Do Home Runs Win Games?

By: Carter, Braden & Joseph

Definitions:

Winning Percentage: Shows what fraction of games the team won during the season.

Calculated By: Winning Percentage = Wins ÷ (Wins + Losses)

Example: A team finishes the season with 92 wins and 70 losses.

Winning percentage = 92 ÷ (92 + 70) = .568 or 56.8%

Home Run: When a batter hits the ball and is able to round all the bases and score a run without being put out, usually done by hitting the ball over the outfield fence in fair territory.

MLB: Major League Baseball

Goal:

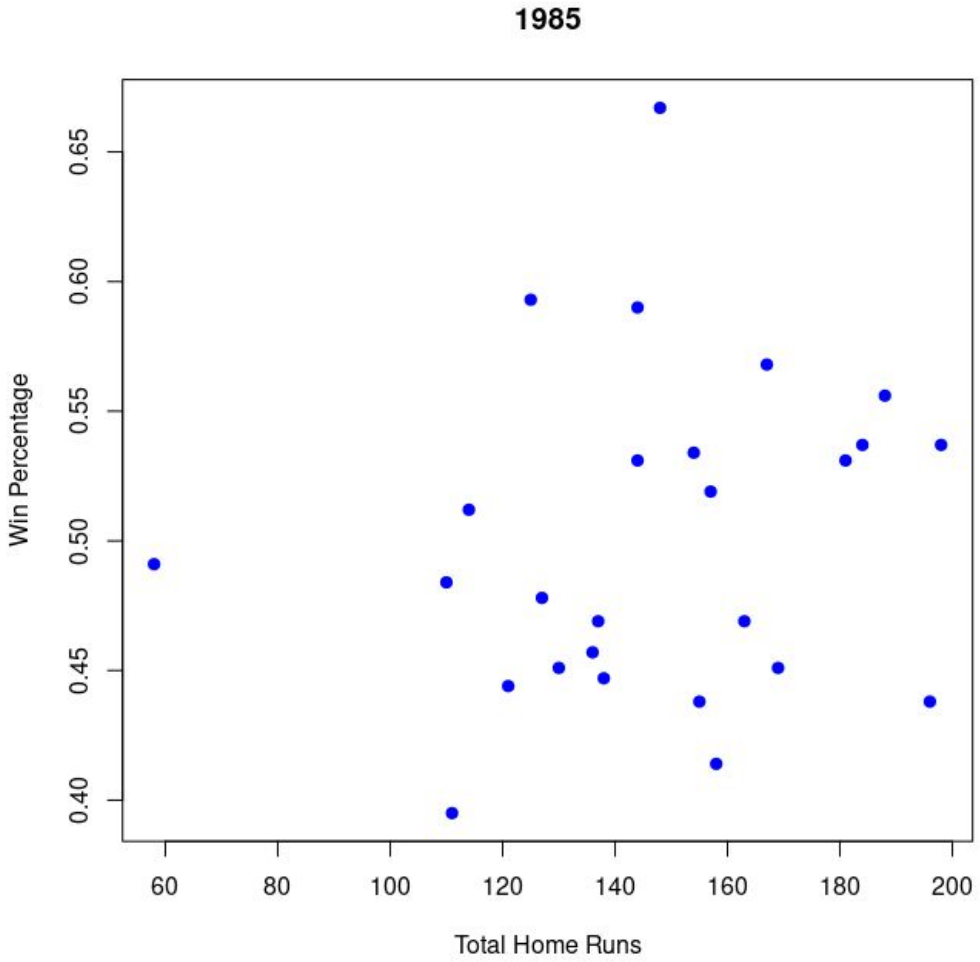
Our goal with this project is we want to see if there is any correlation between Major League Baseball (MLB) teams winning percentages and the number of home runs they hit over the course of a season. We want to use our skills with plots, graphs, and charts to share our results.

What We Did:

- We randomly generated a year from each decade starting with the 1980's all the way up till the 2020's.

- We used the package baseballr, which uses data various sources like FanGraphs, Baseball-Reference, and Baseball Savant.

- And with the data, we put it into R and created scatter plots to visualize our results.



World Series Winner:

Kansas City Royals

154 Home runs

Winning %: .562 or 56.2%

Most Homeruns:

Baltimore Orioles (12th best team)

214 Homeruns

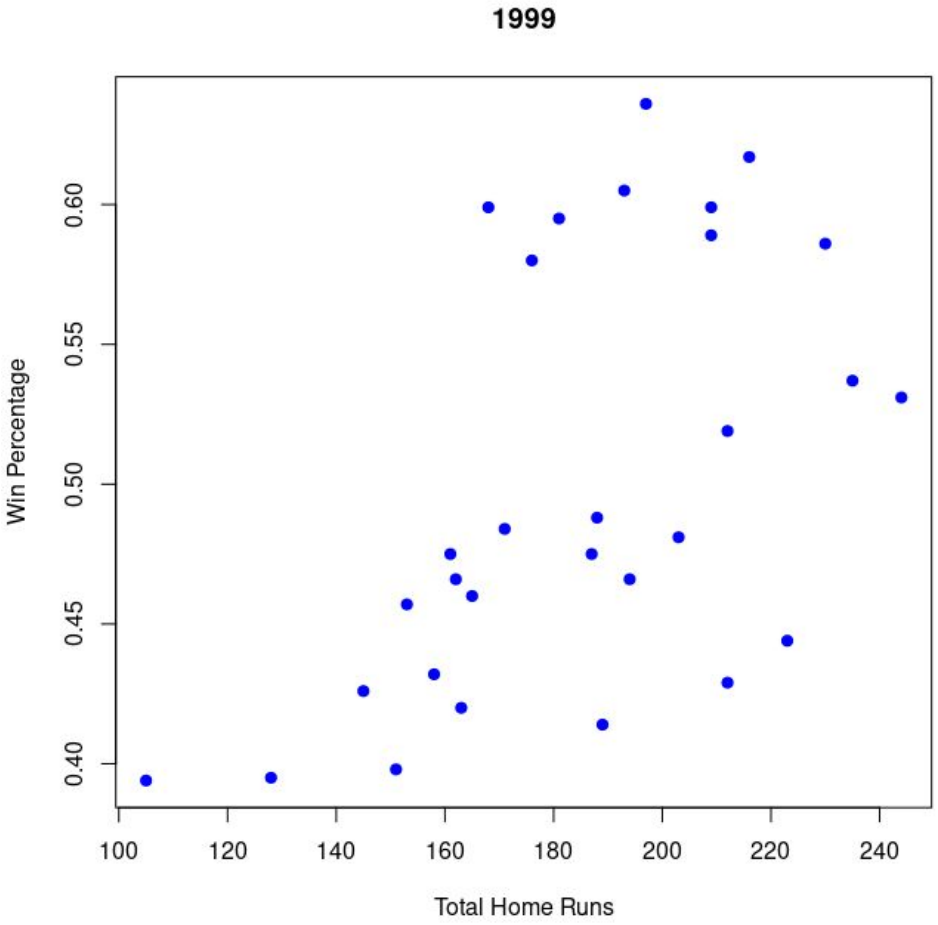
Winning %: .516 or 51.6%

Least Homeruns:

Pittsburgh Pirates (Worst team in MLB)

80

Winning %: .354 or 35.4%



World Series Winner:

New York Yankees

193 Home runs

Winning %: .605 or 60.5%

Most Homeruns:

Seattle Mariners (13th best team)

244 Homeruns

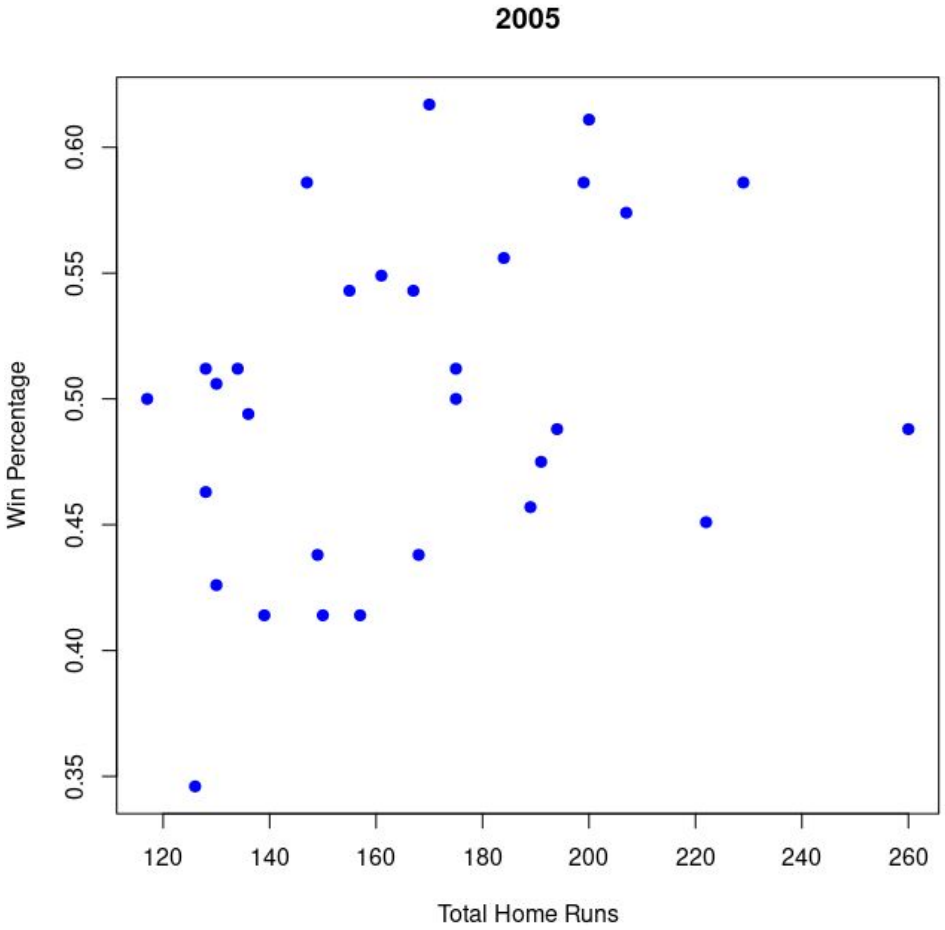
Winning %: .488 or 48.8%

Least Homeruns:

Minnesota Twins (Worst team in MLB)

105

Winning %: .394 or 39.4%



World Series Winner:

Chicago White Sox

194 Home runs

Winning %: .611 or 61.1%

Most Homeruns:

Texas Rangers (18th best team)

260 Homeruns

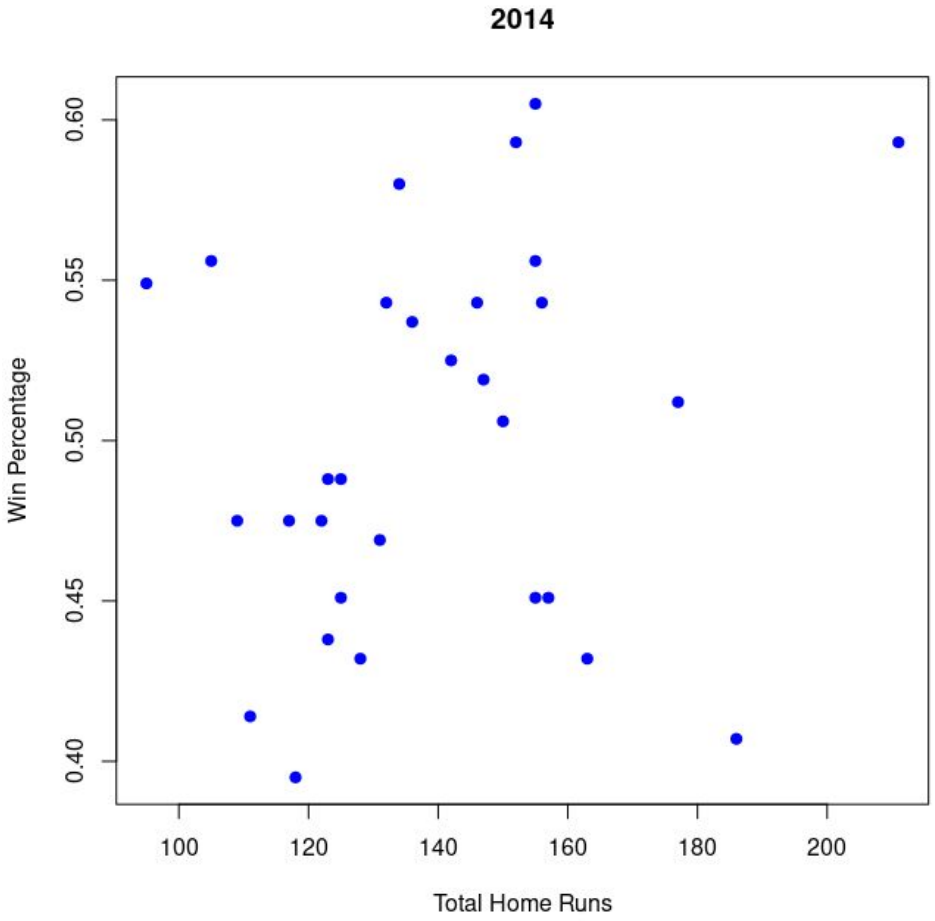
Winning %: .488 or 48.8%

Least Homeruns:

Washington Nationals (16th best team)

117

Winning %: .5 or 50%



World Series Winner:

San Francisco Giants

132 Home runs

Winning %: .543 or 54.3%

Most Homeruns:

Baltimore Orioles (3rd best team)

211 Homeruns

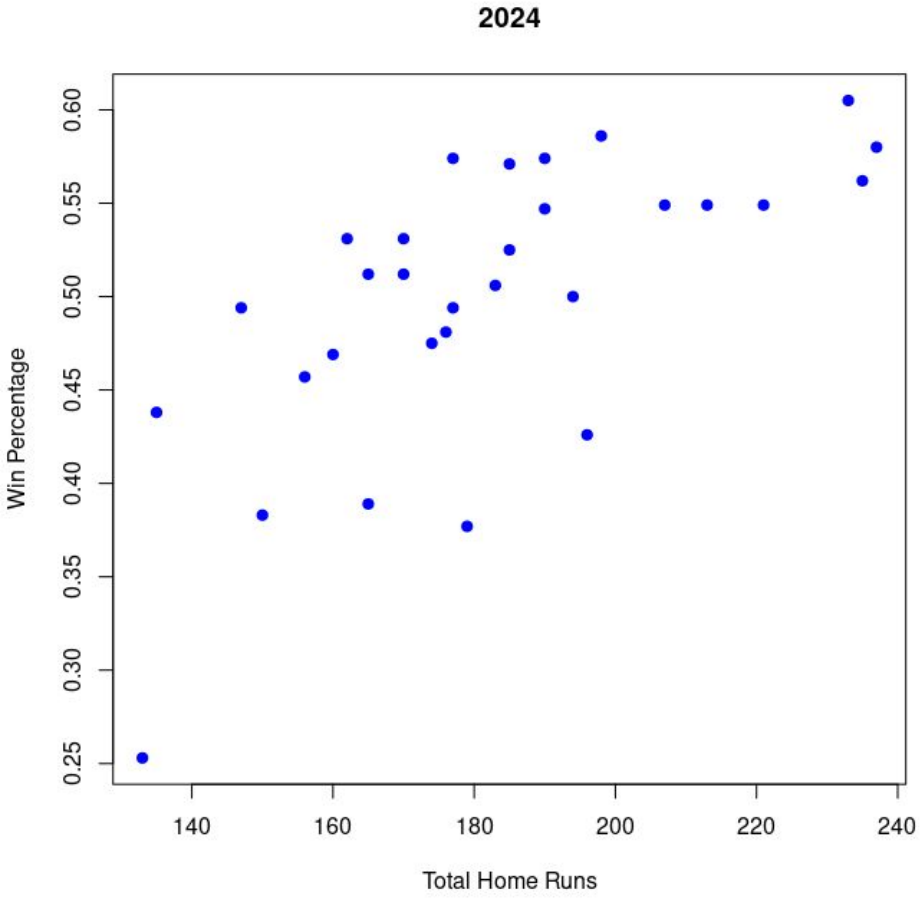
Winning %: .593 or 59.3%

Least Homeruns:

Kansas City Royals (7th best team)

95

Winning %: .549 or 54.9%



World Series Winner:

Los Angeles Dodgers

233 Home runs

Winning %: .605 or 60.5%

Most Homeruns:

New York Yankees (3rd best team)

214 Homeruns

Winning %: .58 or 58%

Least Homeruns:

Chicago White Sox (Worst team in MLB)

133

Winning %: .253 or 25.3%

AI Questions

1. **What was the overall correlation coefficient between home runs and winning percentage, and how would you interpret its strength?**

2. **Were there any teams that hit a lot of home runs but had a low winning percentage, or vice versa? How do you explain those outliers?**

3. **Did your analysis suggest that home runs cause teams to win more games, or just that there's an association?**

4. **Did you look at data from multiple seasons or just one? How might including more seasons affect your results?**

5. **Did you include a trend line or regression model in your plots? What did it reveal about the nature of the relationship?**

Improvements:

1. Color coordinate each dot on the scatter plots so we can tell which dot represents which team

2. Include more years to get a better understanding of the correlation

3. Find a way to include other stats like batting average, on base percentage, slugging percentage etc.

References:

MLB stats, scores, History, & Records. Baseball. (2000). https://www.baseball-reference.com/

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

MLB.com | the official site of Major League Baseball. The Official Site Of Major League Baseball. (1995). https://www.mlb.com/

Chatgpt. (2022, November 30). https://chatgpt.com/