FINANCIAL INSTITUTION
RISK ANALYSIS IN LENDING LOANS

Situation

- Financial institutions finds it hard to lend money to people due to lack of their credit history
- Worst case happen when institutions deny the application from a potential client or lend money to a defaulter.
- In both above case the institution loss the business .
- By using EDA on Data provided by the financial institution we need to find pattern to analyse the risk in lending loans.

- Data provided includes

1. Application data: Various details of the applicant, ‘TARGET ‘ variable indicating applicants status in repaying the loan
2. .Previous data: Details of the loans lend by the financial institution.
3. Column Description: Describing the details of each variables.

- On a particular application 4 decisions are made

1. Approved : Approved by institution
2. Cancelled : Cancelled by the client
3. Refused : Rejected by institution
4. Unused offer **:** Client don’t need

ASSUMPTIONS

- TARGET column in application Data ,combined with other variables will helps to identify the applicants with risk behaviour.
- While conducting univariate , bivariate or multivariate analysis on the top of TARGET variable we can identify the loops for patterns for loan defaulters.
- Previous Data will gives the history of applicants like approved, refused, cancelled etc. and helps to reduce the risk in dealing an existing client.
- By merging the two data frame ,we can analyse the risk in already existing clients those who are applied for the current loans.

Approach

- Study both the files , understand each of the variables and problem statement.
- Create separate data frame by uploading the files Application Data and Previous Data also a merged data frame .
- Go through each of the data frame, columns, note down the properties like shape, info, describe, dtypes, columns etc.
- Clean the data by finding out the missing values, imputting the missing values, or by removing the columns or rows having missing value.
- Also handle outliers if there are any, Standardising the values of columns ,finding out the Data imbalance percentage

- After cleaning the data and standardising the values of the data identify the categorical and numerical columns to perform analysis.
- Based on the nature of column perform univariate , bivariate or multivariate analysis on variables which seems to be of relevance.
- Infer the details obtained through the analysis and note down the patterns followed or correlations between variable.
- Note down the pattern or correlation that clearly indicating the chances that a applicant will default the loan.

Graphs and Inferences

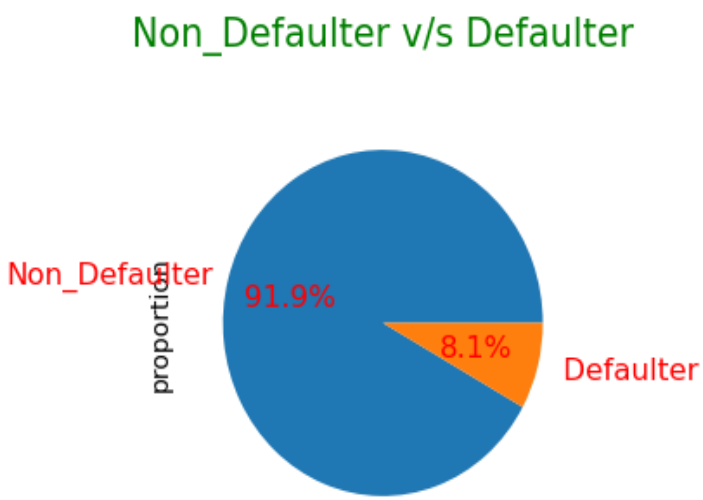


Fig-1

Among the applicants 92% are non_defaulters and only 8% are defaulters. This ensures a secured buisness for the institition.

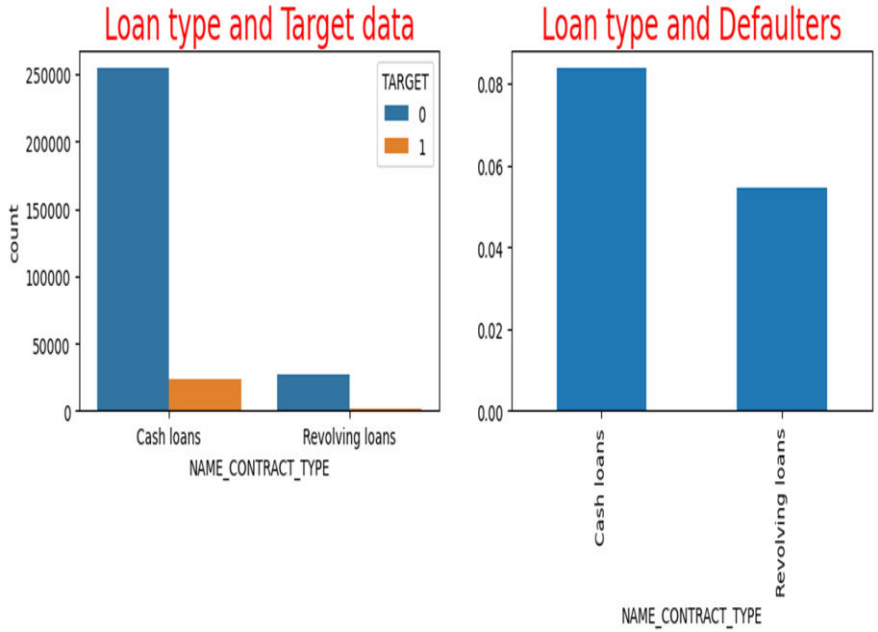


fig2-

Majority of applicants prefer cash loans and those who default loans they also choose cash loans

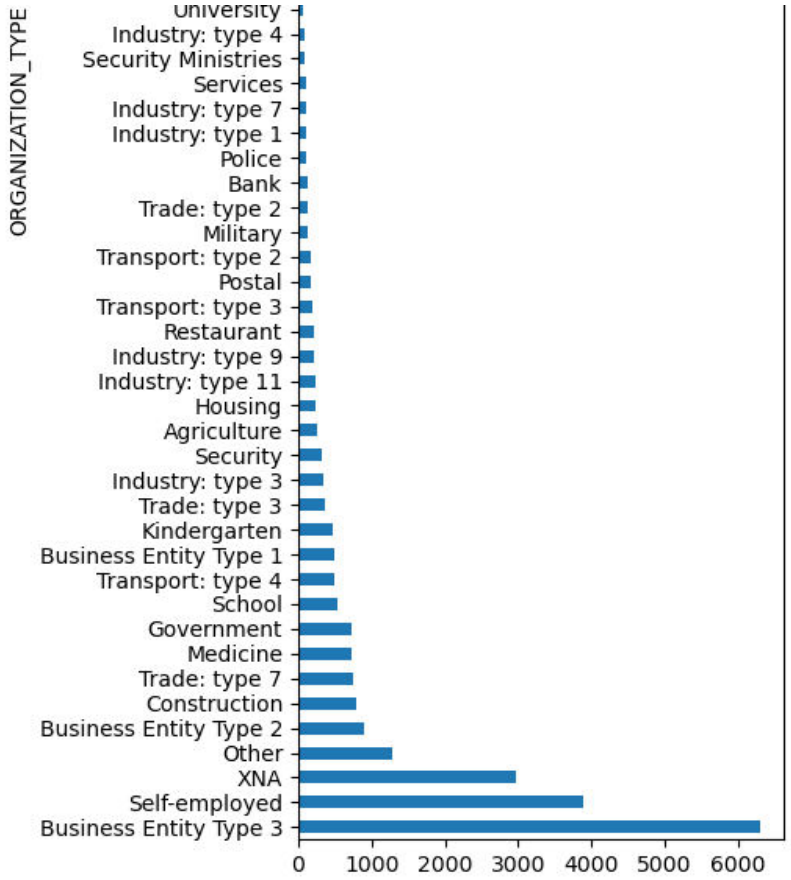


Fig-3

It is clearly seen that those who belongs to 'Business Entity Type 3' in variable 'ORGANIZATION_TYPE' default loan more than any other type

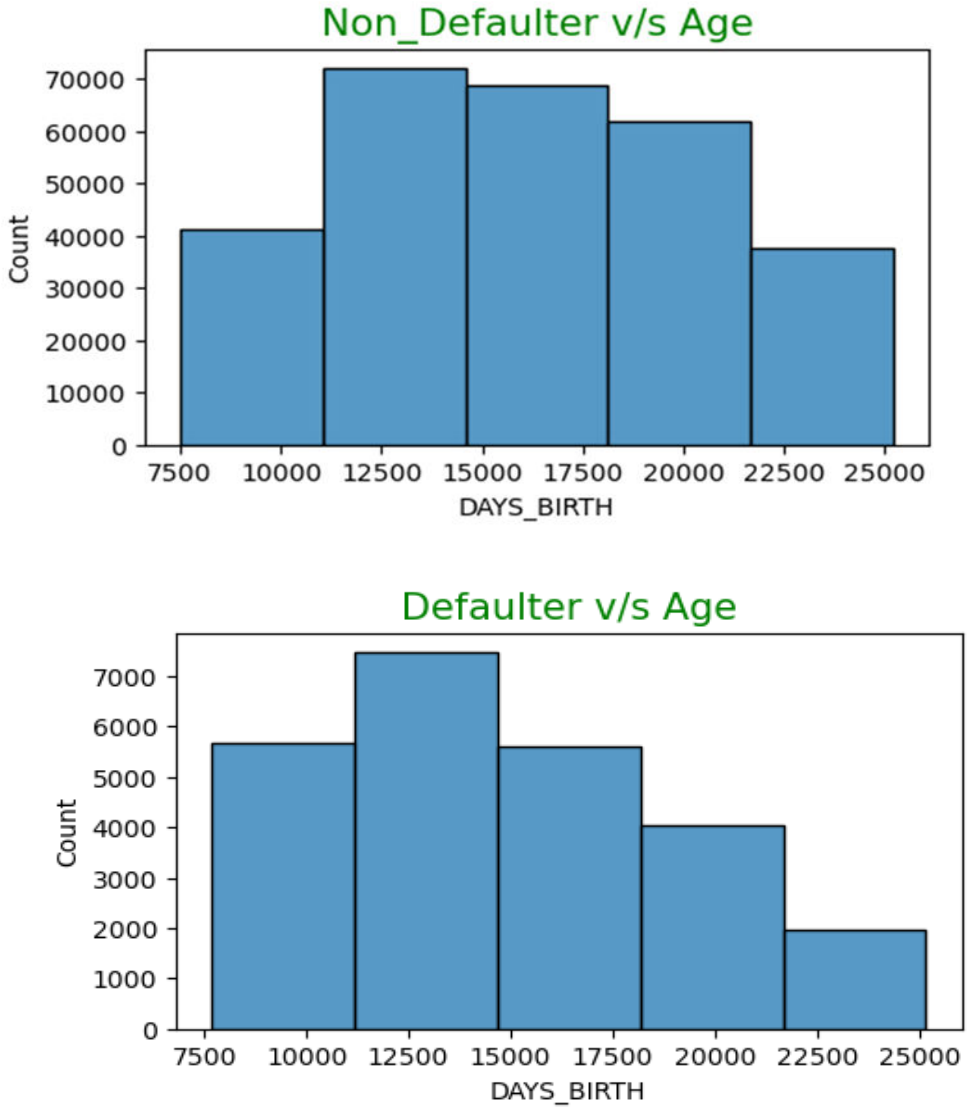
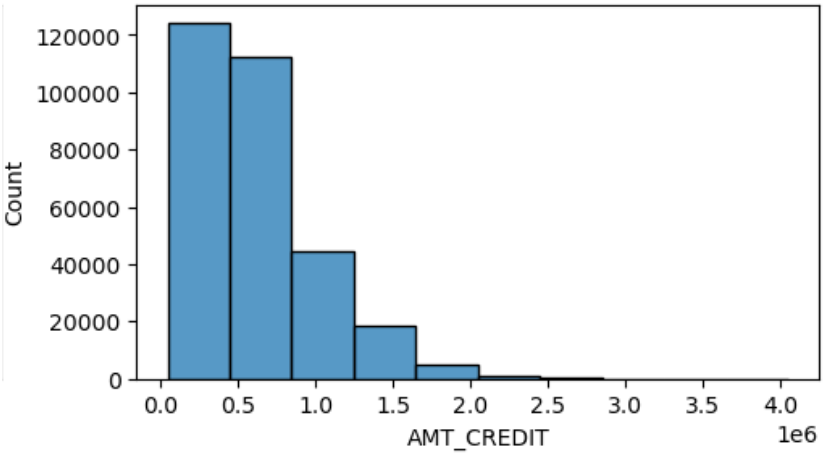


Fig-4

Younger applicants to default loans more than elder applicants.



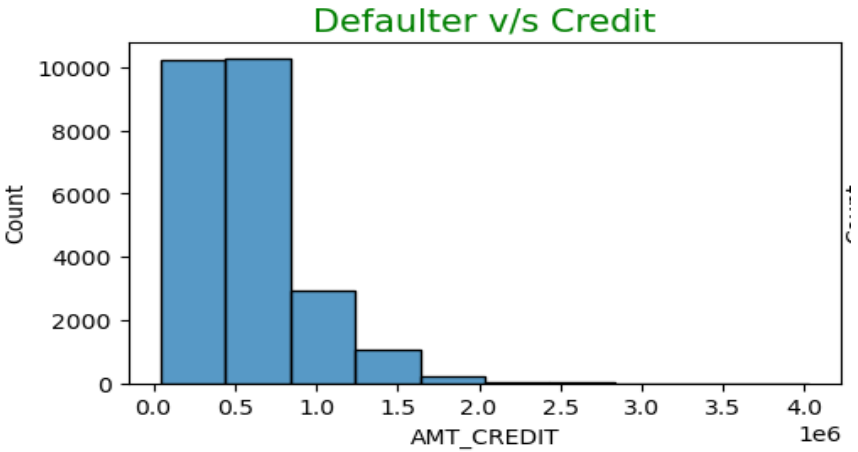
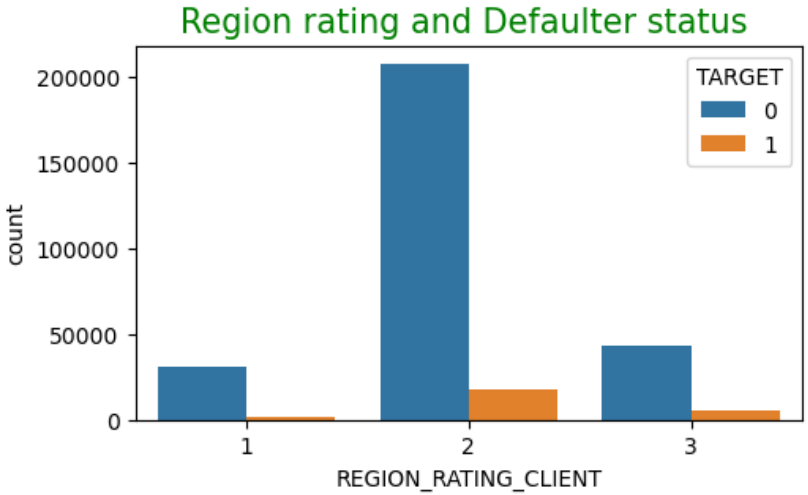


Fig-5

Majority of applicants look for low credit amount, also approximatly 8-9 percentage of total count in each credit bucket makes default payments



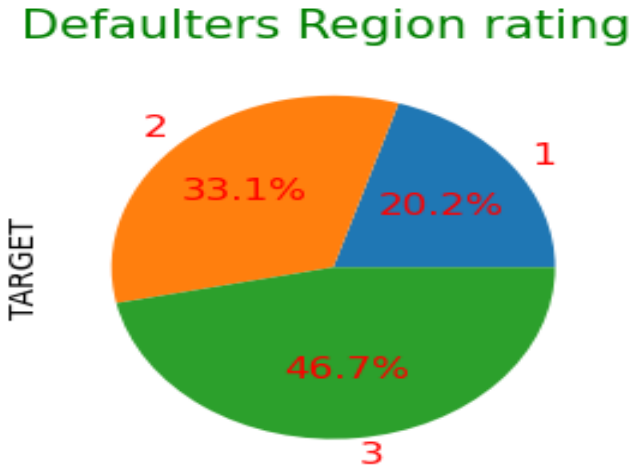
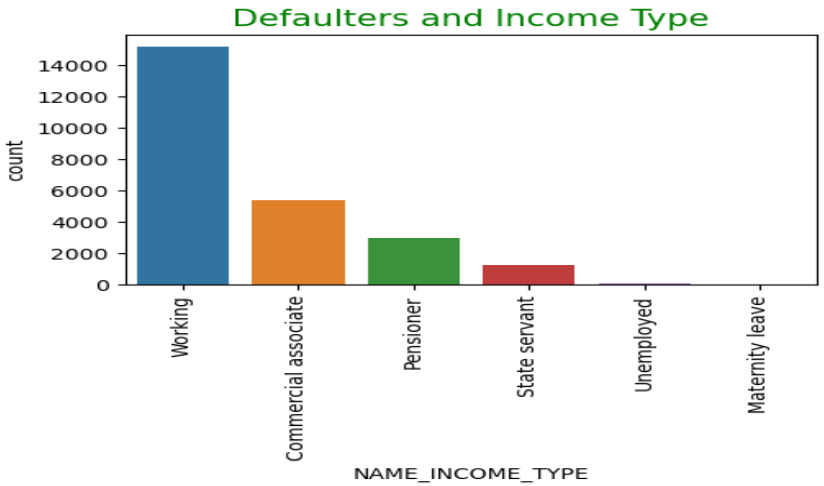


Fig-6

Majority of applicants leave in region-2 , but the mean rate of default is high for region 3



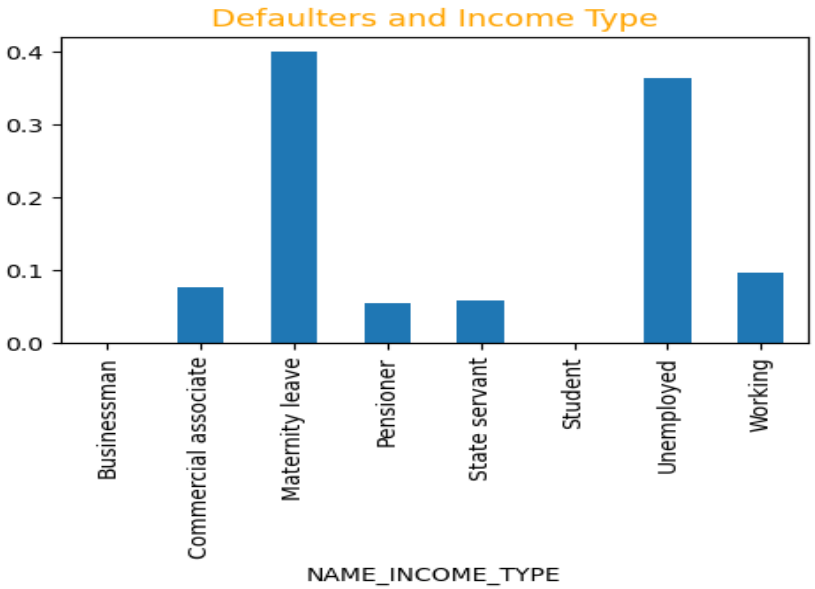


Fig-7

Majority of the defaulters are from 'working ' income type. But the mean rate of default is high for 'Maternity leave' and 'Unemployed'

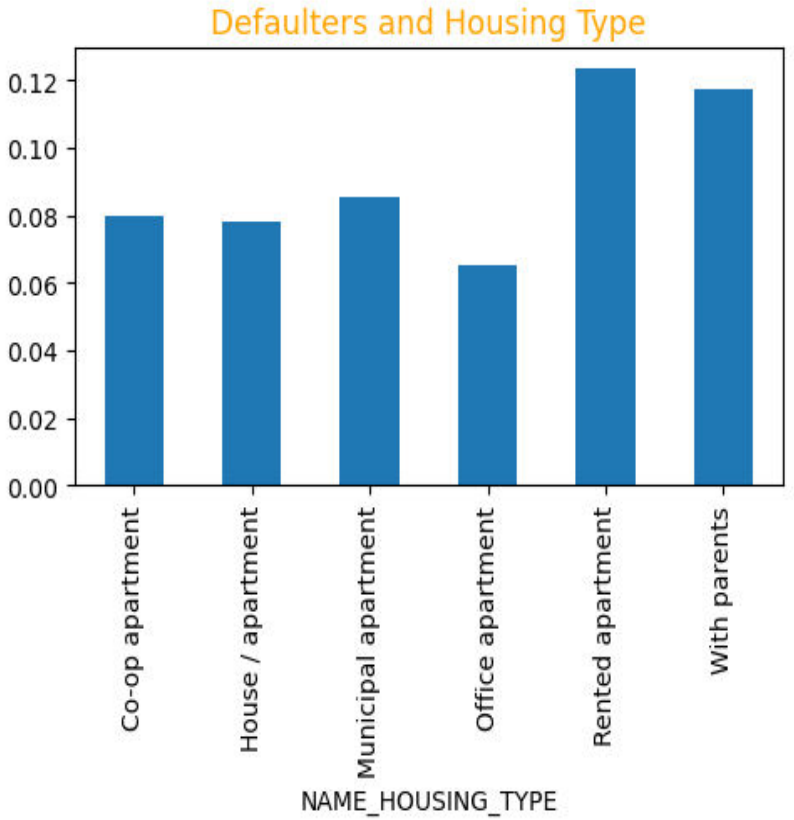


Fig-8

Applicants who lives in 'Rented apartment' Housing_Type has higher mean rate for loan defaulters

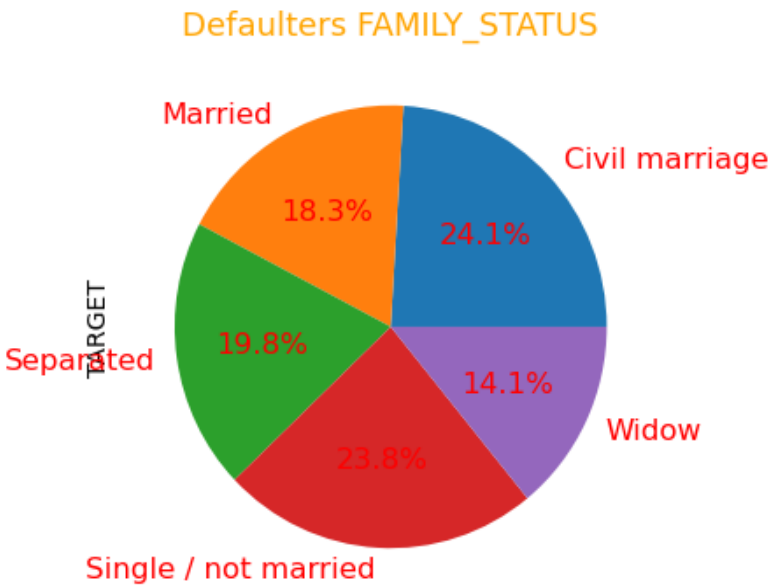


Fig-9

Among the defaulters applicants having 'FAMILY_STATUS' in 'Civil marriage' or 'Single' has higher percentage rate for defaulting loan

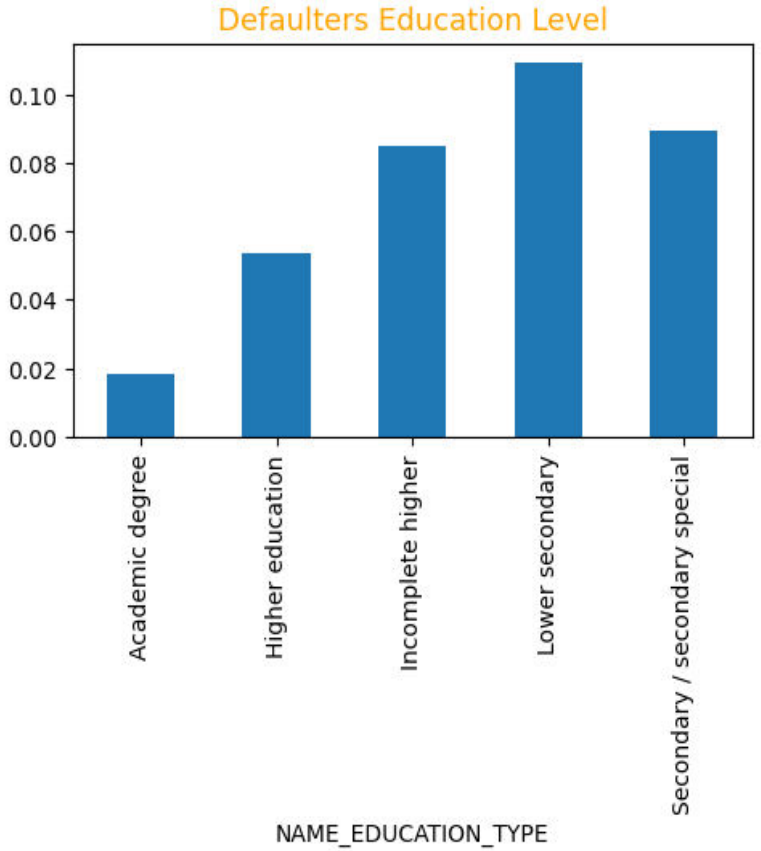
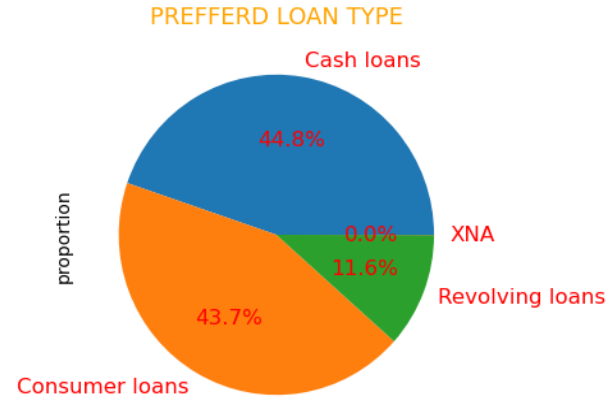


Fig-10

Among the defaulters applicants having 'EDUCATION_TYPE in 'Lower secondary' has higher percentage rate for defaulting loan



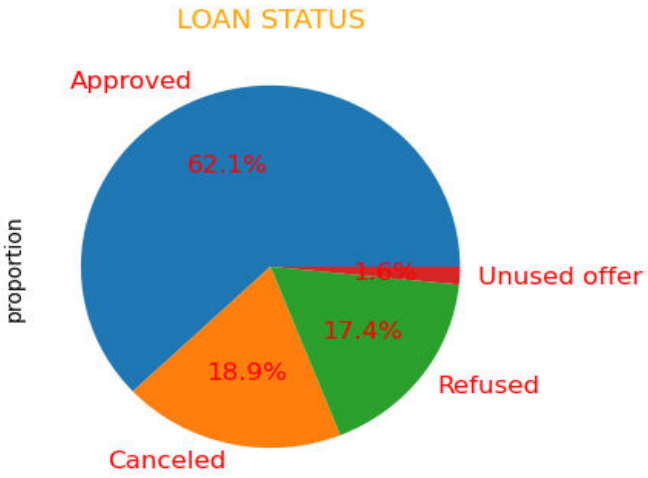


Fig-11

In 'Previous Data' applicants majorly took cash loans and only 62.1% of total applications are approved.

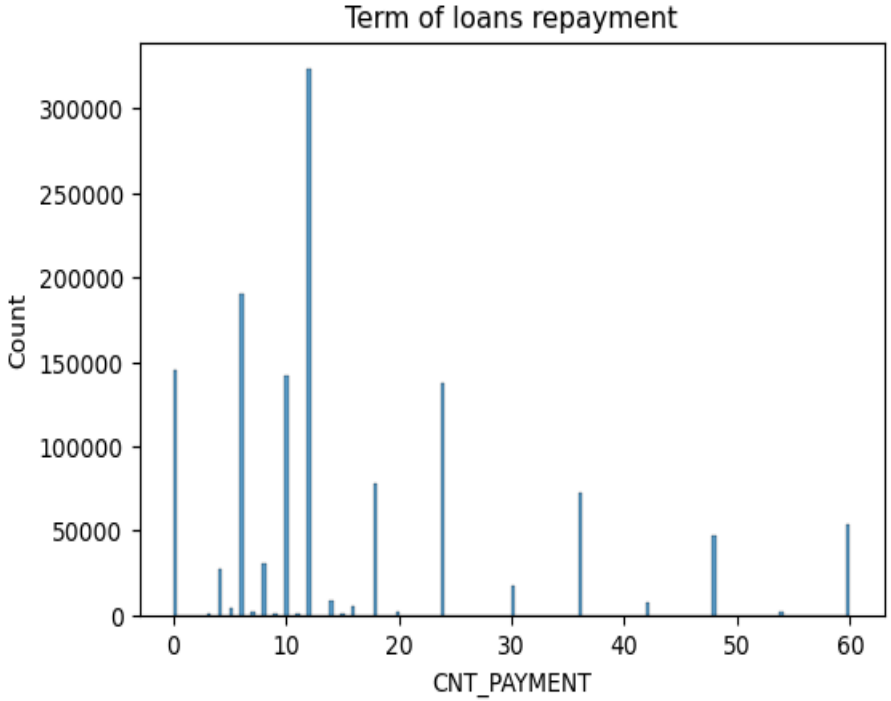


Fig-12

In 'Previous Data' majority of applicants repay the loans with in 15 terms

Recommendation

- Data imbalance between Defaulter and Non_Defaulter clearly shows a safe business.
- In contract type applicants choosing ‘revolving loans ‘ tends to default loans lesser.
- While considering the organization of the applicants, those who come under ‘*Business Entity Type 3’ should undergo more scrutiny as they are large in loan defaulting.*
- *Applicants at younger age tends to default loans compared to elder ones.*
- *Majority of applicants look for low credit amount .*
- *Applicants living in region-3 ,tends to default more loans.*
- *Applicants mentioning ‘Income_Type’ as ‘Maternity leave’ and ‘Unemployed’ have high chances of defaulting loans.*
- *Applicants living in ‘Rented apartment’ have high chances of defaulting loans.*
- *While considering the family status of the applicant ,those who are ‘single’ or ‘Civil marriage’ default loans highly compared to other.*
- *Considering the ‘EDUCATION_TYPE’ , applicants in ‘Lower secondary’ have higher rate of loan defaulting.*