LENDING CLUB CASE STUDY

Group Member & Facilitator : Ravy Chandrasekaran

Problem Definition

- The largest online loan marketplace company facilitates personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.
- Lending loans to 'risky' applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the **defaulters**
- Find the driving factors behind the defaulters
- Doing risk analytics bring the observations of the defaulters to help the lending company to avoid investing in the wrong applicants
- Have to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss

Overall Approach

- Perform cleaning of the data by dropping the unnecessary columns, fill the missing values
- Close to 54 columns doesn’t have any values. Dropping them will provide a clean good analytical data
- Some Target variables are having missing values, filling them with the closest values from the existing column values
- Converting the objects to necessary numeric values to derive a solution
- Removing all the null values and dropping the unnecessary columns will give a compact data for analysis
- Analyse different Target Variable alone and combinedly to derive solution

Observations: using Univariate Analysis

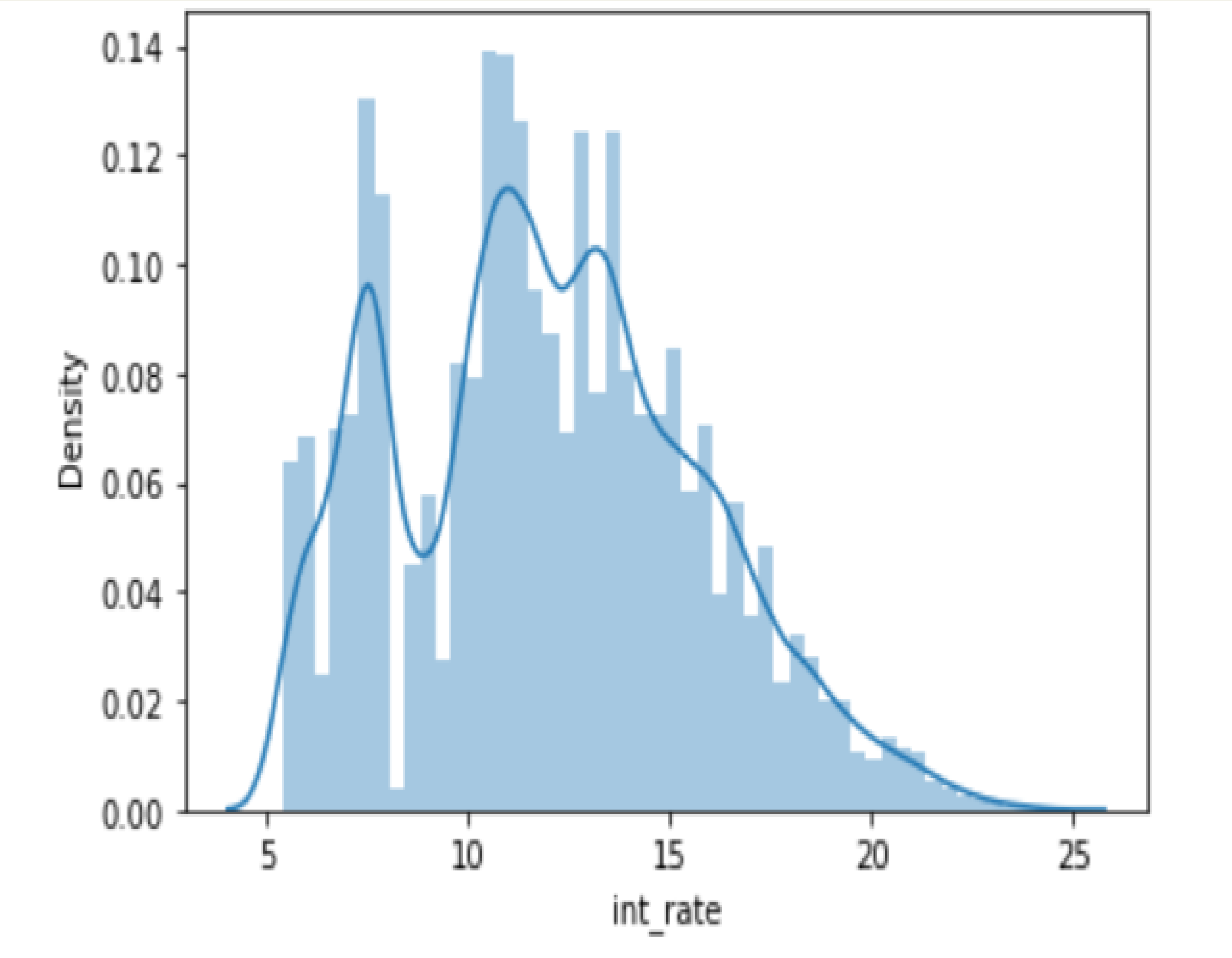
- In the Annual income few salaries are very high range which are not close to the 95% of ones
- After removing the outliers the spread reduced and the median salary value is changed to 57k after removing the outliers
- Analyzing loan amount, funded amount and funded amount invested are funded /invested same way and don’t see any difference
- The peak of Interest Rates on loans are in range between 10 – 14%

Observations: using Bivariate Analysis

- From the Bivariate analysis the purpose of the loan taken is majorly for small business
- Second one is the renewable energy takes the majority of the loan
- From another analysis, Grades F & G takes more loan amount compare to the Grades A, B & C

Plotting for Interest Rates on Loans

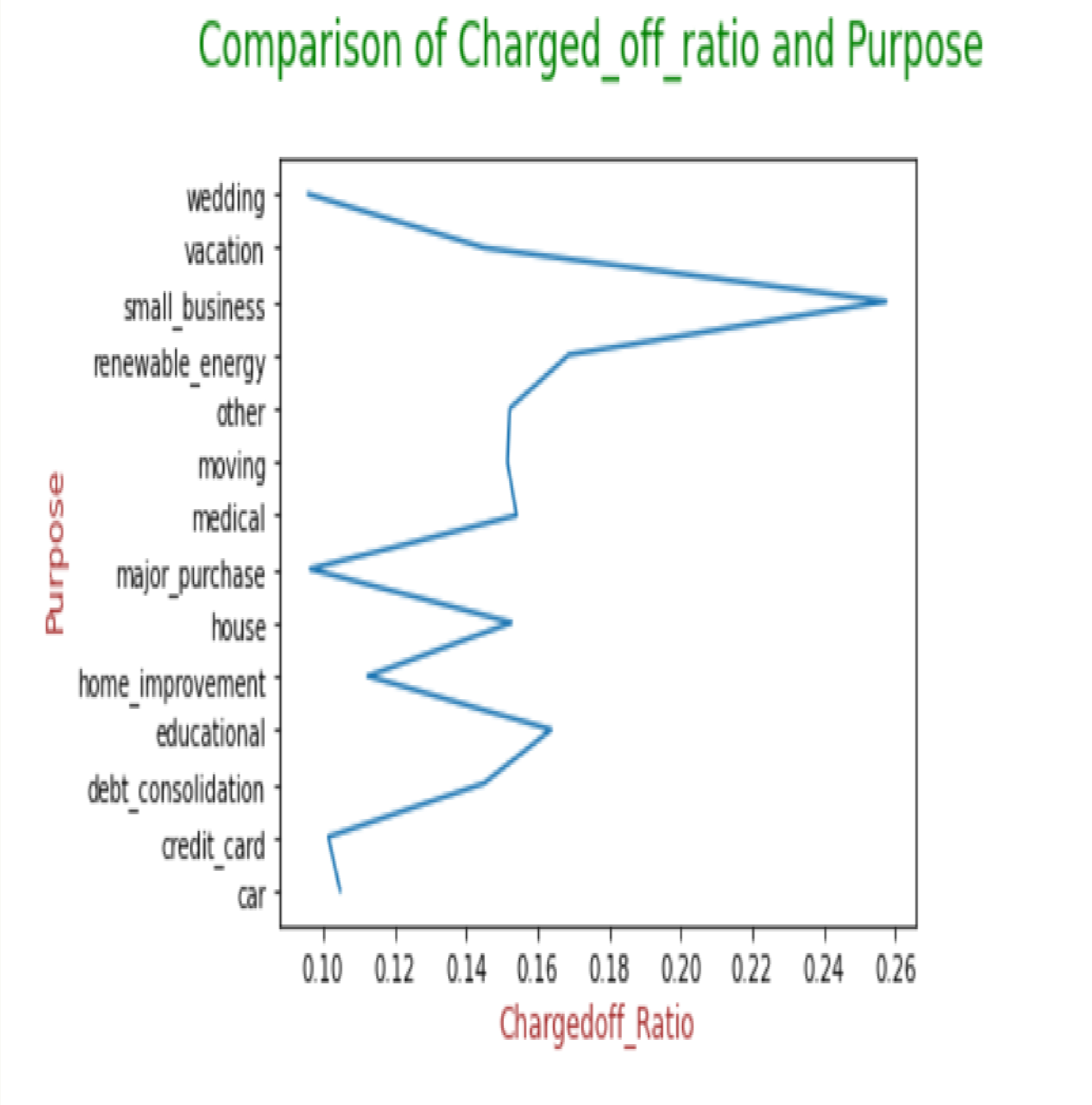
Observation: The distribution plot show that most of the Interest Rates on loans are between 10% - 14%



Comparison of Charged off Ratio and Purpose

Observation : From the above line graph we can find that Small Business category has the highest Ratio

From this Small Business contributes to the higher loan amount



Multivariate Analysis on different Target Variables

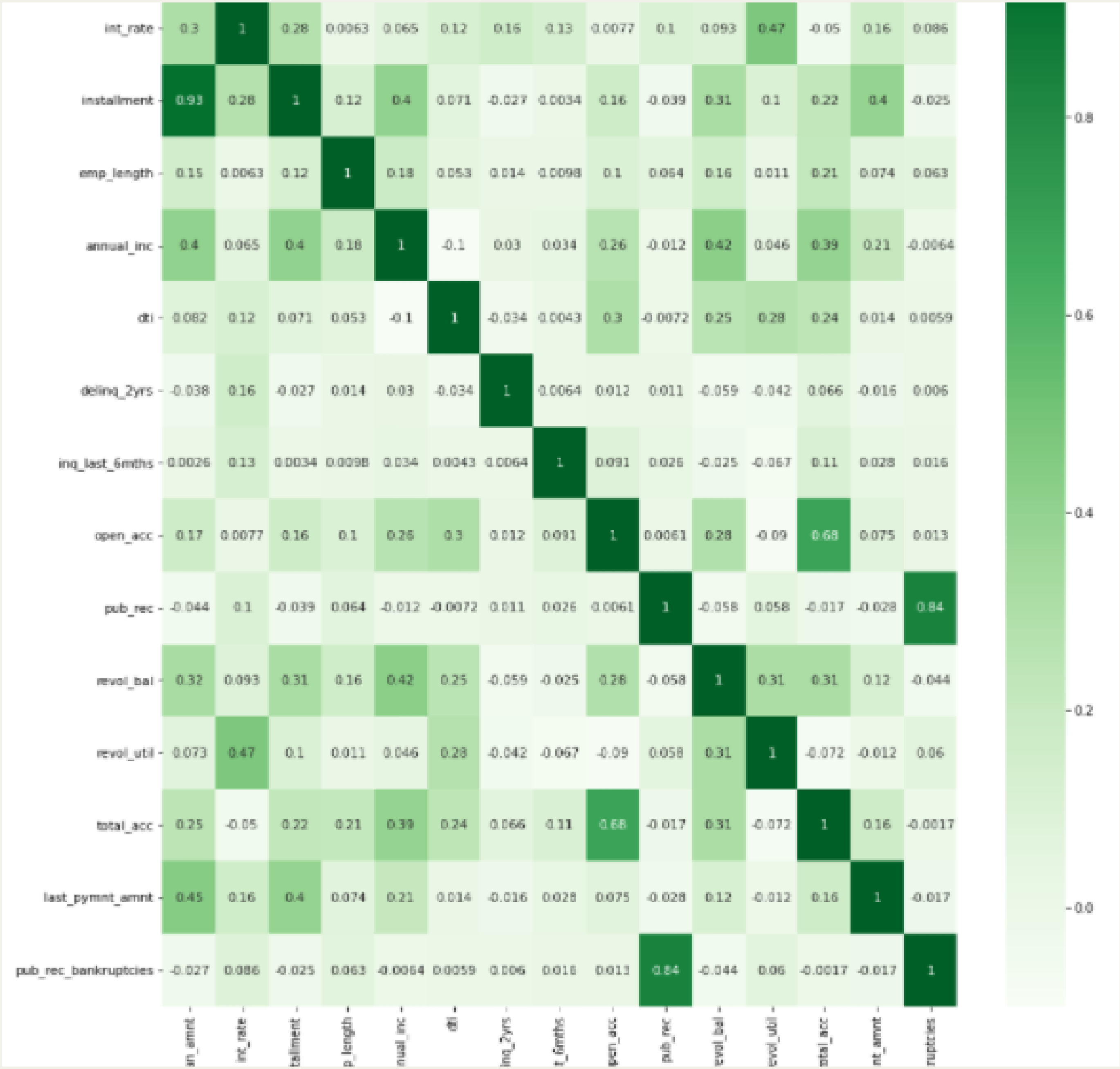
Observations :

Loan Amount and Installments are positively correlated

Derogatory Public Records and Public Record Bankruptcies are positively correlated

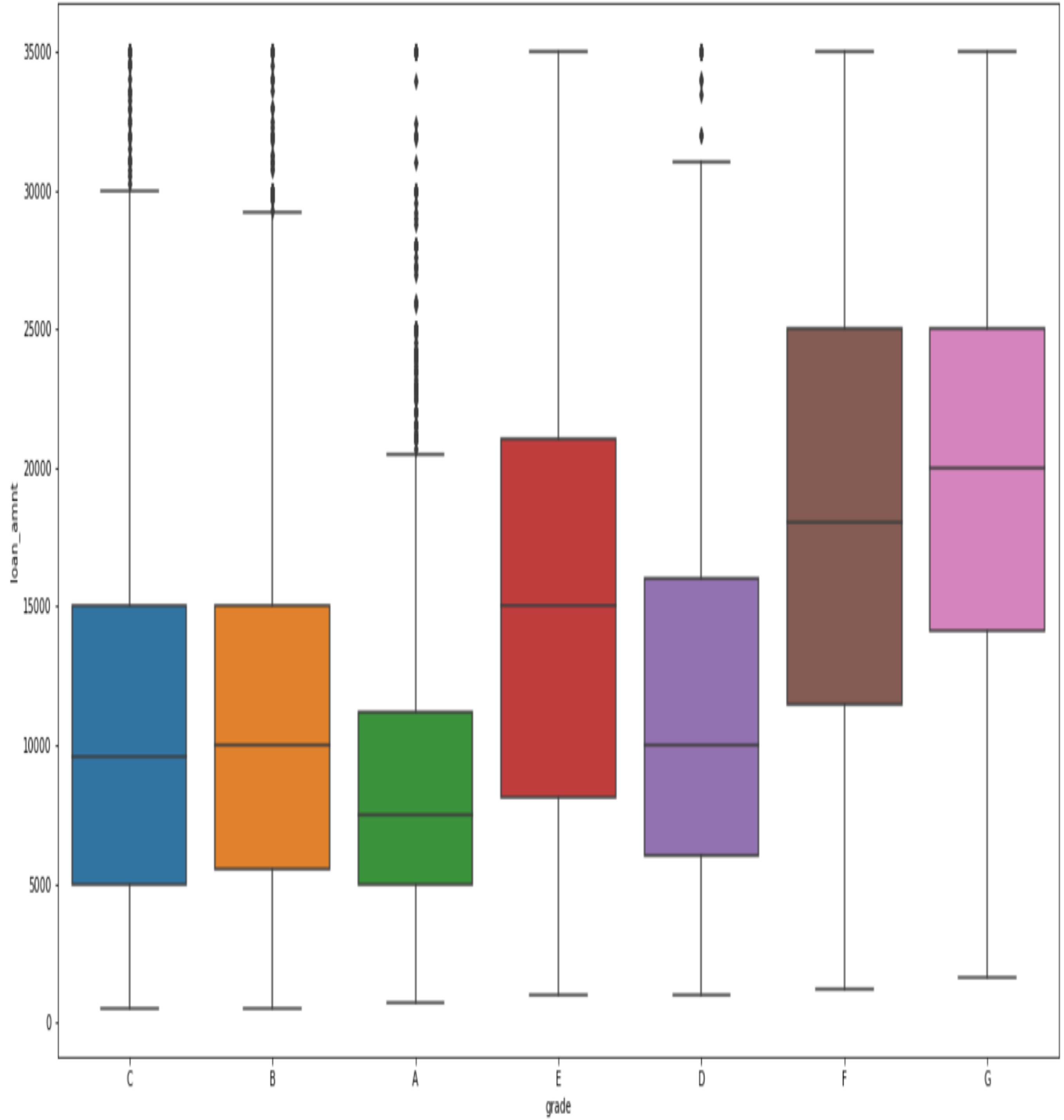
Annual income and Debt income Ratio is negatively correlated

Employee’s years of experience and annual income are positively correlated



Analysis on Grades vs Loan Amounts

Observation: The Grades A,B & C are having lesser loan amounts comparing to Grades F & G. Grades F & G are taking higher loan amounts



Conclusion

- Based on the different analysis and observations
  - *Some grades of people are risky to invest who are defaulters*
  - *Based on Annual salary and other Target factors some people are good in repaying the loan amount*
  - *Lending company to understand the pictures provided and act upon to whom they can lend the money*