Identifying Undervalued and High Performing Major League Soccer Players

Charlie Whorton

DS 745: Visualization and Unstructured Data Analysis

Professor Lauren Mauel

November 5, 2022

# Introduction

Moneyball is a common term associated financial analytics in sports. More specifically, moneyball, is the process of maximizing talent on a sports team’s roster given certain financial constraints. While soccer teams across the world have very high budgets, Major League Soccer (MLS) is still a growing league that has some financial limitations on teams to preserve parity. That means the moneyball approach is key to identifying undervalued players in the MLS. But moneyball doesn’t mean all signings have to be low salary players. Sometimes high salary players are worth the money because of their exceptional talent. In this analysis we will visually explore the MLS market to identify undervalued players, find reliable options high dollar signings, and for discuss intelligent signings made by clubs. As is customary in the soccer world, goalkeepers are left out as the data used does not apply to goalkeepers.

To perform this analysis, I used two csv files from the brilliant minds at American Soccer Analysis <https://www.americansocceranalysis.com/>. The first is a dataset with the Goals Added value for each field player in the league. Mueller, (2021) describes the Goals Added metric as a value that measures a player’s total on-ball contribution in attack and defense. This metric omits actual goal scoring actions from its calculation so a player who scored 20 goals may only have a Goals Added value of 3. This can also be described as the probability of scoring or conceding a goal based on a certain action. Goals Added is broken into six categories: dribbling, passing, shooting, receiving, interrupting, and fouling. Each of these has their own Goals Added value that combine to make the overall Goals Added value.

# Metrics Created

**g+ Avg**: The average Goals Added value per player across the league.

**Goals Added Rank**: A player’s ranking of their Goals Added value across the league.

**Salary Rank**: A player’s ranking of their salary across the league.

**Salary/Goals Added Difference Rank**: A player’s ranking of the difference between their salary ranking and their Goals Added ranking. A greater difference indicates a player has a salary that is lower than what is expected for someone with a given Goals Added value.

# Identifying Undervalued and High Performing Major League Soccer Players

When identifying undervalued players, it’s important to find players who perform well despite having a low salary. A soccer team needs to leave room for high salary players since there elite players demand large contracts.

With this treemap we can see the top 58 players that are performing above expectations based on their salary (the visualization filters out 42 players with a negative value). The first name we are drawn to is Jack Elliott. The size of the mark tells us Jack has the highest Goals Added value of the players shown. The color, with a darker shade being a lower salary and a lighter shade being a higher salary, shows that Jack is one of the higher paid players in this group of players, although he is still right around the league average. Another player that stands out is Mamadou Fall, an 18 year-old center back on a $94,000 salary. This is a low salary given the league minimum is $63,500, but we can see by the size of his mark that Fall has a high Goals Added average. Two other names to remember are center midfielders Sean Davis and Marcelino Moreno. Moreno is third from the top in the first column and Davis is in the third column from the left and second mark from the top. They both have a fairly high Goals Added value. Their salaries are on the higher end of this group of players, but they are still right around the league average.

Chart, treemap chart

Description automatically generated

While this visualization is very effective when it comes to identifying undervalued players league wide, it does not do a good job of identifying specific types of players. We’ll continue to dive deeper into position groups to get a more detailed look at the league as a whole.

# Goals Added for Defenders

To get a more detailed look at undervalued defenders we can view all defenders who have played over 490 minutes in a scatterplot. Here we can see just how much higher a Goals Added value some players have over the rest of the league. Combined with Salary along the x-axis we can see just how expensive or inexpensive a player may be. The size of the marks indicates Goals Added from interrupting the opponents attack and the shape and color indicate whether the player is a center back or full back (outside back). Jack Elliott is once again an outlier. As a player with one of the highest Goals Added values sitting right on the average salary line this chart justifies our findings in the previous visualization. It is worth noting that Jack Elliott’s salary doubled in 2022 recognizing that he was undervalued. We can also see Mamadou Fall as having a fairly high Goals Added value compared to other players at the same position. The top left quadrant is the ideal area for undervalued players as their Goals Added value is above average, but their salary is below average. Mamadou Fall is in a very good location to be a quality prospect. Additional notable names in this are Walker Zimmerman, who won MLS defender of the year. We can see he is high on both Goals Added and salary as he is a quality player, but is well known so he has a salary to match his quality. Lastly, Aljaz Struna is a good example of the effectiveness of this chart. He is a highly paid player at $1.2 million per year but has a below average Goals Added value. Struna only played 18 of 34 games in 2021 which would indicate poor performance. Overall, this shows that he is overvalued and would not be a good signing. To further justify this, he was not re-signed to the team for 2022.

Chart

Description automatically generated with low confidence

# Undervalued Midfielders

When exploring center midfielders we exclude wingers, defensive midfielders and attacking midfielders and we are focusing on players who played over 490 minutes of the season. Additionally, we are filtering out center midfielders who have a Goals Added value less than -.5.

Center midfield requires a wide variety of skills. For this visualization we show Goals Added from interruption on the x-axis and overall Goals Added on the y-axis. This way the location of each mark is more of a reflection of a player’s on-field performance rather than a comparison of their salary to their performance. Salary is noted with the larger circles representing a lower salary. Additionally, darker circles have a higher Goals Added value from passing.

Earlier we mentioned two center midfielders, Marcelino Moreno and Sean Davis. Moreno has a successful Goals Added value from passing which we see by the color of his circle, but not from interruptions. Interacting with the tooltip shows he has a high Goals Added value from dibbling, shooting, and fouling meaning he earns fouls in dangerous areas due to his dribbling ability.

Sean Davis is a different player evident by the fact that he is on a different side of the chart. Sean Davis has a moderately high overall Goals Added value, but that almost entirely comes from interruptions and passing. He has a darker circle indicating he passes well, but has the highest Goals Added from interrupting. This indicates that Sean Davis plays a more supportive and defensive style while Moreno plays more of a creative attacking style.

Chart, bubble chart

Description automatically generated

Since most teams play with two to three center midfielders, we could say that these two players play well together. They each excel at different areas of the game which could make for a complimentary midfield pairing.

João Paulo is shown as reference for an ideal, high salary midfielder. He has a deservedly high salary as he has high Goals Added values for passing and interrupting which are key for a central midfielder.

# Best Attackers

In the final position group, we visualize the best attackers. While this is a different approach from identifying undervalued players in the previous visualizations it can still be a financially advantageous approach. With this approach a team may not look for a diamond in the rough, but instead a team can ensure their high dollar signing will perform as expected. For this reason, we do not consider player salary. This visualization includes forwards, attacking midfielders, and wingers who played over 490 minutes in the 2021 season. Additionally, this chart only shows players who have a Goals Added value from shooting that is greater than or equal to zero.

Along the x-axis there is Goals Added from shooting, which is a key attribute for attackers. The size of each circle represents Goals Added from receiving. This indicates players who receive the ball in dangerous positions. This is another important attribute for an attacker, particularly a forward. The color represents Goals Added from dribbling with blue being a higher value and orange being a lower value. Great dribbling ability is not necessarily a requirement for a great goal scorer, but it is undoubtably a valuable skill to have as a great dribbler can create more goal scoring opportunities.

We can see the effectiveness of this visualization since there are several award winners at the top. Valentín Castellanos scored the most goals of 2021 and it is interesting to see that he was not an effective dribbler. This indicates that he relied on his finishing ability and his positioning to score goals.

An interesting player near the top of this chart is Hany Mukhtar. Hany is an attacking midfielder, so he does not play as high up the field as Castellanos. Hany has a smaller circle because he does not have a high Goals Added value from receiving. Playing further away from goal than players like Castellanos means he relies more on his dribbling to move the ball up field and create goal scoring opportunities. The difference in these two players on this visualization perfectly represents the difference in playing style.

Chart, scatter chart

Description automatically generated

# Conclusion

These visualizations highlight high performing and undervalued players. They also give insight into player styles and which players may not be suited for certain teams. With this information a team can efficiently analyze a large group of players and make data-driven decisions in signing undervalued and high performing players for their club.

# REFERENCES

Muller, J. (2021). *What Are Goals Added (g+)?* American Soccer Analysis. Retrieved from <https://www.americansocceranalysis.com/>

American Soccer Analysis (2021). *Goals Added (g+*) American Soccer Analysis. Retrieved from <https://app.americansocceranalysis.com/#!/mls/goals-added/players>

American Soccer Analysis (2021). *Salaries* American Soccer Analysis. Retrieved from <https://app.americansocceranalysis.com/#!/mls/salaries/players>