

Loan Application **Analysis**

ABC Lending Corporation



Contents



3	Problem Statement
4	Assumptions
5	Approach
6	Univariate Analysis

11	Bivariate Analysis
19	Conclusion
20	Suggestions

Problem Statement

To identify the patterns and extract the features that prove to be the driving factors behind default on payments and loans



Assumptions

- The dataset contains a lot of anomalies
- The naming of the columns of dataset is correct
- XAP and XNA are the missing/invalid values
- 1 year = 365 days

Approach

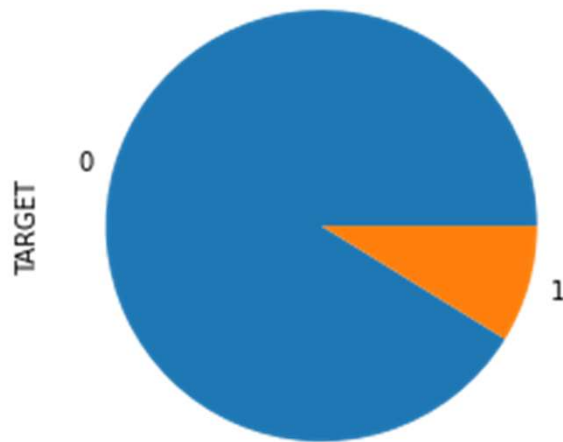
- Understand the dataset, problem statement, aim, and outcome
- Dealing with missing values
 - Removing columns having >40% missing values
 - Removing rows having <5% missing values
- Dealing with invalid values – XAP, XNA
- Sanity check on the dataset
- Standardization of values
- Duplicate record check
- Outlier Analysis
- Univariate Analysis
- Bivariate Analysis



Univariate Analysis

Data Imbalance Check

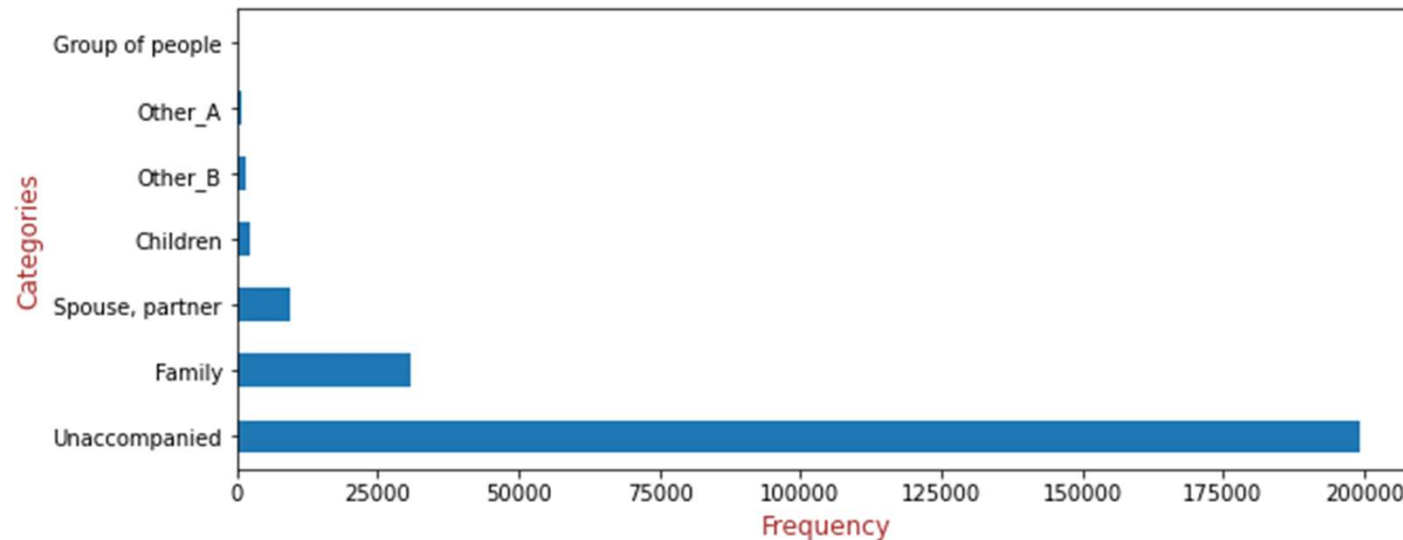
Distribution of TARGET column



Maximum people haven't defaulted on payments or loans taken

Client accompany during application

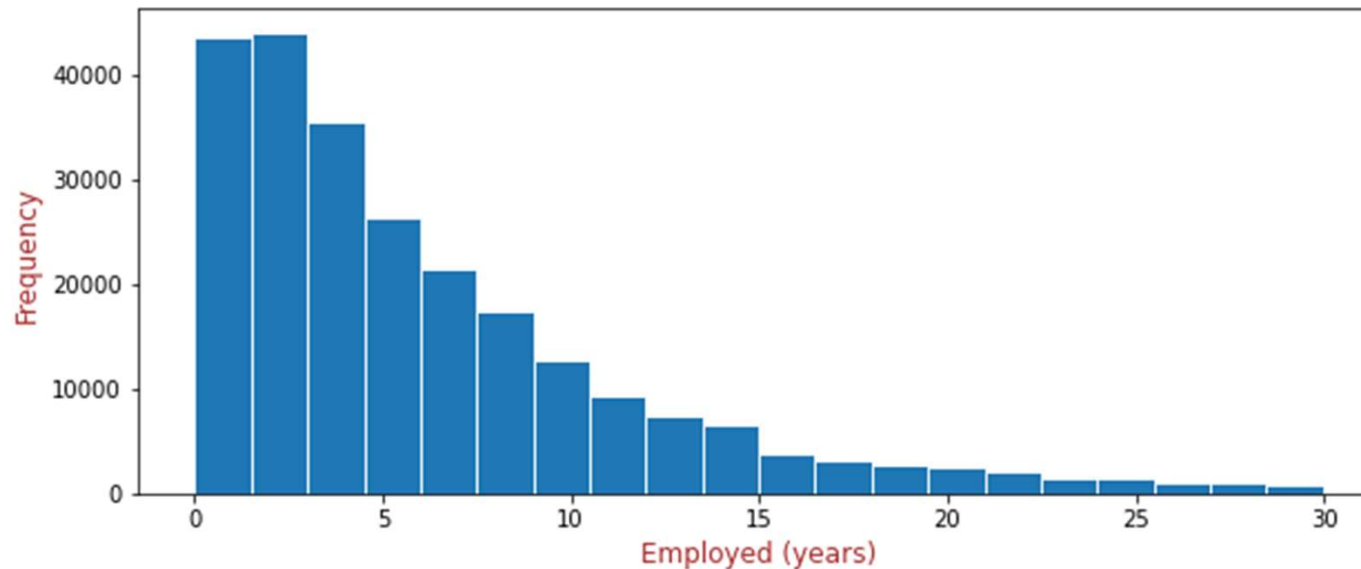
Distribution of who was accompanying the client when applying for the loan



Majority of the loan applicants were unaccompanied by anyone at the time of application

Loan applicants employment years

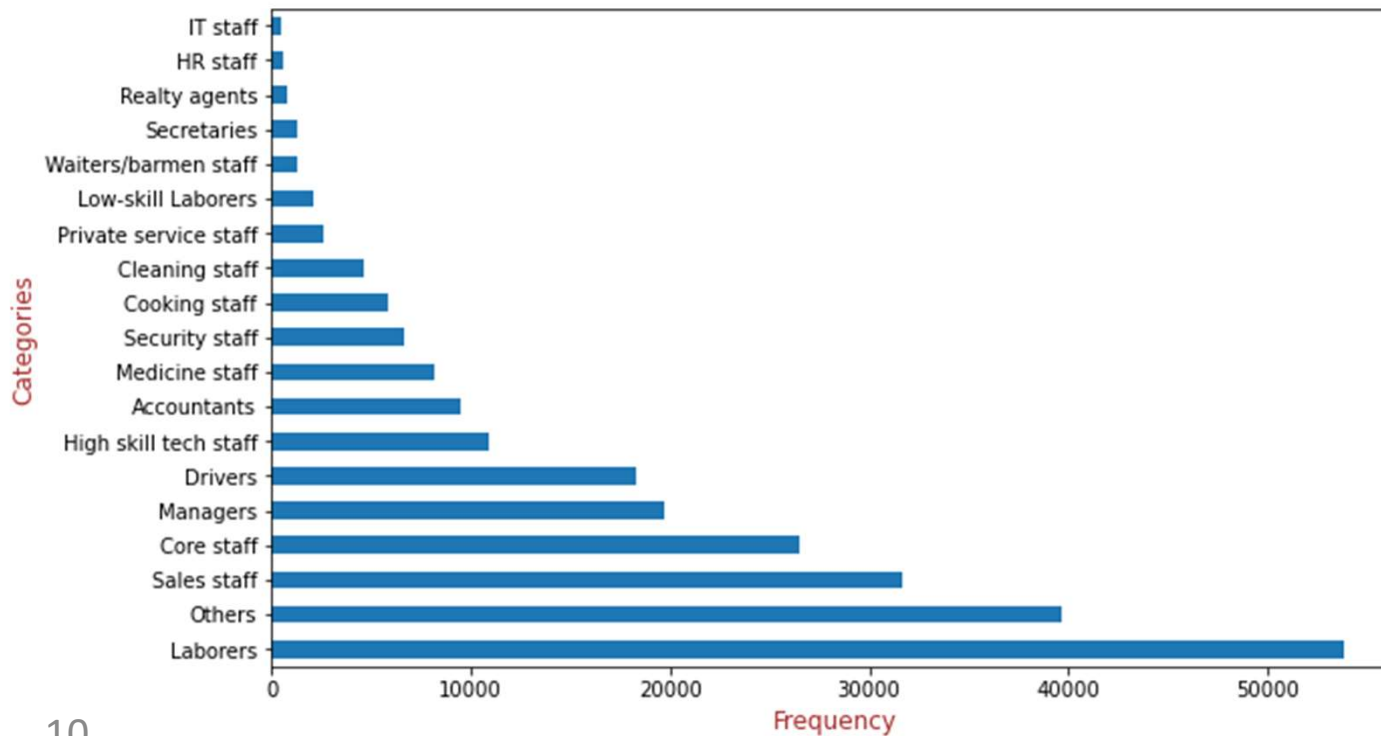
Frequency distribution by employment



Majority of the loan applicants are having work experience less than 5 years

Loan Applicants Job Wise

Loan Applicants Occupation Wise



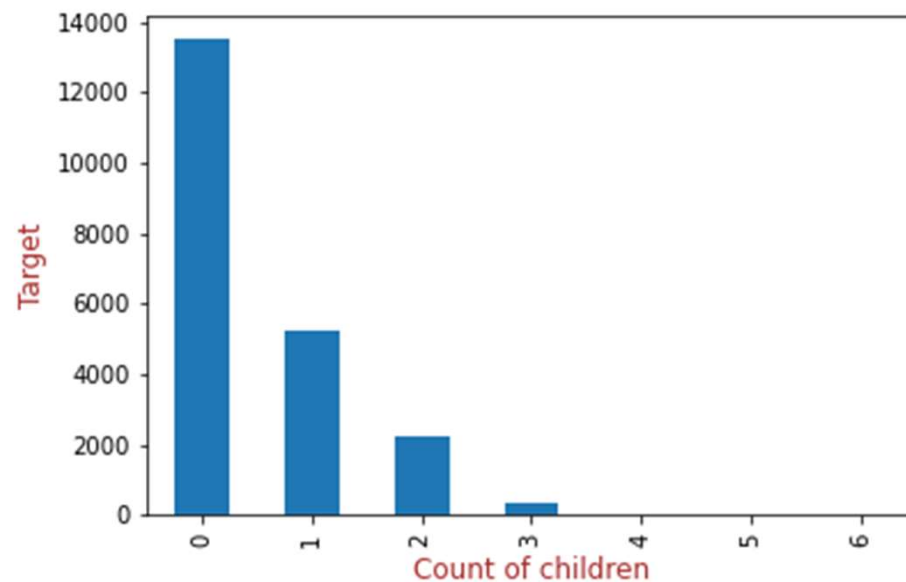
Majority of the loan applicants are laborers



Bivariate Analysis

Children Count v/s Default

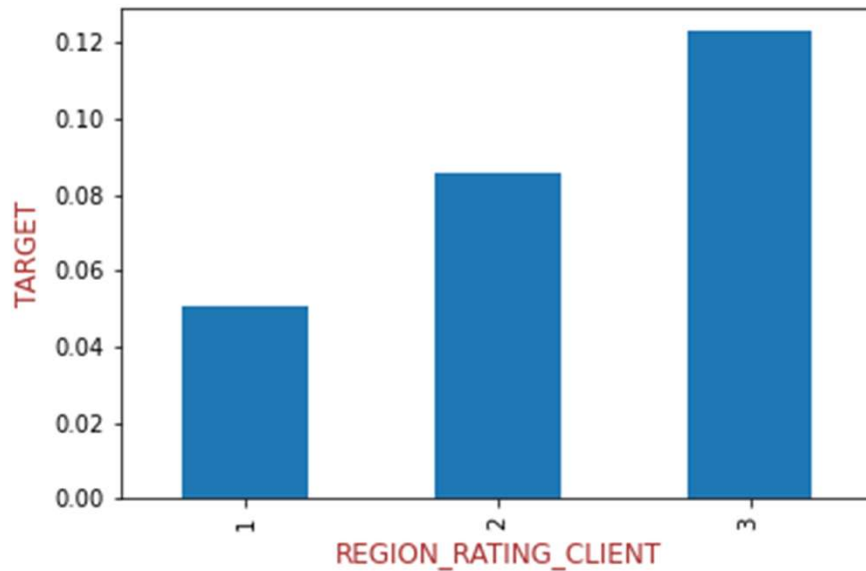
Children Count v/s Target



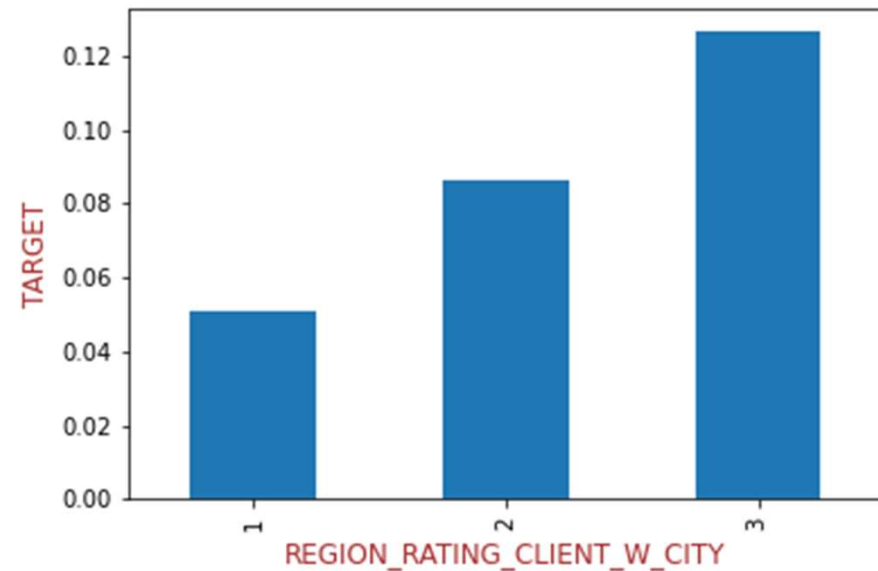
Unmarried loan applicants are likely to default

Regional Rating v/s Default

Region Rating v/s Target

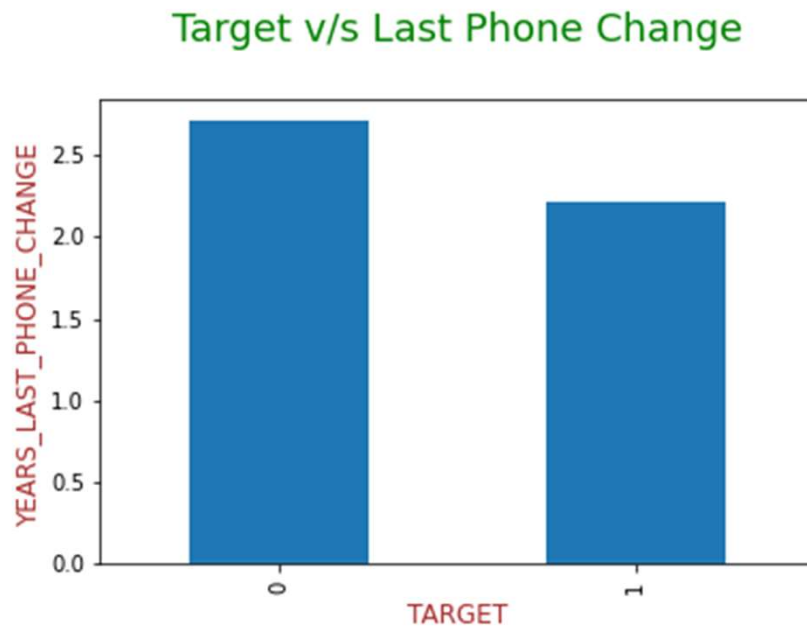


City Rating v/s Target



Loan applicants belonging to regions with rating 3 are likely to default on the loan

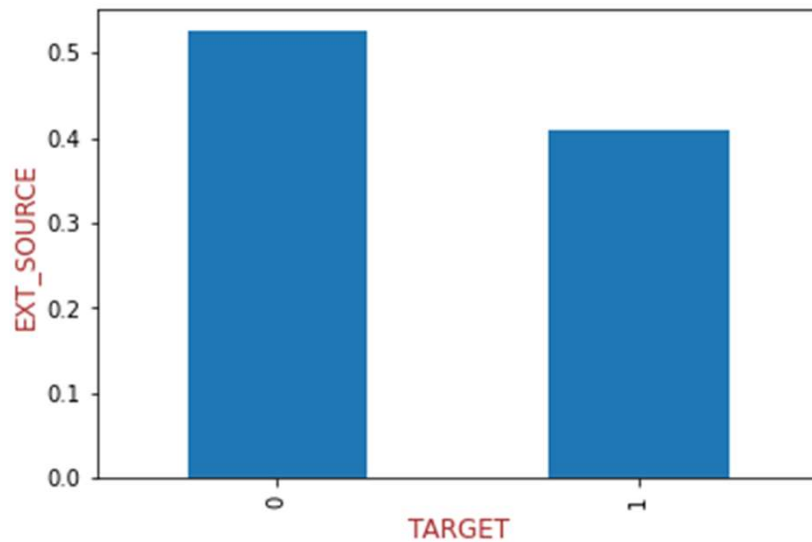
Phone Number Change v/s Default



Loan applicants who have changed their phone numbers before 2 years can be considered as suspicious

External Score v/s Default

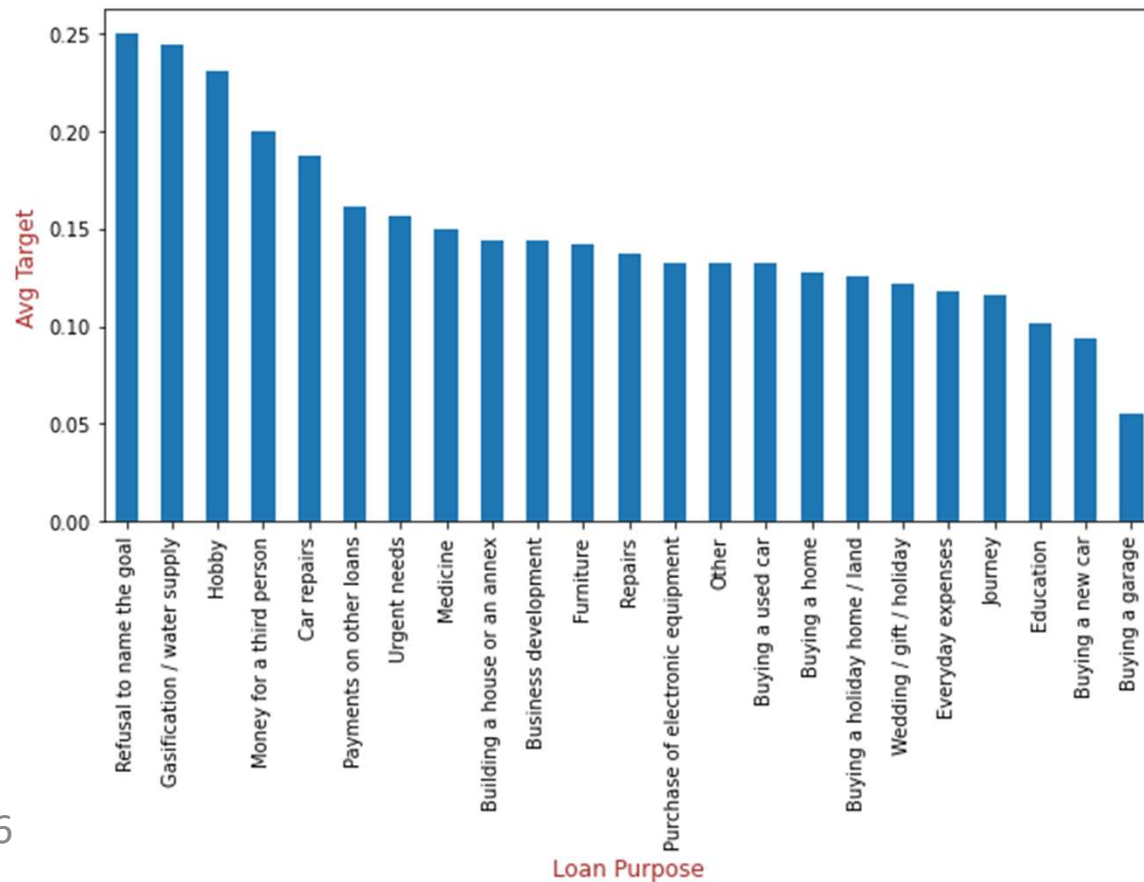
Target v/s Normalized score from external data source



Loan applications whose external normalized score is less than 0.4 have defaulted often

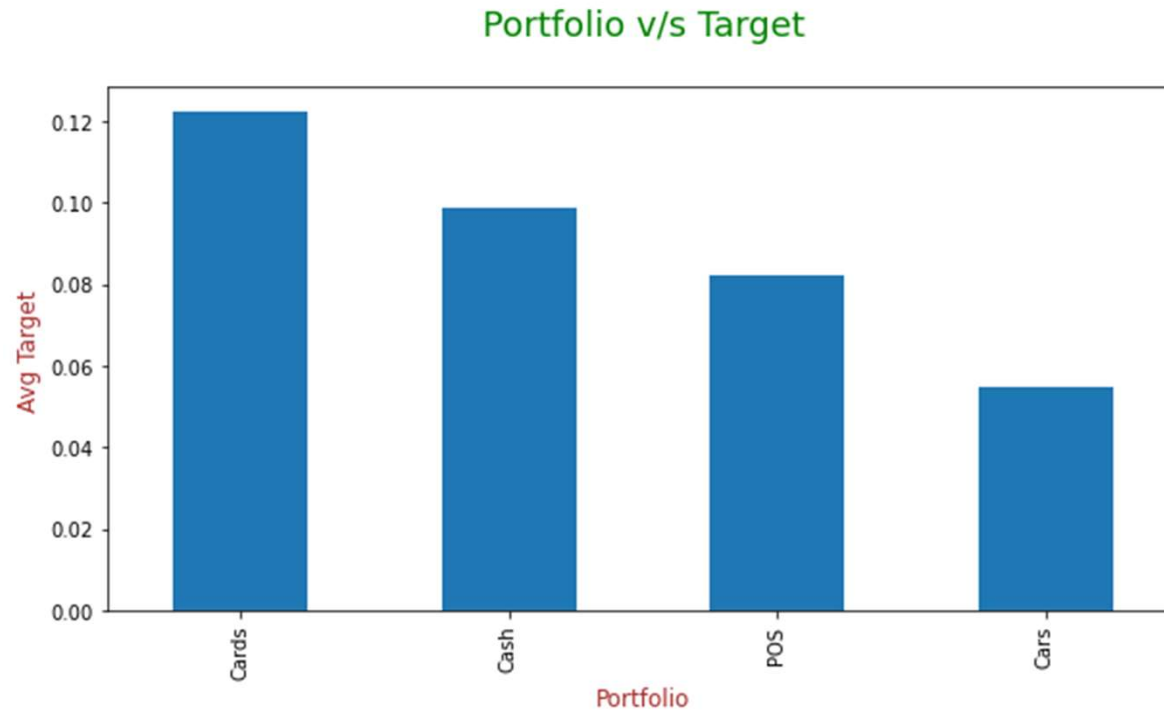
Loan Purpose v/s Default

Loan Purpose v/s Target



Loan applicants who refused to provide the purpose of the loan mostly defaulted

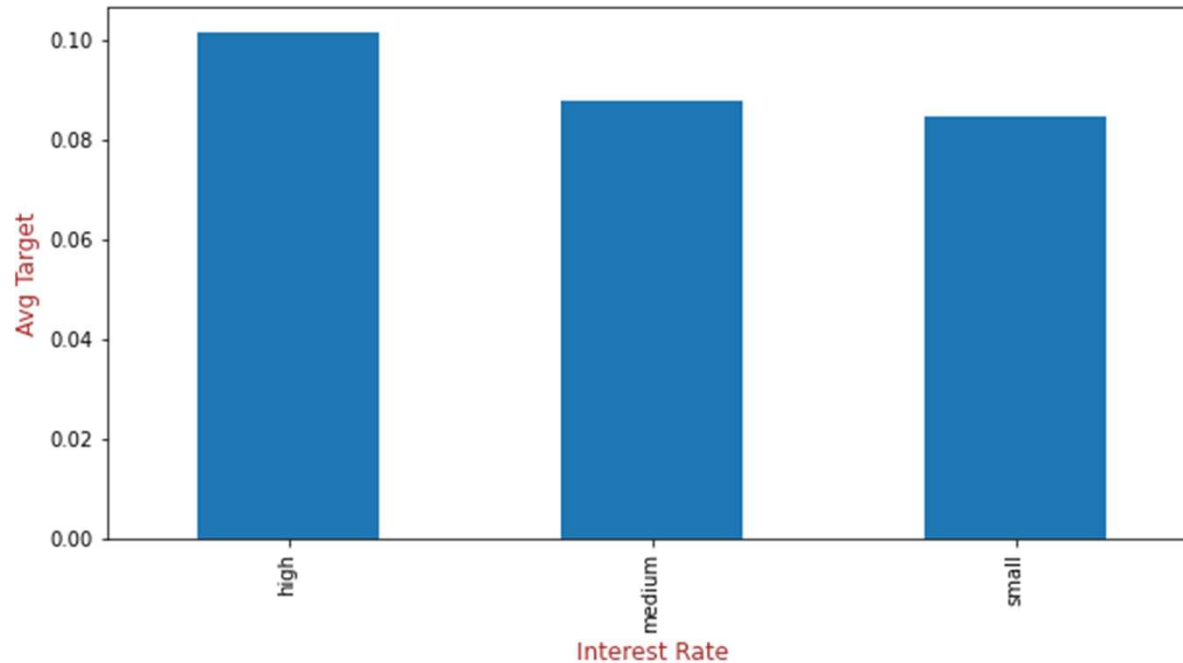
Portfolio v/s Default



Cards category contains the highest default loan cases

Interest Rate v/s Default

Grouped Interest Rate v/s Target



**Higher the interest rate
higher the default chances**

Conclusion

- Overall count of children and number of family members should be less than 8 or 4
- Loan must be disbursed to the client residing in category 1 or 2
- Suspicious clients:
 - Phone number changed in the last 2 years
 - External score less than 0.4
 - Clients refusing to provide the details for the purpose of loan
- Loan type – cards have the highest default
- Higher interest rate increases the chances of default

Suggestions

- Even if the applicant previously applied for loan it is necessary to check the credit history
- Ensure to check the observation of client 30DPD and 60DPD

Thank You

