

LOAN DEFAULT EDA

By Shankar Roshan Valiathan

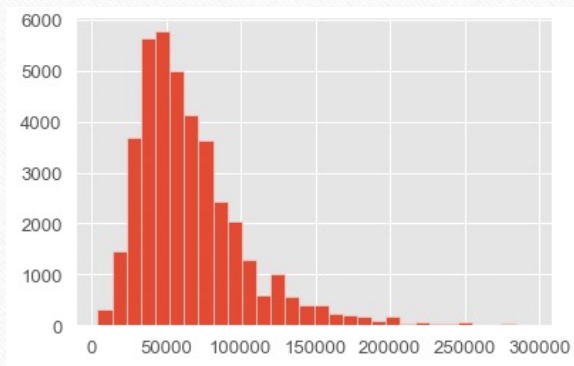
BUSINESS PROBLEM

- CONSUMER'S DEFAULTING ON LOANS CAUSES LOSSES FOR THE BANK/ FINANCIAL INSTITUTION
- HIGHER RISK CONSUMERS – MUST BE IDENTIFIED
- FINANCIAL INSTITUTIONS WANT TO GUARD THEMSELVES IN CASE OF DEFAULT

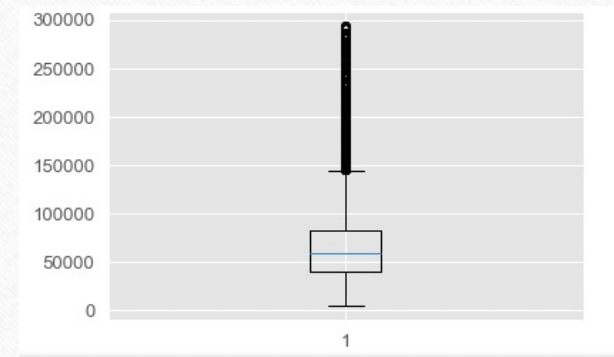
APPROACH TO SOLVING THE PROBLEM

1. Dissecting the variables that contribute to loan defaults.
2. Income, Debt levels, Loan amounts, Installment amounts are a few key variables.
3. Creating ratios from the above
4. Deciphering levels of risk from the given metrics.
5. Illustrating the same through graphs for better readability

Visualization : Income Distribution



HISTOGRAM: INCOME DISTRIBUTION

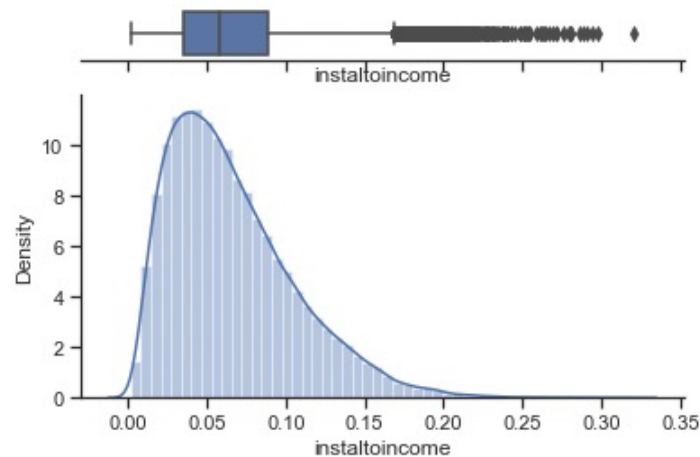


BOXPLOT : INCOME DISTRIBUTION

INFERENCE:

Most Loans are taken by the lower income group

INSTALLMENT TO INCOME RATIO



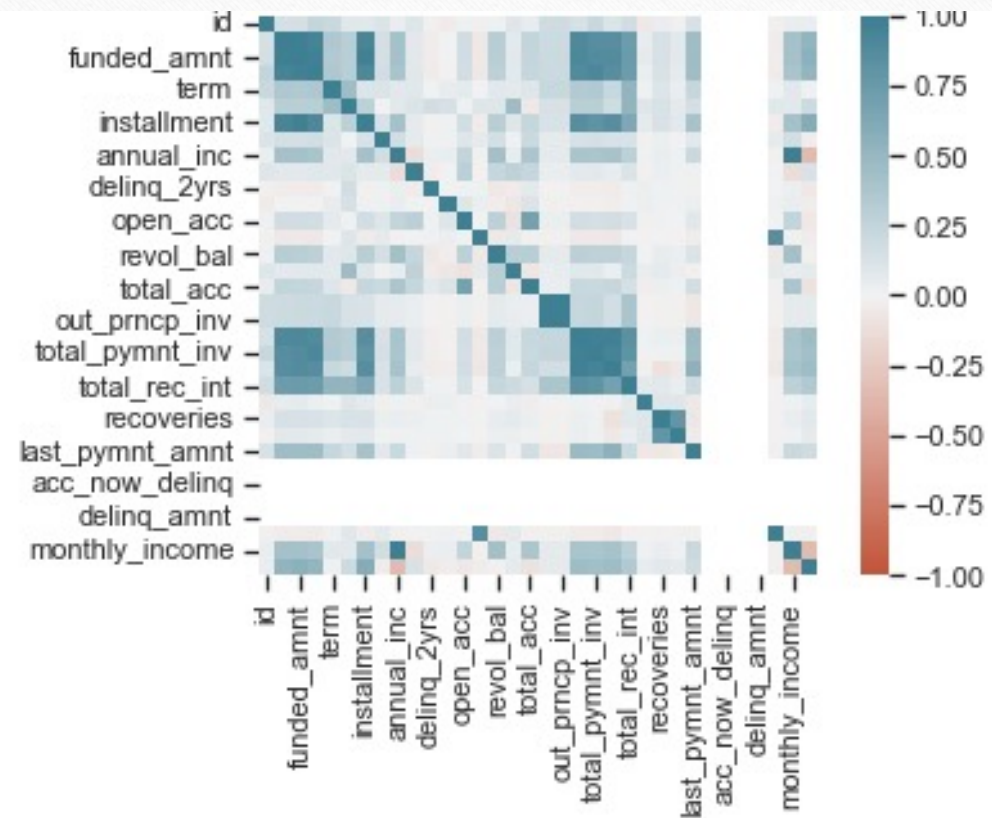
###From the above, we can clearly see the following:

1. When the Monthly installment is greater than 15% of the monthly income, there is a slight cause of concern, as it is to the tail of the distribution, and an anomaly.
2. The next step of analysis would be to check how many of these loans are defaults (Charge offs) compared to the overall.

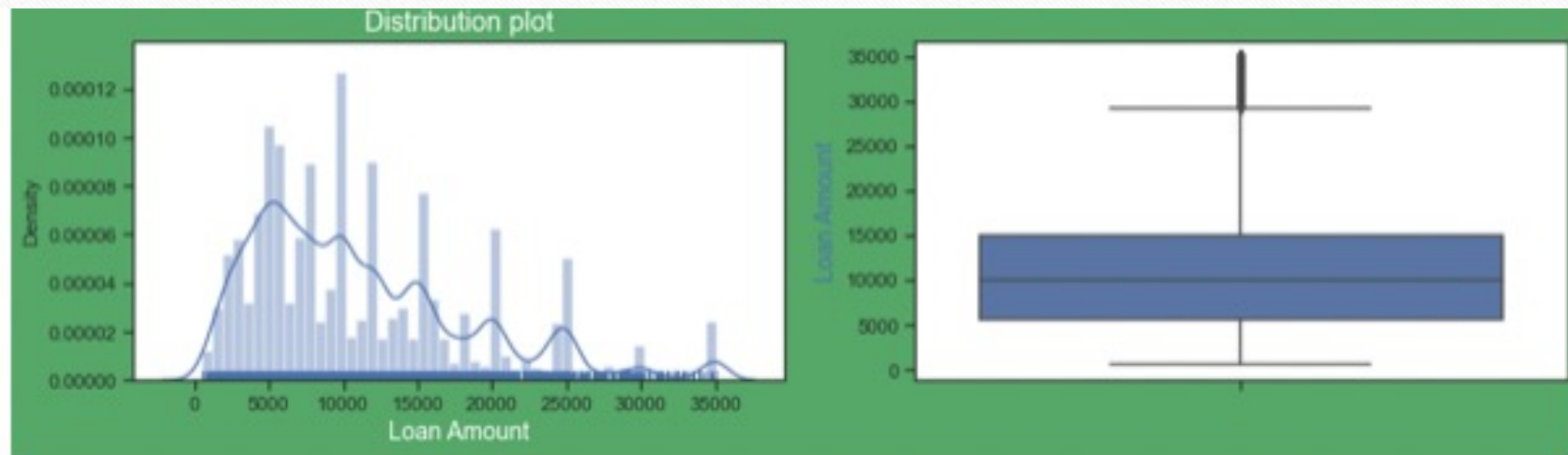
CORRELATION

Correlation matrix shows us the following important relations:

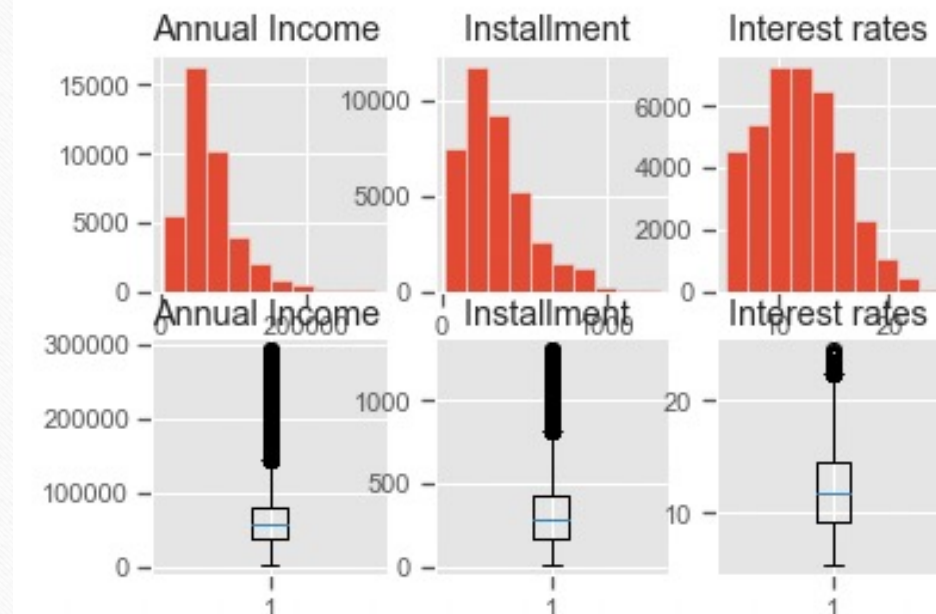
1. Delinquency amt vs Income
2. Recoveries and monthly Inc
3. Installment and monthly inc
4. Loan amt and monthly inc



LOAN AMOUNT



ANNUAL INC, INSTALLMENT, AND RATES



PURPOSE OF LOAN

