



Risk Analytics in Banking and Financial services

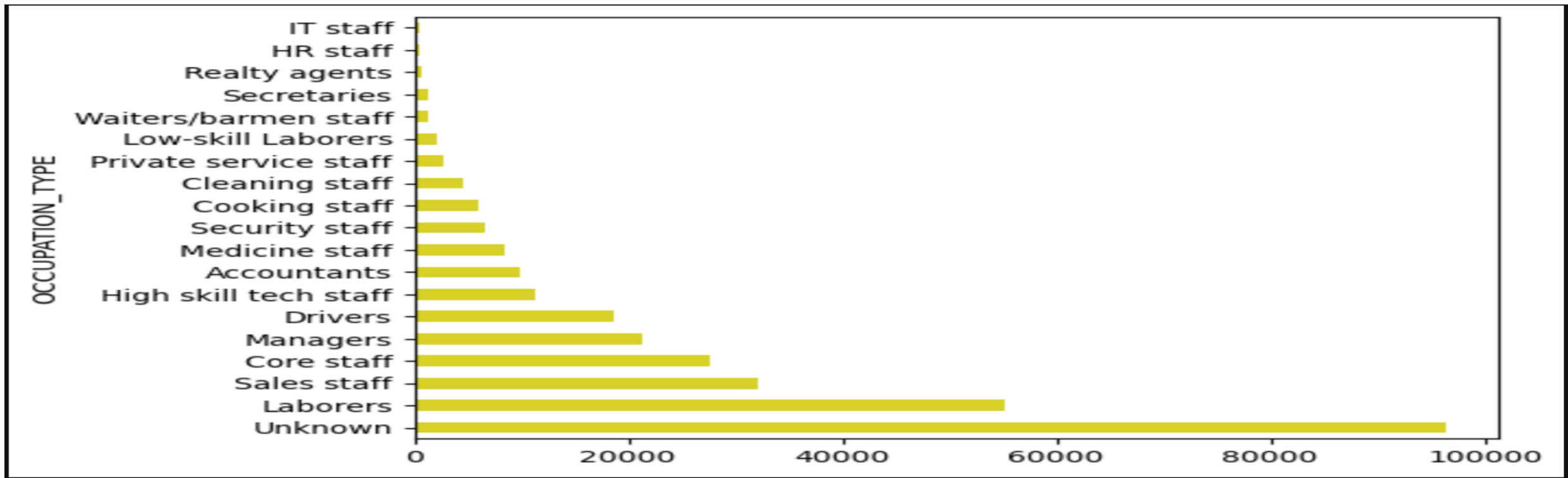
K.RAMA KRISHNA PRASANNA KUMAR

Problem Statement

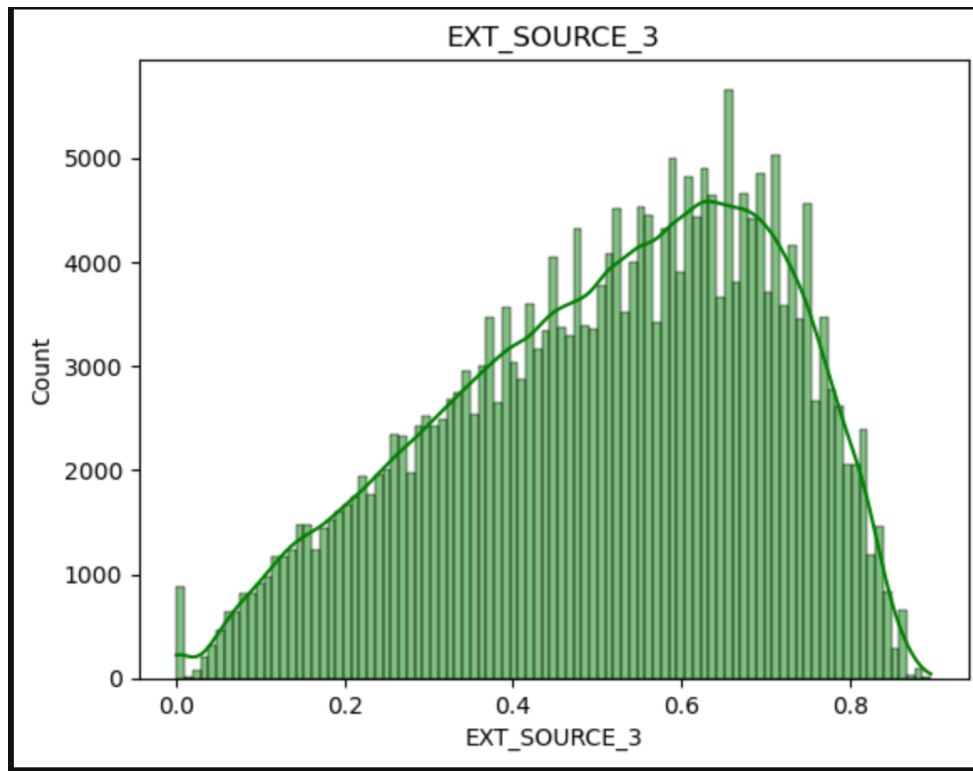
- In the credited EDA assignment we get two data and we perform analysis on the data sets and provide some meaningful insights to bank.
- When a client apply a loan then with help of our analysis that finance company will take discussion either approved or reject the loan.

Application Data Set

After filling the null values in the occupation type we apply bar chart to that variable find visualize the count.

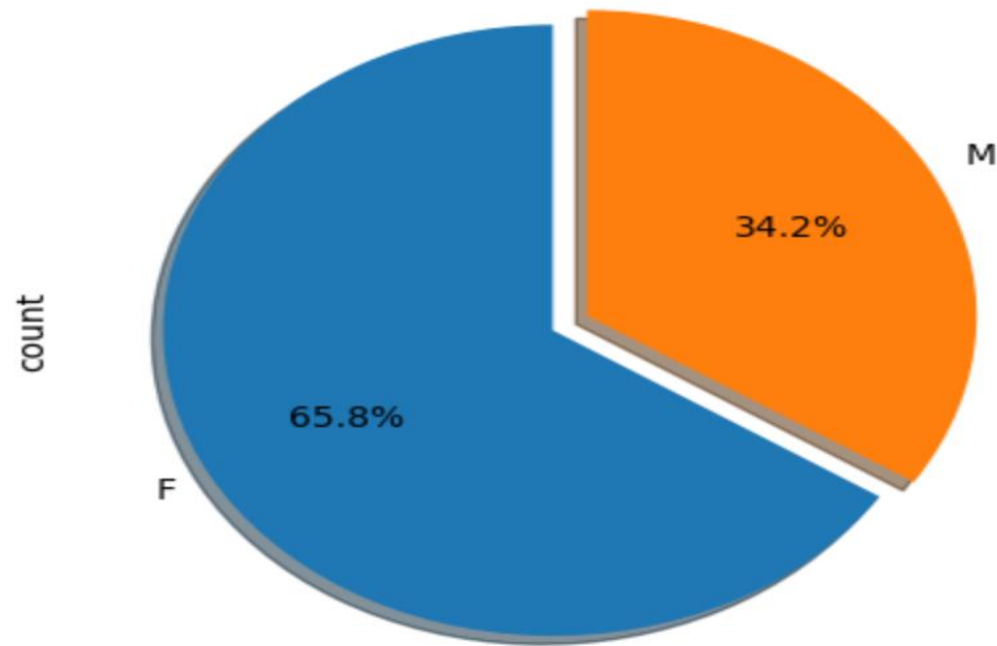


Hist plot of EXT_SOURCE_3

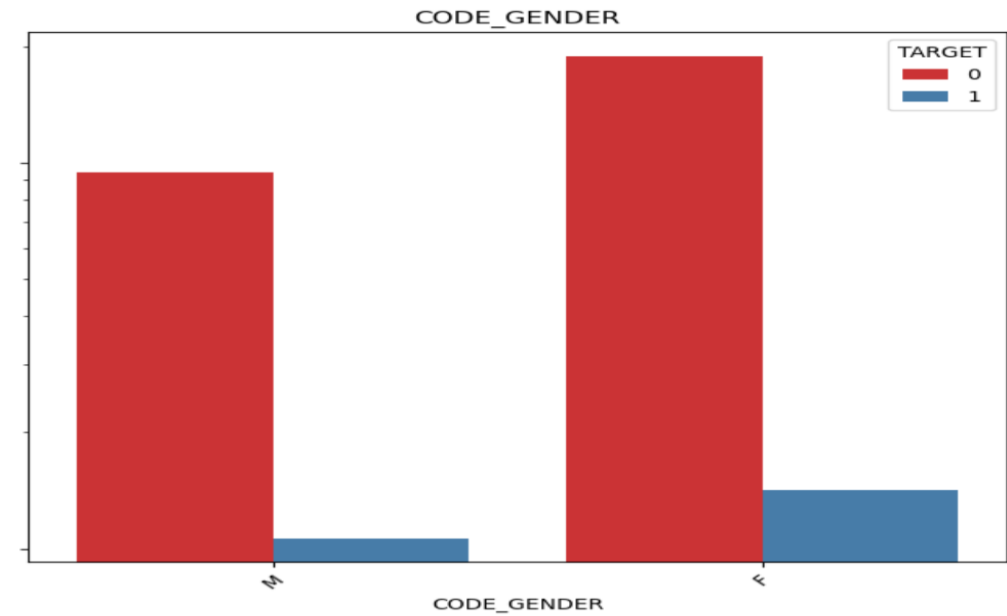


➤ We found null values in that column so we perform the hist plot and fill the nulls with median

Gender Analysis



WE PERFORM GENDER ANALYSIS IN PIE AND BAR CHART WITH RESPECT TO TARGET VARIABLE AND PLOT SHOWN BELOW.

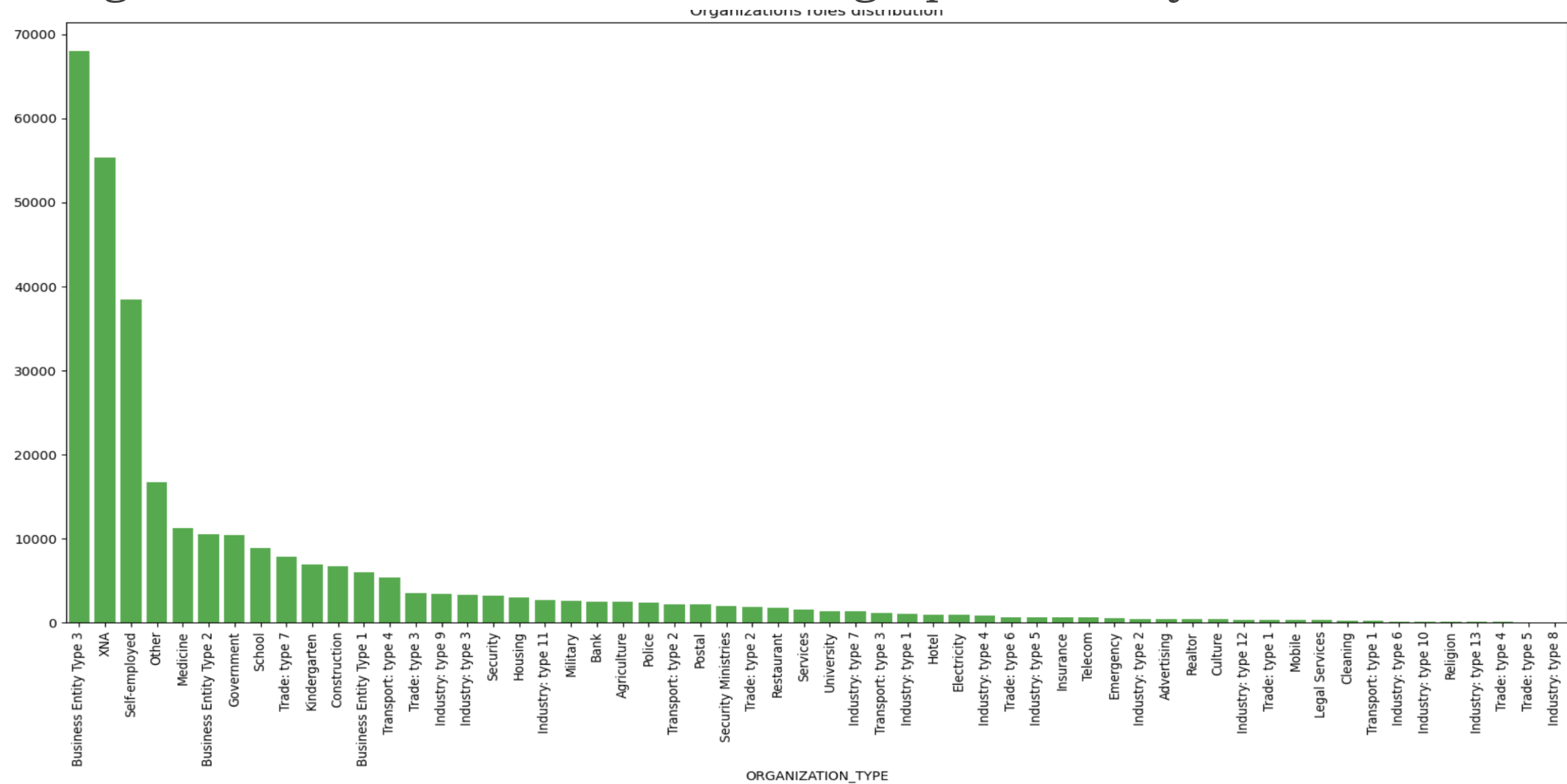


Data Imbalance

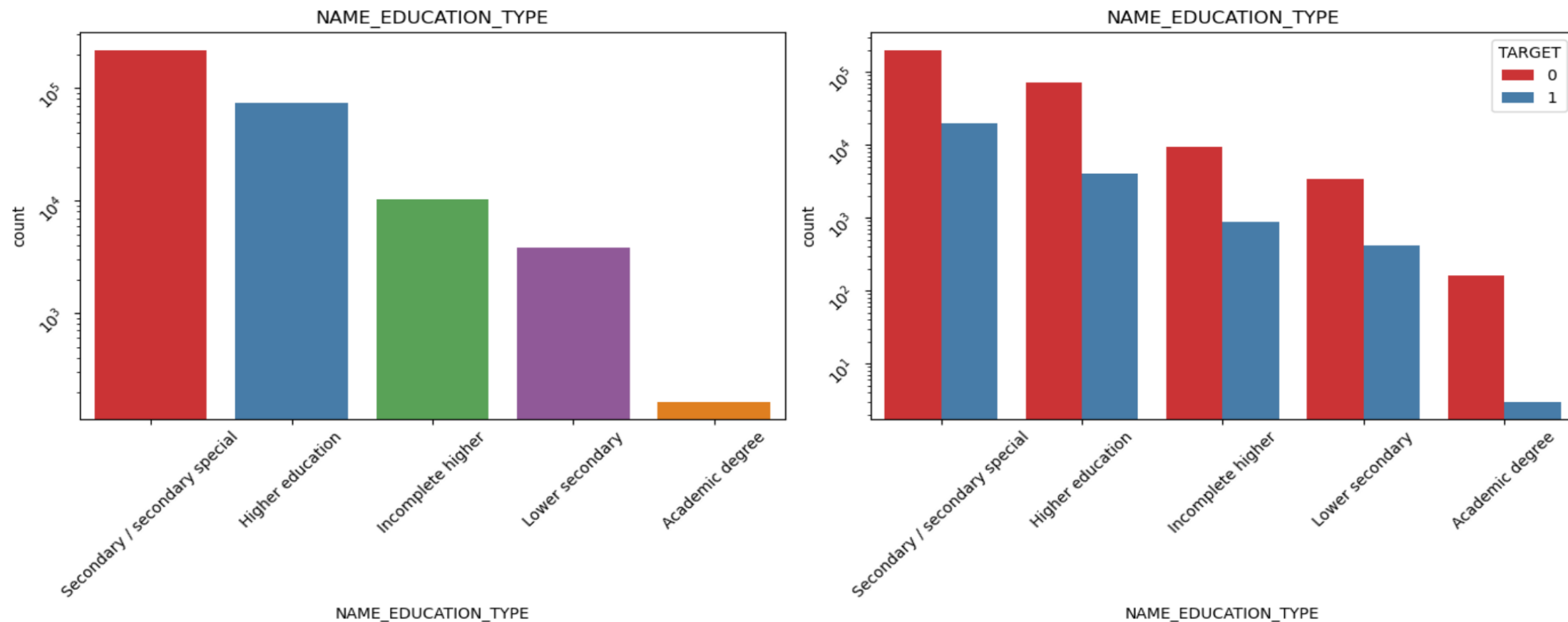
-
- We found the data imbalance in Target column and it clearly visible to us by applying the pie chart.
-



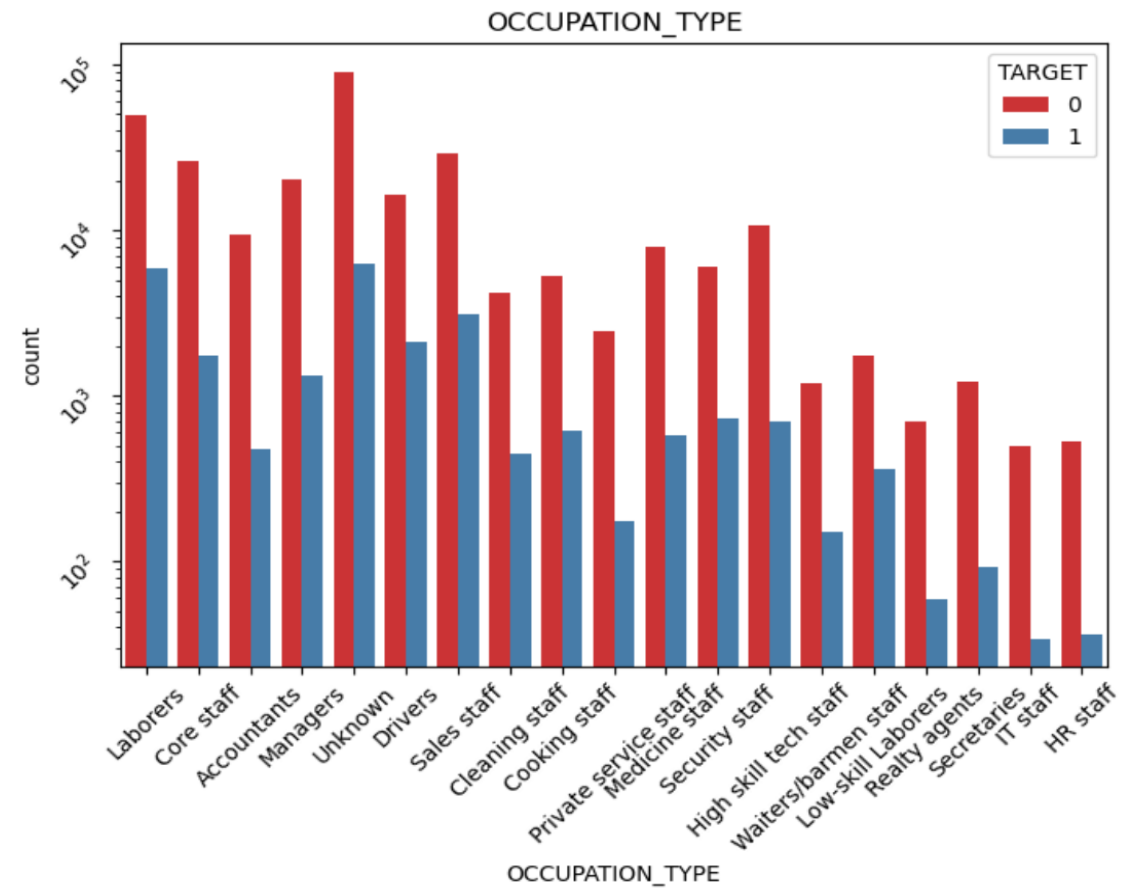
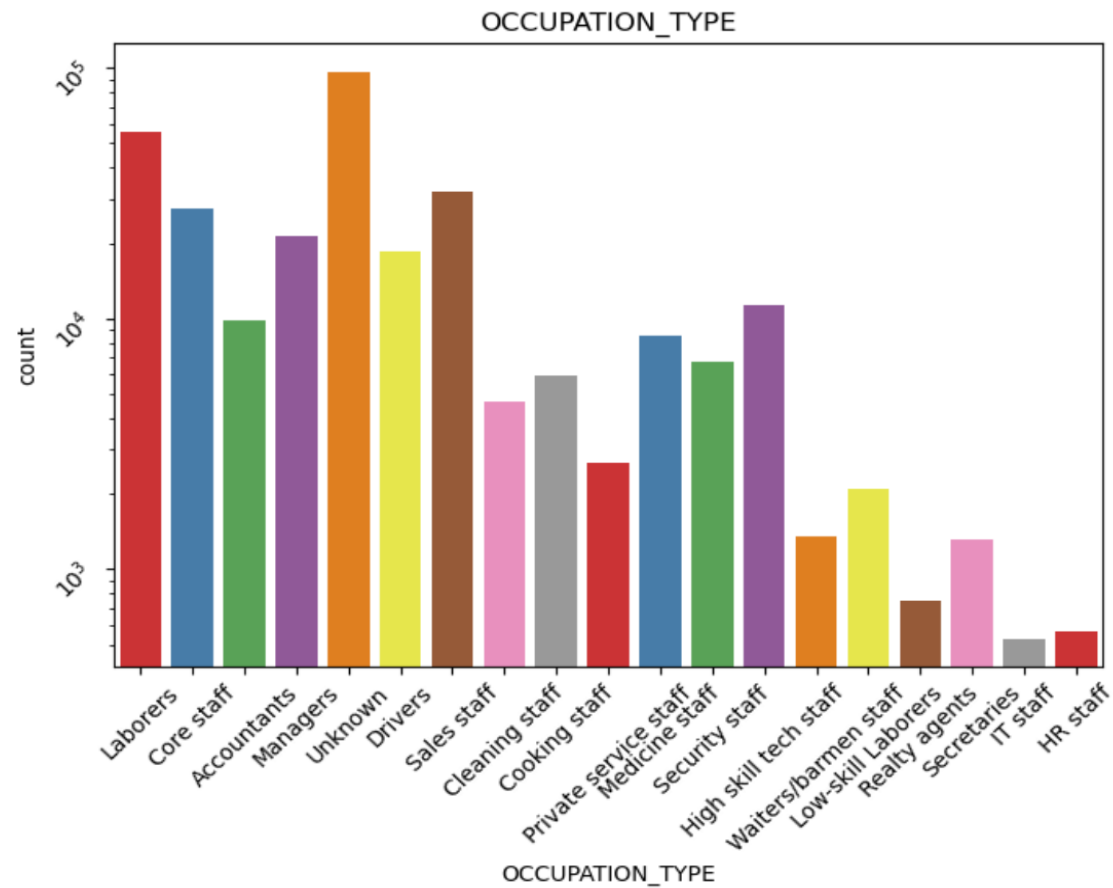
Organizations roles distribution in single plot to Analysis.



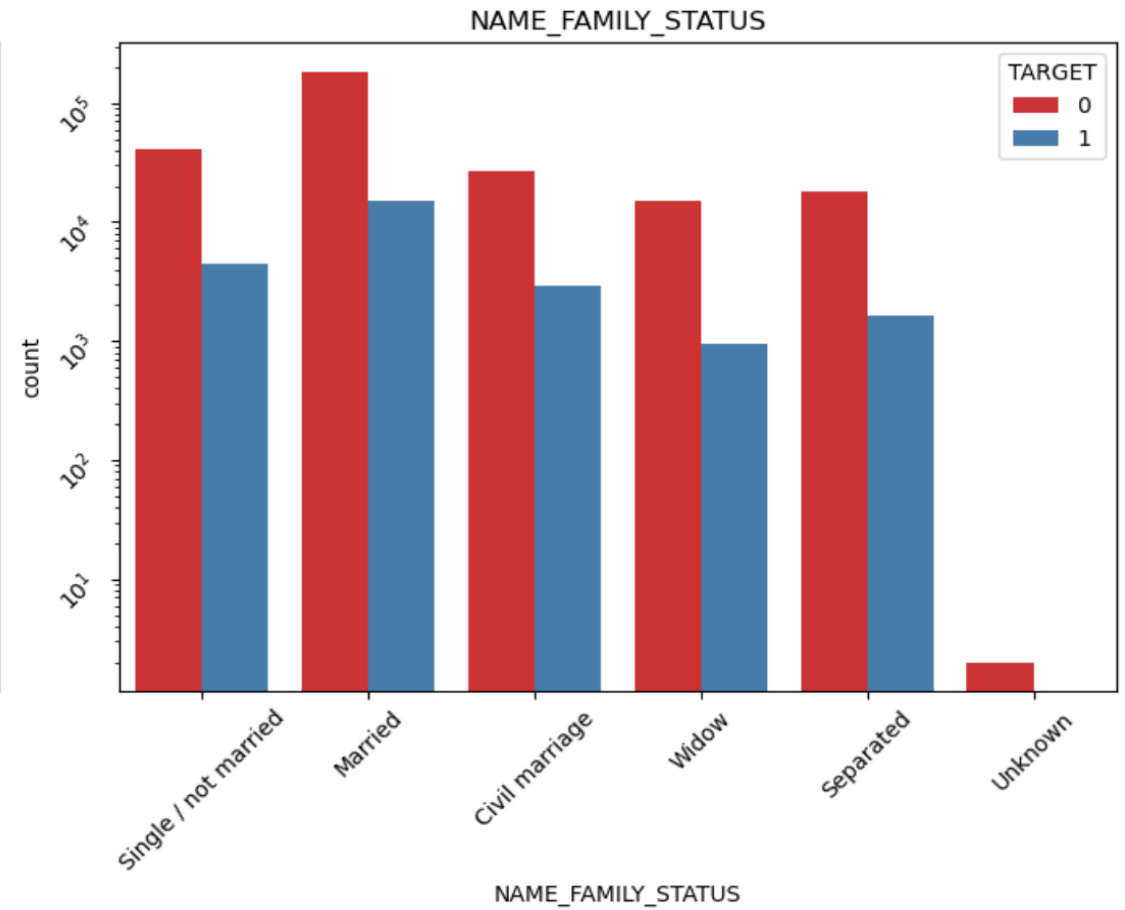
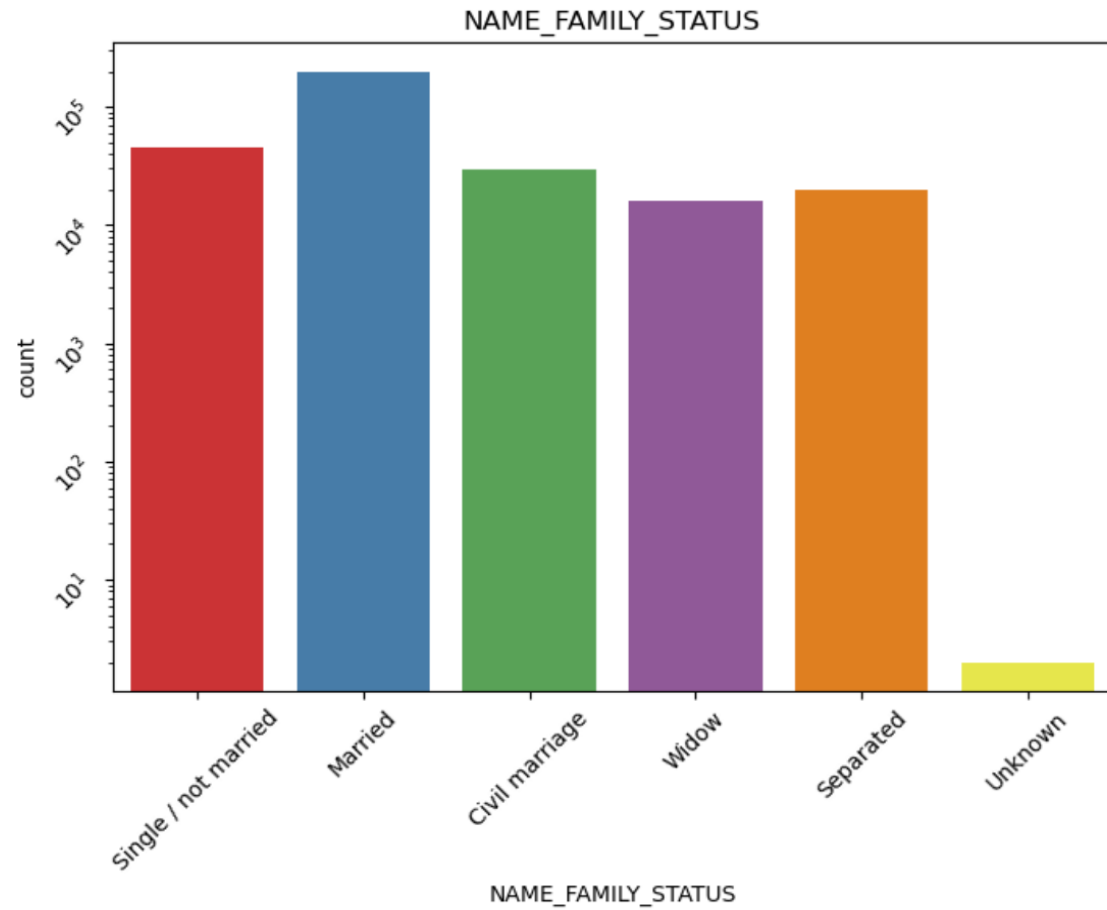
Secondary education as more number of client and defaulter too



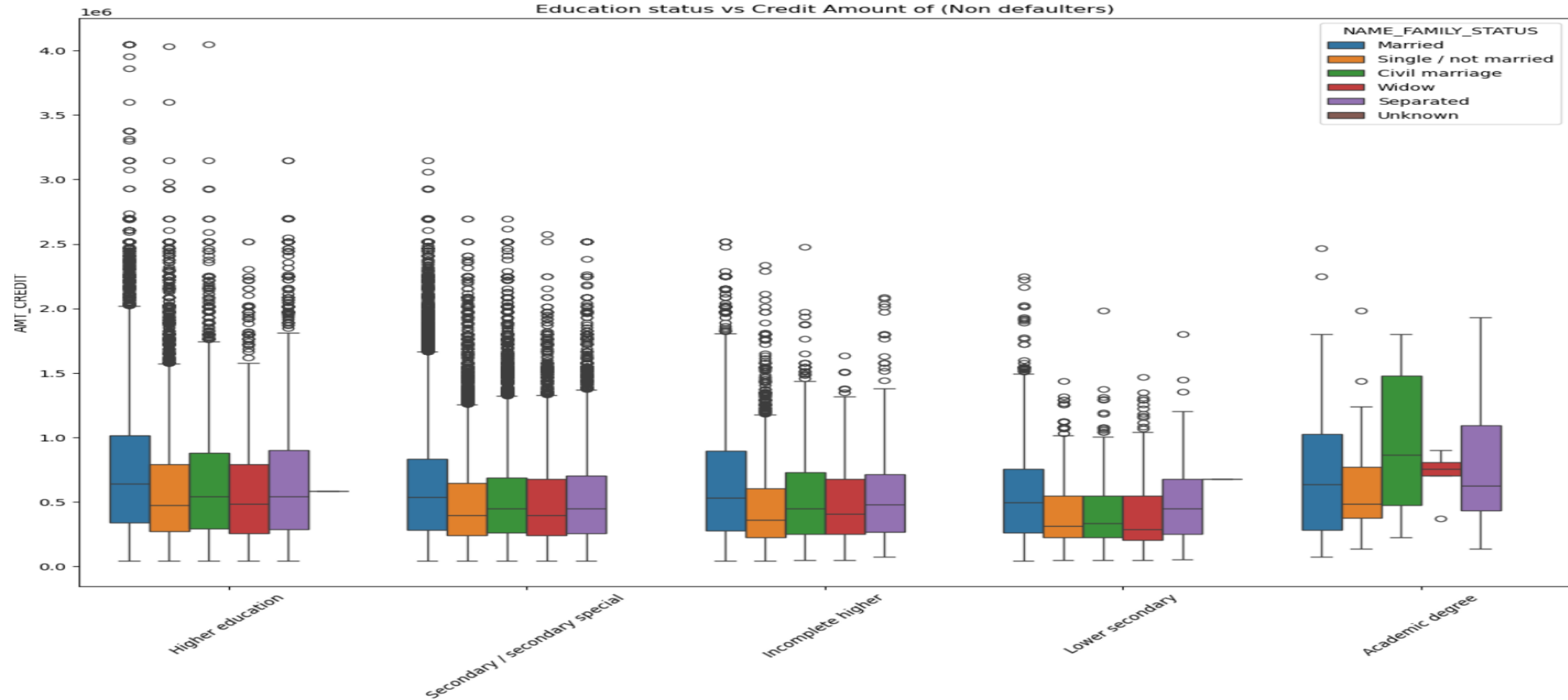
Even though labours has less percentage of accountants both have similar percent of defaulters.



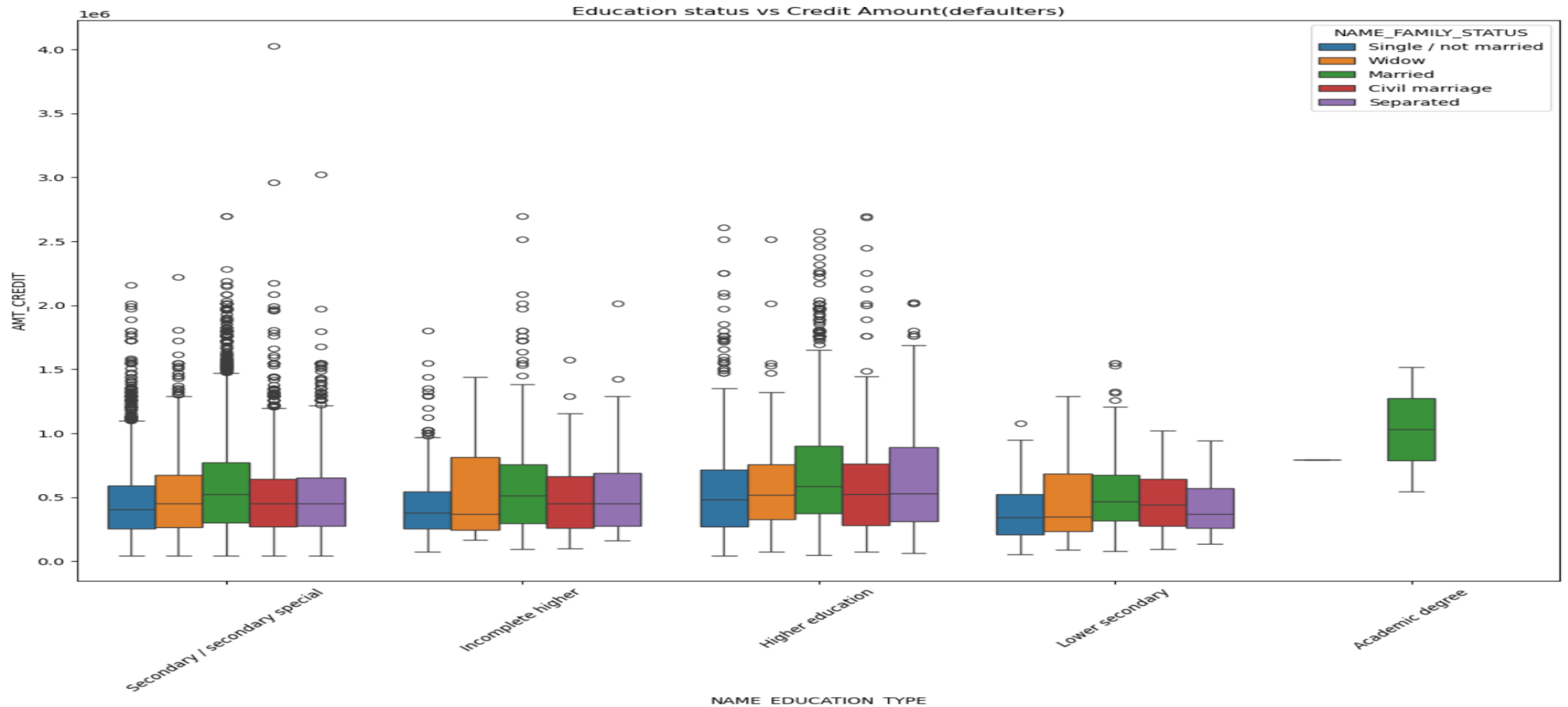
Married has more defaulters and widow have less defaulters



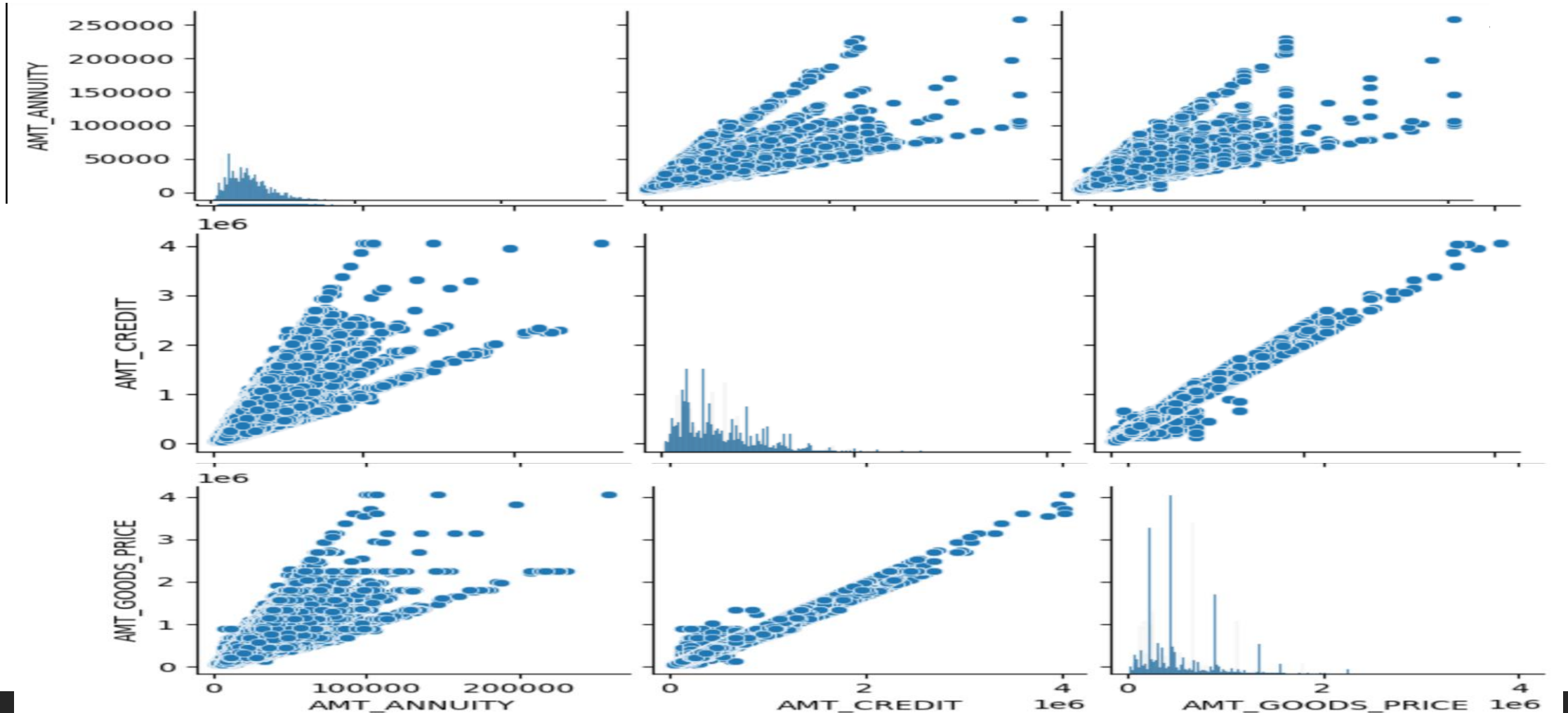
Perform multivariate analysis for both family type and education type wrt target 0(Re payers) and found less outliers in Academic degree



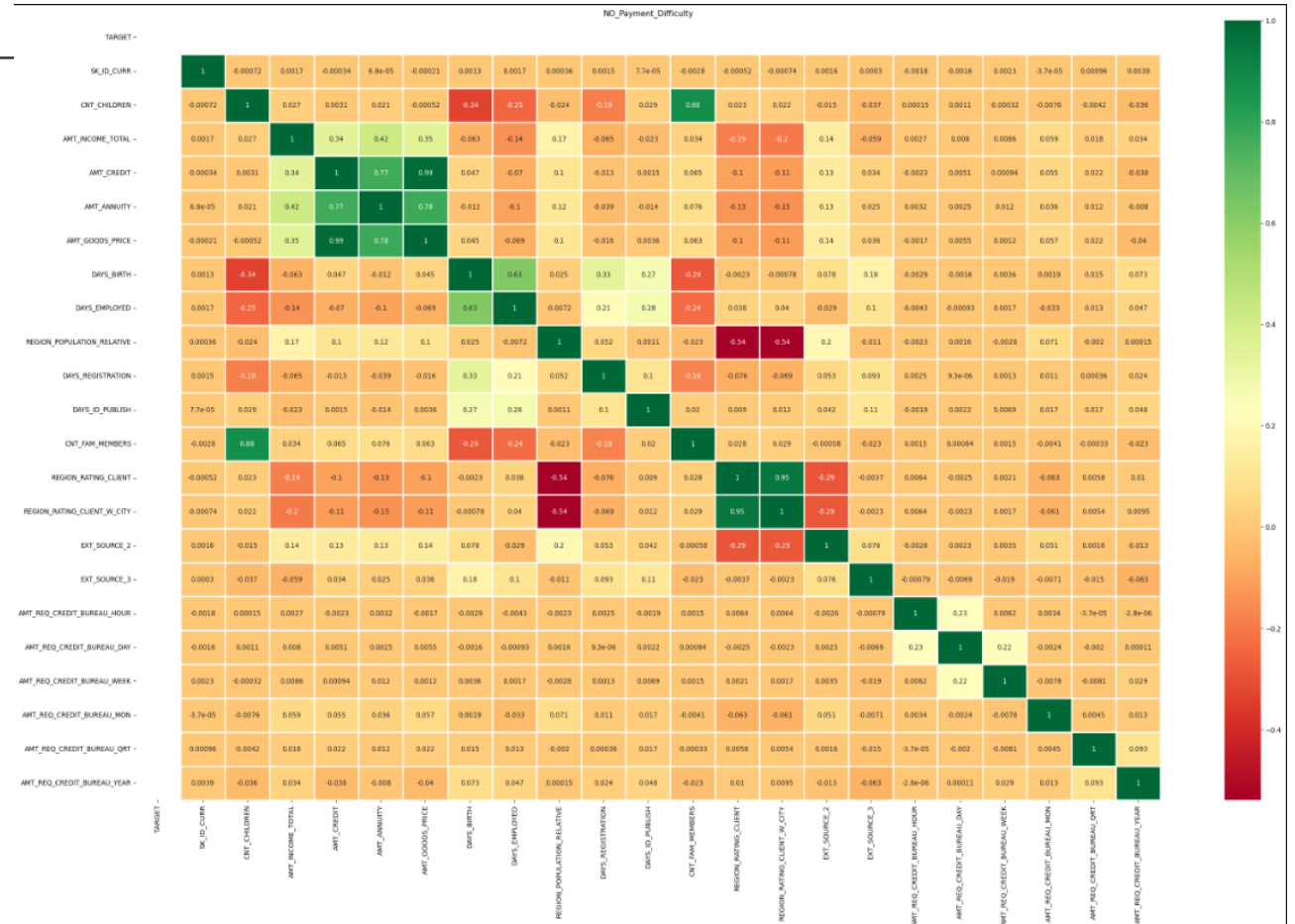
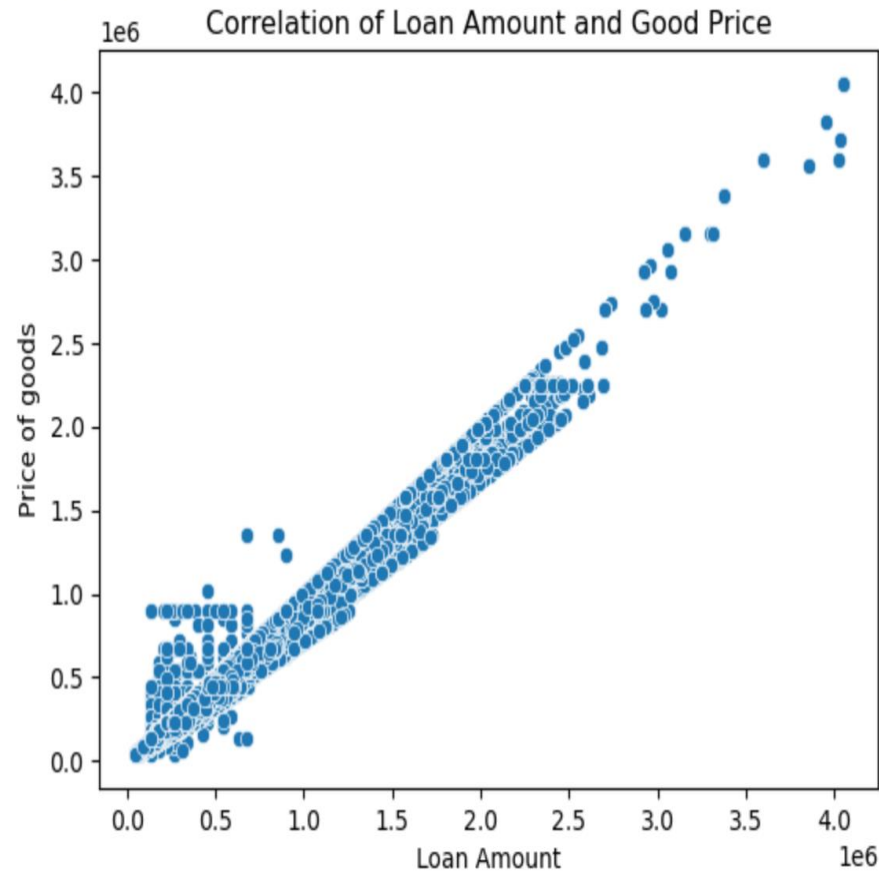
Perform multivariate analysis for both family type and education type wrt target 1(Defaulters) and we found more risk with secondary,incomplete higher and higher education.



We perform multivariate analysis and found strong positive correlation between the loan request ,loan credited and loan for good purchase.

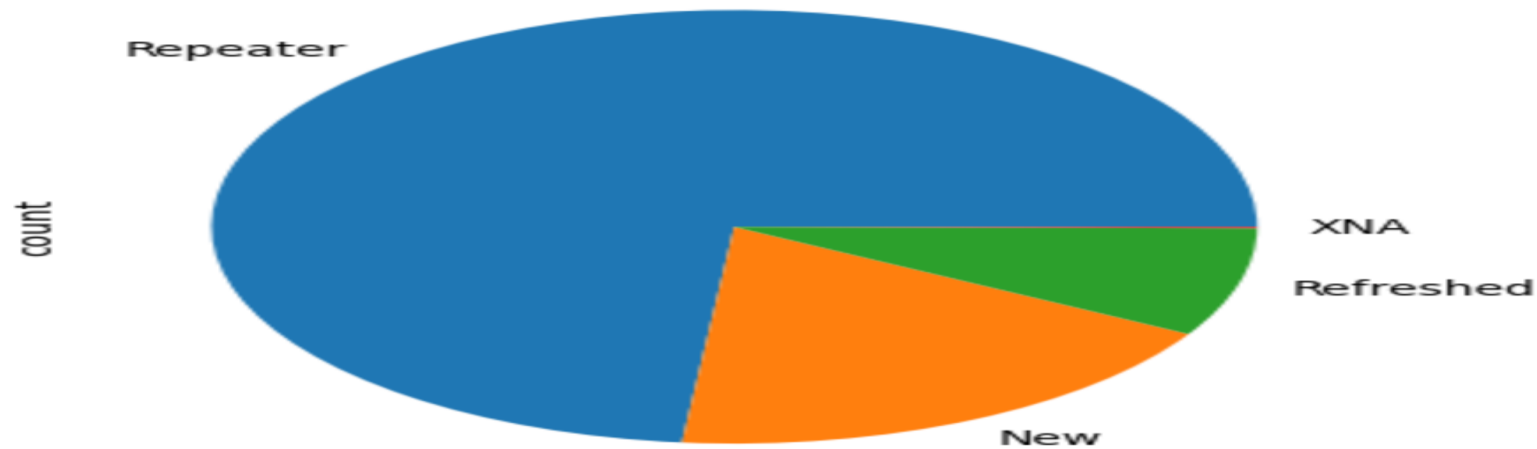


By correlation of target variable wrt other columns and we observing the amount credit and goods price show strong positive correlation

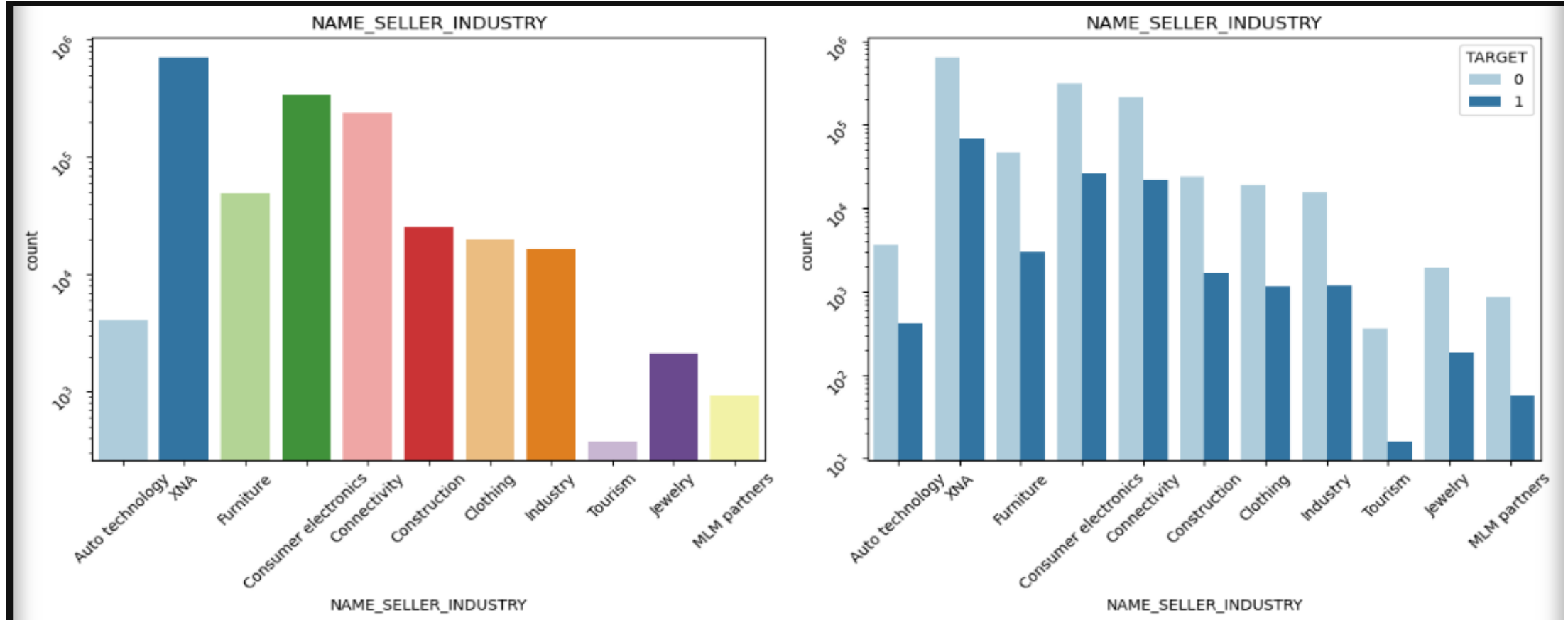


Previous data set

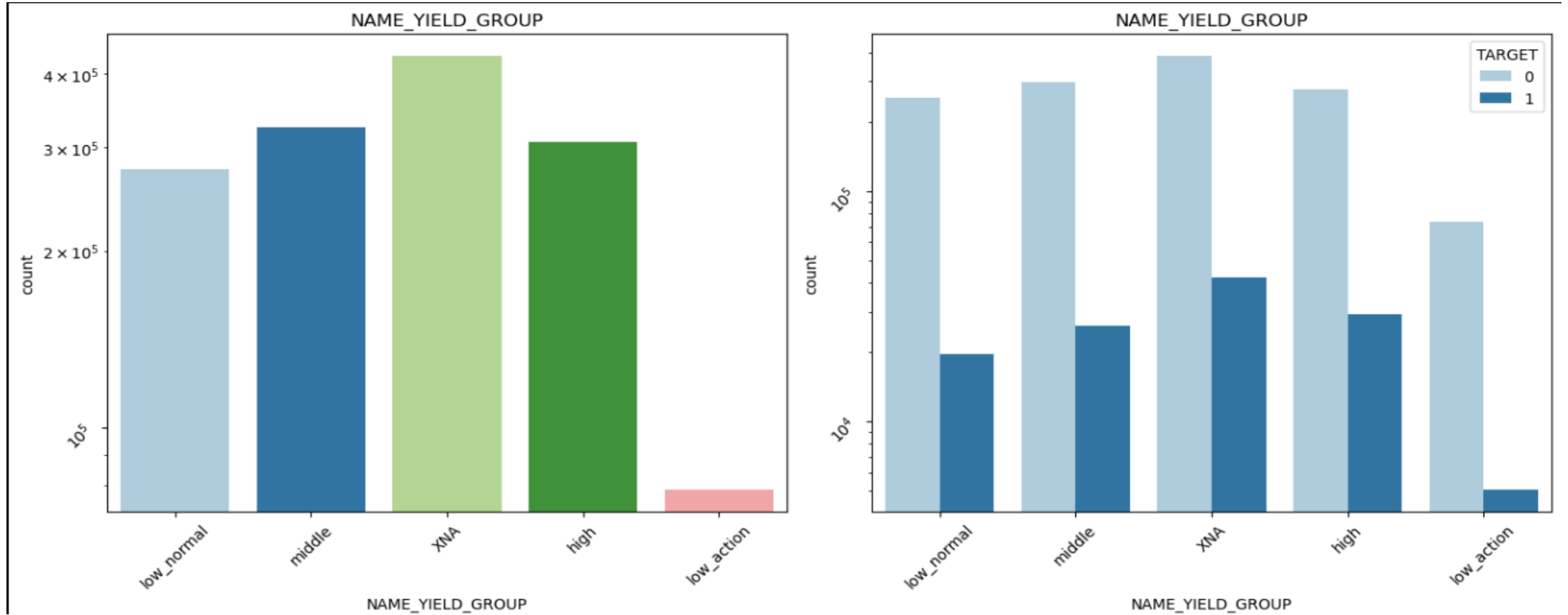
- After read and analyse the data we observed some null values and impute the values with suitable methods.
- Then find the client is new or repeater with pie plot.



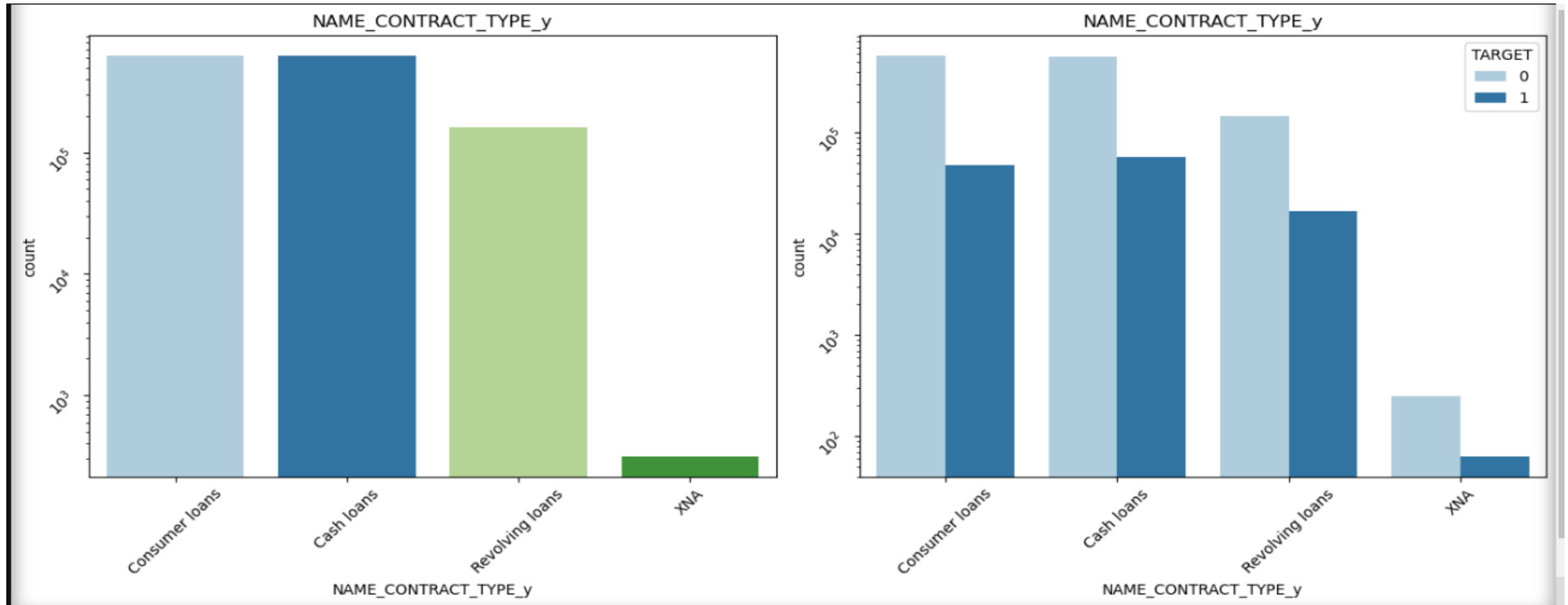
Consumer electronics and connectivity are more targetable industry



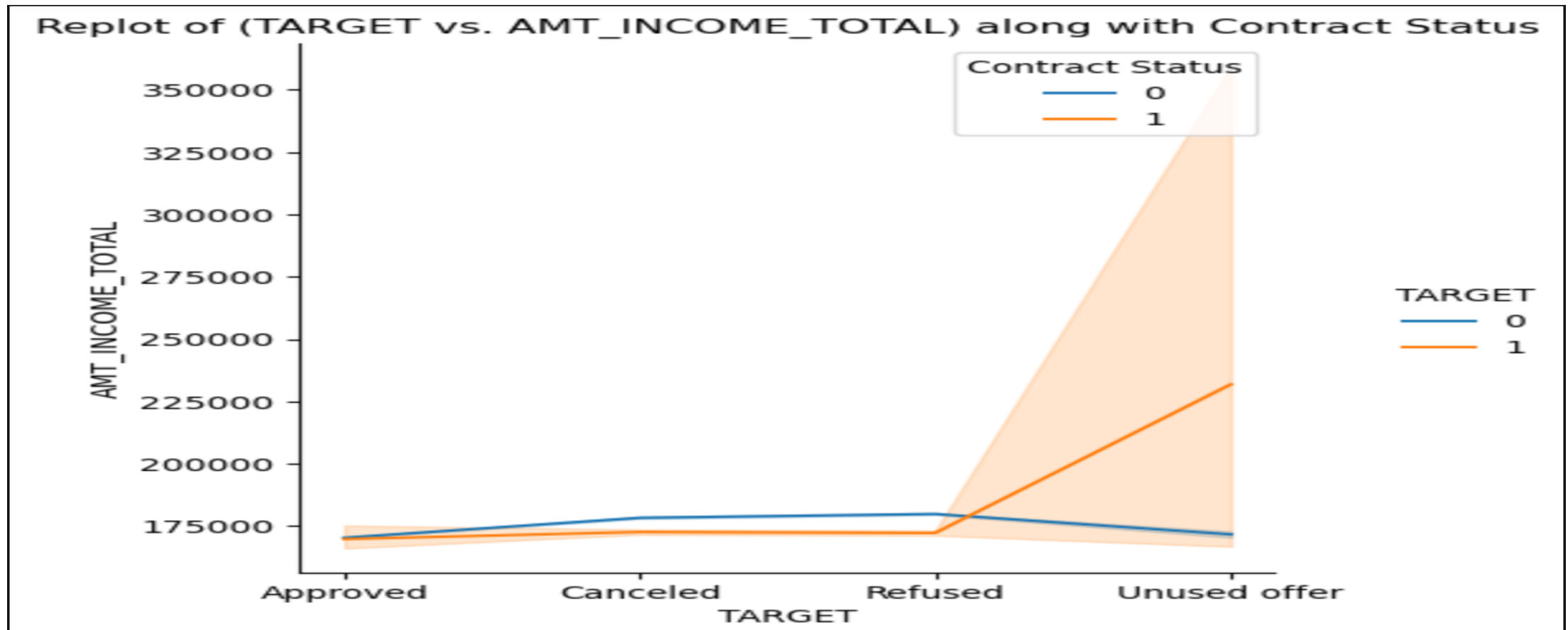
In the previous loan clients with canceled, refused and unused offer are more no of percent are re-payers.



In consumer and cash loans contain more defaulters parallelly double check before process the loan.



The below graph provide a insight of previous unused offer persons contain high defaulter even though they earn more amount of income.



Summary

We analysis each and every details of data sets and get some meaningful insights and also gave some plots.

- Females are more in number to apply loan and widow are less defaulters.
- In category of education Academic degree is preferred than higher and secondary education.
- We get good correlation for taking loan to buy the goods.
- We find more number of repeater in taking loan.
- Consumer electronics and connectivity are more targetable industry
- In previously canceled and refused loans have re-payers then verify twice before cancel the loans.
- While coming to unused offer have more defaulters even though they earn more.



The End