

# EDA Case Study

# ▀ Problem Statement

## Introduction

This assignment aims to give you an idea of applying EDA in a real business scenario. In this assignment, apart from applying the techniques that you have learnt in the EDA module, you will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers.

## Business Understanding

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. Suppose you work for a consumer finance company which specializes in lending various types of loans to urban customers. You have to use EDA to analyze the patterns present in the data. This will ensure that the applicants capable of repaying the loan are not rejected.



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:

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

The data given below contains the information about the loan application at the time of applying for the loan. It contains two types of scenarios:

**The client with payment difficulties:** he/she had late payment more than X days on at least one of the first Y instalments of the loan in our sample,

**All other cases:** All other cases when the payment is paid on time.



When a client applies for a loan, there are four types of decisions that could be taken by the client/company):

Approved: The Company has approved loan Application

Cancelled: The client cancelled the application sometime during approval. Either the client changed her/his mind about the loan or in some cases due to a higher risk of the client, he received worse pricing which he did not want.

Refused: The company had rejected the loan (because the client does not meet their requirements etc.).

Unused offer: Loan has been cancelled by the client but at different stages of the process.

In this case study, you will use EDA to understand how consumer attributes and loan attributes influence the tendency to default.

## Business Objectives

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. This will ensure that the consumers capable of repaying the loan are not rejected. Identification of such applicants using EDA is the aim of this case study.

In other words, 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 utilize this knowledge for its portfolio and risk assessment.

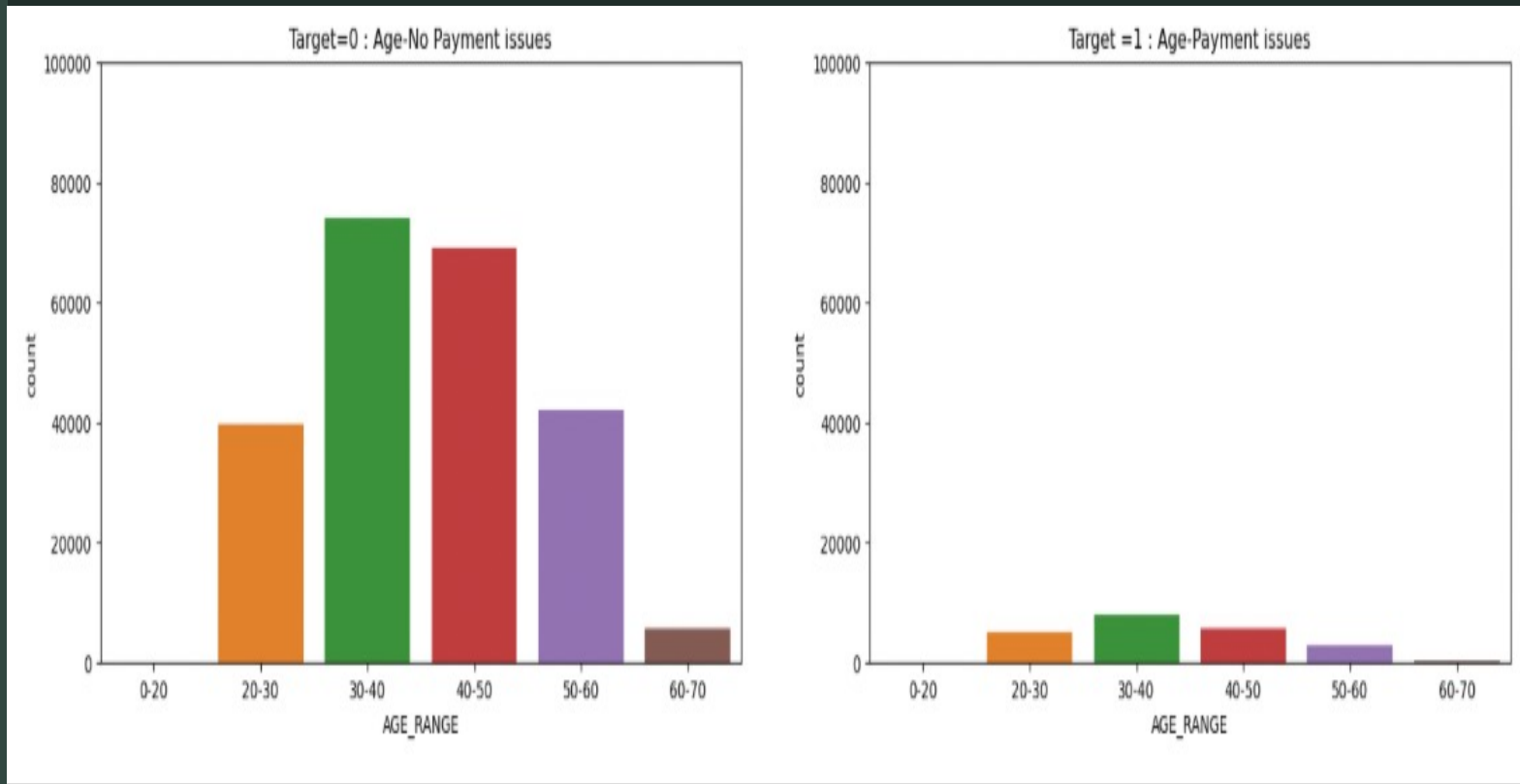
# Categorical univariate analysis for target 0

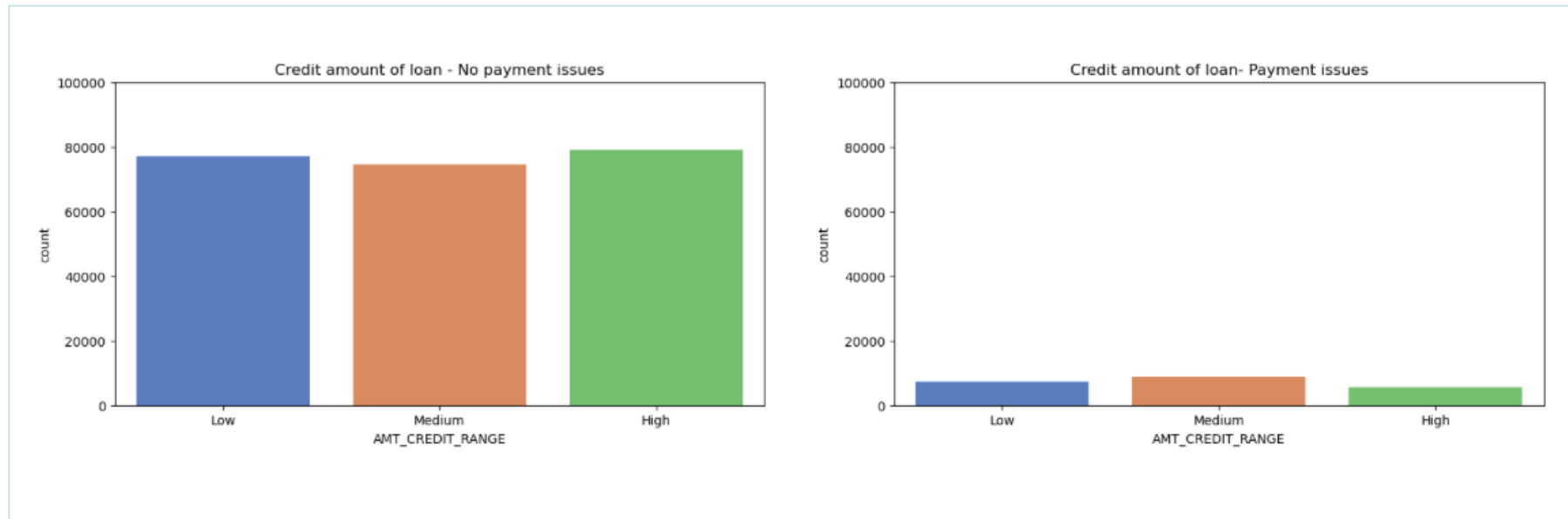


Points to be concluded from the graph:

- Customers belongs to the age group 30-40 are able to repay the loan amount on time
- Also 40-60 age group customers also having good repayment

Bank can target age group 30-40 range customer. Also can target 40-60 age group customers.

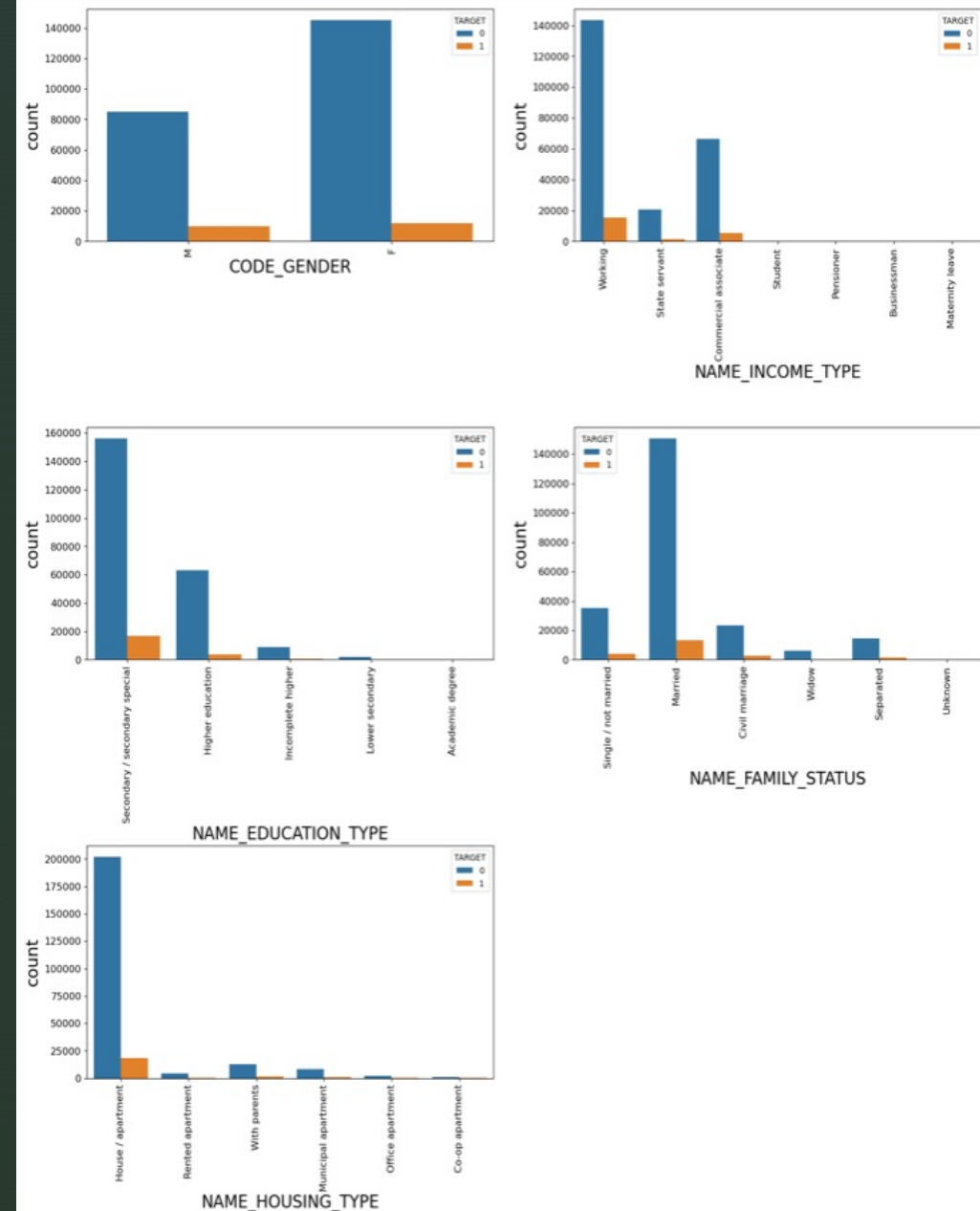




- Customers having medium and high credit can also be considered
- Customers with less credit and most likely to make payment



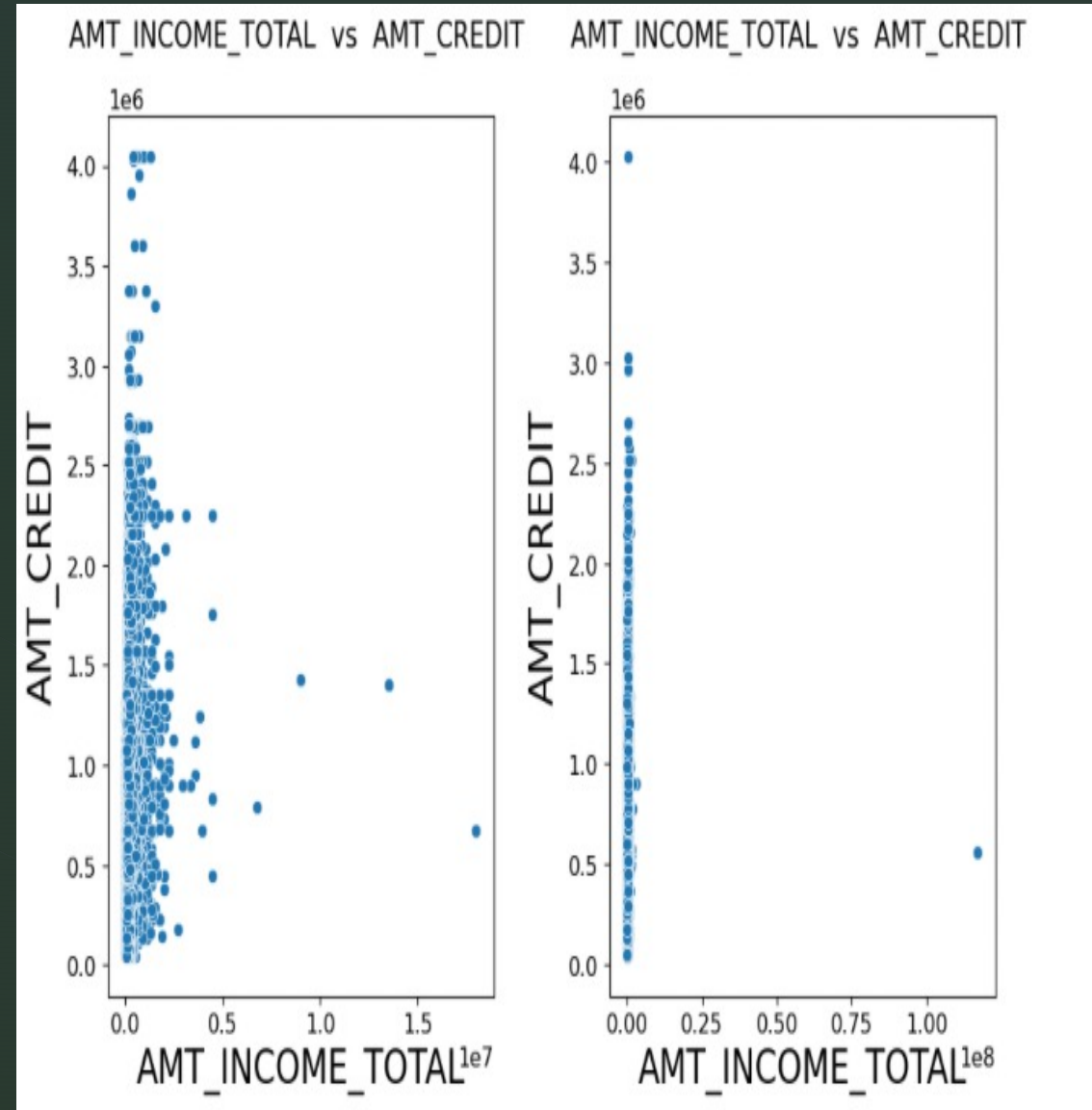
- Female customers are re-paying loan amount on time and banks can target more female customers for lending loan.
- Working customers can be targeted to lending loans as they have higher percentage of making payments on time.
- Customers with secondary education are most likely to make payments when compare to customers with academic degree.
- Married customers have paid loan amount on time when compared to widows.
- Customers owning House/apartment are most likely to make payments on time compare to those living in CO-OP apartment.

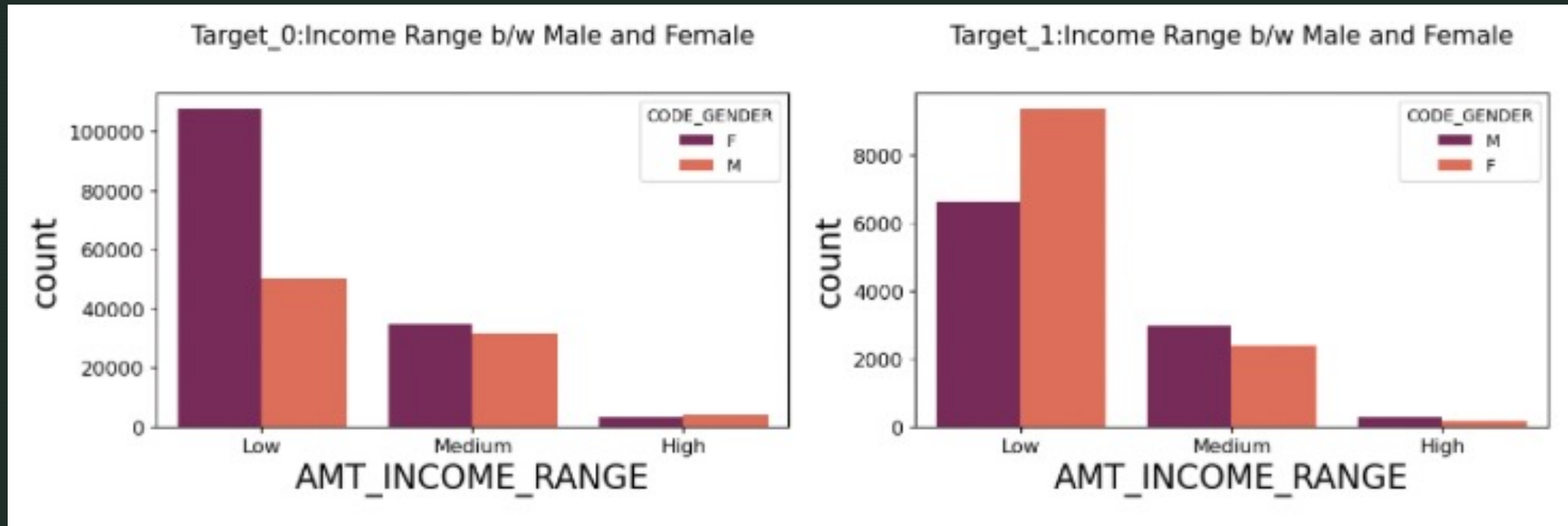




## **Bivariate Analysis for target 0 and target 1**

Those who have paid the loan amount on/within time are more likely to get higher credits than those who didn't pay/did late payments.

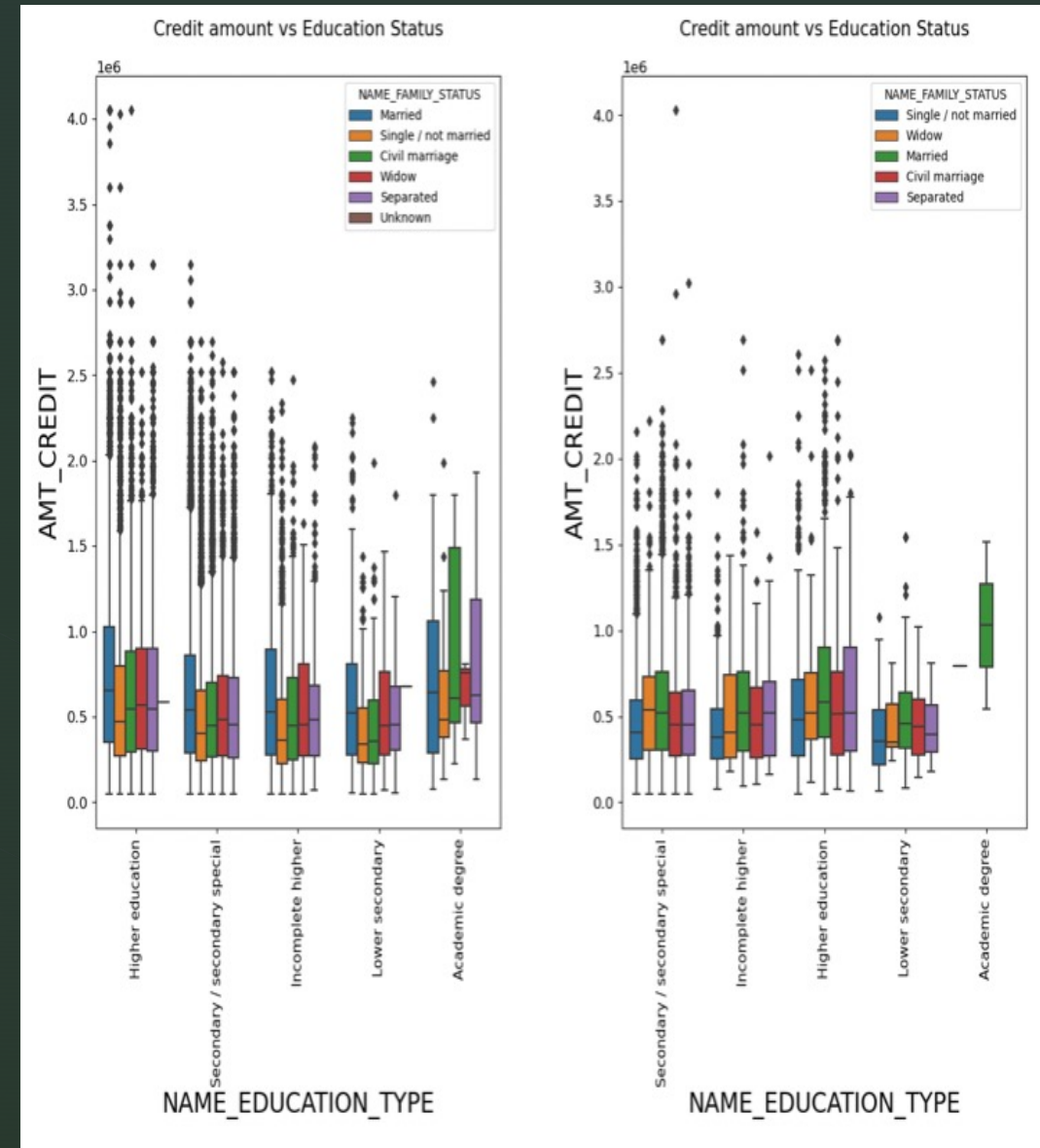




Graph shows that Females with low income don't have any payment issues.

We can observe from the plot is :

- Some of the highly educated, married person are having credits higher than those who have done lower secondary education.
- Those with higher education have higher credits and are more likely to make payments on time.
- More number of outliers are seen in higher education.
- The people with secondary/secondary special education are less likely to make payments on time.

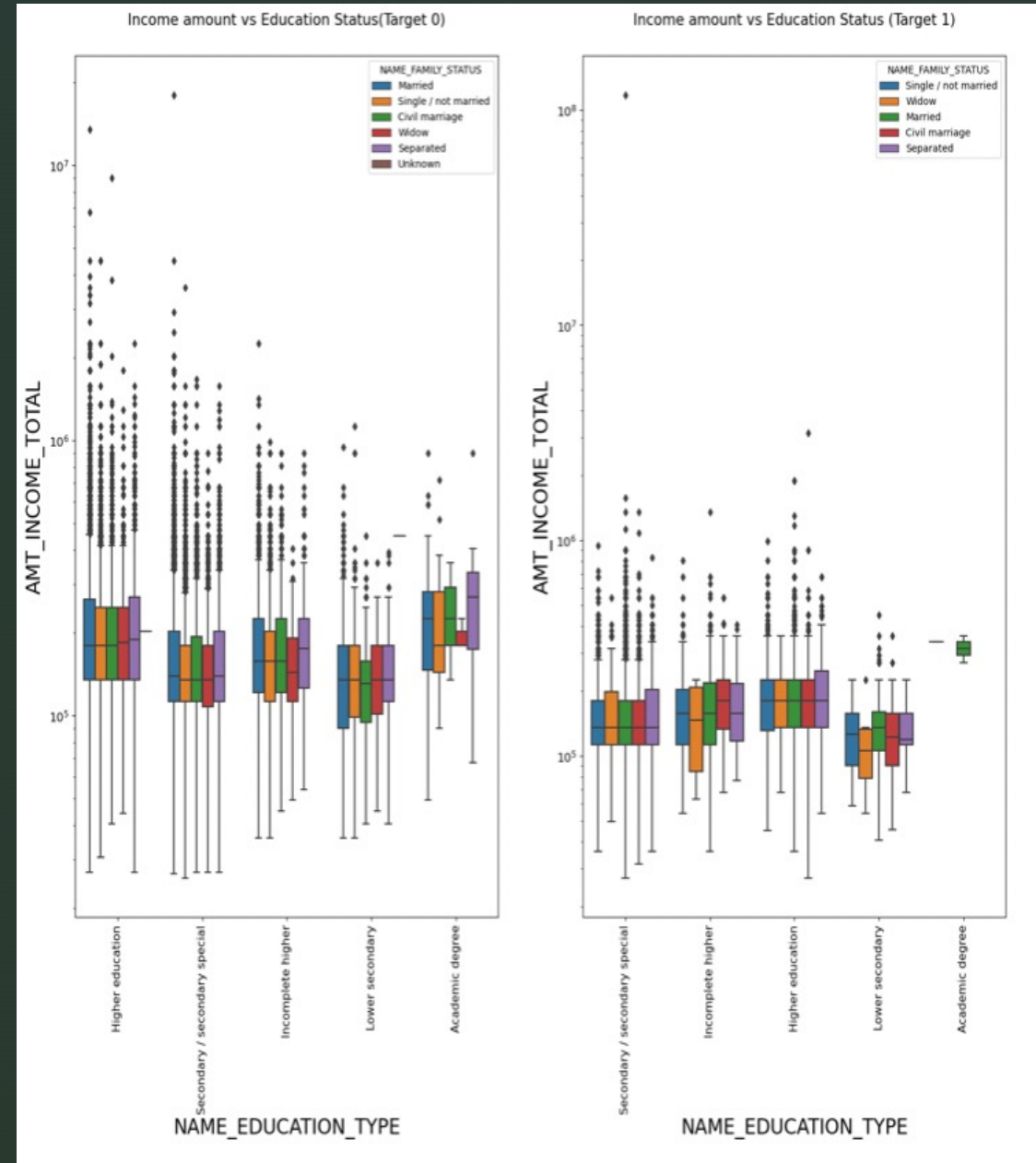




We can observe from the plot is :

- We can see that Higher education has many outliers.
- People with higher education have higher income and don't have difficulties in making loan payment.
- People with higher education who has lesser income are unable to pay the loan.

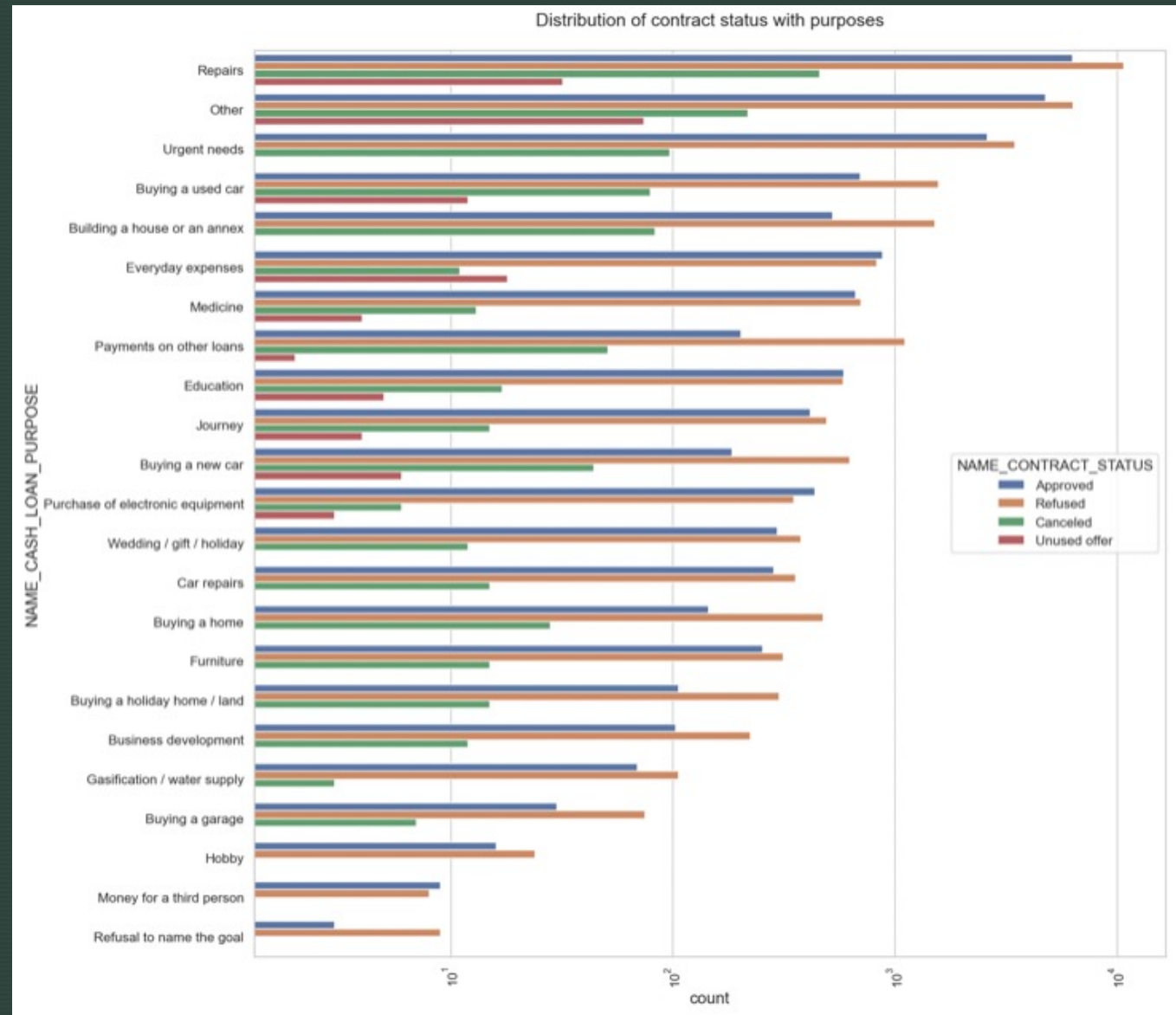
Hence we can conclude that people with Higher income are most likely to make payments.





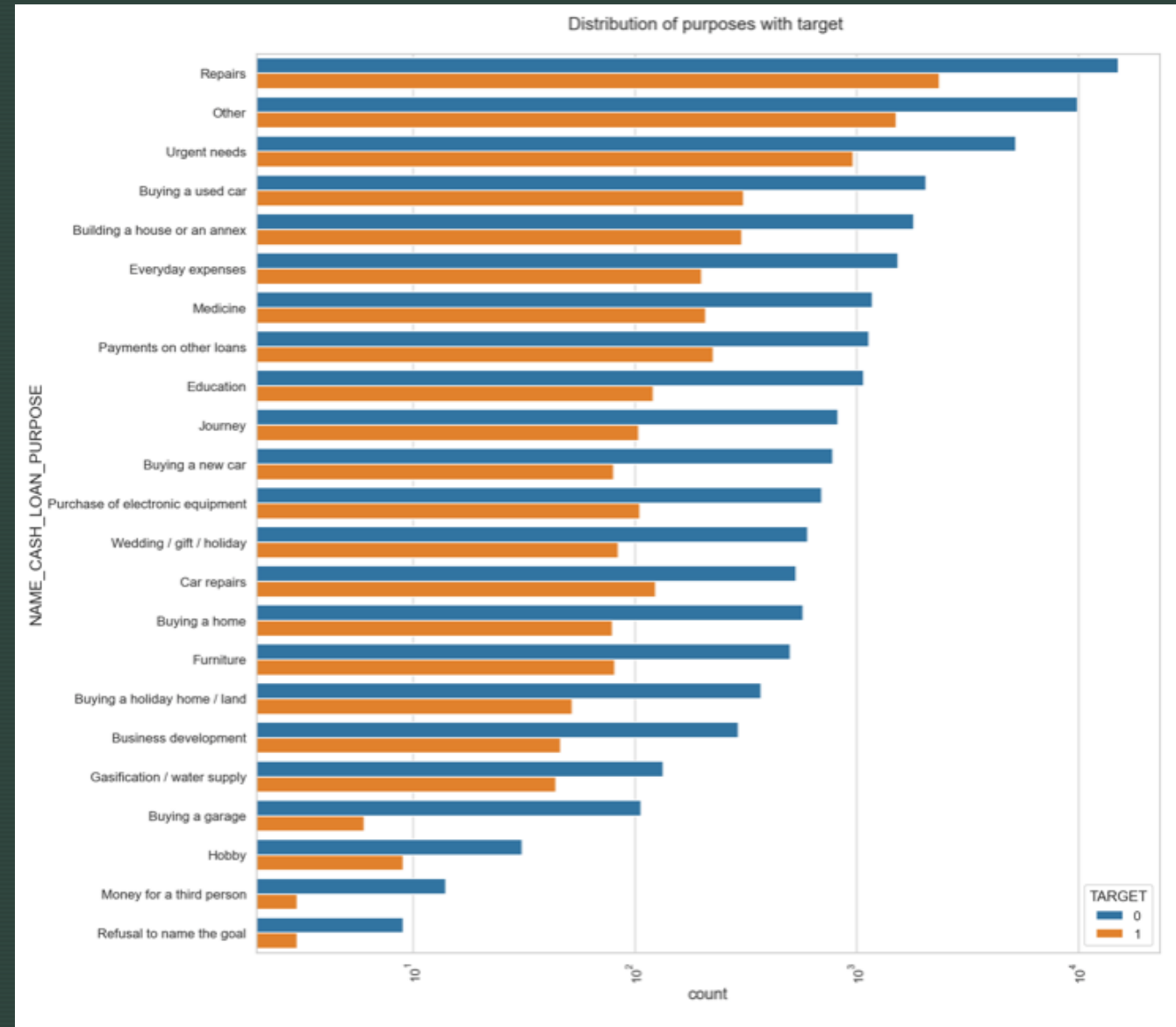
Points to be concluded from above plot:

- Most rejection of loans came from purpose 'Repairs'.
- For education purposes we have equal number of approves and rejection Paying other loans and buying a new car is having significant higher rejection than approves.



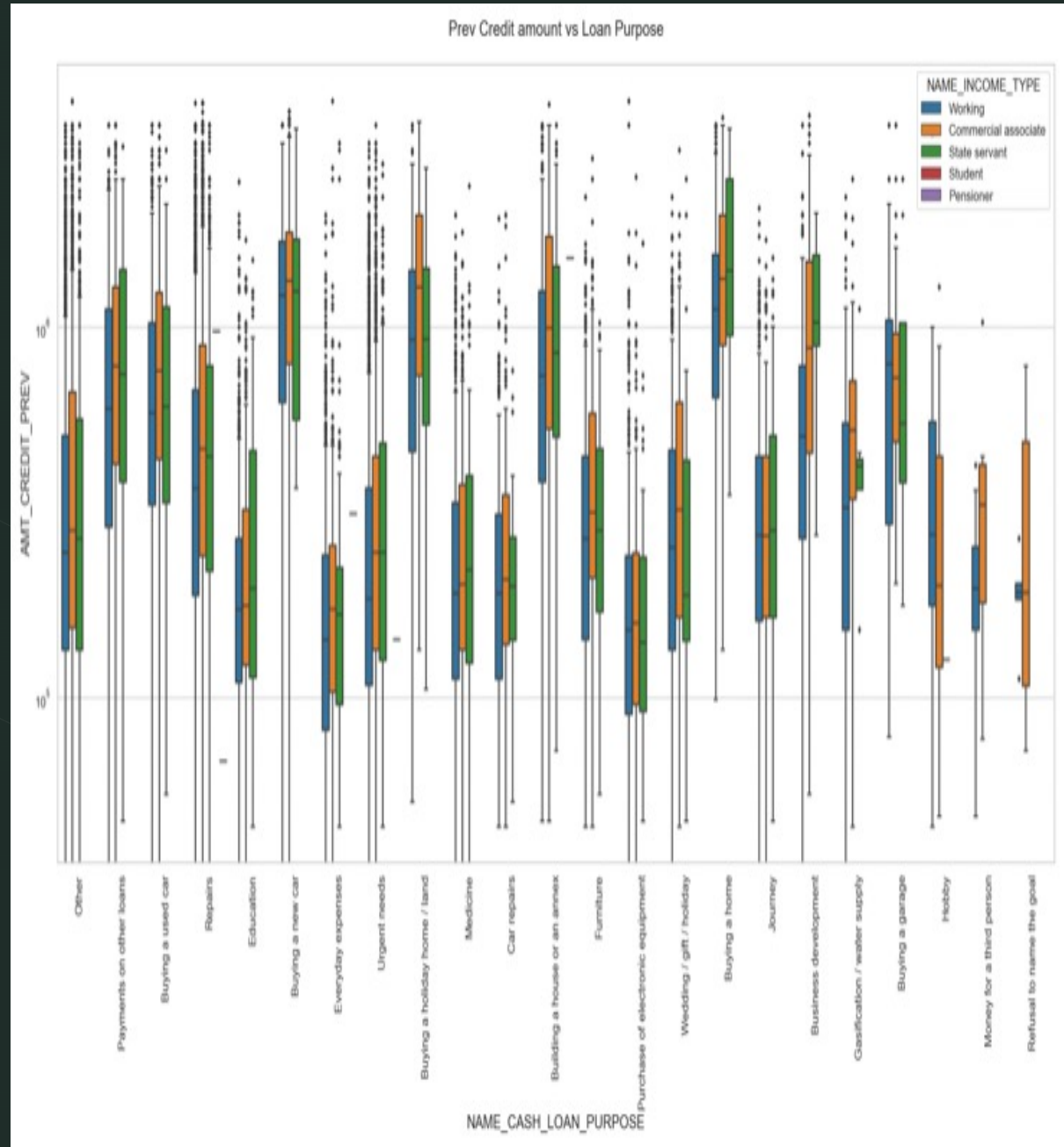
Points to be concluded from above plot:

- Loan purposes with 'Repairs' are facing more difficulties in payment on time.
- There are few places where loan payment is significant higher than facing difficulties. They are 'Buying a garage', 'Business development', 'Buying land', 'Buying a new car' and 'Education'. Hence we can focus on these purposes for which the client is having for minimal payment difficulties

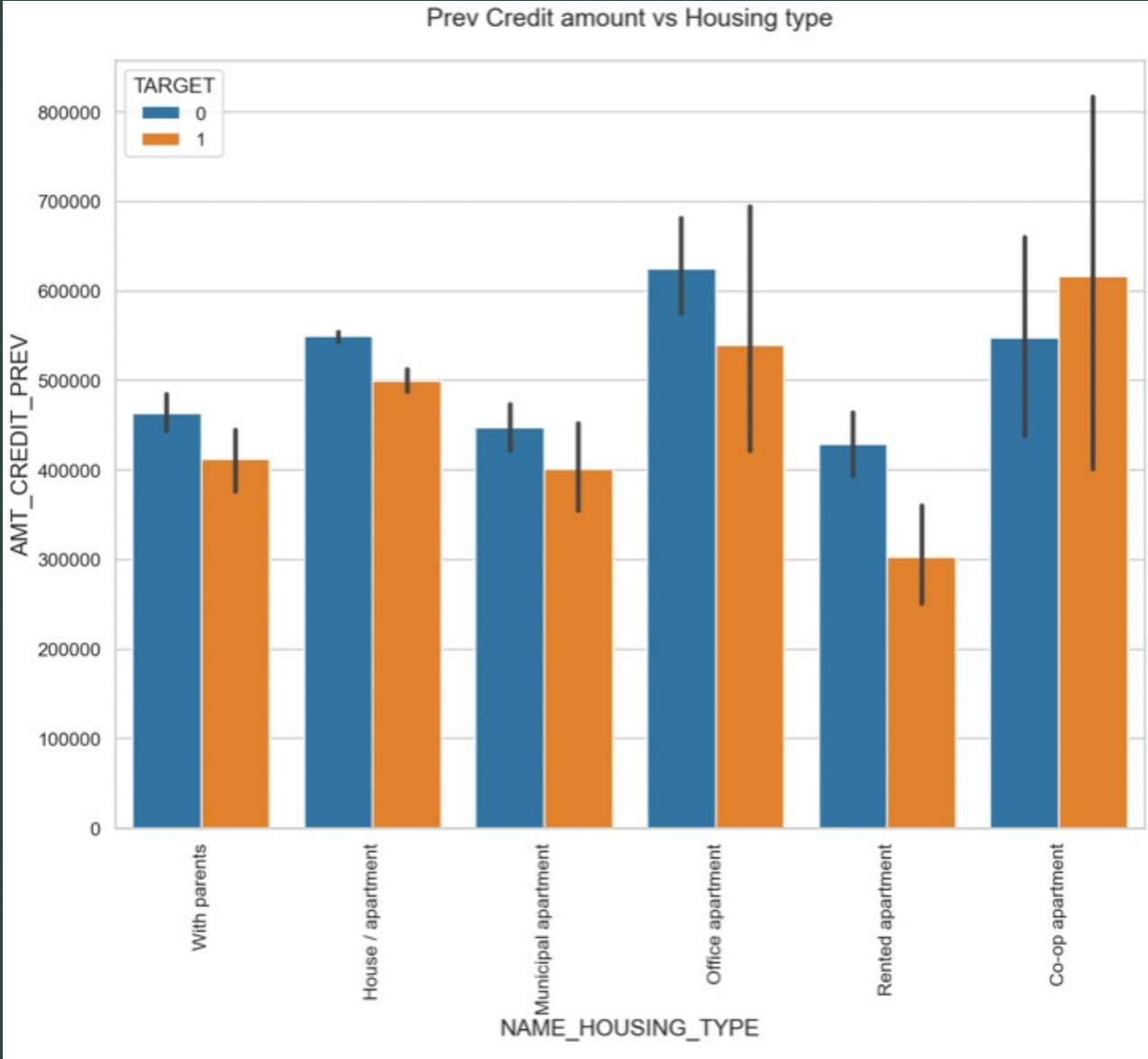


From the above we can conclude some points:

- The credit amount of Loan purposes like 'Buying a home', 'Buying a land', 'Buying a new car' and 'Building a house' is higher.
- Income type of state servants have a significant amount of credit applied.
- Money for third person or a Hobby is having less credits applied for.



Here for Housing type, office apartment is having higher credit of target 0 and co-op apartment is having higher credit of target 1. So, we can conclude that bank should avoid giving loans to the housing type of co-op apartment as they are having difficulties in payment. Bank can focus mostly on housing type with parents or House\apartment or municipal apartment for successful payments.





## Conclusion:

1. Banks should focus more on contract type 'Student' , 'pensioner' and 'Businessman' with housing 'type other than 'Co-op apartment' for successful payments.
2. Banks should focus less on income type 'Working' as they are having most number of unsuccessful payments.
3. Also with loan purpose 'Repair' is having higher number of unsuccessful payments on time.
4. Get as much as clients from housing type 'With parents' as they are having least number of unsuccessful payments.