Does the Home Run Derby Affect a Batter's Swing?

Samantha Nadler

60-Day Baseball Analytics Challenge

Project #1

Overview

- **Myth:** Participating in the Home Run harms a batter's second-half performance.
- **Hypothesis:** There is no difference between a batter's pre-Derby batting performance and their post-Derby batting performance.

Methodology

1. Collect data from the 2024 MLB season.
2. Split data into pre-All-Star Break and post-All-Star Break (All-Star Break was July 15-18).
3. Determine variables needed for appropriate data analysis.
4. Devise a metric for calculating hitting performance between both halves of the season.
5. Calculate this metric on each of the eight 2024 Home Run Derby participants.
6. Visualize any other batting trends necessary for disproof of the "Home Run Derby curse" (e.g. at-bat event distributions, count distributions, at-bat lengths, swing data trends).

Data

Data Used: 2024 MLB Statcast data (provided by Baseball Savant)

Variables Used:

- `game_date`: Splits the season data into two halves (before July 15 and after July 18)
- `player_name` and `batter`: Identify the same player (player_name is a String, batter is a numerical ID)
- `events` and `bb_type`: One-hot encode the final outcome of an at-bat or batted ball
- `balls`, `strikes`, and `pitch_number`: Determine final ball and strike counts and determine plate discipline
- `launch_speed` and `launch_angle`: Determine kinematic shifts in swings on in-play balls

Created a new variable `count` that aggregates `balls` and `strikes`

Data

Eight players' performances were observed and analyzed:

- Pete Alonso, NYM
- Alec Bohm, PHI
- Adolis Garcia, TEX
- Gunnar Henderson, BAL
- Teoscar Hernandez, LAD
- Marcell Ozuna, ATL
- Jose Ramirez, CLE
- Bobby Witt Jr., KC

Post-Break Performance Index (PBPI)

Compares a player's first-half performance to their second-half performance

For pitchers:

$$PBPI_P = 100 \times \left(\frac{\Delta ERA \times w_{ERA} + \Delta FIP \times w_{FIP} + \Delta WHIP \times w_{WHIP}}{w_{ERA} + w_{FIP} + w_{WHIP}} \right)$$

where

$$\Delta ERA = \frac{ERA_0 - ERA}{ERA_0}$$

$$\Delta FIP = \frac{FIP_0 - FIP}{FIP_0}$$

$$\Delta WHIP = \frac{WHIP_0 - WHIP}{WHIP_0}$$

$$w_{ERA} = 0.8,\ w_{FIP} = 1.0,\ w_{WHIP} = 0.6$$

For position players:

$$_S\ PBPI_H = 100 \times \left(\frac{\Delta OPS \times w_{OPS} + \Delta wOBA \times w_{wOBA} + \Delta ISO \times w_{ISO}}{w_{OPS} + w_{wOBA} + w_{ISO}} \right)$$

where:

- $\Delta OPS = \frac{OPS_{After} - OPS_{Before}}{OPS_{Before}}$
- $\Delta wOBA = \frac{wOBA_{After} - wOBA_{Before}}{wOBA_{Before}}$
- $\Delta ISO = \frac{ISO_{After} - ISO_{Before}}{ISO_{Before}}$
- w_{OPS}, w_{wOBA}, and w_{ISO} are the weighting coefficients for each metric. These can be adjusted, but a suggested starting point is to give slightly more weight to the most comprehensive stats. For example, $w_{wOBA} = 1.0$, $w_{OPS} = 0.8$, and $w_{ISO} = 0.6$.

Results

Post-Break Performance Indices for each player:

PBPI for Pete Alonso:	5.69
PBPI for Alec Bohm:	-20.63
PBPI for Adolis Garcia:	-0.62
PBPI for Gunnar Henderson:	-21.07
PBPI for Teoscar Hernandez:	18.95
PBPI for Marcell Ozuna:	-15.02
PBPI for Jose Ramirez:	7.46
PBPI for Bobby Witt Jr.:	21.59
Overall PBPI for the above batters:	-0.98

Results

Alonso
K% increased 9.8% but 6MO% decreased 1.6%
BB% increased 4.6%

Bohm
K% increased 0.9% but 6MO% increased 5.8%
BB% decreased 2.4%

Garcia
K% increased 2.1% but 6MO% decreased 0.5%
BB% decreased 1.1%

Henderson
K% decreased 4.1% but 6MO% increased 0.5%
BB% increased 0.3%

Hernandez
K% increased 3.2% but 6MO% increased 1.5%
BB% decreased 0.1%

Ozuna
K% increased 2.9% but 6MO% decreased 0.5%
BB% decreased 0.5%

Ramirez
K% decreased 2.9% but 6MO% decreased 1.3%
BB% increased 1.3%

Witt
K% decreased 3.9% but 6MO% decreased 4.0%
BB% increased 3.1%

Six-Main-Out rate (6MO%): Calculates a player's rate of ending an at-bat in the following ways (denoted by `strikeout` and `field_out`):

- Striking out
- Grounding out
- Flying out
- Lining out
- Popping out
- Fouling out

Grounding into a double play does not count towards this metric, as the metric is independent of the current base path situation.

Results

Alonso
K% increased 9.8% but 6MO% decreased 1.6%
BB% increased 4.6%

Bohm
K% increased 0.9% but 6MO% increased 5.8%
BB% decreased 2.4%

Garcia
K% increased 2.1% but 6MO% decreased 0.5%
BB% decreased 1.1%

Henderson
K% decreased 4.1% but 6MO% increased 0.5%
BB% increased 0.3%

Hernandez
K% increased 3.2% but 6MO% increased 1.5%
BB% decreased 0.1%

Ozuna
K% increased 2.9% but 6MO% decreased 0.5%
BB% decreased 0.5%

Ramirez
K% decreased 2.9% but 6MO% decreased 1.3%
BB% increased 1.3%

Witt
K% decreased 3.9% but 6MO% decreased 4.0%
BB% increased 3.1%

Four batters' walk rate (BB%) increased:

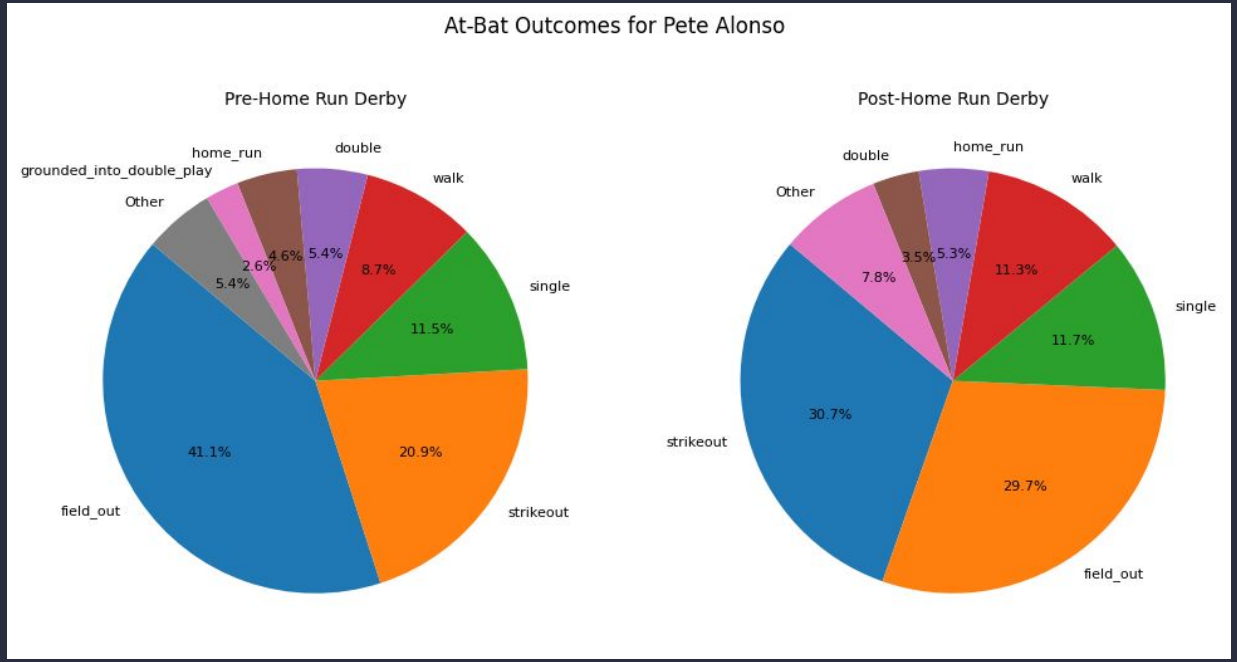
- Pete Alonso
- Gunnar Henderson
- Jose Ramirez
- Bobby Witt Jr.

Of the four, **two** (Ramirez and Witt) decreased both their strikeout rate (K%) and Six-Main-Out rate (6MO%).

Alec Bohm saw undesirable at-bat outcomes between halves, decreasing his walk rate and increasing his 6MO% to the highest degree.

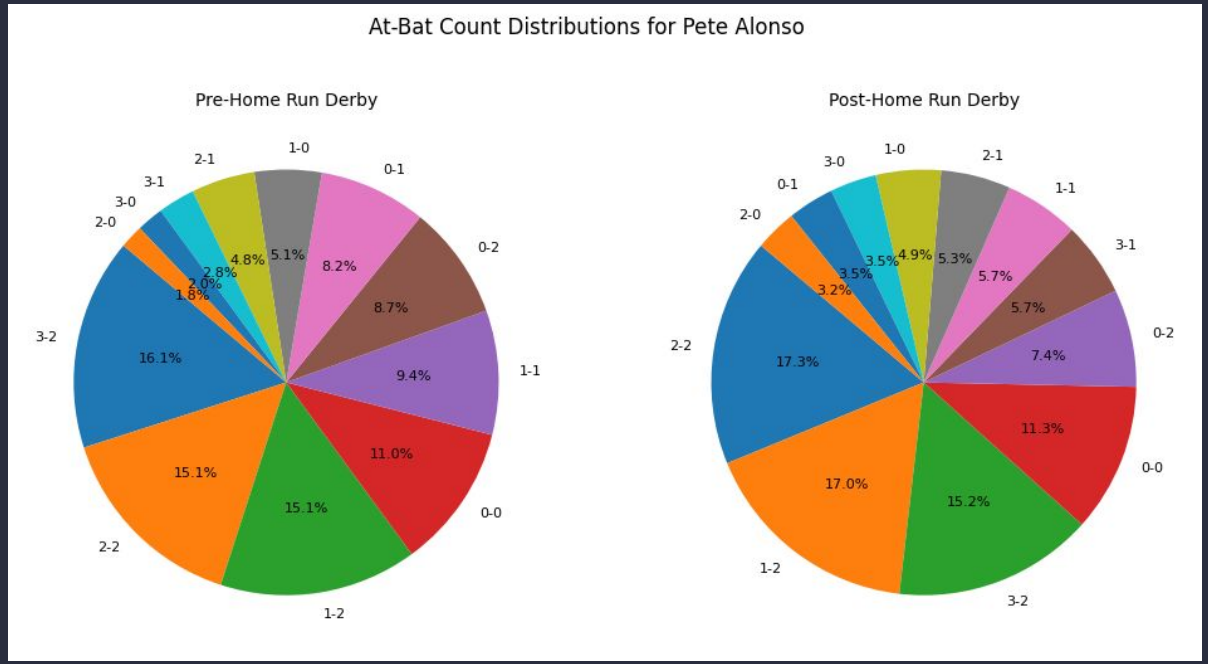
Results

Pete Alonso, New York Mets
At-Bat Events



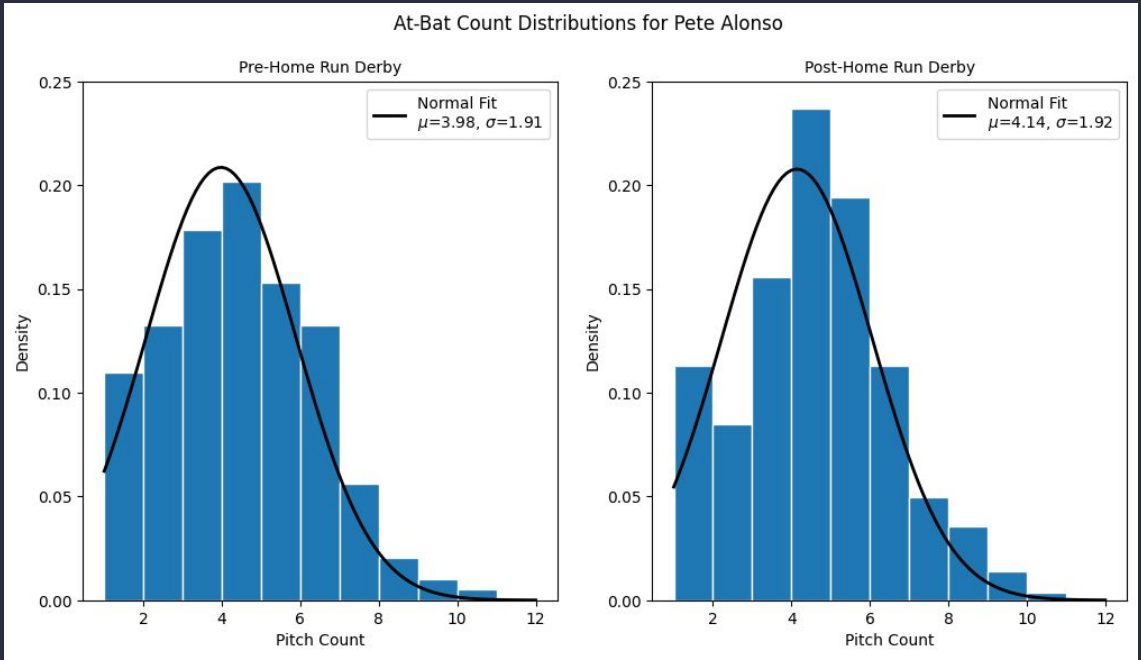
Results

Pete Alonso, New York Mets
Plate Discipline



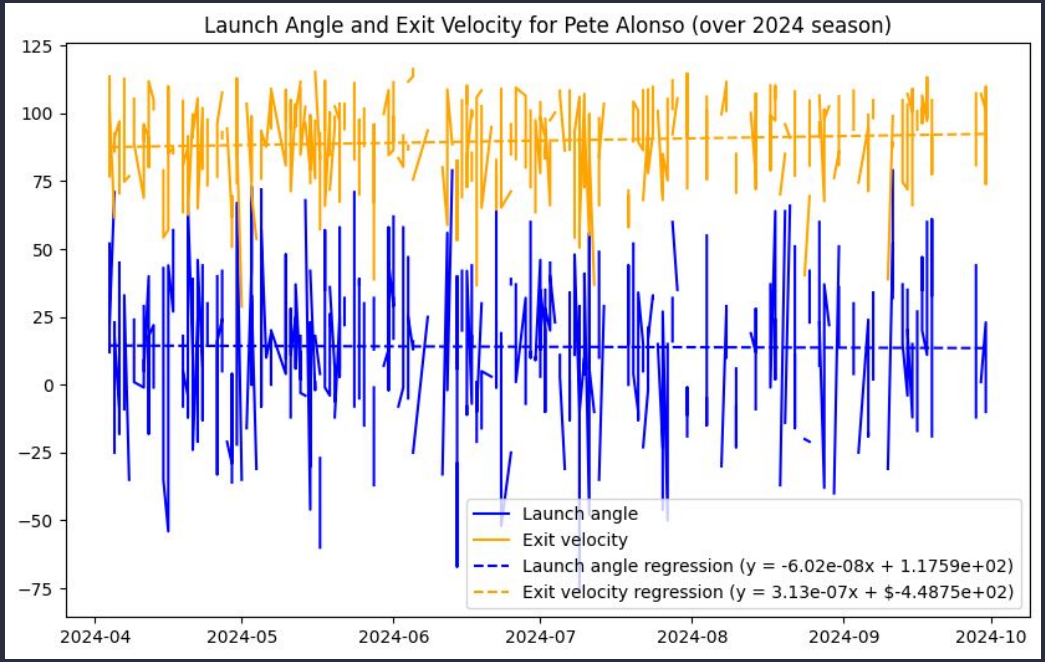
Results

Pete Alonso, New York Mets
Plate Discipline



Results

Pete Alonso, New York Mets
At-Bat Events



Discussion

So does the Home Run Derby affect a batter's swing?

- Collectively, no (as a group, performed slightly worse with no significant change)
- Only about half of participants' performance improved after the All-Star break
- Some batters (e.g. Bohm and Henderson) severely weakened after the All-Star break
- Others (e.g. Garcia) had almost no significant change

Some trends to consider:

- Strikeout rate decreased, walk rate increased for most batters
- At-bat length trended upwards (higher mean length with higher but similar deviation)
- Almost all batters increased exit velocity, decreased launch angle after the break (Bohm)

Some things to consider from the data:

- At-bat distributions between halves (more games before 2024 All-Star break)

The code used for this project has been made publicly available at:

https://github.com/sknadler/60-day-baseball-analytics-challenge