Credit EDA Assignment

LOAN REPAYMENT

TIME TO PAY THE BILLS

PRESENTED BY –
Ms. SONU VASAYA

Problem statement

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.

Exploratory Data Analysis study aims to identify patterns that indicate if an applicant will repay their instalments which may be used for taking further actions such as denying the loan, reducing the amount of loan, lending at a higher interest rate, etc. This will make sure that the applicants capable of repaying the loan are not rejected.

Steps Followed

First Data set (application_data)
❑ **Data Understanding**
  ❑ a. Reading the Data.
  ❑ b. Verifying if the values are null.
  ❑ c. Eliminating the column that contains more than 40% null values.
  ❑ d. Managing Nulls by imputing values.
  ❑ e. Dividing into numerical and category columns.

First Data set (previous_application)
❑ **Data Understanding**
  ❑ a. Reading the Data.
  ❑ b. Verifying if the values are null.
  ❑ c. Managing Nulls by imputing values.
  ❑ d. Merging two Data sets.
❑ **Univariate analysis on Numerical variables**
❑ **Univariate analysis on Categorical variables**
❑ **Univariate analysis of numerical column for TARGET=0 (Applicants paying on schedule)**
❑ **Univariate analysis of categorical column for TARGET=0 (Applicants paying on schedule)**
❑ **Univariate analysis of numerical column for TARGET=1 (Applications that fail to pay on time)**
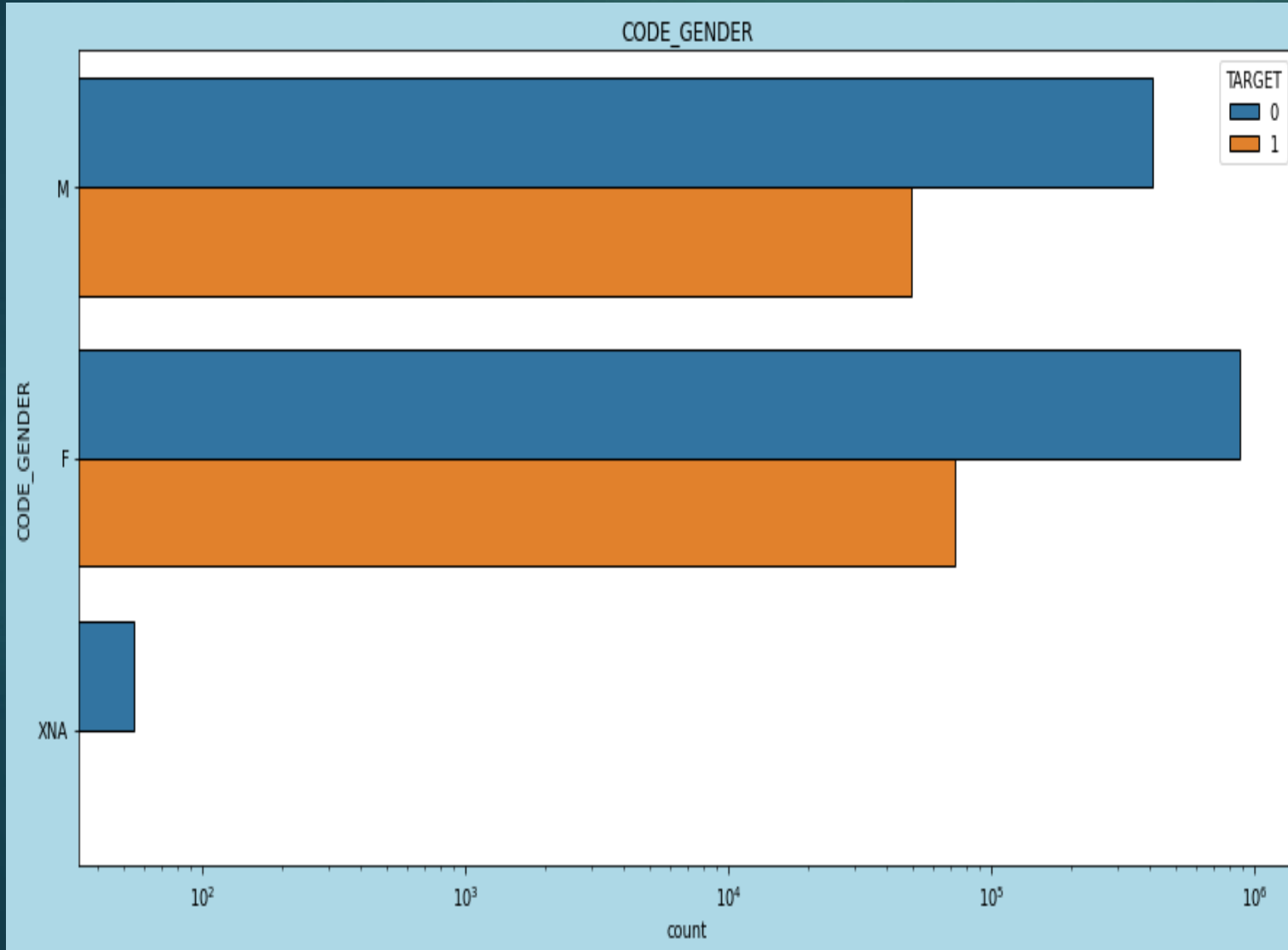❑ **Univariate analysis of categorical column for TARGET=1 (Applications that fail to pay on time)**
❑ **Bivariate Analysis for categorical columns**
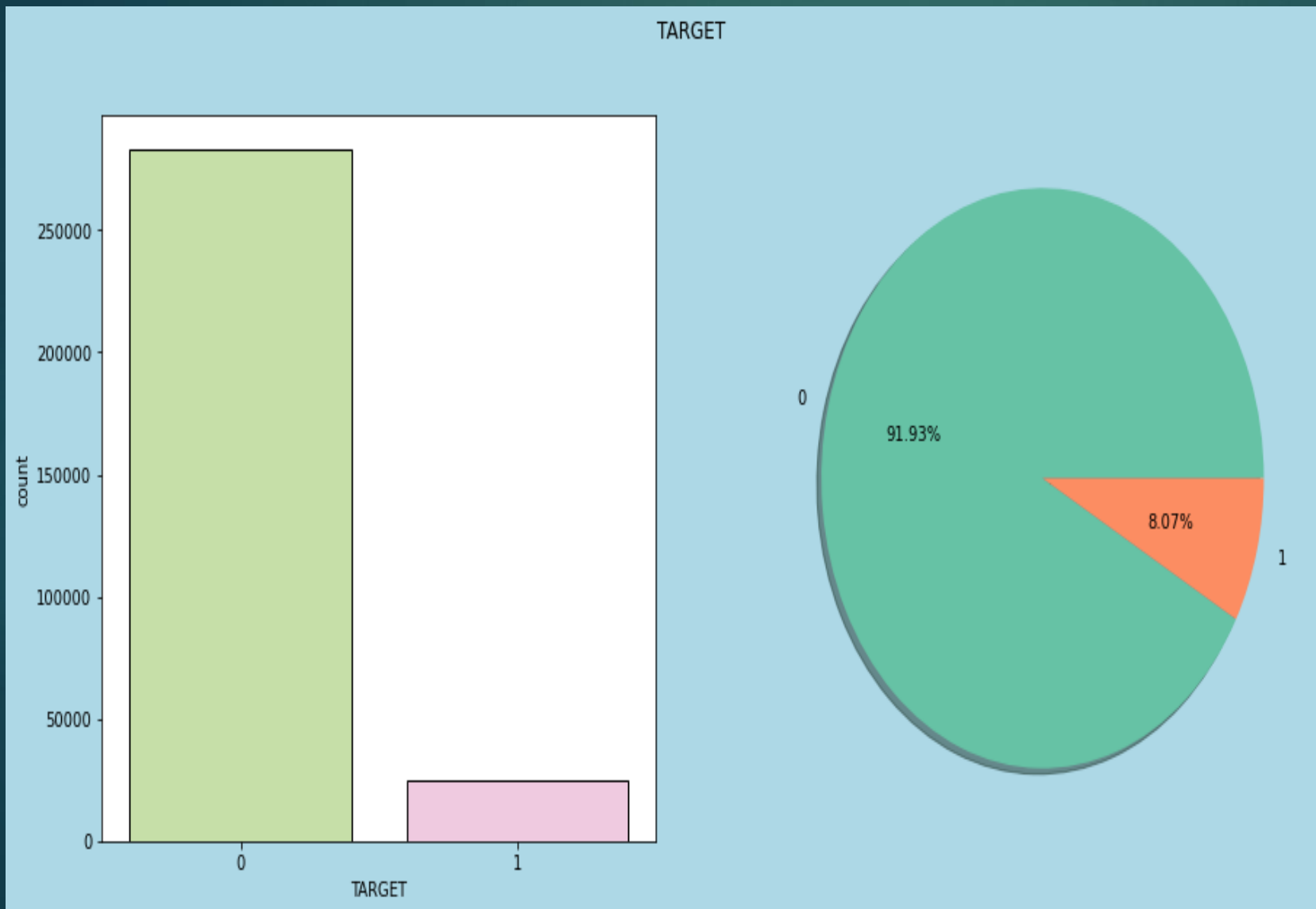❑ **Final Conclusion**

DATA PROCESSING
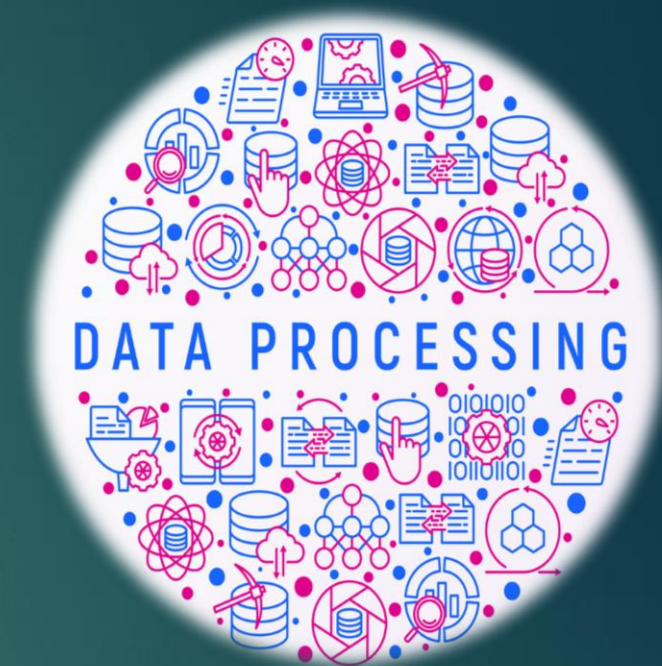
Gender Distribution in Target

Conclusion:
1.Female people have applied more for loan.
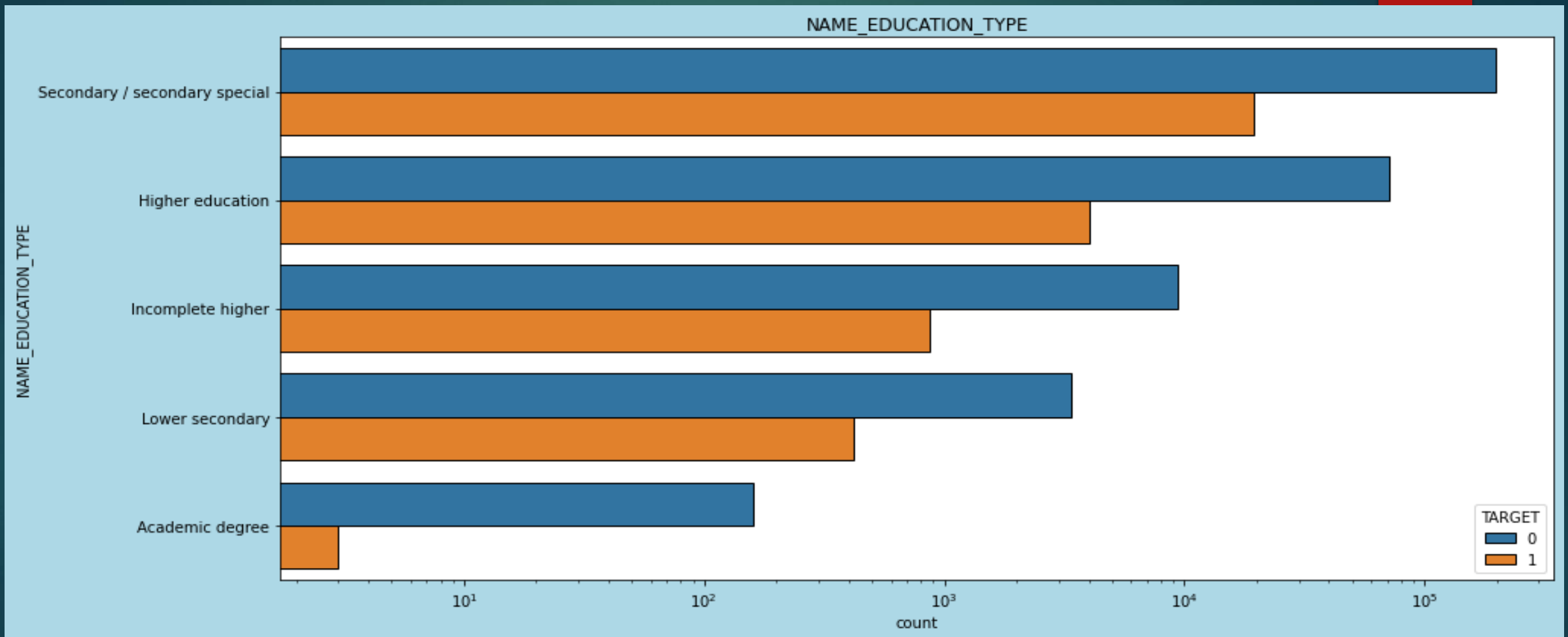2.while male people have high chance to be defaulter.

Analysis on target

Conclusion:
1. We can see 8% of clients face problems in payment.
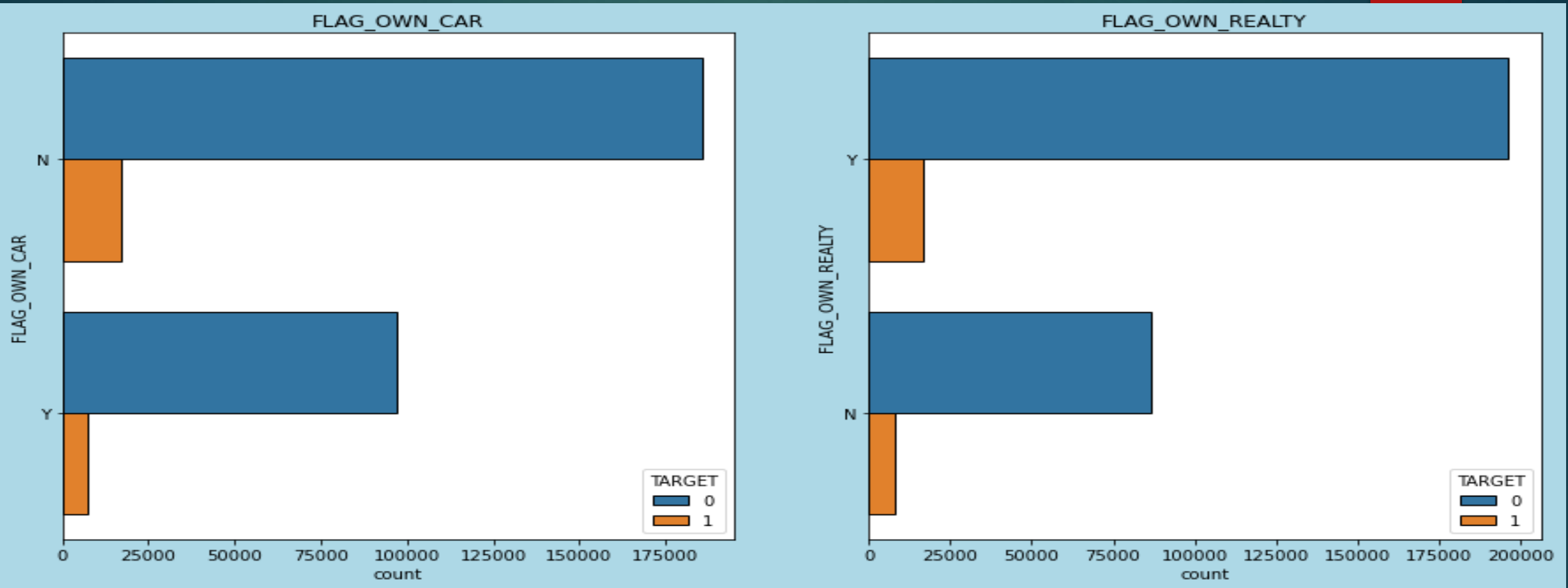2. And 91% of clients do not face problem.

Clients EDUCATION TYPE

Conclusion:
1.Most of the clients are secondary school passout, with higher education and so on.
2. 76% of secondary school passout are Non-defaulters.

Clients with and without house and car

Conclusion :
1.We can say that clients with there own house are more and this can used used as property collateral while providing loan to the costumer and are non defaulters.
2.People with property have more made more payment on time.
3.People with no car have applied more applied for loan.
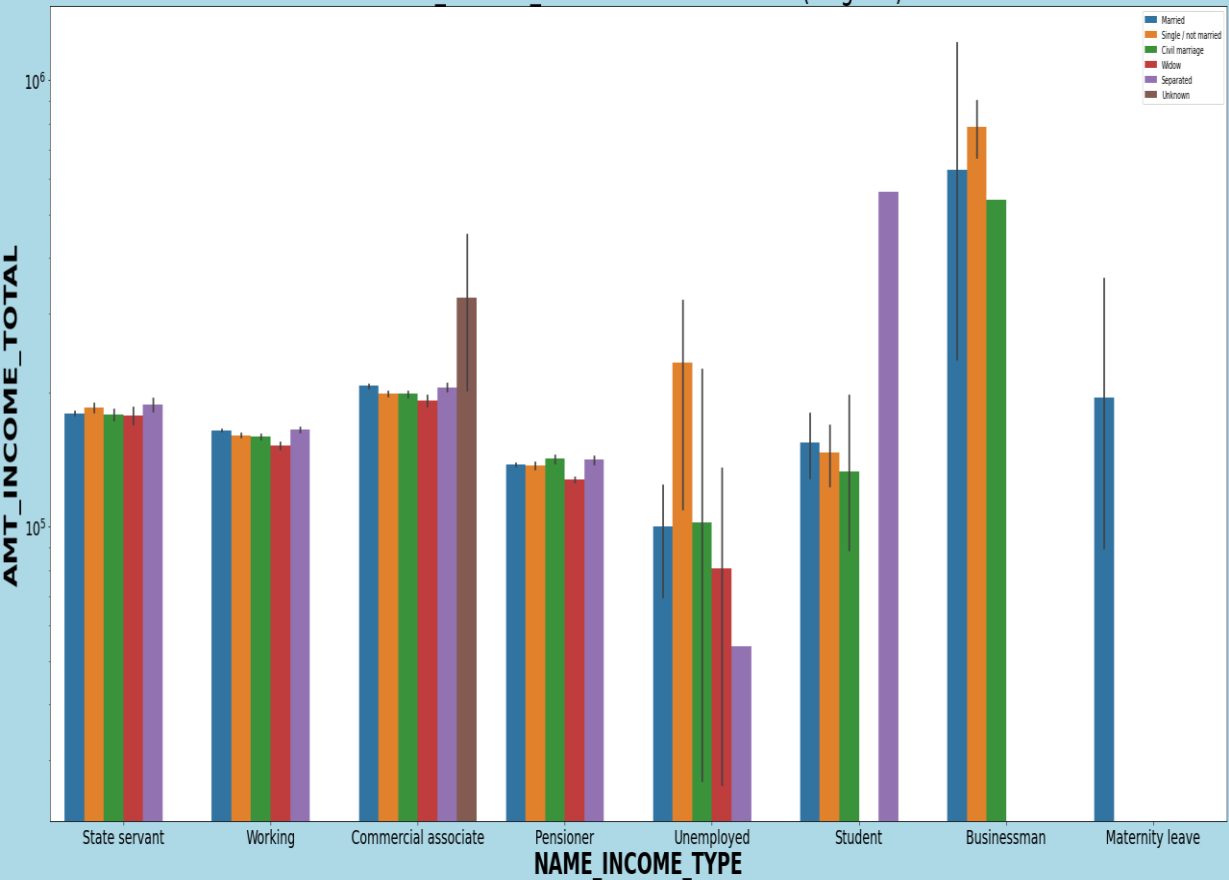
Clients occupation type
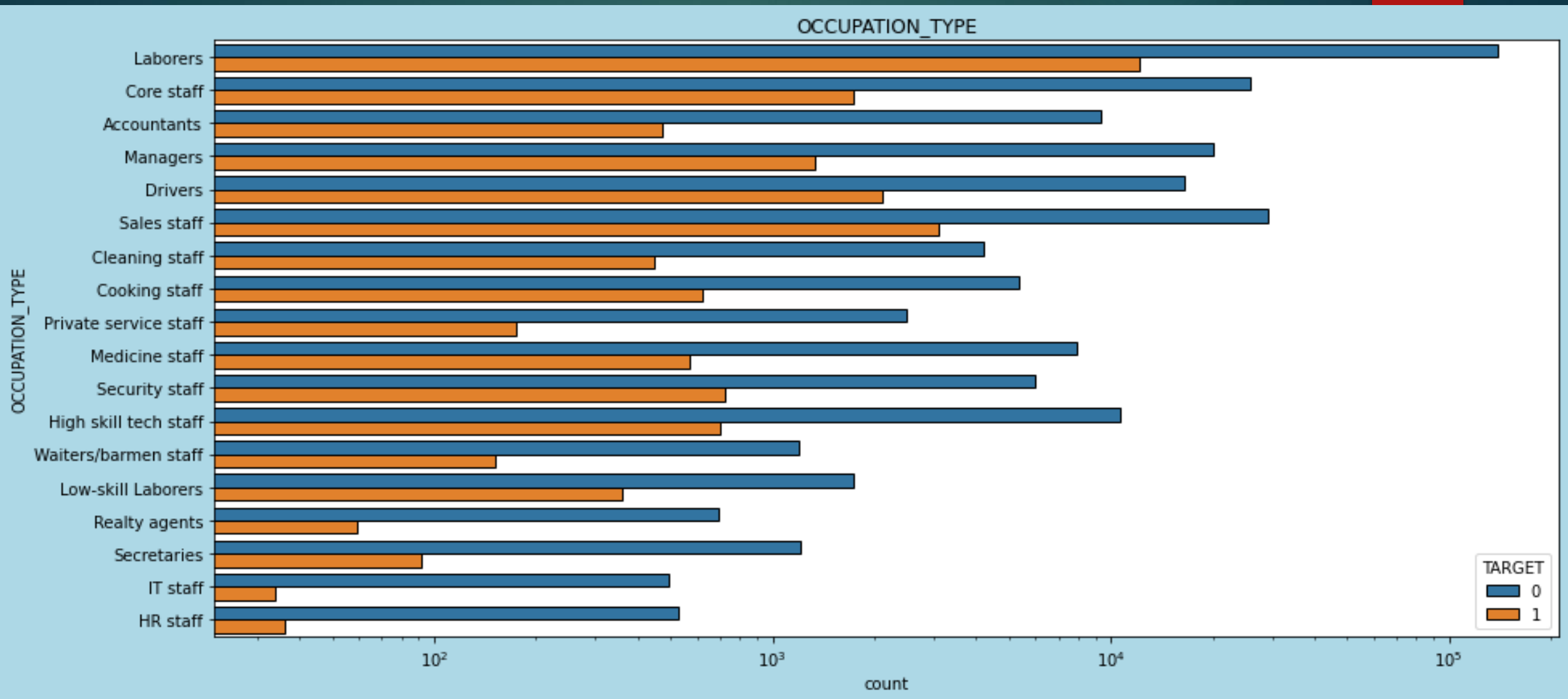
OCCUPATION_TYPE

Conclusion :
1. More people almost 20% Low-skill Laborers have high risk of to be defaulters.
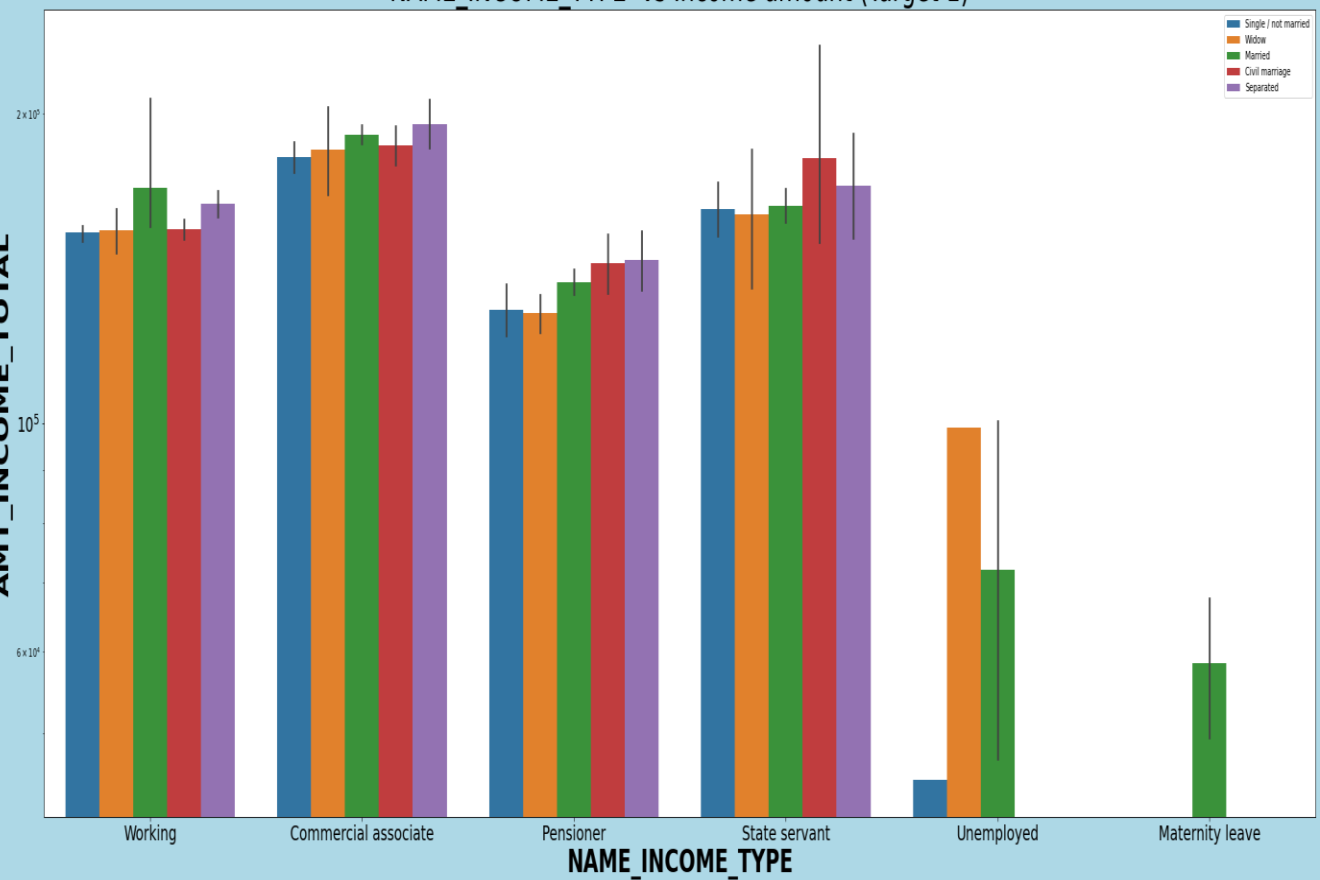2. Security staff, Waiters/barmen staff have high risk of to be defaulters.

Clients income amount vs income type based on family status
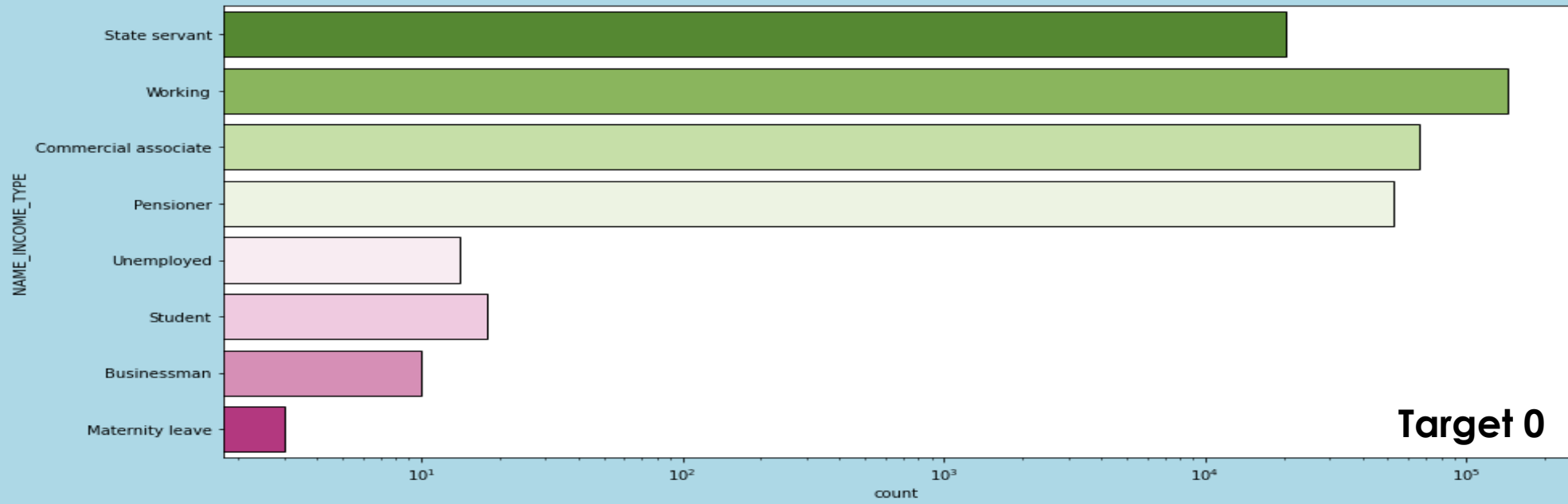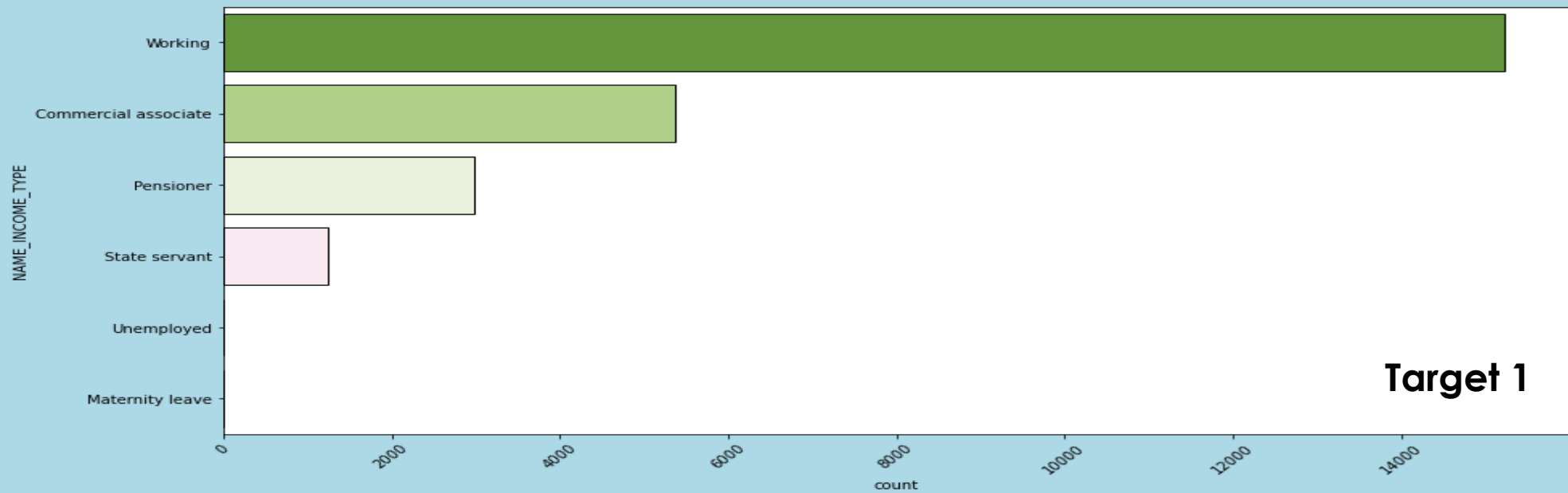

NAME_INCOME_TYPE vs Income amount (Target 0)

NAME_INCOME_TYPE vs Income amount (Target 1)

Conclusion :
1. Due to their low income, pensioners, students, and jobless people are more likely to default on their loans.
2. Unemployed people are less well-off and more likely to be at danger.
3. Married businessmen make more money than other people and never fail on their loans.
4. Due to their low income, pensioners, students, and jobless people are more likely to default on their loans.

Clients income type

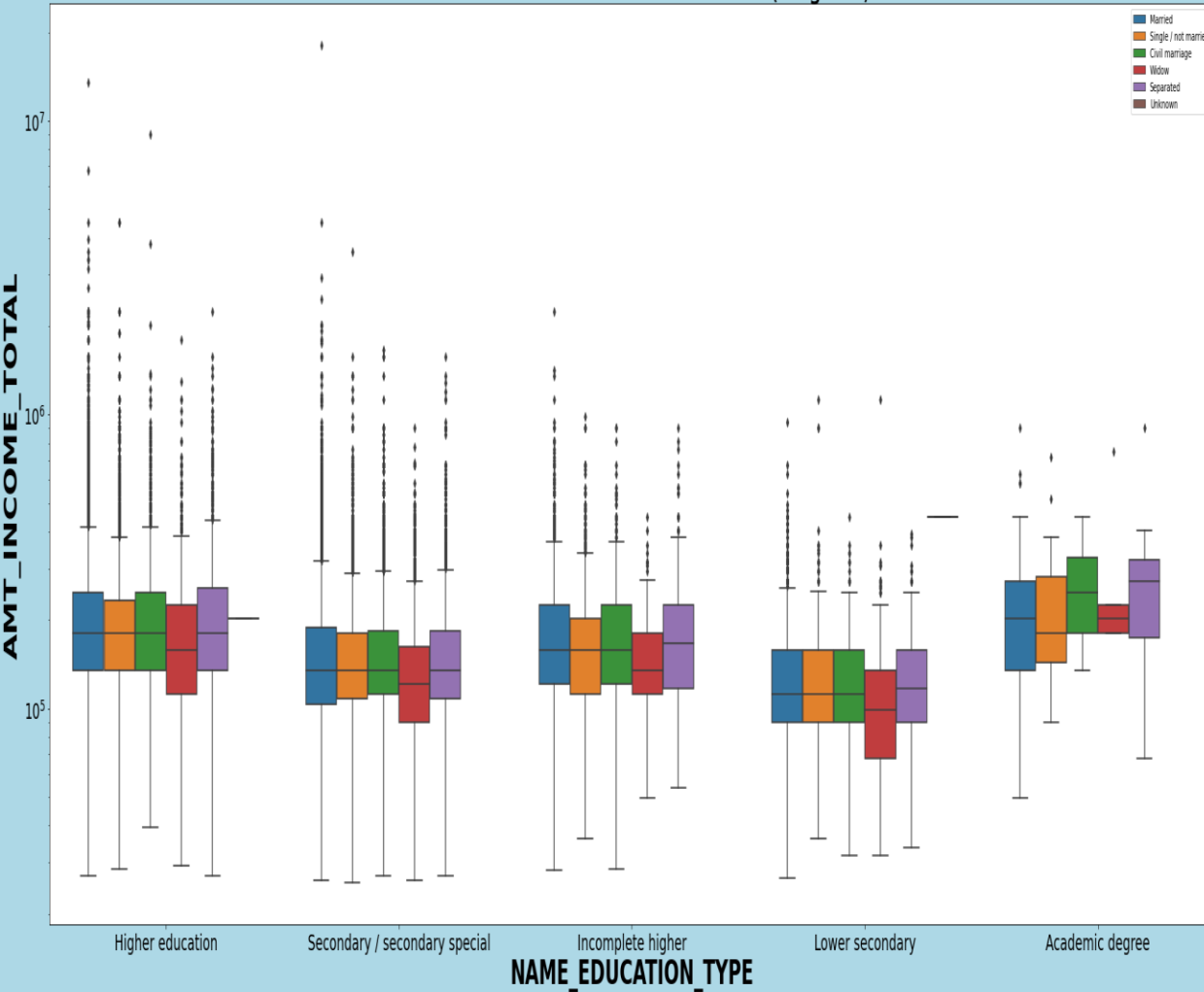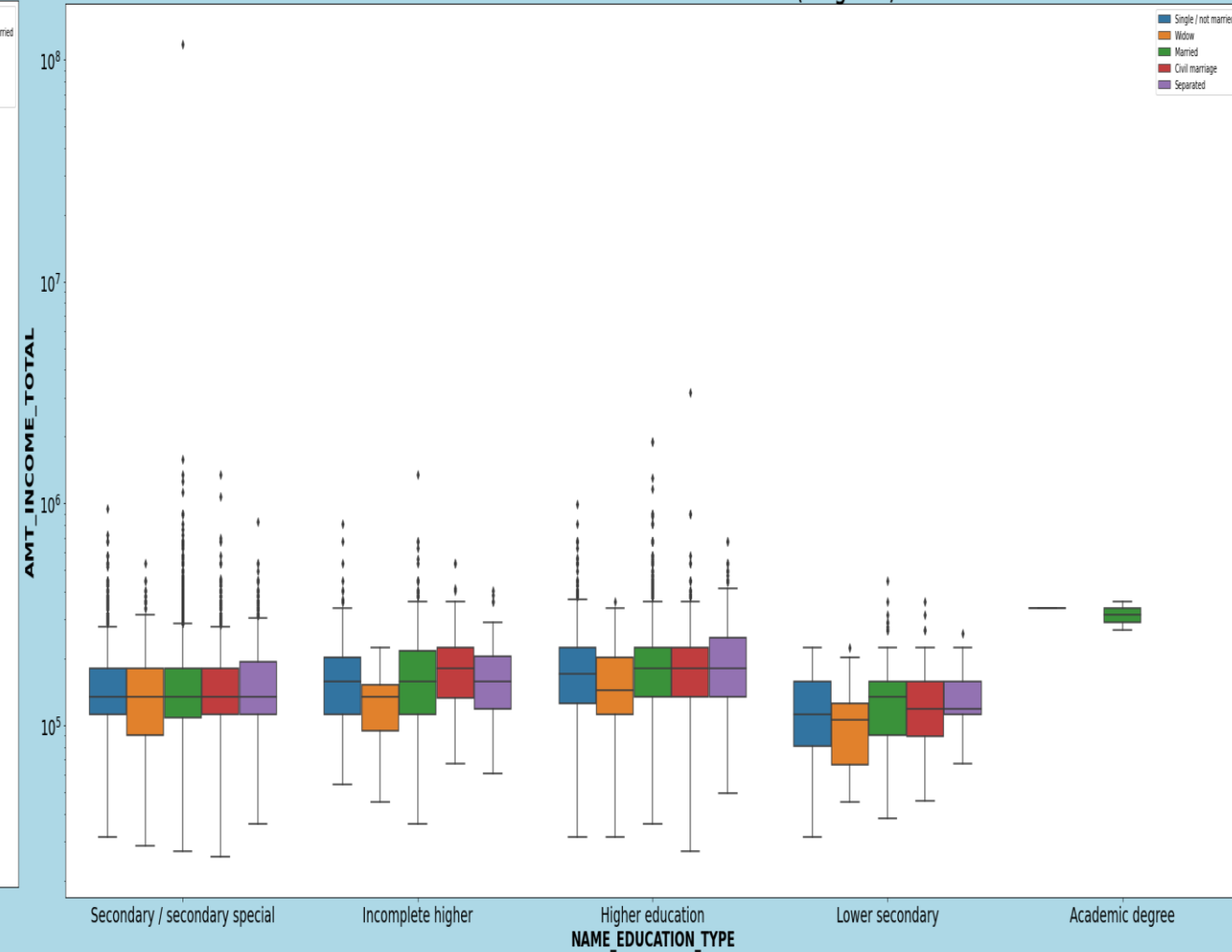

Target 1

Target 0

Conclusion:
1. More of working ,commercial associate and pensioner are Non-defaulters.
2.Businessman have average income of 1001250.00. so they can be given loan.

Clients income amount vs education

**Education Status vs Income amount(Target 0)**

Education Status vs Income amount (Target 1)

Conclusion :
1. Clients having Higher Education, Incomplete Higher Education, Lower Secondary Education and Secondary/Secondary Special have a higher number of outliers.
2. From the above figure, we can say that some of the clients having Higher Education tend to have the highest income compared to others.

INCOME WITH RESPECT TO OCCUPATION TYPE.
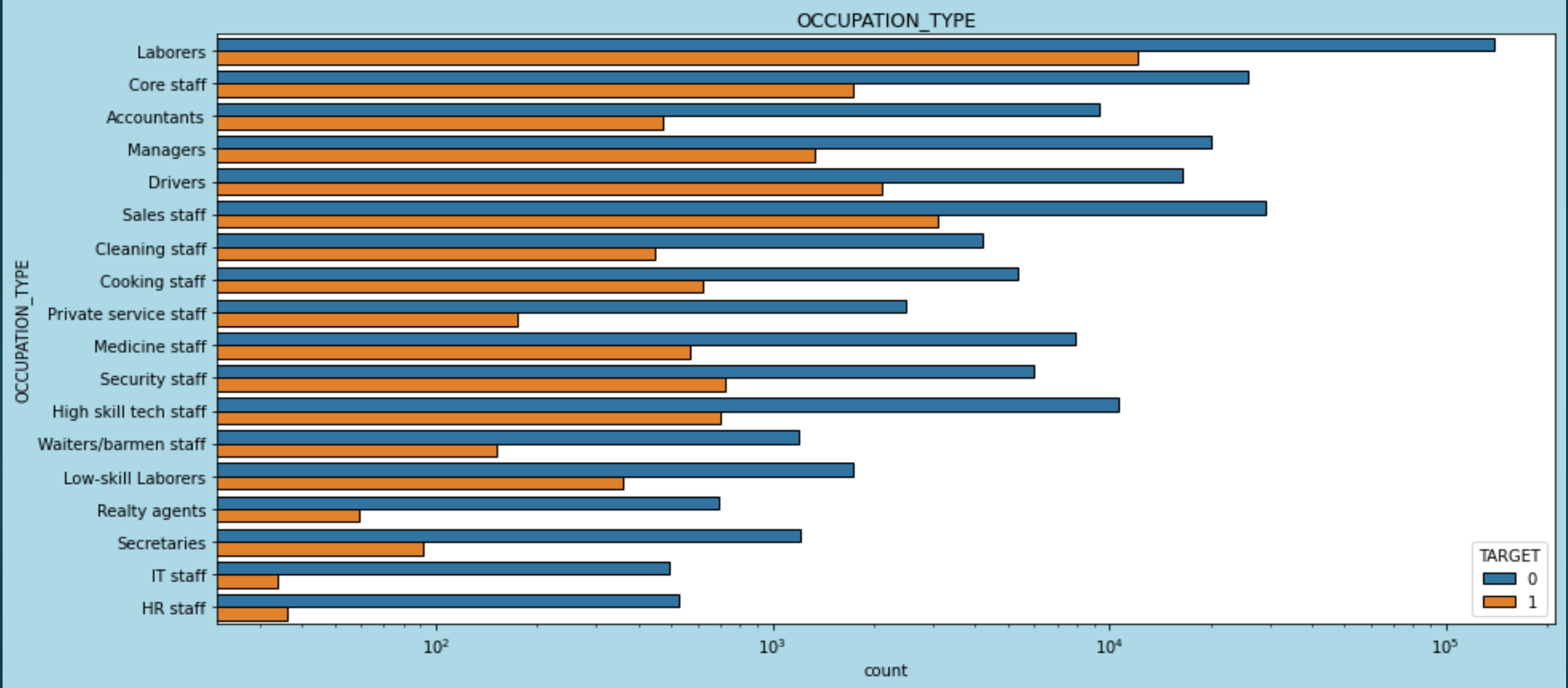
Conclusion :
1. More people almost 20% Low-skill Laborers have high risk of to be defaulters.
2. Security staff, Waiters/barmen staff have high risk of to be defaulters.

Final Conclusion

1.Applicants who are married and have a job have a lower likelihood of defaulting.

2.We can offer loans to new consumers based on their income and occupation, taking into account their average revenue from various occupations.

3.There is a good chance that an applicant with a business or commercial associate income category won't default.

4.30 percent of borrowers who declined to say why they were in arrears.

5.Interest, Cash for an outsider,loan payments on other accounts,gasification and water provision,Auto repairs could have a high default rate.

6.A higher educated individual with an academic degree need to be given preference.

7.There is a significant default risk for male consumers with incomplete education and extremely low earnings.

8.57% of those having an income type of unemployment are at high risk of defaulting.

THANK YOU