

Banking Dataset Analysis

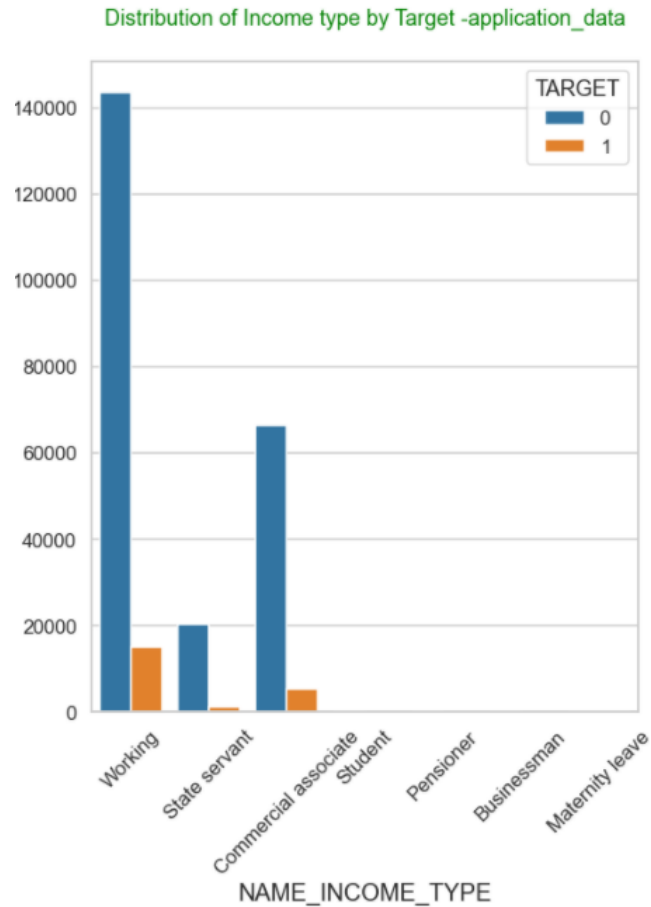
- Sruthi Thabati
Santapuri Jayavardhan Rayudu

Objective:

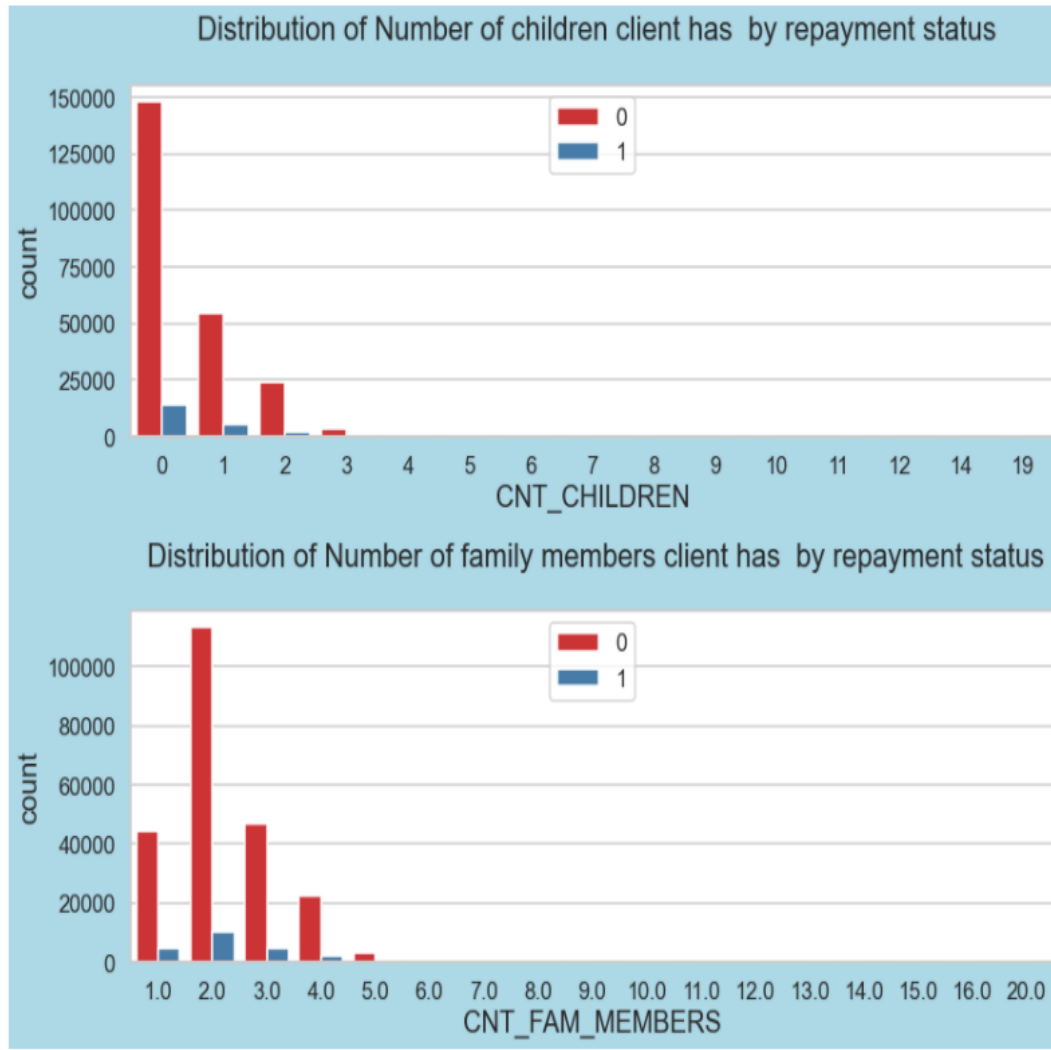
This case study aims to identify patterns which indicate if a client has difficulty paying their installments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. This will ensure that the consumers capable of repaying the loan are not rejected. Identification of such applicants using EDA is the aim of this case study.

Way chosen:

- Analysing the Application data and previous data after cleaning the data.
- Analysing the above datasets using target variable 0 and 1.
- Creating correlation matrix for all the numerical values.
- Conclusion.



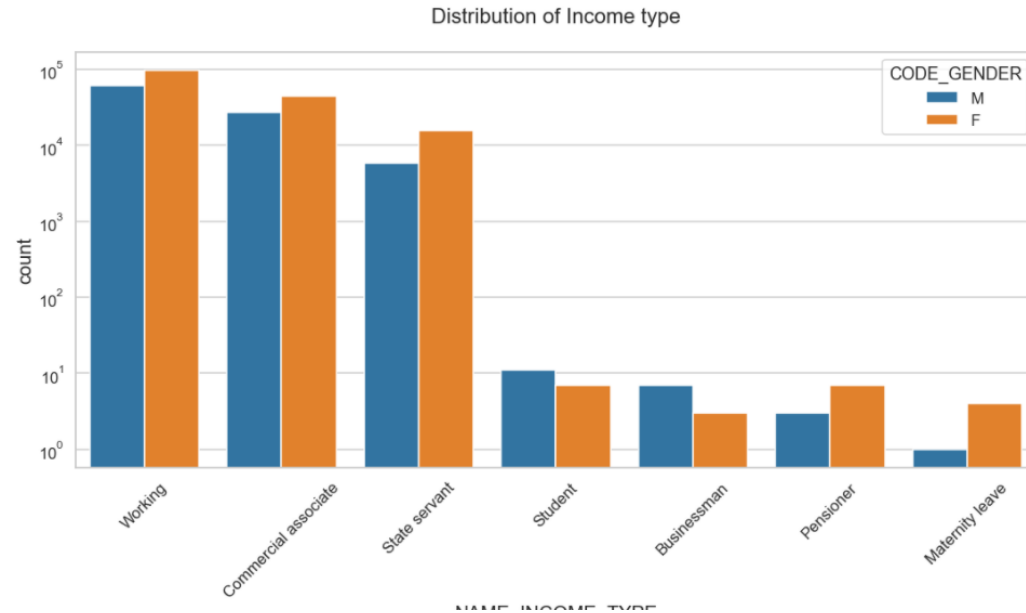
- ❑ In the Application data, defaulters are more in the case of working compared to others. At the same time, repayers are also more in the same income type category.
- ❑ There are less number of defaulters compared to repayers in the Application data basing on income type category.



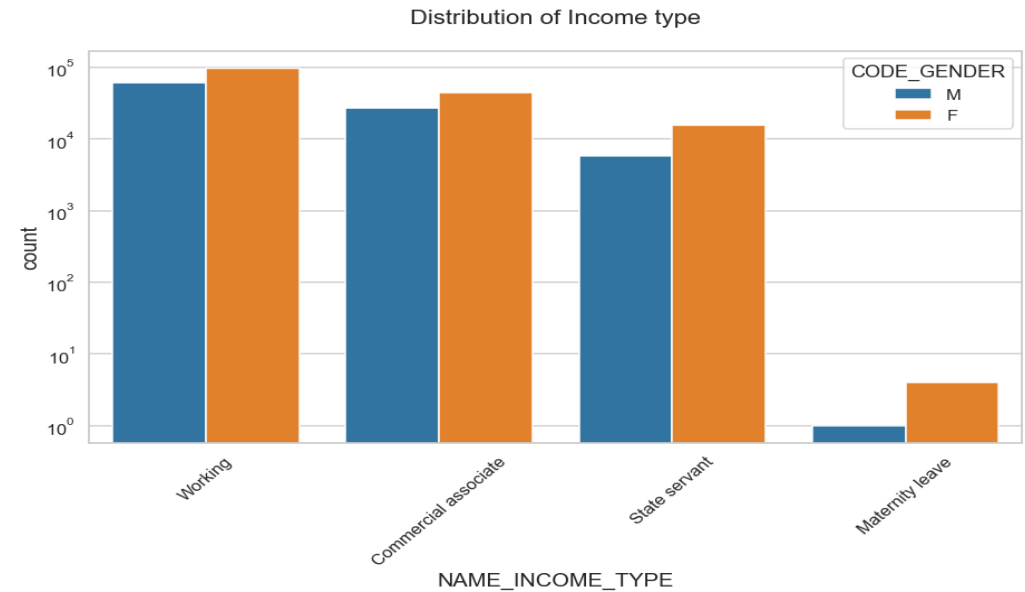
- ❑ Considering the count of children and the family members in the Application data, People with less or no children are more in number compared to people with more children. And, their rate of repayment is also more.
- ❑ Considering the case of Family members, People with 2 Family members are more in number compared to others who applied for loan. And, their rate of repayment is also more.

Distribution of Income Type

Target 0



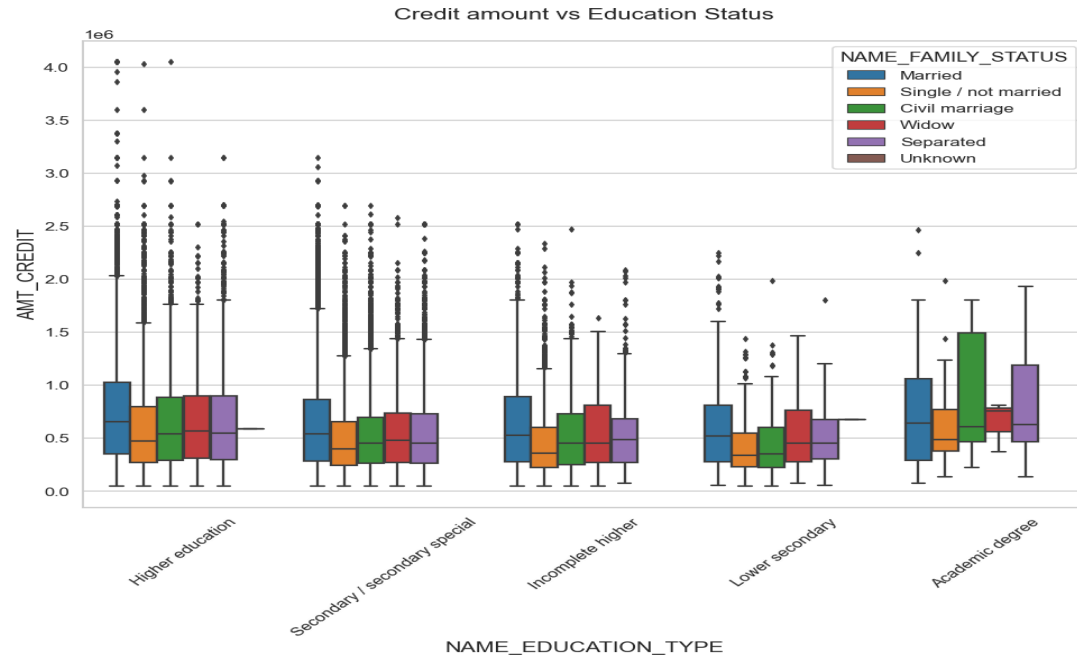
Target 1



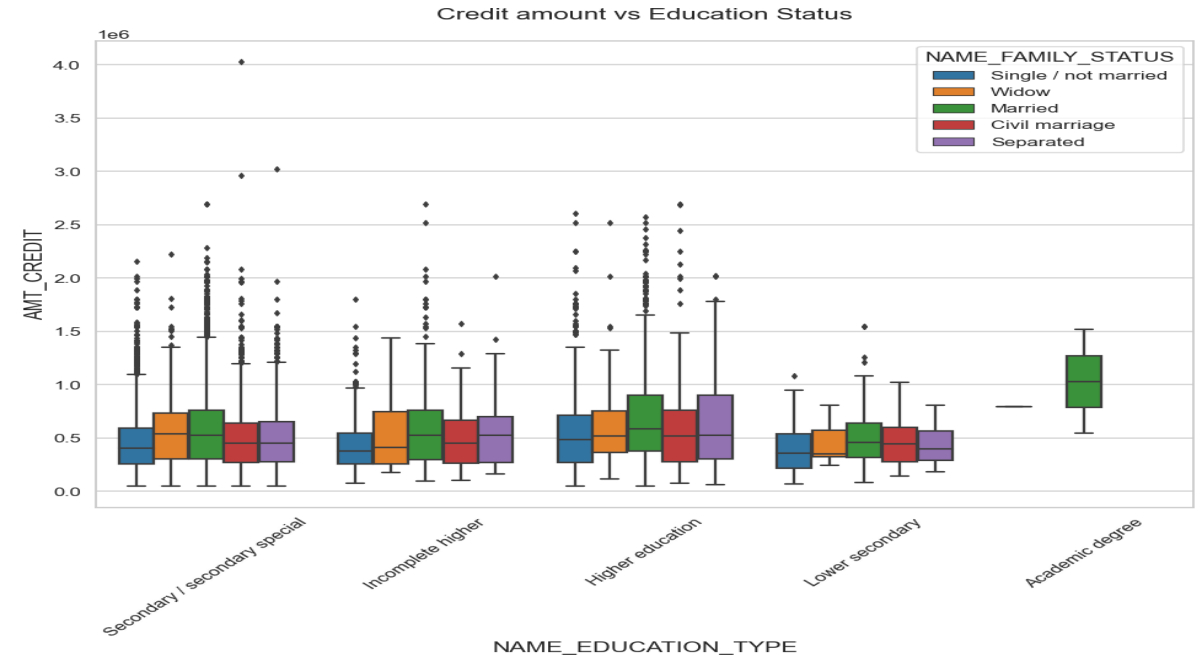
- ❑ Comparing both the situations, we can infer that Student, Businessman and Pensioner categories have repayed the loan compared to others.
- ❑ There is no much difference based on the Gender category to repay or default the loan.
- ❑ Based on the Income and Loan amount, it would be better for banks to approve loan for businessman.

Credit amount vs Education Status

Target 0



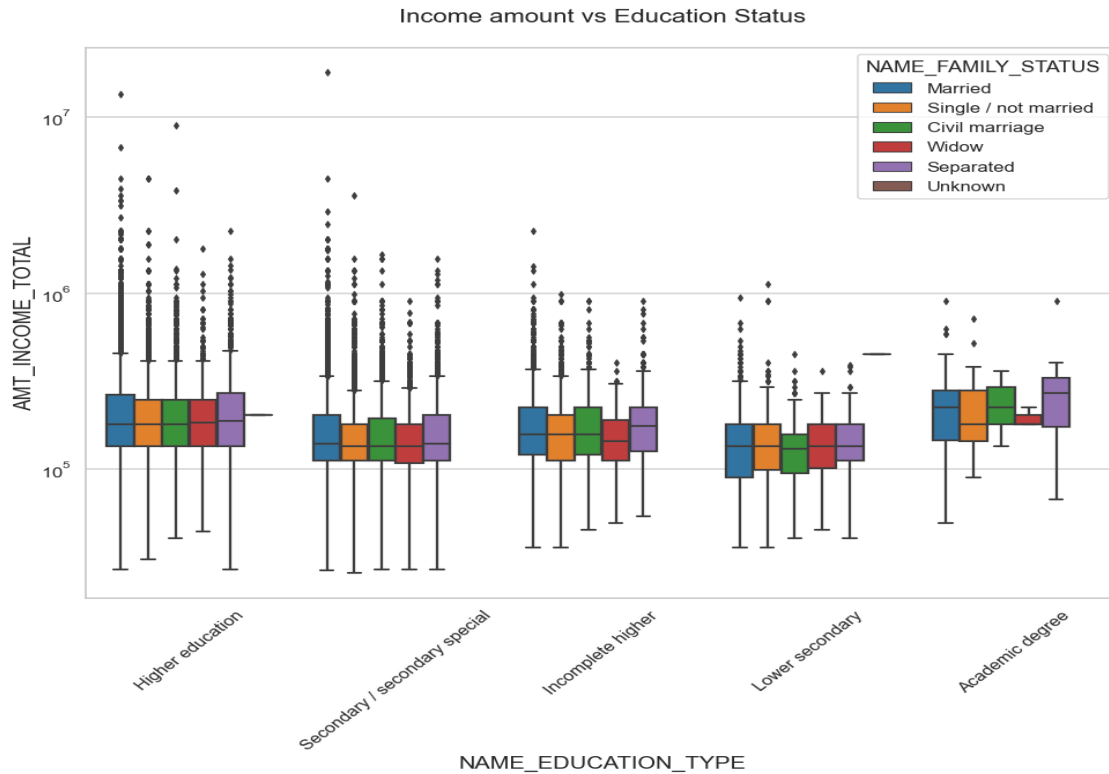
Target 1



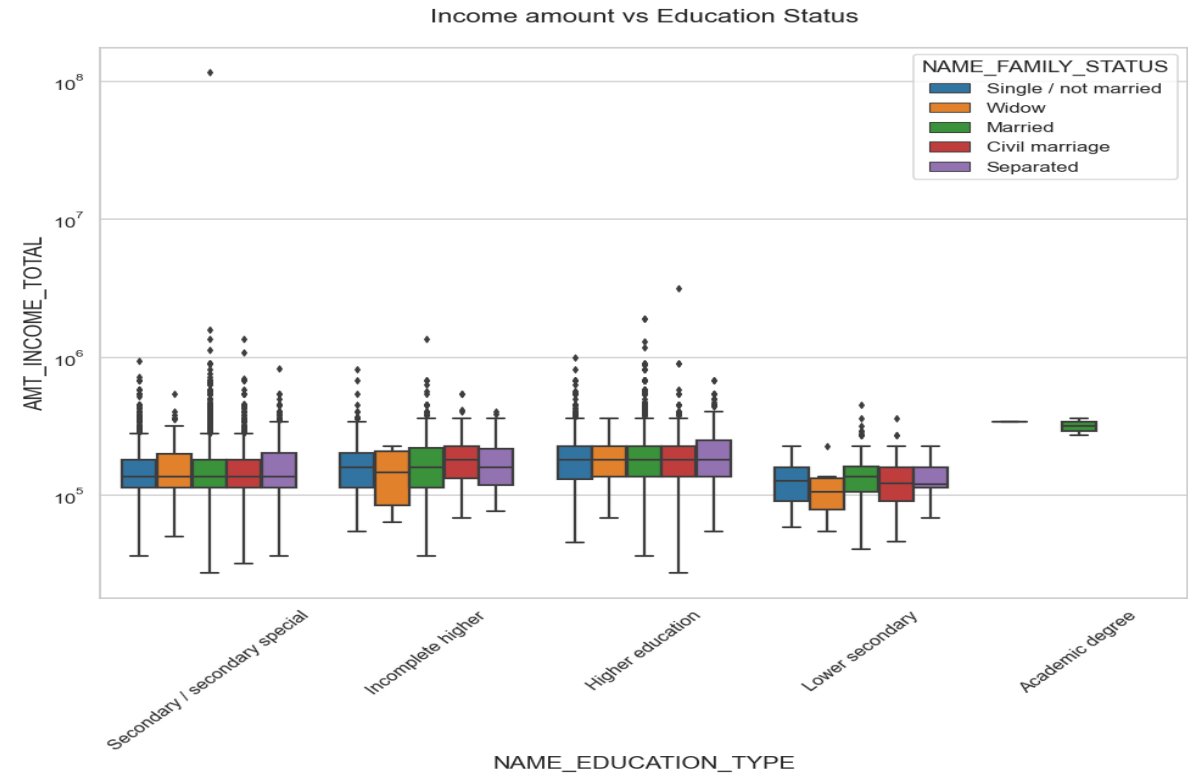
- ❑ People with Family status of Civil marriage with academic degree have chances of not repaying the loan compared to other marital status with academic degree.

Income amount vs Education Status

Target 0

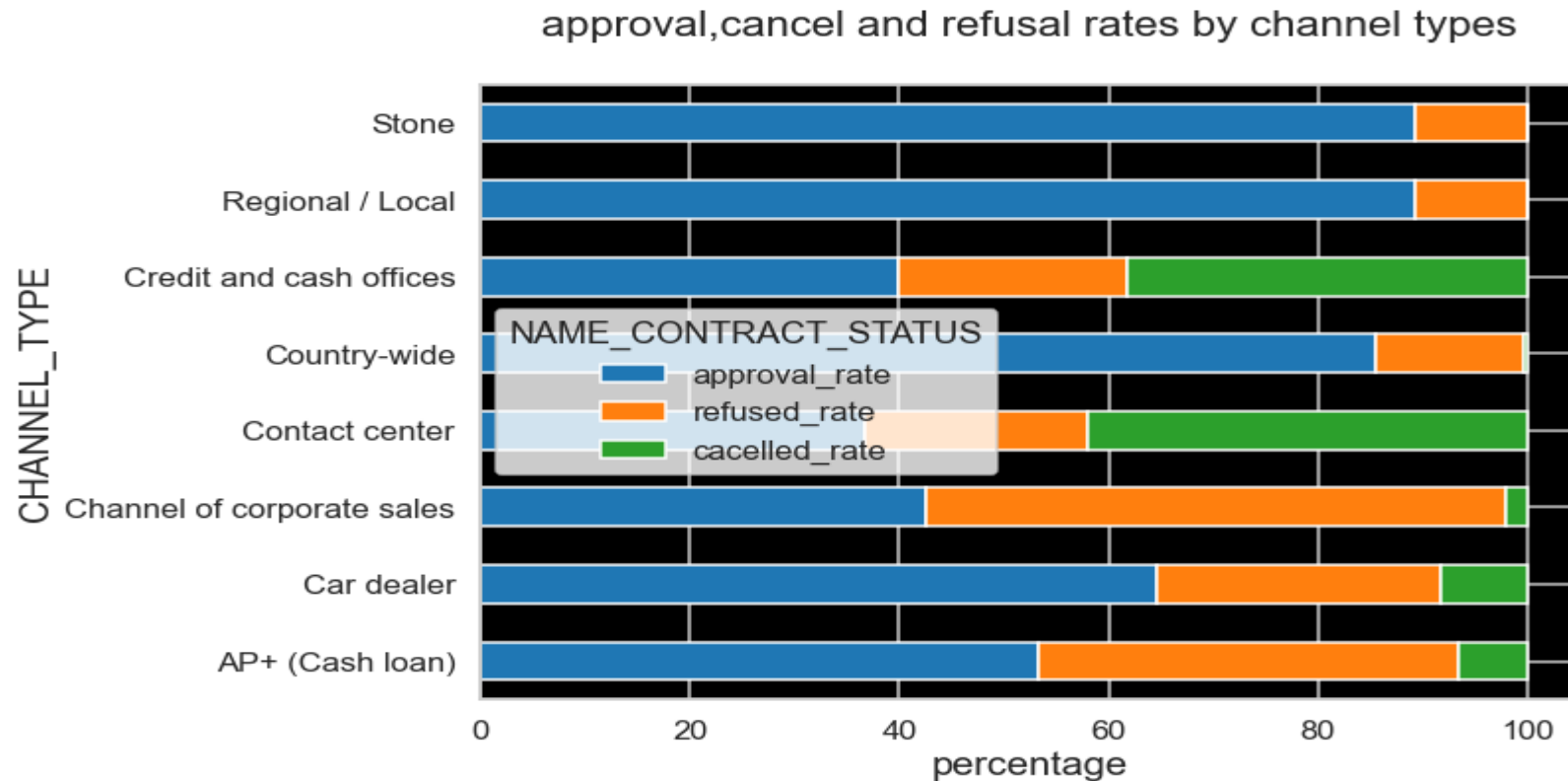


Target 1



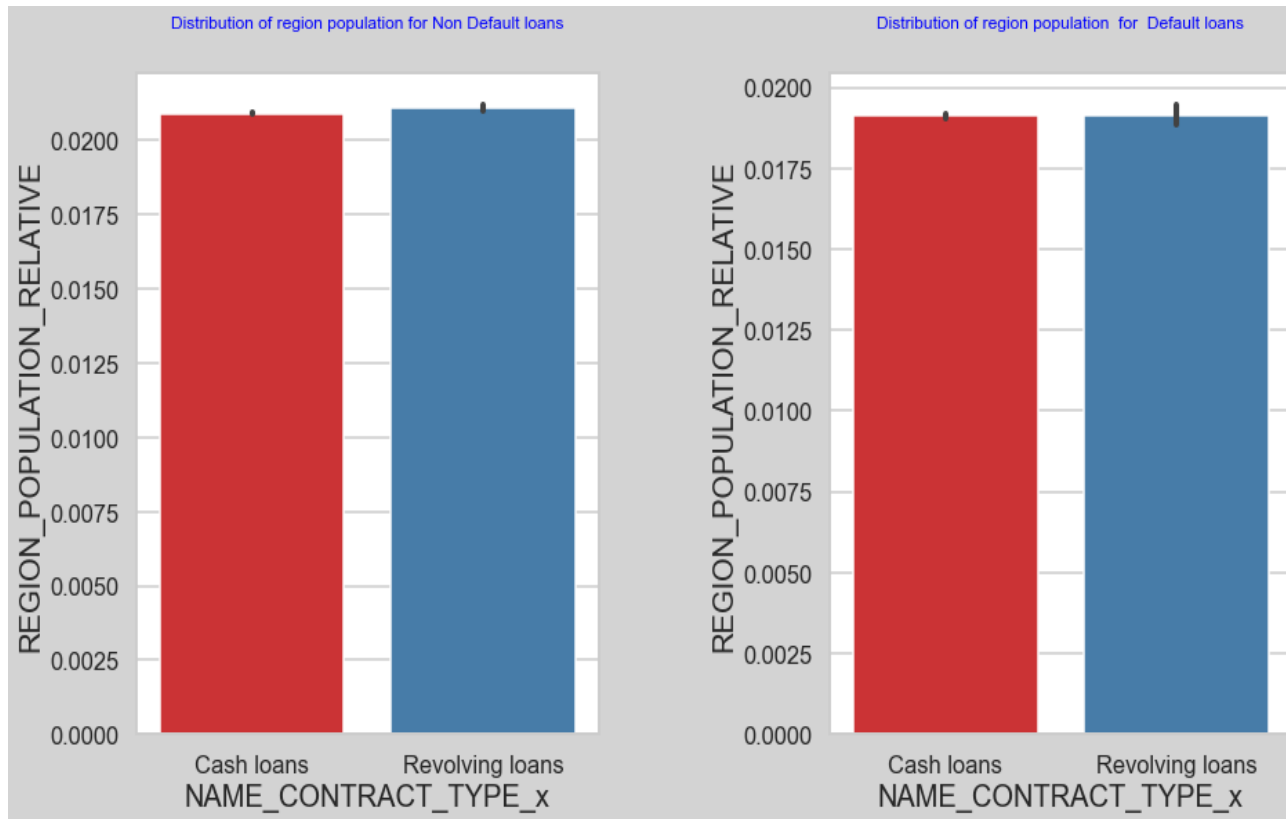
❑ Academic degree holders other than civil marriage can be issued loan irrespective of total income.

Previous_Application dataset



- ❑ Basing on previous application data, approval rate was more for Stone, regional and country wide channels.
- ❑ Refused rate was more for corporate sales and cash loan channel type.
- ❑ Cancelled rate was more for credit and cash office and contact center channel types. Better give them some offer and they may turn good for the bank.

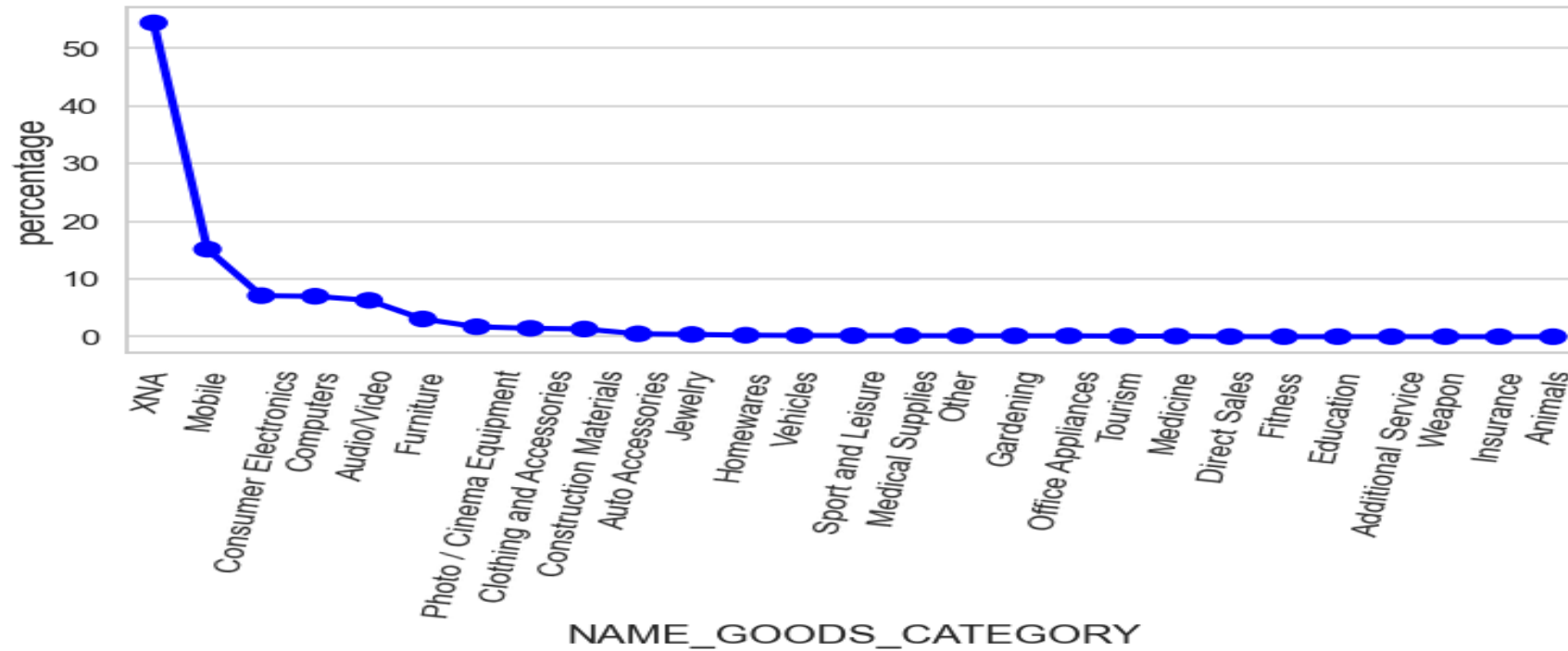
Distribution normalized population of region where client lives by loan repayment status



- ☐ This could be a best catch for the banks as this might provide best possible insight.
- ☐ Location wise like rural and urban loans division can be fruitful to take a decision.

Missing values

popular goods for applying loans

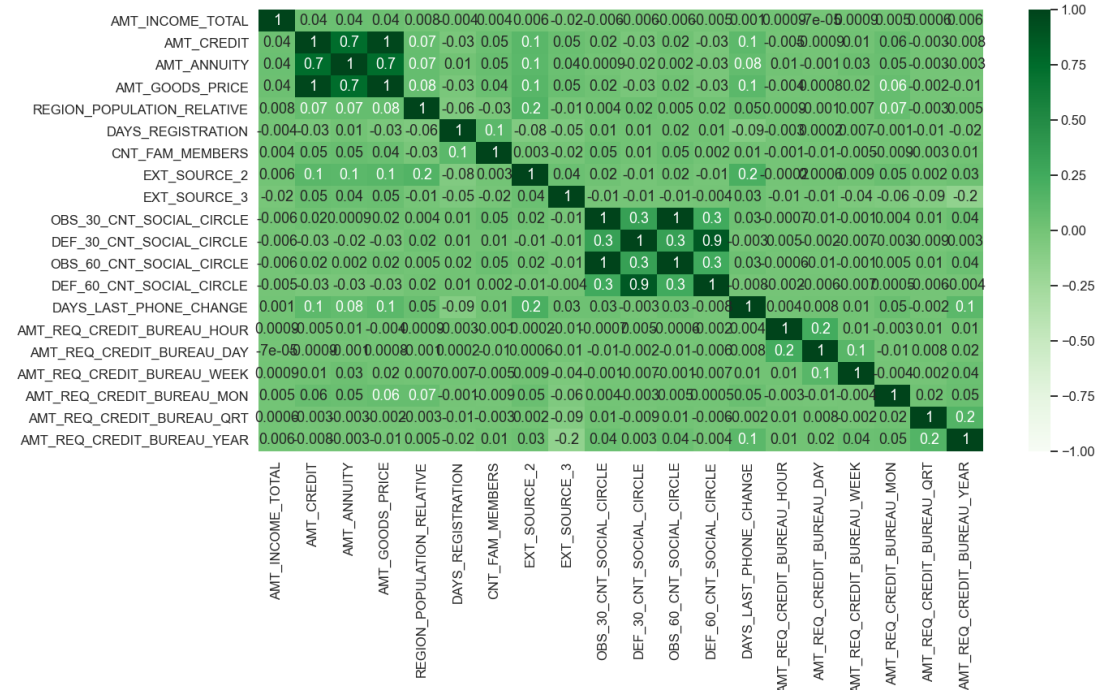
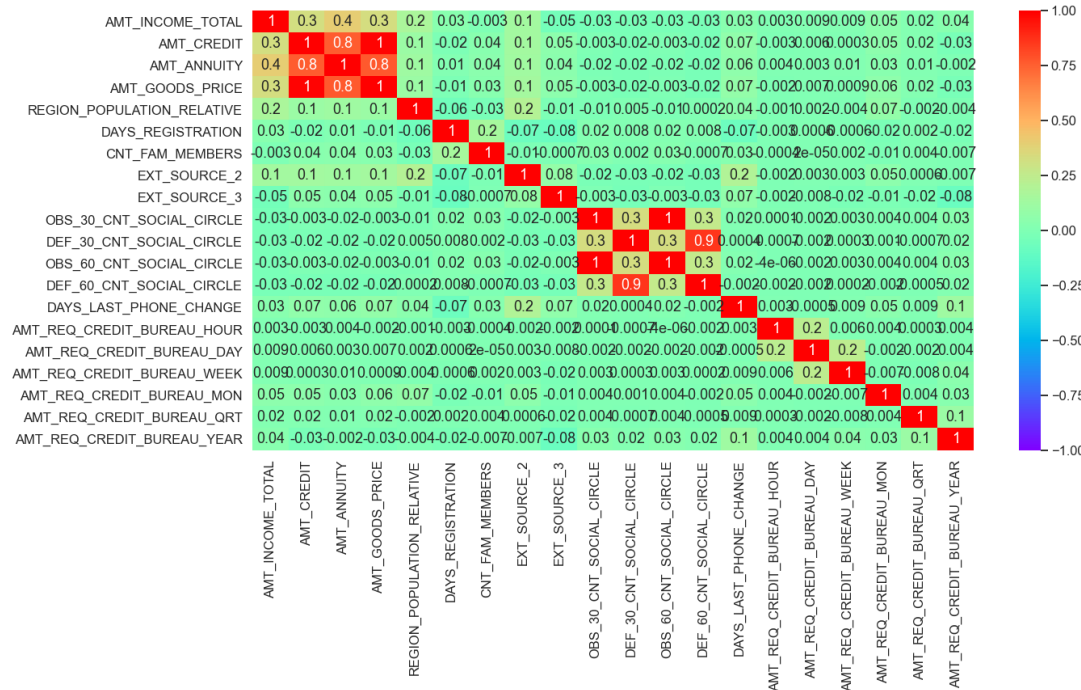


- ❑ Basing on the missing values present in the application_data and previous_application data, it will be advisable for banks to maintain this data as these seems providing valuable info to draw out conclusion .

Correlation matrix

Target 0

Target 1



❑ The correlation matrix looks same for both the targets.

Top 10 correlation variables

	Var1	Var2	Correlation
2494	OBS_60_CNT_SOCIAL_CIRCLE	OBS_30_CNT_SOCIAL_CIRCLE	1.00
5068	AMT_GOODS_PRICE_y	AMT_APPLICATION	0.99
466	AMT_GOODS_PRICE_x	AMT_CREDIT_x	0.98
4991	AMT_CREDIT_y	AMT_APPLICATION	0.97
5069	AMT_GOODS_PRICE_y	AMT_CREDIT_y	0.97
1559	REGION_RATING_CLIENT_W_CITY	REGION_RATING_CLIENT	0.96
5849	DAYS_TERMINATION	DAYS_LAST_DUE	0.94
1388	CNT_FAM_MEMBERS	CNT_CHILDREN	0.90
5692	DAYS_LAST_DUE_1ST_VERSION	DAYS_FIRST_DRAWING	0.89
1871	LIVE_REGION_NOT_WORK_REGION	REG_REGION_NOT_WORK_REGION	0.87

Conclusion

- Banks can be advised to sanction loans for people with high Academic degree and high Income.
- Occupational type of business can be given high priority to issue the loan compared to others as there are no defaulters.
- People with Family status of Civil Marriage with Academic degree shouldn't be on the top priority list.
- Credit amount with Educational status doesn't help in drawing conclusions.
- People with 2 Family members have a highest repayment rate than others.

THANK YOU