



Machine Learning Fall 2021

# 2021 Cleveland Guardians Off-Season Roster Moves: Position Players

Anthony Fundzak

## 2021 Cleveland Guardians Off-Season Roster Moves: Position Players

### The Problem and Objectives:

Problem: The business of sport and the teams within the sport is directly and most impacted directly by the performance of the team. Making trades, signings, cuts, or any other roster move directly impacts the business of the team and can change millions of dollars in profits or losses for the team through ticket sales, merchandise, and other forms of revenue lines. Making roster decisions is something that goes on everyday within the team's front office so that is why I chose this side of business, the sport business world, as my final project. At the end of Major League Baseball seasons, front offices need to cluster and illustrate their players throughout the organization showing their performance and which players stood above the rest. This will help them identify how to make upcoming roster adjustments and anything else to do with the roster. For my case, I am specifically going to be looking at the 2021 Cleveland Guardians and players who recorded stats on the main man roster. The objective of this is to illustrate visual explanations and summaries to better understand where each player compares to others that played on the active man roster throughout the season. Now is the time to compete in the division with not one clear team leading the pack so understanding who to build around on the offensive end will help us set team and business success now and into the future. This is a real-world business problem because as the team goes forward, the business and profits simultaneously go forward with it. The money, profits, and business of the organization starts and stops with the team decisions made by the front office.

### The Approach:

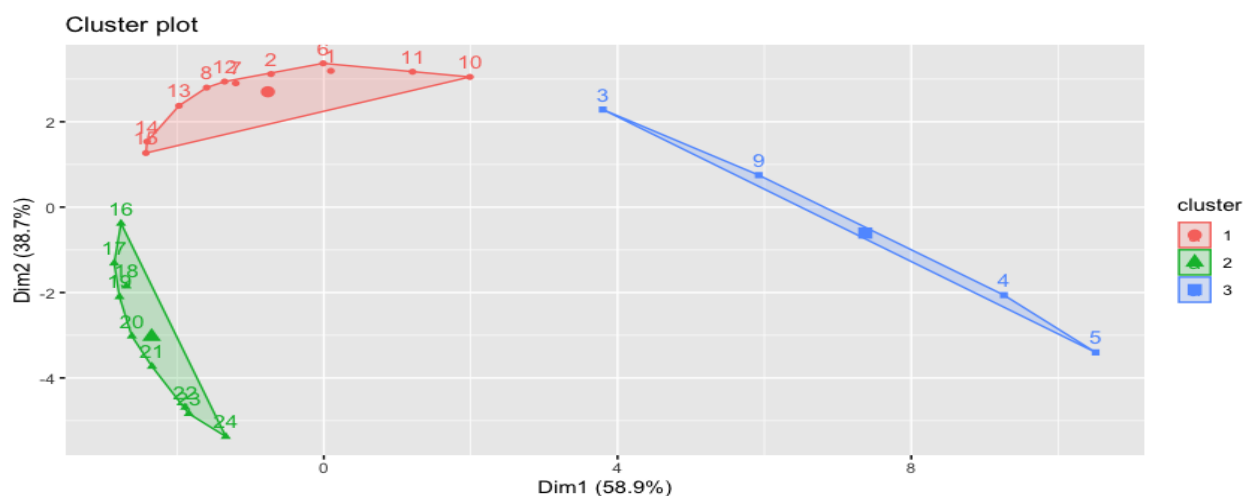
As the games go on, averages and stats are kept for every player that plays in an MLB game and even the minor leagues for the MLB. There are countless stats for the game of baseball but in today's game, one can argue finding the best offensive team as possible will lead you to World Series (as we have seen the last couple of years with the offensive powerhouses winning) and will lead the organization to

## 2021 Cleveland Guardians Off-Season Roster Moves: Position Players

business success stemming directly from the winning and play on the field. My approach is simple, to turn the data into visual representations that anybody can understand and such visuals that can be not just clear, but quick to read and interpret. It was important to list the averages, means, highs, lows of all the stats as that can show where your whole team stands regarding themselves and others around the league. I chose to focus on finding data on individual players as those players are the ones who drive the team success forward leading to business profits for the team and for themselves, also impacting the team. Using k-means clustering, I was able to create accurate and easy to read clusters for how to group the players in terms of offensive performance with the best four players being in the 3<sup>rd</sup> cluster and the rest of the 21 split into the other two clusters. I also chose to use a histogram and a scatter plot as those are two of the most easily readable and approachable plots for stakeholders and others to read and understand. I tested numerous cluster amount to determine the best amount and 3 clusters was proven as the best fit for the data.

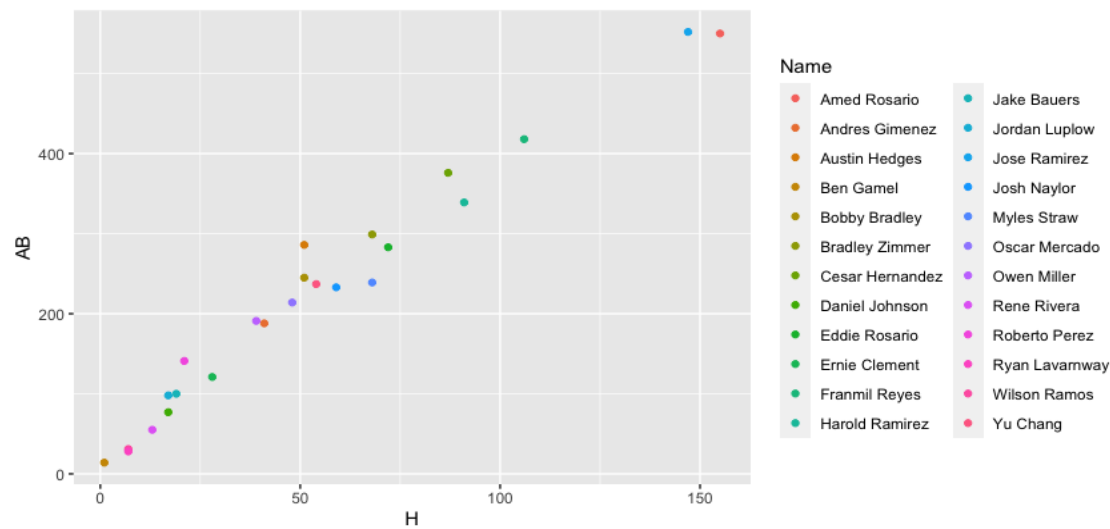
### The Results and Conclusion:

Based on the data summarized and talked about above, we can clearly see the 3 best offensive players that the Guardians need to build around for the 2022 season. A near half majority of the team fell into the bottom/worst cluster indicating some major work is needed to improve the roster to compete during the upcoming season.



2021 Cleveland Guardians Off-Season Roster Moves: Position Players

Building around Franmil Reyes, Jose Ramirez, and Amed Rosario, members of cluster three and the higher of the players on the scatter will be the foundation to our success.



By looking at the stats summary, we can see the Guardians need to improve in batting average, strikeouts, and the ability to hit the long ball home run.

BA		SO		HR	
Min.	:0.0710	Min.	: 6.00	Min.	: 0.000
1st Qu.	:0.2070	1st Qu.	: 27.00	1st Qu.	: 2.750
Median	:0.2275	Median	: 52.00	Median	: 7.000
Mean	:0.2220	Mean	: 57.42	Mean	: 8.458
3rd Qu.	:0.2532	3rd Qu.	: 87.00	3rd Qu.	: 9.250
Max.	:0.2850	Max.	:149.00	Max.	:36.000

As no salary cap exist and unique to the MLB, winning on the field and providing a high-power offense for the fans to spend their money on and do business with the team to increase the salary cap leading to spending on higher caliber players is critical to the success of the franchise now and far into the future.