Tufts BAT - Reds Hackathon

By Ben Mascuch, Elliot Smith-Chauss, Emerson Wang, Kyle Krell, and

Richard Diamond

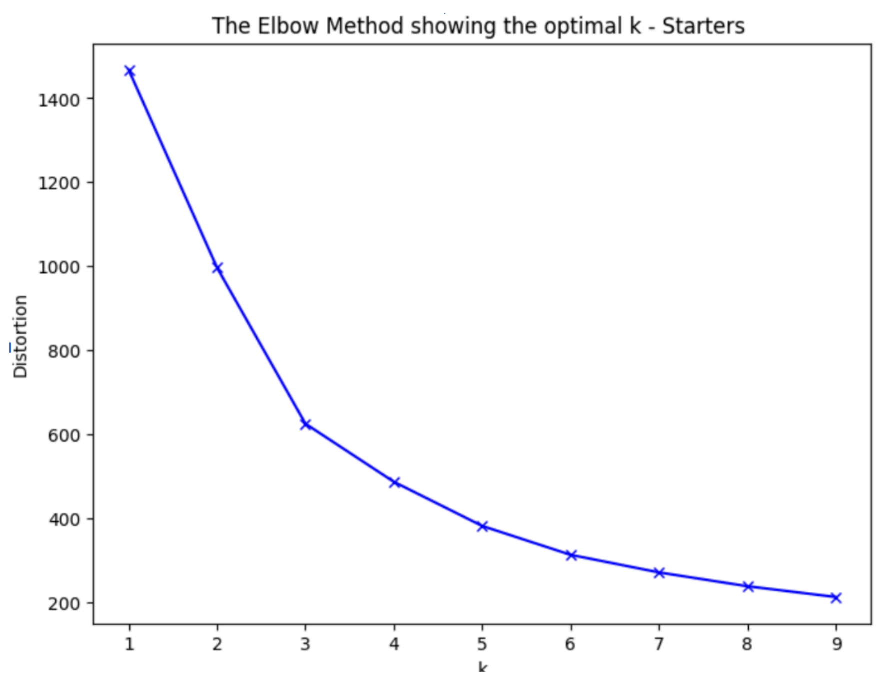
In this project, we were tasked with the following: establish pitcher “roles” and suggest 2-3 pitchers that would be good candidates to switch from one role to another.

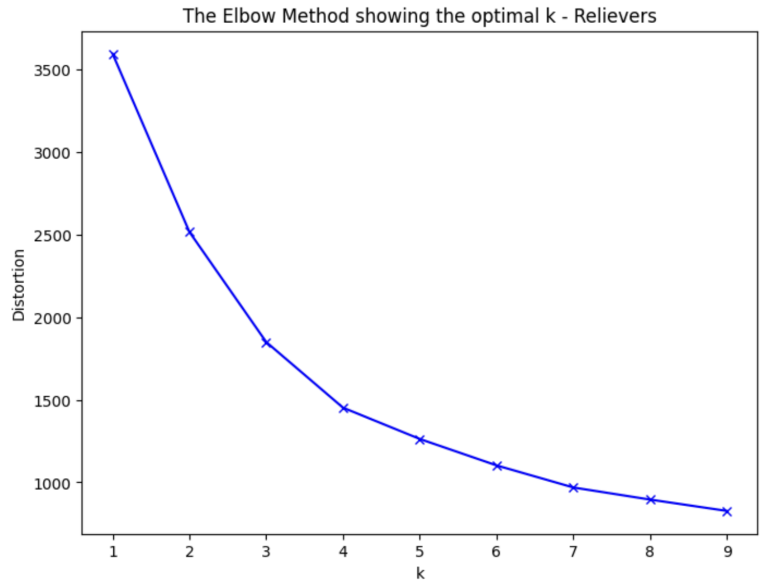
Firstly, we sought to determine what pitcher roles exist. In order to eliminate any bias we might have about these roles, we wanted to use an automated and data-driven method to find existing roles. Then, we wanted to create some metric to measure how close each pitcher is to the various roles that we find.

The first step in this process was to decide what statistics determine what a pitcher’s role is. To find our statistics, we combined Baseball Savant pitch data and Fangraphs total season-level data. Additionally, given that the data went from 2021-2023, we decided to view pitchers in different years as different entities. This decision was made in order to keep the sample size as large as possible and to account for the fact that pitchers are often used in different roles in different years. Next, we created three statistics out of the data given to us: Total Batters Faced per Game (TBF/G) and Platoon+. Platoon+ measures how often a pitcher is used in advantageous platoon situations, relative to other pitchers of the same handedness. Specifically, it takes the percentage of time a pitcher faces a batter of the same handedness and normalizes that percentage by the percentage of plate appearances with a lefty pitcher taken by lefty batters (27%), or the percentage of plate appearances with a righty pitcher taken by righties (45%), depending on which hand the pitcher throws with.

Once we created these statistics, we used the elbow method to determine the number of roles. The elbow method takes unlabeled data and finds the optimal number of categories based on similarities in our input statistics. To run the elbow method, we try each number of categories (k) from 1-9 and measure the distortion, which is the aggregate distance from each point to the centroid of its assigned cluster. To find our optimal number of categories, we observe where the ‘elbow’ in the graph is. This refers to the point at which the line generated by our line graph appears to have a cusp. The elbow can be decently clear or a little subjective, and there may be multiple defensible choices for the number of categories in any case.

We used slightly different inputs for the elbow method for starters and relievers, as what determines a starter’s role is different from what determines a reliever’s role. For starters, we used average exit leverage index (exLI) and TBF/G. We chose to use exLI instead of average leverage index (pLI) because we wanted to capture how long a starter’s “leash” is, which is more accurately measured by exLI. For relievers, we used pLI, TBF/G, and Platoon+. Each input data point for the elbow method was a pitcher’s year-long averages in these categories, with starters and relievers, including pitchers who appeared as both, being considered separate data points. After using the elbow method, we found that the optimal k was 3 for starters and 4 for relievers, meaning we found 3 starter roles and 4 reliever roles. Below are the graphs resulting from the elbow method.





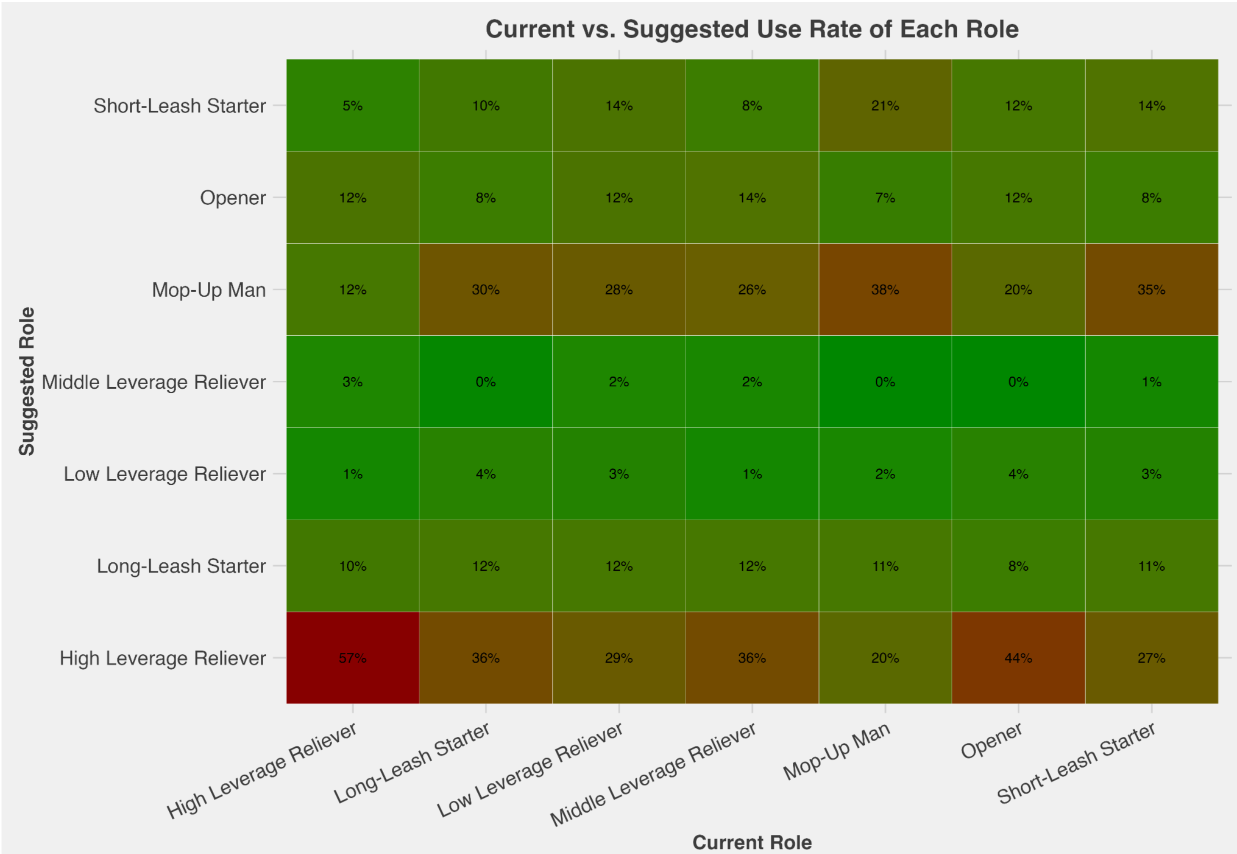
Next, we ran k-means clustering to assign pitchers to categories. k-means clustering is another unsupervised learning algorithm that, given a value k, labels data as part of one of k clusters. We ran k-means clustering on both starters and relievers separately, with k equal to 3 for starters and k equal to 4 for relievers. The categories we found for starters can roughly be broken down as follows: openers, starters with a long leash, and starters with a short leash. For relievers, we found four categories: long relievers, low-leverage relievers, setup/specialist (referred to as "Middle Leverage Reliever" in the figures), and high leverage relievers. Every pitcher was assigned to one of these categories through k-means clustering.

After we categorized each pitcher, we needed to find pitchers who were good candidates to switch roles. To do this, we looked at the average underlying metrics of pitchers in each of our roles and found pitcher's whose underlying metrics more closely matched pitchers of a different role than the one they were assigned to. The underlying metrics we chose were the following: Breaking Ball Stuff+ (BStuff+), Fastball Stuff+ (FStuff+), Offspeed Stuff+ (OStuff+), PitchingBot Command (botCmd), and Average Fastball Velocity (FAv). We normalized each of these columns to a mean of 100 to account for the fact they are on different scales.

After determining the average underlying attributes, we calculated the "distance" from each pitcher to the average values of our five underlying metrics in each role. In this context, "distance" is defined as the square of the category average underlying metrics subtracted from the player's underlying metrics. For instance, if Josh Hader has a 110 in every input stat and the mean of each of these stats for the closer category is 100, then, since we have 5 input stats, his distance from this category is $\sqrt{5(110\ -\ 100)^2} \approx 22.36$. A pitcher's total distance was then just the sum of their distances from each of the 5 roles.

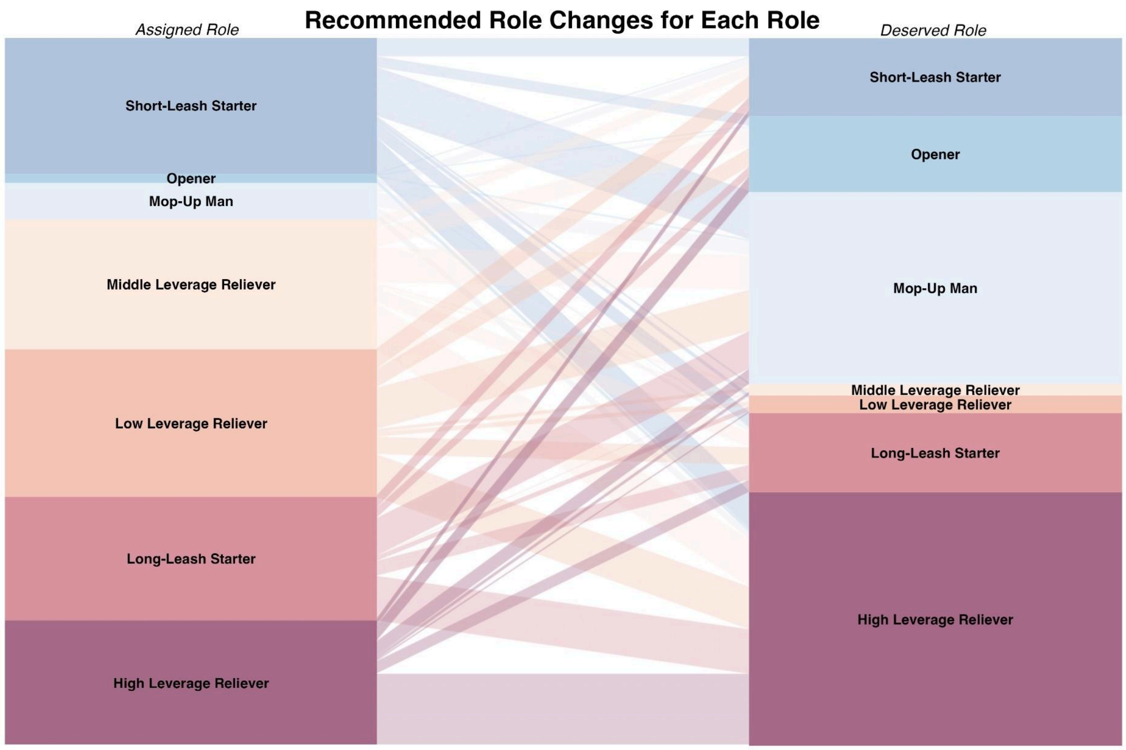
Next, we divided each player's distance from each category by their total distance to get a "score" for each category. This calculation accounts for players who were far from all categories in attributes by making their "distance" relative. One example of this is Aroldis Chapman. His stuff numbers are abnormal, meaning he appeared far from the mean for every category, and appeared to be a bad fit for the closer role. However, when we calculated his closer score, he appeared as a great fit for the role, as his metrics more closely resembled those of a closer than other roles.

Below are several figures that demonstrate how our model categorized pitchers. The first is a matrix that shows the roles in which pitchers are actually used compared to the role that their metrics are closest to.



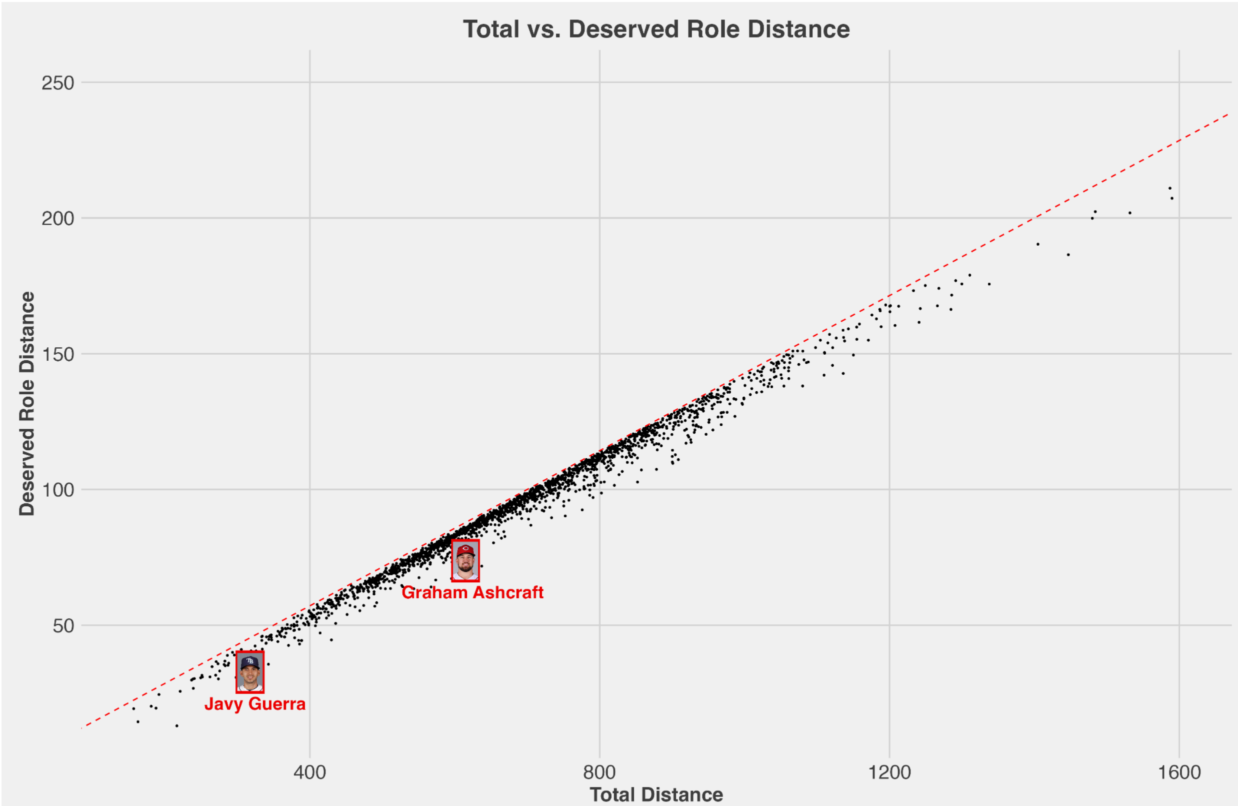
The percentage of pitchers used in each role that should be used in each role. (ex. 57% of pitchers used as High Leverage Relievers should be used as High Leverage Relievers.)

The graphic below shows the proportion of pitchers that are currently used in each role compared to the proportion of pitchers that our model suggested should be used in each role.



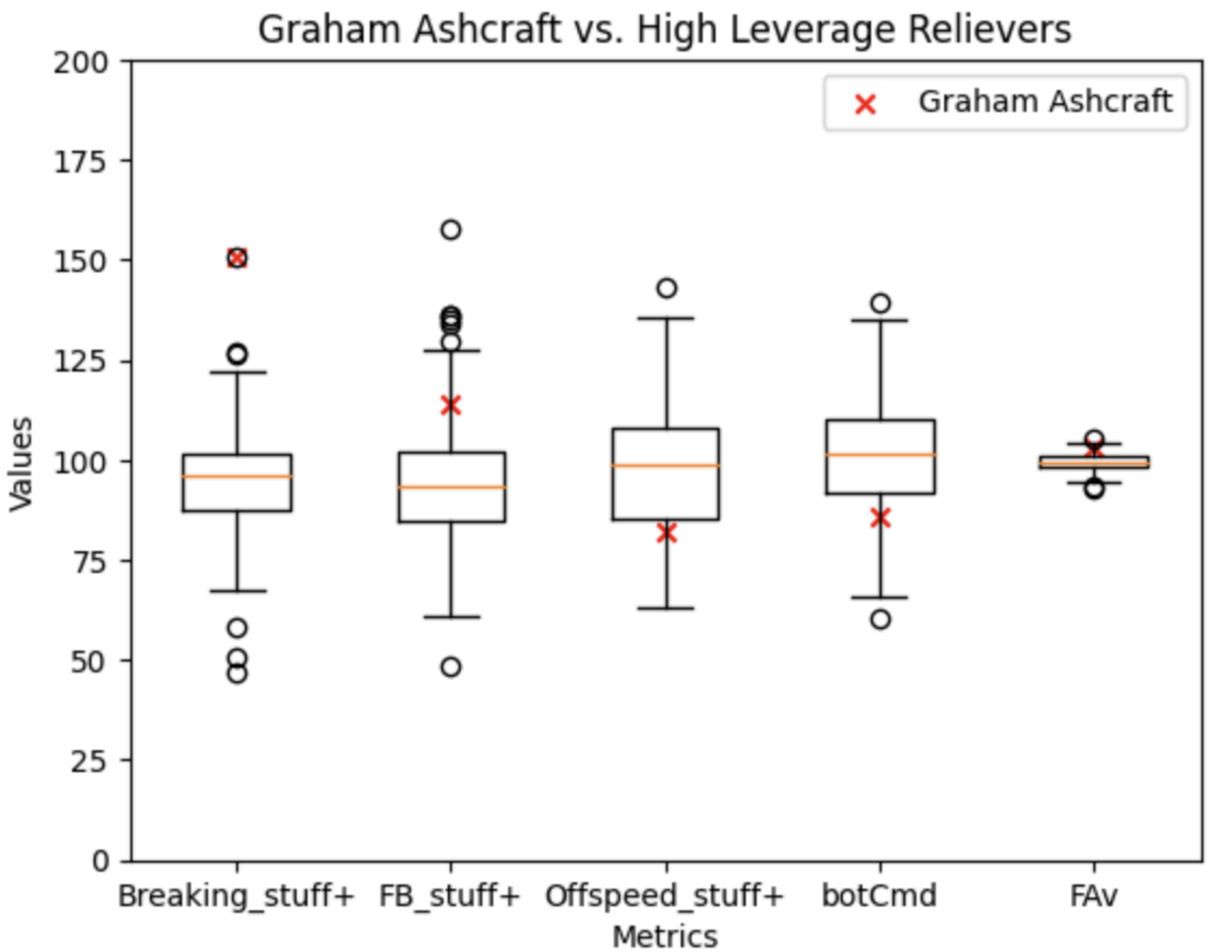
This shows the breakdown of how many pitchers are used in each role currently (on left) and how many our models suggests should be used in each role (on right). The middle visually depicts how many pitchers should be moved from their current role to a new role (in middle). Where the lines start signifies what role those pitchers are currently used in, and where the lines end signifies what role those pitchers should be used in. Thicker lines means more pitchers should be moved from one role to another. (ex. Since there is a thick line from Weak Starter on the left to Mop-Up Man on the right, our model suggests that a lot of weak starters should be moved to mop-up men.)

This next plot shows our justification for choosing to focus on our two players of interest: Graham Ashcraft and Javy Guerra. The plot shows each player's total distance compared to their distance to their closest category. Both Ashcraft and Guerra are far from most pitchers and from the red trendline, meaning they stand out as players that our model strongly suggests being used in a particular role.



The sum of every pitcher's distance to each of the 7 roles plotted against their distance to the role they should be used in. The red dashed line signifies a player that is evenly distributed across all 7 roles, meaning they would fit just as well into any role. Players farther from the red dashed line are players that our model strongly suggests being used in a particular role. (ex. Javy Guerra and Graham Ashcraft are both in the top 12 for 2023 pitchers ranked by distance from the red dashed line, and they are currently used in a different role from their suggested one.)

Our first player of note is Graham Ashcraft. He is categorized as a starter on a short leash, but our model believes that he would be better as a high leverage reliever. This is due primarily to his high FStuff+ of 114 that matches closely to the high leverage reliever average of 11 and his extremely high BStuff+ of 151 that is closest to the high leverage reliever average of 117. Additionally, our model does not view command to be as important for high leverage relievers as starters, furthering the argument that Ashcraft could be used more effectively as a high leverage reliever because his botCmd is below the average for a starter with a short leash. Finally, his FAv is far higher than the average for a starter, matching more closely with the high leverage reliever.



Our second player of interest is Javy Guerra. He is categorized as a low leverage reliever, but our model predicts he should be a high leverage reliever. This is primarily because of his FStuff+ of 123 that matches closest to the high leverage reliever average of 111 and his OStuff+ of 82 that resembles the average of 82 for his new role. Furthermore, his FAv is higher than all average FAv, but closest to the high leverage reliever. Finally, his BStuff+ of 113 is still relatively close to the high leverage reliever average of 117, and while his BotCmd is lower than the high leverage reliever average, the model does not view this as negatively as other attributes. Guerra made several multi-inning relief appearances in 2023, often in mop-up duty, but we suggest he could be an effective pitcher in a shorter relief role. According to our distance measure, Guerra is actually the closest of all pitchers to the high leverage attributes. It is important to note that Guerra is not currently signed to an MLB contract, instead he will pitch for the Hanshin Tigers of the NPB in 2024. In Japan, Guerra will certainly be given an opportunity to pitch in higher leverage situations. We predict that a change in role could benefit Guerra and possibly spark a return to MLB.