



Lending club case study

BY

PRITHVIRAJ V

Data Cleaning

- Following data cleaning process are done
 - Finding Missing data columns
 - Found huge amount of columns are having null values so the column is useless dropped those columns
 - Modifying the data types
 - Some columns data are considered numeric but the type is object because some string values are appended with it. In those columns lambda function is applied.

2. Extracting different data type

- Univariate Analysis
 - This session focuses on analysing variables one at a time, as the term “uni-variate” implies. Before analysing numerous variables together, it’s critical to first analyse each one individually.
 - Continuous Variables
 - When dealing with continuous variables, it’s important to know the variable’s central tendency and spread. Statistical metrics visualization methods such as Box-plot, Histogram/Distribution Plot, Violin Plot, and others are used to measure these.
 - Categorical Variables
 - We’ll utilize a frequency table to study the distribution of categorical variables. Count and Count percent against each category are two metrics that can be used to assess it. As a visualization, a count-plot or a bar chart can be employed.
-

Basic numeric analysis of percentage of defaulters

Percentage of people who paid their loan are: 85.41 %

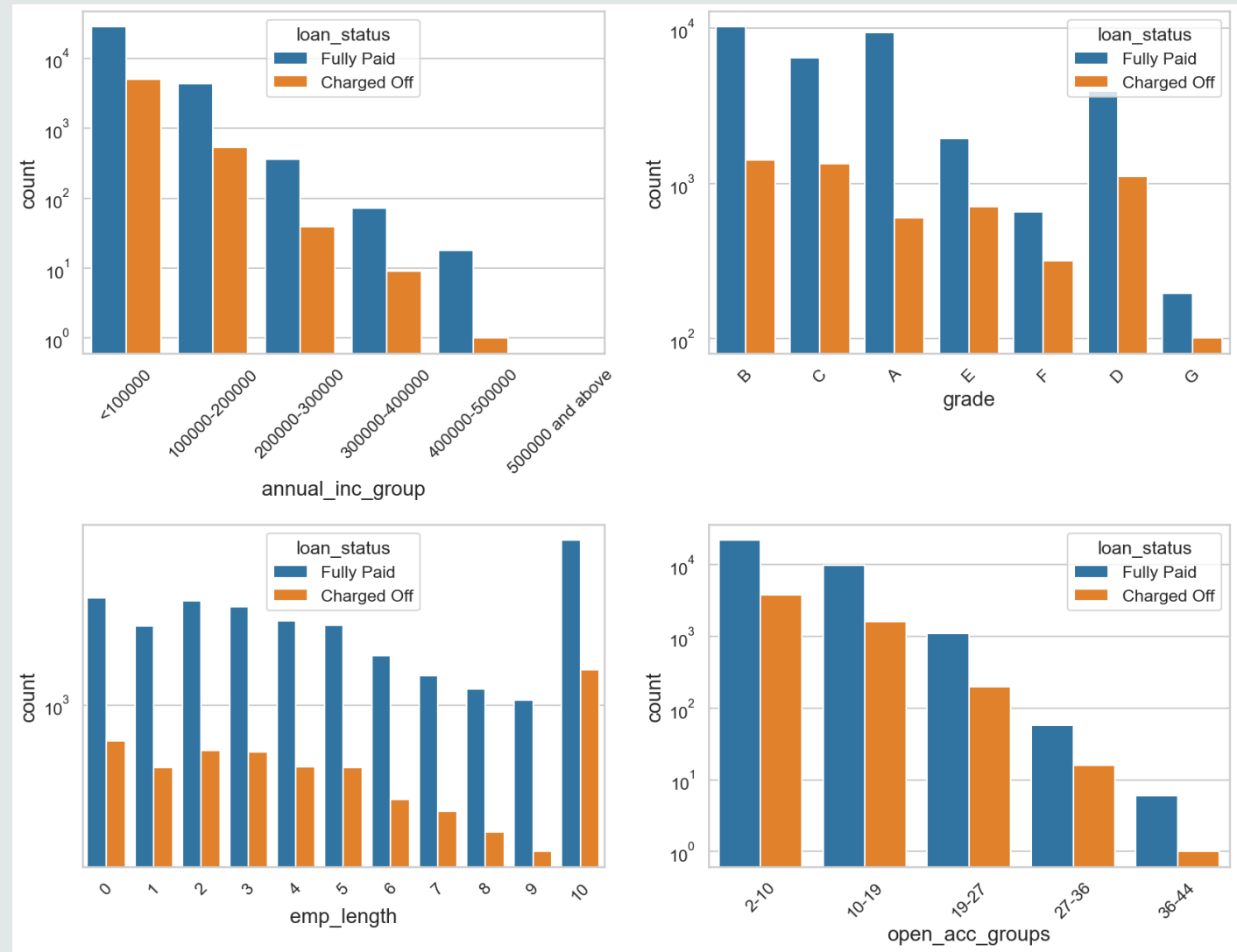
Percentage of people who did not paid their loan are: 14.59 %

Univariate Analysis

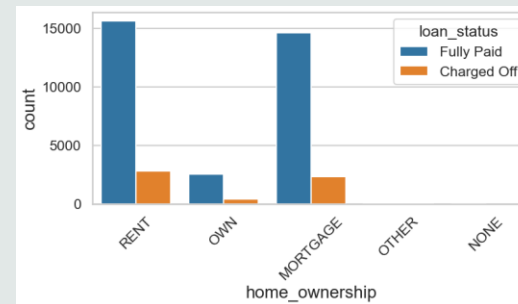
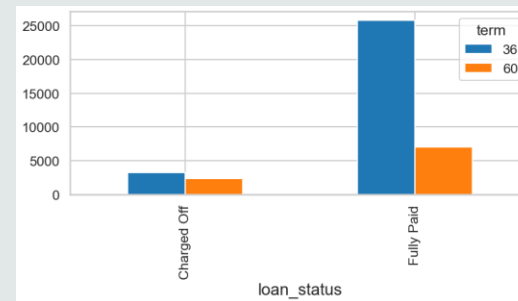
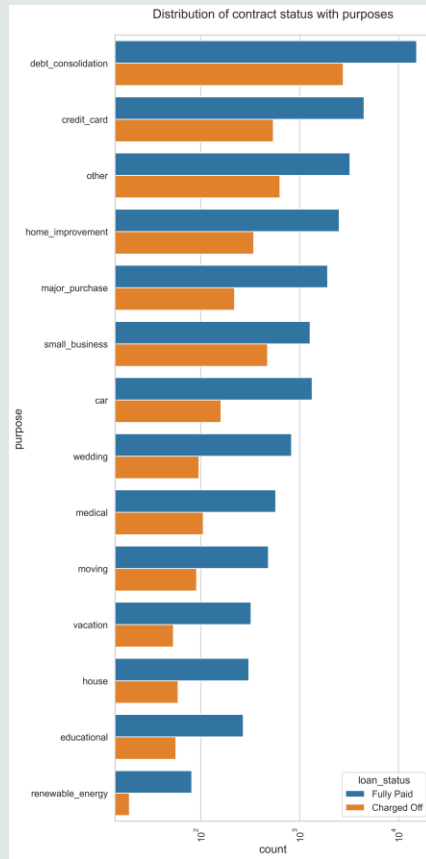
Based on the above comparison

Found that People with `Grade A` and annual salary between `('400000-500000')` and employment of `9 years` are good to provide loan they generally pay loan properly

Note: Should be careful with people on grade D



Univariate Analysis

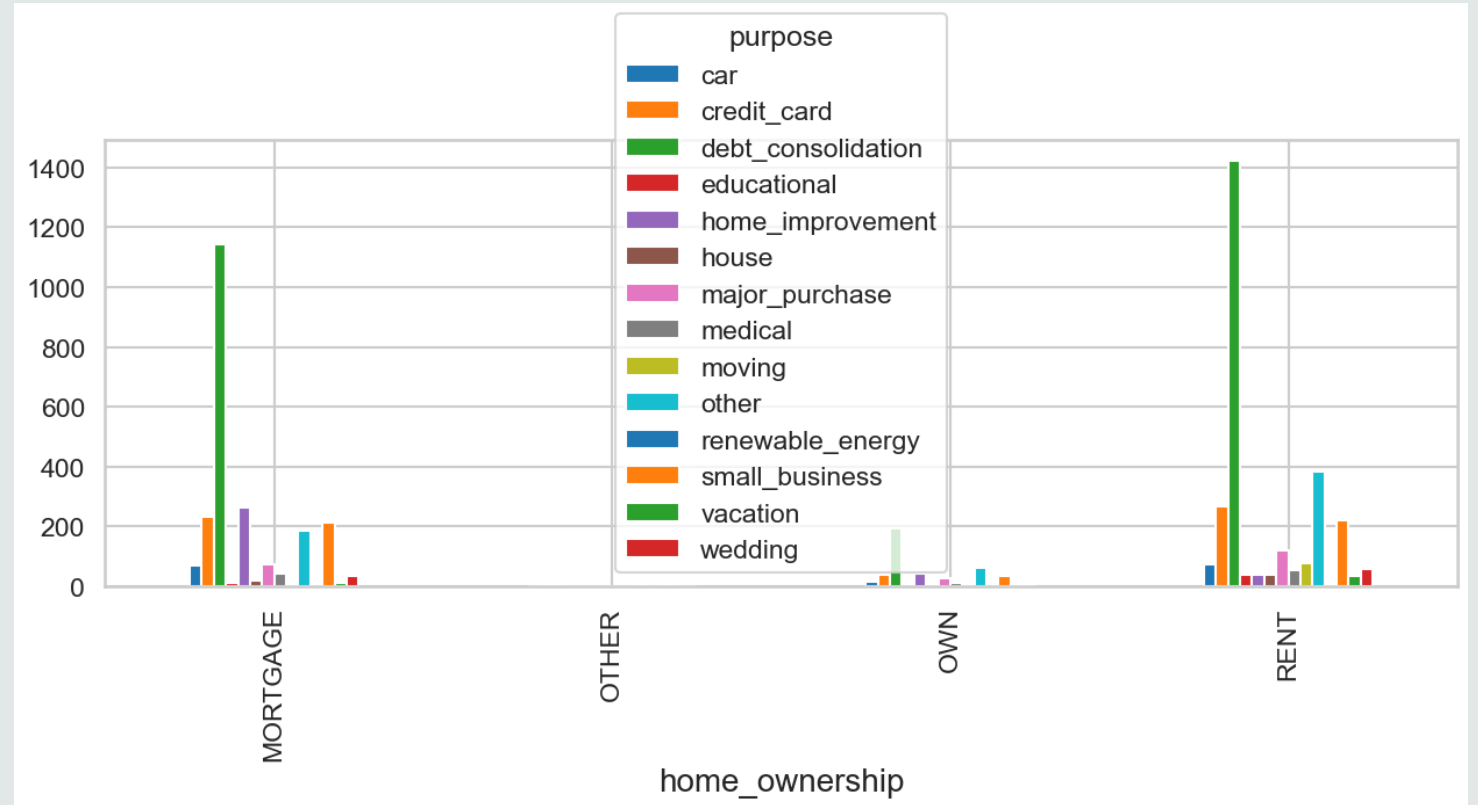


- Based on purpose
 - Education, Renewable Energy, House, and vacation are repaid properly
 - High term of tenure is helping people to repayment of any salary group
- Based Home ownership
 - Found that people owning own house generally pays loan properly

Bivariate analysis

Debt Consolidation are the most defaulted.

Need reduce approving loan to this category of application.



Based on pivot table construction

- Will consider the salary based on the house type
 - - MORTGAGE greater 80000
 - - OWN greater 50000
 - - RENT greater then 50000

Summary

- Will consider the salary based on the house type
 - MORTGAGE greater 80000
 - OWN greater 50000
 - RENT greater then 50000
 - Based on the above comparison
 - Found that People with any grade and annual salary between ('~~400000~~-500000') and employment of **9 years** are good to provide loan they generally pay loan properly
 - Note : Should be careful with people on grade D
 - High term of tenure is helping people to repayment of any salary group
 - Employee above 10+ years of experience are defaulters most probably
 - Higher the interest rate more the 20% are making less repayment
-