Weekly Update #4

This week was spent performing more EDA. I was able to consolidate the transfer market values of the top 50 players in England.Using this data, I was able to merge it with other data containing the stats related to defense, passing, and general goal and assist data. When merging this data, I then ran some correlation plots to find out which values correlated the most with “Value”, which is an indicator of the transfer market value of a player. So far, I’m finding that age is the most negatively correlated variable, meaning that the older the player gets, the more likely that their transfer market price is to fall as well. As expected, goals and assists are the variables that correlate most positively with transfer market value as well as the playing time that the player consistently gets.

Although I thought that starting with 50 players would serve as a less complex, simpler start, it seems that I need more players in order to get more accurate correlations between “Value” and other stats. My next steps will include introducing up to 300-350 more players(from 5-6 different leagues) in order to widen the range of different transfer market prices existing in the sport. Doing this will also allow me to filter the players by position, giving me enough samples to test how different statistics are rewarded in value compared to others. For example, how tackles completed by a defender might differ from goal chances created by a forward player.