Lending Club- Case Study

Contributors:-

Mohammad Asif Bagewadi

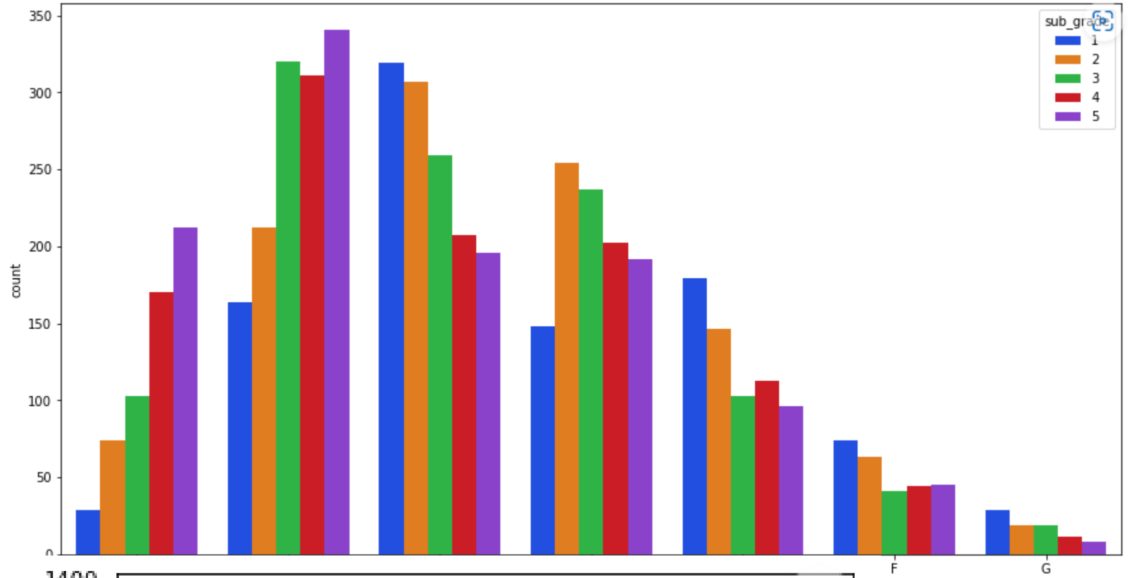
Hemant Singh

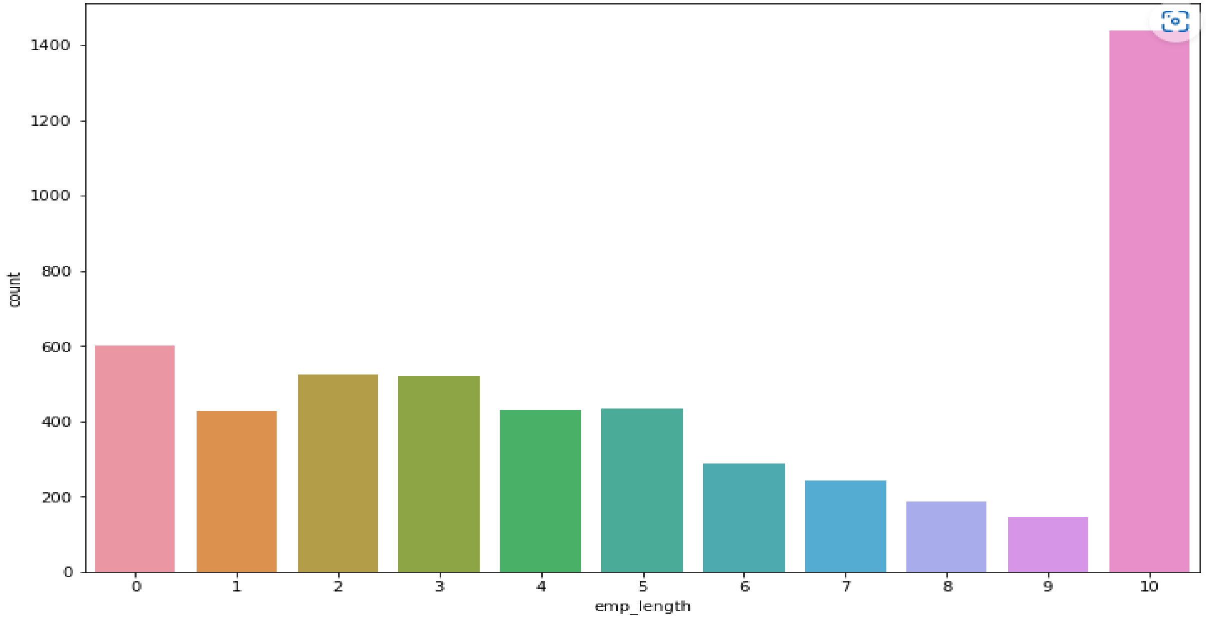
Abstract

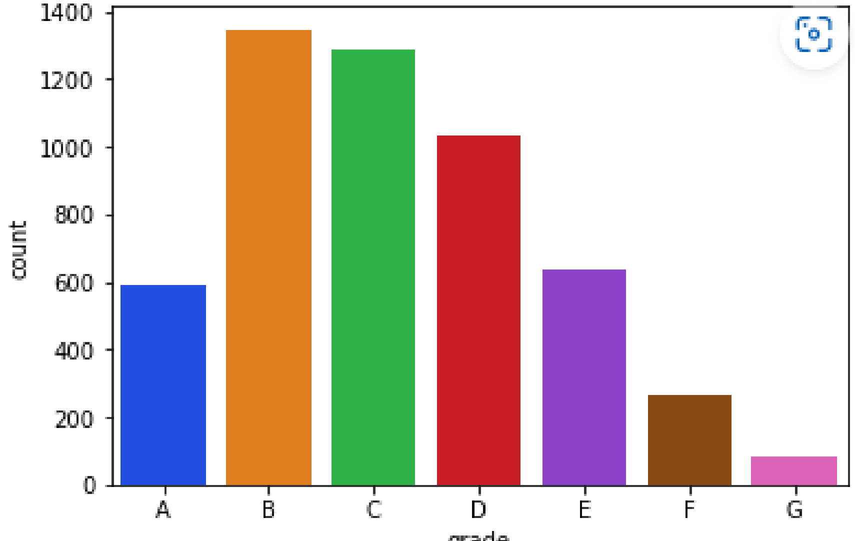
- Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.
- Borrowers can easily access lower interest rate loans through a fast online interface.
- The objective of analysis is to use the information about past loan applicants and find whether they ‘defaulted’ or not.

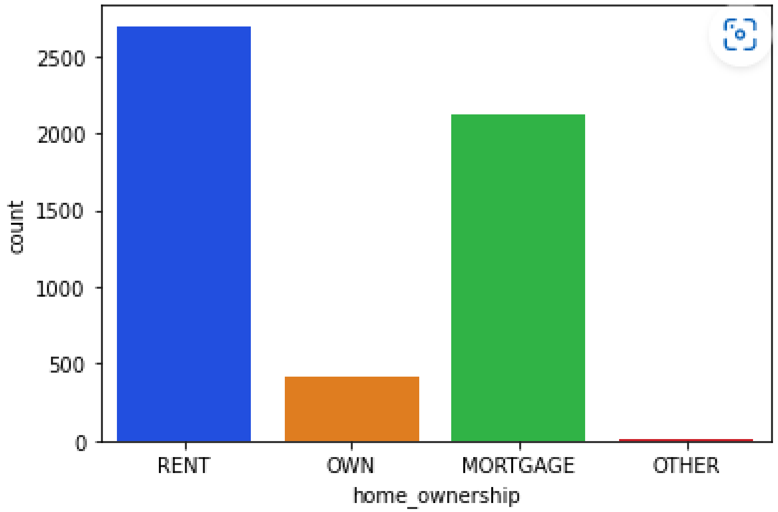
Problem Solving Methodology

Univariate plots- All derived from the dataset containing charged off loans only.





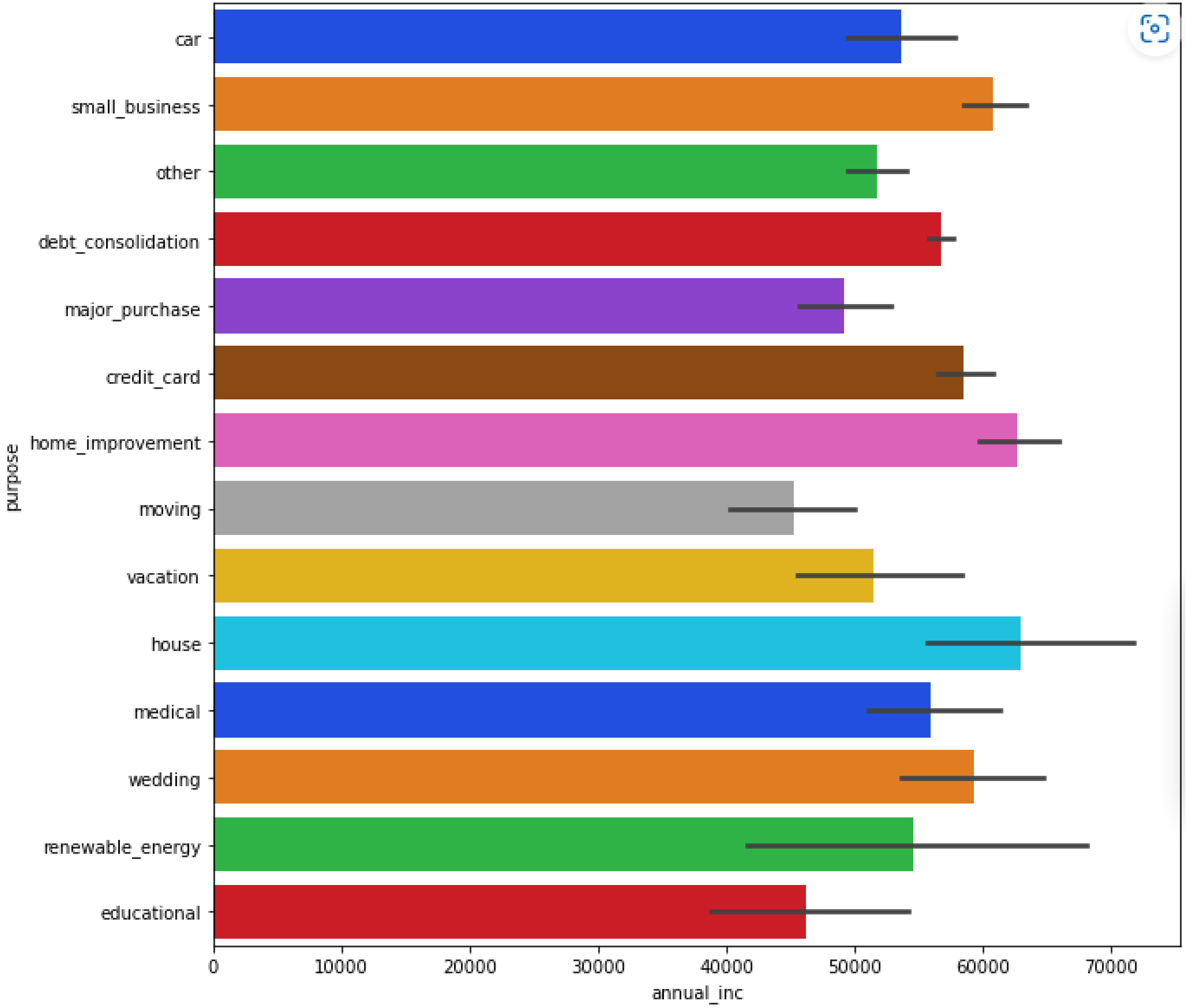


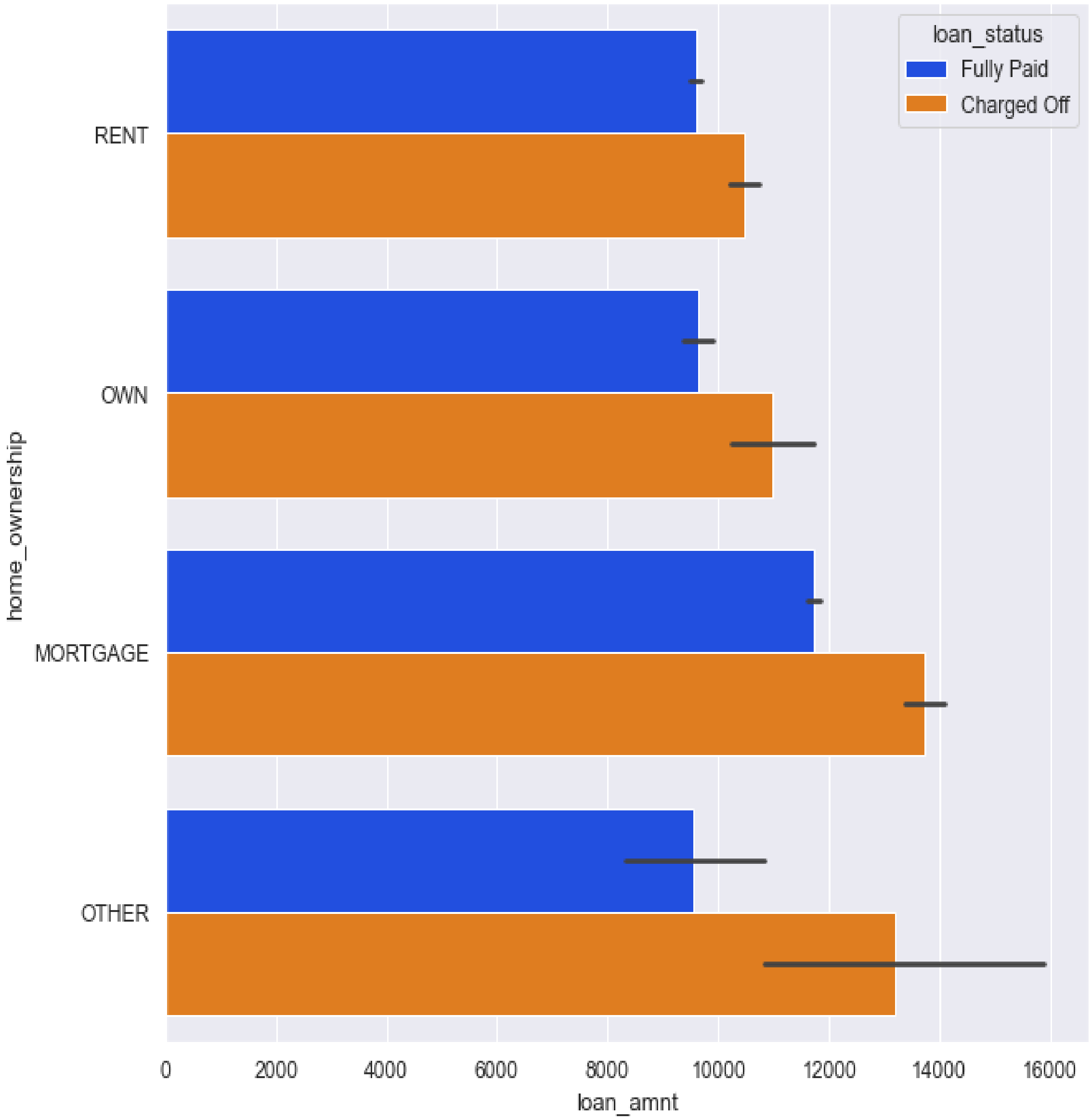


Univariate analysis observations:

- highest number of defaulters have taken small loans (0-10k). It can be said that people consider this a petty amount for a bank take it casually.
- Borrowers with dti between 6 to 24 are more likely to default.
- Majority of the defaulters lie in the income category of 31k to 58k.
- Higher the number of the open credit accounts, higher is the credibility of the borrower.
- Revolving utility rate of 50% or higher poses a risk for the bank as the probability for a default rises based on evidence.
- Ironically, people who are employed for 10 years or more are more likely to default. We might need to dig in further here.
- Public record of bankruptcies does not provide useful data as majority of defaulters have 0 bankruptcies.
- Majority of defaulters have a clean public record as well.
- Bank must ensure that the applicant is source verified to mitigate the risk of a charge off.
- Loans taken for the purpose of debt consolidation carry a huge risk of default.
- Homeowners are less likely to default whereas those on rent or have properties in mortgage pose a considerable risk of charge off.
- Loans with subgrades B3, B4, B5, C1 and C2 are more likely to succumb to a charge off.

Multivariate plots





Multivariate analysis observations:

- Those with verification status as "verified" tend to apply for bigger loans and default
- There is parity in the "Not verified" class of loans.
- However, bank is losing in every segment as the number of charge off's are greater than fully paid.
- If closely observed, the loan amount range for charged off loans seems to be between 8k to 18k.
- This same pattern of charge off's being greater than fully paid is observed when we plot loan amount vs home ownership. Observing the plot reinforces assumption that majority of loans ranging from 8k to 18k are in a risk of default.
- Analyzing the annual income vs loan amount groups chart, we can say that generally speaking people tend to fully pay the loan. It is evident in the graph where the charge off's are less than the fully paid for every income segment.
- It is interesting to note that even though the income level and employment durations of people increase, the probability of loan default still stays higher than getting fully paid.
- People in the income range 60k to 80k tend to seek out a loan for either house, home improvement, or small business. And then tend to default on those loans.

Thank You