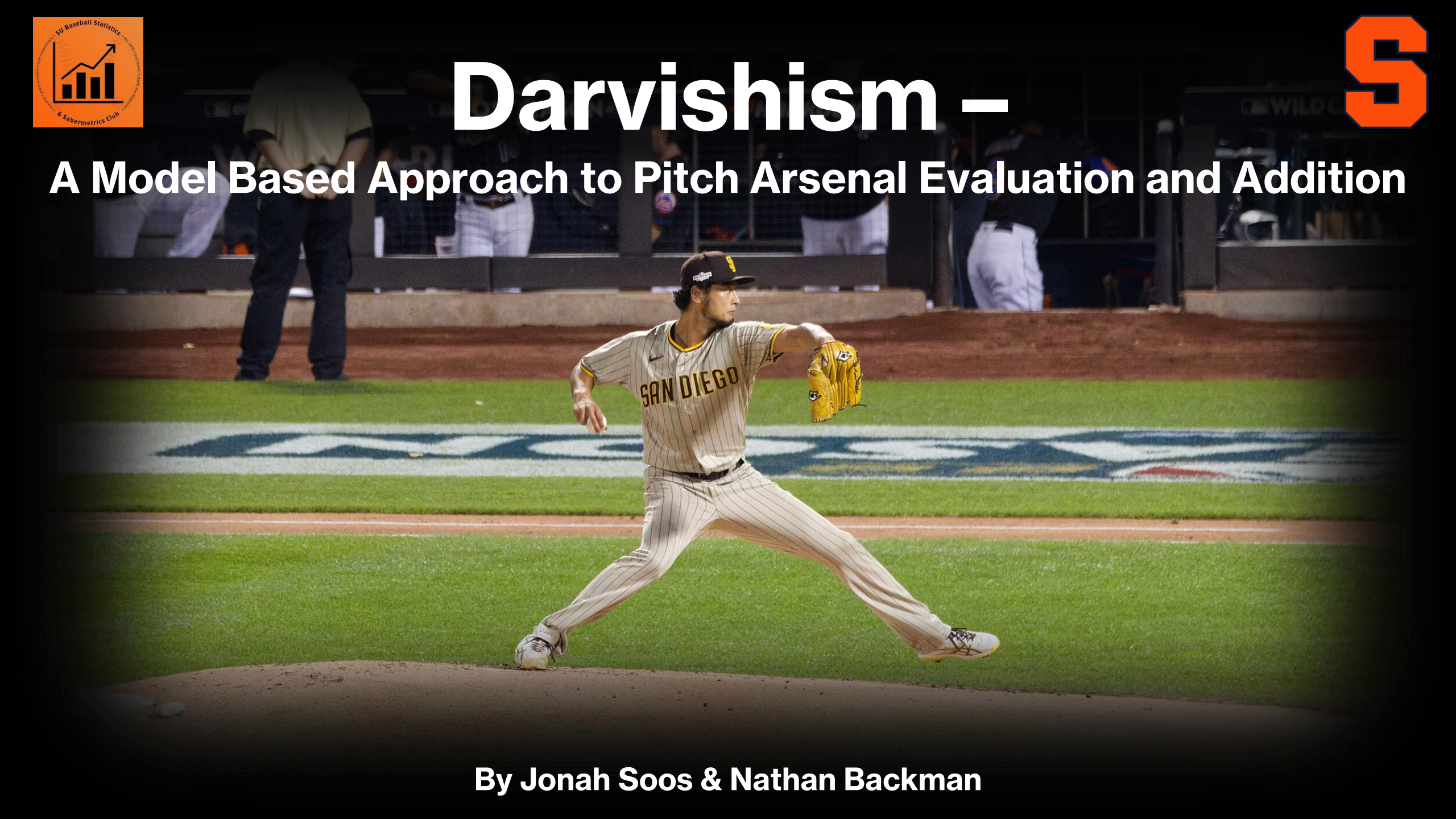
Darvishism –

A Model Based Approach to Pitch Arsenal Evaluation and Addition



By Jonah Soos & Nathan Backman

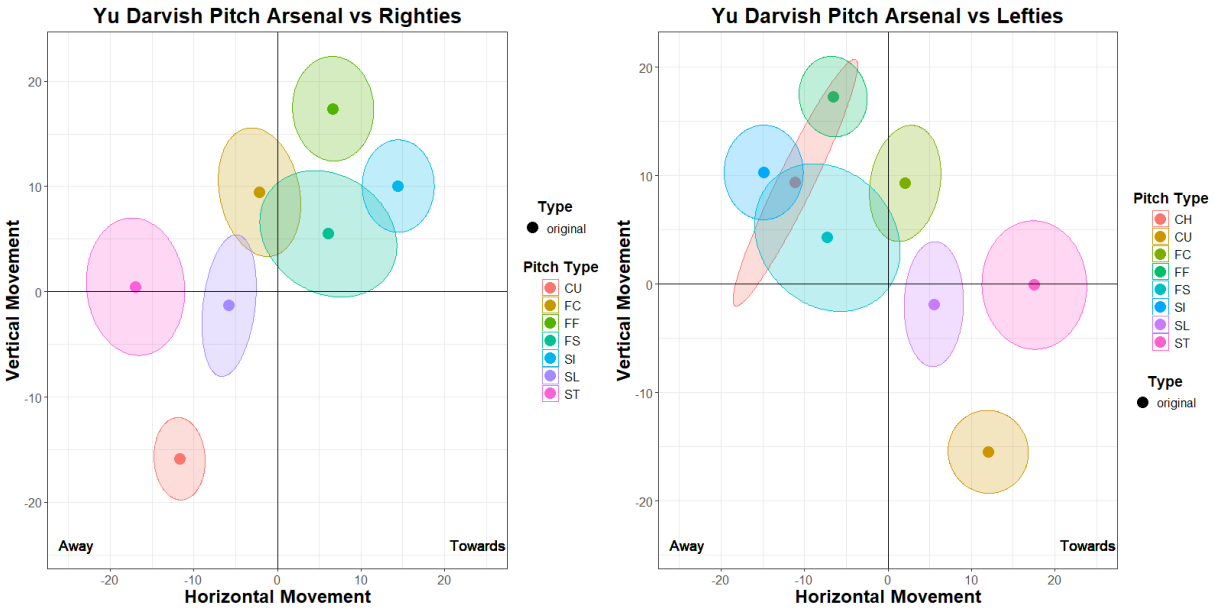
The 11 Pitches of Yu Darvish

Bleacher Nation - 2020

Pitching Stylistically

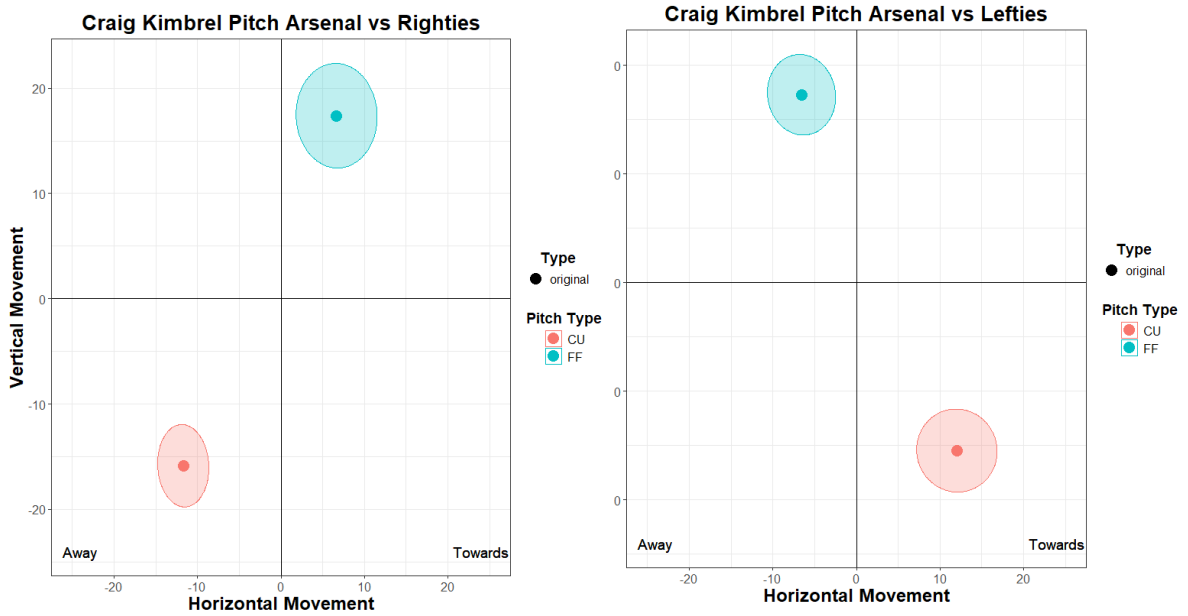
Yu Darvish – The Magician

Uses slight speed and variation to deceive



Craig Kimbrel – Oil & Water

Uses harsh contrast to miss barrels



The dichotomy of pitching: both are effective, but in what cases?

Methodological Overview

Goal: How can we evaluate the **impact of adding or subtracting pitches** from an arsenal on a pitcher's performance?

Analyze the Current

Pitcher Deception Model

Project the Future

Project Pitch Shapes for New Offerings

Analyze the Changes

Marginal Effects of Adding / Subtracting Pitches

Observe the Impact

Recalculate Arsenal Grades and Observe Changes

Key Assumptions:

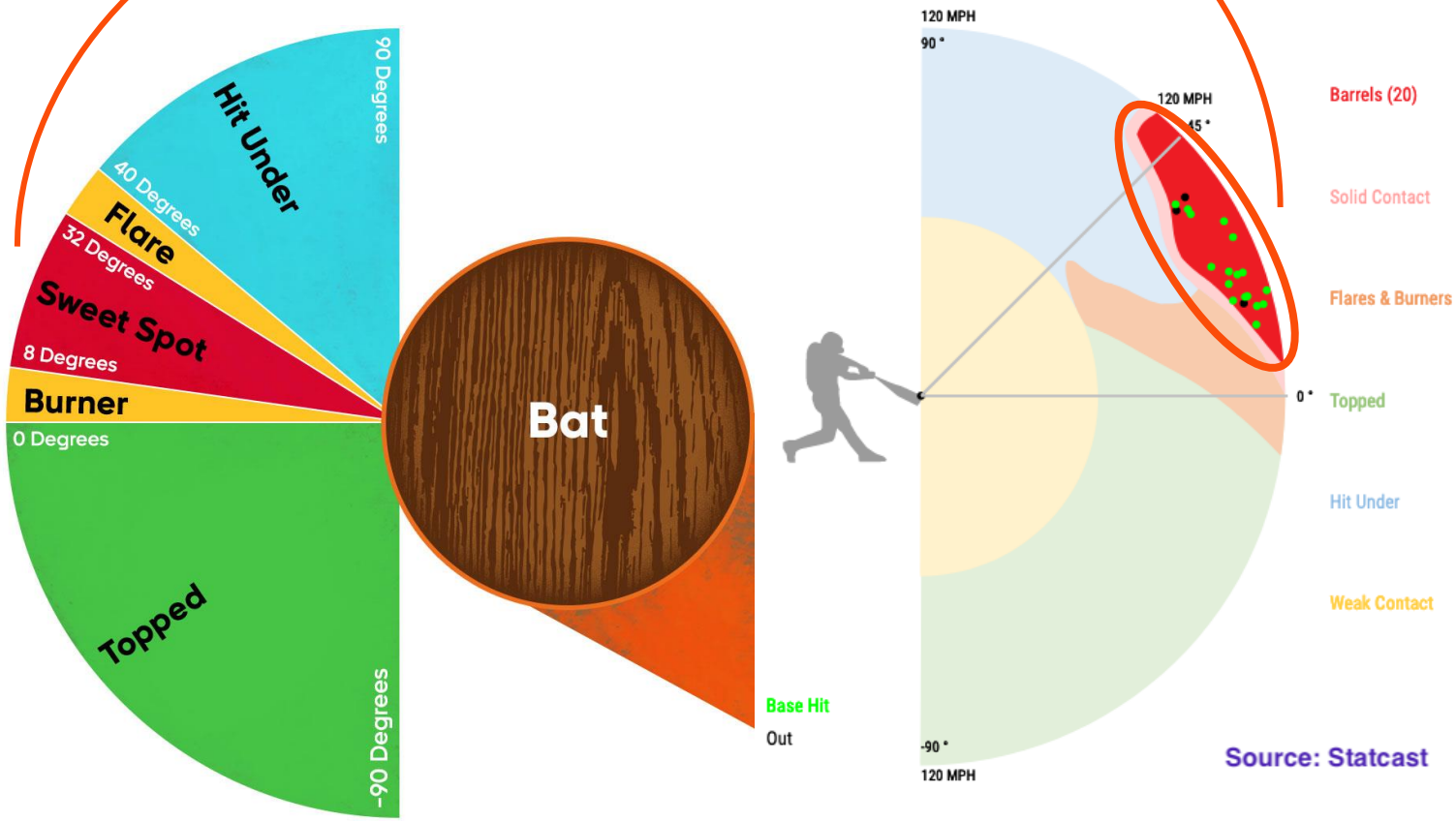
- Pitch shapes remain consistent from year to year in the evaluation of our final arsenal projection
- Pitch types were taken from Statcast classifications with a couple of alterations

Defining Pitch Value

Dependent Variable	xWOBA
Swing & Miss	0.00
Non-Barrel Contact	0.180
Barrel or Sweet Spot	0.712

Evaluating the Value of a Pitch

- **Deception:** The ability to **miss barrels** by inducing **weak contact** and **swings and misses**
- Ordinal variable used to define pitch value
- Final assessment on pitch quality utilizes the associated xWOBA weights
- Foul balls included in non-barrel contact
- Assessment of pitch quality is location independent



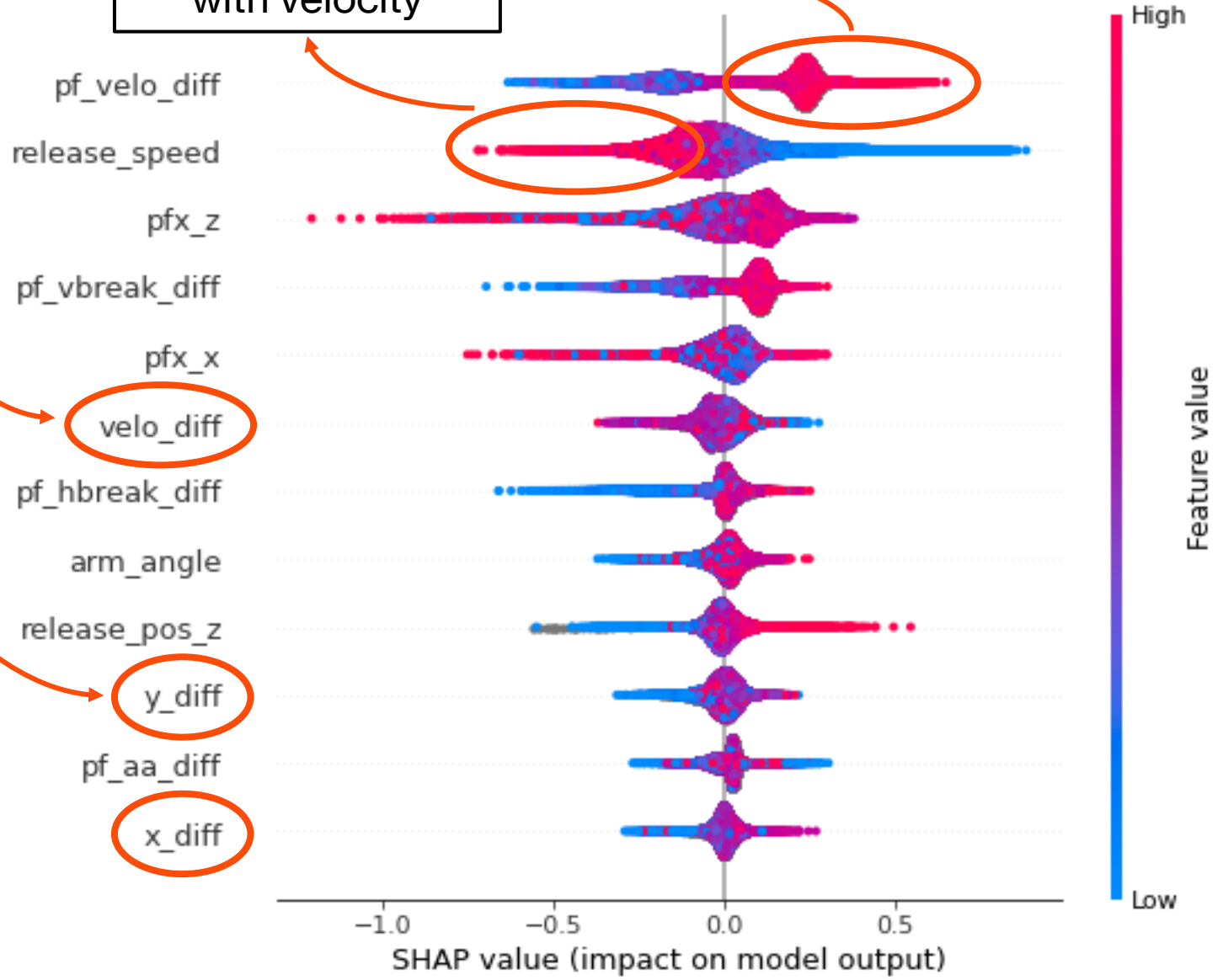
Pitcher Deception Model

Pitcher Deception Model

- Ordinal gradient boosting model framework
- Horizontal movement transformation
- Analyzed previous pitch – allows us to find pitch interaction
- Statcast pitch by pitch data
  - Trained/tested on 2020-2022
  - Validated on 2023
- Random Search Hyperparameter tuning with early stopping and 4-fold cross validation

Strong Correlation with velocity

Capture effects of Previous Pitch



Projecting Pitch Attributes

Projecting Pitch Shapes

- Projections for **pitch velocity, horizontal and vertical movement**
- Similarity scores developed for each pitcher
  - Similarity based upon Euclidean Distance of current pitch shapes
- Decision tree regression
  - Training data utilized based on 60 Closest “Most Similar” Pitchers
  - Similarity scores as model case weights
- Arm angle calculations
  - Fine Tuned with DBSCAN clustering

Projecting Craig Kimbrel’s Sweeper

Top 3 Most Similar Pitchers (Standardized Similarity Score > 90th Percentile)

Hunter Strickland

Matt Brash

Caleb Ort

All have similar arm-angles, shapes of 4-Seam Fastball. Used their Sweepers to project Kimbrel Sweeper movement

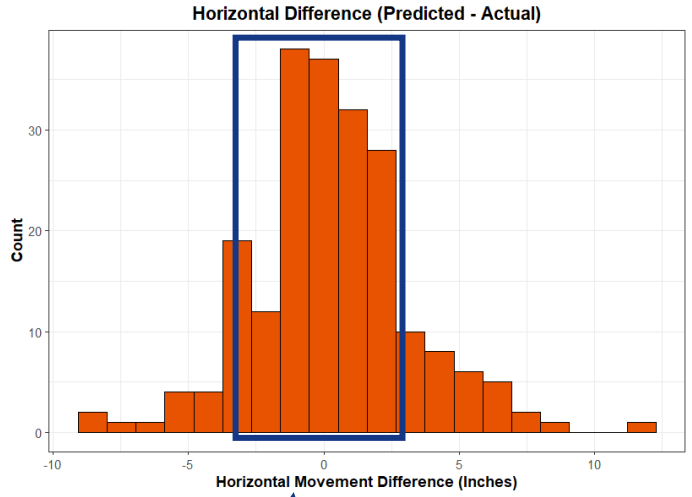
Craig Kimbrel Projected Sweeper

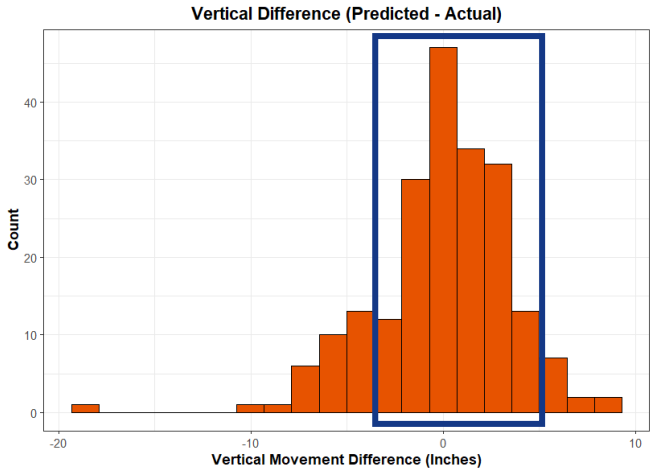
Batter Hand	Velocity	Horizontal Break	Vertical Break	Arm Angle	Release Position
Left	82.05	15.56	3.91	63.83	4.81
Right	82.05	-15.56	3.91	63.73	4.79

Model Validation

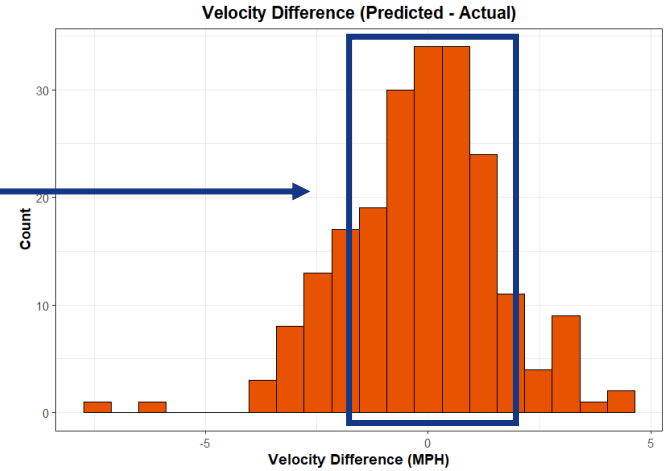
Model Validation Analysis

- Validation set comprised of pitchers that added a new pitch in 2023
- Curveballs and sinkers had the most accurate projections
- Within validation group, **60%** of pitches with at least 20 observations were **within 5 Deception+** points of the actual value





Majority error close to 0

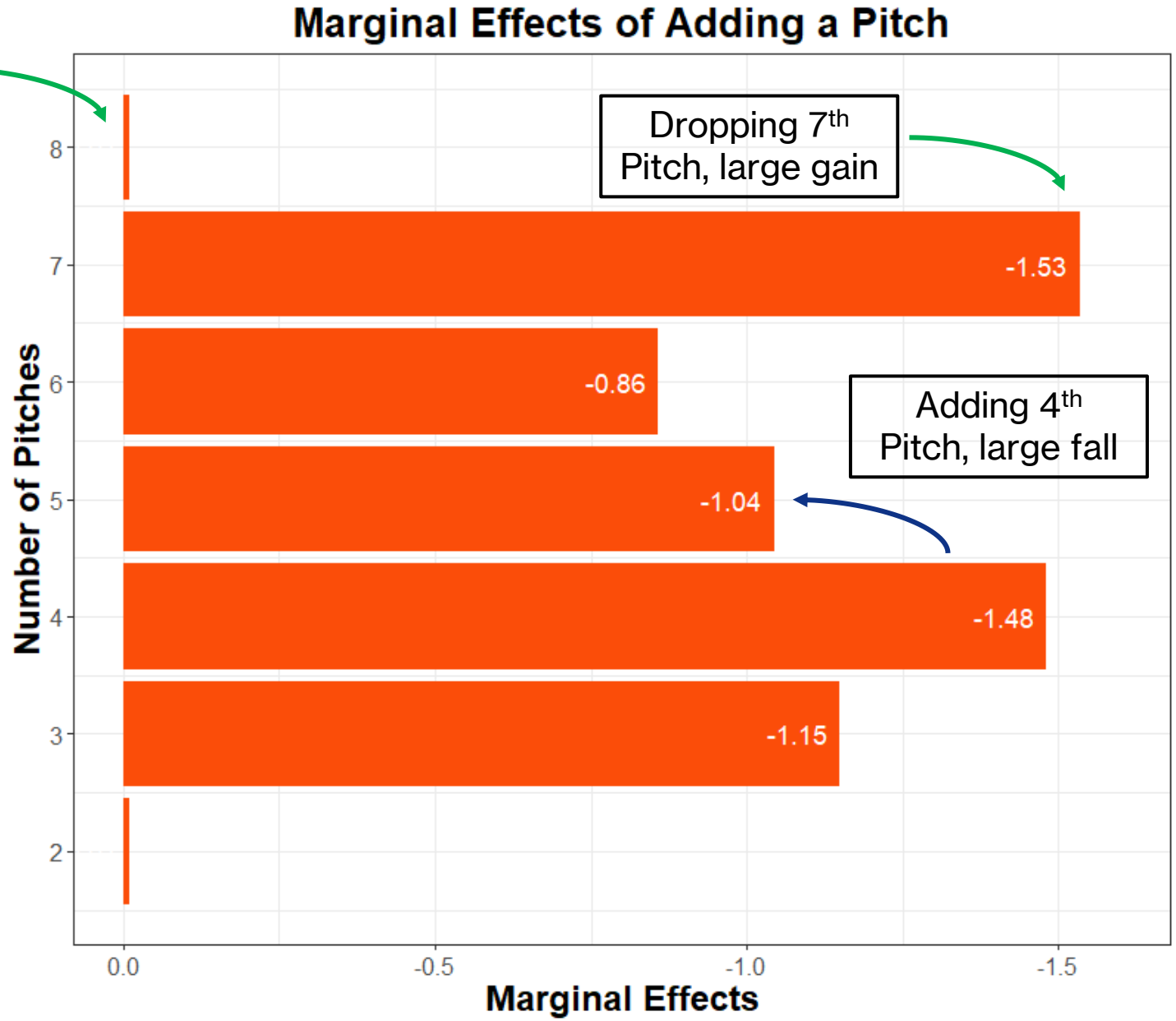


91 Observations

Robert Stephenson Cutter Validation						
Type	Previous Pitch	Batter Hand	Velocity	Horizontal Break	Vertical Break	Deception+
Predicted	Cutter	Right	89.52	0.08	6.44	**92.37**
Actual	Cutter	Right	88.57	0.58	2.72	**97.90**

Marginal Effects

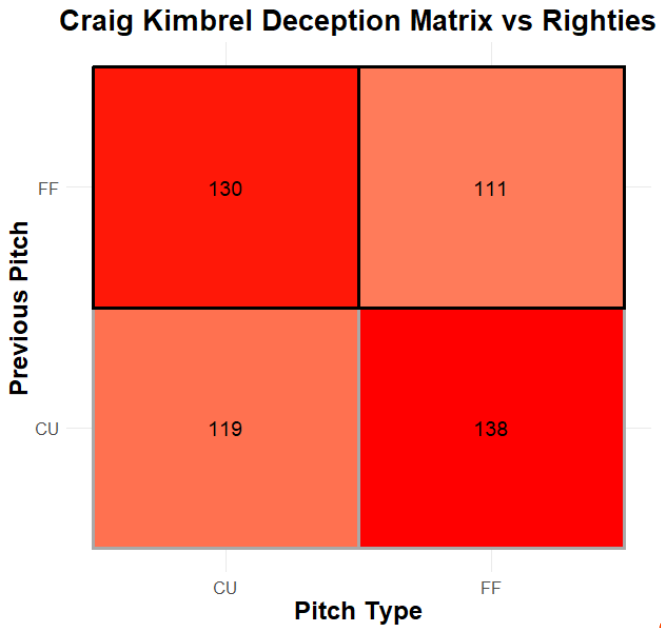
No Statistical Significance



Projecting Pitch Shapes

- Linear OLS model to estimate the effects of adding a pitch to an arsenal
- Effects are **negative** when adding a pitch, **positive** when subtracting a pitch
- Conducive to the idea that less pitches = higher quality
- However, **marginal effects are minimal,** adding a quality pitch will **overcome the effects** in many cases

Deception Matrix Analysis



Marginal Effects: **-1.15**

Pitch Status
Currently Thrown
Predicted Pitch

Arsenal vs Righties

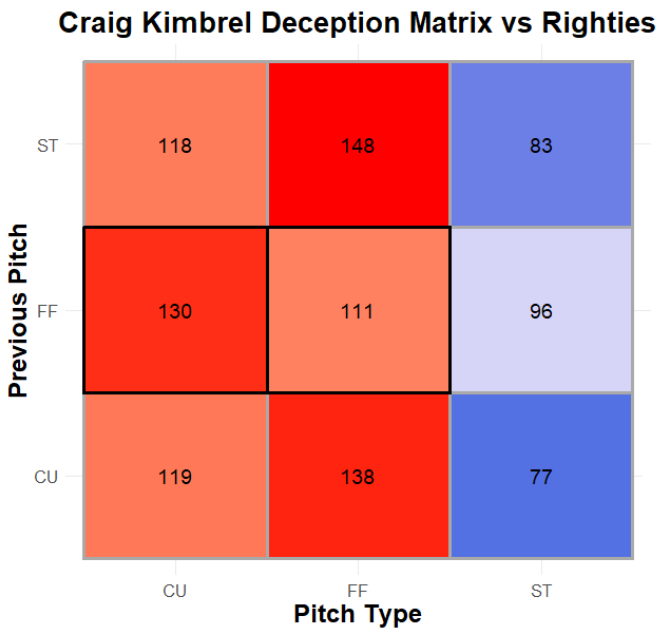
Overall Grade: 124.5

Deception+ off Fastballs: 120.66

Deception+ off Breaking Balls: 128.34

Deception+ off Offspeed: NaN

Net Arsenal Rating: 0



Pitch Status
Currently Thrown
Predicted Pitch

Arsenal vs Righties

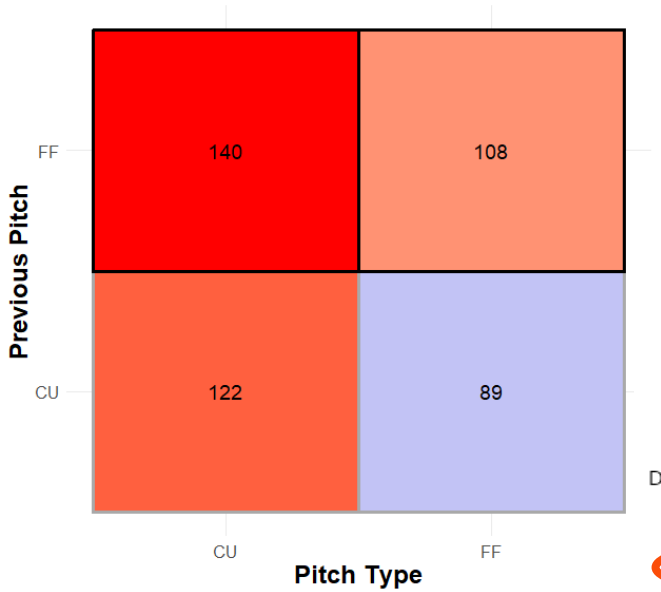
Overall Grade: 112.35

Deception+ off Fastballs: 112.49

Deception+ off Breaking Balls: 114.01

Deception+ off Offspeed: NaN

Net Arsenal Rating: -12.15



Marginal Effects: **-1.15**

Pitch Status
Currently Thrown
Predicted Pitch

Arsenal vs Lefties

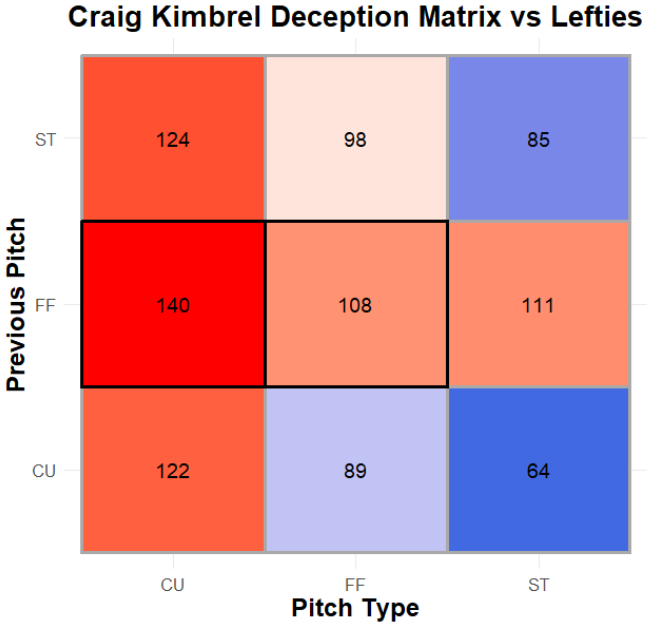
Overall Grade: 114.75

Deception+ off Fastballs: 124.01

Deception+ off Breaking Balls: 105.5

Deception+ off Offspeed: NaN

Net Arsenal Rating: 0



Pitch Status
Currently Thrown
Predicted Pitch

Arsenal vs Lefties

Overall Grade: 103.4

Deception+ off Fastballs: 119.66

Deception+ off Breaking Balls: 97

Deception+ off Offspeed: NaN

Net Arsenal Rating: -11.35

Conclusion

Biggest Arsenal Gainers

Devin Williams

Current Arsenal (R):	Optimized Arsenal (R):	Current Arsenal (R):	Optimized Arsenal (R):
Fastball, Change-Up	4-Seam, Change-Up, **Curveball, Sinker, Sweeper**	Fastball, Change Up, **Cutter**	4-Seam, Change-Up, **Curveball, Sweeper**
113.01 → 121.9 (+8.92)		**99.02 → 108.95** (+9.93)	

Orion Kerkering

Current Arsenal (R):	Optimized Arsenal (R):	Current Arsenal (L):	Optimized Arsenal (L):
Sinker, **Sweeper**	**4-Seam**, Sinker, **Change-Up, Slider**	Sinker, Sweeper	Sinker, **Cutter, Splitter,** Sweeper
106.16 → 123.24 (+17.1)		**109.59 → 118.69** (+9.10)	

Bryan Hoeing

Current Arsenal (R):	Optimized Arsenal (R):	Current Arsenal (L):	Optimized Arsenal (L):
Change-Up, **4-Seam**, Sinker, **Slider**	Sinker**, Splitter, Sweeper**	**Change-Up**, **4-Seam**, Sinker, **Slider**	Sinker**, Splitter, Sweeper**
102.77 → 110.62 (+7.85)		**86.67 → 107.15** (+20.48)	

Erik Swanson

Current Arsenal (R):	Optimized Arsenal (R):	Current Arsenal (L):	Optimized Arsenal (L):
4-Seam, **Splitter**, **Slider**	4-Seam, **Sinker**, **Curveball**	4-Seam, **Splitter**, **Slider**	4-Seam, **Sweeper**, **Curveball**
100.45 → 110.62 (+18.81)		**96.89 → 106.18** (+9.29)	

Key Takeaways

- **High movement profiles** with **unique primary fastballs** tend to deceive batters the most
  - Strong value in the primary fastball and all pitches that move off of it
- Can project with reasonable accuracy pitch characteristics of additional pitches not currently in their repertoire
- **Deception and talent correlate** but missing locational component.
- Hard to value arsenal uniqueness, want to look further into tunnels and how pitches look at the decision point.

ShinyApp

https://jonahsoos24.shinyapps.io/Darvishism_Pitcher_Arsenal_Evaluation_App/

or

@jonahsoos24 on Twitter (Link in Bio)