Exploring Bat Speed: Analyzing Changes in Velocity Between a Baseball Hitter's First Movement and Foot Plant

By Eli Aronson

Denison University

December 13th, 2024

Abstract

Bat speed plays a pivotal role in baseball hitting performance. Using motion capture data from Driveline Baseball, this study investigates kinematic interactions between body segments—pelvis, torso, shoulder and hand velocities—during the period between a hitter's first movement and foot plant. In doing so, the goal was to better understand how these changes in velocities, amongst other factors contribute to bat speed. Four key findings emerged: (1) Changes in velocity at different body parts between first movement and foot plant are statistically significant both individually and in ranking order with other extremities. (2) Hand velocity should produce the greatest change in velocity during this period to maximize bat speed, with a proximal-distal change in velocity sequencing. (3) Body mass significantly influences bat speed, but is not the only factor. (4) Handedness has minimal impact on biomechanics. A mixed-effects modeling approach was used to account for intra-player variability, and MLB Statcast data helped generalize findings. Although handedness is not associated with any significant changes in biomechanics or bat speed, the research found that athletes with heavier body masses exhibit optimal biomechanical sequencing more often, and achieve greater bat speeds. With access to additional data, future research could validate these findings and investigate causality.

Introduction

Bat speed plays a critical part in baseball hitting performance, serving as a key factor in determining hitting outcomes such as exit velocity or distance traveled, which can lead to in-game success (Fortenbaugh, 2011). Traditionally, the baseball swing has been viewed as an art form, but now it has evolved into a science broken down by biomechanical movements. Biomechanics focuses on understanding the physical principles in the movement, which provides athletes and coaches opportunities to optimize their swing through data-driven analysis.

With the integration of motion capture systems, force plates and bat sensors, coaches and biomechanists have found ways to quantify athletic movement to optimize an individual's swing. The introduction of motion capture allows for deeper analysis of measurable components such as movement sequencing, accelerations, and joint angles throughout the swing. At the forefront of private player development lies Driveline Baseball, a state-of-the-art facility leading data-driven individual baseball instruction, working with athletes ranging from high school to Major League Baseball (MLB) players.

Baseball biomechanics is an evolving field and includes a wider scope than just Driveline; there is new research and advancements frequently. However, gaps remain in understanding how specific factors—such as handedness or variations in swing sequencing—affect accelerations, joint angles, and performance metrics such as bat speed. Existing studies on swing mechanics are insightful, but often lack concrete findings regarding biomechanics, as the field is widely theoretical and there is no “one size fits all” approach to the movement. Yet, the integration of machine learning continues to solidify findings in hopes of optimizing player performance and revolutionizing player development.

The overarching goal of this research is to highlight biomechanical interactions that influence bat speed, focusing on the kinematic sequencing between a hitter's first movement and their foot plant. Using data provided by Driveline Baseball, this study analyzed pelvis, torso, shoulder and hand changes in velocity over the period between first movement and foot plant, exploring their relationship with bat speed. While acceleration is often used to evaluate kinematics, the absence of timestamps in the data leaves this research focusing on changes in velocity during this period instead. To account for intra-player variability, the research employed a mixed-effects modeling approach.

The findings unveil that athletes with greater body mass exhibit optimal biomechanical sequencing more often and achieve greater bat speeds. Optimal sequencing, rooted in previous literature, was defined by the kinematic pattern in which distal body segments (further away from the center of mass) accelerate more rapidly than the proximal segments. Although handedness did not significantly influence biomechanics or bat speed, the research suggests the critical role of movement efficiency in generating bat speed.

By addressing the guiding research question of how biomechanical interactions influence bat speed, this study provides insights for athletes or coaches aiming to improve swing mechanics. While the role of handedness held no significant findings, the importance of player mass helps contribute to a deeper understanding in generating bat speed through efficient movements. This study offers valuable insights, however, further research is needed to explore additional factors that may also contribute to bat speed.

Literature Review

Swing Sequencing. The biomechanics of a baseball swing breaks down the swing into seven key phases: waiting, shifting, stepping, landing, swing, impact and follow through (Nakata et al., 2013). These segments outline the movement patterns that a hitter goes through in order to effectively hit the ball. Researchers like Nakata et al. (2013) found that experienced hitters display more electrical activity throughout all phases on the swing, signifying more efficient muscle engagement. Specifically, novice hitters tend to start the movements in their swing much later, which makes it harder to coordinate the sequence of muscle activity necessary to run through the kinetic chain and transfer energy. According to additional studies from Stewart et al. (2020), advanced hitters sequence the stages of hitting in the order that Nakata and previous scholars outlined. The earlier a hitter begins the shifting phase, the more time he has to develop power in the swing that can be transferred throughout the kinetic chain.

Considering both the waiting and shifting phases, the lower body muscle fibers contract in order to generate force against the ground. Creating a ground reaction force allows the hitter to generate power in the shifting phase, aligning with research of lower body muscle recruitment preceding upper body muscle recruitment (Reyes et al., 2011). Before the stride foot plants, the hands and the bat must undergo a pulling action, which generates a resultant force in the swing since the foot moves forward (Crisco et al., 2018). The next two phases, landing and swinging, allows the rotation of the pelvis to transfer much of the ground reaction force throughout the body and direct the bat to the ball (Fortenbaugh, 2011). With greater amount of rotational force into the ball, the velocity at which the ball meets the bat in the impact stage will increase. A greater bat speed can enable the possibility of hitting the ball both harder and farther (Haruna et al., 2023). Sequencing pattern is critical for hitters, and the more experienced and successful batters tend to prepare movement earlier, and fire muscles at appropriate times, allowing for better success hitting the ball.

The timing between segmentations of various body parts peak velocities in the swing is crucial. Fortenbaugh (2011) emphasizes that optimal power generation occurs when the trunk lags behind the rotating pelvis. The intentional segmentation, seen within milliseconds apart, creates a stretch in the torso, leveraging an elastic movement across the upper body to release energy explosively. The trunk separates lower body movements from the bat moving forward, which requires stability of the core.

Rotational sequencing has been scholarly researched across golf. The research of Callaway et al. (2012) highlights the importance of generating high peak pelvis rotational velocities during their swing. Golfers in the study who scored better produced significantly higher pelvis velocities, which the research attributed to stronger gluteus medius and maximus muscles, suggesting the importance of raw strength in rotational velocity (Callaway et al., 2012). When considering other body parts segmented velocities, the maximum velocities should be presented from the body parts more distal, such as the hand (Bourgain et al., 2022; Neal et al., 2007). While the proximal to distal activation sequence to maximize velocity generates the greatest clubhead speed at ball strike, this kinematic principle could also be considered optimal for other rotational sports such as the baseball swing (Bourgain et al., 2022). The peak velocities continue to increase by segmentation in the golf swing, as the proximal body parts closer to the core of the body reach their peak speeds first, decelerate first, and have smaller velocities compared to the distal body parts (Neal et al., 2007). Neal et al. (2007) also finds no significance in the lag time between peak velocities for each segment with regards to well time and mistimed ball strikes in golf. However, because their research focuses on golf where the ball is placed on a tee, that may not carry over to the baseball swing.

Body Mass and Rotational Velocity. Rotational velocity remains a critical velocity in generating bat speed throughout the baseball swing. While baseball players may employ a different sequencing and lag time between peak velocities, the ability to create efficient rotations has a variety of different factors (Fortenbaugh, 2011; Haruna et al., 2023; Szymanski et al., 2009). Specifically, the greatest bat speeds come from the athletes who are strongest and most powerful, while maintaining a lean body mass (Syzmanski et al., 2009). Improvements through resistance training that increase strength and power oftentimes result in a leaner body mass (Syzmanski et al., 2009). Additional research from Symanski et al. (2010) explains how players with greater lean body mass typically have longer levers, which can generate greater segmented velocities and generate more force at ball impact. Specifically for high school baseball players, lean body mass was highly correlated with body mass, suggesting at that age greater body mass can create greater rotational velocities (Symanski et al., 2010). The research of Haruna et al. (2023) grouped bat speed velocities into three groups, where their study found the greatest bat speed group presented significantly higher body weights and lean body masses.

Handedness. In baseball, there is a strategy on how to utilize both left- and right-handed hitters and it does not relate to biomechanics. As Chance (2023) acknowledges in his research, many MLB teams try to utilize "platoon advantage," meaning that teams attempt to match up their hitters to be the opposite handedness as the pitcher. Most hitters find more success off the opposite handed pitcher because they see the ball longer, and most pitchers are worse against the opposite handed hitter (Chance, 2023). Some batters are capable of "switch hitting" where they hit both left- and right-handed, and some players may throw with one hand (their dominant hand), but hit from the other side of the plate. Hitters may learn to bat from the left side as a kid, because the majority of pitchers, just as the population is, are right-hand dominant, and coaches from a youth level believe in the platoon advantage.

The research of Grondin et al. (1999) suggests that those who hit from the left side but are right hand dominant, may have a weaker swing. The authors use an analogy to tennis, where for someone who is right hand dominant, swinging from the left side is like a tennis backhand, a weaker but more reliable and precise shot. Alternatively, players who are left-handed and hit from the left side, their swing is more like a tennis forehand, allowing both an increase in peak velocity, but also more susceptible to miss the ball (Chance, 2023; Grondin et al., 1999). The Grondin et al. (1999) research returns to the Bourgain's et al (2022) sequencing of proximal to distal, where hitters who hit from their dominant side hold their dominant hand as the distal (the top) hand. This allows for more control in the final rotation of the bat, allowing for a greater whip effect, which increases the bat speed. Grondin et al. (1999) found that the proximal hand leads the way, and transfers energy to the distal hand as the bat accelerates in the swing, creating increased bat speed for those who hit from their dominant side.

However, many hitters still utilize the platoon benefit of hitting from the left side, even if that is not the athlete's dominant hand side, suggesting the difference in bat speed is not important enough to impact which side a hitter swings from. Because of the platoon benefit, lefties are overrepresented in the MLB compared to the general population. In 2012, Chu et al. (2016) calculated that 39.4% of batters and 28.4% of pitchers were left-handed. The fighting hypothesis—a theory that left-handed fighters have more success due to the unfamiliarity of left-handed punches to the opposition—also impacts MLB hitters. Regardless of batting side, MLB hitters generally succeed better against right-handed pitchers than their lefty counterparts (Chu et al., 2016). It could be because the platoon advantage left-handed hitters have over them, but lefties still represent under half of the league. Chu et al. (2016) considers the broader familiarity with facing right-handed pitchers for greater success, a variation of the fighting hypothesis. Hitters are more accustomed to facing right-handers on the mound, while right-handed pitchers may also suffer from the fighting hypothesis facing hitters from the left side, as they are less familiar with pitching to them (Chu et al., 2016). The researchers conclude that more of the on-base-slugging (OPS)—a common metric used to measure a hitter's ability to hit for both average and power—results can be explained by the "platoon advantage." Chu et al. (2016) calculates the advantage explains about 15% of a left-handed hitter's OPS, while it explains only 7% for right-handed hitters, suggesting that there should be more left-handed hitters playing in the MLB. Research regarding the biomechanical and anatomical differences of hitters by handedness remains unpublished, however any studies considering the platoon advantage or nuanced biomechanical findings should consider potential differences.

Ethical Considerations

All data including biomechanics comes from the Driveline OpenBiomechanics Project (Wasserberger et al., 2022). This research abides by the Creative Commons Attribution-NonCommerical-ShareAlike 4.0 International License and is used strictly for academic and research purposes. All data is used exclusively with this intent, and all findings will be made publicly available in hopes of furthering research in the field. To ensure integrity and accuracy of the data, any adjustments will be documented in the code to maintain transparency. No manipulation will occur to achieve results to fit a specific narrative or hypothesis.

Additionally, careful consideration has been given to ethical issues surrounding the participant anonymity of the Driveline data. There is no attempt to discover identity using performance, and all of the data is handled with respect for all athletes involved. Lastly, I acknowledge the contributions and tireless work people have contributed to creating both of these invaluable datasets which help serve the study and research. Their work serves as the foundation for a lot of upcoming research in the baseball biomechanics field, and we are fortunate for their contributions to the sports science field.

Data and Methods

Throughout the research, two data sources and a variety of methodological approaches were taken to examine how biomechanical differences may affect an athlete's bat speed. Bat speed is the measurement of the speed that the bat travels through the hitting zone as it approaches hitting the ball, and it plays a pivotal role in this research.

Data

The purpose of using two datasets allows for analyzing hitting on a micro and macro scale. Consider the Driveline biomechanics data as the micro, where there is a smaller sample size, but a plethora of important data that describes movements throughout the swing. On the other hand, the MLB data acts as the macro dataset, where there is a much larger sample, however the data does not pertain to swing mechanics. Rather, this data focuses on performance metrics.

Driveline. The biomechanical analysis comes from the Driveline OpenBiomechanics Project (Wasserberger, 2022). This dataset holds 98 sessions of unique hitters motion capture files, allowing for in-depth analysis and assessment of their swings. Driveline is amongst the leaders in data-driven baseball player development, and by publicizing their data they aim to increase research in the sports biomechanics field. The data collection used both marked data tracking, where the athlete wore markers on his body to measure angles and velocities, and markerless, where cameras analyze each movement at 300 frames per second. The cameras measure movement in a coordinate system as seen in *Figure 1* (Wasserberger et al., 2022).

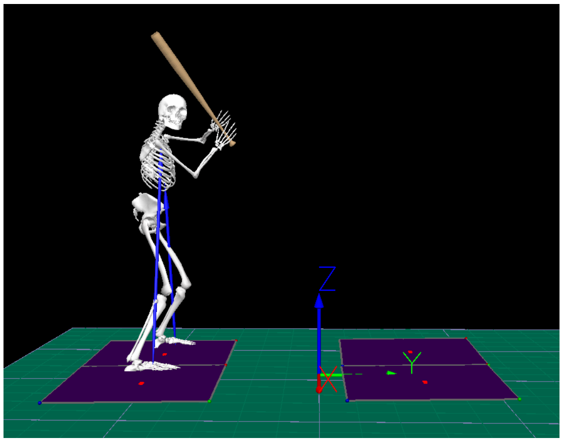


Figure 1: *The coordinate system provided by Driveline to describe their lab setup and label variables. The X axis represents the horizontal direction towards the pitcher's mound. The Y axis is the horizontal direction that is the width of the batter's box. The Z axis acts as a vertical line through the midsagittal plane of the body.*

Driveline's data comprises mainly collegiate baseball players, however there are eleven high school players and ten professionals. While containing an excess of data, biomechanical data often struggles to find concrete conclusions because every human moves differently (Boddy, 2012; Hardgrove, 2013). This research focuses on the changes in velocity from first movement to the foot plant of four main body parts: the pelvis, torso, shoulder and hand. However, even with musculoskeletal differences that will cause no athlete to swing the same, hitters use different techniques to hit the ball hard. One proven influence on striking the ball is the effect of bat speed, as seen from MLB data in *Figure 2*.

MLB Statcast. For the first time in 2024, every MLB hitter had their in-game bat speeds measured and shared publicly on Baseball Savant. The additional data analysis surfaces from Baseball Savant, powered by Google Cloud, and it calculates performance metrics in every game. This research focuses on their calculated expected performance outputs, and how those

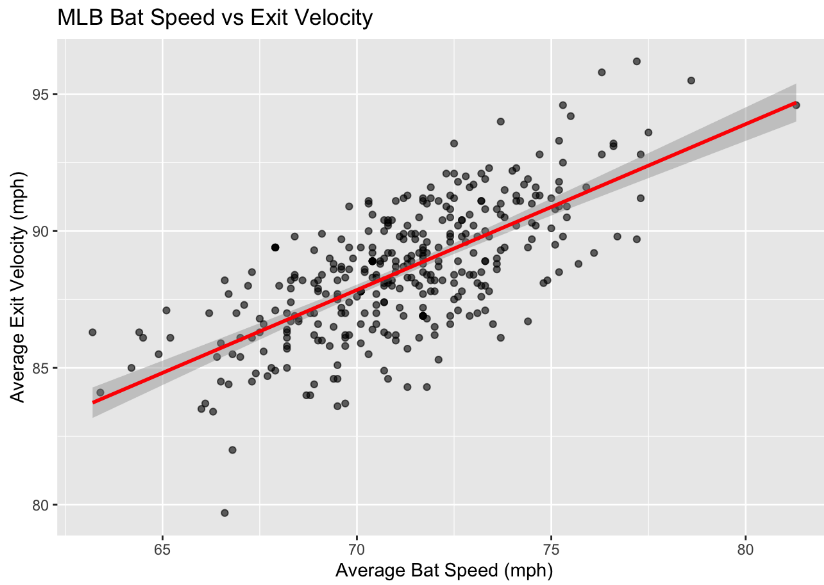


Figure 2: A positive relationship between MLB average bat speed and average exit velocity

influence their bat speed. This dataset contains MLB players who had at least 200 plate appearances in 2024. Further data wrangling used the R Lahman package to add players weights into the dataset to calculate optimal bat speeds given a players weight.

Methods

The biomechanical data analysis employs multinomial logistic regression, binary logistic regression, linear mixed-effects modeling, and ordinal regression to examine factors impacting bat speed. Each model aims to build on the previous finding. A variety of modeling techniques were utilized in order to best fit the variables included in the analysis; however, there are difficulties validating model strength with biomechanics data due to the high variability between individuals and the complex nature of biomechanics that does not meet statistical assumptions. Not every model discussed within the methodology section led to results worth further discussion. Meanwhile, the MLB data uses just two techniques: linear regression and polynomial regression to explore how weight and hitting metrics interact with bat speed.

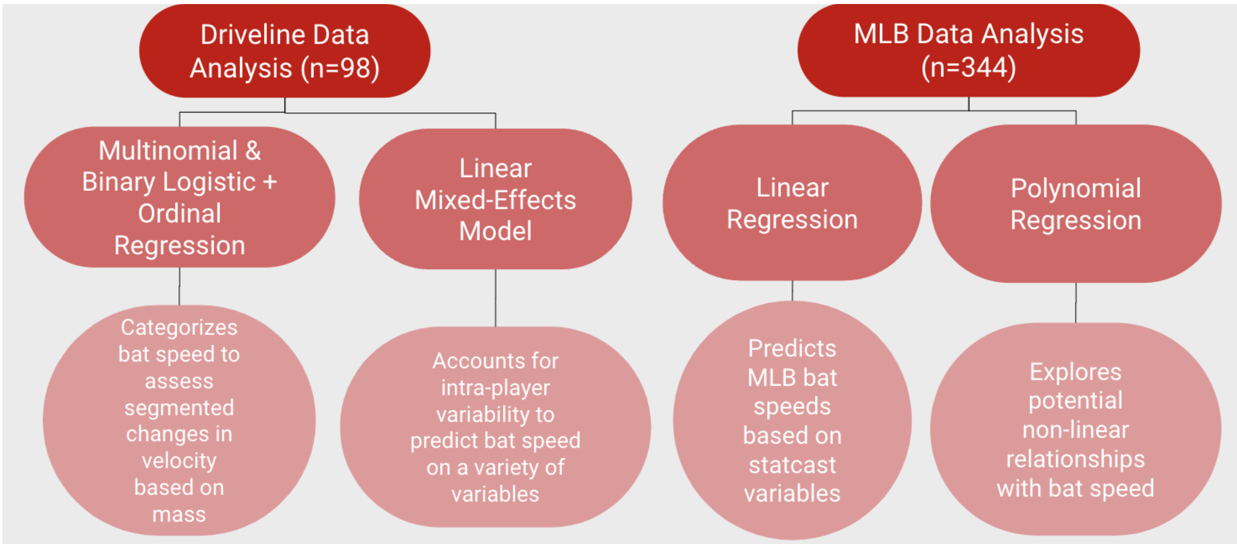


Chart 1: Description of methodology used for each dataset

Driveline. The first method with the biomechanics data uses a multinomial logistic model (1). It categorizes the Driveline bat speeds into thirds: the bottom third, middle third, and top third. After quantifying the groupings, the first model assesses the influence of a given body part (hand, shoulder, torso or pelvis) having the greatest change in velocity in the duration of first movement to the foot plant to produce a bat speed in the upper third. This approach allows for insights into which change in velocity predictor could be associated with greater bat speeds. It also allows for both a categorical dependent variable and multiple predictor variables.

The second technique displays a binary logistic regression (2). This model categorizes bat speed as a binary variable, where the athlete was either above average or below average in the Driveline data. This model takes the entire rank sequence—such as hand, shoulder, torso and pelvis—to predict the probability of bat speed, rather than just the first part of the sequence. The categorization of below or above average creates clear-cut categories to evaluate which sequencing series is associated with better bat speed. While this model contains high coefficients for each body part, it also has high p-values all above .1, suggesting that this model lacks overall significance. As seen in *Appendix Figure 1*, the orders of rank where pelvis or torso are the first or second rank, meaning higher changes in velocity, predict lower bat speeds according to the model.

Linear mixed effect modeling helps account for repeated measurements; since each player took multiple swings in each recorded session, the goal of this modeling is to help account for intra-player variation (Bijman et al., 2021). Mixed-effects models are best at dealing with unbalanced data and inter-subject variability, hence the reasoning in this case. The first mixed effect model (3) used the changes in velocity from the first move to the foot plant of the pelvis for four body parts: pelvis, torso, shoulder and hand. This model explores how the entire body interacts during the swing. This approach did provide insight into overall movement patterns, but the fixed effects for each body part resulted in high t-values and p-values, indicating a lack of statistical significance for this model in the relationship to bat speed.

The second mixed-effects approach (4) focused on the interactions in velocity changes between the adjacent body segments: pelvis and torso, torso and shoulder, and shoulder and hand. Interactions entail that instead of adding the predictors in the variable, they are multiplied in this case—such as pelvis and torso velocities are multiplied and added to the product of torso and shoulder. This method was employed to assess the importance of coordination between nearby body parts. Yet, like the first model, it did not reveal any statistically significant effects on hitter bat speed.

The third mixed-effects model (5) scaled the changes in velocity for each body part to stabilize the model while adding body mass as another predictor. The interactions in change in velocity for this model were separated into upper and lower body, where the terms were pelvis-torso and shoulder-hand velocity changes. Although the interaction terms showed some significance, body mass did emerge as a significant predictor of bat speed. This led to further investigation in the MLB dataset.

An ordinal regression (6) predicted the ranking order for the change in velocity amongst the different body parts based on the athlete's weight. A Kruskal-Wallis test compared the distributions of body mass across the predicted orders, and this test was utilized because it is non-parametric and does not assume normality (Laerd, n.d.) Ordinal regression was utilized because it measures the effect of the weight on each sequencing order; however, this method does not account for any of the distances in between the changes in velocity, strictly the ranking order.

Lastly, a multinomial logistic regression was conducted once again to predict the greatest change in velocity based on the athlete's weight, exit velocity and hip-shoulder separation (7). Hip-shoulder separation was added in an attempt to improve the model accuracy, as the majority of hitters in the dataset have their hand as the greatest change in velocity. This analysis allows the research to generalize into the larger MLB data.

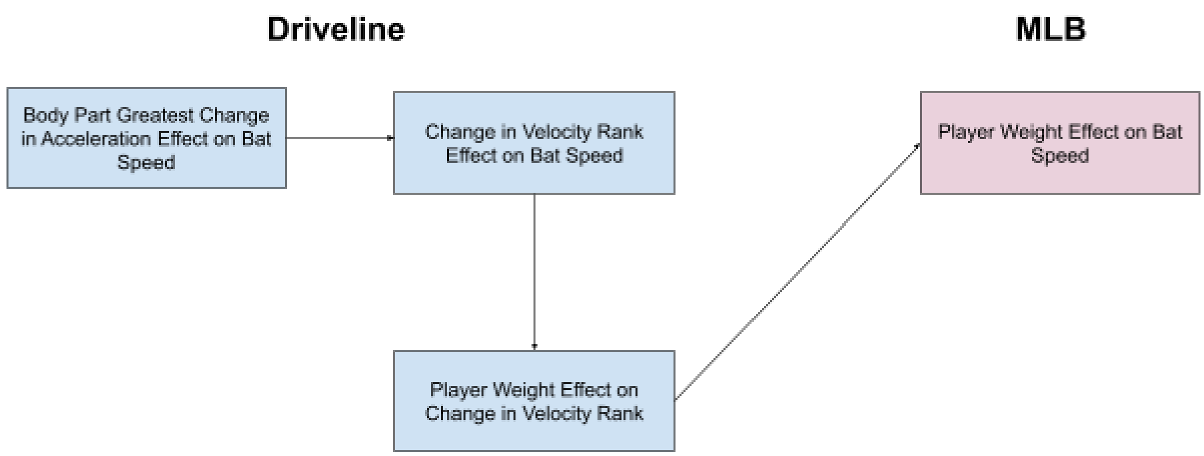


Chart 2: The process of scaling up conclusions from micro to macro data

MLB Statcast. The MLB data from Baseball Savant models the relationship between bat speed and body weight using both linear and non-linear regression models that include polynomial and interaction terms. The variables incorporated in these models include player weight, expected weighted on-base average (xwOBA; a measurement of how well a player makes contact), expected isolated power (xISO; quantifies how well the batter hits for power), average swing length, barrel rate and hard-hit percentage.

The data was manipulated on the Google Cloud website to remove missing values and select the 344 qualified hitters with 150 batted balls in play during the 2024 MLB season. To explore the impact of body weight on bat speed, the research employed a range of body weights—from 170 lbs to 240 lbs, in one pound increments—to assess bat speeds.

1. A linear regression model (8) was utilized to evaluate the association of plater weight on average bat speed. Other predictors in this regression include xwOBA, xISO, swing length, barrel rate, and hard-hit percentage
2. A polynomial regression model (9) was applied to determine potential non-linear relationships between weight and bat speed

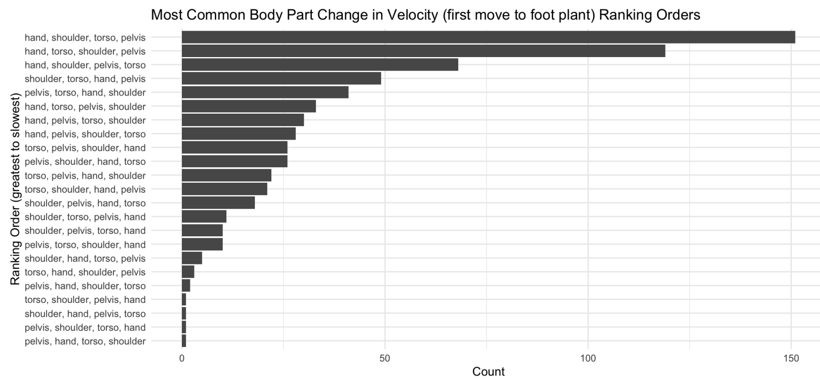
Model performance was evaluated using a train-test 80-20 split. The predictive performance of the models were evaluated through mean squared error and R-squared values.

Results

Driveline

The Driveline biomechanics data does not provide timestamps, yet the two key events of first movement and foot plant give an interval of time. While this timing is not consistent for each athlete, they are distinct moments in the swing. The measure of acceleration is the change in velocity over the change in time, and because time is not a constant measure of time for everyone, this measurement is referred to just as change in velocity in this analysis. *Figure 3* and

Figure 4 highlights that the hand, as it is the smallest of the body parts, holding the bat, and holds the greatest moment arm, has the greatest change in velocity from the first movement



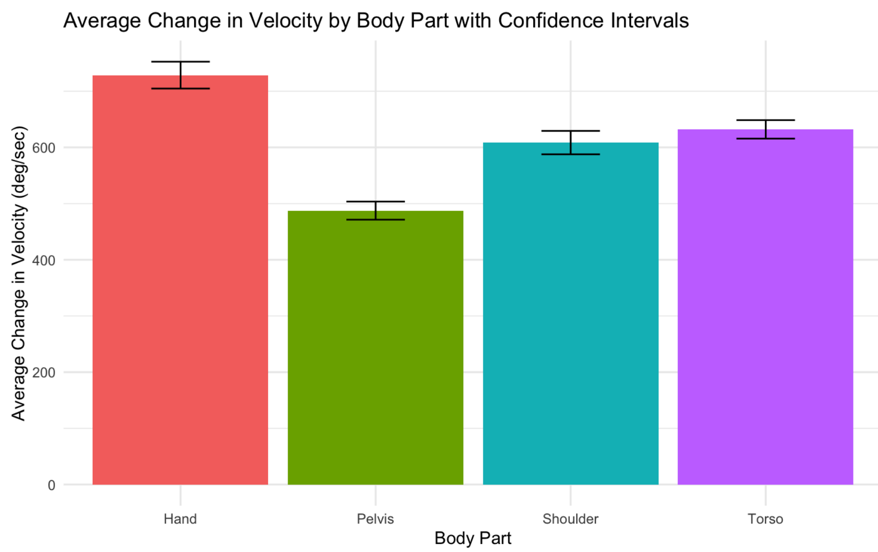


Figure 3 (top): The sequencing of change in velocity by body part from greatest to smallest

Figure 4 (bottom): The average change in velocity in Driveline data by body part

to the foot plant. The first model, the multinomial logistic regression (1) predicts bat speed into a category of thirds based on the body part which had the greatest change in velocity. As *Table 1* shows, the pelvis has a high negative coefficient in the upper third of bat speed, suggesting an association with athletes who have the pelvis act as the greatest change in velocity not having

	Intercept	Pelvis	Torso	Shoulder	Hand
Middle Third	-0.06824425	0.1666771	-0.23185431	-0.16228365	0.1592166
Upper Third	-0.016182190	-0.4974410	-0.04296626	0.06943967	0.3091457

Table 1: Coefficients of Multinomial Logistic Regression for Bat Speed Thirds by which body part has the greatest change in velocity

top end bat speed. Alternatively, hand as the greatest change in velocity also holds an association in the upper third, advising that it may predict a greater bat speed. The remaining coefficients are all between -3. and .3, highlighting slightly weaker relationships when predicting bat speed. Finally, the model converged after ten iterations with a final deviance value of 705.56, indicating an overall strong model fit compared to the significantly higher null deviance value of 1421.83.

The third linear mixed-effects model (5) shows the most significant of the three mixed-effects approaches. This model examined the effects of changes in the four body part velocities when the upper body parts interact together and the lower body parts interact together. Although the model suggests that body mass is a significant predictor of bat speed (Estimate =0.115 and t=4.87), the other predictors (the velocity interactions) did not show statistically significant results. The lack of significance in the velocities of all three mixed-effects models suggest that there are other factors that better associate with bat speed. One indication from the model is the variance (variance = 19.85), which highlights that differences between swings within sessions for a hitter can play a significant role, especially with a low sample size. However, the significance of body mass requires additional exploration.

Handedness was also considered in the analysis of model 5, but the interaction terms involving hand change in velocity and its relationship with bat speed were insignificant. This may suggest that hand movements in the early swing phase have similar effects on bat speed regardless of whether the batter swings left- or right-handed.

In the ordinal regression model (6), the ranking order projection came from three variables: the athlete's weight, the body part with the actual maximum change in velocity, and the bat speed for the athlete. As seen in *Table 2*, the athlete's body mass held a statistically significant slight negative (-.01201) association with the ranking order. Essentially, the heavier

	Value	Standard Error	t value
Player Mass	-0.01201	0.003518	-3.415
Max Change in Pelvis Velocity	.97573	0.198725	4.910
Max Change in Shoulder Velocity	-1.36838	0.191624	-7.141
Max Change in Torse Velocity	0.15187	0.261856	0.580
Bat Speed	.02000	.012343	1.621

Table 2: Coefficients of Ordinal Regression for Predicting Change in Velocity Rank

athletes are less likely to have a higher-ranking extremity—such as pelvis or torso—as the body part with the greatest change in velocity. This trend of having the pelvis or torso as the greatest change in velocity is less common throughout the data as seen in *Figure 3.* The model predicts two unique ranking orders, where a Kruskal-Wallis Test presents a low p-value ($p < 2.2e-16$),

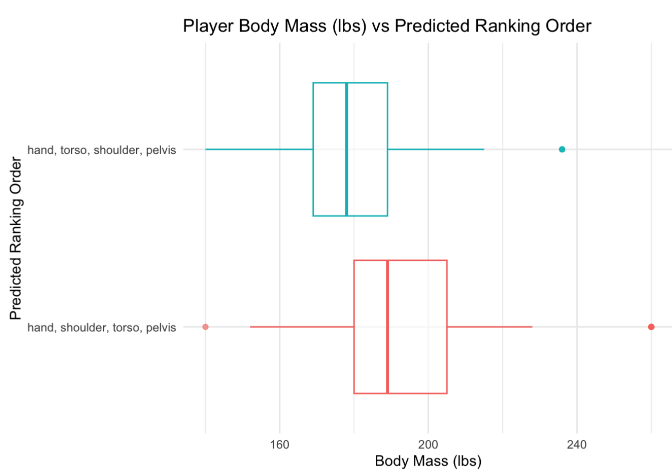


Figure 5: Box plots highlighting body masses of predicted orders

indicating a statistically significant difference in body mass across the two different ranked orders, as seen in *Figure 5*. Amongst the body parts, the pelvis held the strongest positive association with higher ranking positions (i.e. greatest changes in velocity), indicating that athletes with pelvis as the greatest change in velocity were more likely to output pelvis as a higher rank in the prediction. This is likely due to the small number of samples with pelvis as the greatest change in velocity, as seen in *Figure 3* and *Figure 4*. In contrast, the shoulder held a significant negative association, suggesting that the ranking order predicted the shoulder as a lower ranking in the order when given that it would be the greatest change in velocity. This may imply the shoulder's contribution to change in velocity is less impactful and inefficient in the case of model 6. The torso as the maximum change in velocity shows a less significant effect, and the hand is not explicitly displayed because the model treats it as the reference group; where the three other body parts are being compared to the hand, and their associations with the order are measured relative to the hand as the maximum change in velocity. Lastly, bat speed showed a weak positive association on the ranking order, suggesting a slight, non significant influence on the model.

The final Driveline model (7) predicts the hand to have the greatest change in velocity for all but sixteen swings. This strong bias seen in *Appendix Figure 2* suggests the model's tendency to predict hand change in velocity as a key influence in the micro sample.

MLB Statcast. The most significant finding from Driveline that could be applied to the MLB data was the importance of player weight. This was analyzed through both linear (8) and non-linear regression (9) models, while incorporating other predictors from Statcast. The linear regression, seen in *Figure 6* evaluates the relationship between player weight and average bat speed. The results showed that weight had a negative effect on bat speed, although potentially

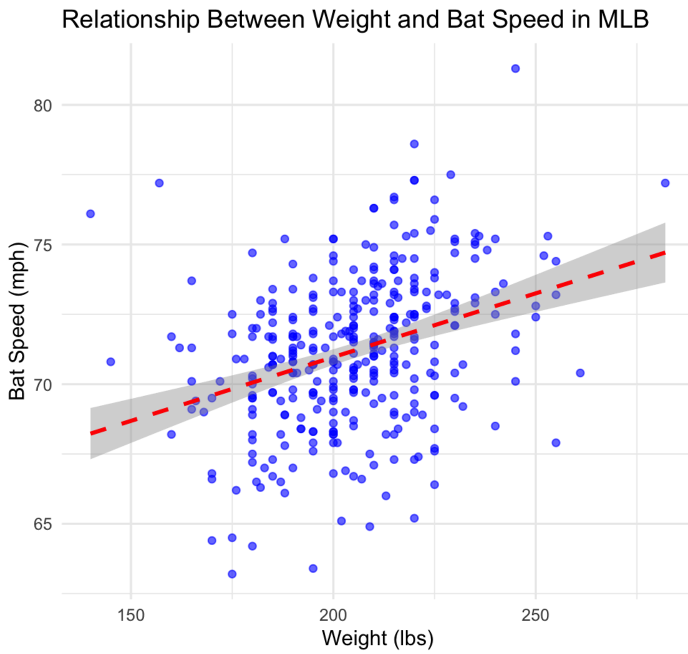


Figure 6: Linear Regression fit of body mass vs bat speed

insignificant (p-value = 0.082). The other predictors, however, such as xISO, average swing length, barrel rate, and hard-hit percentage all had statistically significant positive effects on bat speed. Although there may be some confounding in the variables, the model produced an R-squared of 0.67, indicating that about 67% of the variation in bat speed was explained by this linear regression.

A polynomial regression also produced similar results. The p-value was marginally lower, and the R-squared was slightly higher, suggesting that the polynomial fit may not show much more improvement in terms of explaining variance.

The linear model was evaluated using the train-test split of 80-20. The mean square error was 2.21 and the R-squared was .72, demonstrating predictive capability. *Figure 7* compares the

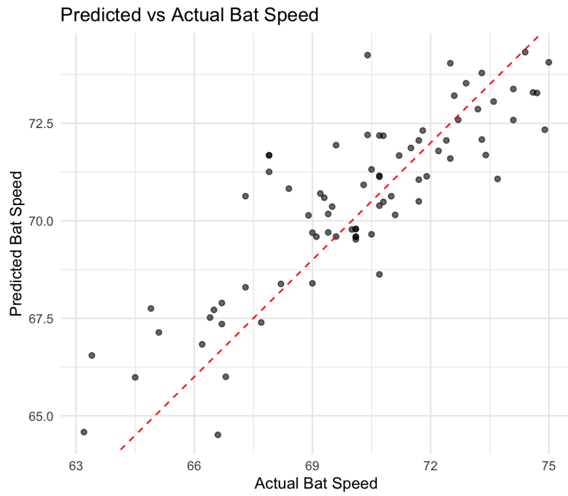


Figure 7: Actual vs predicted bat speed values using linear regression model

actual and predicted bat speeds, revealing a close alignment, supporting the model validity.

Discussion

This research aims to explore the factors that contribute to bat speed. As discussed, biomechanical analysis is highly theoretical; there is not one right way to swing a bat, but rather various strategies that may be beneficial to certain players. Measuring change in velocity between first movement and foot plant allowed the analysis to focus on a specific segmentation of the swing, isolating critical movements in the loading and landing phases.

Swing Sequencing Interpretation

The majority of hitters from Driveline employ the suggested swing sequencing of golfers presented by Bourgain et al. (2022) and Neal et al. (2007). While the exact sequencing cannot be measured without time as units, the body part with the greatest change in velocity during the first movement to footplant sequence likely indicates it is the last body part to decelerate. As seen in *Figure 9* (Signore & Shirazi, 2020), the more distal body parts have a steeper slope from first movement to foot plant, meaning they are increasing acceleration more rapidly. The data in this

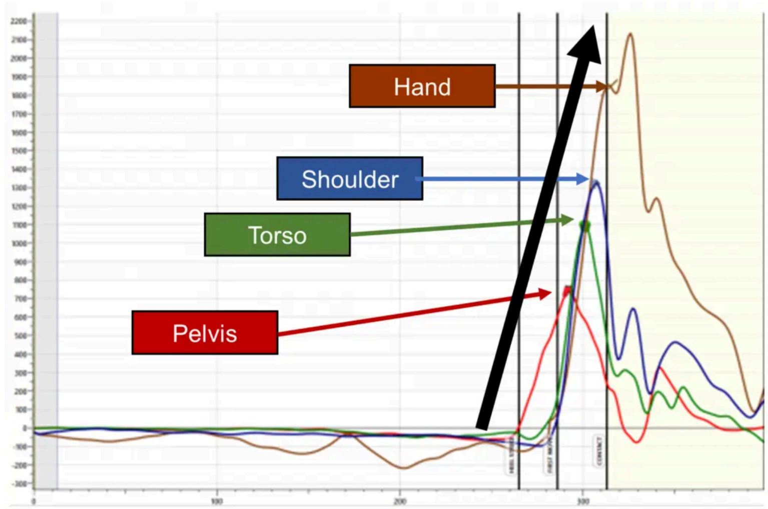


Figure 8: The Optimal Kinetic Sequence where the x-axis is time and the y-axis is acceleration

study cannot represent the precise time intervals acceleration peaks, but ranking the changes in velocity provides insight into the relative contribution of each body part accelerating into generating bat speed. The hand should generally hold the greatest acceleration, or change in velocity throughout the swing as Bourgain et al. (2022) acknowledges because of its distal location. This supports the physics principle of the lever arm—the distance from the pivot point—which presents a greater arm distance resulting in greater net force through increased angular velocity.

Ranking change in velocity can also predict the hitter's timing, meaning they may not be activating their muscles early enough, similar to the research presented by Nakata et al. (2013) and Stewart et al. (2020). If proximal measurements like the pelvis or torso have the greatest change in velocity, while the hand has the slowest change in velocity, the hitter likely has not begun any sort of hand acceleration, or their sequencing is out of place. Both of these outcomes will lead to negative results for the hitter.

While much of the prior research discusses how proximal points will hold smaller accelerations than the distal points, the rate of velocity change for the pelvis remains important.

Between first movement and foot plant, the faster the pelvis can accelerate, the greater the bat speed is on average, as seen in *Appendix Figure 3.*

Impact of Body Mass

The importance of body mass on bat speed was central to both the Driveline and MLB data analysis. In the Driveline analysis, specifically Model 6, it highlights both the findings of Szymanski et al. (2009) and Haruna et al. (2023) in the context of swing sequencing. As established, the proximal to distal sequencing produces the greatest segmented accelerations. While ordered change in velocity rank does not measure the exact sequence, players who have their velocities change from first move to foot plant in the order of proximal to distal likely share similar sequencing throughout the entirety of the swing. The difference amongst athletes' weights in the two predicted ranked orders are statistically significant, as seen in *Figure 5*. The distribution of heavier athletes are predicted to fall into the *hand, shoulder, torso, pelvis* rank in change in velocity, opposed to the *hand, torso, shoulder, pelvis* rank. Without more strength and lean body mass information, it is unreasonable to claim that heavier athletes are better at sequencing. However, stronger athletes, who have greater muscle mass, are more likely able to stabilize throughout the core as Fortenbaugh (2011) suggests, making segmentation within the swing more effective.

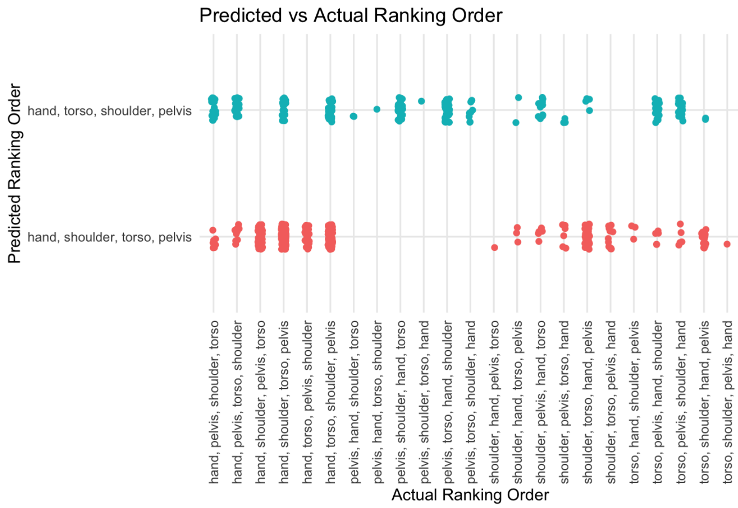


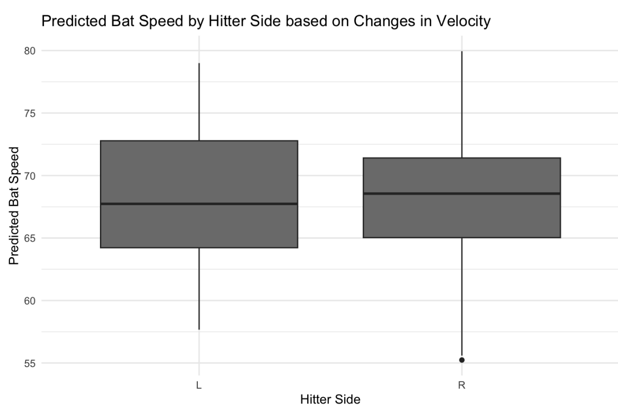
Figure 9: Predicted vs actual change in velocity order

Model 6 offers critical insights into quantifying change in velocity sequencing and body weight to measure the impact on bat speed. The ordinal logistic regression handles the ranking of velocity changes, however there are some limitations to the model. *Figure 9* underscores that while aspects of the model predict the order properly, there are definite flaws presented by the lack of accuracy. The regression does not likely capture the complexity of the non-linear reactions that occur within changes in velocity. Additionally, the risk of multicollinearity between nearby body parts such as torso and shoulder change in velocities likely impacts their contributions. Finally, the relatively small sample size presents the risk of overfitting in all of the models, as they may become tailored to the specific trends within the dataset.

When applied to MLB data, the negative relationship between body mass and bat speed seen in the Driveline analysis remained consistent. The linear regression model (8) showed similar trends, where weight exerted a modest effect on bat speed, as did the other factors. This model showed predictive capability, further indicating the role body mass plays influencing bat speed. However, it remains clear that other factors, such as strength and swing sequencing likely impact bat speed.

The Role of Handedness

When considering handedness and its role in this analysis, the mixed effects models (3 and 5) explored how biomechanics may be affected by handedness, while staying in the scope of changes in velocity and predicting bat speed. However, the results indicate that handedness had no significant result on bat speed. Specifically, Model 5 revealed that the scaled changes in velocity for the hand did not have a significant effect (t= -.497) on bat speed, suggesting that hand acceleration between first movement and foot plant may be less pivotal than other body parts in regard to bat speed. Interestingly, the interaction of shoulder and hand changes in



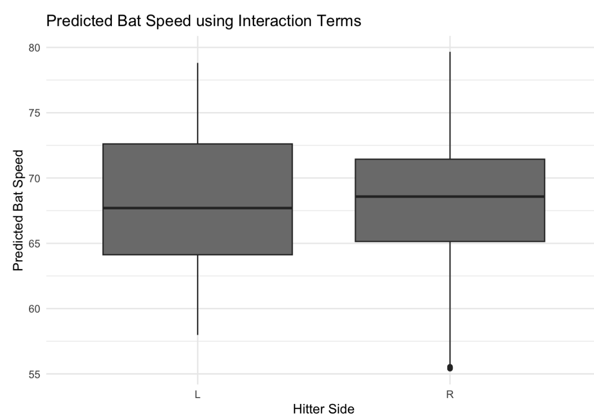


Figure 10 (left) and Figure 11 (right)

velocity also presented an insignificant effect on bat speed, further suggesting that lower body accelerations are more important during the first movement to foot plant phase, similar to the research of Reyes et al. (2011). Further research should consider batters who hit from their non-dominant side, and the role of hand acceleration in the case that Grondin et al. (1999) presented.

Figure 10 takes each change in velocity variable as a separate predictor with an additive linear effect on bat speed, as seen in Model 3. Alternatively, *Figure 11* uses Model 5 to investigate bat speed differences based on handedness presented by the mixed effects of interaction terms of scaled pelvis and torso scaled accelerations with scaled shoulder and hand accelerations. Using interaction terms allows this model to estimate the combined effects between the variables opposed to the individual contributions. Model 5 has a lower REML criterion and variance of the session intercept, indicating the interaction terms better fit the data and better account for variability. The strongest body segment effects between the two models are the interaction between pelvis and torso, suggesting a slight joint contribution to bat speed (t= -1.79). However, these results do not support any sort of difference based on handedness. As discussed, the benefits of handedness in hitting are based on the platoon advantage, not biomechanics (Chance, 2023).

Change in Velocity

Does change in velocity between first movement and foot plant matter? When considering Model 6, previous literature would suggest that both the acceleration of each body part during that duration and how each body part compares to one another is an important factor in producing bat speed. The maximum pelvis acceleration in the model (6) likely held a positive association with being ranked higher because it was less common in the data. For swings which had the pelvis, typically the first body part to accelerate, as the greatest change in velocity, the rest of the accelerations were likely nowhere near their peak. For most swings, the pelvis is fourth in ranking order, indicating that the predictors in the model do not vary significantly from those most popular sequences to other orders, resulting in predictions just ranking two outcomes.

In contrast, the shoulder has a negative association, suggesting the higher the shoulder is ranked (greater change in velocity), the more likely there is inefficient movement. This likely occurs because the model is trained on the most popular rank orders, where shoulder typically ranks as second or third. In instances where the shoulder is accelerating more than the hand, something is likely wrong with the swing. The interactions between first movement and foot plant are dynamic, but keeping the proximal to distal ranking, and maximizing each acceleration seem to contribute to more bat speed.

Conclusion

This research found evidence of swing sequencing influencing bat speed, resulting in an effect on performance. The four main takeaways from the study are the following:

1. *Changes in velocity at different body parts between first movement and foot plant are significant as precise measurements by themselves and compared to other body extremities.*

2. Hand should have the greatest change in velocity during this period if the goal is to maximize bat speed. A proximal-distal sequencing may also be associated with improved bat speed.

3. *Body mass plays a significant role generating bat speed, however it is not the only factor.*

4. *Handedness has a minimal effect on biomechanics.*

Limitations

The models presented in the research provide insights into biomechanics and bat speed, but there are several limitations to be acknowledged. First, Models 1 and 2 both rely on independence among predictors, which is not how the body interacts in the swing. Biomechanics struggle to capture how each body part actually interacts, as it is different for every swing, so models struggle to learn the full relationship. Additionally, as mentioned earlier, the data size and potential over-simplification of acceleration interactions may lead to model overfitting in all of the biomechanics models. This is seen in both *Figure 9* and *Appendix Figure 2*, where the model predictions are not overly accurate or precise. Lack of accuracy does not make the models necessarily insignificant, but the data size impacts the number of plausible acceleration ranks, which leads to overfitting in the models.

The data itself also presents limitations aside from the number of swing sessions. Driveline hitters are predominately college players, with little diversity of player experience. Furthermore, they are all working out of the same hitting facility, likely presented with similar coaching, which is not indicative of a random sample. The lack of time stamps in the dataset makes this analysis more difficult, as hitters will vary the amount of time between first movement and foot plant. Having the measurement of time would allow for more concrete and nuanced analysis regarding the distance in between peak acceleration points. Lastly, public biomechanics are unavailable for MLB players, making any type of conclusions difficult to scale up to larger datasets. The only comparable, guaranteed accurate variables between datasets were bat speed, handedness, and body weight. Even with body weight, previous literature highlighted the importance of measuring strength and lean body mass as well.

Future Work

Because of the quantity of modeling techniques employed to measure different results, employing a more consistent analysis strategy may help solidify results. Exploring models that account for correlated random effects may better capture how performance across sessions is related for each player. Incorporating more data surrounding player health, such as information collected from a smartwatch or Whoop may help influence performance on a given day. In relation to previous literature, conducting analysis of biomechanics with a larger sample of switch hitters would help drive analysis regarding potential biomechanical differences due to handedness.

When attempting to scale up biomechanics research to the MLB level, implementing computer vision could help estimate different joint angles at key points such as first movement or foot plant. This would create more variables to cross analyze between data sets aside from just bat speed, weight and handedness.

References

Adler, D. (2024, May 13). *MLB bat speed leaders for 2024*. MLB.com. https://www.mlb.com/news/mlb-bat-speed-leaders-for-2024

Bijman, E. Y., Kaltenbach, H.-M., & Stelling, J. (2021). Experimental analysis and modeling of single-cell time-course data. *Current Opinion in Systems Biology, 28*, Article 100359. https://doi.org/10.1016/j.coisb.2021.100359

Boddy, K. (2012, January 6). Biomechanics explained: The difficulty of measurement. *Driveline Baseball*. https://www.drivelinebaseball.com/2012/01/biomechanics-explained-the-difficulty-of-measurement/

Bourgain, M., Rouch, P., Rouillon, O., Thoreux, P., & Sauret, C. (2022). Golf Swing Biomechanics: A Systematic Review and Methodological Recommendations for Kinematics. *Sports (Basel, Switzerland)*, *10*(6), 91. https://doi.org/10.3390/sports10060091

Callaway, S., Glaws, K., Mitchell, M., Scerbo, H., Voight, M., & Sells, P. (2012). An analysis of peak pelvis rotation speed, gluteus maximus and medius strength in high versus low handicap golfers during the golf swing. *International journal of sports physical therapy*, *7*(3), 288–295.

Chance, D. M., & Maymin, P. Z. (2023). A new look at the left-handed advantage in baseball. *International Journal of Performance Analysis in Sport, 23*(6), 458–488. https://doi.org/10.1080/24748668.2023.2255806

Chu, C. Y. C., Chang, T., & Chu, J. (2016). Opposite hand advantage and the overrepresentation of left-handed players in Major League Baseball. *Academia Economic Papers, 44*(2), 171–205.

Crisco, J. J., Osvalds, N. J., & Rainbow, M. J. (2018). The Kinetics of Swinging a Baseball Bat. *Journal of Applied Biomechanics*, *34*(5), 386-391. https://doi.org/10.1123/jab.2017-0337

Fortenbaugh, D. M. (2011). The biomechanics of the baseball swing [University of Miami]. https://scholarship.miami.edu/esploro/outputs/991031448086302976

Grondin, S., Guiard, Y., Ivry, R. B., & Koren, S. (1999). Manual laterality and hitting performance in Major League Baseball. *Journal of Experimental Psychology: Human*

Perception and Performance, 25(3), 747–754. https://doi.org/10.1037/0096-1523.25.3.747

Hargrove, T. (2013, August 22). The complexity of biomechanics. *Better Movement*. https://www.bettermovement.org/blog/2013/the-complexity-of-biomechanics

Haruna, R., Doi, T., Habu, D., Yasumoto, S., & Hongu, N. (2023). Strength and conditioning programs to increase bat swing velocity for collegiate baseball players. *Sports (Basel, Switzerland)*, *11*(10), 202. https://doi.org/10.3390/sports11100202

Laerd Statistics. (n.d.). *Kruskal-Wallis H test using SPSS Statistics*. https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php

Nakata, H., Miura, A., Yoshie, M., Kanosue, K., & Kudo, K. (2013). Electromyographic analysis of lower limbs during baseball batting. *Journal of Strength and Conditioning Research*, 27(5), 1179-1187. DOI.org/10.1519/JSC.0b013e3182653ca9.

Neal, R., Lumsden, R., Holland, M., & Mason, B. (2007). *Body segment sequencing and timing in golf*. Golf BioDynamics Pty Ltd; Australian Institute of Sport – Golf Program; Australian Institute of Sport – Sport Science and Sports Medicine Staff.

Reyes, G.F., Dickin, D.C., Crusat, N.J.K., & Dolny, D.G. (2011). Whole-body vibration effects on the muscle activity of upper and lower body muscles during the baseball swing in the recreational baseball hitters. *Sport Biomechanics, 10(4),* 280-293. http://dx.doi.org/10.1080/14763141.2011.629208

Signore, N., & Shirazi, B. (2020, February 4). *Assessing hitters at RPP Baseball*. Rockland Peak Performance. https://rocklandpeakperformance.com/assessing-hitters-at-rpp-baseball/

Stewart, E., Stewart, M., Simpson, J., Knight, A., Chandler, H., & Shapiro, R. (2020). Sequential order of swing phase initiation in baseball. *Journal of Sports Analytics,* 199 – 204. DOI.org/10.3233/JSA-200394

Szymanski, D. J., DeRenne, C., & Spaniol, F. J. (2009). Contributing factors for increased bat swing velocity. *Journal of Strength and Conditioning Research, 23*(4), 1338–1352. https://doi.org/10.1519/JSC.0b013e3181a30d8d

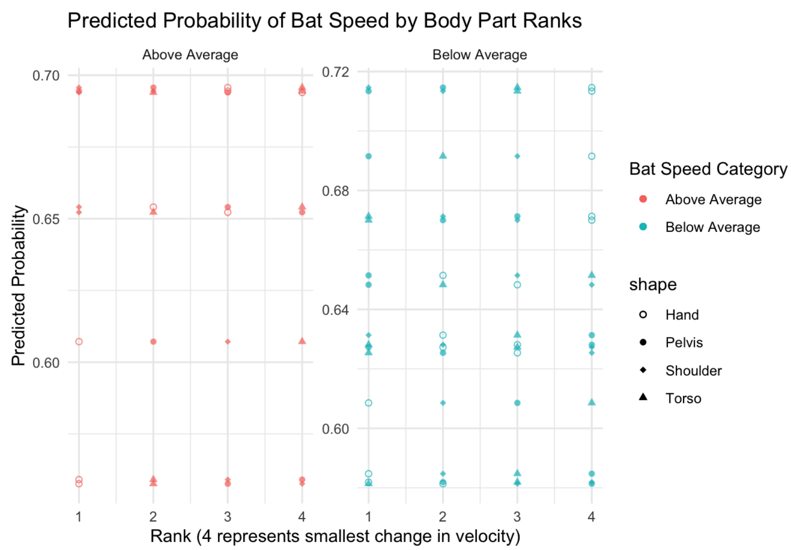
Szymanski, D. J., Szymanski, J. M., Schade, R. L., Bradford, T. J., McIntyre, J. S., DeRenne, C., & Madsen, N. H. (2010). The relation between anthropometric and physiological variables and bat velocity of high-school baseball players before and after 12 weeks of training. *Journal of Strength and Conditioning Research, 24*(11), 2933–2943. https://doi.org/10.1519/JSC.0b013e3181f0a76a

Wasserberger KW, Brady AC, Besky DM, Jones BR, Boddy KJ. The OpenBiomechanics Project: The open source initiative for anonymized, elite-level athletic motion capture data. (2022).

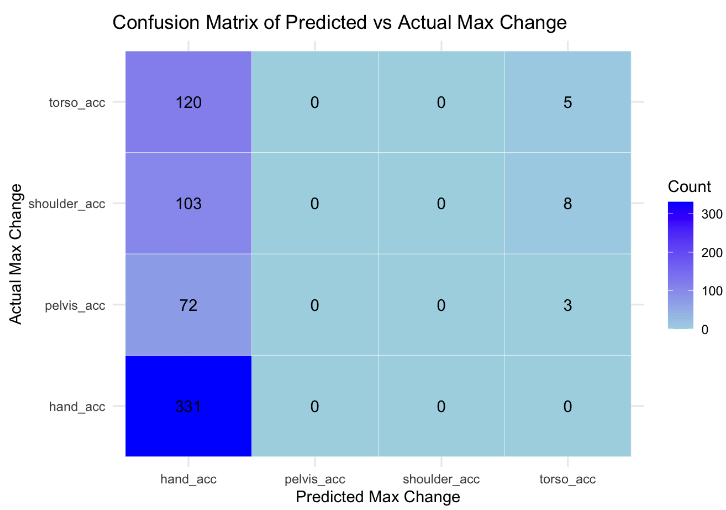
Appendices

https://github.com/aronsoneli/401Project/blob/main/README.md

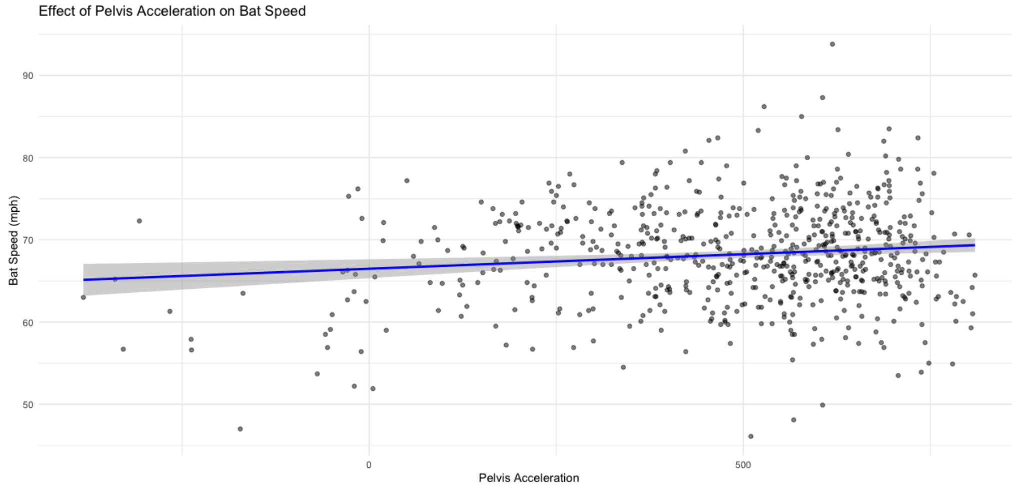
Appendix Figure 1: Using Model 2 to predict bat speed category based on change in velocity sequence



Appendix Figure 2: Predicted max change in velocity body part vs. actual body part



Appendix Figure 3: Pelvis change in velocity relationship with bat speed


Effect of Pelvis Acceleration on Bat Speed
90
80
70
60
50
Bat Speed (mph)
0
500
Pelvis Acceleration