Why the 2020 Dodgers Are the Greatest Team of All Time, at least statistically

Sean Floersch

Faculty Mentor: Chad Vidden, Mathematics and Statistics

ABSTRACT

This paper explores the use of sports analytics in an attempt to quantify the strength of Major League Baseball teams from 1920-2020 and then find which team was the greatest team of all time. Through the use of basic baseball statistics, ratios were personally created that demonstrate the strength of a baseball team on offense and defense. To account for slight year to year differences in the ratios, standard deviations to the yearly mean are used and the personally created Flo Strength metric is used with the standard deviations. This allows the ratios to be compared across all seasons, with varying rules, number of teams, and number of games. These ratios and Flo Strength metric go through statistical testing to confirm that the ratios are indeed standardized across seasons. This is especially important when comparing the strange Covid-19 2020 baseball season. After quantifying all the team strengths, it is argued that the 2020 Los Angeles Dodgers are the greatest team of all time.

INTRODUCTION

The 2020 Covid MLB Season

In a year like no other, with a pandemic locking down society, racial inequalities becoming the forefront of the news, and division across the country mounting, sports were turned to for a sense of normalcy. Slowly, sports returned, but in a way never seen before. The NBA returned, isolated in a bubble. Strict protocol was put in place in the NHL and MLS in order to ensure player and personnel safety. And finally, on July 30, Major League Baseball returned, ready to go with a 60-game season, and season that would see many changes. In order to restrict travel and interactions, the 60-game season schedule was set up so that every team would only play games against teams in their own region. Foreshadowing playoff matchups would become much more difficult without head-to-head play being included. Playoffs would also look different, as MLB expanded to typical 10-team playoff to 16-teams in order to account for teams that started off slow, giving these teams the opportunity to prove themselves in the later half of the season. All this set up a 2020 season that would mimic the year itself: extraordinary.

The Greatest of All Time

The greatest of all time, or GOAT, conversation is a popular debate throughout the sporting world. Fans and analysts are constantly fighting about Lebron James vs. Michael Jordan, Tom Brady vs. Peyton Manning, or the 80s Lakers vs the 90s Bulls. The unfortunate aspect for many of these debates is the subjectivity of them. People put thoughts and feelings into their arguments, making it more of a story about why they like one certain player or team more than the other. In order to provide a crisp, objective argument, statistics are required. This paper will explore the statistical side of the greatest MLB team of all time, an argument that will end with the 2020 Los Angeles Dodgers on top.

Baseball Analytics

Sports analytics has exploded in the 21st century. The NFL, NBA, and MLB have all adopted the use of statistical analysis into how to create the strongest team as well as at the play-to-play level. Although the basic statistics for these sports (Home runs, Touchdowns, points scored) have always been recorded, a deeper analysis and more advanced statistics have been added to these sports that allow statistic lovers everywhere to rejoice. These deeper analytics include pitch-by-pitch and play-by-play changes that the manager can follow to help statistically and probability-wise put the team in the best position to win. Many pro-sports leagues and teams have hired out statistical analysis to find out the best players to bring in, the best position matchups in game, and best strategy in crunch time. In baseball specifically, sabermetrics, advanced baseball statistics, and probability have grown exponentially in the recent years, changing the game forever. [3]

Paper Contribution

This paper will look at a comparison of basic stats across different eras in Major League Baseball. With basic stats, ratios and models will be calculated to show the strength of both the hitting and pitching side of the sport. These ratios are normalized across seasons by using standard deviations to the yearly mean. Correlations to winning percentage are used to weigh out how much these ratios play a part into a winning team. With the combination of ratios and correlations, a singular model called the "FloStrength" Model is calculated to demonstrate how strong a team is. This model is then applied to show that the 2020 Los Angeles Dodgers stack up on top. Then, all the statistics and models are also statistically proven through statistical tests to show that the 2020 Covid-shortened season is comparable to seasons of the past.

DATA

The Lahman data base was used for the years 1920-2019. [4] This data set is famous in the baseball world for the immense database that it holds. Only part of this huge database is used, with a sample of this raw data was in the form seen in table 1. The 2020 data was manually recorded from statistics from MLB.com. [7]

Table 1. Sample raw data from Lahman Database

teamID	yearID	G	W	R	RA	H	HA	HR	HRA	SO	SOA
SLN	1944	157	105	772	490	1507	1228	100	55	473	637

Throughout baseball history, rule changes and league changes have led to different eras to develop. Each era has its own defining characteristics and has also led to differences in statistics. In order to be able to compare two seasons against each other, it is important to understand the eras that the two seasons were a part of. The statistical differences as well as era breakdowns are seen in Table 2. Although these statistics do and will help tell a great story about team's seasons, it is important to note that there are many other aspects of baseball that are not included within this paper. However, these other aspects proved to not be able to tell nearly as much about a team's strength with respect to winning compared to the following statistics and therefore were not included in the methodology of this paper.

Table 2. Basic baseball statistics across MLB eras.

Data	Live Ball	Integration	Expansion	Free Agency	Steroid	Modern
Year	1920-1941	1942-1960	1961-1976	1977-1993	1994-2005	2006-2019
Runs per Game	4.845	4.351	4.041	4.207	4.842	4.459
On Base Percentage (OBP)	0.335	0.324	0.311	0.330	0.328	0.315
Walks + Hits per Inning Pitched (WHIP)	1.447	1.387	1.304	1.348	1.412	1.339
Homeruns (HR) per Game	0.484	0.703	0.771	0.802	1.076	1.036
Hits (H) per Game	9.728	8.829	8.488	8.831	9.163	8.797
Strikeouts (SO) per Game	3.115	4.077	5.518	5.354	6.443	7.399
Walks (BB) per Game	3.162	3.559	3.215	3.245	3.429	3.162

With obvious discrepancies between these statistics by era, it makes it difficult to compare teams from different eras. That is where ratios and normalizing come in.

METHODOLOGY: HOW TO QUANTIFY TEAM STRENGTH

Baseball is not just a game of hitting, nor is it just a game of pitching. In order to create a strong team, a balance between the two need to be met. To easily calculate this balance, ratios were used. By dividing the hitting statistics by pitching statistics for each individual team, such as runs scored divided by runs allowed, ratios can show

how strong a team was. Ratios also reduce differences between seasons. To account for differences in seasons and for the balance of offense and defense, four separate ratios were calculated, each with separate reasoning behind it.

Table 3. Ratios used in the paper.

Ratio	Formula
Run Ratio	$\frac{Runs\ Scored}{Runs\ Allowed}$
On Base Ratio	$\frac{OBP}{4.23\ *\ WHIP}$
Power Ratio	$\frac{4 * Homeruns + 1.5 * (Hits - Homeruns)}{4\ *\ Homeruns\ Allowed\ +\ 1.5\ *\ (Hits\ Allowed\ -\ Homeruns\ Allowed)}$
Strikeout Ratio	$\frac{Pitching\ Strikeouts\ +\ 3\ *\ Batting\ Walks}{Batting\ Strikeouts\ +\ 3\ *\ Pitching\ Walks}$

These four ratios, although they appear complicated, will be explained in depth in the following sections. The main idea behind them, however, is the same. Each one takes a look at major aspects of baseball: runs, getting on base, power hitting, and power pitching. The ideas behind run ratio come from "Analyzing Baseball Data with R" by Max Marchi and Jim Albert, while the other three ratios were a personal addition and creation. [2]

As is the case with all sports, not one single number or statistic is capable of showing how well a team has performed. Instead, a combination of statistics is required. Although win percentage is very good at demonstrating how well a team performs during the season, it is very similar to judging a book just by its cover. Yes, a cover can tell you quite a bit about the contents of the book, but it is impossible to truly understand the plot, theme, and characters of the book without fully reading it. This is the same case with sports, win percentage gives you a decent idea about how solid a team is, but in order to fully understand and know how well a team is built, a deeper statistical analysis is required.

The Basic Element. The easiest way to explain how a team wins is to look at how much the score compared to their opponents. Obviously, in order to win a game, a team has to outscore their opponents, and the more a team outscores their opponent the more they are going to win. To calculate this, the statistic Run Ratio ("RRatio") is created.

$$RRatio = \frac{\text{Runs Scored}}{\text{Runs Allowed}}$$

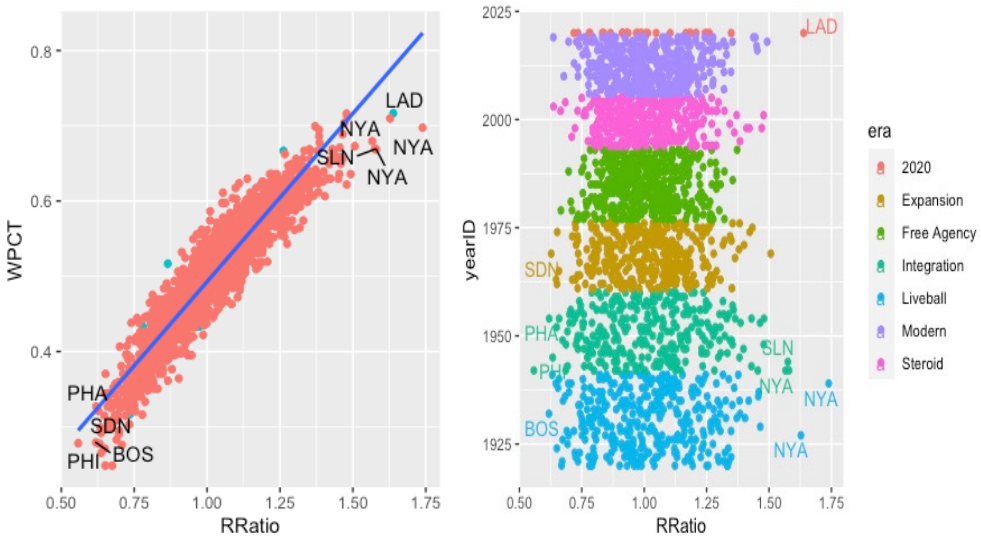

Figure 1 (Left). This graph plots out Run Ratio (RRatio) to win percentage (WPCT). Blue dots are 2020 teams and red dots are teams from 1920-2019.
Figure 2 (Right). The graph plots out Run Ratio with accordance to year. The 6 different baseball eras as well as the year 2020 are color coded with accordance to the key to the right of the graph.

In both graphs, the best and worst teams are shown by their abbreviation. The blue line running through the graph is the line of best fit, showing expected win percentage based on Run Ratio. The line runs through (1.00, 0.5), showing that an average team, with a winning percentage of 50%, is expected to have a run ratio of 1.00. A run ratio above 1 is an above average team, and below one is a below average team. Teams above the line of best fit underperformed according to Run Ratio, while teams below the line overperformed. It is seen through figure 1 that there is a strong linear relationship between run ratio and win percentage. Figure 2 shows that this ratio may favor earlier baseball teams, with the 2020 Los Angeles Dodgers being the only team in the top 5 since the 1944 St. Louis Cardinals. There has also been less spread in general.

The Moneyball Approach. "But does he get on base?" the famous line from Michael Lewis' book and movie "Moneyball" says. [5] This book looks at Oakland Athletics General Manager Billy Beane, who was at the forefront of the analytics movement that has been seen in baseball and the sporting world over the last few years. At the root of this movement is the idea that the more you get on base, the better of a chance to win. To test this, the statistic "On Base Ratio" was created, created by multiplying 4.23 by OBP, and then dividing this by WHIP. The number 4.23 comes from the average number of batters per inning, as this then makes WHIP and OBP on the same scale because both are now registering the number of batters on base per inning pitched. This then shows the ratio of offensive players on base to defensive players on base allowed.

$$OnBaseRatio = \frac{\text{OBP}}{4.23 * \text{WHIP}}$$

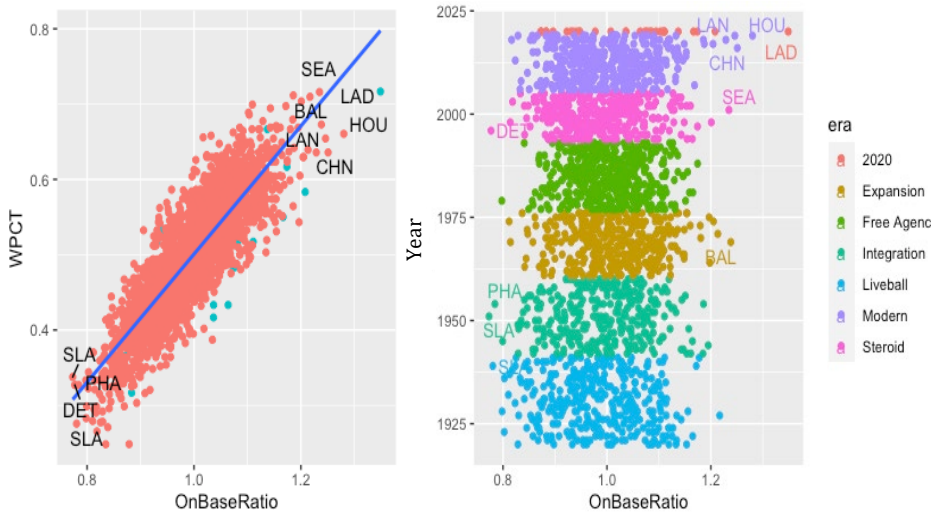

Figure 3 (Left). This graph plots out On Base Ratio to win percentage (WPCT). Blue dots are 2020 teams and red dots are teams from 1920-2019.
Figure 4 (Right). The graph plots out On Base Ratio with accordance to year. The 6 different baseball eras as well as the year 2020 are color coded with accordance to the key to the right of the graph.

Figure 3 shows that there is a strong, linear relationship between On Base Ratio and winning percentage, with the line of best fit running through (1.00, 0.5), showing that an average team, with a winning percentage of 50%, is expected to have an On Base Ratio of 1.00. Figure 4 shows that there is a slight favoring of more recent teams for this ratio. Three of the worst four teams occurred in the oldest two eras, while 6 of the top 7 teams occurred in the two most recent eras, including the 2001 Seattle Mariners, 2016 Chicago Cubs, and the 2020 Los

Angeles Dodgers. This trend towards higher On Base Ratios could be accounted for by rule changes, analytic-driven baseball, and/or expansion of teams.

The Long Ball. The best part of baseball, or worst depending on what side you are on, is the ability for the outlook of a game to change from one swing of the bat. To quantify this, I created the "Power Ratio", a ratio that accounts heavily for homeruns, with a factor of 4 to account for the 4 total bases that are reached with a homerun, while also accounting for double and triple by multiplying 1.5*hits that are not homeruns. 1.5 is calculated by finding the historical average number of bases gained per non-homerun hit. This was done by multiplying singles by 1, doubles by 2, and then triples by 3, and then dividing by the total number of hits that were not homeruns. From this, 1.5 shows that the average bases gained on a hit that is not a homerun. Utilizing ratios is especially important for this ratio, with over double the average homeruns in today's game compared to the Liveball era.

$$PowerRatio = \frac{4 * Homeruns + 1.5 * (Hits - Homeruns)}{4 * Homeruns\ Allowed + 1.5 * (Hits\ Allowed - Homeruns\ Allowed)}$$

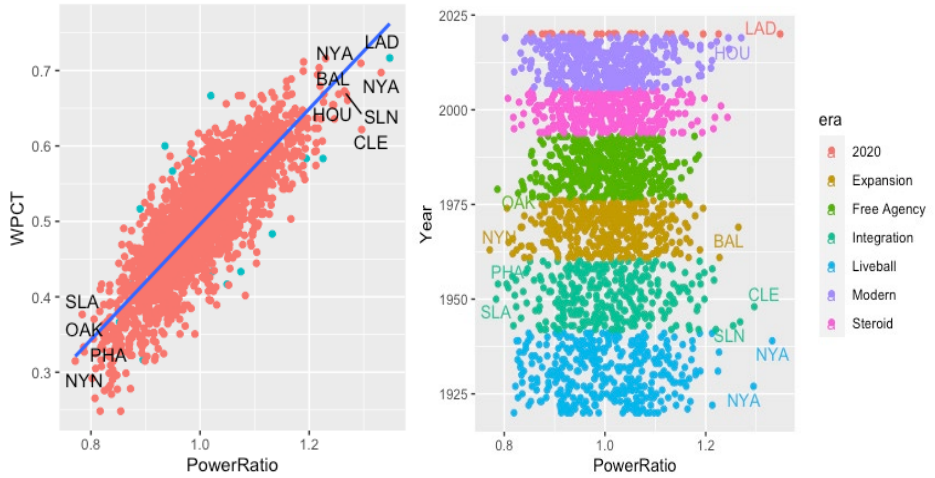

Figure 5 (Left). This graph plots out Power Ratio to win percentage (WPCT). Blue dots are 2020 teams and red dots are teams from 1920-2019.
Figure 6 (Right). The graph plots out Power Ratio with accordance to year. The 6 different baseball eras as well as the year 2020 are color coded with accordance to the key to the right of the graph.

Figure 1 shows the strong, positive, linear relationship between Power Ratio and winning percentage, with the line of best fit running through (1.00, 0.5), showing that an average team, with a winning percentage of 50%, is expected to have a Power Ratio of 1.00. Figure 2 shows that this ratio does not favor any era in particular, with top teams in the power ratio ranging from the 2020 Los Angeles Dodgers to the 1944 St. Louis Cardinals to the 1927 New York Yankees.

The Big Out. Bases loaded. Nobody out. What is the best way to get out of this jam? A strikeout. On the flip side, if you are the team up to bat, what is the worst-case scenario? A strikeout. Baseball is a game of inches, and the moment the ball is in play, anything can happen. Therefore, it makes sense that the more strikeouts you get while pitching, and the less you have while hitting, the better of a chance to win. Going hand in hand with strikeouts in the walk, a free path to first base that does not allow a fielder to make a play nor let the batter get themselves out. In baseball, if a batter gets out 7 out of 10 times, they are considered a great player, meaning that the pitcher has a higher probability of getting someone out if they throw strikes. Walks take this probability, throw it out the window, and let the batter scamper to first base. In order to weight walks, which usually do more damage than a strikeout, the walks and walks allowed were both multiplied by a factor of 3.

$$PowerRatio = \frac{\text{Pitching Strikeouts} + 3 * \text{Batting Walks}}{\text{Batting Strikeouts} + 3 * \text{Pitching Walks}}$$

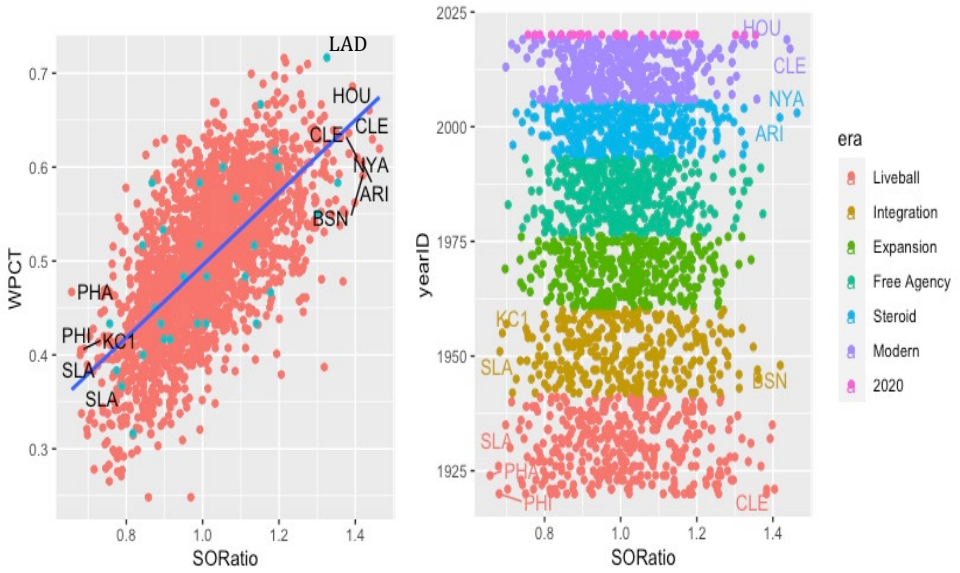

Figure 7 (Left). This graph plots out Strikeout Ratio (SORatio) to win percentage (WPCT). Blue dots are 2020 teams and red dots are teams from 1920-2019. The best and worst teams are shown by their abbreviation.
Figure 8 (Right). The graph plots out Strikeout Ratio with accordance to year. The 6 different baseball eras as well as the year 2020 are color coded with accordance to the key to the right of the graph.

Figure 7 shows that there is a moderate, linear relationship between Strikeout Ratio and winning percentage, with the line of best fit running close to (1.00, 0.5), showing that an average team, with a winning percentage of 50%, is expected to have a Strikeout Ratio of 1.00. Although not as strong as the previous models, there is still a clear correlation. Figure 8, meanwhile, shows that the Strikeout Ratio does not favor old nor new teams. It appears there may be more spread with older teams, with more higher teams as well as more lower teams, including the worst 5 Strikeout Ratio teams all coming from the Integration and Liveball eras. This spread, however, evens out, with the average Strikeout Ratio remaining nearly even across all the eras.

The FloStength Score. As evident from the earlier, seasons vary, and even if that variance is slight, it is difficult to be able to compare the best teams across seasons. That is where the Flo Strength Model comes in, an effort to try and compare teams across seasons by equalizing for slight year-to-year differences. This model also helps aggregate the different aspects of baseball, shown through the ratios, into a single number. The term *sd* is shorthand for standard deviation from the yearly mean, with where this comes from explained later.

$$FloStrength = 93.81 * (sd\ from\ yearly\ mean)RRatio + 84.95 * sdOBRatio + 83.26 * sdPowerRatio + 65.66 * sdSORatio + 1000$$

The FloStrength formula appears a bit complicated at first, but with the following discussion will actually be quite simple. This formula is the main personal contribution.

Table 4. Ratio correlation to win percentage

Data	Correlation
RRatio	0.938
OBRatio	0.849
PowerRatio	0.833
SORatio	0.657

In order to weight each ratio in accordance with their ability to helping a team win a game, the correlation to win percentage is used in the formula. Correlation is the measure of how two variables are linearly related, in this case meaning how much each ratio is a linear predictor for winning percentage. This correlation is than multiplied by 100, for the simple reason of making each number bigger.

Every individual team has their own ratios and ratios do help reduce differences between seasons, but it is not perfect still. It The easiest, most complete way to normalize data across seasons is to utilize standard deviations compared to the season mean. It can be assumed that each season can follow a normal model due to the large sample size and standard deviations are compared to a year-by-year mean for comparative purposes. This helps equalize for slight differences seen in a broader, era-by-era scale seen in table 5. In the model, each team's individual standard deviation from the yearly mean for their ratios is used, and then multiplied by the ratio correlation to weight the ratios equally to their weight of winning a baseball game. Because standard deviations can be negative, usually on a scale of -3 to +3 (meaning 3 standard deviations below the mean to 3 standard deviations above the mean), the number 1000 is then added to make every FloStrength number a positive number. Adding a number to all values does not affect the model's accuracy or standard deviation. Adding 1000 just makes every year's mean 1000, perfect for comparison purposes.

Table 5. Average baseball ratio by era.

Data	Live Ball	Integration	Expansion	Free Agency	Steroid	Modern
Run Ratio	1.013	1.019	1.013	1.009	1.012	1.012
On Base Ratio	0.997	0.992	1.013	1.006	0.985	0.999
Power Ratio	1.004	1.004	1.003	1.002	1.004	1.004
Strikeout Ratio	1.008	1.010	1.009	1.007	1.007	1.007

The idea behind this model is combining several different aspects of baseball, as seen through the 4 previous models, into one model, with this model also helping to account for how much these aspects attribute to a win by using the correlations. The FloStrength model also accounts for slight differences from year to year by comparing to a yearly mean. A final correlation of 0.905 is calculated, showing a strong, linear correlation between the FloStrength score to winning percentage. Although this number is lower than the correlation for run ratio, it is more showing to the overall strength of a team as it includes several different aspects of baseball and into winning.

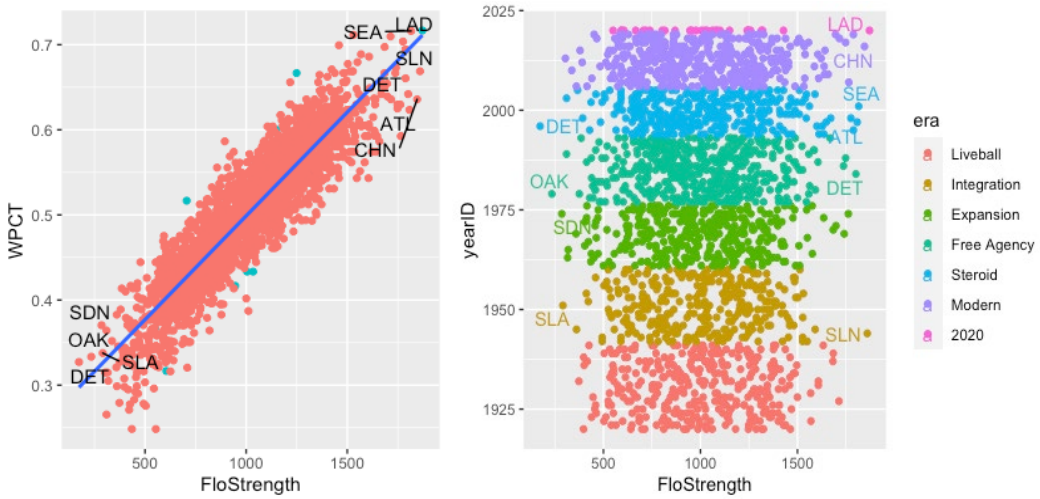

Figure 9 (Left). This graph plots out FloStrength score to win percentage (WPCT). Blue dots are 2020 teams and red dots are teams from 1920-2019. The best and worst teams are shown by their abbreviation.
Figure 10 (Right). The graph plots out FloStrength score with accordance to year. The 6 different baseball eras as well as the year 2020 are color coded with accordance to the key to the right of the graph.

The line of best fit in this figure runs right through (1000, 0.5), meaning the average team with a win percentage of 0.5 will have a FloStrength score of 1000. A final correlation of 0.905 is found, showing a strong,

linear correlation between the FloStrength score to winning percentage, as seen in figure 9. Although this number is lower than the correlation for run ratio, it is more showing to the overall strength of a team as it includes several different aspects of baseball. Figure 10 then shows that this score balances out the eras and different seasons, equally comparing both old and new teams to each other.

THE 2020 LOS ANGELES DODGERS

Going into the season as the favorites to win it all, the Los Angeles Dodgers were ready to break through in 2020. After falling short of lofty expectations the last three years, including a World Series loss in 2017, the Dodgers put together a team that on paper looked like they would be ready to compete.

The Eye Test

Just one glance at the Dodgers could tell you that this team was great. In the outfield was two former MVPs in Mookie Betts and Cody Bellinger, both of whom were just about to hit their prime. Joining them in the outer part of the field was former all-star AJ Pollock, creating a fearful trio roaming the outfield. The power filled infield was led by 2x All-Star and 2x Silver Slugger Corey Seager. Former All-Star Justin Turner manned the hot corner, and reigning All-Star Max Muncy played across the diamond at first base. Playoff hero Enrique Hernandez wrapped up the infield. Standout rookie Will Smith would be playing behind the plate. And then the pitching staff. The rotation was led by future Hall of Famer and 8x All-Star Clayton Kershaw, who also had 3 Cy Young Awards and a MVP to his name. Right behind him in the rotation was All-Star Walker Buehler, Electric rookie Dustin May, and superb southpaw Jose Urias. The bullpen was led by 3x All-Star, 2x Reliever of the Year, closer Kenley Jansen. Backing him up was Rookie of the Year Tony Gonsolin, veterans Dylan Floro, Adam Kolarek, Pedro Baez and Jake McGee, young flame thrower Brusdar Graterol, and another promising rookie in Victor Gonzalez. In total, the Dodgers had a superb pitching staff that was backed up by an unbelievable offense, a deadly mix for opposing teams, and a mix that put the Dodgers in GOAT conversation.

The Stats

The Los Angeles Dodgers, at first glance of statistics, stand out in respect to their counterparts. They ranked in the top 5 in 9 of the 12 basic categories seen in table 6. However, their balance between offense and defense is even more amazing, as evident by their ratios and FloStrength score seen in table 7.

Table 6. 2020 LA Dodgers Basic Stats

Data	Statistic	2020 Ranking
R/Game	5.817	1
RA/Game	3.55	2
OBP	0.338	5
WHIP	1.060	1
HR/Game	1.967	1
HRA/Game	1.1	2
H/Game	8.717	7
HA/Game	7.07	3
SO/Game	7.850	4
SOA/Game	8.617	18
BB/Game	3.800	10
BBA/Game	2.417	1

Table 7. 2020 Dodgers Ratio Statistics

Statistic	2020 Dodgers	2020 Rank
RRatio	1.638	1
OBRatio	1.349	1
PowerRatio	1.348	1
SORatio	1.326	2
FloStrength	1871	1

These ratios and FloStrength score obviously are superior to their season counterparts, however this story dives even deeper. The 2020 Los Angeles Dodgers FloStrength score of 1871 is not only the top score for 2020, but also for every season through 1920. 101 seasons of data were compiled, and this year's Dodgers team came out on top. As seen in table 8, this superiority takes over some of the consensus greatest teams in history, including the drought-breaking 2016 Chicago Cubs, the 1944 St. Louis Cardinals team that had the three most recent MVP's in their lineup, and the 2001 Seattle Mariners who won the most games of all time with 116. Also included in the top 6 Flo Strength teams was the 1997 Atlanta Braves led by their 3 Hall-of-Fame starting pitchers and the 1984 Detroit Tigers, one of the greatest all-around teams of all time [1].

Table 8. Comparison of ratios for greatest FloStrength teams of all time

Data	2020 LAD*	1944 SLN*	2016 CHN*	2001 SEA	1997 ATL	1984 DET*
RRatio	1.638	1.575	1.453	1.478	1.361	1.289
OBRatio	1.349	1.195	1.251	1.235	1.174	1.130
PowerRatio	1.348	1.268	1.246	1.230	1.184	1.169
SORatio	1.326	1.209	1.207	1.214	1.190	1.130
FloStrength	1871	1860	1845	1815	1808	1801

*stands for World Series winner

This domination was not based on one single statistic or ratio, but rather a complete, well rounded team. The Dodgers had the second-best *Run Ratio* of all time at 1.638, although even with this high of a ratio, they underperformed compared to their expected win percentage given the linear regression. Although they scored a lot more runs than their opponents, when it came to one-run baseball games, they "struggled" to a 7-5 record. Although a .583 winning percentage in one run games is nothing to be mad about, when compared to the .750 winning percentage in non-one-run games, it is a bit lackluster. This discrepancy can help explain why the Dodgers underperformed despite outscoring their opponents by so much.

The Dodgers *On Base Ratio* was by far the highest of all time at 1.349, meaning that they got on base 1.349 times per 1 batter allowed on base. Even with such a high on base ratio, they still underperformed when compared to its linear regression expected win percentage. According to TeamRankings, the Dodgers finished to 2020 season with the 7th highest number of double plays hit into, while turning just the 16th most double plays per game. [6] The Dodgers also had the 5th lowest average with the bases loaded, a very low number especially when connected to the 5th highest plate appearances with the bases juiced. This means that some of the Dodgers base runners were erased, something the On Base Ratio does not show and the Dodgers were below average in getting runs in when they had the bases loaded, explaining why their on base ratio was so high compared to their winning percentage.

The 2020 Dodgers finished with the highest *Power Ratio* of all time at 1.348. Such a high ratio compared to expected win percentage points to them actually underperforming when it comes to games won. A flaw of the Power Ratio is that it does not account for players on base when homeruns are hit or other for players on base when other hits are recorded. As seen earlier, the Dodgers had a very high On Base Ratio, meaning they had a lot more guys on compared to their opponents. Combined with the Power Ratio, this means that the Dodgers had a lot of players on base when they got hits and homeruns, and also means they did not have a ton of guys on base when they gave up hits and homeruns. This flaw in the Power Ratio helps attribute to why the Dodgers under-performed compared to their power ratio.

The Los Angeles Dodgers finished off with a *Strikeout Ratio* of 1.326. Although this was well above average, it was not the highest. In fact, when compared to expected win percentage with this ratio, the Dodgers way over performed compared to their Strikeout Ratio. So how did they do it? Although their pitching staff did not strikeout a high number of batters, nor did they walk an extremely low number of batters. What this means is that balls were put in play, and there were usually people on base. Their offense was fantastic, with very few strikeouts and a fair number of walks, leading to their high Strikeout Ratio, even as their pitching staff brought it to a lower standard that it should have been. Luckily for the Dodgers, and the reason why they outperformed their strikeout ratio, is for their defense. The Dodgers had top of the line fielders in Mookie Betts, Corey Seager, Enrique Hernandez, and Cody Bellinger were all top tier defenders, and the team as a whole was ranked second in defensive runs saved and second in defensive efficiency in 2020. Their top-of-the-line defense helped support a pitching staff that was not immune to struggling with walks and also struggled at times to get strikeouts.

COMPARING THE 2020 SEASON TO PAST SEASONS

Despite COVID-19 running through America, the MLB season went smoothly, except for the early season hiccups that the St. Louis Cardinals and the Miami Marlins experienced, hiccups that caused a number of players from these teams to miss games, as well as several other games getting delayed. However, MLB was able to take care of these early issues and flow along the 2 month, 60-game season. Unfortunately, the shortened season brought questions about the validity of the season as a whole. Would the World Series winner go down in the history books with an asterisk? To answer this question. 1-sample-z-tests were performed to compare some basic stats as well as the ratios, as seen in Table 8 and 9. Z-tests are used when there is a population mean, and t-tests are used when there is a sample taken from a population. In this case, the complete population (1920-2019) is being compared, meaning a z-test should be used. The single mean in question is the 2020 season. For significance levels, the alpha levels of

0.05 and 0.01 were chosen. These mean that there is a 5% and 1% chance, respectively, that the differences found by the tests were by chance. These are common alpha levels, as 0.01 is used in high-level research projects to find significance to fully assure that the findings are not by chance and 0.05 is used in more general, low-level research. Extremely low and high z-scores have p-values that are below the alpha levels, meaning that there is significant differences in the findings.

Table 9. Z-test of basic stats for MLB

Data	Z-Score
R/Game	1.5020
OBP	-0.0230
WHIP	-2.2780*
HR/Game	8.2020^
H/Game	-7.9590^
SO/Game	11.0460^
BB/Game	1.1346

* Significant at 0.05 alpha level
^ Significant at 0.01 alpha level

Table 10. Z-test of ratios for MLB

Data	Z-Score
RRatio	0.29180
OBRatio	2.63060^
PowerRatio	0.1942
SORatio	0.20762
FloStrength	0.00000

There are some obvious significant differences seen here. WHIP and Hits per Game were way down compared the mean over the last 100 years, and HR per Game and Strikeouts per Game were way up, but that is where ratios come into play. With the ratios, only on base ratio was significantly different, thanks in part to the WHIP and hits per game being way down. The question then comes whether these differences were truly that different, or if these differences could be explained by simple trends. For this, ANOVA tests were performed to compare the means of the current era, the 2019 season, and the 2020 season, seen in tables 10 and 11. ANOVA tests are used to compare more than 2 means against one another, in this case the Modern Era, the 2019 season, and the 2020 season were all compared to each other.

Table 11. ANOVA test of basic statistics for MLB

Data	Modern Era	2019	2020
R/Game	4.459	4.83*	4.642
OBP	0.315	0.311	.322*^
WHIP	1.339	1.335	1.327
HR/Game	1.036	1.395*	1.282*
H/Game	8.797+	8.654+	8.035
SO/Game	7.399	8.815*	8.669*
BB/Game	3.162	3.272	3.387*

* Significantly greater than Modern
^ Significantly greater than 2019
+ Significantly greater than 2020

Table 12. ANOVA test of ratios for MLB

Data	Modern Era	2019	2020
RRatio	1.012	1.021	1.022
OBRatio	0.999	0.995	1.034
PowerRatio	1.003	1.007	1.007
SORatio	1.007	1.022	1.011
FloStrength	1000	1000	1000

There were some slight differences in the more basic stats, but the ratios are not significantly different, meaning that the differences seen earlier were explained more by recent trends. Through these tests, it is concluded that the are no differences between the 2020 season and past seasons, allowing a comparison to be made.

CONCLUSION AND FUTURE WORK

Through statistical analysis, it can be concluded that the Los Angeles Dodgers of 2020 are among the best, if not the best, team of all time. Although the 2020 season was strange for many reasons, the use of ratios, correlations, and standard deviations to yearly means allows the 2020 season to be comparable. Critics will oppose this idea by saying that the shortened season does not allow for outliers to come closer to the average. Critics will also point out that the Dodgers only had to play 9 teams instead of the typical 24 teams. These 9 teams could, hypothetically, be the worst 9 teams in the league, inflating the Dodgers numbers. In reality, the Dodgers actually had a slightly above average strength of schedule at 0.08, good for 14th out of 30 teams. [6] This means they played an average amount of good, decent, and bad teams along their way. All this criticism of the abnormal season is fair; however, statistic tests showed there was no evidence of statistical differences between the 2020 season, the 2019

season, or other recent years. In future expansions of this model, adding in a strength of schedule aspect could improve the model. Adding in a smaller aspect to account for coaching, whether with career length or career win percentage, could also improve the model. Future applications of the FloStrength model could include simulations, predictions, and sports betting. The FloStrength model could also easily be translated to the NFL and NBA, with offensive and defensive ratios that highlight the most important parts of these sports swapping out the baseball ratios.

EXTRA THANKS

An extra thanks would like to be given to Dr. Douglas Baumann for proofreading and giving feedback of the statistical analysis work performed, Eric Eager from Pro Football Focus for input on the paper, and Mike and Sarah Floersch for helping me find the love for mathematics and statistics.

WORKS CITED

[1] “1997 Atlanta Braves Statistics.” *Baseball*, www.baseball-reference.com/teams/ATL/1997.shtml.
[2] Albert, Jim, et al. *Analyzing Baseball Data with R*. CRC Press/Taylor and Francis Group, 2019.
[3] Costa, Gabriel B., et al. *Understanding Sabermetrics: an Introduction to the Science of Baseball Statistics*. McFarland & Company, Inc., Publishers, 2019.
[4] Lahman, Sean. *Lahman's Baseball Database* , 2020.
[5] Lewis, Michael, and Bennett Miller. *Moneyball*. Sony Pictures Home Entertainment, 2012.
[6] “MLB Baseball.” *MLB Baseball on TeamRankings.com*, www.teamrankings.com/mlb/.
[7] “MLB Stats: Baseball Stats.” *MLB.com*, www.mlb.com/stats/team.