

Draft:

https://public.tableau.com/profile/joseph7072#!/vizhome/BaseballData_v1/Story1?publish=yes

Final:

https://public.tableau.com/profile/joseph7072#!/vizhome/BaseballData_v2feedback/Story1

- **Summary:** in no more than 4 sentences, briefly introduce your data visualization and add any context that can help readers understand it

Fans idolize players with high batting averages (hits/at-bats) and those who can score home runs are always a crowd favorite. One would imagine these players to be the biggest in order to generate enough power to consistently get home runs or even hits against faster pitchers. The story shows the distribution of pro players, their stats, and their corresponding weight/height. Each visualization shows a piece of the surprising story of most top players (by Avg./HR count) don't get their advantage from their size (height/weight) and are actually in the middle of the pack for these factors.

- **Design:** explain any design choices you made including changes to the visualization after collecting feedback

Initially, the story came about from just exploring the data and asking questions that could be analyzed from what was available. After exploring a bit, I decided to focus on the relationship between a player's stature and performance.

To start with, I thought it would be good to give the viewer a bit of context with distributions of the data for HR, Avg., height, weight, and BMI. I initially split these up by handedness since there were a surprising amount of left handed and ambidextrous players, then later decided this detracted from the central story. Rather than just providing the box and whiskers, I also added circles for each observation to better view the distributions and outliers. I felt the box and whisker plot the best for a general, high-level introduction to the data rather than using a scatterplot or histogram, which would present more information and more to distract the viewer. In addition, I wanted to point out the extremes at both ends of the spectrum of HR count and added in the two bar graphs. I included titles for easy context.

Next, I wanted to explore and understand the relationship of stature specifically on home runs as this intuitively seemed to be what would be impacted most. I created a scatterplot/bubble chart with height, weight, a height/weight ratio, and BMI by HR count. I double encoded color and size with HR count for emphasis and included the color range in the legend. I modified the color pallet in response to feedback, which also gave the added benefit of more clearly pointing out the outliers. In addition, I modified the axes to better focus on the bulk of the data. I also added a trend line to

more clearly identify any relationships. I included a filter to allow the viewer some context into which observations were shown as well as to allow them to modify and explore.

Next, I wanted to do the same exploration with Avg. and reconstructed the chart and replaced HR with Avg. on the y-axis. This time I decided to keep the size encoded with HR count and change the color to show the Avg. so that both factors could be viewed in the same chart, especially so since double encoding the Avg. did little to better understand the data.

The final scatter plot focuses specifically on HR count, height, and weight in order to drive the point that most home run leaders didn't particularly stand out among other major league players in terms of size. Again, I decided to double encode HR for emphasis. Here, it also had the effect of reducing some of the clutter in dense areas. It's simple and allows the viewer to focus on the mass in the center, with high HR counts.

- **Feedback:** include all feedback you received from others on your visualization from the first sketch to the final visualization

Feedback:

The handedness doesn't really contribute much to the central point.

If you filtered some of the graphs, you should probably show that somewhere. What about the players that scored less than 151 HRs?

Is there a way to tell who each player is?

I'm surprised the top hitters aren't much bigger than normal.

There's a lot of blue.

- **Resources:** list any sources you consulted to create your visualization

<https://www.bcbst.com/providers/MPMTools/BMICALculator.shtm>

<https://www.baseball-reference.com/players/b/brownmi02.shtml>

https://www.tableau.com/learn/tutorials/on-demand/additional-filtering-topics?product=tableau_desktop&version=tableau_desktop_2018.2&topic=visual_analytics