Analyzing Baseball Platoon Strategies: Integrating Bat Speed, Swing Length, Spray Angle, Clustering, and Performance Metrics

Wei Chieh Chen
Syracuse University

2025-04-30

Introduction

Purpose: To present a data-driven analysis of platoon strategies in baseball using 2024 Statcast data

- What are Platoon Strategies?
  - Leverage batter-pitcher handedness matchups to boost offensive output
  - Right-handed batters (RHB) typically excel vs left-handed pitchers (LHP)
  - Left-handed batters (LHB) perform better vs right-handed pitchers (RHP)
- Why Use Statcast Data?
  - Provides advanced metrics: bat speed, swing length, attack angle
  - Enables deeper analysis of swing mechanics and performance
- Objective:
  - Cluster batters by swing characteristics
  - Simulate game outcomes with Monte Carlo simulation
  - Measure platoon performance metrics (BA, OBP, SLG, OPS, OPS+)

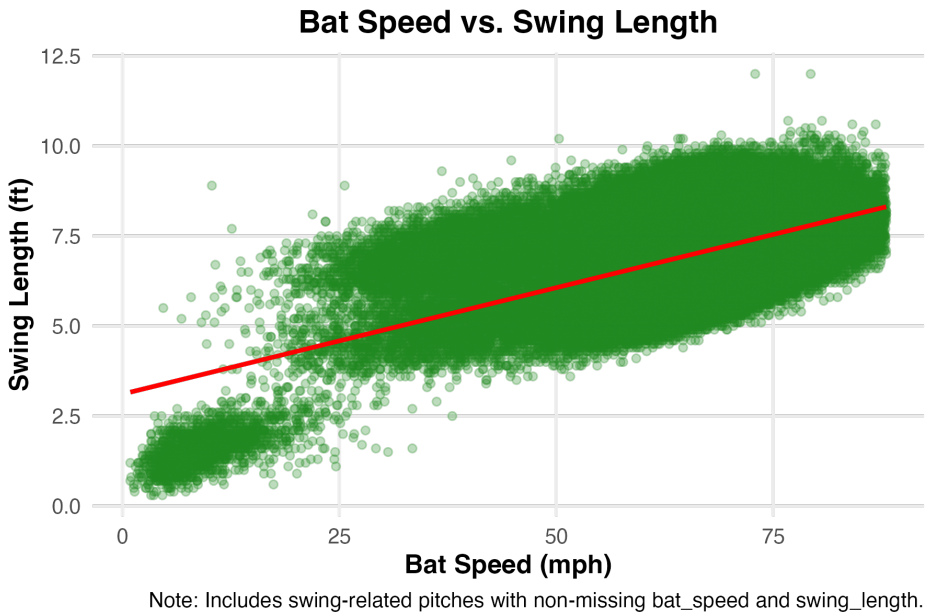
Research Questions and Goals

- Key Questions:
  - How do swing metrics (bat speed, swing length, attack angle) vary across batter clusters?
  - What is the expected run production of a platoon-optimized lineup?
  - How do RHB, LHB, and Switch Hitters perform against LHP and RHP?
- Goals:
  - Identify distinct batter profiles using clustering
  - Estimate offensive output with simulation
  - Provide actionable insights for lineup optimization in professional baseball

Data Source

- Dataset: Statcast Pitch Swing Data (April 2, 2024 – October 30, 2024)
  - 701,557 pitches from MLB regular season and playoffs
  - Key variables:
    - Bat speed: mean = 71.1 mph, SD = 4.8
    - Swing length: mean = 7.0 ft, SD = 0.6
    - Attack angle (derived), launch angle, launch speed
    - Batter/pitcher handedness (stand, p_throws), pitch type, event outcomes
- League Context: 2024 MLB averages (Baseball Reference)
  - BA: 0.243, OBP: 0.312, SLG: 0.399

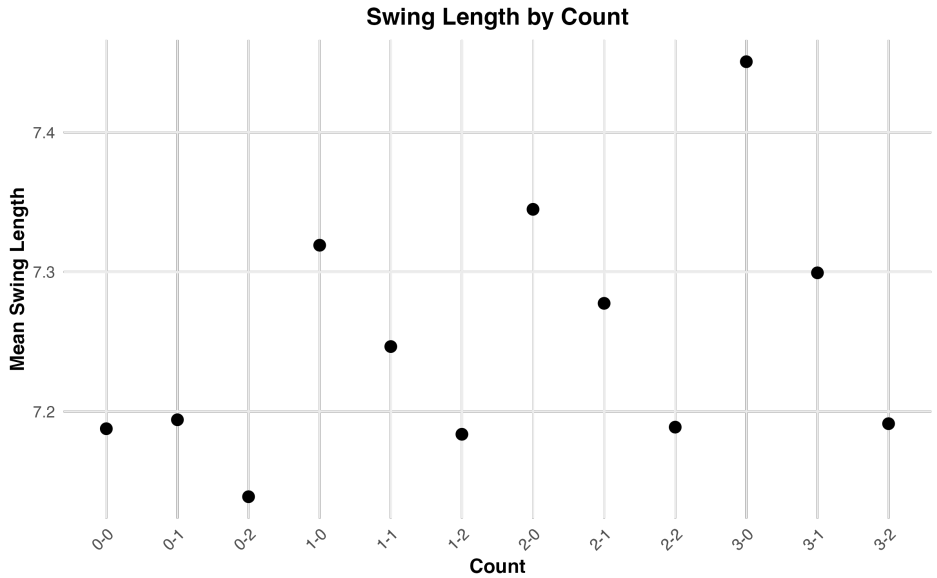
Exploratory Data Analysis (1/3)



- The plot of Swing Length vs Bat Speed shows wide variability.

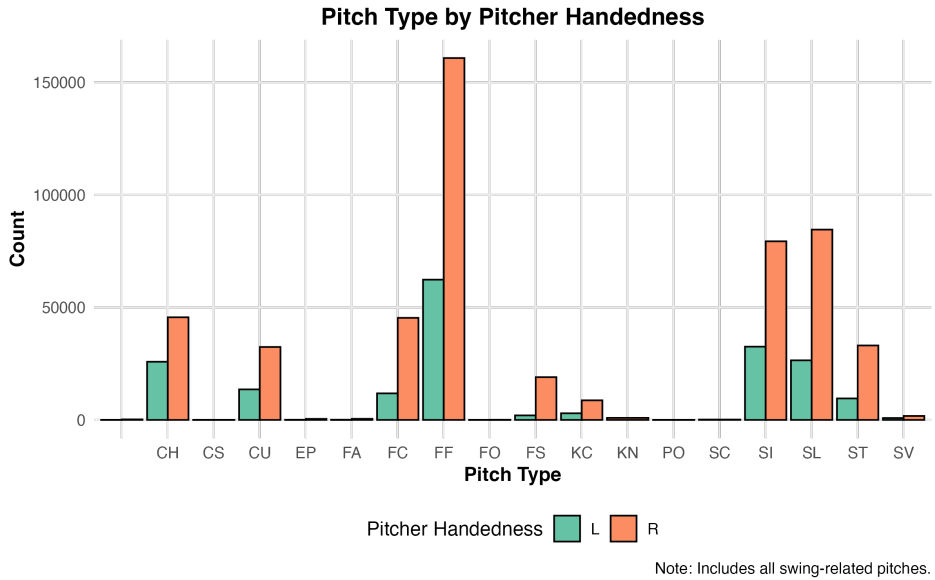
Exploratory Data Analysis (2/3)

- Next step is to pinpoint which game-situation inputs (like pitch count) influence it.



- Hitters noticeably shorten their swings in two-strike counts (especially 3-2), while they lengthen them in hitter-friendly counts such as 1-0, 2-0, and 3-1.

Exploratory Data Analysis (3/3)



- Pitcher handedness reflects Right-Handed Pitchers > Left-Handed Pitchers.
- Pitch type highlights the prevalence of four-seam fastballs.

Preprocessing

- Steps:
  - Filter for swing-related pitches (e.g., “hit_into_play”, “foul”)
  - Calculate spray angle:
    $sprayangle = atan2(hc_y - 200, hc_x - 125) \times 180/\pi$
  - Adjust for batter handedness (LHB: negative, RHB: positive)
- Compute attack angle:
  - Average launch angle of top 20% swings by launch speed
  - Use parabolic model for precision
- Calculate plate discipline metrics:
  - Z-Swing% (in-zone swing rate), O-Swing% (out-of-zone swing rate), Swing%

Clustering Analysis – Methodology

- Aggregation: Summarize by player, pitch type, pitcher handedness
  - Metrics: mean bat speed, swing length, attack angle, spray angle, eBA, eSLG
- Dimensionality Reduction: Apply PCA, retain first 4 components
- Clustering: K-means (k=4) separately for RHB and LHB
  - Labels:
    - RHB-1 to RHB-4 (right-handed batters, clusters 1–4)
    - LHB-1 to LHB-4 (left-handed batters, clusters 1–4)

Clustering Analysis – Cluster Characteristics

- RHB Clusters:
  - RHB-1: Contact-oriented (moderate bat speed: 70.9–71.2 mph, low attack angle: 6.2–6.5°)
  - RHB-2: Aggressive power hitters (high bat speed: 72.4–72.8 mph, high attack angle: 15.9°)
  - RHB-3: Compact swingers (low bat speed: 67.0 mph, low attack angle: 7.6°)
  - RHB-4: Balanced hitters (moderate bat speed: 69.4–69.5 mph, attack angle: 14.8–15.3°)
- LHB Clusters:
  - LHB-1: Compact swingers (low bat speed: 65.2–66.0 mph, attack angle: 10.2–10.7°)
  - LHB-2: Versatile hitters (moderate bat speed: 69.1–69.9 mph, attack angle: 10.5–12.2°)
  - LHB-3: Power hitters (high attack angle: 18.6–20.1°, strong vs RHP)
  - LHB-4: Aggressive hitters (highest bat speed: 72.3–73.4 mph, attack angle: 13.7–13.8°)

Monte Carlo Simulation

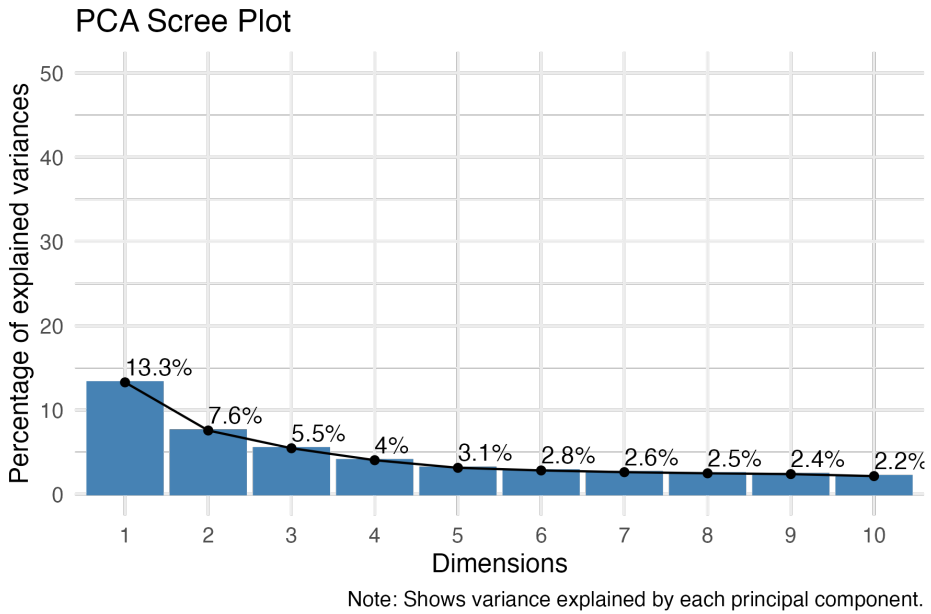
- Setup:
  - Estimate event probabilities (e.g., HR, strikeout) per cluster using eBA, eSLG, wOBA
    - E.g., HR probability = 0.2 × wOBA, strikeout = 0.3 × (1 - eBA)
  - Sample lineup: RHB-1, LHB-2, RHB-4, LHB-1, RHB-3, LHB-3, RHB-2, LHB-4, RHB-1
    - Alternates RHB and LHB for platoon balance
- Simulate 10,000 games, 9 innings each, park factor = 1
- Result: Average 4.8664 runs/game
- Above 2024 MLB average (4.39 runs/game), reflecting effective platoon strategy

Platoon Performance Metrics

- Aggregation: Last pitch of each plate appearance (PA)
- Batter Classification: RHB, LHB, Switch Hitters based on stand
- Metrics Calculated:
  - BA: Hits / At-Bats (excludes walks, HBP, sac flies, sac bunts)
  - OBP: (Hits + Walks + HBP) / (At-Bats + Walks + HBP + Sac Flies)
  - SLG: Total Bases / At-Bats (1 for single, 2 for double, etc.)
  - OPS: OBP + SLG
  - OPS+: $100 \times \left(\frac{\text{OBP}}{\text{lgOBP}} + \frac{\text{SLG}}{\text{lgSLG}} - 1 \right)$
  - lgOBP $= 0.309$, lgSLG $= 0.397$ (from dataset)

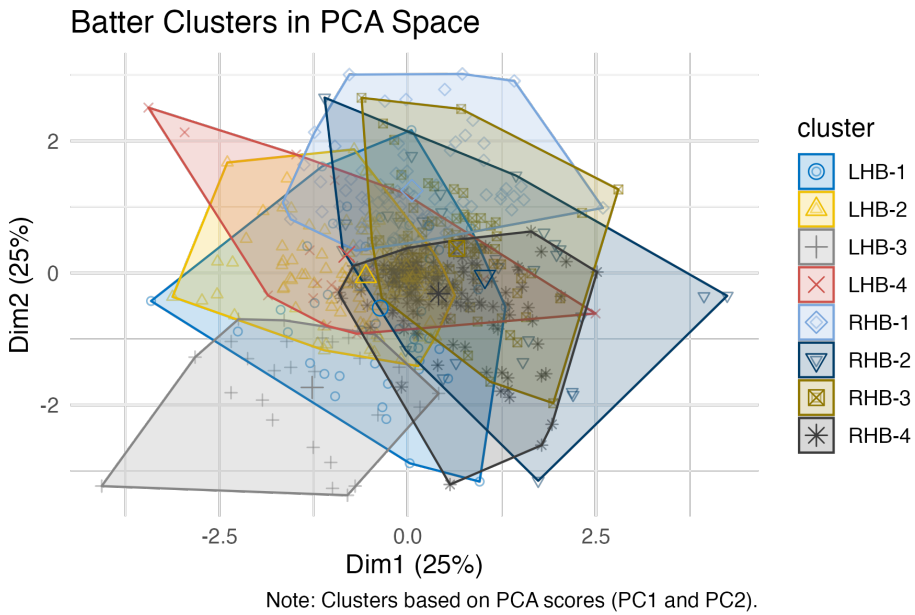
Results – Clustering Insights (1/4)

- PCA Scree Plot (Figure 2): First 4 components capture significant variance



Results – Clustering Insights (2/4)

- Cluster Visualization (Figure 3): Distinct RHB and LHB groups



Note: Clusters based on PCA scores (PC1 and PC2).

Results – Clustering Insights (3/4)

Table 1: Platoon Summary by Cluster for Four-Seam Fastballs (FF)

Cluster	Pitcher	Bat.Speed..mph.	Swing.Length..ft.	Attack.Angle..deg.
LHB-1	LHP	65.2	6.3	10.7
LHB-1	RHP	66.0	6.3	10.2
LHB-2	LHP	69.9	6.7	10.5
LHB-2	RHP	69.1	6.7	12.2
LHB-3	LHP	70.2	6.8	18.6
LHB-3	RHP	70.6	6.9	20.1
LHB-4	LHP	72.3	6.8	13.7
LHB-4	RHP	73.4	6.9	13.8
RHB-1	LHP	71.2	6.9	6.2
RHB-1	RHP	70.9	7.0	6.5
RHB-2	LHP	72.8	7.3	15.9
RHB-2	RHP	72.4	7.3	15.9
RHB-3	LHP	67.0	6.4	7.6
RHB-3	RHP	67.0	6.4	7.6
RHB-4	LHP	69.5	6.8	14.8
RHB-4	RHP	69.4	6.8	15.3

Results – Clustering Insights (3/4)

- RHB-2 vs RHP: High bat speed (72.4 mph), attack angle (15.9°)
- LHB-3 vs RHP: High attack angle (20.1°), power-oriented
- LHB clusters show higher attack angles vs RHP (e.g., LHB-3: 20.1° vs 18.6° vs LHP)

Results – Platoon Performance (1/5)

Table 2: Swing Metrics vs LHP

Batter_Type	Bat.Speed..mph.	Swing.Length..ft.	Attack.Angle..deg.
LHB	69.1	7.1	12.8
RHB	70.6	7.3	11.4
Switch Hitter	69.3	7.2	9.6

Results – Platoon Performance (2/5)

Table 3: Performance Metrics vs LHP

Batter_Type	Batting_Avg	On_Base_Pct	Slugging_Pct	OPS	OPS_plus
LHB	0.236	0.301	0.365	0.666	89
RHB	0.246	0.313	0.406	0.719	104
Switch Hitter	0.252	0.312	0.412	0.724	105

Results – Platoon Performance (3/5)

Table 4: Swing Metrics vs RHP

Batter_Type	Bat.Speed..mph.	Swing.Length..ft.	Attack.Angle..deg.
LHB	70.3	7.2	12.7
RHB	70.1	7.3	11.5
Switch Hitter	69.0	7.1	13.6

Results – Platoon Performance (4/5)

Table 5: Performance Metrics vs RHP

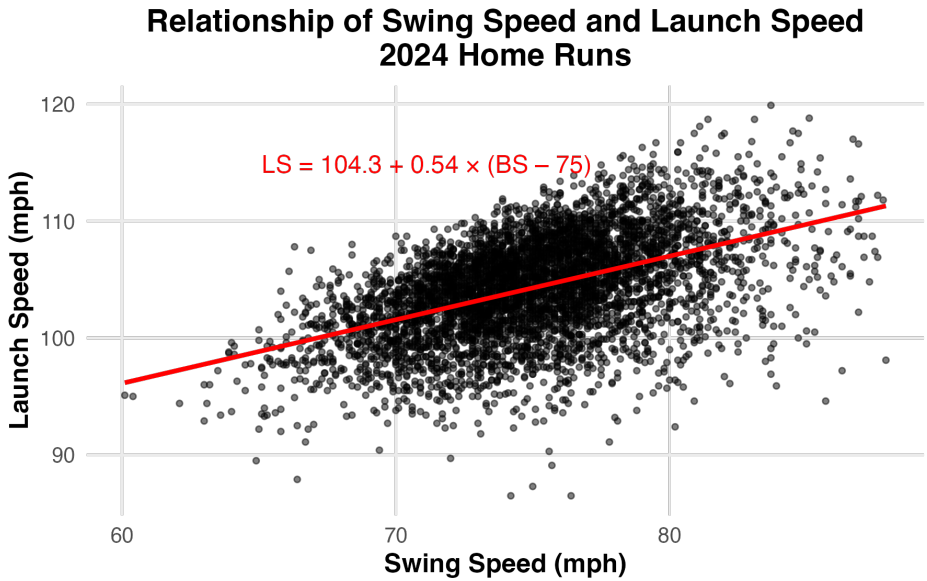
Batter_Type	Batting_Avg	On_Base_Pct	Slugging_Pct	OPS	OPS_plus
LHB	0.244	0.318	0.416	0.734	108
RHB	0.239	0.301	0.386	0.687	95
Switch Hitter	0.239	0.311	0.393	0.704	100

Results – Platoon Performance (5/5)

- vs LHP (Tables 2 & 3):
  - RHB outperform LHB (OPS+ = 104 vs 89)
  - Switch Hitters: OPS+ = 105 (slightly above average)
  - RHB: Higher bat speed (70.6 mph vs LHB's 69.1 mph)
  - Switch Hitters: Lower attack angle (9.6°), flatter swing
- vs RHP (Tables 4 & 5):
  - LHB excel (OPS+ = 108, SLG = 0.416)
  - RHB: OPS+ = 95, Switch Hitters: OPS+ = 100
  - LHB: Higher attack angle (12.7° vs RHB's 11.5°)

Results – Swing Dynamics

- Figure 1: Bat Speed vs Launch Speed (Home Runs)



- Positive correlation: Launch Speed $= 104.4 + 0.54 \times$ (Bat Speed – 75)
- Higher bat speed $\rightarrow$ harder hits $\rightarrow$ more home runs
- Implication: Bat speed is a key driver of power hitting, supporting aggressive clusters like RHB-2 and LHB-4

Discussion – Applications

- Lineup Optimization:
  - Prioritize LHB vs RHP (OPS+ = 108), RHB vs LHP (OPS+ = 104)
  - Use swing metrics (e.g., LHB attack angle vs RHP) for matchups
- Scouting & Development:
  - Target RHB-2/LHB-3 for power, RHB-1/RHB-3 for contact
  - Train to optimize bat speed (see Figure 1)
- Game Strategy:
  - Simulation (4.8664 runs/game) informs pinch-hitting, lineup adjustments
  - League Context: 2024 MLB averages (Baseball Reference): 4.39 runs/game
  - Extend simulation with real-time pitcher data
- Player Loading Management:
  - Reduce Injuries Probability

Discussion – Limitations & Future Work

- Limitations:
  - Simulation simplifies dynamics (fixed probabilities, park factor = 1)
  - Missing data in Statcast may bias clustering
  - Playoff games lower offensive metrics (BA = 0.242 vs 0.243)
  - Assumes independence in regression (Figure 1)
  - Lacks pitcher-specific metrics (e.g., arm angle)
- Future Work:
  - Incorporate pitcher metrics (e.g., arm angle variations)
  - Model situational factors (e.g., pitcher fatigue, defense)
  - Sports Medicine (e.g. Integrate Injury Datasets)

Conclusion

- Key Findings:
  - Clustering identifies distinct batter profiles (e.g., RHB-2, LHB-3 for power)
  - Platoon advantages: LHB vs RHP (OPS+ = 108), RHB vs LHP (OPS+ = 104)
  - Simulation estimates 4.8664 runs/game, supporting platoon strategy
- Impact: Provides a framework for lineup optimization using Statcast data
- Next Steps: Enhance simulation with pitcher data for more precise strategies

References (1/2)

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433–459.
- Adler, D. (2024). The rainbow of arm angles in the playoffs. Retrieved from https://www.mlb.com/news/the-range-of-pitcher-arm-angles-in-the-2024-mlb-postseason (Accessed: 2025-04-30).
- Albert, J. (2024). Bat Speed and Swing Length. Retrieved from https://baseballwithr.wordpress.com/2024/05/17/bat-speed-and-swing-length/ (Accessed: 2025-04-30).
- Baseball Reference. (2024). 2024 MLB Standard Batting. Retrieved from https://www.baseball-reference.com/leagues/majors/bat.shtml (Accessed: 2025-04-30).

References (2/2)

- Marshall, D. (2017). Reverse engineering swing mechanics from Statcast data. Retrieved from https://community.fangraphs.com/reverse-engineering-swing-mechanics-from-statcast-data/ (Accessed: 2025-04-30).
- Petriello, M. (2024). Everything to know about Statcast's new bat-tracking data. MLB.com. Retrieved from https://www.mlb.com/news/what-you-need-to-know-about-statcast-bat-tracking (Accessed: 2025-04-30).
- Winston, W. L., et al. (2022). Mathletics: How Gamblers, Managers, and Fans Use Mathematics in Sports (2nd ed.). Princeton University Press.