

CREDIT EDA CASE STUDY

Credit Risk Analysis

By Pallabita Ghosh

Purpose

This analysis will help the bank to take decisions on loans whether bank should approve or reject his client for requested loan amount and how this decision will impact bank's profit or loss for current loan applicant.

Problem Statement

Application Data contains two type of scenario:

- The client with payment difficulties: He had late payment more than X days on at least one of the first Y installments of the loan in our sample
- All other cases: When payment are paid on time.

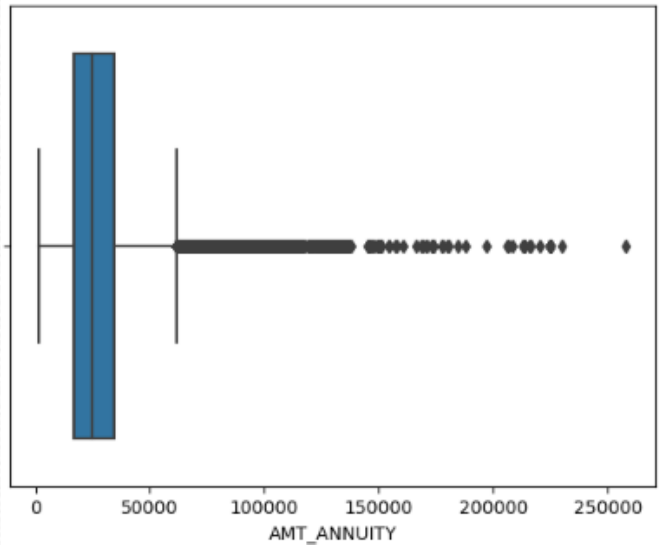
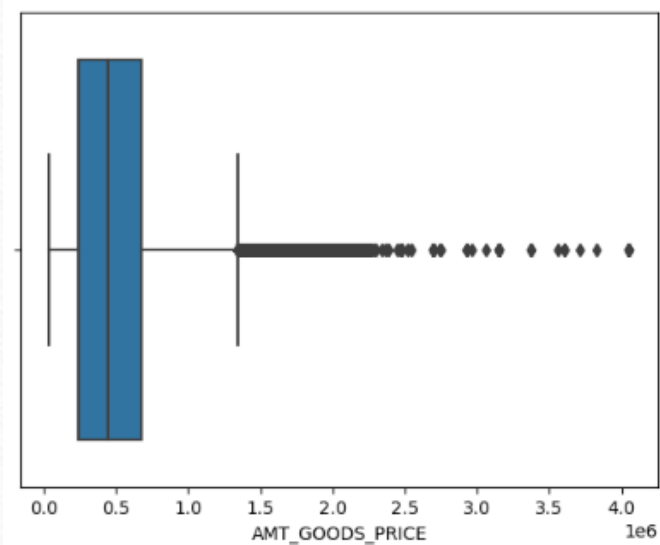
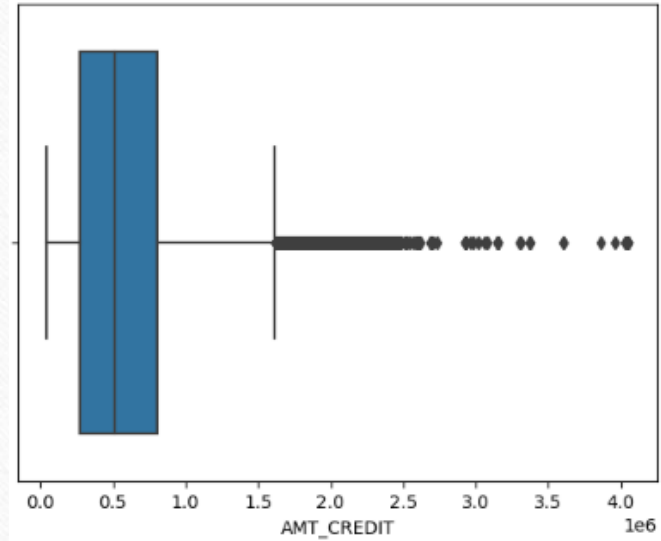
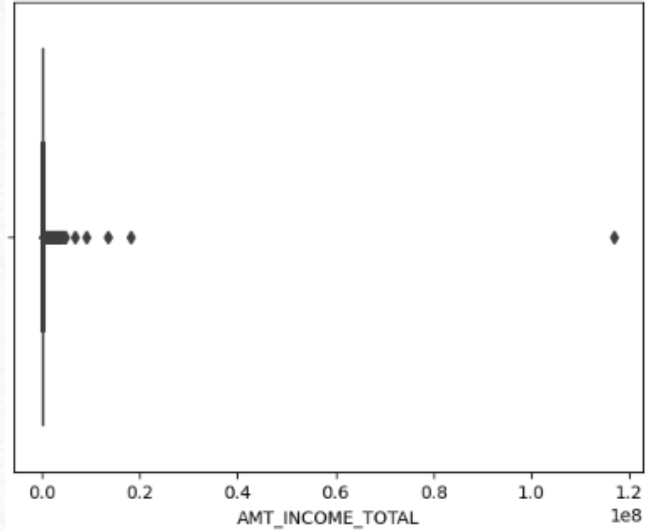
Previous Application Data contains four types of contract status:

- Approved
- Cancelled
- Refused
- Unused Offer

Steps Involved

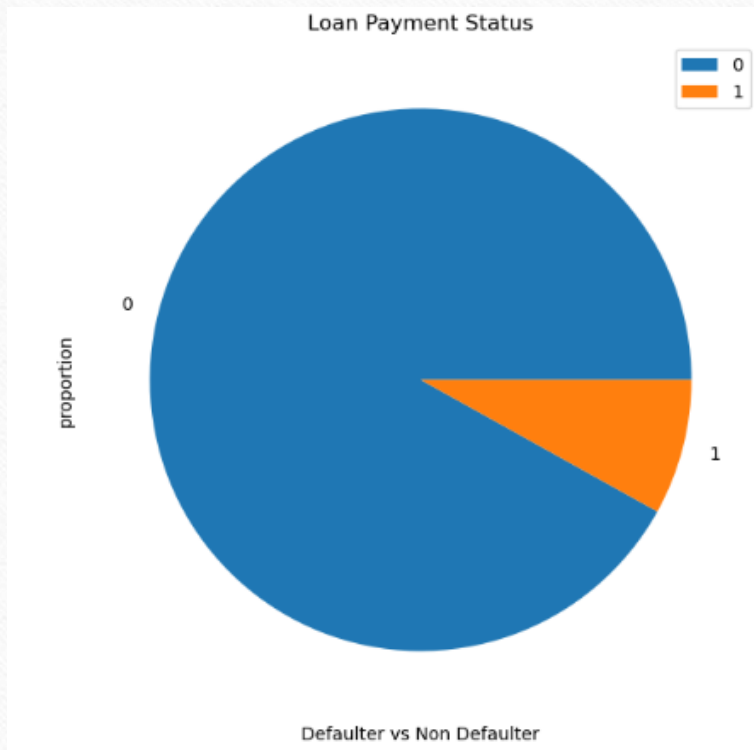
- Understanding and Importing Data
- Checking Data Quality and Missing Value
- Identifying Outliers
- Identifying Data Imbalance and checking for Univariate, Bivariate and Correlation
- Merging Current Application Data with Previous Application Data
- Analyzing Data through Univariate, Bivariate and Correlation
- Conclusion

Analysis of Application Dataset



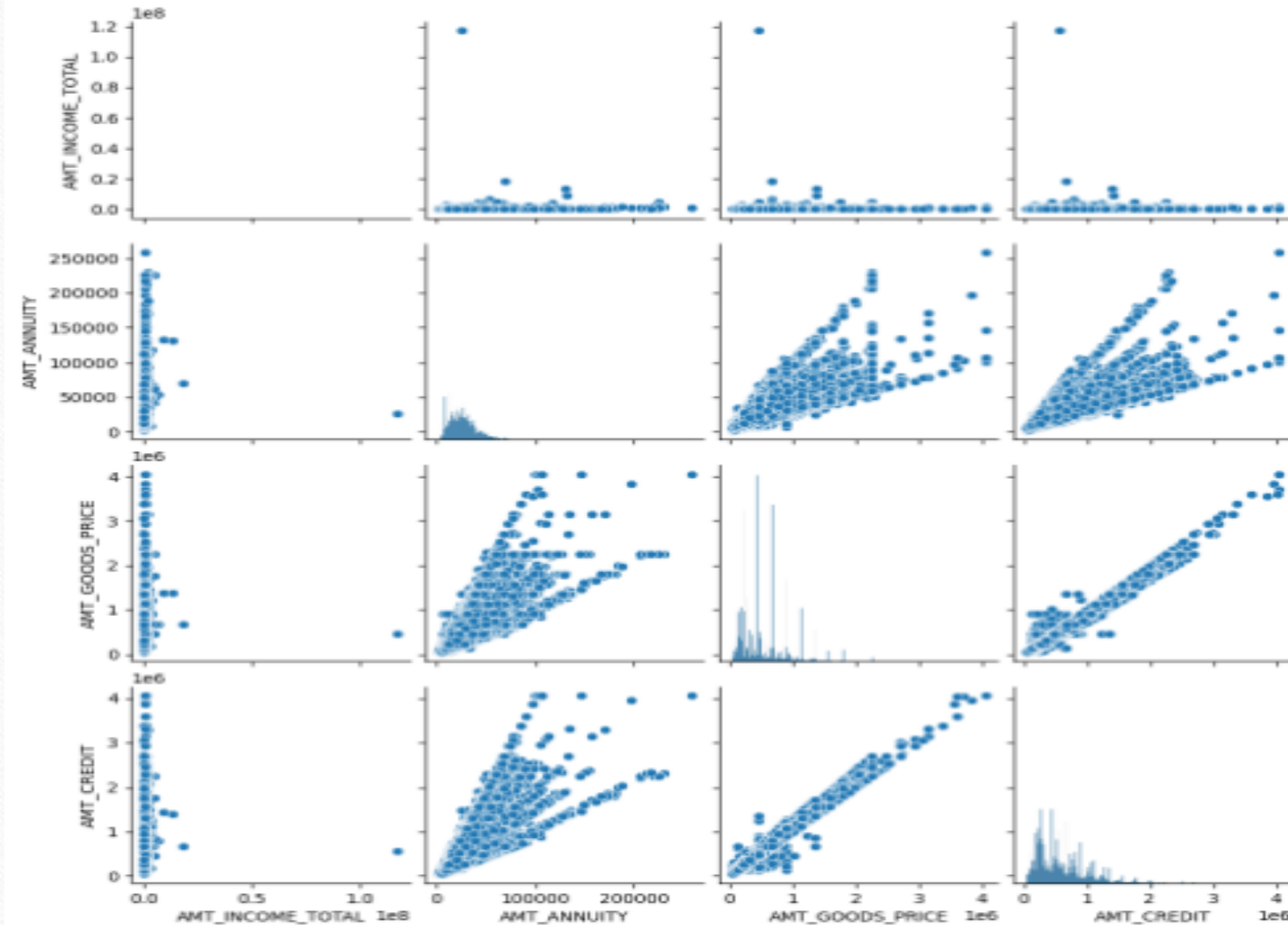
From this graph we can see total income, credit amount, annuity amount and goods price all these plots have outliers because of higher income range, high amount of credit, high annuity amount and high goods price.

Imbalance Percentage of Target Variable



This pie chart depicts number of non defaulter is greater than no of defaulter people.

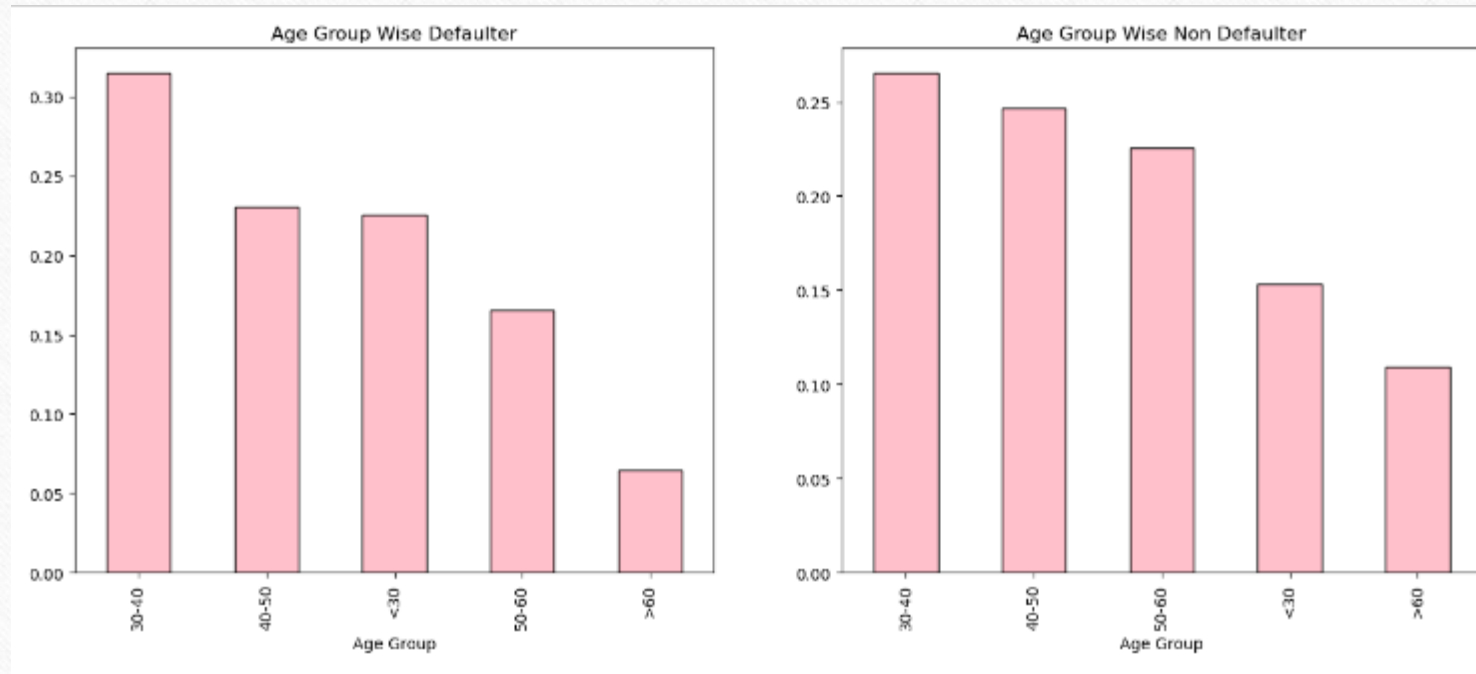
- a. 0 indicates Non Defaulter
- b. 1 indicates Defaulter



We can see that credit amount and amount goods price are highly correlated with one another. People with higher credit amount related to people to higher goods price.

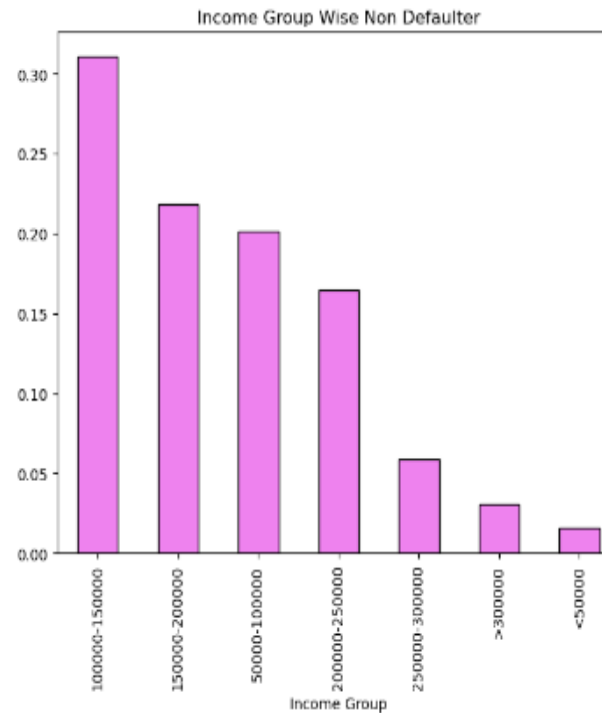
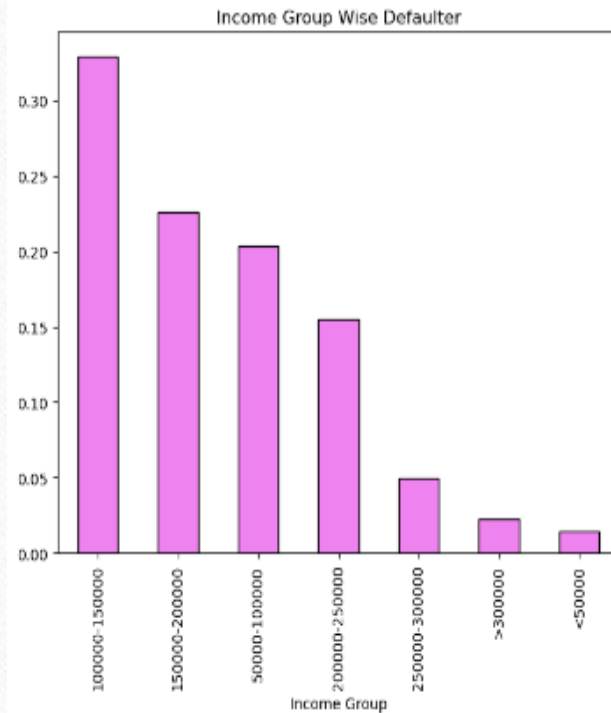
People with more income face less difficulty in paying credit amount .

Age Distribution Based on Target Variable



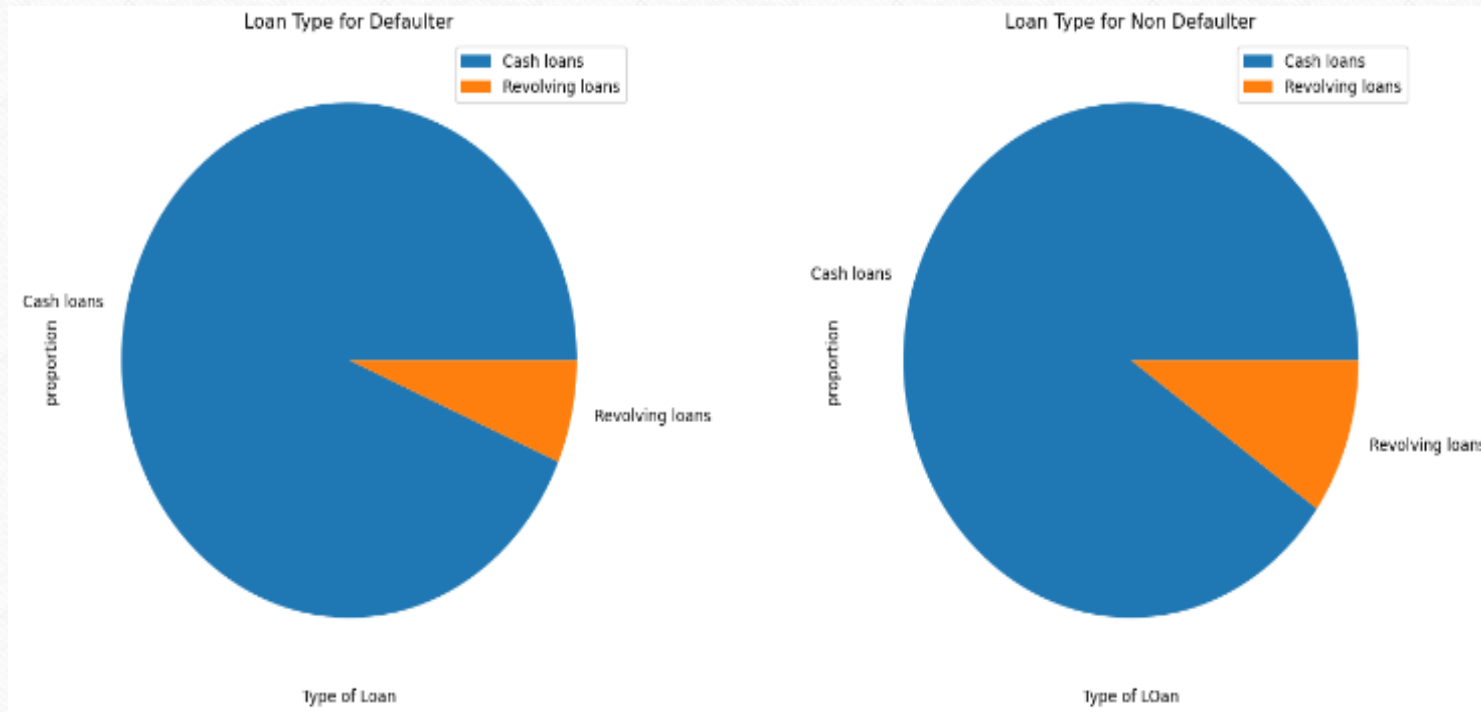
Age group of 30-40 people are applying for most of loans and age group <30 take less loan

Income Distribution Based on Target Variable



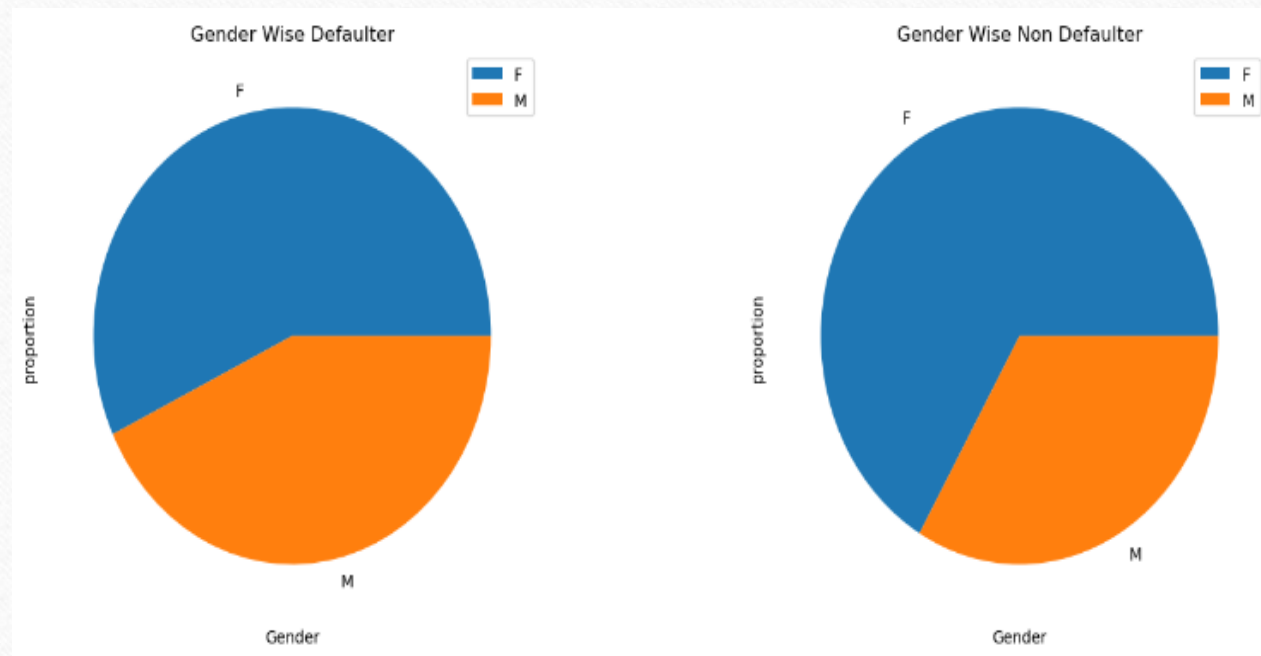
Income group of 100000-150000 people tend of most of the loan amount and >300000 and <50000 income group people apply for less loan

Loan Type Based on Target Variable



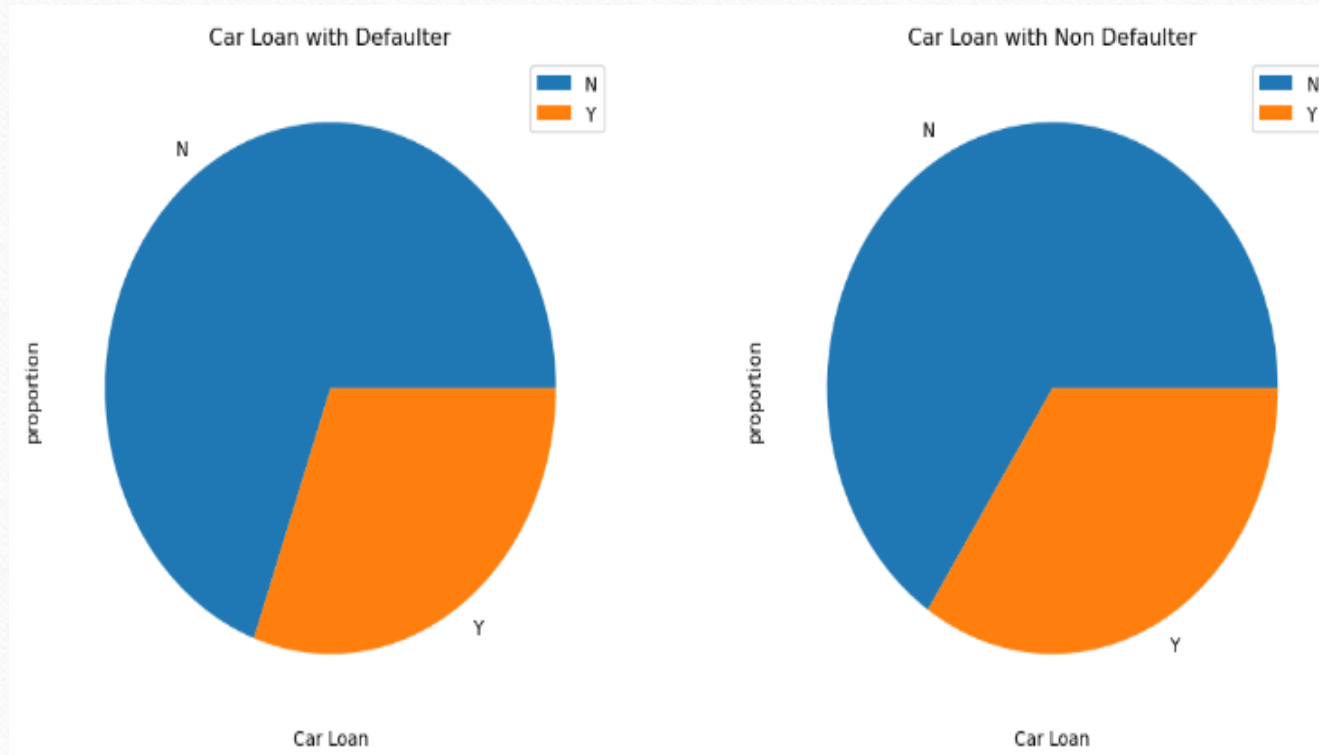
People are more likely to take cash loans compared to revolving loans

Gender Based on Target Variable



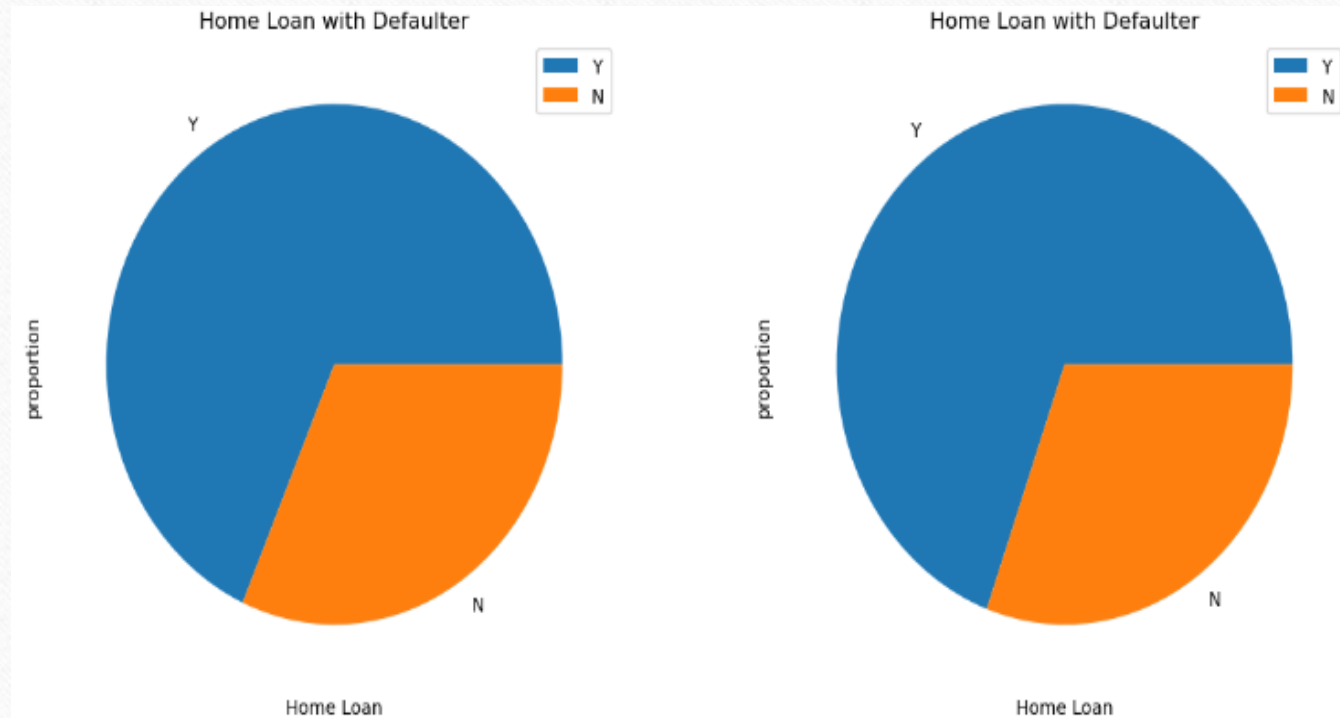
Female apply for more loans compared to male.

Car Loan Based on Target Variable



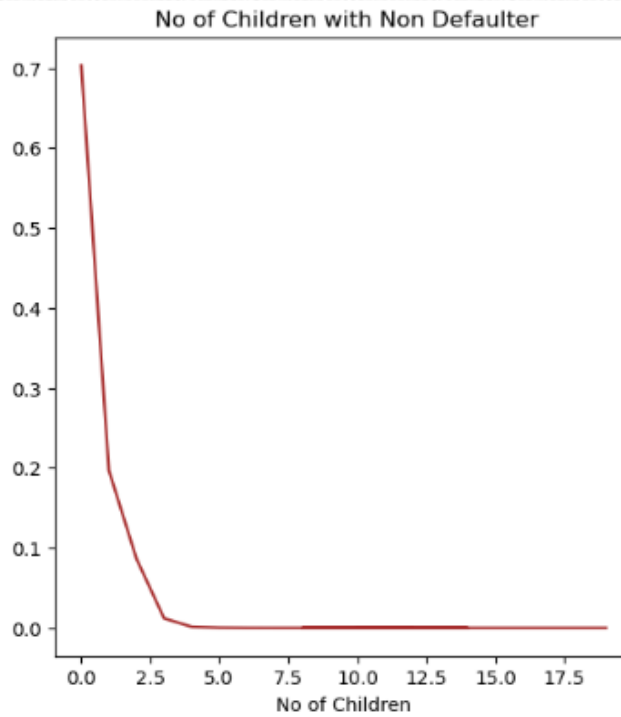
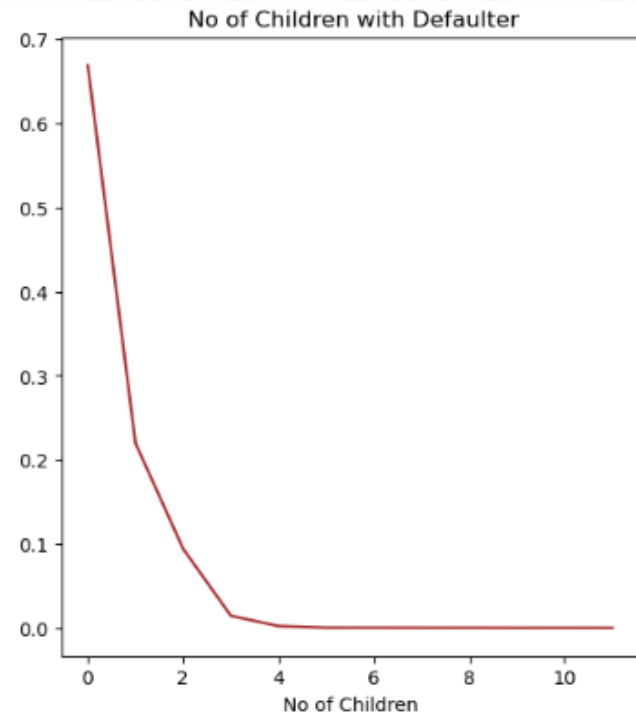
People like to apply for less loans for buying car.

Home Loan Based on Target Variable



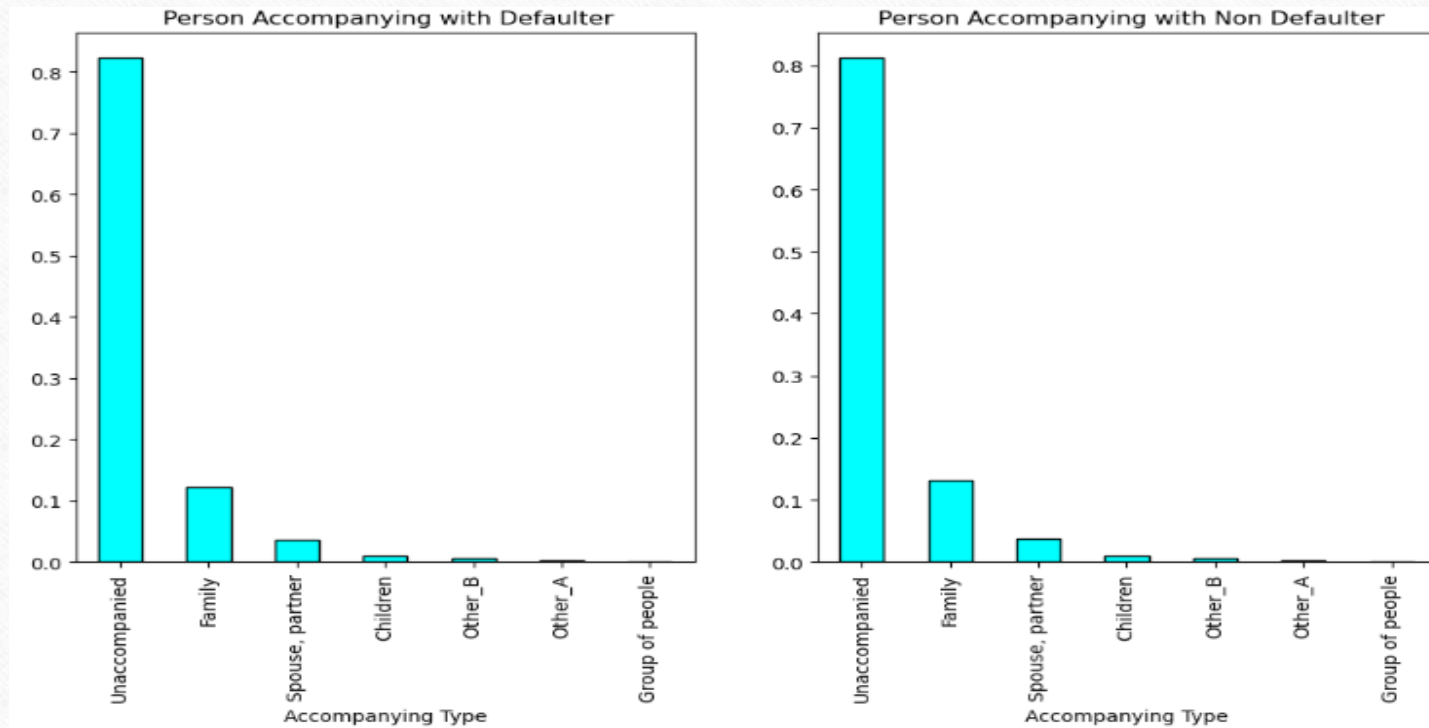
People are applying for home more home loan.

No of Children Based on Target Variable



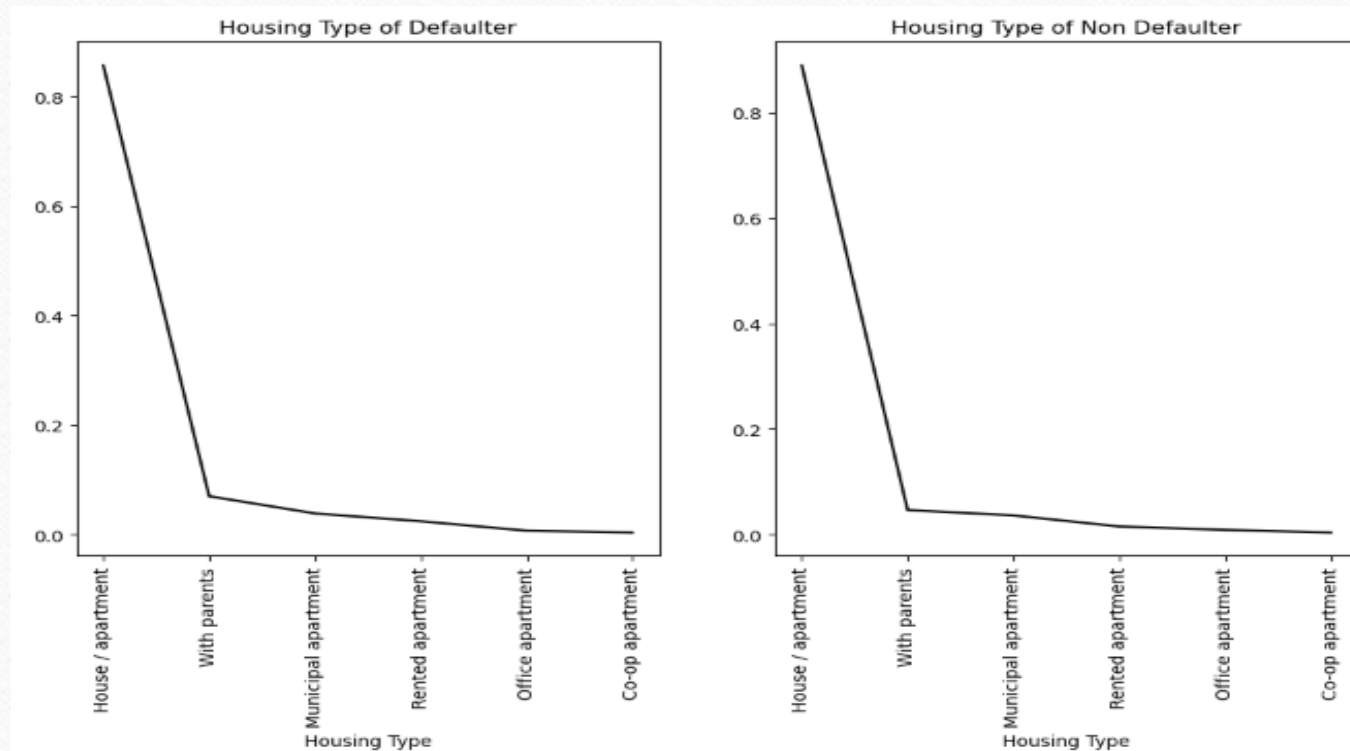
People with no child are applying for more loans compared to people with one or two children.

Person Accompanying Based on Target Variable



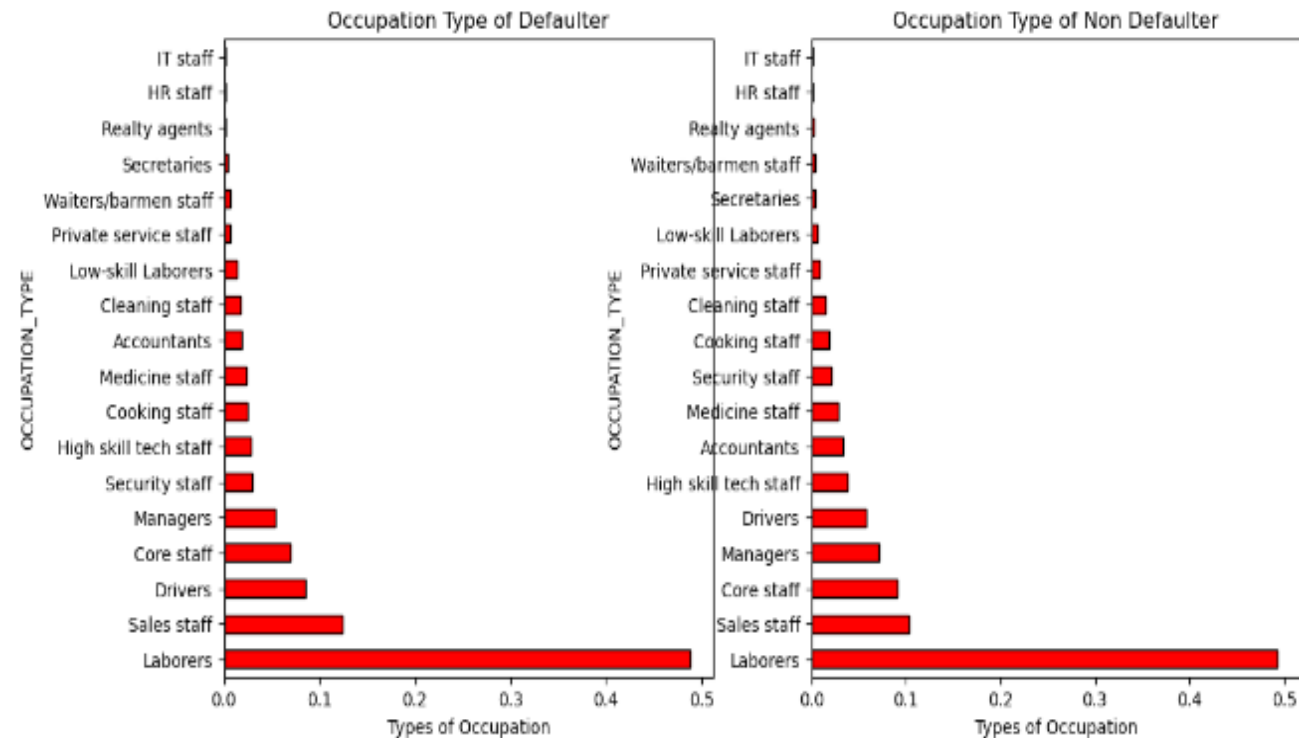
Unaccompanied people are applying for more loan compared to family person.

Housing Type Based on Target Variable



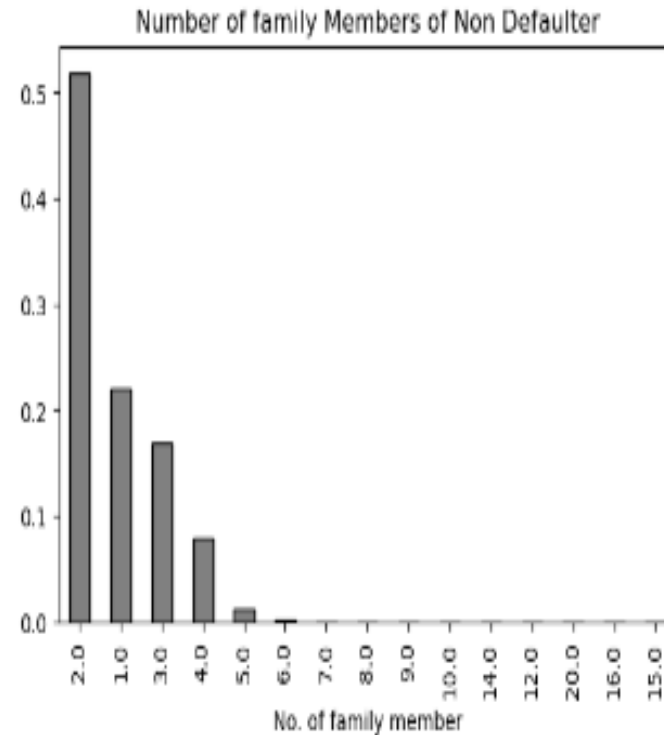
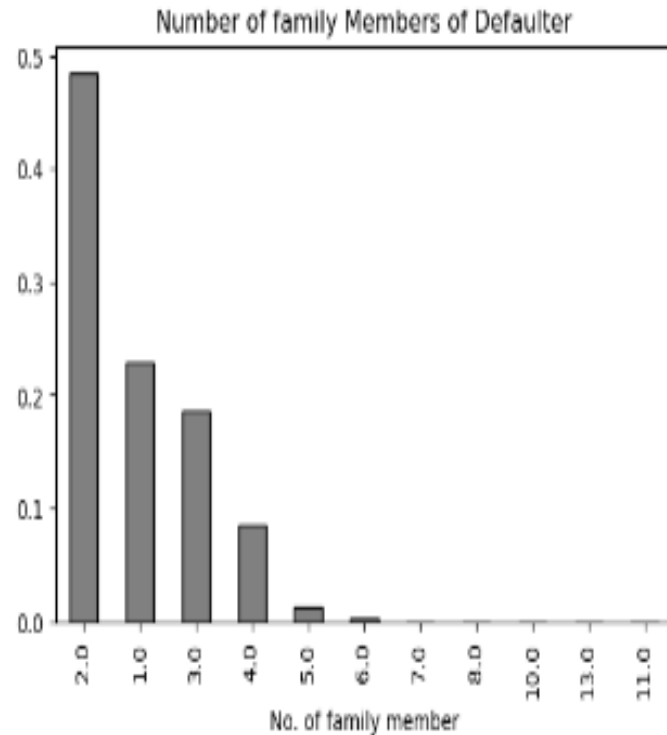
People are apply for loan for owning house/apartment compared to rented or govt subsidized house.

Occupation Type Based on Target Variable



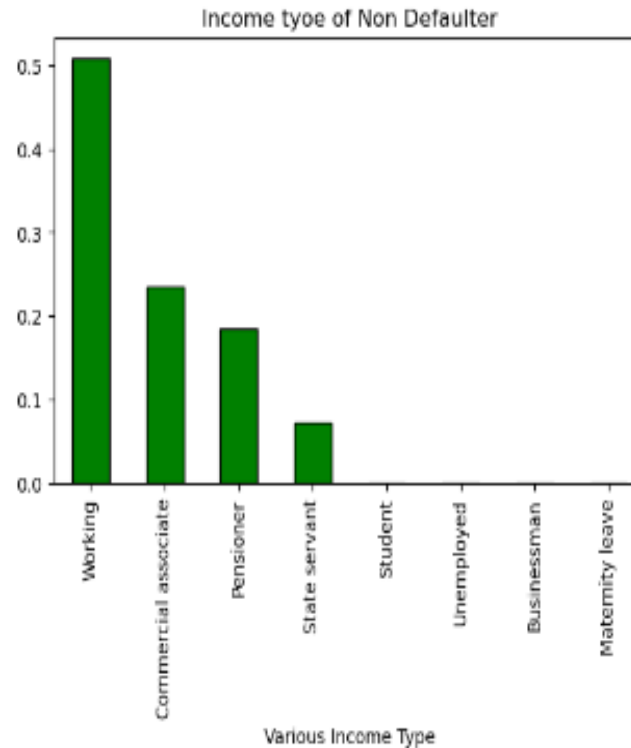
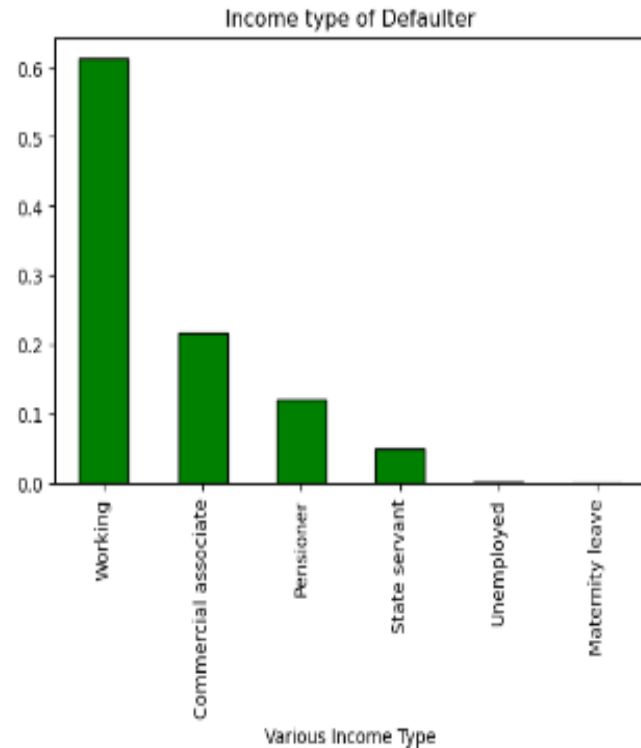
Laborer are taking most of the loan and HT staff and IT staff are applying very low loan amount.

No of Family Member Based on Target Variable



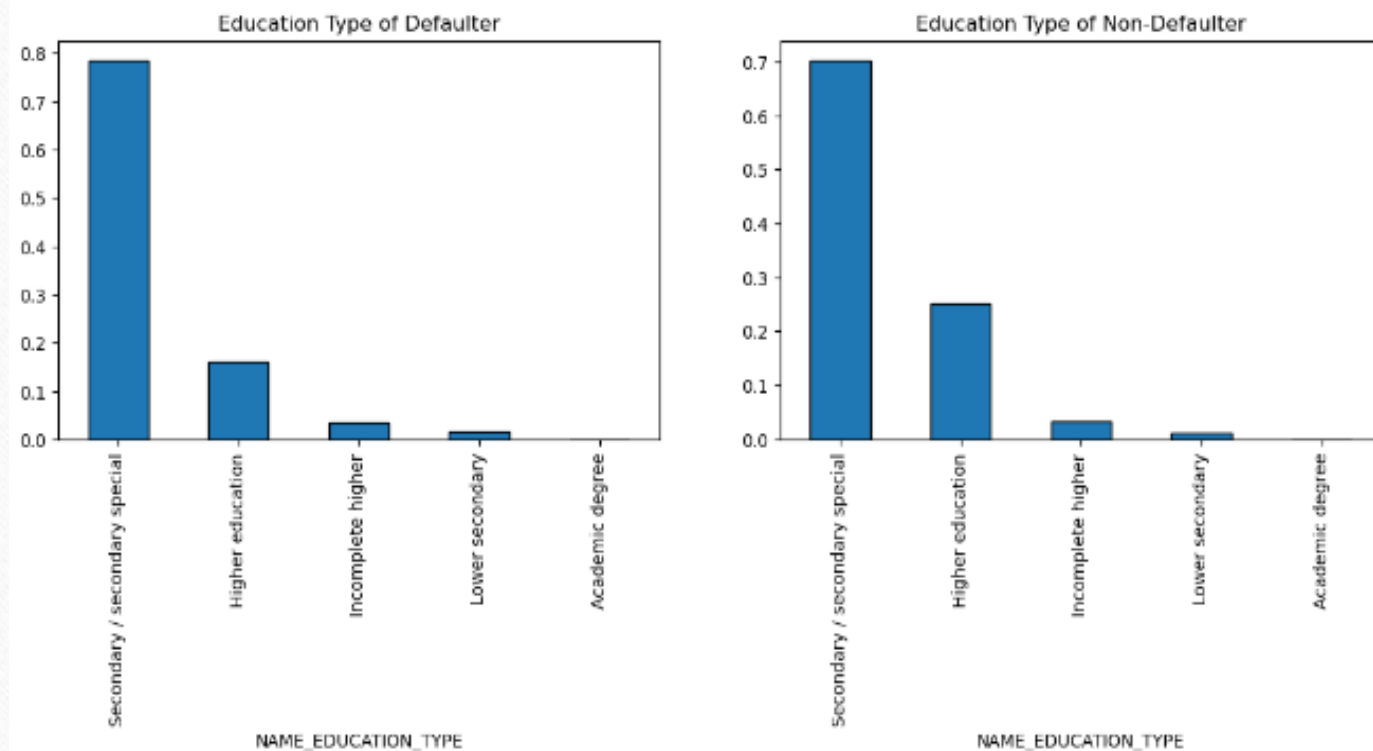
People with family size 2 are applying for more loan compared to people with big family like 10 or 11 family member.

Income Type Based on Target Variable



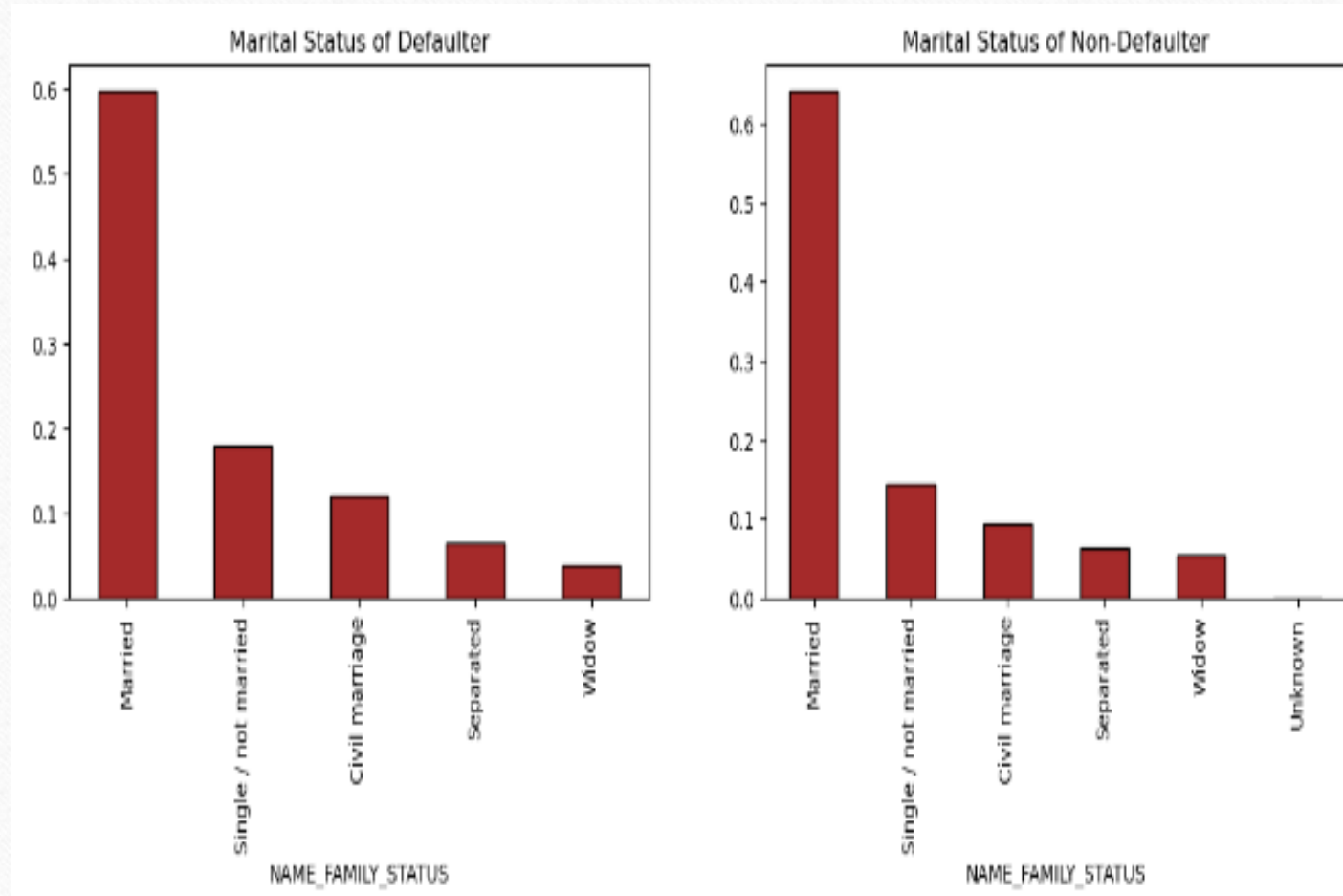
Working class people like to apply for loans compared to the people who are student, unemployed or under maternity leave.

Education Type Based on Target Variable

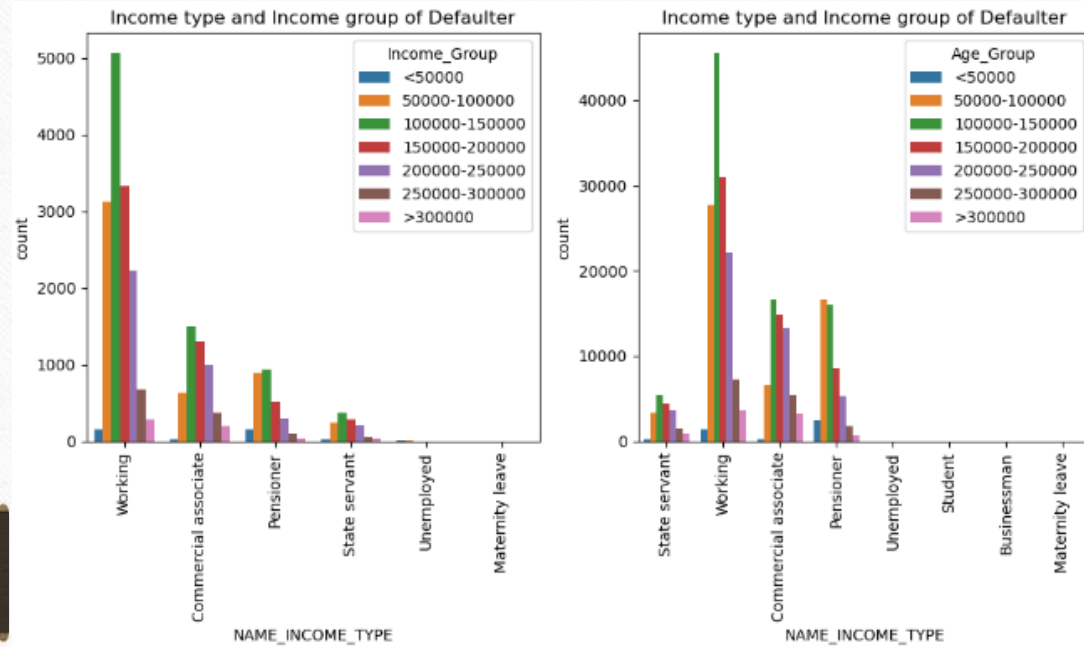


Secondary educated people apply for loan compared to academic educated people.

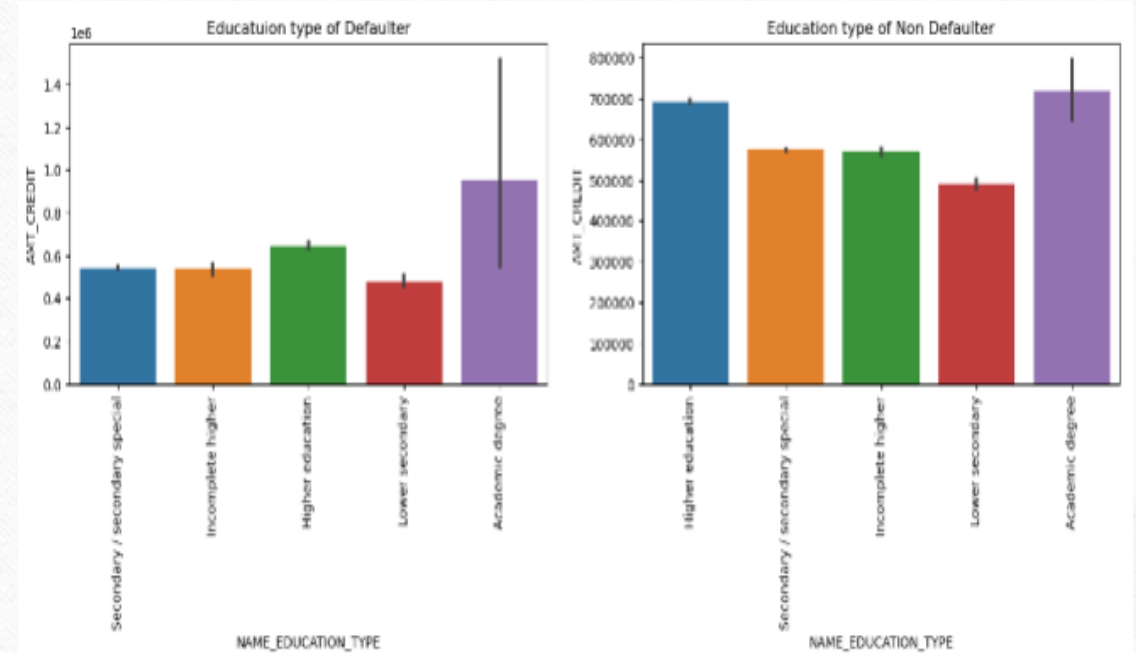
Marital Status Based on Target Variable



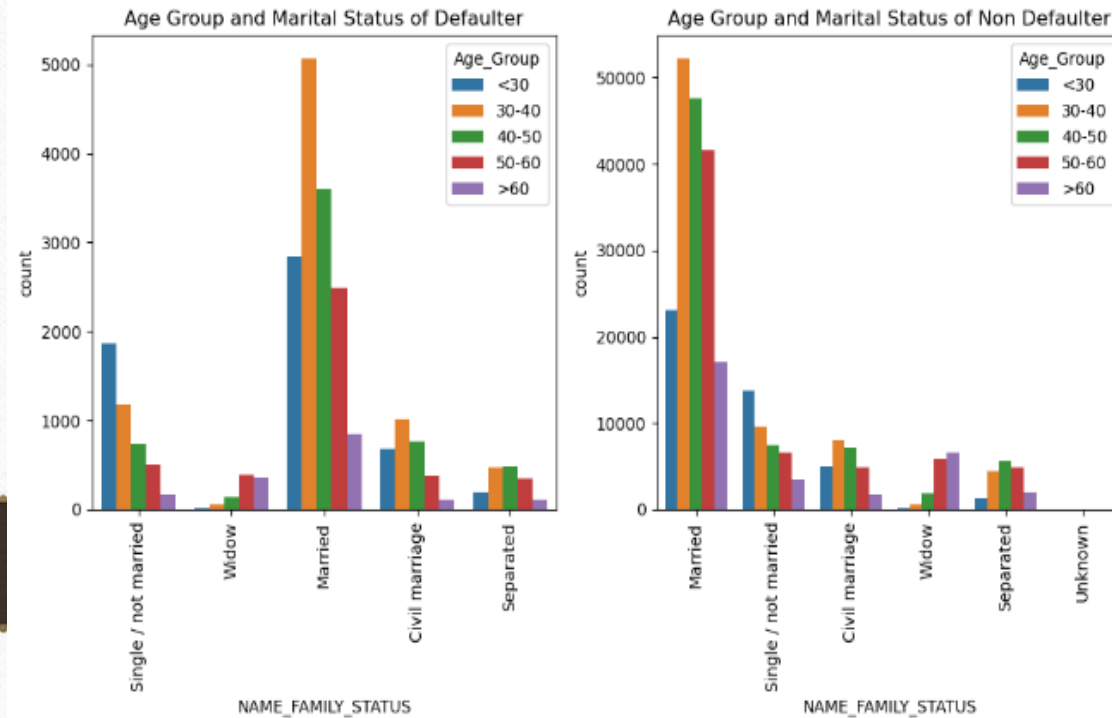
Married people apply for more loans compared to separated or widow.



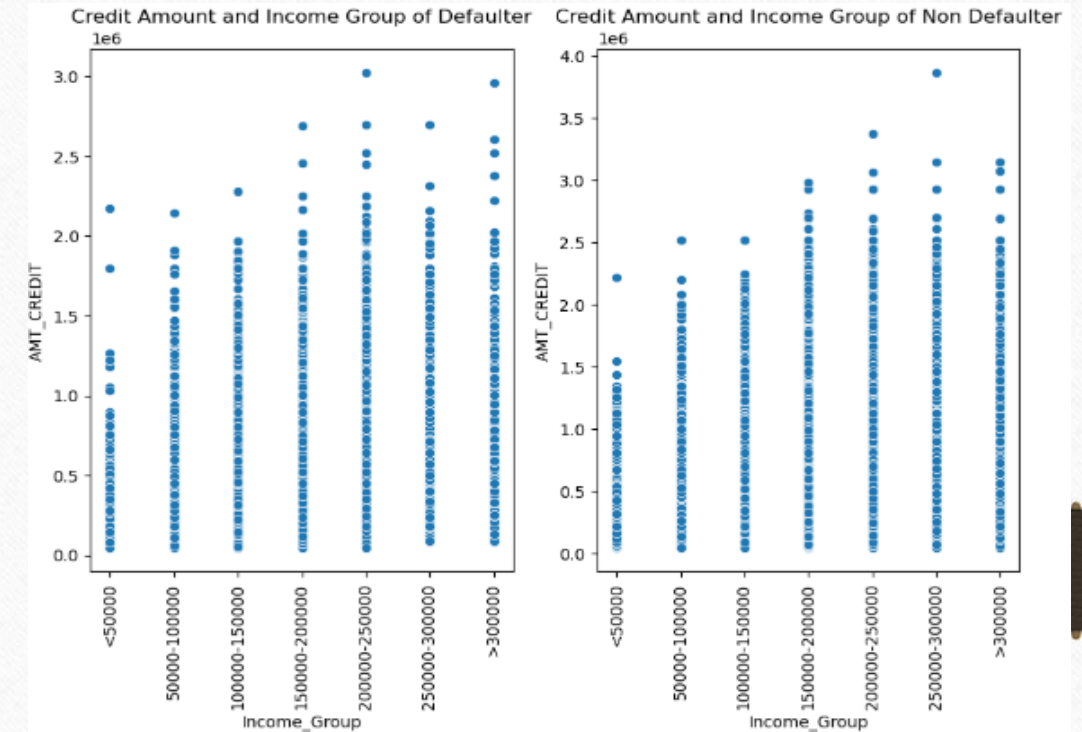
Working class people who are around the income group of 100000-150000 are eligible for most of the loan amount on the other side, student, unemployed, maternity leave persons they are not coming under any income category and not applying for any loan amount.



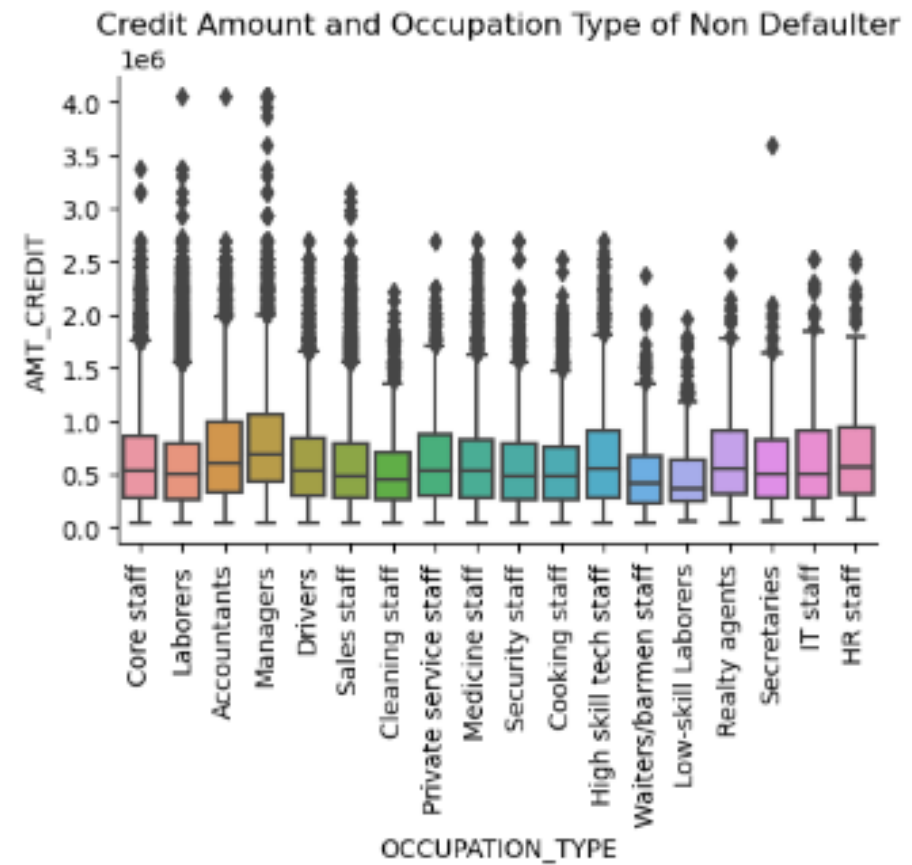
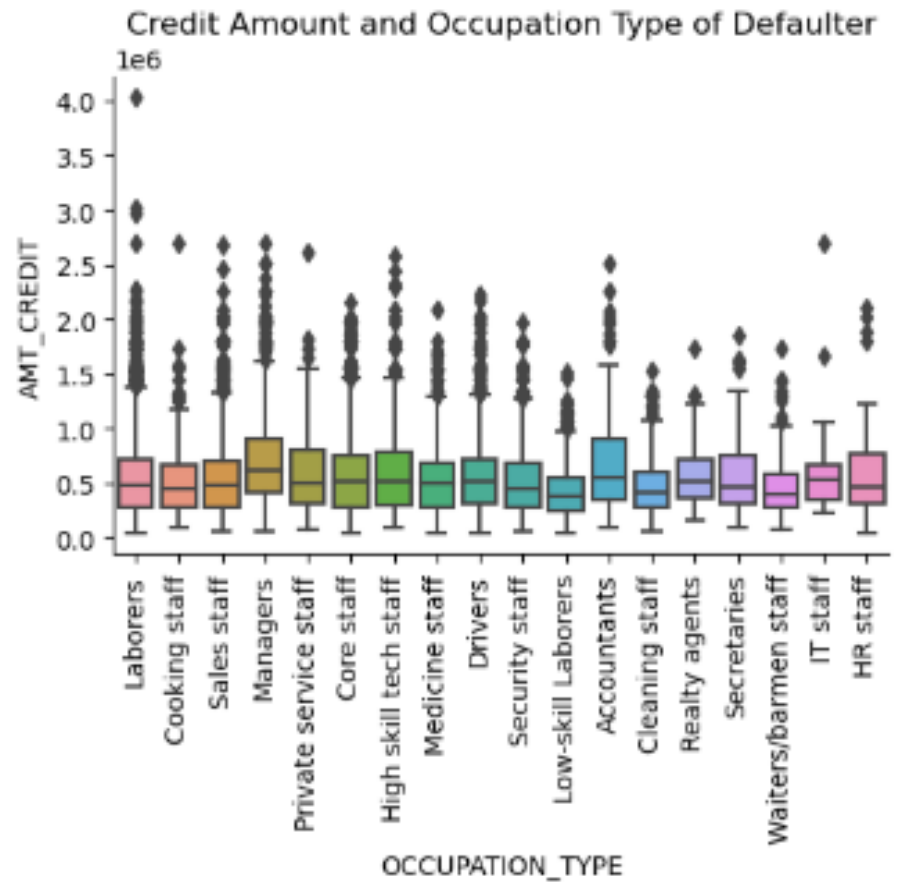
Academic degree people are getting more loan compared to lower secondary or other education type people. As they are highly educated people so, they will repay the loan amount to bank on time which is profitable for bank.



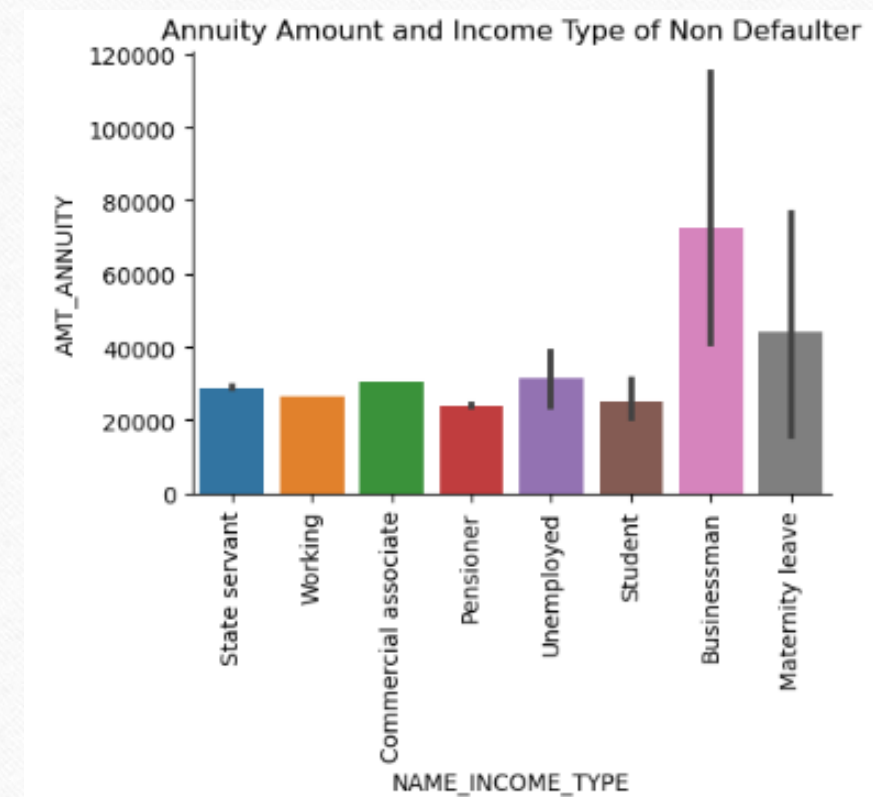
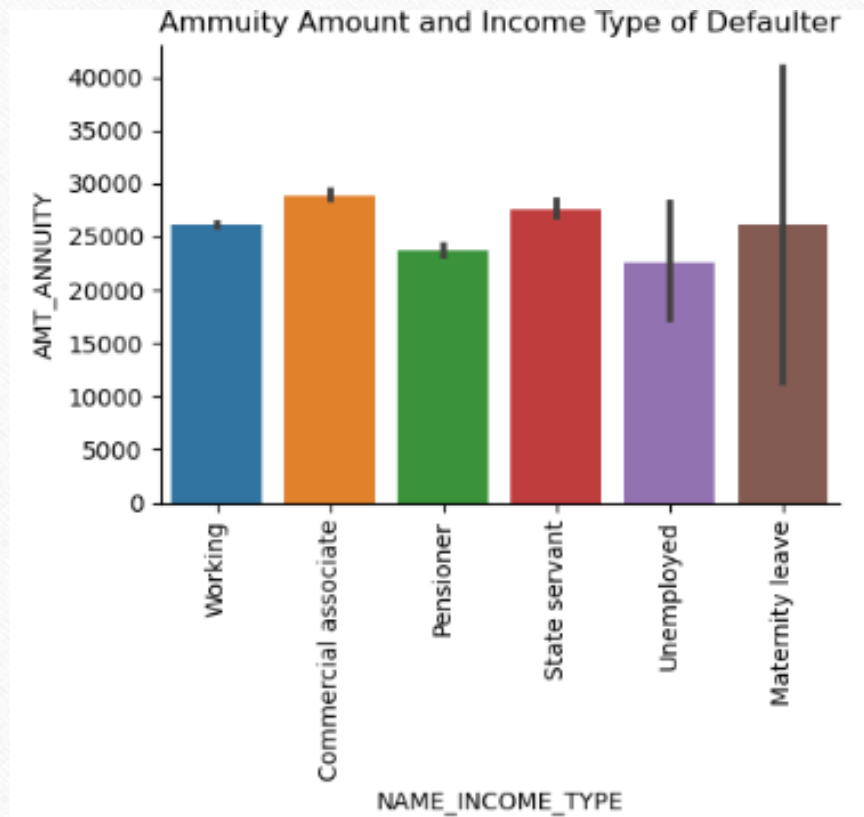
Age group of 30-40 who are married, take most of the loan amount and unmarried people who are less than 30 take most of the loan compared to separated, widow people.



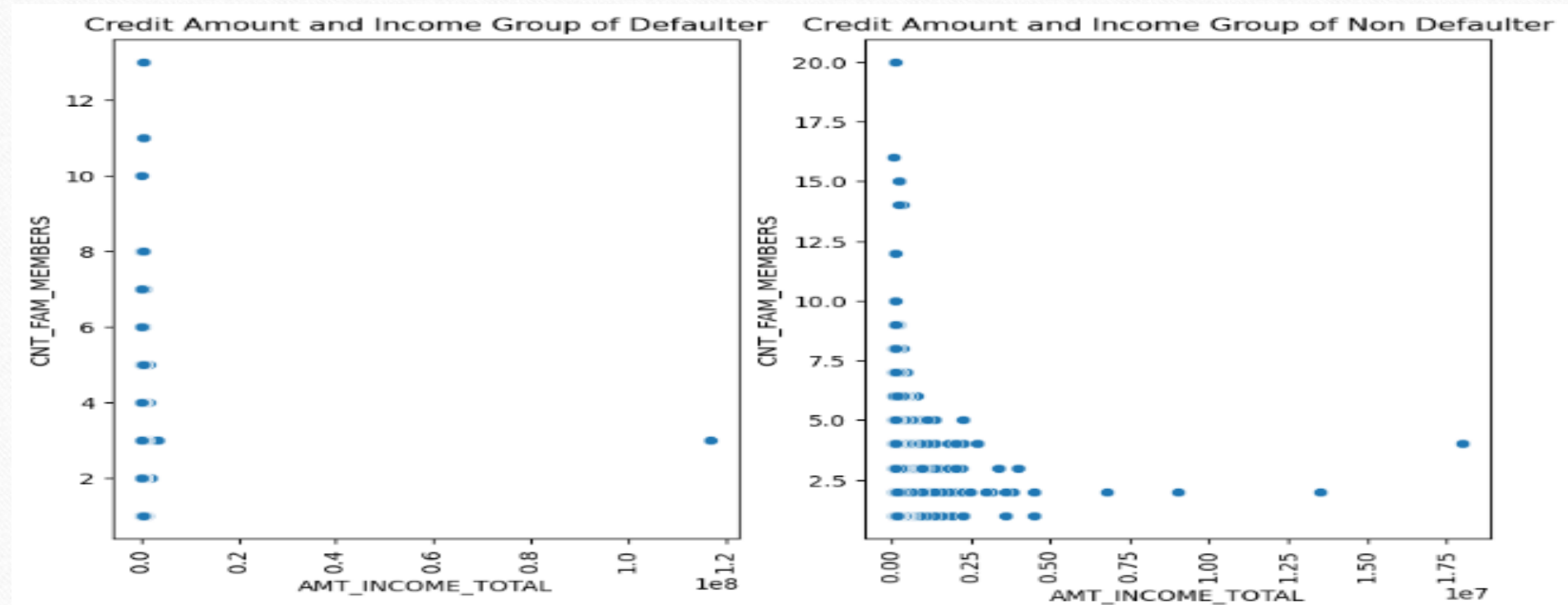
Higher income group people take more loan and lower income group people take less loan as their income range in very high so its expected that they will repay loan on time.



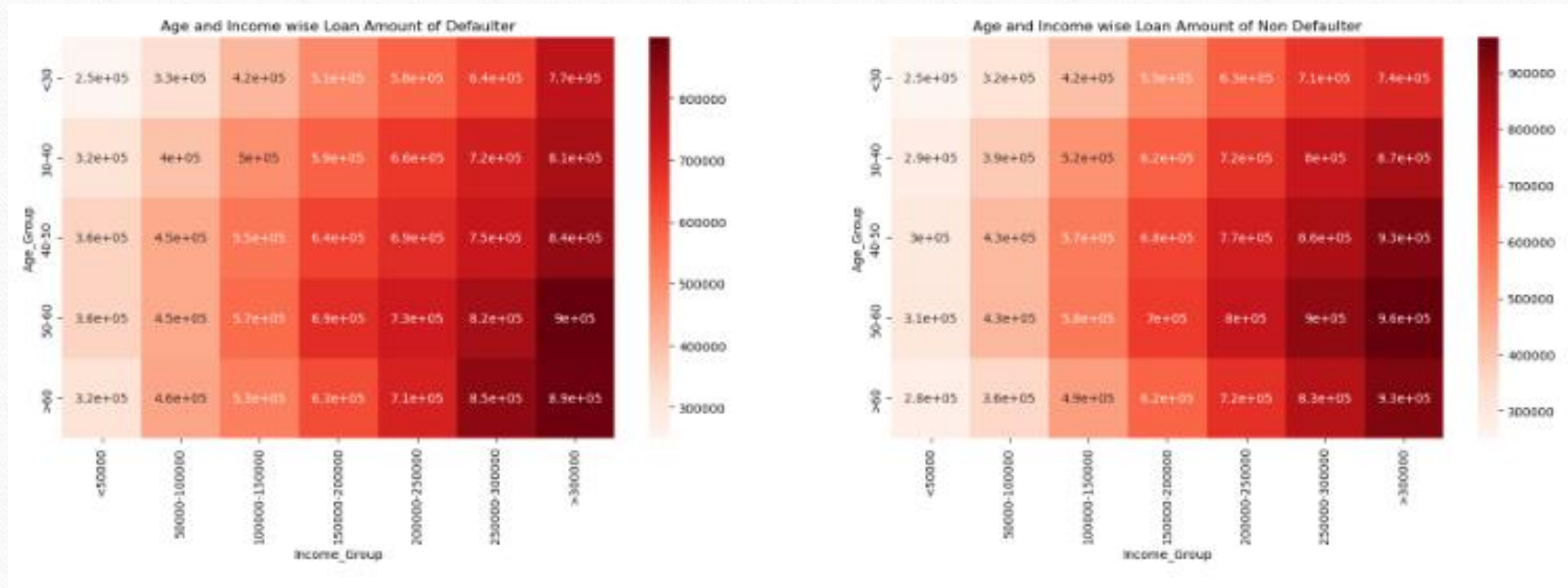
Manager, Accountants, HR staff, IT staff they are getting loan compared to low skilled labor, cleaning staff and labor class people due to their income level and they are delaying to repay their loan.



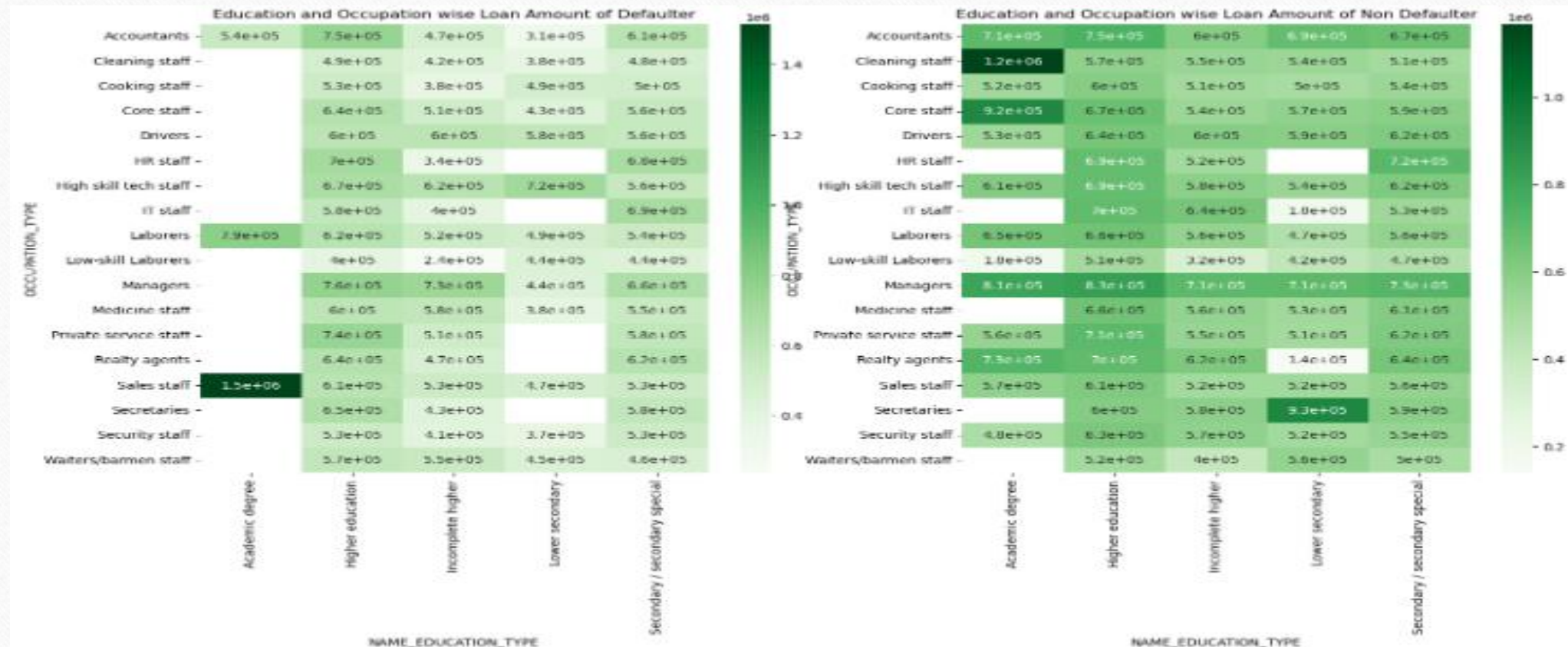
From this plot as per income type, businessman are mostly non defaulter people and they paying annuity amount on time. From defaulter perspective, most of the working professional are defaulters except pensioner and unemployed people.



People with less family member and higher income will face less payment difficulty compared to people with more family member.

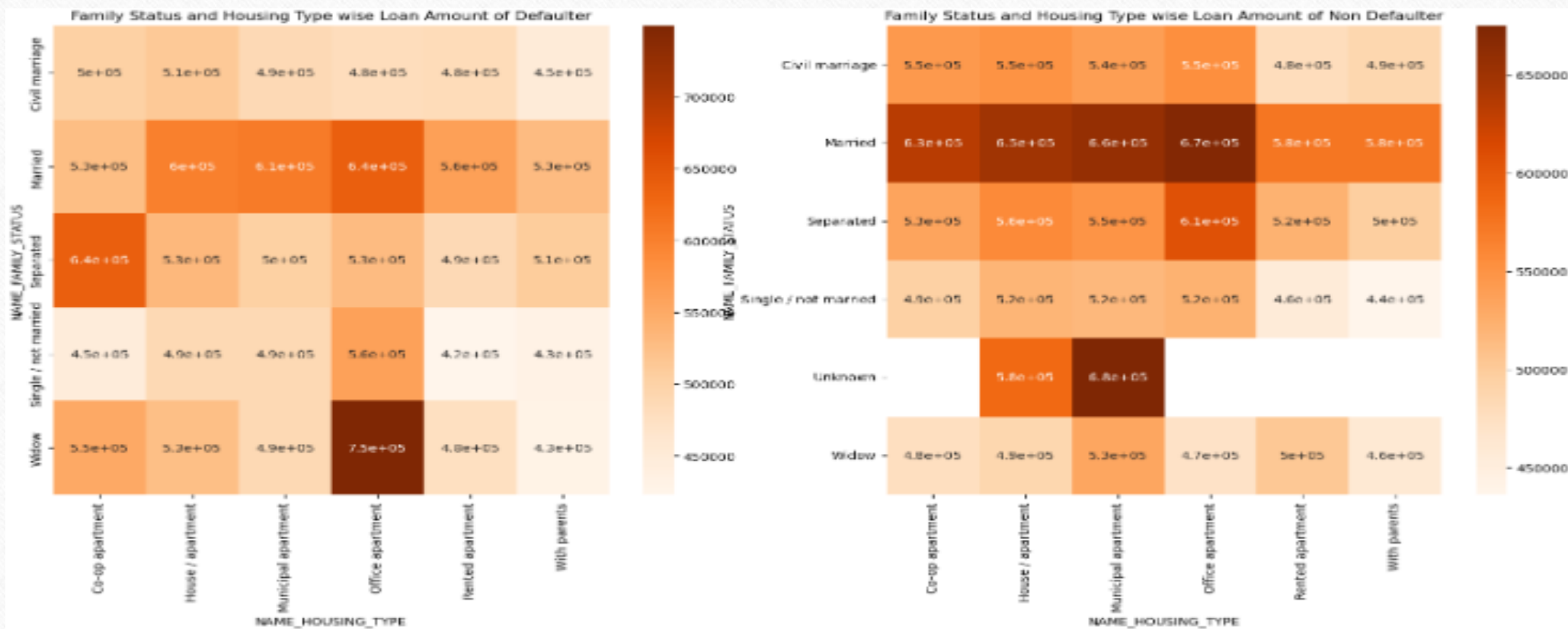


People around the age group of 40-50, 50-60 and 60 above are the highest earning group so they will get more loan amount and will face less difficulty in repaying the loan compare to the people who are around the age of below 30 and 30-40 with lower income level.



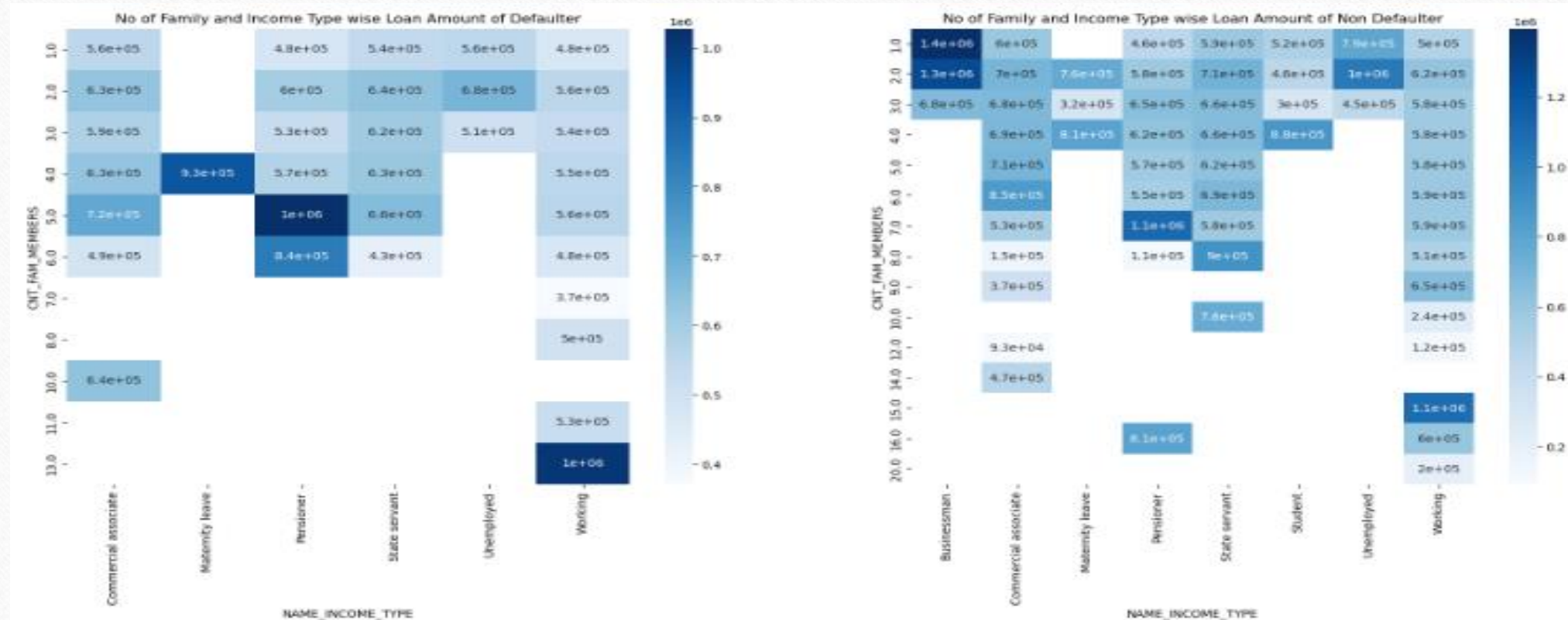
People who have academic degree and sales staff they are facing challenges in repaying the credit loan amount.

Cleaning Staff with academic degree are paying loan on timely basis.



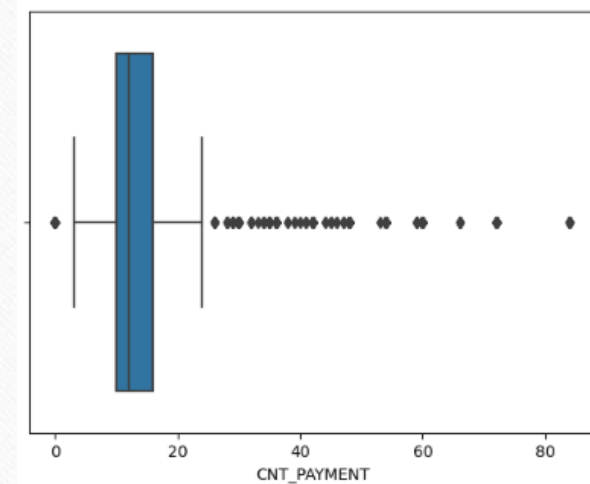
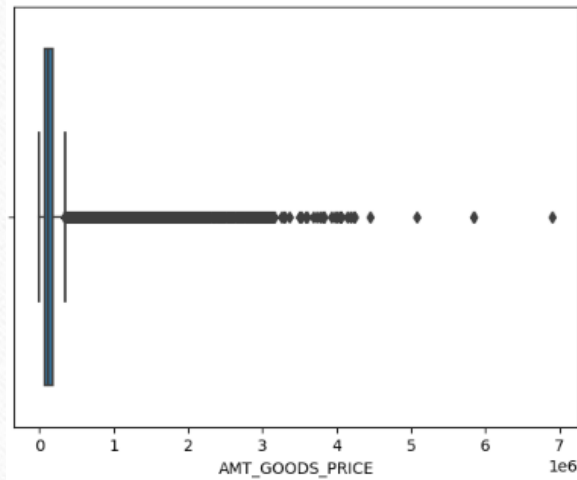
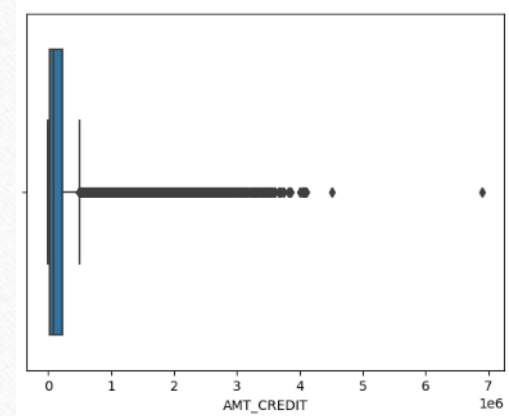
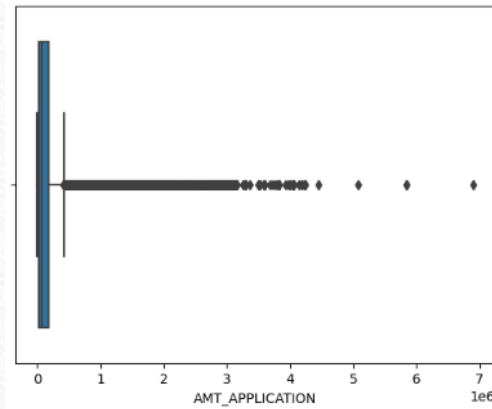
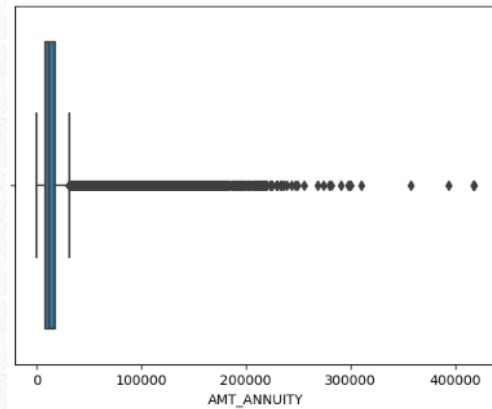
Widow people are more likely to go for office apartment because they are single earning member and they get more difficulty in repaying the loan.

Married people take loan for co-op apartment, house/apartment, municipal apartment and office apartment due to their higher income source and they can repay on time.



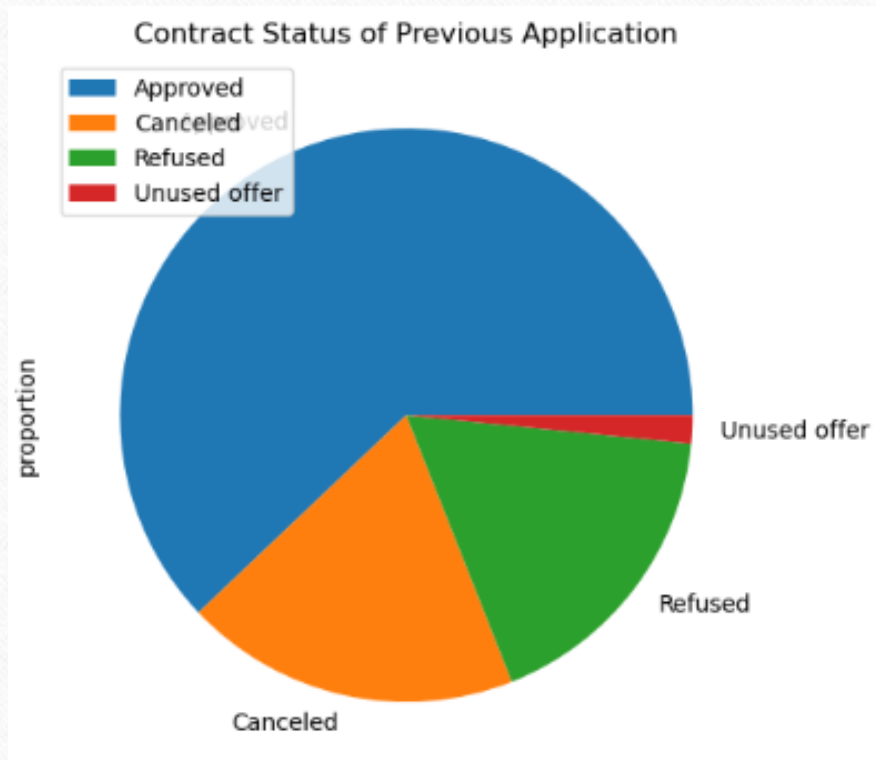
Maternity leave , pensioner take more loans with 4 or 5 family member, working people take loan with 12 family member, all of them face payment difficulty in repaying their loan.
 Businessman with 1 or 2 family members take more loans and repay on time due to less expense for family and more income source.

Analysis of Previous Application Dataset

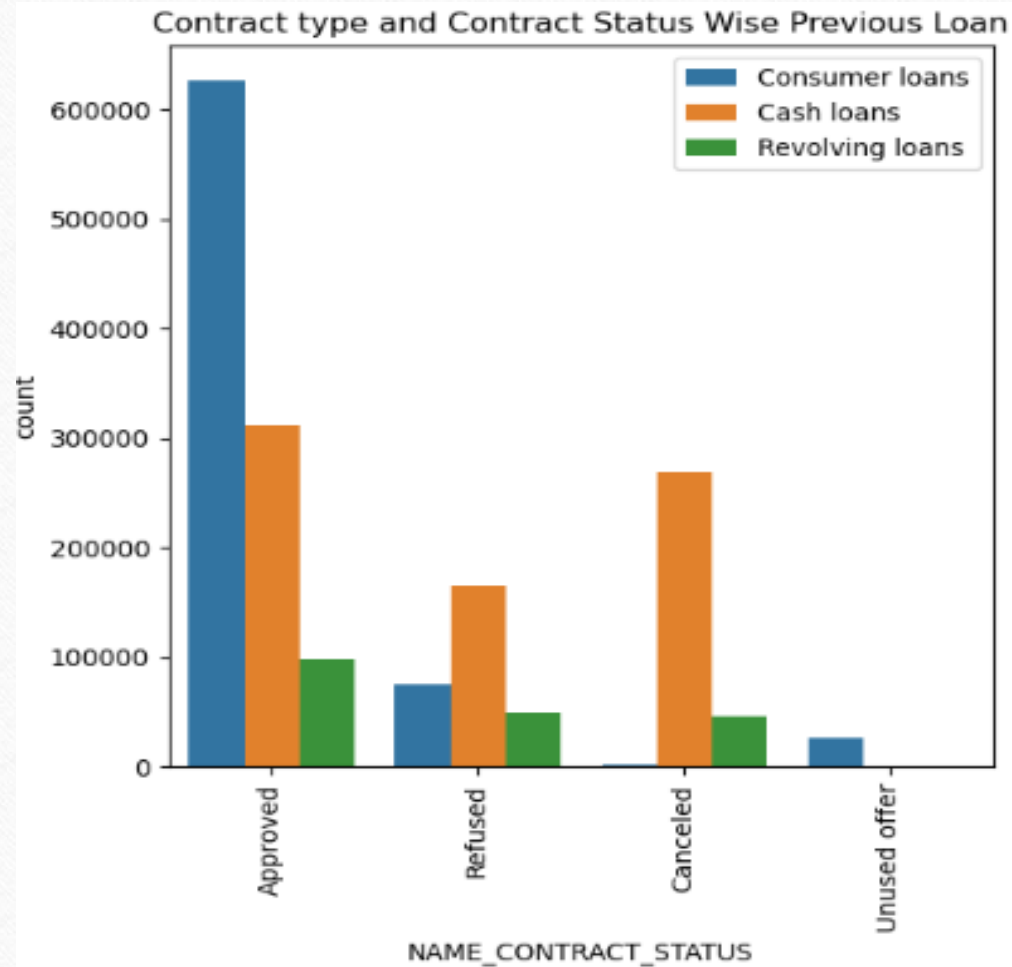


No of application, credit amount, goods price, annuity amount, cnt payment they all have outlier as per previous data because they have higher range of income, higher credit amount, annuity amount.

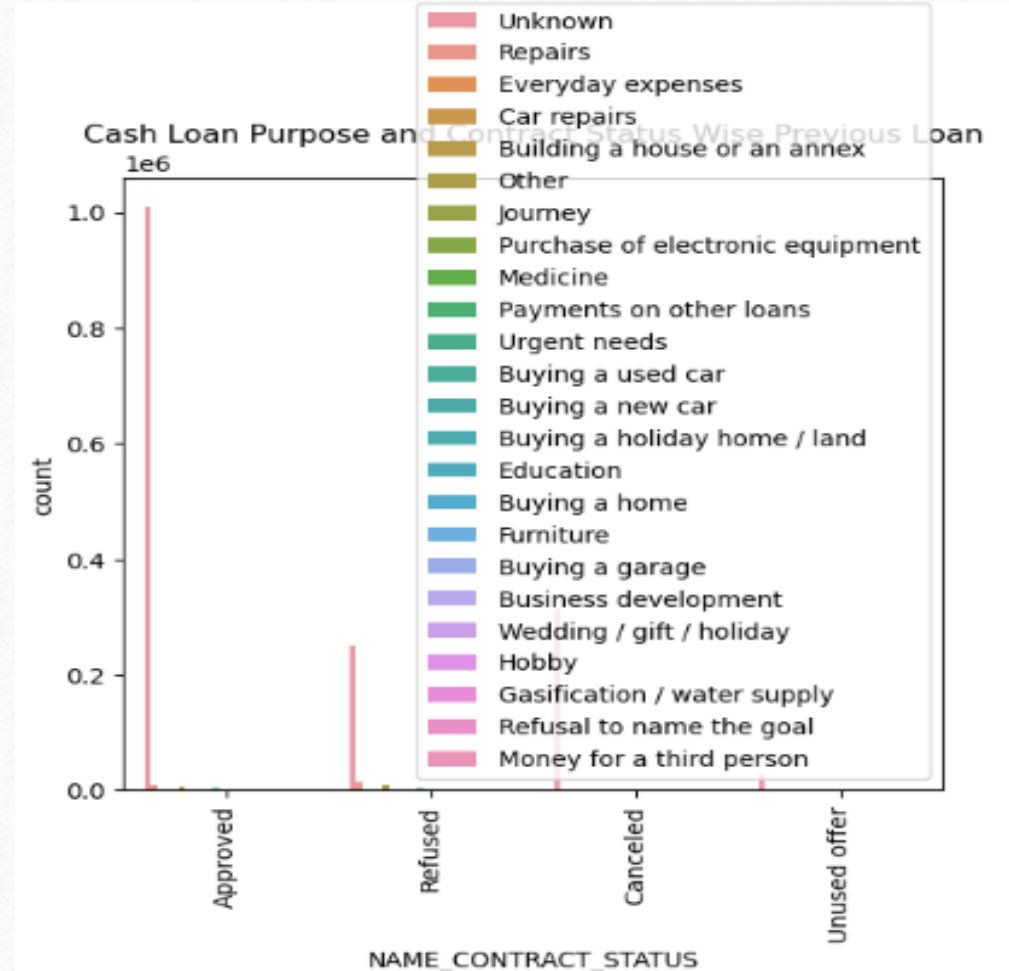
Contract Status Based on Previous Application



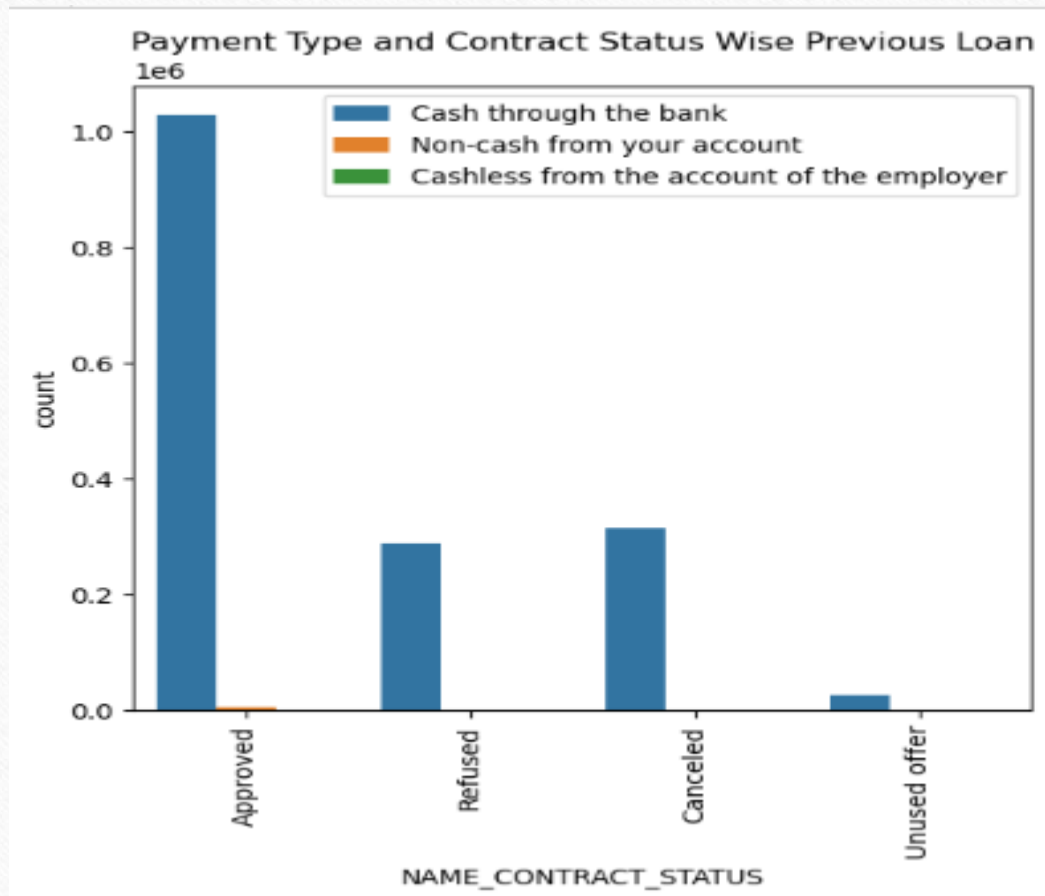
As per previous application, most of the loan got approved.



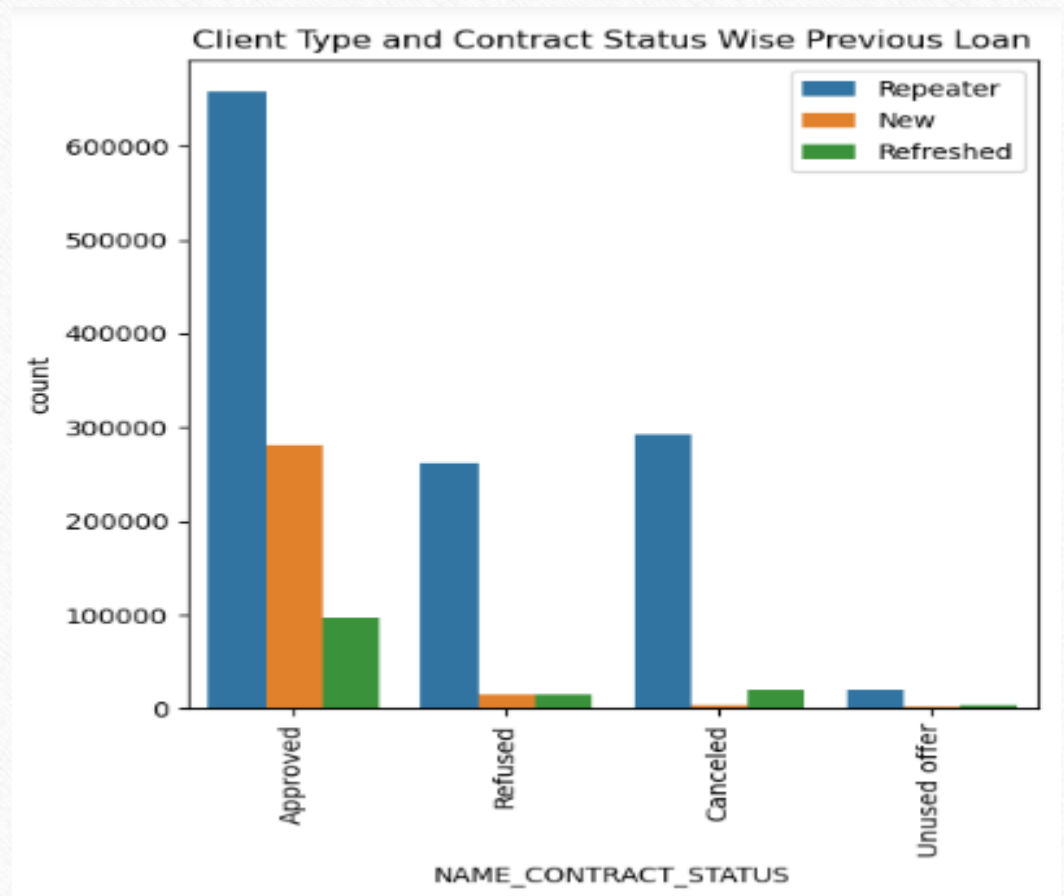
Majority of consumer loans were approved and more of cash loans were cancelled.



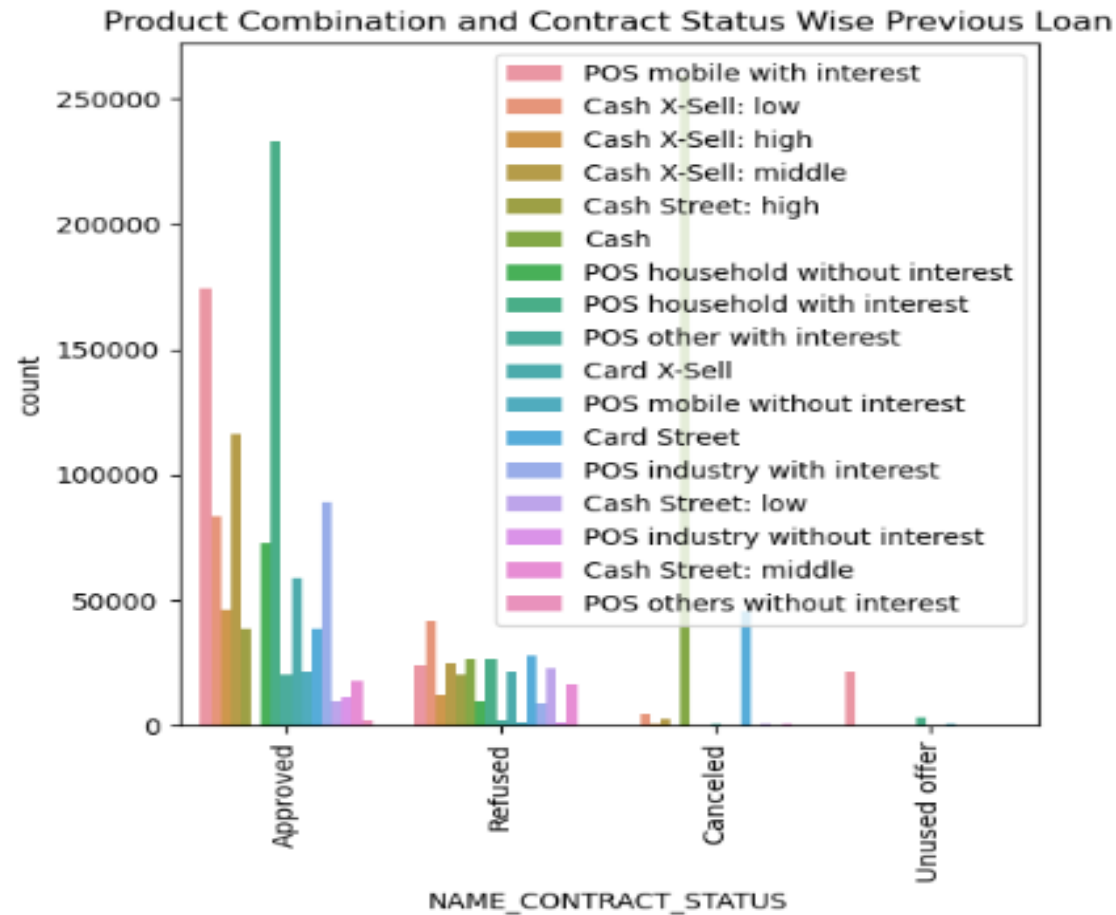
Cash loan is approved but reason of loan is unknown for most of the people



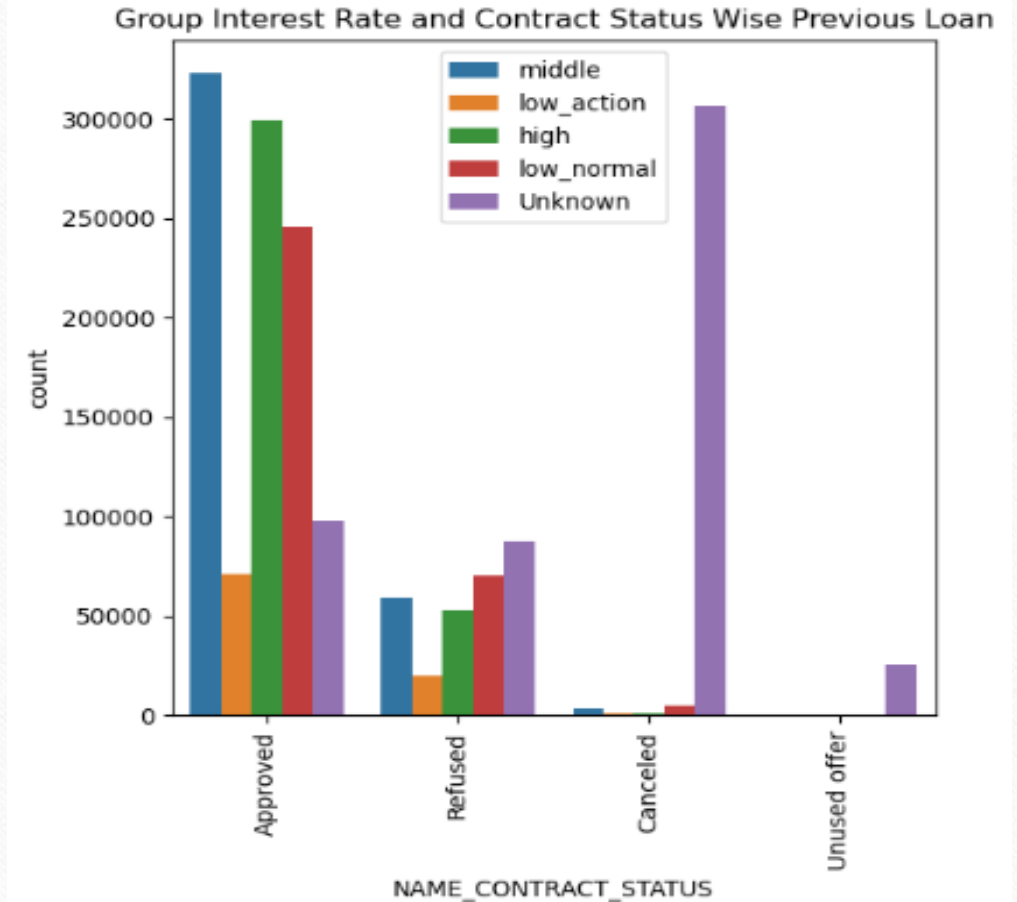
For most of the approved loan, people like to pay cash through the bank.



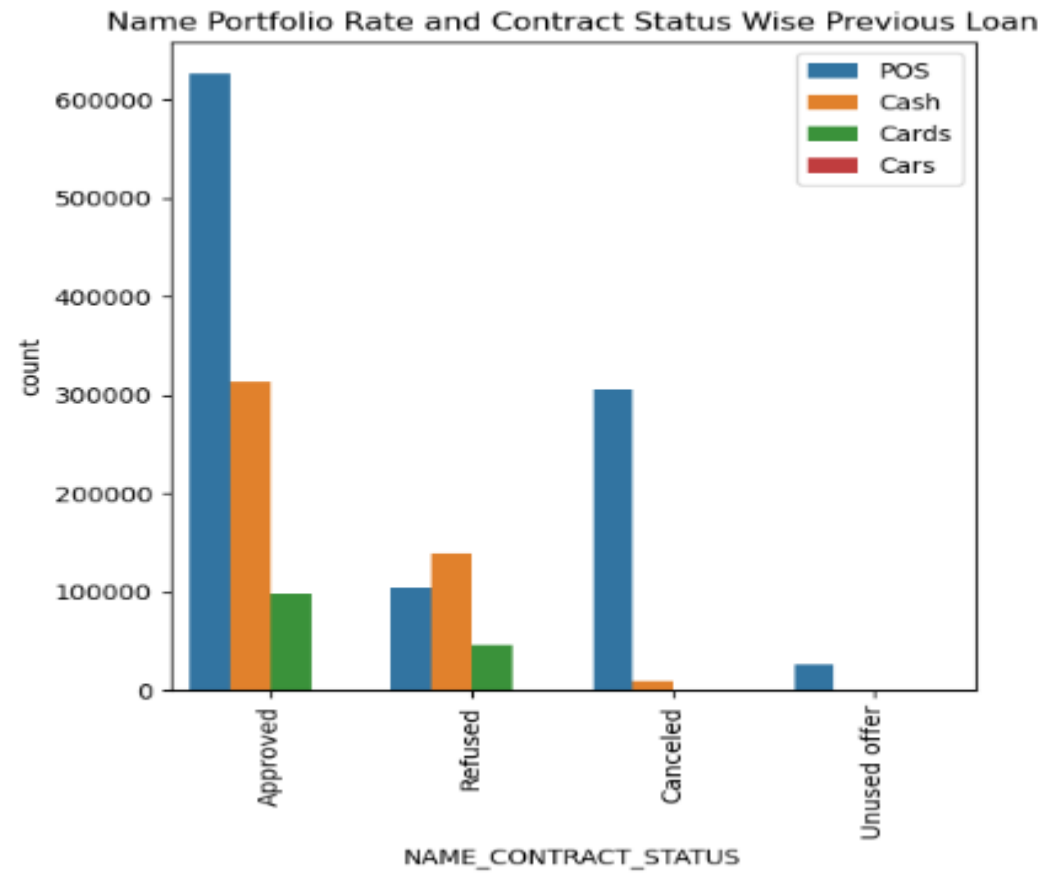
Most of the people are repeater for taking loan.



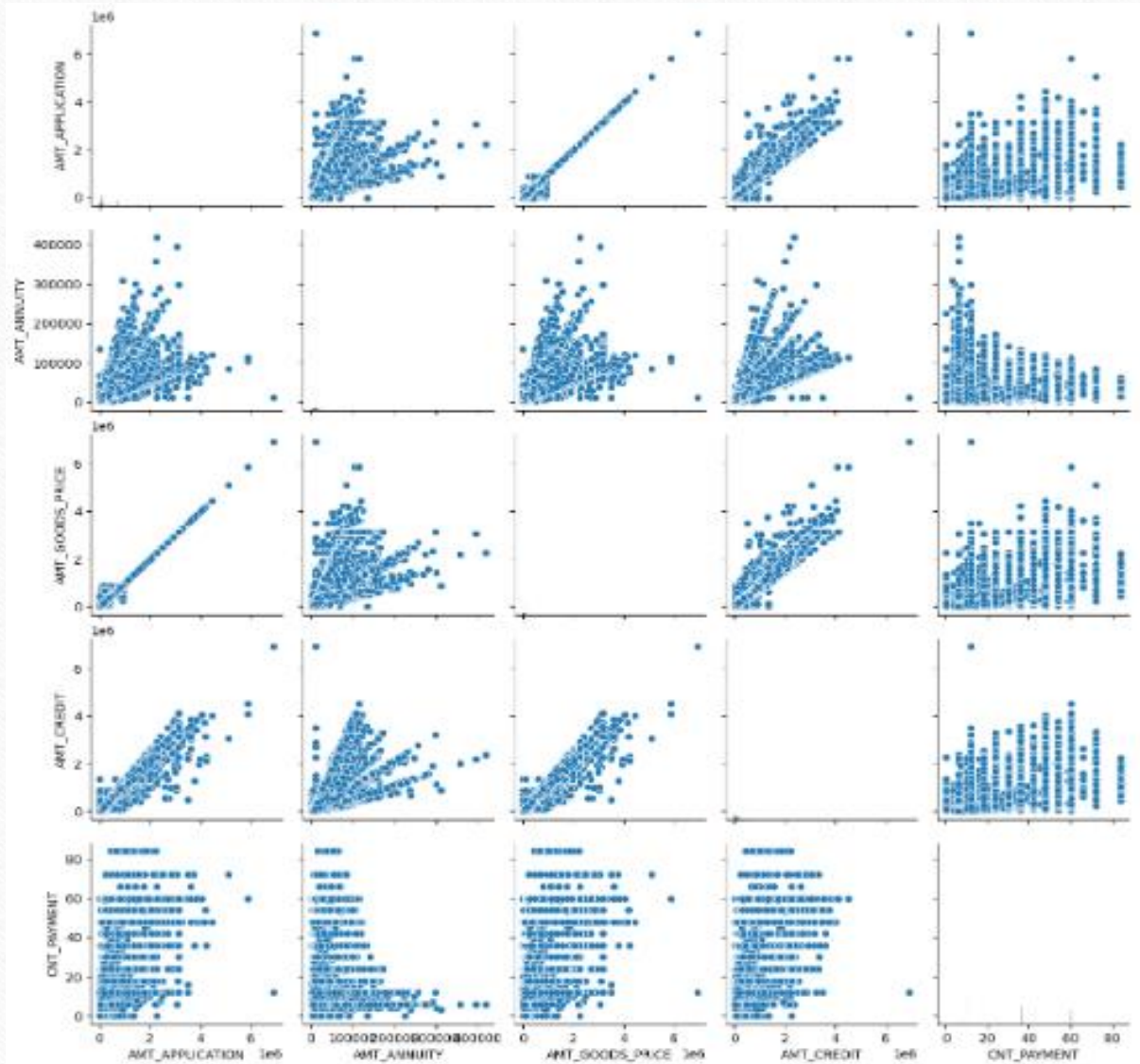
Most of the cash loan got cancelled and approved loans are mostly for POS household with interest.



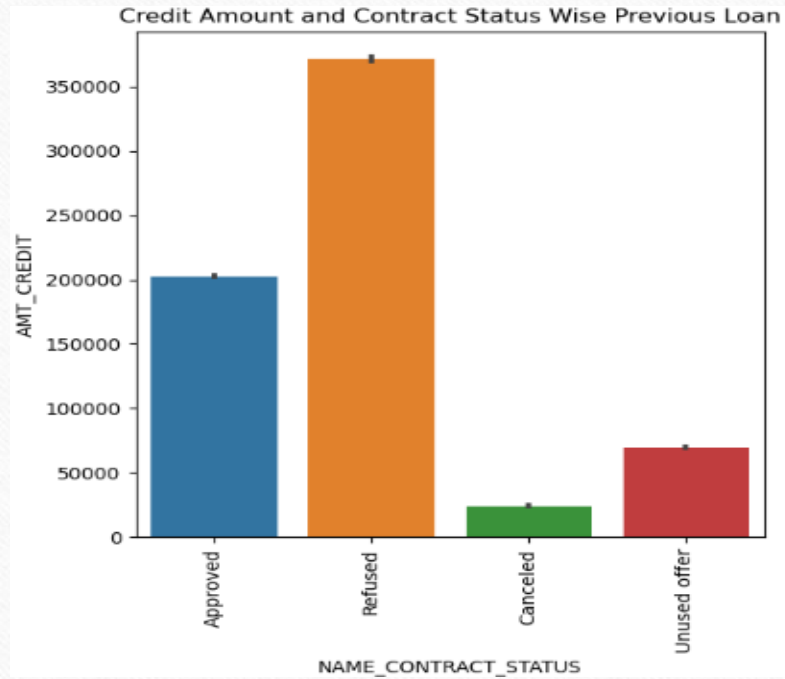
Most of the approved loans are for middle and high interest rate and cancelled loan is for unknown.



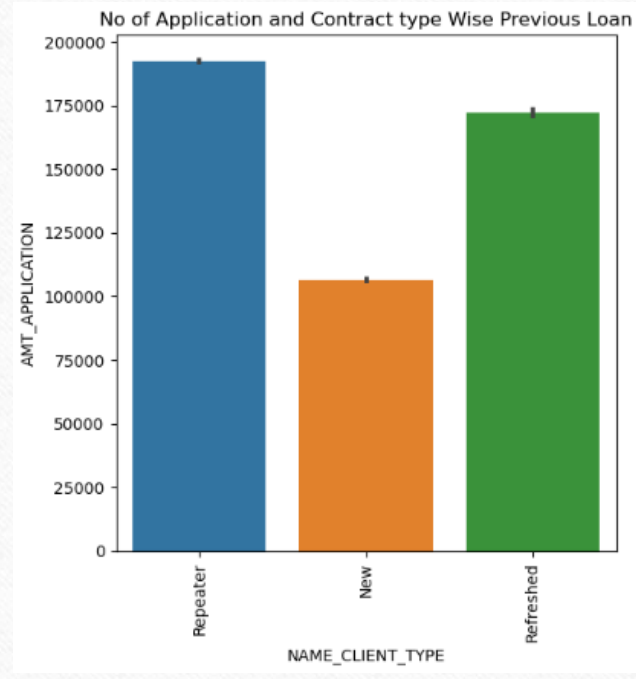
Most of the approved loans is POS



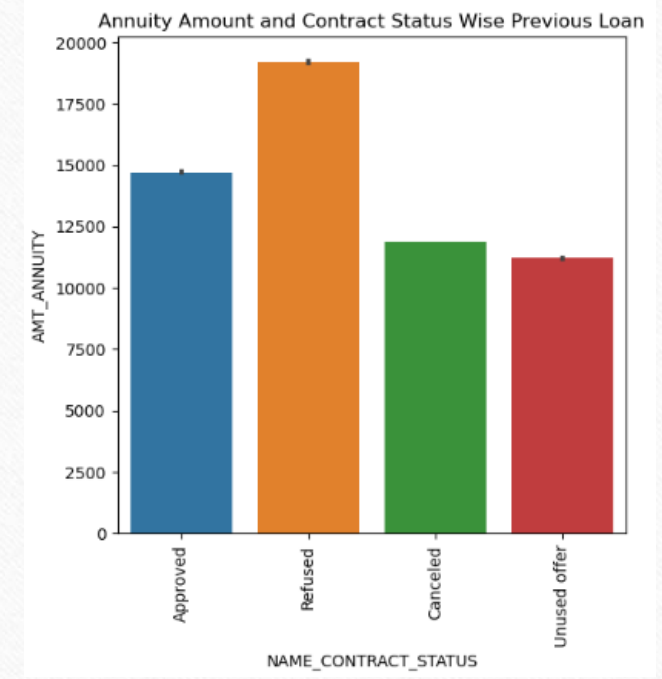
As per previous data, people with credit amount, people with annuity amount, goods price, no of application are positively correlated with each other.



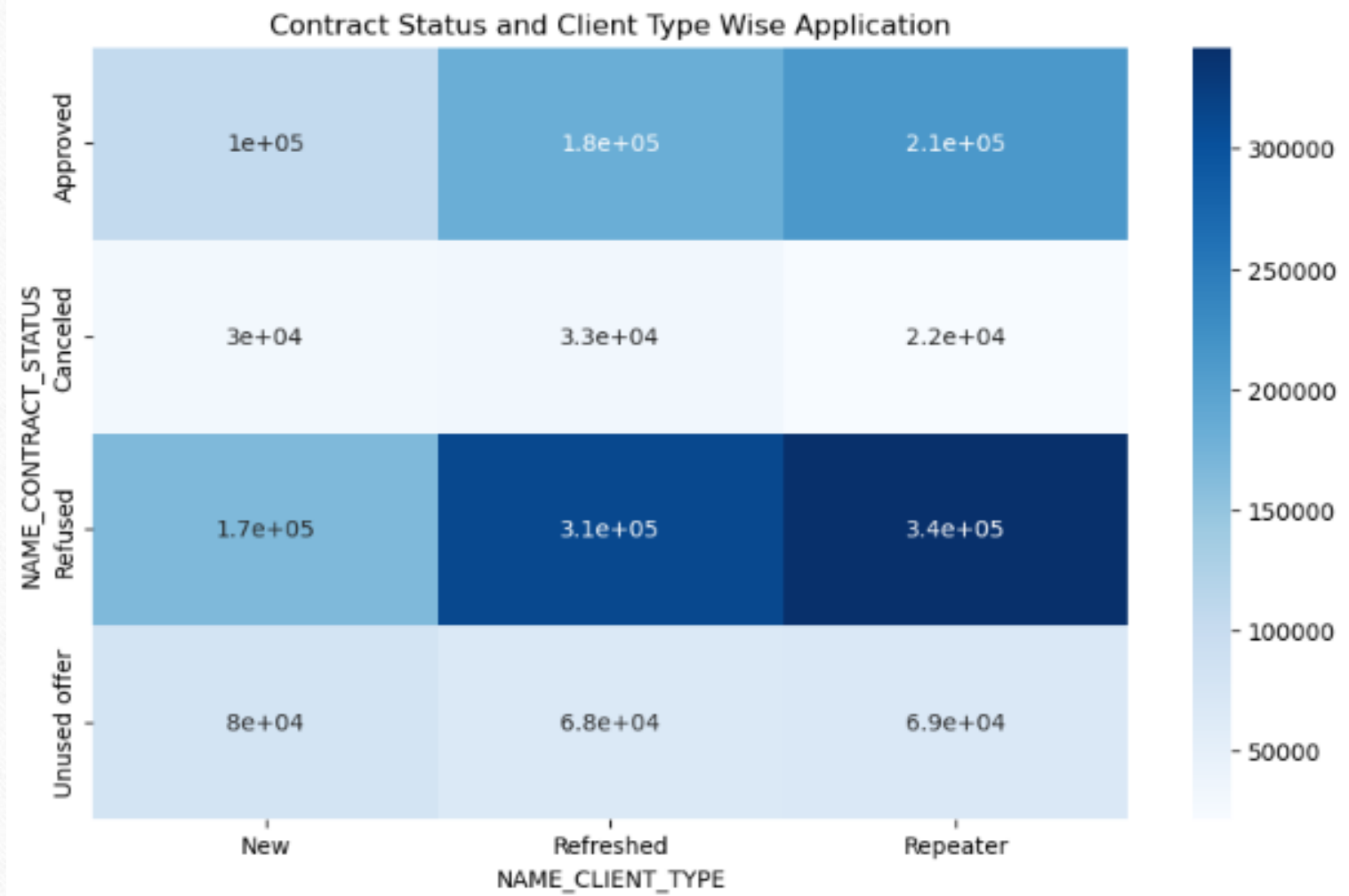
Clients got higher refusal for demanding high credit amount.



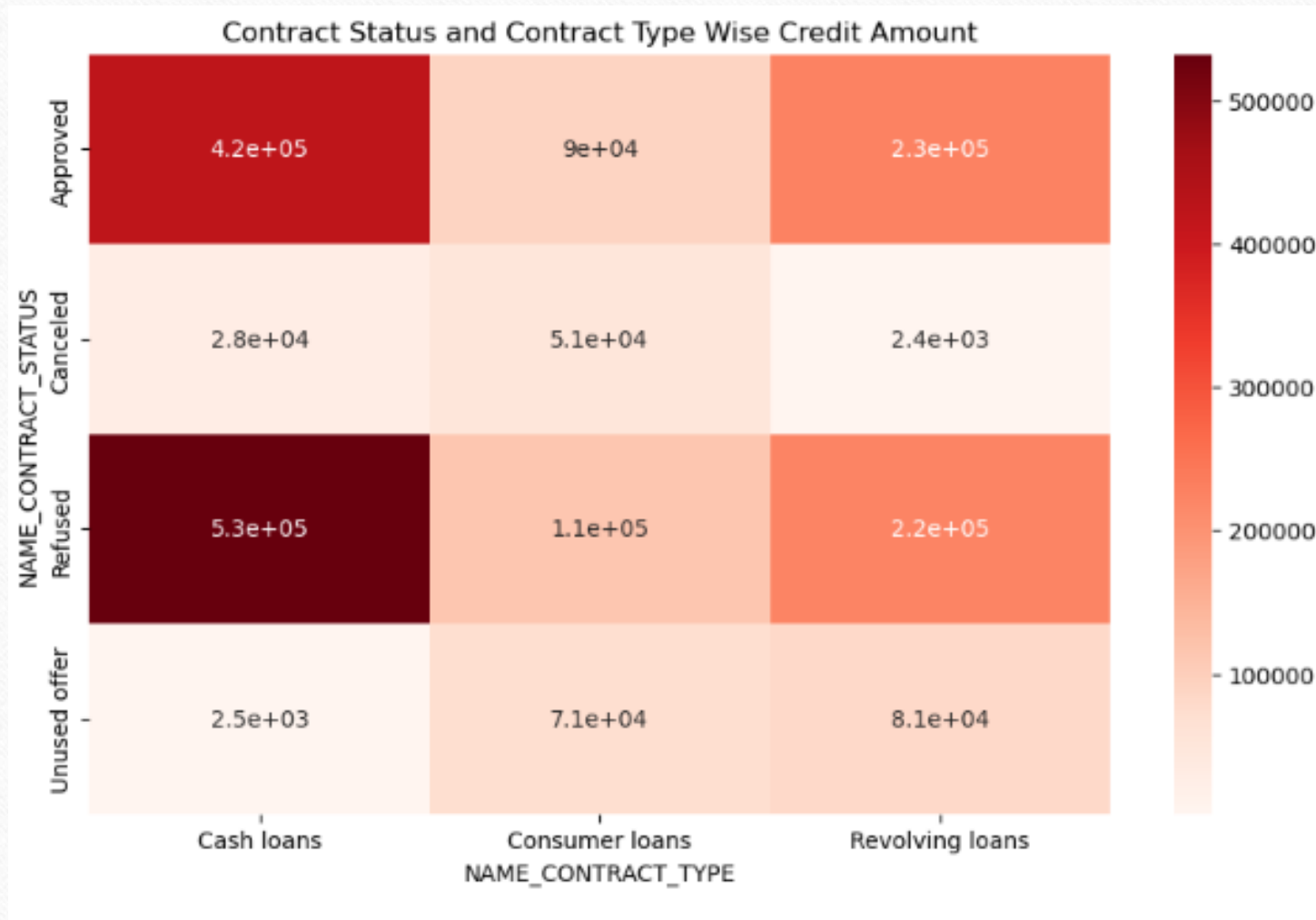
Most no of applicants are repeater.



People are also getting higher refusal as per annuity amount



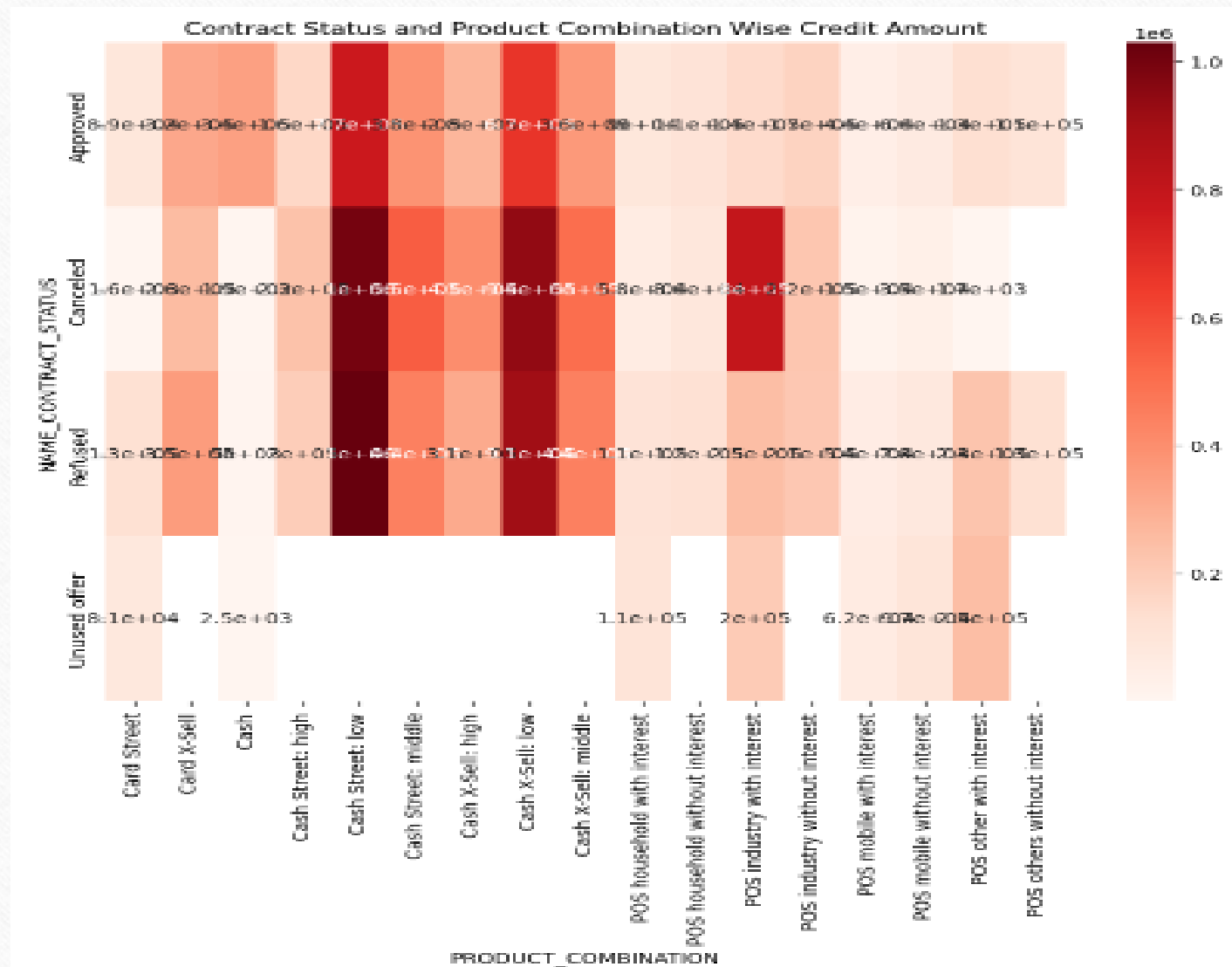
Clients who are applying for loan amount most of them are repeater and due to high loan demand they are getting rejection from bank.



Clients who are asking for cash loan with higher credit amount are getting more rejection from bank compared to the people who are asking for reasonable loan amount. Those people are getting loan approval.

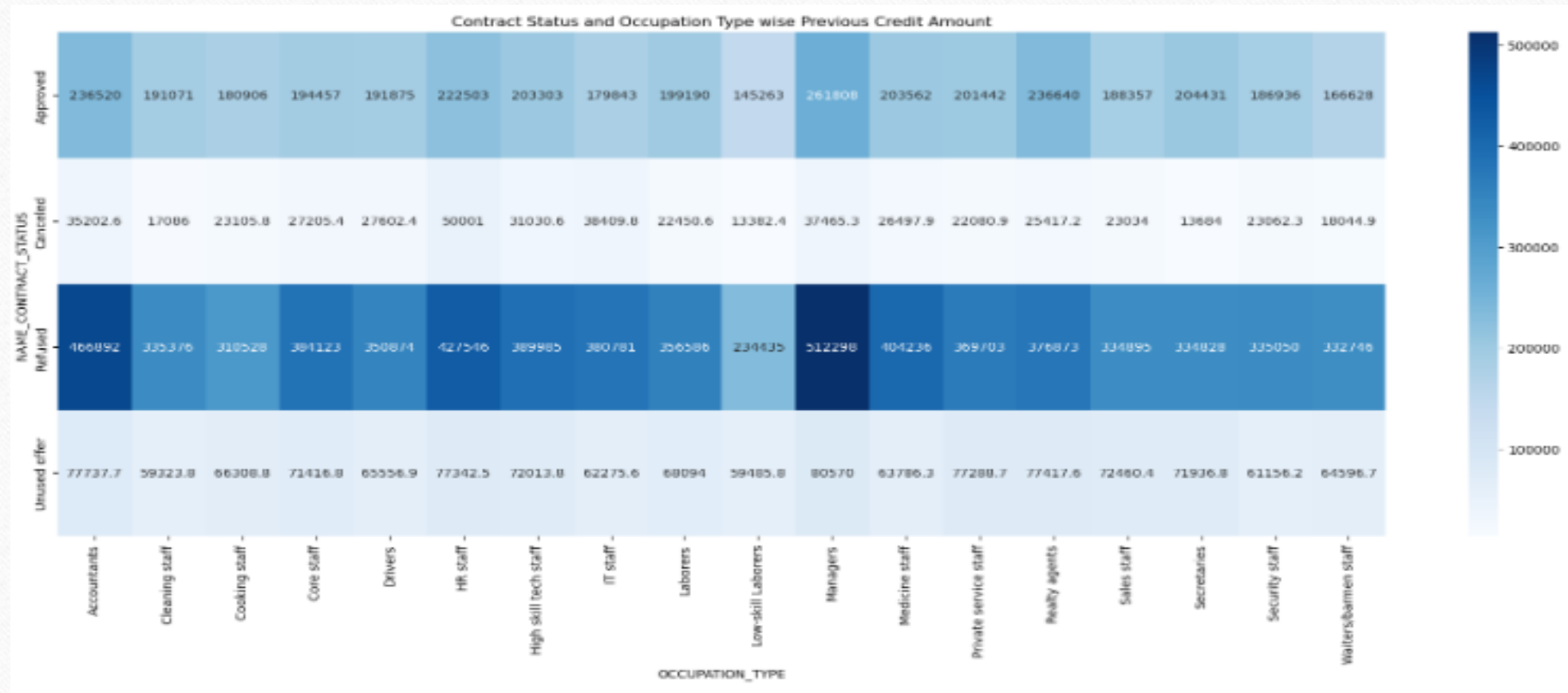


People who are asking for higher credit amount they get refusal from bank though most of them are repeater.

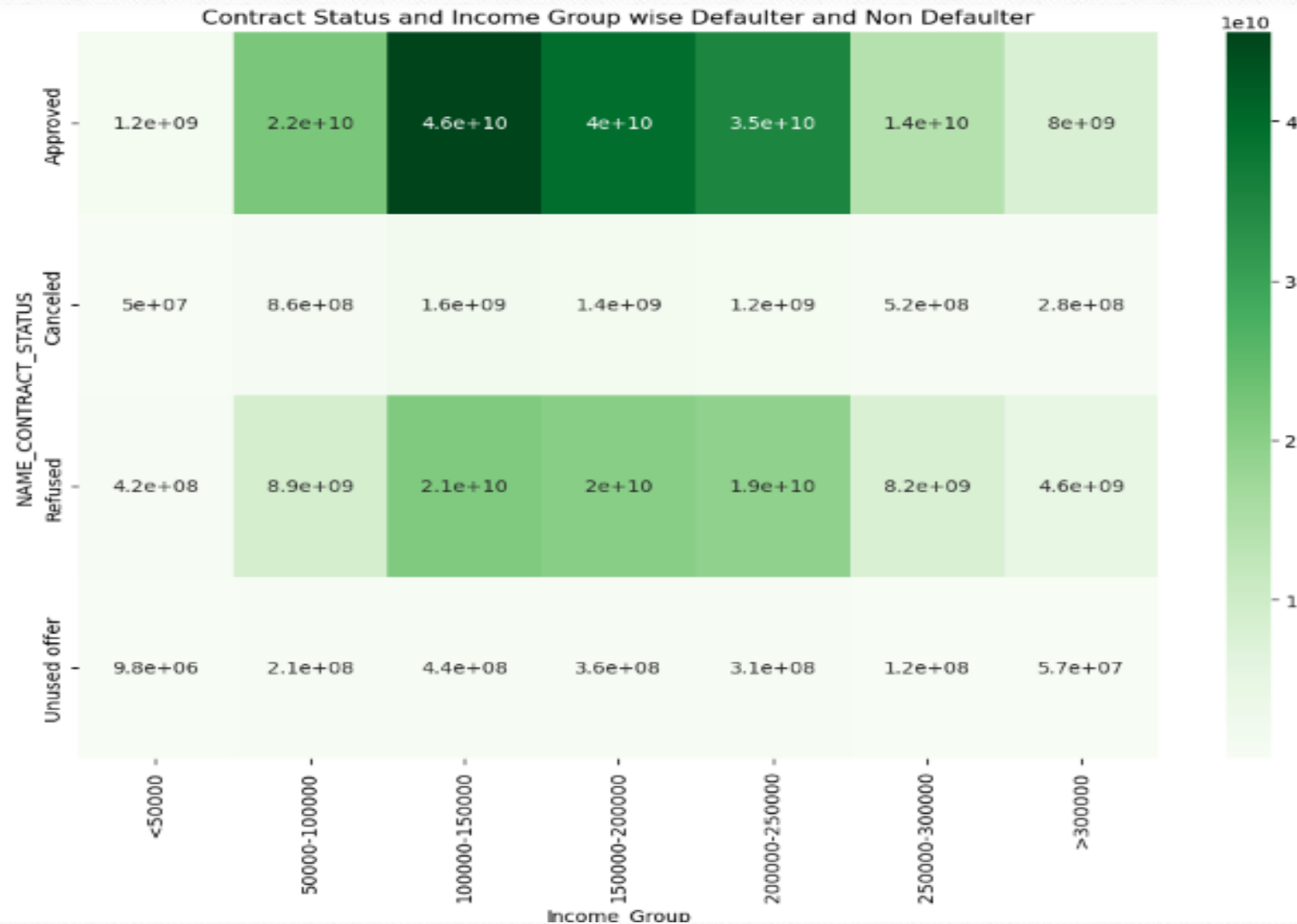


Clients who are asking for higher credit amount for product details like cash street low and cash x sell low got rejection for loan amount and for few of then loan got cancelled.

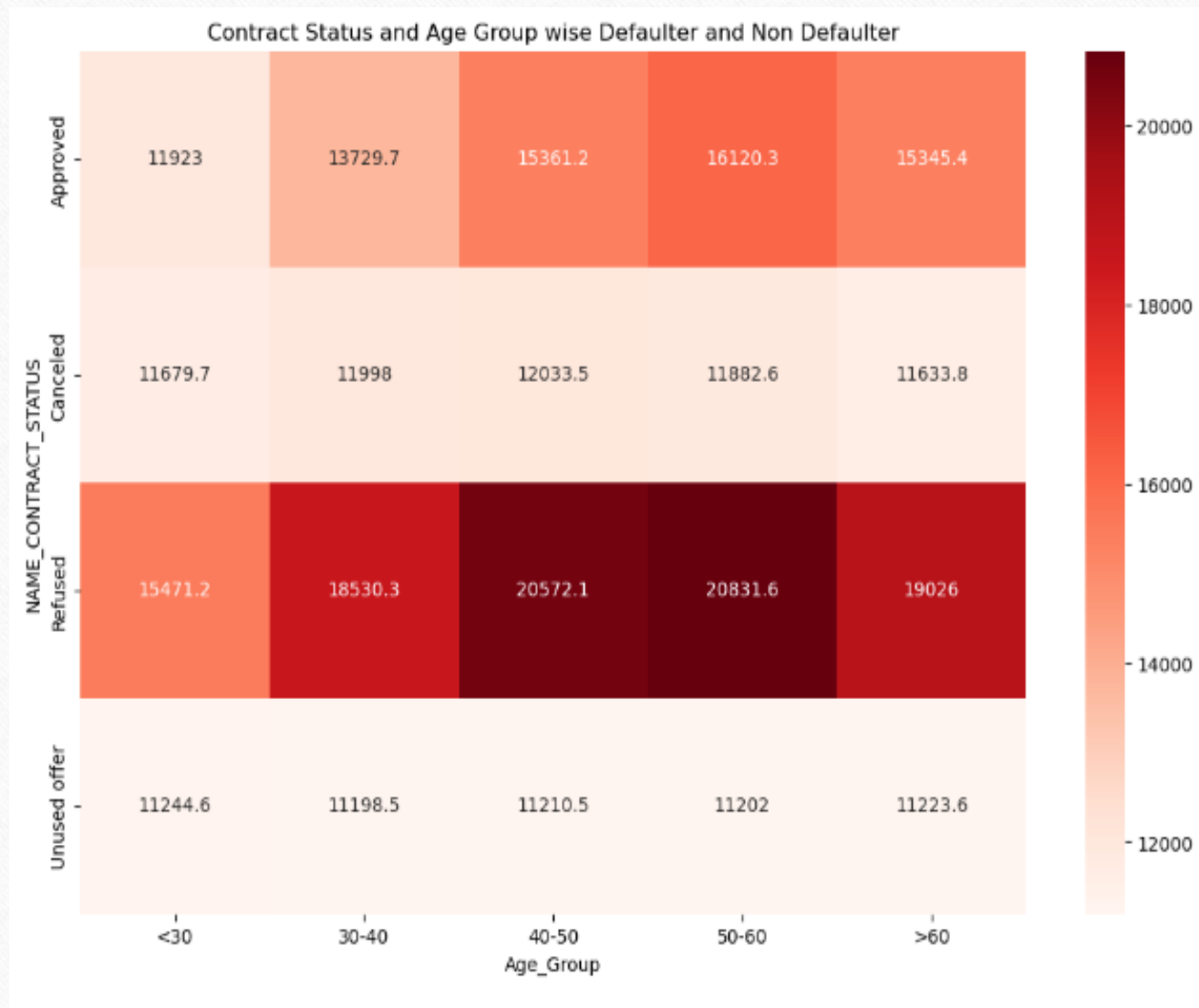
Analysis of Merged Dataset



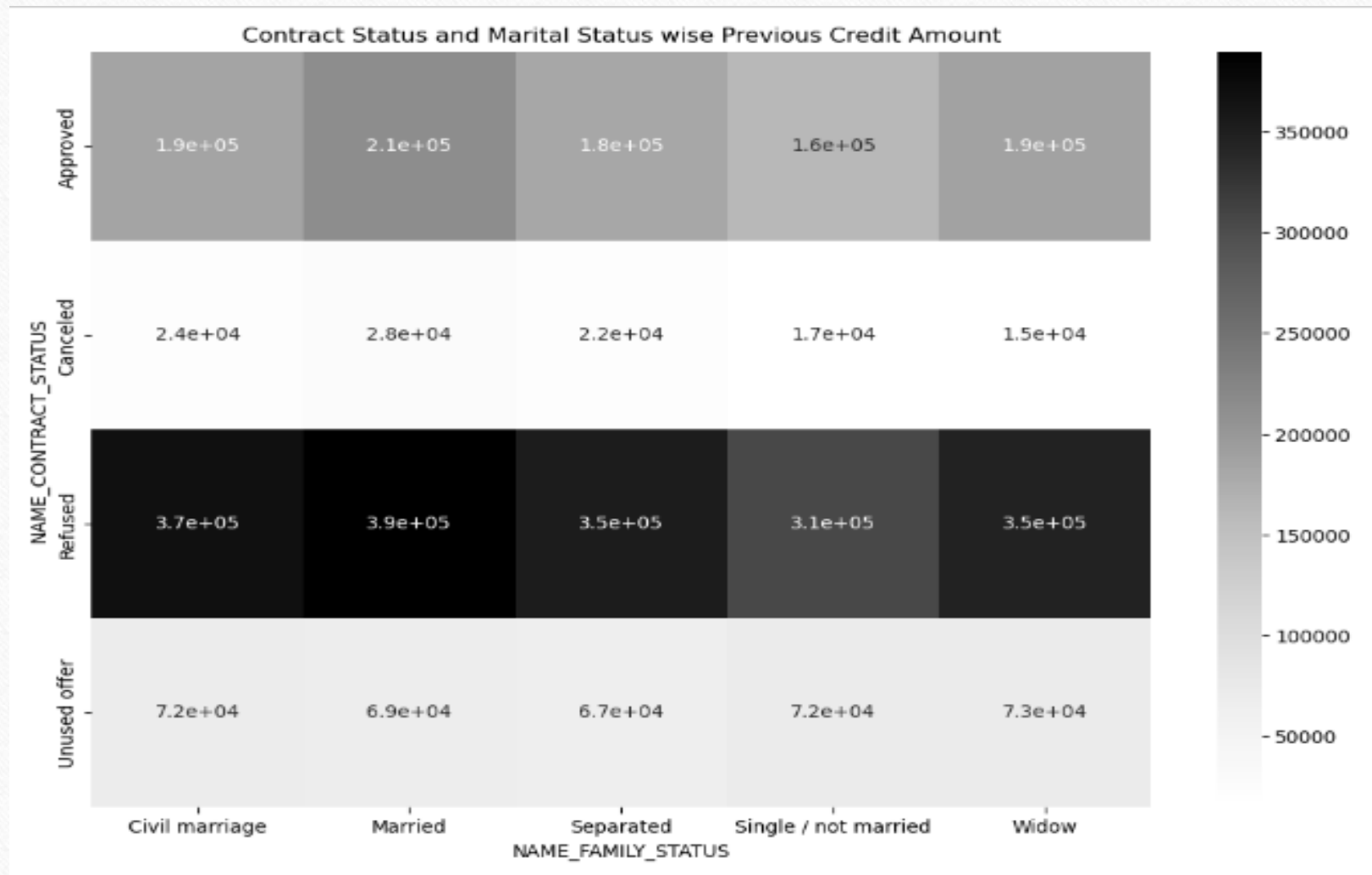
People who are managers and accountants are getting more rejection for loan when asking for credit amount as per previous data. So we can conclude that they are defaulter and have not paid loan on time which is loss for bank.



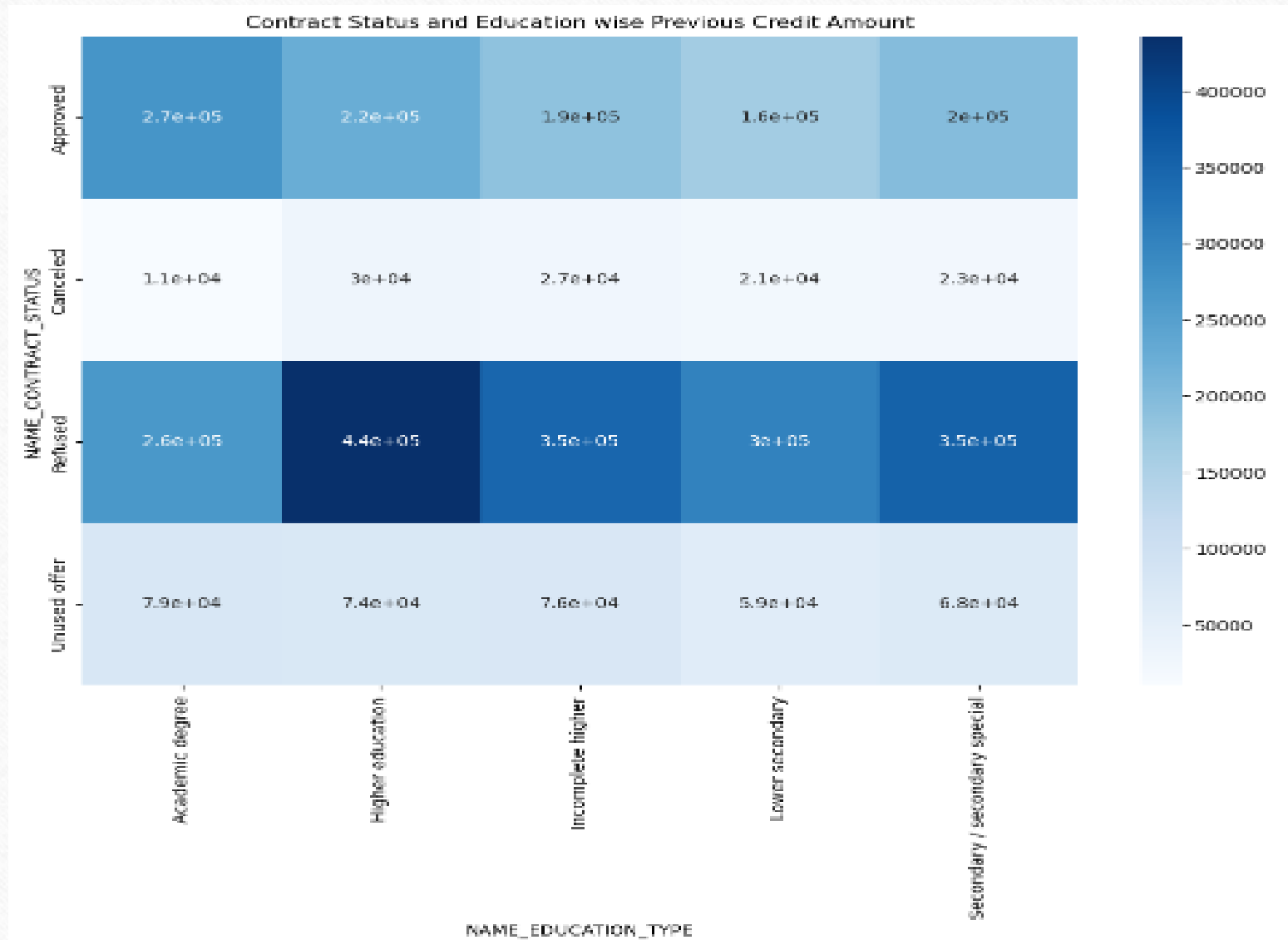
People who are around the income group of 100000-150000, 150000-200000, 200000-250000, mostly ask for credit amount and they got banks approval as per previous data. So we can say they face less difficulty in repaying the loan amount to bank.



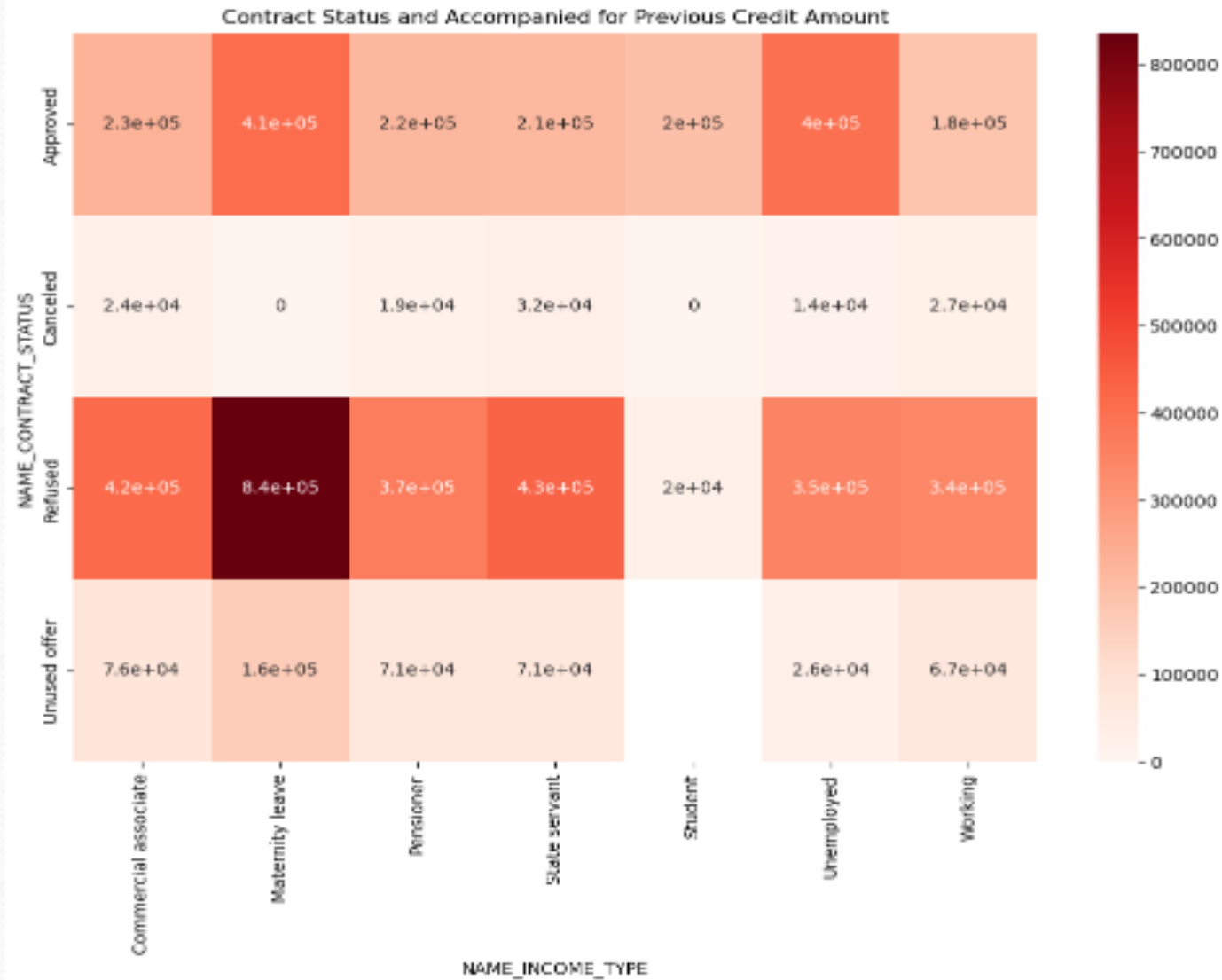
As per previous application, annuity amount for the age group of 40-50 and 50-60 are getting more rejection because they are the defaulter people who face difficulty in repaying their loan .



As per previous data, Married and civil marriage people get rejection for credit loan amount because they are mostly defaulter people.



As per previous data, highly educated people getting more rejection from bank because they are the defaulter and have not paid loan on time.



As per previous data, people are getting rejection for credit amount who are under maternity leave because they are the defaulter.

Conclusion

• Based on Application Data:

- ❖ People are more interested in cash loan.
- ❖ No of defaulter is lesser than non defaulter.
- ❖ Females are mostly defaulter.
- ❖ People like to take more home loan than car loan.
- ❖ Laborer like to take loans but they are defaulter. Managers, Accountants take most of the loan amount.
- ❖ Mostly married people who are around the age group of 30-40 and unaccompanied like to take more loan.
- ❖ Higher income group people with less family member apply for more credit amount.
- ❖ Commercial associate, state servant they ask for more credit amount and they are the defaulter.
- ❖ Secondary educated people take less loans and academic degree people apply for more loans and they are defaulter.
- ❖ Businessman take most of the loan and clear loan on time and pensioner face difficulty in repaying loan.

• Based on Previous Application Data

- ❖ As per previous report most of the loan got approved
- ❖ Most of the people are repeater and they are interested in consumer loan.
- ❖ Refusal rate is high for higher credit and annuity amount. That means people face difficulty in repaying loan.
- ❖ Approved loan means who were better client as per previous data.
- ❖ People who are repeater mostly get rejection from bank due to their payment difficulty.
- ❖ Most of the approved loans are based on middle and higher group interest.
- ❖ Most of the client like to pay through bank.

• Based on Merged Data

- ❖ Mostly manager , accountants are defaulter due to their high credit amount so bank refused their loan because of payment difficulty.
- ❖ People around the income group of 100000-150000 are mostly non defaulter and repay on time so they get approval of bank.
- ❖ Married people apply for more loan and they get rejection from bank because they are also defaulter.

- ❖ As per previous record, age group of 50-60 and 60 above people get more rejection due to their high credit value and annuity amount, so they face difficulty in repaying the loan which is loss for bank.
- ❖ Specially widow women are more defaulter when because they have single income source and face difficulty in repaying loan, so bank refused their loan.
- ❖ Bank will Approve loan for those clients who has clear loan record as per previous data.

Thank You