

EDA Assignment

Bank Loan – Risk Analytics

Business Objectives

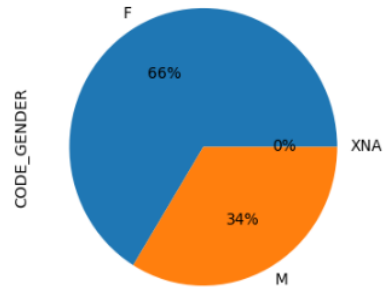
This case study aims to identify patterns which indicate if a client has difficulty paying their instalments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. This will ensure that the consumers capable 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.

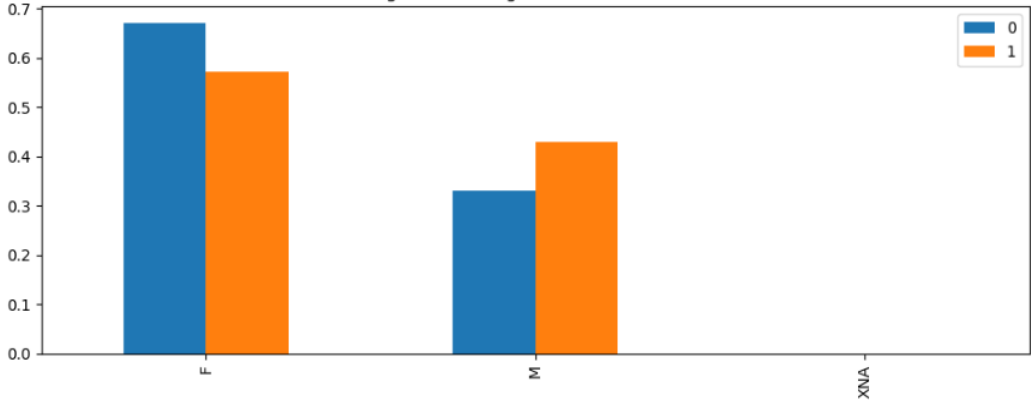
Insight



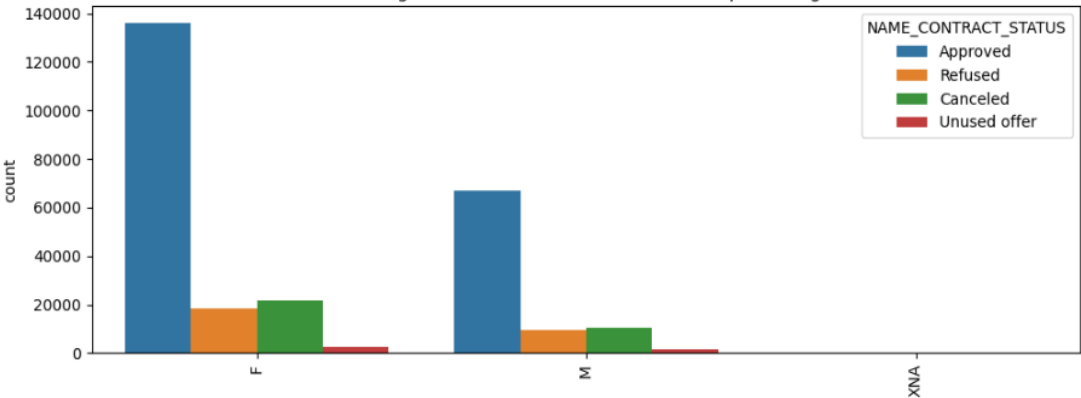
Plot for the column: CODE_GENDER on overall data



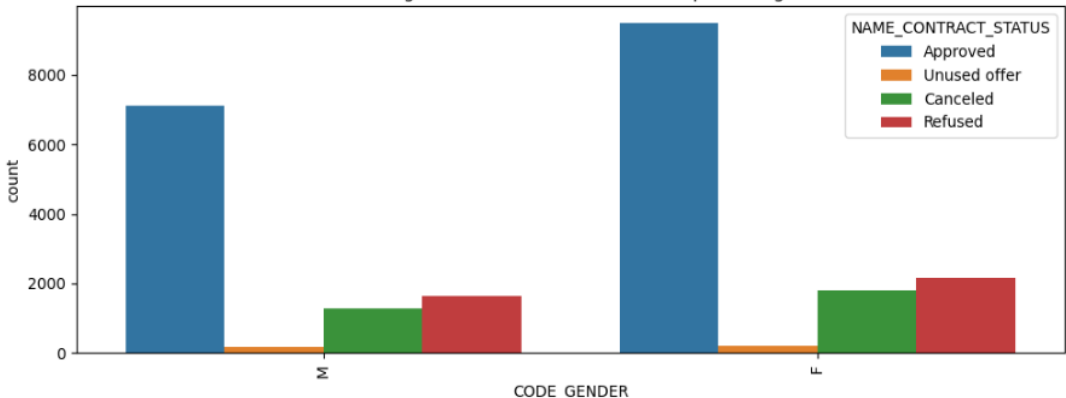
Plotting data for target in terms of total count



Plotting data for Non-Defaulter in terms of percentage



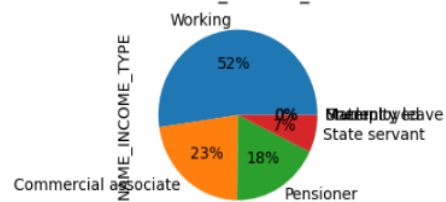
Plotting data for Defaulter in terms of percentage



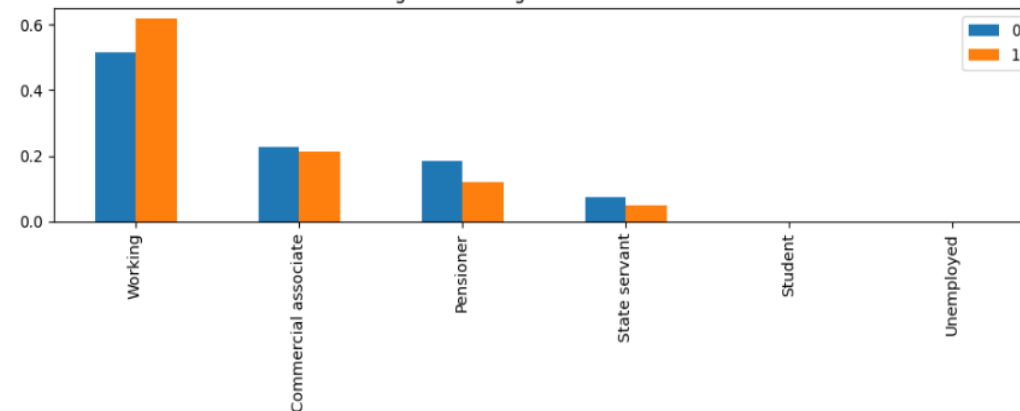
Females got more loan got approved when compared to men and they have less repayer percentage when compared to male



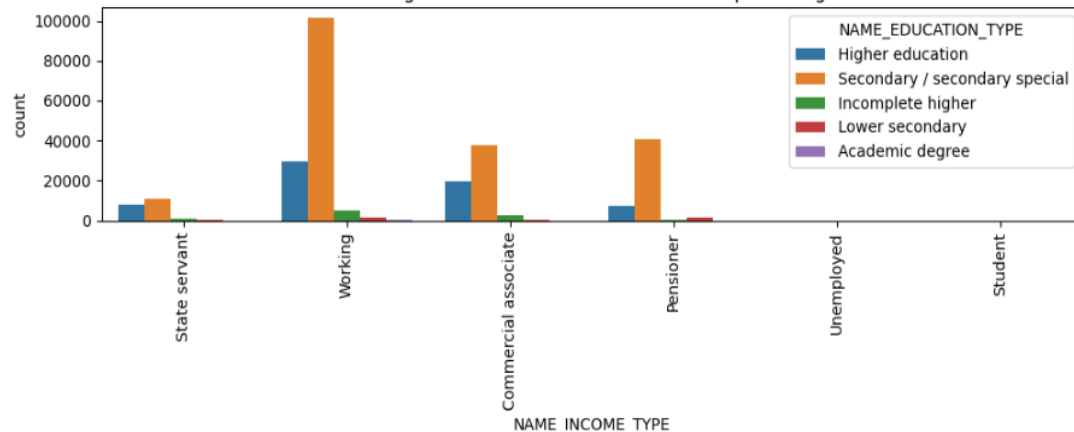
Plot for the column: NAME_INCOME_TYPE on overall data



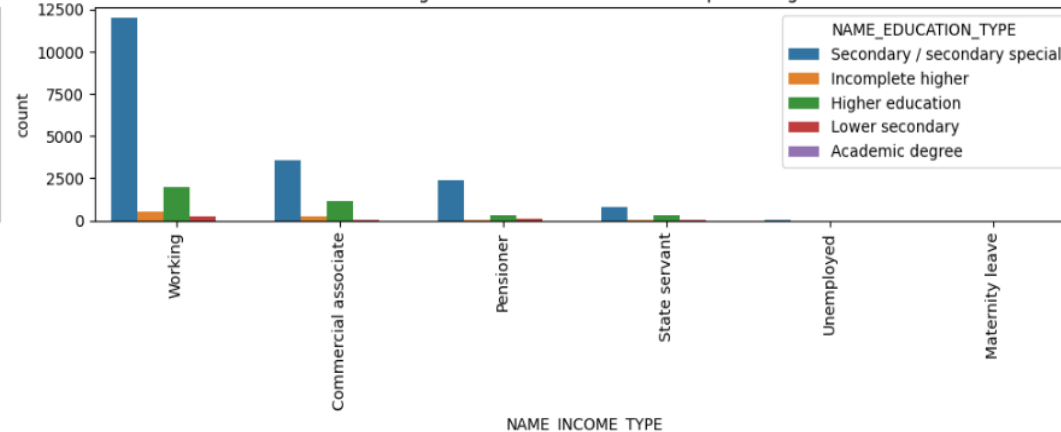
Plotting data for target in terms of total count



Plotting data for Non-Defaulter in terms of percentage



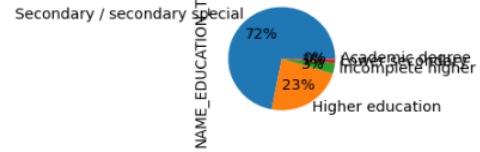
Plotting data for Defaulter in terms of percentage



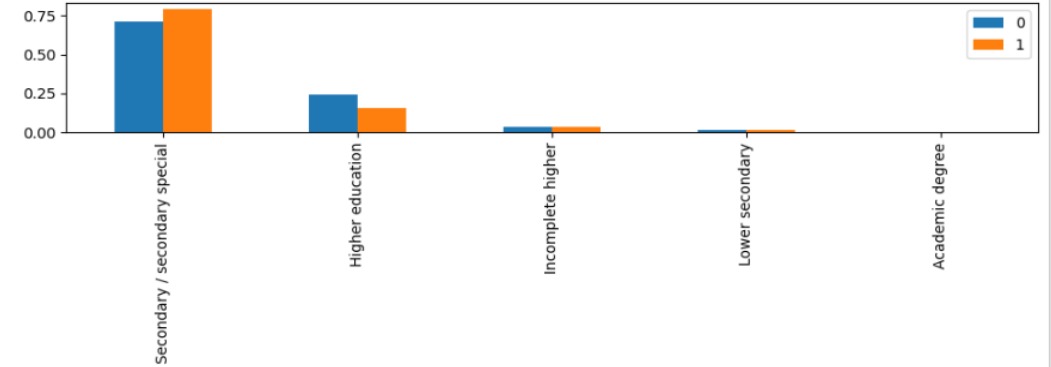
50% of loan applicants are income earners through working and they have more customer with payment difficulties and the defaulter and non defaulter working customer are highly educated bank can give loan to other customer like commercial associate , pensioner they have quite less payment difficulties.



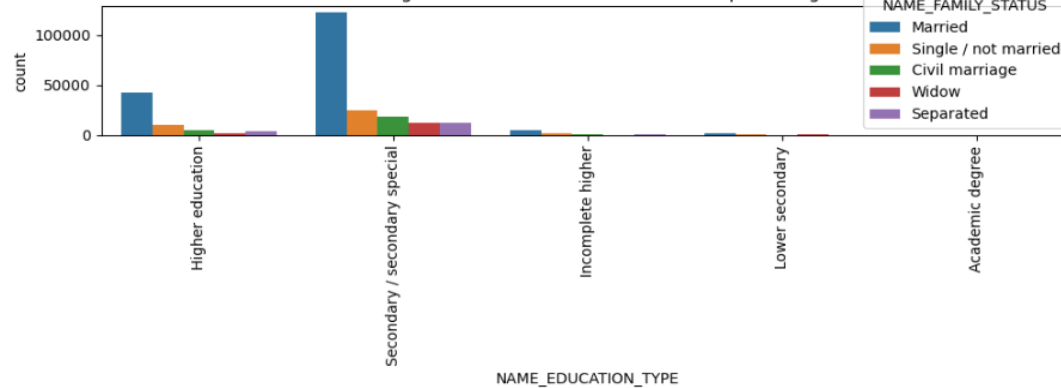
Plot for the column: NAME_EDUCATION_TYPE on overall data



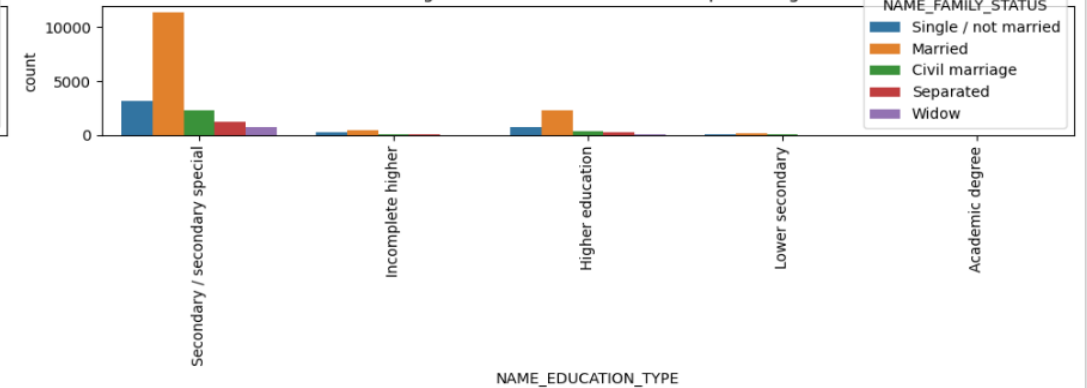
Plotting data for target in terms of total count



Plotting data for Non-Defaulter in terms of percentage



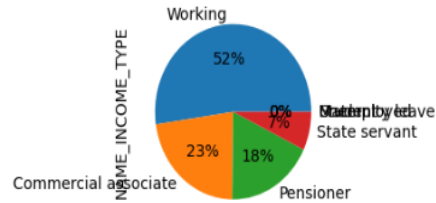
Plotting data for Defaulter in terms of percentage



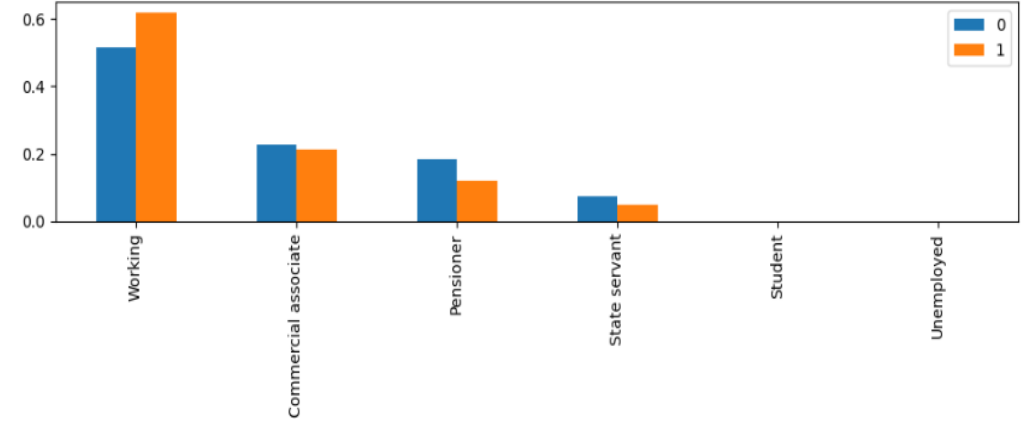
Around 3/4 of the loan application customer are well educated people of secondary special and most of them are married in both defaulter and non-defaulter but they faces payment difficulties bank can increase the loan offer to higher education they are quite less in payment difficulties.



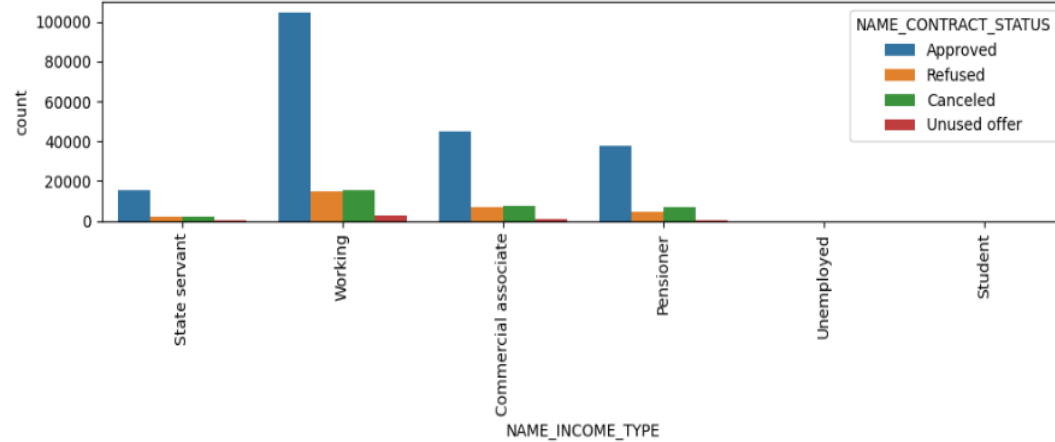
Plot for the column: NAME_INCOME_TYPE on overall data



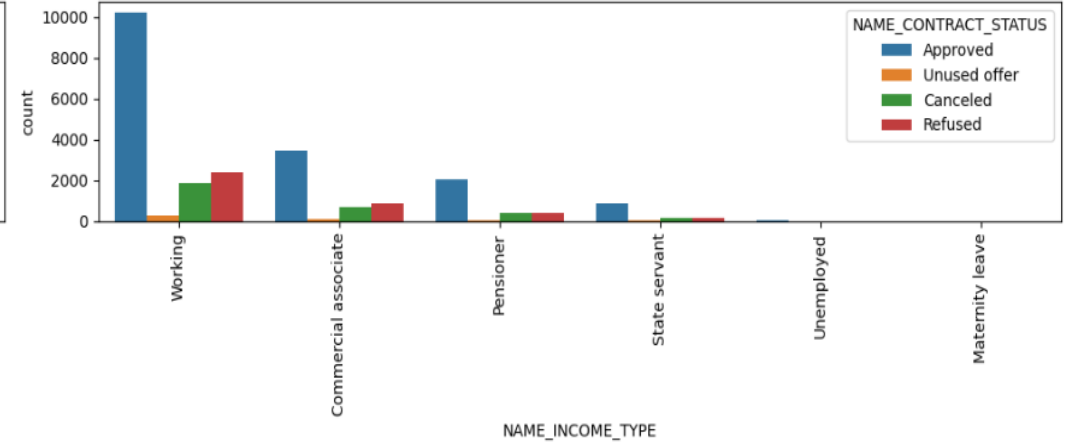
Plotting data for target in terms of total count



Plotting data for Non-Defaulter in terms of percentage



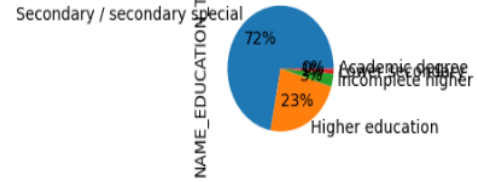
Plotting data for Defaulter in terms of percentage



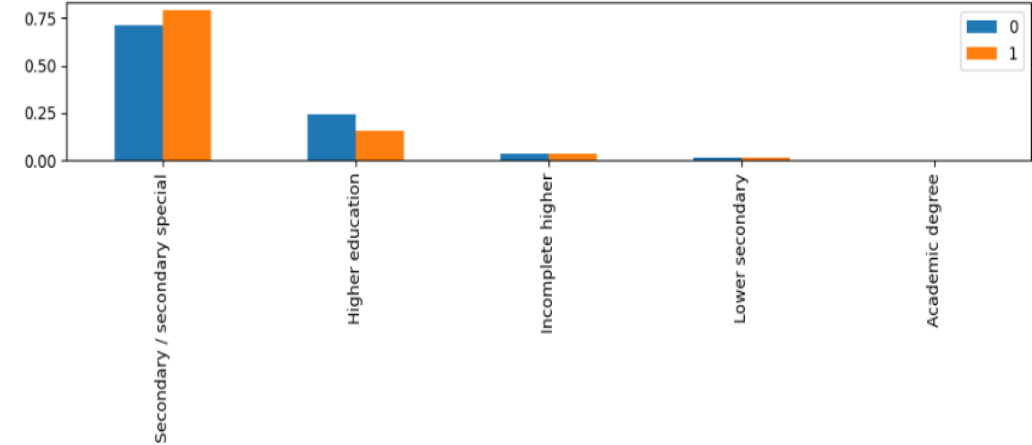
For both the defaulter and non-defaulter the loan is being approved mostly for the working customer when compared to other customer but they have payment difficulties and can increase the loan application for others.



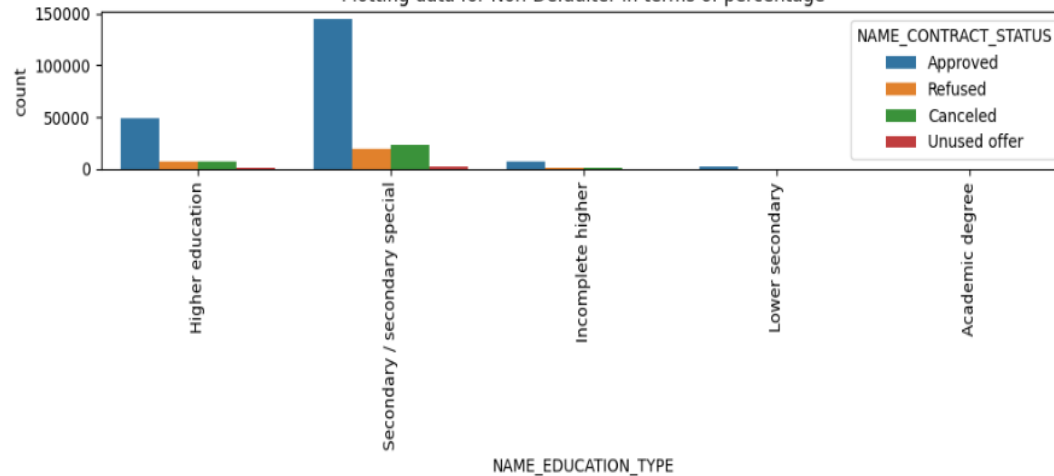
Plot for the column: NAME_EDUCATION_TYPE on overall data



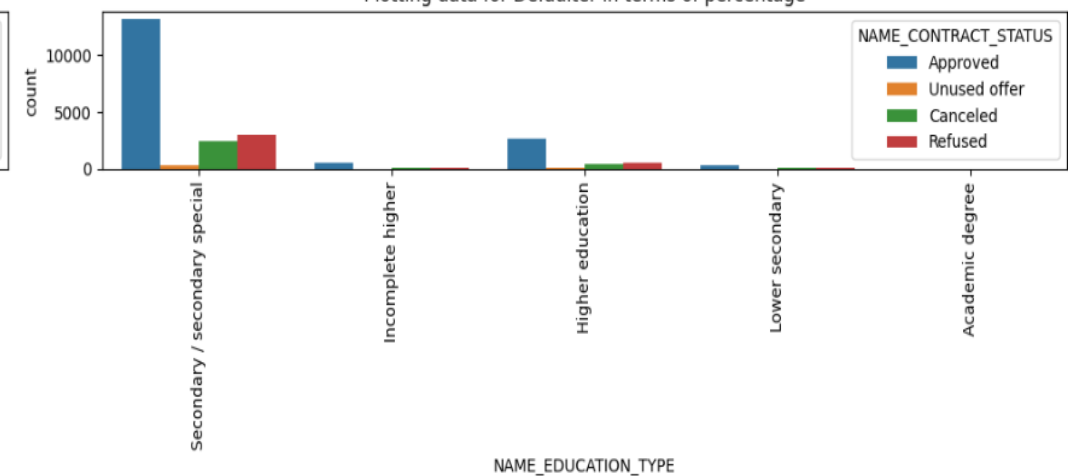
Plotting data for target in terms of total count



Plotting data for Non-Defaulter in terms of percentage



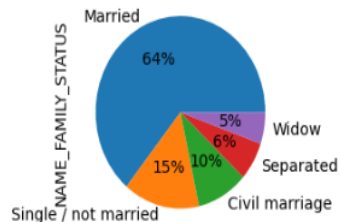
Plotting data for Defaulter in terms of percentage



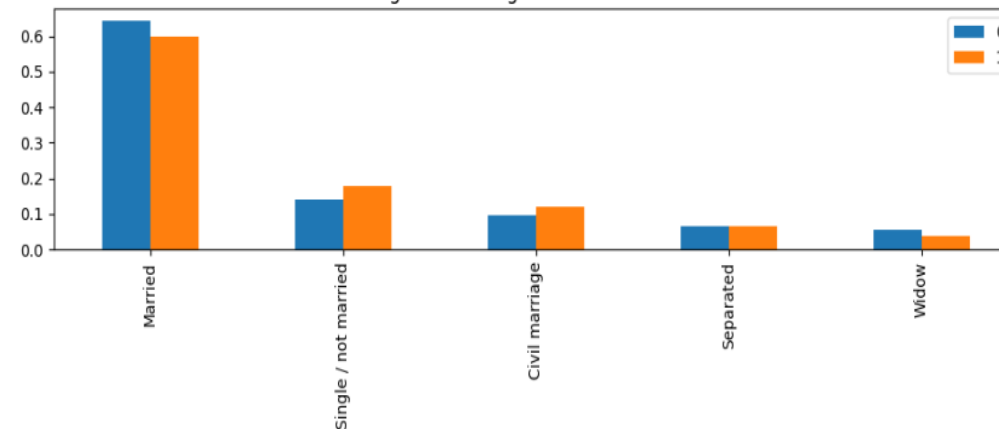
Similar to the above analysis the loan for the both defaulter and non-defaulter are approved highly for the secondary special educated customer but they have more payment difficulties so bank can give loan to higher education customer.



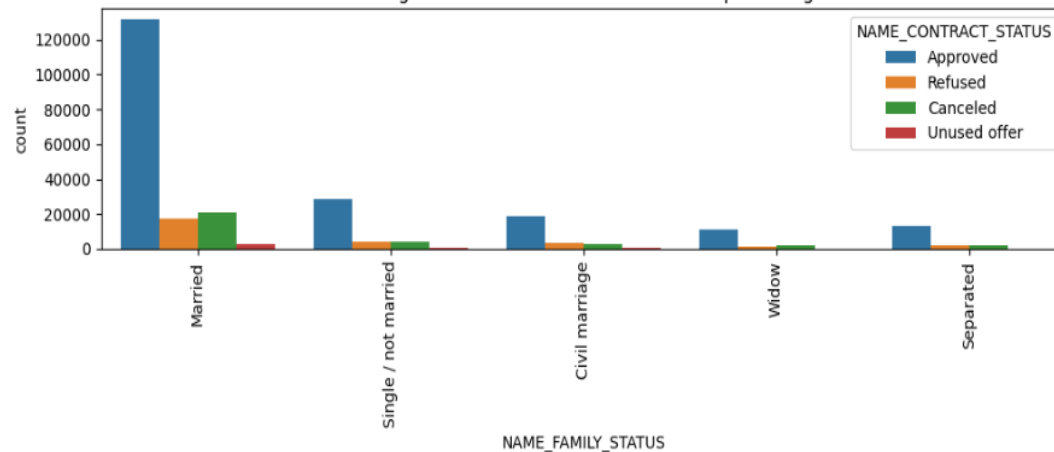
Plot for the column: NAME_FAMILY_STATUS on overall data



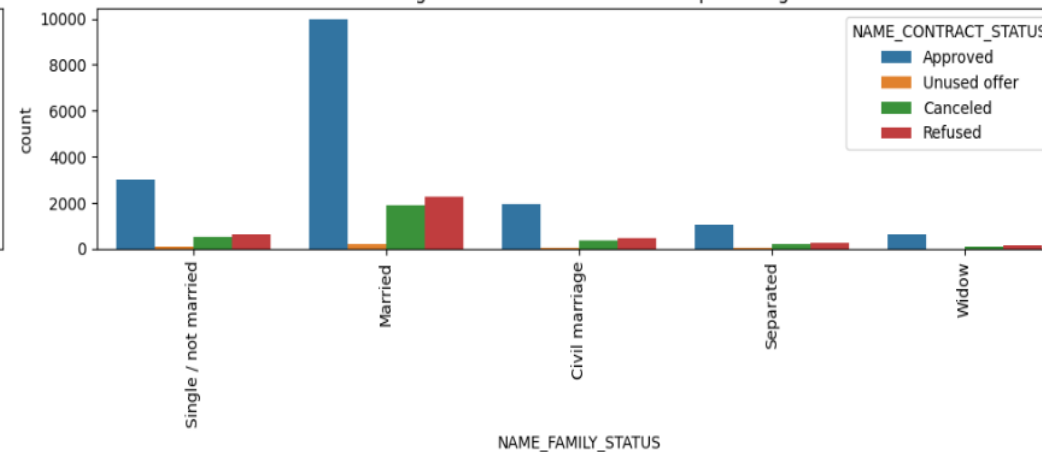
Plotting data for target in terms of total count



Plotting data for Non-Defaulter in terms of percentage



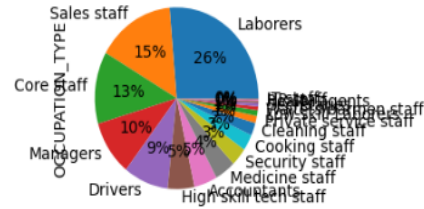
Plotting data for Defaulter in terms of percentage



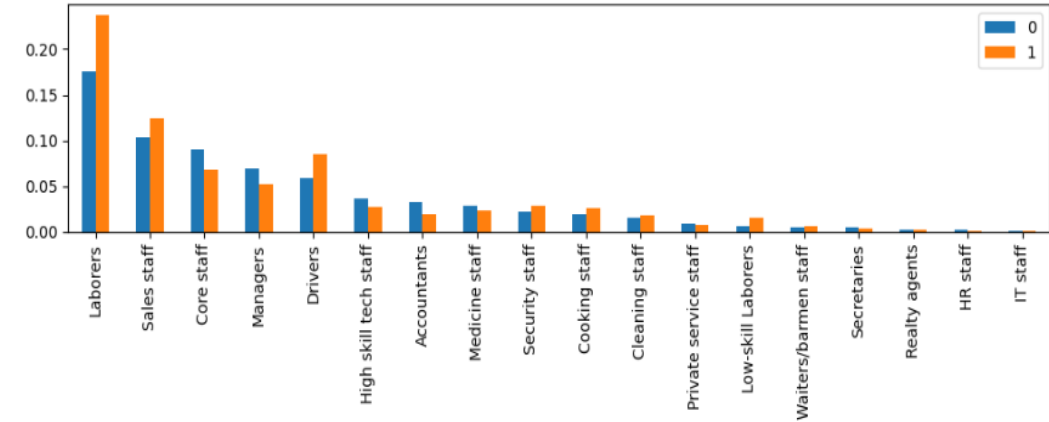
Married customer are approved loan more in percentage when compared to other customer in both the cases and they are highly facing payment difficulties when compared to others but the percentage of repayer is quite less in all the case when compared to payment difficulties.



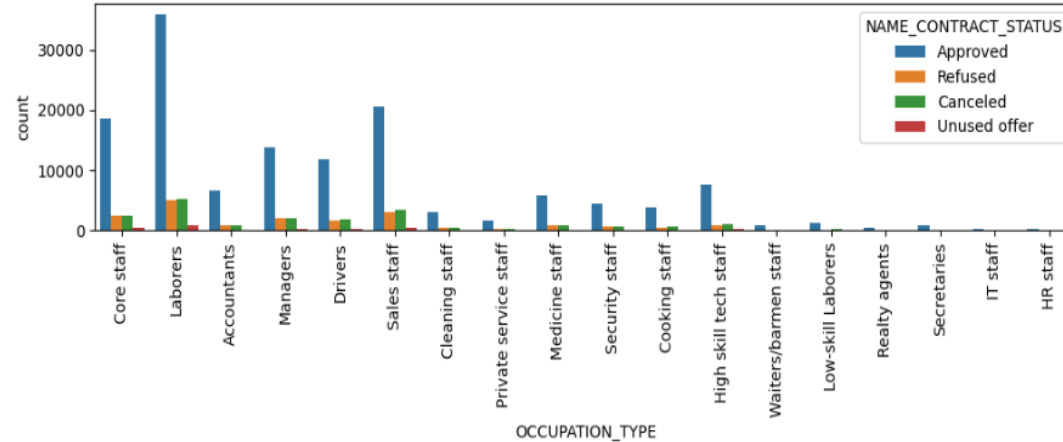
Plot for the column: OCCUPATION_TYPE on overall data



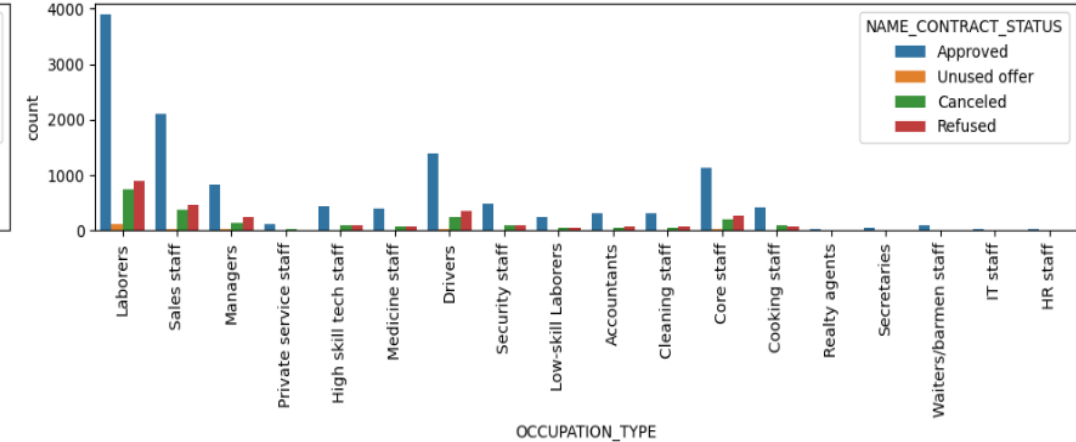
Plotting data for target in terms of total count



Plotting data for Non-Defaulter in terms of percentage



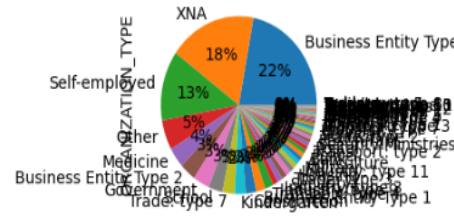
Plotting data for Defaulter in terms of percentage



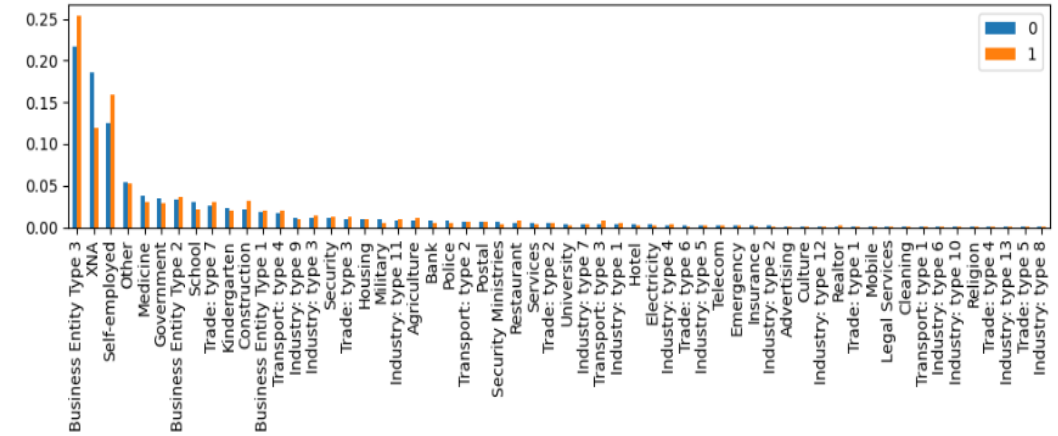
Laborers are the more customer who get loan and they have more payment difficulties and next comes the sales staff and core staff bank can increase the loan offer to core staff ,managers ..they have less payment difficulties when compared to others.



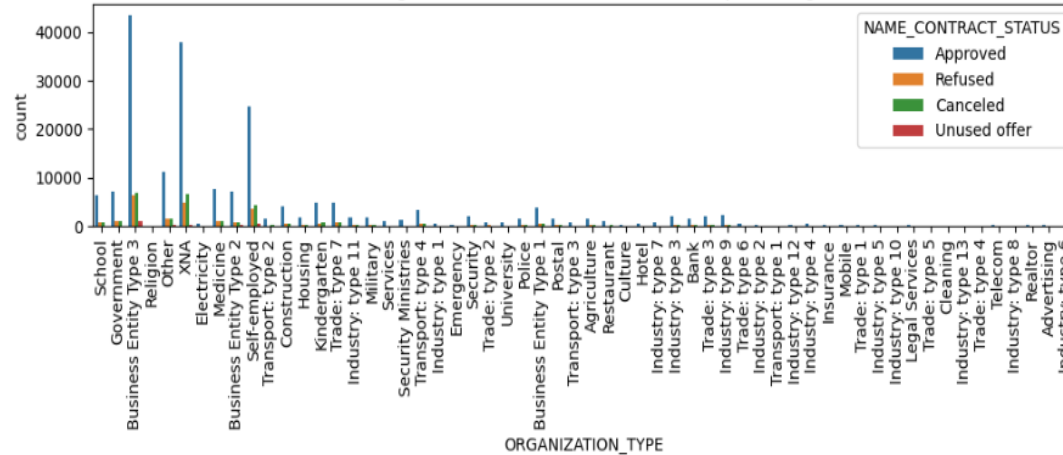
Plot for the column: ORGANIZATION_TYPE on overall data



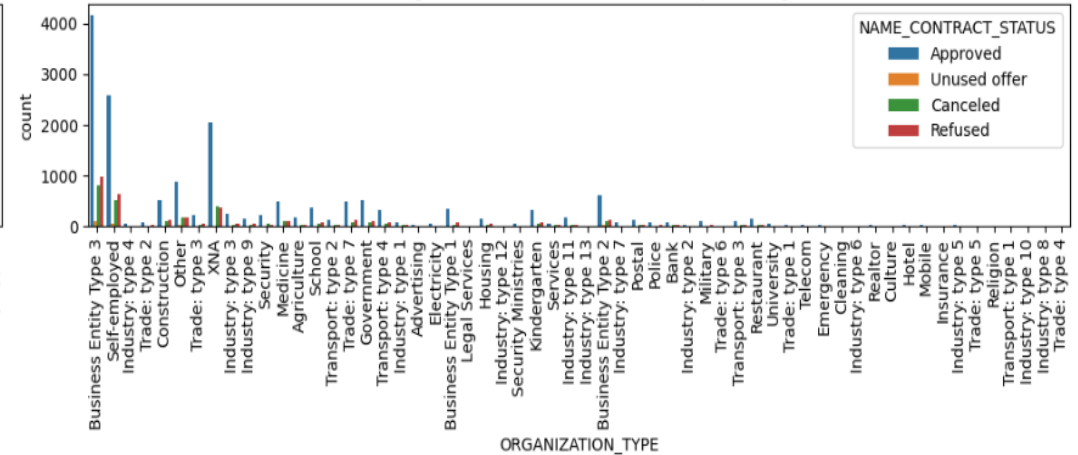
Plotting data for target in terms of total count



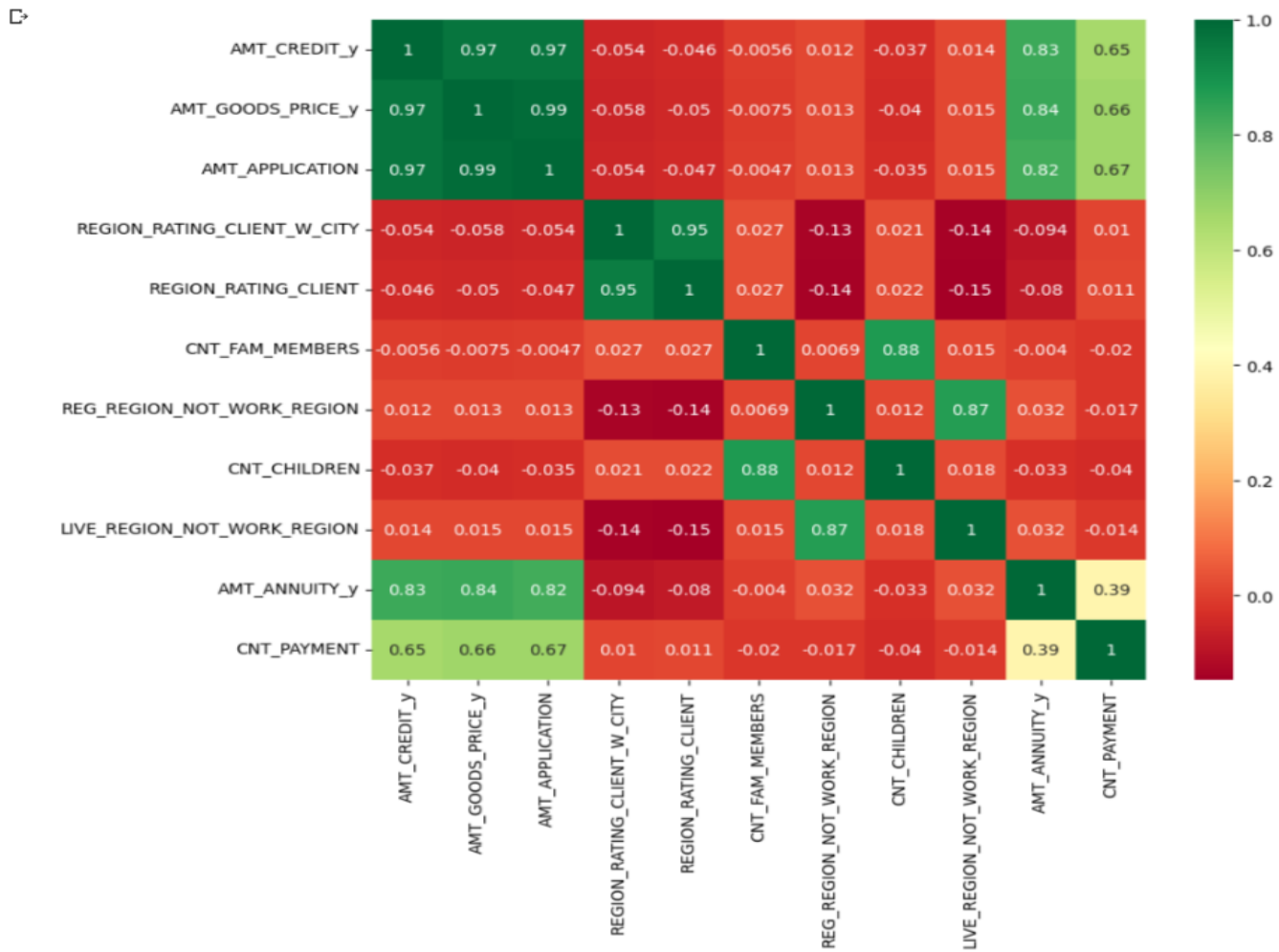
Plotting data for Non-Defaulter in terms of percentage



Plotting data for Defaulter in terms of percentage



Bank provided more loan for the business entity type 3, self-employed categories but they are the cases have more payment difficulties..which causes loss for the bank ..so they can choose other feild(like medicine, govt, etc.. for providing loan.



cnt_fam_members and cnt_children are highly correlated
amt_annuity_y and cnt_payment are correlated with amt_goods_price_y and
amt_credit_y

Conclusion

- Most of the customer are married , working labours who has taken the loan but they face difficulties in repaying the amount . So in these case bank can reduce the loan amount to avoid more risk on bank losses.
- Bank can target female customer more as they are good in repayment when compared to male customer.
- Those customer who has payment difficulties , bank can either cancel the loan or limit the loan amount to avoid bank loss.