

Aisha Mahmoud and Diego Mendez

DSC 101 Project

## QUESTION:

Who is the best player for each position in the English Premier League and in the American Major League Soccer league?

Subquestion: Which players are better, those in the English Premier League or those in the Major League Soccer?

We, Diego and Aisha, both play soccer, so we knew that we could do a project on soccer and it be something we were both interested in and passionate about. Seeing as how we both currently live in the United States of America, it makes sense that we would investigate into the Major League Soccer. We also found a nice data set on the English Premier League, so we decided to do it for that too.

Anyone who follows soccer would care about this; especially those who would want to see just how different are the players who play in the MLS and the EPL are.

“Despite its young age (22 years old) as a league relative to some of the European leagues that have been around for over 100 years, MLS still is compared to the quality of some of the top leagues in the world, like England’s Premier League, Spain’s La Liga, Germany’s Bundesliga, and Italy’s Serie A. Consistently, we hear all the criticisms about the league: that the quality of play is low, that it’s a retirement league, that it’s a Mickey Mouse league. In spite of all the criticisms and the belief that many soccer fans may have about MLS, the league play has improved dramatically, it has attracted much better and younger players, and attendance has skyrocketed as more money is invested into the clubs and their academies.”<sup>1</sup>

For primary data, we have two data sets in Excel that tell many various statistics about many players, one for MLS and one for the EPL. We just need to choose which stats to focus on.

---

<sup>1</sup> II, Donald Wine. “Community Corner: How Can MLS Become a Top 10 League?” Stars and Stripes FC. Stars and Stripes FC, September 5, 2018.  
<https://www.starsandstripesfc.com/mls-usmnt/2018/9/5/17816714/community-corner-mls-top-10-league-epl-la-liga-bundesliga-liga-mx-serie-a-ligue-un>.

## Second Sprint Review: METHOD

We want to find who are the best players in each league--but different players have different roles, and with that, you can't judge them all over a blanket criterion. Thus, we naturally knew that we would divide the players up into positions, which three main positions: defense, midfield, and offense. After this came the time to decide what makes a defender, a midfielder, and an offensive player so good.

We decided that the stats we would use to compare would be: for a defensive position: challenges won, tackles, ball recoveries; for a midfielder: assists, accurate pass percentages, key passes, and dribbles; and for offensive players: expected goals versus goals, percentage of shots on target, and dribbles. These are all stats that are there in the data set, that we will use R to compare and find the statistics for.

Additionally, since we're only finding the top players in each of the two leagues, we don't need the data of the bottom-most players. We will decide the cutoff for this by the number of minutes played. Through R, we will use the mean function to find the middle of the data set, and anyone who played less than the mean number of minutes played will be removed from the data. Clearly, if someone doesn't play much, they won't have as much of a chance to participate and have a meaningful impact on their matches, meaning their stats won't be near the top. This excludes skillful players who are injured and can't play.