

Presentation On Credit EDA Case study

Mr Nishchay Yadav Mr Akshay Siraswar



Purpose

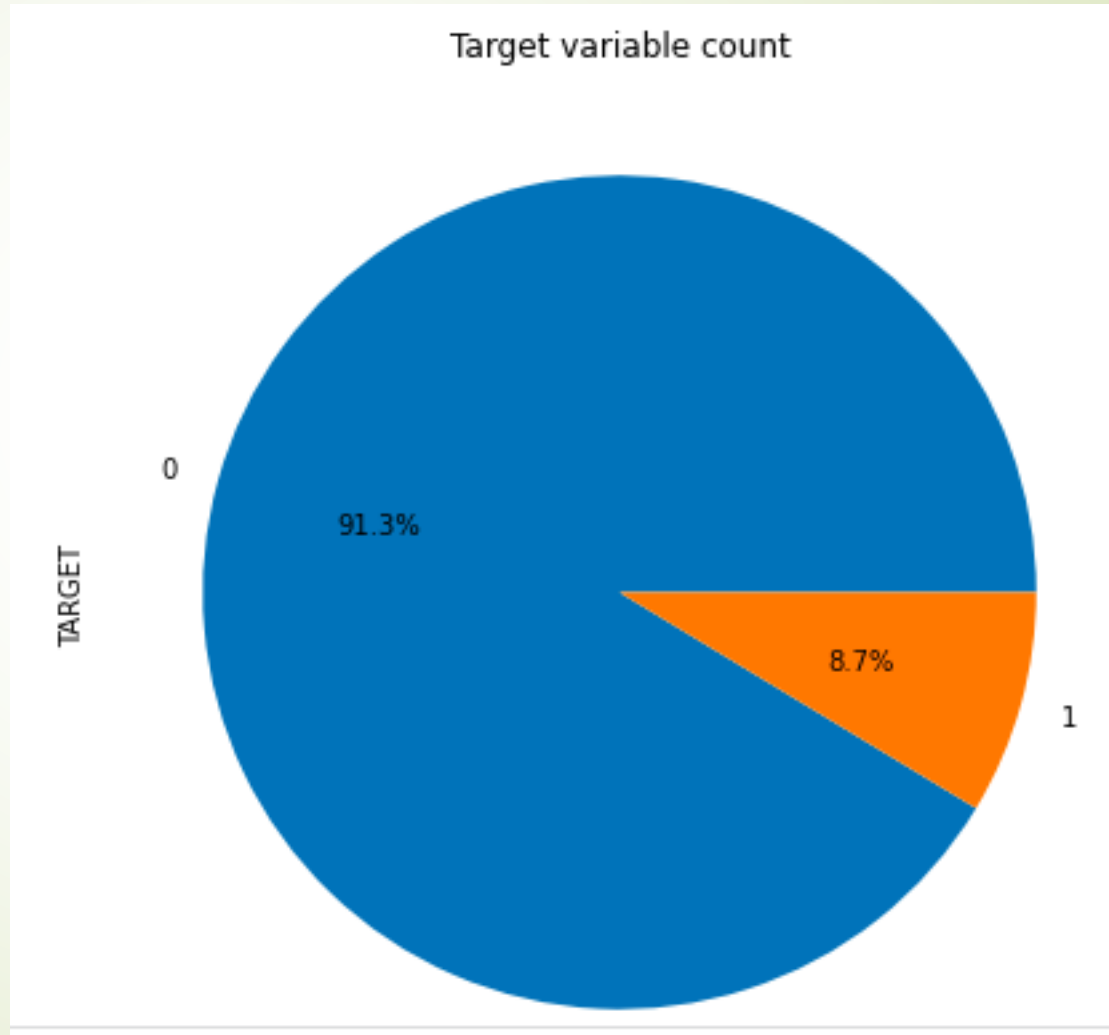
- Credit risk analysis will help the company to make a decision for loan approval based on the applicant's profile. Which controls loss of business to the company and avoid financial loss for the company.

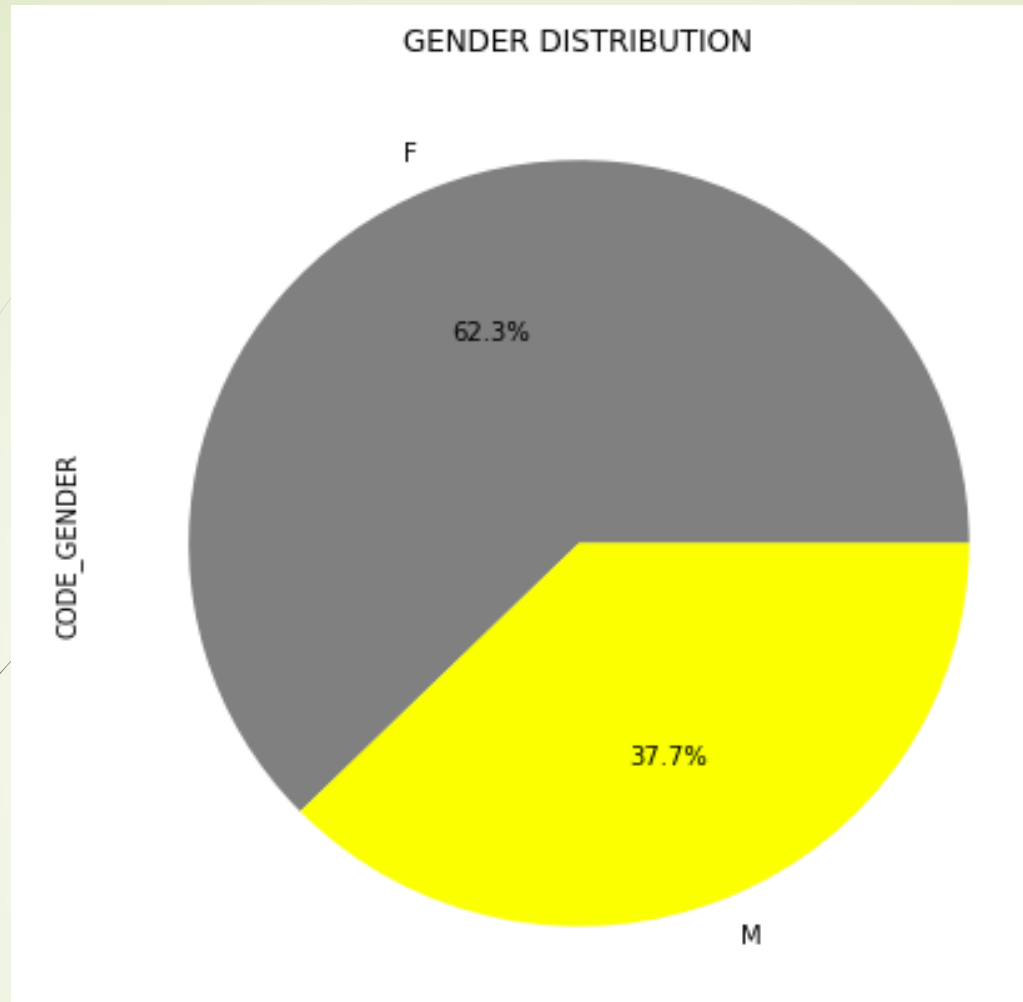
Steps

- 1.Dataunderstanding and sourcing
- 2.check for Data quality issues and Binning
- 3.check for Data imbalance and univariate,segmented univariate & Bivariate analysis,correlation
- 4.Merging of application data with previous application data
- 5.Data analysis by univariate, segmented univariate , Bivariate analysis and correlation
- 6.Recommendations and Risks

UNIVARIATE ANALYSIS

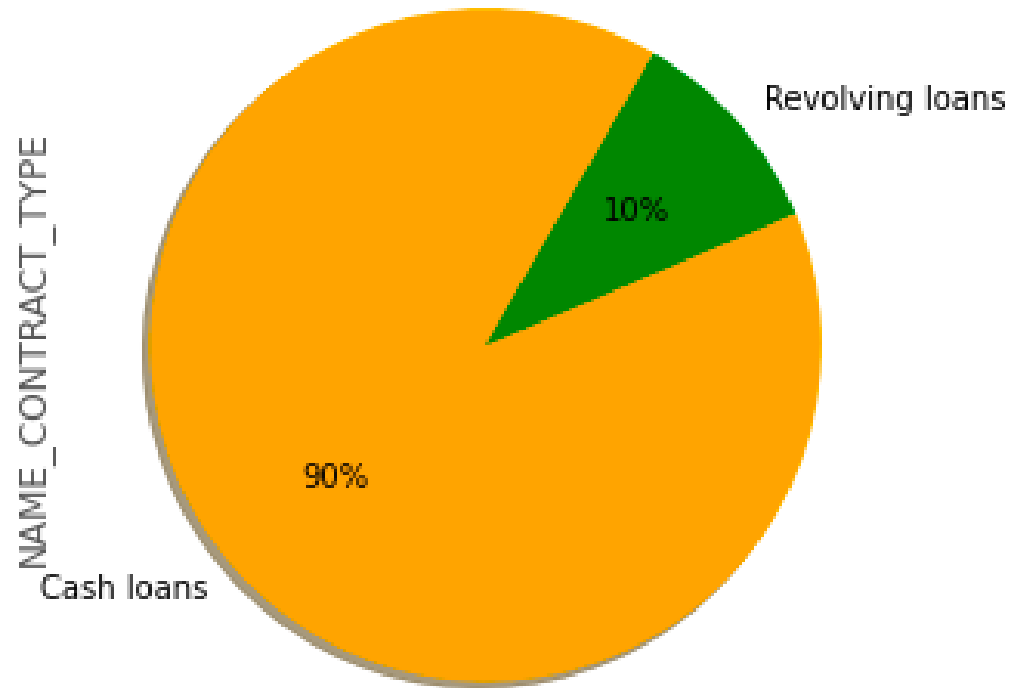
From the pie chart we can see that 8.7% people are defaulters while 91.3% people are non-defaulters.



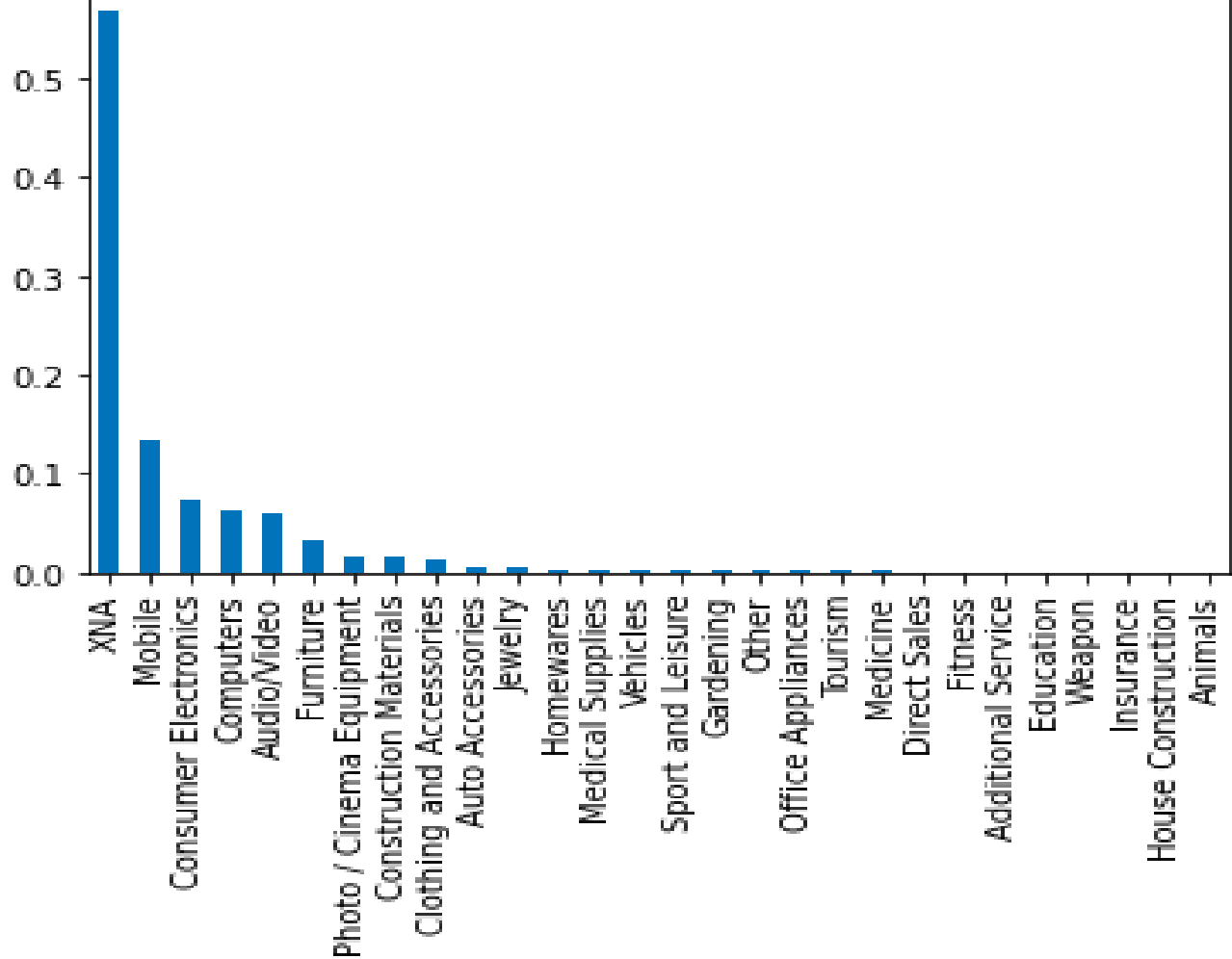


Females are more interested in taking loans as compared to men.

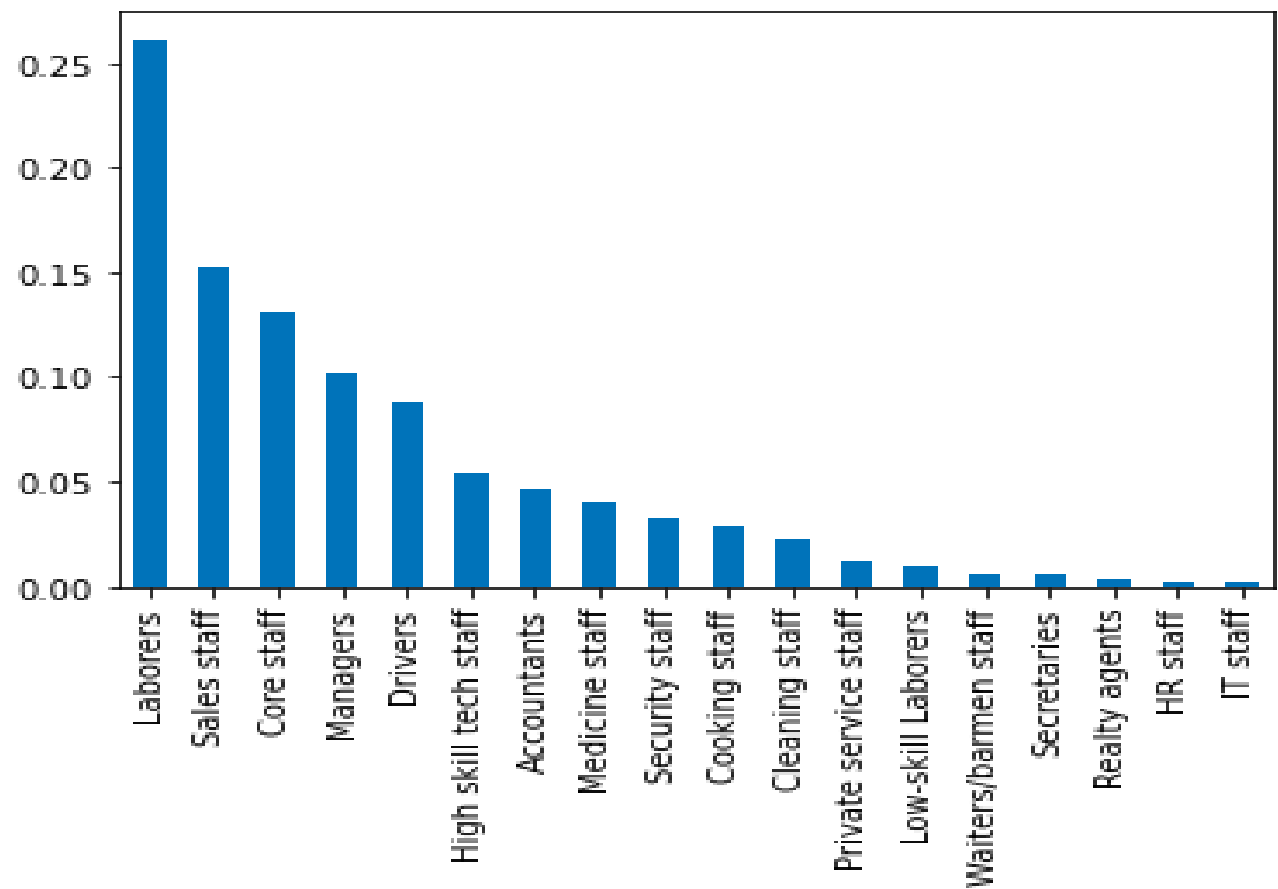
Distribution of contract types in current_application



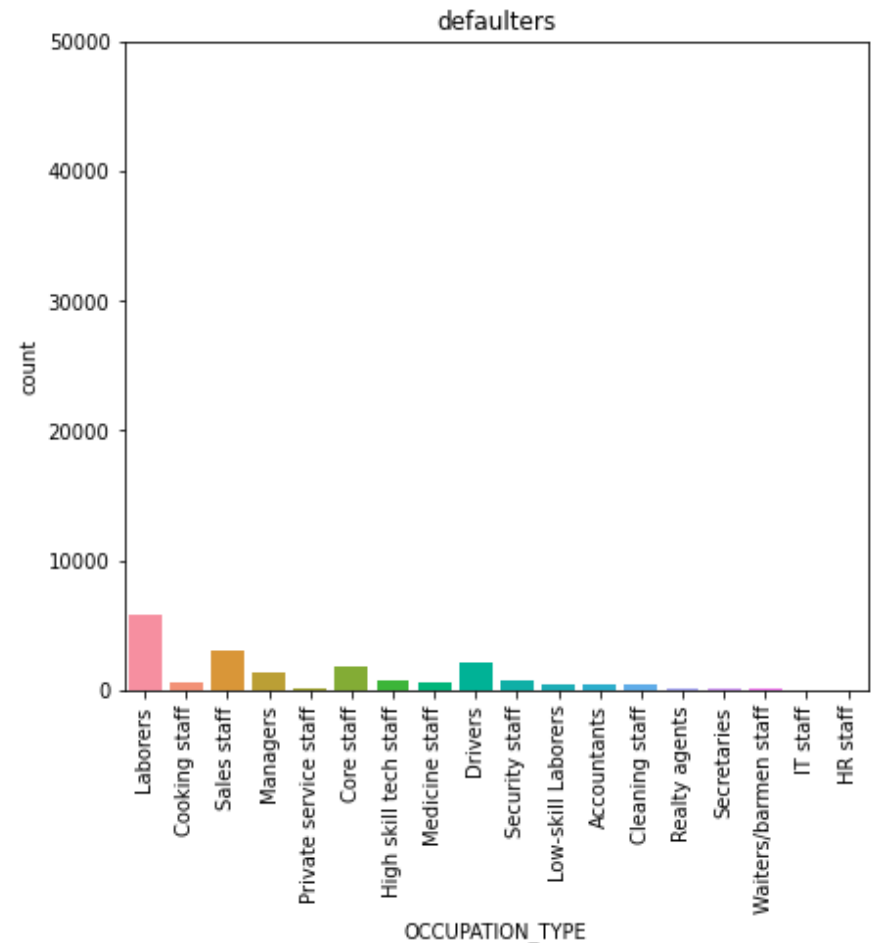
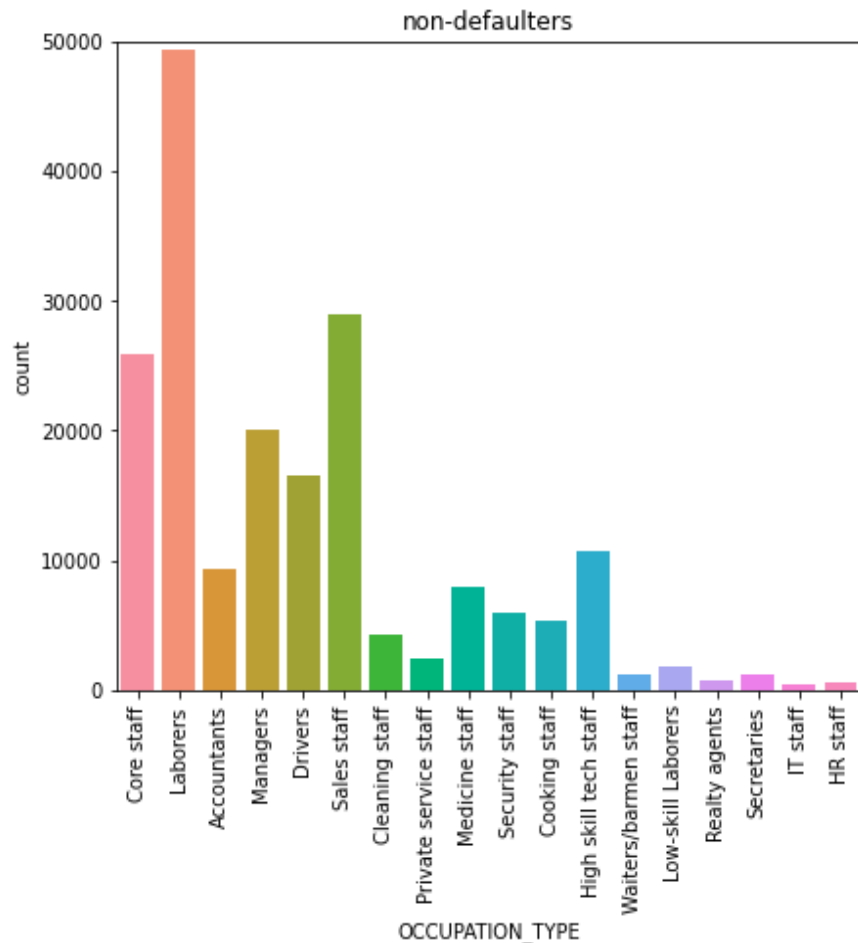
90% loans belong to cash category



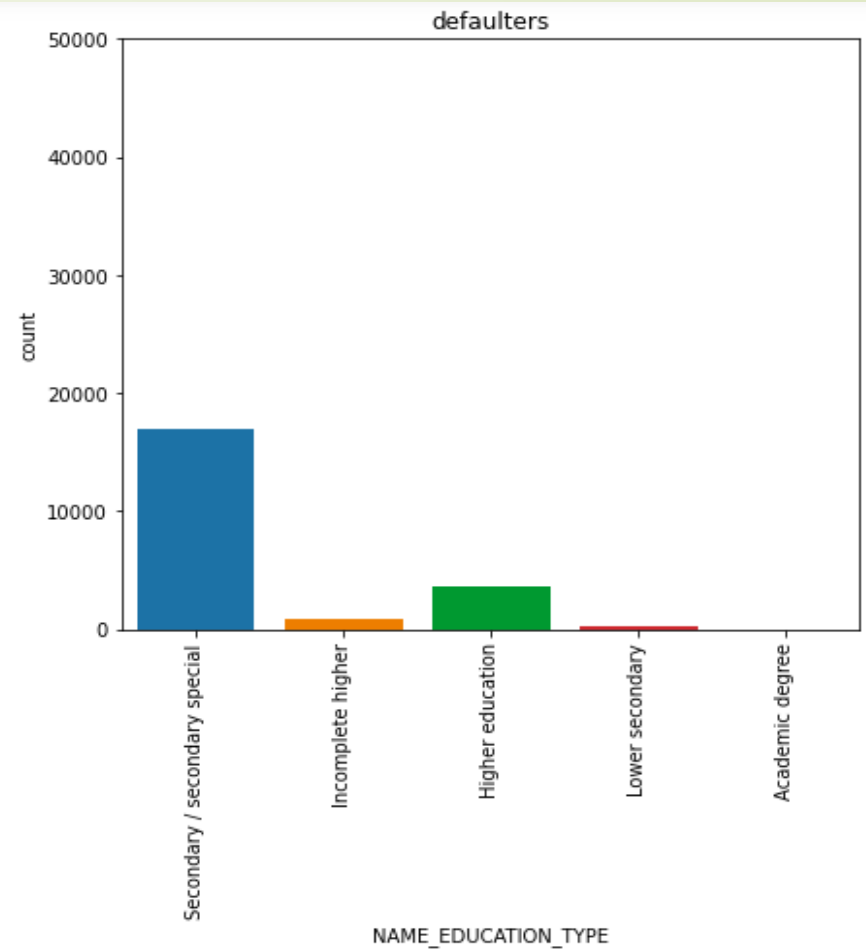
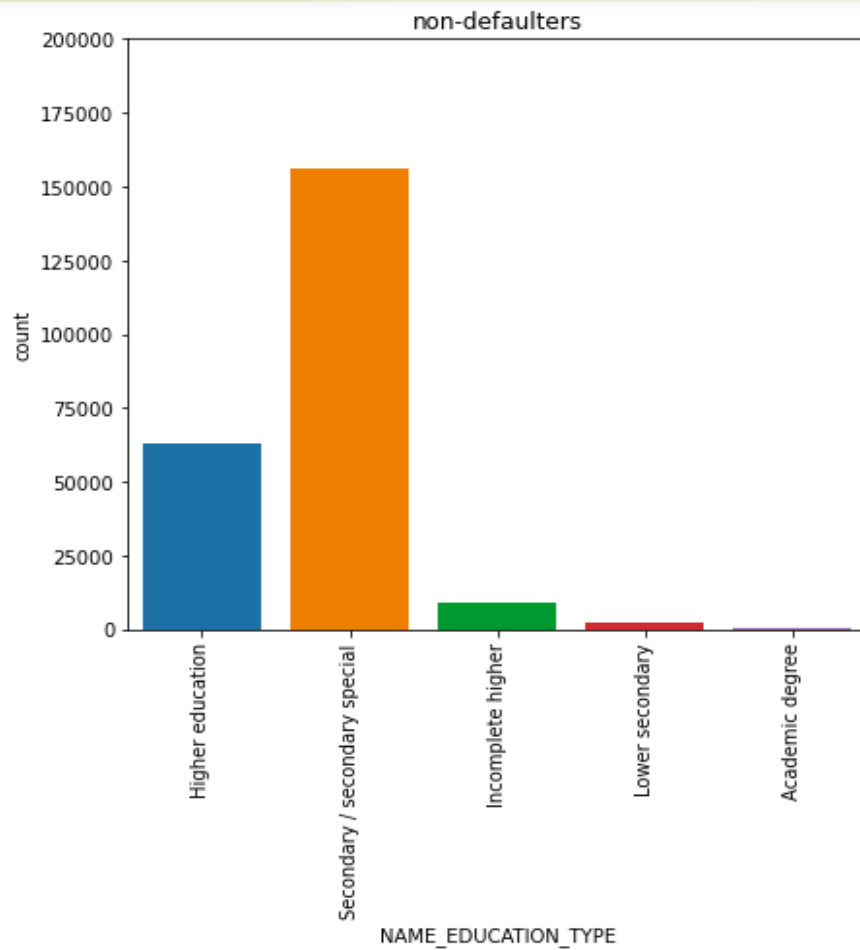
Mobiles ,Computers and consumer electronics are popular goods categories for which people prefer to take loans



Occupation types like Laborers, Sales staff
Core staff are more dependent on loans.

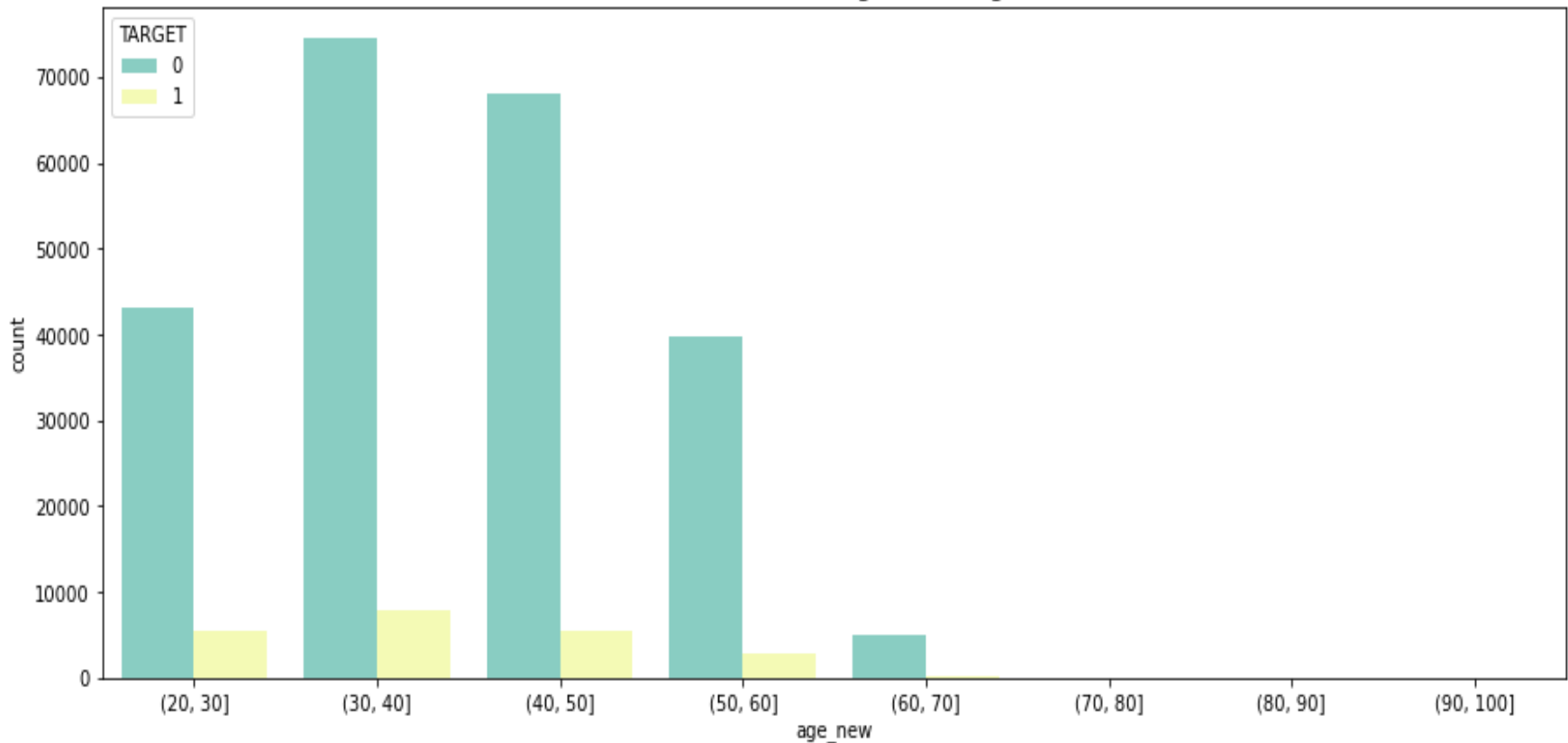


Laborers are most likely to repay back their loan as well as forms the most defaulters category



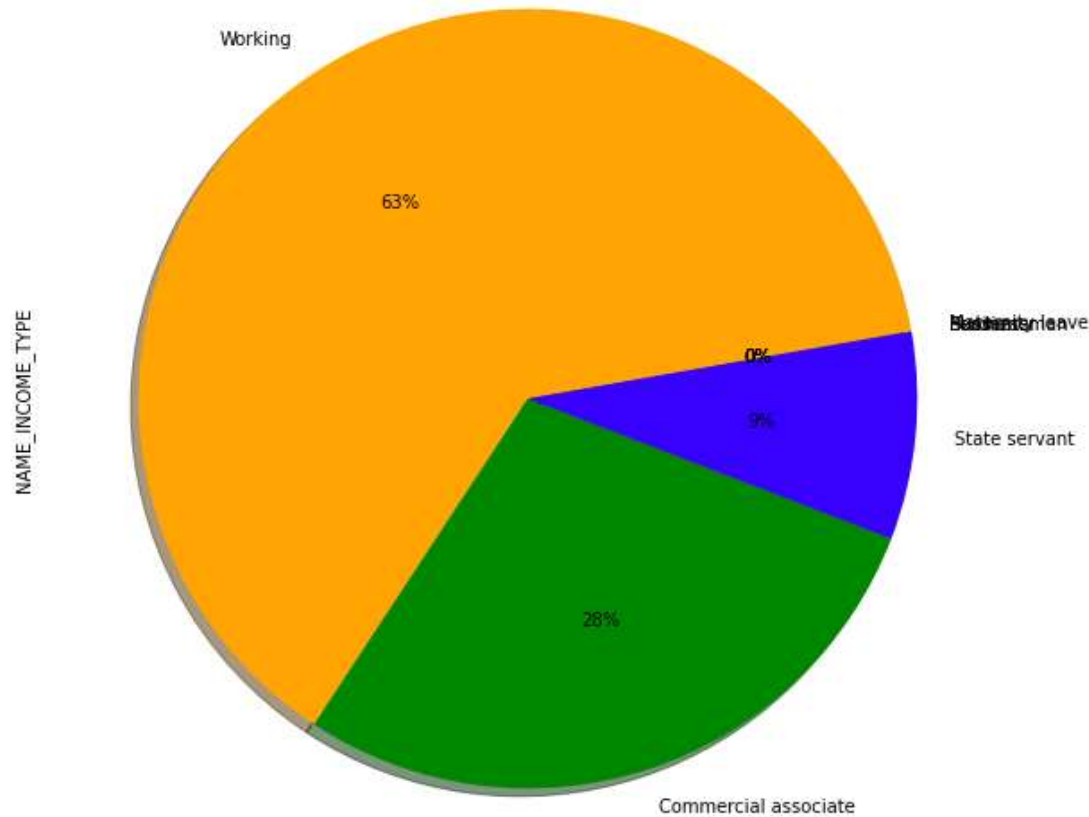
Highest loan defaulters category belong to client with secondary education as the highest level of education followed by higher education ones. Lower secondary and academic degree holders are less likely to take loan.

Distribution of client's age with target status

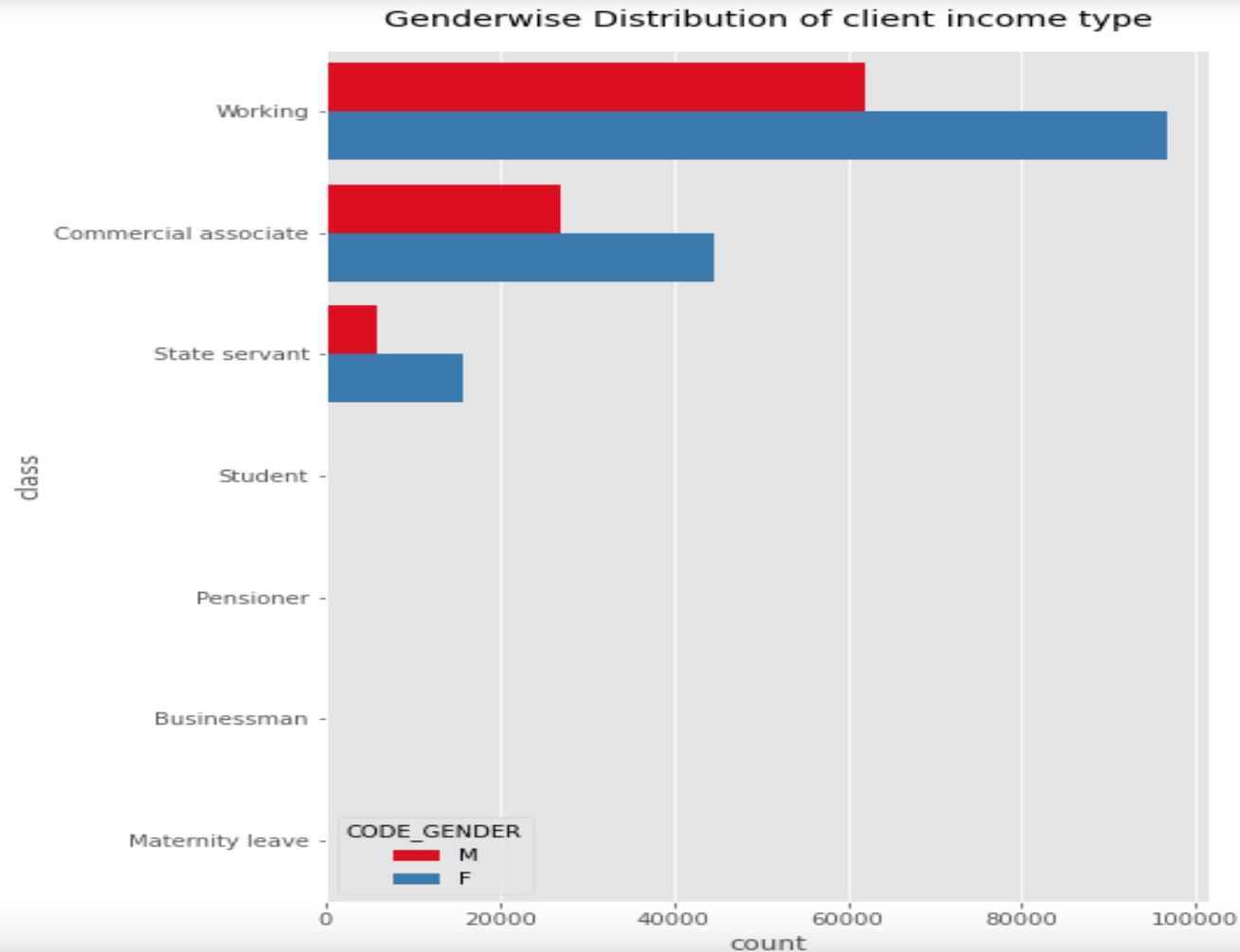


1. maximum defaulters pertains to people belonging to age group 30–40.
2. people above 50 years age are less likely to become defaulters.

distribution of income types



Most loan takers belong to workers class followed by commercial associates and state servants

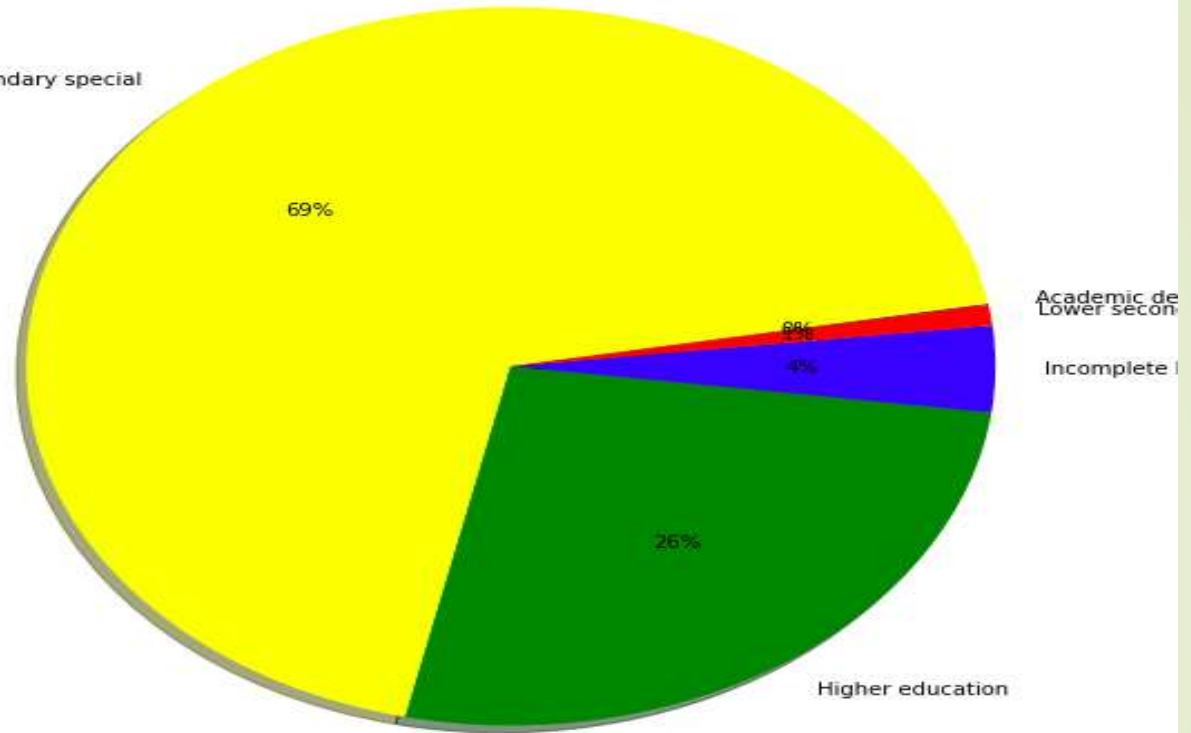


Among various classes female have taken loans more than male.

distribution of education types

Secondary / secondary special

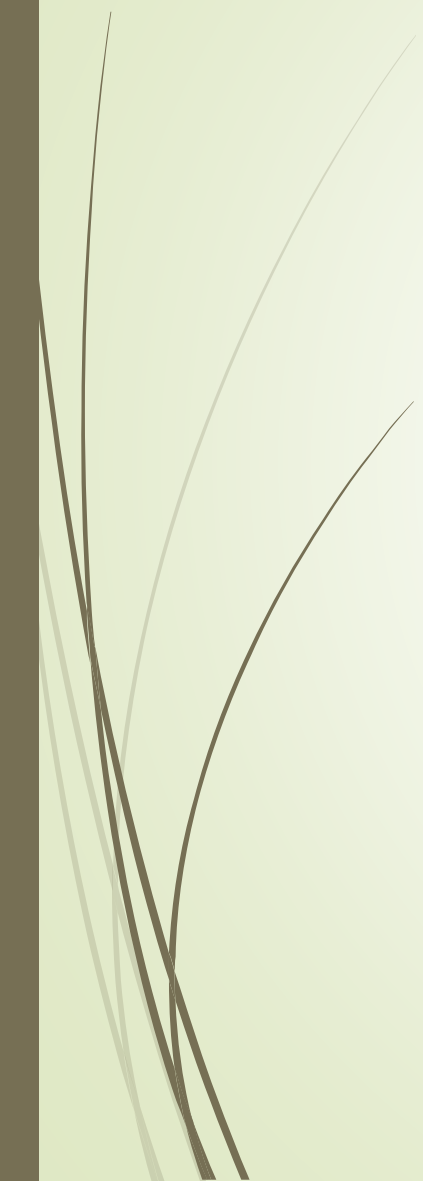

NAME_EDUCATION_TYPE



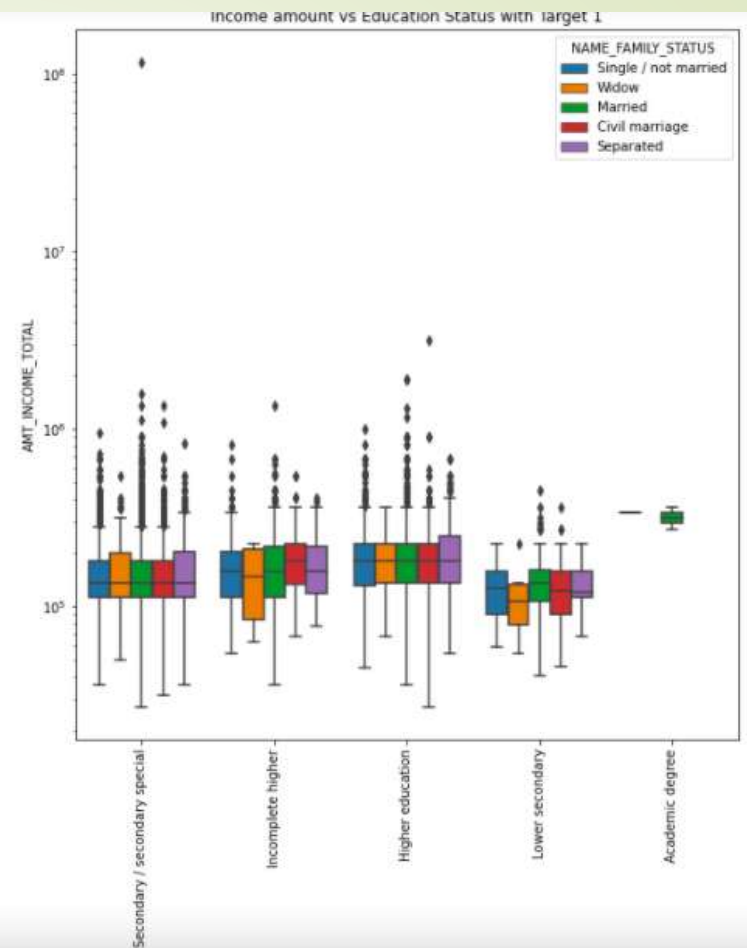
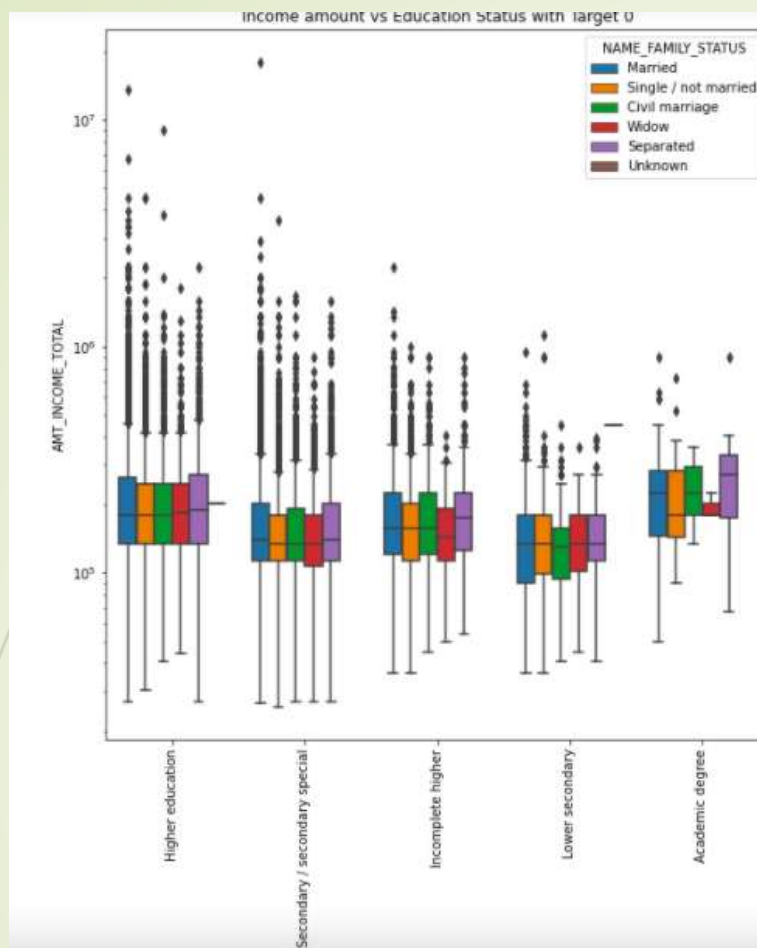
It can be seen that client with higher education level and secondary education level constitutes maximum percentage category for loan seekers.

BIVARIATE AND MULTIVARIATE ANALYSIS

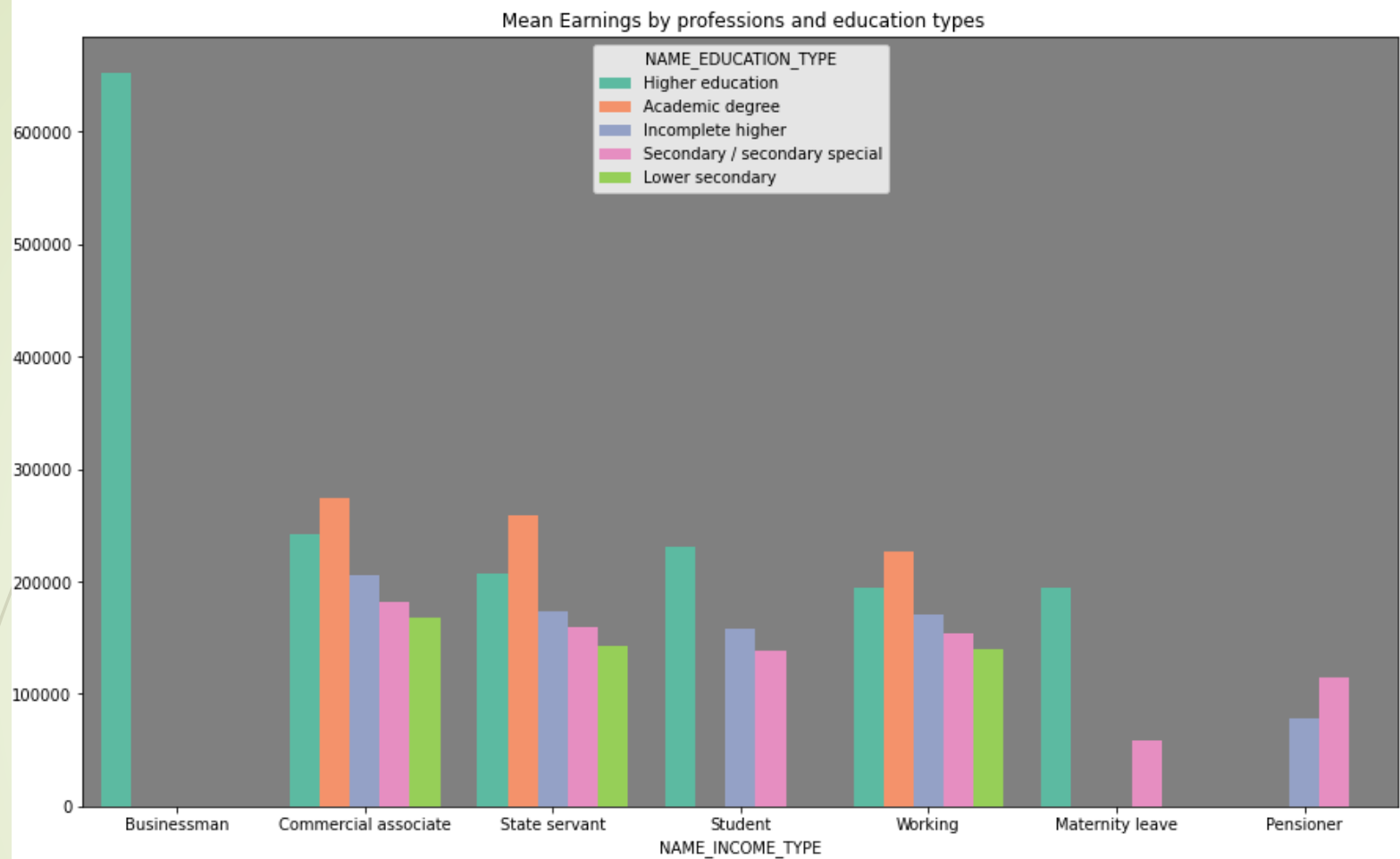




We can see that the density in the lower left corner is similar in both the case, so the people are equally likely to default if the family is small and the AMT_CREDIT is low. We can observe that larger families and people with larger AMT_CREDIT default less often .

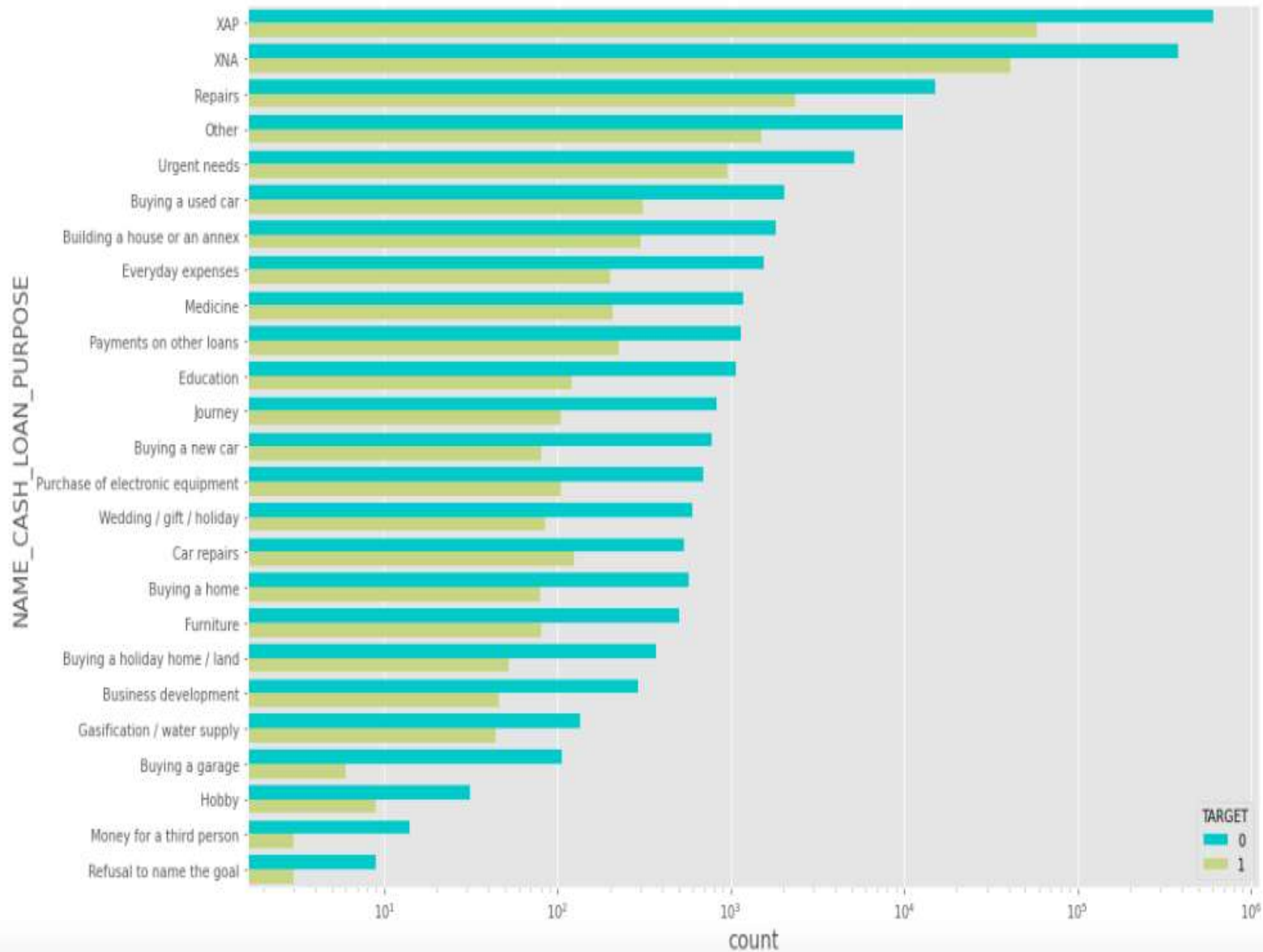



People with higher education with less income are defaulters while with high income are non defaulters. Higher and secondary Education have many outliers



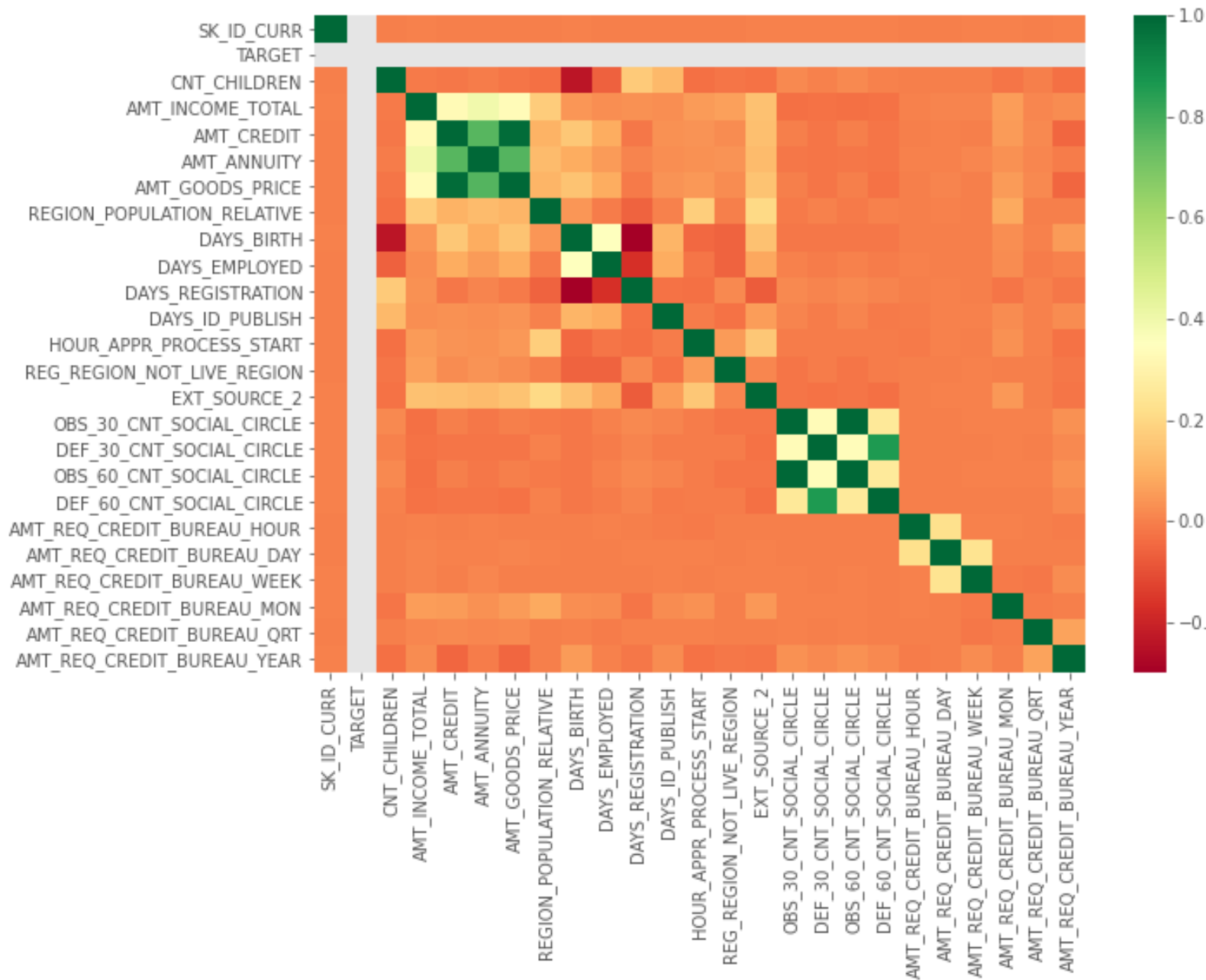
Businessman only with higher education level apply for loan. Pensioners least apply for loan. Students with higher education level mostly seek for loan as compared to students with other education levels

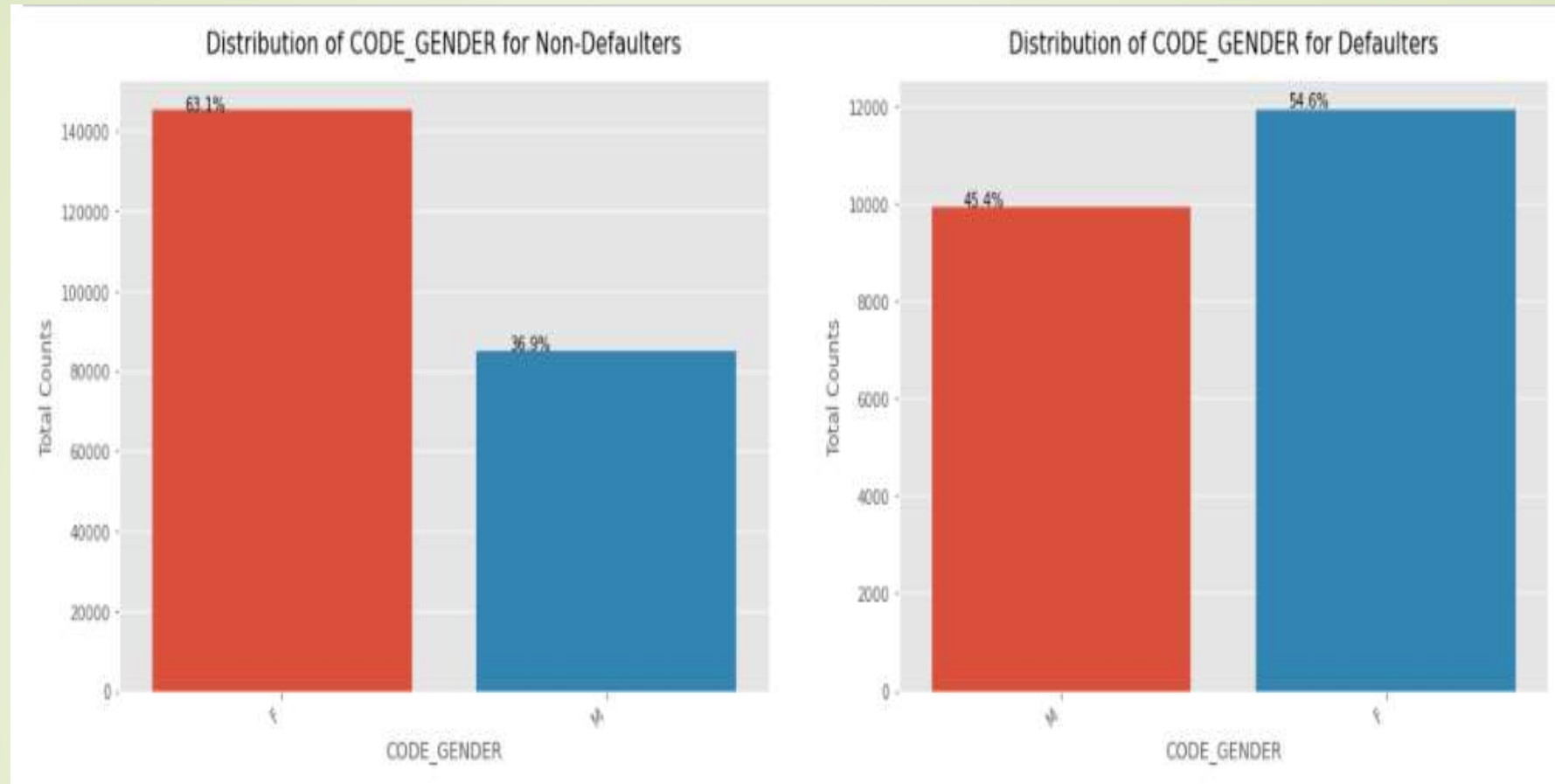
purpose of loan with payment status





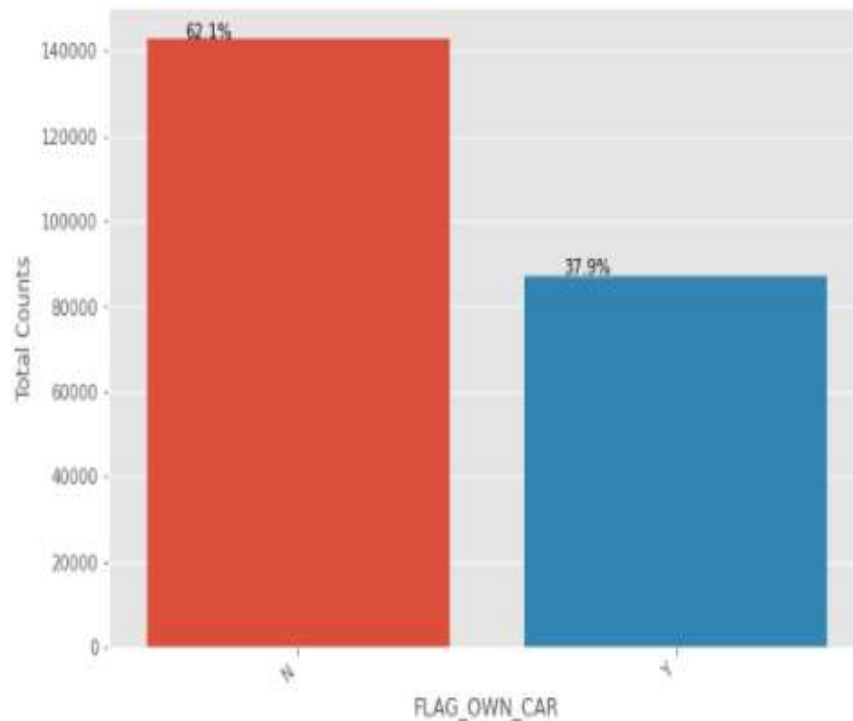
People who took the loan for repair purposes are unable to pay the loan purposes like Buying a garage, Buying a new car have high ratio of loan paid to loan unpaid therefore set a high eligibility for loan approval



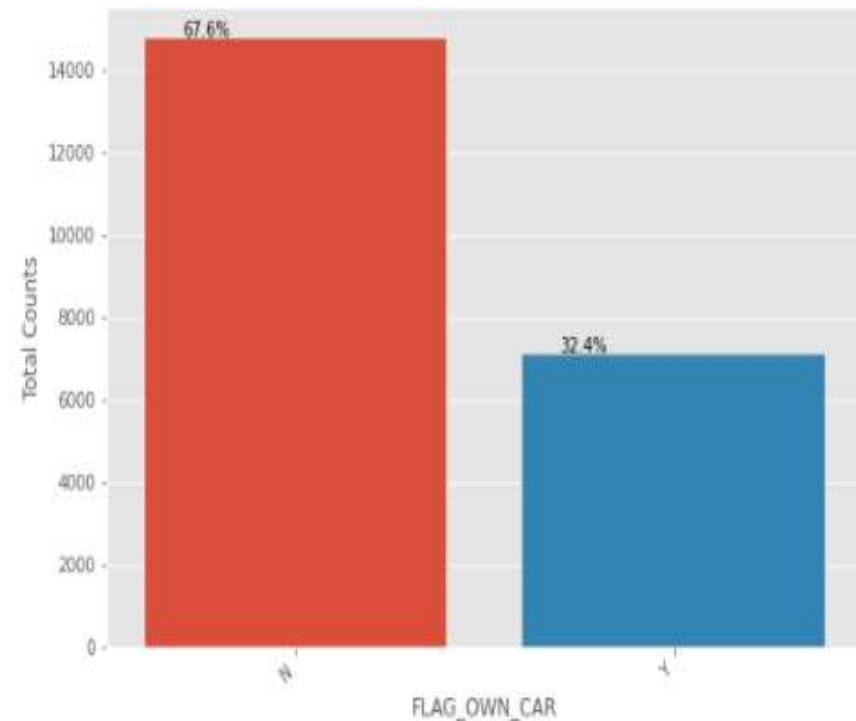


We can see that Female contribute 67% to the non-defaulters while 57% to the defaulters. We can conclude that We see more female applying for loans than males and hence the more number of female defaulters as well. But the rate of defaulting of FEMALE is much lower compared to their MALE counterparts.

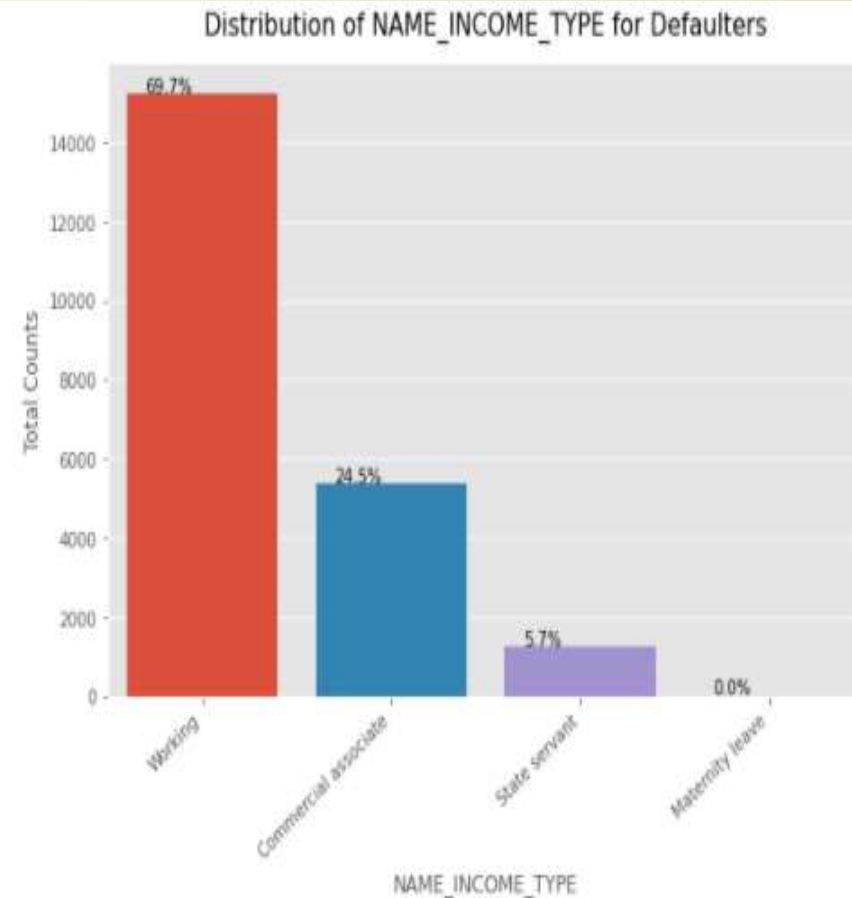
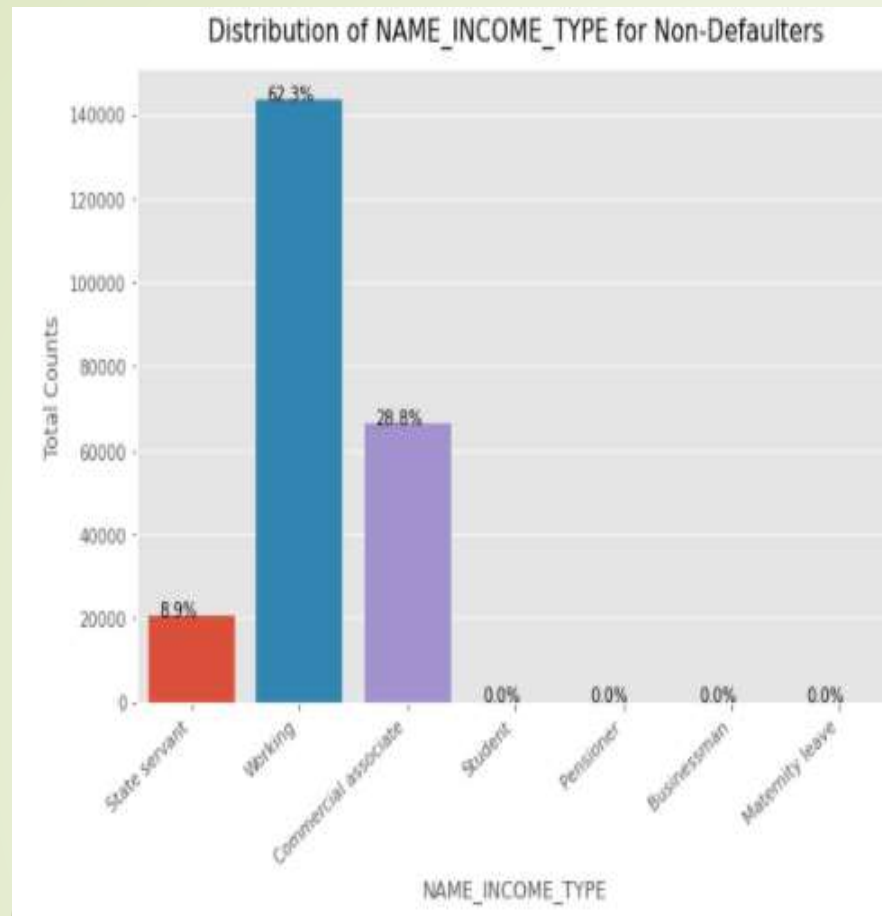
Distribution of FLAG_OWN_CAR for Non-Defaulters



Distribution of FLAG_OWN_CAR for Defaulters

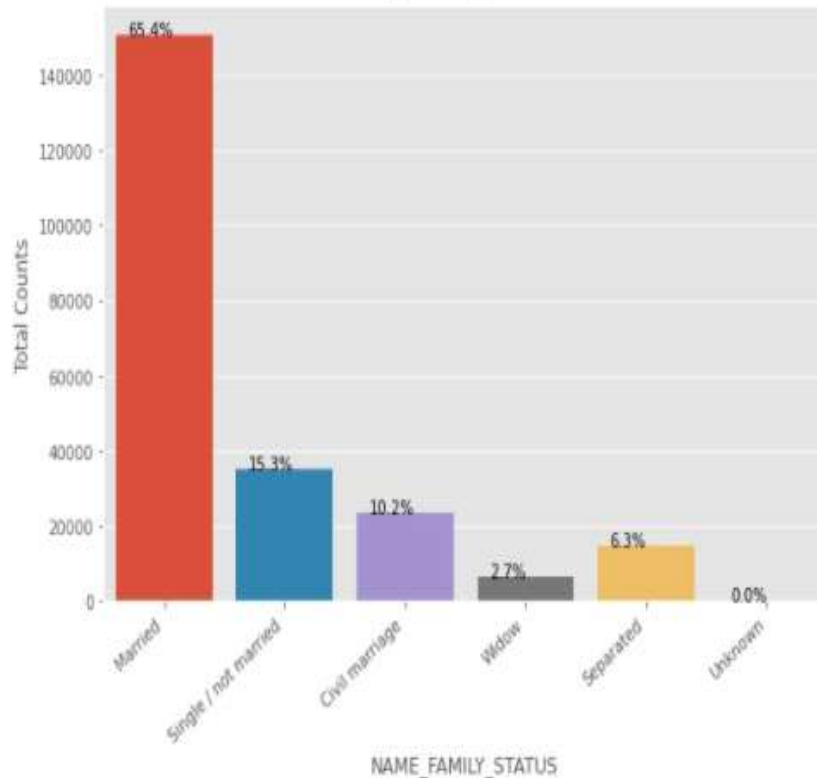


We can see that people with cars contribute 65.7% to the non-defaulters while 69.5% to the defaulters. We can conclude that While people who have car default more often, the reason could be there are simply more people without cars Looking at the percentages in both the charts, we can conclude that the rate of default of people having car is low compared to people who don't.

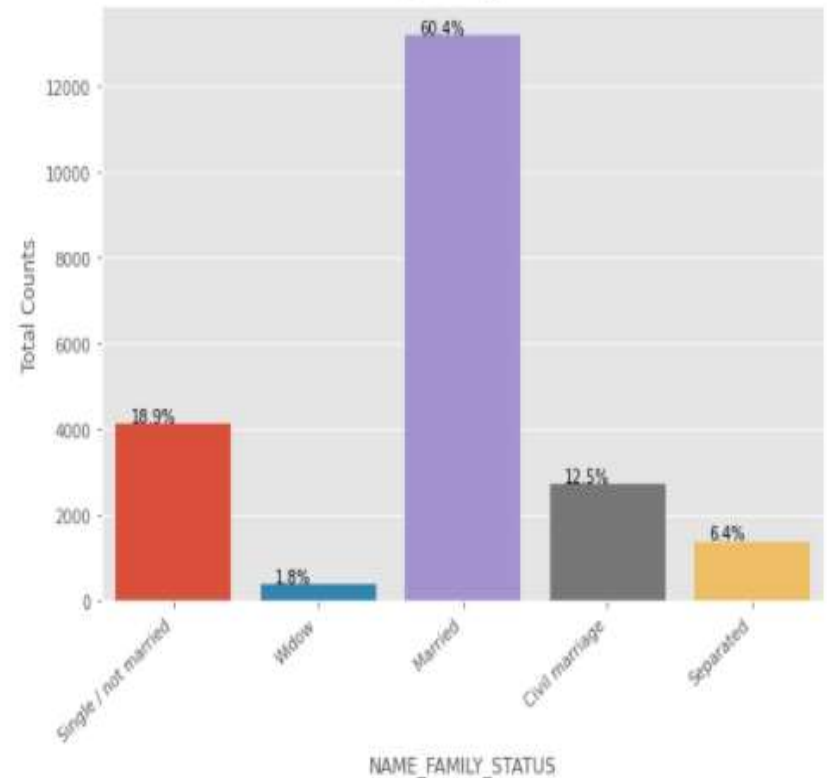


We can notice that the students don't default. The reason could be they are not required to pay during the time they are students. We can also see that the Business Men never default. Most of the loans are distributed to working class people. We also see that working class people contribute 51% to non defaulters while they contribute to 61% of the defaulters. Clearly, the chances of defaulting are more in their case.

Distribution of NAME_FAMILY_STATUS for Non-Defaulters

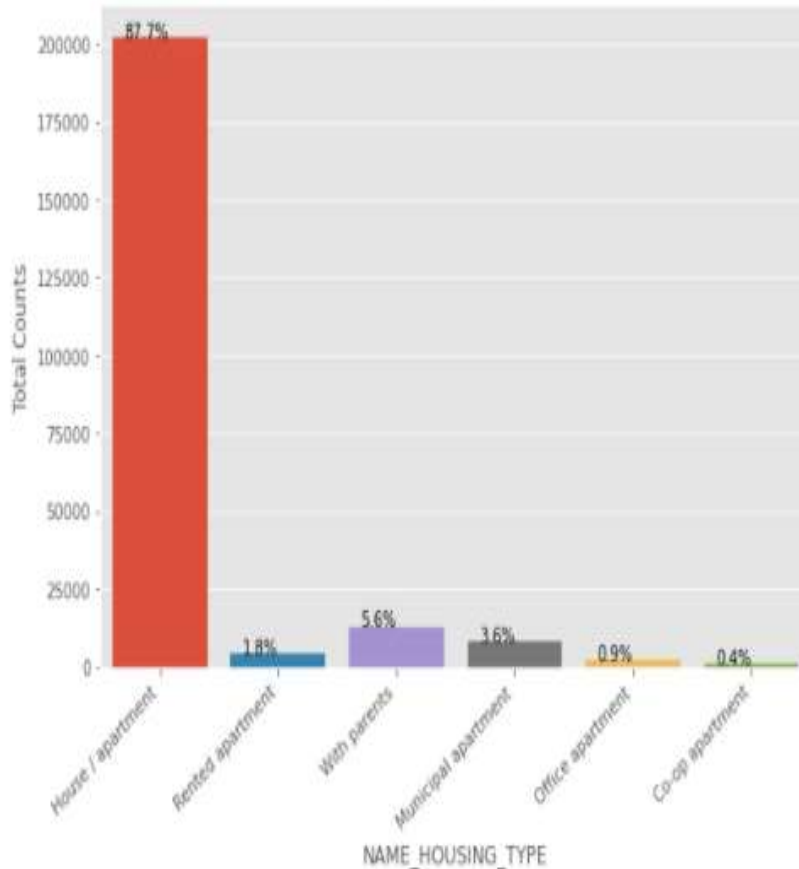


Distribution of NAME_FAMILY_STATUS for Defaulters

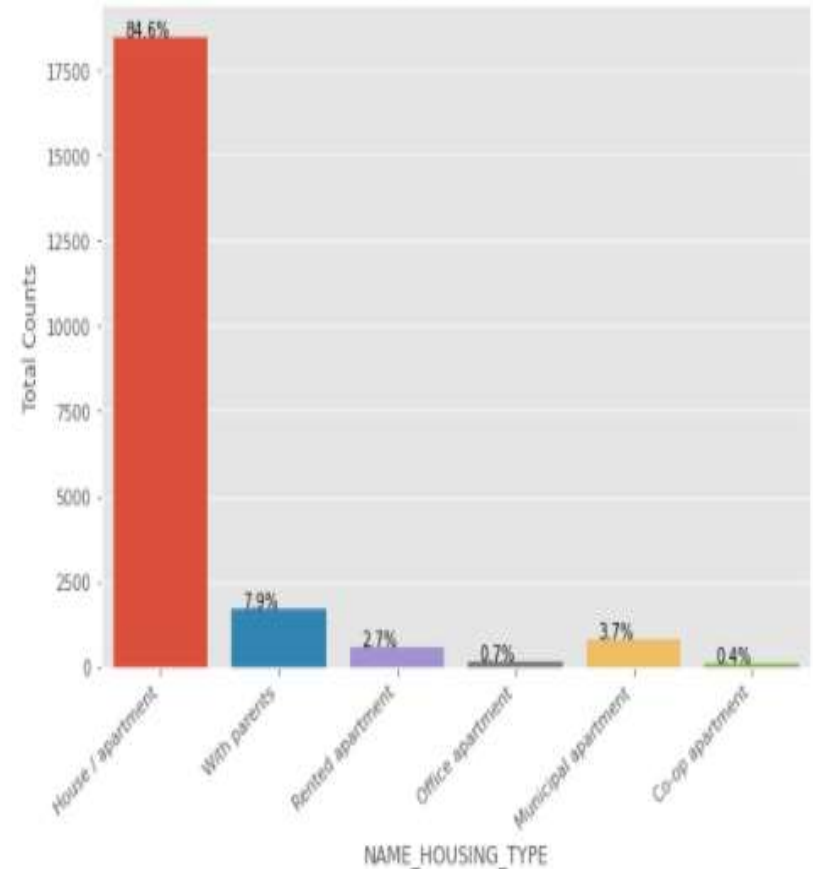


Married people tend to apply for more loans comparatively. But from the graph we see that Single/non Married people contribute 14.5% to Non Defaulters and 18% to the defaulters. So there is more risk associated with them.

Distribution of NAME_HOUSING_TYPE for Non-Defaulters

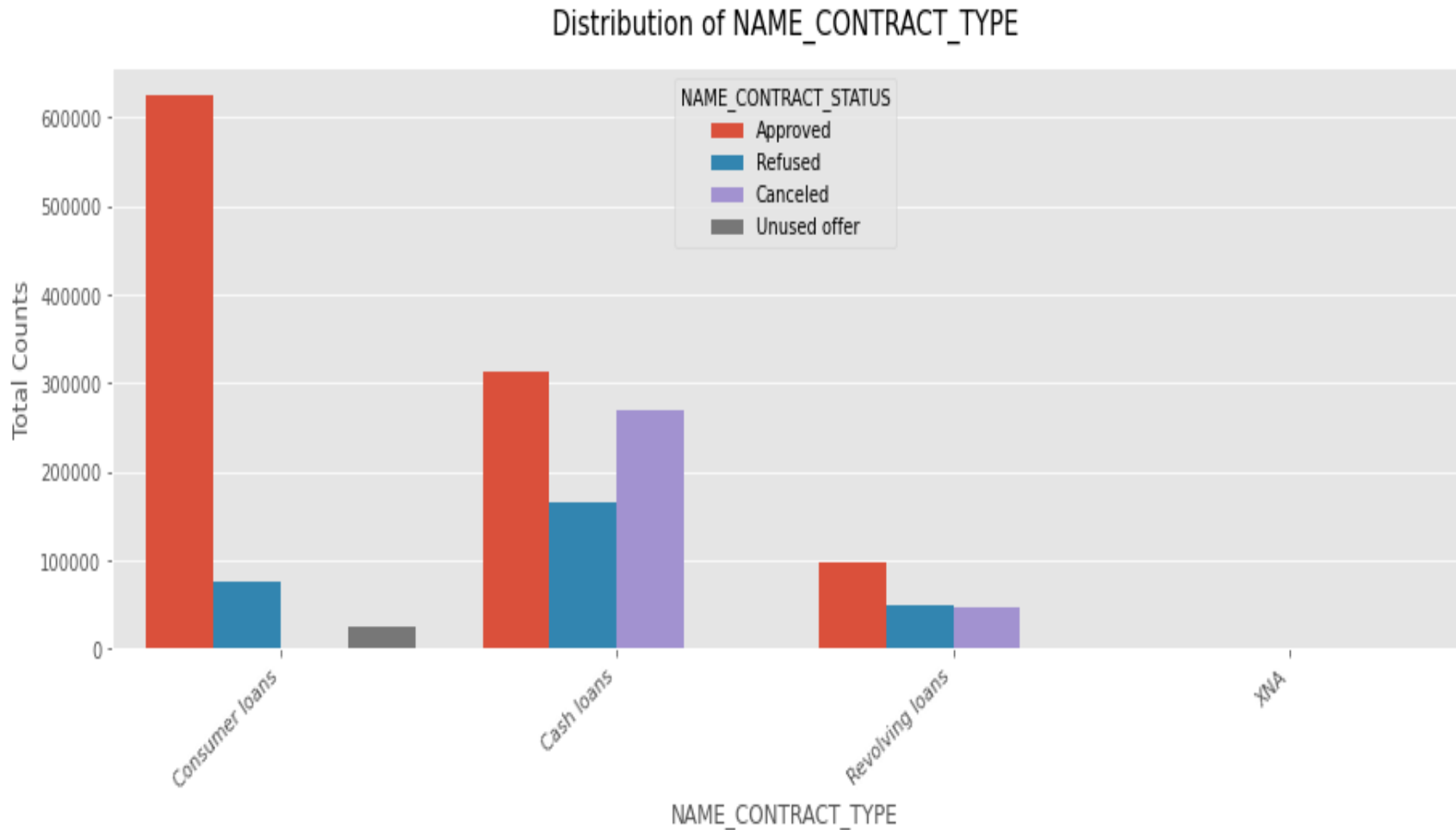



Distribution of NAME_HOUSING_TYPE for Defaulters



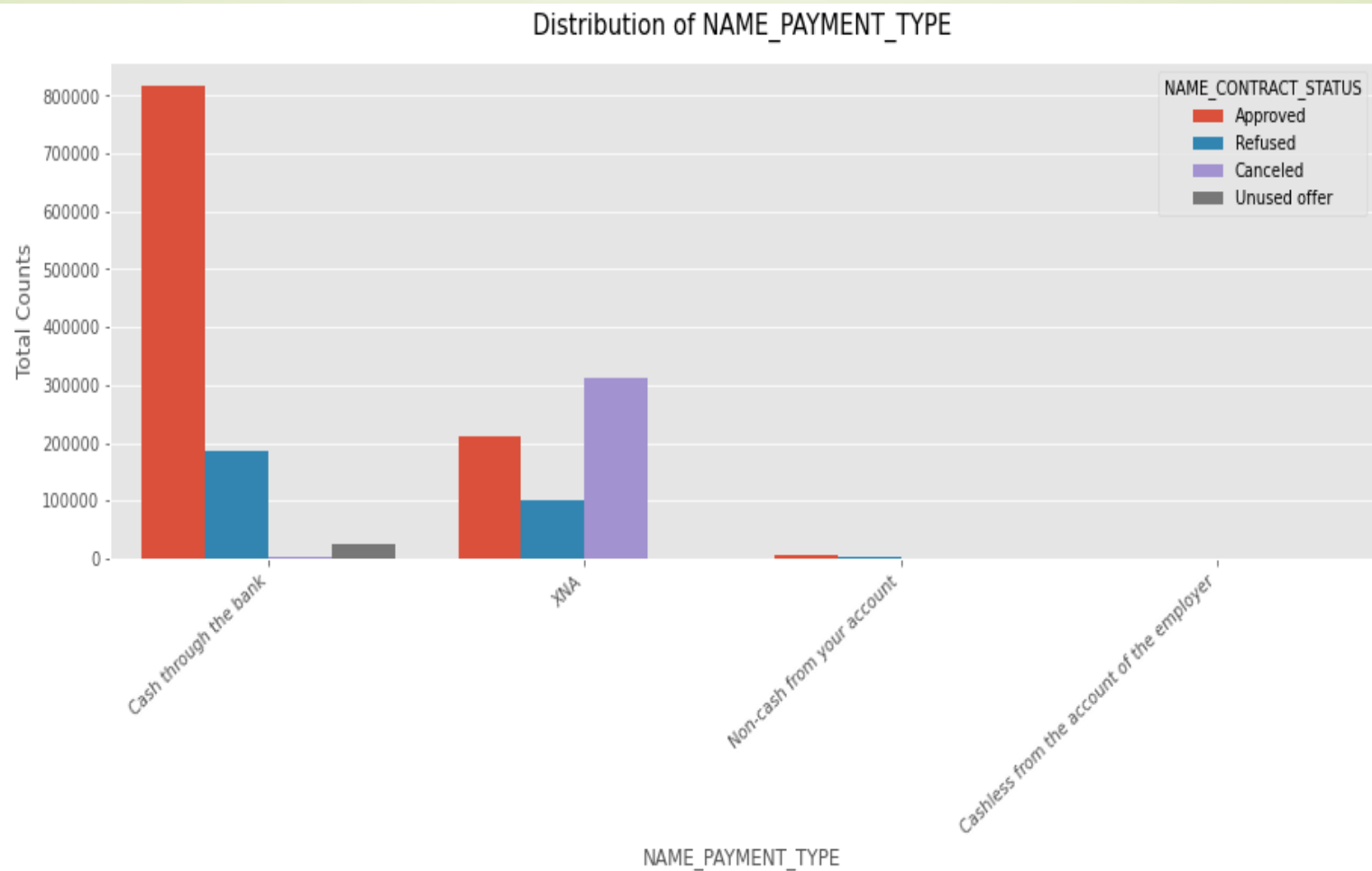
It is clear from the graph that people who have House/Apartment, tend to apply for more loans. People living with parents tend to default more often when compared with others. The reason could be their living expenses are more due to their parents living with them.

Data Analysis For Previous Application Data



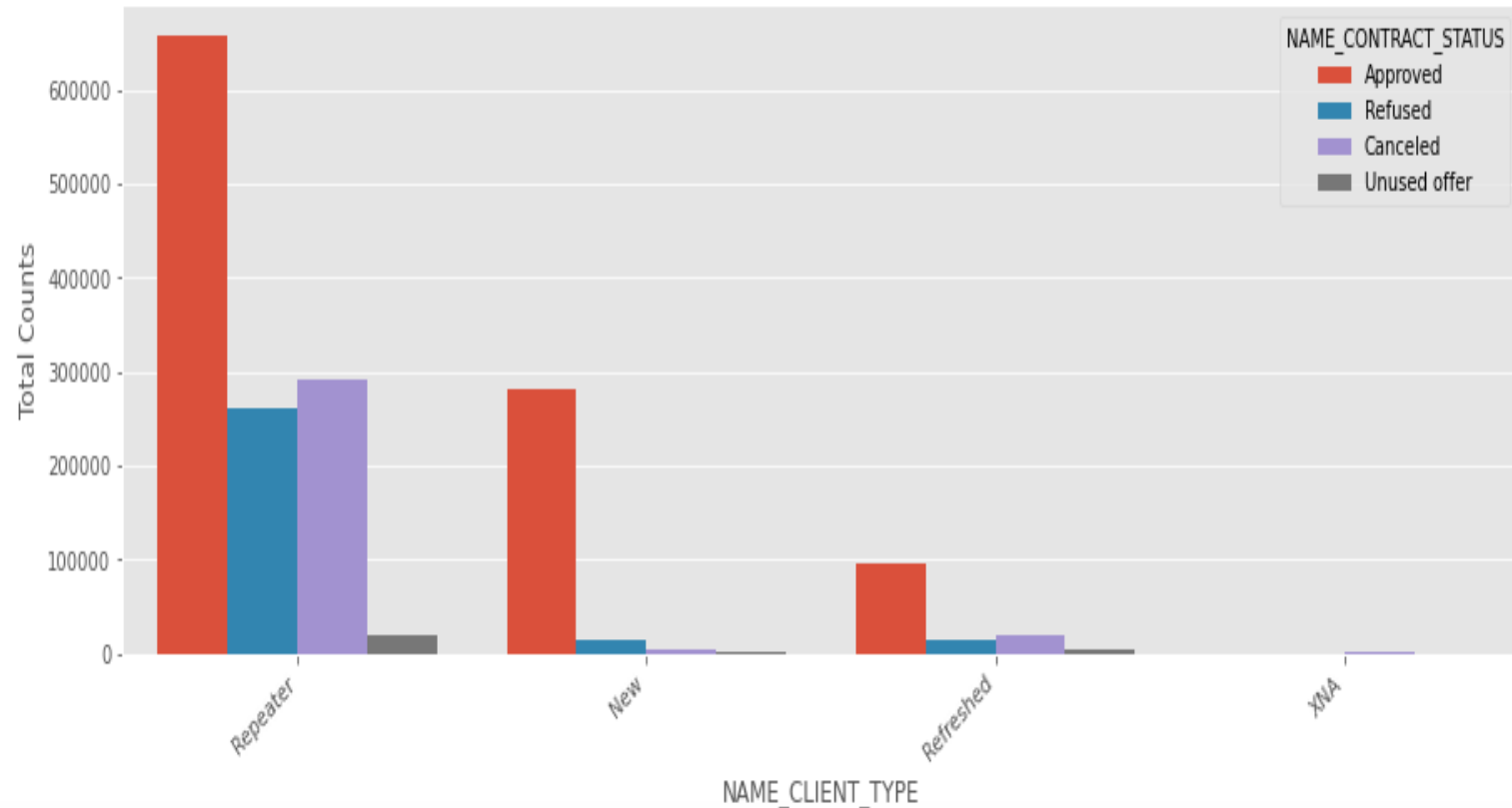


From the above chart, we can infer that, most of the applications are for 'Cash loan' and 'Consumer loan'. Although the cash loans are refused more often than others.

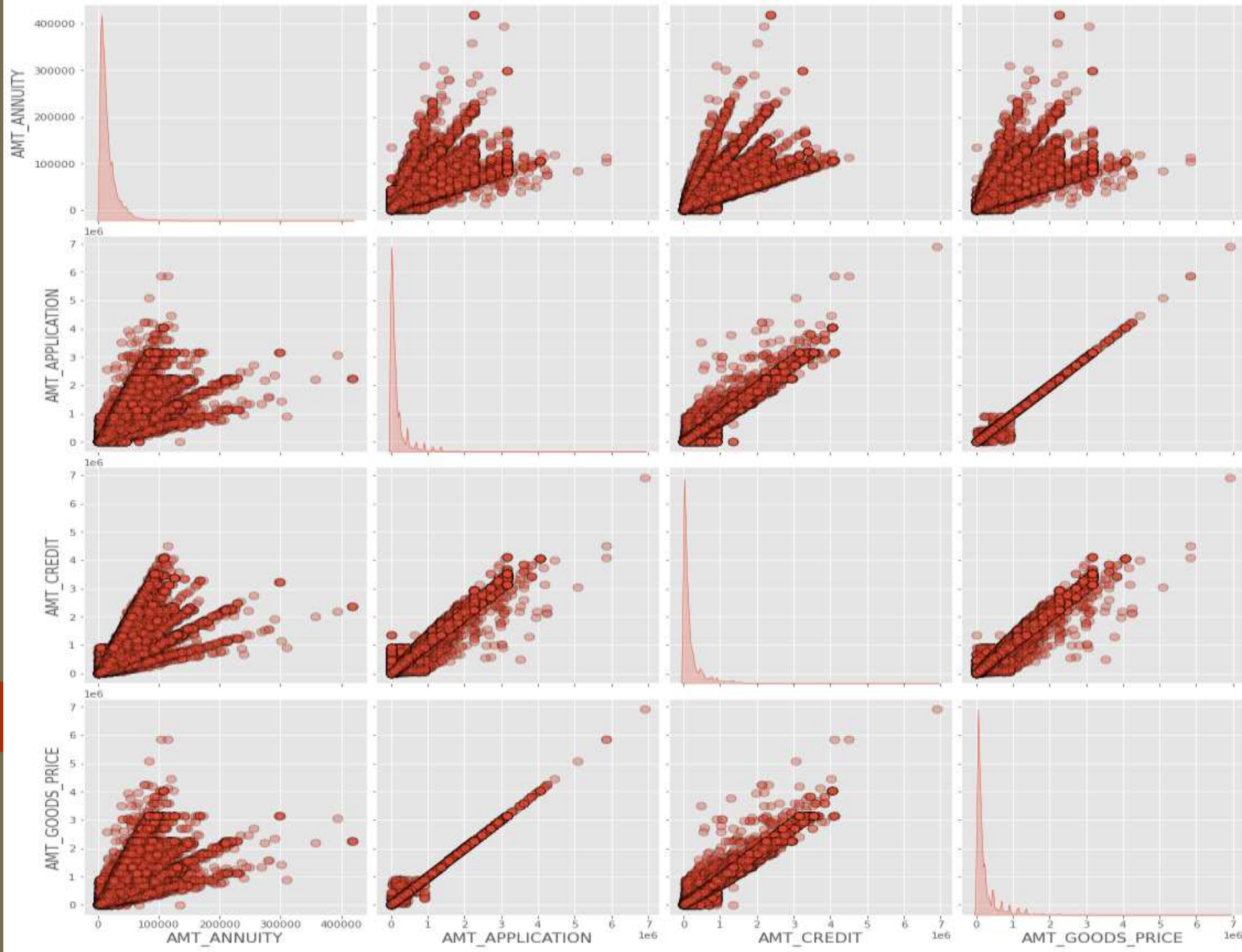



From the above chart, we can infer that most of the clients chose to repay the loan using the 'Cash through the bank' option. We can also see that 'Non-Cash from your account' & 'Cashless from the account of the employee' options are not at all popular in terms of loan repayment amongst the customers.

Distribution of NAME_CLIENT_TYPE



Most of the loan applications are from repeat customers, out of the total applications 70% of customers are repeaters. They also get refused most often.






Annuity of previous application has a very high and positive influence over: (Increase of annuity increases below factors)

- (1) How much credit did client asked on the previous application
- (2) Final credit amount on the previous application that was approved by the bank
- (3) Goods price of good that client asked for on the previous application.

For how much credit did client ask on the previous application is highly influenced by the Goods price of good that client has asked for on the previous application
Final credit amount disbursed to the customer previously, after approval is highly influence by the application amount and also the goods price of good that client asked for on the previous application.



Recommended group where loan can be credited. (Less chance to be defaulter) •

- Clients who are working as a state servant.
- Old people of any income group.
- Client with high income category.
- Old female client.
- Client with higher education (female).
- Any client who's previous loan was approved.
- Widow who has unused previous loan status.
- Refreshed client who has unused loan status previously.



Risky group

- Lower secondary educated clients are the most in number to be defaulted when their previous loans were cancelled or refused.
- Male clients with civil marriage.
- Previously refused loan status group



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.



THANK YOU

