT EDA

❖ Anand Bhosale

Problem Statement

When the company receives a loan application, the company has to decide on loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

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

➤ If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

Solution Approach

We are using the following approach for the insights:

➢ The required libraries need to be import for data cleaning and visualisation purpose.

➢ Identifying the missing data and using the appropriate method (Remove/ Replace).

➢ Identifying the outliers (if any) in the dataset.

➢ Identifying if there is a data imbalance and finding the ratio of data imbalance
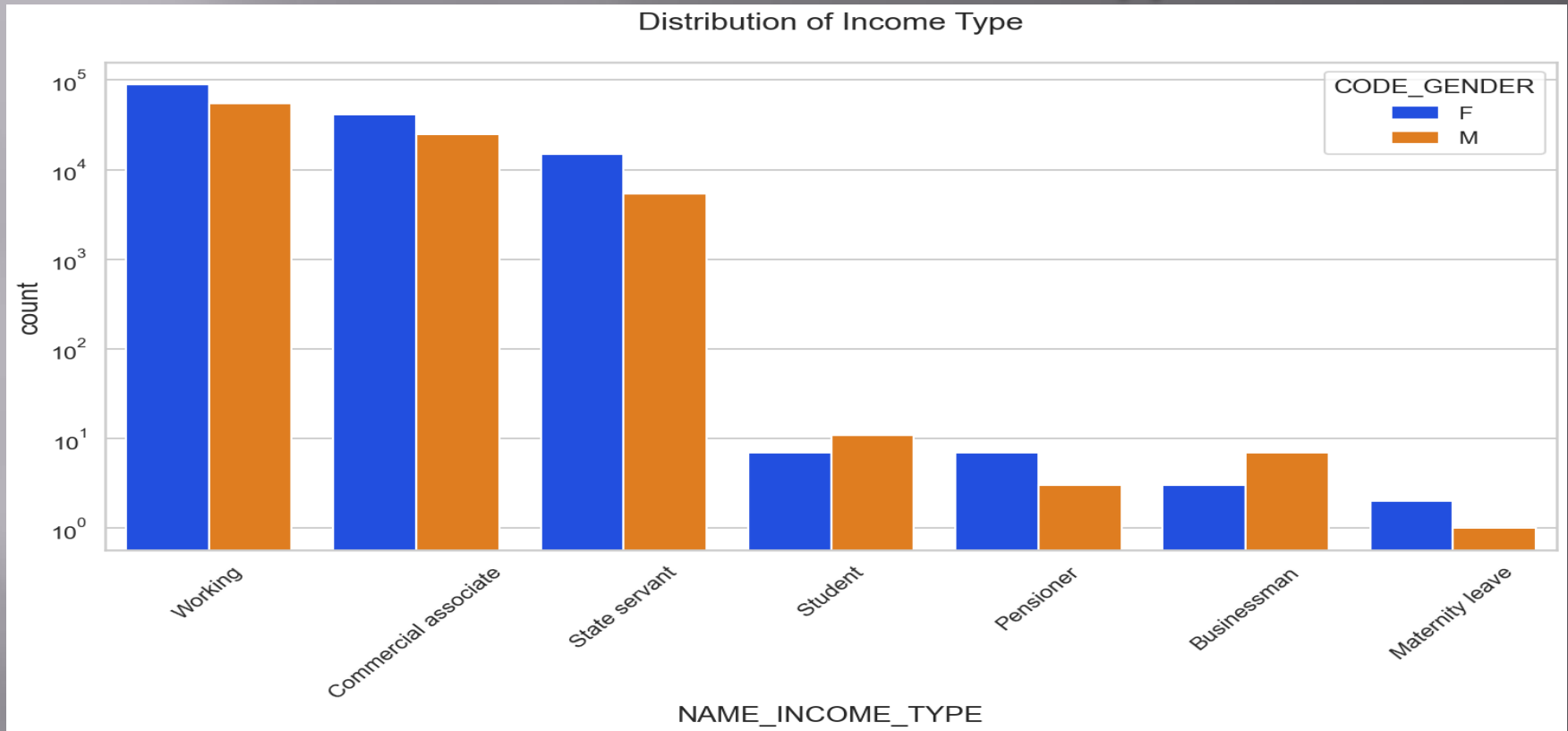
Univariate Analysis For Categorical Columns For Target = 0
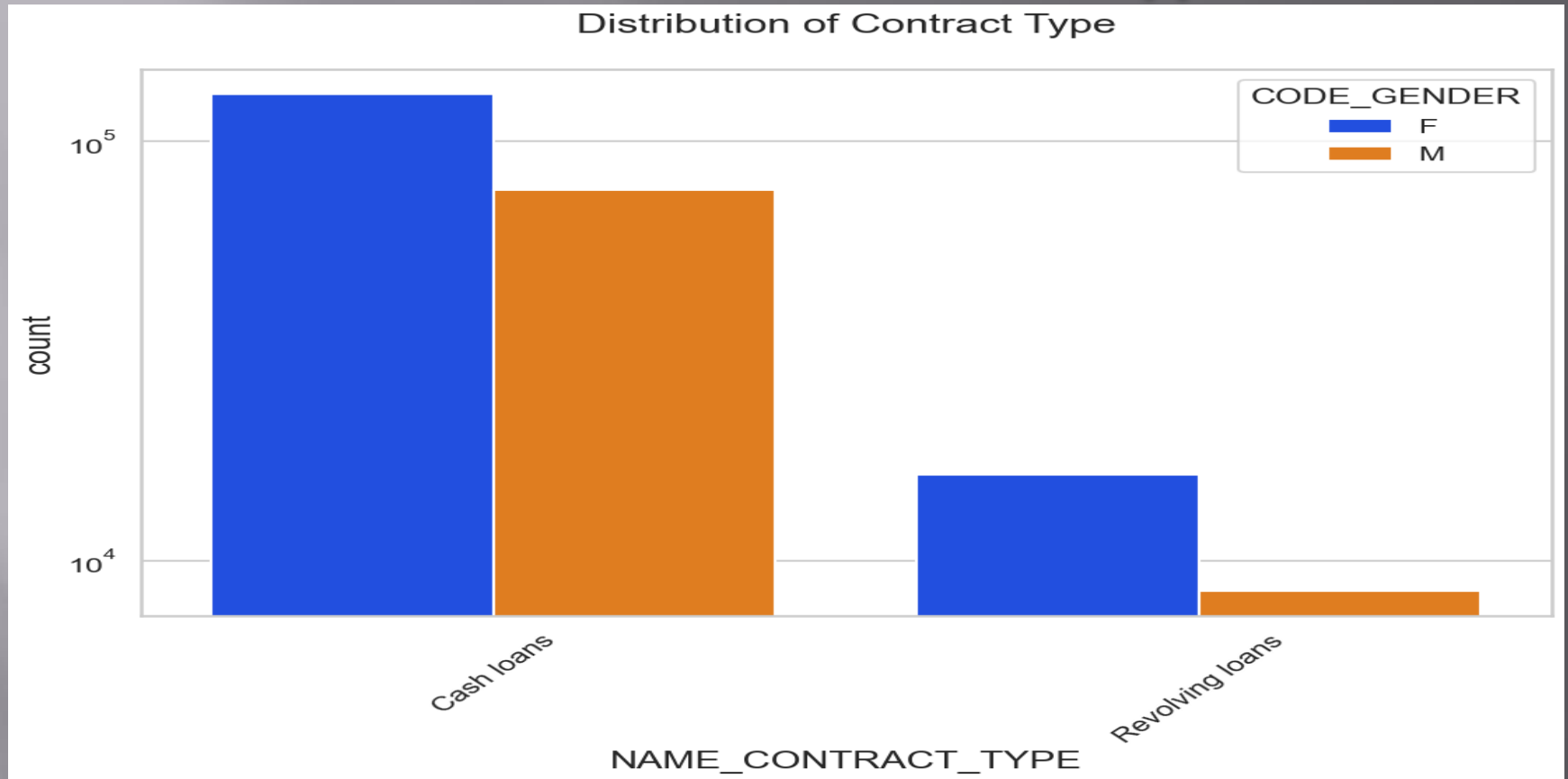
Distribution of Income Range
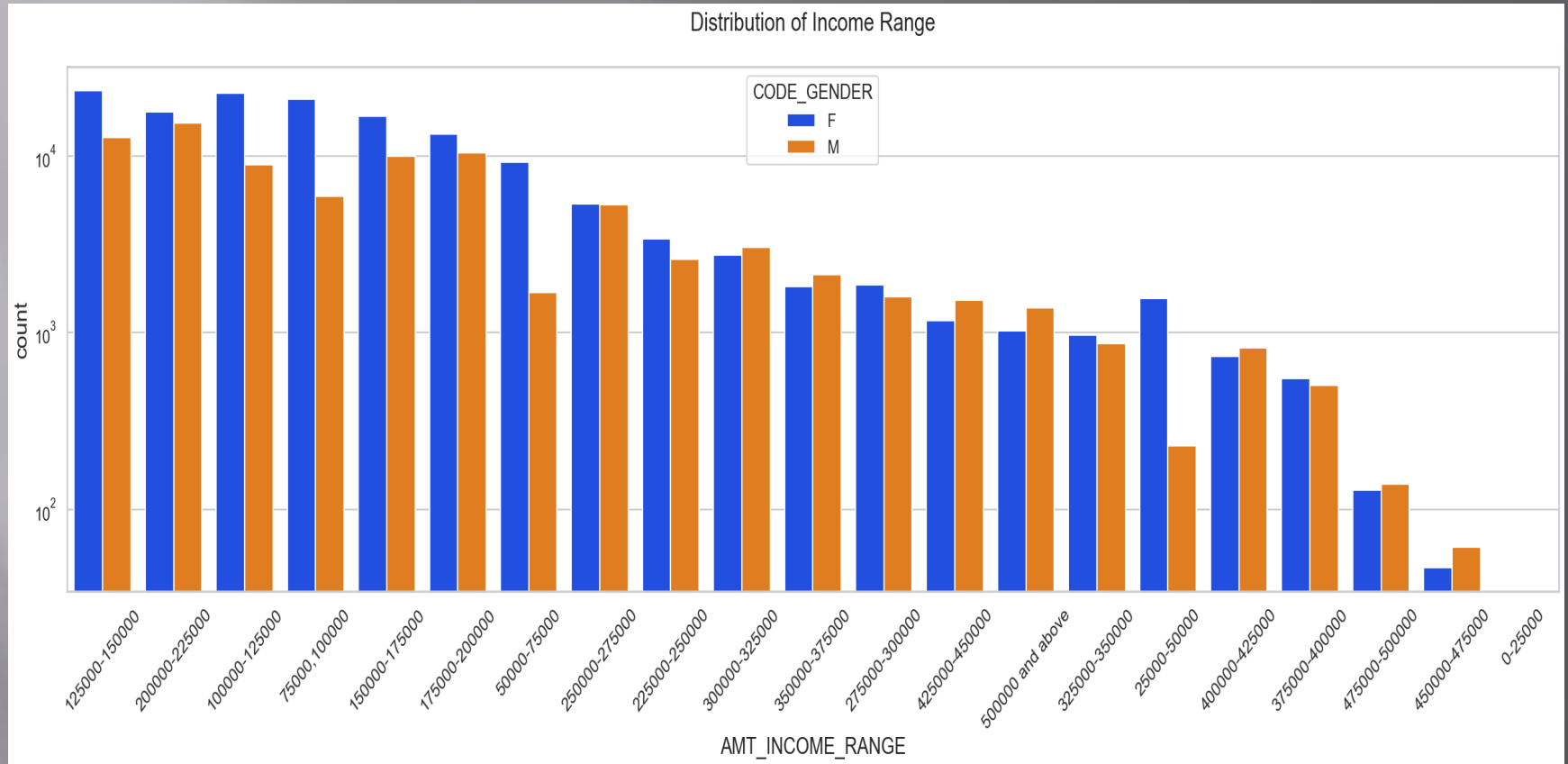
- The count of Females are more than Males
- From 1,25,000 to 2,00,000 of income range having more number of credits
- From 1,25,000 to 2,75,000 of income range the Females having more credit than Males

Distribution of Income Type

Distribution of Income Type

- The number of credits are higher for type of Working, Commercial Assiciates and State Servants

- Females are having more credits than the Males in Income Type of Working, Commercial Associates and State Servants

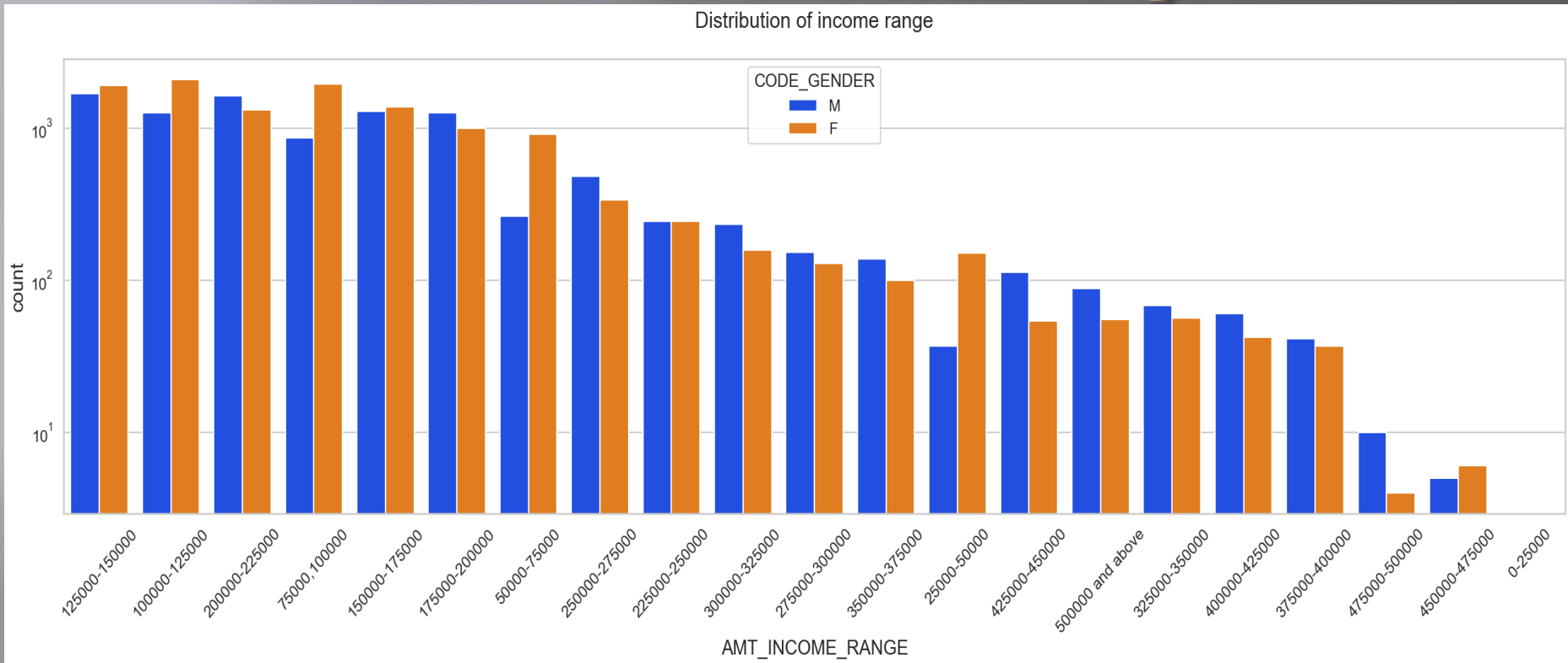# Distribution of Contract Type

Distribution of Contract Type

- In Contract Type the Cash Loans having the higher number of credits as compared to Revolving Loans

- In both the cases Females are having more credits than Males

# Univariate Analysis For Categorical Columns For Target = 1

Distribution of income range
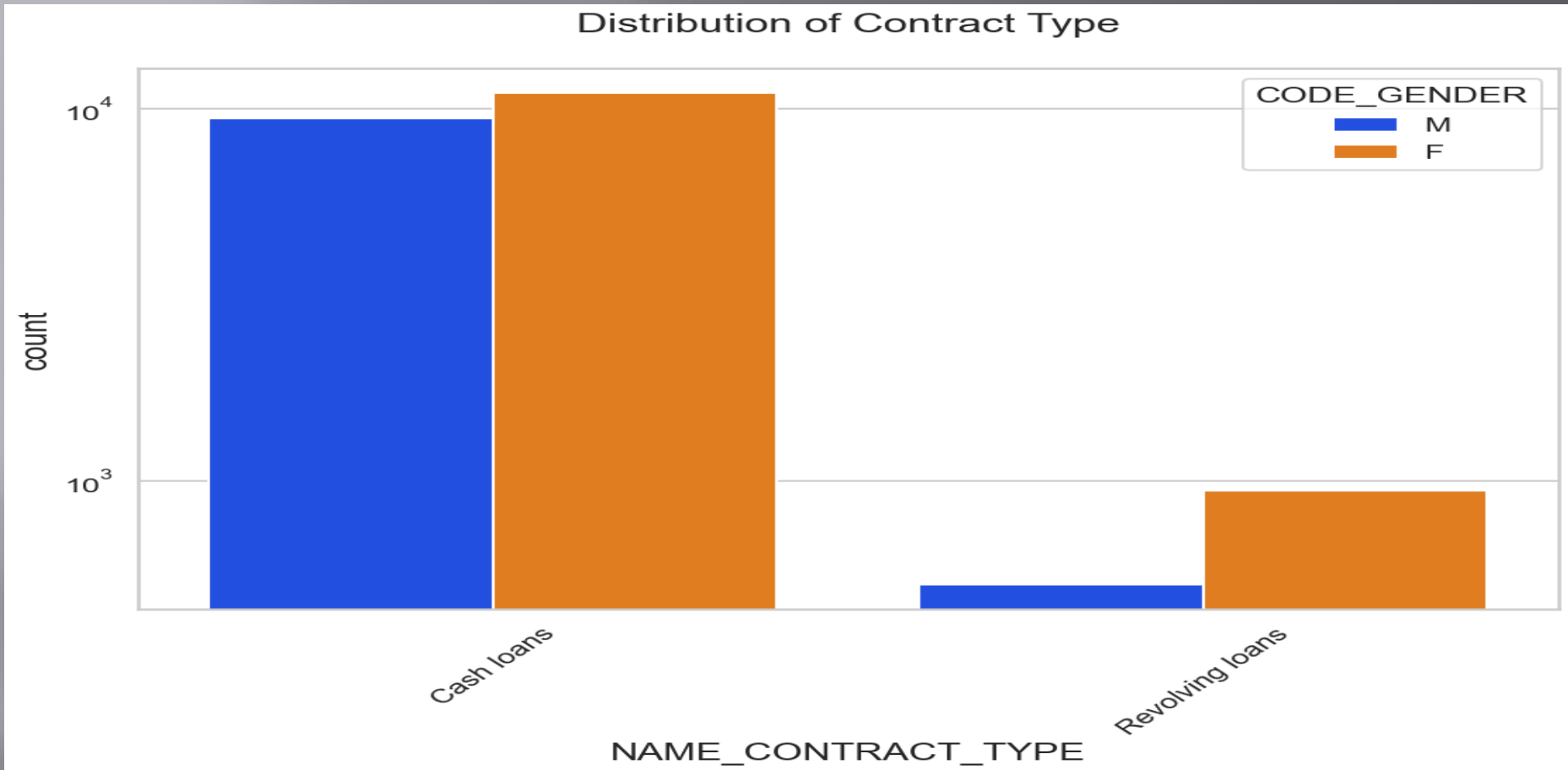

Distribution of income range

- The count of Females are more than Males

- From 1,25,000 to 2,00,000 of income range having more number of credits

- From 1,25,000 to 2,75,000 of income range the Females having more credit than Males

Distribution of Income Type

Distribution of Income type

- The number of credits are higher for type of Working, Commercial Assiciates and State Servants
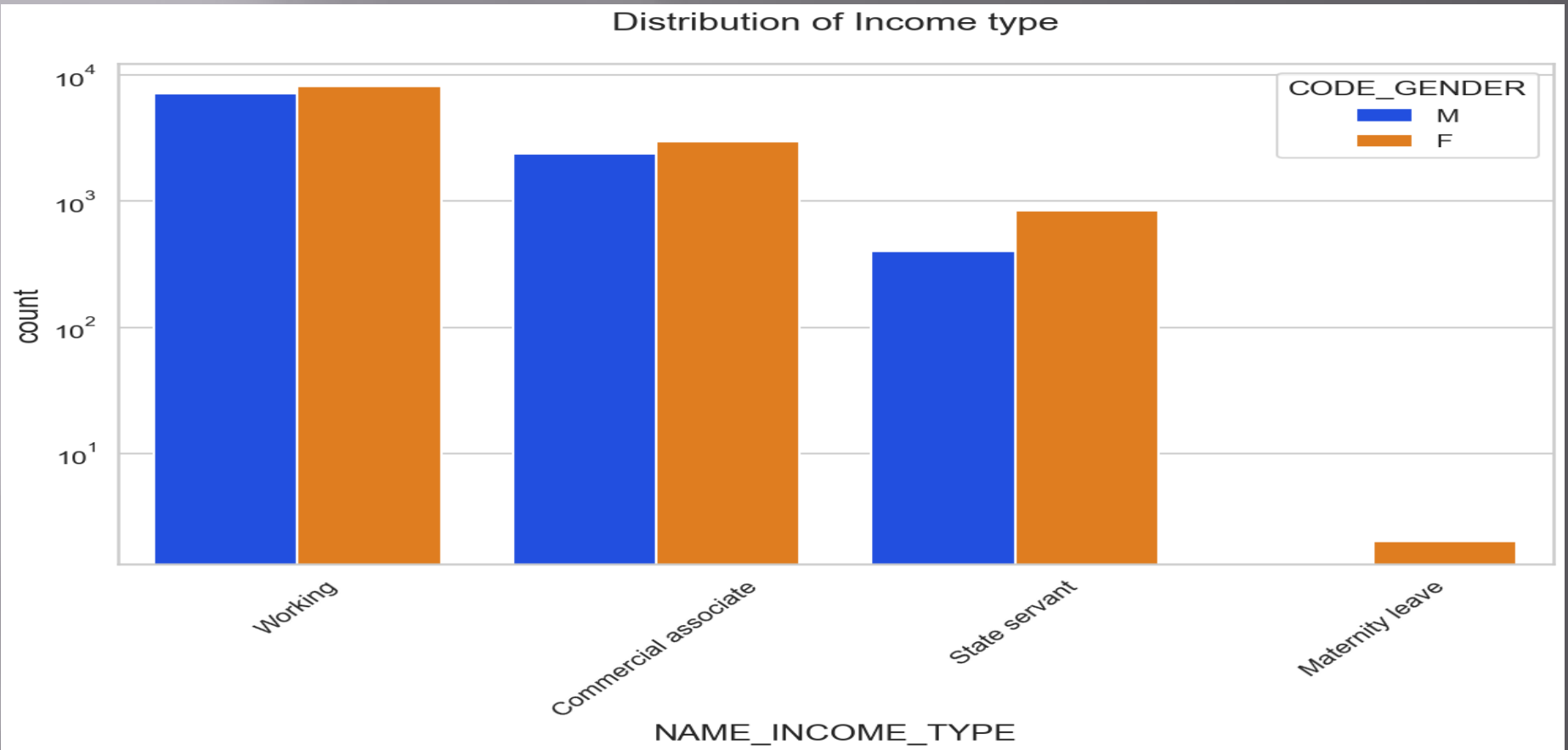
- Females are having more credits than the Males in Income Type of Working, Commercial Associates and State Servants
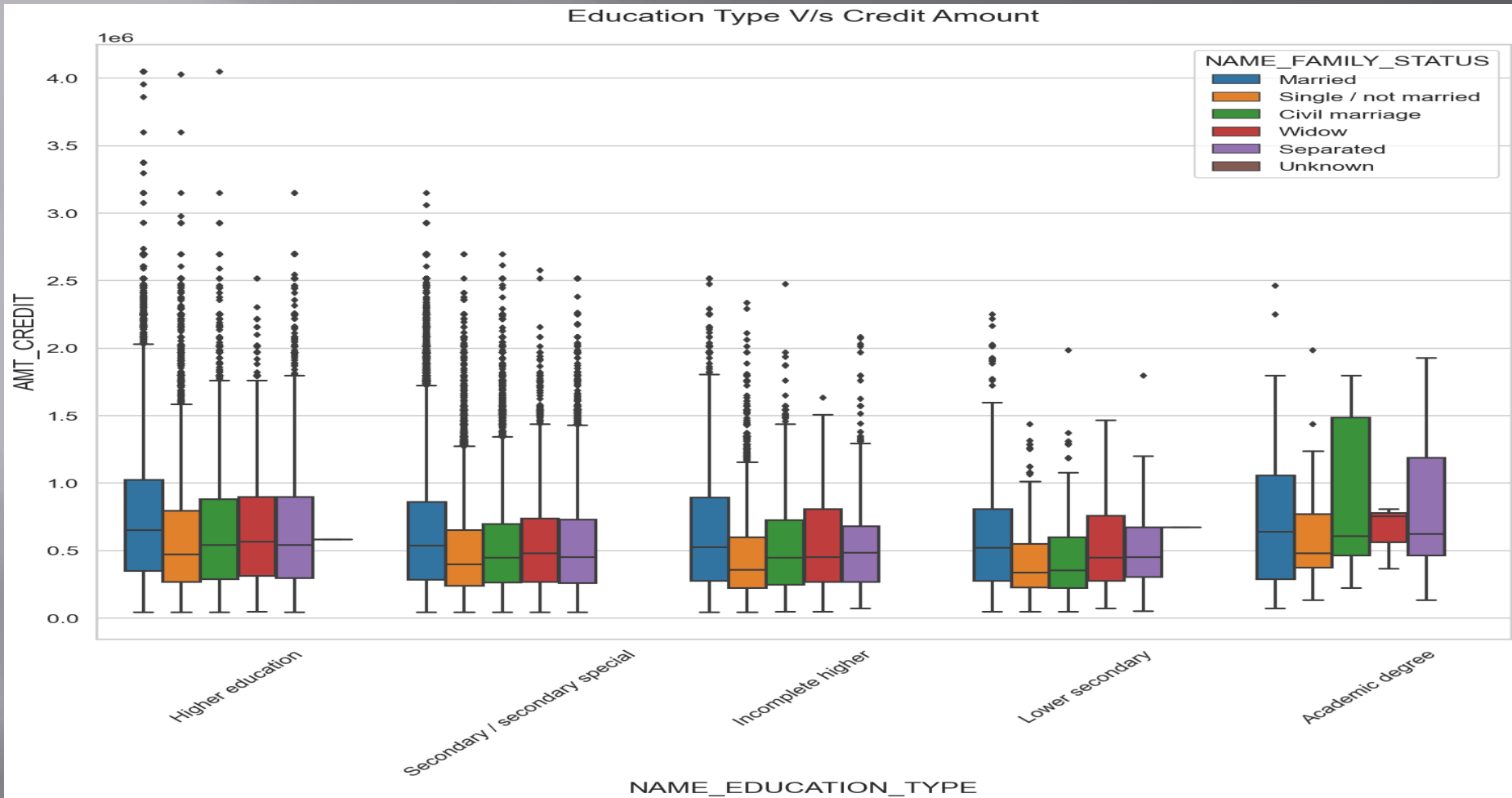
Distribution of Contract Type

Distribution of Contract Type

- In Contract Type the Cash Loans having the higher number of credits as compared to Revolving Loans

- In both the cases Males are having more credits than Females
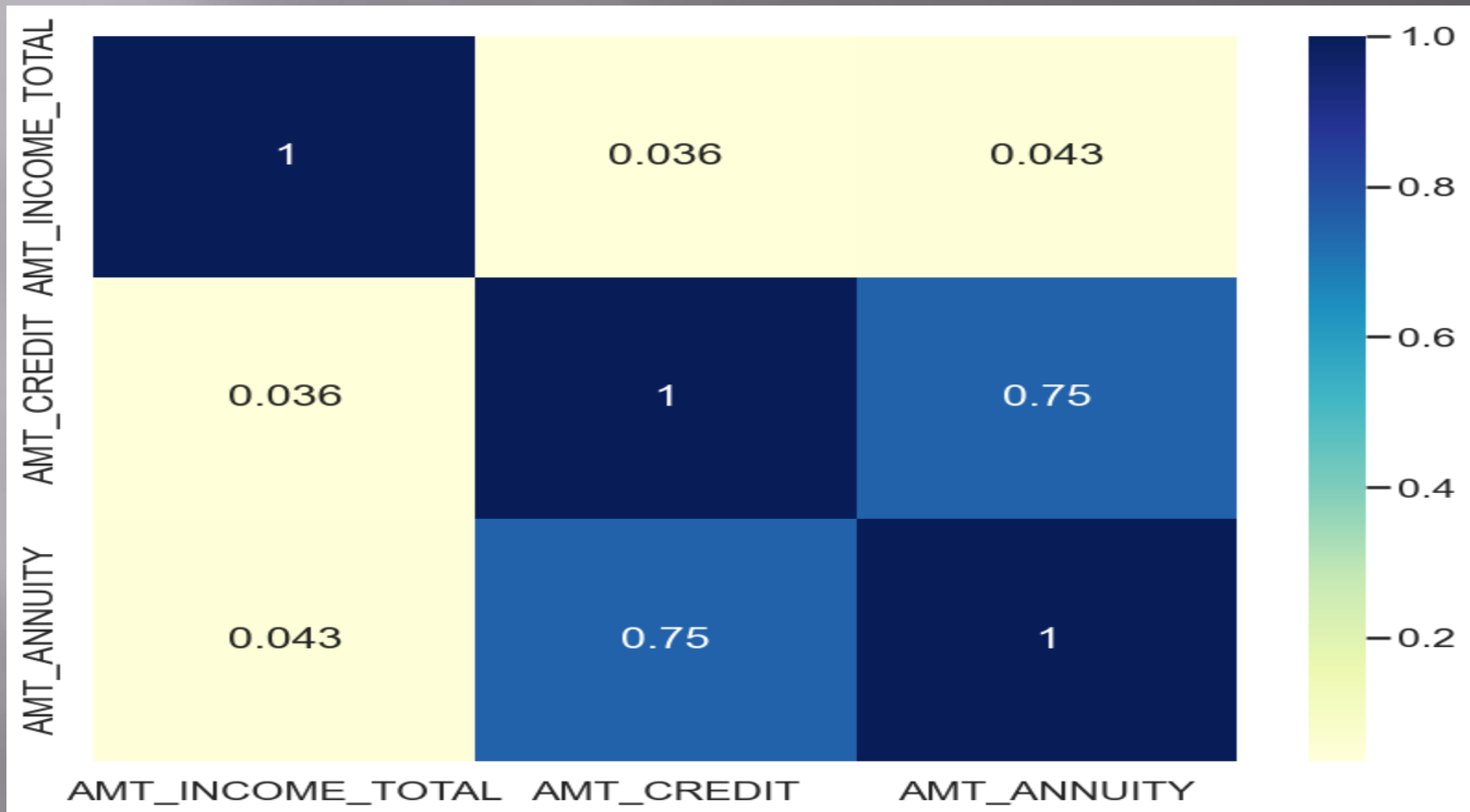
Bivariate Analysis

Education Type V/s Credit Amount

- In all Categories of Education Type : Married, Single/Not Maiired and Civil Marriage types are having larger number of Outliers

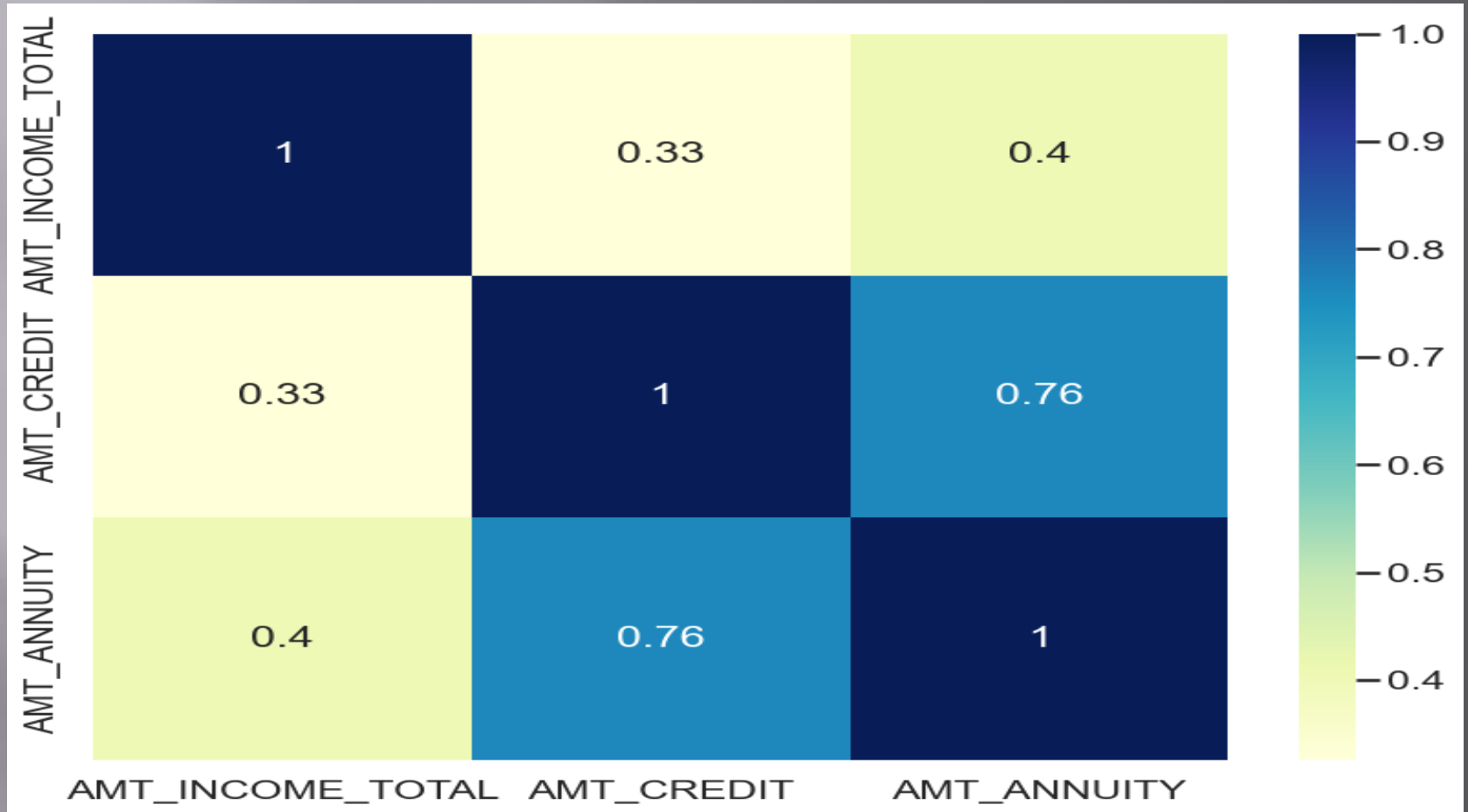- Civil Marriage in Academics Degree Category is having more credits

Coralation

For Defaluters

Corelated Columns are :

- AMT_CREDIT & AMT_ANNUITY
- AMT_INCOME_TOTAL & AMT_ANNUITY
- AMT_INCOME_TOTAL & AMT_CREDIT

For Non-Defaluters

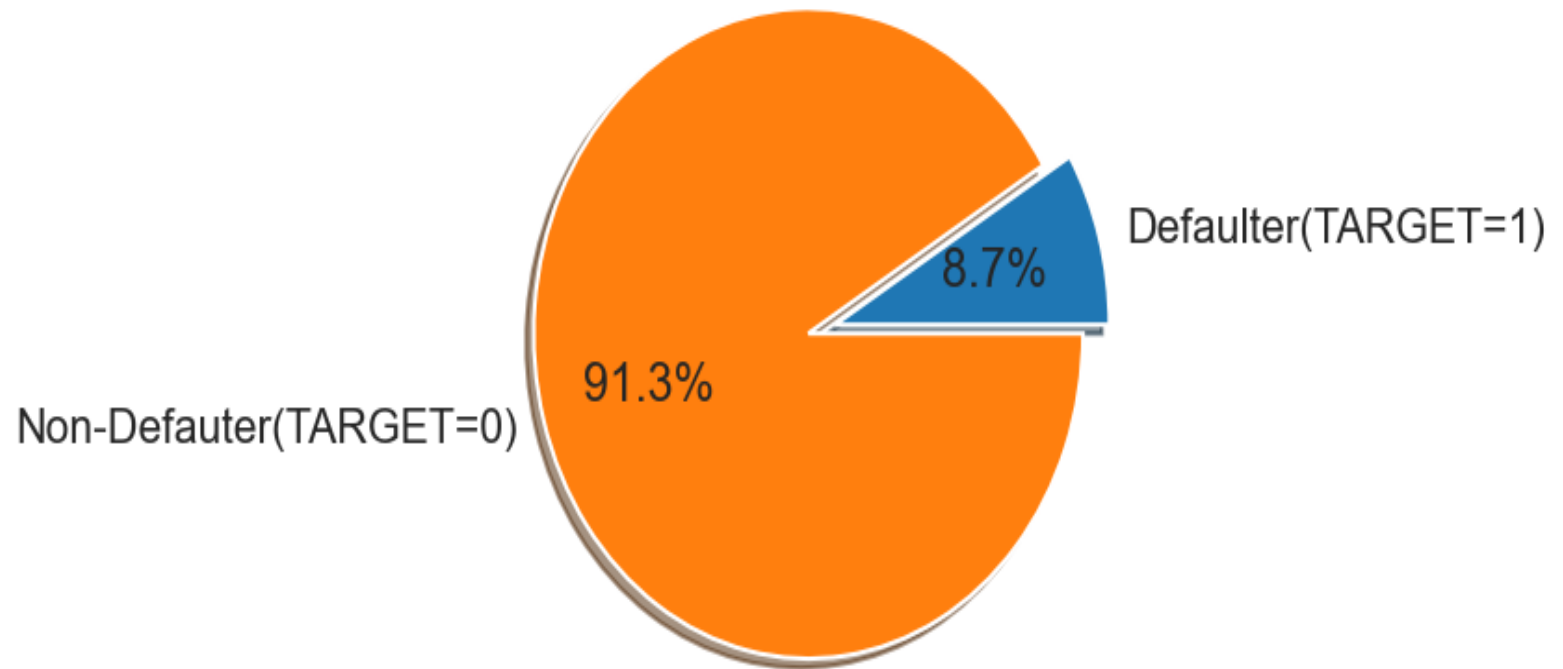**Corelated Columns are :**

- AMT_CREDIT & AMT_ANNUITY
- AMT_INCOME_TOTAL & AMT_ANNUITY
- AMT_INCOME_TOTAL & AMT_CREDIT

Data Imbalance

Data imbalance- Pie Chart

Defaulter(TARGET=1) 8.7%

Non-Defauter(TARGET=0) 91.3%

Conclusion

- Students, Pensioner and Businessman in contract type are doing payments successfully, hence banks should focus more on these factors.

- Working category in income type are unsuccessful in doing payments, hence banks should have less focus on these.

THANK YOU