Capstone Project 2 Proposal

NBA

In more recent times, The NBA has more and more players come from overseas to come play professional basketball. It is becoming evidently harder for coaches, scouts and general managers to evaluate talent. My plan is to compare common statistics on average per game (Points, Assists, Rebounds, Blocks, and Steals) of NBA players who went to college versus those who came overseas. I will be using data visualization tools as well as creating predictive models in order to predict player performance. I want coaches to be able to have informed decisions when drafting a University player versus an overseas player. I am planning to get the data from www.basketball-reference.com . There is no API for this data, so I plan on scraping it using the Beautiful Soup package. I will then wrangle the data to make sure that all the data is relevant as well clean things up such as names, colleges, stats etc. I will then do an EDA by using visualization such as violin plots, residual plots, and multiple linear regression. I will then use various different Machine Learning Models to train and test my data. I will conclude and then make recommendations based on my findings. A couple things to note this will only be data from the past 20-30 years because it involves when more and more international players came into the league. I also want to note that some players went from highschool straight to the NBA. **My model evaluation is to compare the average points per game, rebounds per game, assists per game etc. with the American Players to the International players using predictive models.**