

Examining the Third Time-Through-The-Order: A Statistical Approach

Description: In baseball there have been numerous instances where a pitcher exhibits exceptional performance through the initial five innings. Then seemingly nowhere, the manager decides to remove that pitcher out from the game, right as the leadoff man was due up to face the starter for the third time. This analysis looks to explore this phenomenon and evaluate whether or not there is enough empirical evidence to support the theory: That when a pitcher faces a batter for the third time, his performance is prone to decline.

Step 1: Data Collection

By utilizing the Baseball Savant API through the package *baseballr* in R Studio, I will be able to gather data containing all play-by-play information. The importance of having play-by-play information is so I can get a full picture of what is happening once a pitcher faces a lineup for the third time. Cumulative statistics like FIP or wOBA against will be utilized in this study, however, understanding why the pitcher ended up with a certain FIP or wOBA is paramount.

By looking pitch by pitch, I will be able to understand the thought process of both the pitcher and hitter and be able to make a more informed decision on the theory.

Below is the code I used to gather data from the Baseball Savant API.

```
5 library(baseballr, dplyr)
6 # Data collection using the baseballr package
7 # Due to some loading issues I had to break up the data collection in 3-5 day intervals and didn't select the entire season.
8 # These still provide 107,398 pitches which is enough data for the study.
9
10 data1 <- scrape_statcast_savant(start_date = '2023-04-01', end_date = '2023-04-07', player_type = 'pitcher')
11 data2 <- scrape_statcast_savant(start_date = '2023-05-01', end_date = '2023-05-04', player_type = 'pitcher')
12 data3 <- scrape_statcast_savant(start_date = '2023-06-01', end_date = '2023-06-04', player_type = 'pitcher')
13 data4 <- scrape_statcast_savant(start_date = '2023-08-01', end_date = '2023-08-04', player_type = 'pitcher')
14 data5 <- scrape_statcast_savant(start_date = '2023-08-10', end_date = '2023-08-14', player_type = 'pitcher')
15 data6 <- scrape_statcast_savant(start_date = '2023-05-10', end_date = '2023-05-14', player_type = 'pitcher')
16 data <- data1 %>% bind_rows(. , data2,data3,data4,data5,data6)
```

The function *scrape_statcast_savant* gathers every pitch from the entire league on any given day. This results in a lot of information being delivered to my machine and unfortunately results in some loading issues. That is why I broke up the data collection to 3-5 days over the course of the regular season to eliminate biases from one particular point in the season.

Utilizing the *dplyr* package I am able to bind all the gathered data frames seamlessly into one large data frame.

The first part of this study will be looking at all pitchers as a whole and then I will explore a handful of pitchers of varying skill level.

Step 2: Data Preprocessing

The data that was returned to my new data frame **data** is the raw data collected by MLB Statcast. Now I have to go through the process of cleaning and manipulating some variables so I can perform my statistical analysis. Below is the code I used to clean and create/edit some columns which will be useful in my study.

```

20 # 2. Data Preprocessing:
21
22 new_names <- c('Date' = 'game_date', 'RelSpeed' = 'release_speed', 'RelSide' = 'release_pos_x', 'RelHeight' = 'release_pos_z', 'PitcherName' = 'player_name',
23 'Zone' = 'zone', 'Batterside' = 'stand', 'PitcherThrows' = 'p_throws', 'Balls' = 'balls', 'Strikes' = 'strikes', 'HorzBreak' = 'pfx_x',
24 'InducedvertBreak' = 'pfx_z', 'PlateLocSide' = 'plate_x', 'PlateLocHeight' = 'plate_z', 'on3B' = 'on_3b', 'on2B' = 'on_2b', 'on1B' = 'on_1b',
25 'outs' = 'outs_when_up', 'Inning' = 'inning', 'Top.Bottom' = 'inning_topbot', 'ExitSpeed' = 'launch_speed', 'Angle' = 'launch_angle',
26 'SpinRate' = 'release_spin_rate', 'Extension' = 'release_extension', 'PitchofPA' = 'pitch_number', 'TaggedPitchType' = 'pitch_name',
27 'SpinAxis' = 'spin_axis')
28
29 Pitch_calls <- c('swinging_strike' = 'Strikeswinging', 'foul' = 'Foulball', 'ball' = 'ballcalled', 'called_strike' = 'Strikecalled',
30 'hit_into_play' = 'InPlay', 'blocked_ball' = 'ballcalled', 'foul_tip' = 'Foulball', 'swinging_strike_blocked' = 'Strikeswinging',
31 'hit_by_pitch' = 'HitbyPitch', 'foul_bunt' = 'BuntFoul', 'missed_bunt' = 'BuntSwingingStrike', 'bunt_foul_tip' = 'BuntFoul')
32
33 Pitch_Types <- c('4-Seam Fastball' = 'Fastball', 'Split-Finger' = 'Splitter')
34
35 Play_Results <- c('Field_out' = 'out', 'Grounded_into_double_play' = 'out', 'Force_out' = 'out', 'Home_run' = 'Homerun',
36 'Hit_by_pitch' = 'HitbyPitch', 'Strikeout_double_play' = 'Strikeout', 'Field_error' = 'Error', 'sac_fly' = 'Sacrifice',
37 'Sac_bunt' = 'Sacrifice', 'Fielders_choice' = 'Fielderschoice', 'Sac_fly_double_play' = 'Sacrifice',
38 'double_play' = 'out', 'Fielders_choice_out' = 'Fielderschoice', 'other_out' = 'out')
39
40
41 DF <- data %>% rename(any_of(new_names)) %>% mutate(Count = paste0(Balls, '-', Strikes)) %>%
42 mutate(PitcherThrows = recode(PitcherThrows, 'L' = 'Left', 'R' = 'Right'), Batterside = recode(Batterside, 'L' = 'Left', 'R' = 'Right'),
43 TaggedPitchType = recode(TaggedPitchType, !!!Pitch_Types), PitchCall = recode(description, !!!Pitch_calls), PlayResult = str_to_title(events),
44 PlayResult = recode(PlayResult, !!!Play_Results)) %>%
45 filter(, !(TaggedPitchType %in% c('other', 'Screwball', 'Pitch out')) & PitchCall != 'pitchout' & RelSpeed > 0) %>%
46 mutate(ScoreDiff = case_when(Top.Bottom == 'Top'~away_score-home_score, T~home_score-away_score)) %>%
47 mutate(Swing = case_when(PitchCall %in% c('Strikeswinging', 'Foulball', 'InPlay', 'BuntFoul', 'BuntSwingingStrike')~'Yes', T~'No')) %>%
48 mutate(ManOn1 = case_when(is.na(on1B) == TRUE~'No', T~'Yes'), ManOn2 = case_when(is.na(on2B) == TRUE~'No', T~'Yes'),
49 ManOn3 = case_when(is.na(on3B) == TRUE~'No', T~'Yes')) %>%
50 mutate_at(vars(ManOn1:ManOn3), ~ifelse(. == 'No', 0, 1)) %>%
51 mutate(PCAT = case_when(TaggedPitchType %in% c('Sink', 'Cutter', 'Fastball')~'FAST',
52 TaggedPitchType %in% c('Changeup', 'Splitter', 'Eephus')~'OFF', T~'BREAK')) %>%
53 arrange(Date, Inning, at_bat_number, PitchofPA) %>% group_by(Date, PitcherName, batter) %>%
54 mutate(TimeSeeing = rleid(at_bat_number)) %>% ungroup()

```

The following are explanations of the changes and additions I made to the raw data frame **data**.

1. *new_names* is a vector that allows me to seamlessly rename the column names in the data frame. I do this so it follows an ordered pattern without the underscores and so I can remember what each column is called.
2. *Pitch_Calls* is a vector that allows me to utilize the function *recode* from *dplyr* so I can rename all of the values in the description column. There are a couple instances where some values are named differently but for the purposes of this study can be classified as one value. For example, 'swinging_strike_blocked' is the same as 'swinging_strike' which is why I recoded them both to be 'StrikeSwinging'.
3. *Pitch_Types* is also a vector that allows me to recode some of the pitch types in the file.
4. *Play_Results* is the last vector that allows me to recode the descriptions of the play outcomes such as Single, Double, Triple or Homerun.
 - a. I have broken up Pitch Calls and Play Results into two different columns so they are separated which will be helpful in seeing how a pitcher got to a certain play result.
5. The last important addition to the data frame is the *TimeSeeing* column. This column tracks the number of times a pitcher has seen a batter in any given game. This isn't included in the raw data, so I had to utilize the *data.table* package with the function *rleid*. I arranged the data frame, so it was in order from the start of the game to finish, then I grouped the data frame by Date, Pitcher, and Batter and on the 'at_bat_number' column I ran *rleid* to count the number of times a pitcher faced a batter in any one game.

Step 3: Statistical Analysis

The first step in my statistical analysis was to get a feel for the data and run some preliminary calculations. I wanted to first confirm that I had enough data and with 1,371 games with over 27,000 individual plate appearances, I felt that was sufficient. Additionally, there were 189 total starters included in this study.

My findings from the preliminary analysis were that the average starter threw 6 and 1/3 innings with 20 total plate appearances each game. Also, my initial hypothesis that starters in the MLB don't normally go through the lineup a 3rd time seems to be holding true as the average starter faces the same batter 2.7 times per game.

Below is a table displaying these statistics:

Preliminary Statistics							
Total Games	Total PAs	LHP Starters	RHP Starters	Total Starters	AVG Starter IP	AVG Starter PAs	AVG Starter Lineup
1371	27393	55	134	189	6.1	20.4	2.7

Next I analyzed some performance statistics in which I feel tell a good story of pitcher performance. All of the statistics I looked at here are averages from all the pitchers grouped together based on their time through the order.

Pitch/PA : The number of pitches a pitcher threw per plate appearance.

Strike % : The number of strikes a pitcher threw divided by the total number of pitches thrown.

Zone % : The number of pitches a pitcher throws inside the strike zone divided by the total number of pitches thrown.

Chase % : The number of pitches a swung at on pitches thrown outside the strike zone divided by the total number of pitches thrown outside the strike zone.

Strikeout % : The number of strikeouts a pitcher had divided by the total number of plate appearances.

Walk % : The number of walks a pitcher had divided by the total number of plate appearances.

xwOBA : The average expected wOBA from a pitch thrown.

RA : The average number of runs allowed

FIP : The average FIP pitchers had which is a cumulative statistic comparable to ERA.

Below is a table of my findings:

Performance Statistics									
Time Seeing	Pitch/PA	Strike %	Zone %	Chase %	Strikeout %	Walk %	xwOBA	RA	FIP
1	3.9	64%	49%	28%	24%	9%	0.366	0.7	3.42
2	3.8	64%	50%	28%	20%	8%	0.379	1.0	3.67
3	3.8	64%	48%	30%	20%	7%	0.402	0.6	3.54
4	3.7	66%	54%	28%	17%	6%	0.427	0.3	3.41

After going through the data, I realized that I should be comparing the 1st time a pitcher sees a particular batter to the 3rd time a pitcher sees the batter. The first time should be the benchmark where a batter should not have expectations of what will happen compared to his 3rd time seeing the pitcher.

What surprised me in these statistics was how close several of this descriptive statistics turned out to be. For example, Strike %, Zone %, and Chase % all were virtually the same from the 1st time through to the 3rd time through. Those statistics can be used as explanatory variables for the xwOBA and FIP statistics.

If you notice, the 1st time seeing a batter a pitcher's xwOBA is on average .366, then the 3rd time through his xwOBA goes to on average .402. From first glance, that jump of nearly .040 seems like a pretty large gap which would support the initial hypothesis that pitchers shouldn't face a batter more than twice.

I ran a t-test on the xwOBA when pitchers face a batter for the 1st time versus when they face one a 3rd time and to my expectations, I was able to reject the null hypothesis at a significance level of 5%. Since I failed to reject the null hypothesis that the two means were the same, based on my random sample of 2000 rows for each category, there is a statistically significant difference.

Paired t-test

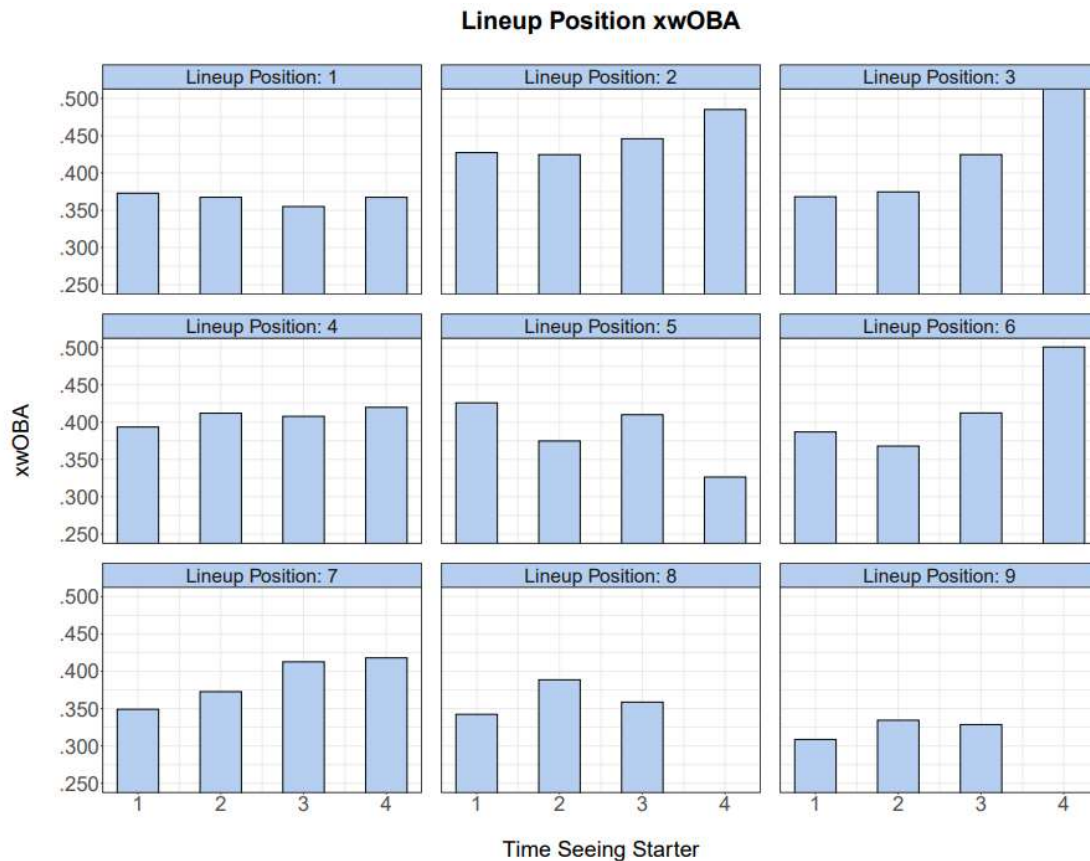
```
data: sample_FirstTime$estimated_woba_using_speedangle and sample_ThirdTime$estimated_woba_using_speedangle
t = -3.2241, df = 1999, p-value = 0.001284
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 -0.06844994 -0.01667206
sample estimates:
mean difference
 -0.042561
```

Only looking at the xwOBA and FIP increases, I would have expected Strike % and Chase % to decrease as when those statistics are higher it tends to end up with favorable outcomes for the pitcher. Therefore, I will now start to look at what is causing that increase in xwOBA and determine if there are any relationships between those factors and the number of times a pitcher faces a particular batter.

I decided to look at hitters specifically in their lineup order. So, there are 9 different lineup positions, and as to be expected, throughout a game the 1-3 hitters tend to get more plate appearances. In the table below, I show the percentages as to how often a certain lineup position gets a plate appearance. Then I also show the expected wOBA of those lineup positions. Teams strategize their lineups, so their stronger hitters are in the beginning and their weaker hitters tail off at the end. In the table it can be inferred that the 2-5 hitters are typically the strongest.

Lineup Position		
Lineup	Seen	xwOBA
1	13%	0.366
2	12%	0.432
3	12%	0.388
4	12%	0.404
5	11%	0.402
6	11%	0.385
7	10%	0.369
8	10%	0.364
9	9%	0.322

Next I look at how well those hitters in the order of their lineup perform against starting pitchers facing them more than once.



Looking at these graphs, one can see that the 2-4 positions really continue to get better as the game goes on. Their expected wOBA steadily increases the more they see the starting pitcher. This brings me to my next point, where I am now starting to believe managers will remove a pitcher prior to their 3rd time facing the 2-4 hitters.

The expected wOBA as well as the other statistics I have mentioned before and in the table on the next page do show clear evidence that hitters performed better the 3rd time around.

Lineup Position Statistics					
Lineup	Time	xwOBA	wOBA	BABIP	ISO
1	1	.373	.337	.202	.156
	2	.367	.339	.203	.129
	3	.355	.319	.212	.146
	4	.367	.385	.214	.262
2	1	.427	.325	.185	.142
	2	.425	.357	.217	.183
	3	.446	.376	.190	.229
	4	.485	.477	.250	.350
3	1	.368	.326	.191	.139
	2	.375	.349	.204	.184
	3	.425	.338	.219	.150
	4	.702	.690	.100	.900
4	1	.393	.338	.187	.151
	2	.412	.345	.196	.174
	3	.408	.330	.190	.175
	4	.420	.239	.222	.111
5	1	.426	.339	.198	.159
	2	.375	.308	.199	.126
	3	.410	.356	.177	.231
	4	.326	.417	.333	.333
6	1	.387	.325	.170	.192
	2	.368	.297	.166	.145
	3	.412	.333	.205	.147
	4	.501	.537	.500	.250
7	1	.349	.328	.184	.150
	2	.373	.319	.195	.134
	3	.413	.342	.195	.162
	4	.418	.000	.000	.000
8	1	.342	.300	.184	.134
	2	.389	.334	.182	.182
	3	.359	.271	.170	.128
	4				
9	1	.309	.269	.171	.098
	2	.334	.305	.191	.123
	3	.329	.286	.202	.146
	4				

Step 4: Specific Pitcher Analysis

Now that I have evidence starters as whole tend to perform worse against hitters the more often they face them, I want to look at specific starters and see if this pattern persists. My initial thought is that the rule “Starters should not face a hitter more than twice in a game” should not be a general rule of thumb. I believe it can pertain to some pitchers, but others would perform just as well the 3rd time around as they did the 1st.

For this part, I have selected 6 pitchers who had the highest WAR amongst starters in 2023, as well as 6 pitchers who were in the middle of the pack in terms of WAR and 6 from the bottom of the pack.

Group A:

Zack Wheeler: WAR – 5.9

Spencer Strider: WAR – 5.5

Kevin Gausman: WAR – 5.3

Sonny Gray: WAR – 5.3

Gerrit Cole: WAR – 5.2

Zac Gallen: WAR – 5.2

Group A Pitcher Statistics										
Pitcher	Time Seeing	Pitch/PA	Strike %	Zone %	Chase %	Strikeout %	Walk %	xwOBA	RA	FIP
Zack Wheeler	1	4.2	67%	51%	34%	33%	4%	0.332	0.7	2.70
	2	3.9	65%	50%	31%	22%	4%	0.333	1.1	3.29
	3	3.9	63%	48%	31%	24%	7%	0.402	0.6	3.43
Spencer Strider	1	4.0	71%	55%	35%	40%	7%	0.409	0.7	2.57
	2	4.2	65%	49%	31%	35%	9%	0.381	1.1	2.98
	3	3.8	68%	50%	38%	34%	5%	0.378	0.7	3.43
Kevin Gausman	1	4.2	64%	47%	31%	36%	10%	0.431	0.8	3.12
	2	3.8	67%	47%	35%	30%	5%	0.402	1.0	3.16
	3	4.0	66%	48%	32%	25%	7%	0.381	0.4	3.18
Sonny Gray	1	3.7	66%	51%	29%	29%	5%	0.322	0.5	2.57
	2	3.7	64%	49%	28%	23%	8%	0.378	0.6	3.10
	3	3.8	61%	47%	25%	20%	10%	0.361	0.6	3.37
Gerrit Cole	1	4.2	65%	50%	28%	29%	7%	0.368	0.8	3.47
	2	3.9	68%	52%	30%	27%	4%	0.361	0.4	2.92
	3	3.9	65%	45%	32%	24%	6%	0.359	0.5	3.26
Zac Gallen	1	3.8	67%	50%	32%	30%	5%	0.389	0.6	2.85
	2	3.9	63%	49%	29%	25%	6%	0.368	0.7	3.20
	3	3.9	63%	48%	29%	22%	6%	0.453	1.0	3.52

To my surprise, these 6 pitchers seemed to follow the pattern that removing a pitcher after they have faced the lineup twice is the right decision. These were the 6 best pitchers in MLB last season and all of them have worse Strikeout % and higher FIPs than when they faced a lineup the first time.

Group B:

Merrill Kelly: WAR – 3.2

Logan Gilbert: WAR – 3.2

Miles Mikolas: WAR – 3.1

Freddy Peralta: WAR – 3.0

Jose Berrios: WAR – 3.0

Sandy Alcantara: WAR – 2.9

Group B Pitcher Statistics										
Pitcher	Time Seeing	Pitch/PA	Strike %	Zone %	Chase %	Strikeout %	Walk %	xwOBA	RA	FIP
Merrill Kelly	1	4.1	63%	45%	32%	26%	11%	0.332	0.6	3.63
	2	3.7	65%	46%	34%	25%	8%	0.398	0.9	3.24
	3	3.7	65%	46%	32%	27%	8%	0.417	0.7	3.57
Logan Gilbert	1	4.1	67%	55%	27%	27%	4%	0.380	1.0	3.65
	2	3.6	68%	54%	33%	22%	5%	0.340	0.9	3.58
	3	3.7	65%	50%	30%	25%	5%	0.409	0.6	3.33
Miles Mikolas	1	3.8	68%	57%	27%	18%	5%	0.395	1.1	3.64
	2	3.6	69%	55%	31%	14%	3%	0.352	0.6	3.28
	3	3.7	65%	53%	25%	15%	5%	0.460	1.2	3.92
Freddy Peralta	1	4.3	64%	49%	29%	32%	8%	0.398	0.9	3.89
	2	4.1	63%	47%	30%	33%	7%	0.360	1.2	3.46
	3	4.1	61%	45%	27%	25%	10%	0.330	0.2	3.28
José Berrios	1	3.8	65%	51%	28%	24%	7%	0.373	0.8	3.56
	2	3.9	65%	48%	32%	24%	5%	0.382	0.8	3.52
	3	3.5	65%	48%	32%	23%	7%	0.445	0.8	3.48
Sandy Alcantara	1	3.6	69%	53%	35%	20%	4%	0.363	0.8	3.63
	2	3.6	68%	53%	33%	21%	6%	0.363	0.9	3.40
	3	3.5	68%	48%	36%	18%	10%	0.377	1.1	3.73

This group of pitchers don't follow a clear-cut pattern as we saw with Group A. For example, from these statistics, it would appear that Peralta actually improves the more he faces a batter. The interesting part of that is out of these 6 pitchers he gets pulled before he faces a lineup for the 3rd time the most. Still, he faced 140 batters at least 3 times over the 2023 season, but it would be interesting to see if this boils down to a team philosophy choice.

While Merrill Kelly's FIP slightly improved the 3rd time around, his xwOBA went from .332 to on average a .417. That's a huge gap between performance which can explain why a manager may be inclined to remove him after he's gone through the lineup twice.

Group C:

Dean Kremer: WAR – 1.5

Lucas Giolito: WAR – 1.0

Patrick Corbin: WAR – 0.9

JP Sears: WAR – 0.6

Lance Lynn: WAR – 0.5

Jordan Lyles: WAR – 0.2

Group C Pitcher Statistics										
Pitcher	Time Seeing	Pitch/PA	Strike %	Zone %	Chase %	Strikeout %	Walk %	xwOBA	RA	FIP
Dean Kremer	1	4.2	64%	52%	26%	21%	7%	0.383	0.9	4.47
	2	3.9	64%	50%	28%	21%	6%	0.432	1.1	3.74
	3	3.8	62%	47%	24%	22%	10%	0.367	0.4	3.28
Lucas Giolito	1	4.0	64%	49%	27%	29%	8%	0.440	1.1	4.26
	2	4.0	62%	47%	27%	23%	11%	0.423	1.4	4.40
	3	4.1	64%	49%	29%	25%	6%	0.320	0.8	3.48
Patrick Corbin	1	3.6	63%	49%	30%	19%	6%	0.452	1.1	4.08
	2	3.8	60%	48%	24%	15%	9%	0.372	1.1	4.15
	3	3.6	61%	48%	28%	12%	7%	0.426	1.0	3.89
JP Sears	1	4.0	63%	49%	29%	25%	9%	0.398	0.8	4.43
	2	3.9	67%	53%	30%	20%	5%	0.380	1.2	4.09
	3	3.8	66%	52%	30%	19%	7%	0.374	0.6	3.59
Lance Lynn	1	4.1	63%	47%	28%	26%	9%	0.415	1.3	4.85
	2	3.7	64%	48%	28%	23%	8%	0.368	1.0	3.70
	3	3.9	64%	47%	28%	22%	7%	0.438	1.2	4.09
Jordan Lyles	1	3.9	64%	49%	28%	19%	7%	0.357	1.1	4.30
	2	3.7	61%	46%	27%	14%	7%	0.370	1.9	4.48
	3	3.6	63%	49%	29%	14%	4%	0.420	0.9	3.81

Similar to Group B, these pitchers don't follow a clear-cut pattern that they should be removed from the game prior to the 3rd time around. However, this is where game management comes into play because these pitchers nearly all have their best FIPs that are higher than Group A's worst FIPs. Even though it would appear that these pitchers are improving as the game continues, they still pose a threat of allowing a run.

Step 5: Summary and Conclusion:

This study sought out to examine whether managers should remove their starting pitcher before the third time through the lineup. Utilizing Baseball Savant, I was able to retrieve raw pitch by pitch data and examine players as a whole as well as by themselves.

The main driver of 'success' I used was expected wOBA which has been used widely throughout baseball as a measure of success. By looking at xwOBA we are able to see what should have happened given a particular hit profile. I like to look at this rather than wOBA because defenses

around the league are not made equal. Some pitchers have the luxury of having an all-star defense, while others do not. Expected wOBA is meant to remove that variable from the equation as a whole. Several other statistics such as Strikeout%, Chase %, and FIP were used to measure a pitcher's success.

Overall, this was an interesting study, and my conclusion is that there is merit to pulling a pitcher prior to the 3rd time through the lineup. We saw the best pitchers in MLB struggle against hitters for their 3rd time. However, I do not believe this should be a rule of thumb for all pitchers. As I showed when I analyzed pitchers individually, some pitchers showed improvement as the game progressed. Freddy Peralta routinely gets pulled from the game prior to the 3rd time through. This may be due to his pitch count, however, the times he does face a batter for the 3rd time he showed improvement.

For these reasons, I believe this should not be a strategy a team poses to all their pitchers but takes a deep look at their rotation and make that call on a case-by-case basis.