

# EDA Case Study and Bank Loan Database

---

*Submitted by*

*Vismaya VT*

# Objective

---

Based on the applicant's profile, credit risk analysis will assist the in deciding whether to approve the loan.

This prevents the company from losing money and manages the loss of business.

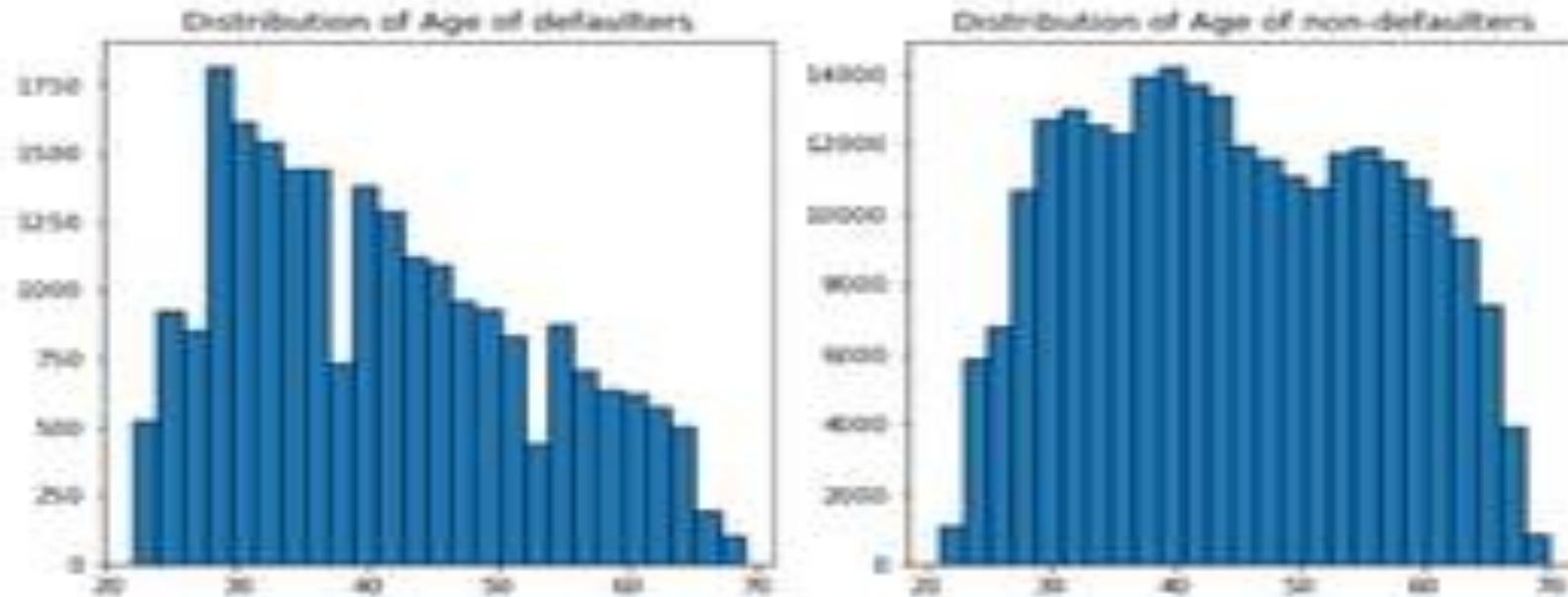


# Steps:

---

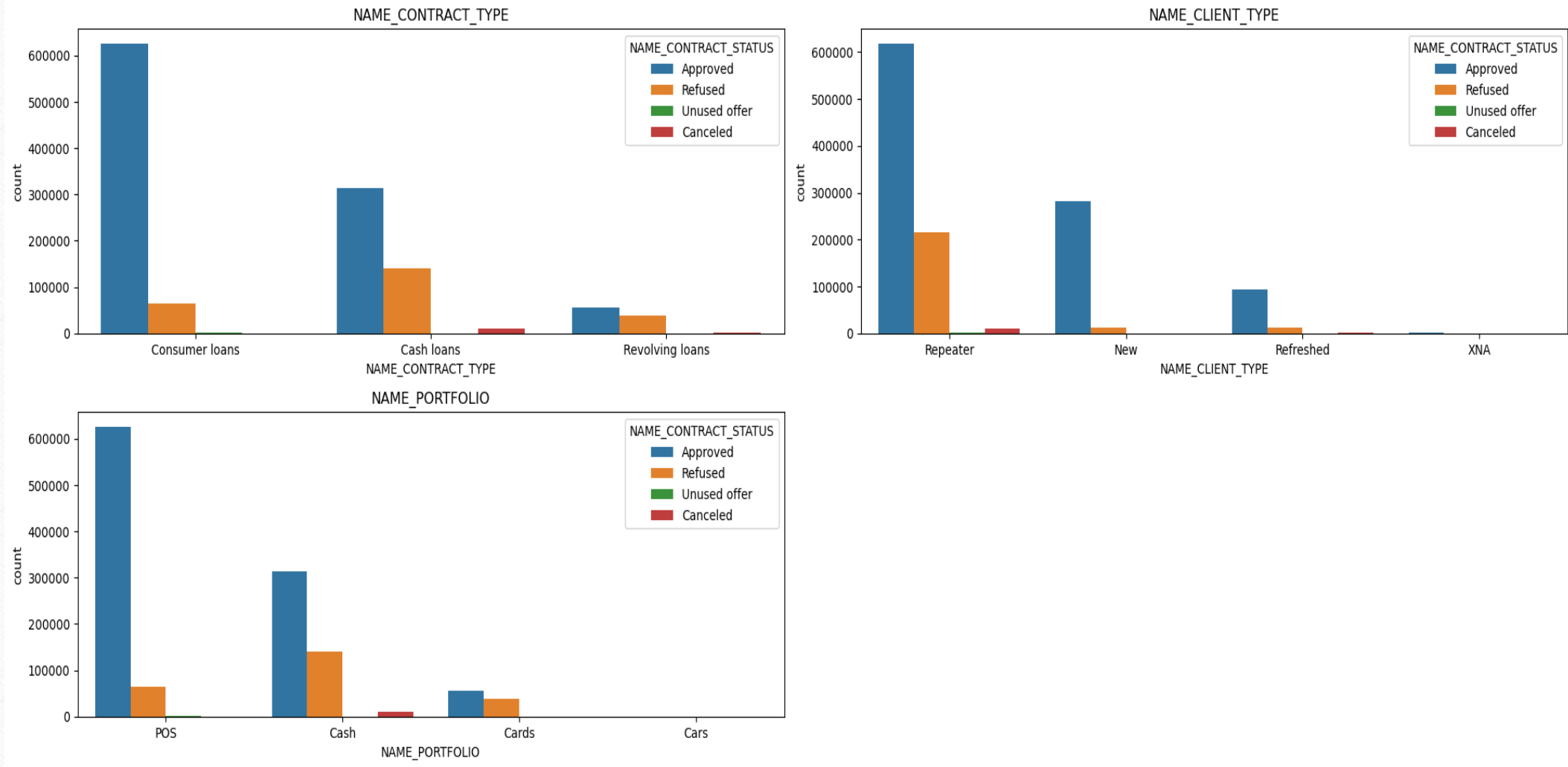
- Knowledge of data and sources.
- Examine the data for binning and quality problems.
- Examine the data for imbalance and run correlation, univariate, segmented univariate, and bivariate analyses.
- Combining application data from earlier iterations
- Univariate, segmented univariate, bivariate, and correlation data analysis
- Suggestions and Dangers

# Univariate Analysis of continuous data in Application data



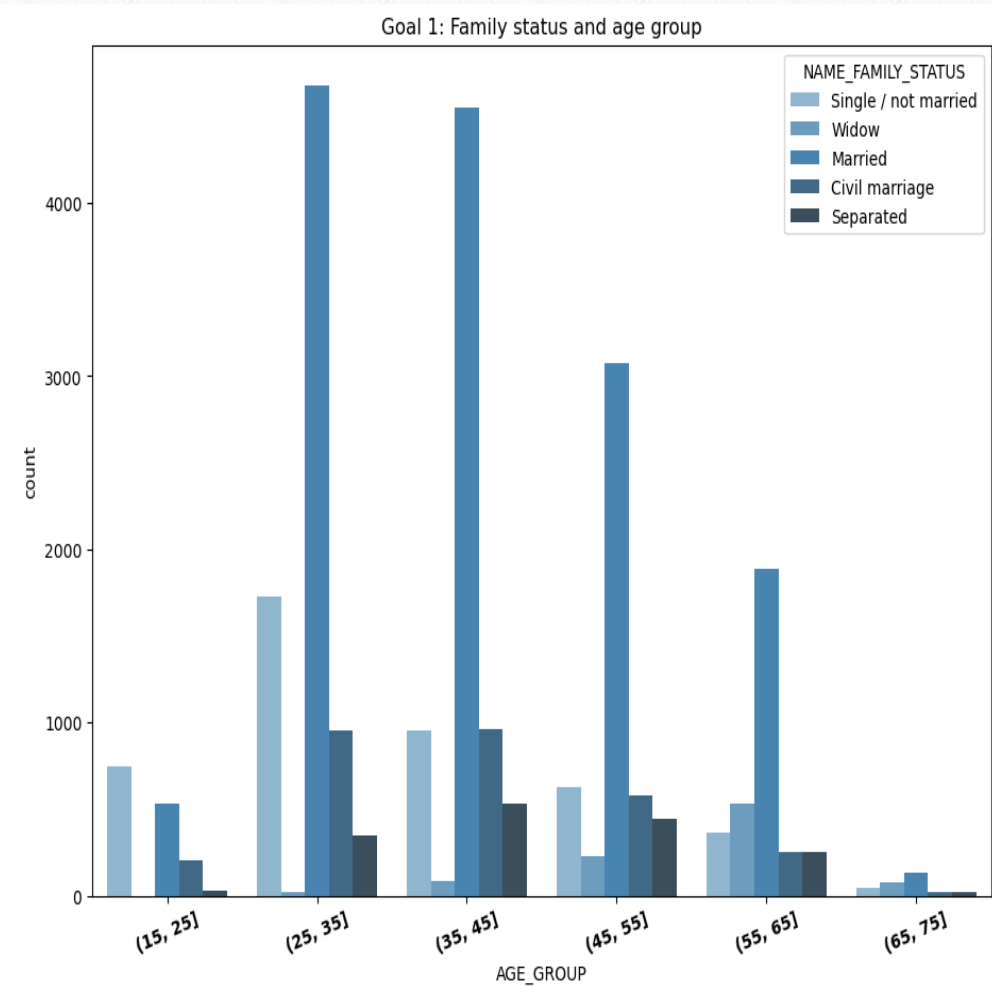
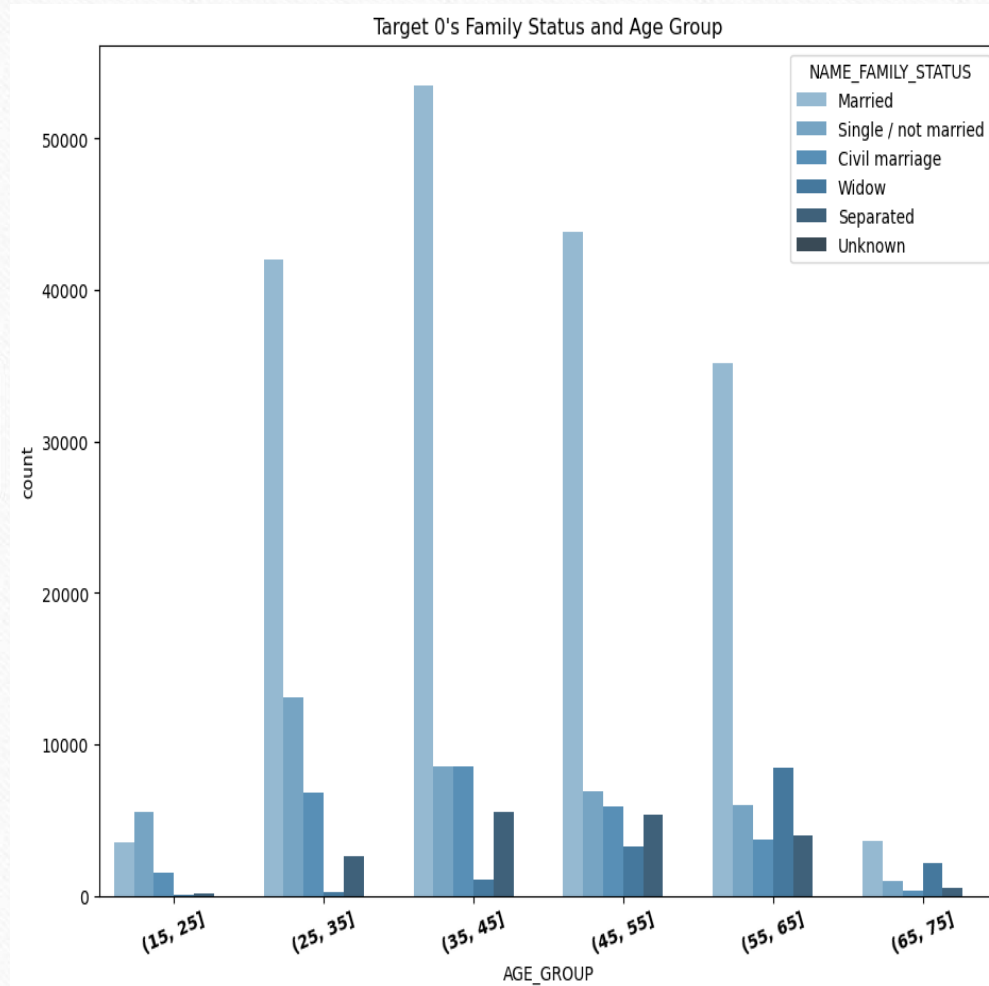
Around 29 years to 40 years people are more defaulters.  
There is high chance to be defaulted of the young people.  
Non-defaulted people are almost equally distributed

## Categorical and Categorical Variable

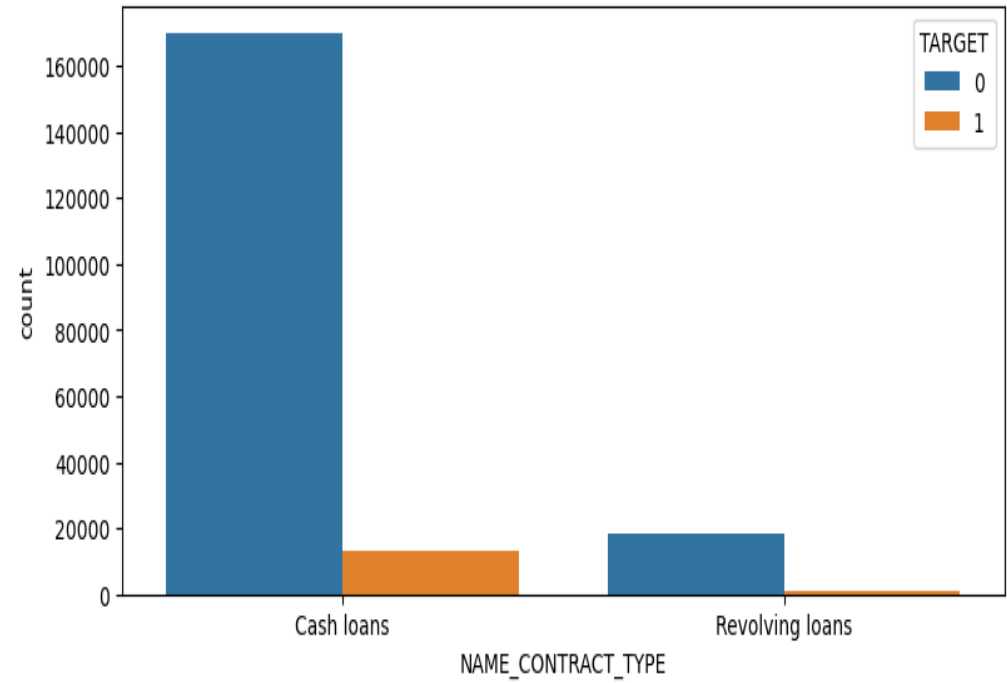
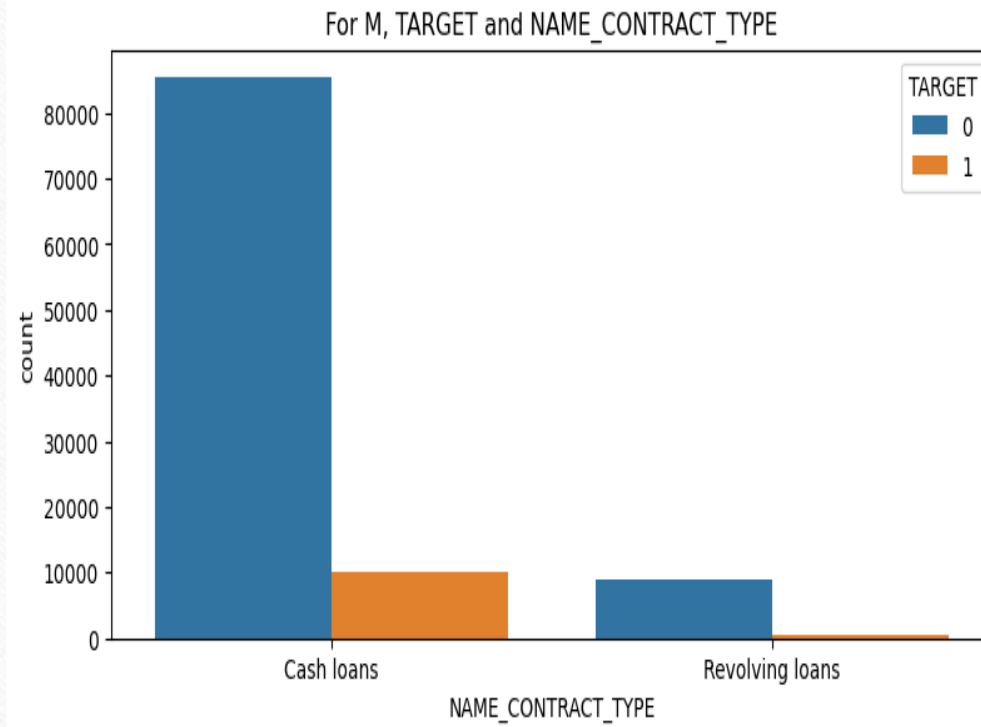




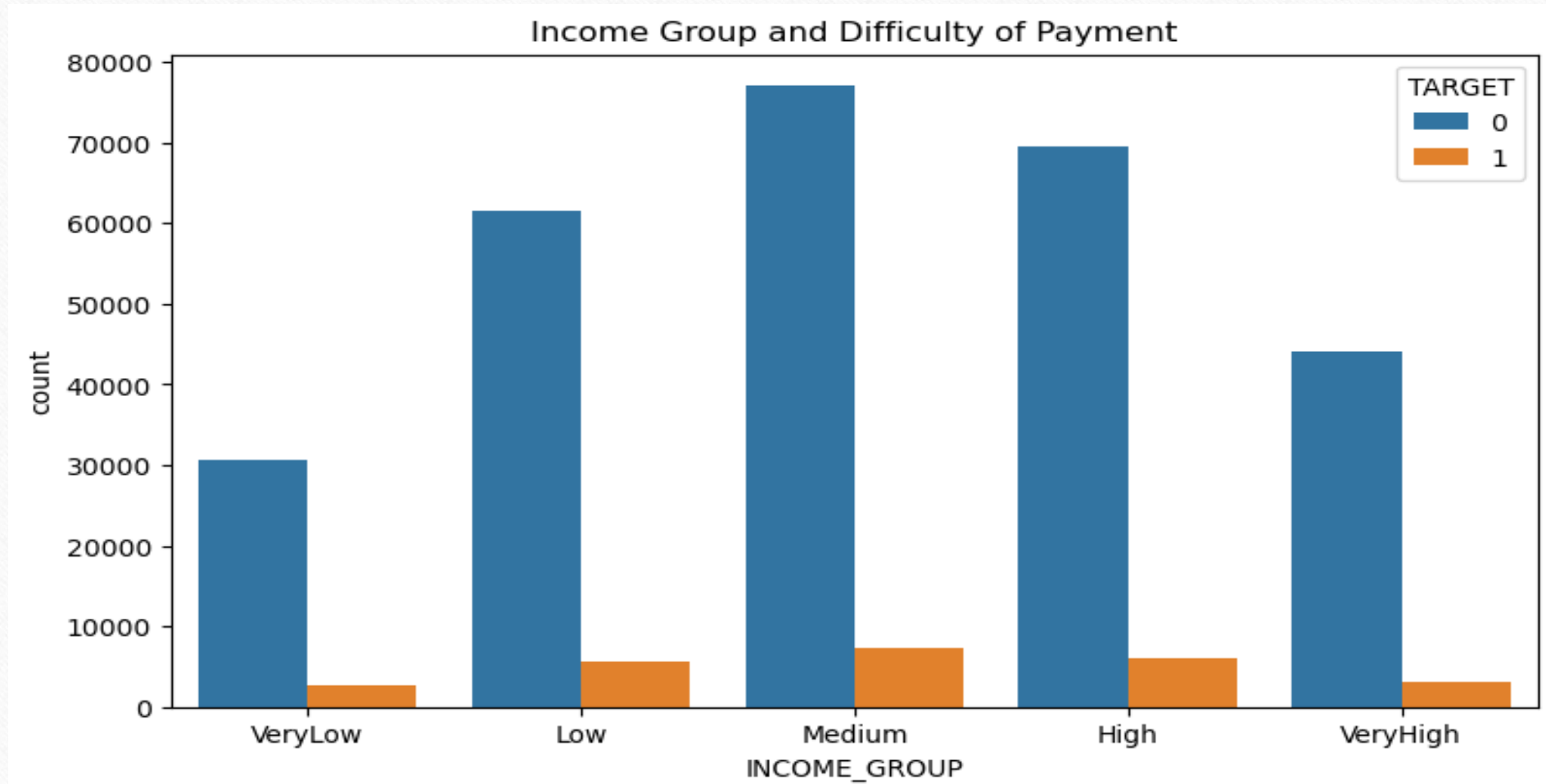
# Bivariate analysis



# Bivariate analysis

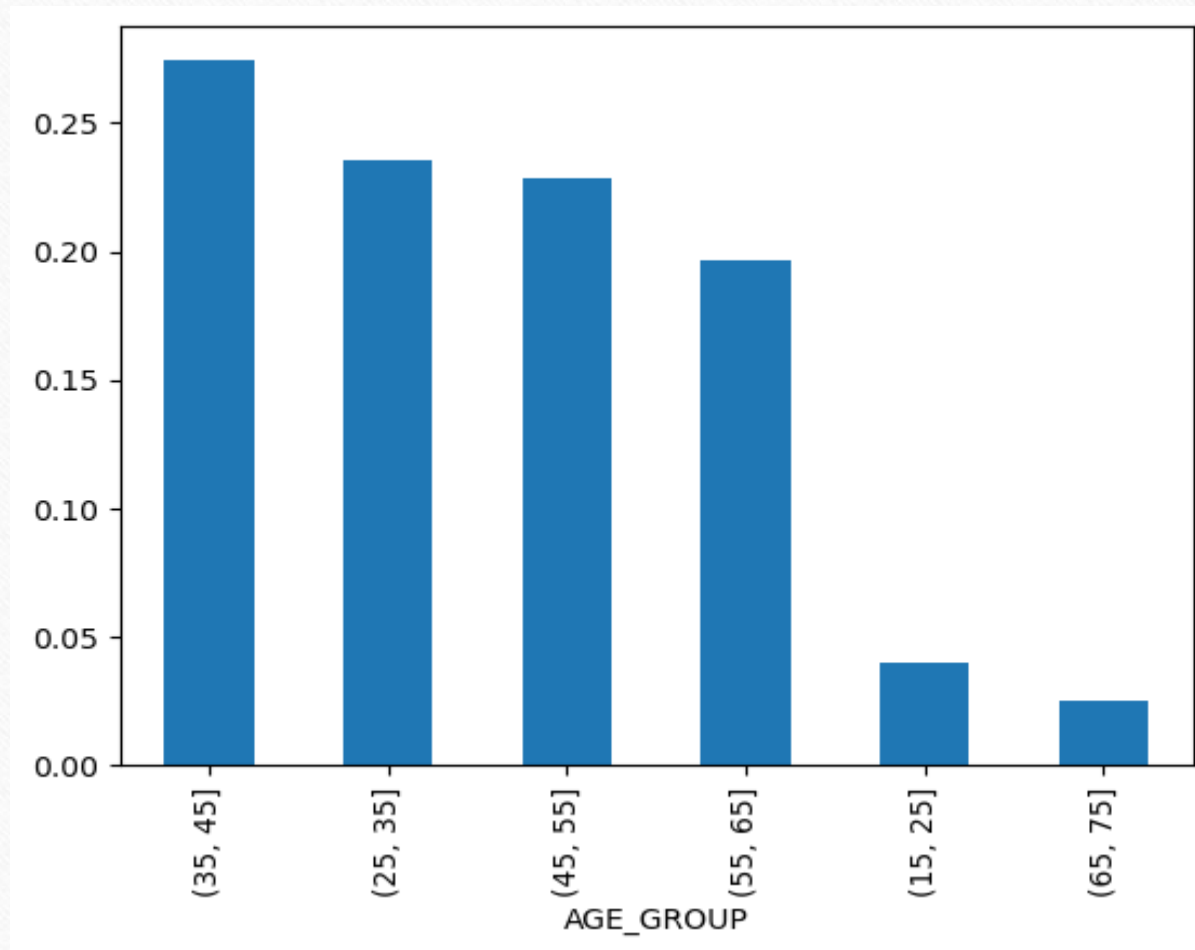


## Continuous and Categorical Variable



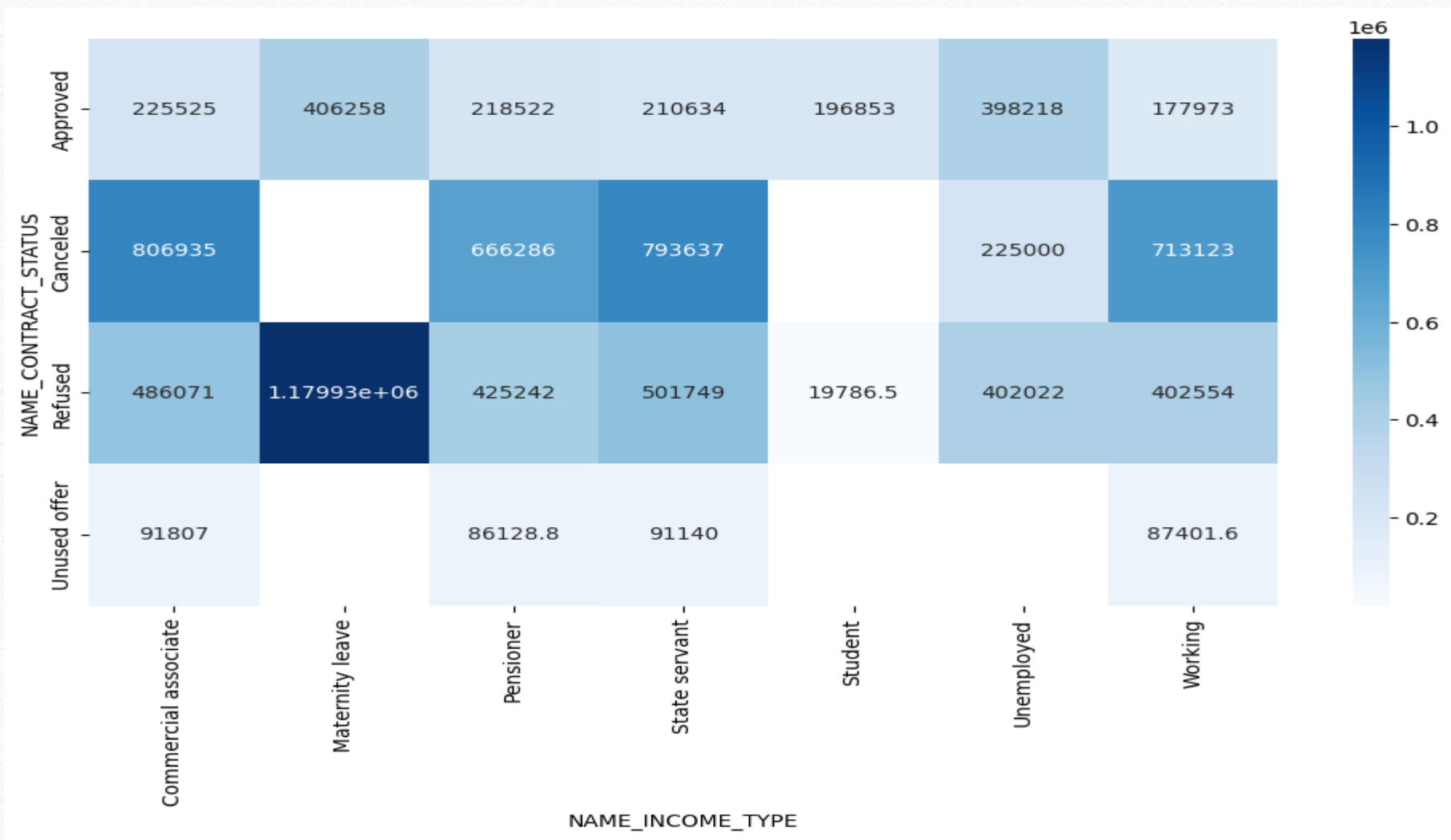


# Binning



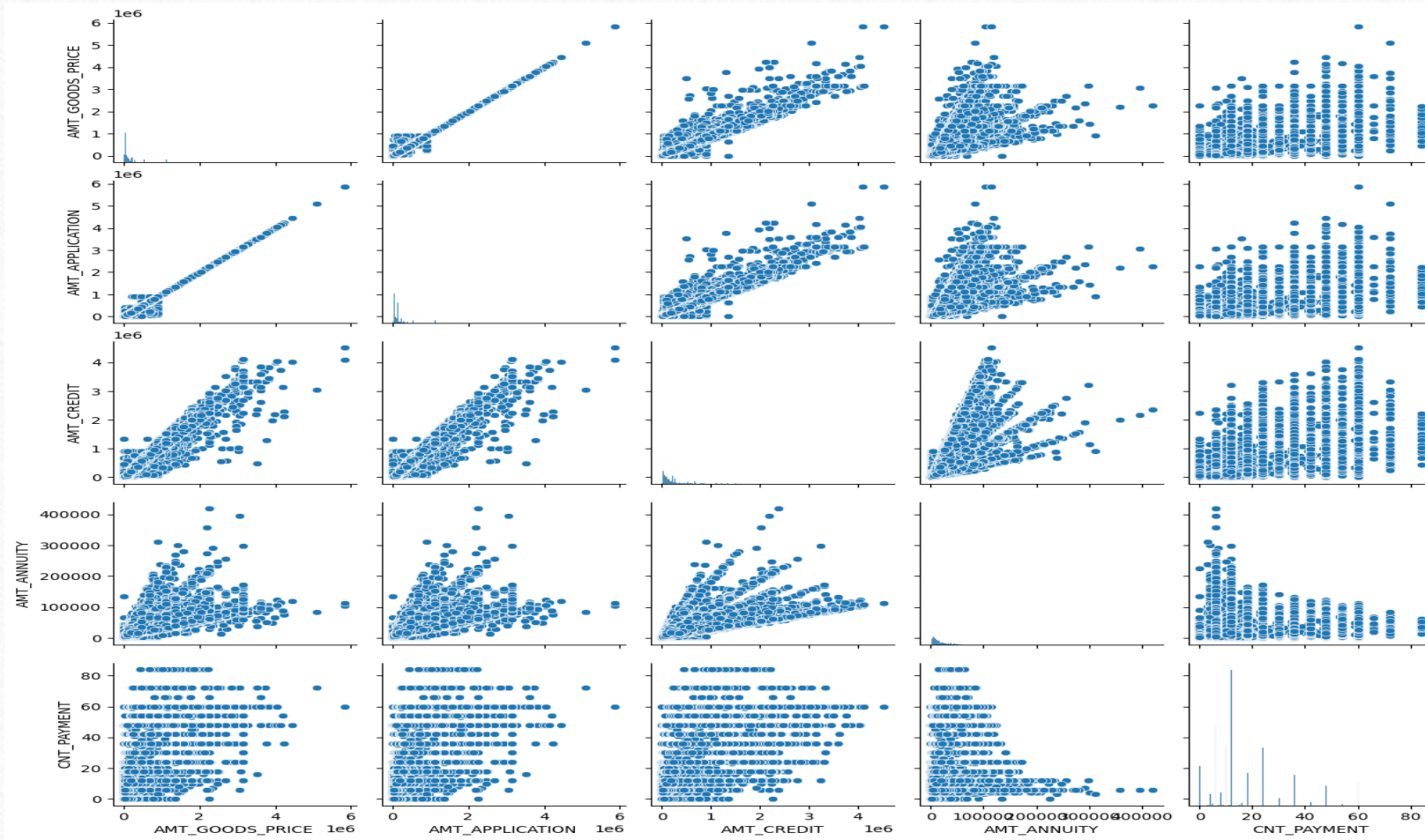
## Multiple correspondences:







## Highest Correlations



# Suggested organization to which credit can be applied.(reduced likelihood of default).

---

- Customers who serve as state employees.
- Older adults in any economic bracket.
- A client in the area of high income.
- A senior lady customer.
- A female client with a higher academic level.
- Any customer whose prior loan was authorized.
- A widow whose prior loan status was unused.
- Refreshed the loan status for a client who has previously unused it.

## ***Dangerous Collective:***

---

- Clients with lower secondary education levels are more likely to default on loans after having prior loans canceled or denied.
- Male clients wedded in a civil union.
- Group with a prior loan rejection status.