Swing+, Power+, and Contact+: Composite Metrics for Evaluating Hitter Performance in Major League Baseball

Introduction

Data for this project has been sourced from BaseballSavant, a MLB data clearinghouse that provides various offensive and defensive statistics for both teams and individual players. Player performance, pitch tracking, and game events are made publicly available by the MLB's Statcast technology. We have chosen to use the site's offensive statistics for swing speed, as well as other metrics, to evaluate the effects that different swing patterns have on a player's offensive performance. Additional individual performance metrics have been downloaded from FanGraphs, an independent data platform that sources its MLB data from official MLB sources such as Retrosheet and Baseball-Reference.

With these data, we are trying to determine the factors necessary for a 'perfect' swing for each player, dependent on the desired outcome. This will allow managers and players to know and alter their behaviors during the game based on specific situations, giving themselves the best chance to win. The 'desired outcome' for each at-bat is dependent on both the game situation and the player at bat, as some players are more likely to swing for base hits while others often swing harder in an attempt to hit home runs.

A common belief is that a faster swing is better *(Driveline Baseball, 2020)*, while shorter swings are more desirable for base-hits and contact hitters. Based on these theories, the factors we initially want to evaluate for a correlation with individual performance are bat speed and swing length. Bat speed is defined as "an individual swing's speed measured at the point six inches from the head of the bat (sweet spot)" while the swing length is calculated "in terms of feet the bat traveled during the swing, from the start of the bat tracking data up until impact point" *(Baseball Savant, n.d.; MLB.com, n.d.)*. The data for these metrics is given by BaseballSavant for the complete 2023-2024 season. The question we initially set out to evaluate was can measures of a batter's swing, such as bat speed and swing length, predict the success of the hitter?

Thus far, we have looked at correlations between our variables of interest. We have noticed a significant change in contact rates between hitters in an 'elite' bat category and those in a 'poor' bat speed category, with generally higher contact rates for players with slower swings. This exploratory data analysis appears to corroborate the commonly held belief that slower swings are more desirable for base hits.

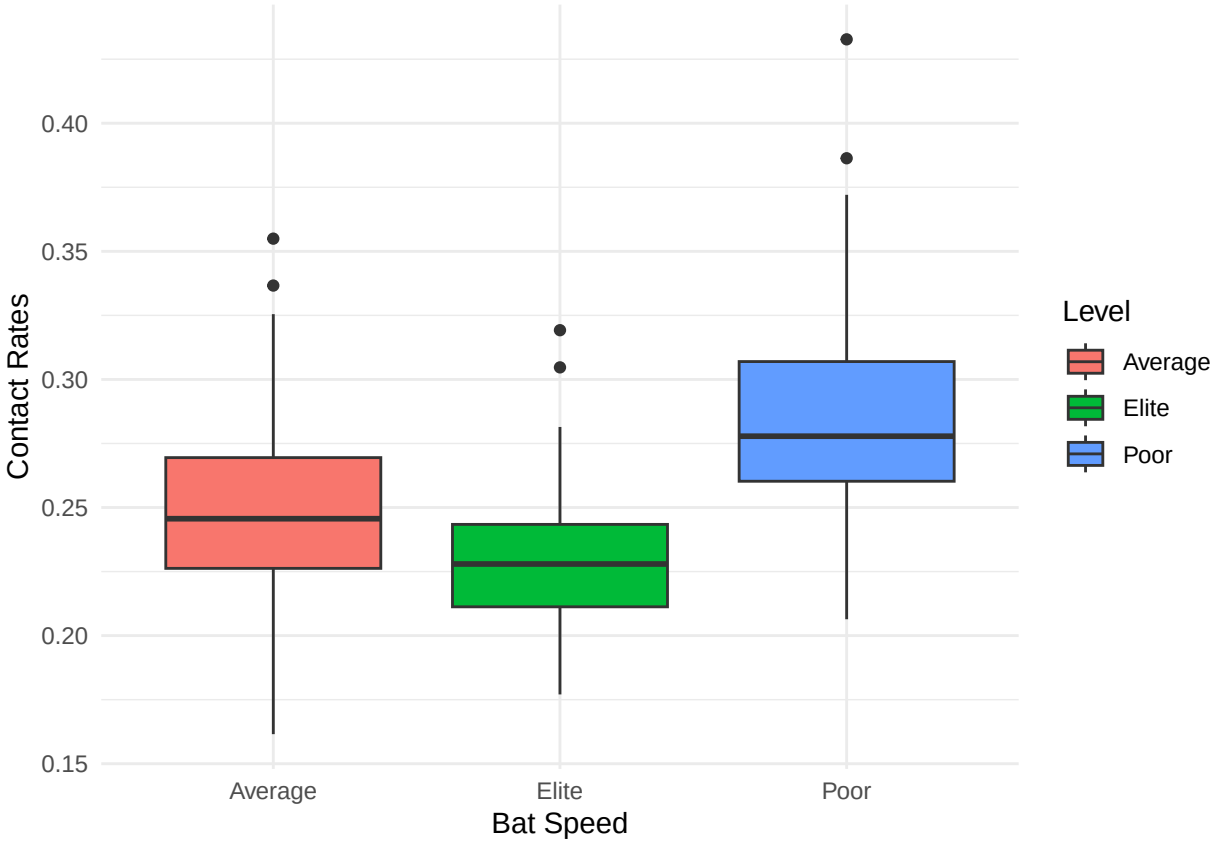

Figure 1: Bat Sped Comparison

Motivation

To better understand the importance of swing metrics like bat speed and swing length, we began by exploring both academic literature and baseball-specific articles to evaluate whether there were established correlations or key gaps in existing research. However, we quickly found that most metrics we hoped to work with, including bat speed, swing length, squared-up rate, and blast rate, had only recently become publicly available through Baseball Savant following the 2023 All-Star break. As a result, there are few academic studies directly evaluating these metrics in the context of in-game performance. Instead, we turned to trusted baseball media and data visualization platforms for insights.

One of the earliest and most influential pieces we reviewed was an *NBC Sports* article that introduced the Statcast bat speed leaderboard and highlighted concerns with interpreting bat speed in isolation. They emphasized that players with elite bat speed, such as Giancarlo

Stanton, often have high whiff rates. In contrast, players with more moderate swing speed and shorter swings, such as Luis Arraez, make more consistent contact *(Samulski, 2024)*. These findings led us to investigate the balance between speed and control and explore how combining these traits could lead to more complete hitters.

Our exploratory data analysis has revealed a strong positive correlation between bat speed and expected weighted on-base average on contact (xwOBACON), a Statcast metric that excludes walks and strikeouts and focuses solely on quality of contact.

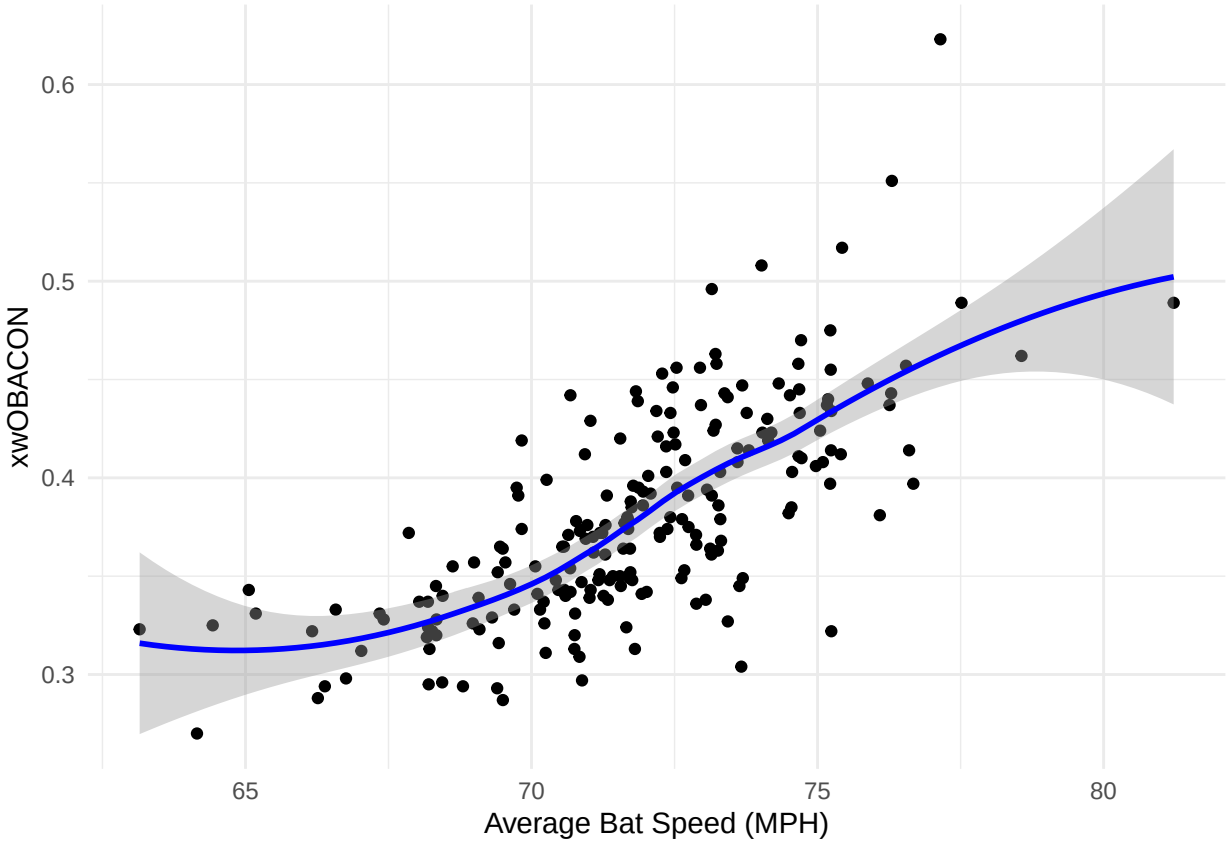

Figure 2: Average Bat Speed and xwOBACON

This aligns with findings from industry sources such as Blast Motion, which emphasized that bat speed is a prerequisite for succeeding at higher levels of baseball but needs to be interpreted within the context of swing timing and mechanics *(Prep Baseball Report, 2024)*.

We also found evidence that hitters alter their swings depending on the pitch count. According to Blast Motion's internal swing studies, bat speed tends to be significantly lower in two-strike counts compared to three-ball counts. Their data shows swings in three-ball

counts average nearly 7 mph faster than in two-strike situations. This supports claims in *The Athletic* that hitters adopt a more compact, contact-focused approach when behind in the count and tend to increase aggression when ahead *(Sarris, 2024)*.

Despite these trends, there is no current consensus in academic or sports reporting on what the "ideal" bat speed is. The variability in approaches based on hitter profile and game context suggests that maximizing performance may come not from targeting universal benchmarks, but from tailoring swings to individual strengths and situational needs. These findings raise several important strategic questions that our ongoing analysis aims to address. Should hitters invest more time in increasing bat speed, or should training focus on maximizing efficiency, contact rate, or swing adaptability? How can swing metrics be adapted to match specific player profiles, particularly between contact-oriented and power-oriented hitters?

To explore these questions, we are building three composite swing metrics: Swing+, Power+, and Contact+. These metrics integrate standardized z-scores for bat speed, swing length, launch angle, blast rate, squared-up rate, and whiff rate to evaluate swing quality both overall and by hitter type. The goal is to quantify what a "good" swing looks like in different contexts and identify which hitters excel or struggle under specific game conditions.

Data

We have collected several datasets focused on individual MLB batters from the 2023 and 2024 seasons. These include player-level swing data such as average bat speed, swing length, squared-up rate, blast rate, and whiff rate. In addition, we compiled hitting performance metrics such as wOBA, OPS, and wRC+ from FanGraphs. We also merged count-specific bat speed data (e.g., 2-strike and 3-ball situations) using Baseball Savant filters. This allows us to analyze not only overall swing traits but also how players adjust their swings in specific situations.

To better interpret swing effectiveness, we created several new variables by calculating z-scores relative to league averages. These include z-scores for bat speed, swing length, launch angle, squared-up rate, blast rate, and whiff rate. Specifically, we had to treat launch angle as unique because the range of 'good' angles is wide. Instead of taking the distance the value is from the mean, we took the distance the data point is from the extremes of the accepted range. For example, under 10 degrees of launch angle is considered 'bad' since it results in ground balls rather than line drives, therefore a launch angle of 9 degrees would be penalized

slightly in the created metric whereas 11 degrees would have no effect. Vise versa, a launch angle greater than 30 degrees would be a pop fly, resulting in a penalty (negative value) as well. We then developed three composite metrics:

- **Swing+** integrates all swing-related z-scores into a single, standardized score representing overall swing quality.
- **Power+** emphasizes metrics important for power hitters such as bat speed, launch angle, squared-up rate, and blast rate.
- **Contact+** emphasizes traits more relevant to contact hitters such as swing length, squared-up rate, launch angle, and whiff rate (inversely).

These metrics allow us to evaluate how a player's swing performs relative to their archetype, as well as how effectively they match their mechanical profile to their role or outcome goals.

All data were sourced from Baseball Savant and FanGraphs. Baseball Savant is operated by MLB and provides advanced Statcast tracking data, including swing mechanics and pitch-level outcomes. FanGraphs is a highly regarded analytics site that aggregates official MLB statistics and provides standardized metrics like wRC+ and OPS. Both platforms are widely trusted within the baseball community and are regularly used by teams, analysts, and fans alike.

There are several limitations in terms of data availability. First, bat speed data has only been released for one full season, so we are currently unable to examine long-term trends, aging curves, or historical comparisons. Second, while bat speed data can be filtered by count, exit velocity data cannot, which restricts our ability to study how quality of contact changes across different pitch situations. We also do not have access to individualized pitch tracking or biomechanical markers such as swing plane or time to contact, which would be valuable for understanding fine-grained mechanical adjustments.

Our current dataset allows us to explore our major area of interest:

- **Building the 'perfect' swing for different desired outcomes**
  - How can swing metrics such as bat speed and swing length be used to predict a hitter's success?
  - How do bat speed, swing length, and related variables contribute to exit velocity and offensive production?

- Can composite metrics like **Swing+**, **Power+**, and **Contact+** help identify which swing profiles are most effective for different types of hitters?

These questions focus on both performance prediction and in-game strategy, using publicly available Statcast data to bridge player mechanics and outcomes.

We are using descriptive statistics, correlation matrices, and data visualization tools including scatterplots, box plots, and bar charts to explore and communicate relationships in the data. For our composite metrics, we applied z-score standardization across all relevant variables. We have also considered using linear models to support interpretation, especially when analyzing the impact of swing traits on performance outcomes like wRC+.

These tools provide interpretable, accessible insights that can be understood by both analysts and non-technical audiences. Our emphasis is on clarity and practical application rather than complex machine learning, which may obscure the meaning behind individual predictors. By using simple statistical techniques and visual analysis, we can more effectively communicate how specific swing traits contribute to success.

Our analysis is limited by the scope and availability of Statcast data. Because some swing metrics are only available at the player-season level, we cannot conduct pitch-by-pitch modeling. Additionally, our composite metrics are based on equal weighting of z-scores, which may not reflect the actual contribution of each metric to performance. More advanced techniques, such as principal component analysis or regression-based weighting, could improve future versions of these metrics.

Our approach assumes that the swing data collected by Statcast is accurate and representative of game situations. We also assume that seasonal averages are sufficient to evaluate player tendencies, even though some in-game context is not captured. For correlation analysis and linear modeling, we assume approximate linear relationships between swing traits and performance metrics, which generally holds true based on preliminary scatterplots and EDA.

Analysis

We completed the creation of our composite swing metrics using standardized z-scores based on league-wide distributions. Each metric was built from a different combination of swing

components (e.g., bat speed, swing length, squared-up rate, blast rate, whiff rate), depending on its intended emphasis. After calculating these metrics, we assessed their predictive power relative to both individual traits and overall performance.

Table 1: Top 10 Players by Swing+

Player	Swing+
Giancarlo Stanton	3.227859
Juan Soto	2.987501
Aaron Judge	2.954631
Vladimir Guerrero Jr.	2.494678
Shohei Ohtani	2.401150
Kyle Schwarber	2.388373
Gunnar Henderson	2.265243
Yordan Alvarez	2.252849
Yandy Díaz	2.186635
Oneil Cruz	2.100127

Table 2: Top 10 Players by Power+

Player	Power+
Juan Soto	4.9625
Giancarlo Stanton	4.7931
Yordan Alvarez	4.7092
Shohei Ohtani	4.5943
Aaron Judge	4.1572
Kyle Schwarber	3.4558
Bobby Witt Jr.	3.2666
Vladimir Guerrero Jr.	2.9609
Vinnie Pasquantino	2.5203
Oneil Cruz	2.5092

Table 3: Top 10 Players by Contact+

Player	Contact+
Luis Arraez	5.3609
Steven Kwan	4.0413
Ernie Clement	3.0073
Mookie Betts	2.8704
Alex Verdugo	2.8624
Nolan Schanuel	2.6254
Brendan Donovan	2.5737
Isiah Kiner-Falefa	2.4679
Alex Bregman	2.4331
Keibert Ruiz	2.3868

From these leaderboards we wanted to evaluate first how good these metrics are in predicting the performance of players relative to the type of hitter. To evaluate the validity of Power+ as a measure of a hitter's ability to drive the ball, we plotted it against ISO (Isolated Power), a common offensive statistic that captures a player's extra-base hit potential. ISO is calculated as slugging percentage minus batting average and is often used to assess raw power independent of contact rate.

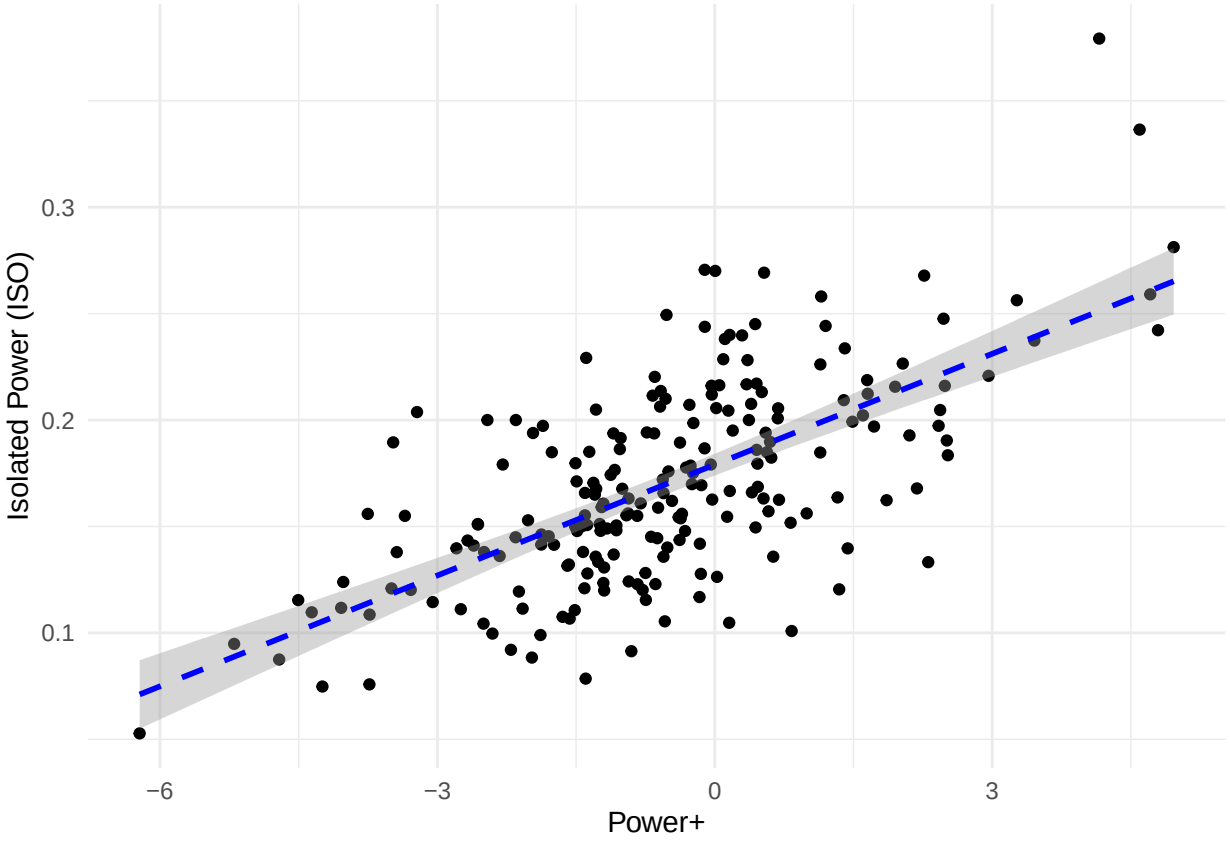

Figure 3: ISO vs. Power+

The positive linear relationship supports the construction of Power+ as a meaningful swing metric. As Power+ increases, ISO also tends to rise, with a correlation coefficient of 0.657. This suggests that hitters with high Power+ scores are more likely to generate extra-base hits. While there is variation at the extremes, this visualization confirms that Power+ effectively captures power traits relevant to game outcomes and aligns well with established hitting metrics like ISO.

To assess the value of Contact+ as a composite metric, we examined its relationship with two core indicators of contact-focused hitting: BB/K ratio (walk-to-strikeout rate) and batting

average (AVG). Because Contact+ is constructed from traits like reduced whiff rate and compact swing length, we expect positive relationships with both BB/K and AVG.

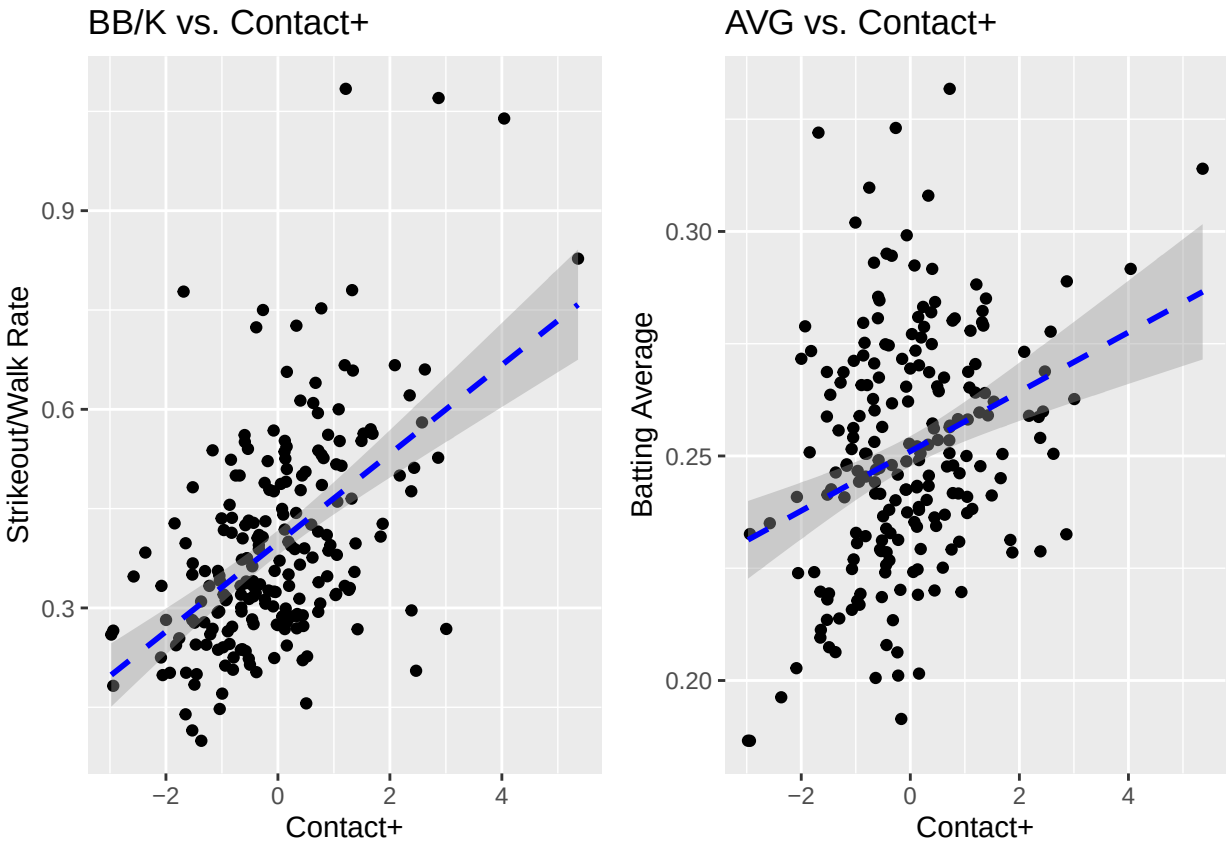

Figure 4: Contact+ Indicators

The analysis shows a correlation of 0.518 between Contact+ and BB/K, and a correlation of 0.313 between Contact+ and AVG. These results suggest that Contact+ captures important swing traits associated with disciplined, contact-oriented hitting. Hitters with higher Contact+ scores tend to strike out less and walk more, which improves their BB/K ratio. Similarly, those with compact, efficient swings tend to generate more frequent base hits, contributing to higher batting

To determine whether our swing metrics correspond to overall offensive output, we tested how each of the three composite scores relates to wRC+. This metric, weighted runs created plus, is one of the most comprehensive indicators of a hitter's offensive value. It adjusts for park and league effects and allows comparison across all player types.

Since wRC+ captures total run contribution, we expect strong relationships with our swing-

based metrics if they are meaningfully constructed. Power+, Contact+, and Swing+ were each plotted against wRC+, and we calculated Pearson correlations to assess the strength of their linear relationships.

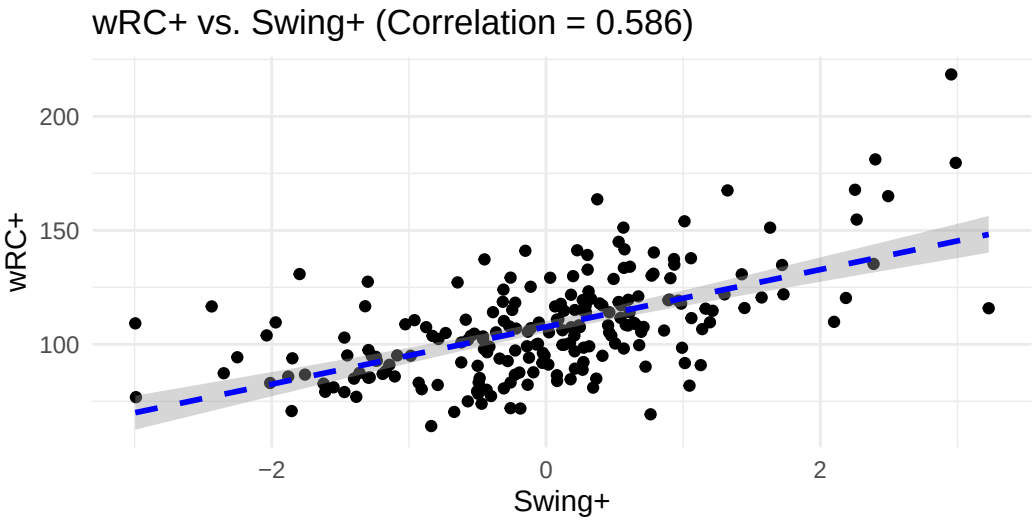

Figure 5: Swing+ vs. wRC+

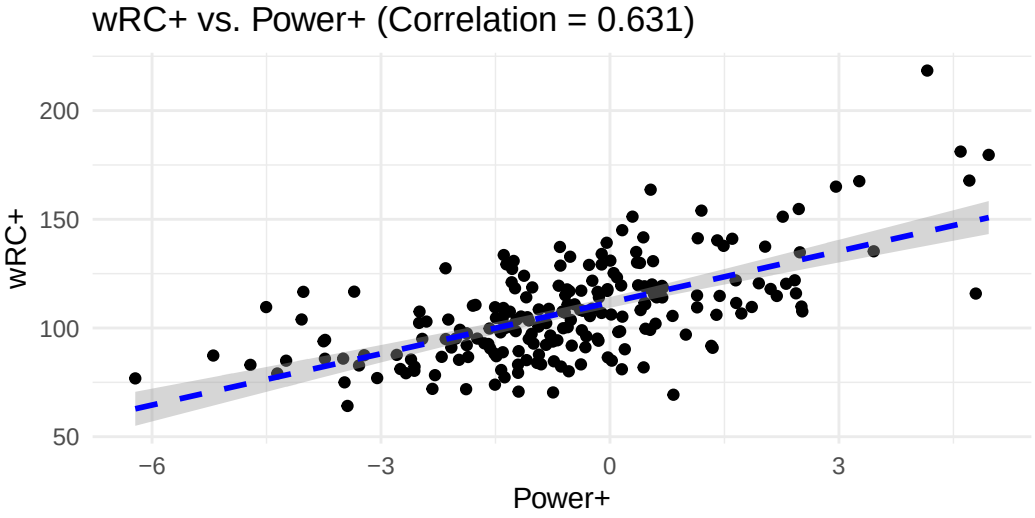

Figure 6: Power+ vs. wRC+

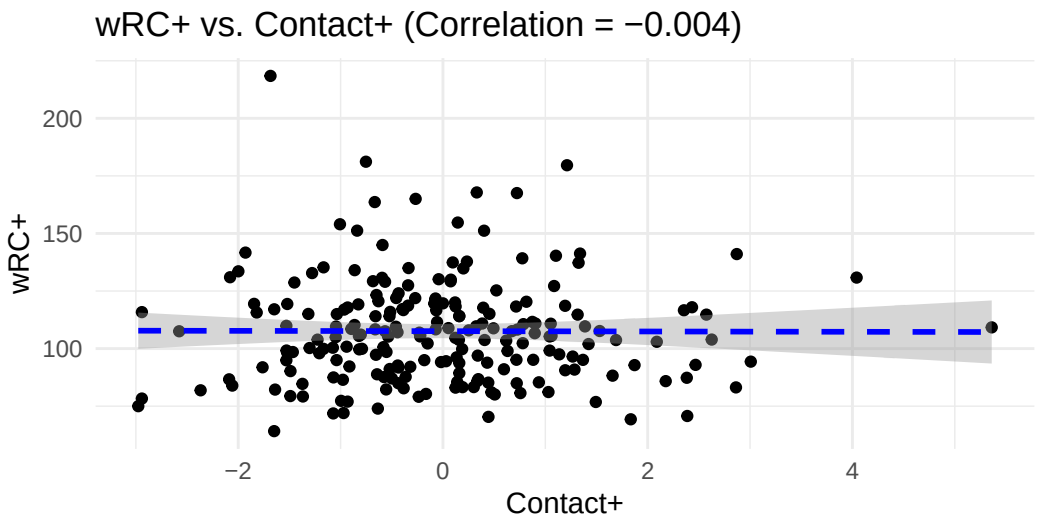

Figure 7: Contact+ vs. wRC+

The results above reveal distinct relationships between each composite metric and overall offensive performance, as measured by **wRC+**.

- **Power+** showed the strongest correlation with **wRC+** (r = 0.631), suggesting that traits such as bat speed, launch angle, squared-up rate, and blast rate are highly predictive of a hitter's run-producing potential. This confirms that **Power+** is a meaningful metric for identifying players who generate value through impactful contact.
- **Swing+** also displayed a strong positive correlation with **wRC+** (r = 0.586). Because **Swing+** integrates all component metrics without prioritizing one swing style, this result indicates that our all-around swing metric is a solid proxy for offensive effectiveness across player types.
- In contrast, **Contact+** exhibited virtually no relationship with **wRC+** (r = –0.004). This suggests that swing traits associated with consistency and control do not directly translate to run production. While **Contact+** remains useful for contextualizing approach and in-count strategy, it may not be a reliable predictor of overall value.

These findings support the notion that power-based metrics more closely align with traditional measures of offensive output. While **Contact+** may still matter for extending at-bats or setting up lineup depth, its limited correlation with outcomes like **wRC+** implies a narrower role in performance modeling.

Next we wanted to look at some specific players to see how they are positioned across the different metrics. We chose Juan Soto, one of the best hitters in the league; Mookie Betts, the only player in the top 25 of both Power+ and Contact+; Luis Arraez, known for being the league's best contact hitter; and Keibert Ruiz, a player who struggled last season.

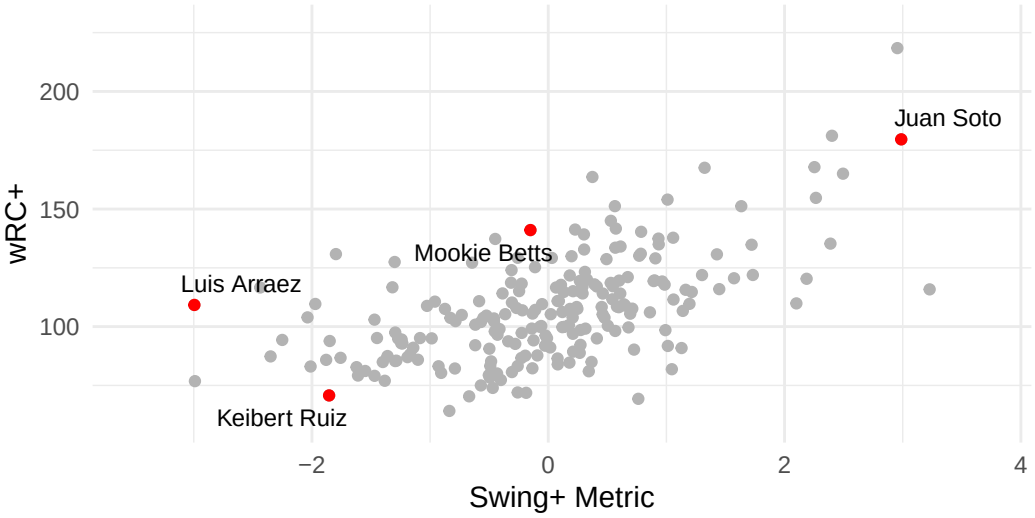

Figure 8: Highlighted Swing+ vs. wRC+

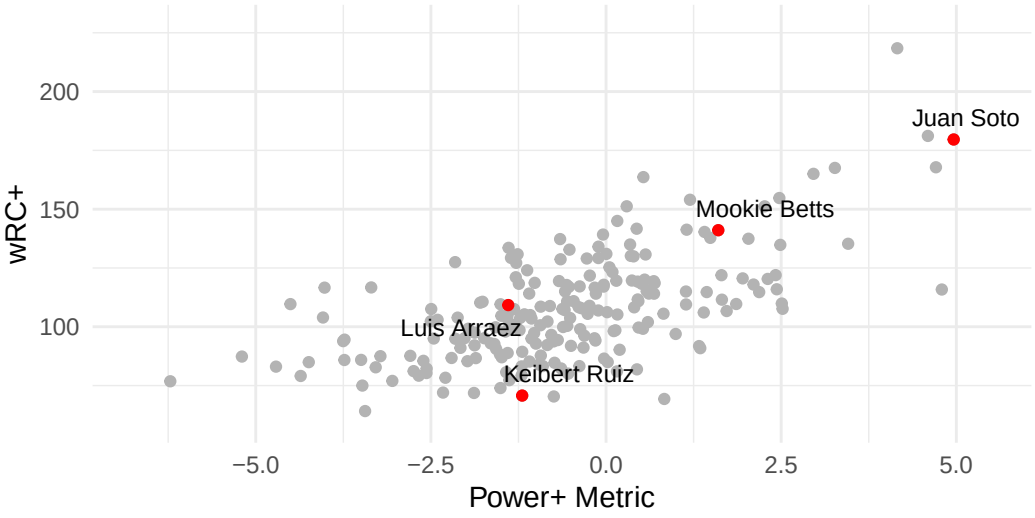

Figure 9: Highlighted Power+ vs. wRC+

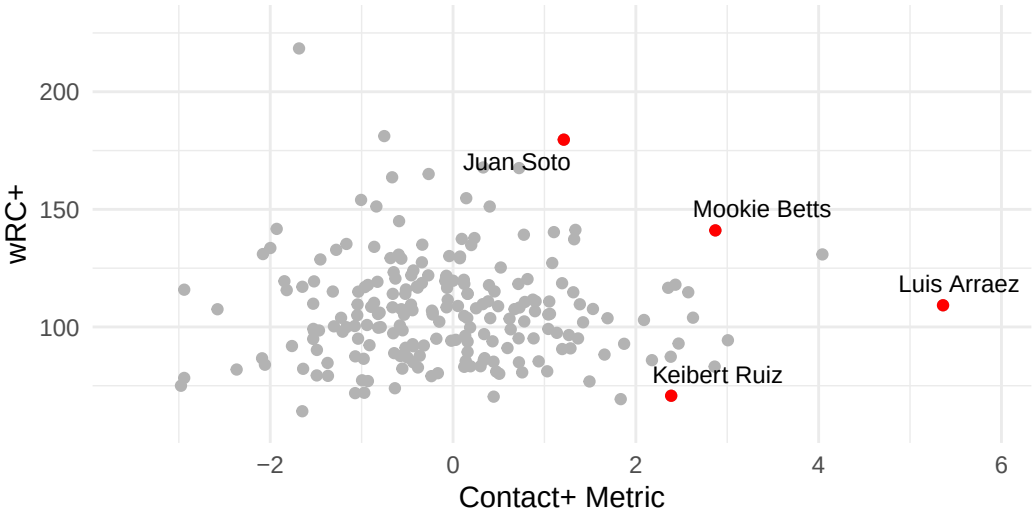

Figure 10: Highlighted Contact+ vs. wRC+

By visualizing these individual players across the Swing+, Power+, and Contact+ metrics, we are able to evaluate how well each metric reflects on-field performance for different types of hitters.

Juan Soto stands out across all visualizations, particularly in Power+ and Swing+, where he ranks among the highest. His wRC+ is also elite, supporting the validity of these metrics as indicators of offensive value. His strong showing across power- and swing-focused metrics confirms the value of these traits in elite hitters.

Mookie Betts is uniquely balanced—he is the only player in our dataset ranked in the top 25 for both Power+ and Contact+. In the visualizations, he appears as a high performer across all three charts. His well-rounded placement reinforces the idea that swing consistency and compact mechanics can enhance the impact of raw power. This helps explain Betts's long-term offensive success.

Luis Arraez performs well in Contact+ but only moderately in Swing+ and Power+. Despite this, he maintains a solid wRC+, illustrating that contact-oriented hitters can still be productive. However, his outlier status also highlights the limitations of Contact+ in predicting broader offensive output. It captures contact efficiency, but not power-driven production.

Keibert Ruiz has below-average values across all three metrics, with especially weak performance in Power+. This aligns with his below-average wRC+ and underwhelming offensive season, making him a useful case for validating the metrics' ability to flag struggling hitters.

Together, these examples show how Swing+ and Power+ are more tightly associated with overall offensive success, while Contact+ captures a narrower dimension of hitting. These metrics can provide decision-makers with actionable insights to evaluate performance, guide development, and identify undervalued players or inefficiencies.

Application

One practical application of our analysis is optimizing batting lineups using the composite swing metrics **Power+**, **Contact+**, and **Swing+**. By quantifying key swing traits through standardized z-scores, these metrics allow managers to assign hitters to lineup positions based on their underlying mechanical strengths rather than relying solely on traditional performance metrics.

For example, hitters with high **Contact+** scores and lower **Power+** values may be ideal candidates for leadoff or second in the order, where the priority is consistent contact and on-base percentage rather than slugging. In contrast, players with elevated **Power+** scores, even if they exhibit moderate strikeout rates, can be positioned in the third, fourth, or fifth spots to maximize their run-producing potential. Hitters with a high **Swing+**, which integrates both power and contact components, may offer lineup flexibility and can be inserted into various roles depending on the opposing pitcher's profile.

Beyond lineup configuration, these metrics provide a data-driven approach for player acquisition and roster construction. Front offices can identify gaps in the current batting order and target players in free agency or trade discussions who rank highly in the swing metric that is underrepresented. For instance, if a team lacks consistent contact producers, **Contact+** scores can guide them toward acquiring undervalued players with high bat-to-ball skills.

Additionally, these metrics can support player development initiatives. Coaching staffs can use **Swing+** as a diagnostic indicator to evaluate whether a hitter's mechanics are optimized for both power generation and contact quality. A player who underperforms in **Power+** may need to adjust swing path or timing to improve barrel efficiency, while a low **Contact+** score could indicate excessive swing length or vulnerability to certain pitch types.

Lastly, the standardized nature of these metrics enhances comparability across player archetypes, seasons, and even minor league levels, making them valuable tools for scouting and long-term player evaluation.

Future Work

Looking into the future more data would be the most helpful to adding to the research. More years of data would allow investigations into player changes from season to season. With that data we could see how players that change their approaches and if it helped or hurt their performance. Also, in the future it would be great to have exit velocity data sorted by pitch count to match the bat speed data capabilities. With that data we would be able to investigate players that have more aggressive approaches (increased bat speed and length in hitter counts the most) to see if that helps their performance overall. We have graphics showing that hitters' overall changes do change but some players change more than others.

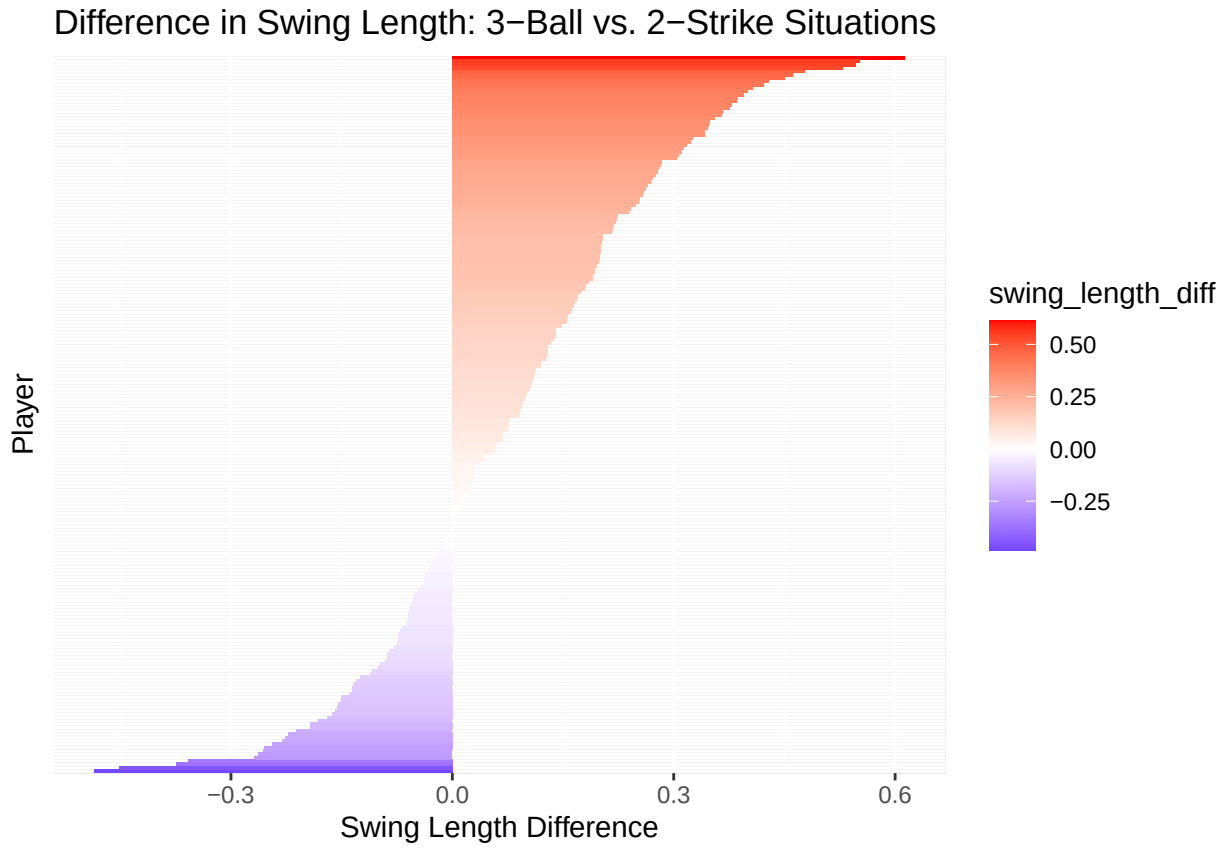

Figure 11: Change in Bat Speed

Further granularity in the data, especially at the pitch-by-pitch level, would substantially enhance the ability to evaluate swing decisions in context. A player's swing is not only influenced by their overall approach but also by situational variables such as the count, pitch type, location, and game state. Being able to directly connect each swing to these variables

would allow for the development of situational swing profiles. For instance, linking metrics like exit velocity, launch angle, and squared-up rate to specific pitch counts would make it possible to measure how effectively players adjust in two-strike counts compared to hitter-friendly counts like 3-0 or 3-1. Our current findings show that players do change their swing mechanics between 3-ball and 2-strike situations, but without knowing the outcomes of those swings, we cannot fully assess whether these changes are successful or consistent. A more detailed dataset could reveal whether players who make more count-dependent adjustments tend to be more productive overall or more efficient in specific game states.

This line of analysis could also extend to organizational evaluation. With access to richer datasets, we could study how front offices build rosters based on player swing characteristics. For example, a team that signs multiple players with high bat speed but poor contact quality might be targeting a high-variance, power-focused approach. Conversely, a team prioritizing players with consistent swing length and low whiff rates may be emphasizing contact and ball control. Over multiple seasons, these strategies could be evaluated against actual team performance to identify whether swing-based roster construction correlates with offensive success.

From a modeling standpoint, there is value in further decomposing the composite swing metrics we introduced. While Swing+, Power+, and Contact+ summarize multiple inputs, future work could assess the marginal contribution of each component, such as bat speed, blast rate, or swing length, to offensive output. This would allow for a better understanding of which swing traits are most predictive of run creation and might inform more optimized weightings in composite scores. It may also help differentiate between players with similar aggregate scores but fundamentally different swing profiles.

Incorporating more advanced situational context, such as inning, score margin, leverage index, or postseason pressure, would also be a valuable extension. These contexts often shape player behavior and may influence swing decisions in ways that season-aggregated metrics cannot detect. Identifying how players alter their swings under pressure, or how certain traits hold up in high-leverage situations, could be useful for both player development and tactical deployment in-game.

One of the current limitations of our analysis is its reliance on wRC+ as a global performance indicator. While wRC+ is useful for summarizing a player's offensive value over a season, it does not capture the contribution of individual swings. Developing or applying a performance metric that can be estimated on a per-swing or per-plate appearance basis would allow

for more precise validation of swing quality and provide immediate feedback for coaching, player evaluation, and in-game decision making. Overall, deeper contextual data and more refined modeling approaches could unlock new insights into the relationship between swing mechanics and offensive value, with applications ranging from individual player development to roster strategy and in-game optimization.

Appendix - Data Dictionary

Variable	Code Name	Description
Player Name	Name	Name of player
Competitive Swing Percentage	percent_swing_competitive	% of swings that qualify as the fastest 90% of a player's swings, plus any 60+ MPH swings resulting in an exit velocity of 90+ MPH
Average Bat Speed	avg_bat_speed	Average bat speed measured at the sweet-spot of the bat
Hard Swing Rate	hard_swing_rate	A fast swing is one that has 75 MPH or more of bat speed.
Squared Up per Swing	squared_up_per_swing	A swing that generates an exit velocity that is at least 80% of the maximum attainable exit velocity based on the speed of the swing and pitch
Blast per Swing	blasts_per_swing	[squared-up percentage] x 100 + [bat speed] >= 164, or a ball that was squared up with a fast speed
Swing Length	swing_length	the sum distance traveled by the head of the bat in XYZ space from the start of data until contact point
Whiffs per Swing	whiff_per_swing	The numbers of pitches swung at and missed by a batter divided by the total number of swings
On Base plus Slugging	OPS	A metric that combines how often a players gets on base (on base percentage) with how much power they have (slugging)
Weighted Runs Created +	wRC.	wRC+ is a comprehensive statistics that that quantifies a player's total offensive value in terms of runs created, while adjusting for external factors like ballpark and era
Weighted on base Percentage	wOBA	wOBA is a version of on-base percentage, where the value for each method of reaching base is determined by how much that event is worth in relation to projected runs scored
Swing+	Swingplus	Swingplus is a created variable that takes z scores for the different metrics and either the positive or negative depending on if it's better for a player to be above or below the mean. Uses the batters' bat speed, swing length, launch angle, squared up rate, and blast rate to measure a relative good or bad swing compared to other players.
Contact+	Contactplus	Contactplus is a created variable that takes z scores for different metrics and either the positive or negative depending on if it's better for a player to be above or below the mean. Uses the batters' launch angle, squared up rate, and whiff rate to measure a relative good or bad swing compared to other players but only using stats that contact hitter should focus on.

(continued)

Variable	Code Name	Description
Power+	Powerplus	Powerplus is a created variable that takes z scores for different metrics and either the positive or negative depending on if it's better for a player to be above or below the mean. Uses the batters' bat speed, launch angle, squared up rate, and blast rate to measure a relative good or bad swing compared to other players but only using stats that power hitter should focus on.

References

"Blast Motion on Bat Speed: Training, Assessing, and Understanding Your Own Swing." *Prepbaseballreport.com*, 2024, https://www.prepbaseballreport.com/news/PBR/blast-motionon-bat-speed--training--assessing--and-understanding-your-own-swing?r=%2Fall-pbrnews. Accessed 1 May 2025.

Driveline Baseball. "Measuring and Improving Baseball Swing Efficiency - Driveline Baseball." *Driveline Baseball*, 14 Apr. 2020, https://www.drivelinebaseball.com/2020/04/measuring-and-improving-baseball-swingefficiency. Accessed 1 May 2025.

Samulski, Eric. "Bat Speed Leaderboard: What Is It and Who Stands Out?" *NBC Sports*, 13 May 2024, https://www.nbcsports.com/fantasy/baseball/news/bat-speed-leaderboard-what-is-it-and-why-do-we-care. Accessed 1 May 2025.

Sarris, Eno. "MLB's New Bat Speed Numbers: The Good, the Bad and the Uncertain Future." *The New York Times*, https://www.nytimes.com/athletic/5477905/2024/05/13/mlb-new-bat-speed-numbers.

"Statcast Bat Tracking." *Baseballsavant.com*, https://baseballsavant.mlb.com/leaderboard/bat-tracking.

"Swing Length | Glossary." *MLB.com*, https://www.mlb.com/glossary/statcast/swing-length.