Home Runs in the MLB

William Gephart

SN: 010897778

**Introduction:**

The intention behind this study is to take a look at statistics in Major League Baseball to determine how home runs have affected the overall scoring in the major leagues from the years 2005 to 2019. The problem statement for this analysis is whether or not home runs have had a positive effect on overall scoring in the MLB. Has the recent surge in home runs across the MLB had a positive effect on run output or is this change in offensive approach misguided in the results being shown on the field?

Over the years, player approach at the plate in MLB games has drastically changed. In the 1990s and early 2000s, there was a huge emphasis on the offensive approach of small ball baseball. This generally looks like getting people on base and then doing whatever it takes to move those runners into scoring position so that they can score runs more easily. This strategy was heavily used by the Oakland Athletics when they coined the term “Moneyball”. There was a well-known movie and book made about this story of how the Athletics searched for offensive players using metrics that hadn’t been seriously looked at in previous years. As far as the offensive statistics go, they looked quite heavily at the on-base percentage (OBP) to determine whether a player was worth the money they were going to spend on them. Having players that were able to get on base suddenly became a priceless commodity in the modern game, and cause other teams to take notice after the Athletics' tremendous success in the 2002 season. Over the past 5 years, however, the game has begun to change again to an approach that has seen the home run ball skyrocket as part of the game's offensive playstyle. Teams are now not only looking for guys who can get on base, and effectively move players around, but they are also looking at players who can drive the ball into the outfield and have tremendous power across multiple statistics.

Part of the reason that this approach has changed in recent years is due to the make of athletes at the professional level. Players nowadays have become more powerful and athletic across the board and it is starting to become a consistent part of the MLB where players have multiple tools that they can draw from (hitting, hitting for power, running, fielding, and throwing). Another reason for this change in approach is the change in how baseballs are made at the major league level. According to Alan M. Nathan at the University of Illinois, the increase in home runs was due primarily to how the ball was made. The air drag in the ball was changed so that the aerodynamic properties were improved overall. This meant that given certain conditions in the launch of the ball, the ball would travel further resulting in a home run. This change was so drastic and caused such a surge in home runs that the MLB introduced a new ball in 2021 that they hoped would decrease the properties of the ball that were leading to such a large number of home runs in recent years.

**Research Questions and Hypothesis:**

The two questions I will be answering in this analysis are the following: “How much have Home Runs increased from year to year over the years 2005-2019?” and “Has the home run surge in the last 5 years had a positive impact on the number of runs scored in the average season?” After taking a look at these questions I have come up with the hypothesis that home runs have affected the scoring output in the MLB in a positive way rather than the average runs per year decreasing.

I decided to go this route in my hypothesis based on my own experiences in watching baseball and noticing trends throughout the league over the years. I have been an avid baseball fan throughout my life, and I can quite honestly say that I have never seen statistics on home runs being put out on broadcasts this frequently in my entire life. Now more than ever am I seeing stats about average home runs hit per game, team records when they hit at least 1 home run in a game, or teams record when they hit 2 or more home runs in a game. I chose this time frame because I know that the 2000s were a time when small ball baseball was as prevalent as ever, and the past 5 or so years have changed that. I intend to look at different statistics such as home runs per game, average home runs per season, average runs per season, and home runs on batted balls over all of these seasons to develop an analysis that will allow us to get a small glimpse into what the changes in the MLB have done to affect the game in the past 15 years.

**Method:**

The data that I collected came from the baseball reference website. I began by going year to year for each MLB team and looking at the number of runs they scored and the amount of home runs that they hit from the years 2005 to 2019. I then took the averages of each of these numbers to create a separate table telling me the average number of runs that each team scored in that respective season as well as the average number of home runs that each team hit in each season as well.

I used Excel to create these tables and also used the functions within excel to categorize and average this data into their respective categories. After doing that, I then used R-Studio and began writing code that would allow me to create various graphs and plots to visualize the data in an easier-to-understand type of way. I created graphs for the home runs hit per batted ball each year, and I also created a plot that gave me an idea of how many home runs were hit and how that affected how many runs were scored as well.

After analyzing this data within R, I began taking a look around the internet to get a better idea as to what factors could potentially be causing this surge, how those factors might affect the outcome of these factors, and how these could potentially affect the hypothesis that I made for this research study.

**Results:**

Chart, bar chart

Description automatically generatedChart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generatedChart, bar chart

Description automatically generated

After looking at the graphs that I was able to create with the data I collected, I was able to confirm my initial hypothesis that Home Runs had a positive effect on Runs scored on average per season. We saw that there was a strong increase as predicted in the home runs hit per batted ball and that the uptick in the number of home runs in the past 5 or so years was reciprocated by the number of runs that were scored as well.

The first graph that I would like to take a look at is the graph titled, “HR/BIP By Year”. This graph shows the percentage of balls in play by year that ended up being home runs. As can be seen in the first 9 years of this graph, there is a somewhat steady decline in terms of the percentage of balls in play that ended up being home runs. Based on the information I gathered from BaseballTrainingWorld.com which received their baseline statistics from Baseball-Reference.com, I was able to see that there was a decrease in what percentage of balls in play were home runs from 11.4% to 10.06% in the years 2005 – 2014. This was followed by a somewhat steady increase from that same 10.06% to 16.12% in the years 2014 – 2019. This might not seem like a drastic change to the naked eye, but a 6% difference in MLB can be the difference between someone hitting .300 and someone hitting .240 as their batting averages. This did a great job of showing how balls in play changed a considerable amount in the past 15 years.

The next two graphs I would like to take a look at are the “Average Home Runs per Year”, and “Average Runs per Year” bar graphs shown above. As was expected in the home runs bar graph we see an increase over time in the average number of home runs that were hit per season. When looking at the number of average home runs in 2005 starts at a somewhat reasonable level sitting at 167 home runs. That steadily drops off hitting its lowest point in 2014 when it reached 139 home runs on average per season. That’s just .86 home runs per game which are staggering when compared to the modern game. After 2016 there was a dramatic increase like we saw in the first graph up to 225 home runs on average in a season. This changes the average home run per game total to 1.39 home runs per game. Just over half a home run more per game in comparison to the 2014 total. This might not sound like a lot on paper, but it has shown results in the run totals as seen in the second bar graph where we see the average run totals per season. In looking at just the same two years with the Average runs bar graph above we can see that 2014 produced the lowest average runs per season of this 15-year data set, and the 2019 season produced the second-largest run total in this data set as well. This data set also shows us that from the years 2005 – to 2014 the Average runs scored per game in those 10 seasons was 728.1. From 2015 to 2019 the average was 736.2 which isn’t as large as I expected it to be, but that could partially be a result of the lower totals in 2016 and 2017.

Lastly, I wanted to create a graph that would show the individual plot points as well as a trendline when looking at Home Runs and Runs Scored in individual seasons across these 15 years. This is shown in the graph titled, “Home Runs vs Runs Scored”. As we look at this graph we get a good idea of how unique the years 2015-2019 have been in comparison to the previous 10 years. As we look at the graph we can see that the years 2015-2019 all fall above the trendline in this graph with 2015 being the closest to the trendline and 2019 being the farthest away with the highest home run and 2nd highest run total in this data set. Every plot point from 2005 to 2014 falls under the trendline showing that there is a significant difference in the effectiveness of the offensive approach after the year 2015 in the MLB.

**Discussion:**

There are several implications that I believe can be taken out of this research study. First off, we can take a look at player makeup to get an idea of whether or not that has played a role in the ball traveling further. We know that there is more analytics now than ever, and baseball was one of the pioneering sports that brought analytics into the athletic world. As mentioned before, there are all sorts of tools that make up a major league baseball player and power is one of the top five areas that scouts and coaches have been looking for in players on the offensive side of the ball. Having the ability to hit a line drive, and not only go for extra bases but hit home runs is an invaluable trait in today’s game. It will be interesting to see how changes to the game in the future could potentially impact the way scouts look at and analyze players that they could potentially draft or sign for their organization.

A second implication from this data is that there would need to be a change in the ball that could make it fairer for pitchers to participate at the major league level competitively. I remember back in 2019 when the home run surge it its peak that there were all sorts of conversations about changing the distance from home plate to the pitcher’s mound, or changing the height of the pitcher's mound to make it more competitive for pitchers. These are all ideas that are currently being tested in the minor league baseball system among other things but the most immediate approach that the MLB took was to change the design of the baseball. According to mlb.com, there were some inconsistencies in the seam height on the ball which was causing there to be less drag through the air. This, in turn, caused the ball to travel farther through the air. In response, Rawlings (the company that manufactures MLB’s baseballs), loosened the tension on the wool windings that are in the ball. They expected this to cause the ball to travel several feet less on balls hit over 375 while also not causing any change in pitcher velocity.

The last implication that I believe can be made is that professional sports are going to be continuously evolving from year to year and will never be the same from decade to decade. Looking at a sport like football, we have seen a dramatic shift in emphasis from defensive heavy schemes to offense heavy causing more scoring and more exciting play all around. Similarly, with the invention and evolution of the three-point shot, basketball has continuously moved closer to a model that promotes three-point shooting, and strong paint presence over a mid-range game which was seen for most of the early parts of basketball history. Like those two sports, baseball will continue to grow and evolve as it goes on. There has been a massive surge in home runs, and home run hitters in the past decade, and that could potentially change in the coming years as well. However, I find it fair to state that there was a change in offensive approach in the 2010s that cause such a drastic rise in home run rate, and effectiveness.

There were a few limitations I noted while conducting this research. First of all, it would be nice to be able to look into the future and see how the changes that the MLB has made to the ball could potentially affect home runs going forward. I would love to have been able to take a look at the 2020, 2021, and 2022 seasons to get an idea of what results have come from these changes, but the pandemic and the drastic change to ball makeup made it tough to get an accurate idea of what the statistics are saying. I will say that in the year 2022 there has been a dramatic decrease in the home run rate so far. This could be due to the shortened offseason from the lockout, or could also be a result of the game continuing to change but only time will tell. Another limitation to this study was not having the ability to research team makeup over these 15 years. There’s something I would like to call the “Barry Bonds Effect” in a study like this wherein any given year there could be a player that hits 50-70 home runs causing a massive outlier in home run data. It would be fascinating to get a better idea as to what these statistics would look like while also having player, team, and even ballpark data to go off of.

**Appendix:**

Chart, bar chart

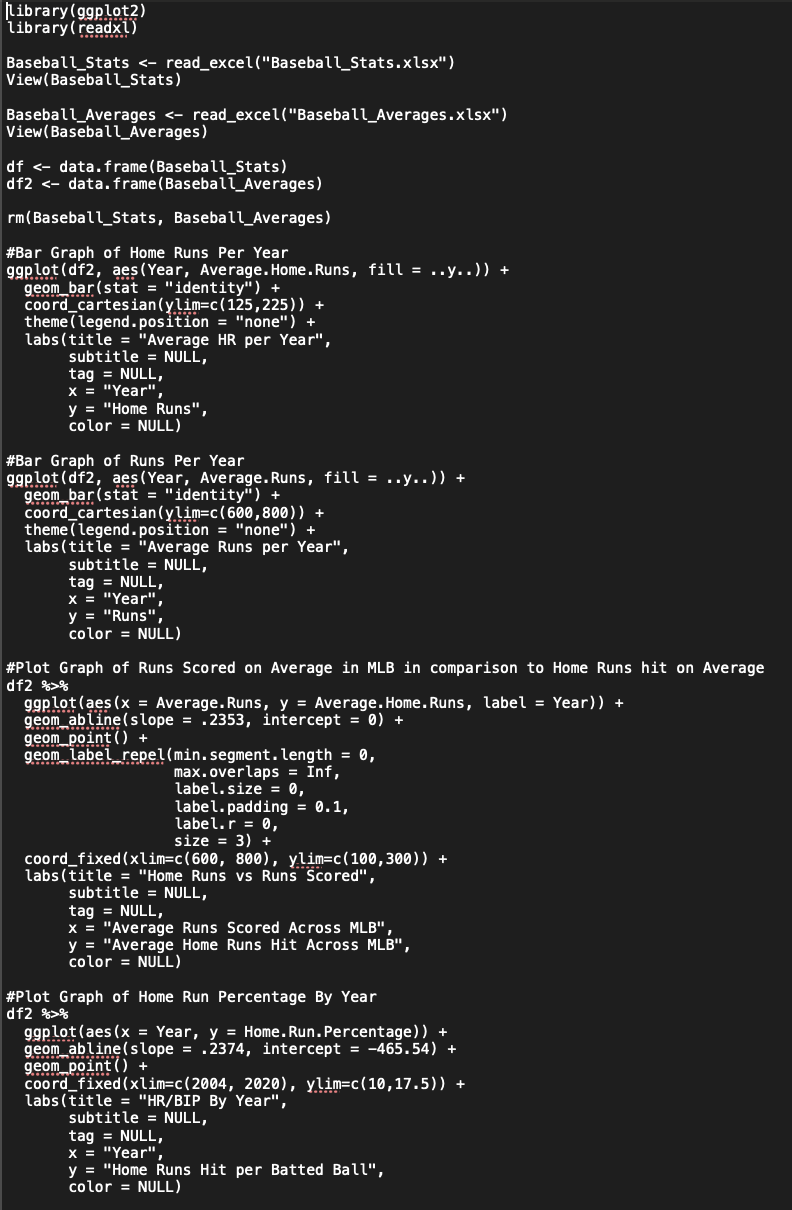
Description automatically generatedChart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generatedChart, bar chart

Description automatically generated



**References:**

Feinsand, M. (2021, February 9). *MLB to alter baseballs for ’21 season*. MLB.Com. Retrieved May 3, 2022, from https://www.mlb.com/news/mlb-to-alter-baseballs-for-2022

Long, J. (2019, September 26). *R Cookbook, 2nd Edition*. Rc2e.com. Retrieved March 5, 2022, from https://rc2e.com/index.html

*Major League Batting Year-by-Year Averages*. (n.d.). Baseball-Reference.Com. https://www.baseball-reference.com/leagues/majors/bat.shtml

Nathan, A. M. (2019, December 12). *The Home Run Surge*. The Physics of Baseball. Retrieved May 2, 2019, from http://baseball.physics.illinois.edu/homeruns.html

Nelson, S. (2021, April 18). *What Percentage of Hits are Home Runs in the MLB?* Baseball Training World. Retrieved May 3, 2022, from https://baseballtrainingworld.com/what-percentage-of-hits-are-home-runs-in-the-mlb/