Boris Jokic

Report

This dataset contains experimental data about willingness to compete. Shortly: subjects play several rounds of experimental game and each round has different payment scheme and/or different incentives. Then we measure their willingness to compete based on their decision to compete and earn double or not to compete and earn less. Since risk aversion and some other factors affect their decision to compete or not, we control for these factors.

Aside from experimental data, this dataset contains data collected from surveying the subjects. Survey contains more than 100 questions about their demographics, income and many other things that should be considered for this research.

To find some interesting and meaningful results, modeling is necessary. So, I didn't find some very interesting results by doing just exploratory analysis of data, but there are some things worth of mentioning:

- Since each round of experimental game has very similar tasks, subjects tend to have higher scores in each new round, as a result of learning/practicing process.
- Even though men and women are performing equally well, women choose to compete less than men.
- Women have much higher risk aversion than men, which is one of the main factors why they choose to compete less. In terms of this particular dataset, I actually measured risk tolerance, which is inverse from risk aversion. That is why we can observe higher values for men than women, because men are more risk tolerant (less risk averse).
- There are some highly correlated variables that we can observe in correlation matrix (heat map). They might seem to have too high correlation (close to 1), but this is not concerning because these variables are supposed to be highly correlated. For example, each subject's score in round 1 is highly correlated with their score in round 2, which makes perfect sense because they are doing same type of tasks under different payment schemes.
- Distribution of many variables is skewed to right, mainly due to outliers. Many of these outliers are just result of false reporting.