Weapons of Best Production: Predicting the Optimal Pitch Arsenal Adjustment for Superior Stuff+

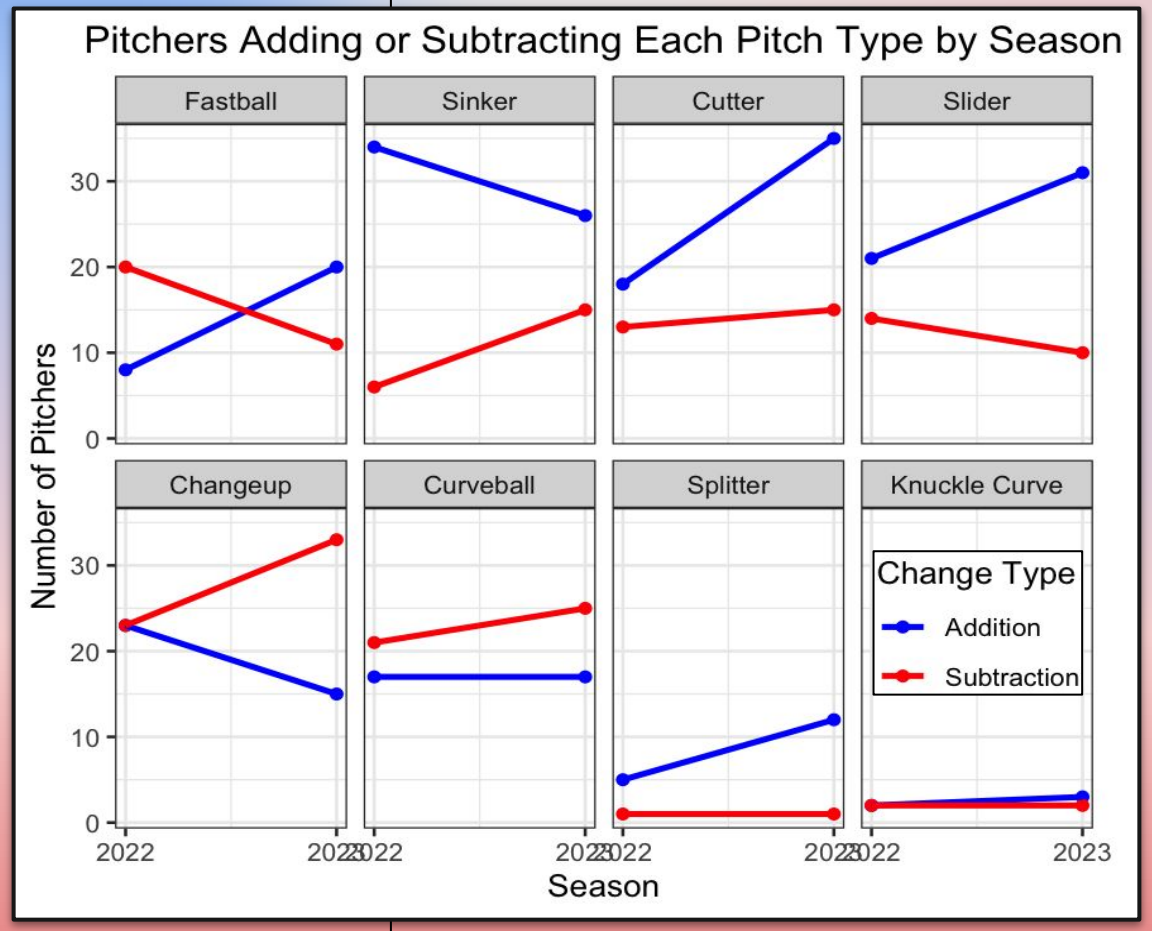
Gabriel Eze, Neha Kotha, Danny Nolan
Advisor: Dr. Sam Fleischer, Los Angeles Dodgers

Motivation

- One way pitchers attempt to improve performance is by changing their **pitching arsenal**, i.e., the types of pitches they throw.

Main Question:

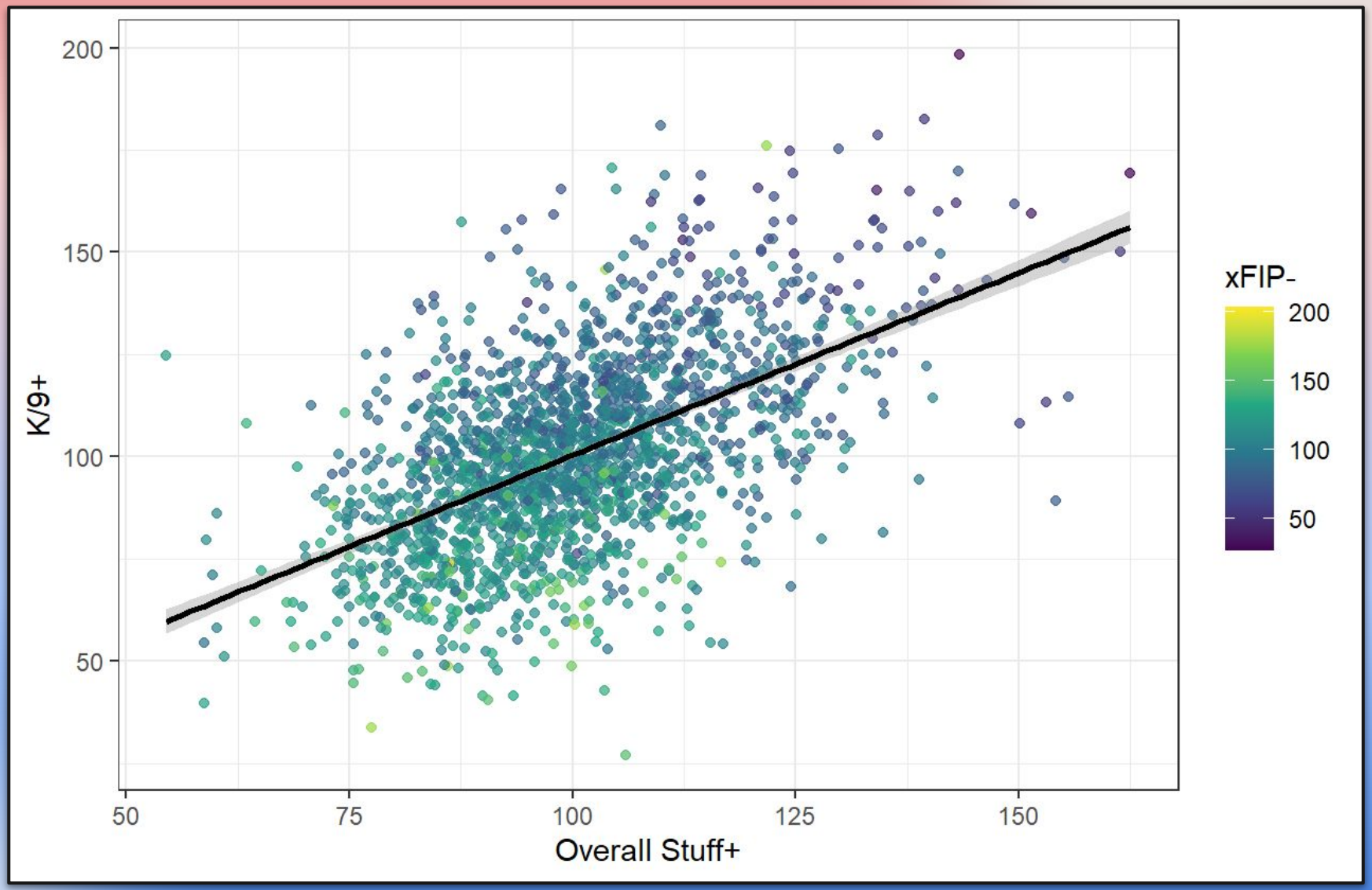
- Can we create a pitch recommendation system that suggests, with conviction, which pitch a player should add and how effective it might be?



Data

- MLB Seasons: 2021, 2022, 2023
- Pitch data from **FanGraphs**
- Spin rate and release point data from **Baseball Savant**
- 5% pitch usage cutoff to eliminate any misclassification or one-off occurrences
- **Stuff+** accounts for only the physical traits of a pitch:
    - Velocity
    - Horizontal and Vertical Movement
    - Spin Rate
    - Release Point

Stuff+ is a Useful Predictor for Performance



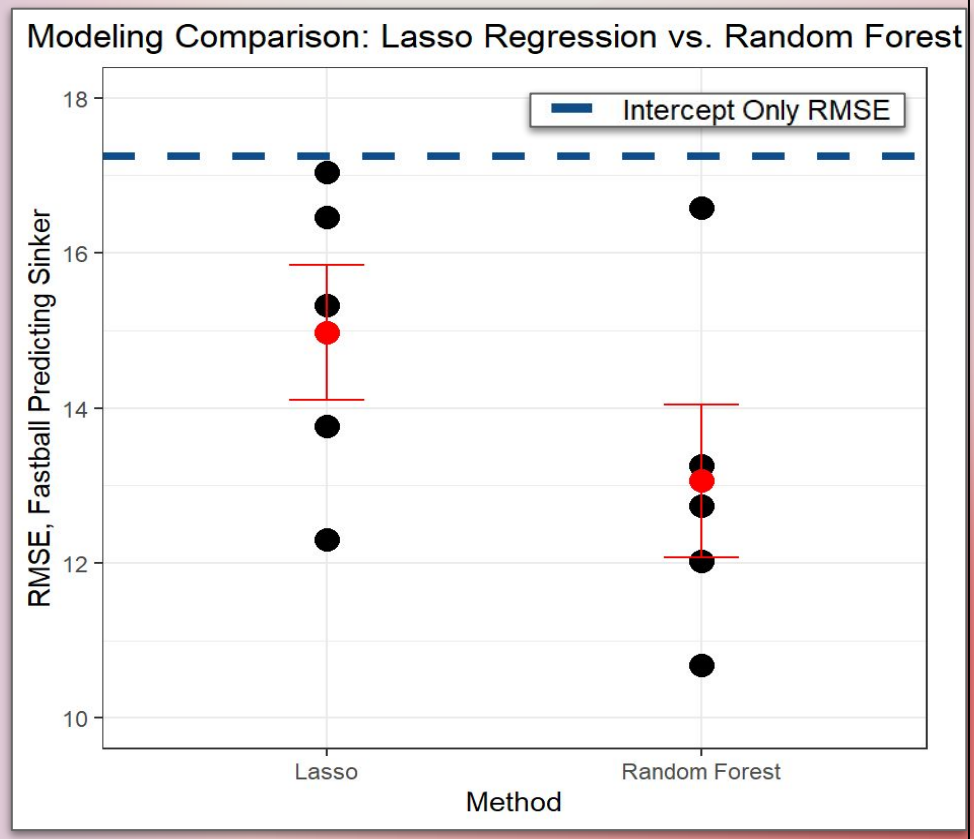
Modeling

Approach:

- Used pitch characteristics of other pitches to model Stuff+ values for a certain pitch
- Decided to model the relationship between **pairs of pitches**
- Began by using Fastball traits to predict Sinker Stuff+

Methods:

1. **Lasso Regression**
2. **Random Forest Regression**

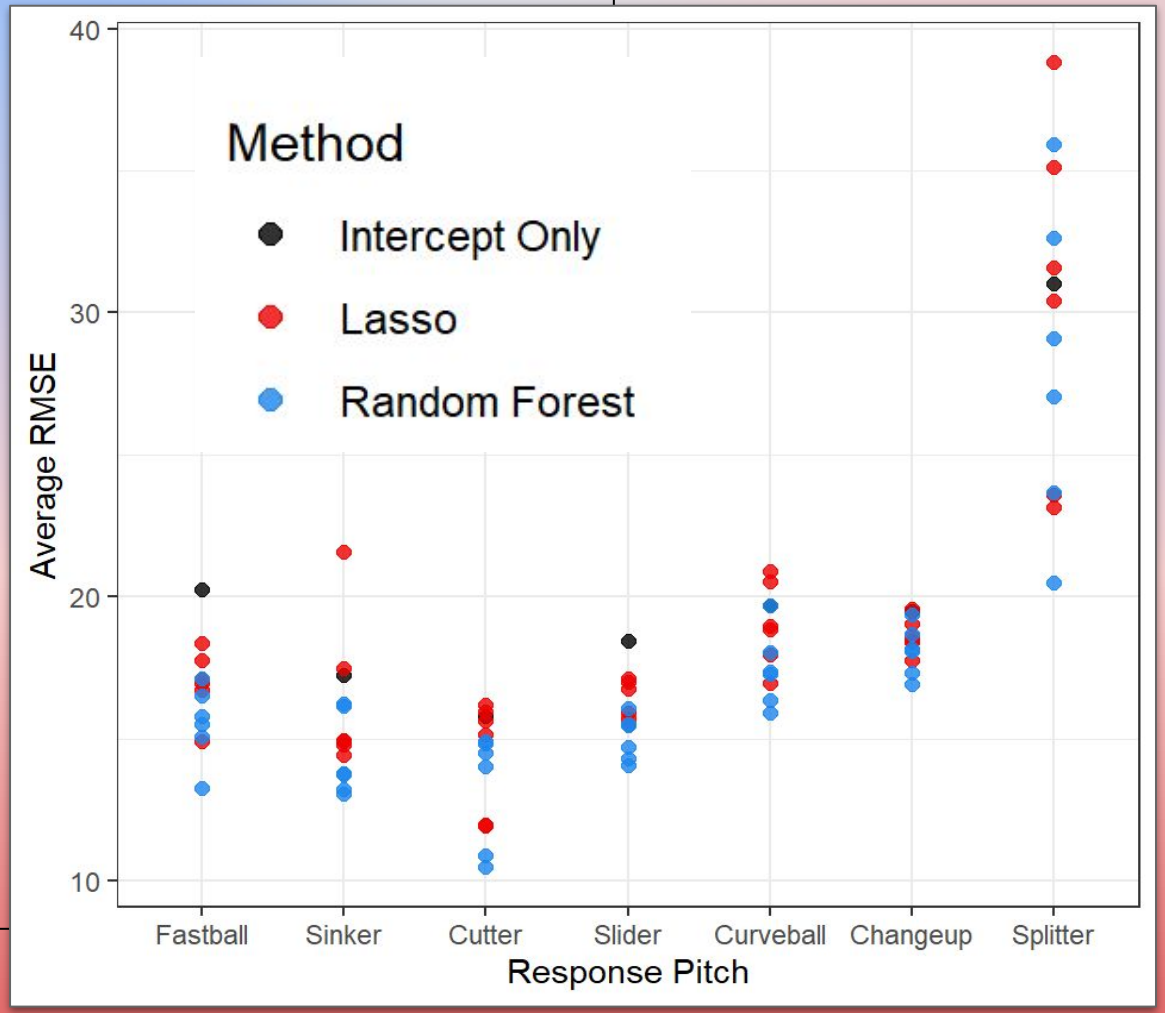


Used **5-fold cross-validation** to compute average RMSEs for each pair of pitches (red points)

Results

Which Model is the Winner?

- The random forest model beats the basic intercept-only model and generally clears the lasso model.
- For each pitch type, a random forest process produces the lowest average RMSE.
- Tells us that the interactions among the predictor variables are significant.



Applications

- We took the pairing yielding the smallest RMSE and built predictions for each pitch.
- Players can use these predictions to add an above-average pitch to their arsenal.

Real-life stories:

Ryan Pressly

Nabil Crismatt

Story 1: Ryan Pressly

- Our model suggests that Ryan Pressly would have had a great changeup in 2021 and 2022.
- He added one in 2023 and it had a Stuff+ of 136, among the best in baseball.

Predicting Changeup Stuff+ using Curveball Traits

Top Five Changeup Stuff+	Bottom Five Changeup Stuff+
1. Ryan Pressly (2021, 2022): 109.94, 109.93	1. Ross Detwiler (2021): 67.63
2. Justin Verlander (2023, 2022): 107.52, 106.29	2. Grant Dayton (2021): 69.01
3. Aaron Civale (2023, 2022): 103.61, 100.54	3. Bryse Wilson (2023): 72.59
4. Tyler Glasnow (2023, 2021): 102.98, 100.94	4. Robert Dugger (2021): 73.12
5. Keegan Thompson (2023): 100.79	5. Jimmy Herget (2021): 73.2

Story 2: Nabil Crismatt

- Nabil Crismatt's fastball in 2021 and 2022 suggested that he would have a poor slider if he decided to throw one.
- He added the pitch in 2023 and it had a Stuff+ of 69, well below the average slider value.

Predicting Slider Stuff+ using Fastball Traits

Top Five Slider Stuff+	Bottom Five Slider Stuff+
1. Robert Suarez (2022): 129.35	1. Nabil Crismatt (2021, 2022): 85.01, 86.32
2. Joe Kelly (2022): 129.21	2. Cody Bradford (2023): 88.93
3. Thomas Hatch (2023): 128.95	3. Jared Koenig (2022): 89.12
4. Oliver Ortega (2022): 128.87	4. Hyun-Jin Ryu (2021): 90.64
5. Felix Bautista (2023): 128.69	5. Brent Suter (2022): 91.96

Conclusions

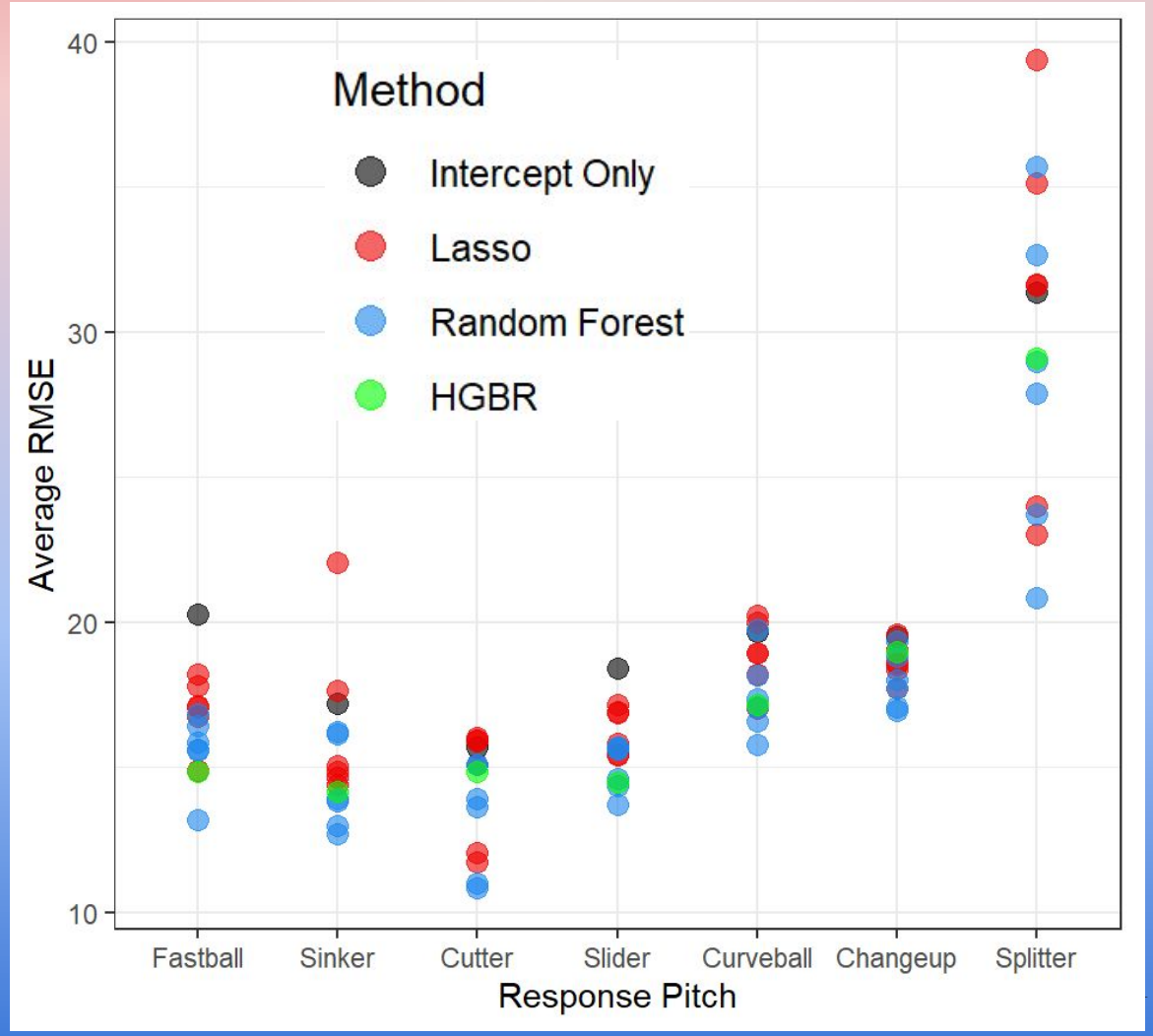
- Constructed relatively reliable Stuff+ predictions that teams and players can utilize for a competitive advantage.
- Discovered which pitches are most useful for predicting others.
- Indicated important interactions between pitch traits through the model selection process.

Limitations

- Predictions do not take into account a pitcher's whole arsenal.
- Inconsistencies between data sources in lumping certain pitch types together.
- Location and pitch sequencing also play roles in a pitch's success.

Predicting Stuff+ via Multiple Pitches

- **Method: Histogram-Based Gradient Boosting Regression Tree (HGBR):**
  - A tree-based regression tool that has native support for missing values.
  - Advantage: Allows us to incorporate all pitches into the predictive model.
  - Disadvantage: Higher RMSE values than the random forest.



Future Work

- Fix the discrepancies in pitch classification between FanGraphs and Baseball Savant
- Examine how a pitch works in conjunction with other offerings with a statistic such as Runs Above Average
- Adjust for factors outside of pure stuff, such as location and pitch sequencing

Questions?

Scan the QR code to access our Shiny app!

Works Cited

Acquavella, Katherine. *"MLB Study Says Balls Weren't Intentionally Juiced in 2019; Home Run Spike Credited to Seams, Launch Angles."* CBSSports.com, December 11, 2019. https://www.cbssports.com/mlb/news/mlb-study-says-balls-werent-intentionally-juiced-in-2019-home-run-spike-credited-to-seams-launch-angles/.

baseballsavant.com. *"Statcast Pitch Arsenals Leaderboard."* Accessed July 25, 2024. https://baseballsavant.mlb.com/leaderboard/pitch-arsenals?year=2024&min=250&type=avg_spin&hand=.

FanGraphs Baseball. *"Major League Leaderboards - 2023 - Pitching."* Accessed July 25, 2024. https://www.fangraphs.com/leaders/major-league?pos=all&stats=pit&lg=all&qual=y&type=14&season=2023&month=0&season1=2023&ind=0.

McGrattan, Owen. *"Stuff+, Location+, and Pitching+ Primer."* Sabermetrics Library, March 10, 2023. https://library.fangraphs.com/pitching/stuff-location-and-pitching-primer/.