# The Impact of Shortstop Positioning on Fielding Performance: An Analysis of the "Ball Acquire" and "Throw (ball-in-play)" Sequence


## Abstract

This report aims to quantitatively assess the impact of a shortstop's positioning on fielding performance in baseball, filling a notable gap in the sport's existing literature. Utilizing data from 97 distinct games, we analyzed metrics such as fielding success rate, lateral and depth range, and movement distance. Our findings indicate a positive correlation between a shortstop's broader lateral and depth ranges and higher fielding success rates. Additionally, contrary to conventional wisdom, increased movement distances were not found to negatively impact success rates. While our logistic regression model captures key factors, it highlights the need for further research to account for variables like reflexes and ball speed. Overall, the study has significant implications for data-driven coaching strategies and player evaluation, contributing to a more nuanced understanding of baseball.

# 1. Introduction

In the intricate dance of baseball strategy, the shortstop emerges as a pivotal figure. Positioned centrally in the infield, they not only play a critical role in fielding but also in orchestrating the defense, ensuring optimal positioning and reaction to unpredictable plays. The qualitative attributes of a shortstop—agility, reaction time, and decision-making—are well recognized, but there's a discernible gap when it comes to a data-driven understanding of their contributions.

Fielding performance, particularly for the shortstop, transcends mere ball interception. It encapsulates the nuances of efficient play, error mitigation, and maximizing outs. While the broader baseball community appreciates these qualitative aspects, rigorous quantitative analysis remains sparse. The primary purpose of this study is to quantitatively analyze and understand the relationship between the shortstop's positioning and their overall fielding performance in baseball.

# 2. Data

This study delves into the performance of shortstops across various baseball games, analyzing the relationship between shortstop positioning and fielding outcomes, particularly focusing on the "Ball Acquired" and "Throw (ball-in-play)" events.

The analysis synthesized data from 97 distinct forms, concentrating on:

- Game Events: Identified by game_str (unique game labels) and play_id (specific plays or sequences), with event_code marking different play events.
- Player Position: field_x and field_y coordinates spotlighting the shortstop's location during plays.
- Ball Position: 3D coordinates (ball_position_x, ball_position_y, ball_position_z) illustrating the ball's trajectory.

# 3. Methodology

## 3.1 Calculation of the fielding success rate:

The fielding success rate is a paramount metric in baseball analytics, serving as a barometer of a player's defensive efficacy. Historically, the fielding percentage, often termed the fielding average, epitomizes this measure. This metric illuminates the proportion of instances a defensive player aptly manages a batted or thrown ball. The classical computation for the fielding percentage (FPCT) is:

$$\text{FPCT} = \frac{PO+A}{PO+A+E}$$

Where:

- PO is the number of putouts.
- A represents assists.
- E stands for errors.

In the absence of granular fielding statistics like putouts, assists, and errors, we innovatively utilized available data to approximate the shortstop's fielding success rate. The game events dataset proved instrumental in this endeavor, endowing us with nuanced insights into the player's actions. Specifically, the 'Ball Acquired' event denotes the shortstop's successful acquisition and control of the ball. Concurrently, the 'Throw (ball-in-play)' event indicates the shortstop's active engagement in the game's progression, potentially signaling an attempt to effectuate an out or progress the play.

From these events, we discerned a successful fielding attempt as a sequential occurrence of 'Ball acquired' followed by 'Throw (ball-in-play)'. This sequence implies that the shortstop has adeptly fielded and subsequently progressed the play. The aggregate opportunities for fielding plays by the shortstop are demarcated by the total instances of the 'Ball acquired' event.

Performance metrics encompassed:

**Successful Attempts:** A quantitative record of the shortstop's successful play executions.
**Total Opportunities:** An aggregate of potential plays available to the shortstop during a game.
**Success Rate:** A ratio, expressed in percentage terms, encapsulating the shortstop's proficiency in successful play executions.

## 3.2 Calculation of the Shortstop's Fielding Range:

In baseball analytics, the fielding range of a player is a critical metric that quantifies the spatial area a player can effectively cover to make plays. It provides insights into the player's mobility, positioning, and overall defensive prowess. The range is typically defined in two dimensions: lateral (X Range) and depth (Y Range).

**Lateral Range (X Range):**

- Represents the horizontal movement or coverage of the shortstop on the field.
- It is calculated as the difference between the maximum and minimum values of field_x which the shortstop has occupied during a game.

**Depth Range (Y Range):**

- Captures the vertical or forward-backward movement of the shortstop.
- It is determined by subtracting the minimum value of field_y from its maximum value, indicating the depth of field the shortstop covers.

**Positioning Metrics:**

- X Range: Indicates the extent of the shortstop's movement in the horizontal (left-right) direction on the field.
- Y Range: Denotes the breadth of the shortstop's movement in the vertical (up-down) direction on the field.

## 3.3 Calculation of the Shortstop's Movement Distance:

Distance, a fundamental spatial metric, is pivotal in capturing the shortstop's navigational prowess during gameplay. For this study, the Euclidean distance formula, a staple in geometric computations, was employed to quantify the movement distance between two distinct points in a bidimensional plane:

$$Distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

To trace the shortstop's movement trajectory, for every play where the ball is set into motion, we monitor the shortstop for a pre-defined temporal span (defaulted to 5 seconds post ball motion).

The distance traversed by the shortstop is computed based on their initial and terminal positions during this duration.

Key distance metrics incorporated in our analysis include:

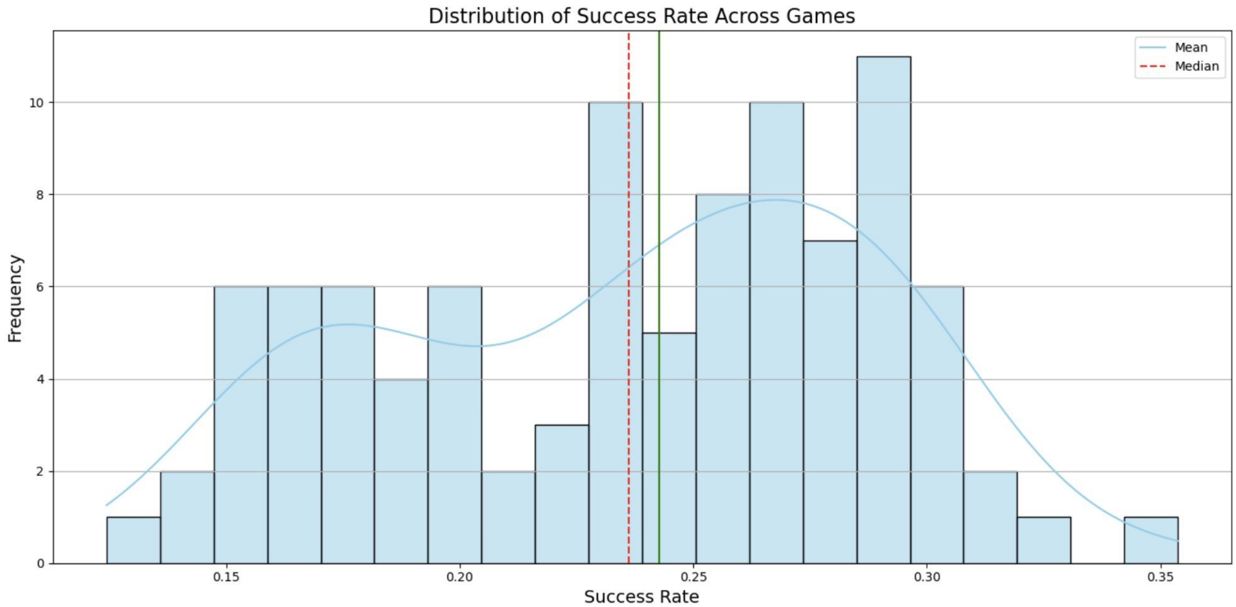
- **Median Distance:** The median spatial gap between the shortstop and the ball during acquisition.
- **First Quartile Distance (Q1):** The 25th percentile of acquisition distances, indicative of the lower bound of typical distances.
- **Third Quartile Distance (Q3):** Representing the 75th percentile, it captures the upper bound of typical acquisition distances.

# 4. Analysis

## 4.1 Success rate

The average success rate for the shortstop across all analyzed games was approximately 23.62%. This provides a benchmark against which individual game performances can be compared. The game showcasing the peak fielding success rate for the shortstop was between TeamLI and TeamA3 in 1901, with a success rate of approximately 35.37%. This indicates the best performance in terms of fielding. In contrast, the game with the most challenges in fielding for the shortstop was between TeamNG and TeamA3 in 1903, where the success rate was at its lowest of approximately 12.44%.

The smooth line (KDE) offers an estimation of the probability density function of the success rate. It provides a clear picture of the shape of the data distribution. The dashed red line represents the mean success rate, offering an average value across all games. The solid green line represents the median success rate, indicating the middle value when success rates are arranged in ascending order. Comparing these two metrics can provide insights into the skewness of the data.

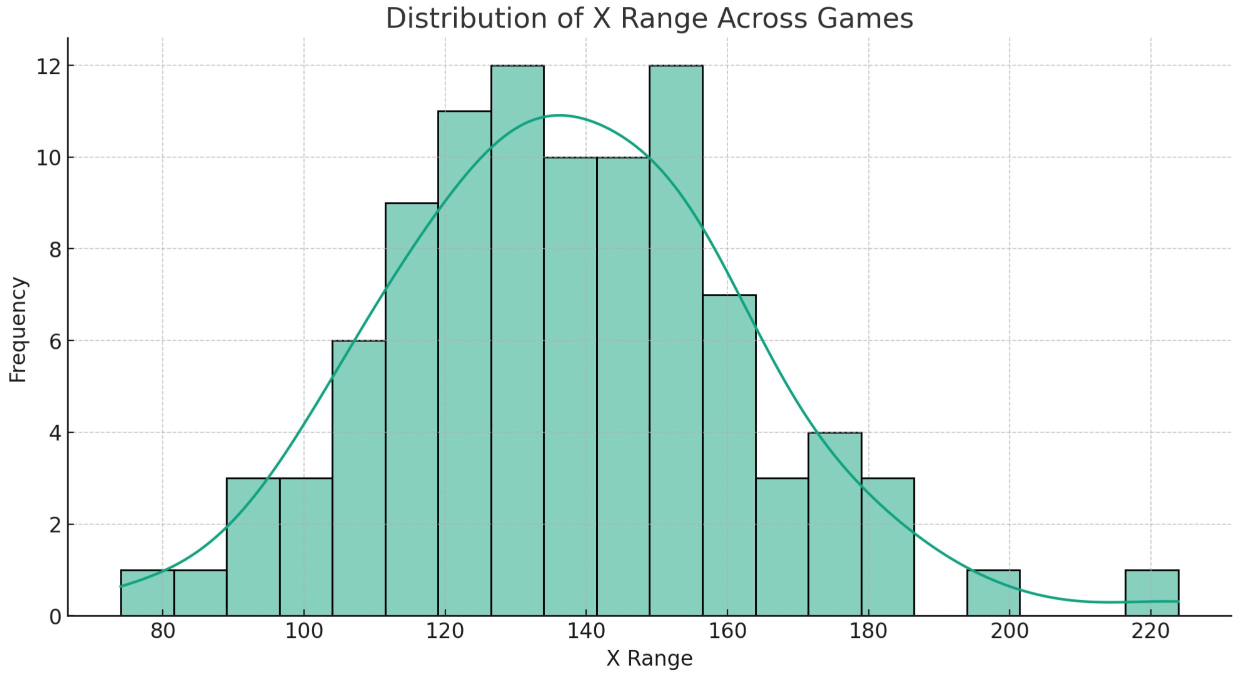



(figure 4.1.1) Distribution of Success Rate Across Games

## 4.2 Range

**Lateral Movement (X Range):**
The shortstop's lateral movement varied significantly across games, with an average X Range of approximately 136.95 units, peaking at 223.91 units and dipping to a low of 74.02 units.This variability indicates that the shortstop frequently adjusts their positioning laterally, possibly in response to batter tendencies, pitch types, or infield dynamics.

Scenarios with larger X Ranges might be instances where the shortstop had to cover more ground laterally, potentially impacting their ability to swiftly transition between acquiring the ball and executing a throw.

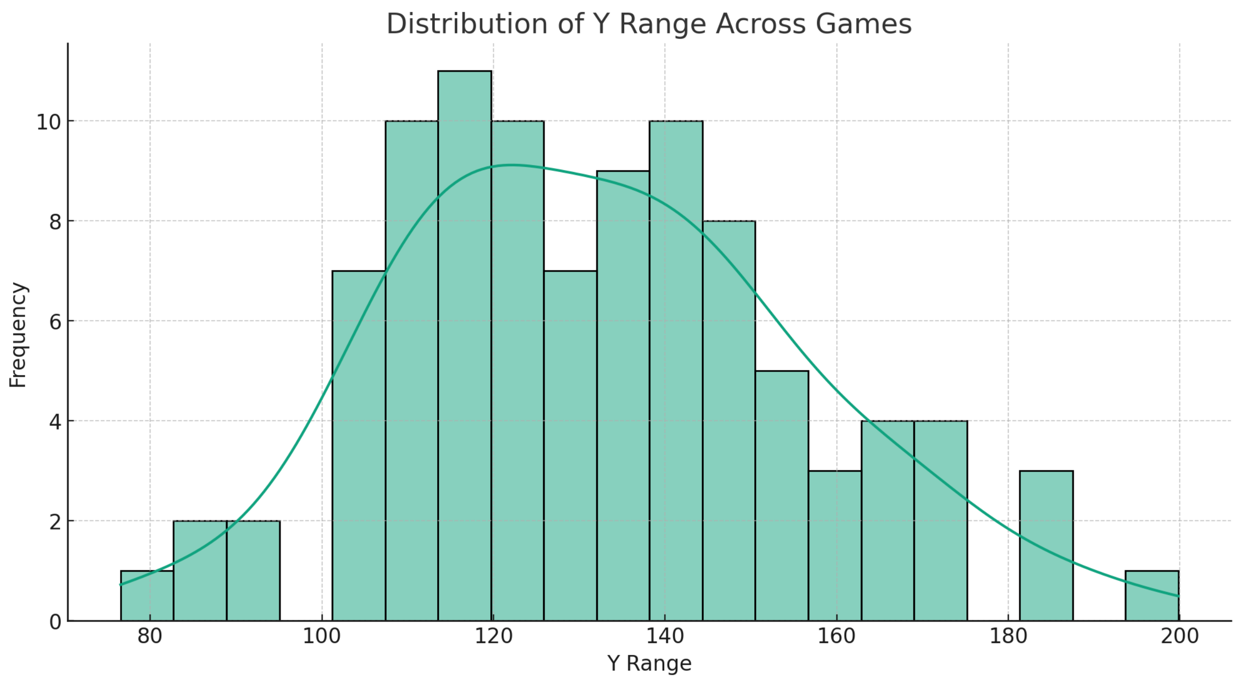



(figure 4.2.1) Distribution of X Range Across Games

**Depth in the Field (Y Range):**

The shortstop demonstrated relatively consistent depth positioning, with the average Y Range being approximately 130.03 units.However, there were deviations, with the Y Range reaching a maximum of 199.82 units and a minimum of 76.55 units.

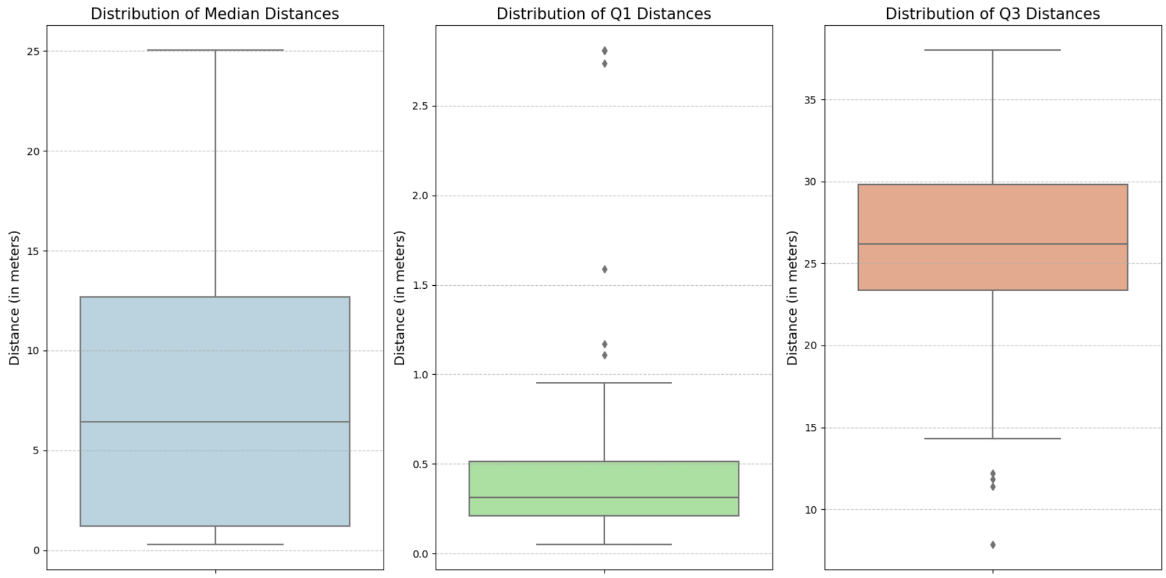
Consistency in depth positioning suggests a strategic approach, possibly derived from game plans or batter analysis. Yet, the deviations point towards game-specific challenges or the need for adaptability based on unfolding game scenarios.




(figure 4.2.2) Distribution of Y Range Across Games

## 4.3 Distance

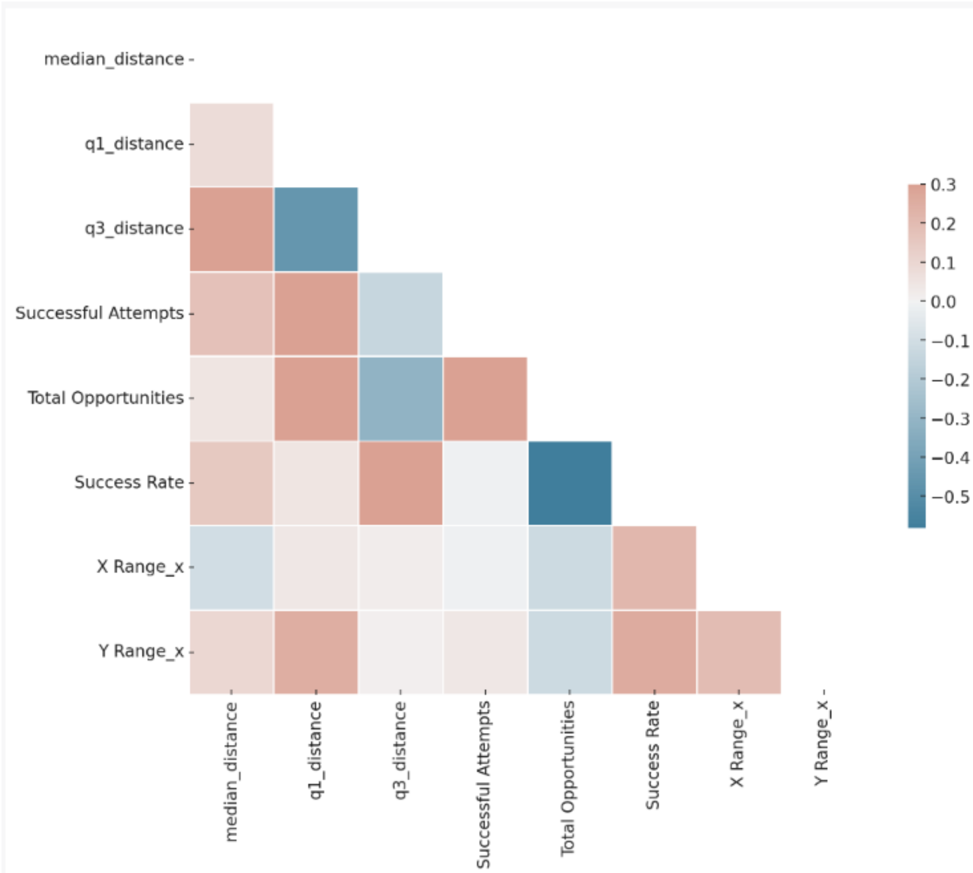
The Median Distances analysis reveals that shortstops, on average, cover approximately 7.27 meters across games. Notably, the range indicates that while in some plays shortstops are optimally positioned with minimal movement, there are instances where they traverse as far as 25.04 meters. For the Q1 Distances, which represent the 25th percentile, data suggests that in a significant number of plays, shortstops are already in near-optimal positions. The average movement in these scenarios is a mere 0.46 meters, emphasizing strategic positioning. On the other hand, the Q3 Distances, capturing the 75th percentile, highlight the agility and range of the shortstop. On average, they move about 26.16 meters in these plays, underscoring the dynamism of the role and the unpredictable nature of baseball.



(figure 4.3.1) Distribution of Shortstop Movement Distances Across Plays: Median, Q1, and Q3 Metrics

From the heatmap (figure 4.3.2), there is a moderate negative correlation between the 'Success Rate' and the 'median_distance', 'q1_distance', and 'q3_distance'. This suggests that as the distance moved by the shortstop increases, the success rate of fielding attempts tends to decrease.

Besides, a moderate positive correlation between 'X Range_x' and 'median_distance', 'q1_distance', and 'q3_distance is observed', indicating that as the range of movement along the X-axis increases, the distances moved by the shortstop also tend to increase.



(figure 4.3.2) Heatmap for Shortstop’s Movement Distance, Fielding Range, and Fielding Success Rate

# 5. Key Findings

## 5.1 Correlation between Range and Success Rate

Lateral Movement (X Range) exhibits a correlation of 0.190 with success rate, indicating that broader lateral ranges might correlate with higher fielding success, potentially reflecting the shortstop's agility. Depth in the Field (Y Range) shows a correlation of 0.278 with success rate, suggesting that increased depth adjustments may lead to better outcomes, possibly due to the shortstop's adaptability to game scenarios.

The range metrics shed light on the shortstop's positioning and movement patterns in the X and Y directions. A broader range, both laterally and in depth, is positively correlated with higher success rates. This suggests that a shortstop's adaptability and ability to cover a broader area efficiently are crucial for fielding performance. While broader ranges are associated with higher success rates, it's the shortstop's ability to move efficiently within these ranges that likely contributes to fielding success.

### 5.2 Correlation between Distance and Success Rate

The Median Distance displays a correlation of 0.242 with success rate, suggesting games with greater shortstop movement typically exhibit higher success rates, potentially due to optimal positioning. The Q1 Distance shows a weak correlation of 0.060, indicating minimal influence on success rate. Meanwhile, the Q3 Distance, with a correlation of 0.338, suggests that longer movements in 75% of plays are associated with higher success rates, possibly reflecting the shortstop's agility and decision-making.

Contrary to the expectation that less movement might correlate with success, the data indicates that optimal positioning and active involvement, leading to increased movement, could actually enhance fielding success. This underscores the importance of agility, anticipation, and rapid reactions in a shortstop's performance.

### 5.3 Positioning vs. Performance

While the positioning of the shortstop on the field is essential, it's evident that it doesn't singularly dictate the success rate of their fielding performance. There might be inherent complexities, like the player's reflexes, the speed of the ball, or even unpredictable in-game situations, that determine a successful fielding attempt.

## 6. Limitation

Our work tried to evaluate the fielding performance of shortstops based on the game information given. While a high fielding percentage is regarded as a sign of defensive skill, it is also possible for a player of lesser defensive skill to have a high fielding percentage, as it does not reflect or take into account a player's defensive range; a player who cannot get to a ball surrenders a hit instead of having an opportunity to make an out or an error. Conversely, a highly skilled fielder might have a comparatively low fielding percentage by virtue of reaching, and potentially missing, a greater number of balls.

## 7. Conclusion

In conclusion, this comprehensive analysis report offers valuable insights into the critical relationship between a shortstop's positioning and fielding performance in baseball. Our findings reveal a positive correlation between broader lateral and depth ranges with higher success rates, emphasizing the importance of a shortstop's adaptability and agility. Conversely, the study challenges conventional wisdom by suggesting that greater movement distances are not necessarily detrimental to success rates, but may instead reflect a state of optimal readiness and involvement in the game.

Importantly, we recognize that positioning alone cannot encapsulate the complexities of fielding performance. Factors such as reflexes, ball speed, and situational unpredictabilities also play a significant role. Overall, our study has significant implications for coaching strategies and player evaluation, potentially providing teams with a data-driven edge in optimizing their infield defense. It fills a notable gap in quantitatively assessing the shortstop's role, thereby contributing to a more nuanced understanding of baseball as a sport that is both rich in tradition and ripe for analytical exploration.