

EDA CASE STUDY

BANK LOAN ANALYSIS

Akshay Mohan S

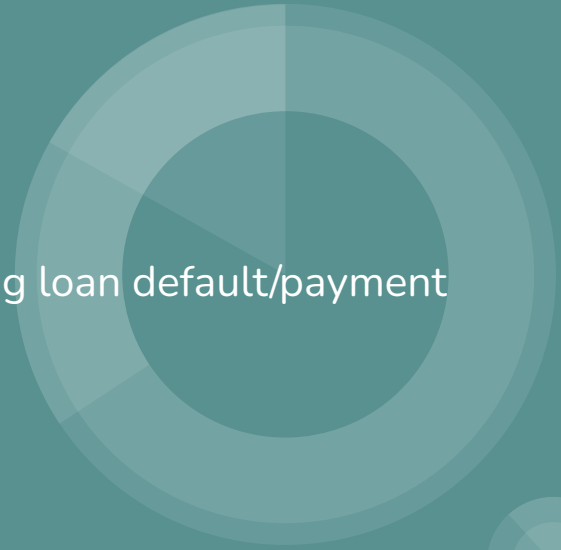


Problem Statement

1. Assignment objective: Application of EDA in a Business Scenario
2. Analyse loan application data to minimise the risk in lending
3. Loan companies facing challenge due to insufficient credit history of applicants
4. Risk Types:
 - Rejecting applicants
 - Approving defaults
5. Identify Patterns indicating payment difficulties or defaults

Problem Statement

- Problem: Identifying critical factors indicating loan default/payment difficulties
- Objective: Improve risk Assessment



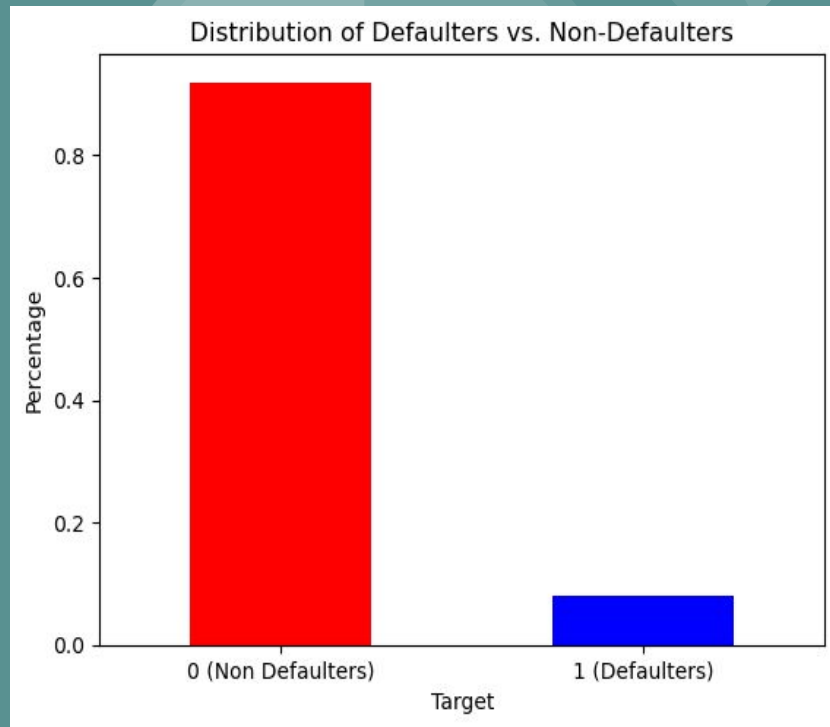
Analysis Steps And Approach

- Understanding Problem Statement
- Exploring the data
- Data Cleaning
- Outlier Analysis
- Data Analysis
- Conclusin

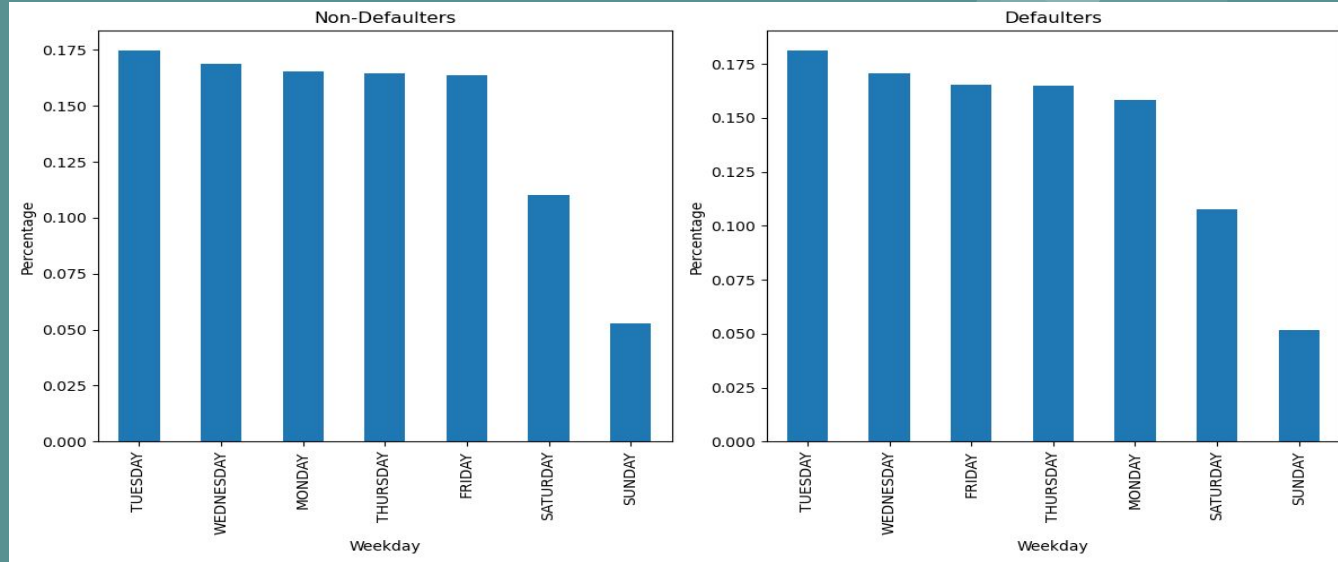


Data Imbalance

- 91.9% are non defaulters
- 8.07% are defaulters
- Non Defaulters vs Defaulters make a ratio of 11.38



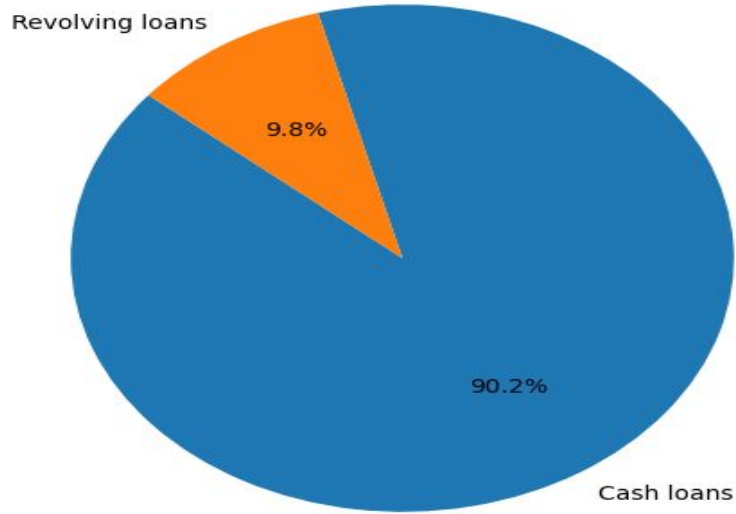
Univariate Analysis



- Application starting processes are comparatively less in weekends

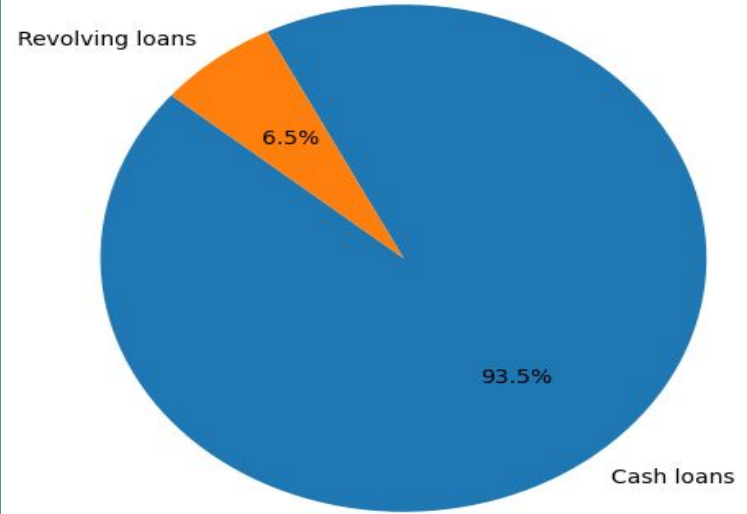
Name Contract Type

Distribution of NAME_CONTRACT_TYPE (TARGET 0)



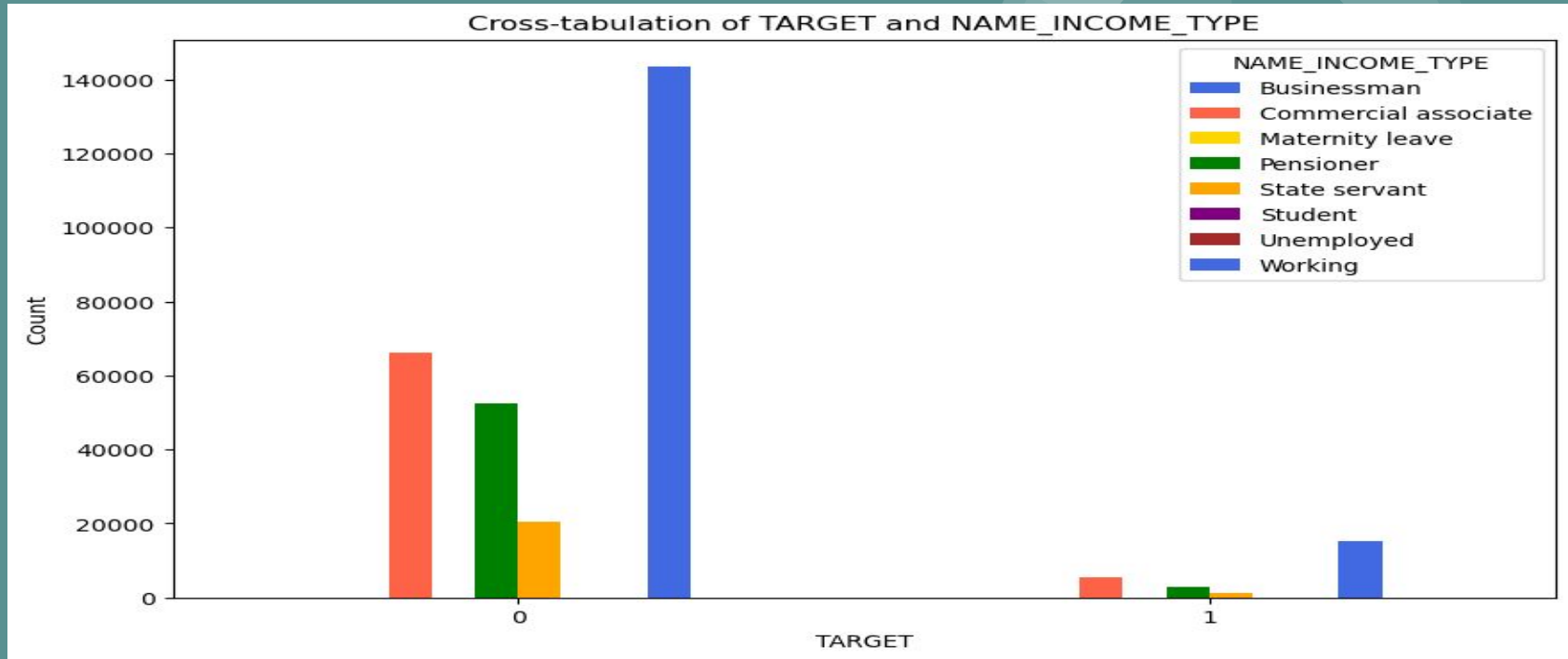
Cash loans were taken by 90.2% of the people!
and only 9.8% people took Revolving Loans

Distribution of NAME_CONTRACT_TYPE (TARGET 1)



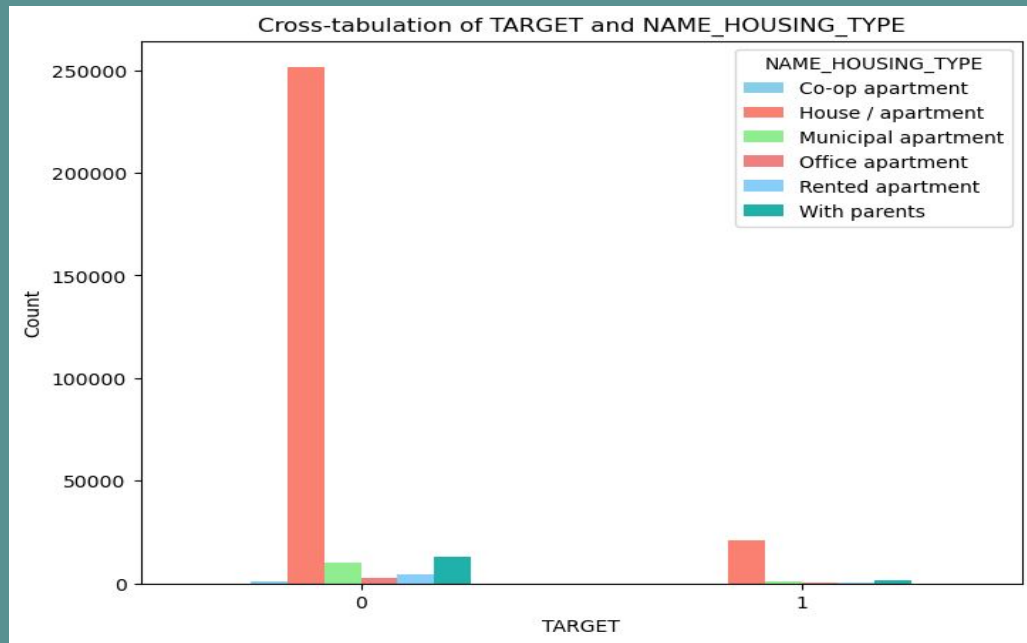
Cash loans were taken by 93.5% and rest
were take by 6.5%

Name Income Type



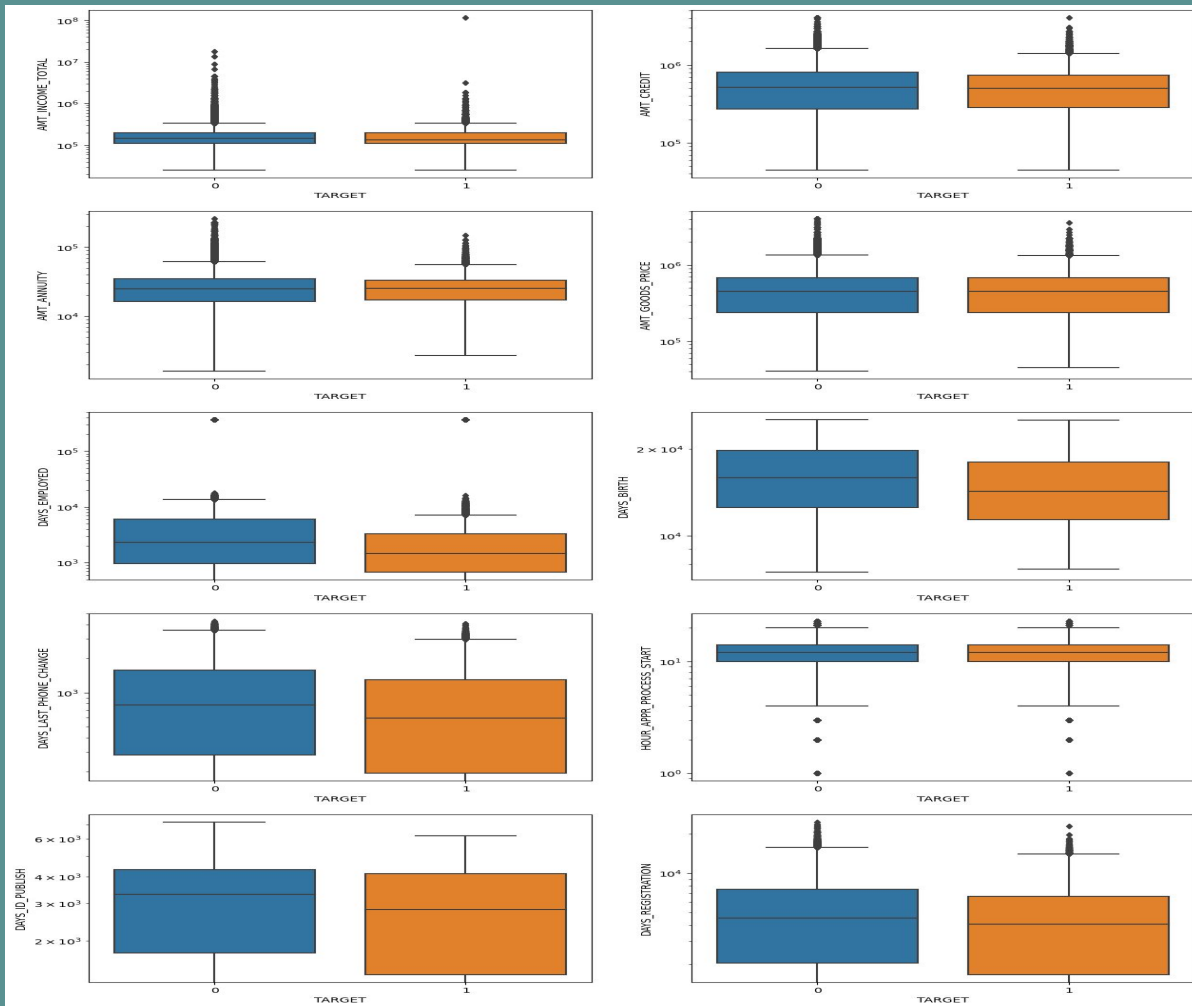
- Most loans are taken by working people

Name Housing Type



- People who owns a house/apartment tends to take more loans

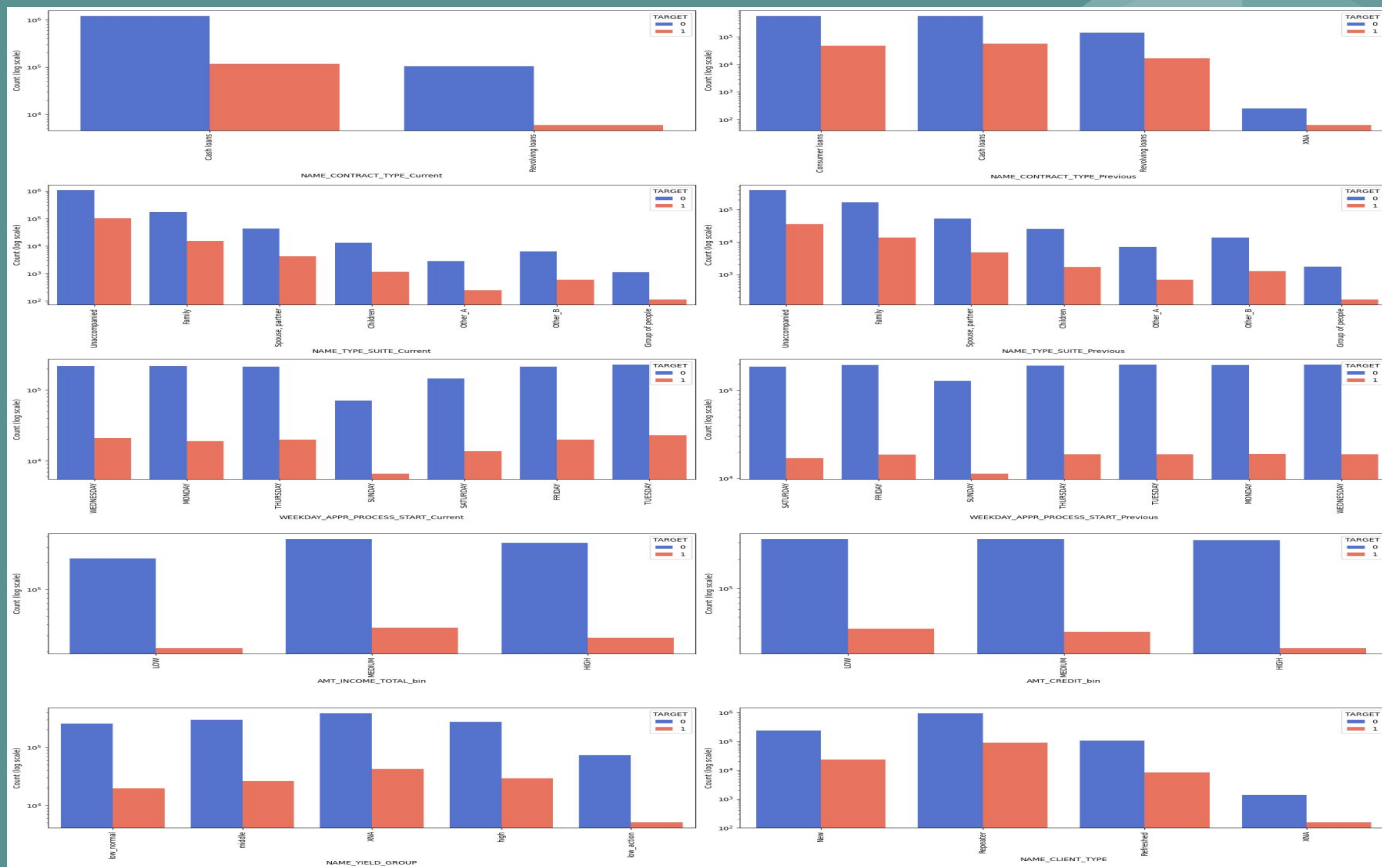
Bi variate continuesplot



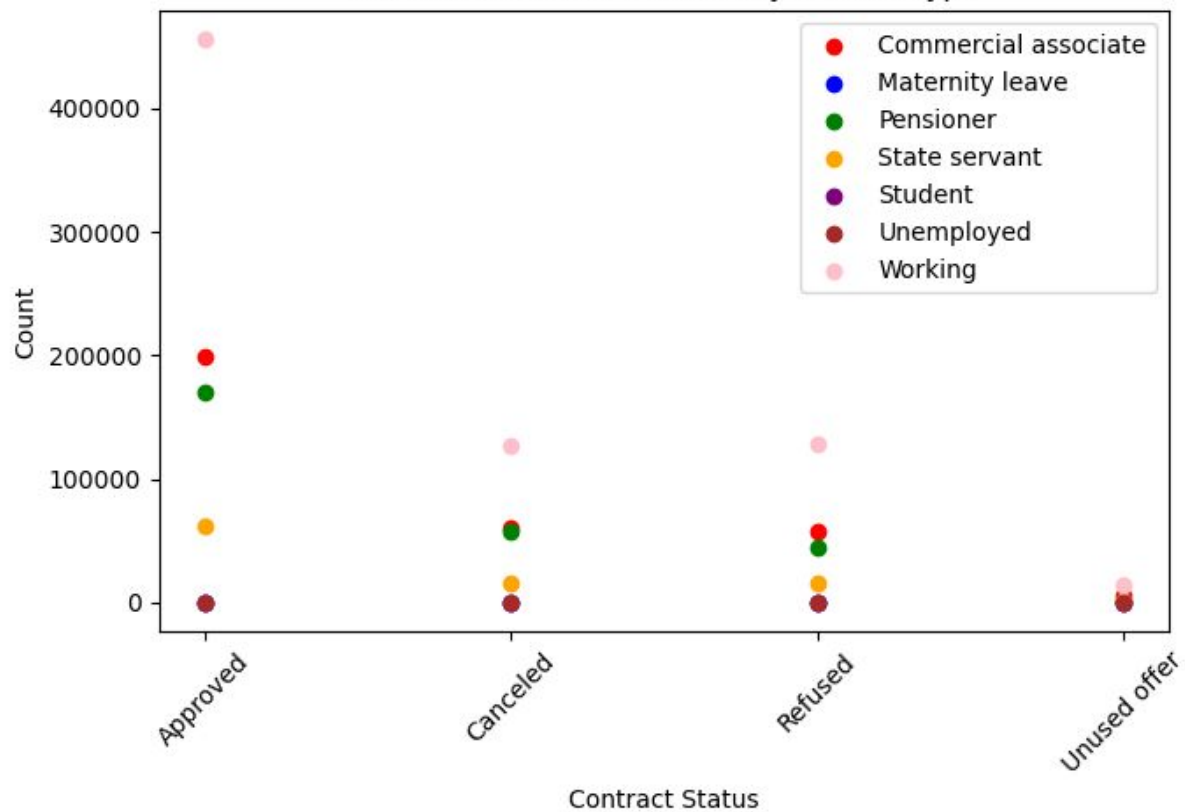
Insights

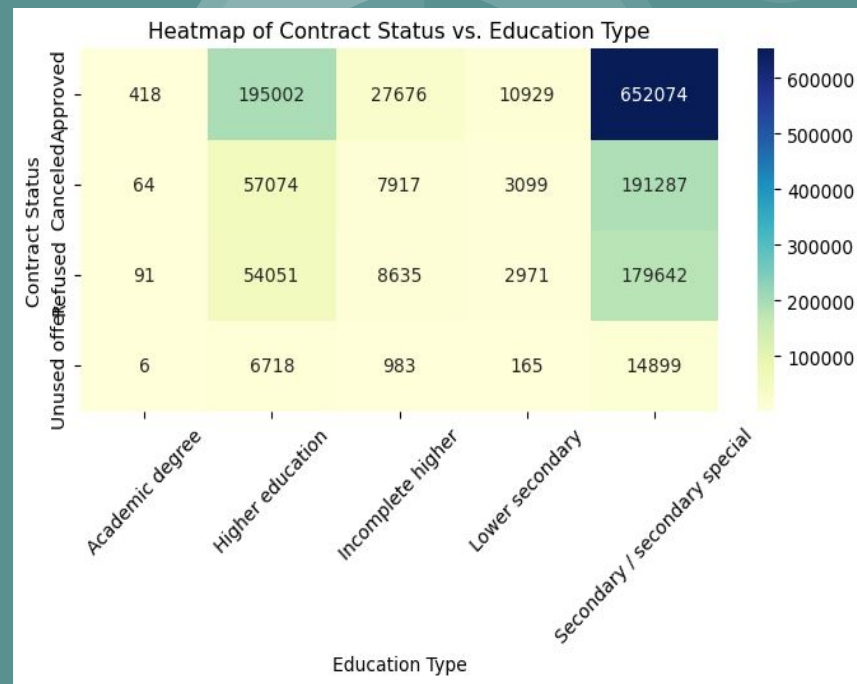
- Most of the people who fails to repay the loan do have a low income
- People with have high number of employment days are able to repay the loan
- For non default cases, they keep their phone numbers for a greater time
- Amt_Income_Total has a very high outlier
- Age colum is proper, as it does not have any outlier

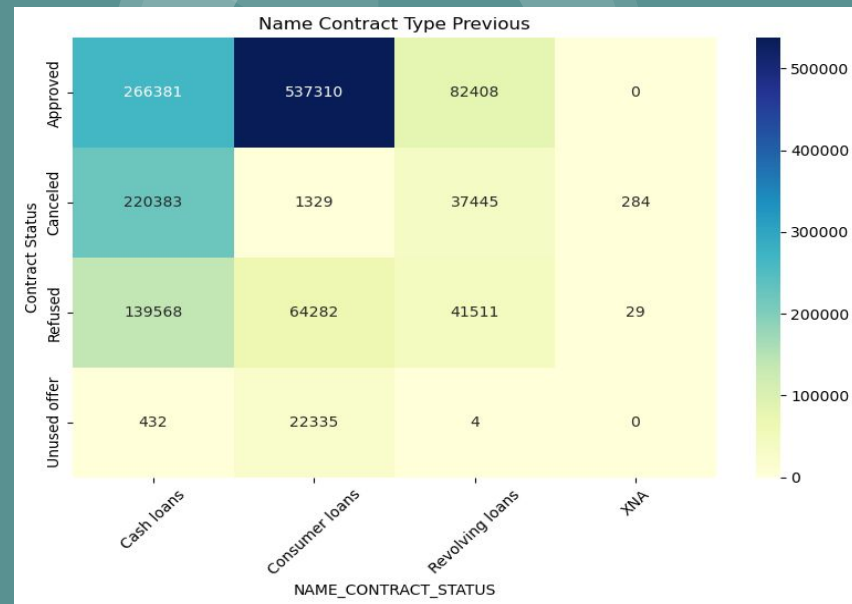
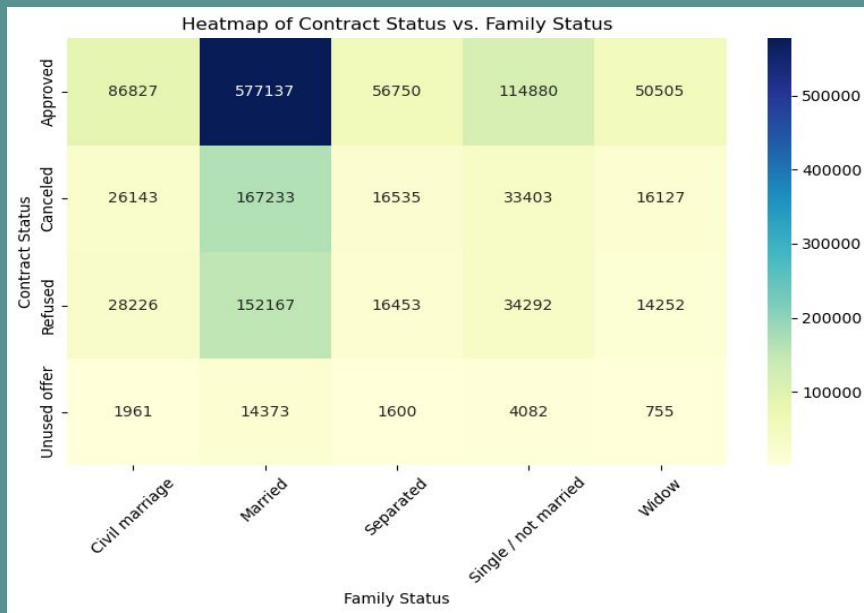
Target Variable Analysis

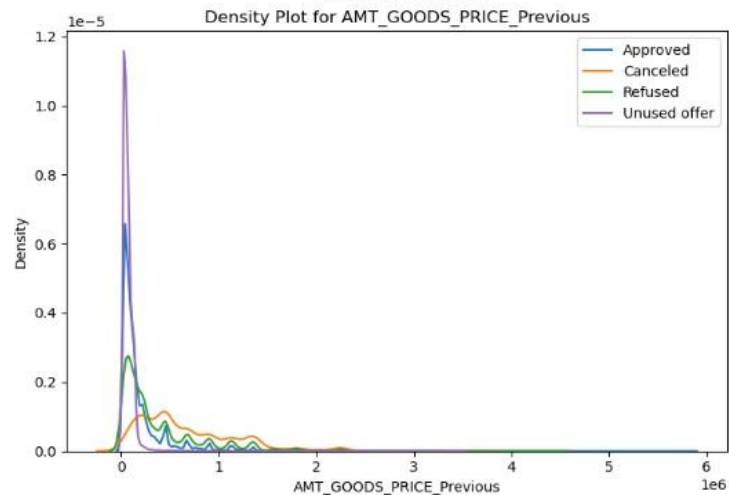
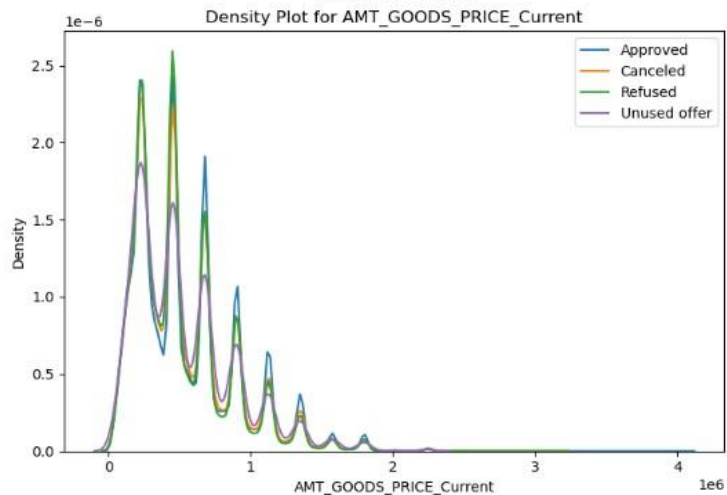
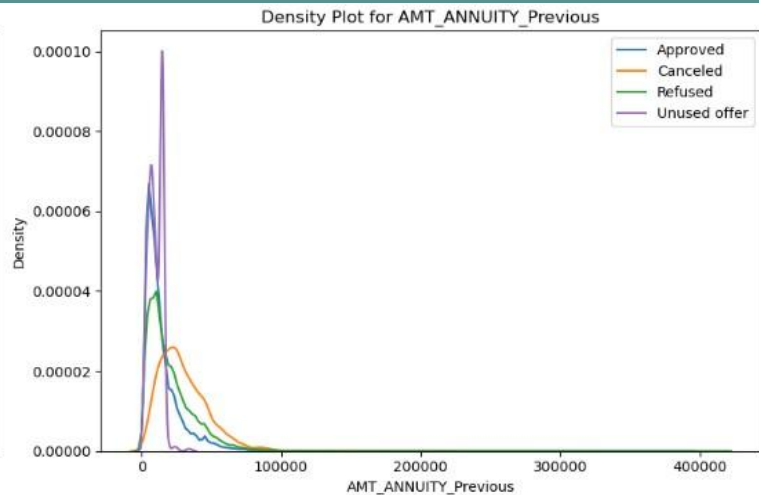
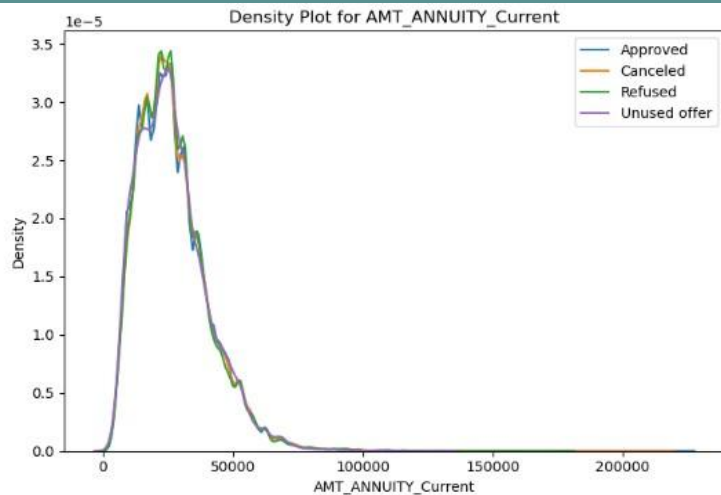


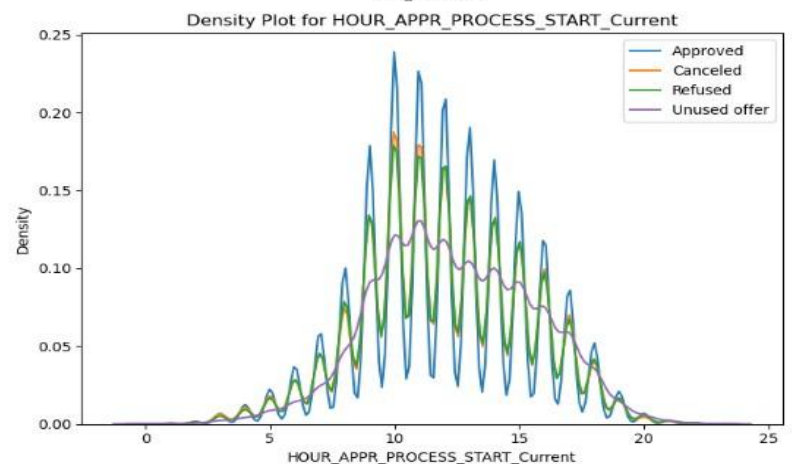
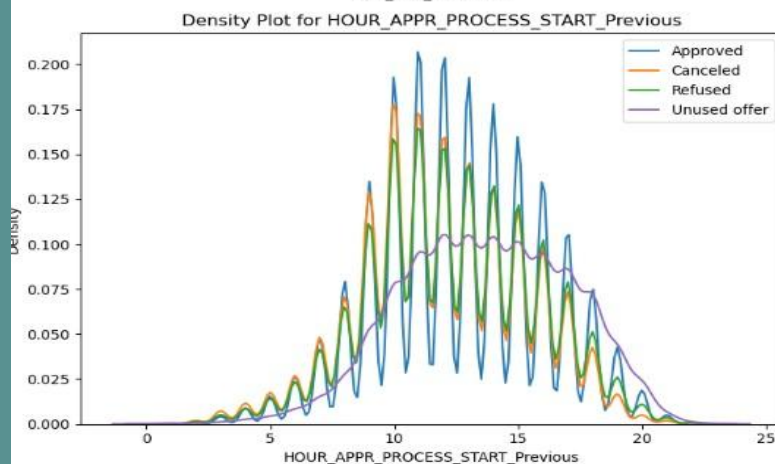
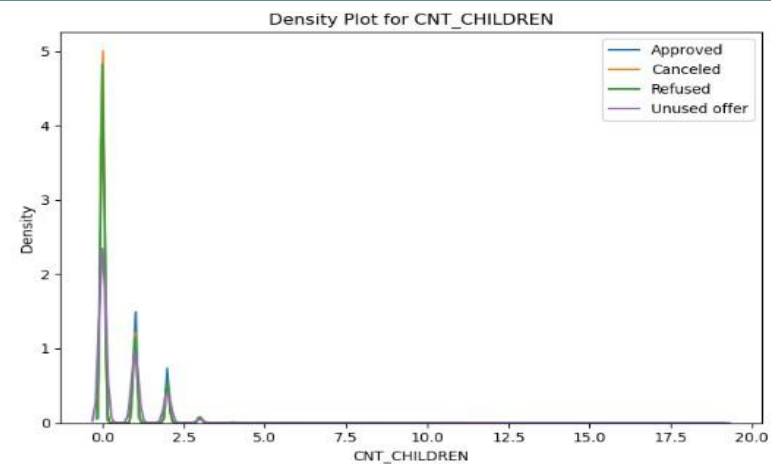
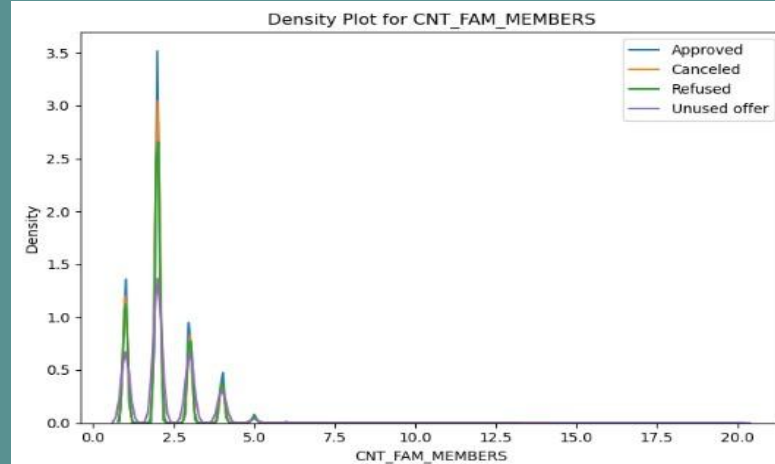
Dot Plot of Contract Status by Income Type

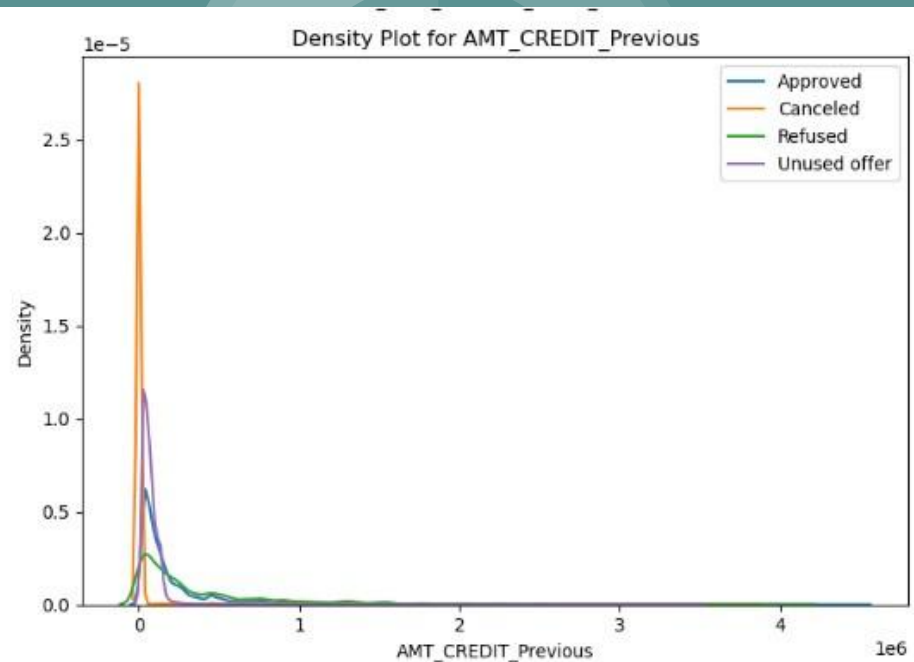
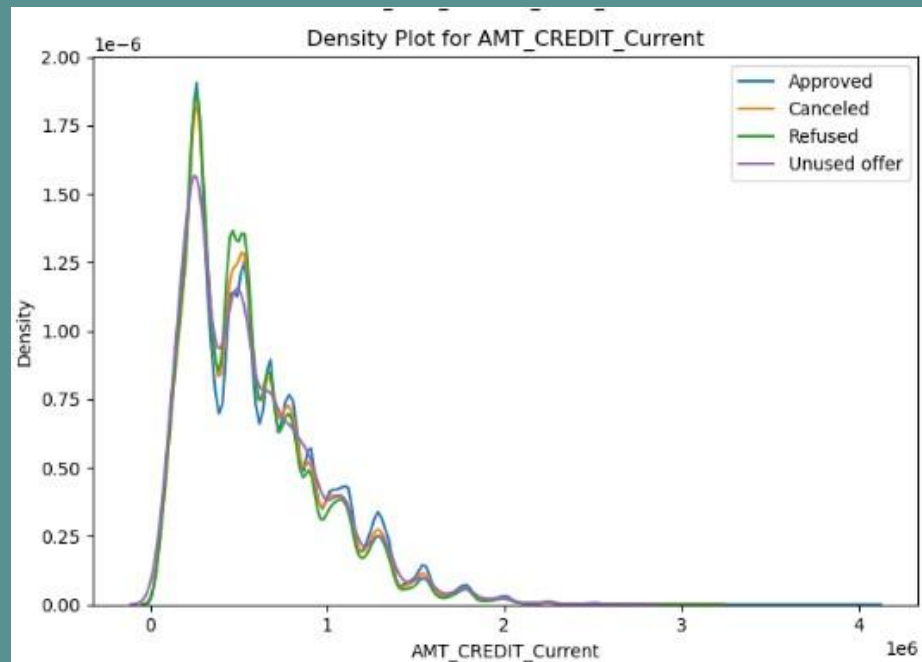












Conclusions

- Students are able to repay loans and are high chance of non defaults
- Clients with high income are able to repay the loans
- Banks must target students and businessmen as they have very less default rate
- Income type with “Working” has the most number of unsuccessful payments.