Credit EDA Case Study

-by **Rajendra Thammaneni**

PROBLEM STATEMENT & BUSINESS OBJECTIVE

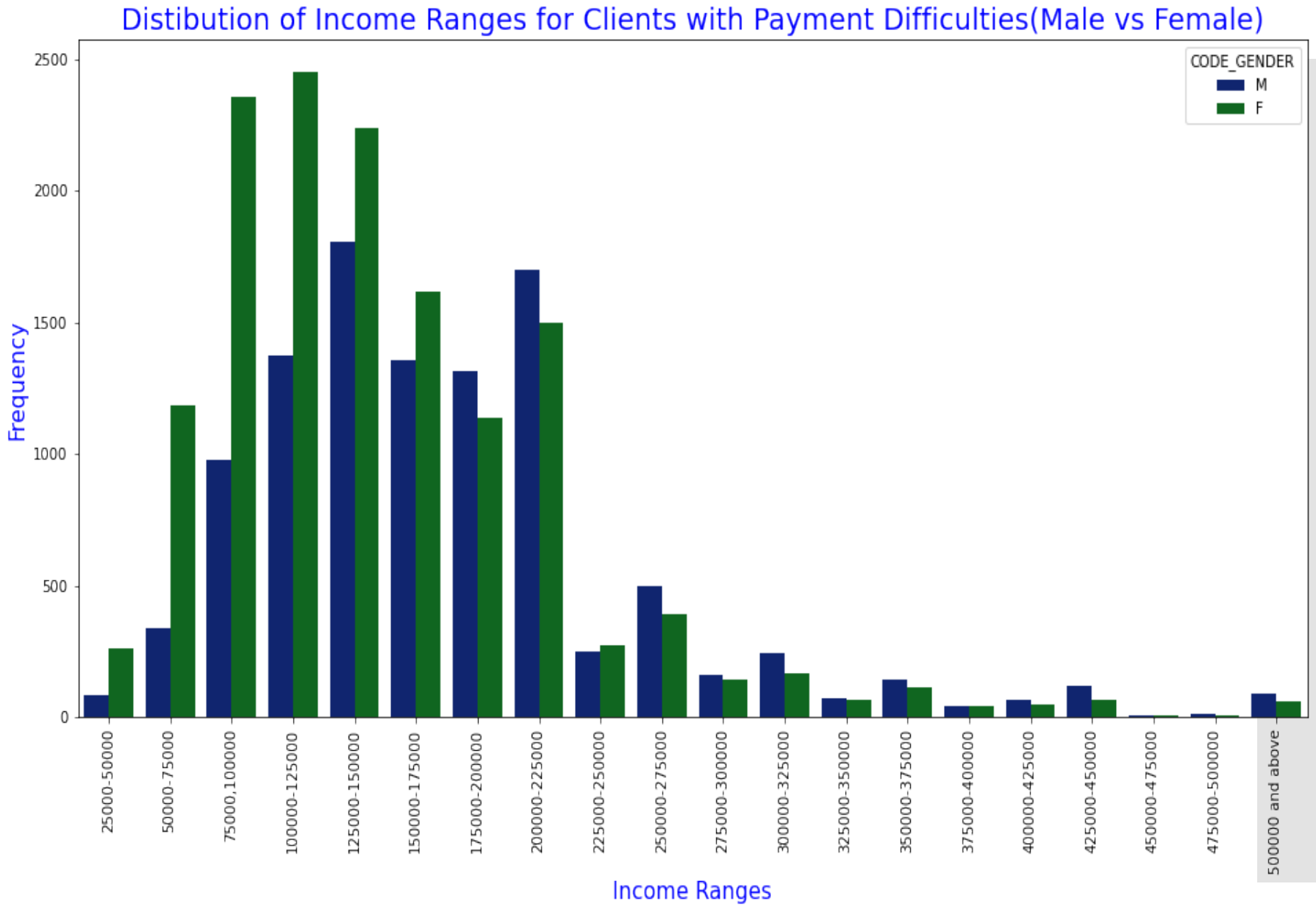
- The loan providing companies find it hard to give loans to the people due to their insufficient or non-existent credit history. Because of that, some consumers use it to their advantage by becoming a defaulter.
- This case study aims to identify patterns which indicate if a client has difficulty paying their instalments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.
- By conducting univariate, bi-variate and multivariate analysis on given data set, patterns and relationships between variables will be established

Analysis of clients facing payment difficulties(Target=1) class of data

Analysis of Distribution of Income ranges for clients who face payment difficulties

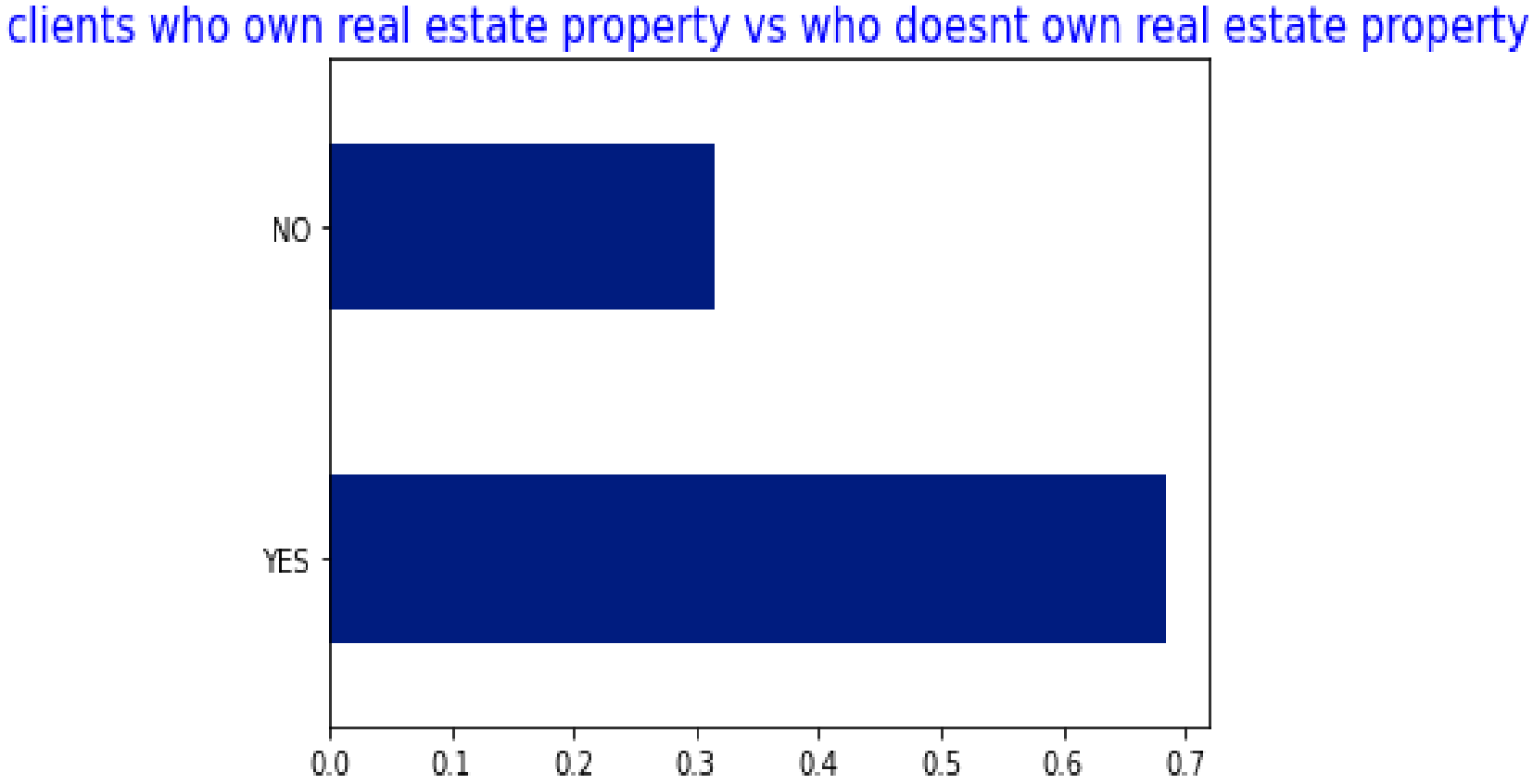
More Females with income ranging from 25000 to 175000 are facing payment difficulties compared to Males.

For income ranges above 175000, More Male clients are facing payment difficulties than Female clients.



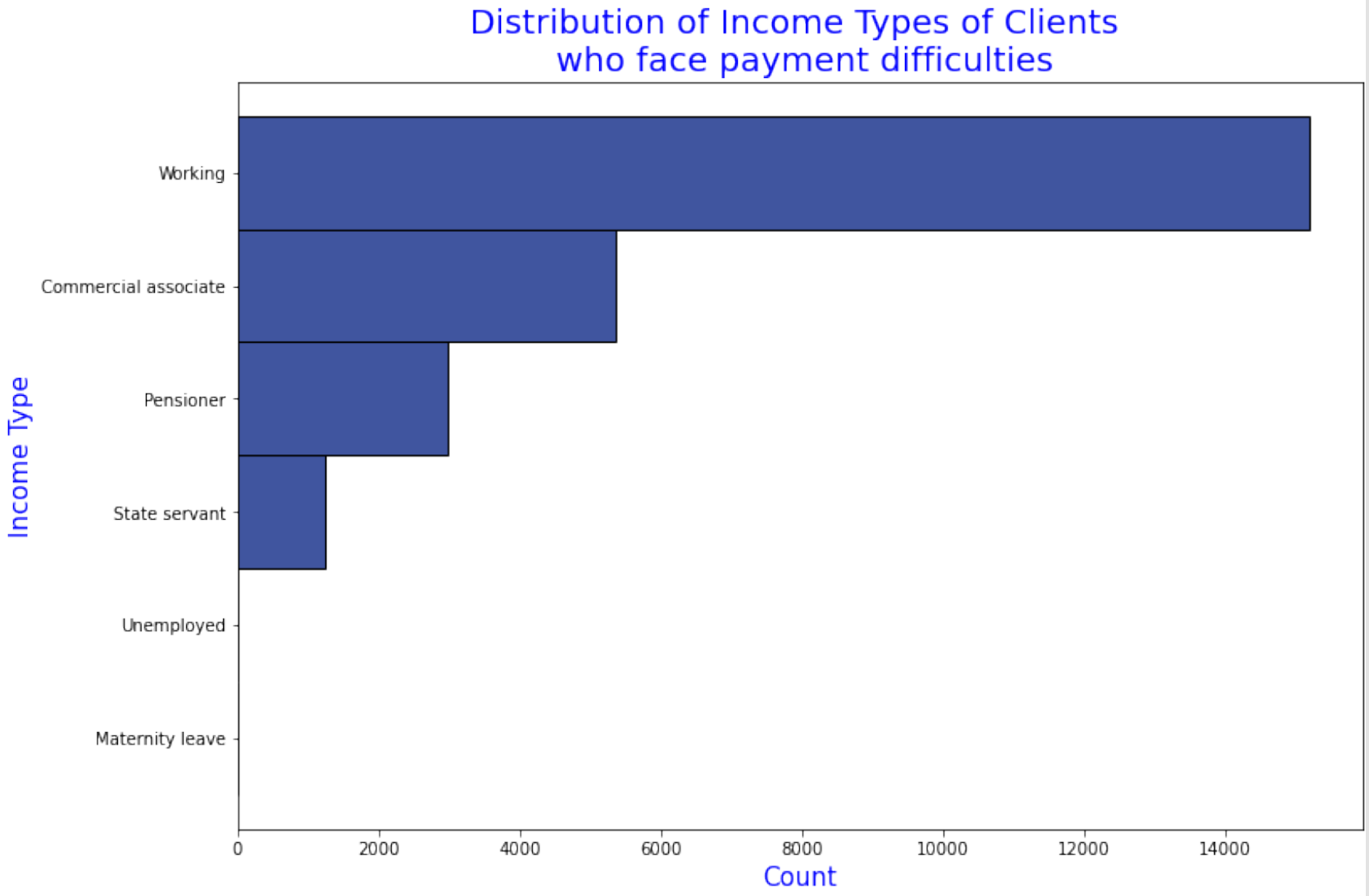
Univariate analysis on clients who own real estate property vs who doesn't own real estate

clients who own real estate property are more prone to payment difficulties compared to who doesn't own a real estate property.



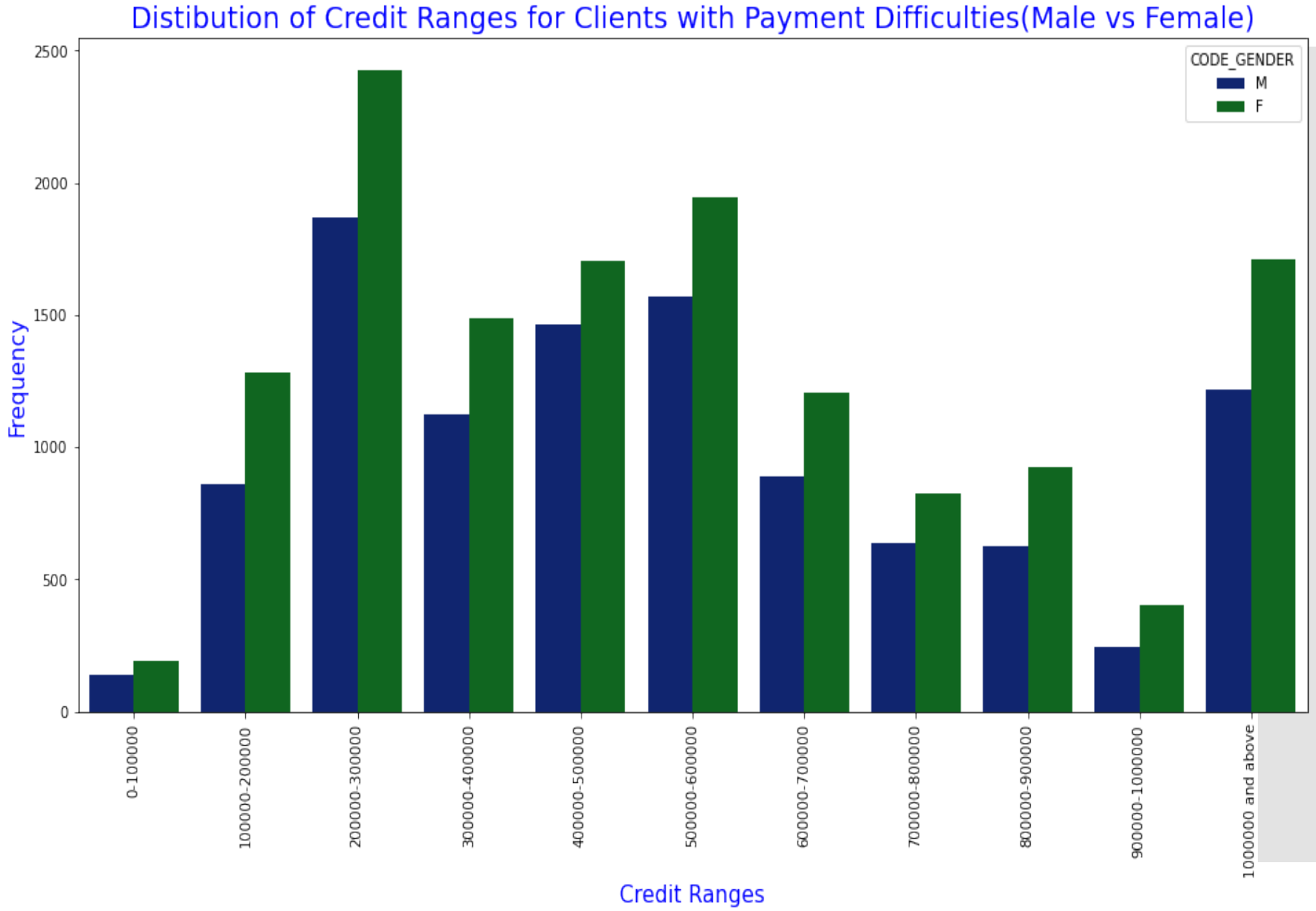
Univariate analysis on Income Types categorical variable.

This distribution indicates, those clients whose income type is 'working' face higher difficulties in payment among other income types.



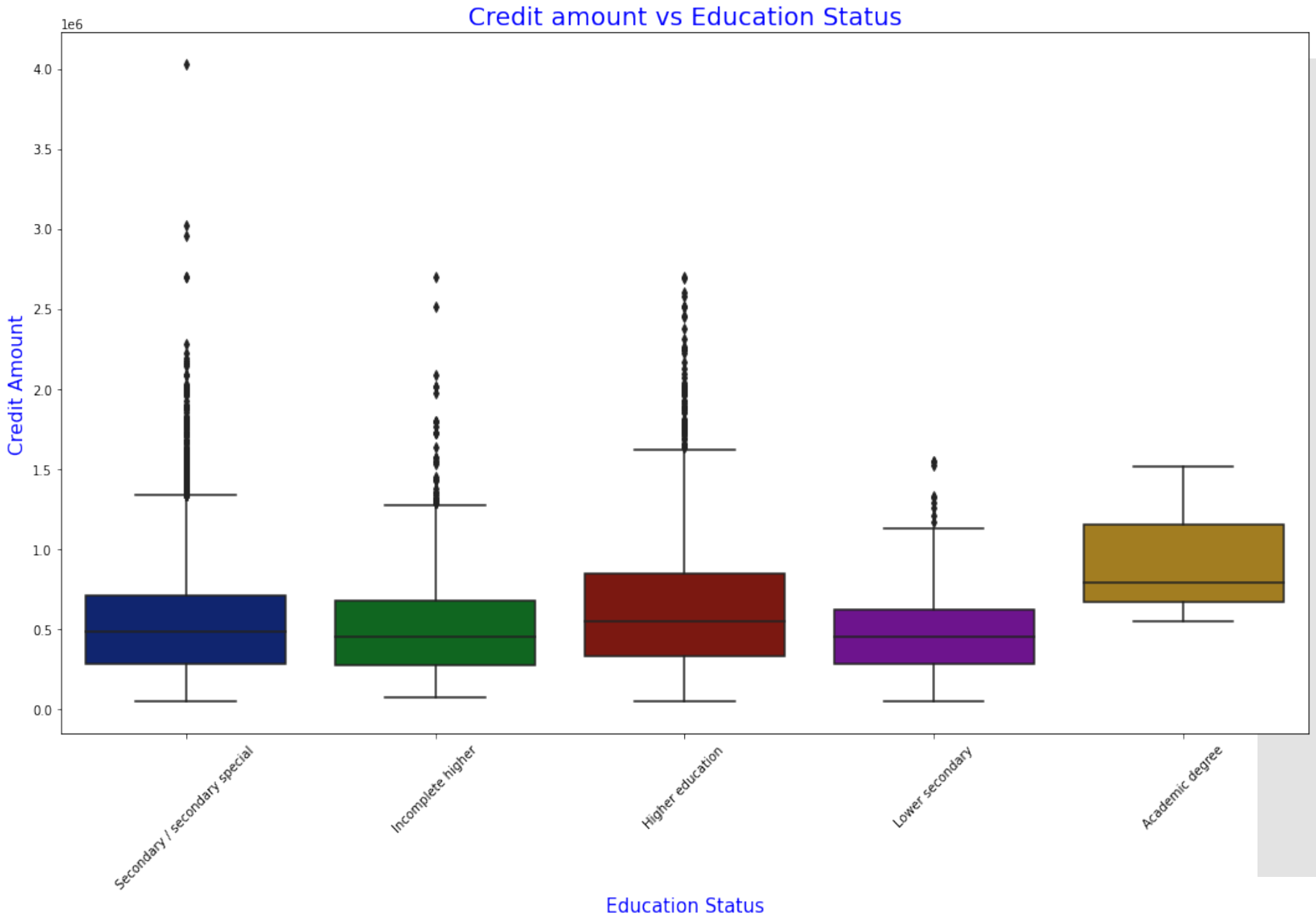
Analysis of Credit ranges of Clients facing payment difficulties(MALE vs FEMALE)

In all Credit ranges , more female clients are facing payment difficulties compared to male clients



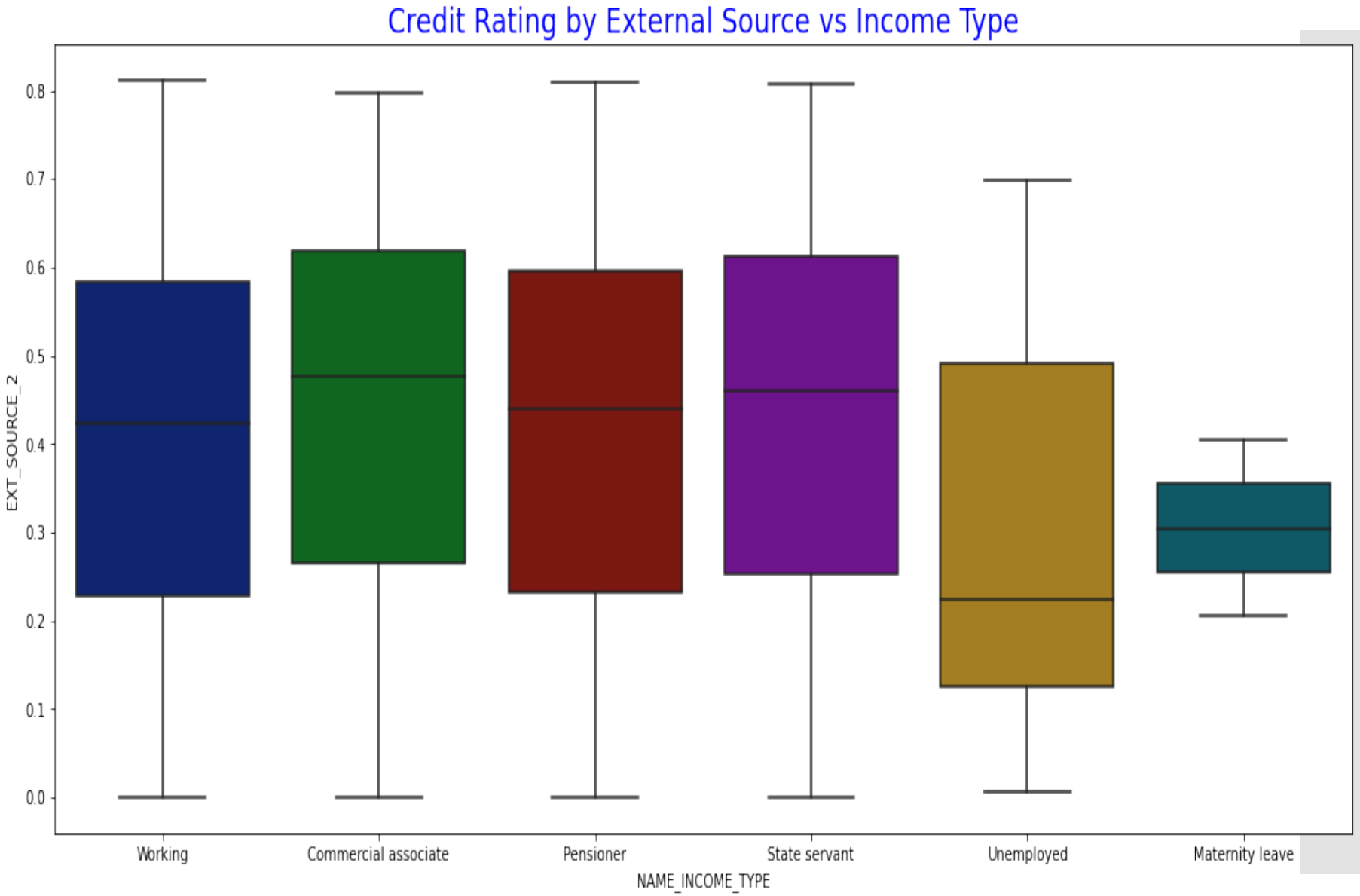
Bivariate analysis of Credit amount and Education Status variables

From above Boxplot analysis, Median Credit amount is higher for clients with academic degree who are facing payment difficulties and there are more clients above median credit amount in this category facing payment difficulties.



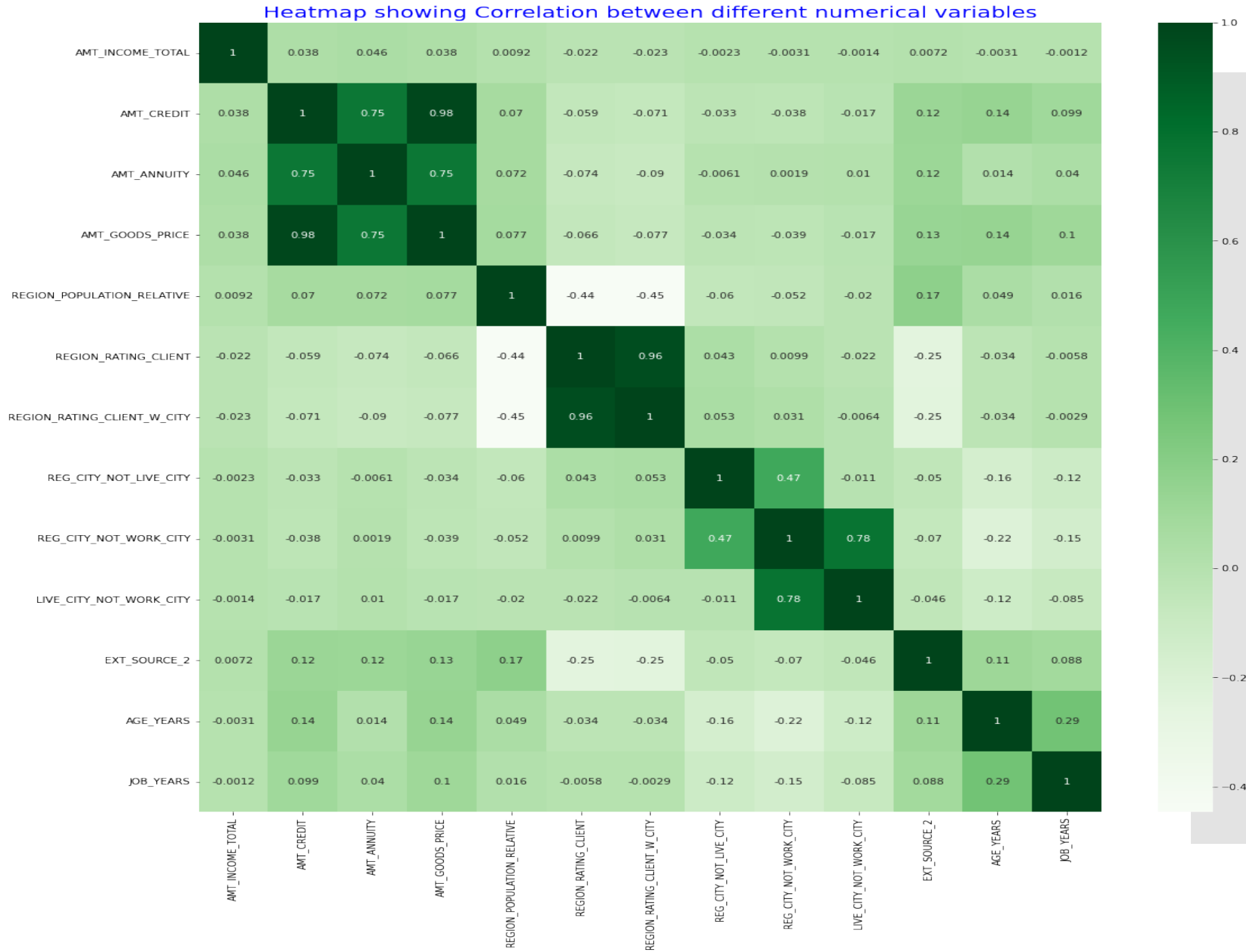
Credit Rating vs Income Type

Among the clients facing payment difficulties, Unemployed category of clients have lower median credit ratings among all other categories.



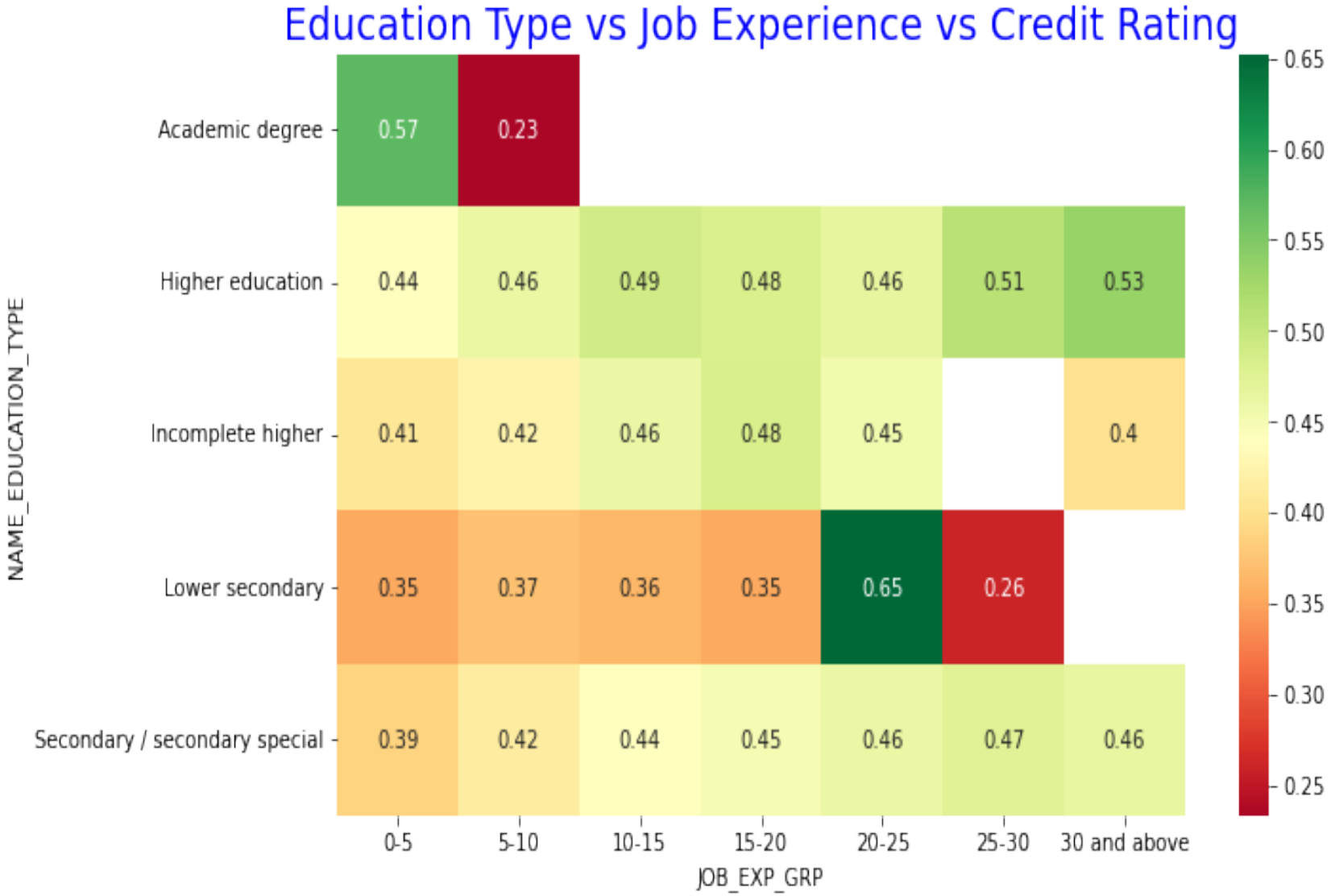
Heat map showing correlation between different numerical variables for clients facing payment difficulties

There is strong positive correlation between annuity amount and credit amount. Also between annuity amount and Goods price amount, there is high positive correlation. This correlation matrix shows the top correlations between different variables.



Heat map for Education type vs job experience vs Credit rating of clients facing payment difficulties

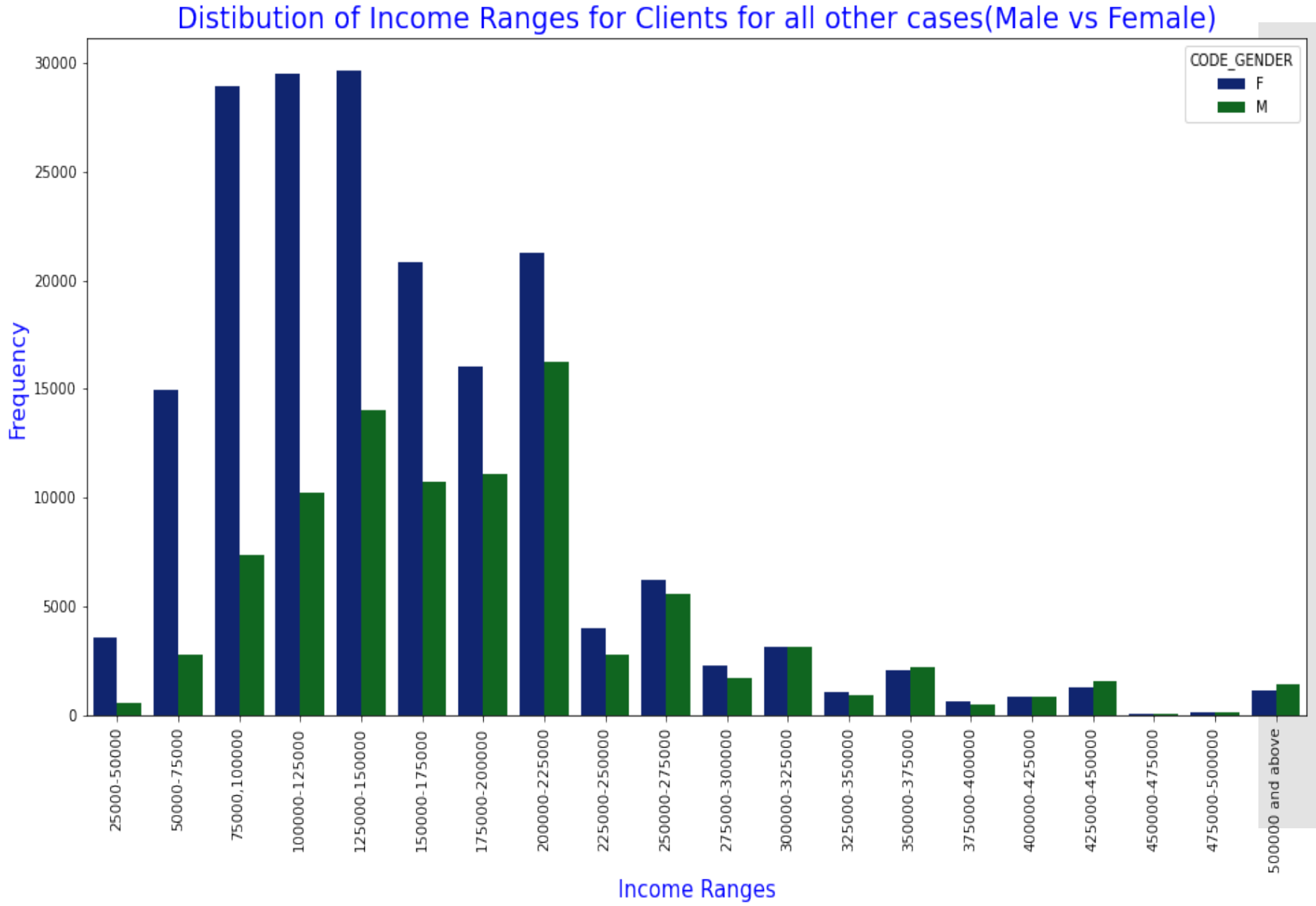
Clients with lower secondary education as a group tend to have lower than median credit ratings compared to all other categories by education type



Analysis of all other cases data class(Target=0 class)

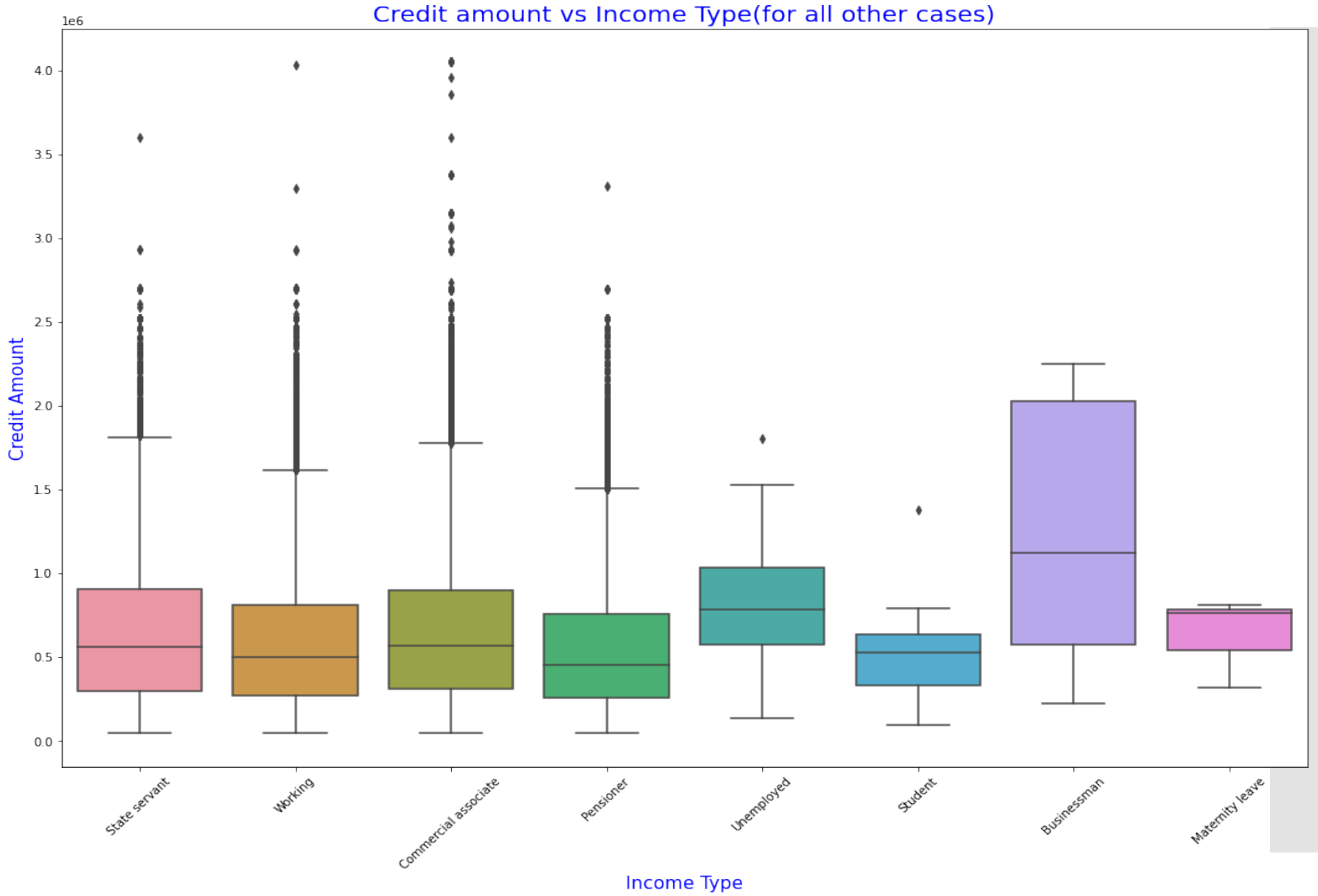
Histogram for income ranges of client(all other cases)

From this distribution , it is understood that females count is significantly higher in lower income groups ranging from 25000 to 225000 who are not missing payments.



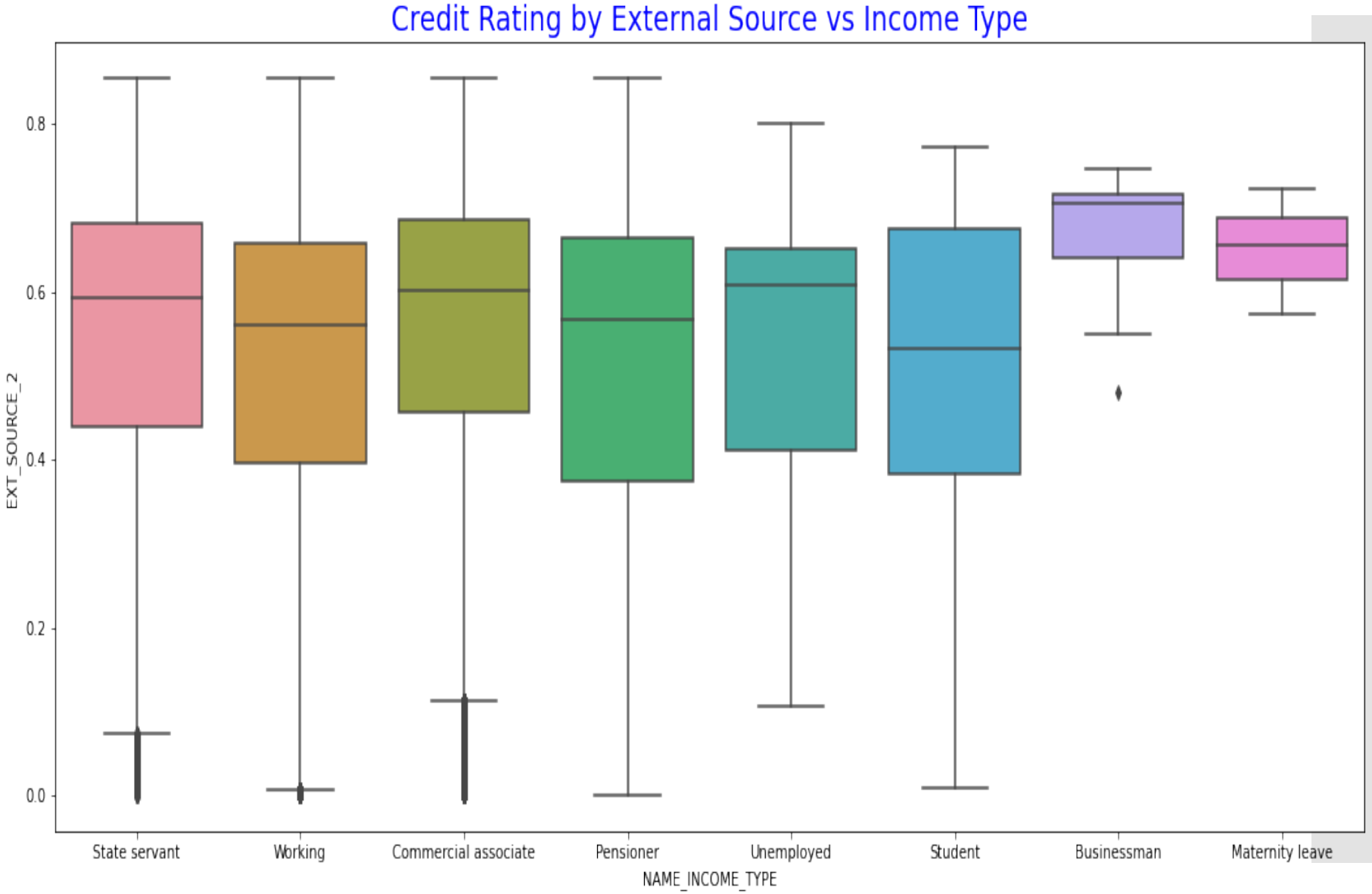
Distribution of credit amount for different income types(for all other cases)

Businessman category has higher median credit amount with large inter quartile range compared to all other categories



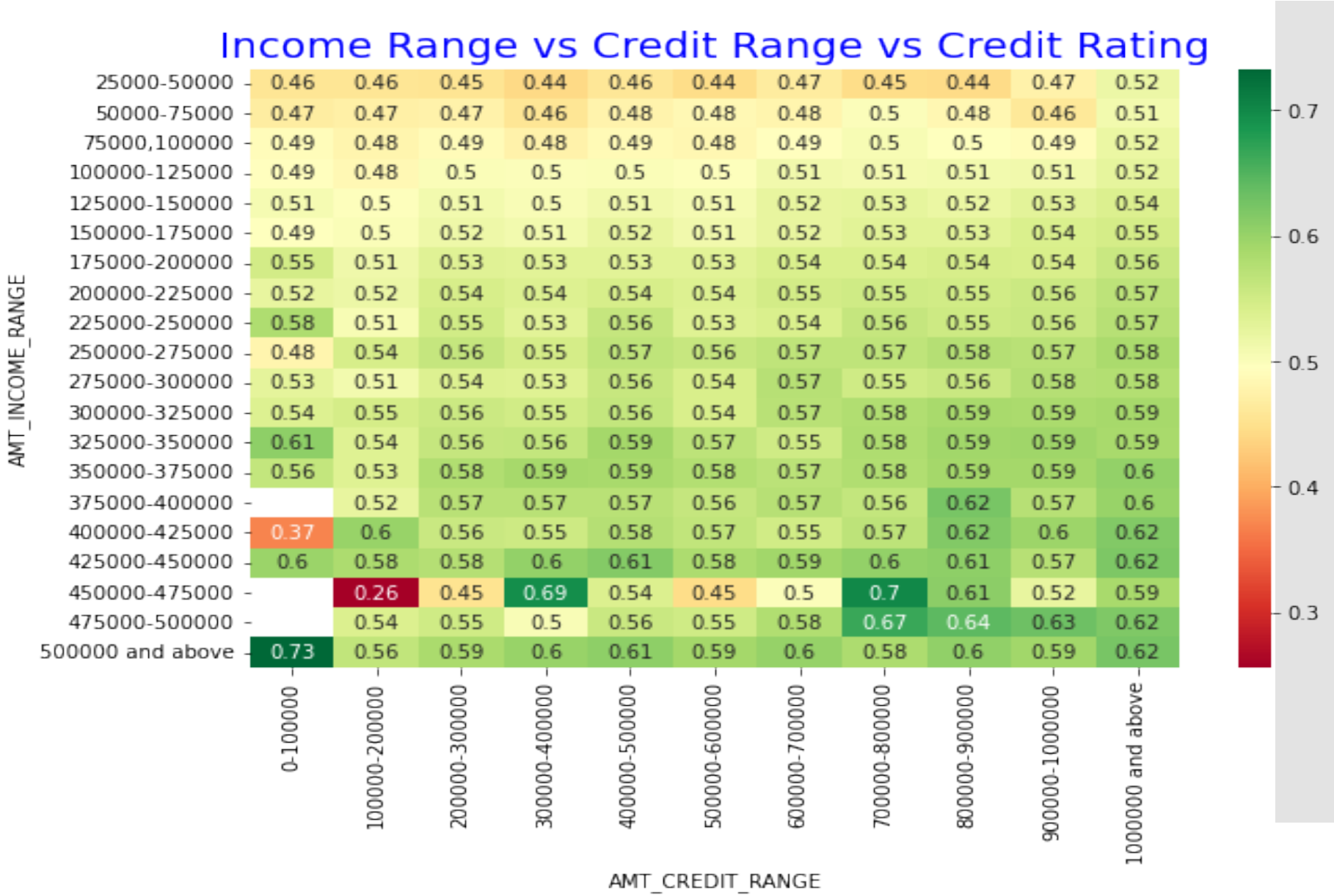
Box plot for credit rating vs income type(for all other cases)

Businessman category of clients has higher median credit ratings compared to other categories who are not facing any payment difficulties



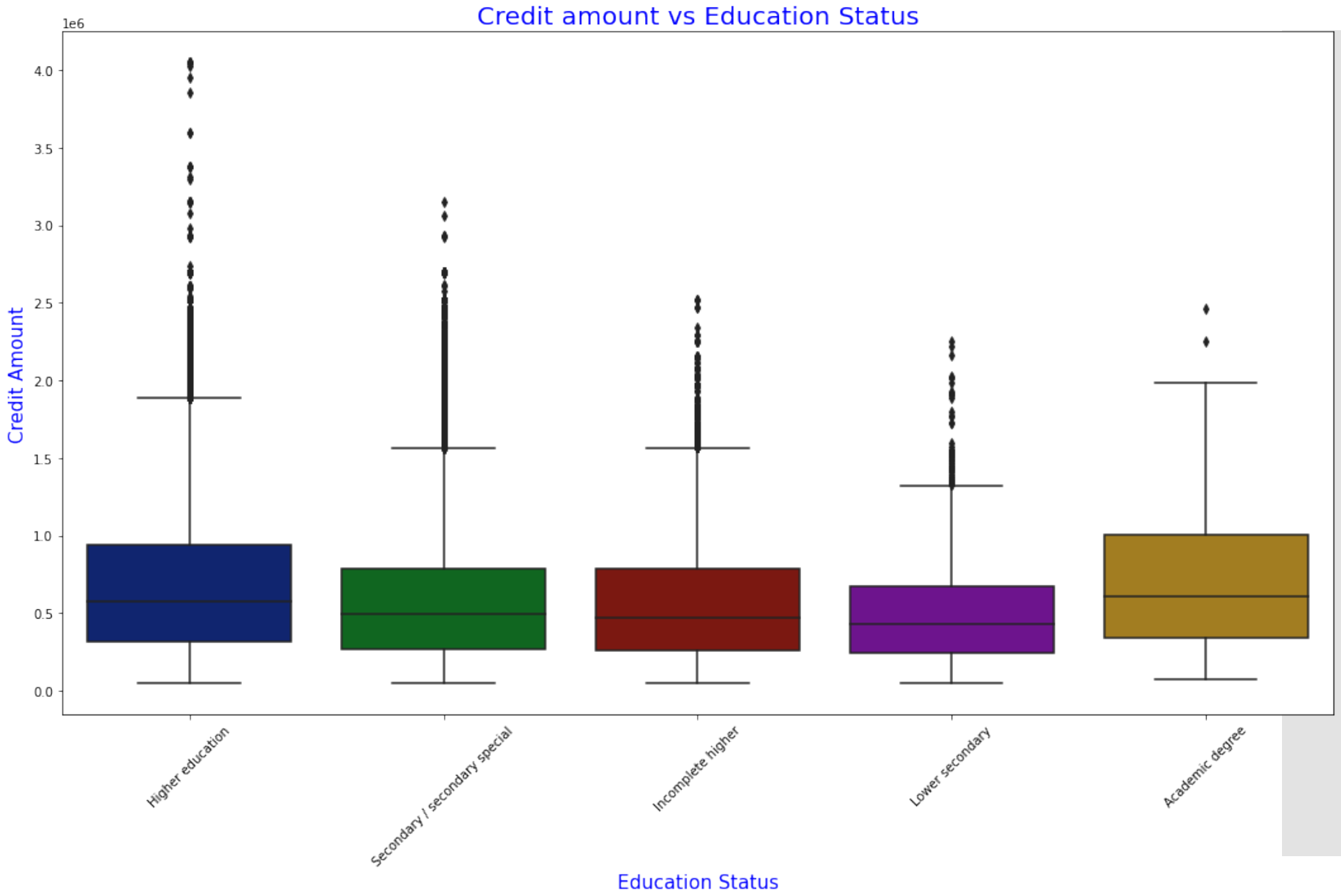
Heat map for Income Range vs Credit Range vs Credit Rating of client(for all other cases)

Clients in all credit ranges, tends to have higher credit ratings in the income ranges above 175000, among the clients who don't miss any payments



Box plot showing credit amount vs education status(for all other cases)

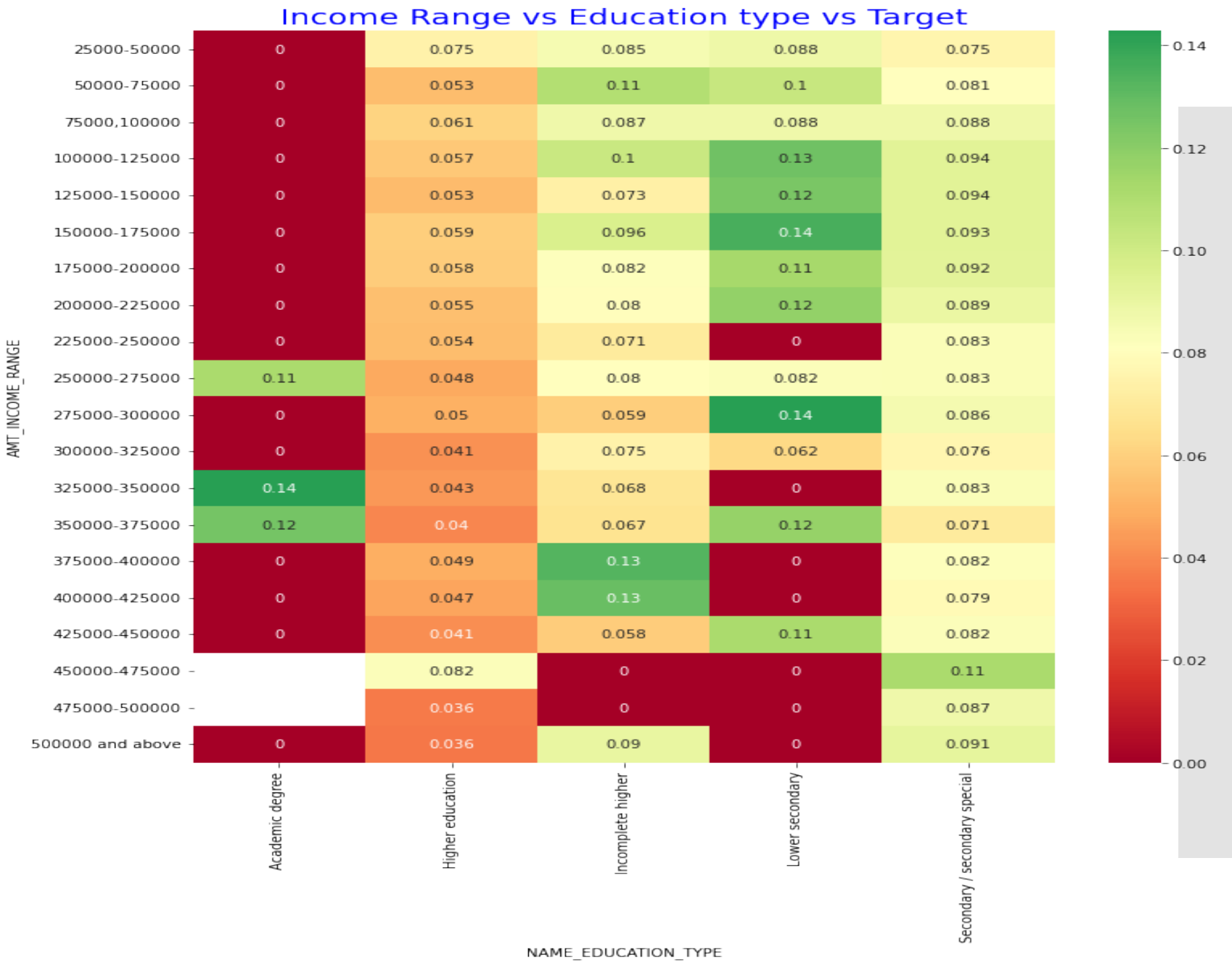
Academic degree holders and clients with Higher education have higher median credit amount compared to other categories.



Analysis of combined dataset(Target=1 and Target=0 classes)

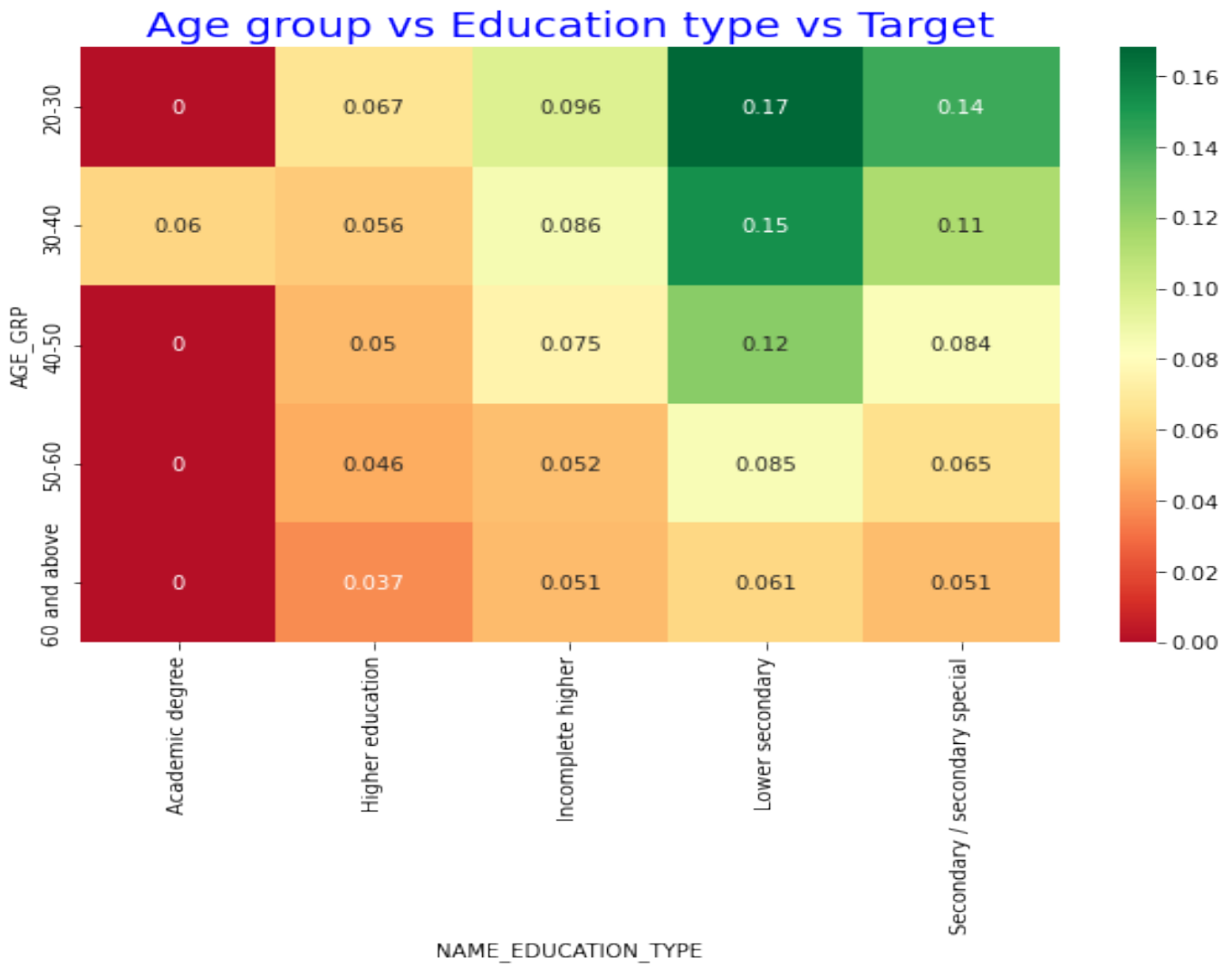
Heat map income ranges , education type with aggregation of Target variable

This heat map shows the clients with lower secondary education as a category are more likely to miss their payments compared to other categories



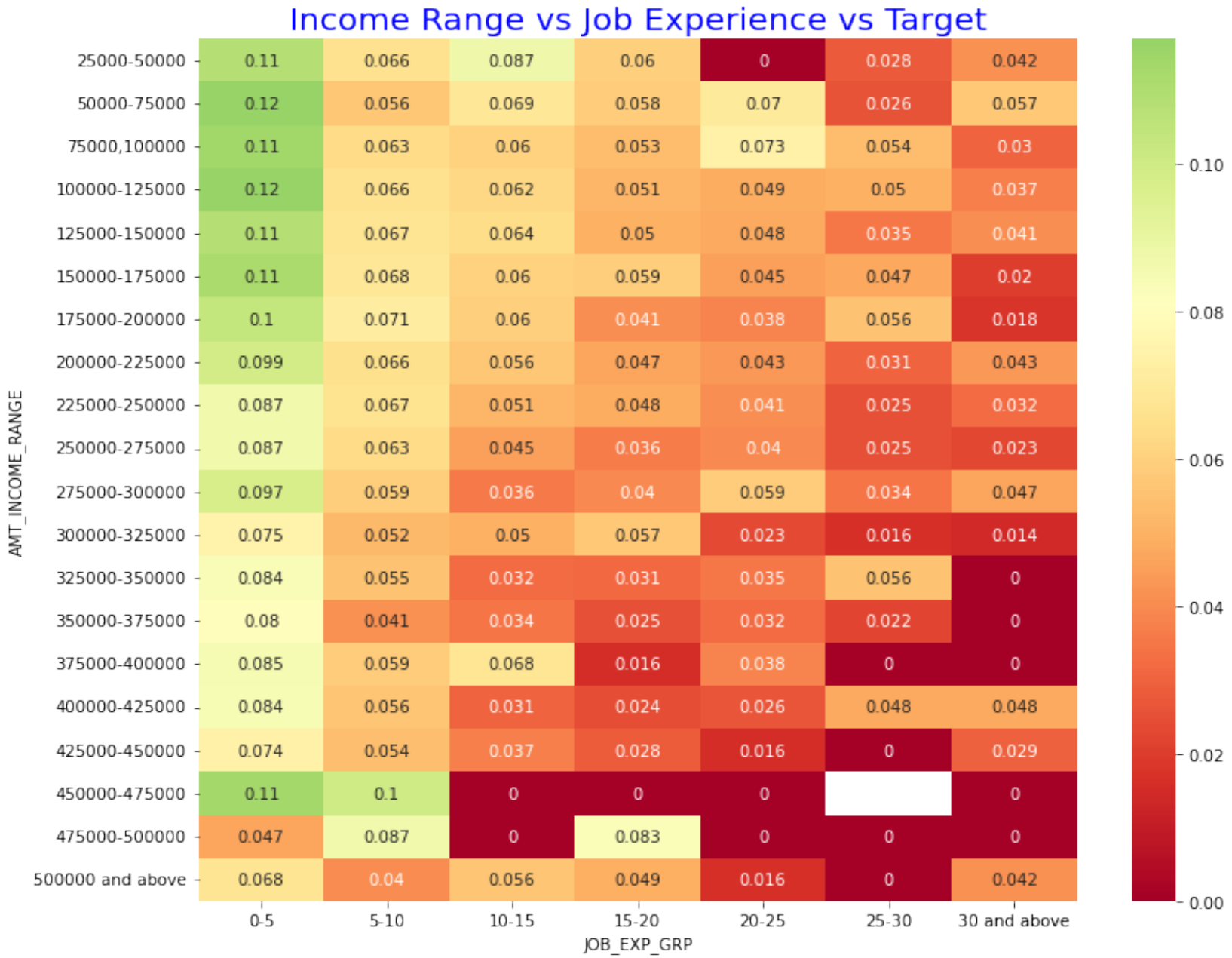
Heat map of Income ranges , Education type with aggregation of mean of Target variable

Clients with lower secondary education and in age groups of 20-30 and 30-40 are more prone to payment difficulties



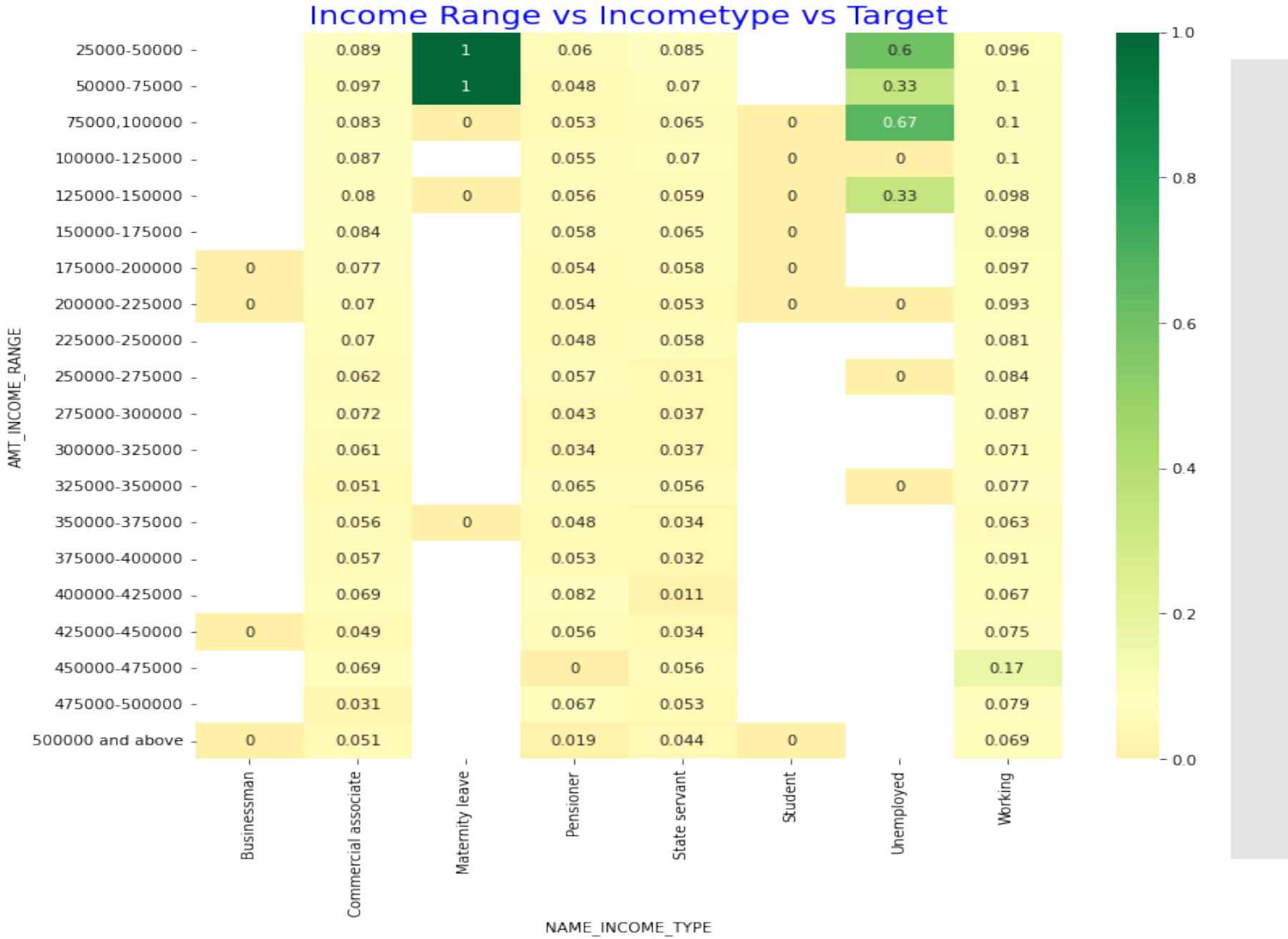
Heat map for income ranges vs job experience vs Target variable

Clients with job experience of less than 5 years are more likely to face payment difficulties compared to clients with more than 5 years experience



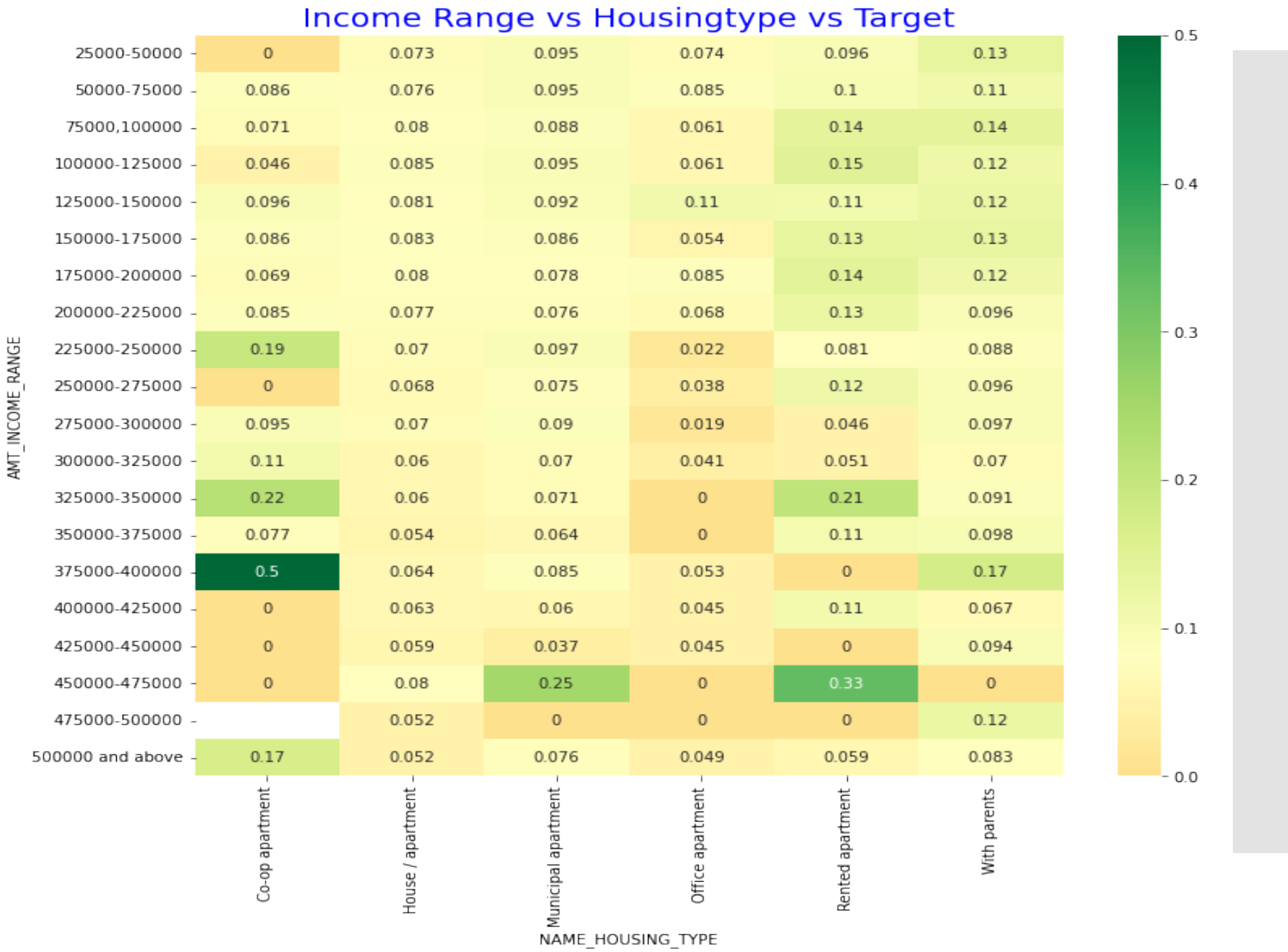
Heat map for income ranges vs job experience vs Target variables

Unemployed clients with lower income ranges up to 150000 are facing payment difficulties. Client with maternity leave category also facing payment difficulty.



Heat map for income ranges vs Housing type vs Target variables

Clients living in rented apartment and living with parents facing payment difficulty compared to clients living in other category of housing types



Observations from the Analysis

- Clients with lower secondary education as a category are more likely to miss their payments compared to other categories.
- Within the lower secondary education group, clients in age groups of 20-30 and 30-40 are more prone to payment difficulties.
- Clients with job experience of less than 5 years are more likely to face payment difficulties compared to clients with more than 5 years experience.
- Unemployed clients with lower income ranges up to 150000 are facing payment difficulties. Client with maternity leave category also facing payment difficulty.
- Unemployed clients in higher credit ranges of 300000 to 800000 as a category are facing payment difficulties.
- Working clients in credit ranges of 100000 to 800000 are likely to payment difficulties.
- Businessman and student category clients are not facing any payment difficulties.

Observations from the Analysis

- Clients who own real estate property are most prone to payment difficulties compared to who doesn't own a real estate property.
- Median Credit amount is higher for clients with academic degree who are facing payment difficulties and there are more clients above median credit amount in this category facing payment difficulties
- Among the clients facing payment difficulties, Clients with lower secondary education as a group tend to have lower than median credit ratings compared to all other categories by education type
- Clients living in rented apartment and living with parents facing payment difficulty compared to clients living in other category of housing types.
- Businessman and student category of income type have less payment difficulties.
- In all Credit ranges , more female clients are facing payment difficulties compared to male clients.