The Art of Swing: You Are a Slugger?

Wang Ma, Jingjuan Huang, Qiang Hu, Junjie Qiu, Leping Li
STATDS, SUSTech, Shenzhen, China
{12012424,12112847,12111214,12111831,12112627}@mail.sustech.edu.cn

Abstract

The objective of this project is to establish precise metrics that accurately assess a player's batting ability, while systematically eliminating extraneous influences such as luck, defense, and field conditions. This report aims to summarize and critically analyze the key aspects of the newly-proposed batting metrics by Statcast, focusing on the proposed metrics and their implications for determining a batter classification rule.

1 Introduction

Baseball, often referred to as America's pastime, holds a unique position in the world of sports due to its rich history and the intricate statistical analysis it enables. This sport's complexity offers a fertile ground for applying advanced statistical and machine learning techniques, making it an ideal subject for data science projects. The objective of this report is to delve into the intricacies of evaluating batting performance using modern metrics proposed by Major League Baseball's Statcast system.

Context and Motivation. Evaluating a batter's performance in baseball is challenging due to a myriad of influencing factors such as luck, the defensive prowess of fielders, and the characteristics of different ballparks. Traditional metrics like batting average and home runs provide a snapshot but often fail to capture the complete picture of a player's abilities. The need for more precise metrics is paramount for both performance analysis and financial decisions, such as player salaries. This project aims to establish precise metrics that accurately assess a player's batting ability, systematically eliminating extraneous influences such as luck, defense, and field conditions.

Advances in Baseball Analytics. Statcast, a state-of-the-art tracking technology, has revolutionized baseball analytics by providing detailed measurements of player movements and actions on the field. On May 13, Statcast introduced six key metrics to better quantify a batter's abilities:

- **Bat Speed**: The velocity of the bat as it moves through the hitting zone.
- **Fast-swing Rate**: The frequency at which a batter can achieve high bat speed.
- **Squared-up Rate**: The consistency with which a batter makes solid contact with the ball, leading to optimal hitting outcomes.
- **Blast**: A measure of the quality of contact, typically resulting in extra-base hits.
- **Swing Length**: The distance the bat travels during a swing, affecting the timing and control of the hit.
- **Swords**: Instances where a batter is completely fooled by a pitch, often leading to weak contact or misses.

Categorization of Batters. Our project also highlights the categorization of batters based on their skills and performance. Batters are classified into various groups such as Sluggers, Elite Batters, Contact Batters, and Weak Batters. Each category exhibits distinct characteristics:

Figure 1: Four Categories of Batters

- **Sluggers**: Known for their power and ability to hit home runs, but with high strikeout rates and a low batting average (e.g., Giancarlo Stanton, Kyle Schwarber).
- **Elite Batters**: Combine power with high contact rates, making them versatile threats at the plate. Their most notable features are exaggerated expected weighted on-base average (xwOBA) and on-base plus slugging (OPS) (e.g., Mike Trout, Shohei Ohtani).
- **Contact Batters**: Focus on consistently making contact with the ball, often resulting in high batting averages. They are often referred to as "The Hit Makers" (e.g., Ichiro Suzuki).
- **Weak Batters**: Struggle to make effective contact and often lack power, limiting their overall impact. These players are at the margin of MLB and may be demoted to Minor League Baseball (MiLB) or play defensive roles (e.g., fringe MLB players).

Dataset: Statcast In recent years, the analysis of sports data has gained significant traction within the machine learning community. One particularly rich source of data is the **Statcast system**[1], developed by Major League Baseball (MLB). Statcast is a state-of-the-art tracking technology that collects an extensive array of metrics for every play in an MLB game. This system utilizes high-resolution cameras and radar equipment to measure the precise movements and physical characteristics of players and the ball. The resulting dataset provides granular insights into the sport, capturing details such as pitch velocity, player positioning, and batted ball trajectories.

The Statcast dataset, accessible via the **Baseball Savant platform**[2], offers a comprehensive collection of baseball statistics from the 2015 MLB season onward. This dataset encompasses various dimensions of the game, including pitching, batting, and fielding. For instance, it records pitch type, spin rate, exit velocity of batted balls, and the sprint speed of players. These detailed metrics enable nuanced analysis and modeling, facilitating research in areas such as player performance evaluation, game strategy optimization, and injury prevention. By leveraging the Statcast dataset, researchers can apply advanced machine learning techniques to uncover patterns and insights that were previously unattainable with traditional baseball statistics.

Objectives of the Study. As mentioned, result-based metrics can be greatly influenced by uncontrollable factors such as luck and field conditions. Therefore, this study aims to measure a batter's true batting ability purely based on their "swinging," which can be seen as the true batting ability of a player. The objectives of this study include:

- Measuring True Batting Ability: This involves eliminating the influence of external factors such as luck, defensive skill of the opponents, and the peculiarities of different fields.
- Determining Player Salary: By using refined metrics, teams can make more informed financial decisions, ensuring that contracts reflect a player's true value.

Contributions. Our contributions can be summarized as follows:

1. Analysis of Statcast Metrics: We analyze the six newly proposed metrics by Statcast using this year's real player data, uncovering their relationships with traditional metrics.

2. Development of Classification Rules: We create a workable classification rule to categorize different types of batters based on the new metrics.

Through this project, we aim to provide a more comprehensive and accurate assessment of player performance, leveraging advanced metrics and data analysis techniques to enhance the understanding of batting abilities in baseball.

2 Bat Speed and Fast Swing Rate

2.1 Bat Speed

Bat speed, a critical metric in baseball performance analysis, refers to the average speed of a player's bat during the top 90% of their swings. This measurement provides a focused look at a player's most effective and powerful swings, which are crucial for assessing their hitting capabilities. Higher bat speeds are generally associated with better performance, as they enable players to hit the ball harder and farther.

Visualization techniques, such as histograms and kernel density plots, reveal insightful patterns in bat speed distribution. For instance, an analysis of player data shows that average bat speeds typically range from 62 to 80 mph, with a significant concentration around 71 mph. This indicates that most players' swings fall within this range, suggesting a common target for optimal performance. The distribution of average bat speeds approximates a normal distribution, highlighting that extreme values are less common.

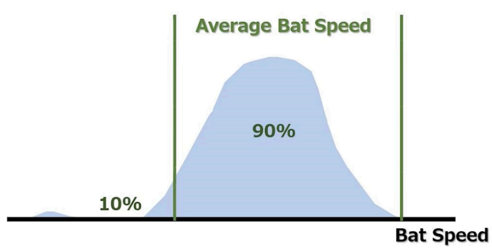

Figure 2: Definition of Average Bat Speed

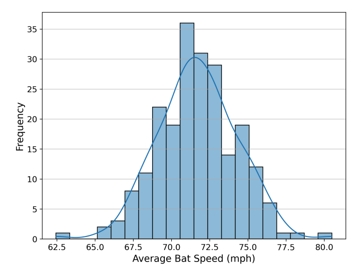

Figure 3: Average Bat Speed Distribution

2.2 Fast Swing Rate

Statcast defines a "fast swing" as one that achieves a swing speed of at least 75 mph. The "fast-swing rate" is calculated as the percentage of a player's total swings that reach this speed. During the first month of 2024, it was observed that 23% of all swings qualified as fast swings. The 75 mph threshold is significant because it marks the point at which per-swing production aligns with the league average. As swing speed increases beyond this threshold, players typically see enhanced performance metrics. This distinction is crucial as the difference between a swing that reaches 75 mph and one that does not is considerable.

From the table we calculated, fast swings are associated with higher batting averages, slugging percentages, weighted on-base averages (wOBA), hard-hit rates, and better overall run production per 100 swings. This highlights the value of increasing swing speed to achieve better offensive outcomes.

Table 1: Fast Swing vs Not a Fast Swing

Metric	Fast Swing	Not a Fast Swing
BA	0.311	0.219
SLG	0.612	0.374
wOBA	0.392	0.267

Metric	Fast Swing	Not a Fast Swing
hard-hit rate	0.535	0.360
RV/100	0.722	−2.868

2.3 Correlation Between Average Bat Speed and Fast Swing Rate

To explore the relationship between average bat speed and fast swing rate, a linear regression model was utilized. The analysis revealed a strong positive correlation between the two variables, with an R-squared value of 0.879. This high R-squared value indicates that a significant portion of the variability in fast swing rate can be explained by average bat speed, underscoring the importance of bat speed in determining swing efficiency.

For instance, Giancarlo Stanton, known for his powerful hitting, exhibits an average bat speed of 80.60 mph and an exceptionally high fast swing rate of 0.969. This combination highlights his ability to consistently generate high swing speeds, contributing to his formidable offensive output. Stanton's high bat speed allows him to make powerful contact with the ball, resulting in higher slugging percentages and overall better performance at the plate.

In contrast, Luis Arraez, who has an average bat speed of 62.50 mph and a fast swing rate of 0, demonstrates a different approach to hitting. Despite his lower power metrics, Arraez's superior hitting technique is reflected in his higher expected wOBA compared to Stanton. This suggests that while bat speed is a critical factor, hitting skill and technique also play vital roles in a player's overall performance.

Consequently, there is a need for metrics like the Squared-Up Rate to provide a more comprehensive assessment of a player's hitting proficiency beyond just power and bat speed. Such metrics can capture the effectiveness of a player's ability to make solid contact with the ball, accounting for factors such as timing, pitch selection, and precision. This holistic approach to evaluating hitting performance can help identify players who excel through technique rather than sheer power, offering a more balanced view of their offensive capabilities.

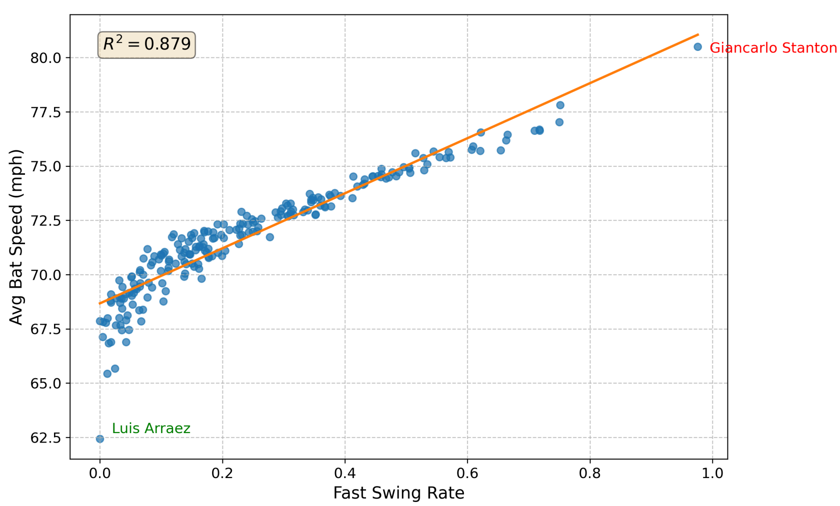

Figure 4: Correlation between Avg Bat Speed and Fast Swing Rate

3 Squared-Up Rate

Firstly, for each ball, we can determine if it is "Squared-Up" by comparing the actual exit velocity to the max exit velocity. The max exit velocity is calculated based on pitch velocity and bat speed. A ball is considered "Squared-Up" if the ratio of actual exit velocity to max exit velocity exceeds 80%. After determining the "Squared-Up" status for each ball, the Squared-Up Rate is calculated by dividing the number of "Squared-Up" instances by the total number of swings.

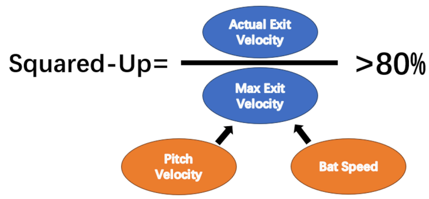

Figure 5: Squared-Up

Generally, the swing speed conversion rate for “Squared-Up” hits is higher than for non-"Squared-Up" hits. From the batter’s perspective, achieving a “Squared-Up” hit requires striking the ball with the sweet spot of the bat, located 4-9 inches from the tip. Therefore, in baseball, players with high contact attributes are better at controlling the bat and hitting the ball with the sweet spot. With consistent pitch speed, the exit velocity of a “Squared-Up” ball can be higher even with a lower swing speed.

$$\begin{array}{c}\text{Squared-Up}\\ \text{Rate (Swing)}\end{array} = \frac{\text{Squared-Up Qty}}{\text{Total Swing Qty}}$$

Figure 6: Squared-Up Rate

The speed of the pitch significantly influences the maximum exit velocity of the hit, a phenomenon commonly referred to as “borrowing power.” This is why many batters prefer to wait for a four-seam fastball to hit home runs.

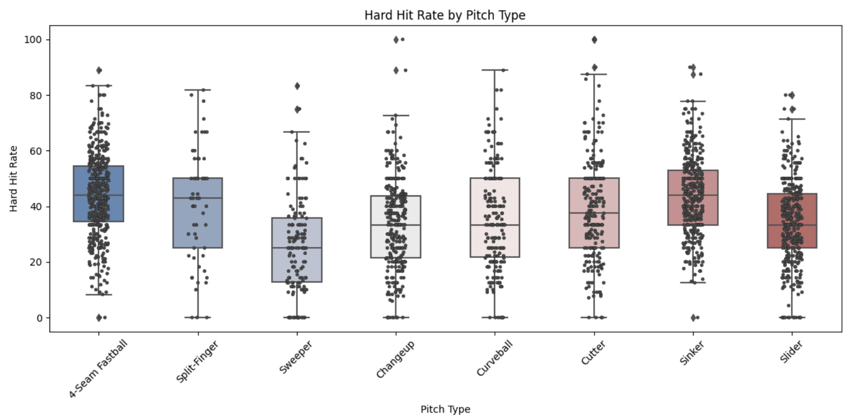

Figure 7: Fast Hit Rate by Pitch Type.

Currently, the average Squared-Up Rate for each swing is around 25%. The hitting outcomes for Squared-Up hits significantly outperform those for non-Squared-Up hits in terms of batting average, slugging percentage, and run production. Despite having less power, Luis Arraez leads the league with a Squared-Up Rate of 43.6%, while Giancarlo Stanton’s rate is only 21.3%. Consequently, Arraez’s expected weighted on-base average (xwOBA) is higher than Stanton’s, illustrating the value of consistent contact and technique over sheer power.

Table 2: Squared-Up vs Not Squared-Up

Metric	Squared-Up	Not Squared-Up
BA	0.3708	0.1261
SLG	0.6602	0.1430
wOBA	0.4382	0.1197
hard-hit rate	0.5976	0.0133
RV/100	11.1873	−6.6557

Table 3: Giancarlo Stanton vs Luis Arraez

Metric	Giancarlo Stanton	Luis Arraez
Bat Speed	***80.6***	***62.5***
Fast Swing Rate	97.2	0.0
xBA	0.250	0.326
xSLG	0.512	0.409
xwOBA	0.342	0.340
Squared-Up Rate	***0.213***	***0.436***

The distribution charts for all Major League Baseball batters with more than 150 swings show interesting trends. In the left chart, the horizontal axis represents the Fast Swing Rate, while in the right chart, the horizontal axis represents the Squared-Up Rate. The vertical axis in both charts represents the expected batting average.

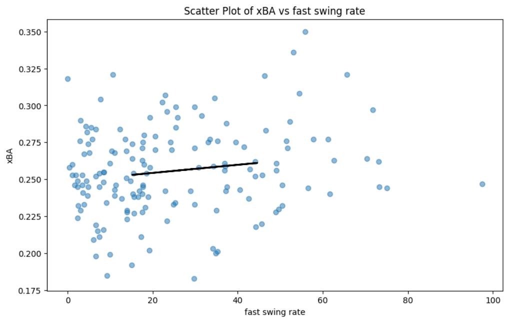

Figure 8: xBA vs. Fast Swing Rate

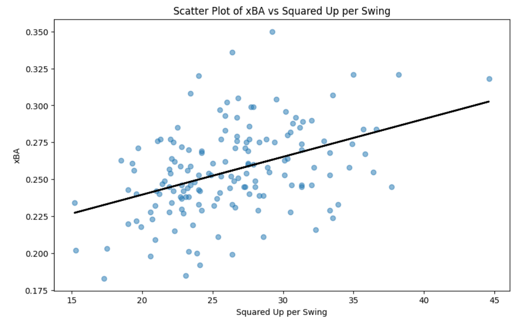

Figure 11: xBA vs. Squared-Up Rate

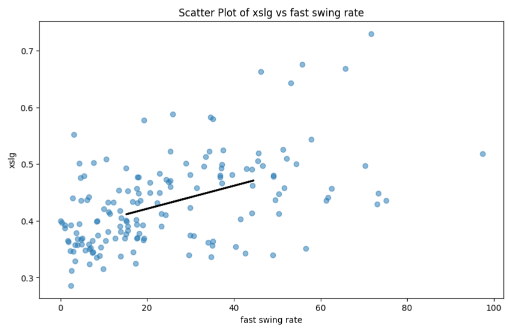

Figure 9: xSLG vs. Fast Swing Rate

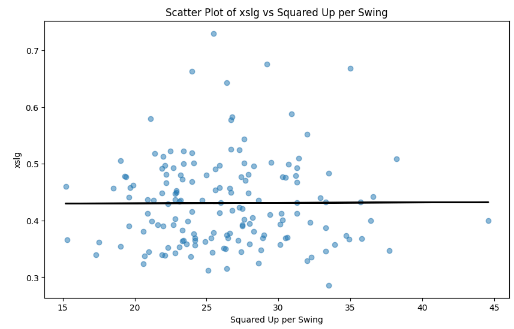

Figure 12: xSLG vs. Squared-Up Rate

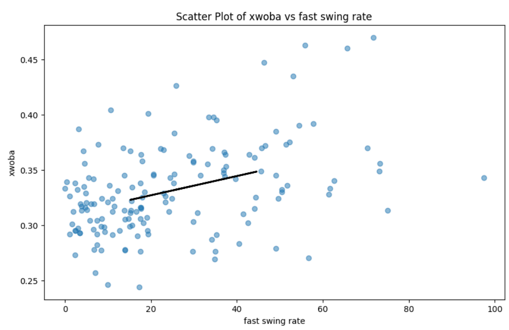

Figure 10: xWOBA vs. Fast Swing Rate

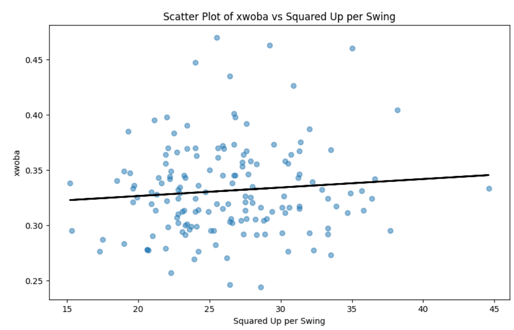

Figure 13: xWOBA vs. Squared-Up Rate

The data indicate that the Squared-Up Rate has a more significant impact on the expected batting average compared to the Fast Swing Rate. Conversely, in terms of slugging percentage, the Fast Swing Rate has a greater influence than the Squared-Up Rate. When considering the final expected weighted on-base average (xwOBA), the impact of both rates is comparable, with the Fast Swing Rate having a slight advantage.

It is important to note that swing speed is not always optimal when maximized. Comparing the swing speed distribution of Giancarlo Stanton and Shohei Ohtani, we observe that Stanton's swing speed distribution is very narrow, whereas Ohtani's is more dispersed. When facing inside pitches, batters need to swing faster to pull the ball, while for outside pitches, they need to delay their swing to push the ball. Stanton's swing speed is almost identical for both inside and outside pitches, resulting in a significantly lower batting average for outside pitches compared to inside pitches. In contrast, Ohtani adjusts his swing speed based on the pitch location, swinging slower for outside pitches. This adjustment allows him to better follow the ball, leading to a higher batting average on outside pitches.

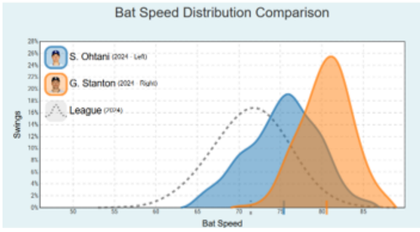

Figure 14: Bat Speed distribution

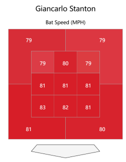

Figure 16: Stanton BS

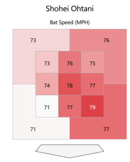

Figure 18: Shohei BS

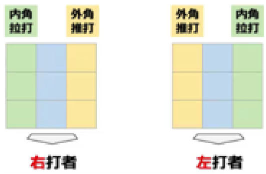

Figure 15: Pull & Push

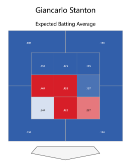

Figure 17: Stanton xBA

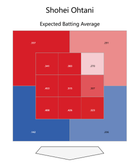

Figure 19: Shohei xBA

By plotting Squared-Up Rate on the horizontal axis and Fast Swing Rate on the vertical axis, we observe an inverse relationship between the two; as swing speed increases, it becomes more challenging to control the point of contact on the bat. Examining the heat maps of expected batting average, expected slugging percentage, and expected wOBA, we conclude that the combination of Squared-Up Rate and Fast Swing Rate can effectively define the four types of batters mentioned earlier.

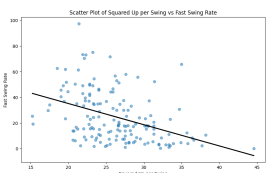

Figure 20: Fast swing

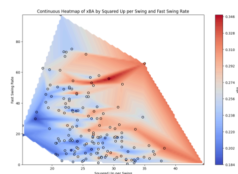

Figure 21: xBA

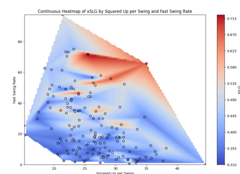

Figure 22: xSLG

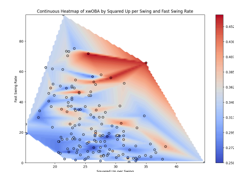

Figure 23: xwOBA

As of June 16th, which players fall into these quadrants on the distribution chart? The top-right corner, with three players far ahead of the pack, features the Yankees' Juan Soto, Brewers' catcher William Contreras, and Shohei Ohtani. The bottom-right quadrant, representing skillful hitters, includes Luis Arraez and Mookie Betts. The top-left quadrant, known for power hitters with less control, includes notable sluggers such as Giancarlo Stanton, the Phillies' Kyle Schwarber, and Aaron Judge.

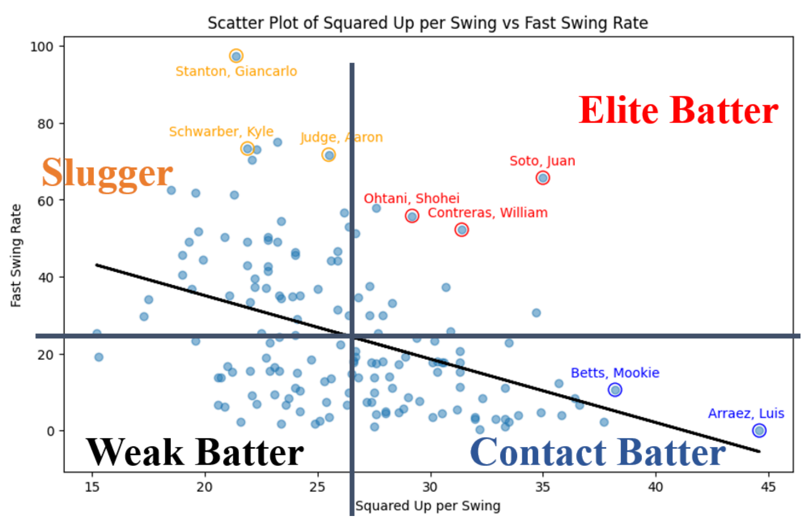

Figure 24: Four kinds of batters.

4 Blast and Swing Length

In the realm of baseball analytics, two key metrics have emerged as critical indicators of a player's hitting performance: "blast" and "swing length." A "blast" is defined as a hit where the combined value of 100 times the squared-up percentage and the bat speed equals or exceeds 164 mph. "Swing length," on the other hand, measures the distance the bat's sweet spot travels during a hit. These metrics provide insight into a player's power and technique, respectively, and have significant implications for their overall effectiveness at the plate.

4.1 Blast

According to recent data, the distribution of blast rates among batters is centered around 10%, with only a select few players exceeding this mark. Notably, Juan Soto, William Contreras, and Shohei Ohtani are the only batters to achieve a blast rate over 20%. This highlights their exceptional ability to generate high-speed, well-squared hits consistently. Heatmap visualizations comparing bat speed and exit velocity reveal that blasts consistently result in higher exit velocities than non-blasts, despite similar bat speeds. This indicates that the quality of contact, rather than just bat speed, plays a crucial role in generating blasts.

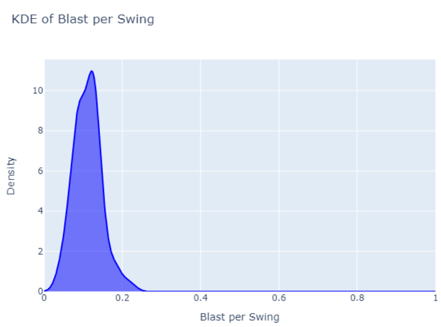

Figure 25: Distribution of Blast Rates of Batters

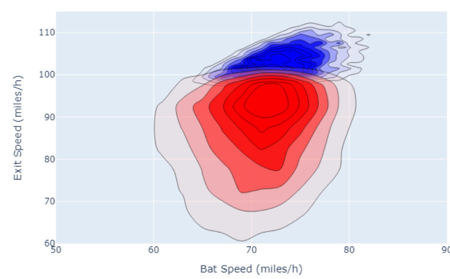

Figure 26: Distribution between Bat Speed and Exit Speed

The statistical differences between blasts and non-blasts are stark. Blasts result in a weighted On-Base Average (wOBA) of 0.705 compared to 0.177 for non-blasts, a batting average (BA) of 0.547 versus 0.174, and a slugging percentage (SLG) of 1.120 against 0.221. Furthermore, the expected run value for blasts is a staggering 31.720, while non-blasts carry a negative value of −5.820. These figures underscore the immense impact blasts have on a player's offensive contribution and the overall game dynamics.

4.2 Swing Length

Swing length also shows significant positive correlations with various performance metrics. Distributions of swing length against wOBA, BA, SLG, and expected run value all demonstrate that longer swing lengths tend to correlate with better performance outcomes. Additionally, there is a notable positive correlation between swing length and bat speed, indicating that players with longer swings often generate more bat speed.

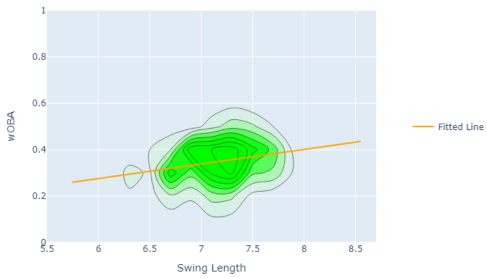

Figure 27: Distribution of Swing Length and wOBA

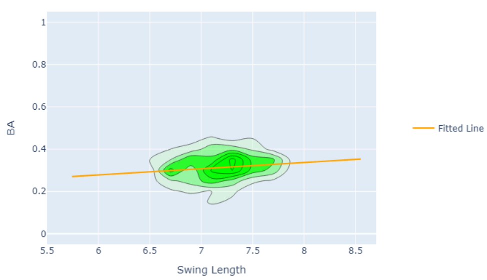

Figure 28: Distribution of Swing Length and BA

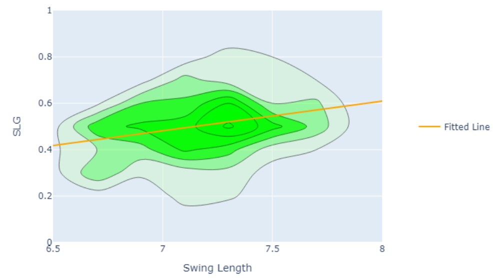

Figure 29: Distribution of Swing Length and SLG

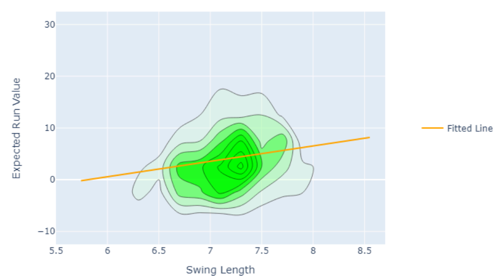

Figure 30: Distribution of Swing Length and Expected Run Value

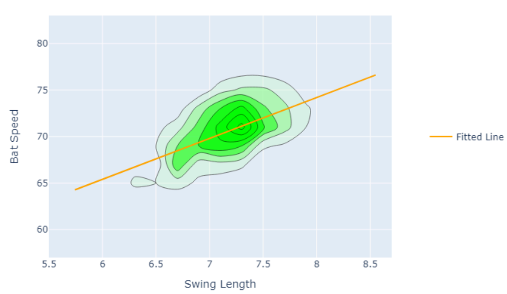

Figure 31: Distribution between Swing Length and Bat Speed

A closer examination of two players with contrasting swing lengths—Giancarlo Stanton and Luis Arraez—provides further insights. Stanton, with a swing length of 8.4 inches and a bat speed of 80.6 mph, has an expected batting average (xBA) of 0.250, a SLG of 0.512, and a hard hit percentage of 53.4%. In contrast, Arraez, with a shorter swing length of 5.9 inches and a bat speed of 62.5 mph, boasts an xBA of 0.326, a SLG of 0.409, and a hard hit percentage of 21.4%. Additionally, Stanton has a higher strikeout rate (K%) of 30.8% compared to Arraez's 5.7%. This comparison illustrates how swing length and bat speed can influence different aspects of a player's hitting profile, from power to contact quality.

5 Sword: *It Stopped Like a Sword*

5.1 Introduction of Sword

The MLB official describes **sword** as *when a pitcher fools a hitter so badly that he forces a non-competitive swing, one where a batter either regrets his choice or can't stop himself from taking a hack that looks so ugly it ends up going viral on social media* [3]. Please pay attention that a sword is made by the hitter but induced by the way a pitcher pitches.

A "sword" refers to a swing that resembles the action shown in Figure 32. Our goal is to gain insights into the factors that contribute to a sword. This analysis can provide pitchers with statistical evidence to better deceive hitters and achieve better performance.

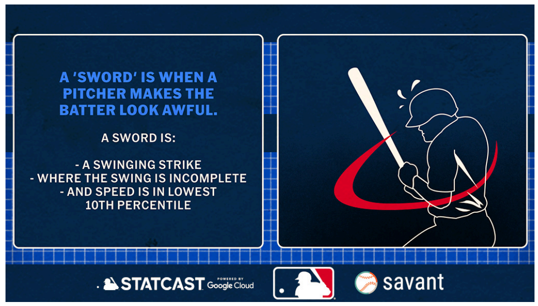

Figure 32: Visual Description of Sword by MLB.

5.2 Pitch Type

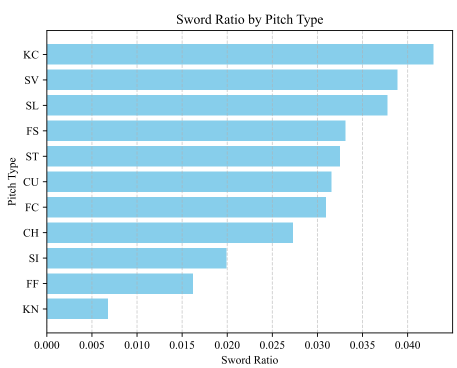

Figure 33: Sword Ratio by Pitch Type. The higher the better pitch is easy to fool the hitter. Table 4 Shows what those abbreviation stands for.

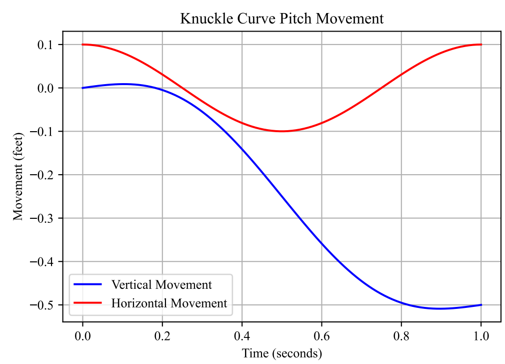

Figure 34: Knuckle Curve. The curve shows how a knuckle pitch moves with respect to time.

From Figure 33 we observe that a Knuckleball (Pitch Type = KN) is the most possible kind of pitch that can induce a hitter to react with sword. Figure 34 shows how a Knuckleball moves over time, plotted on a graph with the x-axis representing time (in seconds) and the y-axis representing movement (in feet).

The blue curve represents the vertical movement of the pitch. It starts slightly above zero and decreases steadily, reaching a minimum of about −0.5 feet at around 0.6 seconds, before leveling off. The red curve represents the horizontal movement of the pitch. It starts at about 0.1 feet, decreases to zero at around 0.3 seconds, continues to drop to a minimum of about −0.2 feet at around 0.6 seconds, and then rises back up to around 0.1 feet by the end of the time interval. These movement patterns are characteristic of a well-thrown knuckle curve, making it an effective pitch due to its late and sharp downward break along with a slight horizontal shift.

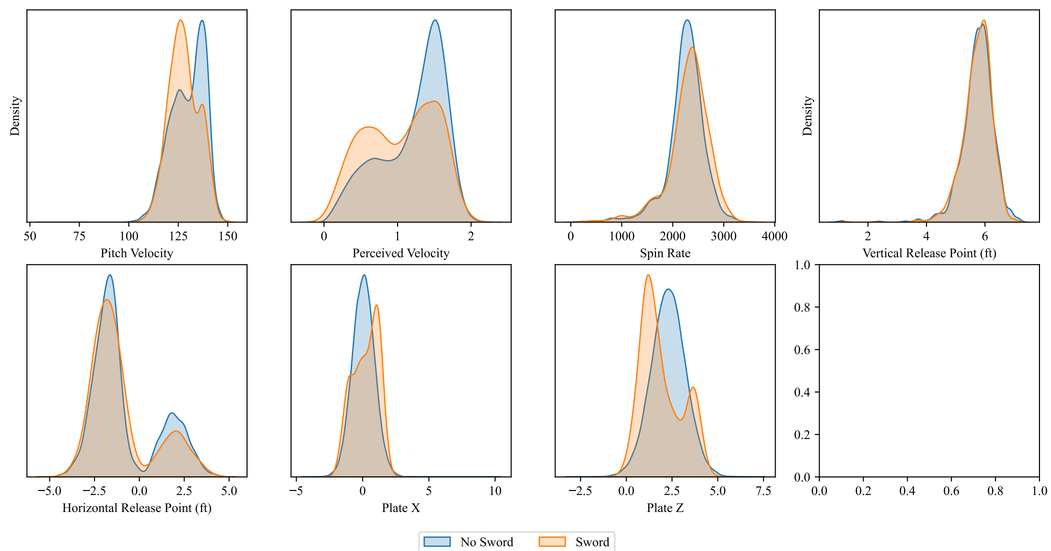

Figure 35: Kernel Distribution Plot of Pitch Attributions.

5.3 Summary

We collected MLB data from April 3, 2024, to May 31, 2024, totaling 93,505 pitching instances. We analyzed the distribution of various attributes by dividing the data into two groups: those pitches that induced a "sword" (henceforth referred to as "Sword" pitches) and those that did not (referred to as "No Sword" pitches). Here's an analysis of each subplot:

1. **Pitch Velocity**: "Sword" pitches tend to have higher velocities than "No Sword" pitches. The distribution for "Sword" peaks around 125-135 mph, while "No Sword" has a broader distribution with a peak slightly lower.
2. **Perceived Velocity**: "Sword" pitches show a higher perceived velocity distribution compared to "No Sword". Both categories have peaks around similar perceived velocities, but "Sword" has a noticeable secondary peak indicating variability.
3. **Spin Rate**: "No Sword" pitches generally have a higher spin rate, peaking around 2500-3000 RPM. "Sword" pitches have a broader distribution with a slightly lower peak.
4. **Vertical Release Point (ft)**: Both categories have similar vertical release points, with peaks around 6 feet. The distributions for "Sword" and "No Sword" are almost identical, indicating no significant difference in vertical release height.
5. **Horizontal Release Point (ft)** : "No Sword" pitches have two distinct peaks, one around −2.5 feet and another near 2.5 feet. "Sword" pitches show a more centered distribution with a peak around 0 feet, indicating less variability in horizontal release points.
6. **Plate X**: No difference.
7. **Plate Z**: "No Sword" pitches have a broader distribution with peaks around 0 and 2.5 feet. "S-word" pitches are more concentrated around 0 and 2.5 feet, indicating tighter control over the vertical position at the plate.

In conclusion, "Sword" pitches are **generally faster** with slightly higher perceived velocities but have **lower spin rates** compared to "No Sword" pitches. Both categories have similar vertical release points, but "Sword" pitches exhibit **more consistency in horizontal** release points. "Sword" pitches are **more precisely located vertically** at the plate compared to "No Sword" pitches.

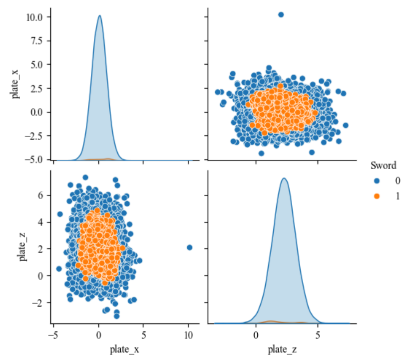

Figure 36: Perceived Final Position. This position is the perceived position of the last observation by the observer.

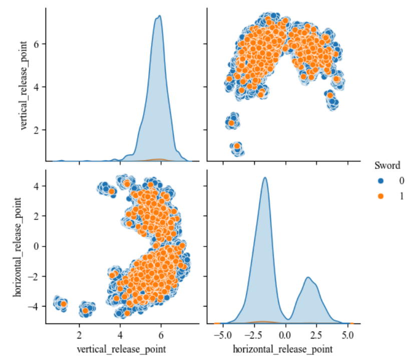

Figure 37: Release Position. This position is perceived when the ball is released from the pitcher's hands.

6 Conclusion

6.1 Categorization of Batters

After conducting a detailed analysis of the newly introduced six batting metrics by Statcast, we return to the initial question: how to classify batters. Reflecting on our findings during the investigation of the squared-up rate, we observed that the scatter plot of squared-up rate versus fast swing rate effectively characterizes different types of batters.

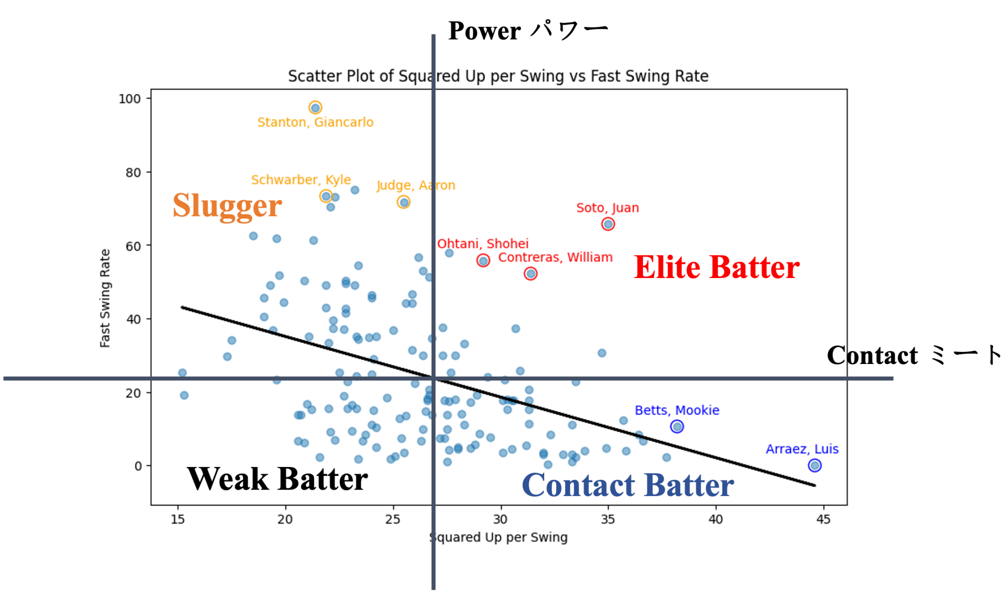

Figure 38: Our Classification based on Squared-up rate Vs Fast Swing rate

6.2 Key Findings

In this project, we have undertaken a rigorous analysis of contemporary baseball metrics to better understand and evaluate a player's batting ability. By focusing on metrics introduced by Statcast, such as bat speed, fast swing rate, and squared-up rate, we have demonstrated how these advanced measurements can provide a clearer picture of a batter's true performance, beyond traditional statistics that often fail to account for external factors like luck and defensive skills.

- **Bat Speed and Fast Swing Rate**:
  - **Bat Speed**: Our analysis confirmed that higher bat speeds are generally associated with improved offensive outcomes, including higher batting averages and slugging percentages. The distribution of bat speeds across players suggested a normal distribution, with most players' average bat speeds clustering around 71 mph.
  - **Fast Swing Rate**: We identified a significant correlation between bat speed and fast swing rate, with a linear regression model revealing that nearly 88% of the variability in fast swing rate can be explained by average bat speed. This underscores the critical role of bat speed in achieving higher swing efficiency and better offensive performance.
- **Squared-Up Rate**: The squared-up rate, defined by the ratio of actual exit velocity to maximum exit velocity, emerged as a critical metric for evaluating a player's ability to make solid contact with the ball. Players with higher squared-up rates, like Luis Arraez, demonstrated superior hitting outcomes despite lower power metrics, highlighting the importance of contact quality and technique.
- **Categorization of Batters**:

  We categorized batters into four distinct groups: Sluggers, Elite Batters, Contact Batters, and Weak Batters. Each category exhibited unique characteristics, with Sluggers known for power but high strikeout rates, and Contact Batters excelling in consistency and technique.
- **Impact of Metrics on Player Evaluation**:

  Our study showed that advanced metrics like fast swing rate and squared-up rate provide a more nuanced and accurate assessment of a player's true batting ability. These metrics help to eliminate the noise introduced by external factors, thus offering a more reliable basis for performance evaluation and financial decision-making, such as player salaries.

6.3 Future Directions

Our project lays the groundwork for several future research directions:

- Integration with Machine Learning: Further research could involve the integration of these metrics with machine learning models to predict player performance and career trajectories more accurately.
- In-depth Analysis of Defensive Metrics: Complementing our focus on batting metrics with an in-depth analysis of defensive performance could provide a holistic view of player value.
- Application to Minor Leagues: Extending these metrics to evaluate minor league players could help in identifying talents and predicting their success at the major league level.

6.4 Final Thoughts

In conclusion, this project has highlighted the importance of using advanced, precise metrics to evaluate a player's true batting ability. By moving beyond traditional statistics and incorporating sophisticated measurements from Statcast, we can achieve a more accurate and comprehensive understanding of what makes a successful batter. This approach not only enhances the analytical capabilities within baseball but also contributes to the broader field of sports analytics, offering valuable insights that can be applied to other sports and domains.

7 Appendix

Table 4: Reference of Pitch Type and Corresponding Abbreviation.

Abbrev	Pitch Type
CH	Changeup

Abbrev	Pitch Type
FF	Four-Seam Fastball
SL	Slider
CU	Curveball
KN	Knuckleball
SV	Slurve
FC	Cutter
KC	Knucle-curve
FS	SPlitter
EP	Eephus
SC	Screwball
ST	Sweeper
FO	Forkball
SI	Sinker

Reference

[1] M. L. Baseball, “Statcast.” [Online]. Available: https://www.mlb.com/statcast

[2] M. L. Baseball, “MLB Savant.” [Online]. Available: https://baseballsavant.mlb.com/

[3] “Swords | Glossary | MLB.com.” [Online]. Available: https://www.mlb.com/glossary/statcast/sword