Name: Vy Vo, Hieu (Calvin) Hoang, Khoa Tran
Date: 6/10/2024

Github Link: https://github.com/calvinhoang203/MLB-Analysis

Pitch by pitch: An analysis of Shohei Ohtani pitching evolution over the years

Introduction:

Baseball is a fairly popular sport in American culture, yet, amusingly, our group is completely in the dark about it and doesn't even know the rules of the game. However, we have heard of a Japanese pitcher entering the MLB, and we decided to focus our research on the only knowledge we have about baseball: Shohei Ohtani. Using MLB data gathered from Savant, we aim to analyze the various pitches used by Ohtani. As a baseline for comparison, we decided to compare his record to one of the top current pitchers, Gerrit Cole. Our general goal is to learn a little more about baseball and its players. Ultimately, through data analysis, we want to understand Shohei Ohtani's pitches, track his progress over the years, and identify potential areas of focus for improvement.

Background Information:

From the B/R Report, we identified the current standing of Shohei Ohtani as ranked 21st. We accessed the players' game records through CSV files obtained from the Savant website. For this report, we will be looking at Ohtani's data from the years 2018, 2020, 2021, 2022, and 2023. A downloadable CVS file can be found on our GitHub called Ohtani_2018-2023. This file contains essential variables such as pitch types, release speed, and more, which were utilized in our analysis.

We want to address a crucial point: we do not have data for Ohtani in 2019. Following our research, we discovered that Ohtani suffered from "a tear in the ulnar collateral ligament in the elbow of his pitching arm", as reported by The New York Times. Subsequently, in 2020, the year following his injury, Ohtani pitched in only 80 games, likely indicative of his ongoing recovery process. This doesn't significantly affect our analysis, however, it is still very interesting to see his pitch style from before the injury and after.

Methodology:

1. Data Scraping: To get the Top 25 pitchers of 2023, we scraped the data from the B/R website.

2. Contingency Table: We used to analyze pitch distribution and which pitch was most used per year.
3. Line Graph: We used to find general trends of pitches over the years.
4. Histogram: We utilized a 3D histogram to look at speed distribution for different pitches.
5. HeatMap: We used heatmap to look for different trends in pitches for right-handed batters vs left-handed batters
6. Scatterplot: We used 3D scatter plots to look at the relationship between release position VS speed = pitch's result

Data Scraping:

**** Please run the code for all top 25****

We used data scraping to scrape the top 25 pitchers in 2023 from the B/R website. We noticed that the fan-favorite Shohei Ohtani is ranked 21. This peaks our interest in how Shohei Ohtani's pitches differ from the best pitcher, according to the source, Gerrit Cole.

Contingency Table and Line Graph Analysis:

We created a contingency table analyzing Ohtani's different pitch types over the years. We found that his pitching style changed a lot over the 5 years. Firstly, in 2018, Ohtani only used four pitches, but in recent years, he has developed a larger variety of pitches. Furthermore, Ohtani's main pitch in 2018 was a fastball, followed by the sweeper. The trend continued until 2022, when his main pitch switched to sweeper, with fastball becoming secondary. While contingency tables do a good job showing us trends, we decided to create a line plot for better visualization (see Figure 1). From the line plot, we can directly see the switch in Ohtani's main pitch in 2022, where the lines cross: the gray line drops while the orange line rises. Additionally, we can see his preferences for his less frequently used pitches.

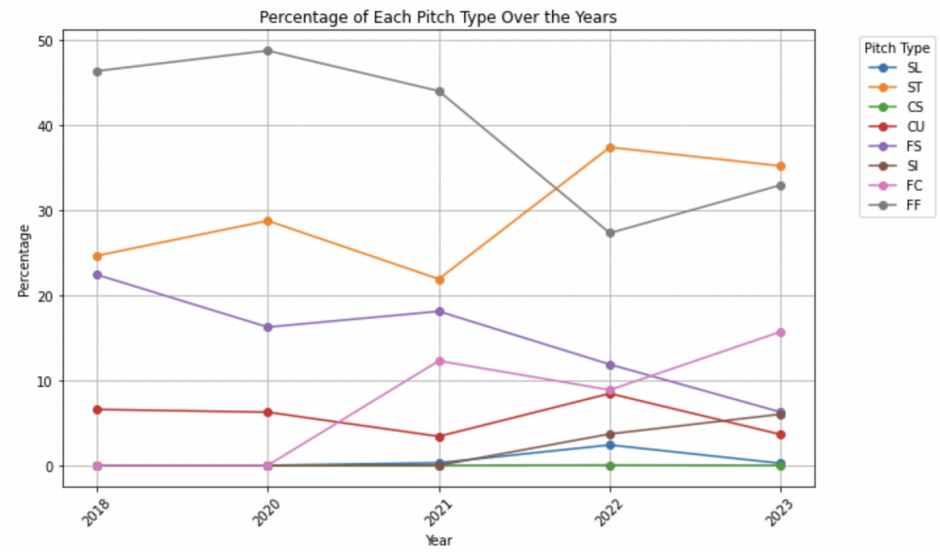

Figure 1

****Please run the Contingency Table code. It features an interactive dropdown table.****

To further analyze possible reasons for Ohtani's change in pitch style, we decided to create columns Strike and Non-Strike. The Strike column consists of pitches that game rules consider strikes such as swinging strike, called strike, strike, swinging strike blocked, foul, foul tip, and foul bunt. The Non-Strike columns are ball, blocked ball, hit by pitch, foul bunt, and hit into play. The columns record the percentage of pitches that are strikes and non-strikes for each pitch in the year. From the Strike percentage, we can see the reason why Ohtani transitioned from fastball to sweeper is because of the higher percentage. In 2018, Ohtani's fastball struck 45.82% of the time while his sweeper does much better at 50.95%. The trend of his sweeper outperforming his fastball is consistent in 2018, 2021, 2022, and 2023. From this, we can conclude that Ohtani's decision to switch to sweeper was a well-calculated transition.

Additionally, from our contingency table, we can see Ohtani's general growth. In 2018, his strike percentage was 45.82%, increasing to 46.15% in 2020, 47.42% in 2021, slightly decreasing to 47.21% in 2022, and rising to 52.9% in 2023. From this glance, we can see his consistent accuracy growth as a pitcher over the years.

Histogram Analysis:

We created a 3D histogram where X represents the year, Y represents speed, and Z represents density to examine Ohtani's overall gameplay over the years. The 3D plot provided valuable insight into his pitch speed, allowing us to observe another growth Ohtani had over time.

*** *Please run the code for the 3D Histogram graph* ***

The histogram shows a lot of the same information as the contingency table, such as which pitches Ohtani used each year and how much he used them. However, the addition of speed provides another dimension for comparison. Notably, Ohtani's fastball, represented by the yellow bell curves, stands out the most. In 2018, Ohtani's fastball averaged around 96 mph. In 2019, this number significantly decreased to around 93 mph due to his recovery from injury and lack of gameplay. By 2021, he had returned to an average of 95 mph. In 2022 and 2023, his fastball speed significantly grew to an average of 97 mphThe histogram shows that Ohtani has been steadily improving his pitch speed over the years, further highlighting his consistent growth as a pitcher.

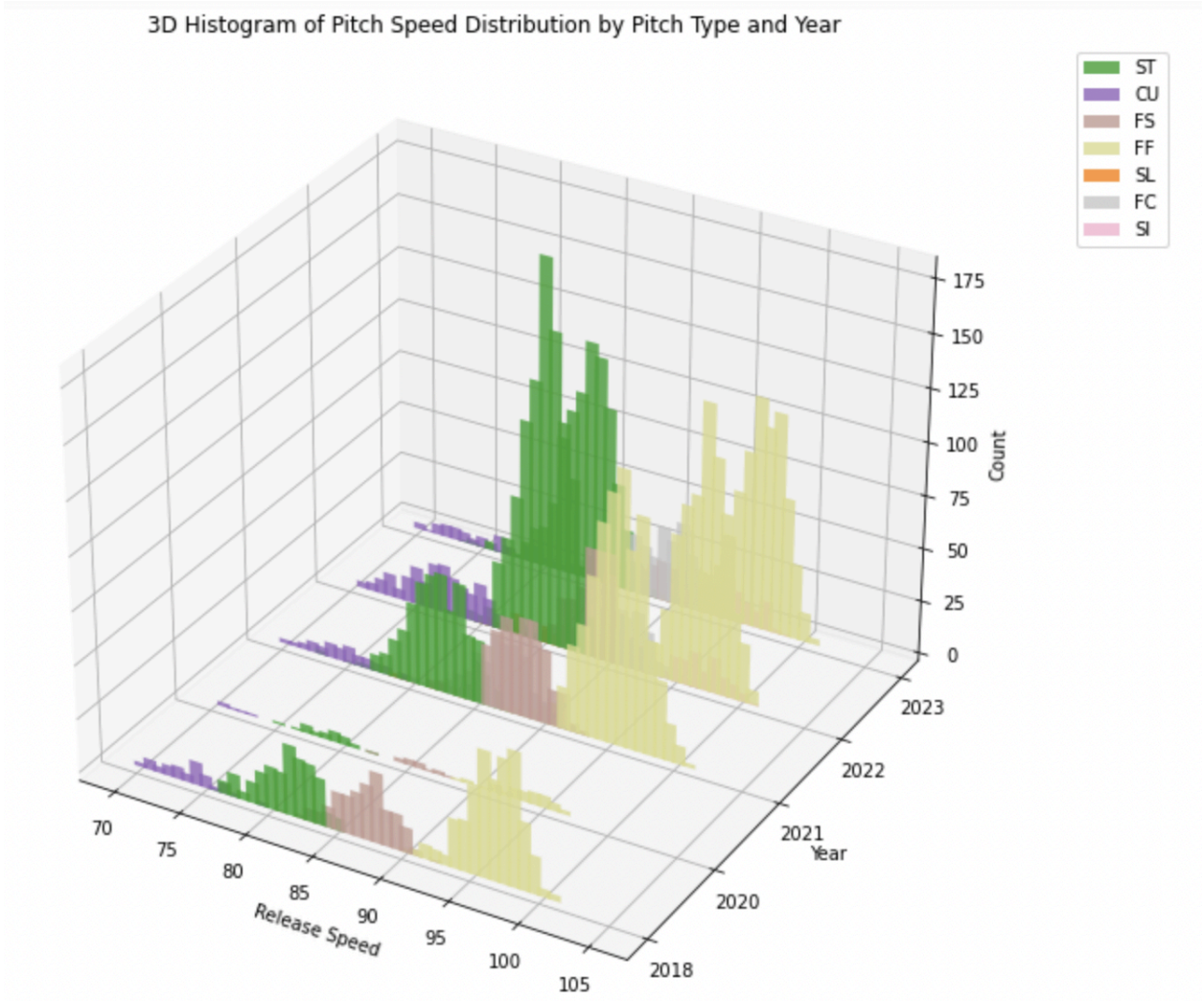

Figure 2

Scatterplot:

**** Please run the code for the 3D graph. There are two interactive dropdown selections. There is also the selection of the year on the right side.****

Now, we want to delve deeper into Ohtani's pitches by examining the 3D scatter plot of his pitch release positions over the years. We aim to determine if changes in these positions have contributed to his increasing strike percentage. In the 3D plot, the X-axis represents the vertical release position, the Y-axis represents speed, and the Z-axis represents the horizontal release position. Additionally, the plot can be filtered to display either strike or non-strike pitches but we will only be looking at strike.

When beginning our analysis of Ohtani's pitch release, we were fascinated by the potential changes in his performance from 2018, the year before his injury, to 2021, when he had fully recovered. The scatter plot analysis excludes 2020 game data due to Ohtani's injury recovery and lack of recorded data and we focused specifically on his fastball and sweeper, his main pitches. In 2021, we observed a slight decrease in fastball speed compared to 2018, as previously shown in the histogram. Both the horizontal and vertical release positions for the fastball increased, shifted left (see Figure 3). This indicates a significant change in his release post-injury, although his speed has not returned to its previous levels. Similarly, Ohtani's sweeper also showed increased horizontal and vertical release positions, shifting left between 2018 and 2021. From this analysis, we can conclude that Ohtani's pitch release underwent significant changes during his recovery period. These changes were positive, as data from the contingency table reveals that Ohtani's fastball strike rate increased from 2018 to 2021.

Interestingly, we observed that in 2022, Ohtani's fastball pitch release position decreased, shifting right, instead of remaining in the same position as in 2021. It returned to similar horizontal and vertical release positions as in 2018. At the same time, we can see his speed steadily increases and the contingency table reveals a continued improvement in his strike record, suggesting a possible correlation between his increased speed and enhanced strike performance. In 2023, Ohtani's fastball position decreased again, shifting back to left. Similarly, Ohtani's sweeper pitch release position also displayed similar changes, in 2022, there was a decrease in both vertical and horizontal positions, shifting the red cluster right. Then, in 2023, it shifted back to the left, as illustrated in Figure 5 in the purple cluster.

These changes in Ohtani's pitches indicate continuous improvement, particularly in his speed, while he continues to refine his release position to find the perfect release. Nevertheless, it

appears he is nearing the optimal release position, as evidenced by the consistent increase in his pitch strike record.

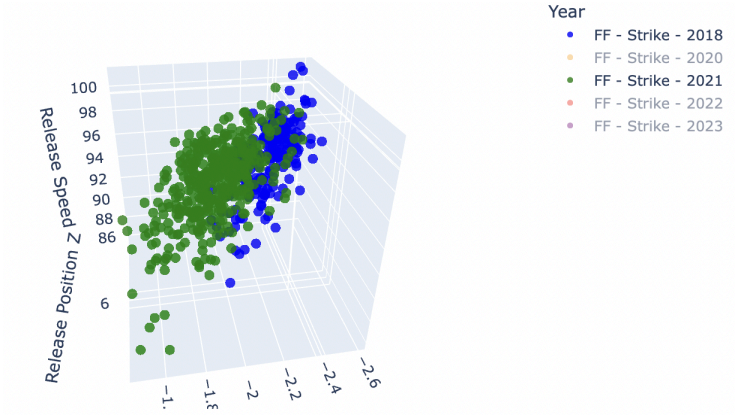

Figure 3

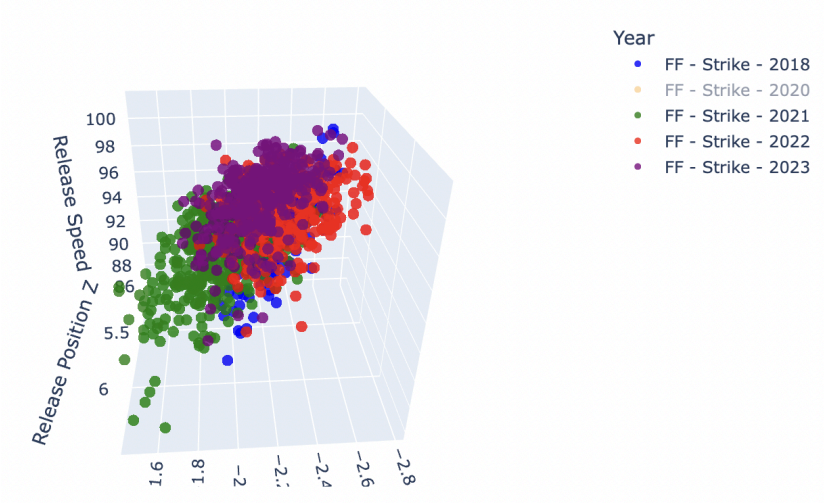

Figure 4

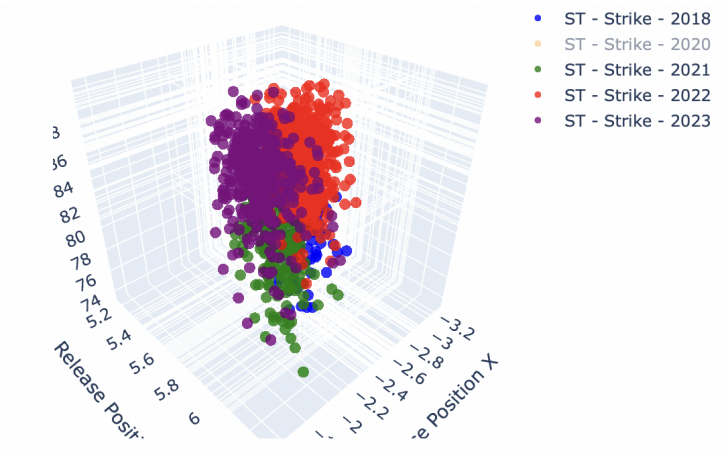

Figure 5

Heatmap:

There have been several changes in Ohtani's pitching style over the years from 2018 to 2023 (except 2019). The heatmaps provide insights into the horizontal and vertical release positions of his pitches. You can select the type of pitch and the batter's handedness (left or right).

Over the years, Ohtani has shown increasing consistency in his release points, especially against right-handed batters. In 2018, Ohtani's fastballs showed consistent release points for both left and right-handed batters. Then, some changes started to become noticeable over the years. In 2020, the release positions for left-handed batters showed a broader spread horizontally. This might indicate possible adjustments in his pitching strategy. The release positions for right-handed batters remained more concentrated. However, this was clearly not Ohtani's best pitch, likely because he only played 80 games that season due to his injury. Later on, he began to show better progress with his pitches. As we look closely at 2021 and 2022, the heatmaps for Ohtani's fastballs show tighter clustering of release points for both left and right-handed batters, indicating that he improved in maintaining his pitching consistency. In 2023, Ohtani showed significant improvement in his release points. Both his release points for left and right-handed batters are consistent and concentrated, showing a more defined clustering. In conclusion, Ohtani has demonstrated an ability to adapt and refine his techniques, leading to better consistency in his pitching.

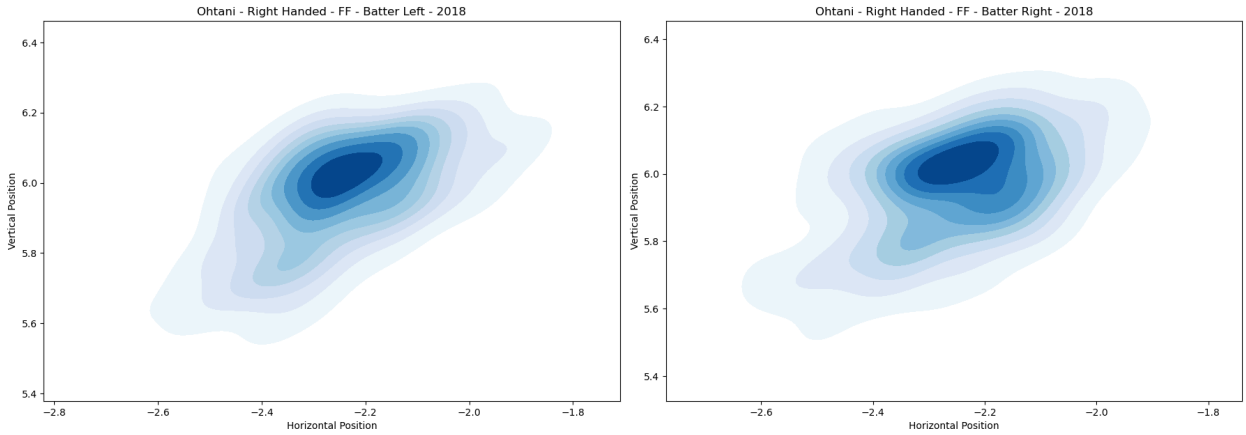

Figure 6

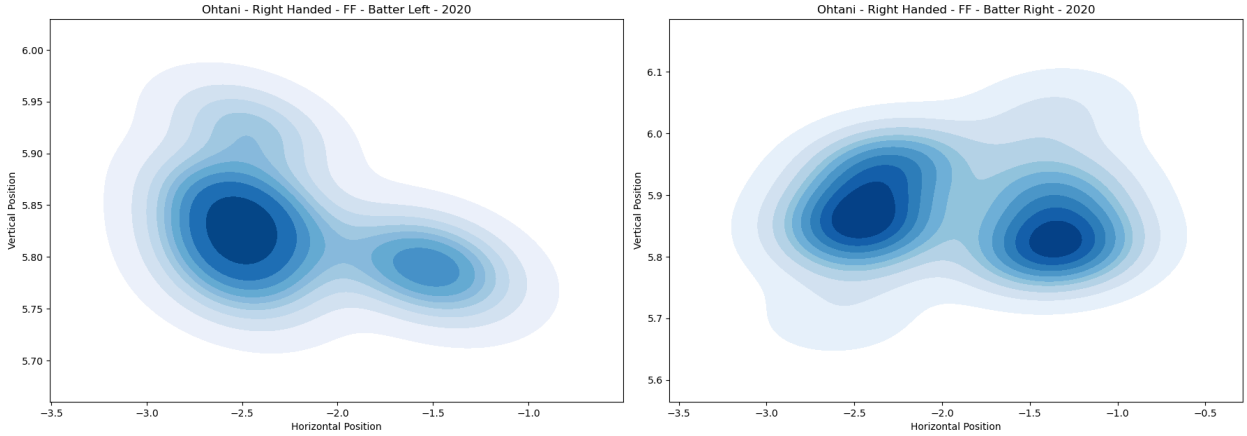

Figure 7

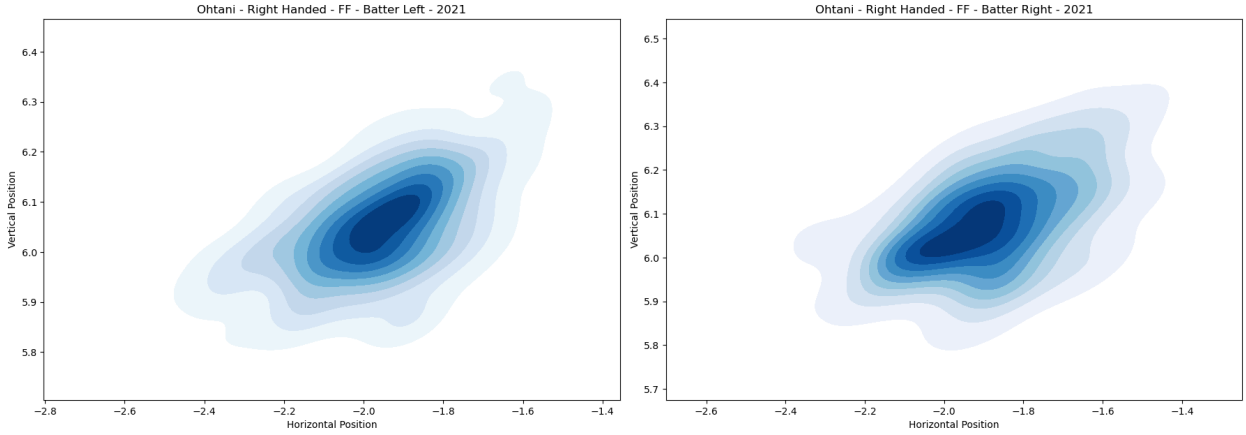

Figure 8

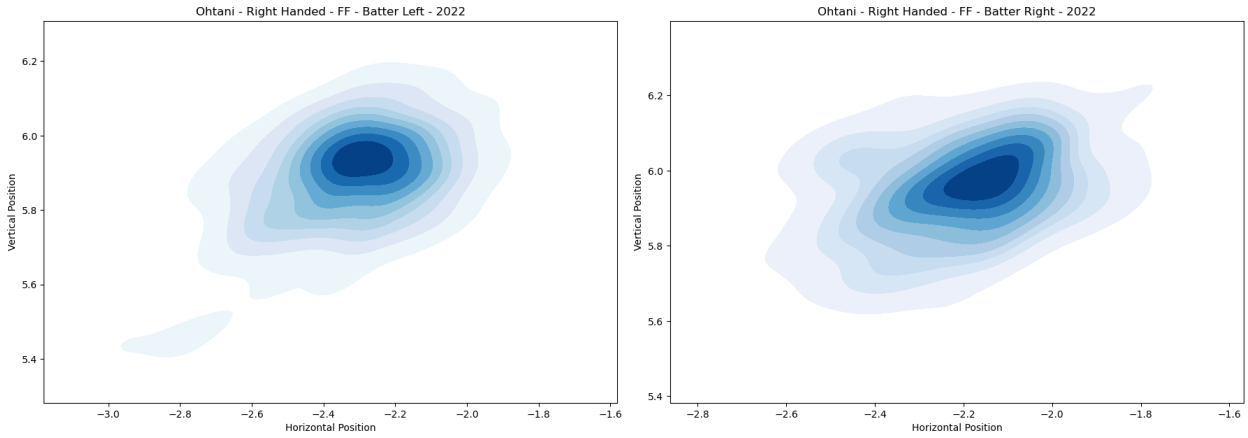

Figure 9

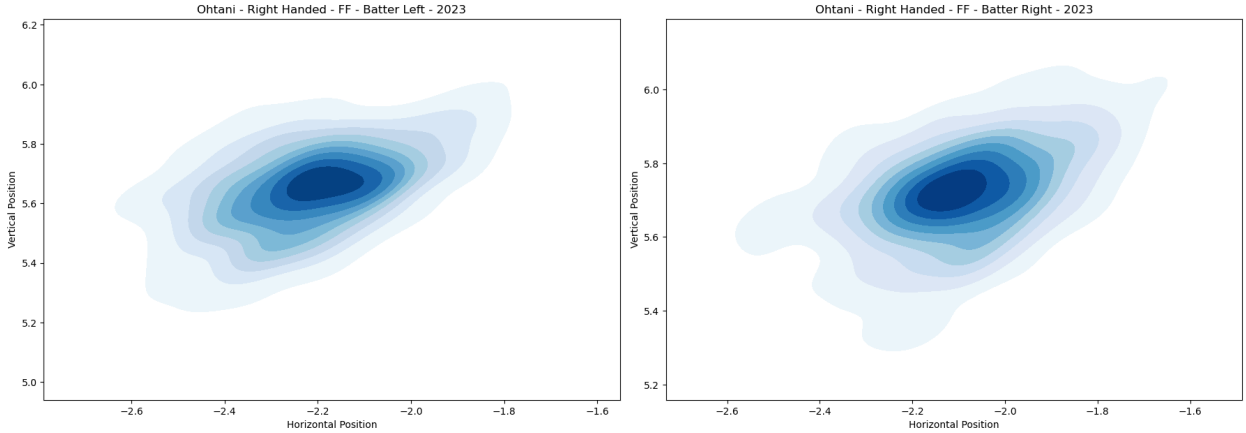

Figure 10

Cole VS Ohtani and Contingency Table:

While it is very interesting to look at Ohtani's growth over the years, we also want to compare his performance to that of the top MLB player, Cole Gerrit. We decided to do a simple comparison using contingency tables and 3D scatter plots. To be fair, we only compare using Ohtani's and Cole's 2023 data.

From the contingency table, we can see that their reference in pitch type is very different. While Ohtani has transitioned to the sweeper as his primary pitch over the years, he still enjoys mixing in various other pitch types. Two possible conclusions can be drawn from Ohtani's large pitch variation: either his gameplay involves mixing up his pitches to confuse the batter, or he may still be in the process of finding his specialized pitch. In contrast, Cole specializes in fastballs, utilizing them in 53% of his pitches. Regardless, comparing all Cole and Ohtani's pitches directly

is challenging. Therefore, we've chosen to focus solely on their fastball data because it is Cole's specialization, and Ohtani has also specialized in fastballs in previous years.

We can see directly from the contingency table that Cole's Fastball pitch strike percentage is greater than Ohtani's by 0.87% (see Figure 11 for Cole's contingency table and Figure 12 for Ohtani's). Not significantly different, however, it shows that Ohtani still has room to grow in his fastball accuracy to reach the top player.

	Pitch Type	Count	Percentage	Strike Percentage	Non-Strike Percentage
0	ST	737	35.2%	53.6%	46.4%
1	FF	690	32.95%	52.9%	47.25%
2	FC	329	15.71%	48.33%	52.28%
3	FS	131	6.26%	29.77%	70.23%
4	SI	126	6.02%	42.06%	57.94%
5	CU	76	3.63%	34.21%	65.79%
6	SL	5	0.24%	60.0%	40.0%

Figure 11

	Pitch Type	Count	Percentage	Strike Percentage	Non-Strike Percentage
0	FF	1739	53.0%	53.77%	46.18%
1	SL	683	20.82%	44.07%	55.93%
2	KC	396	12.07%	46.72%	53.28%
3	CH	233	7.1%	36.48%	63.52%
4	FC	230	7.01%	50.87%	49.13%

Figure 12

Scatterplot Analysis: Ohtani (2023) vs Cole (2023)

The 3-D scatterplot comparing Shohei Ohtani's and Gerrit Cole's four-seam fastball (FF) pitches in 2023 highlights some key differences and similarities in their pitching styles. Ohtani's pitches show a consistent horizontal release point between -2 and -1.5 feet and a vertical release point around 5.6 to 6 feet. His fastball speeds range from 88 to just over 100 mph, with clustering between 93 and 97 mph, indicating his precision and consistency in delivering high-velocity pitches. For example, Ohtani's ability to maintain a tight release point range allows him to better

control the trajectory and movement of his fastballs, making it harder for batters to predict the pitch location.

On the other hand, Cole's pitches have a broader horizontal release range from -2.4 to -1.6 feet and a vertical release point around 5.8 to 6.2 feet. His fastballs also reach speeds up to 100 mph, with clustering between 94 and 98 mph. Cole's broader release range suggests a more adaptable pitching style, allowing for slight variations in his delivery. Thus, these variations can make it more difficult for batters to predict the exact location and movement of his pitches.

Overall, while both pitchers achieve impressive fastball speeds, Ohtani's pitching is marked by precise and consistent release points, enhancing his control and accuracy. Cole's slightly higher average speed and broader release range highlight his ability to adapt and vary his pitches. Therefore, Ohtani can improve his pitching by incorporating some of Cole's adaptability into his own style. He could experiment with increasing his average pitch speed, taking advantage of the benefits that come with higher velocity while maintaining his precision and control. By doing this, Ohtani could further enhance his pitching skills and remain one of the top pitchers in the MLB.

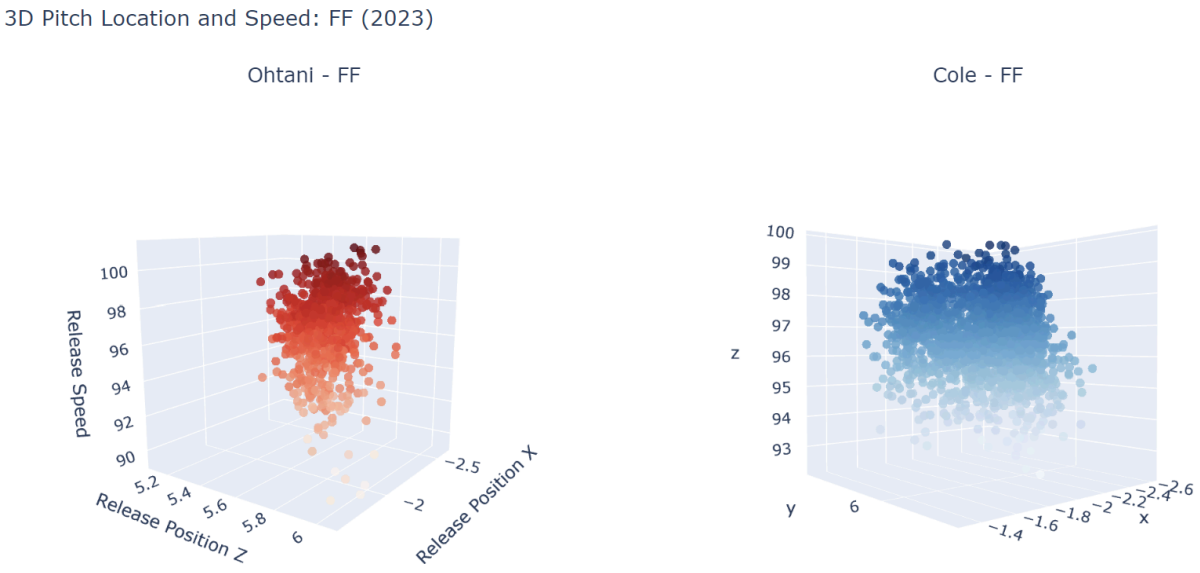

Figure 13

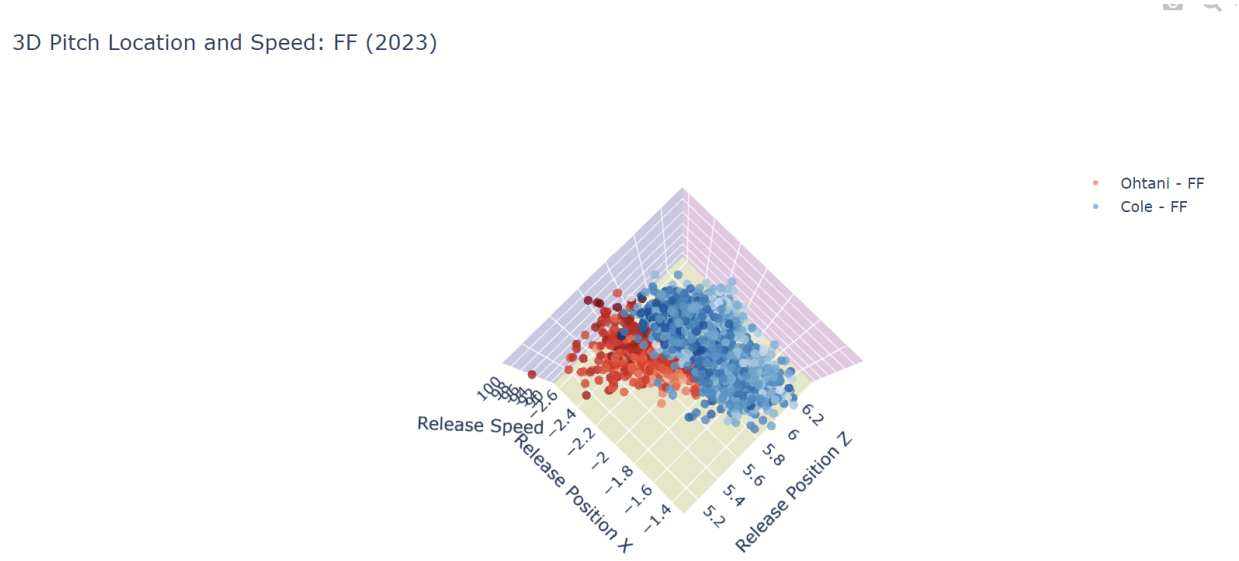

Figure 14

Conclusion:

Through our analysis, we observed Shohei Ohtani's growth over the 5 years period in his MLB career. In 2018, Ohtani mainly relied on his fastball, showing his raw talent and speed. In 2019, his performance was interrupted by a significant injury. In 2020, still recovering from injury, he was limited to only playing 80 games. Despite this setback, Ohtani showed remarkable resilience and adaptability. Also in 2020, he began to use more varied pitches along with his fastball. This trend continued in 2021 and 2022, with improvements in his strike rates and pitch consistency. In 2023, Ohtani's consistent release position and stable pitch speeds showed his dedication to refining his technique and improving his control and accuracy. Additionally, we identified key aspects of Ohtani and Gerrit Cole's pitching styles. Ohtani's consistency and precision are evident in his tight release points and stable pitch speeds, enhancing his control and accuracy. Cole's adaptability and higher average velocity make his pitching style unpredictable and effective, allowing him to adjust his mechanics for different batters.

By incorporating elements of Cole's adaptable approach, such as varying his release points and experimenting with higher pitch speeds, Ohtani can further enhance his pitching skills. This could help him rise to be one of the top pitchers in the MLB. Our analysis highlights the strengths of these two elite pitchers and suggests areas for improvement. Understanding their differences and similarities highlights their unique styles and shows how Ohtani can further develop as a top pitcher in the league.

Future work:

For future analysis, we could look at how Ohtani performs against different types of batters to see how he adjusts his pitching strategy. We could also examine his off-season training and its impact on his in-season performance. Additionally, comparing Ohtani's pitching mechanics with other top pitchers could reveal best practices and new techniques he could use to improve. This approach will give us a better understanding of Ohtani's growth as a pitcher and areas where he can further improve.