Optimizing Baseball Batting Orders: A Data-Driven Approach to Lineup Efficiency

Benjamin Buman, Kiran Shay, Jake Rasmussen
Mentor: Dr. Dahbura

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Introduction

We built a lineup optimization tool that uses advanced baseball analytics to find the most run-productive batting orders. Our model evaluates 4-player combinations using probabilistic run expectancy to maximize overall offensive output.

Objectives

- Develop a model to evaluate the effectiveness of a baseball batting lineup.
- Introduce and implement Baserunner-dependent Run Production (BRP) as a context-aware performance metric.
- Identify the lineup order that maximizes expected run production based on player statistics and probabilistic outcomes.

Materials and Methods

Player statistics, including PA, H, 2B, 3B, HR, BB, HBP, SB, and IBB, were collected from public databases and used to generate outcome probability distributions for each batter. Run values were assigned using pre-calculated BRP (Baserunner-dependent Run Production) values derived from RE24 tables, which account for baserunner context. Each lineup was evaluated by simulating all 4-batter tuples and summing their expected contributions. The optimization process used exhaustive search across all 9-player permutations, with options to apply constraints such as fixed batting positions and handedness limits.

Results and Findings

- Our algorithm consistently places the best hitters at the top of the lineup to maximize plate appearances.
- Optimized lineups prioritize expected run production over traditional roles like the “cleanup hitter.”
- High-BRP players (e.g., Kyle Schwarber or Bryce Harper) were often placed in the leadoff or #2 spot rather than later in the order.
- This strategy increases the number of times elite hitters come to the plate, improving run output over the course of a game.
- Compared to traditional or randomly ordered lineups, our optimized lineups yielded 5-20% more expected runs per game (depending on roster and constraints).
- Results demonstrate the effectiveness of probabilistic modeling and data-driven decision-making in challenging conventional baseball strategy.

Original Phillies Lineup

1. Bryson Stott | BRP: 0.55
2. Trea Turner | BRP: 1.42
3. Bryce Harper | BRP: 0.93
4. Kyle Schwarber | BRP: 0.70
5. Nick Castellanos | BRP: 0.62
6. Jacob Realmuto | BRP: 1.24
7. Max Kepler | BRP: 0.60
8. Alec Bohm | BRP: 0.41
9. Brandon Marsh | BRP: 0.61

Expected Runs per Game: 7.2

Our Phillies Lineup

1. Kyle Schwarber | BRP: 0.89
2. Bryce Harper | BRP: 1.03
3. Brandon Marsh | BRP: 1.30
4. Jacob Realmuto | BRP: 0.79
5. Trea Turner | BRP: 1.3
6. Nick Castellanos | BRP: 0.72
7. Max Kepler | BRP: 0.70
8. Bryson Stott | BRP: 0.59
9. Alec Bohm | BRP: 0.51

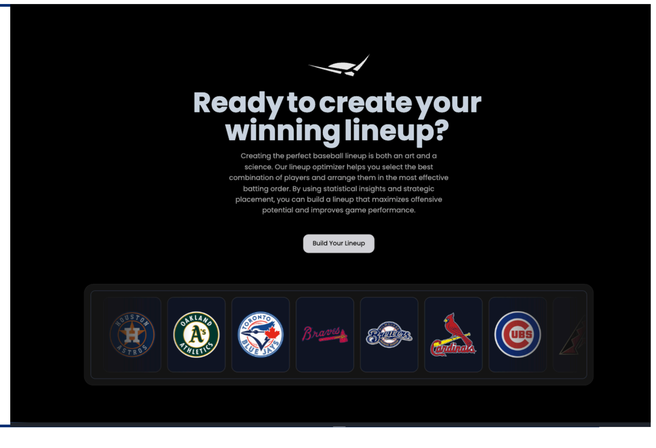
Expected Runs per Game: 7.8

Key Differences Between the Lineups

- Expected Runs per Game Increased by 8.3%
- Kyle Schwarber moved from 4th to 1st. Frontloading his power and walk rate maximizes his plate appearances.
- Bryce Harper moved from 3rd to 2nd. Higher up to take advantage of his elite BRP value.
- This highlights the strategic value of maximizing plate appearances for top performers and challenges conventional baseball lineup construction.

Key Features of the Website

- Expanded support to both MLB and the ALPB, enabling league-specific lineup optimization.
- Introduced lineup saving to allow users to revisit and compare strategies over time.
- Added constraint features including:
  - Fixing players to specific lineup positions
  - Limiting the number of consecutive same-handed batters



What’s Next

Next, we’re deploying the lineup optimizer as a fully functional web app for coaches, analysts, and front office staff to use in real time. We’ll continue refining the algorithm—incorporating more advanced metrics, team-level strategy inputs, and real-world feedback to improve accuracy and usability.

This is just the beginning. As the game evolves, so will our tools—bringing smarter, data driven decision-making to the dugout.