LENDING CLUB CASE STUDY

Bhargava Teja Maddula

PROBLEM STATEMENT

- A largest Company which facilitates personal loans, business loans, fainancial loans and also for medical procedures.
- Understanding the Customer Loan Data, Getting the Defaulters from the variables.

STEPS

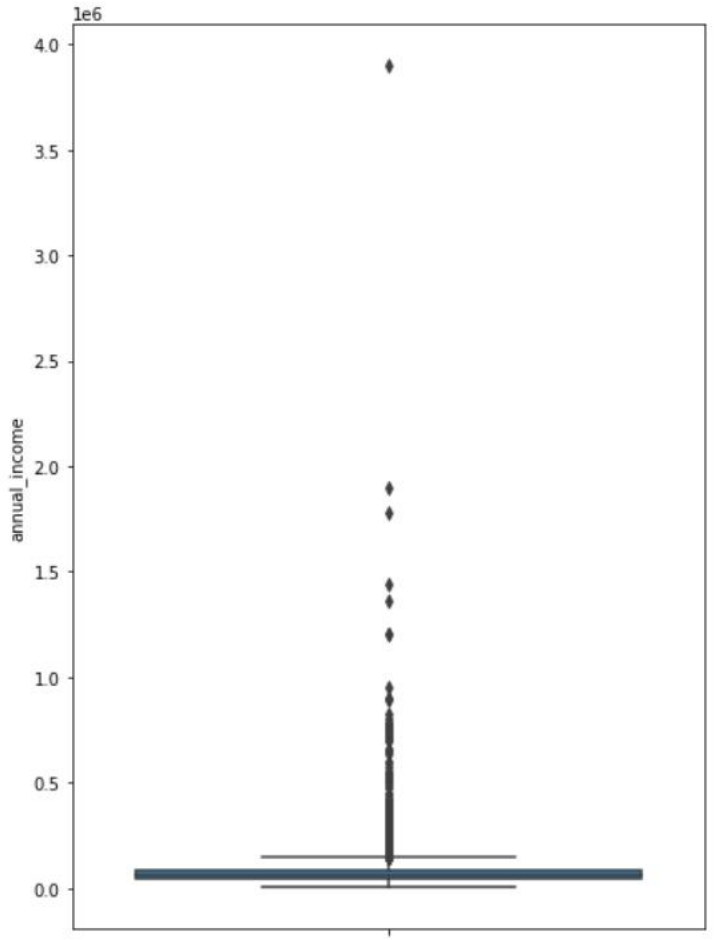
- Importing Modules
- Loading Data
- Data Cleaning
- Detecting Outliers
- UniVariate Analysis
- BiVariate Analysis
- Observation

Data Cleaning

- Removed Duplicate , Unwanted Observations for both rows and columns
- Outliers was removed

Handling Outliers

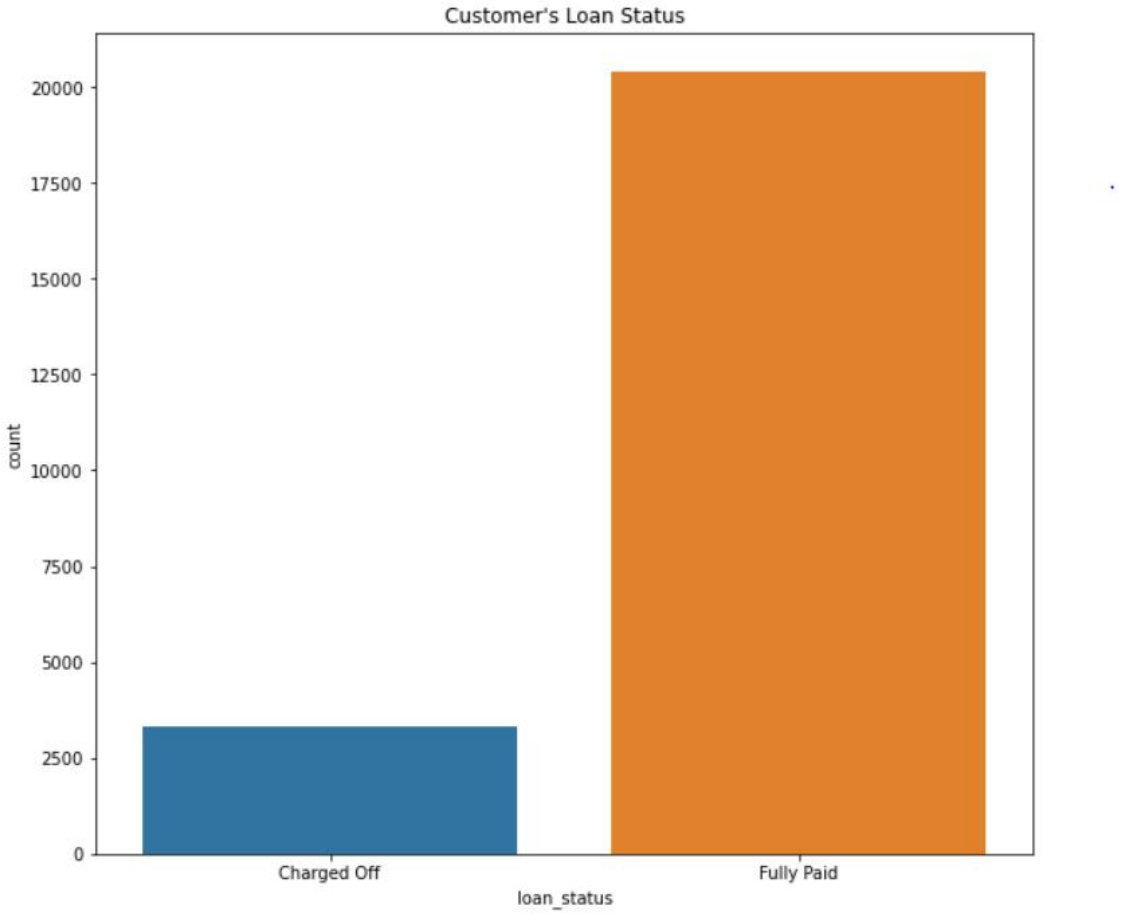
Taken the annual income column
in the loan data
and only one value is outlier,
so have removed that



Univariate Analysis

Observations :

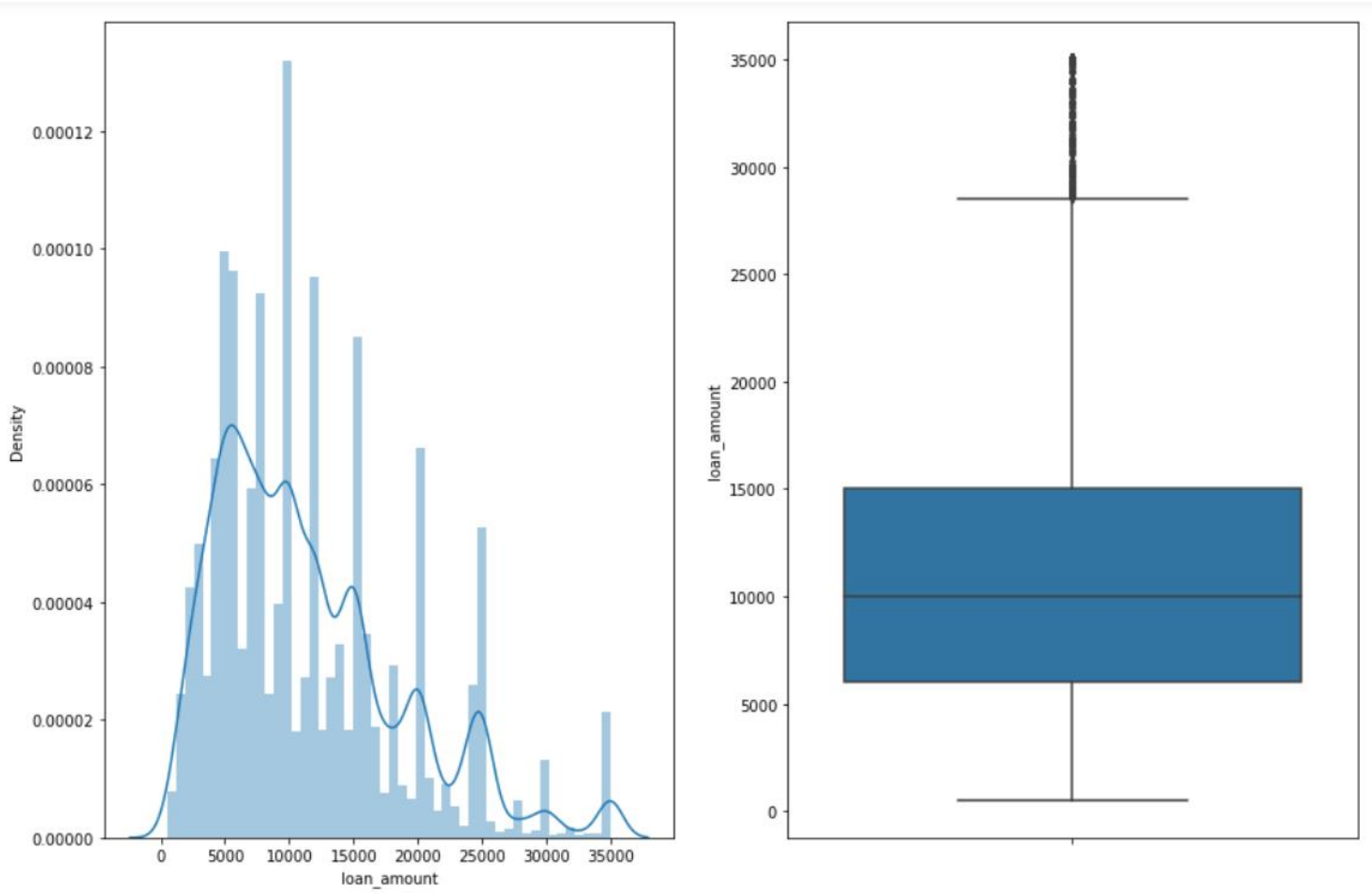
most of the customers are fully paid
and
most charged off are around
3000 to 5000



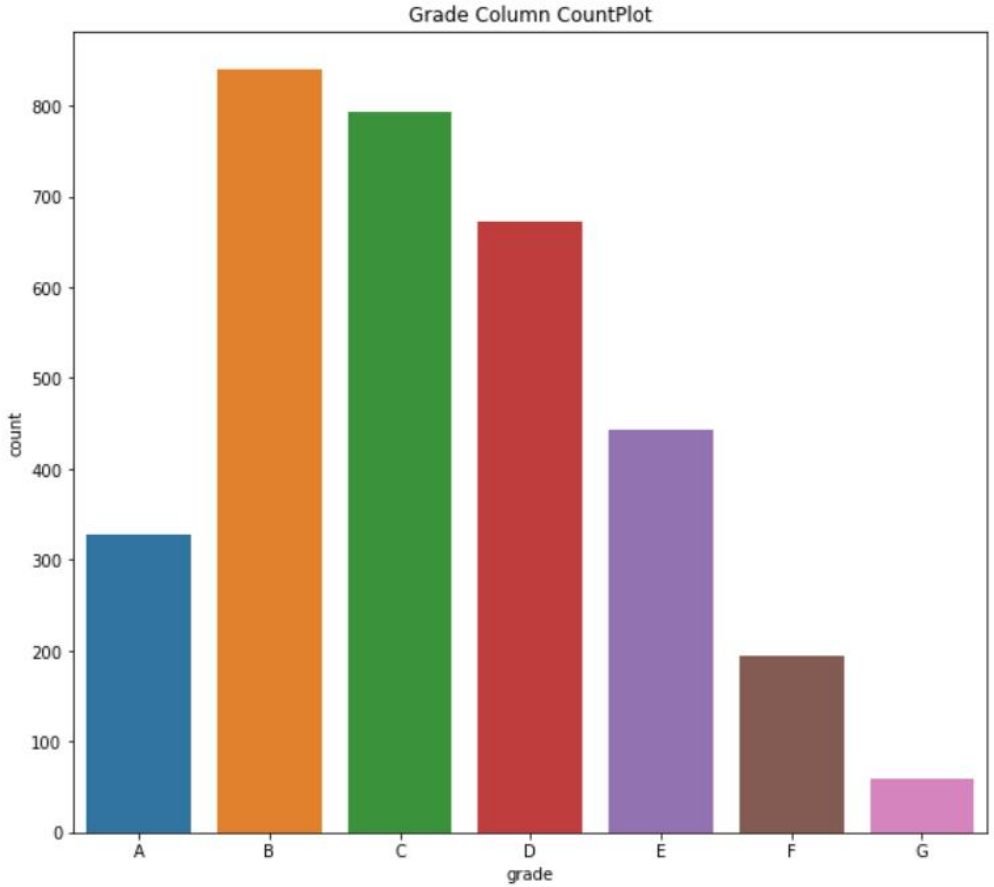
Univariate Analysis

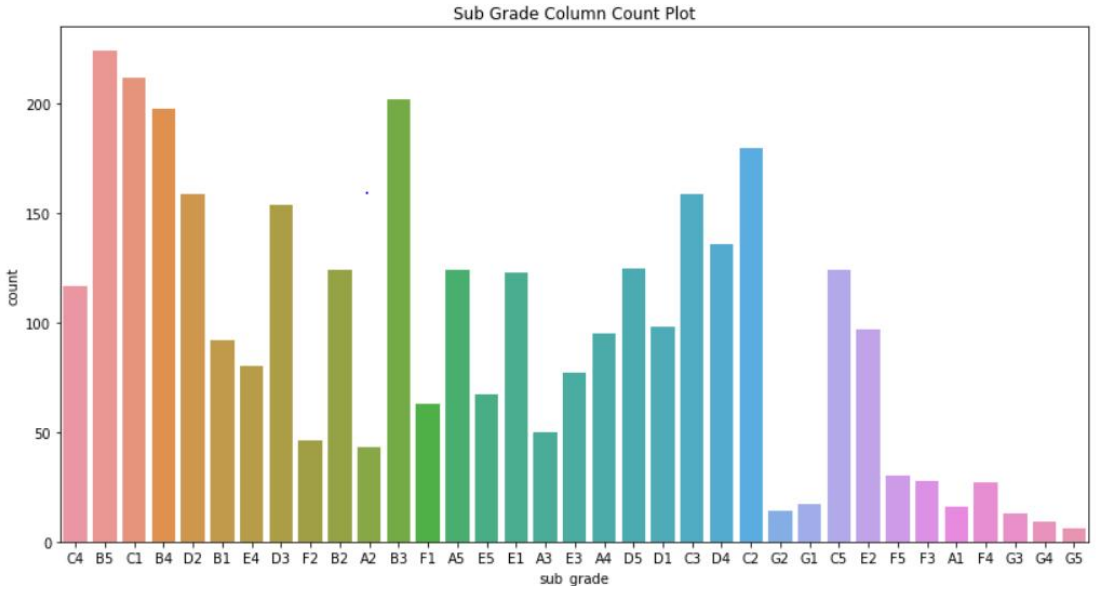
Loan Data:

Most of the customers have taken loan of amount 10,000 and also median distribution is 10000 only



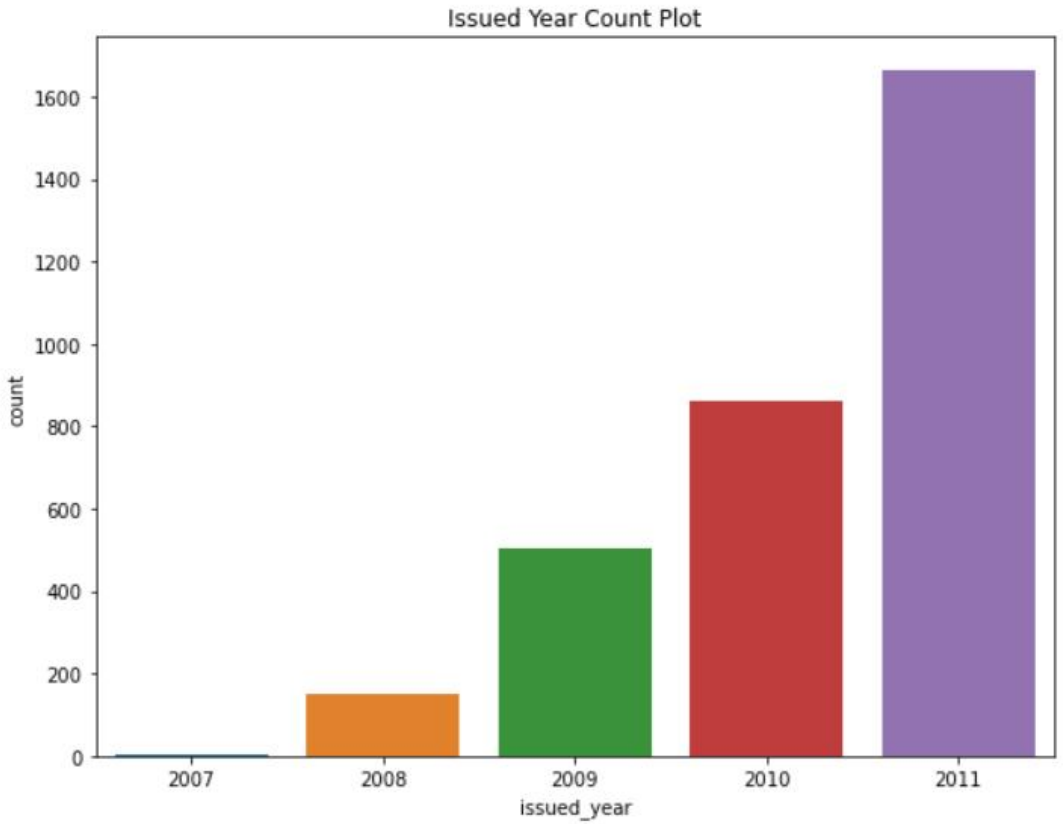
Grade and Subgrade :





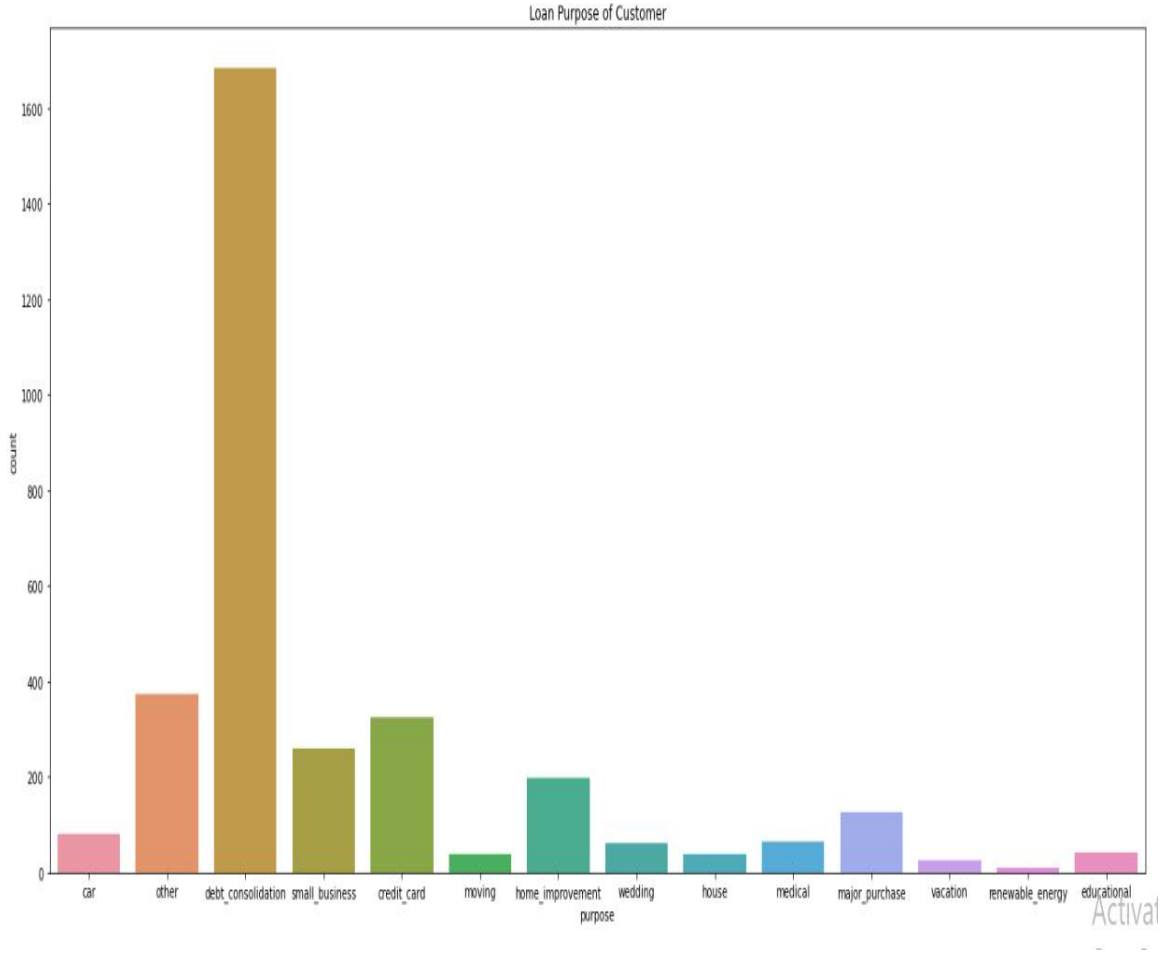
In the Loan Data Majority of the Defaulters are B Grade which are highly possible.

Issued_Date:



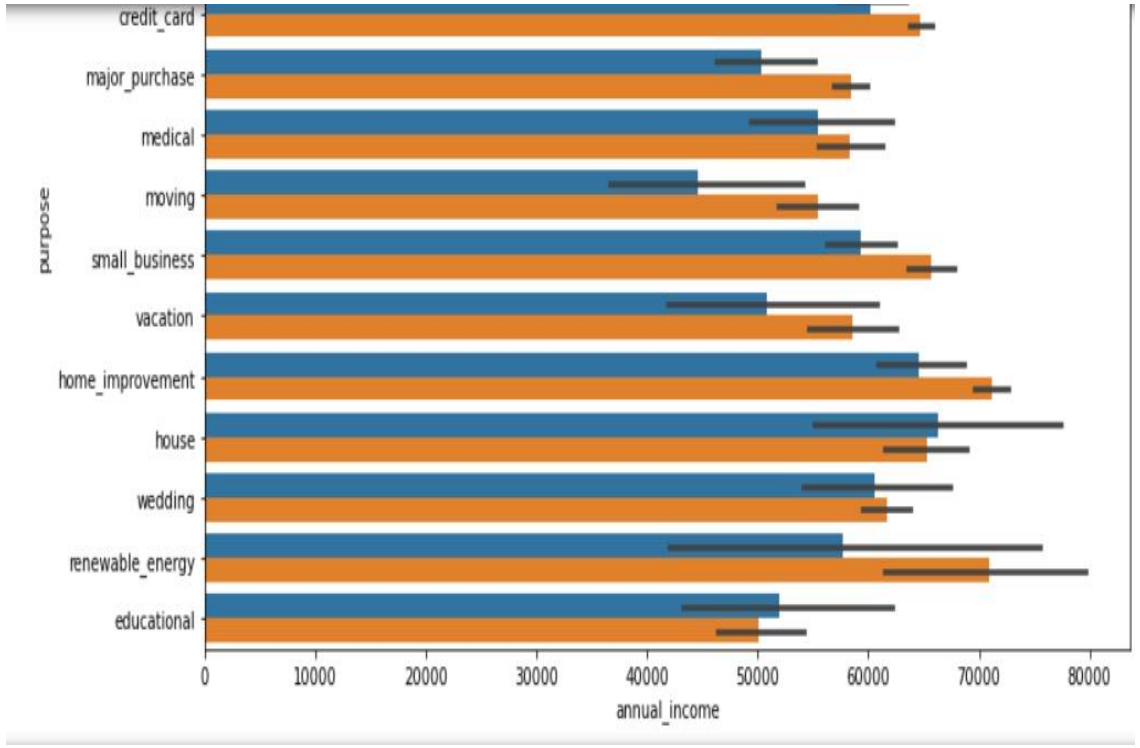
- Based on the Issue date taken Years from that we got to know at 2011 most of them are defaulters
- Also Observed Linear Increase of Defaulters

Purpose Distribution :



Loan Purpose of the customers are mainly due to debt_consolidation.

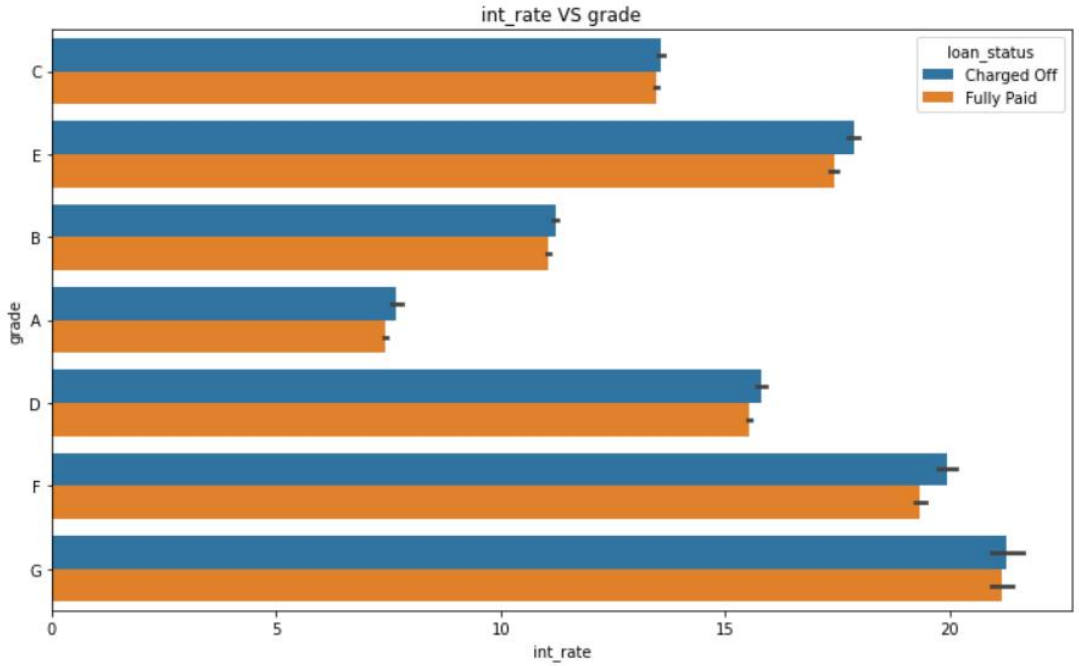
Bivariate Analysis



Annual income vs Purpose :

- 60k -70k annual income , who has taken loan for house and home_improvement are defaulters

int_rate vs grade :



- G grade customers having most interest rate as well so these are the Defaulters

Observations & Conclusions

- Customers who are rentals, annual incomes who have around 30k to 60k and purpose is debt_consolidation are moslty defaulters.
- Average laon amount is 10000.
- Grading System is working as expected.
- Most of the Data has Fully Paid Customers.
- In observations we have concluded who is the defaulters.

THANK YOU