Analysis on Bat Speed and Swing Length

Matthew Bressler, Gyeongwu Kim, Kyle Zhou

2025-1-15

Introduction

The release of the book *Moneyball: The Art of Winning an Unfair Game* in 2003, followed by the movie in 2011 shed light on the surge of the use of analytics in the game of baseball. Since then, the field has grown dramatically, with organizations and players doing everything they can to leverage analytics to gain insight on in-game strategy, help with making strategic roster construction decisions, and enhance player development. However, an untapped area of baseball analytics is tracking player biomechanics in game.

In recent years, private sector companies have begun implementing biomechanics assessments for their athletes during training to provide insight on mechanical improvements. For example, pitchers use motion capture during bullpens to track their movements to see if there is a deficiency in a certain part of the throw. This information has been largely inaccessible for in-game tracking until recently. With the implementation of Hawk-Eye and Statcast features, we are now beginning to have access to real time tracking metrics.

Statcast has been effective in tracking the movement of the baseball, capturing aspects like exit velocity, launch angle, pitch spin axis etc. Recently, they upgraded their platform, allowing for tracking data on a player's swing. These enhancements provide us insights into swing mechanics, including bat speed and swing length.

In this report, we use new Statcast data on in-game player bat tracking from the 2024 MLB season. Specifically, we use two key metrics:

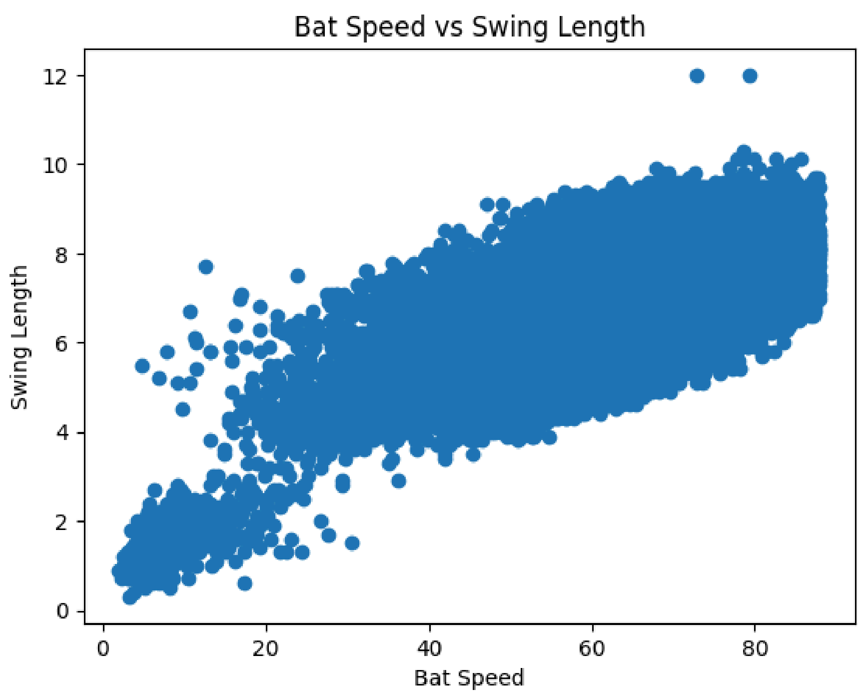
1. Swing Length: The total (sum) distance (in feet) traveled by the barrel of the bat in X/Y/Z space, from the start of bat tracking data (generally around 150 ms), up until impact point.

2. Bat speed: how fast the sweet spot of the bat is moving, in mph, at the point of contact with the ball (or where the ball and bat would have met, in case of a swing-and-miss).

In this report, we aim to use data on bat speed and swing length to assess players' performance metrics. We started by observing league wide trends, then shifted to doing analysis on individual players. We looked at performance in different situations, and created a model to predict player performance.

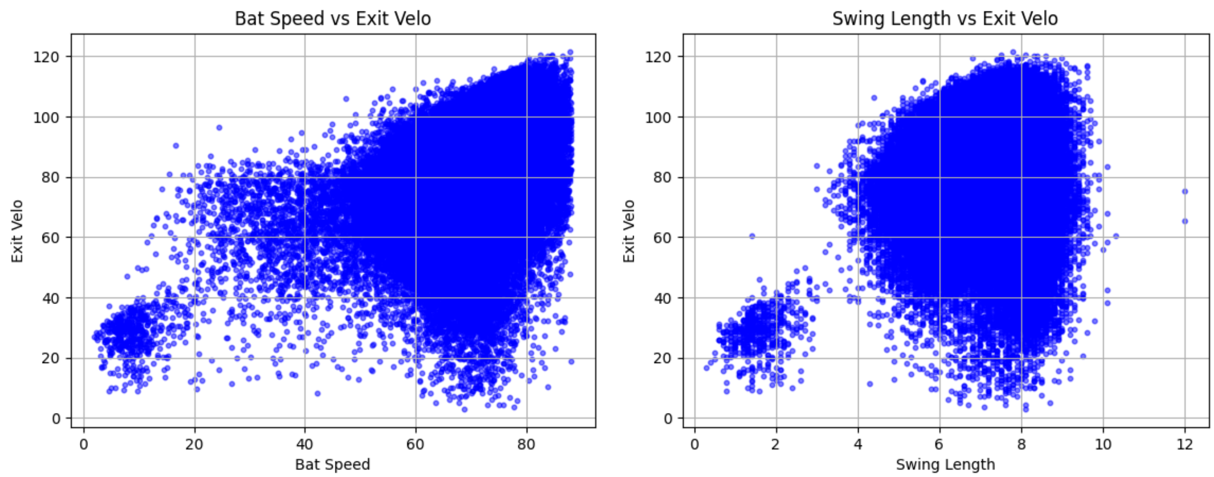
Exploratory Data Analysis

We begin our analysis by exploring general trends when it comes to these bat tracking statistics, and how they relate to each other. First, we investigate how bat speed and swing length are related. We use a scatter plot to plot every single tracked swing from the season, where swing length is measured in feet and bat speed is measured in miles per hour. This allows us to assess how these two variables interact across the dataset.



We notice a clear positive correlation between these two variables. As swing length increases, bat speed tends to increase as well. Intuitively, this trend makes sense and is explained by simple physics. The longer a batter's swing, the more time the barrel has to accelerate, allowing for a higher bat speed at the point of contact.

This highlights the importance of how swing length can contribute to bat speed. However, further analysis is needed to determine how this relationship translates to performance metrics. We observe how each of these factors contribute to exit velocity, which can be an indicator of batter effectiveness. Exit velocity measures how fast the ball leaves the bat after contact.



We notice that for all batted balls throughout the season, there is not a strong association between either bat speed or swing length with higher exit velocities. The points are relatively scattered. The points with high bat speed and swing length but poor exit velocity could indicate poor contact quality. There is however, a strong association with bat speed. With this insight, we can conclude that bat speed is more of a contributing factor to exit velocity than swing length is. This can be due to a variety of factors, such as greater variability in terms of precision when a batter takes a longer swing, or a longer swing taking more time to swing and less to react to the pitch.

So far, all of our analysis has been done on each individual pitch throughout the season. This allows us to look for general trends across the league. However, we want to break the data down and do further analysis on individual players.

Creating a Player Profile

To deepen our analysis, we shift our focus from league-wide trends to individual players. To do this, we created a player profile. For each player, we aggregated all the pitches that a player swung at to calculate their average swing length and average bat speed. We also took their weighted on-base average (wOBA) from those pitches in order to do some performance analysis. The structure is listed out in the table below.

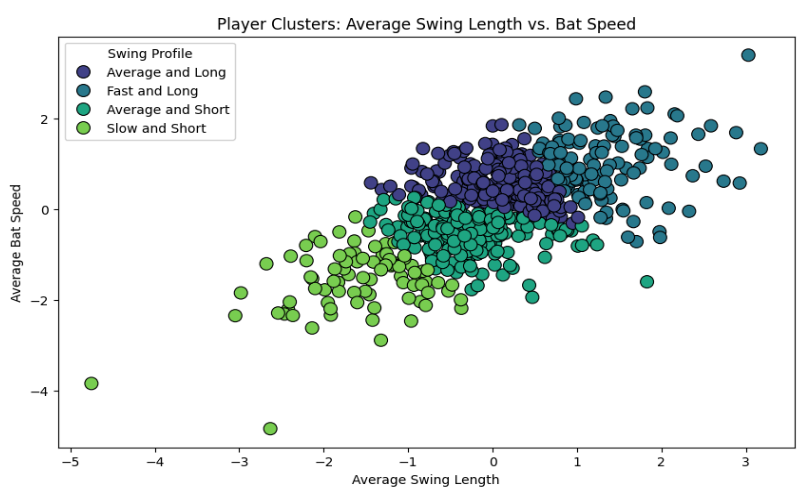
	player_name	avg_swing_length	avg_bat_speed	woba_value
0	Abrams, CJ	7.365126	70.224533	0.365176
1	Abreu, José	6.994067	71.232184	0.177670
2	Abreu, Wilyer	7.321706	72.818641	0.362255
3	Acuña Jr., Ronald	7.512827	74.580044	0.329803
4	Adames, Willy	8.087973	72.580638	0.322585
...	...	...	...	...
554	Yelich, Christian	7.404644	71.339515	0.381116
555	Yoshida, Masataka	7.189295	69.138296	0.309426
556	Young, Jacob	6.515986	63.034302	0.318219
557	Zavala, Seby	7.185108	65.931641	0.237500
558	d'Arnaud, Travis	7.718819	69.944367	0.324854

Characterizing players by swing

Now that we have all the players' individual data on swing length and bat speed, we wanted to categorize them based on their swing tendencies. We wanted to create 4 specific groups for players:

1. Average bat speed and long swing
2. Average bat speed and short swing length
3. Fast bat speed and long swing length
4. Slow bat speed and short swing length

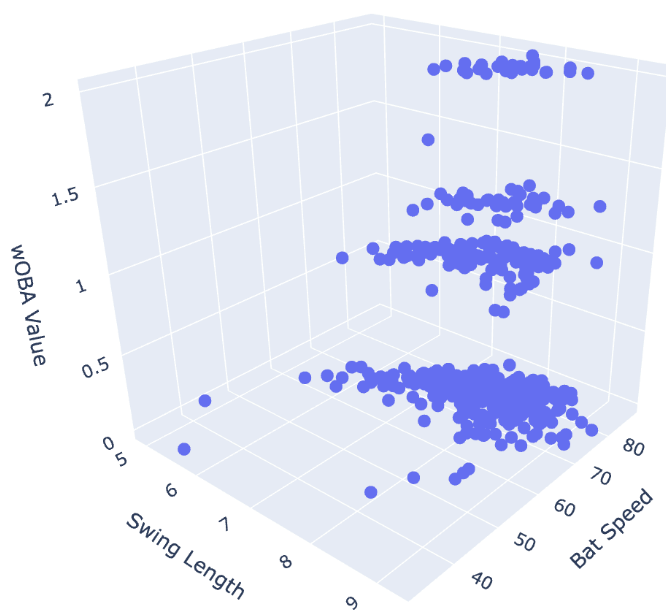
In order to group the players, we used k-means clustering with four clusters to create our desired groups. For a better visualization, we normalized swing length and bat speed.



With players being put into one of four categories, managers can make more informed decisions about situational hitting, lineup construction and player development. For example, in a late game pinch hit situation, if a hard throwing pitcher is on the mound, a manager may opt to use a fast bat speed, short swing length player. It may also be more optimal to have a player with a certain swing tendency at a certain part of the lineup.

Relating to Performance Metrics

To evaluate performance for each player, we wanted to find relationships between bat speed and swing length with other performance statistics. We started by using our player profile to relate these metrics with wOBA. We created visualizations for each individual player. The 3D graph below shows Vladimir Guerrero Jr.'s bat speed on the x-axis, swing length on the y-axis, and wOBA per pitch on the z-axis.



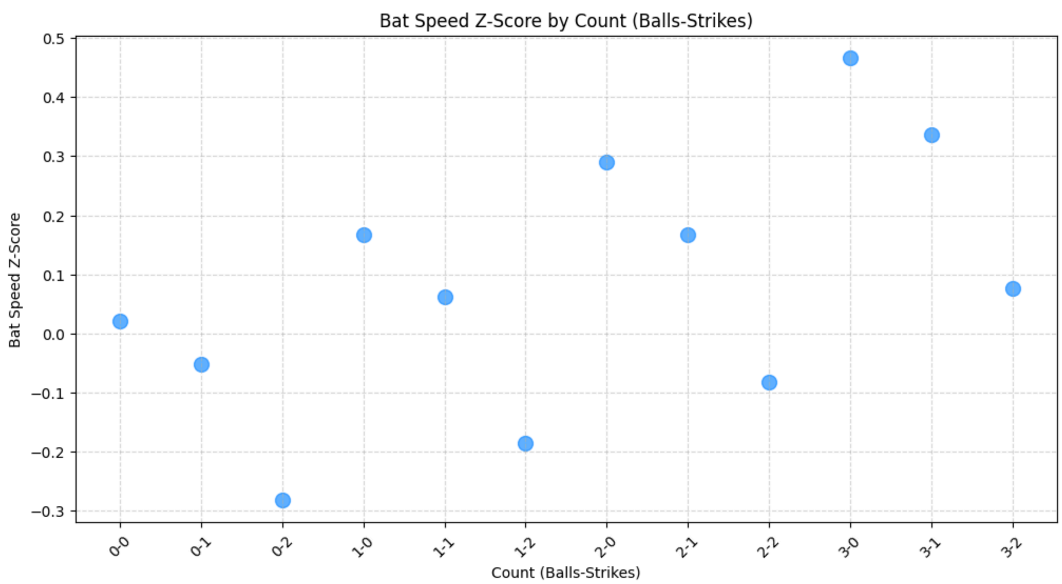
There is a staircase pattern since the wOBA value for each pitch is fixed based on a specific result (out, single, double, etc.) For Vladdy specifically, we can see that for his higher wOBA values, his bat speed and swing length are relatively high. For other players, we can use this information to see if they have a high concentration of certain data points based on their swing tendencies, providing insight on how they should swing.

Hitters in Different Situations

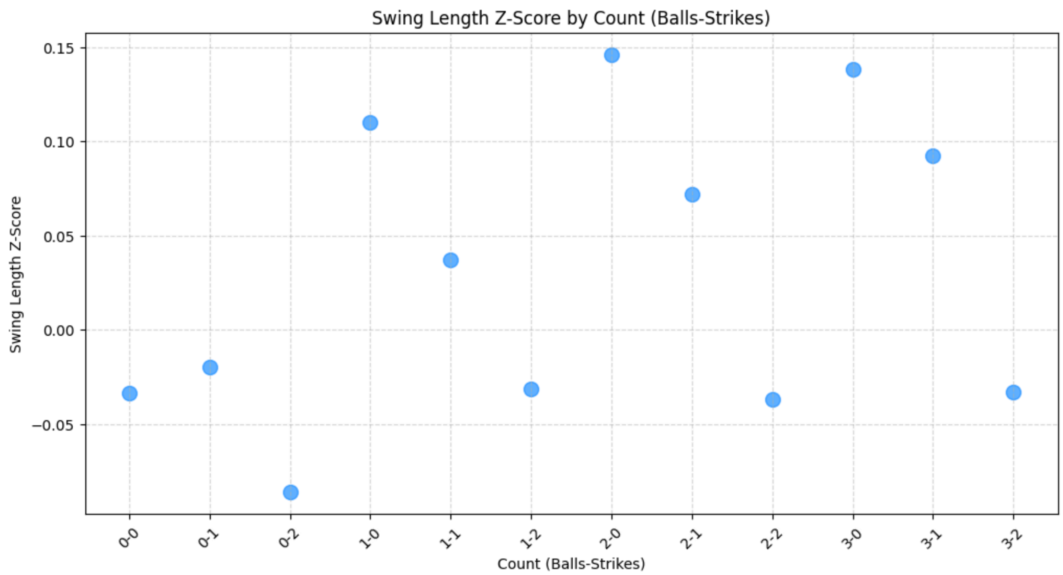
Next, we aim to evaluate player performance based on different situations. The two we focused on were different counts, and seeing different pitch types. We wanted to analyze how batter's may change their swing tendencies.

Analysis By Count

We took the count for each pitch that an individual batter swung at, and found the average bat speed and swing length. We then normalized the data to create a better visualization. We first looked at bat speed for each count.



We observe that for each count with the same amount of balls, there is a general pattern. Bat speed always decreases when there are more strikes. Bat speed decreases even more when there are two strikes, suggesting that the current strategy is to not swing as fast when a batter has more strikes on them. Additionally, when balls increase, players swing increasingly faster. When ahead in the count, players see less risk of swinging faster, potentially trying to produce greater exit velocities. When down in the count, they are more conservative, trying to put the ball in play. We see a similar pattern when evaluating different swing lengths based on the count.



Similarly, when a batter has more strikes on them, they decrease the length of their swing. When there are more balls and when they are ahead in the count, they lengthen their swings, potentially trying to generate more power at contact. Again, when down in the count, players seem to be more conservative, and take shorter swings compared to when they are up in the count. However, there is a discrepancy for the first pitch of the at bat. Players are not as aggressive, and don't take very long swings.

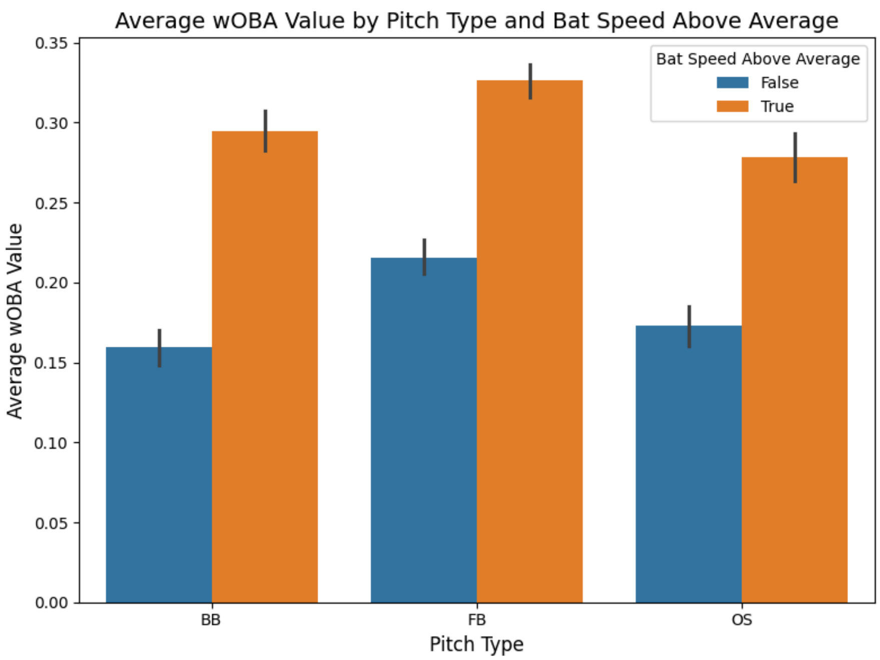
Both bat speed and swing length generate similar patterns when viewed from a certain count, reinforcing the idea that the count greatly impacts a player's swing strategy. Batters tend to be more aggressive in plus counts, while they exhibit conservatism in minus counts.

Analysis By Pitch Type

We took the pitch type of each pitch that an individual batter swung at, and created three separate categories of pitches. We considered fastballs (FB), breaking balls (BB) and off-speed (OS). We considered pitches such as curveballs and sliders to be BB's and change-ups and knuckleballs to the OS. We then separated each pitch type into two categories, one for if the individual player's bat speed was above their average, and one for if it was below their average. To measure performance, we used wOBA and split it based on our categories. The structure is shown in the table below, using CJ Abrams as an example.

	player_name	pitch_type	bat_speed_above_average	avg_woba_value
0	Abrams, CJ	BB	False	0.185000
1	Abrams, CJ	BB	True	0.222685
2	Abrams, CJ	FB	False	0.255634
3	Abrams, CJ	FB	True	0.480195
4	Abrams, CJ	OS	False	0.103846
5	Abrams, CJ	OS	True	0.221094

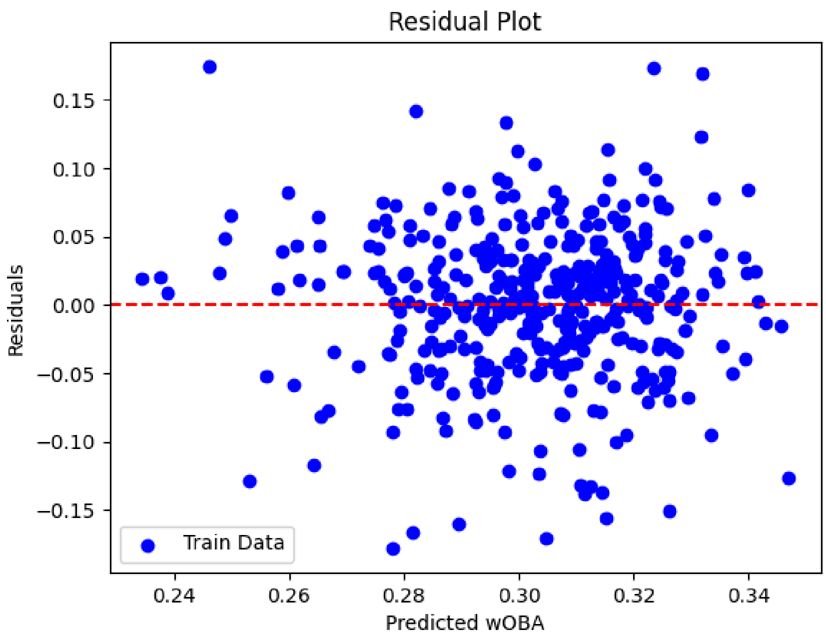
Looking at individual player's, we can determine which type of swing is optimal based on which pitch is thrown. We recognized that several players have higher wOBA values when swinging with an above average bat speed, so we decided to compare the entire league based on pitch type.



We notice that for all pitches, an above average bat speed will generate a higher wOBA value, no matter what type of pitch is thrown. This analysis shows the value of swinging with a faster bat speed in optimizing performance across all pitch types.

Creating a Model for Performance Metrics

In the final step of our analysis, we created a model based on bat speed and swing length to predict wOBA. We used four features: average bat speed, average swing length, standard deviation of bat speed, and standard deviation of swing length. These features were used to train a linear regression model, with wOBA as the target variable. To evaluate the model, we used mean squared error and residuals.



Summary

Our goal was to use biomechanics data, specifically bat path and swing length, to gain insight on batter performance. We analyzed trends across the entire league and evaluated individual players' performance. Additionally, we clustered players based on swing tendencies, and examined how players performed in different situations. With the growth of in-game biomechanics tracking data, further analysis can be done to find patterns in all aspects of a player's mechanics. As we continue to bridge the gap between biomechanics and performance data, our ability to enhance player performance will grow with it.

Works Cited

Lambert, Jack. "Using MLB Bat Tracking Data to Better Understand Swings." *Driveline Baseball*, 1 Aug. 2024, www.drivelinebaseball.com/2024/07/using-mlb-bat-tracking-data-to-better-understand-swings/?srsltid=AfmBOooMJAYFsRPJzZzO2kO8dQ4X3je3Cbu6pkp81doUrqSR4unkyCsw.

"Statcast Search CSV Documentation." *Baseballsavant.Com*, baseballsavant.mlb.com/csv-docs. Accessed 15 Jan. 2025.

"Statcast: Glossary." *MLB.Com*, www.mlb.com/glossary/statcast. Accessed 15 Jan. 2025.