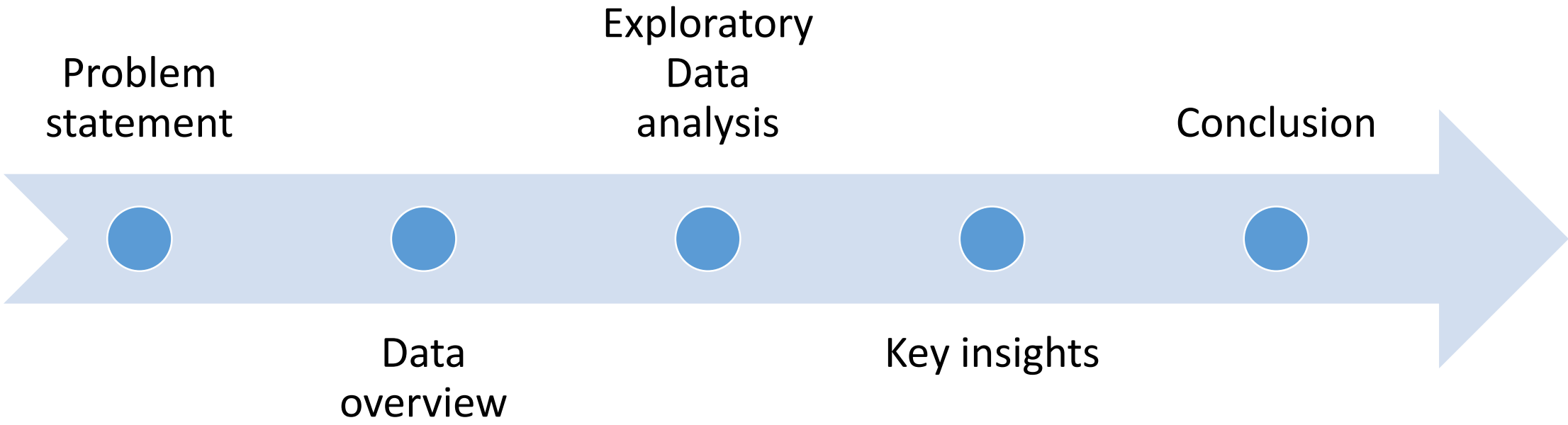
RISK ANALYTICS CASE STUDY

WHAT WE WILL SEE



PROBLEM STATEMENT

The loan providing companies find it hard to give loans to the people due to their insufficient or non-existent credit history. Because of that, some consumers use it to their advantage by becoming a defaulter.

When the company receives a loan application, the company has to decide for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

1. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company

2. If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

DATA OVERVIEW

This dataset has 3 files as explained below:

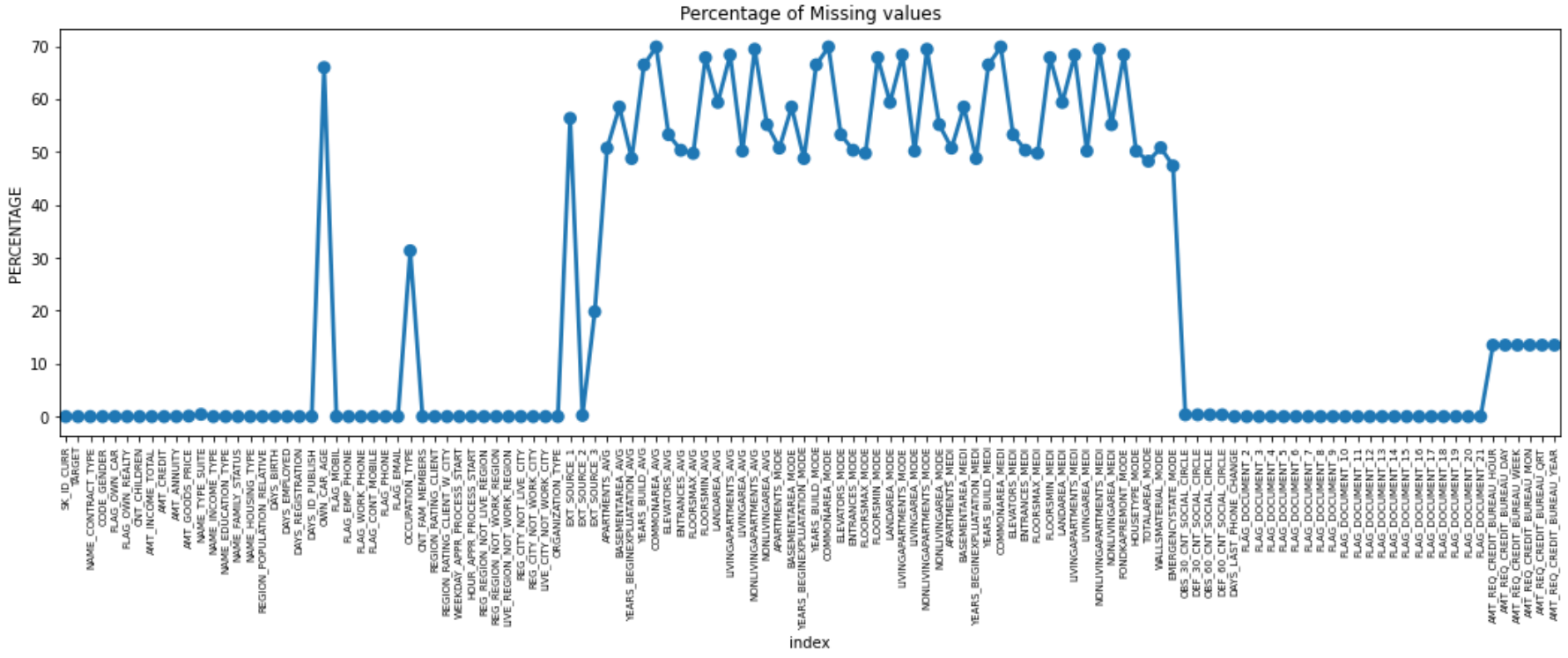
1. 'application_data.csv' contains all the information of the client at the time of application. The data is about whether a client has payment difficulties.

2. 'previous_application.csv' contains information about the client's previous loan data. It contains the data whether the previous application had been Approved, Cancelled, Refused or Unused offer.

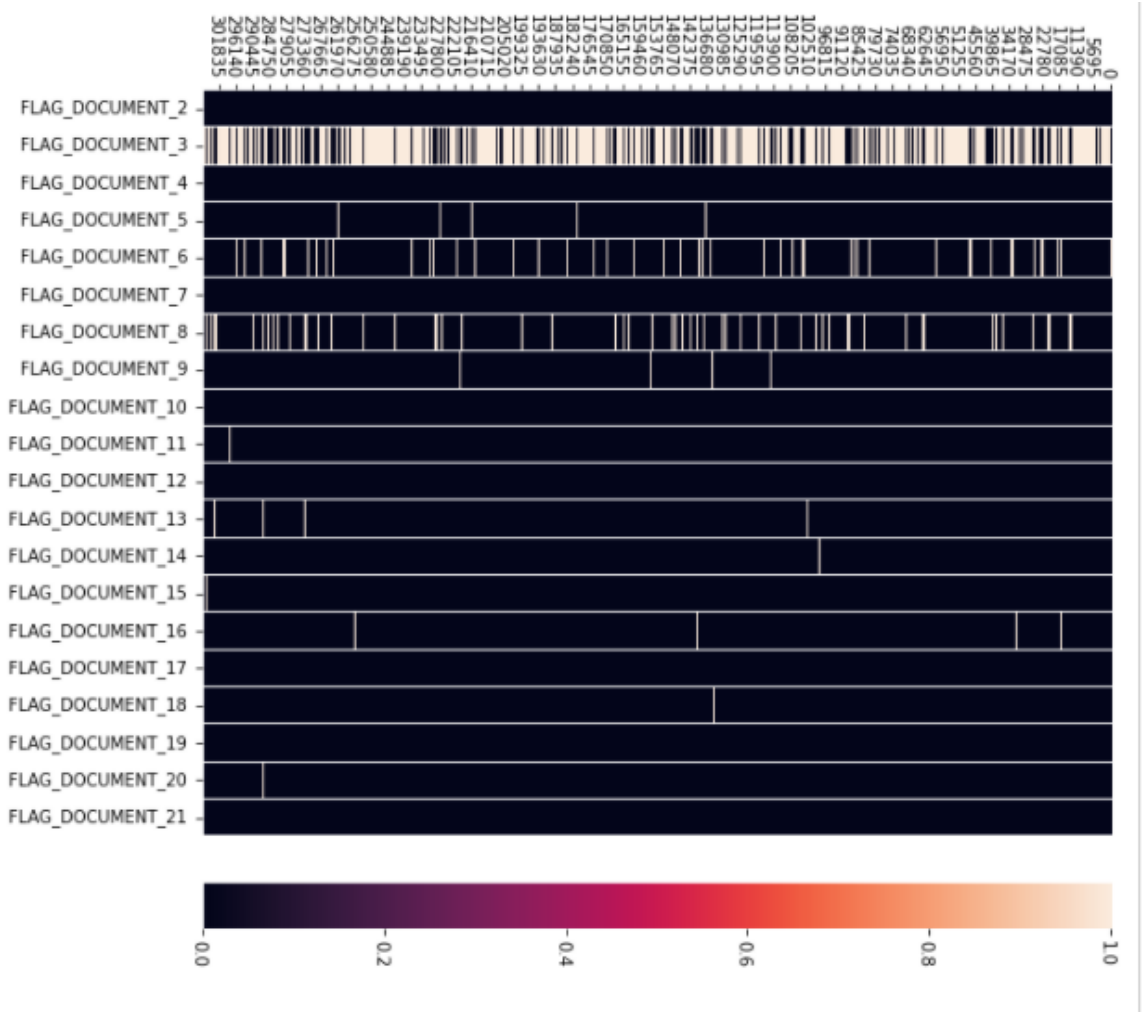
3. 'columns_description.csv' is a data dictionary which describes the meaning of the variables.

EXPLORATORY DATA ANALYSIS

- Here, missing data percentage is very high.
- Maximum missing data percentage is 70%



EXPLORATORY DATA ANALYSIS

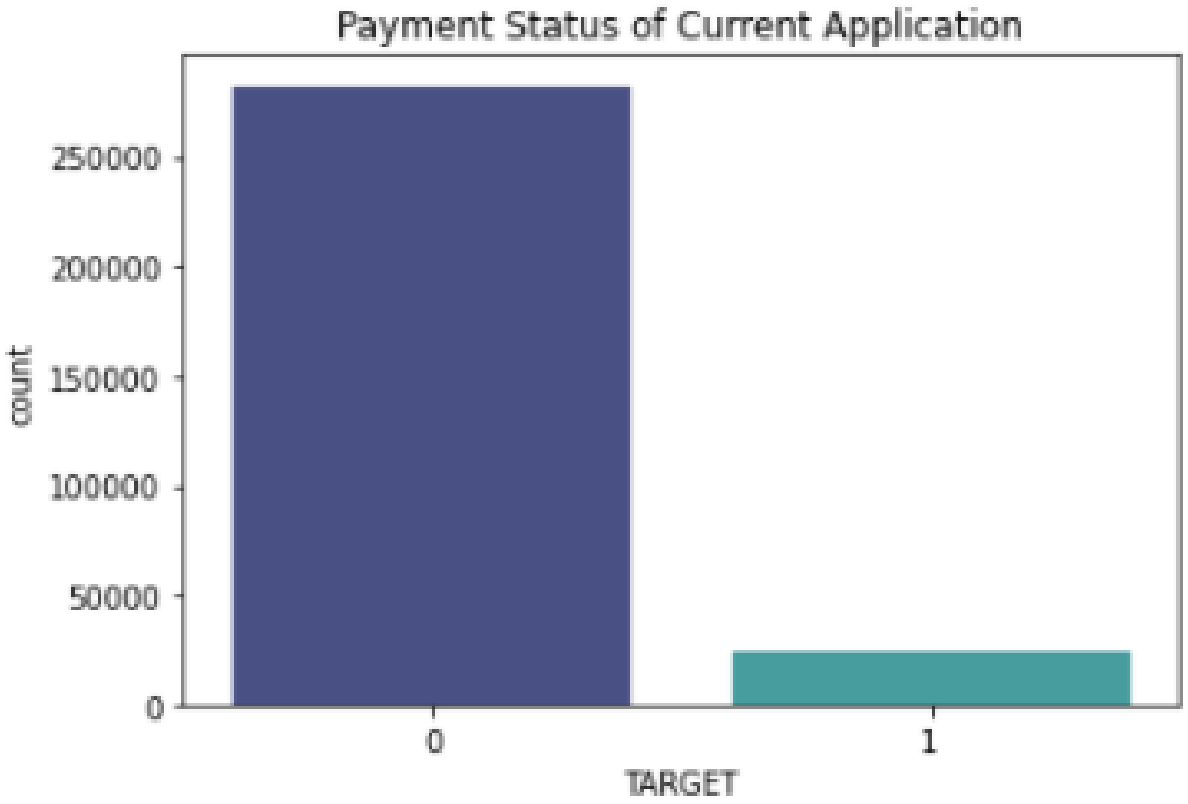


The heat map showing us that all of the Flag documents were not provided by maximum applicants. Hence we dropped all this columns.

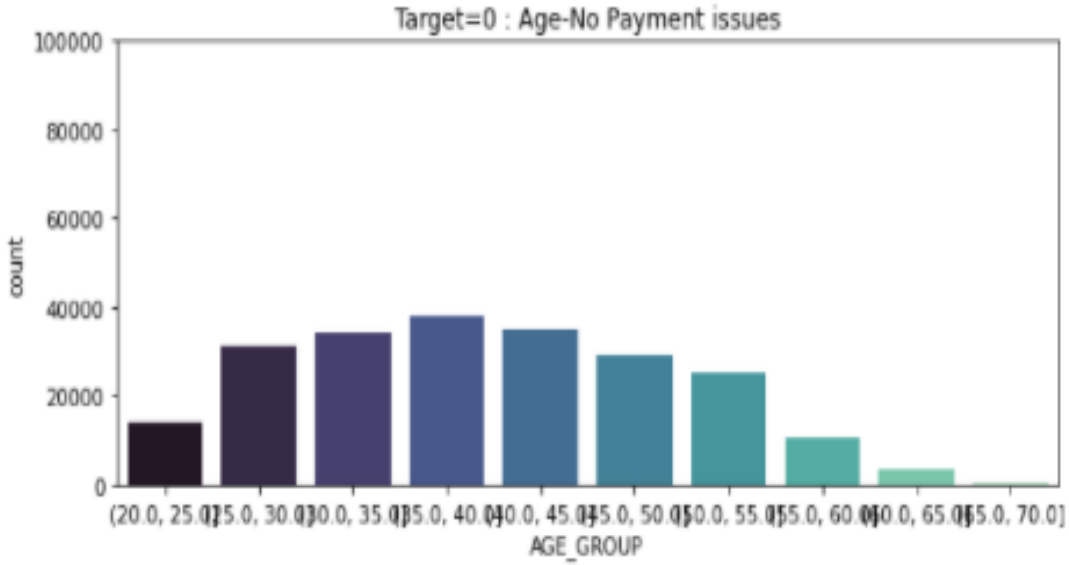
EXPLORATORY DATA ANALYSIS

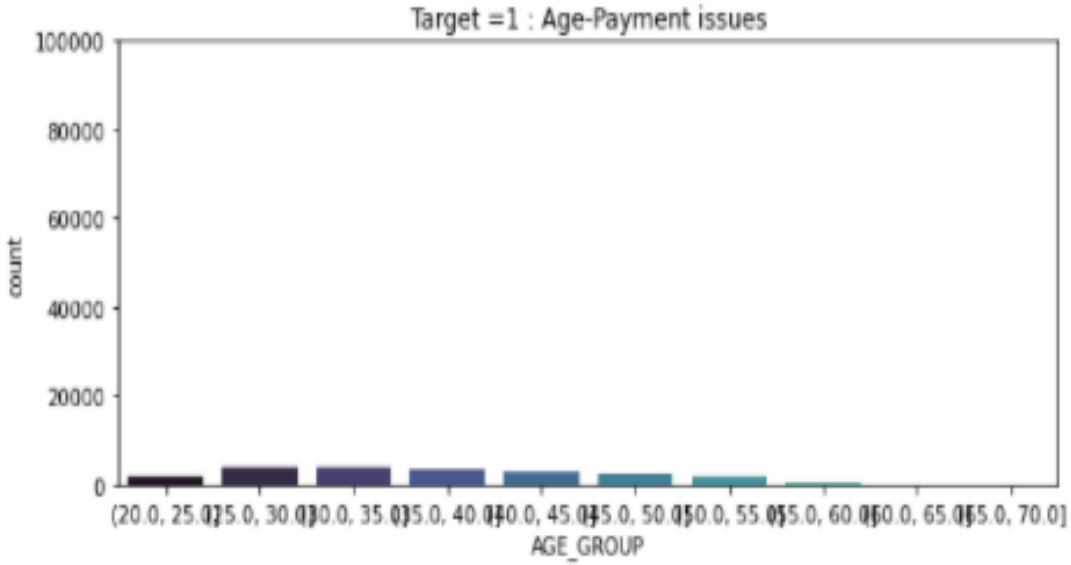
Target variable: (1 - client with payment difficulties: he/she had late payment more than X days on at least one of the first Y installments of the loan in our sample, 0 - all other cases).

This data is highly imbalanced as number of defaulter is very less.



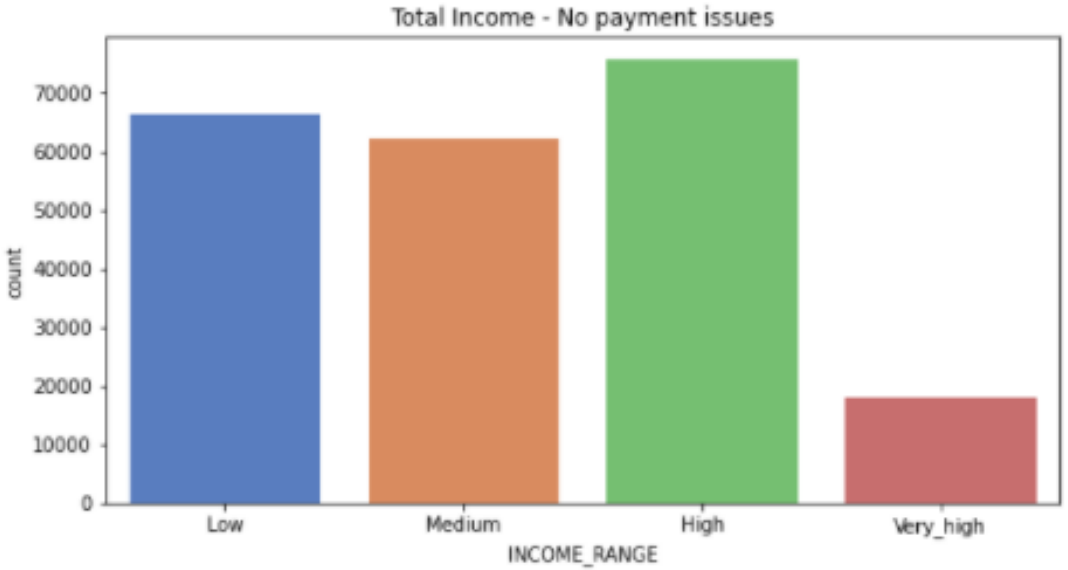
EXPLORATORY DATA ANALYSIS

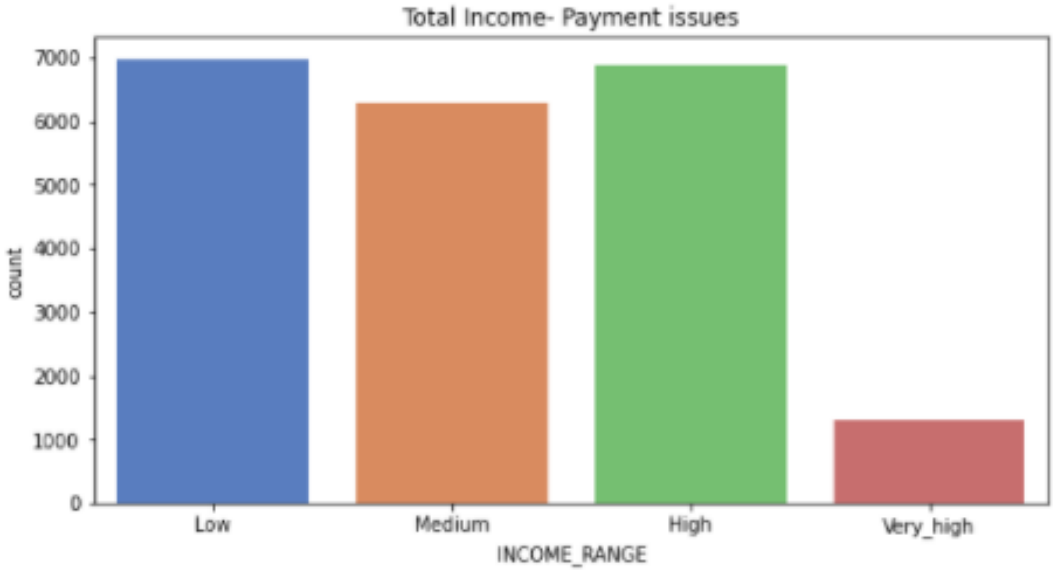




Maximum customers belongs to age group 35-40 are able to make payment on time.

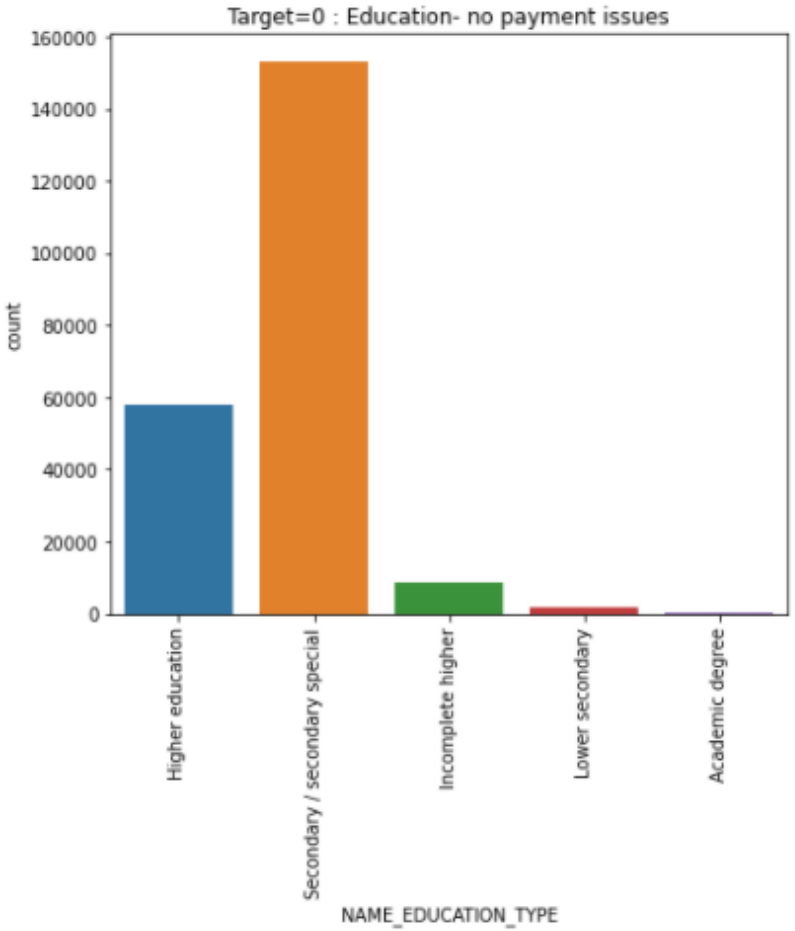
EXPLORATORY DATA ANALYSIS

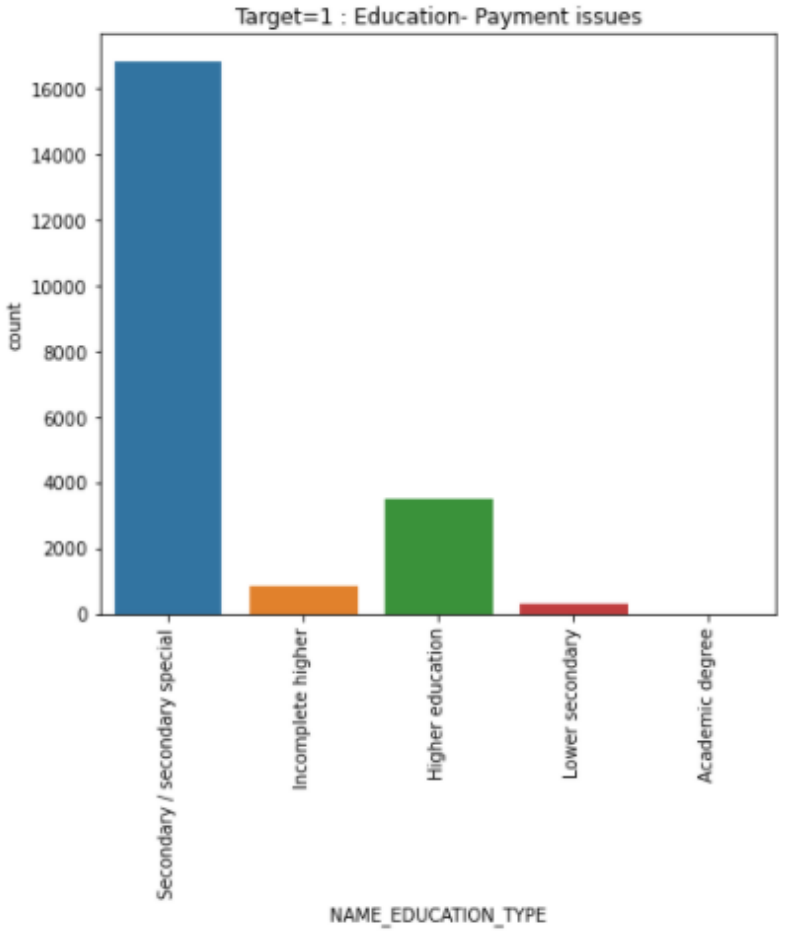




People from Low Income group have payment difficulty on time.

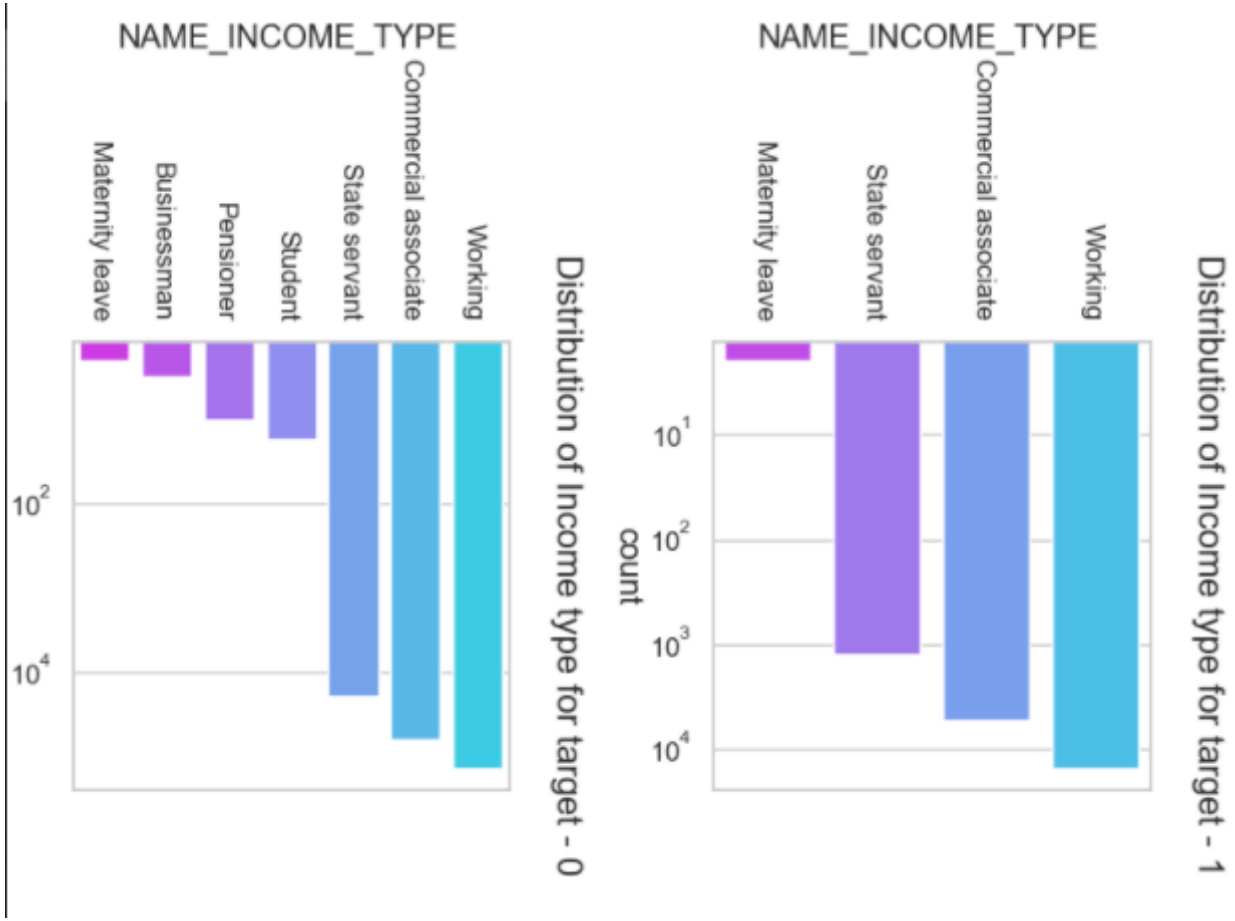
EXPLORATORY DATA ANALYSIS





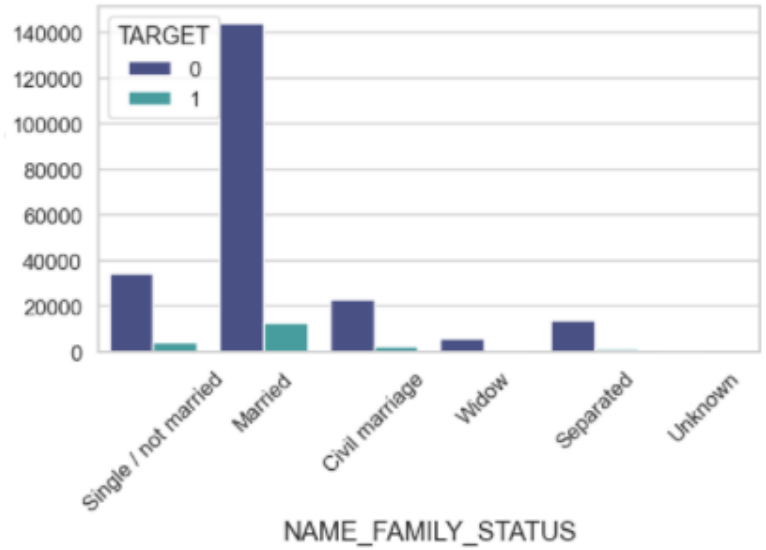
Low Income with Secondary Degree people have payment difficulty.

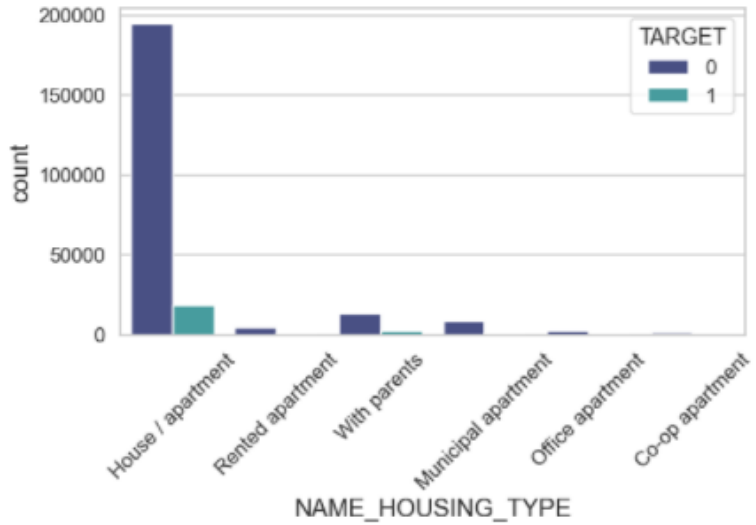
EXPLORATORY DATA ANALYSIS



People belong to WORKING class have maximum payment difficulties.

EXPLORATORY DATA ANALYSIS





Female customers pay loan amount on time and banks can target more female customers for lending loan.

Married customers have paid loan amount on time.

Customers owning House/apartment are most likely to make payments on time.

EXPLORATORY DATA ANALYSIS

EXPLORATORY DATA ANALYSIS

CNT_CHILDREN	CNT_CHILDREN	1.000000
AMT_GOODS_PRICE	AMT_CREDIT	0.999801
REGION_RATING_CLIENT	REGION_RATING_CLIENT_W_CITY	0.998617
CNT_CHILDREN	CNT_FAM_MEMBERS	0.981805
AMT_CREDIT	AMT_ANNUITY	0.970854
AMT_GOODS_PRICE	AMT_ANNUITY	0.970412
AMT_INCOME_TOTAL	AMT_ANNUITY	0.691833
	AMT_GOODS_PRICE	0.626691
AMT_CREDIT	AMT_INCOME_TOTAL	0.623808
DAYS_EMPLOYED	AGE	0.471438
REGION_POPULATION_RELATIVE	AMT_INCOME_TOTAL	0.280997

dtype: float64

Top 10 Correlation from Target - 0

CNT_CHILDREN	CNT_CHILDREN	1.000000
AMT_GOODS_PRICE	AMT_CREDIT	0.999451
REGION_RATING_CLIENT_W_CITY	REGION_RATING_CLIENT	0.998903
CNT_CHILDREN	CNT_FAM_MEMBERS	0.980040
AMT_ANNUITY	AMT_CREDIT	0.965858
	AMT_GOODS_PRICE	0.964497
AMT_INCOME_TOTAL	AMT_ANNUITY	0.646049
AMT_GOODS_PRICE	AMT_INCOME_TOTAL	0.588497
AMT_CREDIT	AMT_INCOME_TOTAL	0.583571
AGE	DAYS_EMPLOYED	0.363670
AMT_GOODS_PRICE	AGE	0.188893

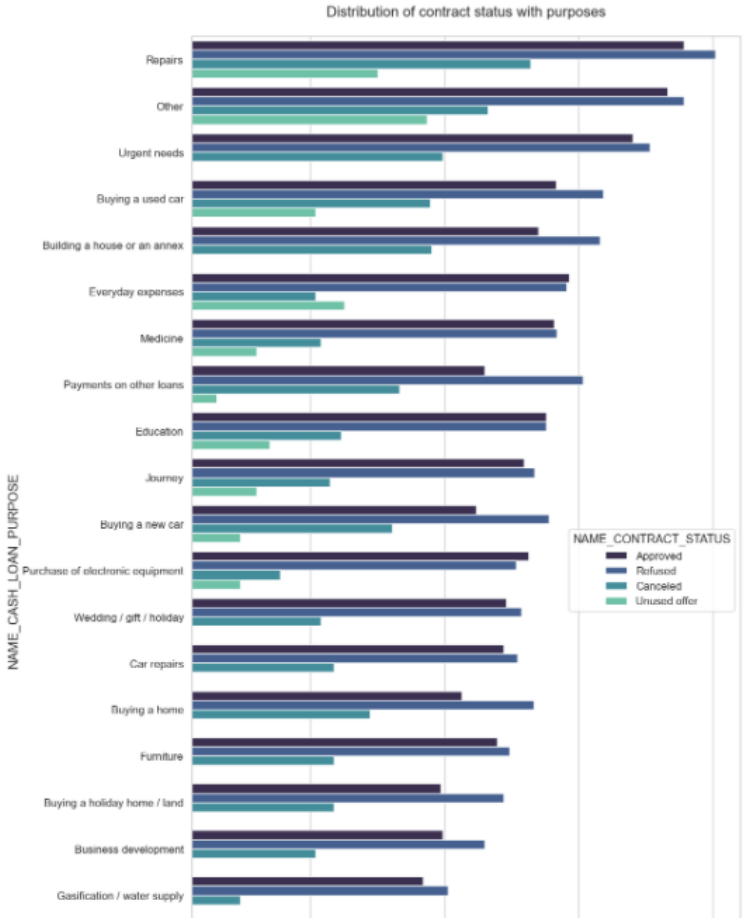
dtype: float64

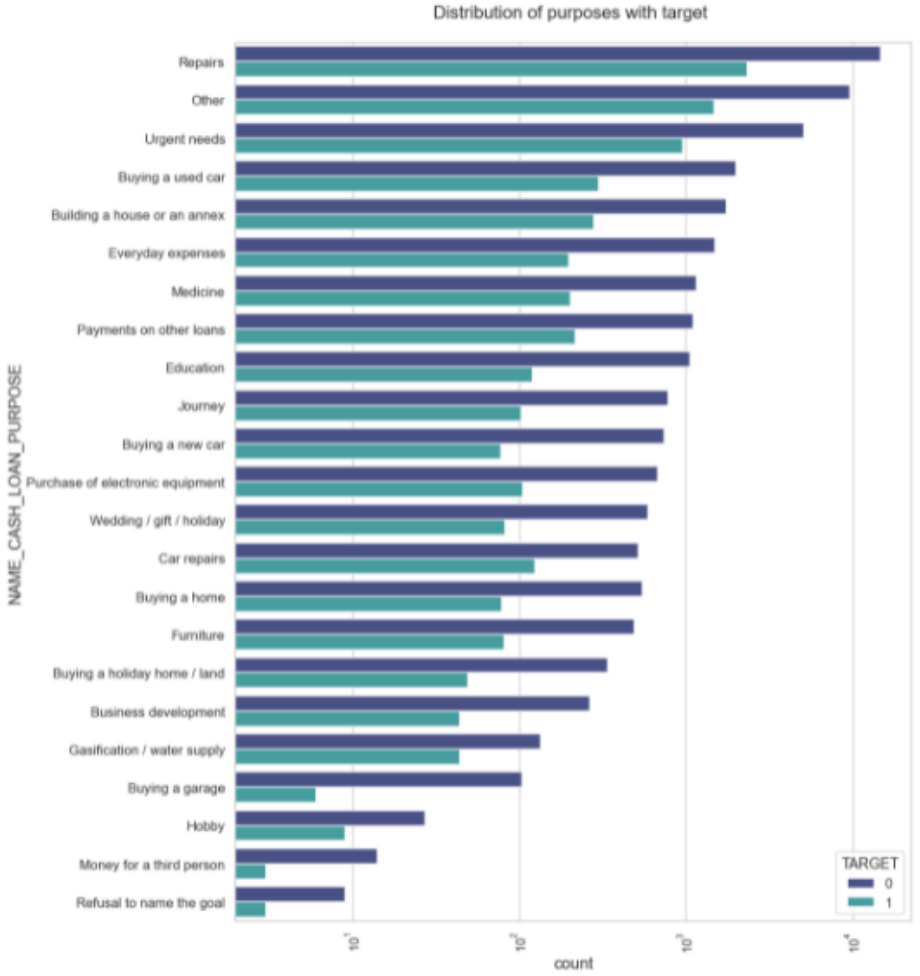
Top 10 Correlation from Target - 1

EXPLORATORY DATA ANALYSIS

Those who have paid the loan amount on/within time are more likely to get higher credits

EXPLORATORY DATA ANALYSIS



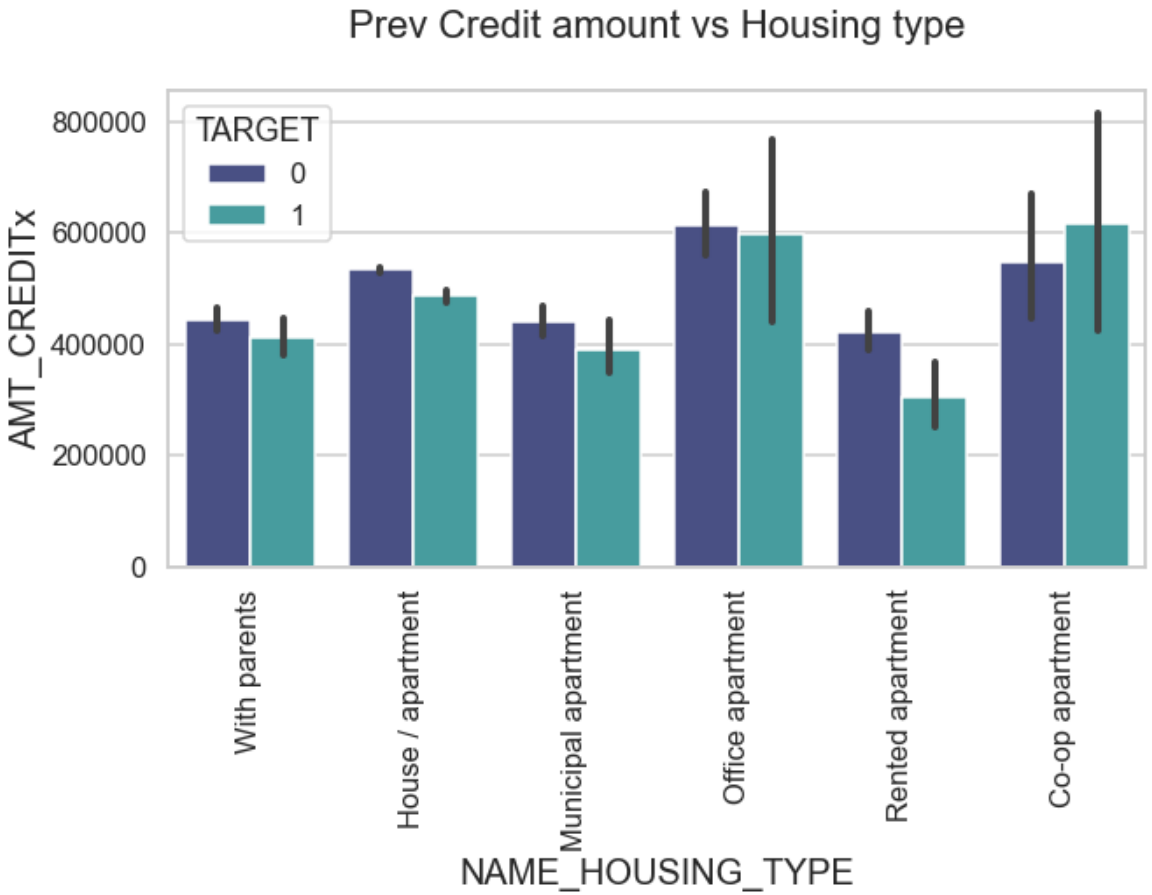


Most rejection of loans came from 'Repairs'.

Rejection rate is high for buying a car, building a house, & Buying a home

Loan given for Repairs purpose having more difficulty in payment on time.

EXPLORATORY DATA ANALYSIS



Here for Housing type, office apartment is having higher credit of target 0 type customer and co-op apartment is having higher credit of target 1 type customer. So, we can conclude that bank should avoid giving loans to the housing type of co-op apartment as they are having difficulties in payment. Bank can focus mostly on housing type with parents or House\ apartment or municipal apartment for successful payments.

KEY POINTS & CONCLUSION

1. Maximum customers belongs to age group 35-40 are able to make payment on time.
2. Low Income range customers have more payment difficulties.
3. Revolving Loan percentage is less for Target - 1 customers.
4. Low income working people have payment difficulties.
5. Female customers pay loan amount on time and banks can target more female customers for lending loan.
6. Married customers have paid loan amount on time.
7. Customers owning House/apartment are most likely to make payments on time.
8. Married with academic degree having higher rate of default.
9. Giving loans to the housing type of co-op apartment is risky.