STAT 410: Project Proposal

Alexander Xiong

I am interested in exploring the relationship between player salary and various player performance metrics and non-performance statistics such as player age, and position, and team dynamics. Given the advanced analytics of basketball, I can also explore the relationship between advanced statistics such as win shares or usage percentage and the non-performance statistics with player salary. I enjoy watching and following basketball so this would be a quantitative approach to determining whether a player is “worth” their contract.

I would need to organize yearly performance data from players as well as their salary. I believe that I can extract all necessary data from basketball-reference.com. I would need to establish a lower bound to filter players that play for a 10-day contract or are released before the season ends. Moreover, I would also need to filter for injuries that would prevent a player from completing a season or are suspended. I plan on collecting data from 2010 to the present; however, if it is necessary, I will collect more.

I plan on testing residuals to establish the normality of the response variable, although I do suspect that there would be a heavy tail, as big salaries would skew the distribution right. Since I plan on modeling the association between salary and performance and non-performance metrics of a player, I plan on using multiple linear regression to do so. I may determine that some metrics are non-crucial predictors so hypothesis tests would be necessary to determine how to best simplify the model. Furthermore, after producing a linear model, I plan on analyzing current younger players that have not received a contract outside their rookie contracts to determine what contract their performance could entail. This model would also allow me to plot which players are outplaying their salaries and which players are underperforming.