

Lending Club case study

Problem statement

- In researching this loan data set, our intent is to better understand, and eventually model, the outcome of a loan. More specifically, we wish to develop an understanding of the factors that affect the overall performance or outcome of a LendingClub loan. We expect a number of interesting business questions can be explored through our analysis, both broad and narrow in scope.
- We will explore loan outcomes as a component of estimating the total value of a loan portfolio. More narrowly, we also wish to understand those factors that may have the most influence on the overall performance of a given loan.

Variable Analysis

- Most of our distributions are right or positively skewed. Respective to a majority of variables in our dataset, it makes sense that these observations have a greater mean than median. For example, from annual income standpoint most borrowers typically make less than 100,000 dollars with some borrowers making quite a bit more ranging from 100,000 to 200,000 dollars.

- Conclusion 1: Most of the defaulters are having employment experience between 4-6 years, company should be careful while lending loans to the employees having less experience.
- Conclusion 2: customers between 10K-20K\$ annual income can likely to be instable in managing their finances and hence fall under defaulters.
- Conclusion 3: There are more outliers on revolving balance which means the credit health is more on them.
- Conclusion 4: As expected the revol_util % is high for most of the defaulters. Average percentage being $> 60\%$ which is a high value and shows bad credit score.