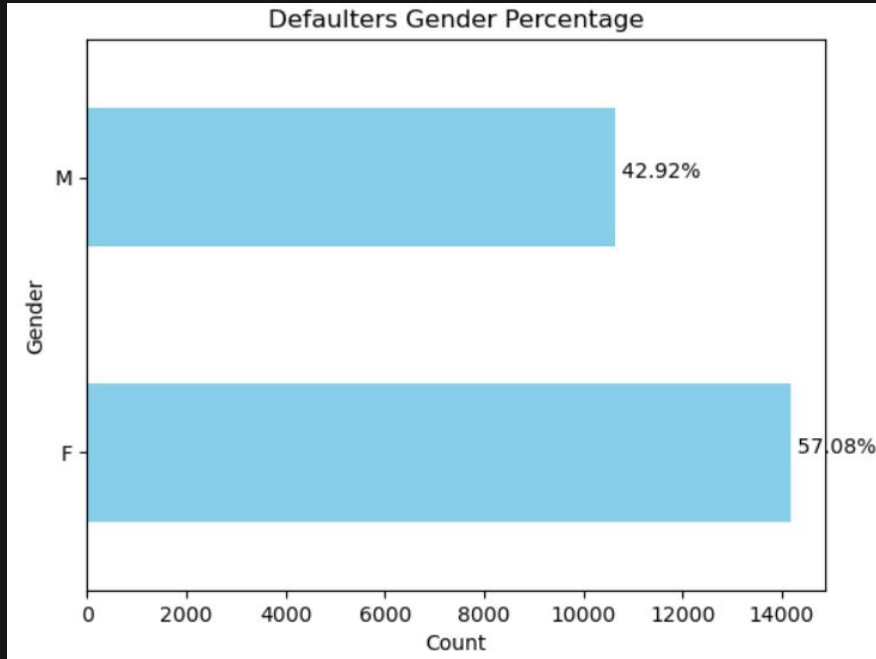


Exploratory Data Analysis - Assignment

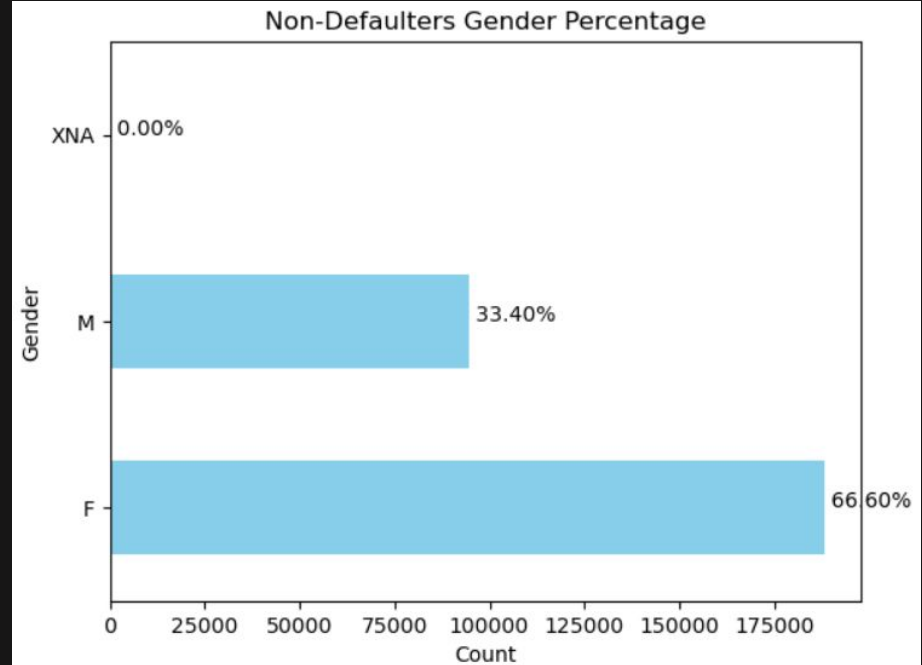
Univariate Analysis

Gender

Defaulters Gender Percentage



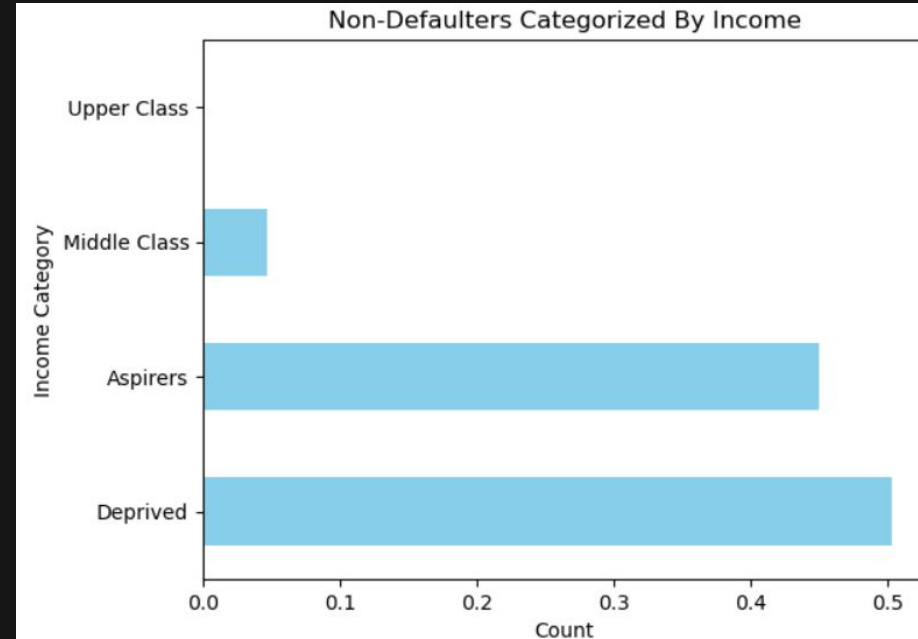
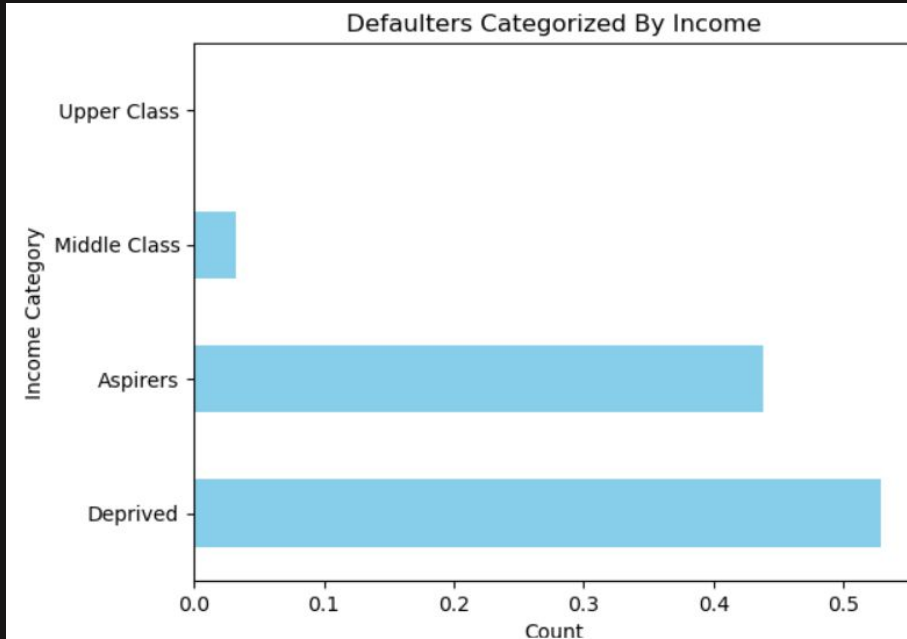
Non-Defaulters Gender Percentage



- Defaulters have 42.9% males and 57.08% females
- Non - Defaulters have 33.4% males and 66.6% females

Univariate Analysis

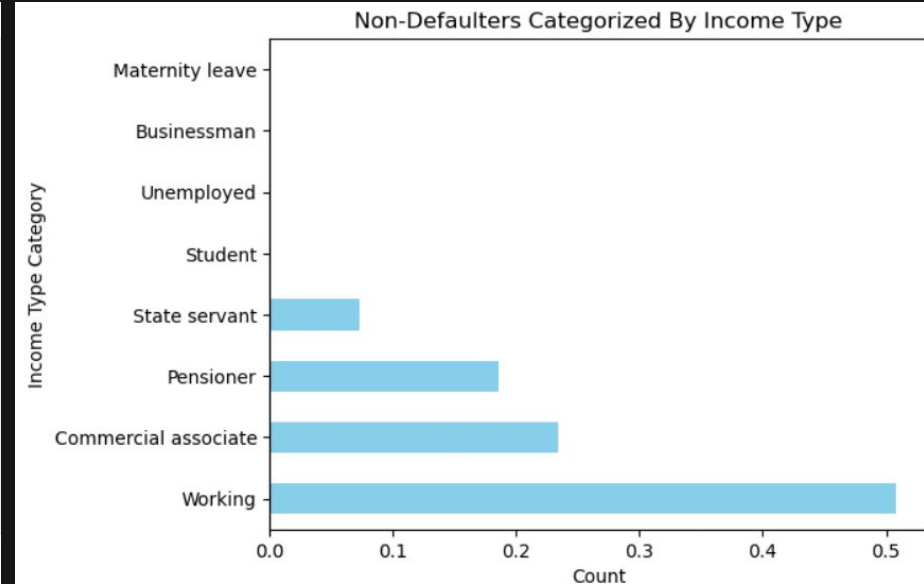
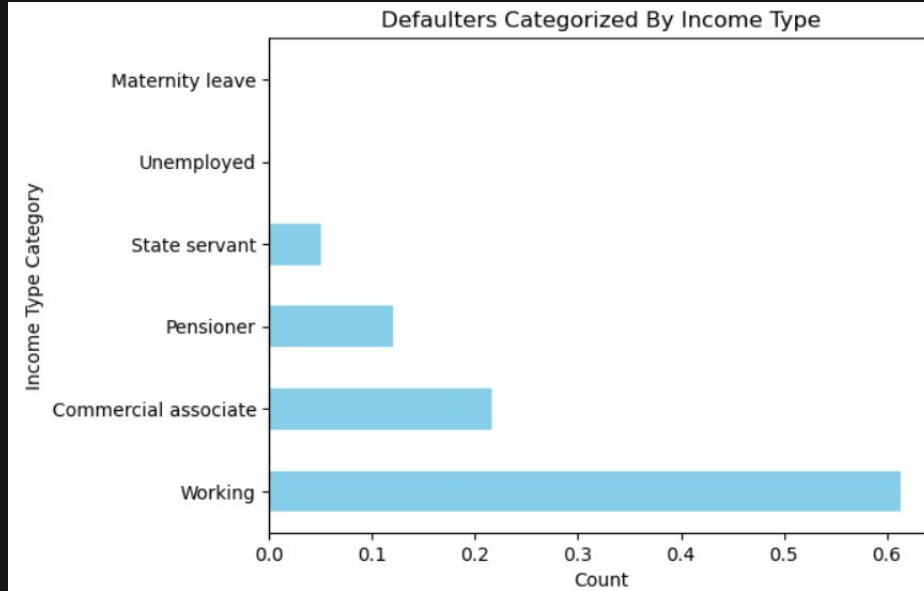
Income



- In Defaulters category most of the people comes under “Deprived” category
- In Non-Defaulters category most of the people comes under “Deprived” category

Univariate Analysis

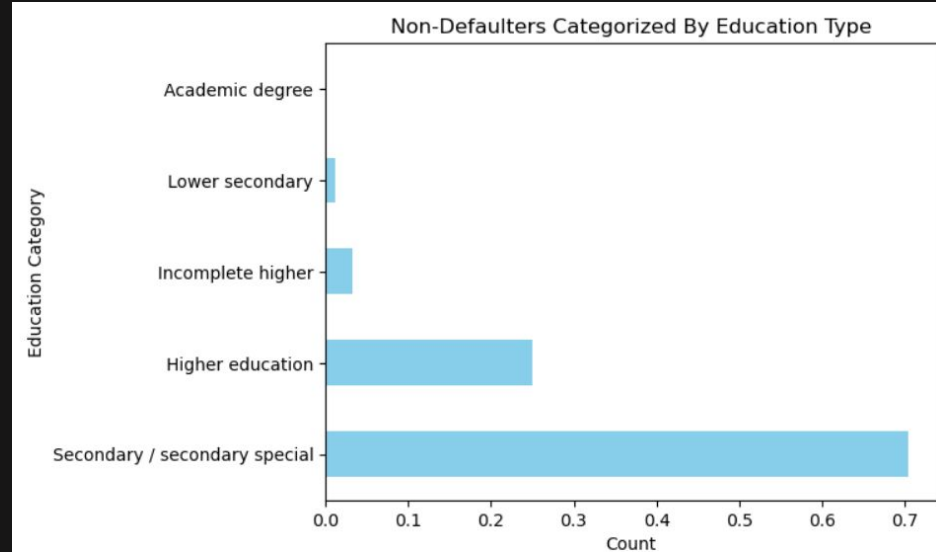
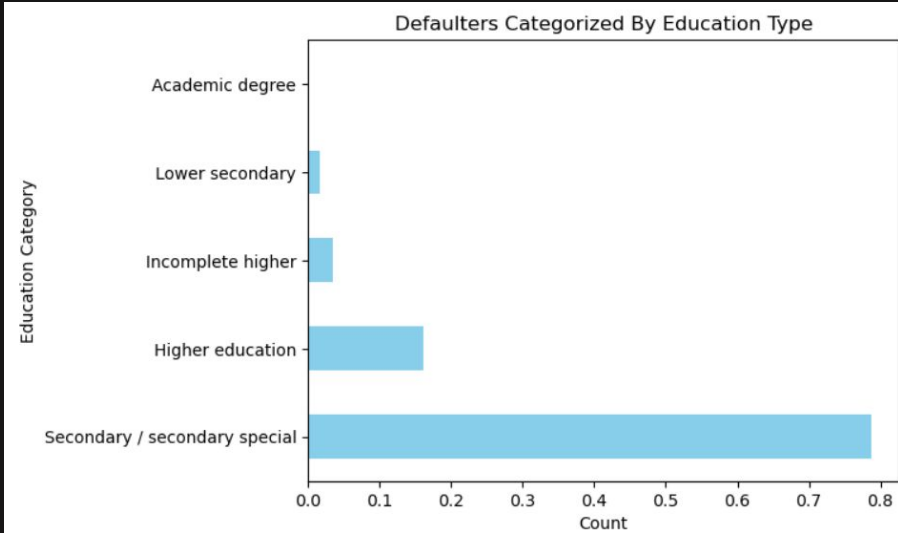
Income Type



- Working class are the maximum in both Defaulter and non-defaulter category.
- State servants and Pensioners Default their loan less compared to working class.

Univariate Analysis

Education Type

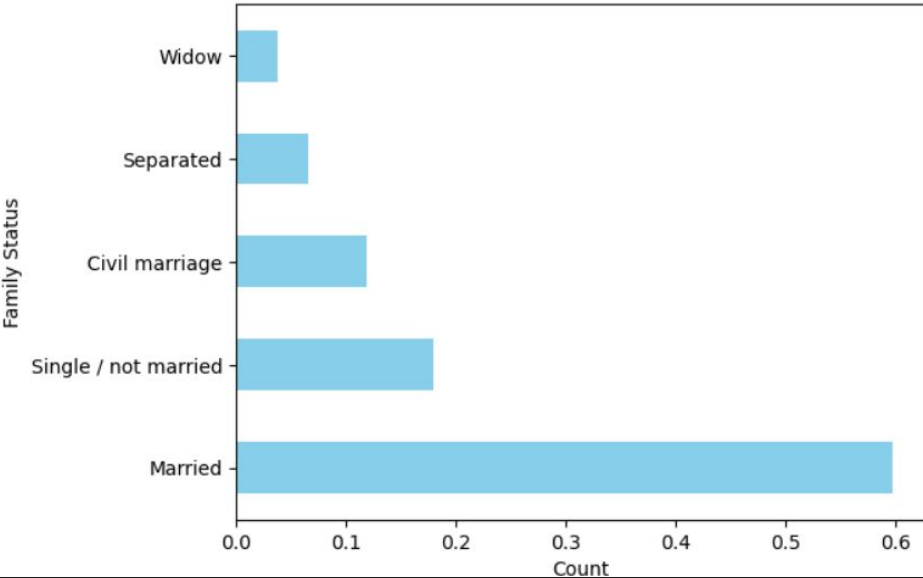


- Secondary / secondary special default their loans more

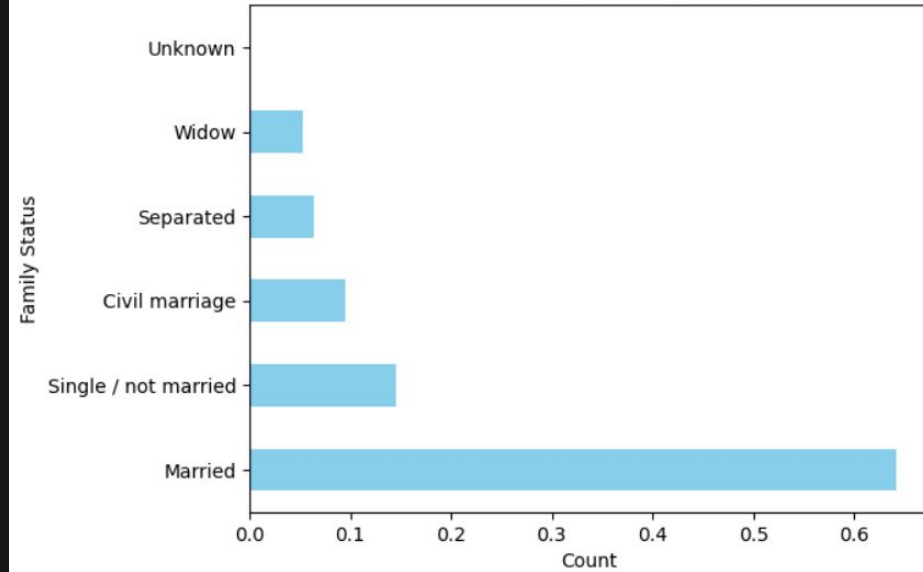
Univariate Analysis

Family Status

Defaulters Categorized By Family Status



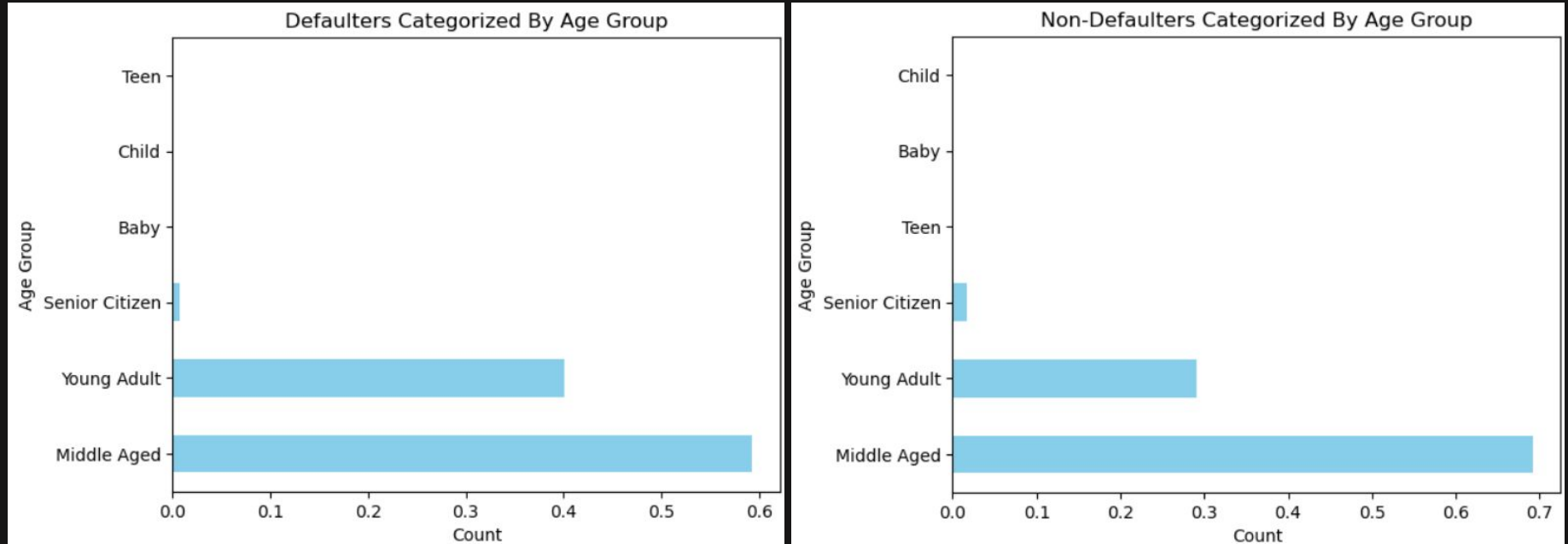
Non-Defaulters Categorized By Family Status



- Married people default their loans the most

Univariate Analysis

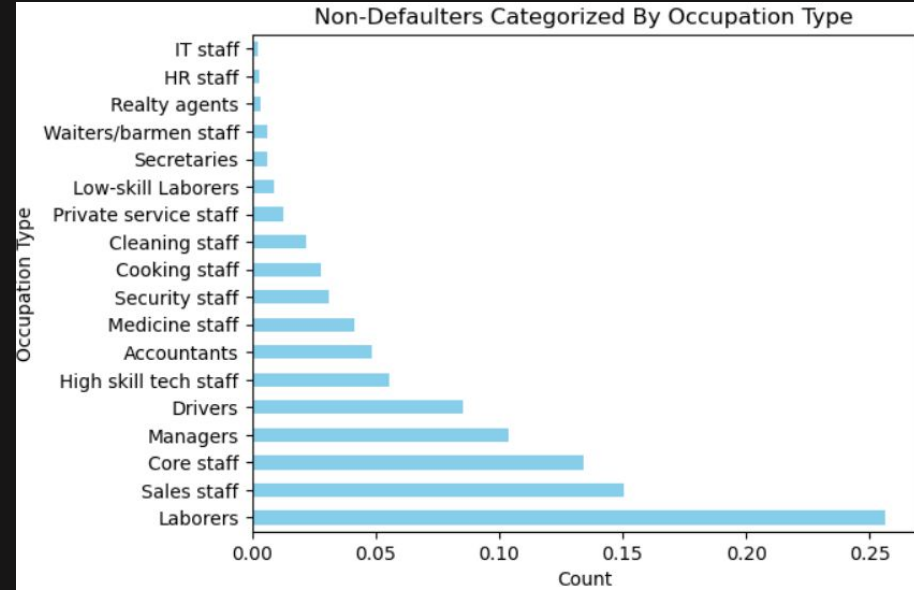
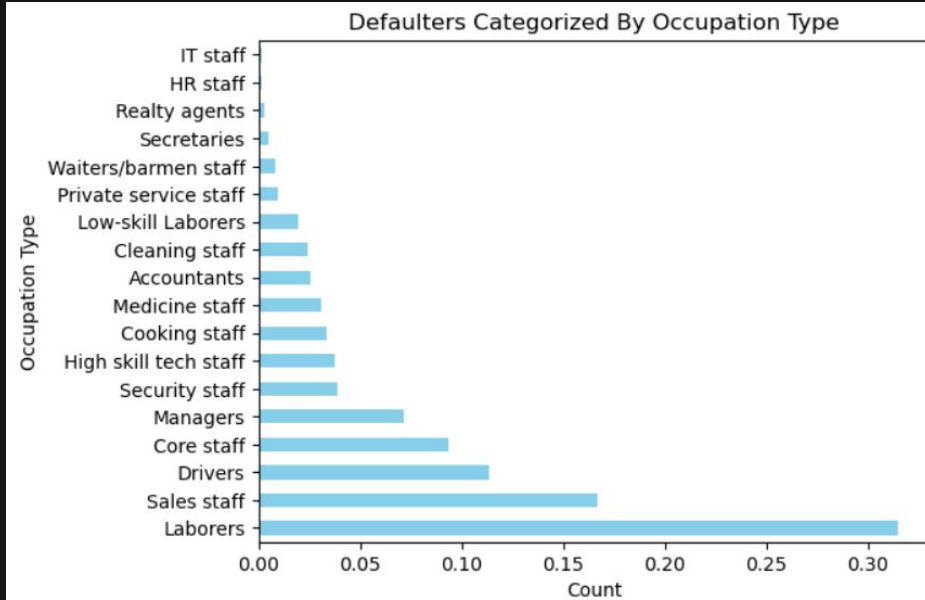
Age Group



- Middle Aged people are the most defaulting loans

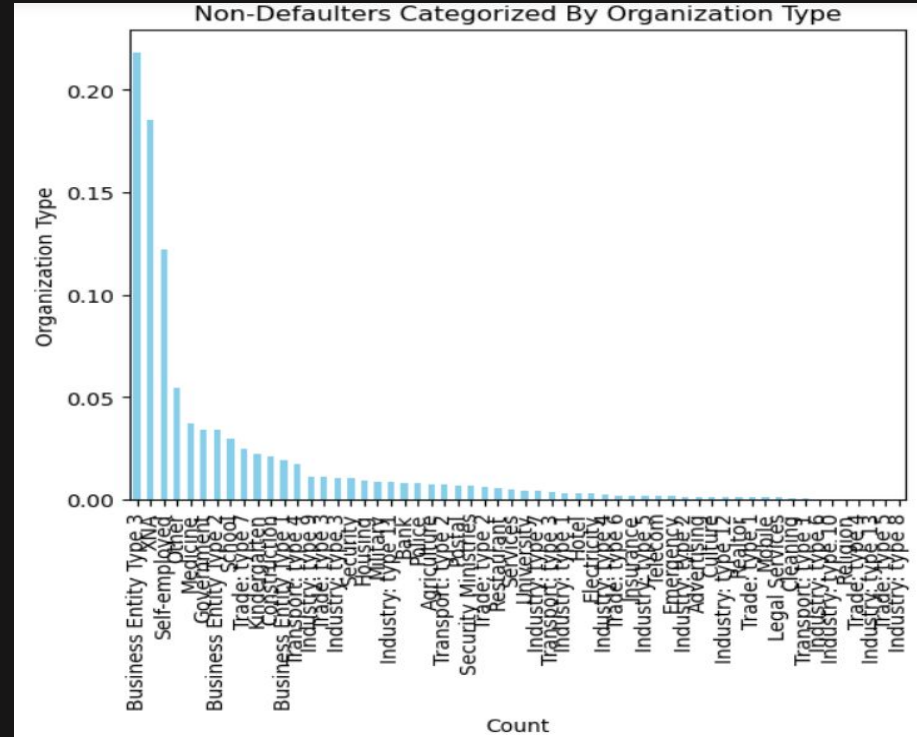
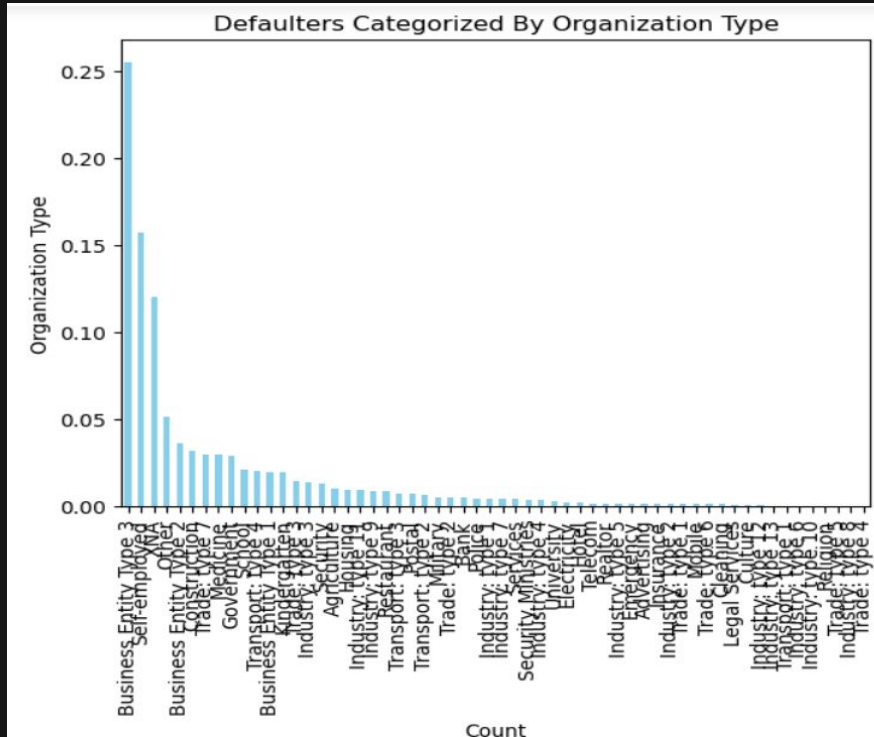
Univariate Analysis

Occupation Type



- Middle Aged people are the most defaulting loans

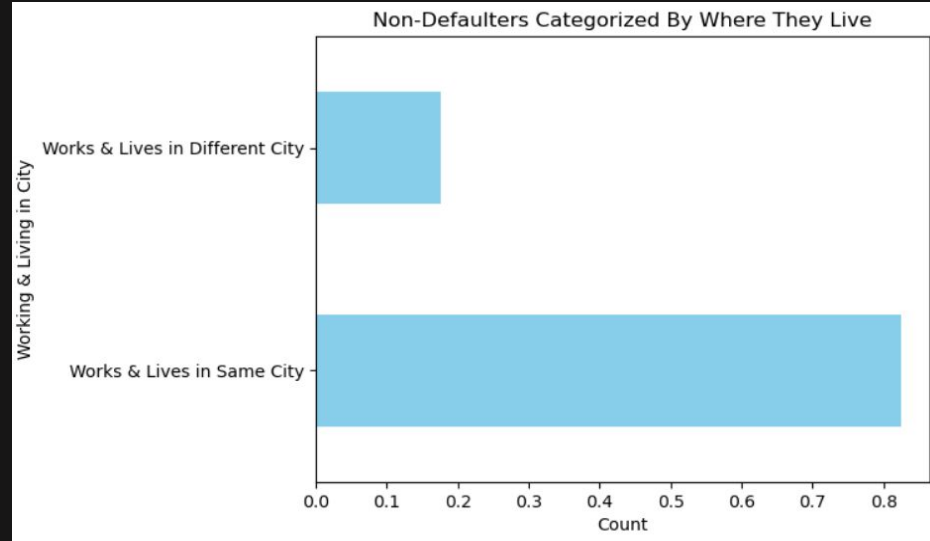
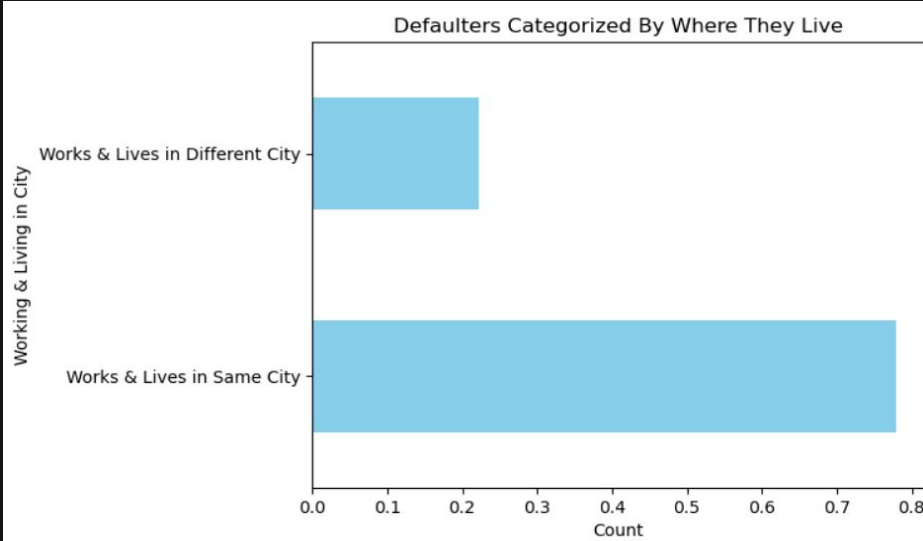
Organization Type



- People with & self employed default the least

Univariate Analysis

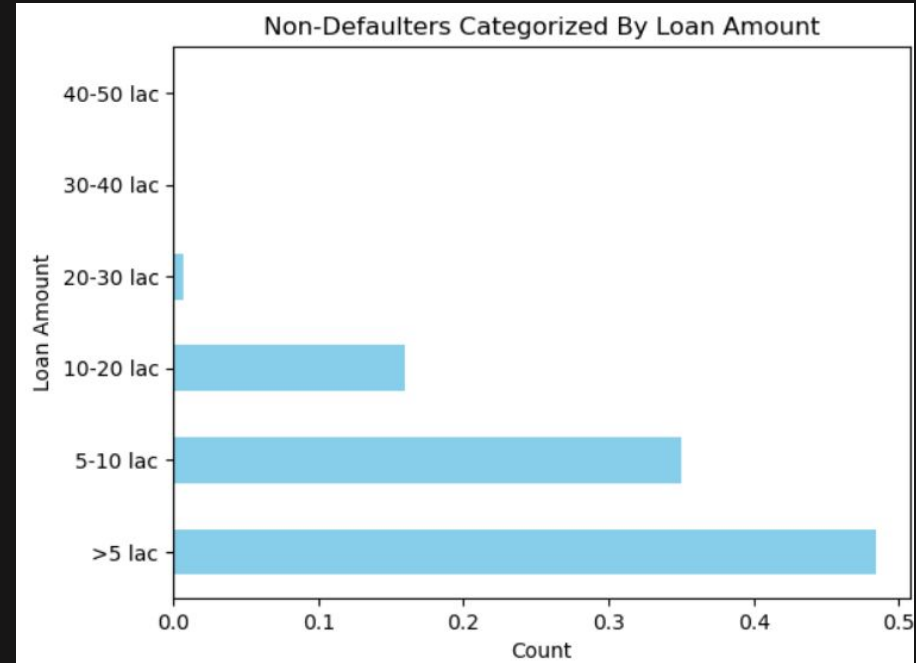
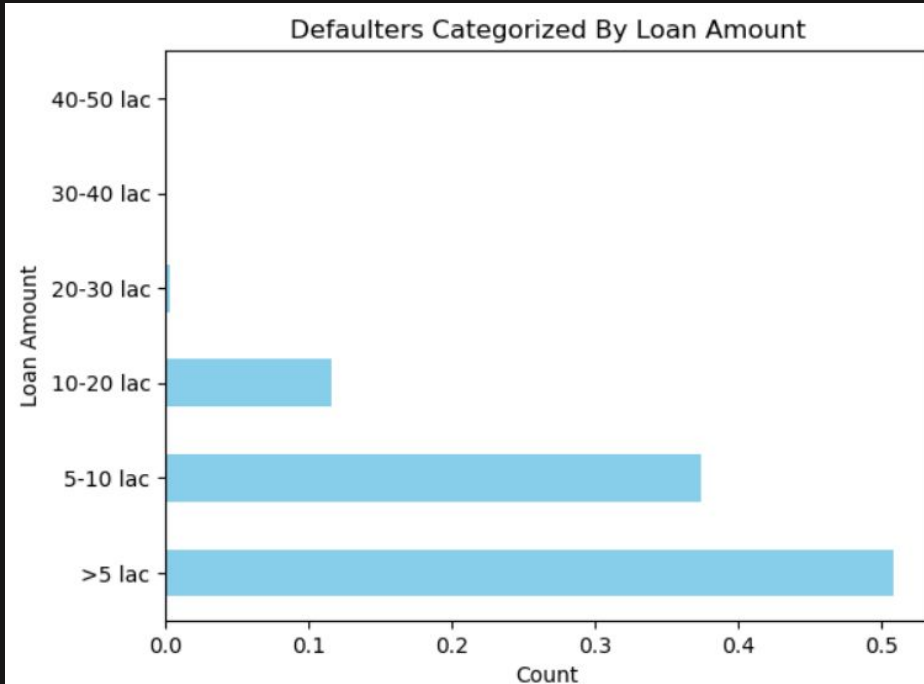
Organization Type



- People who work and live in different city default more

Univariate Analysis

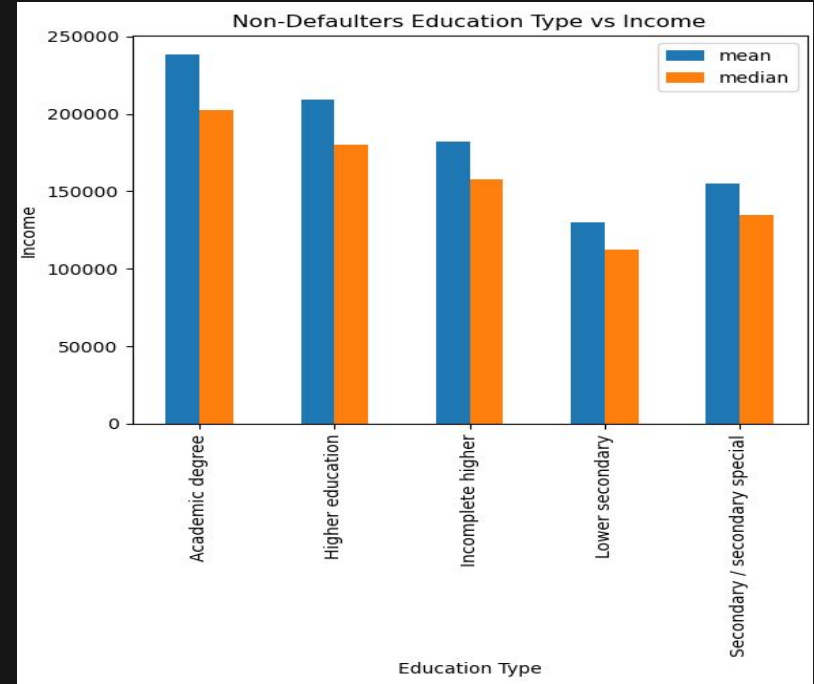
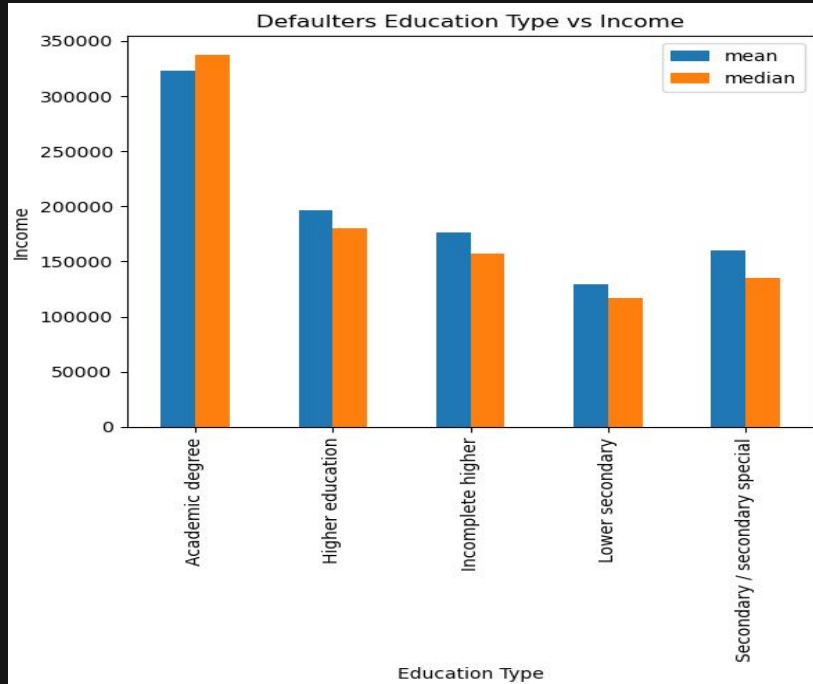
Organization Type



- People taking loan for less than 10 lac default the most

Bivariate Analysis

Education Type vs Income

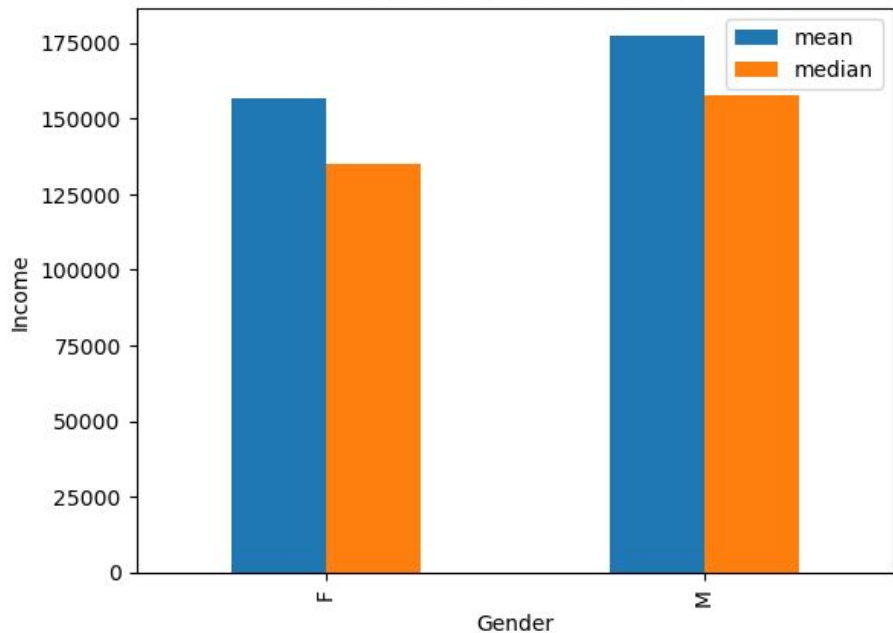


- People with lower income for the same education type default more

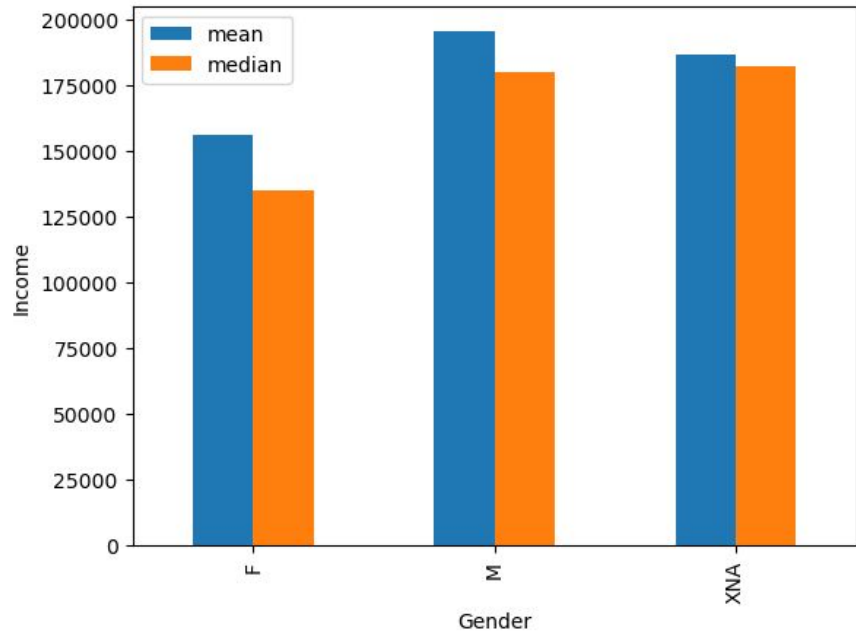
Bivariate Analysis

Gender vs Income

Defaulters Gender vs Income



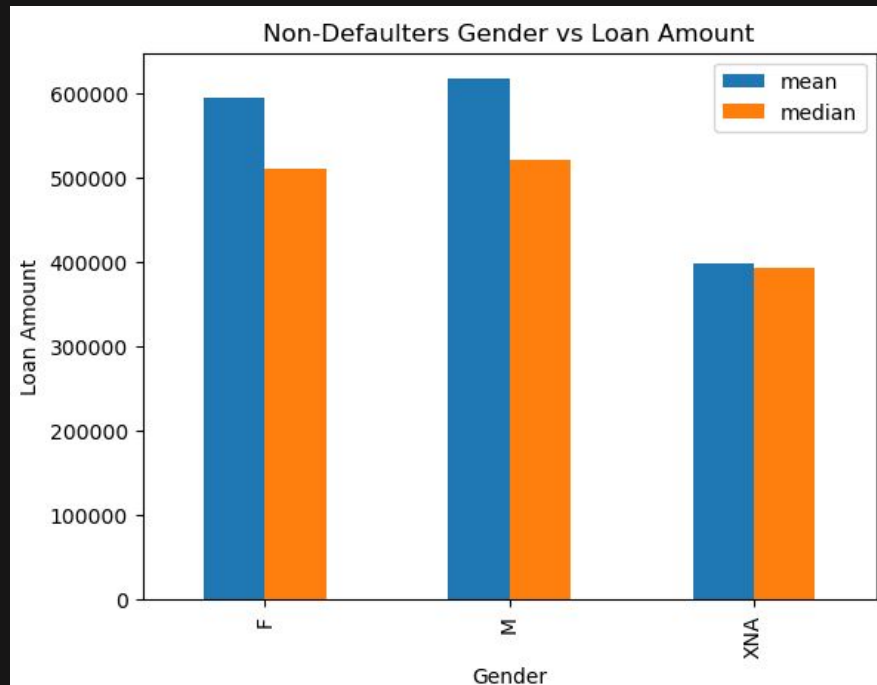
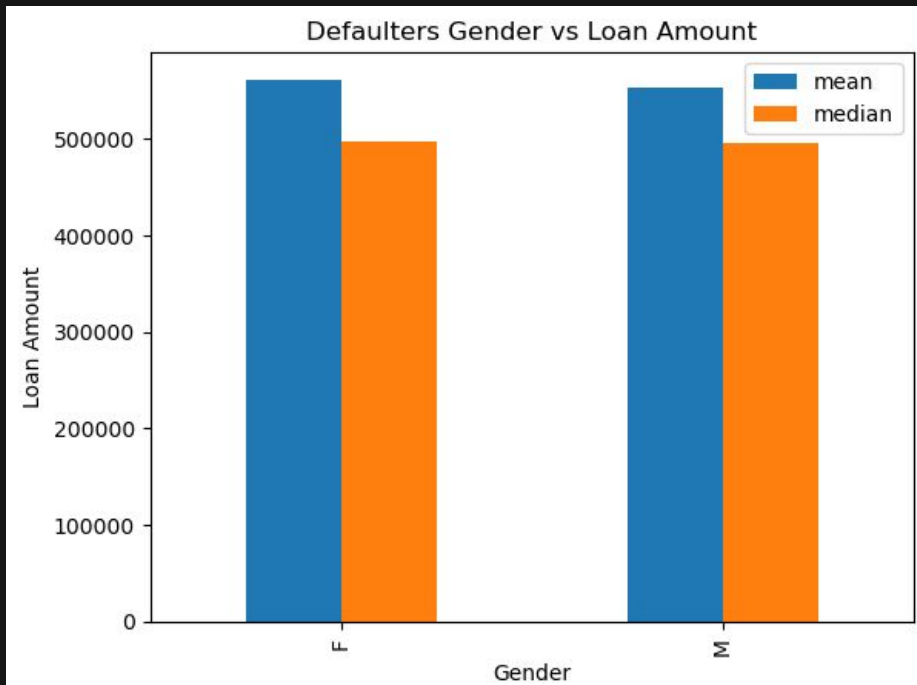
Non-Defaulters Gender vs Income



- Males with lower income default more

Bivariate Analysis

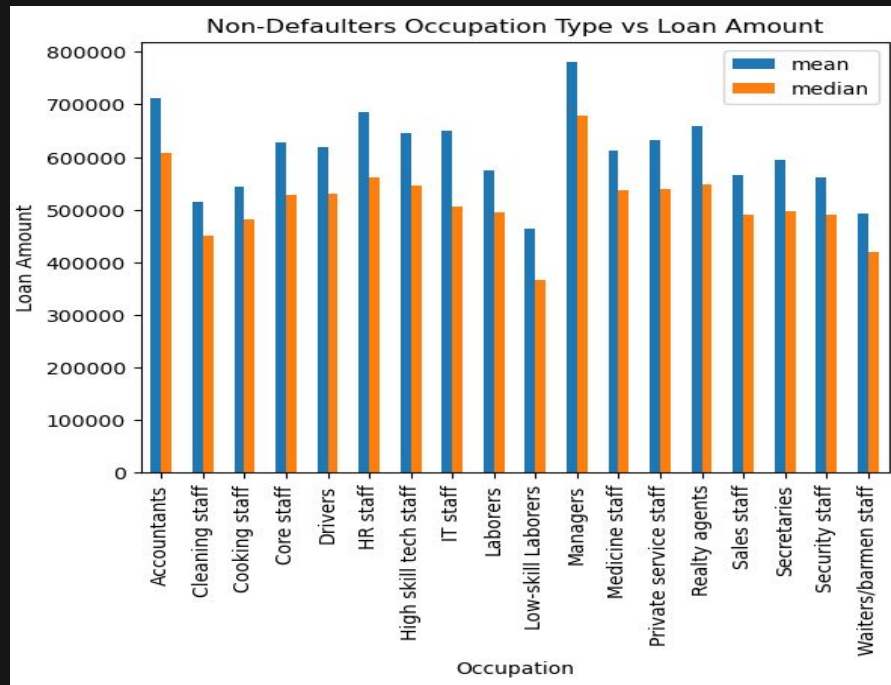
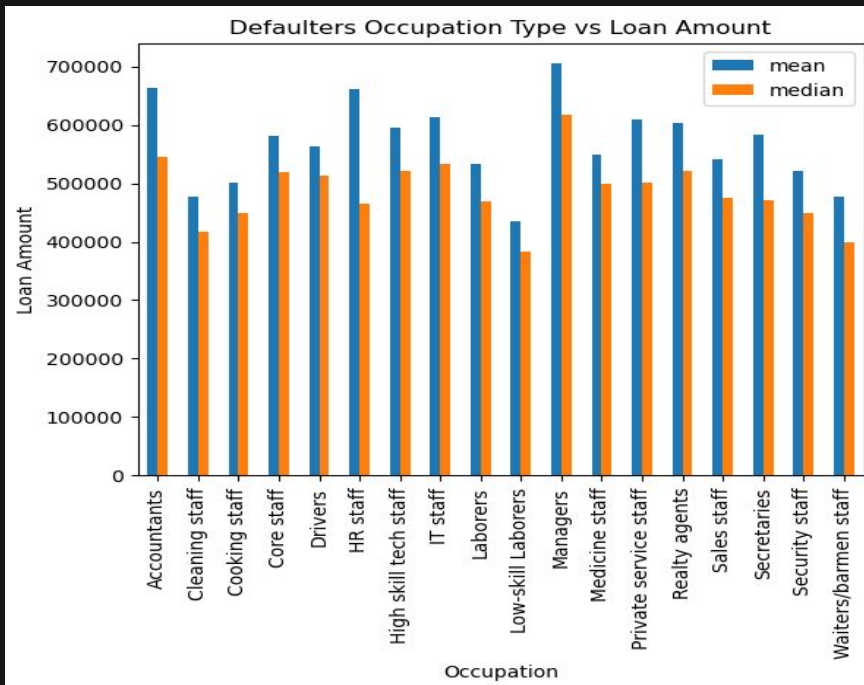
Gender vs Loan Amount



- Females who have defaulted have taken slightly higher loan than males.

Bivariate Analysis

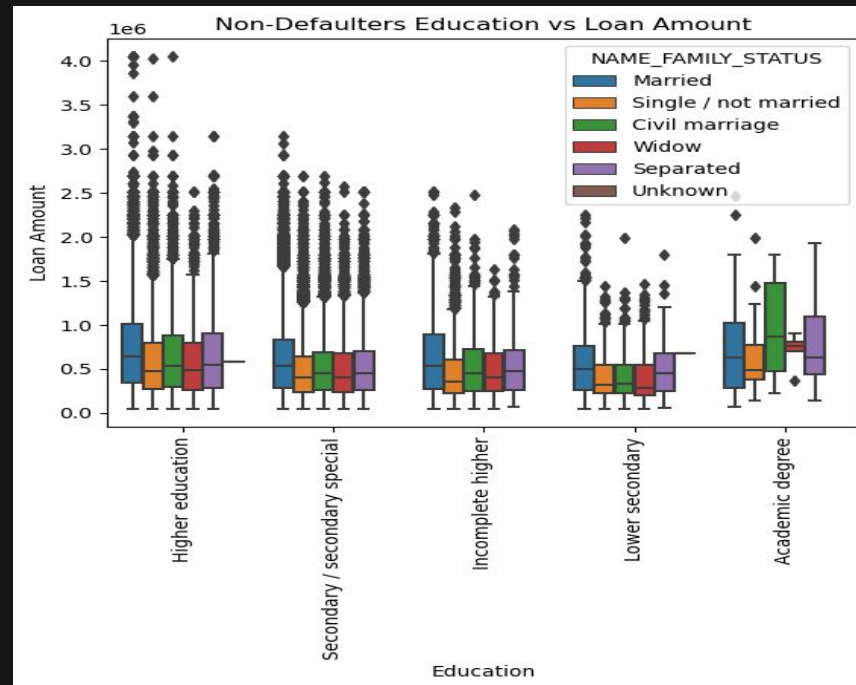
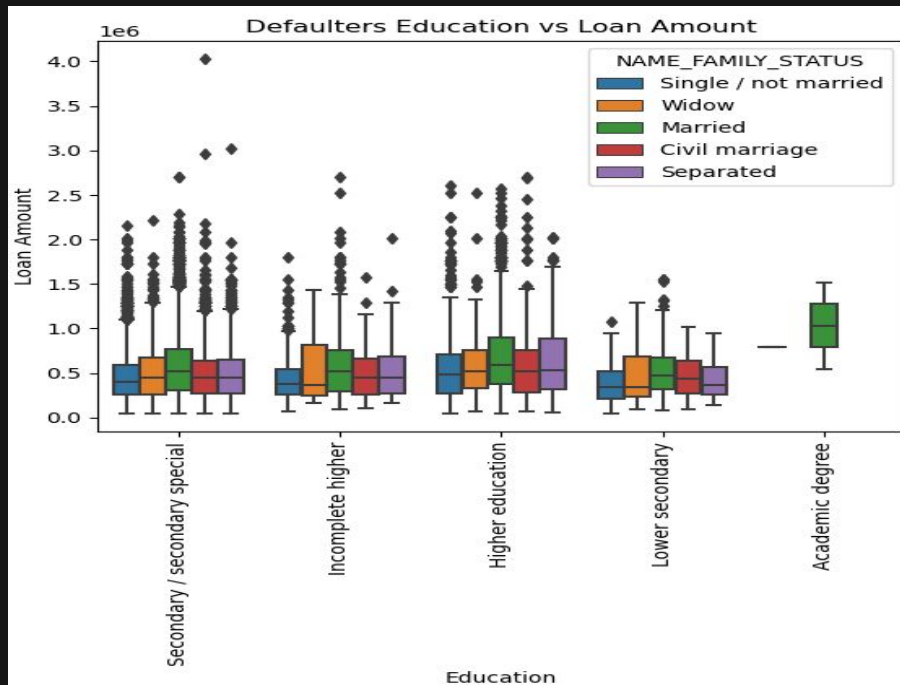
Occupation Type vs Loan Amount



- Managers take higher loans in both defaulters and non-defaulters data set.

Bivariate Analysis

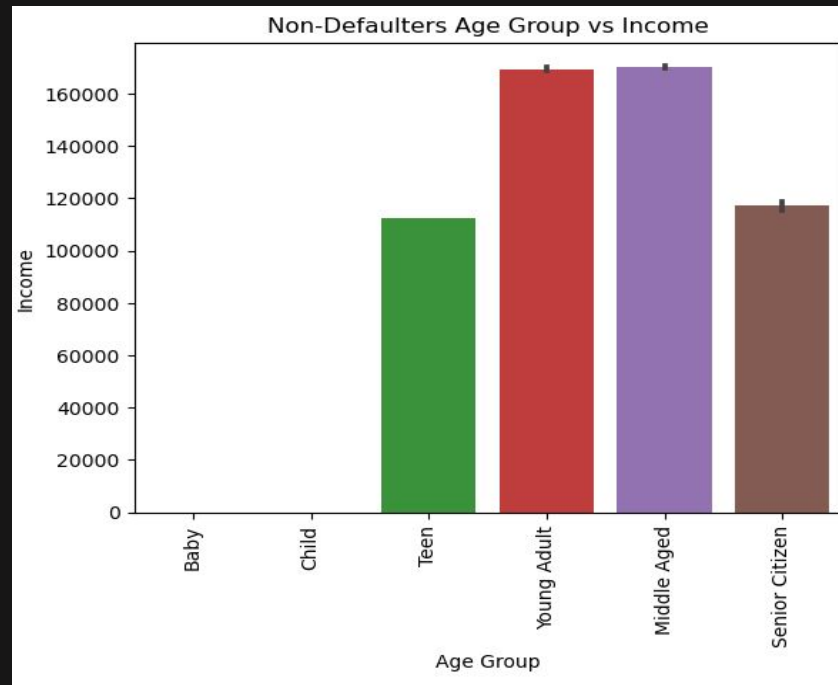
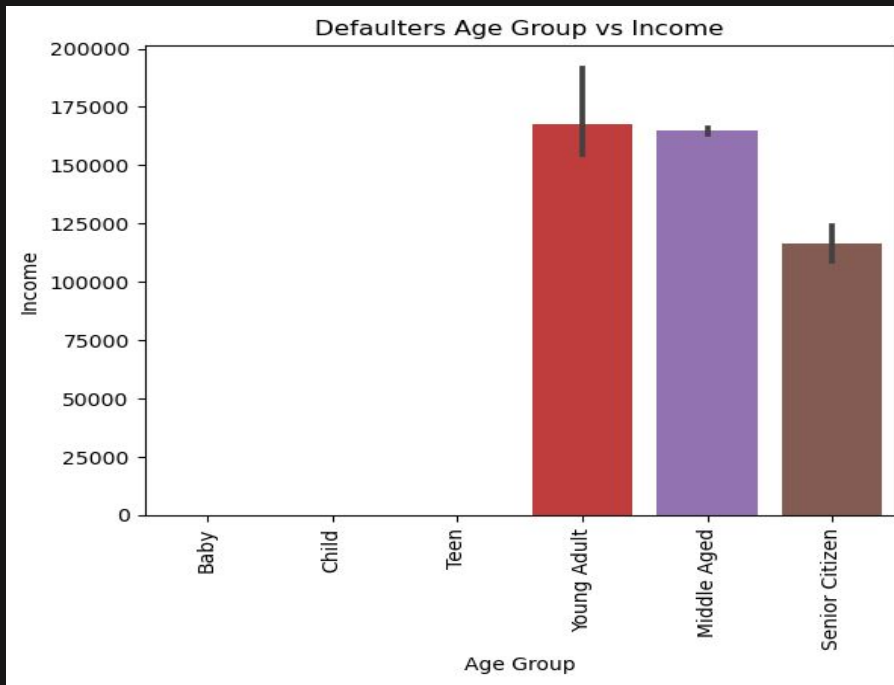
Education vs Loan Amount



- People who are married with higher education take higher loans and default their loans less

Bivariate Analysis

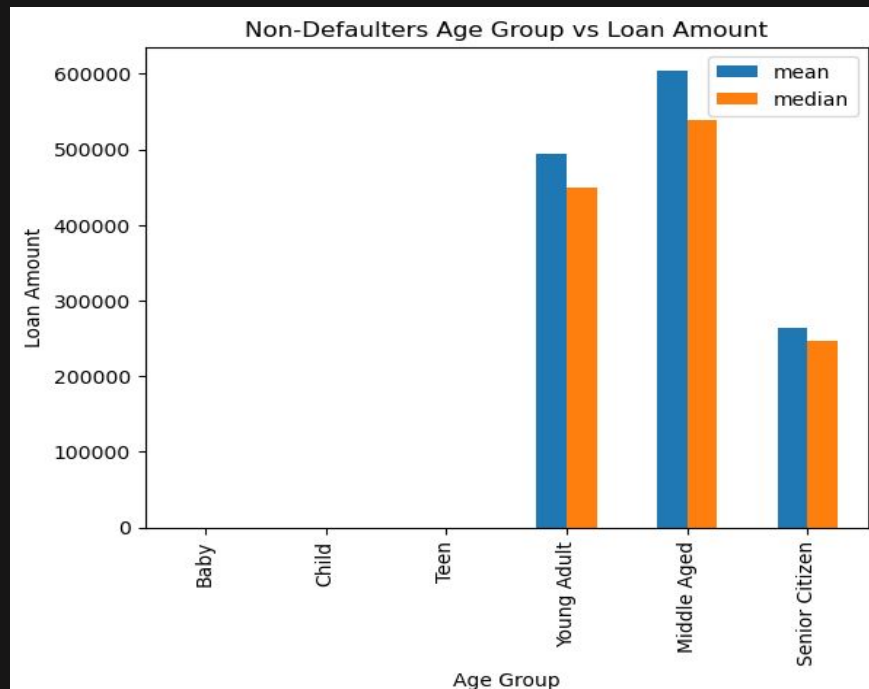
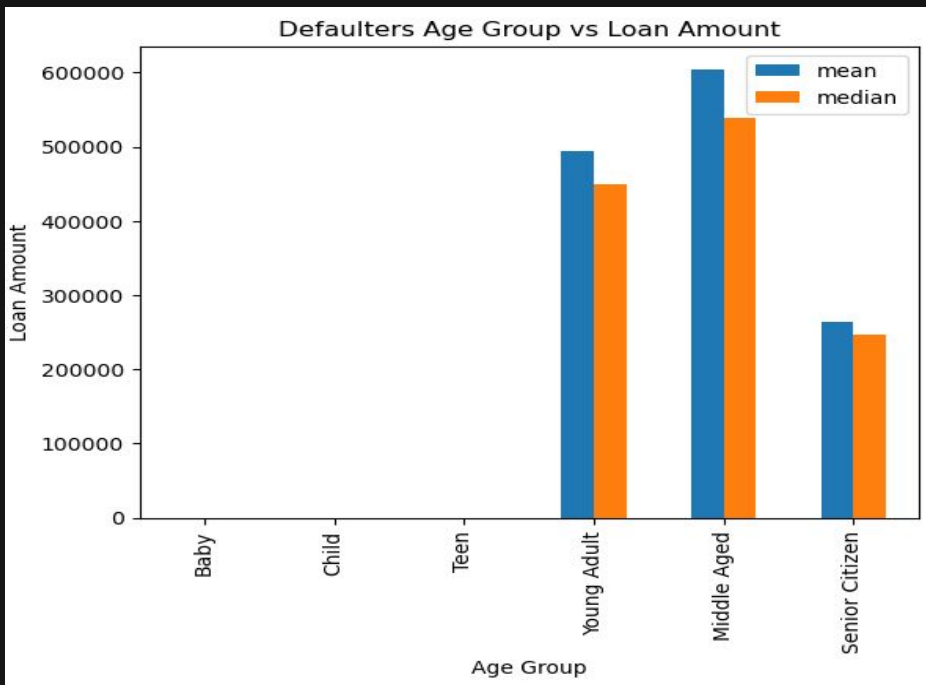
Age Group vs Income



- In defaulters Young adults have more income followed by Middle aged people

Bivariate Analysis

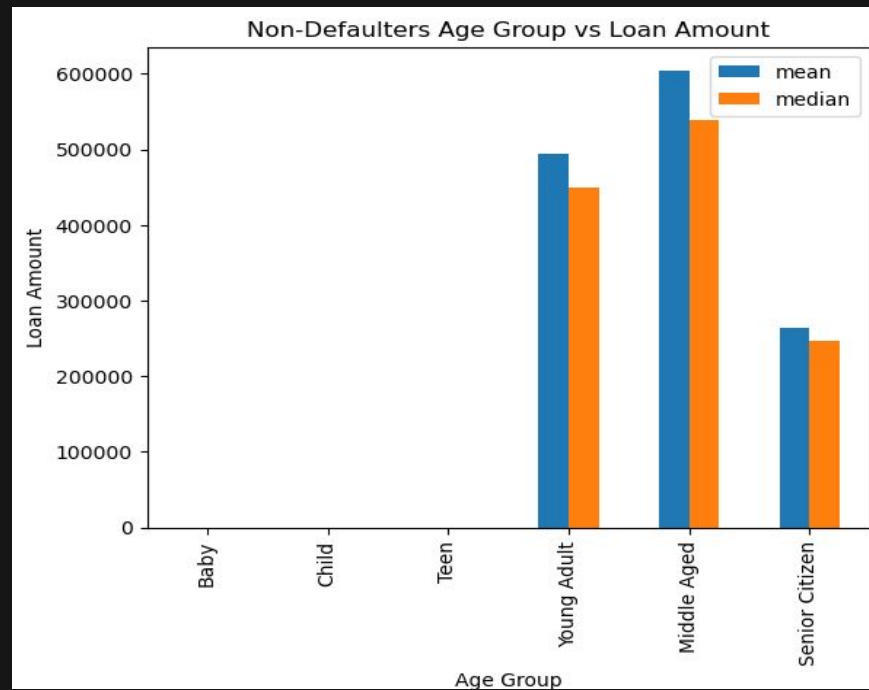
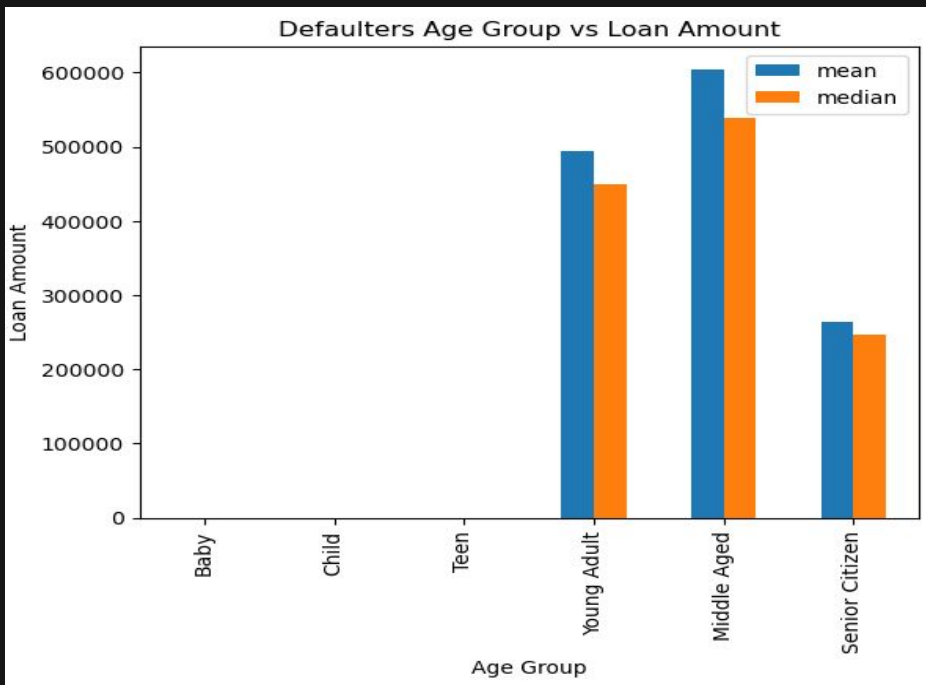
Age Group vs Loan Amount



- Middle aged people take more loans

Bivariate Analysis

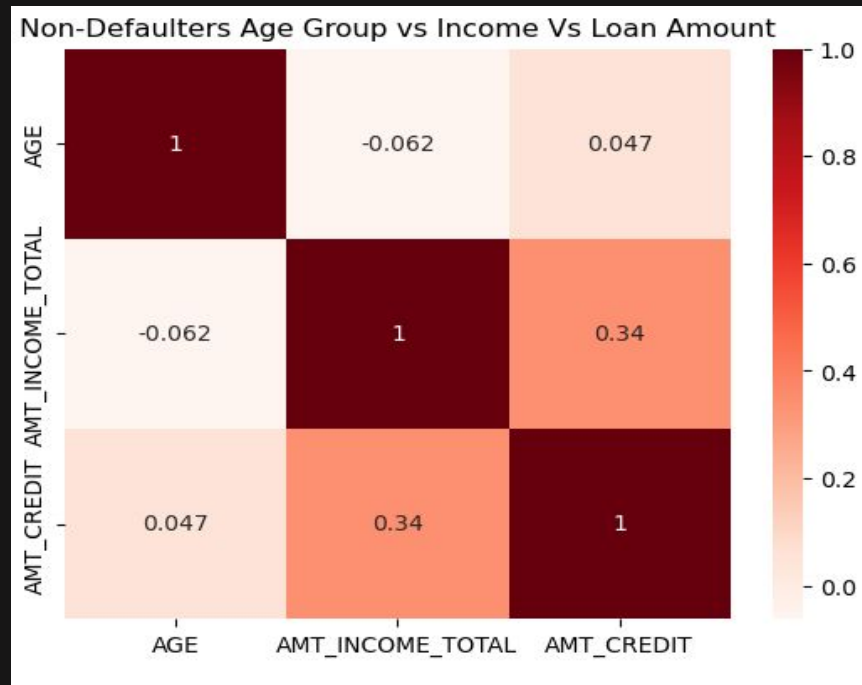
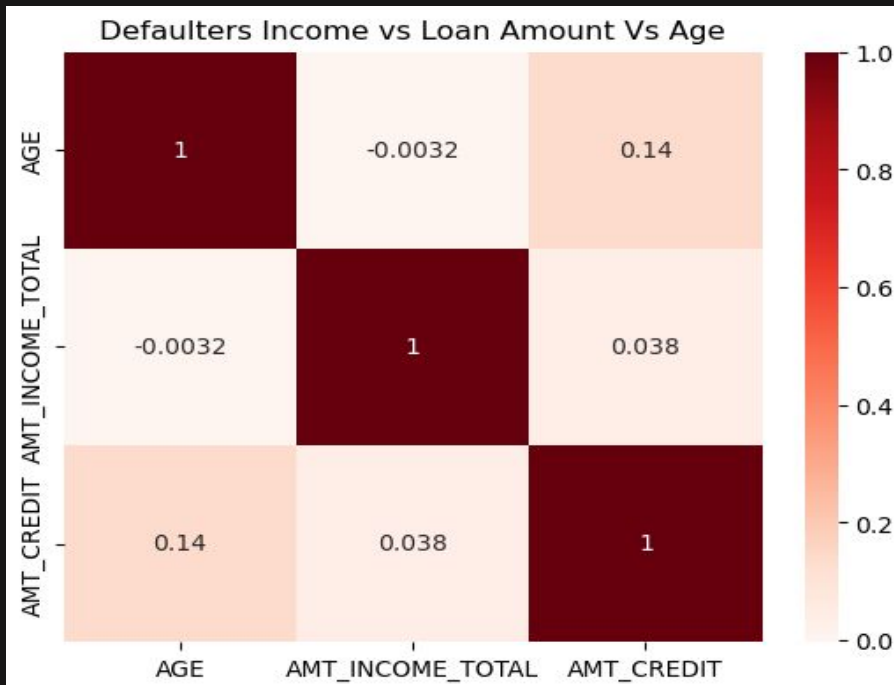
Age Group vs Loan Amount



- Middle aged people take more loans

Bivariate Analysis

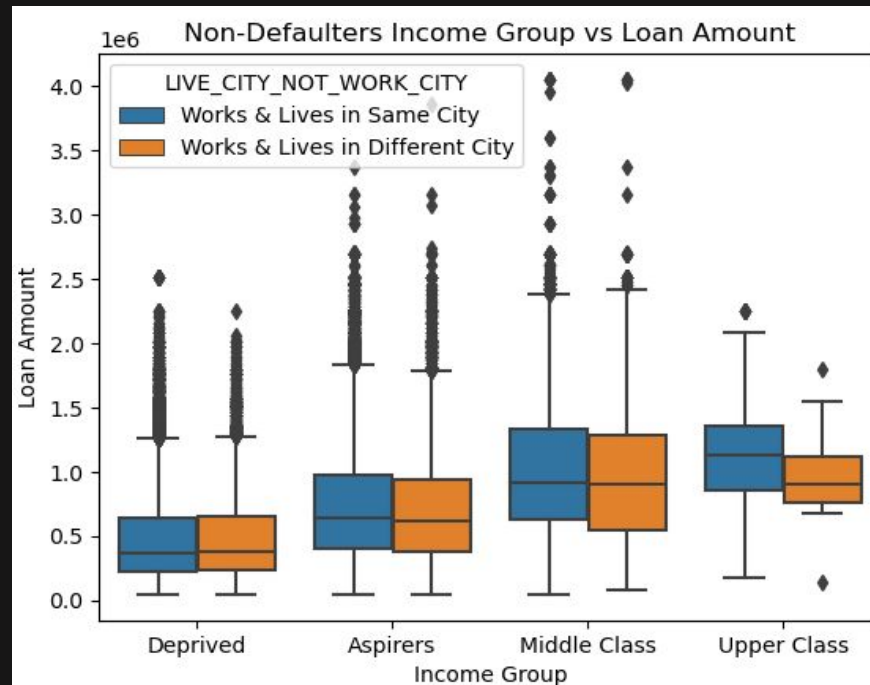
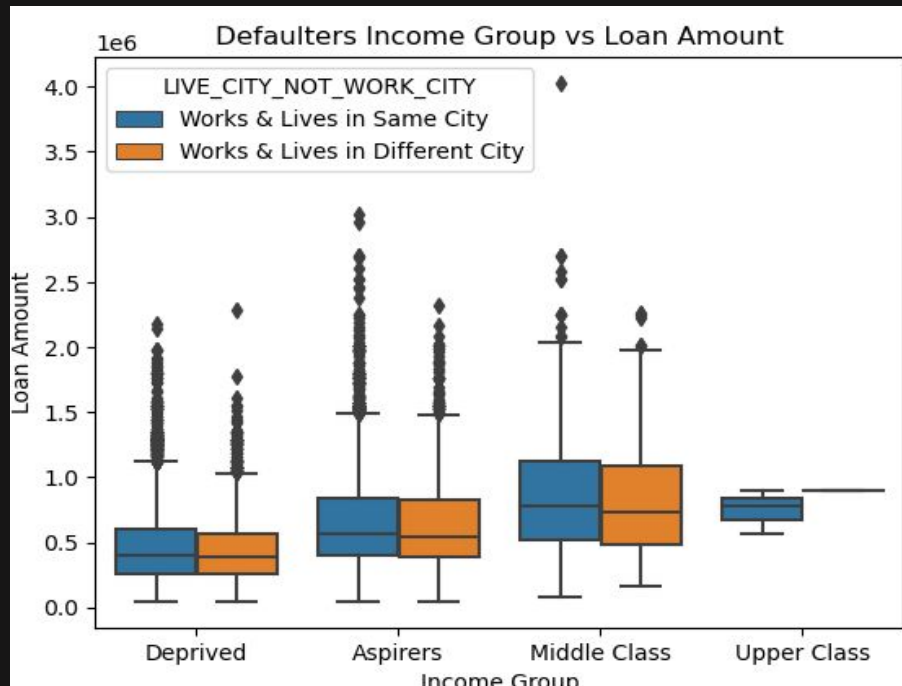
Income vs Loan Amount Vs Age



- Comparatively high correlation between income and loan amount

Bivariate Analysis

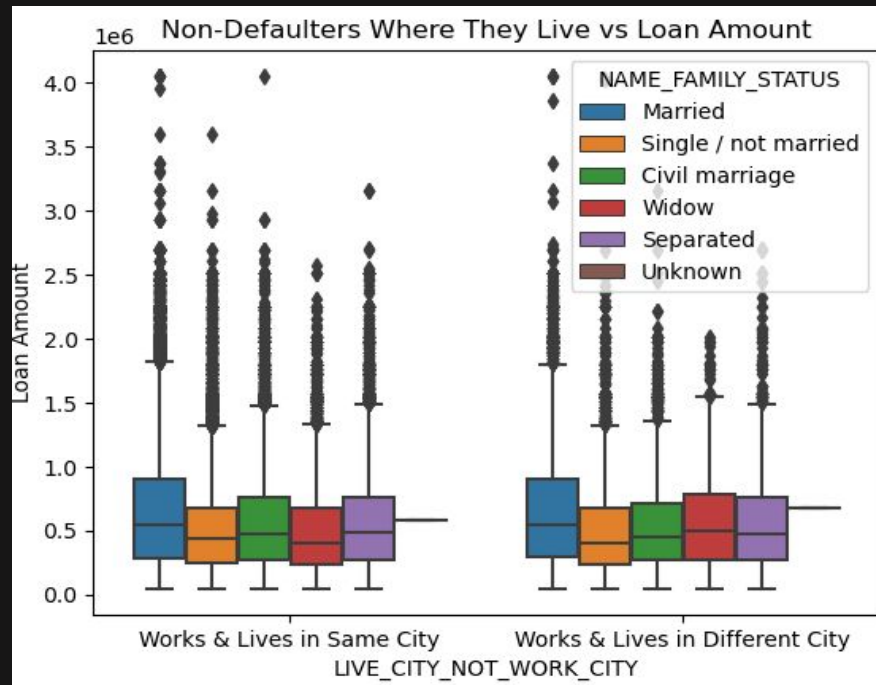
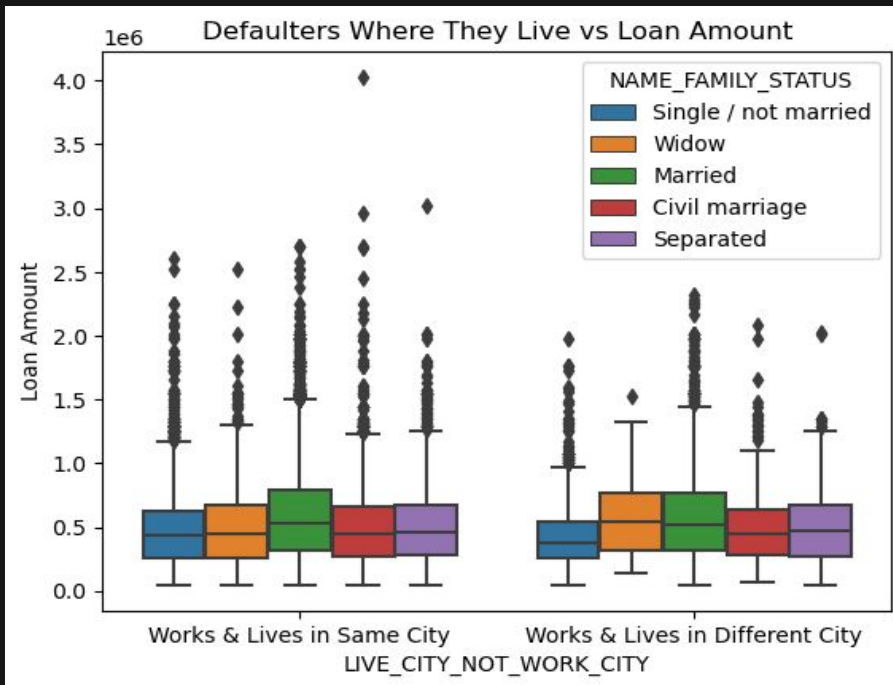
Income Group vs Loan Amount



- Upper class people work and live in the same city and default their loan the least

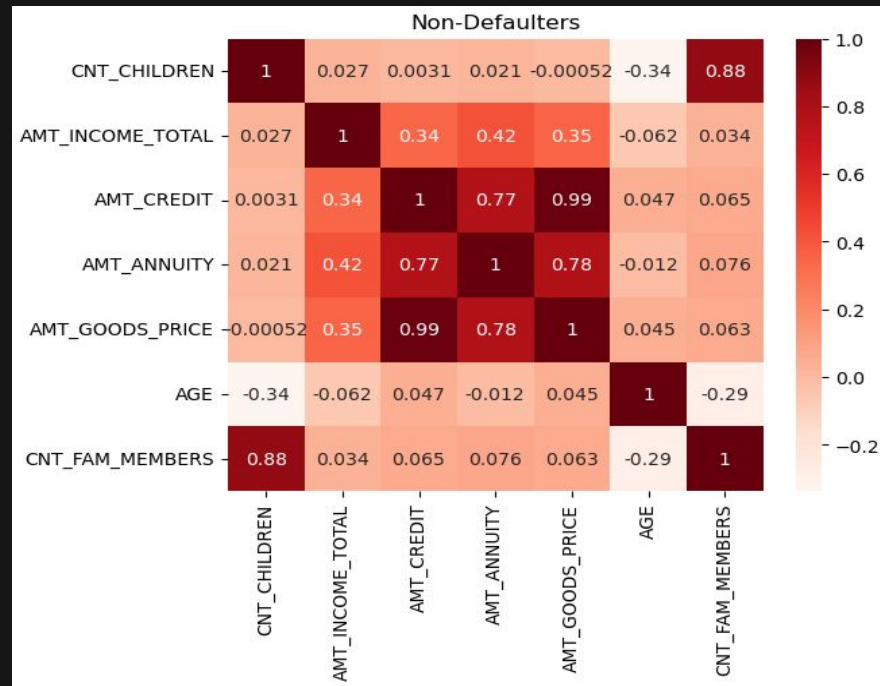
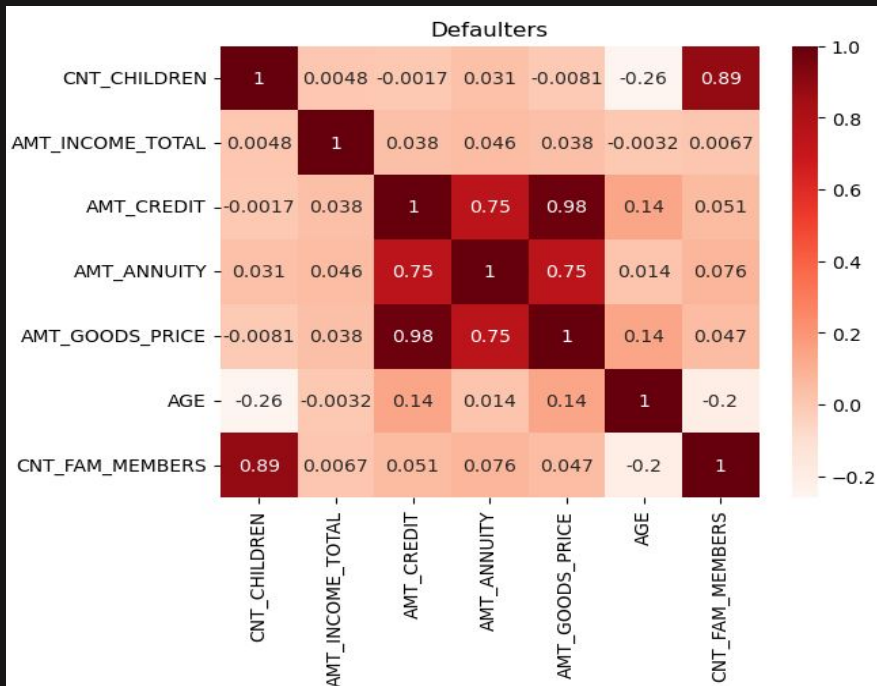
Bivariate Analysis

Defaulters Where They Live vs Loan Amount



- People who don't default their loans are married doesn't matter where they live also they take higher loans

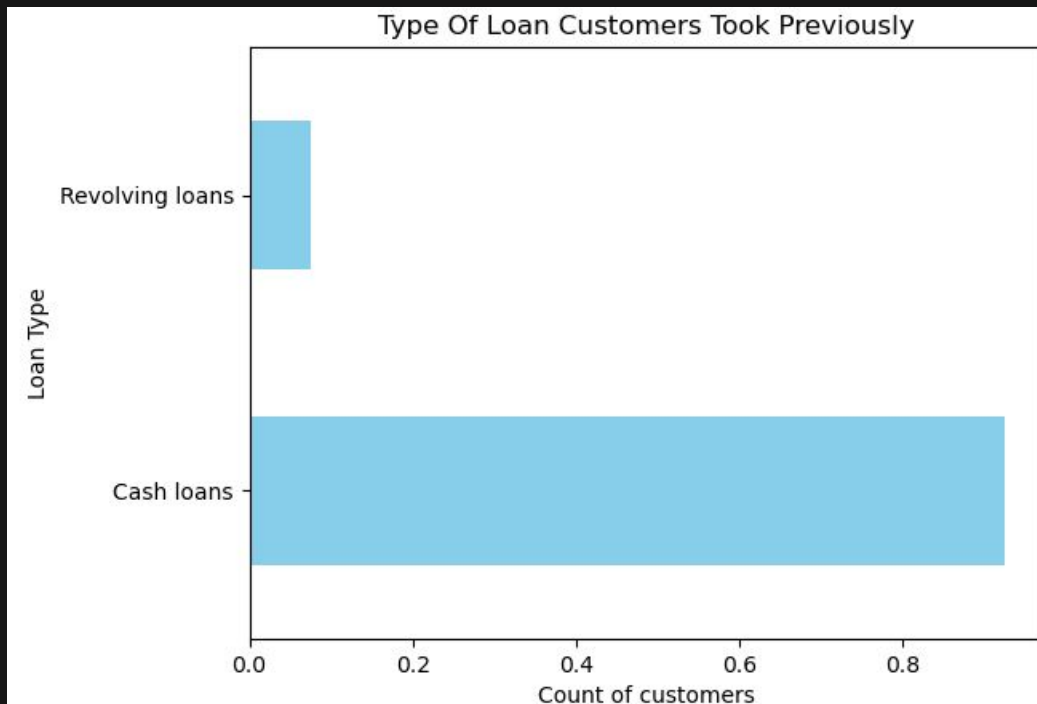
Multivariate Analysis



- People who do not default their loans have high correlations between Amount annuity and loan amount

Combined Data Set - Univariate Analysis

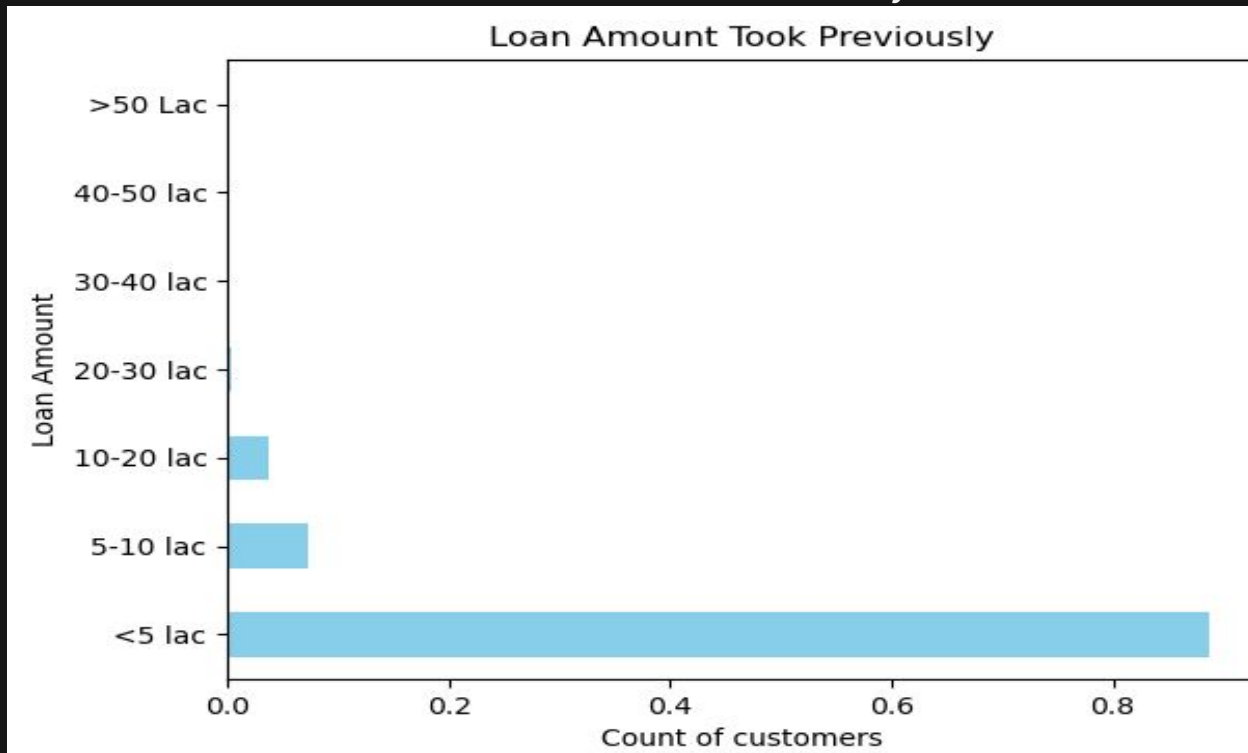
Type Of Loan Customers Took Previously



- Majority of people have taken cash loans

Combined Data Set - Univariate Analysis

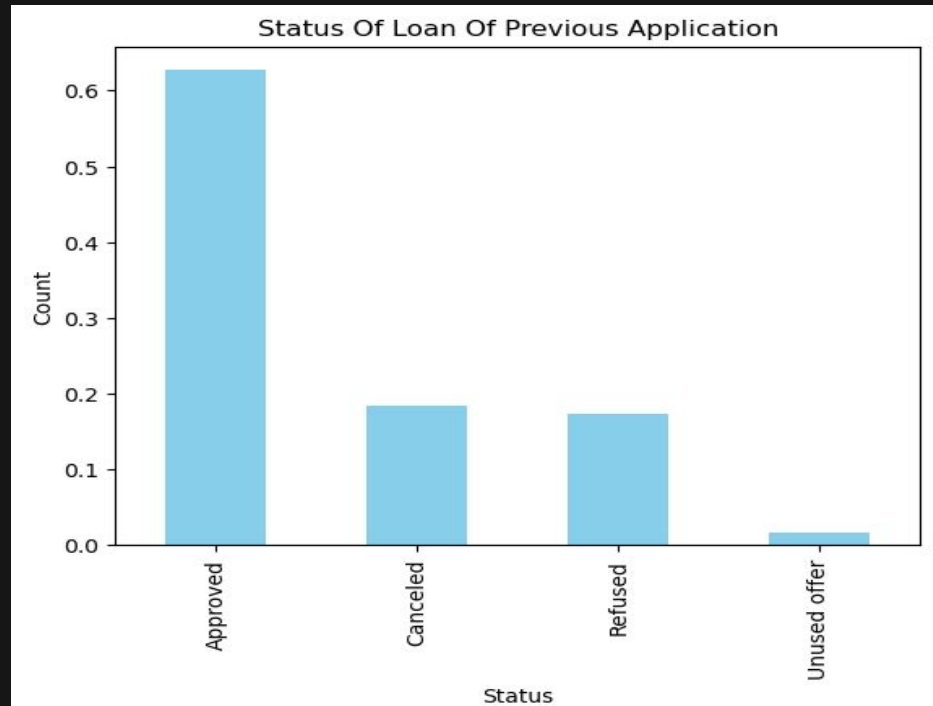
Loan Amount Took Previously



- Most of the customers have taken a loan of less than 5 lacs previously

Combined Data Set - Univariate Analysis

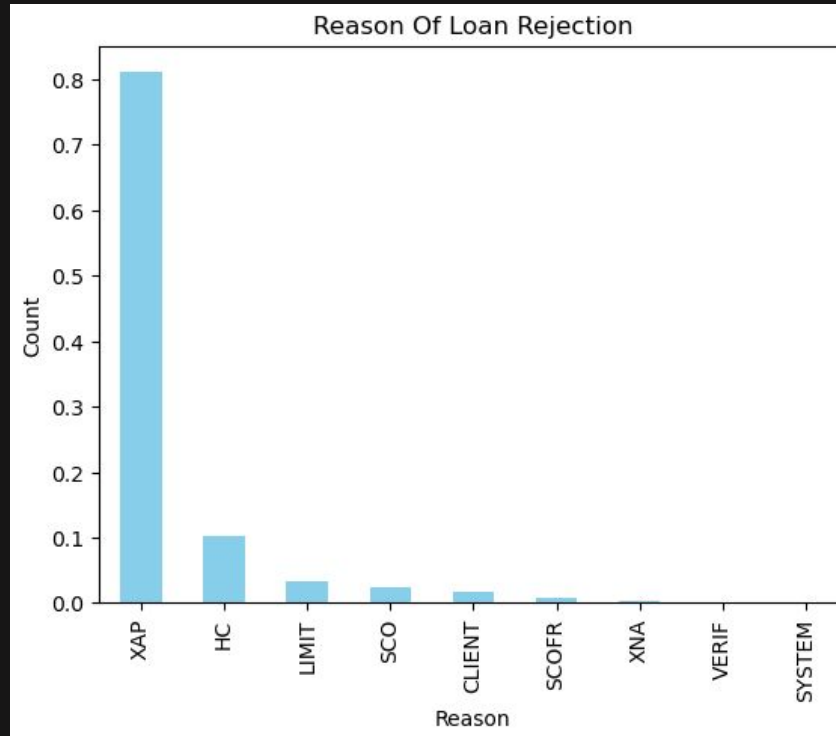
Status Of Loan Of Previous Application



- Most of the loans were approved previously followed by cancelled

Combined Data Set - Univariate Analysis

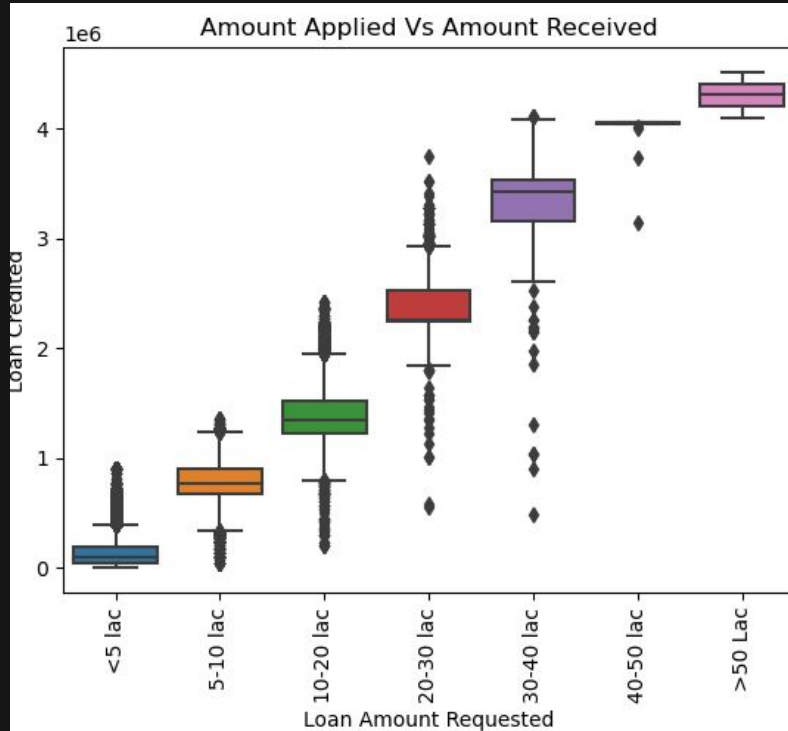
Reason Of Loan Rejection



- Though there is a value “XAP” but according to the data two reasons following the it is HC and LIMIT

Combined Data Set - Bivariate Analysis

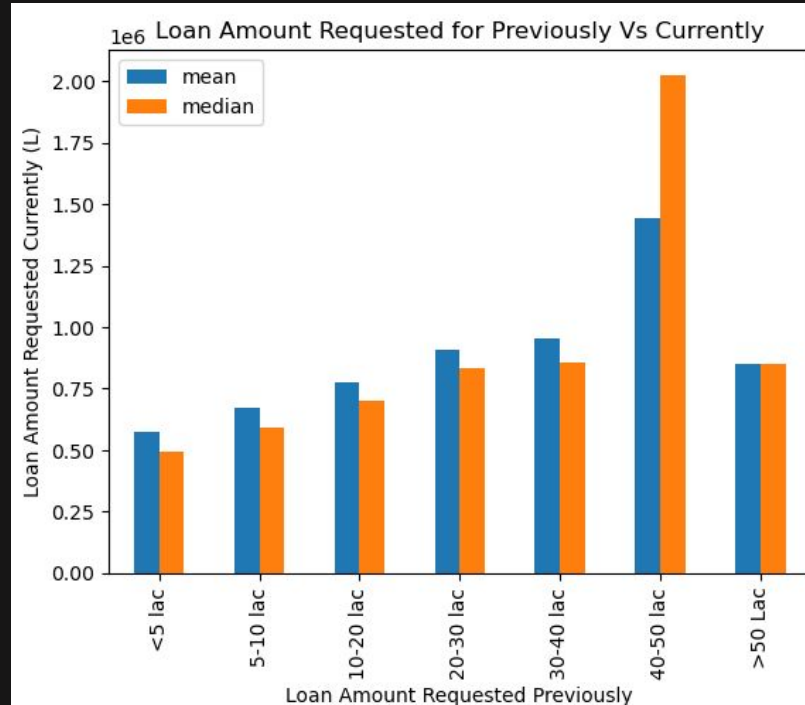
Amount Applied Vs Amount Received



- As the request amount increase, amount credited increases but people who requested for 30-40 lacs have most outliers in terms of getting lesser amount than what they requested for.

Combined Data Set - Bivariate Analysis

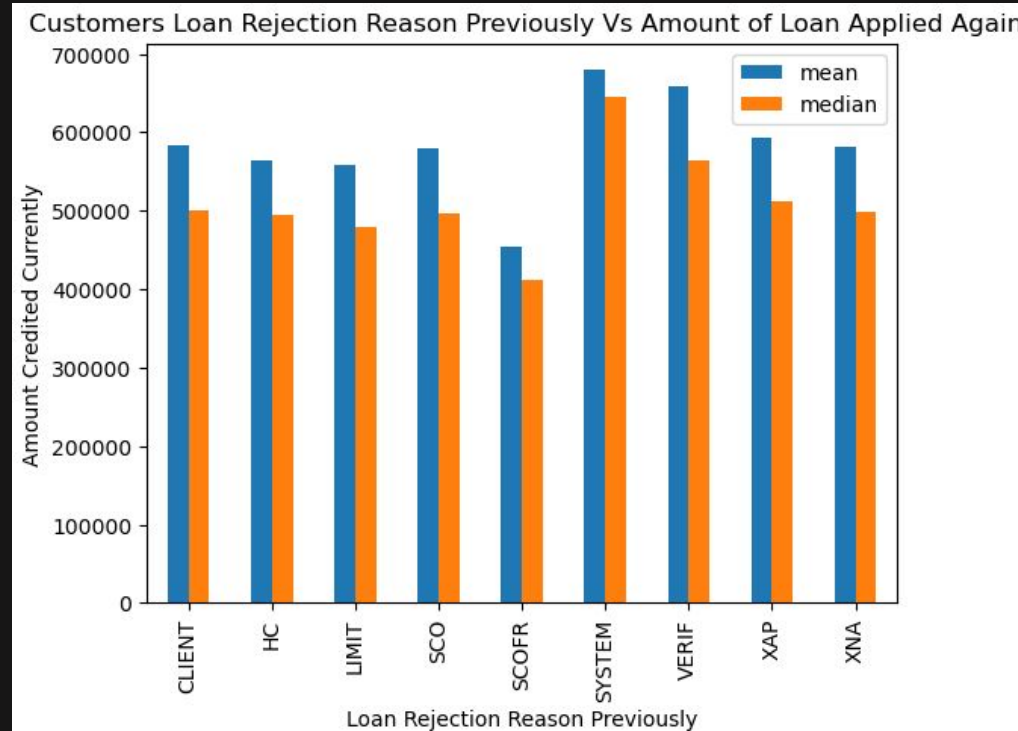
Loan Amount Requested for Previously Vs Currently



- People who requested for 40-50 lacs amount previously requested for higher loan amount

Combined Data Set - Bivariate Analysis

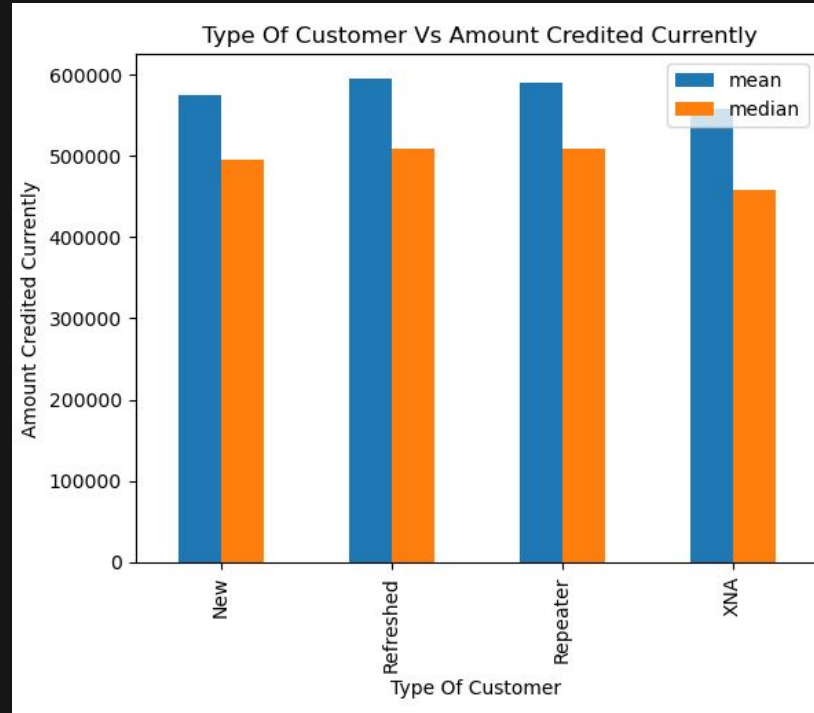
Customers Loan Rejection Reason Previously Vs Amount of Loan Applied Again



- People for whom the loan were rejected due to SYSTEM & VERIF had applied were loan again the most.

Combined Data Set - Bivariate Analysis

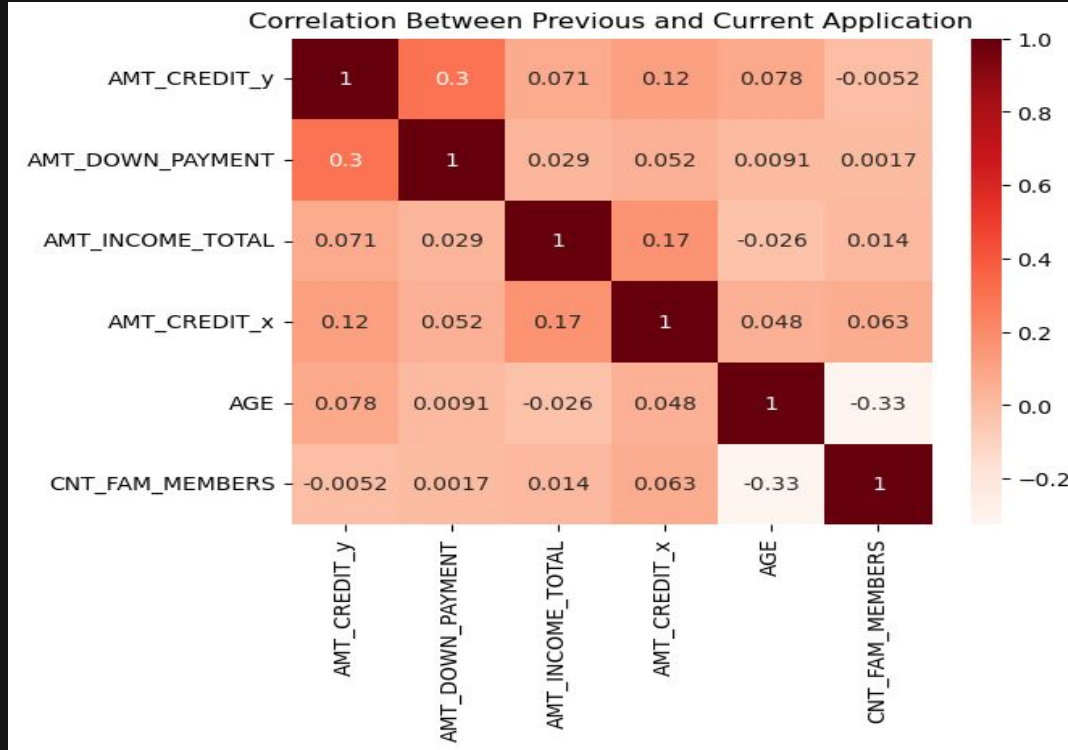
Type Of Customer Vs Amount Credited Currently



Refreshed or repeat customers received higher loans

Combined Data Set - Multivariate Analysis

Correlation Between Previous and Current Application



Amount of down payment done previously has high correlation between amount credited currently

THANKS