Noisy Judgments:

A probability surface-based analysis of umpiring variability

James Stockton, PhD
and
Emily-Anne Patt

SABR Analytics Conference

12 March 2023

James Stockton, PhD

- Lead Data Scientist, Altamira Technologies

- Currently supporting USAF Chief Data and AI Office

- Prior data science work supporting private industry and defence/intelligence clients
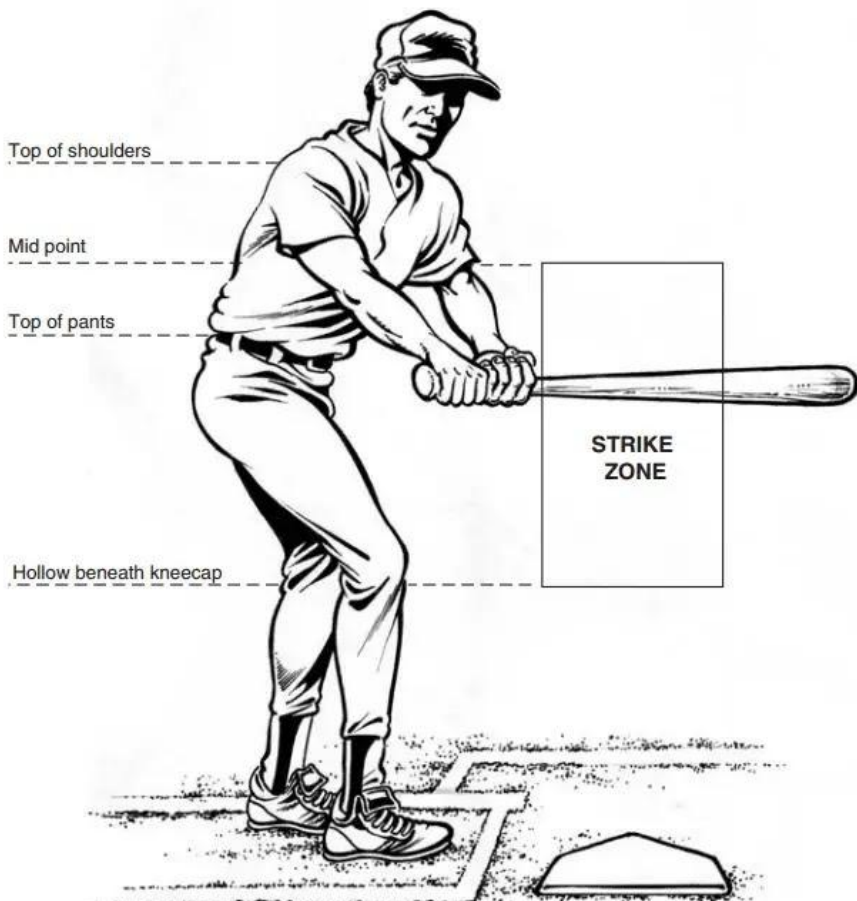
- Can't hit a curveball

Manager,
Quantitative Intelligence & Methodologies
Alphabet, Inc.

Emily-Anne
Patt

Coach at Washington Nationals
Youth Baseball Academy

Top of shoulders

Mid point

Top of pants

STRIKE ZONE

Hollow beneath kneecap
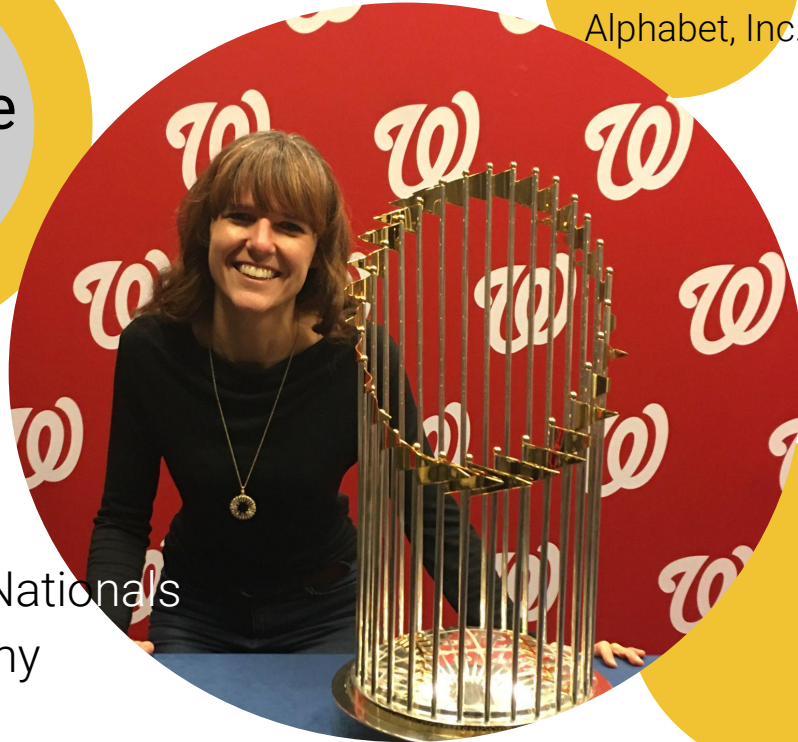
Motivation

- The strike zone is the fundamental measurement of baseball.

- A measurement is only as good as its ruler.

- The umpire is the ruler, not the MLB rule book.
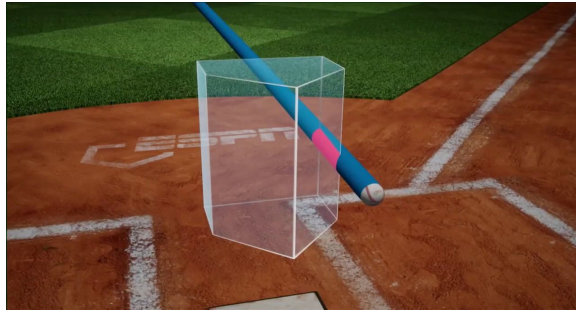
- So...how good is that ruler?

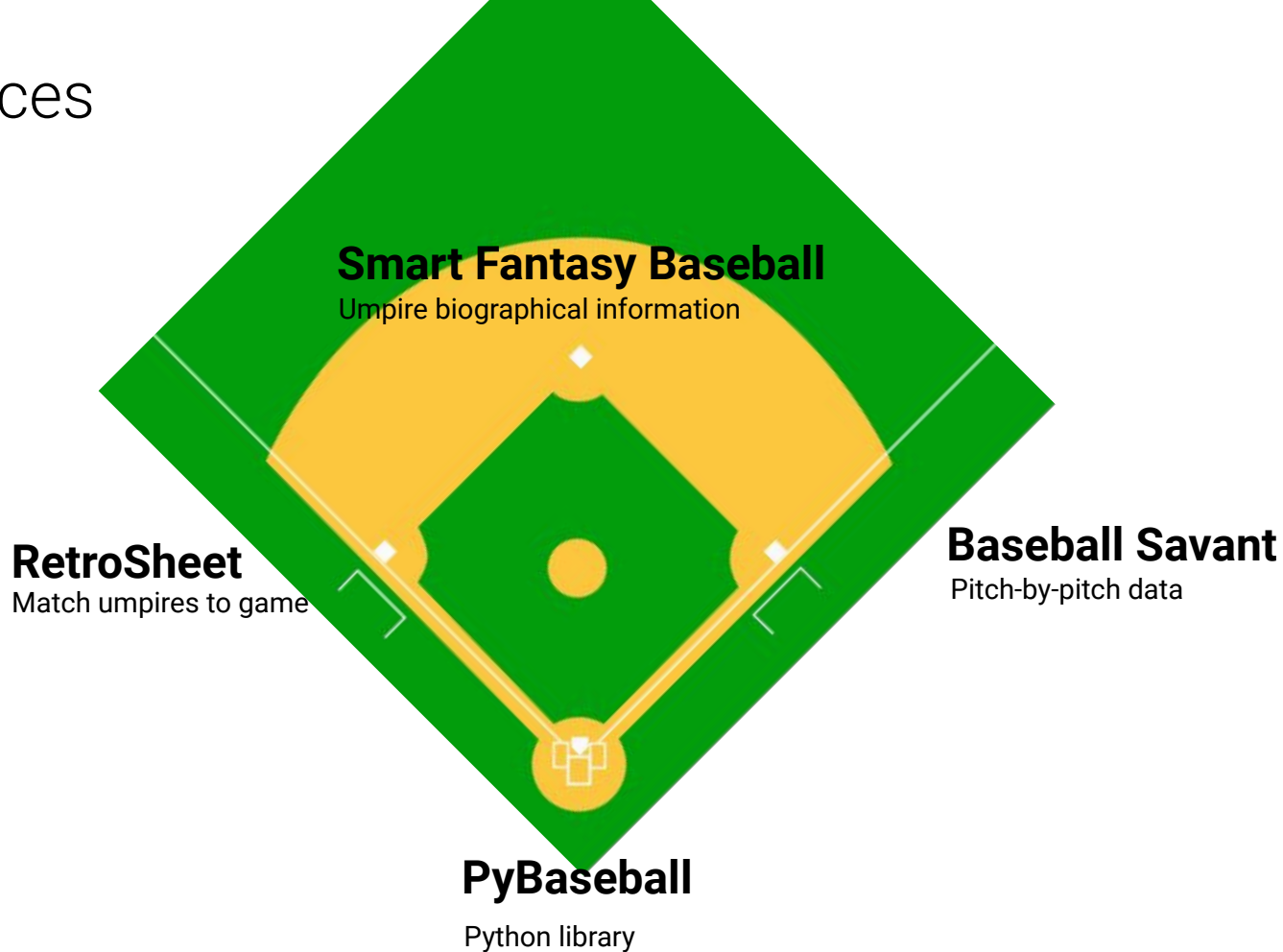# Analytical Outline

1. Collect data on called balls and strikes.

2. Use the data to model the strike zone as it's actually called.

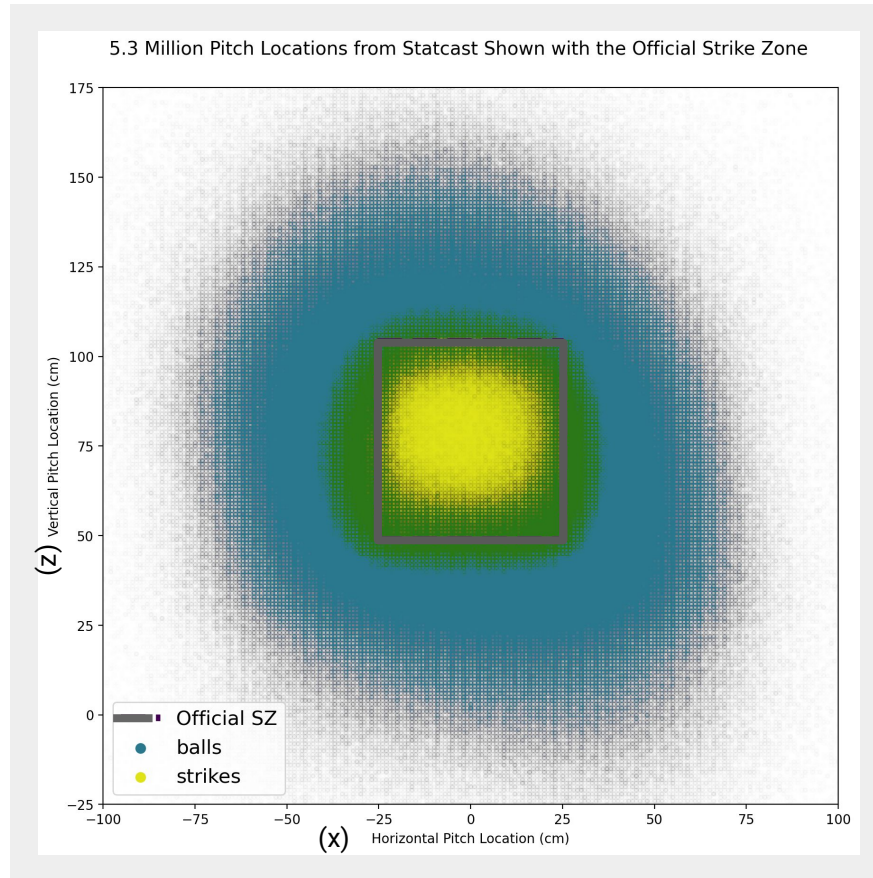3. Explore how that model changes across subsets of the data.

Data Sources

Smart Fantasy Baseball
Umpire biographical information

RetroSheet
Match umpires to game

Baseball Savant
Pitch-by-pitch data

PyBaseball
Python library

Combined Dataframe

x	z	strikezone(z)$_{top}$	strikezone(z)$_{bottom}$	Umpire	Pitcher	Pitcher handedness	Batter	Batter handedness
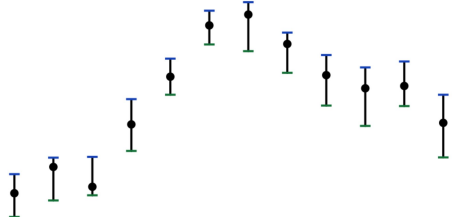
Pitch location for 5.3 million called pitches

5.3 Million Pitch Locations from Statcast Shown with the Official Strike Zone

Building a Strike Zone Probability Surface: Step 1

5.3 Million Pitch Locations from Statcast Shown with the Official Strike Zone

71%	63%	54%	47%
72%	65%	55%	43%
78%	65%	57%	47%
76%	64%	55%	41%

1 cm

Note: Reported pitch location accuracy and precision changes over time with evolution of pitch tracking technology.

Modeling

The Useful Kind of Errors

- The number of pitches falling into each grid cell varies.

- The strike percentage in each cell is measured with a different level of confidence.

  - Sampling and survey statistics when measuring proportions

- The pitches in the cell vote on the likelihood a new pitch will be a ball or a strike.

- Accounting for measurement errors in the fitting process improves results.

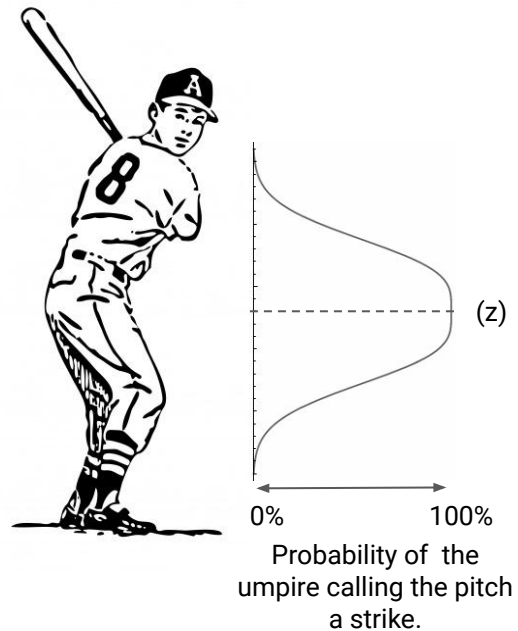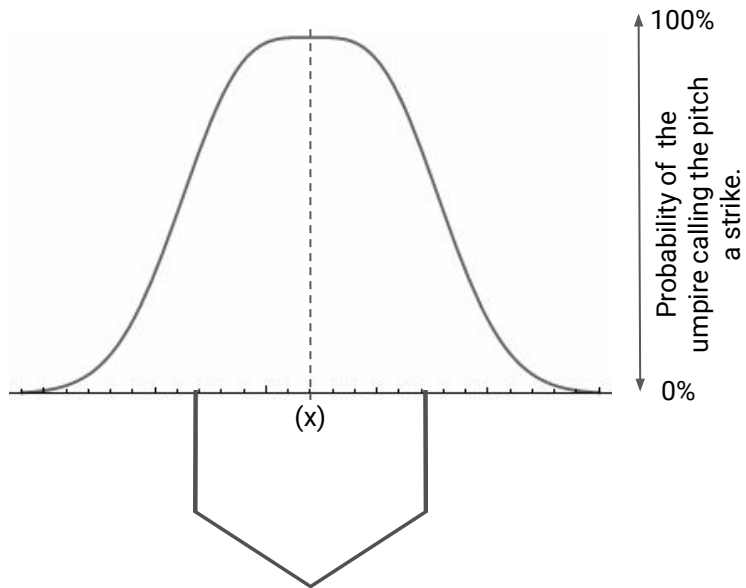$$\hat{p} \pm a^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

E9

Image credit: Houston Chronicle
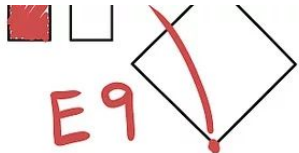
Height Normalization

- Used StatCast reporting of the sz_{top} and sz_{bottom} for every pitch

- Averaged (mean) values for each defines a normalized strike zone height

- Proportionally scaled each pitch location relative to this standardized size

Modeling ⚾⚾⚾〇〇〇〇
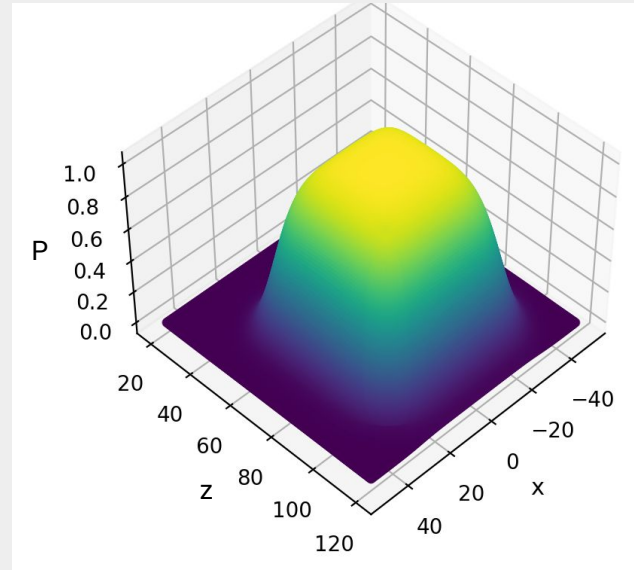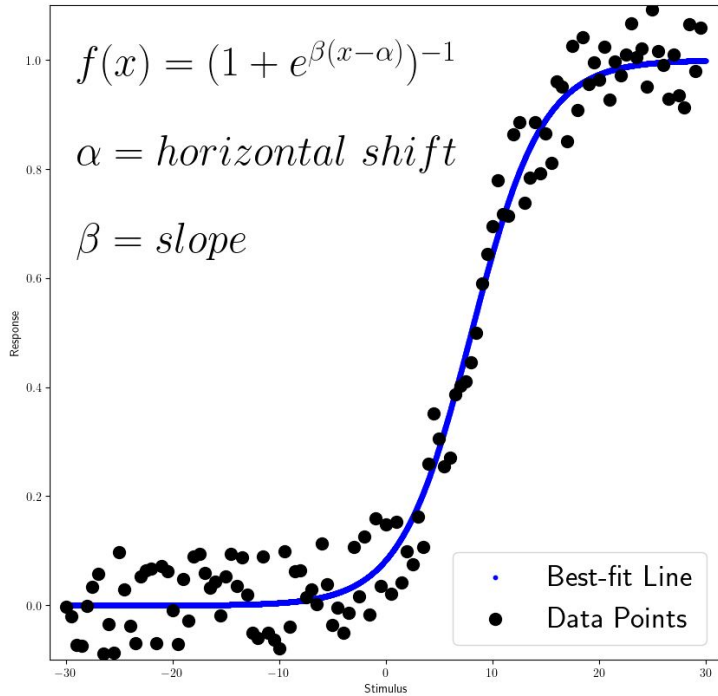
The Anatomy of a Strike Zone Sigmoid

Takeaway: Each probability distribution is made up of two *sigmoids*, curving from the middle of the strike zone where the pitch will nearly always be called a strike to the edge of the strike zone where the strike calls are rare.

Modeling

Building a Strike Zone Probability Surface: Step 4
The Anatomy of a Strike Zone Sigmoid

Example Sigmoidal Fit to Data Points

$$f(x) = (1 + e^{\beta(x-\alpha)})^{-1}$$

$$\alpha = horizontal\ shift$$

$$\beta = slope$$

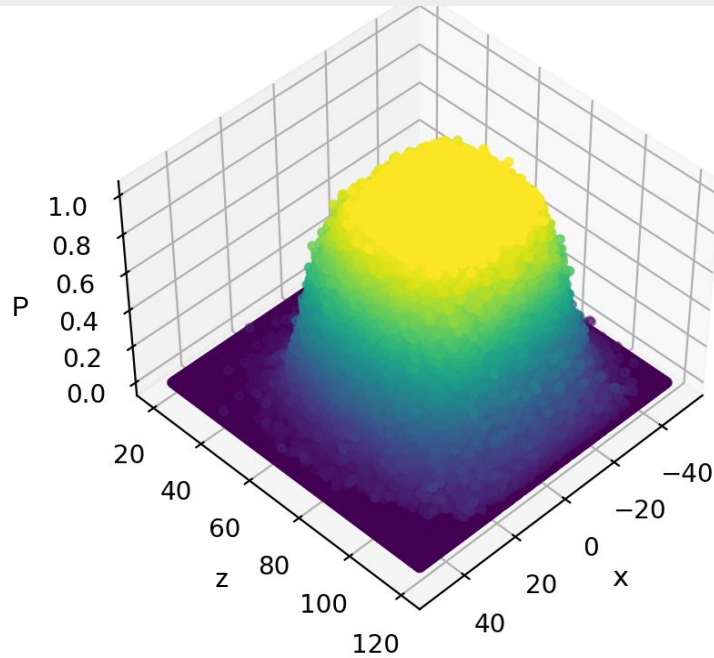Legend:
- Best-fit Line
- Data Points

$$F(x, y) = \frac{(1 + e^{-\beta_0(x-\alpha_0)})^{-1}}{1 + e^{-\beta_1(x-\alpha_1)}} * \frac{(1 + e^{-\beta_2(x-\alpha_2)})^{-1}}{1 + e^{-\beta_3(x-\alpha_3)}}$$

Modeling

Building a Strike Zone Probability Surface: Step 4
The Anatomy of a Strike Zone Sigmoid

Modeling

Example Sigmoidal Fit to Data Points

$f(x) = (1 + e^{\beta(x-\alpha)})^{-1}$

$\alpha = horizontal\ shift$

$\beta = slope$

Best-fit Line
Data Points

$$F(x,y) = \frac{(1 + e^{-\beta_0(x-\alpha_0)})^{-1}}{1 + e^{-\beta_1(x-\alpha_1)}} * \frac{(1 + e^{-\beta_2(x-\alpha_2)})^{-1}}{1 + e^{-\beta_3(x-\alpha_3)}}$$
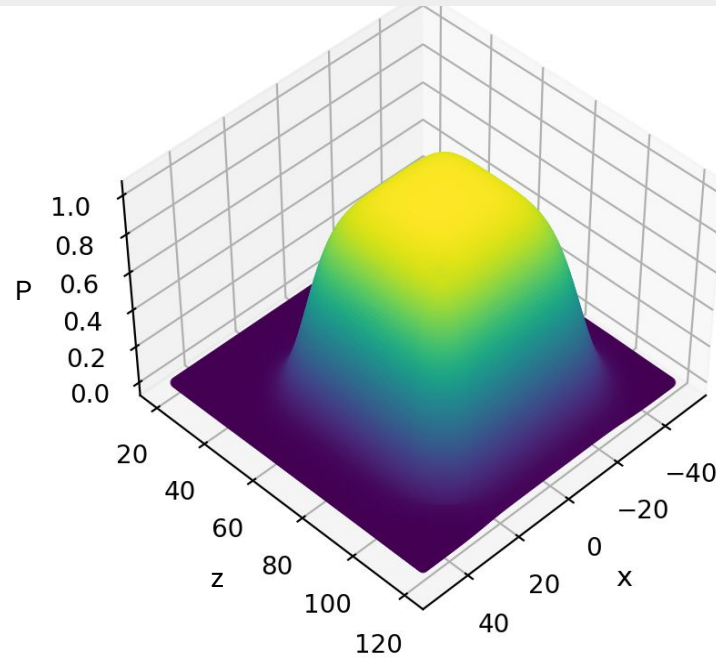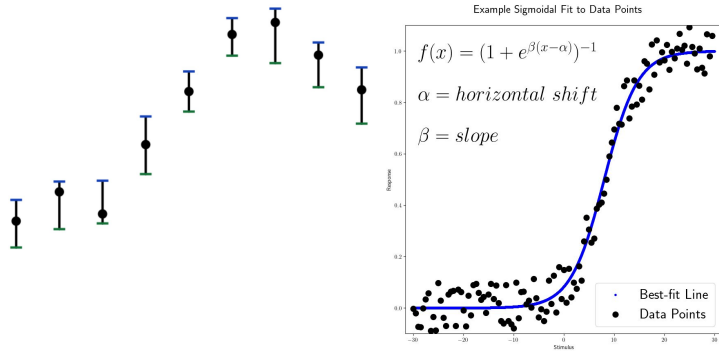
Best-fit Strike Zone and Associated Probability Surface for All Available Data in the 2008-2022 Seasons
The 50% Probability Contour Defines the Strike Zone and is Shown with its Dimensions

56.1 cm (22.1 in)

— 57.7 cm (22.7 in) —

Vertical Pitch Location (cm)

Horizontal Pitch Location (cm)

Strike Probability

Key Benefits of the Method

1. Accounts for sampling errors in the data

2. Well-tuned underlying functional form matches the shape of the data

3. The fitting process smooths out noisy data using only a handful of free-parameters
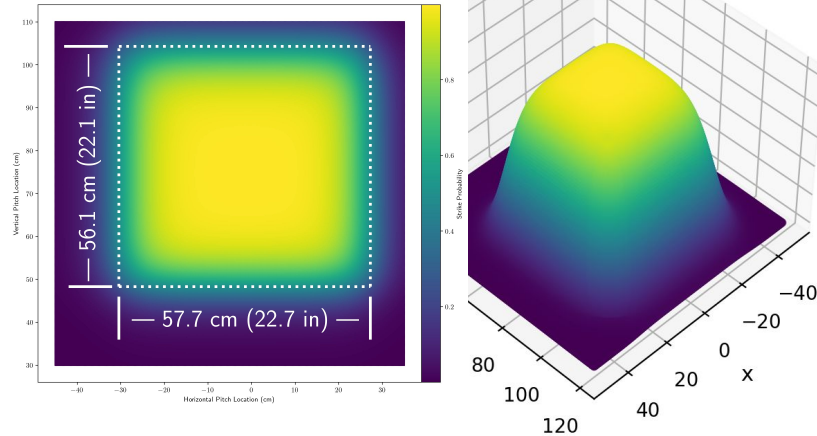
Takeaway: The results are reliable to smaller sample sizes than prior efforts, which opens up access to questions about specific matchups and changes over time.

Modeling

The Umpires' Strike Zone

- The best-fit strike zone for all umpires across all seasons in the data set is 57.7 cm along the horizontal axis and 56.1 cm tall.

- The best-fit strike zone is shifted 1.6 cm to the left of the center of the plate.
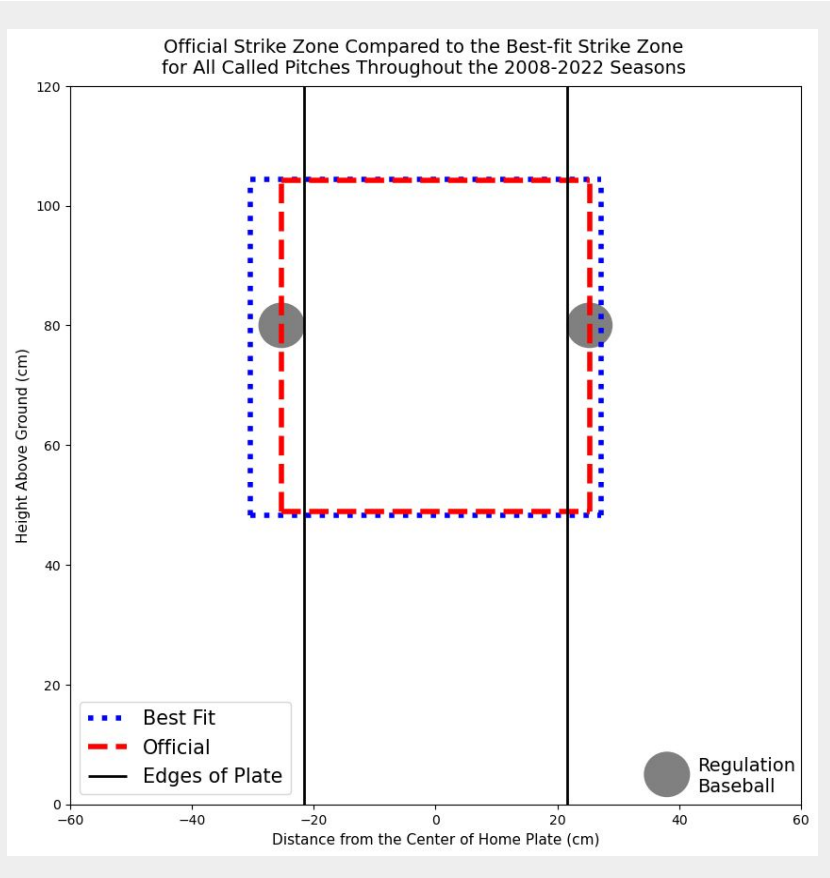

Best-fit Strike Zone and Associated Probability Surface for All Available Data in the 2008-2022 Seasons
The 50% Probability Contour Defines the Strike Zone and is Shown with its Dimensions

Validating the Ruler
The Umpires' Best-Fit Strike Zone Compared to the Official Zone

Official Strike Zone Compared to the Best-fit Strike Zone
for All Called Pitches Throughout the 2008-2022 Seasons

- **The best-fit strike zone** (blue dotted box) **is 7.1 cm wider than the official strike zone** (red dashed box).
  - Red dashed box passes through the center of a baseball that is tangent to home plate.

- **The best-fit strike zone is shifted 1.6 cm to the left of the center of the plate.**
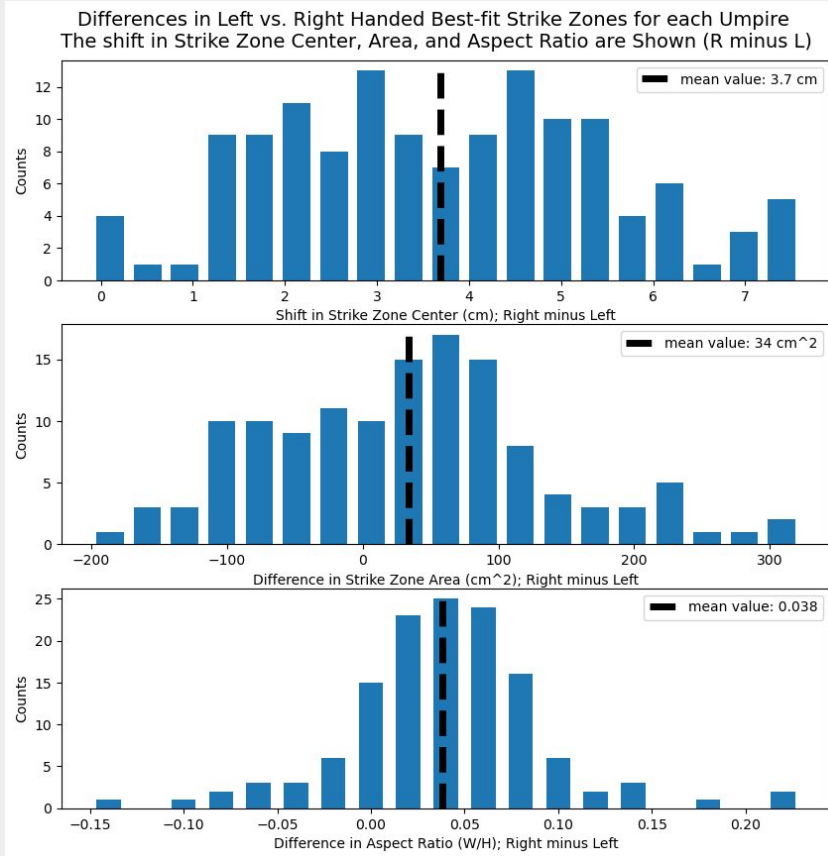
Validating the Ruler
The Umpires' Best-Fit Strike Zone Compared to the Official Zone

Best Fit Strike Zone Across All Umpires
for Left vs. Right-handed Batters
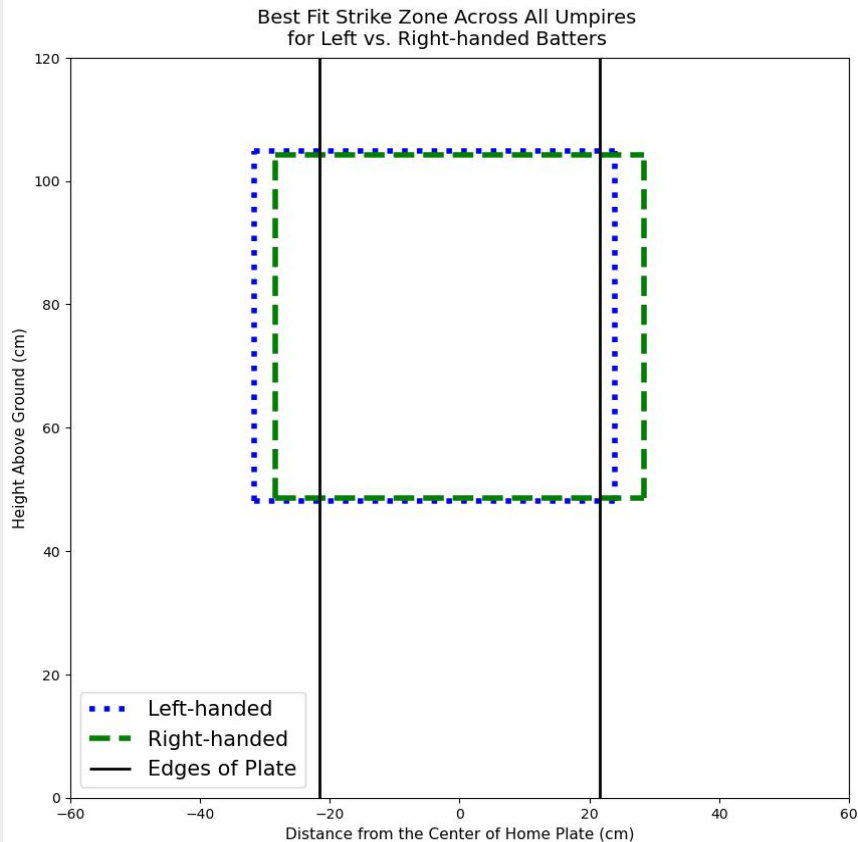
- The best-fit strike zone for right-handed batters (green dashed box) and left-handed batters (blue dotted box) have nearly identical areas

- The right-handed strike zone is centered roughly over the plate, while the left-handed strike zone is shifted 3.9 cm to the outside of the plate
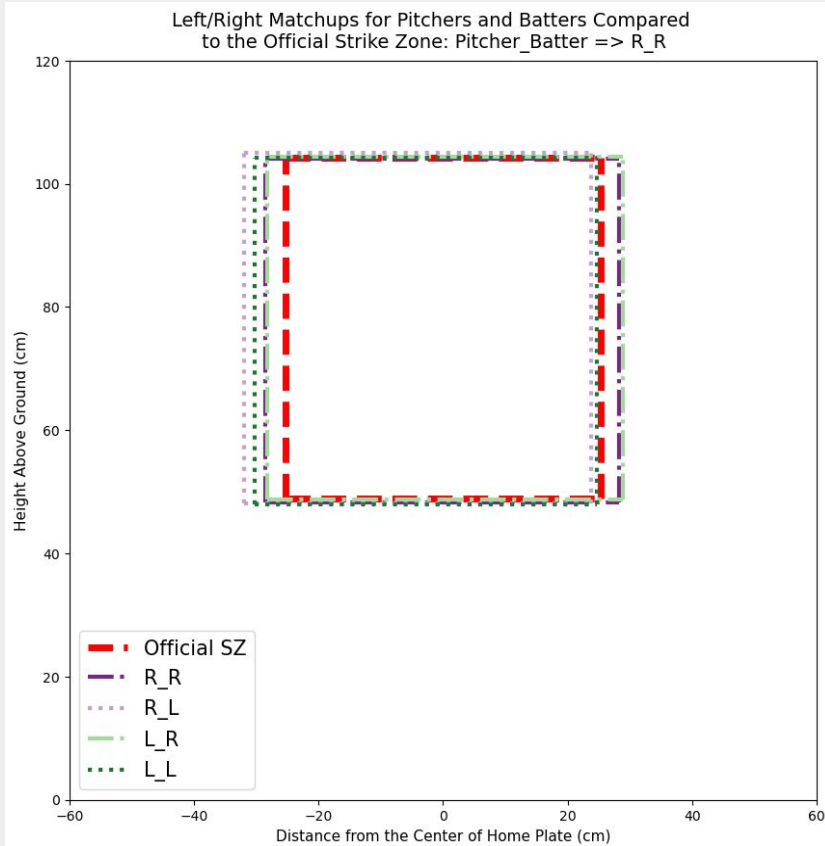
Results

The Lefty Strike
The Presence of the Shift is Uniform across the Umpire Corps

Differences in Left vs. Right Handed Best-fit Strike Zones for each Umpire
The shift in Strike Zone Center, Area, and Aspect Ratio are Shown (R minus L)

- [Top] **The shift in the strike zone for left-handed batters holds true across the umpire corps, with the shift measuring between 1 and 7 cm off-center.**

- [Middle] **Centered roughly on zero, the histogram shows the area of umpires' strike zones generally does not shift between right- and left-handed batters.**

- [Bottom] **We see a clear difference in the aspect ratio, with the lefty strike zone taller and narrower than the right-handed batters' zone.**
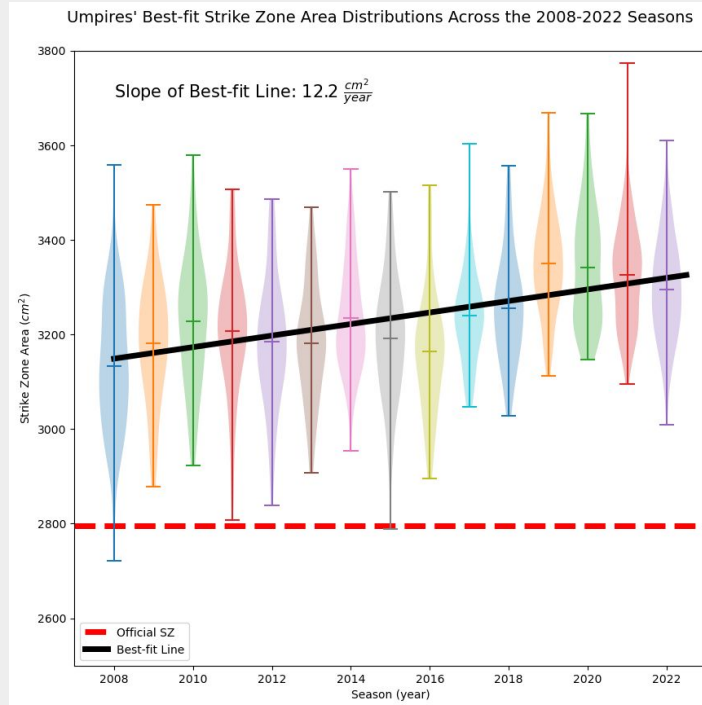
The Lefty Strike
The Lefty Shift More Pronounced for Right-Handed Pitchers

Left/Right Matchups for Pitchers and Batters Compared to the Official Strike Zone: Pitcher_Batter => R_R
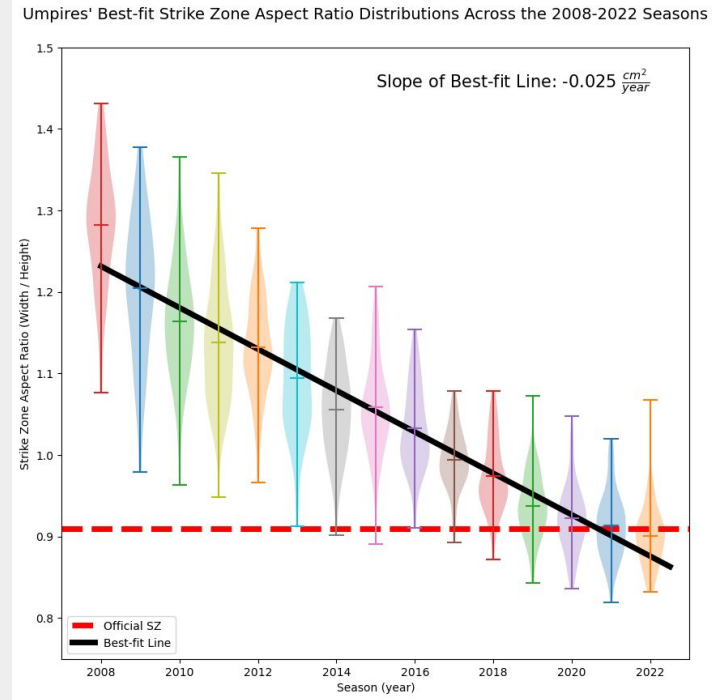
- The lefty-strike is even more pronounced when there is a RHP (purple dotted rectangle).

- The magnitude of the strike zone shift between left and right-handed batters for right-handed pitchers (purple) is roughly double the shift for left-handed pitchers (green).

Results

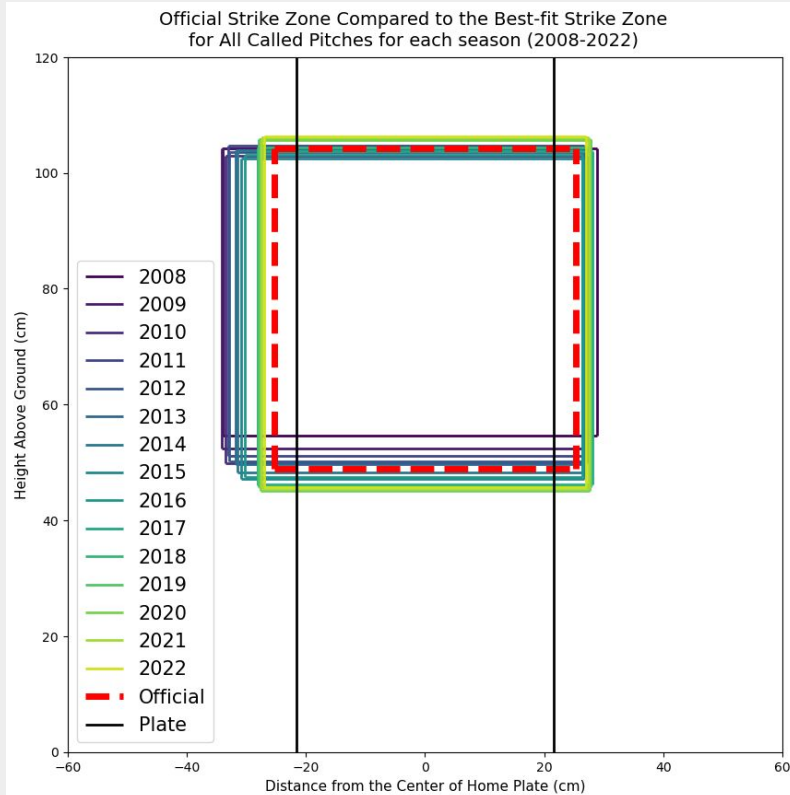# Seasonal Shifts

Umpires' Best-fit Strike Zone Area Distributions Across the 2008-2022 Seasons

Slope of Best-fit Line: 12.2 $\frac{cm^2}{year}$

Takeaway: The area of the strike zone has increased gradually since the 2008 season.

Umpires' Best-fit Strike Zone Aspect Ratio Distributions Across the 2008-2022 Seasons

Slope of Best-fit Line: -0.025 $\frac{cm^2}{year}$

Takeaway: The strike zone has become narrower and taller since the 2008 season.

Results

Seasonal Shifts

Official Strike Zone Compared to the Best-fit Strike Zone for All Called Pitches for each season (2008-2022)

- The strike zone narrowed nearly 14%, from 63 to 54 cm between the 2008 and 2022 seasons.

- The left side of the strike zone moved towards the plate about 7 cm.

- The upper limit of the strike zone has remained relatively consistent across the umpire corps since 2008, rising by 2 cm (2%) during that period.

- The bottom of the strike zone, however, dropped by nearly 9 cm—more than 16%—between 2008 and 2022.
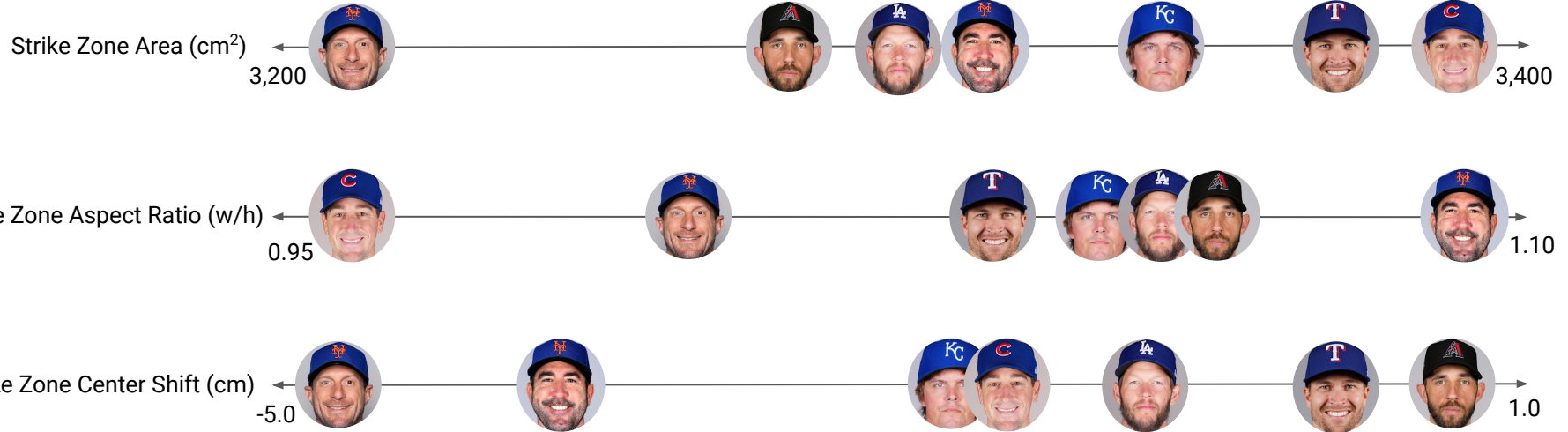
Around the League: Umpire Accuracy

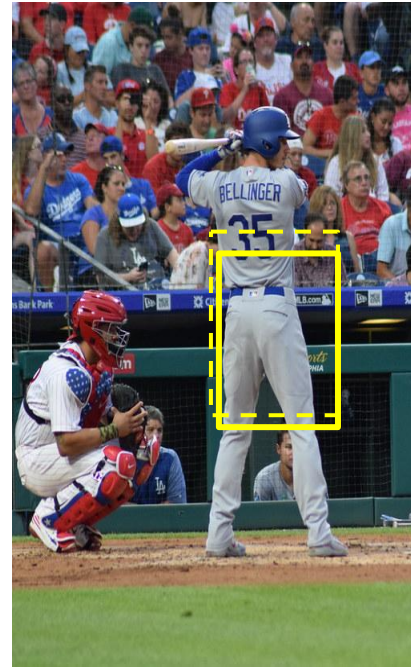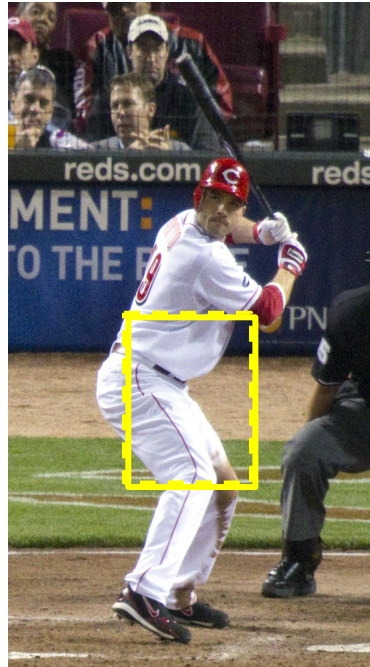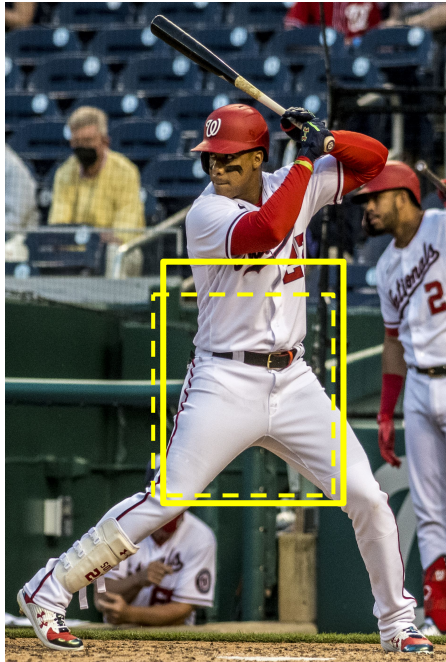$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}}$$

The IoU ratio for the umpire corps ranges from 0.70 to 0.90, with the majority of umpires clustering around the 0.85 mark.

Intersection Over Union of Best-fit Strike Zones and the Official SZ for Umpires in the 2008-2022 Seasons
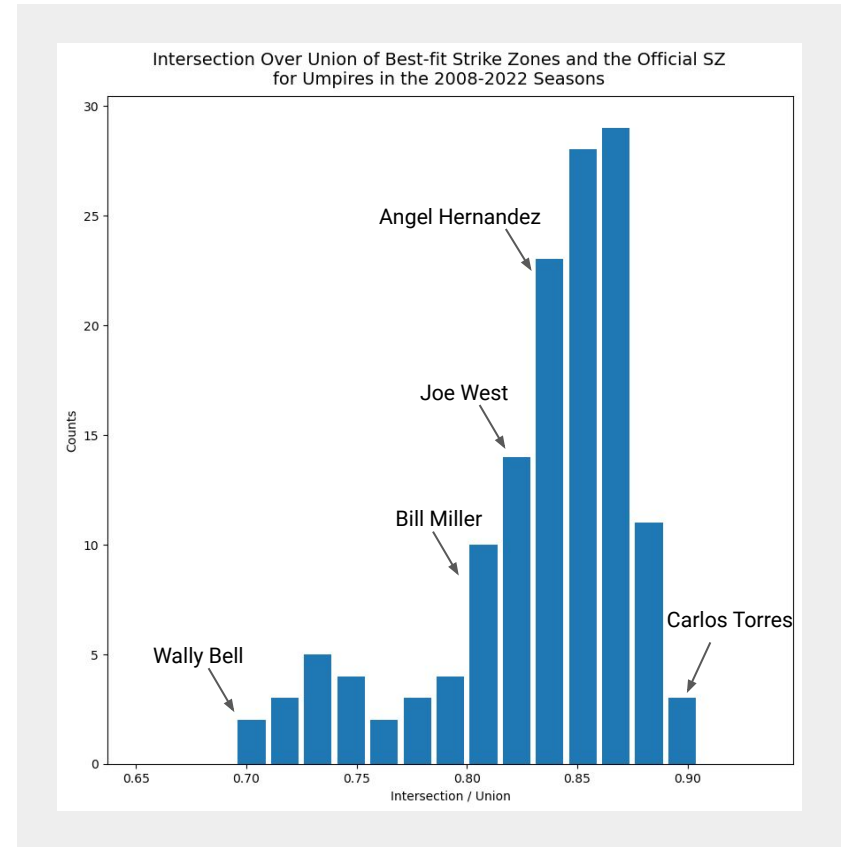
Around the League: Elite Pitchers

Around the League: Notable Hitters

Best Fit Strike Zone   MLB Strike Zone

Discussion

What's next?

What insights emerge when we split the data 2+ ways? Add pitch selection? Swing rate?

How do MLB's 2023 rule changes affect the shape of the strike zone?

How have pitchers and batters adjusted to the strike zone's changing dimensions over time?

How can the model help us measure a catcher's framing performance?