

# CREDIT EDA ASSIGNMENT



PRESENTED BY –  
Mr. MONISH LALANI





# 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)**

### **A. Data Understanding**

1. Reading Data
2. Checking for the null values
3. Removing the column with more than 35% of null values
4. Handling Nulls
5. Separating in categorical and numerical columns

### **B. Univariate analysis for Numerical variables**

### **c. Univariate analysis for Categorical variables**

### **D. Univariate analysis of numerical column for TARGET=0 (Applicants making payments on time)**

### **E. Univariate analysis of categorical column for TARGET=0 (Applicants making payments on time)**

### **F. Univariate analysis of numerical column for TARGET=1 (Applicants not making payments on time)**

### **G. Univariate analysis of categorical column for TARGET=1 (Applicants not making payments on time)**

### **H. Bivariate Analysis for categorical columns**

## **Second Data set (previous\_application)**

### **A. Data Understanding**

1. Reading Data
2. Checking for the null values
3. Removing the column with more than 35% of null values
4. Handling Nulls

### **B. Merging both data sets.**

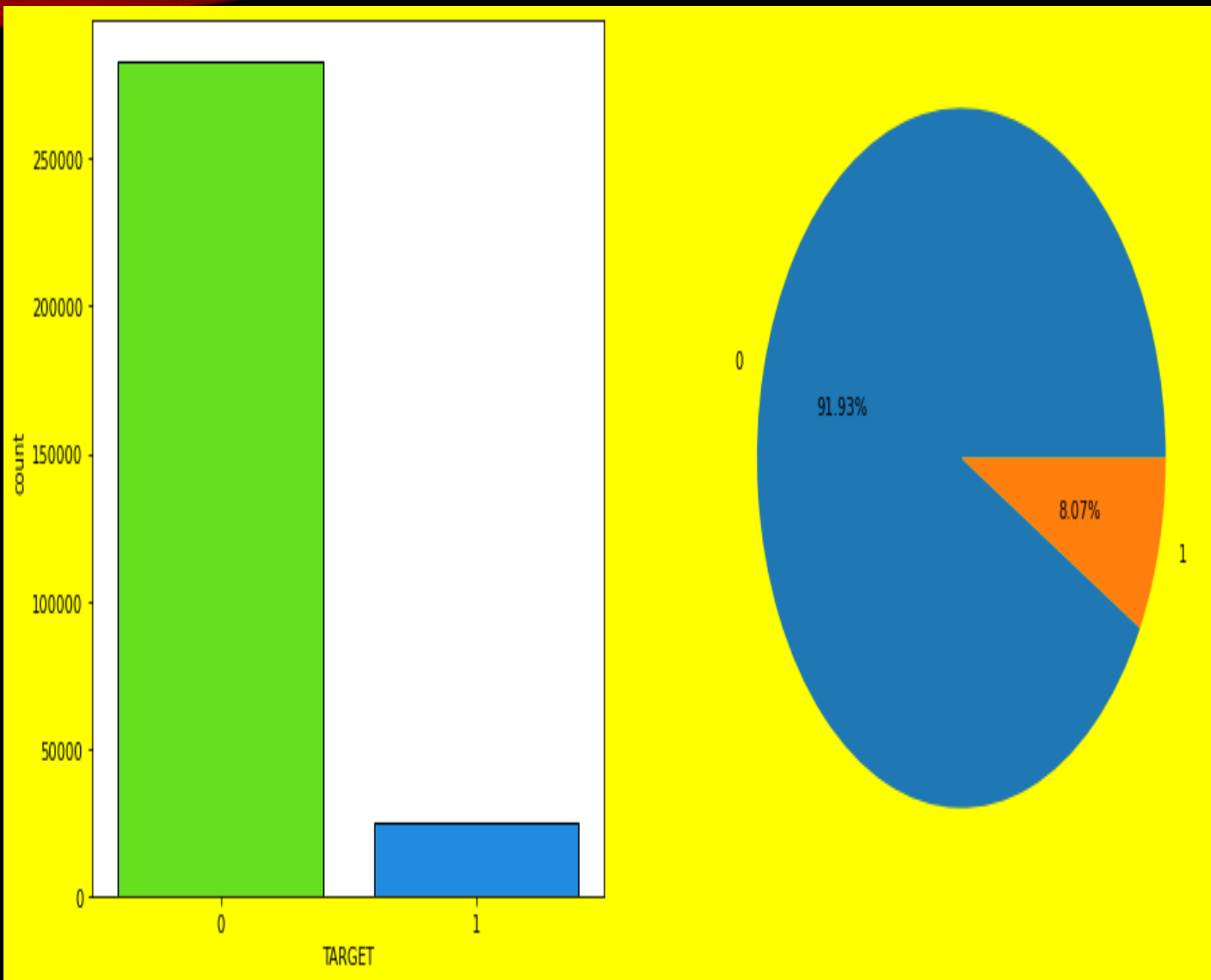
### **C. Separating in categorical and numerical columns**

### **D. separating on the basis on target column.**

### **E. Univariate and Bivariate analysis**

### **F. Final Conclusion.**

# ANALYSIS ON TARGET

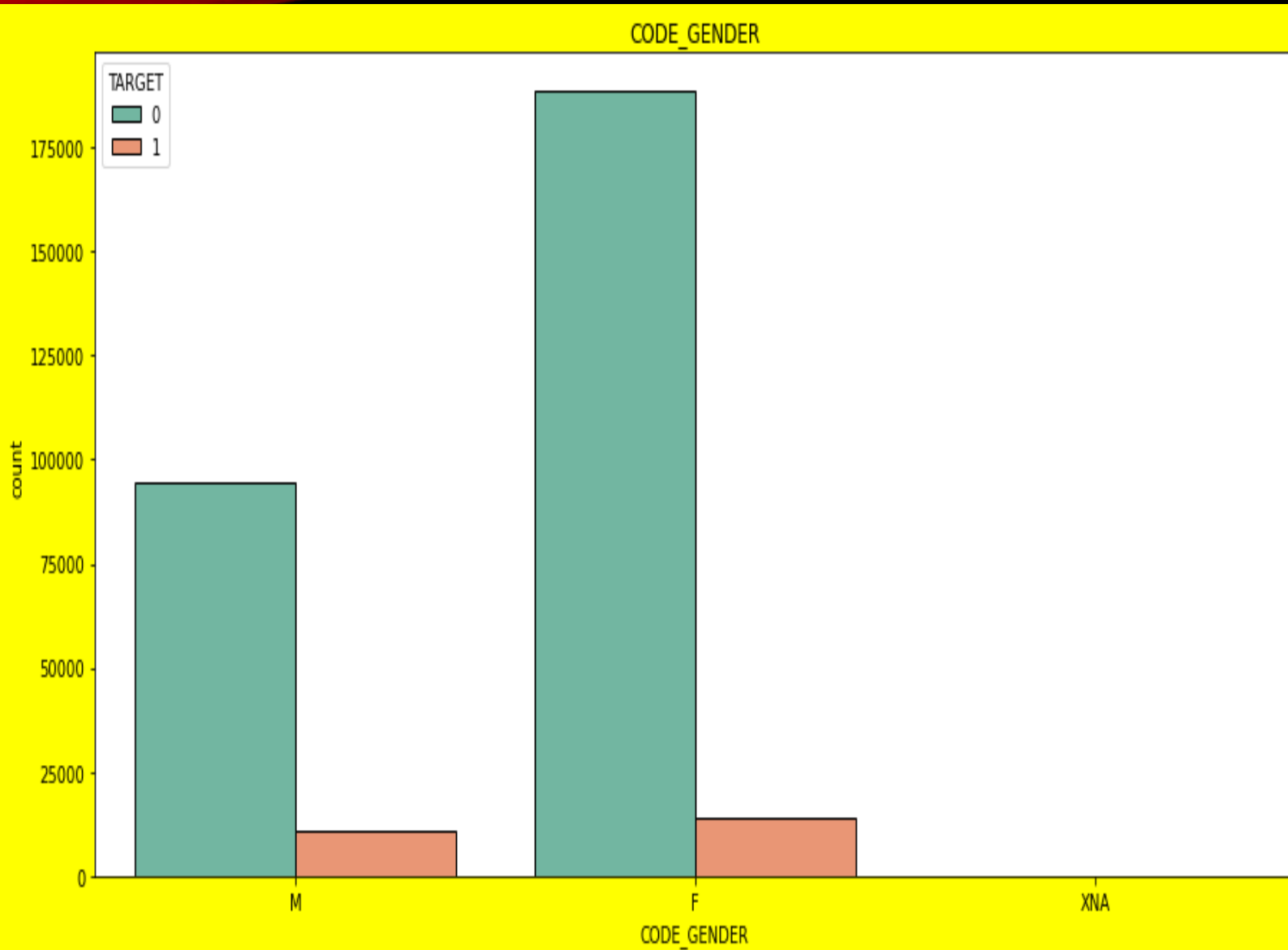


TARGET	Count of clients
TARGET 1	24825
TARGET 0(NON DEFAULTER)	282686

Conclusion:

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

# GENDER DISTRIBUTION IN TARGET

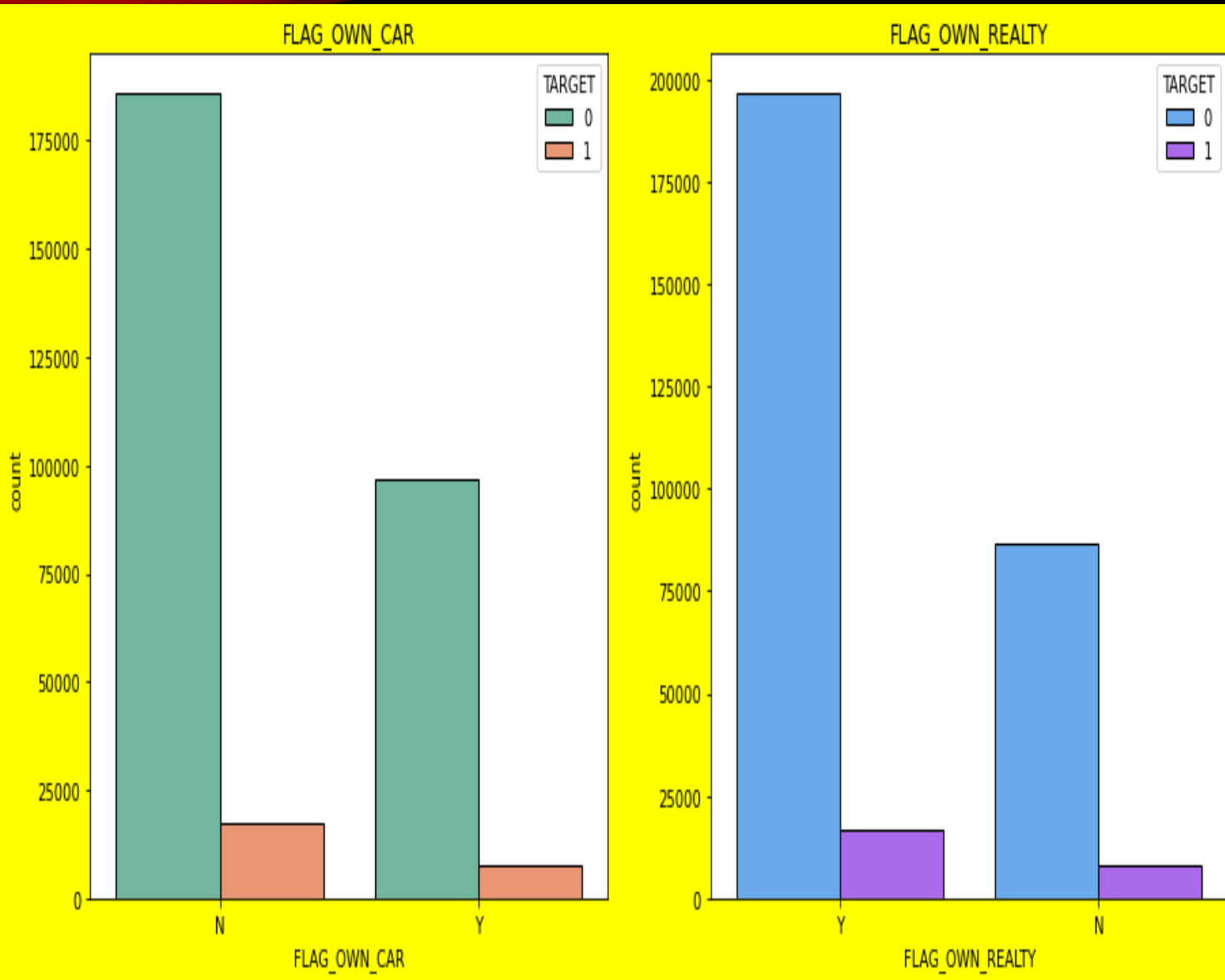


## **conclusion:**

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



# CLIENTS WITH AND WITHOUT HOUSE AND CAR

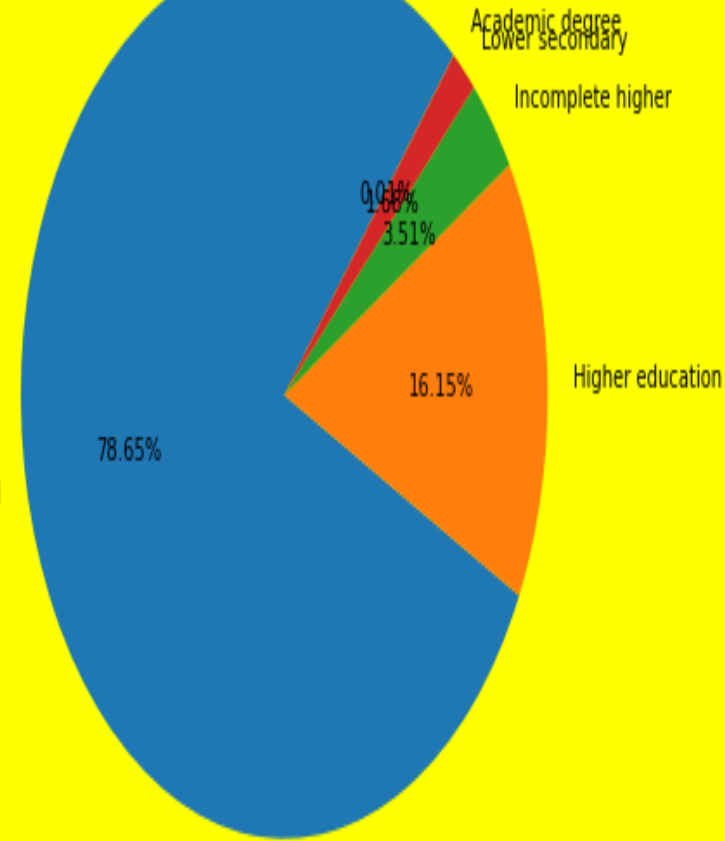
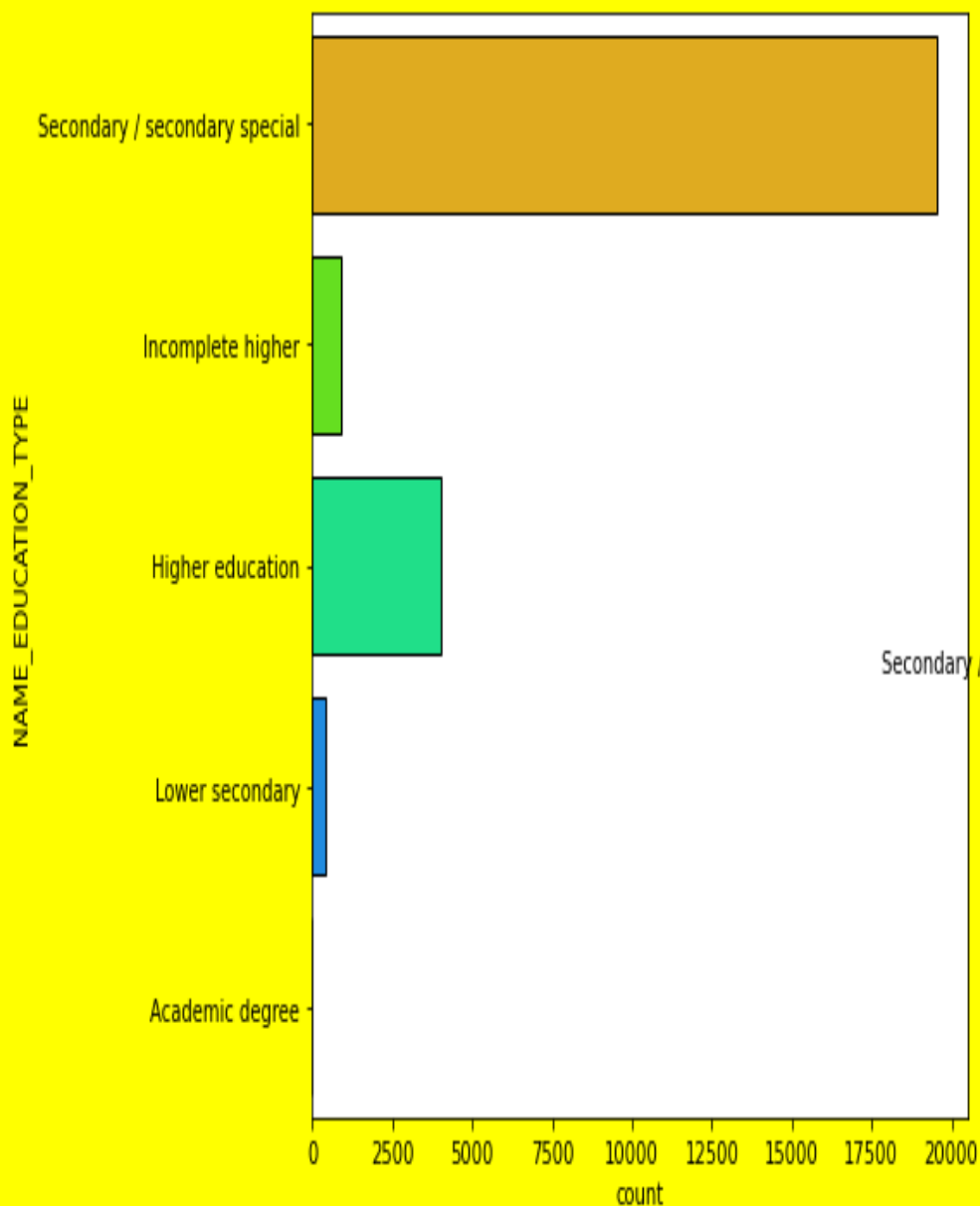


ASSETS	YES	NO
HOUSE	196329	30672
CAR	97011	185675

Conclusion:

1. We can say that clients with their own house are more and this can be used as property collateral while providing a loan to the customer. And are non defaulters.
2. People with property have made more payments on time.
3. People with no car have applied more for a loan.

# CLIENTS EDUCATION TYPE



Secondary / secondary special	19524
Higher education	4009
Incomplete higher	872
Lower secondary	417
Academic degree	3

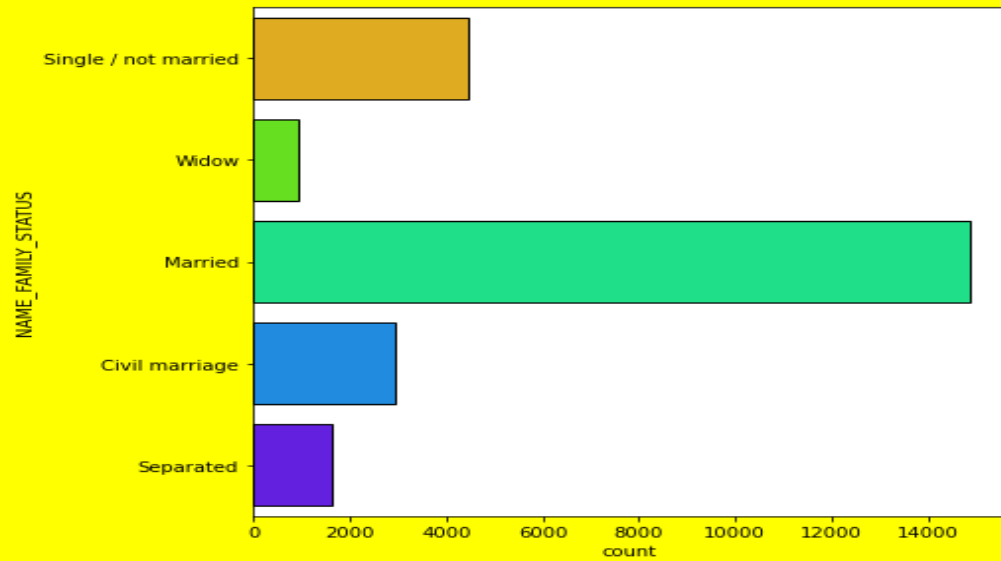
## Conclusion:

1. Most of the clients are secondary school passout, with higher education and so on.

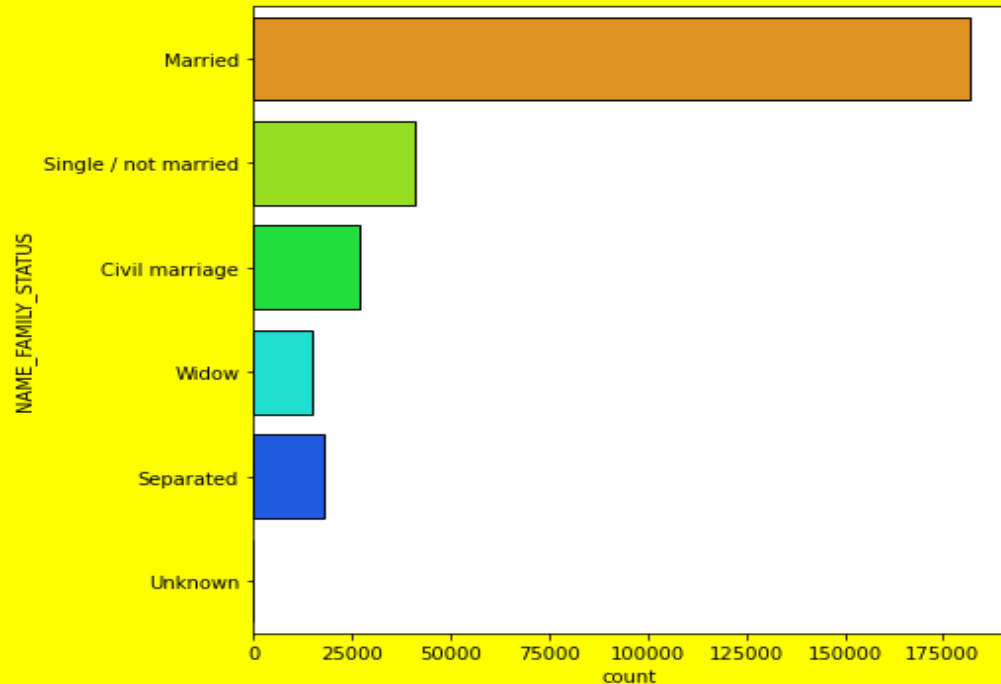
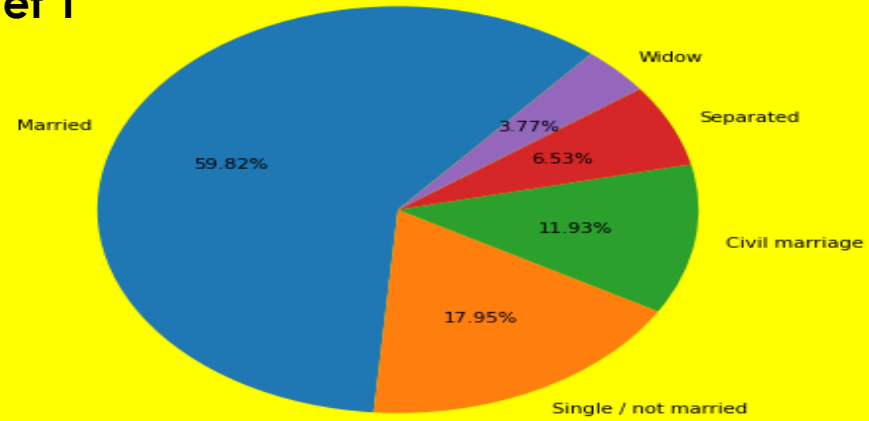
2. 76% of secondary school passout are **Non-defaulters**.



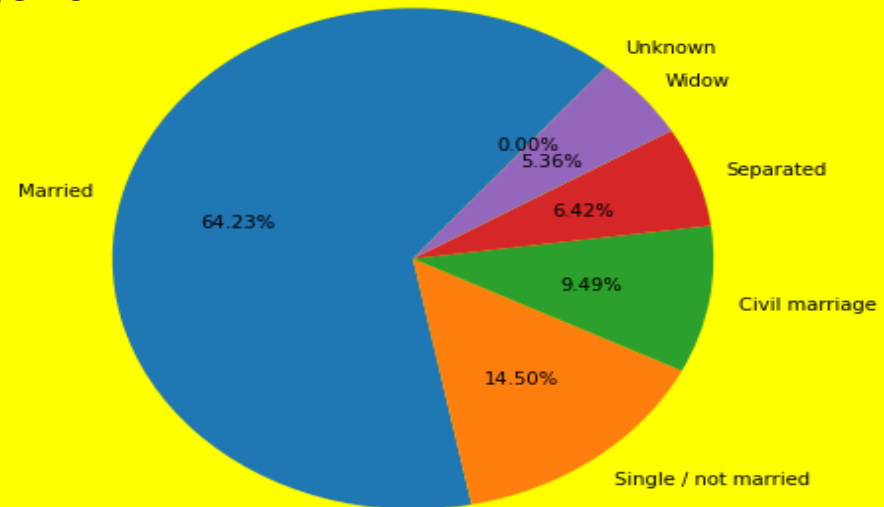
# CLIENTS FAMILY TYPE



Target 1



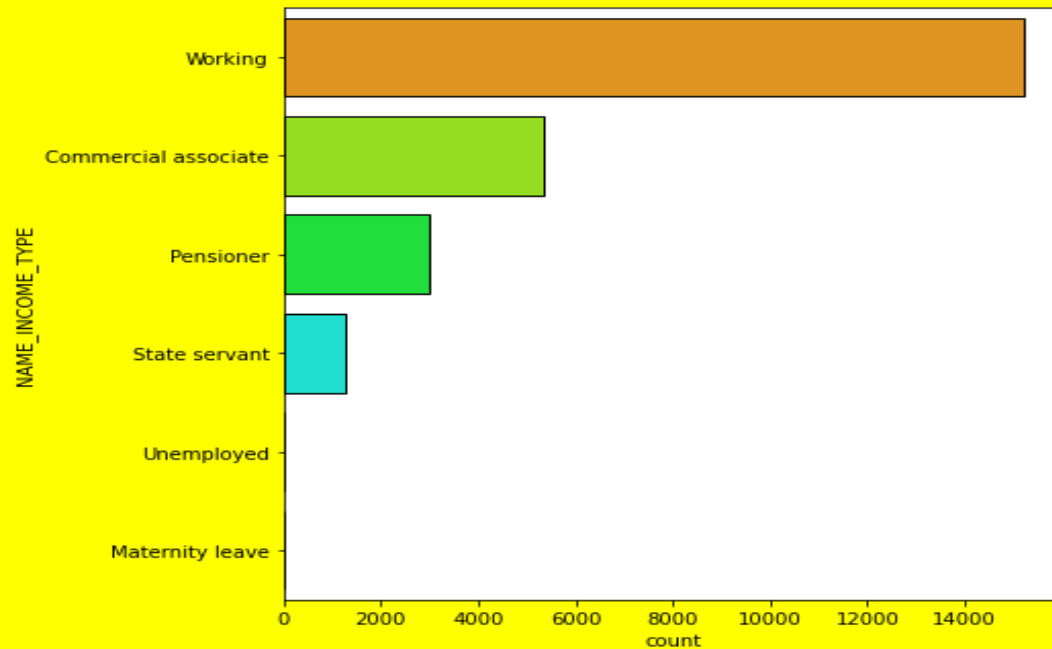
Target 0



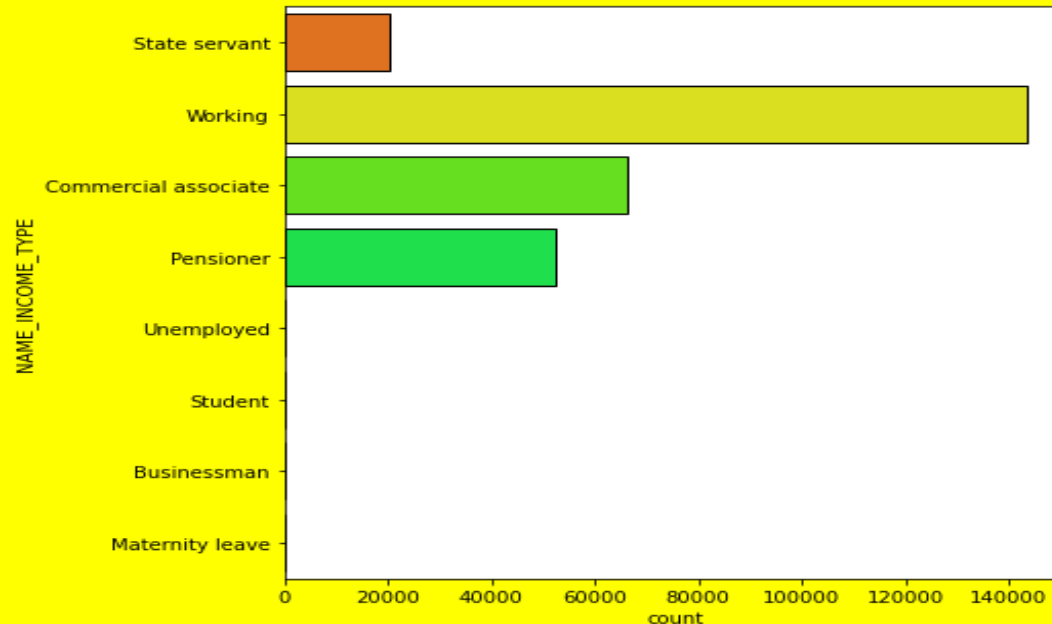
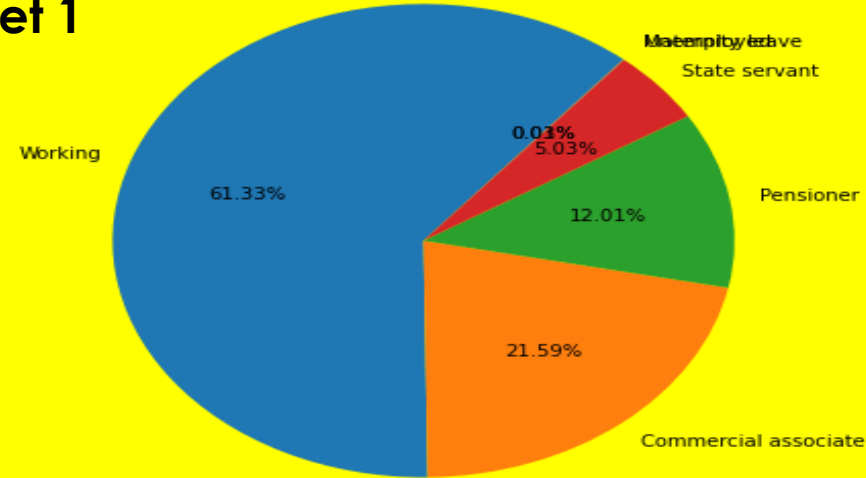
## Conclusion:

1. 64 % of married clients are **Non-defaulters**.
2. 14 % of single and 9.77 % are **Non-defaulters**.

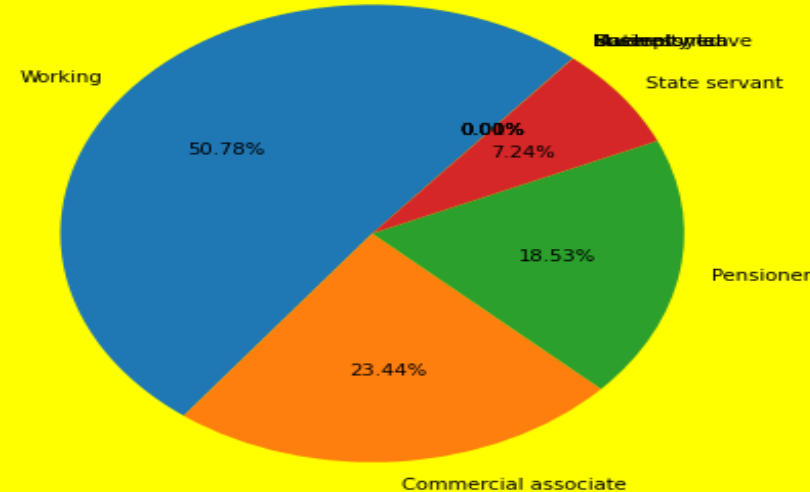
# CLIENTS INCOME TYPE



Target 1



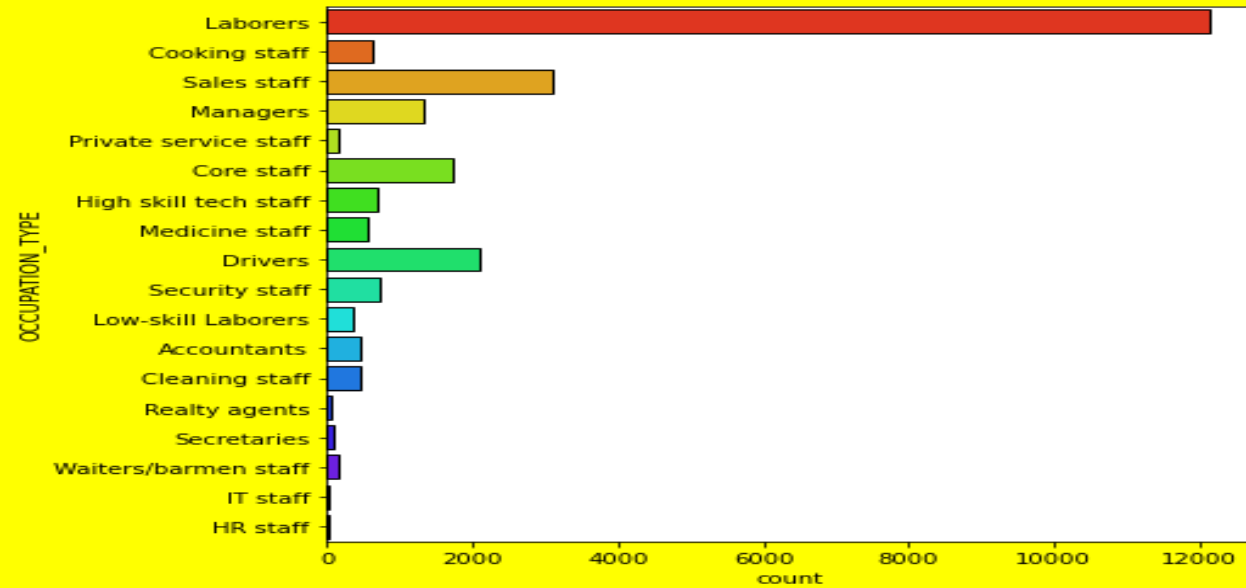
Target 0



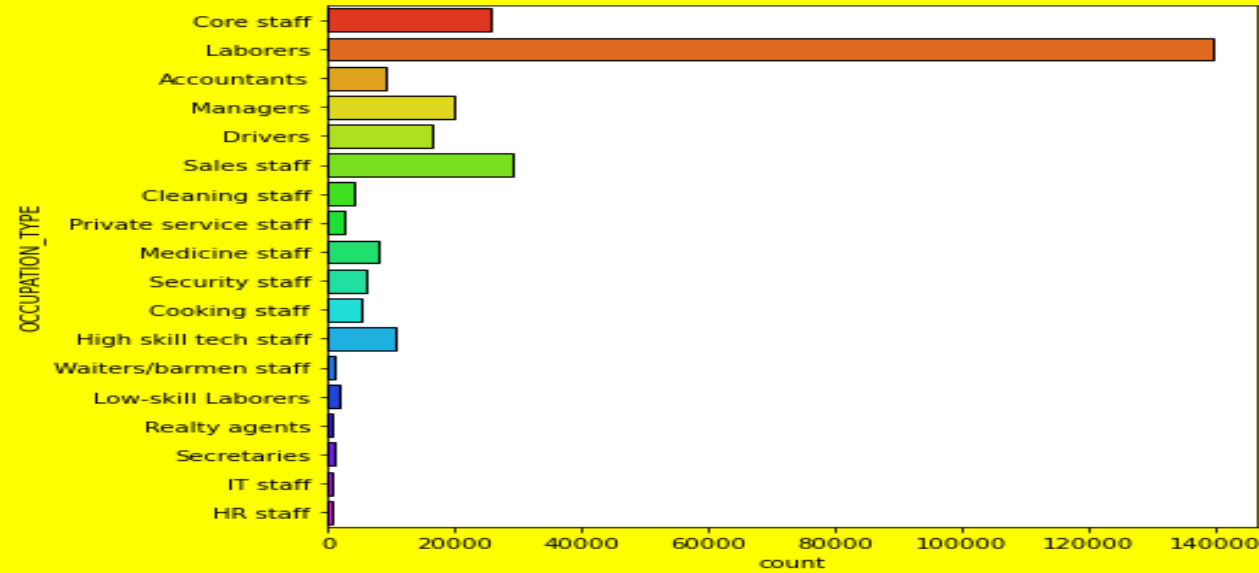
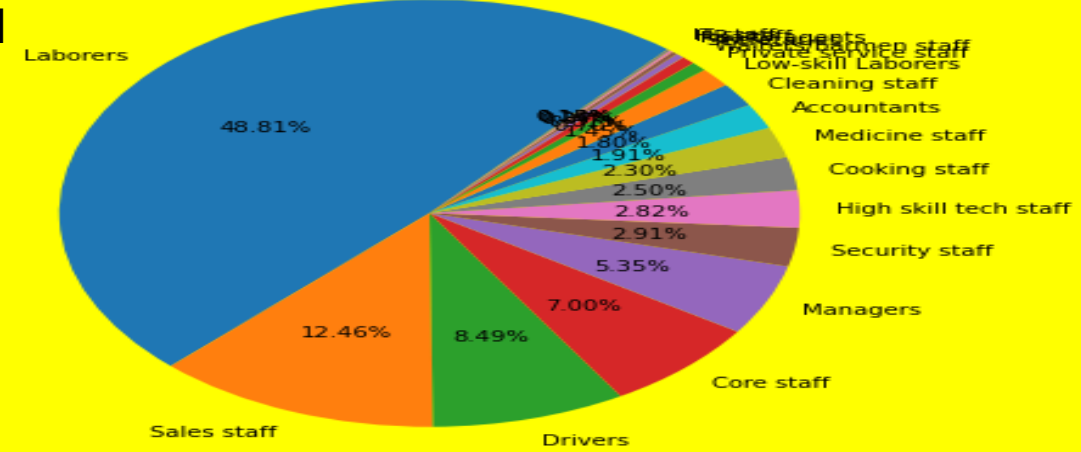
Conclusion:

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.

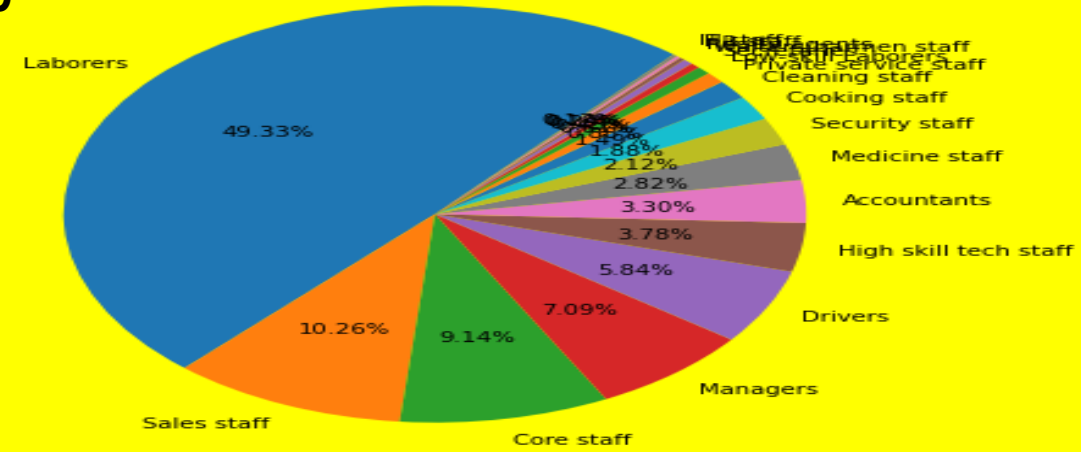
# CLIENTS OCCUPATION TYPE



Target 1

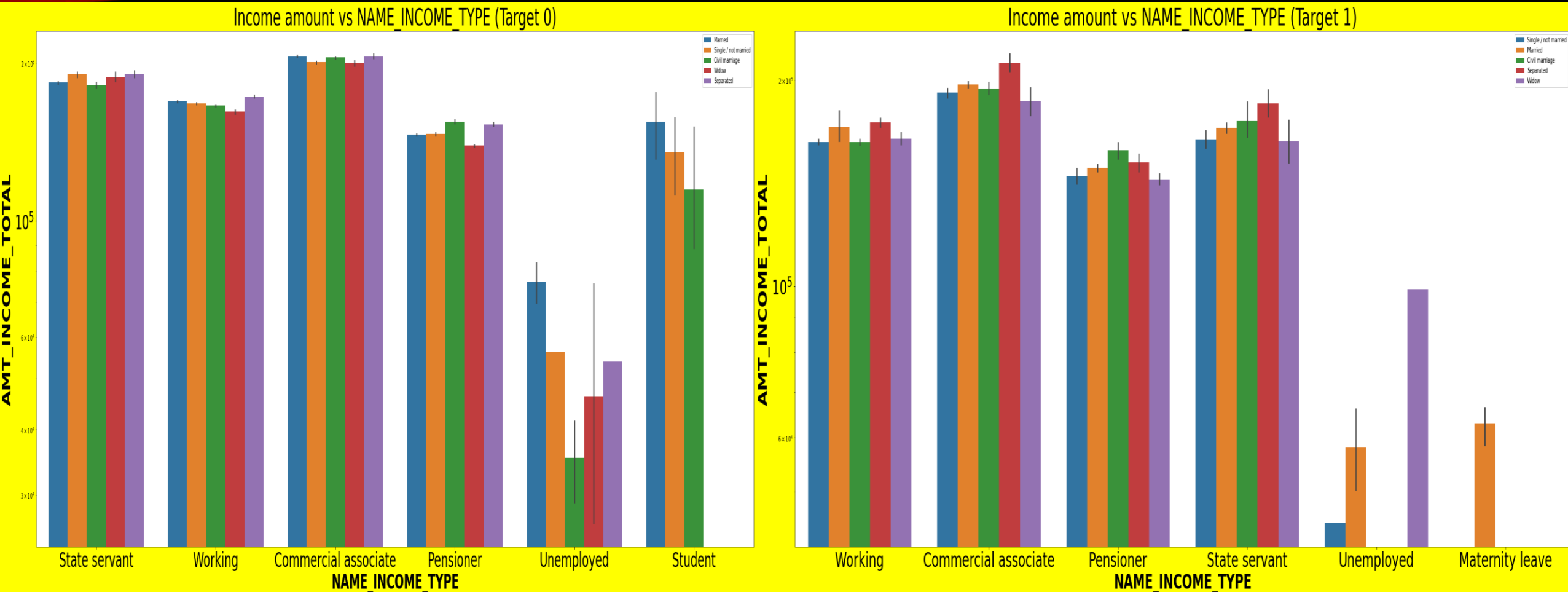


Target 0



1.49% of people have OCCUPATION\_TYPE of laborers.  
2.Sales staff are 12%

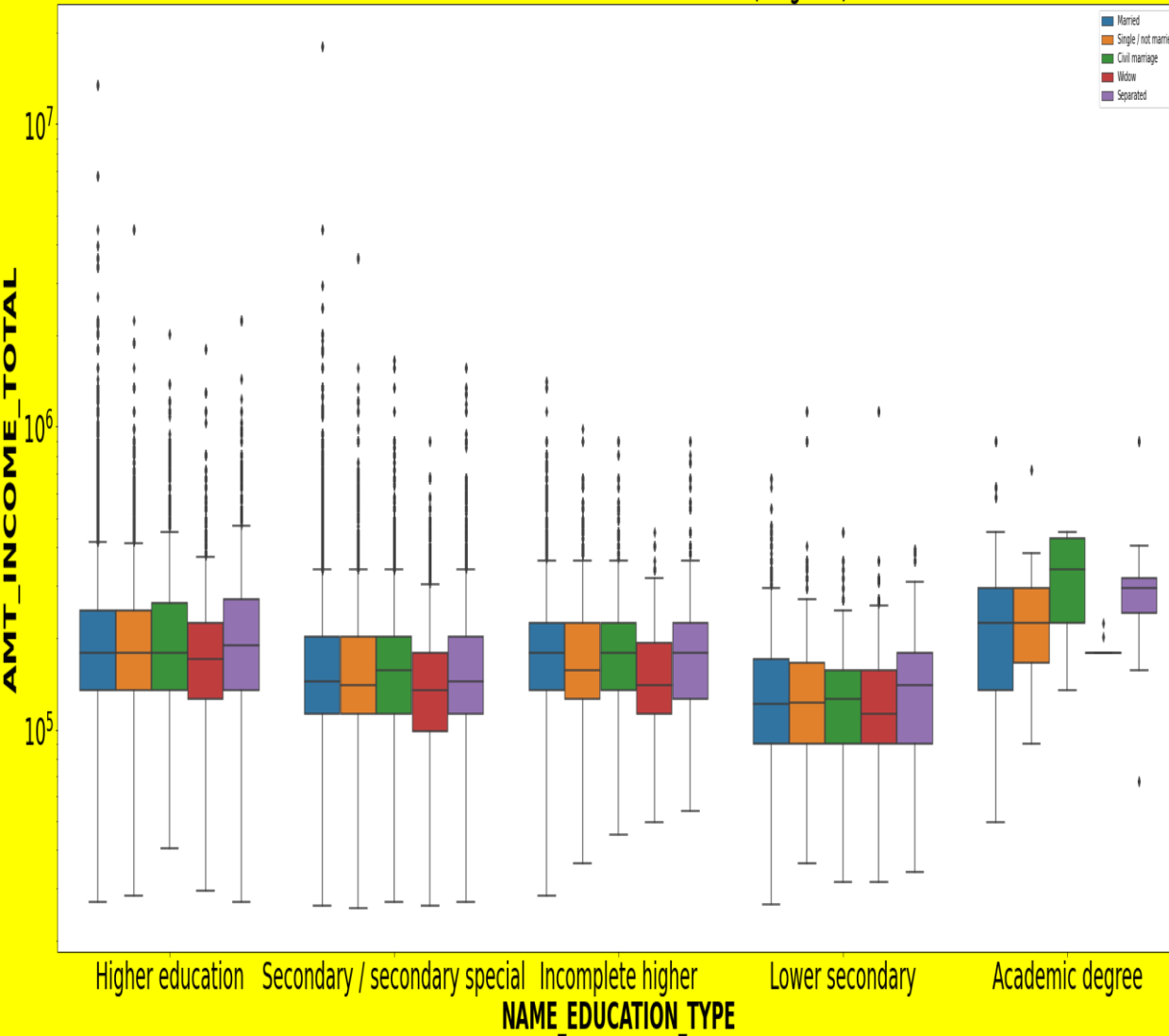
# CLIENTS INCOME AMOUNT VS INCOME TYPE BASED ON FAMILY STATUS



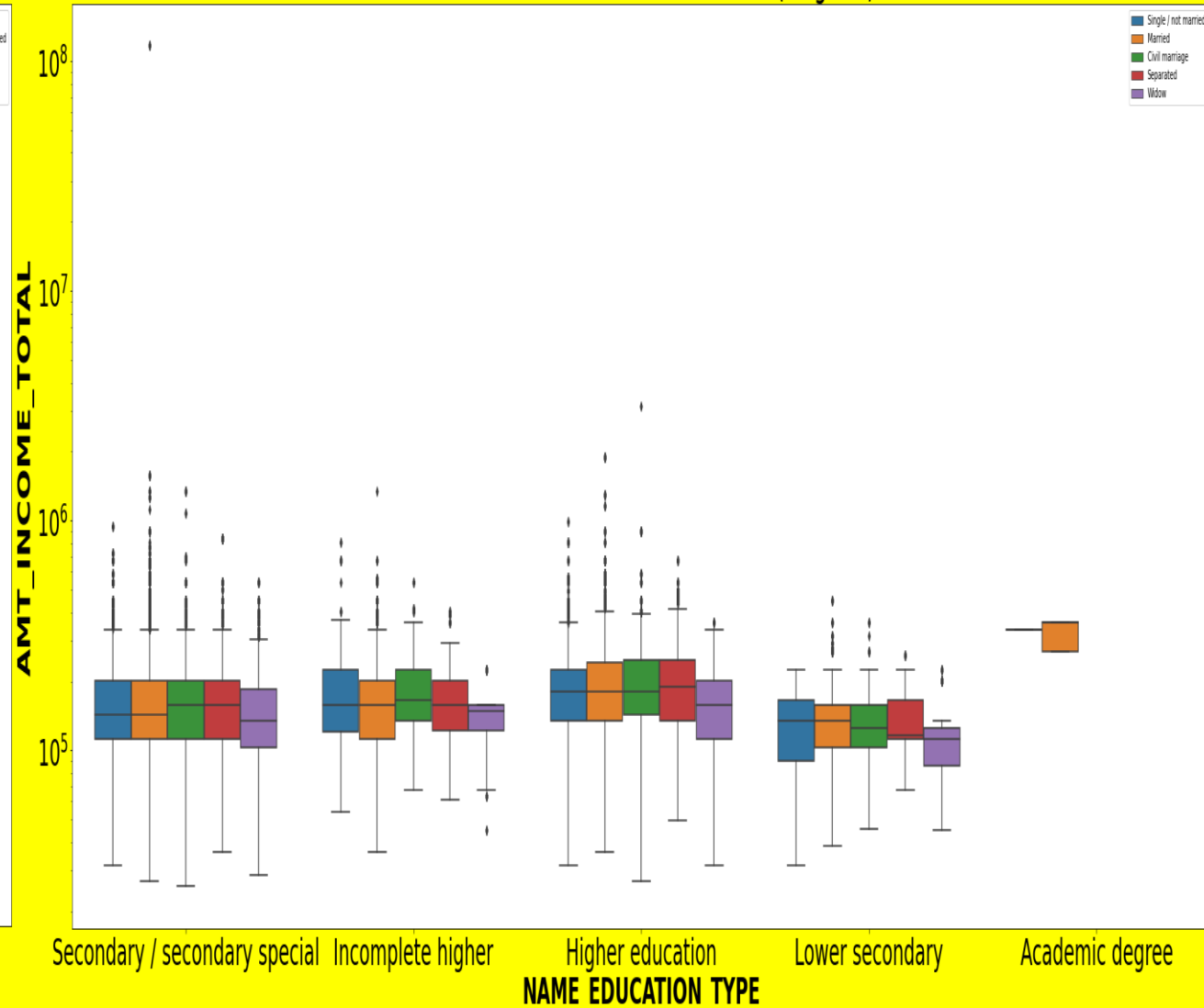
1. Business man , and commercial associate are likely to apply more for loan and are **Non-defaulters**
2. Married Business man have more income compared to others and are non defaulters.
3. While pensioner ,student and unemployed have high risk of defaulters due to low income.

# CLIENTS INCOME AMOUNT VS EDUCATION

Income amount vs Education Status(Target 0)



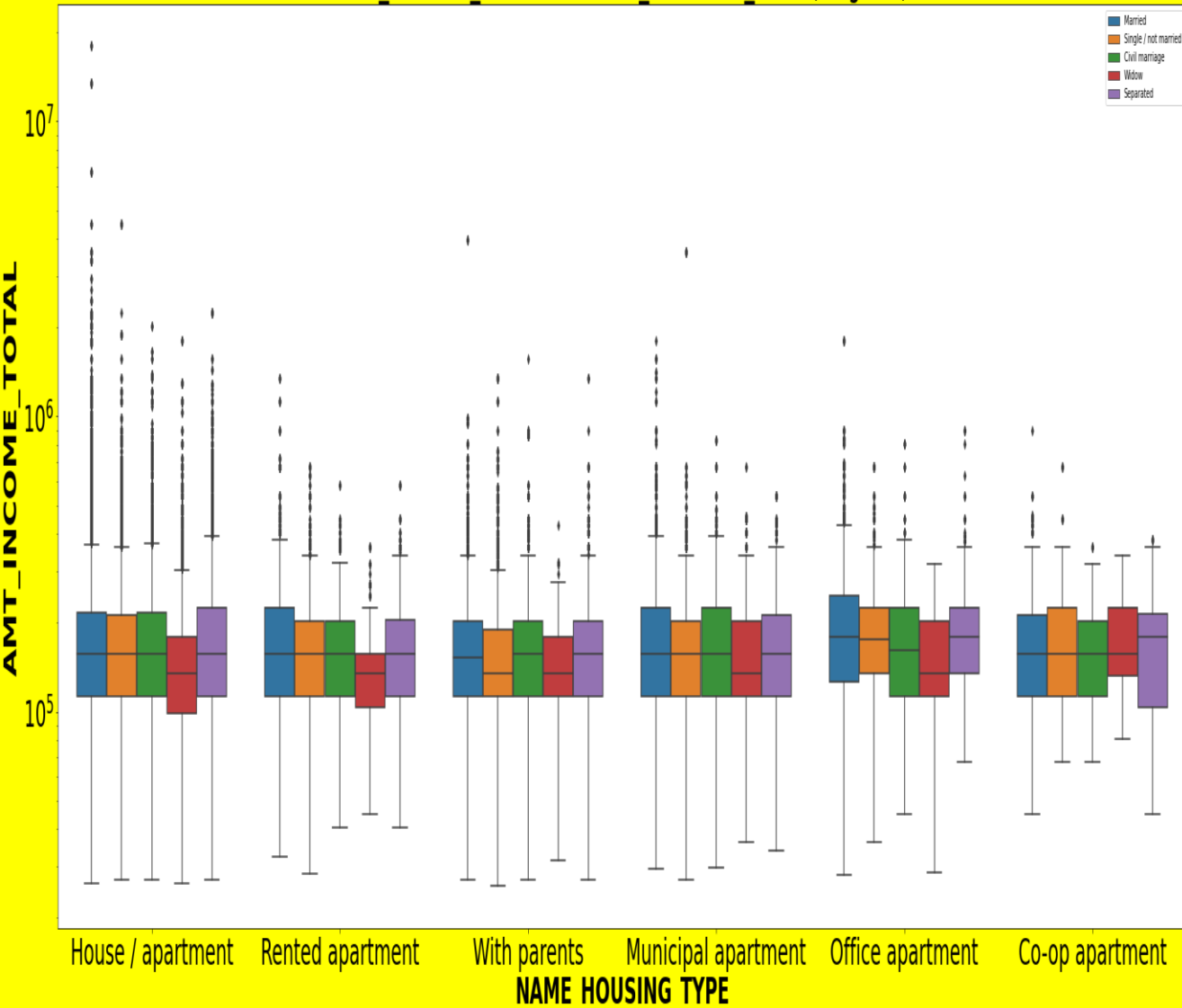
Income amount vs Education Status(Target 1)



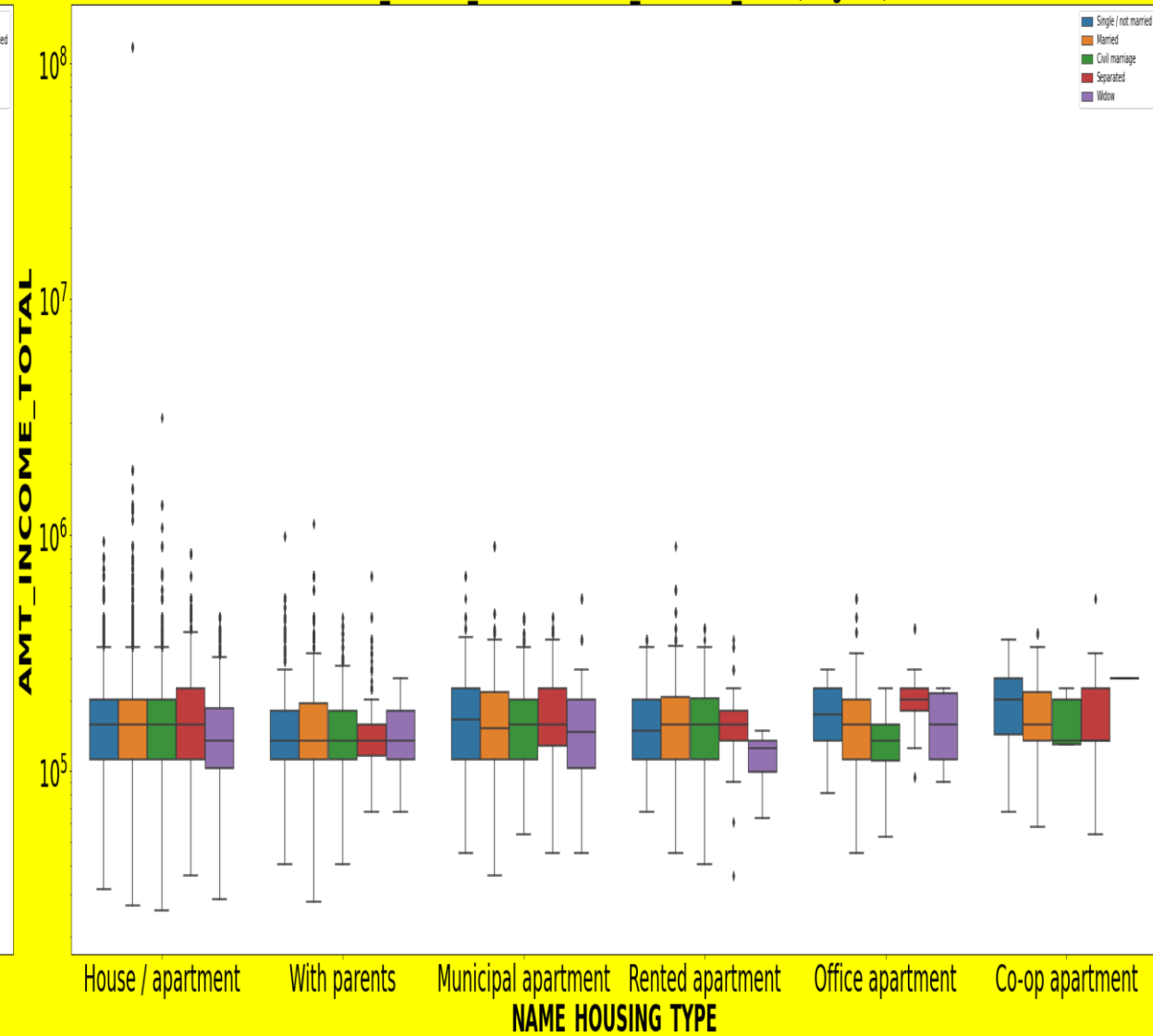
1. Clients with Academic degree have no outliers. While other every clients
2. From the above figure, we can say that some of the clients having Higher Education tend to have the highest income compared to others.

# AVGERAGE INCOME WITH RESPECT TO OCCUPATION TYPE.

AMT\_INCOME\_TOTAL vs NAME\_HOUSING\_TYPE (Target 0)



AMT\_INCOME\_TOTAL vs NAME\_HOUSING\_TYPE (Target 1)



**BASED ON THERE AVGERAGE INCOME OF DIFFERENT OCCUPATION WE CAN PROVIDE LOAN TO THE NEW CLIENTS AS PER AS THERE INCOME AND OCCUPATION.**



# CONCLUSION

1. Person with **Higher education** , **Academic degree** should be preferred.
2. Applicant with income type of **Business** and **commercial associate** are likely to be non default.
3. Person with **House** as an asset can be preferred as a asset collateral and can we given loan.
4. BASED ON THERE AVGERAGE INCOME OF DIFFERENT OCCUPATION WE CAN PROVIDE LOAN TO THE NEW CLIENTS AS PER AS THERE INCOME AND OCCUPATION.
5. Applicants who are employed and are **married** have lesser chance to be defaulter.
6. Male clients with **Incomplete Education** having **very low salaries** have a **high risk of default**.
7. 30% of people of loan purpose Refusal to name the goal were defaulters.
8. Hobby, Money for a third person, Payments on other loans, Gasification / water supply ,Car repairs may have high chance to be defaulters
9. 57% people with income type Unemployed have high chance to be defaulters.



**THANK YOU**