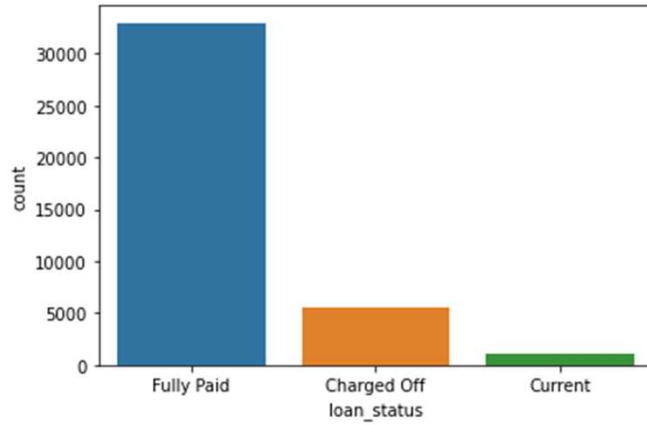


Problem statement

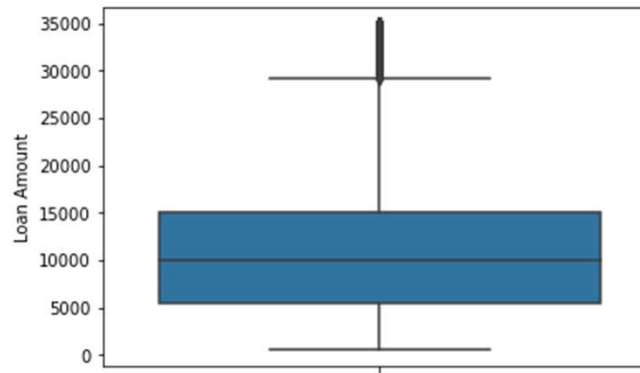
- Identify and predict risky applicants to loan who may default causing loss to the lenders
- Approach
 - Load Data
 - Remove columns contains empty and non useful information
 - Review the percentage of people defaulted in the datasheet
 - Understand the characteristic of the people who are defaulted

Univariate Analysis

Loan_status

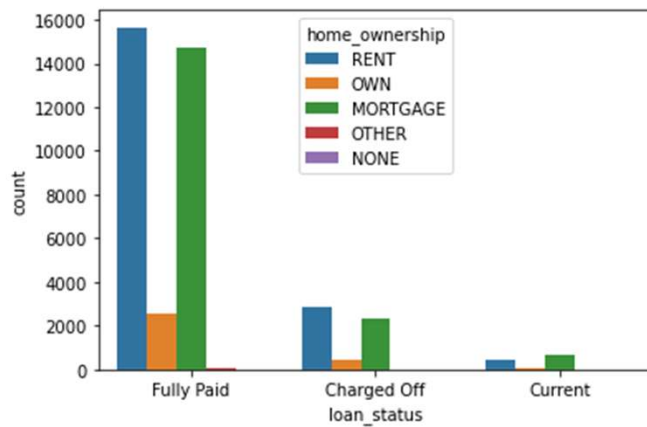


Avg Loan Amount -
10000

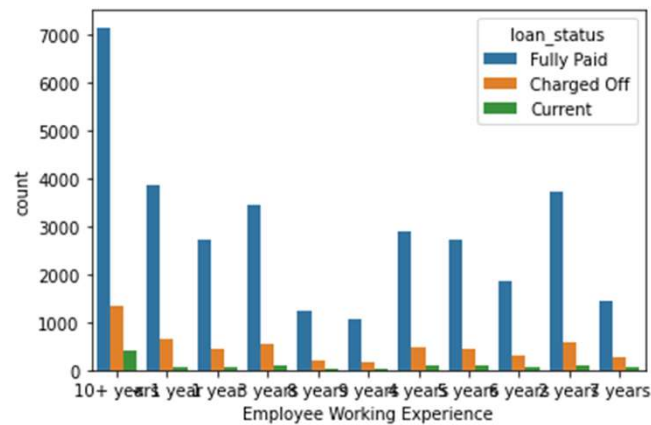


Bivariate Analysis

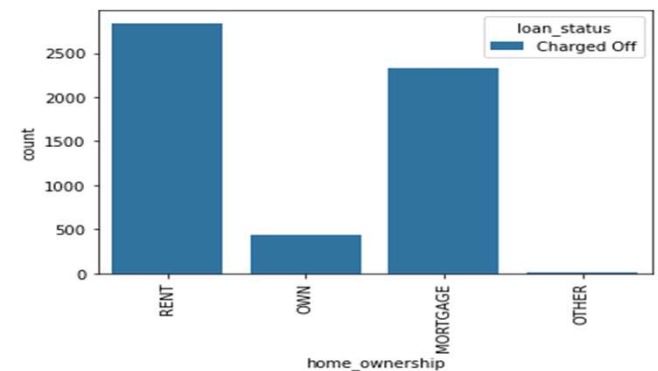
Loan_status & home_ownership



Loan_status & Emp Experience



Loan_status = Charged off & Home ownership



Observations

- People who owns a house will less probability to default
- Members from state CA have probability of default

