

Pitch Type and Location Effects on Swings and Hits

Selah Dean



Introduction

Background

- Data is revolutionizing the game of baseball at all levels of the sport
- It is included in all aspects of the game including pitching, fielding, and hitting
- Pitch design and pitch selection are two areas of focus when it comes to pitching and using data

Research Question

- How do different pitches and different components of a pitch affect the outcome of the hitter for different pitchers?
- This project aims to answer this question by looking at data collected from the 2023 fall season for the CWRU baseball team

Data Description

Data Collection

- I collected the data over 6 weeks during the fall baseball season for 18 CWRU pitchers
- Data came from a Shiny App and Rapsodo
- For each week each pitcher had a separate csv file for data collected from the Shiny App and Rapsodo

Rapsodo

- A monitor that is placed on the field that collects data on the movement of the pitch
- Includes features such as pitch type, velocity, spin rate, and location
- Was unable to record every pitch with Rapsodo so there was some missing data

Shiny App

- Recorded the hitters response to the pitch (ball or strike and if there was a play made)
- Data was inputted while rapsodo was running and including pitch type and velocity
- If Rapsodo was not working was still able to record hitting data but pitching data was either missing or velocity came from a radar gun

* see slides 10 and 11 for more detailed information on the attributes collected from Rapsodo and the Shiny App

Approach

Initial Data Cleaning

- By hand I created a new csv file that matched up the Shiny App and Rapsodo data
- Loaded the files into R, added a column for the pitcher name, week, and date which came from the file name and folder name
- The final data frame contained 2780 rows (each row represented a different pitch) and 31 columns

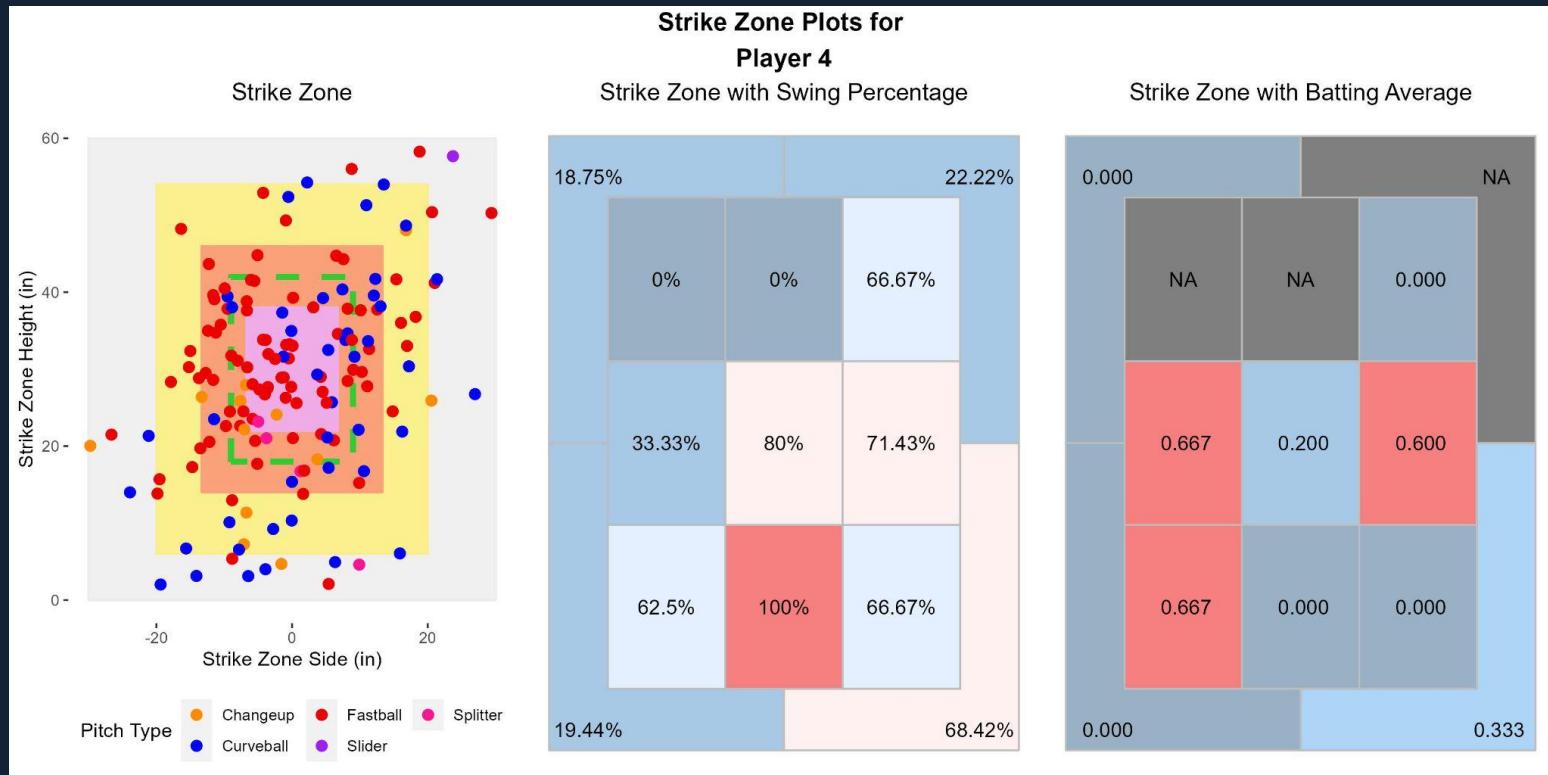
Handling Missing Data

- Due to issues experienced with the Rapsodo there was a large amount of missing data, I only did analysis on the five pitchers who pitched the most innings and had the most complete the dataset

Analysis

- Performed a Fisher's Exact test for 9 different contingency tables for each of the 5 pitchers selected
- Considered swings, hits, play results with pitch types and locations
- With the results I determined what significant associations exist for each pitcher

Initial Visualizations



*players 1,2,3,5 graphs on are slide 12, dimensions of the zones are on slide 13

Results - Command

Contingency Tables for Player 1

Swing vs Strike	
NO	YES
FALSE	59 20
TRUE	36 39

Fisher's Exact Test p-value: 0.00087

Pitch Type vs Strike	
NO	YES
Changeup	24 4
Fastball	45 37
Slider	26 18

Fisher's Exact Test p-value: 0.01026

Contingency Tables for Player 2

Swing vs Strike	
NO	YES
FALSE	87 36
TRUE	39 71

Fisher's Exact Test p-value: 0

Pitch Type vs Strike	
NO	YES
Changeup	15 11
Fastball	67 51
Slider	44 45

Fisher's Exact Test p-value: 0.53207

Contingency Tables for Player 3

Swing vs Strike	
NO	YES
FALSE	72 29
TRUE	29 62

Fisher's Exact Test p-value: 0

Pitch Type vs Strike	
NO	YES
Changeup	20 13
Curveball	1 2
Fastball	43 46
Slider	37 30

Fisher's Exact Test p-value: 0.56167

Contingency Tables for Player 4

Swing vs Strike	
NO	YES
FALSE	66 31
TRUE	25 37

Fisher's Exact Test p-value: 0.00094

Pitch Type vs Strike	
NO	YES
Changeup	7 5
Curveball	33 15
Fastball	47 46
Slider	1 0
Splitter	3 2

Fisher's Exact Test p-value: 0.24725

Contingency Tables for Player 5

Swing vs Strike	
NO	YES
FALSE	54 20
TRUE	26 40

Fisher's Exact Test p-value: 8e-05

Pitch Type vs Strike	
NO	YES
Changeup	10 4
Fastball	42 31
Slider	28 25

Fisher's Exact Test p-value: 0.46139

Results - Swings

Contingency Tables for Player 1

Swing vs Pitch Type

	Changeup	Fastball	Slider
FALSE	16	41	31
TRUE	12	56	13

Fisher's Exact Test p-value: 0.00672

S/B Call vs Pitch Type

	Changeup	Fastball	Slider
Ball	16	30	21
Called Strike	0	11	10
Foul	6	26	4
Swinging Strike	6	30	9

Fisher's Exact Test p-value: 0.00457

Swing vs Pitch Location (Zone)

	1	2	3	4	5	6	7	8	9	11	12	13	14
FALSE	2	0	3	6	5	0	0	0	1	3	23	18	16
TRUE	2	3	4	0	6	7	2	6	4	2	9	4	26

Fisher's Exact Test p-value: 5e-04

S/B Call vs Pitch Type

	Changeup	Fastball	Slider
Ball	13	47	30
Called Strike	1	19	15
Foul	3	16	15
Swinging Strike	9	41	29

Fisher's Exact Test p-value: 0.60677

Contingency Tables for Player 2

Swing vs Pitch Type

	Changeup	Fastball	Slider
FALSE	14	66	45
TRUE	12	57	44

Fisher's Exact Test p-value: 0.89624

Swing vs Pitch Location (Zone)

	1	2	3	4	5	6	7	8	9	11	12	13	14
FALSE	3	2	4	5	2	4	3	5	5	12	25	31	16
TRUE	2	2	7	9	6	13	10	11	9	1	13	10	16

Fisher's Exact Test p-value: 5e-04

Contingency Tables for Player 3

Swing vs Pitch Type

	Changeup	Curveball	Fastball	Slider
FALSE	17	1	49	36
TRUE	16	2	42	31

Fisher's Exact Test p-value: 0.94263

S/B Call vs Pitch Type

	Changeup	Curveball	Fastball	Slider
Ball	15	1	36	23
Called Strike	2	0	13	13
Foul	5	1	12	7
Swinging Strike	11	1	30	24

Fisher's Exact Test p-value: 0.74923

Contingency Tables for Player 4

Swing vs Pitch Type

	Changeup	Curveball	Fastball	Slider	Splitter
FALSE	8	37	59	1	1
TRUE	4	11	50	0	4

Fisher's Exact Test p-value: 0.01091

Swing vs Pitch Location (Zone)

	1	2	3	4	5	6	7	8	9	11	12	13	14
FALSE	5	3	2	8	2	2	3	0	2	13	21	29	6
TRUE	0	0	4	4	8	5	5	3	4	3	6	7	13

Fisher's Exact Test p-value: 5e-04

Contingency Tables for Player 5

Swing vs Pitch Type

	Changeup	Fastball	Slider
FALSE	10	44	37
TRUE	6	54	20

Fisher's Exact Test p-value: 0.03894

Swing vs Pitch Location (Zone)

	1	2	3	4	5	6	7	8	9	11	12	13	14
FALSE	0	1	5	2	1	1	2	2	4	9	13	19	12
TRUE	4	3	4	2	7	4	5	5	5	1	14	6	6

Fisher's Exact Test p-value: 0.00235

Results - Hits

Contingency Tables
for Player 1

Hit vs Pitch Type		
Changeup	Fastball	Slider
FALSE	4	19
TRUE	0	7

Fisher's Exact Test p-value: 0.56688

Hit vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
FALSE	0	2	2	1	4	4	1	2	3	1	3	2	5	
TRUE	0	0	0	1	1	0	1	0	1	1	1	1	1	

Fisher's Exact Test p-value: 0.99449

Play Result vs Pitch Type		
Changeup	Fastball	Slider
flyout	2	5
groundout	1	10
hitbypitch	0	2
lineout	0	1
popout	0	1
single	0	7
strikeout	1	3
walk	1	2

Fisher's Exact Test p-value: 0.78311

Play Result vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
flyout	0	1	2	0	1	3	0	1	0	0	0	1		
groundout	0	1	0	0	2	1	1	1	0	3	1	2		
hitbypitch	0	0	0	0	0	0	0	0	0	1	0	0		
lineout	0	0	0	0	0	0	0	0	0	0	0	0		
popout	0	0	0	0	0	0	0	0	0	0	0	0		
single	0	0	0	1	1	0	1	0	1	1	1			
strikeout	0	0	1	1	0	0	1	0	1	0	1	2		
walk	0	0	0	0	0	0	0	0	0	1	3	0		

Fisher's Exact Test p-value: 0.67766

Contingency Tables
for Player 4

Hit vs Pitch Type		
Changeup	Curveball	Fastball
FALSE	3	4
TRUE	0	1

Fisher's Exact Test p-value: 0.35526

Hit vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
FALSE	0	0	1	1	4	2	1	3	1	0	4	2		
TRUE	0	0	0	2	1	3	2	0	0	0	1	0		

Fisher's Exact Test p-value: 0.36558

Play Result vs Pitch Type					
Changeup	Curveball	Fastball	Slider	Splitter	
double	0	1	2	0	1
flyout	0	1	4	0	1
groundout	1	1	5	0	0
hitbypitch	0	0	2	0	0
homeroon	0	0	1	0	0
lineout	0	0	1	0	0
single	0	0	5	0	1
strikeout	2	2	5	0	0
walk	0	0	6	0	0

Fisher's Exact Test p-value: 0.86657

Contingency Tables
for Player 2

Hit vs Pitch Type		
Changeup	Fastball	Slider
FALSE	4	24
TRUE	1	11

Fisher's Exact Test p-value: 0.22235

Hit vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
FALSE	0	0	5	3	3	4	6	3	0	6	5	5		
TRUE	0	0	1	0	0	2	1	4	2	0	2	1	1	

Fisher's Exact Test p-value: 0.83823

Contingency Tables
for Player 3

Hit vs Pitch Type		
Changeup	Curveball	Fastball
FALSE	5	0
TRUE	3	0

Fisher's Exact Test p-value: 0.56606

Hit vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
FALSE	1	2	3	1	5	3	4	4	1	0	0	6	5	
TRUE	2	0	1	3	1	2	0	3	0	2	1	1		

Fisher's Exact Test p-value: 0.25115

Contingency Tables
for Player 4

Hit vs Pitch Type		
Changeup	Fastball	Slider
FALSE	4	17
TRUE	0	8

Fisher's Exact Test p-value: 0.59742

Hit vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
FALSE	0	0	1	1	4	2	1	3	1	0	4	2		
TRUE	0	0	0	2	1	3	2	0	0	0	1	0		

Fisher's Exact Test p-value: 0.36558

Play Result vs Pitch Type					
Changeup	Fastball	Slider			
double	0	0	1	0	1
flyout	0	0	0	1	0
groundout	0	1	0	3	0
hitbypitch	0	0	0	0	0
homeroon	0	0	0	1	0
lineout	0	0	0	0	0
single	0	0	1	1	1
strikeout	0	0	1	0	2
walk	0	0	6	0	0

Fisher's Exact Test p-value: 0.12844

Play Result vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
FALSE	1	2	3	1	5	3	4	4	1	0	0	6	5	
TRUE	1	0	1	1	2	2	1	0	1	0	0	0	0	

Fisher's Exact Test p-value: 0.30544

Play Result vs Pitch Location (Zone)														
	1	2	3	4	5	6	7	8	9	11	12	13	14	
FALSE	0	0	0	1	1	0	0	0	0	0	0	0	0	
TRUE	0	0	0	0	0	0	0	0	0	0	0	0	0	

Fisher's Exact Test p-value: 0.08796

* see slide 16 for description of the results

Conclusion

Main Takeaways

- For all players there is an association between where the pitch was thrown to and if the batter swung
- For all players there was no significant association between where the pitch was thrown to and if a hit was given up
- For all other results, the association depended on the pitcher and some had significant results and others did not

Pitfalls

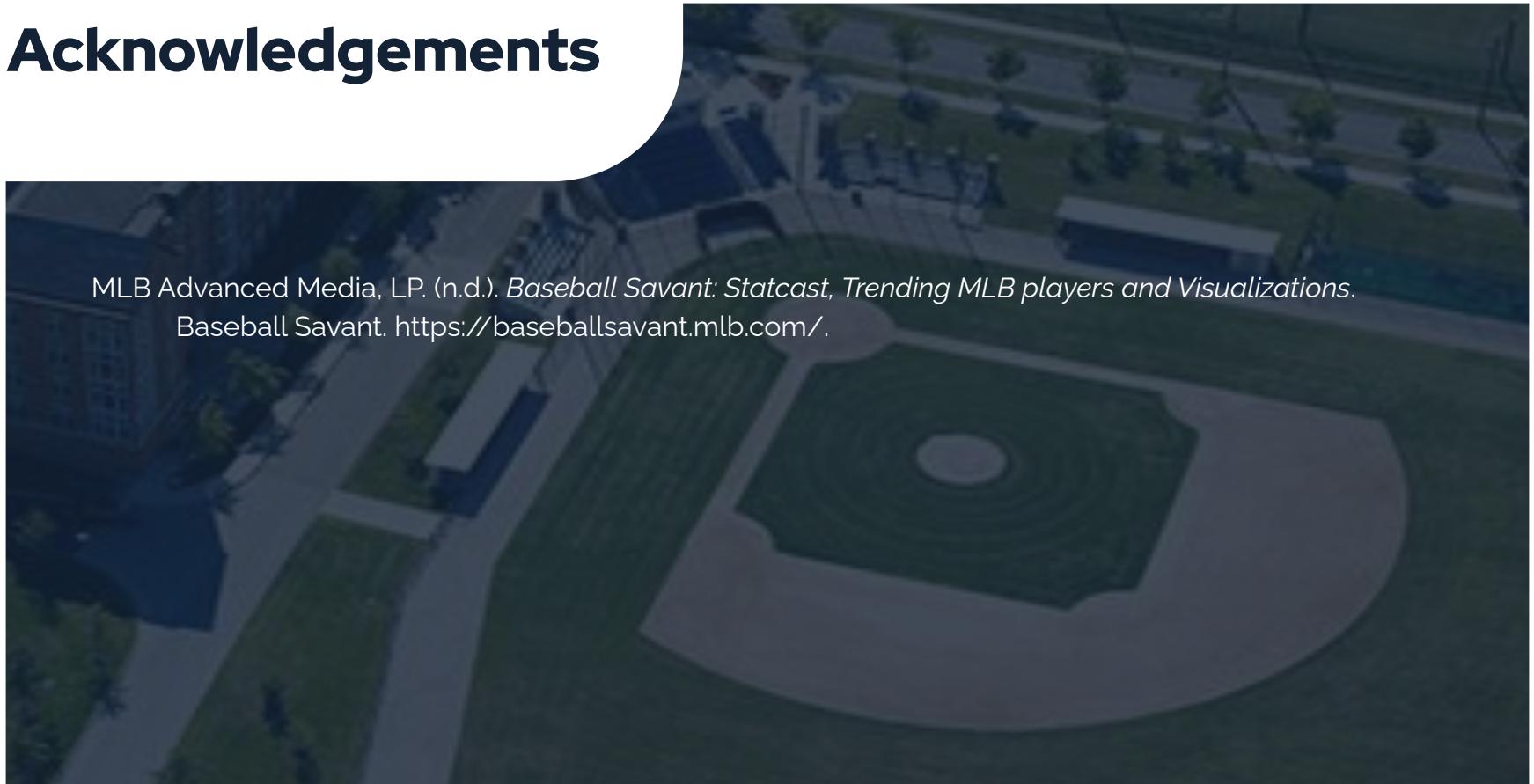
- The data set is relatively small which affects the results of the analysis
- For each pitcher approximately 200 pitches were considered but the number decreased when looking at hits given up where approximately 40 at bats were considered
- Did not take into account handedness of pitcher and batter

Future Research

- Consider more details of the movement of the pitch such as velocity, spin rate, vertical and horizontal break and the effects on swings induced and hits given up
- Compare the results to find ways for a pitcher to improve if a different pitcher is getting better results

Acknowledgements

MLB Advanced Media, LP. (n.d.). *Baseball Savant: Statcast, Trending MLB players and Visualizations.*
Baseball Savant. <https://baseballsavant.mlb.com/>.



Data Description

Variable Name	Possible Values	Description
S.B	Ball, Called Strike, Foul, Swinging Strike	The strike call as called by the umpire during the game. Recorded on the Shiny App.
Type	Fastball, Changeup, Curveball, Cutter, Slider, Splitter	The pitch type thrown. Recorded on the Shiny App (Rapsodo classification matches in the column Pitch.Type)
Result	A few examples include K, 1B, 4-3, F8, BB which were abbreviations used in baseball scoring	Denotes the result of the at-bat. If the value is NA then the at-bat did not end on that pitch. Recorded on the Shiny App
Is.Strike	Yes, No	Whether the ball landed in the strike as determined by Rapsodo
Strike.Zone.Side	Float value in inches	The horizontal location of the landing of the ball. From Rapsodo. 0 inches denotes the middle of the plate.
Strike.Zone.Height	Float value in inches	The vertical location of the landing of the ball. From Rapsodo. 0 inches denotes the ground.
Result_Descrip	strikeout, groundout, flyout, popout, single, double, triple, homerun, walk, hitbypitch, error	Simplifies the Result column and places values in groups. NA denotes the at-bat did not end on the pitch.
Hit	True, False	True if the Result Description is a hit (single, double, triple, homerun), and false if the Result Description is an out (strikeout, groundout, flyout, popout) or an error. NA denotes the at-bat did not end on that pitch or there was a walk or hitbypitch which is not included when calculating batting average.
Swing	True or False	True if the batter swung (either swinging strike or foul in the S.B column) and False if the batter did not swing (either ball or called strike in the S.B column)
Zone	Integers 1 through 14 (excluding 10)	Breaks the strike zone and outside the strike zone into categories. The dimensions for each category come from Baseball Savant. See slide 12 for a visual of the zones.

Shiny App and Rapsodo

[Save Chart](#) [New Pitcher](#)

New Pitch Pitches Thrown: 29
 New Batter Current Batter:
 New Inning Current Inning:

Balls vs Strikes

Pitch Type	Count
Balls	10
Strikes	5

Pitch Types Thrown

Pitch Type	Count
Changeup	74
Fastball	103.45
Slider	74.22

Current Pitcher:

MPH	SIB	Type	Result	Batter	Inning
71	Ball	Changeup		23	6
85	Swinging Strike	Fastball	PS	23	6
74	Swinging Strike	Slider		37	6
75	Swinging Strike	Changeup		37	6
85	Foul	Fastball		37	6
73	Foul	Changeup		37	6
76	Ball	Slider		37	6
84	Foul	Fastball		37	6
74	Ball	Changeup		37	6
75	Foul	Changeup		37	6
75	Called Strike	Slider	K	37	6
76	Ball	Slider		0	6
84	Foul	Fastball		0	6
74	Ball	Slider		0	6
78	Ball	Changeup		0	6
81	Swinging Strike	Fastball		0	6
74	Ball	Slider	Bb	0	6
Ball		Fastball		34	6
85	Called Strike	Fastball		34	6
83	Called Strike	Fastball		34	6
74	Swinging Strike	Changeup	F9	34	6
83	Swinging Strike	Fastball	1b	0	7
73	Ball	Slider		22	7
83	Swinging Strike	Fastball	F9	22	7
82	Called Strike	Fastball		6	7
73	Swinging Strike	Slider	F8	6	7
73	Ball	Slider		9	7
83	Called Strike	Fastball		9	7
74	Swinging Strike	Changeup	F7	9	7

Rapsodo PITCHING No Group [HOME](#) [SETTINGS](#) [TEAM MANAGEMENT](#) [Gyro](#) [Gyro Calibration](#) [CONNECTED](#) [98%](#)

SPIN DIRECTION
87.3 mph 01:06

Avg: 83 Max: 87 Gyro Degree: 18°

VERTICAL BREAK 17.8" **HORIZONTAL BREAK** 11.8"

SPIN EFFICIENCY

TOTAL SPIN	TRUE SPIN
1966	95 % 1872

STRIKEZONE

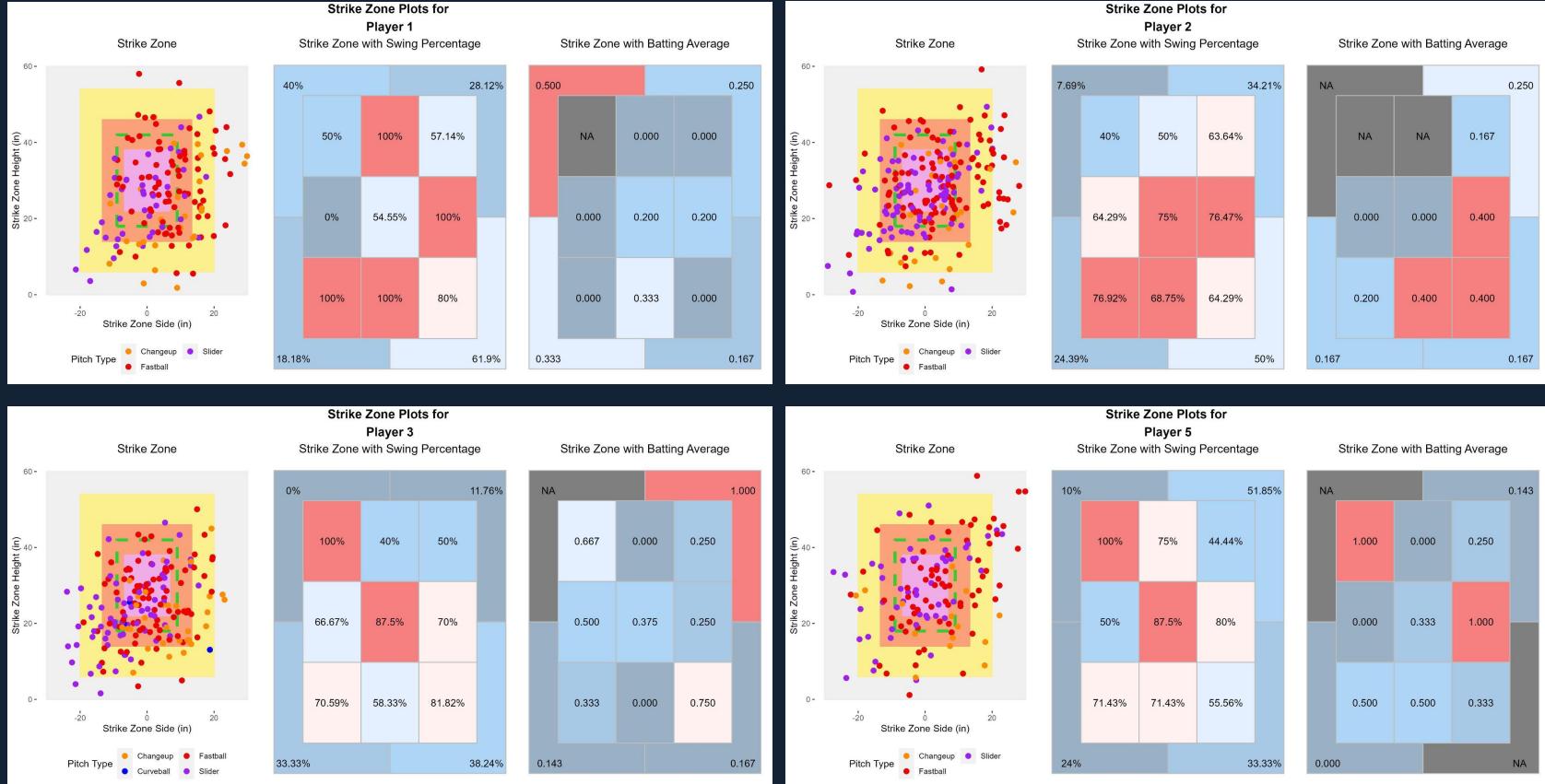
RELEASE
RELEASE ANGLE: 0° RELEASE HEIGHT: 5.9'

HORIZONTAL ANGLE -2° RELEASE SIDE: 0.6'

PITCH COUNT 37 **PITCH TYPE** FB 2S CT CB SL CH SP KN OTH

FB 4 SEAM 2S 3 SEAM FB CUTTER CURVE BALL SLIDER CHANGE UP SPITTER KNUCKLE BALL OTHER

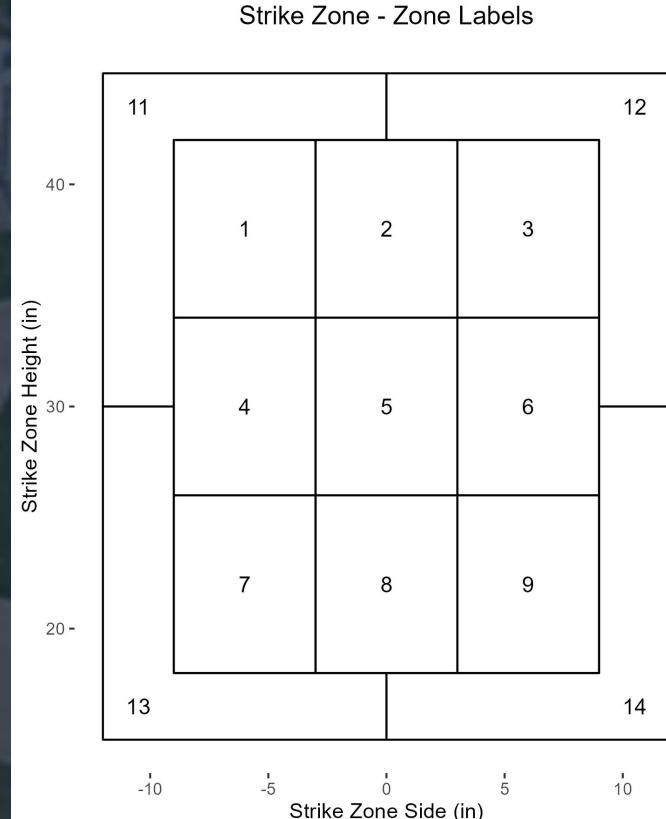
Initial Visualizations - Remaining Players



Layout of Zones

Notes

- Zones 1-9 are located within the strike zone
- Zones 11-14 are located outside the strike zone
- Zones 11-14 extend past the rectangle shown
- There is no zone 10



Results - Command (details)

Swing vs Strike

- At a significance level of 0.05 all five players had a statistically significant association between whether the pitch landed in the strike zone and if the batter swung.
- When looking at the contingency tables, all the players induced fewer swings on pitches outside of the strike zone and more swings on pitches inside the strike zone.

Pitch Type vs Strike

- Players 2,3,4, and 5 did not have a significant association between pitch type thrown and if it landed in the strike zone.
- Player 1 did have significant association between pitch type and strike.
- Looking at the contingency table, Player 1 threw 4 changeups for strikes and 24 not for strikes. This most likely had a large impact on the p-value of the Fisher's Exact Test.
- While Player 4 did not have a statistically significant association, the p-value was smaller than the other players with non significant results. It can be seen that that Player 4 had fewer curveballs that were strikes.

Results - Swings (details)

Swing vs Pitch Type

- Players 1, 4, and 5 had significant associations between what pitch type was thrown and whether the batter swung
- Players 2 and 3 did not have a significant association

S/B Call vs Pitch Type

- Players 1, 4, and 5 had significant associations between what pitch type was thrown and whether the batter swung
- Players 2 and 3 did not
- The results are the same as comparing swing vs pitch type
- S/B call gives more detailed information than just whether the batter swung or not

Swing vs Pitch Location

- All players had a significant association between pitch location and whether a batter swung
- This mirrors the results from the swing vs strike tables
- However, pitch location gives a more detailed description of where the ball landed

Results - Hits (details)

Hit vs Pitch Type

- No player had a significant association between pitch type and whether a hit was given up

Hit vs Pitch Location

- No player had a significant association between pitch location and whether a hit was given up

Play Result vs Pitch Type

- Player 2 was the only player that had a significant association between the result of the play and the pitch type thrown

Play Result vs Pitch Location

- Only player 5 had a significant (at a level of 0.1) association between the result of the play and the location of the pitch thrown

