Introduction

Baseball has become a sport all about velocity. The obsession with how hard a pitcher can throw is hardly new; I love hearing all the wacky stories about the strange ways scientists would try measuring a pitcher's speed in the days before radar guns and the tall tales about how hard certain pitchers may have thrown, such as the absurdity that Bob Feller threw 107.9 mph. However, the obsession has only increased in recent years. The radar gun's velocity reading is displayed on a baseball broadcast after every pitch, and it is something every fan pays attention to. Over the past decade, the sport of baseball has noticeably changed. Hits have gone down, strikeouts and home runs have gone up, and every team seemingly has an endless number of pitchers that throw harder than the average man could imagine. This brief article does a great job of explaining just how hard it is to hit Major League pitching.

The mission of this project comes down to two research questions: How has fastball velocity changed over the years? How has the type of pitches that pitchers throw changed over the years? To address these questions, I examine how fastball velocity has changed on a year-by-year basis as well as the motivation to throw harder. I also look at how the style of pitching has changed over the years and compare it to the rest of the league, and once again view the motivation to do so.

As a huge baseball fan, these questions are topics that are constantly molding the talent and strategy of the game. Every fan knows that pitchers are throwing harder, that the number of pitchers who can throw hard is extremely higher than before, but to what extent? What do these changes look like? This project is going to focus on the New York Yankees, the team that I am a fan of. While these changes are present throughout the baseball world, the Yankees are an especially interesting team to further examine. Starting in the mid-2010's, the Yankees were known as a team that employed a lot of high-velocity pitchers, and despite that, a team that was throwing fastballs less often. This direct strategy deployed by my favorite team has had this topic on my mind for many years now.

Methodology

All of my data sources are from Baseball Savant, the official website of Major League Baseball's StatCast technology. The website is a treasure of baseball related data. When it comes to pitch tracking, there are two technologies that have shaped the world of baseball data. PITCHf/x was the first pitch tracking technology, introduced in 2006, that measured the velocity, type, and location of every pitch thrown in Major League Baseball. PITCHf/x was ultimately replaced by StatCast, a doppler radar technology that is even more detailed than PITCHf/x, in 2015. The Baseball Savant website contains every single pitch thrown in MLB since 2008, using PITCHf/x for 2008-2014 and StatCast for 2015 to present day. Querying and exporting data from their website is easily done through their search function.

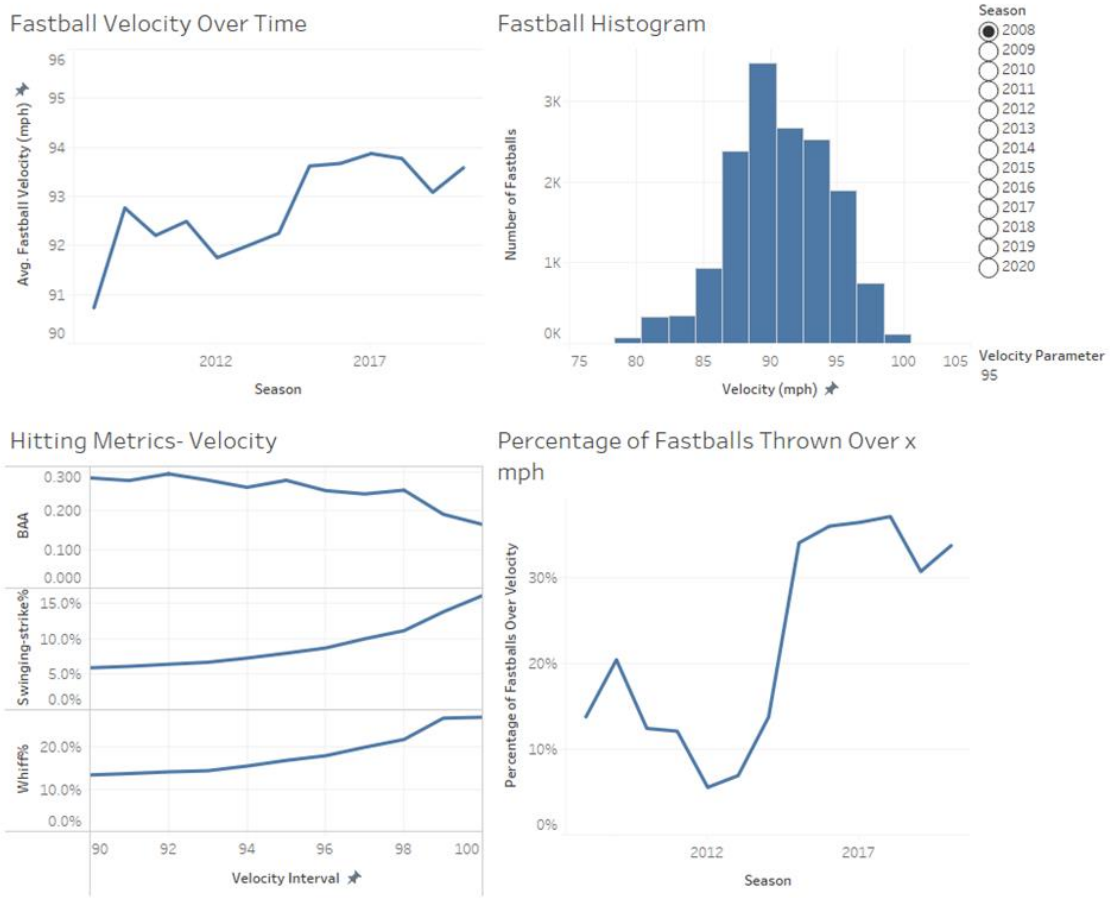
My main dataset, simply put, is every single pitch thrown by the New York Yankees from 2008 to 2020. After taking out null values and pitches that should not be included in analysis (pitch-outs and intentional walks, which I would describe as non-competitive pitches), that adds up to 291,122 pitches over 13 seasons. The crucial columns are the pitcher's name, the

date of that pitch, the pitch's velocity, the type of pitch, and the result of that pitch (e.g., called strike, hit into play). I created this dataset by searching for every pitch thrown by the Yankees in an individual season, repeating this 13 times, and then merging them together.

My secondary datasets are much smaller. One looks at the percentage of pitches thrown of three pitch categories (fastballs, breaking balls, and offspeed) across all of MLB for each season from 2008 to 2020. This will be used to compare the Yankees to the league average in one of my visualizations. The other dataset contains stats on every season thrown by a Yankee pitcher in this time period, which 365 seasons for 187 different pitchers. This gets used to look at the pitching styles of the best Yankees pitchers in one of my visualizations.

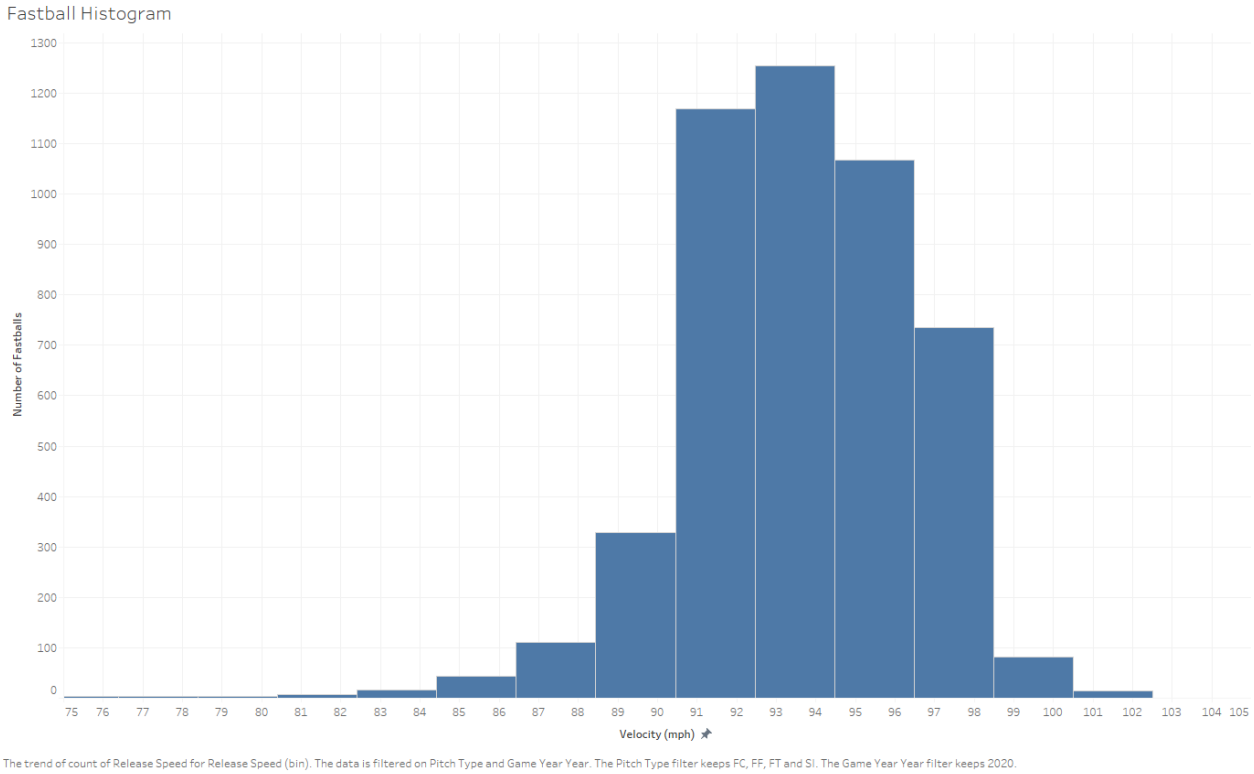
Analysis

For my visualizations on fastball velocities, I had to filter each worksheet by the pitch type in order to look at fastballs only. The four following pitches are types of fastballs and are included: four-seam fastball, two-seam fastball, cut fastball (cutter), and sinker. I will explain more later about what makes these four pitches unique, but for now I should clarify what a fastball is for any non-baseball fans. If you imagine just picking up a ball and throwing it as hard as you can, that would be a fastball in baseball terms. When examining changes in velocity in baseball, we are really only focused on fastball velocity.



To visualize the change in velocity, I first focused on three graphs that have the baseball season on the x-axis to show change over time. The goal of the first chart is very simple: how has the average fastball velocity changed over time for the Yankees? We see a general trend upwards, with some dips and dives, to let us know that the average really has increased. However, this probably is not the most descriptive way to view this change.

Next, I looked at a histogram that shows the number of fastballs thrown by different intervals of velocity (88-90 mph, 90-92 mph, etc.), where we can select which season we want to view. In the dashboard above, we see the year 2008 selected. A large portion of fastballs thrown by the Yankees in 2008 are between 88 mph and 92 mph, while a very small portion is thrown above 96 mph. We can contrast this with the most recent season, 2020:

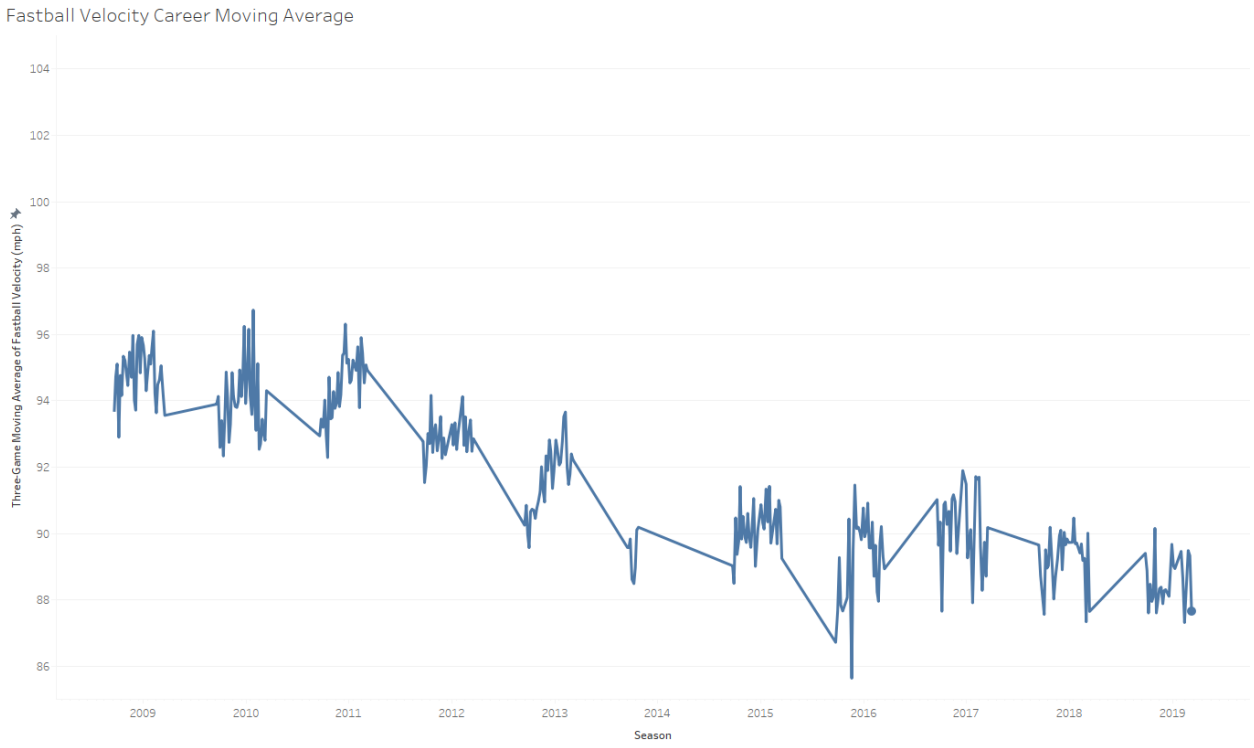


The trend of count of Release Speed for Release Speed (bin). The data is filtered on Pitch Type and Game Year Year. The Pitch Type filter keeps FC, FF, FT and SI. The Game Year Year filter keeps 2020.

The third velocity-over-time chart looks at the percentages of all fastballs thrown in a season that were over a certain velocity. For example, in the pasted dashboard output, the percentage of fastballs thrown over 95 mph is displayed. Along with the histogram, I think this is the best way to truly grasp how velocity has changed. This graph was created using a parameter, where the viewer can select which velocity threshold they want to view. Next, a calculated field looked at every pitch and assigned it a value of 1 if it was greater than the velocity parameter, and a 0 otherwise.

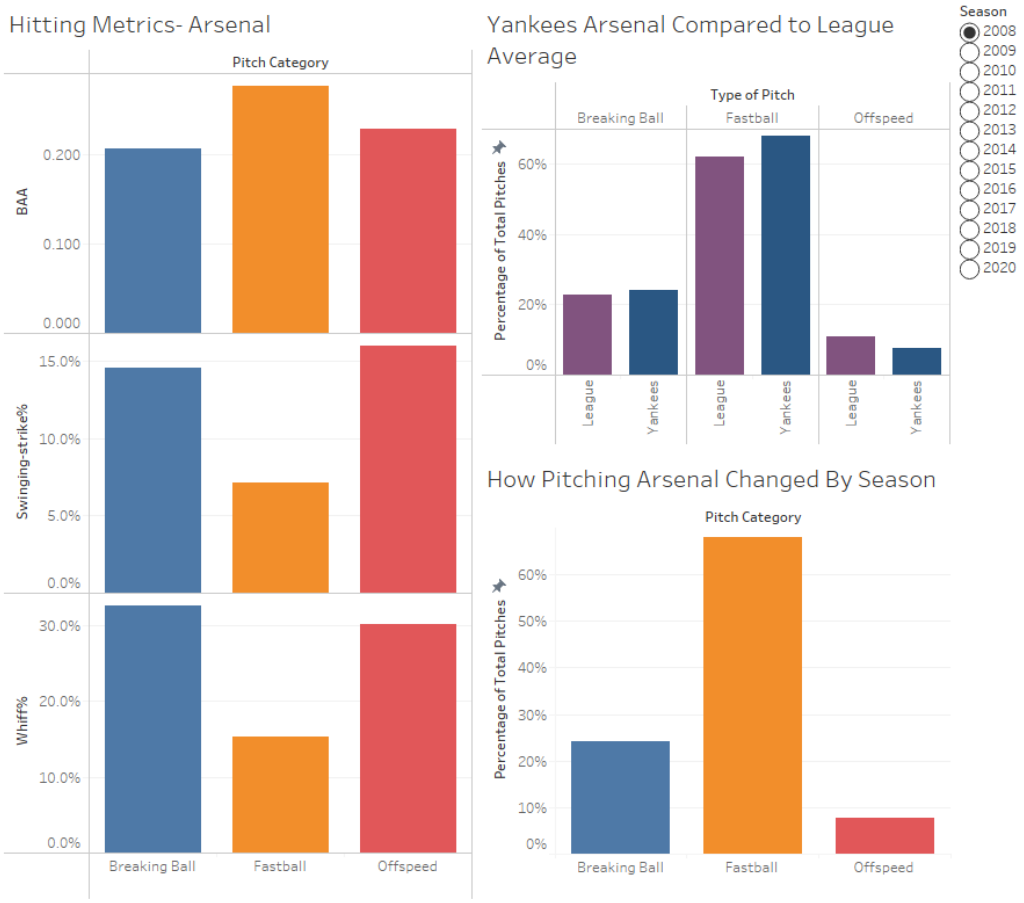
Also seen in the dashboard is a visualization titled “Hitting Metrics- Velocity.” The purpose of this is to convey the motivation behind increased velocity in the game. The idea is to see how hitters perform at different velocities to see if that encourages higher velocity in the game. The three metrics I chose are batting average against (BAA), swinging-strike percentage,

and whiff percentage. BAA is not my favorite baseball stat because I find it a bit misleading, but because of the simplicity and prevalence of batting average (hits / at-bats) in the sport, it is popular among a lot of fans. BAA is calculated by dividing total hits by total pitches that end the at-bat. Since it is only applicable to at-bat-ending pitches and not all pitches, that is why I find it misleading. Swinging-strike percentage and whiff percentage are preferred by me in the context of evaluating how effective a pitch is. They are very similar formulas: swinging-strike percentage looks at swings-and-misses divided by all pitches, while whiff percentage looks at swings-and-misses divided by pitches that resulted in a swing. Some people prefer one over the other, so I included both. To calculate these metrics in Tableau, I had to create more calculated fields for binary variables to denote whether a pitch was an at-bat-ending pitch, a hit, a swing, and a swing-and-miss.



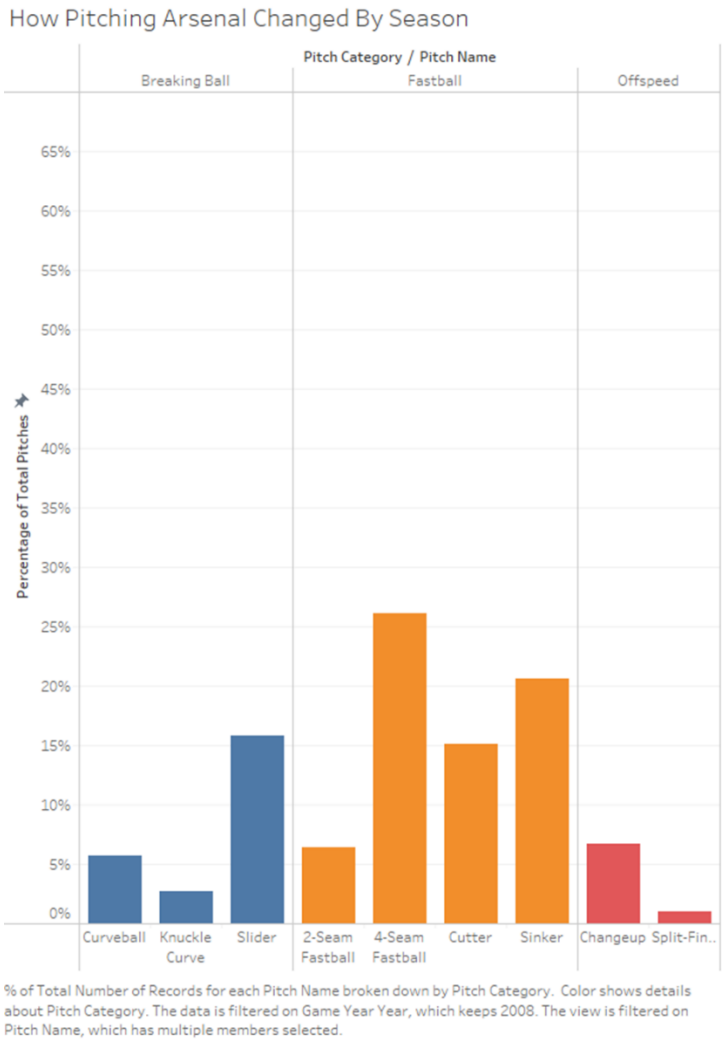
The plot of Moving Average of Avg. Release Speed for Game Date Day. The data is filtered on Pitch Type and Player Name. The Pitch Type filter keeps FC, FF, FT and SI. The Player Name filter keeps Sabathia, CC.

Finally, I included an animated moving average visualization that shows a pitcher's fastball velocity throughout their career. This doesn't quite answer the research questions about velocity on a grand-scheme level, but I think it is really interesting to see how velocity changes for a pitcher throughout a season and their career. I created the visualization with two specific pitchers in mind: CC Sabathia and Aroldis Chapman. Sabathia is a pitcher who threw very hard at the beginning of his career but experienced lots of problems with diminished velocity as he aged. Chapman is one of the hardest throwing pitchers ever, and he has a reputation of building velocity throughout the season, as well losing some mph on his fastball in recent years. The visualization above is for CC Sabathia.



The next series of visualizations relates to the second research question: how has the style of pitching changed? In baseball, they often call this a pitcher's arsenal; what pitches does he throw and how often does he throw it? There are three main categories of pitches: fastballs, breaking balls, and offspeed. Fastballs include four types. A four-seam fastball is thrown more for velocity, while a two-seam and sinker are thrown more for downward movement and to pitch to contact. A cutter is for horizontal movement in a fastball. For breaking balls, there are curveballs, a type of curveball called a knuckle cure, and sliders. Breaking balls are slower with a focus on movement. Curveballs have more vertical drop, while a slider has horizontal break. Offspeed pitches are meant to be slower pitches to throw off the hitter's timing. There are two types of offspeed: a changeup and a split-finger (splitter).

I have a drilldown visualization that shows the percentage of total pitches thrown by the Yankees that were classified as the three pitch categories. This can be drilled down to the specific pitch types within those categories. With this chart, we can see year-to-year how the style of the Yankees pitching changed and if they were throwing more or less of certain pitch categories and types. I also included an animated version of this chart to quickly see year-to-year changes in percentages.



A second chart compares the Yankees to Major League Baseball as a whole. Now we can see if the style of Yankee pitching changed in a similar trend to the rest of baseball, and if the Yankees particularly had more drastic or unique changes. This chart uses side-by-side bars to directly make a comparison of the Yankees' percentage compares to the league average.

I included another hitting metrics chart, very similar to before, but now for pitch categories and types. The chart uses the same calculations for batting average against, swinging-strike percentage, and whiff percentage. Now we can see how effective hitters are against the categories and types of pitches, which might explain some motivation for changes in pitching style. Since batting average is represented in baseball as a decimal to the thousandths place, and not a percentage, I have the three metrics on separate axes.

Top Ten Seasons by wOBA

Player Season	Pitch Category ..	wOBA	BAA	Swingin..	Whiff%
Rivera, Mariano 2008	Fastball	0.189	0.167	11.8%	23.3%
Green, Chad 2017	Breaking Ball	0.203	0.180	11.5%	27.0%
	Fastball	0.203	0.129	15.7%	33.8%
	Offspeed	0.203		0.0%	0.0%
Betances, Dellin 2014	Breaking Ball	0.204	0.075	16.9%	48.2%
	Fastball	0.204	0.231	9.7%	18.1%
Warren, Adam 2017	Breaking Ball	0.212	0.170	10.1%	22.6%
	Fastball	0.212	0.157	7.6%	16.7%
	Offspeed	0.212	0.206	14.9%	29.2%
Miller, Andrew 2015	Breaking Ball	0.218	0.098	25.4%	51.6%
	Fastball	0.218	0.241	9.0%	23.1%
Rivera, Mariano 2010	Fastball	0.220	0.186	8.7%	17.0%
Robertson, David 2011	Breaking Ball	0.232	0.156	16.6%	33.9%
	Fastball	0.232	0.183	9.0%	22.4%
	Offspeed	0.232	0.000	25.9%	58.3%
Pineda, Michael 2014	Breaking Ball	0.234	0.188	17.4%	29.3%
	Fastball	0.234	0.225	7.1%	15.1%
	Offspeed	0.234	0.133	14.0%	31.7%
Rivera, Mariano 2011	Fastball	0.235	0.215	8.5%	16.6%
Betances, Dellin 2015	Breaking Ball	0.240	0.098	17.4%	48.2%
	Fastball	0.240	0.223	13.1%	27.7%

WOBA, BAA, Swinging-strike% and Whiff% broken down by Player Season and Pitch Category (copy). The view is filtered on Player Season, which has multiple members selected.

The final chart lists the top ten seasons by a Yankees pitcher during this time frame, ranked by their wOBA for the season. Without getting into too much detail, wOBA is a metric that encompasses offense into one statistic. I wanted to view the best seasons by Yankees pitching and see what their style of pitching was. What pitches did they throw? How effective were those pitches? This question doesn't directly relate to a research question but it informative on how the best pitchers pitched.

Conclusion

Velocity has certainly increased since 2008. My favorite of my charts to emphasize this is the "Percentage of Fastballs Thrown Over x mph" chart. With the parameter set to 95 mph, we see that only 6% of fastballs in 2012 were thrown over 95 mph. Since 2015, every year has been over 30%. If we increase the parameter to a more extreme velocity, like 98 mph, it is even more drastic. This was below 1% for every season between 2009 and 2013 and peaked at 14% in 2017. This makes it clear that the number of pitchers that throw hard is not only increasing, but the number of pitches that throw at extreme velocities is also increasing. Teams are certainly

targeting pitchers who throw at these higher velocities, and younger pitchers are training themselves to throw harder now that they know how it is valued. This change was noticeable for Yankees fans as the team clearly was targeting high-velocity pitchers starting in 2015. That year, they added Andrew Miller to a team already with hard-throwing Dellin Betances. Then in 2016, they added Aroldis Chapman, the fastest pitcher in the game. Not only were the relief pitchers throwing hard, but the starting pitchers, who typically throw slower, were among the best in the game. These pitchers include Luis Severino, Nathan Eovaldi, James Paxton, and Gerrit Cole.

What is the motivation for throwing so hard? In the "Hitting Metrics- Velocity" chart, we see that it is true that hitting is much more difficult against higher velocity fastballs. Pitches that were between 90 and 91 mph had a batting average of 0.285 and a whiff rate of 13.4%. At 100 to 101 mph, BAA decrease all the way to 0.165 and whiff rate goes up to 27.1% It is just a matter of having a better understanding what makes a pitcher better; velocity is a huge component of this.

Starting in 2017, I noticed that despite the Yankees employing so many high-velocity pitchers, they were throwing fastballs less often. My curiosity was really struck after reading an article that year about the Yankees being the leaders in a new pitching philosophy (unfortunately I can't find the article). It is true. In 2008, the Yankees threw fastballs 68.8% of the time, and that went all the way down to 50.65% in 2017. What is the motivation behind this? On "Hitting Metrics- Arsenal," we see that hitters are much better against fastballs than any other pitch. Traditionally in baseball, your fastball is your main pitch and everything else would get called "secondary" pitches.

With the analytical mindset now so present in baseball, the Yankees saw that breaking balls and offspeed pitches were more effective, so they decided to throw those pitches more often. The whiff percentage against fastballs is 15.4%, but 32.6% and 30.2% against breaking balls and offspeed, respectively. This has become a trend across baseball, but were the Yankees leaders in the trend? In 2017, compared to the Yankees fastball percentage of 50.65%, the league average was much higher at 61%. The league didn't start throwing less than 60% of its pitches as fastballs until 2019, so I would argue the Yankees were part of the trendsetters in this style of pitching.

My additional research questions would look at other trends in baseball and relate them back to the trends discussed in this paper. The number of balls in play as gone down in baseball. Strikeouts have gone up. Hits have gone down. Home runs have gone up. I have noticed lots of fans point their finger at the mentality of hitters and strategy of baseball to explain these trends. These fans usually say that a hitter's all-or-nothing power approach is why or blame it on defensive shifting. Personally, I believe those factors are miniscule compared to the effects of velocity. Hits are down and strikeouts are up, in my opinion, because is so much better now than it used to be. Can we attribute how much of these changes are explained by velocity? Can we visualize the trends in home runs, batting average, strikeouts, and defensive shifts? That would be the next step in research.