Fair or Forgotten? Forecasting and Evaluating 2024 Playing Time

Brett Cerenzio, Aidan Gurvitch, Danielle Napierski, Zach Silverstein, and Dylan Stafeil

2025 Cincinnati Reds Baseball Analytics Hackathon

Objective:

Predict Total Playing Time Players Will Receive in 2024

Batters → Plate Appearances

Pitchers → Batters Faced

Using 2021-2023 Baseball Savant and Lahman Data

Filtering Players

Filtering the Dataset

Remove Pitchers from PA & Batters from BF Datasets

Pitchers aren't signed to hit and batters aren't signed to pitch (minus Ohtani), helps reduce noise in the data

Pitchers from PA

Check players who appear in BF data and has <10 PAs in 2022 and 2023

Universal DH implemented in 2022

Batters from BF

Check players who appear in PA data and has <50 BFs across 2021-2023

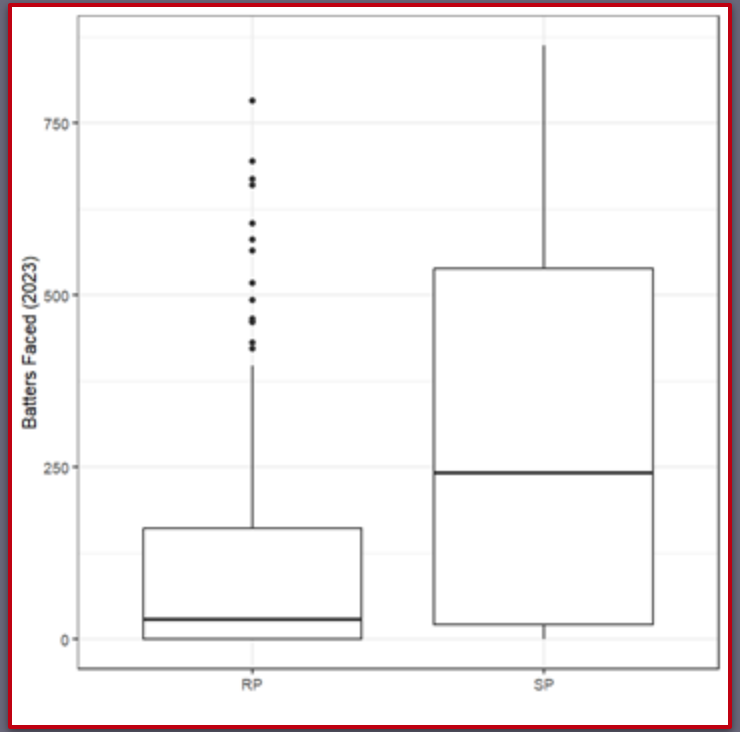
This is AFTER pitchers are removed

SPs Versus RPs

Differentiation

Number of Appearances
Based on Each Role

0.15 * (# 2021 Appearances) +
0.85 * (# 2022 Appearances)



Methodology

Position Players

Workload

- PA in Previous 2 Seasons
- Lineup + Batter Position
- Second Half PA
- Games Started After Debut
  - Rookies
- PA Per Game

Experience

- Debut Month (Rookies) and Year
- Age

Performance

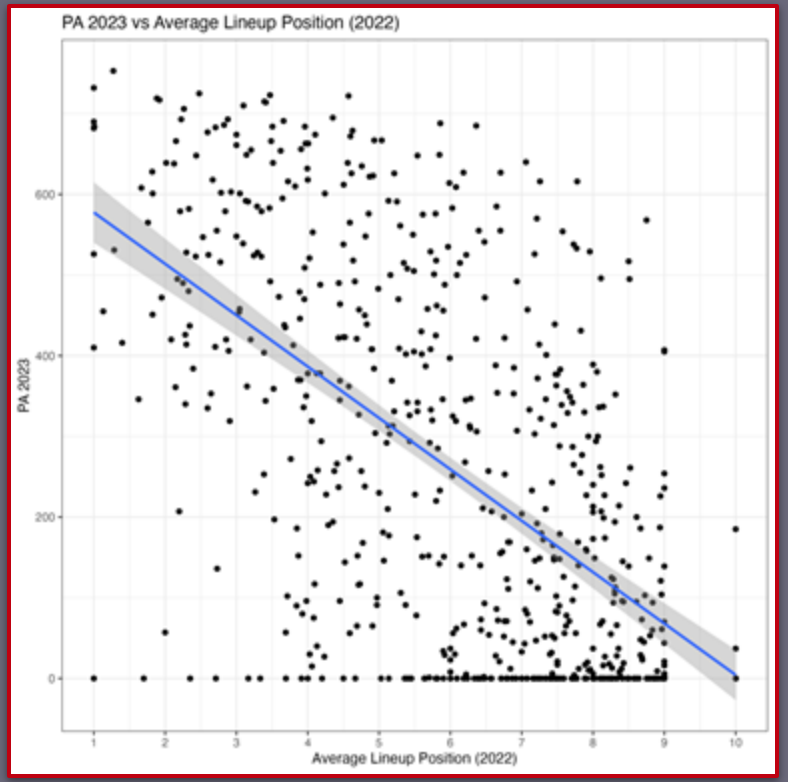
- Avg and 90th % EV
- Total Bases, OBP, SLG
- Fielding/Framing RV
- Other Count/Rate Stats

Lineup Position

Found the Average Batter Position in the Starting Lineup Across all 3 Years

Players who Didn't Start a Game in a Given Year Were Assigned 10 in the Lineup Position

Players Batting Earlier in the Lineup Tend to Receive More Playing Time



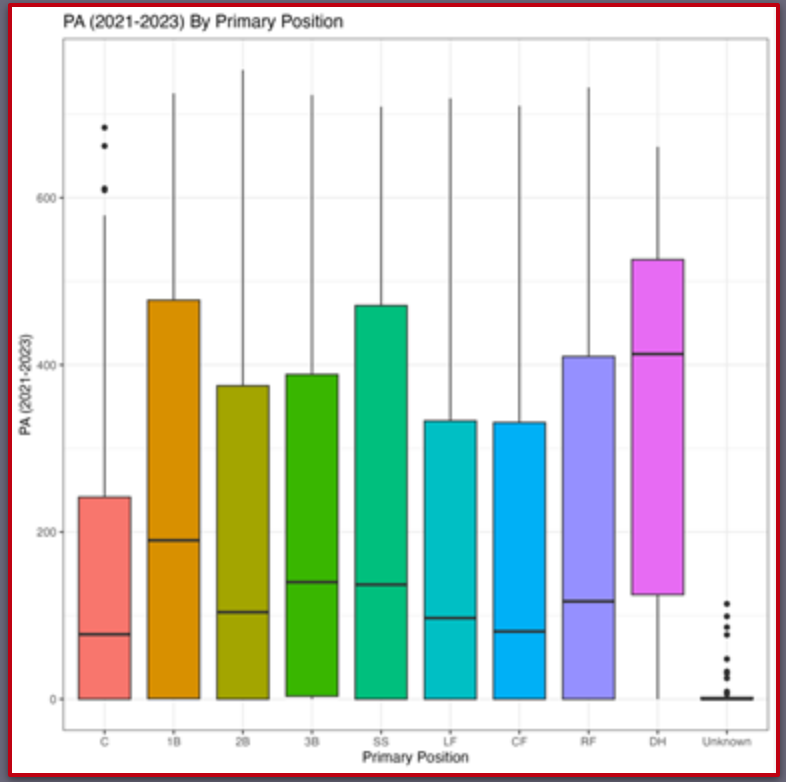
Batter Fielding Positions

Identified Primary Position by Where They Started the Most Over 2021-2023

Catchers Typically see Less Playing Time

1B and DHs get a lot More Playing Time Relative to Other Positions

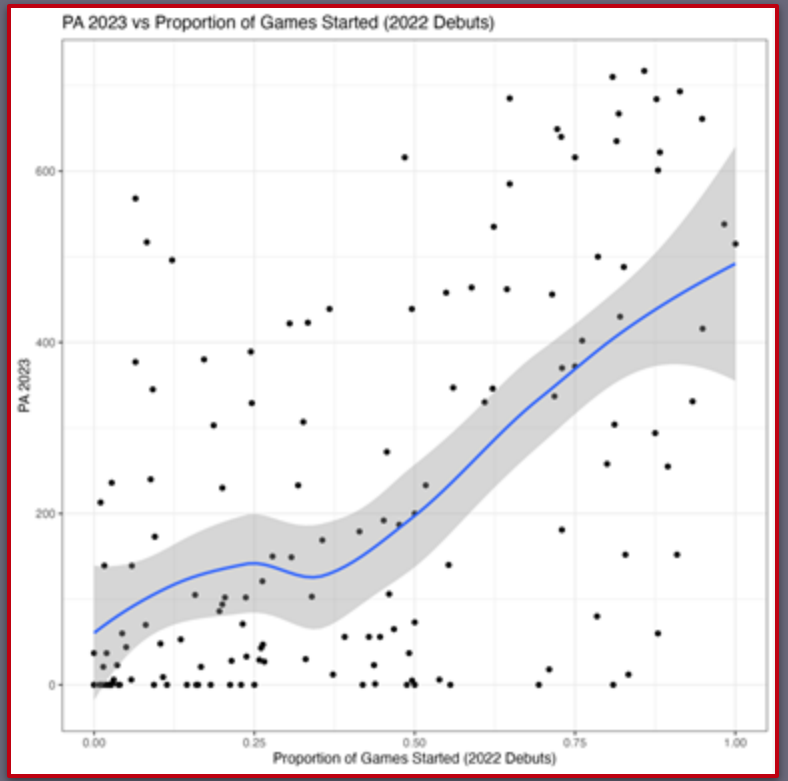
“Unknown” Represents Players That Never Started in the Field



Proportion of Games Started (Debuts)

Identified Number of Games Remaining on a Team's Schedule and Number of Games a Player has Started Using Debut Date

If a Rookie Starts Roughly 35% of Games, Their Playing Time Should Increase



Pitchers

Workload

- BF In Past 2 Seasons
- Pitches per Appearance
- BF per Appearance
- Avg. Times Through Order
- Pitches per PA
- BF in Second Half of Season
- BF Splits

Performance

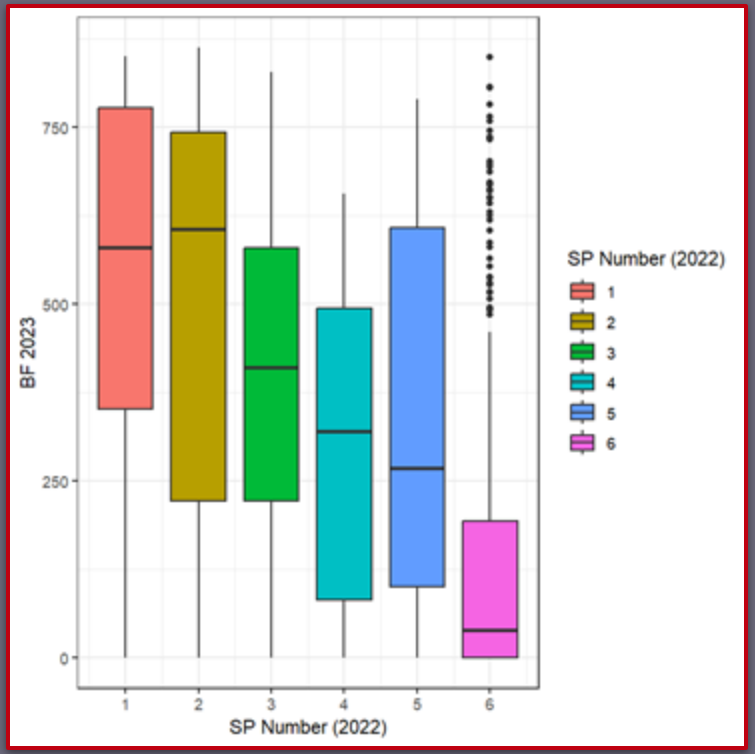
- 90th Percentile Fastball Velocity
- Strikeouts and K%
- Other Count and Rate Statistics

SP: Number in Rotation on Opening Day
RP: Number of 9th Inning Appearances

Number in Starting Rotation (Opening Day)

Opening Day Starters Face the Most Batters on Average Throughout the Season

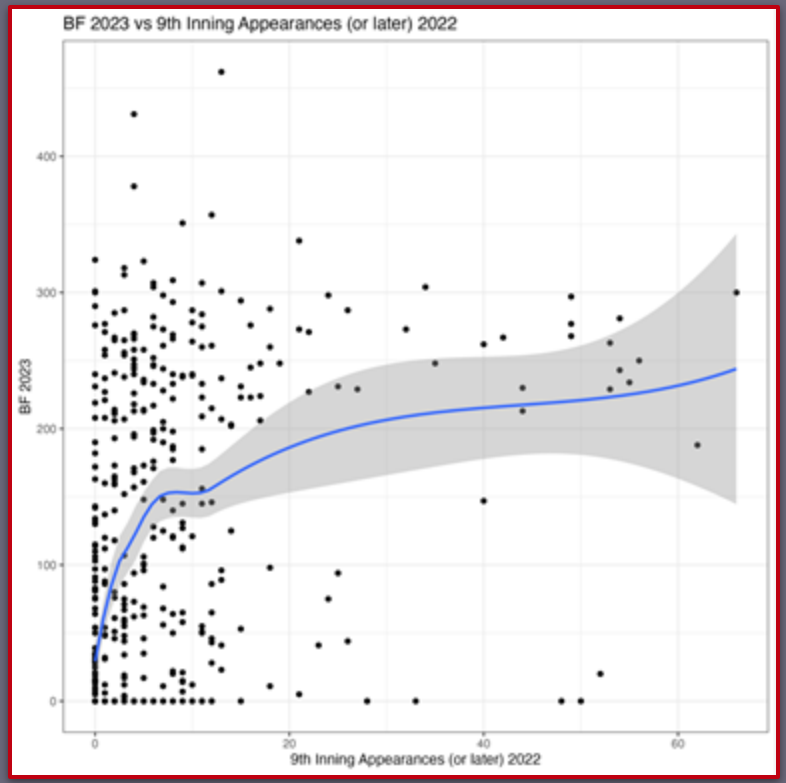
Inverse Relationship Between Rotation Number and Workload



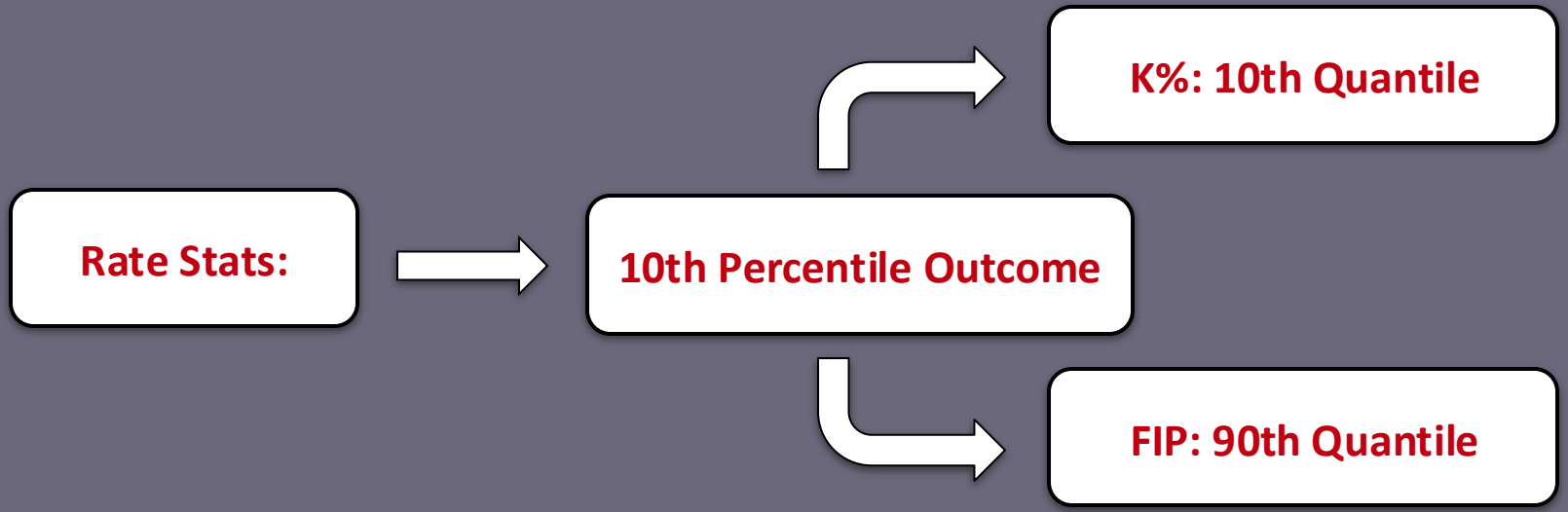
Number of 9th Inning Appearances

High Leverage Relievers are Going to Pitch in the 9th Often

Logarithmic Relationship Between 9th Inning Appearances and BF the Following Season



NA Imputation



Lack of PAs in a Season is Most Likely due to an Injury or Poor Ability

Clustering Analysis

Why This Approach?

Identify Playing Consistency Throughout the Season (Monthly Playing Time) and Evaluating its Impact on Future Playing Time

Insights to Whether a Player got Less PA/BF due to Being Benched or Injured

These Metrics are Only Valuable When Used Together, Rather Than Individually

K-Means Clustering

Batters and Relief Pitchers

- Monthly Playing Time
- Proportion of Total/Season Playing Time That Occurred in Each Month
- Whether the Player was a Rookie

**all for the prior 2 seasons*

Batters: 18 Clusters

RPs: 14 Clusters

K-Means Clustering

Starting Pitchers

Algorithm 1

Monthly Playing Time Throughout the Past 2 Seasons

Algorithm 2

Monthly Proportion of Total Playing Time Throughout the Past Season

And Whether a Rookie in 2021 or 2022 (as separate variables)

Algorithm 1: 7 Clusters

Algorithm 2: 6 Clusters

Models

Finding the Right Fit

Model Construction

Removed players who debuted in 2023, didn't play in past or current season, or didn't exist in Lahman database

35 Fold CV	Limited Data, Model Was Efficient Despite Numerous Folds
BART	Applies Non-Parametric Approach to Bayesian Modeling Less Prone to Overfitting, but Computationally Expensive and Limited Interpretability
XGBoost	Boosting Algorithm that Minimizes the Residuals Efficient and High Performance, But Prone to Overfitting

Weighting Characteristics

Individualized Weights Were Created Based on Player Characteristics

Group 1

- 100+ PA/150+ BF 2 Seasons Ago
- Less Than 30 PA/50 BF in Past Season

(Weight 2)

Group 2

- 100+ PA/150+ BF in Past Season
- Less Than 30 PA/50 BF 2 Seasons Ago
- NOT a Rookie

(Weight 3)

Group 3

Rookies Who Debuted Before May

Assigned 2022 Stats

Group 4

Rookies Who Debuted in May or June

Counting Stats Only

(Weight 4)

Group 5

Rookies Who Debuted in July or August

Counting Stats Only

(Weight 5)

Group 6

Rookies Who Debuted in September or October

Counting Stats Only

(Weight 6)

Group 7

All Other Players Who Don't Meet Groups 1-6

(Weight 1)

Results

PA Weights

Weight 1

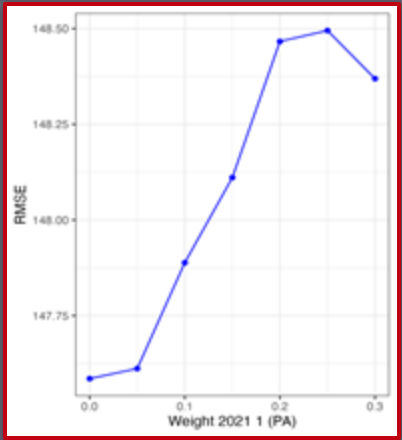
All Other Players Who Don't Meet Groups 1-6

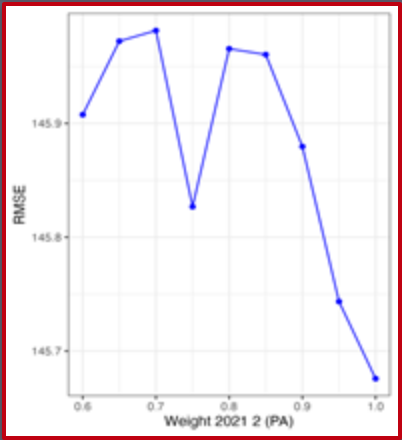
Weight 2

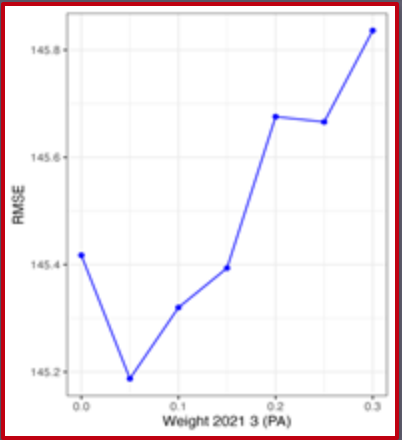
- 100+ PA/150+ BF 2 Seasons Ago
- Less Than 30 PA/50 BF in Past Season

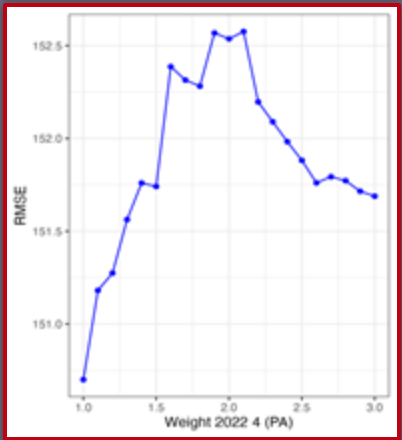
Weight 3

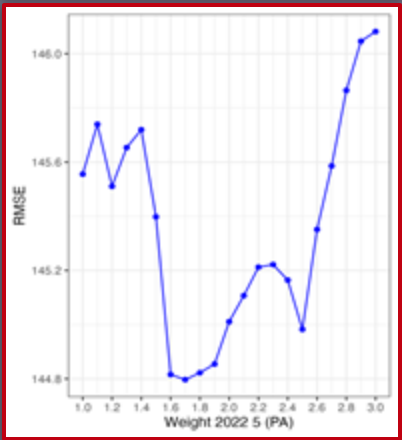
- 100+ PA/150+ BF in Past Season
- Less Than 30 PA/50 BF 2 Seasons Ago
- NOT a Rookie

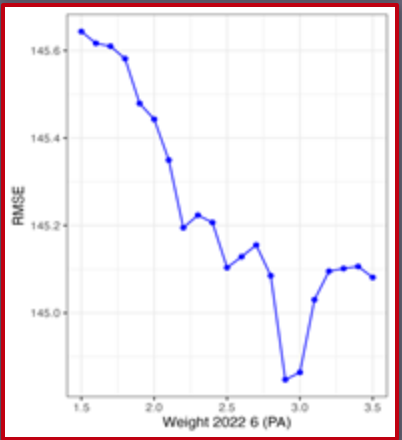












Weight 4

Rookies Who Debuted in May or June

Counting Stats Only

Weight 5

Rookies Who Debuted in July or August

Counting Stats Only

Weight 6

Rookies Who Debuted in September or October

Counting Stats Only

SP Weights

Weight 1

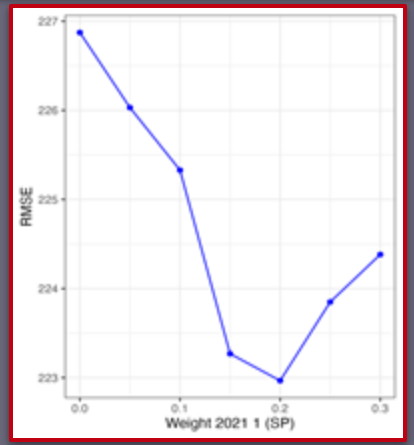
All Other Players Who Don't Meet Groups 1-6

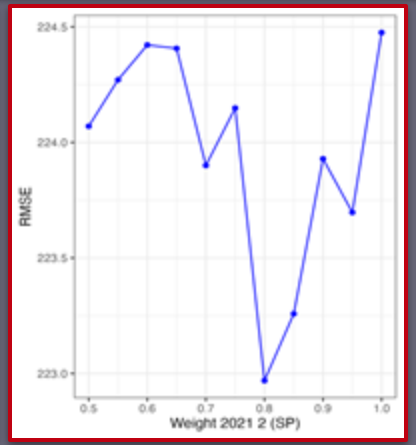
Weight 2

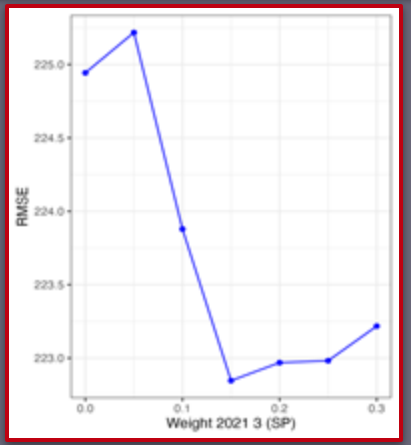
- 100+ PA/150+ BF 2 Seasons Ago
- Less Than 30 PA/50 BF in Past Season

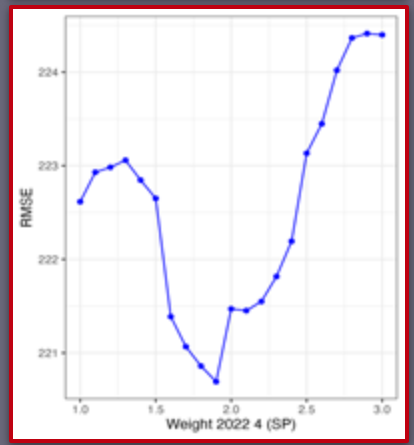
Weight 3

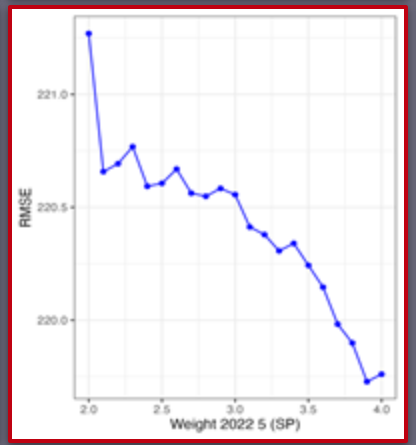
- 100+ PA/150+ BF in Past Season
- Less Than 30 PA/50 BF 2 Seasons Ago
- NOT a Rookie

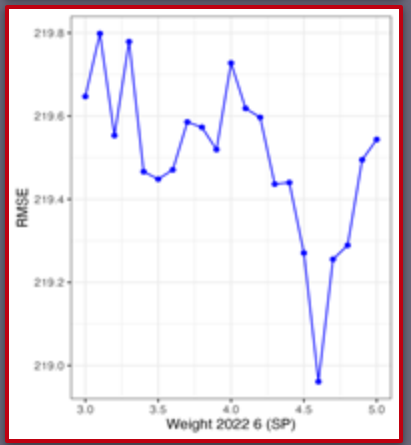












Weight 4

Rookies Who Debuted in May or June

Counting Stats Only

Weight 5

Rookies Who Debuted in July or August

Counting Stats Only

Weight 6

Rookies Who Debuted in September or October

Counting Stats Only

RP Weights

Weight 1

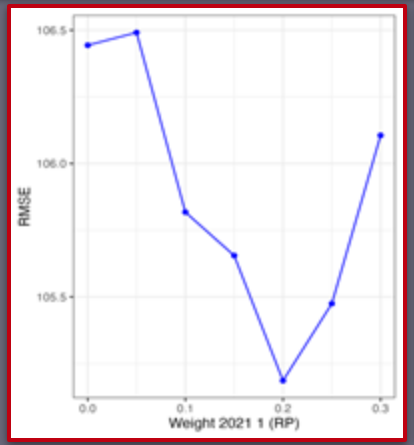
All Other Players Who Don't Meet Groups 1-6

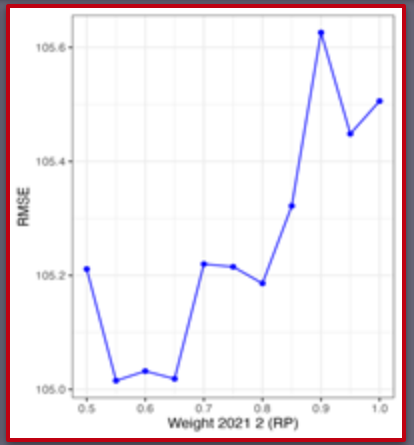
Weight 2

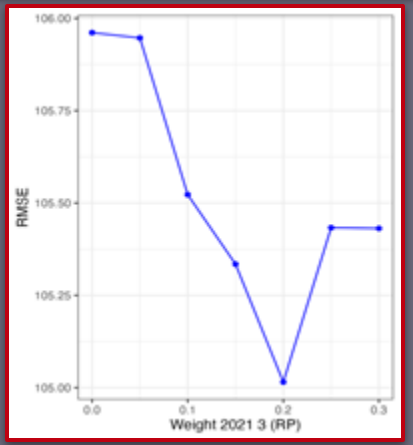
- 100+ PA/150+ BF 2 Seasons Ago
- Less Than 30 PA/50 BF in Past Season

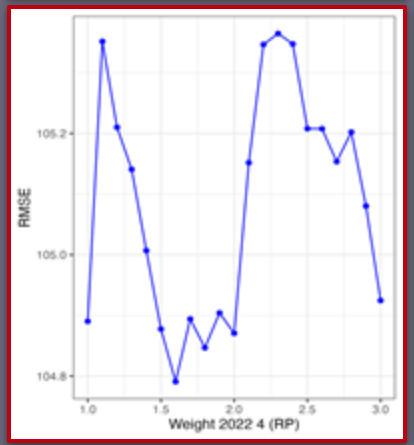
Weight 3

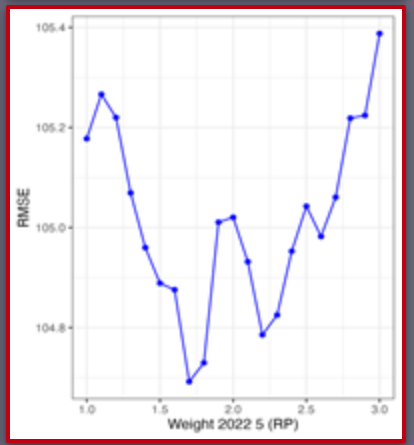
- 100+ PA/150+ BF in Past Season
- Less Than 30 PA/50 BF 2 Seasons Ago
- NOT a Rookie

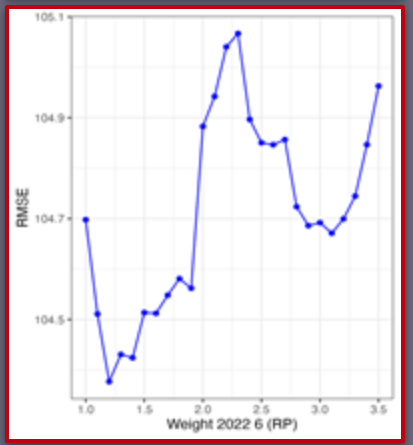












Weight 4

Rookies Who Debuted in May or June

Counting Stats Only

Weight 5

Rookies Who Debuted in July or August

Counting Stats Only

Weight 6

Rookies Who Debuted in September or October

Counting Stats Only

Batter Distribution → RMSE: 144.3

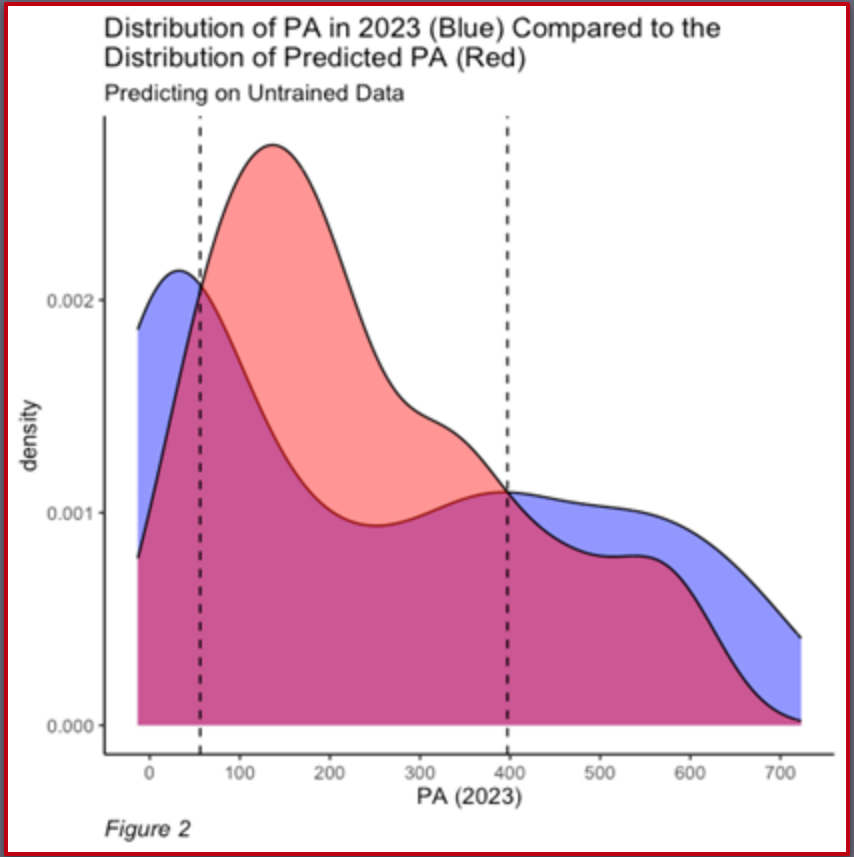


Figure 2

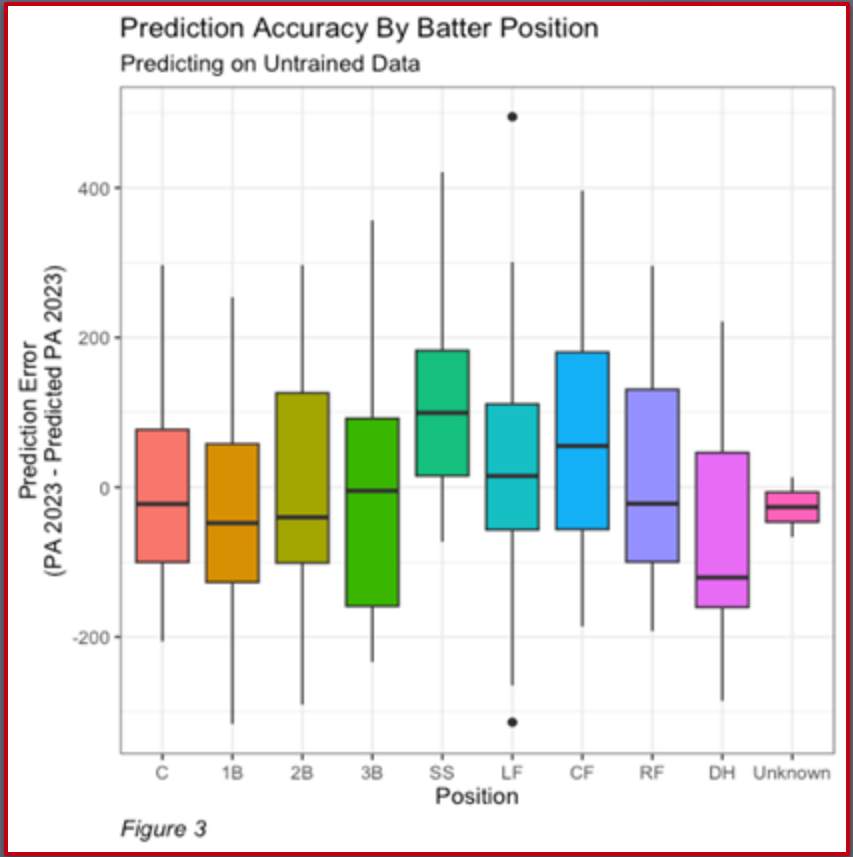


Figure 3

Batter Distribution

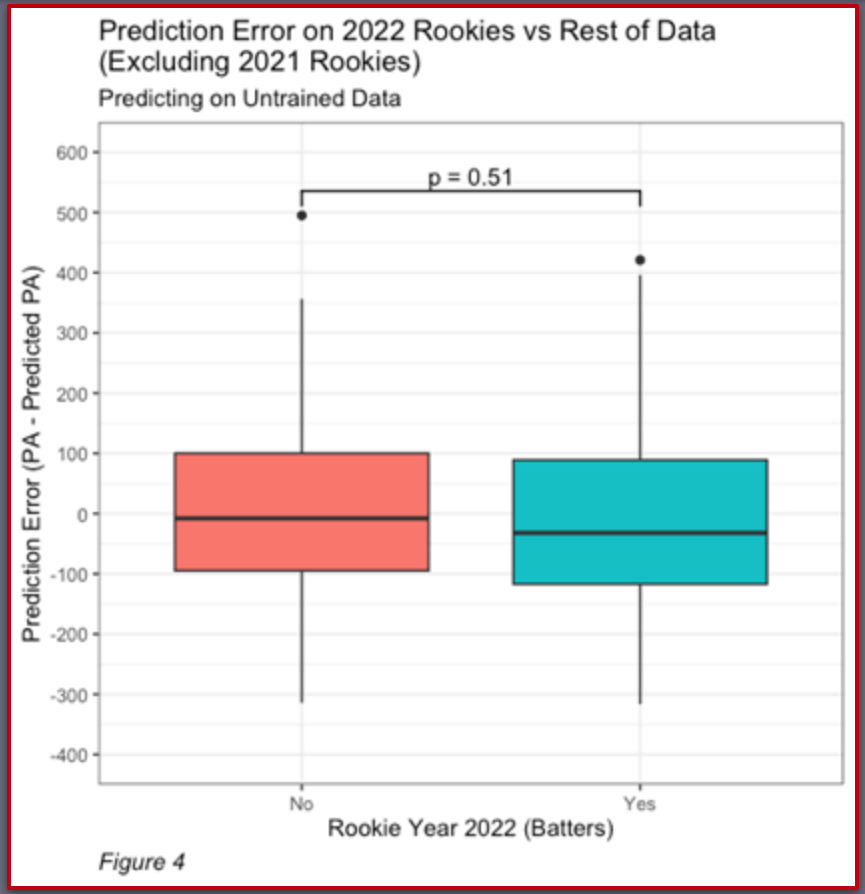


Figure 4

SP Distribution → RMSE: 220.4

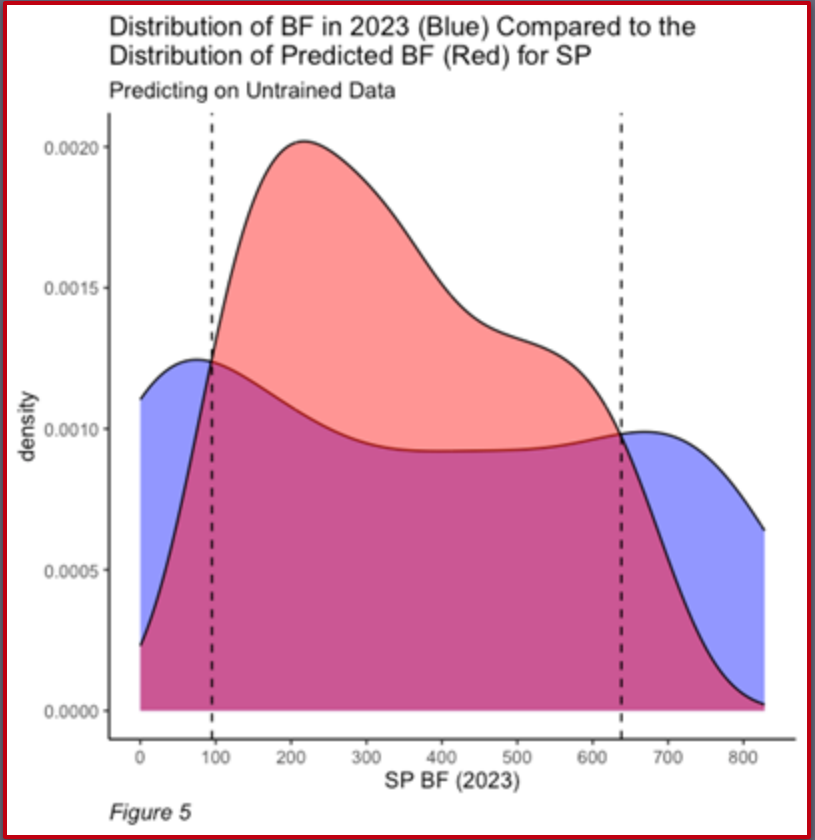


Figure 5

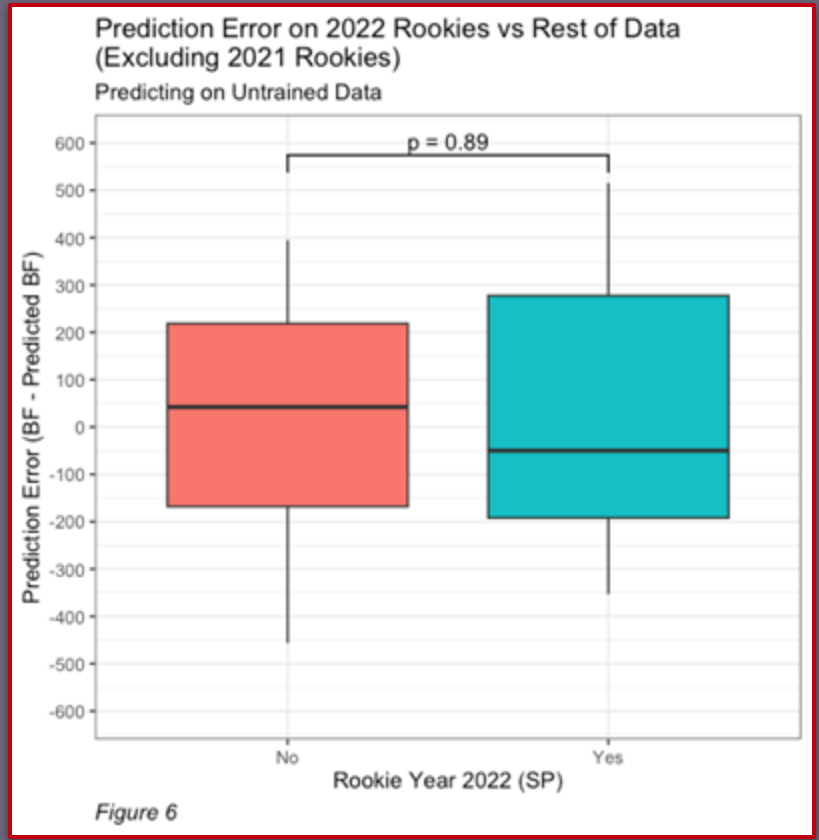


Figure 6

Batters Faced Outliers (2023)			
Pitcher	**2021 BF**	**2022 BF**	**2023 BF**
Pitcher 1	586	574	0
Pitcher 2	779	608	7
Pitcher 3	538	646	0
Pitcher 4	774	780	19

RP Distribution → RMSE: 103.9

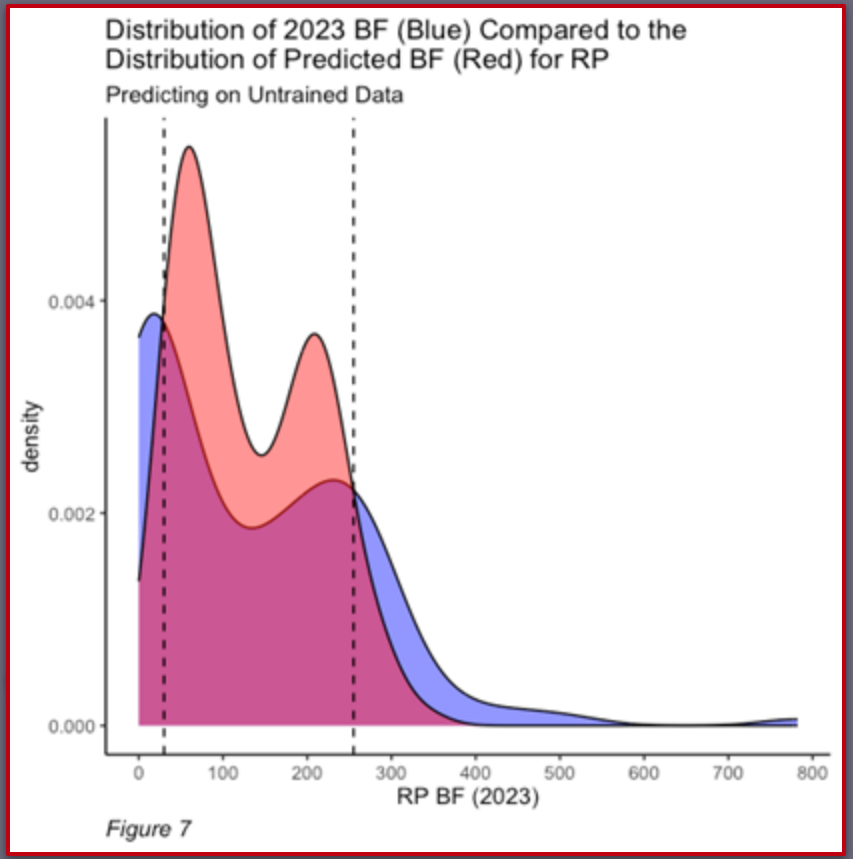


Figure 7

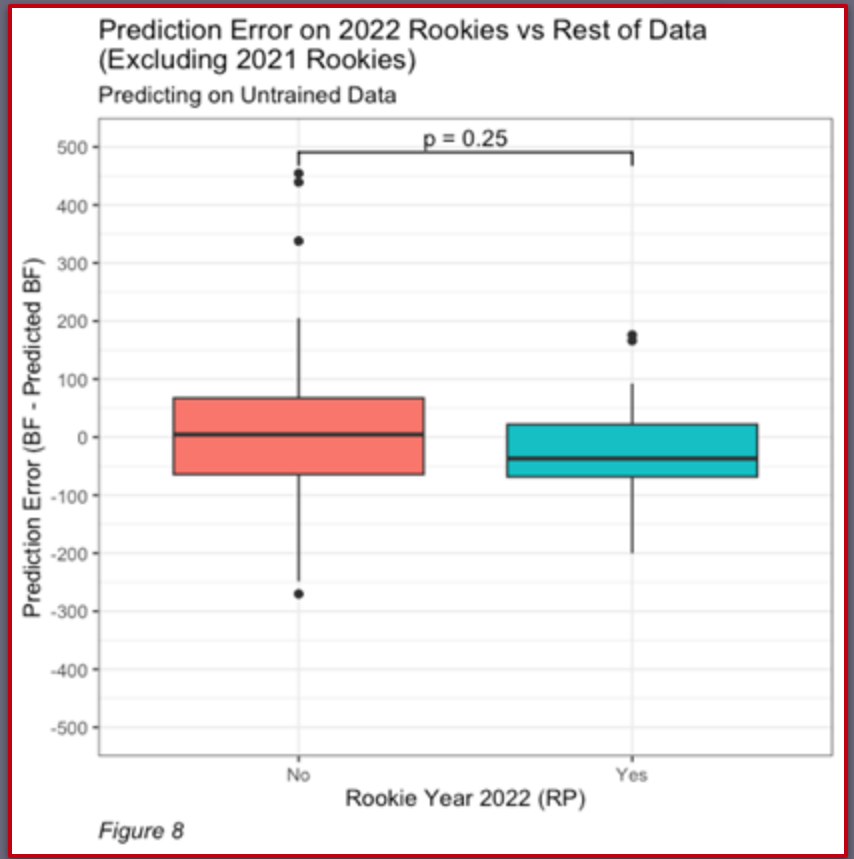


Figure 8

STARTING PITCHERS:

- BRADY SINGER (640)
- HUNTER GREENE (574)
- ANDREW ABBOTT (394)
- NICK LODOLO (314)
- RHETT LOWDER (274)

RELIEF PITCHERS:

- NICK MARTINEZ (357)
- ALEXIS DÍAZ (184)
- SCOTT BARLOW (189)

Limitations

- Preseason and Offseason Injuries
- Lack of Information on Contracts and Service Time
- Doesn’t Account for Team Quality
- Don’t Know Time Spent on 26-Man Roster

Questions?