

2024 Cincinnati Reds Hackathon - Mendoza Line

February 5, 2024

**Eda Xu
Louis Chen
Julian Sotelo**



1. Introduction

Context

The game of baseball is changing rapidly every decade. Since 2013, the average starting pitcher's innings dropped from 5.9 innings to 5.1 innings per appearance. Leading to an approximate 20% increase in relief appearances per team in 2023, jumping to 526 relief appearances from 478 relief appearances in 2013. Among these appearances, the additional innings these relievers had to cover - in 2013, were 3,048 multi-inning relief appearances (~102 per team); in 2023 this was up to 3,941 multi-inning appearances (~131 per team).

The aforementioned statistics illuminate a significant shift in the utilization of pitchers. The declining innings pitched by starters and the corresponding surge in relief appearances underscore a strategic metamorphosis within the game. This tactical shift brings to the fore the increased reliance on the bullpen, spotlighting the critical role of relievers in contemporary baseball.

Importance of Analysis

The analysis presented in this report is pivotal for several reasons. Foremost, it enables us to dissect the factors contributing to the increased burden on relievers. By understanding these factors, we can better prepare our pitching staff for the rigors of modern baseball, ensuring that they are equipped both physically and mentally to shoulder this load.

Additionally, this analysis is instrumental in assessing pitcher effectiveness. With the role of relievers becoming increasingly critical, evaluating their performance becomes a more complex and nuanced task. It is not simply a matter of innings pitched but the quality and strategic timing of those innings. By dissecting the data, we aim to distill the essence of what makes a reliever successful in this new era of baseball.

Data Sources

Fangraphs Overview: Fangraphs serves as the bedrock of our statistical analysis. Their comprehensive data on pitcher performance, including innings pitched, strikeout rates, walks, ERA, and much more, provide us with a quantitative backdrop for our qualitative assessments. This data is crucial for benchmarking our pitchers against league trends and understanding where they fit within the broader landscape of MLB pitching.



Baseball Savant Data: Baseball Savant enriches our analysis by offering pitch-by-pitch data that reveals the intricacies of each pitcher's game. From pitch velocity to spin rate, from exit velocity to launch angle on batted balls, Baseball Savant provides us with the tools to conduct a microscopic examination of pitcher effectiveness. This level of detail is vital for understanding how subtle changes in a pitcher's delivery or pitch selection can have outsized effects on their performance.

Baseball Reference: Finally, the historical data from Baseball Reference allows us to contextualize our analysis within the game's evolutionary timeline. By considering all position players who have pitched in official MLB games, we gain a holistic view of pitching trends and the changing role of player versatility. With the exclusion of the anomalous Shohei Ohtani, our focus remains sharp on traditional pitching roles, ensuring that our analysis is as targeted and relevant as possible.

In conclusion, the data from these sources, when synthesized, will not only provide us with a snapshot of the current state of pitching in MLB but also offer foresight into how the game might continue to evolve. This foresight will be instrumental in guiding our strategic decisions, ensuring that our team remains at the vanguard of baseball's unfolding narrative.



2. Methodology

Data Filtering

In our pursuit of clarity and precision, we meticulously filtered our dataset using the historical archives available at Baseball Reference. This repository served as our sieve, allowing us to separate position players from our primary subjects of study: the pitchers. In this filtration process, we consciously omitted the multifaceted Shohei Ohtani to maintain the purity of our pitcher-specific analysis.

Feature Engineering

Venturing deeper into the realm of Baseball Savant, we embarked on an expedition to engineer features that would best encapsulate pitcher performance across selected scenarios. We considered various situations that can significantly impact the outcome of a game:

- *Specific Scenario Metrics:* We quantified the frequency of 2-out walks, 0-2/1-2 walks, and lead-off walks per season. These scenarios are crucial as they often lead to high-leverage situations that can alter the course of an inning or the game itself.
- *Average Batters Faced and Innings Pitched:* We calculated the average batters faced per appearance and the average innings pitched per appearance to gauge the workload and staying power of a pitcher.
- *Inning-Specific Averages:* By examining the average inning a pitcher enters and exits the game, we gained insights into the strategic use of pitchers and their endurance.
- *Mode Inning Entry and Exit:* Identifying the most common inning for a pitcher's entry and exit allowed us to understand bullpen management trends and pitcher stamina.
- *Inning-Based Role Classification:* We classified pitchers into roles based on the innings they most frequently pitched: Innings [3,4] as Long Relievers, Innings [5,6] as Middle Relievers, Innings [7,8] as Setup Pitchers, and Inning [9] as Closers.

Analytical Techniques

Our analytical arsenal was equipped with sophisticated techniques designed to cut through the complexity of raw data and unearth underlying patterns:

- *Principal Component Analysis (PCA):* We implemented PCA across our feature set, including the newly engineered metrics. This dimensionality reduction technique allowed us to uncover the principal components accounting for at least 80% of the total variance.



More than a mere statistical exercise, PCA provided a lens through which we could discern the most impactful features, guiding our subsequent analysis.

- *Feature Loadings*: The loadings from our PCA became the cornerstones of our analysis, giving us a clear indication of which features were most significant in defining pitcher performance.
- *Z-Score Analysis*: Utilizing the feature loadings, we developed a Z-score framework to evaluate each pitcher. This scoring system categorized starting pitchers into Elite_SP, Strong_SP, Average_SP, and Suboptimal_SP, while relief pitchers were classified as Elite_RP, Strong_RP, Average_RP, and Suboptimal_RP. This stratification was anchored in our PCA findings and gave us a robust benchmark for assessing performance relative to a pitcher's role.

$$z = \frac{(X - \mu)}{\sigma}$$

X: A Specific Measured statistic for a pitcher, such as ERA, WHIP, and K/9. Mainly used from our feature loadings from the prior PCA Analysis.

μ (mu): The mean value of that specific statistic for the group of pitchers being considered, which could be all classified roles (Starting Pitcher, Middle Relief, and Setup/Closer)

σ (sigma): The standard deviation of that statistic for the group of pitchers.

- *Cosine Similarity*: We split the pitchers into two distinct cohorts: elite/strong and average/suboptimal. Within these groups, we analyzed the roles classified during the feature engineering phase to determine the pitchers' past performances. Employing Cosine Similarity, we ventured to align each average/suboptimal pitcher with a new role, predicated on their similarity to the performance metrics of their elite/strong counterparts



- *Elite Means*: After receiving our cosine similarity metrics we are then able to examine all pitching metrics across the elite level pitchers and compare them in our list of suboptimal pitchers (pitchers belonging to Average_SP, Average_RP, Suboptimal_SP, Suboptimal_RP). The Elite pitcher metrics we have aggregated belong to the Elite level pitching we categorized (Elite_SP, Elite_RP, Strong_SP, Strong_RP).



Player One: Luis L. Ortiz PIT



Current Role: Starting Pitcher

Common Inning Pitched:

Suggested Role: Setup/Closer

Throws: R

Age: 25

2023 Stats

Wins	5
Losses	5
ERA	4.78
G	18
IP	86.2
SO	59
WHIP	1.70

Analysis & Discussion:

Luis L. Ortiz has come off an average season as a starting pitcher. Our analysis proves he is suitable in a closing role since his statistical distributions are close to the mean of closers who meet our 'Elite Pitcher' criteria.

Swing Percentage (0.9% Above Elite Mean):

- Analysis: Closers typically need to be able to induce swings, especially in high-leverage situations. A higher Swing% suggests that Ortiz is effective at making batters swing at his pitches. This could be due to his ability to deceive batters or to throw pitches that are challenging to lay off. In a closing role, this skill is crucial for quickly getting outs and preventing the opposition from gaining momentum.
- Context: Being above the elite mean in Swing% indicates that Ortiz has a tendency to engage batters more effectively than many elite closers. This is



beneficial in a closing role, where the ability to generate swings (and ideally misses) can be the difference between a save and a blown opportunity.

2. LOB% (9.6 Percent Above Elite Mean):

- a. Analysis: A high LOB% is indicative of a pitcher's ability to work out of jams and strand runners on base. For a closer, who often enters the game in high-pressure situations with runners on base, a higher LOB% is particularly valuable. It suggests that Ortiz has a knack for maintaining control in critical moments and can prevent inherited runners from scoring.
- b. Context: Being significantly above the elite mean in this statistic implies that Ortiz excels in high-pressure situations, a hallmark of a successful closer. It's a skill that can have a direct impact on the outcome of games.

3. FAv (5.74 Percent Above Elite Mean):

- a. Analysis: A higher fastball velocity can be a significant advantage for a closer. It often correlates with higher strikeout rates and can be intimidating for hitters, especially in the late innings of a game.
- b. Context: Ortiz's fastball velocity being notably higher than the elite mean suggests that he has the raw power often sought in a closer. This ability to overpower hitters can be crucial in securing the final outs of a game.

In conclusion, Luis L. Ortiz's statistical profile, characterized by a higher than average Swing%, LOB%, and FAv compared to elite closers, supports the idea that he could excel in a closing role. His ability to induce swings, effectively strand runners, and deliver a powerful fastball align well with the key attributes of successful closers in high-pressure situations. These statistics, when taken together, provide a strong analytical basis for considering Ortiz for a closer position.



Player Two: Chase Silseth LAA



Current Role: Starting Pitcher

Common Inning Pitched:

Suggested Role: Setup/Closer

Throws: R

Age: 23

2023 Stats

Wins	4
Losses	1
ERA	3.96
G	16
IP	52.1
SO	56
WHIP	1.28

Analysis & Discussion:

Chase Silseth has come off an average season as a starting pitcher. Our analysis proves he is suitable in a setup pitcher closing role since his statistical distributions are close to the mean of closers who meet our 'Elite Pitcher' criteria.

1. Strikeouts (SO 1.82 percent below the Elite Mean):
 - a. Analysis: Strikeouts are a critical measure of a pitcher's ability to control the game by directly eliminating hitters. For a setup man or closer, high strikeout rates are valuable because they reduce the chance of balls in play, which can lead to runs, especially in high-leverage situations.
 - b. Context: Silseth being slightly below the elite mean in strikeouts indicates he's nearly on par with elite closers in terms of his ability to get batters out without relying on fielding. His strikeout capability, while marginally lower, still suggests effectiveness in situations where a strikeout is particularly crucial.



2. WAR (Wins Above Replacement) (1.8 percent below the Elite Mean):

- a. Analysis: WAR is a comprehensive statistic that estimates a player's overall contribution to their team in terms of wins. While traditionally used more for everyday players, WAR can still provide valuable insight into a pitcher's overall impact.
- b. Context: Silseth's WAR being slightly below the elite mean suggests that his overall contribution to his team is nearly at the level of elite closers. This indicates a solid all-around performance, taking into account not just his pitching but the overall impact of his appearances.

3. Pitchbot Fastball (<1 percent below the Elite Mean):

- a. Analysis: The fastball is a crucial pitch for any reliever, especially in a setup or closing role. Its effectiveness can be a major factor in a pitcher's success, as a dominant fastball can overpower hitters and set up other pitches.
- b. Context: Silseth's fastball quality being just below the elite mean indicates that he possesses a strong and effective fastball, nearly on par with those of elite closers. This can be a significant asset in high-pressure situations where fast, effective pitching is paramount.

In conclusion, Chase Silseth's profile, as indicated by his strikeout rate, WAR, and fastball quality, suggests he is well-suited for a role as a setup man in closing situations. His ability to generate strikeouts, overall contribution to team success, and the effectiveness of his fastball are all nearly at the level of elite closers. These attributes, combined with the role-specific demands of a setup pitcher in closing situations, make a strong case for his suitability in this role. His performance in these key areas, despite being slightly below the elite mean, shows he possesses the necessary skills and effectiveness for high-leverage late-game situations.



Player Three: Jovani Moran MIN



Current Role: Middle Reliever

Suggested Role: Setup/Closer

Throws: L

Age: 26

2023 Stats

Wins	2
Losses	2
ERA	5.31
G	43
IP	42.1
SO	48
WHIP	1.46

Analysis & Discussion:

Jovani Moran has come off an average season as a starting pitcher. Our analysis proves he is suitable in a closing role since his statistical distributions are close to the mean of closers who meet our 'Elite Pitcher' criteria.

1. Swinging Strike Percentage (8.59 percent above the elite mean):
 - a. Analysis: A high swinging strike percentage is a key indicator of a pitcher's ability to miss bats, which is crucial for a closer. The ability to generate swings and misses leads to strikeouts and reduces the chance of batters making contact, thereby lowering the risk of hits and runs.
 - b. Context: Being significantly above the elite mean in swinging strike percentage suggests Moran has an exceptional ability to fool hitters and get



them to chase or miss pitches. This skill is especially valuable in high-pressure, late-inning situations where avoiding contact is often critical.

2. Pitch Bot Stuff Rating (2.72 percent above the elite mean):

- a. Analysis: A high 'stuff' rating indicates that Moran possesses an effective and perhaps diverse range of pitches. This is important for a closer, who needs to be able to reliably get outs in short, high-leverage appearances. Quality pitches are essential for keeping hitters off balance.
- b. Context: Being above the elite mean in this category implies Moran's pitches are not just effective, but potentially more deceptive or harder to hit compared to those of many elite closers. This could be a significant advantage in a closing role.

3. Opponent Batting Average (1.08 percent below the elite mean):

- a. Analysis: A low opponent batting average indicates that hitters generally struggle to make successful contact against Moran. For a closer, minimizing hits is essential, as every base runner represents a significant threat in the late innings of a close game.
- b. Context: Performing better than the elite mean in this metric shows Moran's effectiveness in limiting hits. This ability is crucial for a closer, who often faces the best hitters in high-pressure situations.

In summary, Jovani Moran's statistical profile — marked by a high swinging strike percentage, an above-average pitch bot stuff rating, and a lower-than-average opponent batting average compared to elite closers — strongly suggests he could be well-suited for a closing role. His ability to generate swings and misses, coupled with the quality of his pitches and effectiveness in limiting hits, aligns well with the key attributes of successful closers, particularly in critical game situations. These statistics collectively provide a compelling analytical basis for considering Moran as a potential closer.

