

Gold Atlantis: Repaying the Credit

MENTOR – JAYA PANDEY

NAME – SAKSHI

INTRODUCTION

In this Case Study, I will be presenting my analysis of the credit data of Gold Atlantis, aimed at identifying patterns in the data and making predictions on the likelihood of a customer becoming a defaulter.

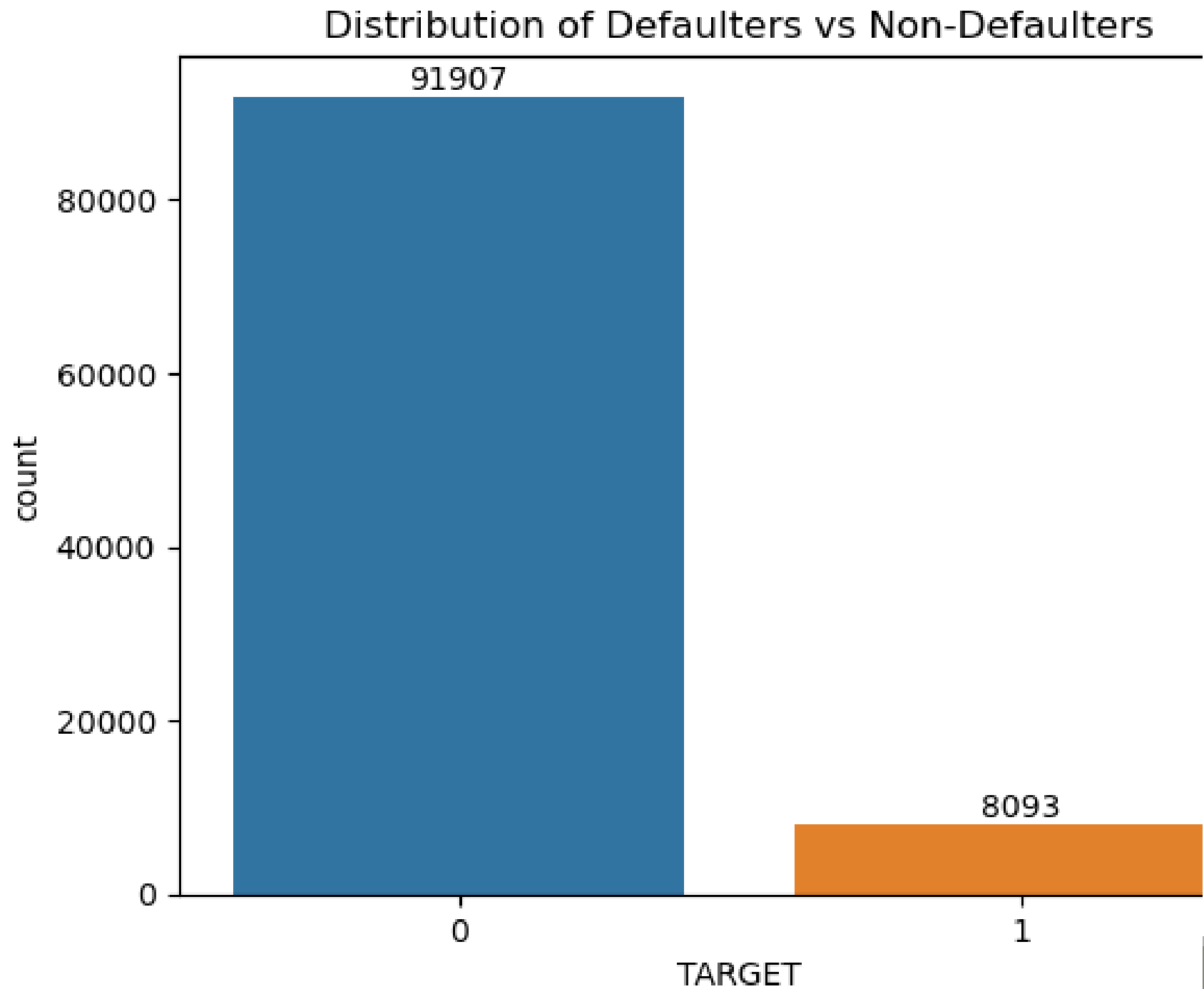
I conducted extensive data analysis using various statistical techniques such as univariate, bivariate, and multivariate analysis to identify patterns and develop insights. The goal of this analysis is to help the firm identify customers who have a lower probability of becoming a defaulter.



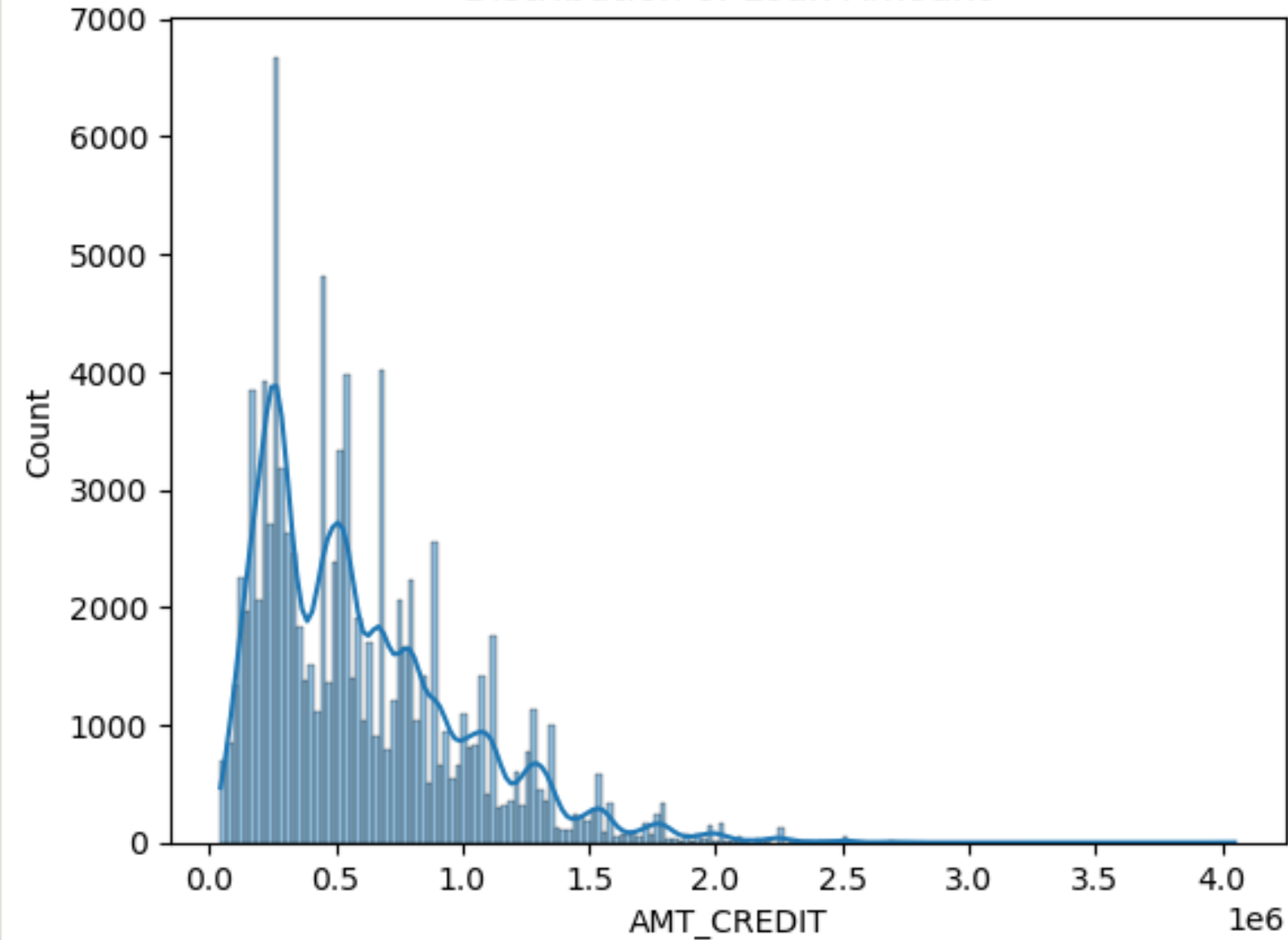
Univariate Analysis:

Interpretation

We can see that the number of non-defaulters is much higher than the number of defaulters, indicating that the majority of applicants have repaid their loans.



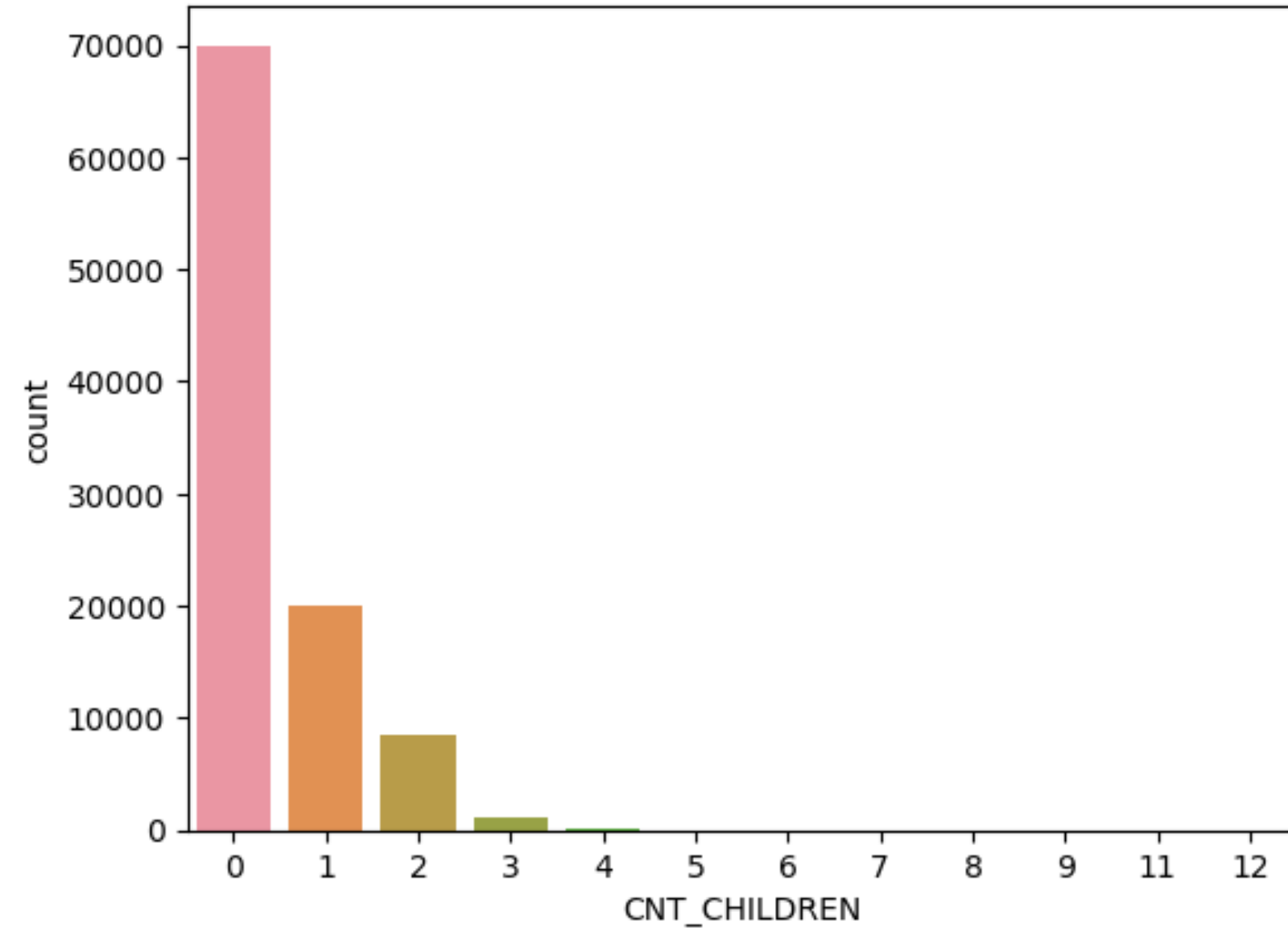
Distribution of Loan Amount



Interpretation

- * The distribution of the loan amount is right-skewed, meaning that the majority of the loans are smaller and fewer loans are larger.
- * There are some extreme loan amounts, which can be seen from the long tail on the right side of the plot. These are likely to be the larger loans, which are few in number but higher in amount than the majority of the loans.

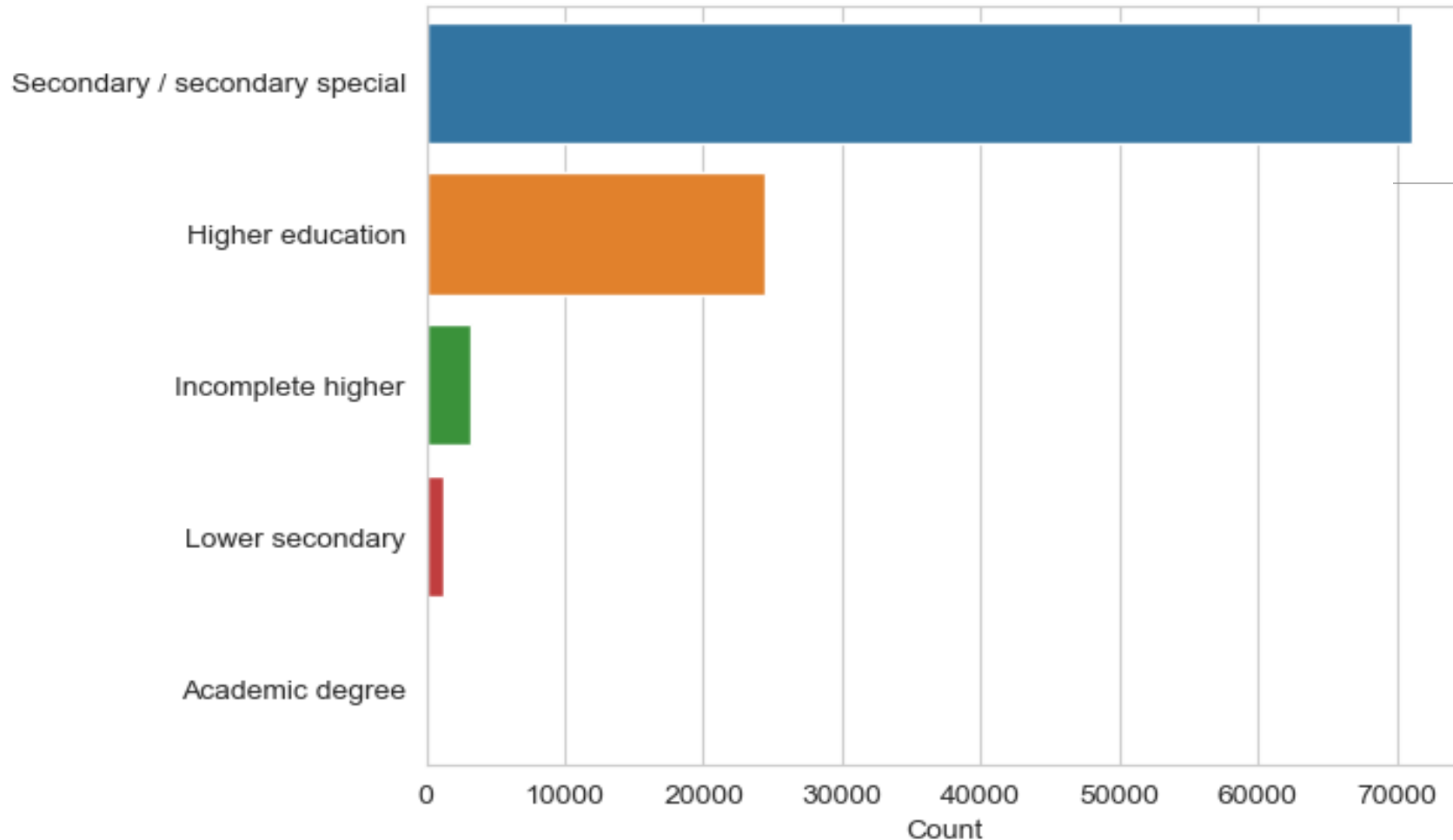
Distribution of Number of Children



INTERPRETATION

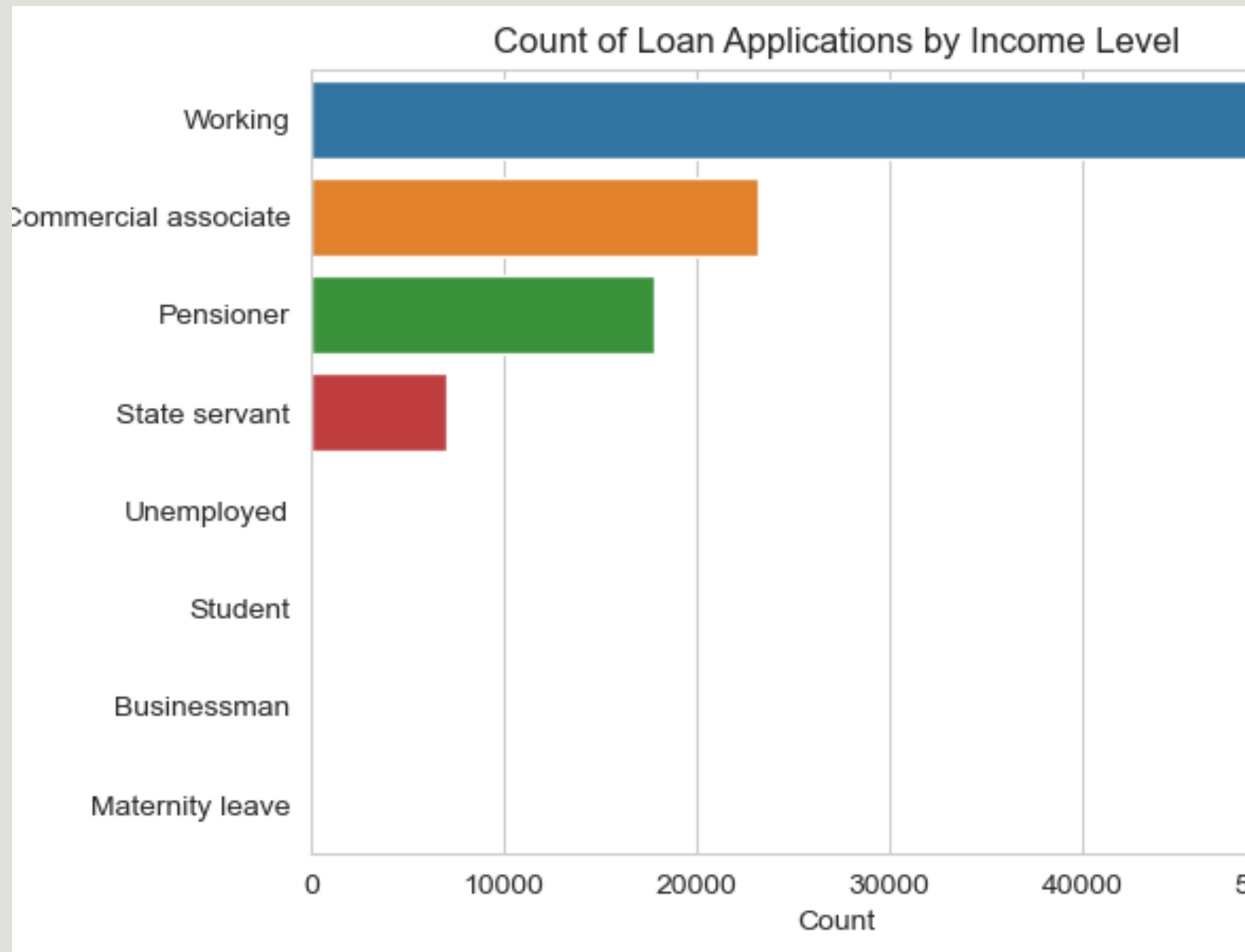
This plot shows the distribution of the number of children. We can see that the majority of the applicants do not have any children, with a few applicants having one or two children.

Count of Loan Applications by Education Level



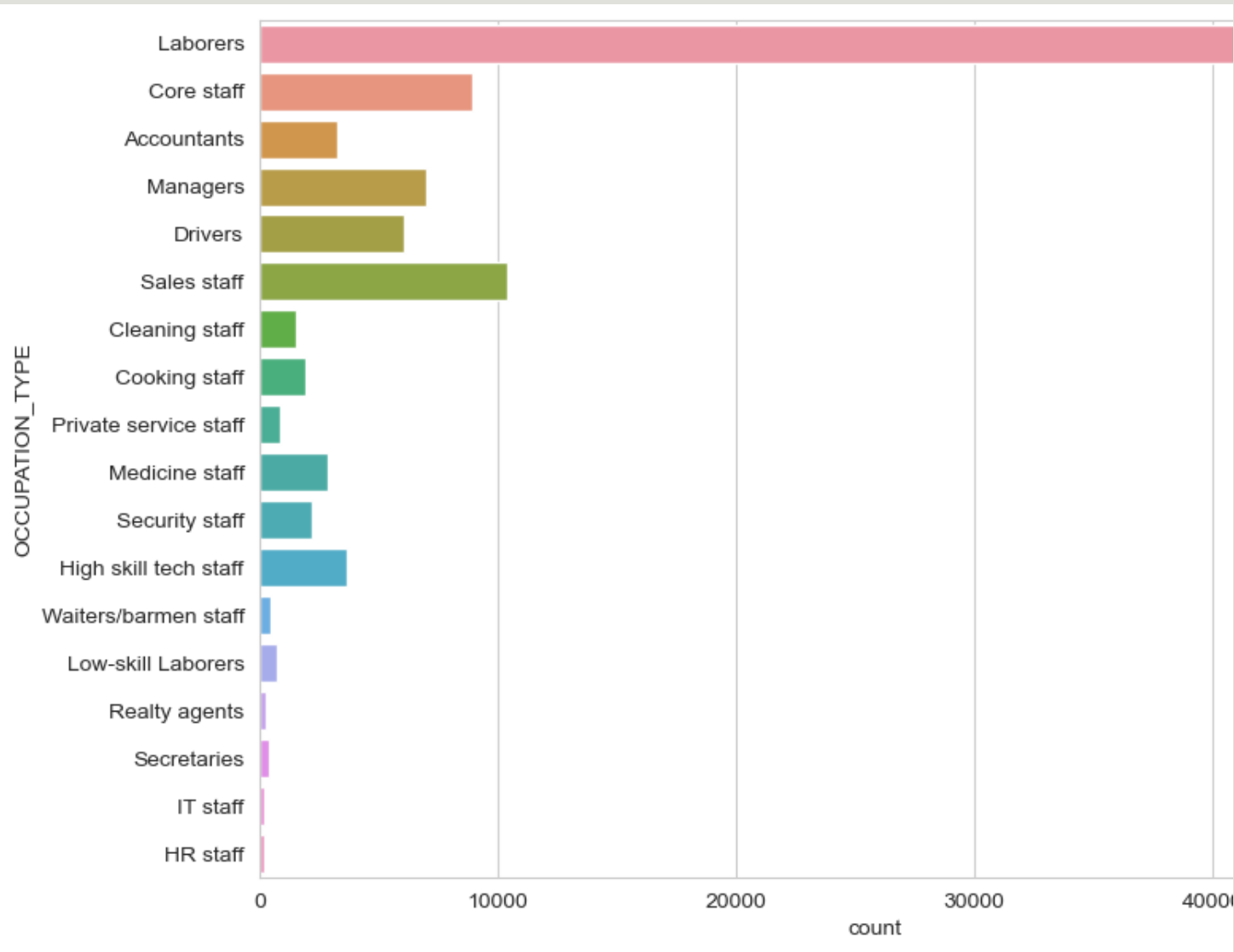
Interpretation

- It shows that most of the loan applications are from applicants who have completed Secondary / secondary special education.
- * The next largest group of loan applications is from applicants with Higher education. This suggests that applicants with higher education are more likely to apply for loans.



Interpretation

The majority of loan applicants in this dataset have an income level of "Working" followed by "Commercial associate" and "Pensioner".



Interpretation

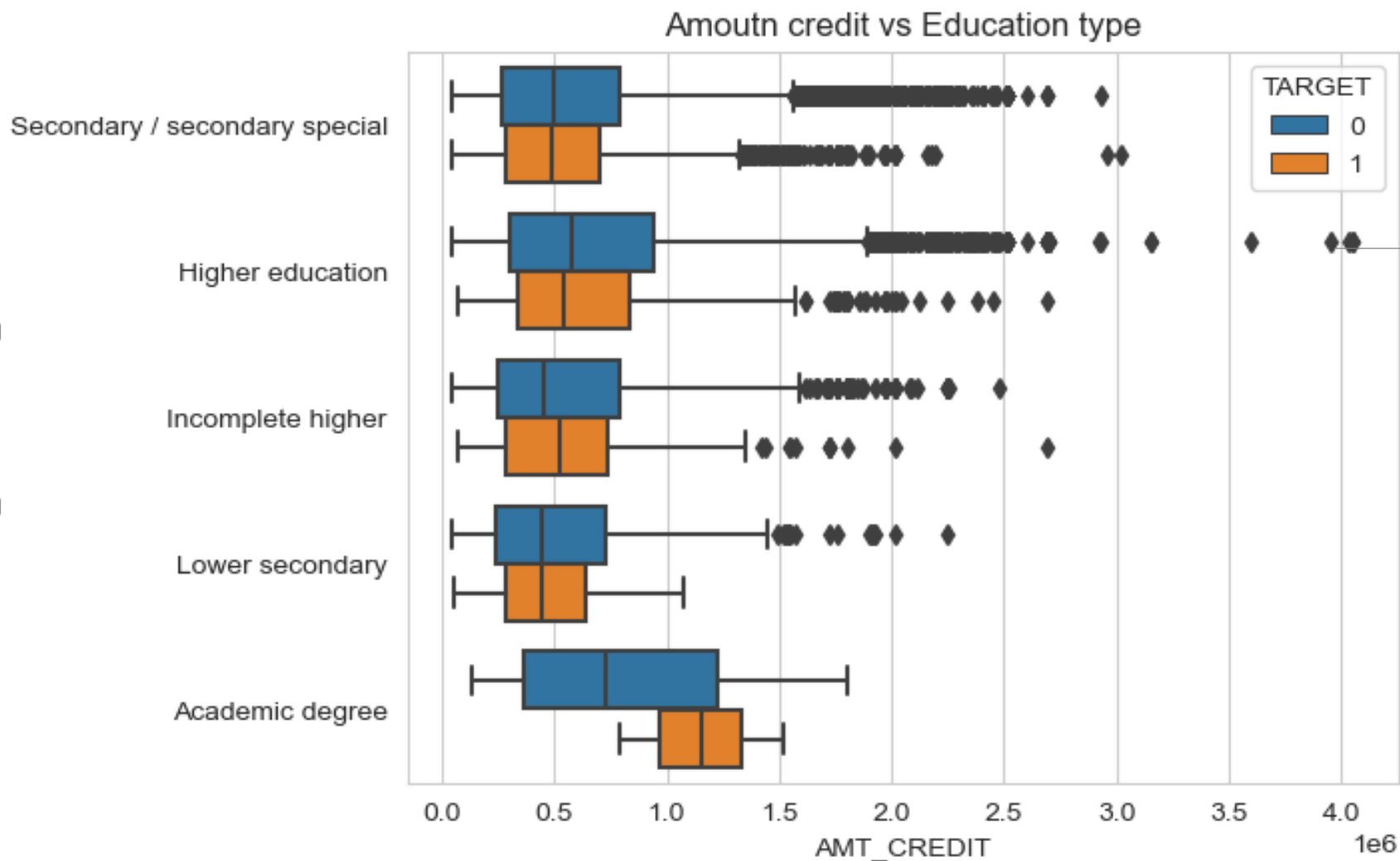
* THERE ARE SIGNIFICANT NUMBERS OF LOAN APPLICANTS ACROSS MOST OF THE OCCUPATION TYPES.

- The highest number of loan applicants have an occupation type of Laborers or Sales staff.

* IT staff and HR staff have the lowest number of loan applications.



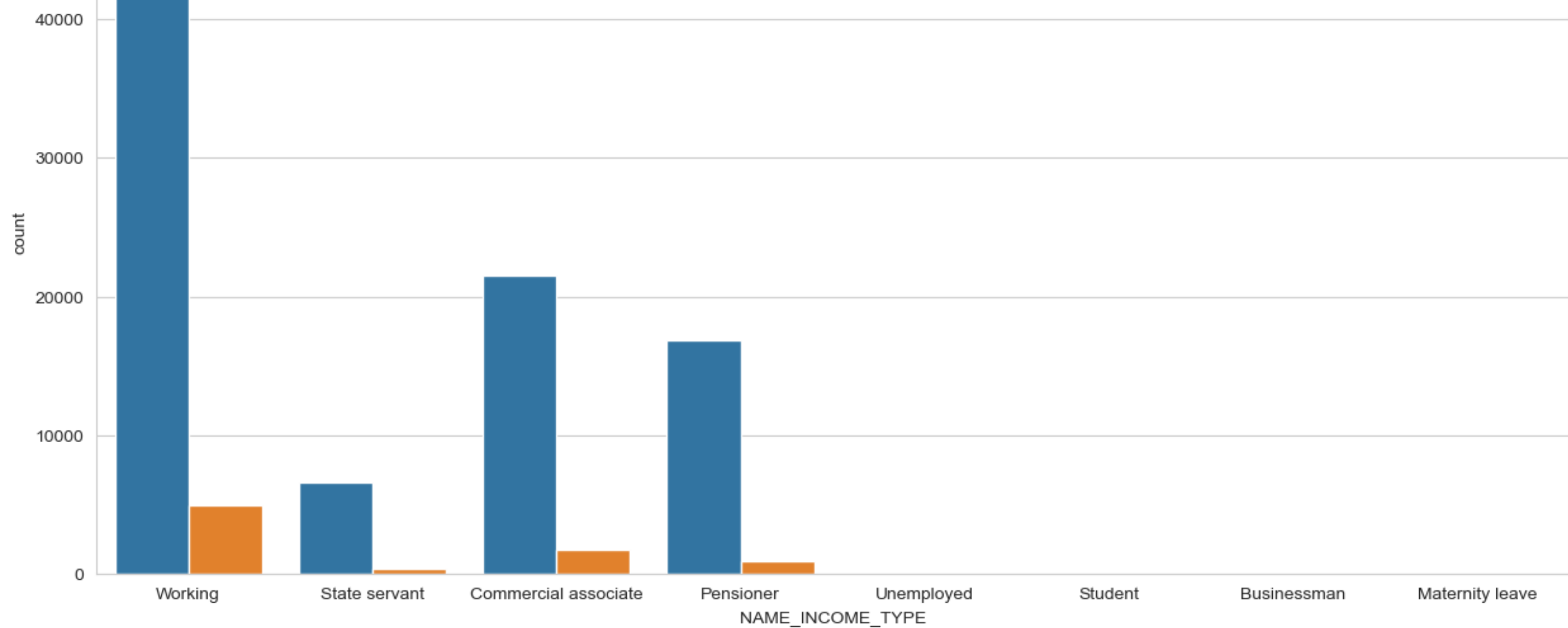
Bivariate Analysis:



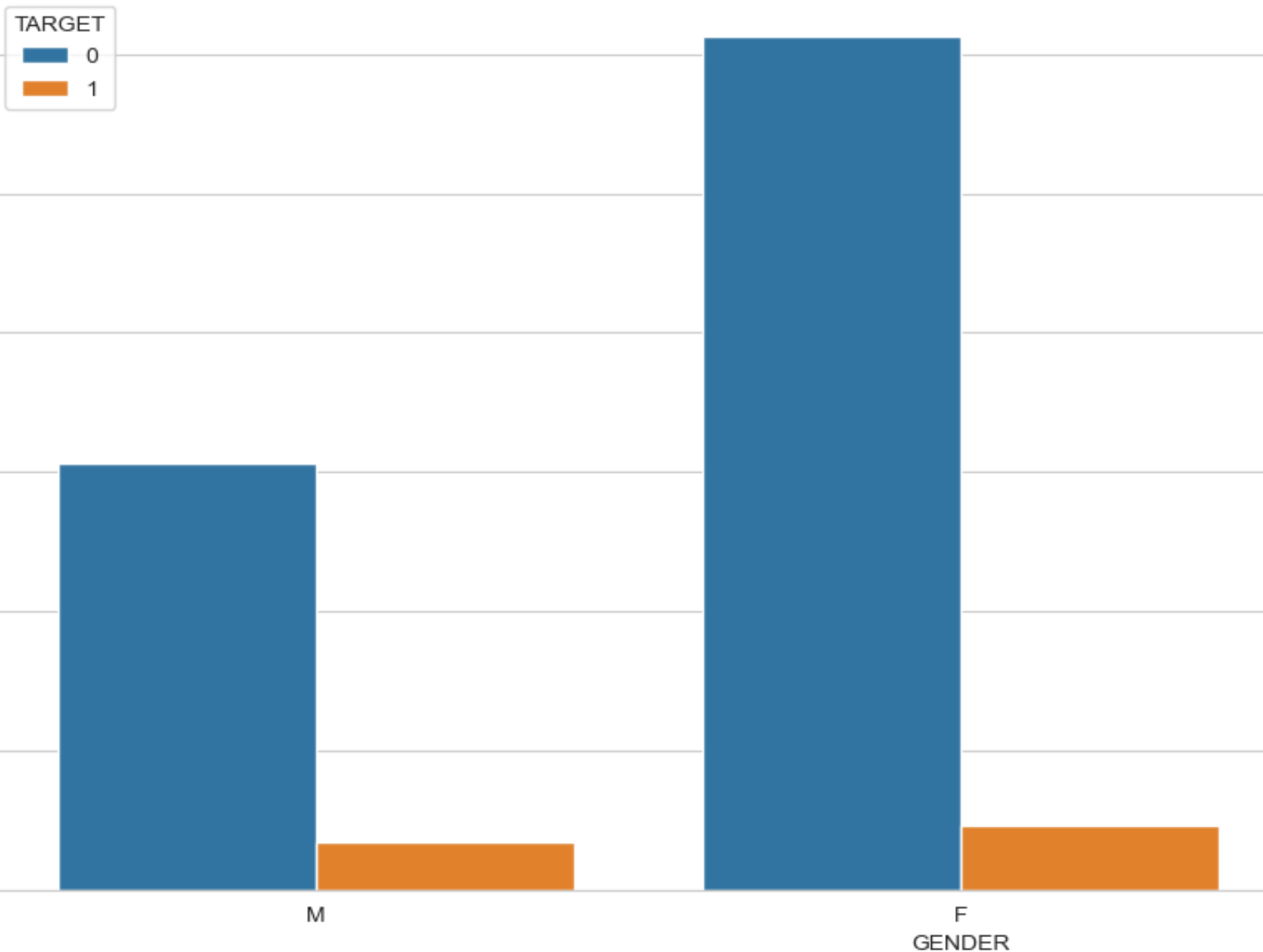
Interpretation

* The median loan amount for those with an academic degree is the highest among all education levels.

* Overall, there are more non-defaulters than defaulters across all education levels.



Interpretation : Majority of the applicants belong to Working category and it has the highest number of defaulters as well.

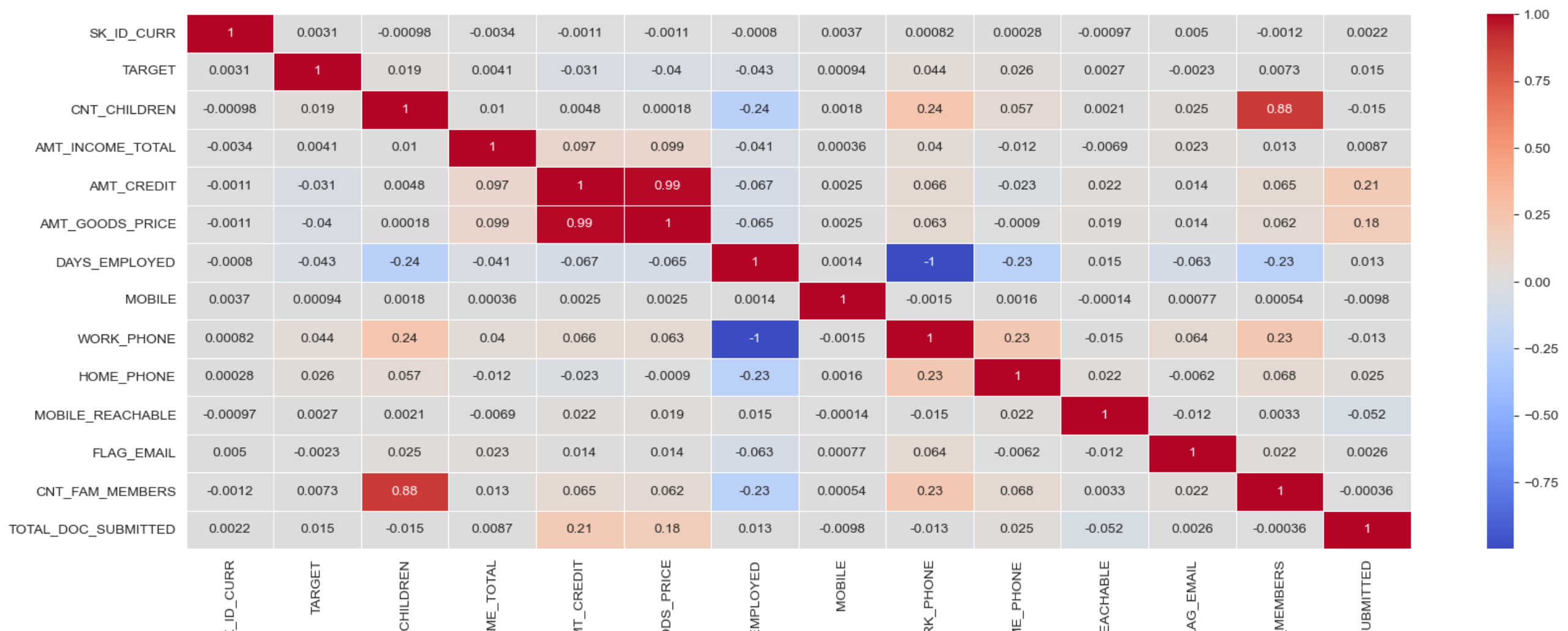


Interpretation

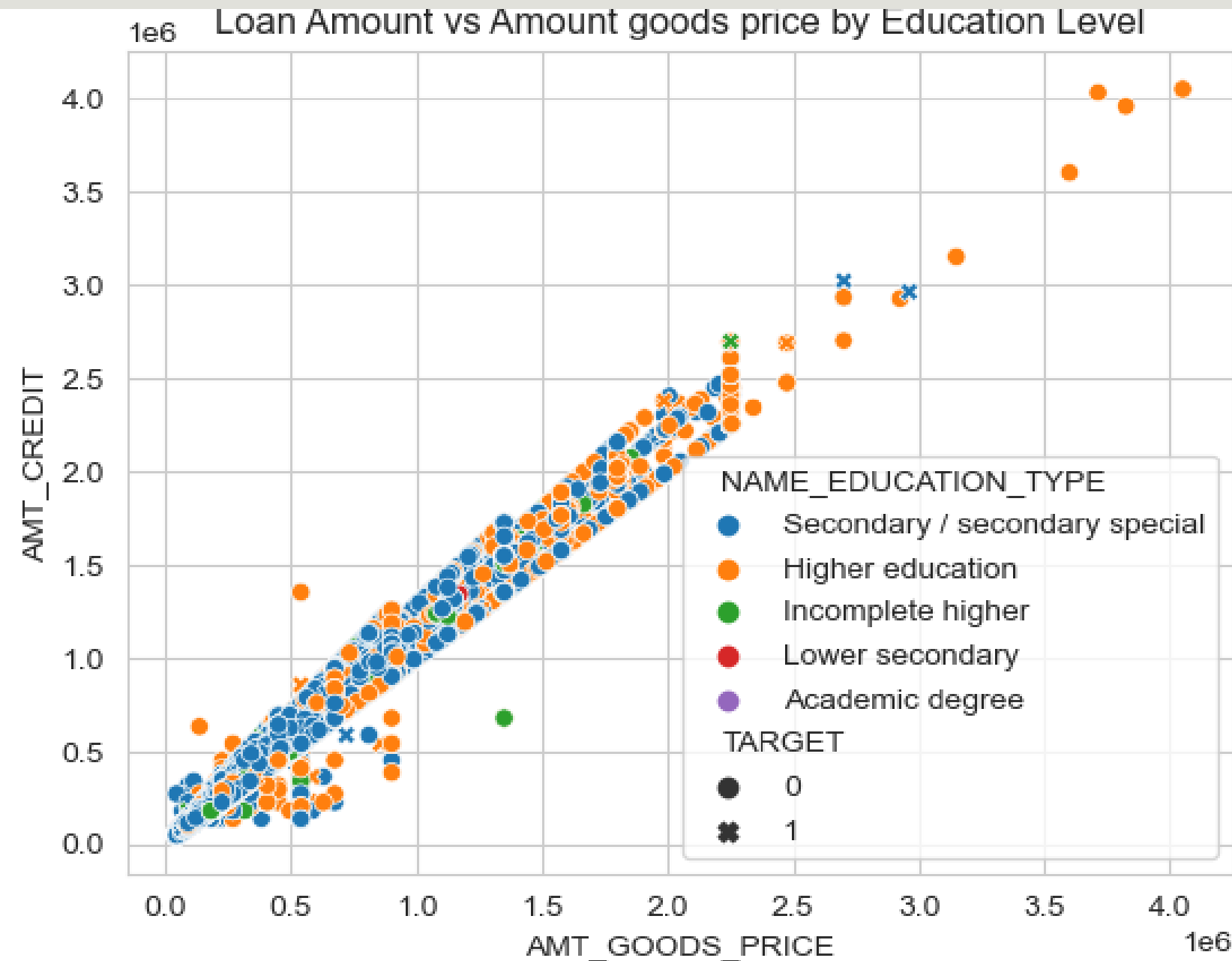
- * The count of females is higher than males for loan applications.
- * For both genders, the count of non-defaulters is higher than the count of defaulters.
- * The count of male non-defaulters is almost the same as female non-defaulters.
- * This indicates that the gender of the applicant does not have a significant impact on loan repayment.



Multivariate Analysis



Interpretation : The heatmap shows a positive correlation between the amount of credit and the amount of goods price .



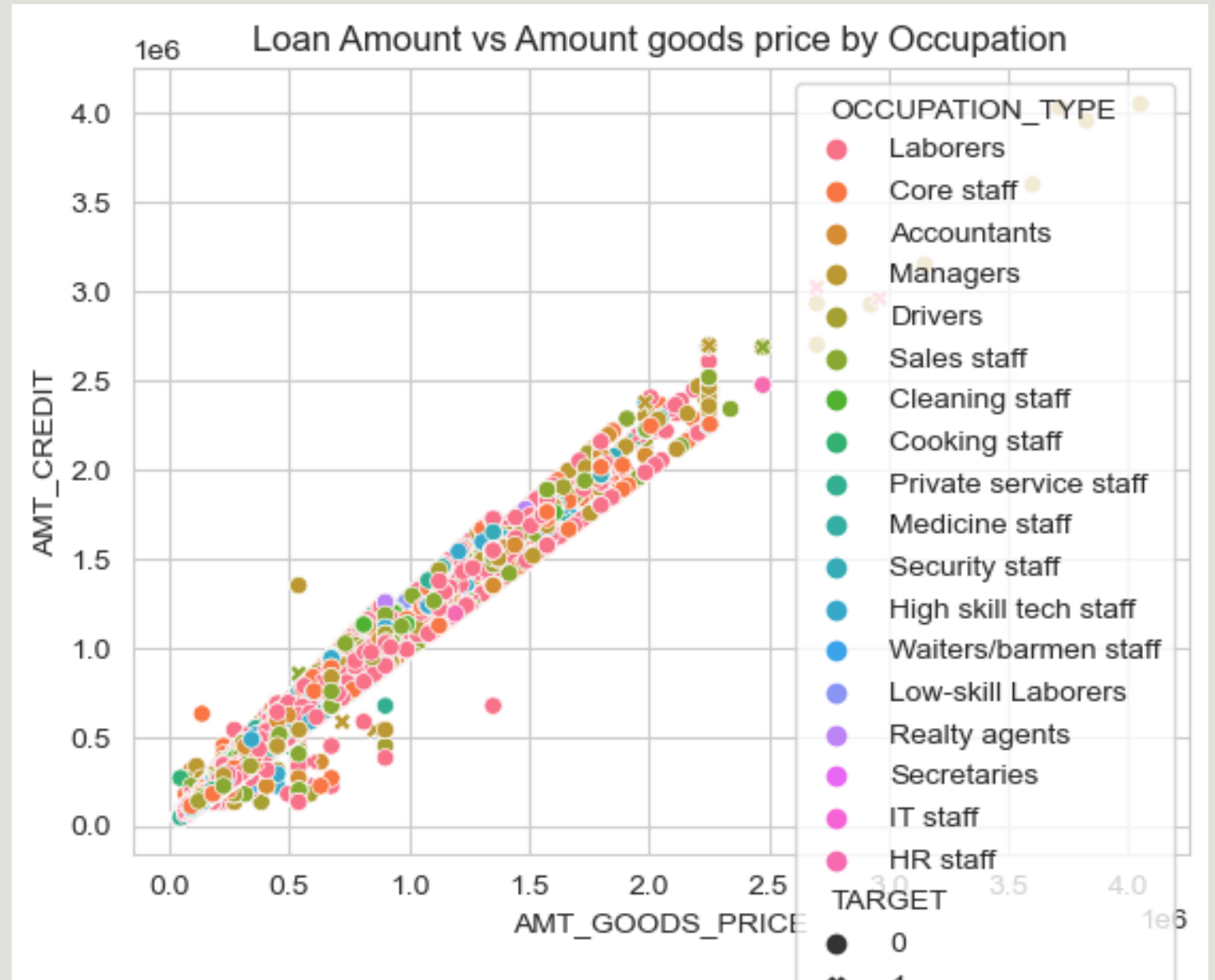
Interpretation

* Majority of the loan applications are from people with Secondary/secondary special education, followed by those with higher education.

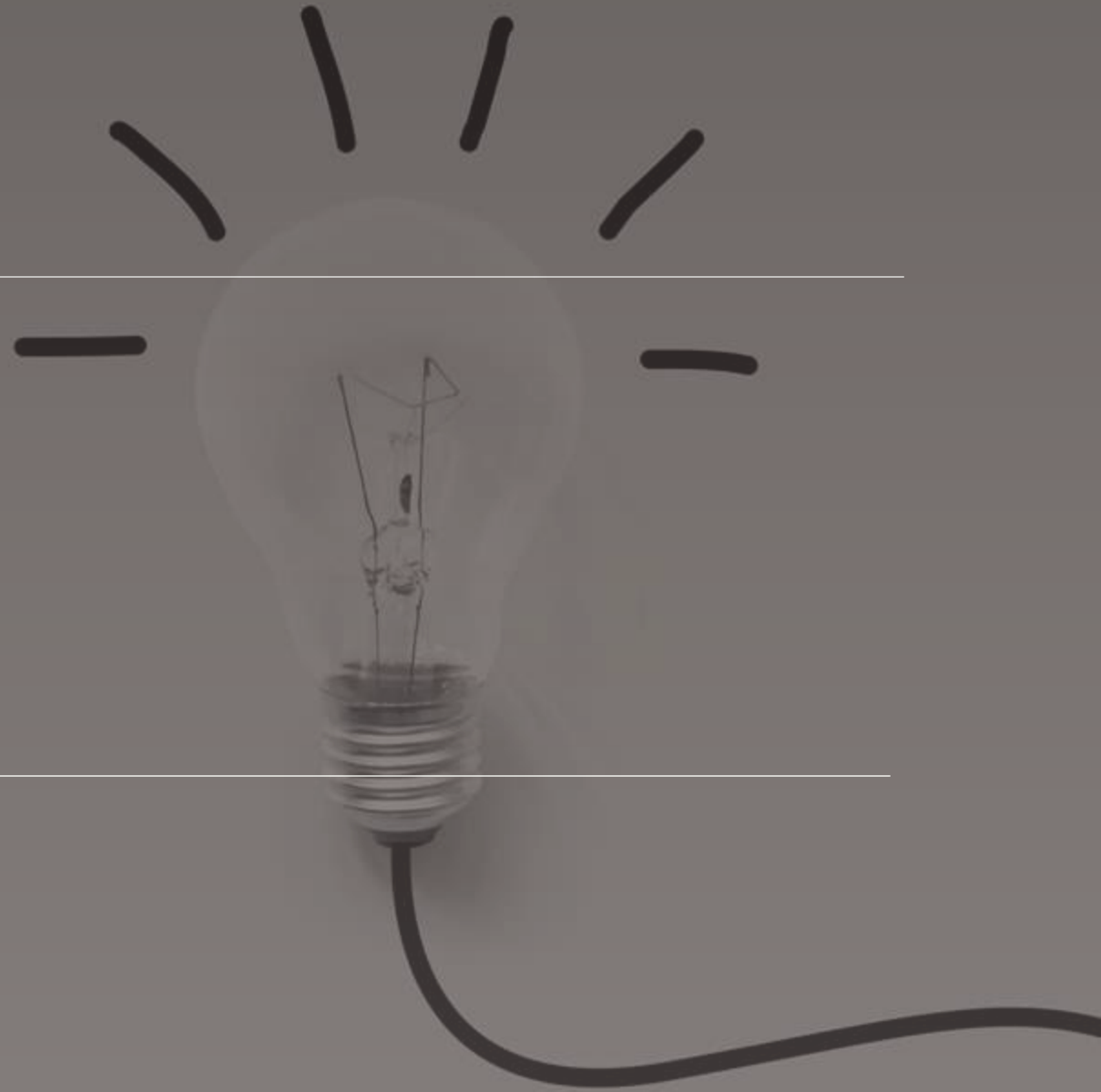
* Applications from people with incomplete higher education are the lowest.

Interpretation

The majority of loan applications are from applicants with occupation types of "Laborers" and "Sales staff".



Conclusions



Defaulters

- * Defaulters are more likely to have lower income levels and education levels, as well as to be unemployed or working in lower-paying industries.

- * Defaulters are more likely to have higher loan amounts compared to their income levels and to have more children.

Non-defaulters

- * Non-defaulters are more likely to have higher income levels and education levels, as well as to be employed in higher-paying industries.

- * Non-defaulters are more likely to have lower loan amounts compared to their income levels and to have fewer children.

- 50-50 chances

* Applicants in this category have a mix of characteristics of both defaulters and non-defaulters, making it difficult to predict whether they will default on their loans.

* Generally, applicants in this category have moderate income and education levels, as well as moderate loan amounts and number of children.



Thank you!