

Tableau Visualization WriteUp

Summary

This visualization is exploring data about baseball players, including height/weight/handedness and seeing what impact, if any they have on baseball statistics. Baseball statistics of interest are # of homeruns and the batting average. Visualization of the data helped establish that being left handed, about 5'10" and having a weight about 170 pounds makes you more likely to a baseball legend.

Design

Initial Design

One of the challenges with this data set is that there is no indication of timeframe – is this lifetime or the stats for one season? Since there is only one height/weight value per player, we will work with the assumption that it is from a single season. I decided to focus on exploring which characteristics had a correlation with better stats – HR and Batting Average. In other words, what is the sweet spot to be a baseball legend?

With over 1000 items in the data set, initial visualization of factors impacting the stats were quite busy. I decided to create categories for Home Runs and Batting Average. To do this, I referenced Wikipedia and the MLB website. Historically a batting average > 300 (0.3 in the data set) is considered a sign of greatness and <200 a sign of a bad player. The additional categories between the 2 were added using the halfway point as a divider. Home Run categories were not as easily determined. Again, using Wikipedia, the current career homerun leader is Barry Bonds with 762. Babe Ruth (3rd place) has 714. The top 50 players had no fewer than 427 home runs so anything above 425 was marked as legendary. The player in 100th place had a HR count of 342 so that was used to differentiate between Excellent and Average players. Finally, the 300th player (last listed) had a HR count of 221 so anything below that was determined to be a bad player.

Looking at the factors affecting the player' status, Height and Weight are both continuous numeric values so plotting them on a scatter plot to show the relationship between the two made perfect sense. Handedness however, is a categorical variable and thus a bit more of a challenge. In the end, I decided to reverse the concept used for height and weight. The two continuous variables in this look were home runs and batting average so they became our x and y axis with the categorical variable, handedness providing color. In order to show the concept of additional player categories (Bad, Average, Excellent and Legendary), reference lines for the average home run score and batting average were added. Finally, I added red lines to show the starting point of legendary status. These lines served as visual reminders of the additional categories.

[Initial Visualization](#)

Updated Design

After reviewing the feedback received, I updated the design to be clearer to a reader who wasn't familiar with the raw data used by adding legends, labels and additional text boxes. I updated the colors and shapes on the scatter plot slides to be consistent. Updated feedback indicated that this helped make the intent of the visualization clearer.

[Updated Visualization](#)

Feedback

Shared the story on Tableau public and with some colleagues. Feedback on initial design was as follows:

1. Move the textbox ("Batting average and home runs...") so it doesn't hide the data table. Can it be fixed at the top somewhere?

2. Are your greatness categories ("Bad"... "Legendary") defined somewhere? Can you provide a caption on the second and third vizes to explain them?
3. What do the different shapes mean on the second viz? Categories, just like the colors? I'd drop shape and experiment with using size to show batting average.
4. What does size mean on the third viz? Needs a legend.
5. Use the same display options on the second and third vizes (color and shape or color and size).
6. On the fourth viz, can the grey lines be labeled like the red lines are?
7. Last viz needs a caption, or an opening statement at the top. What am I looking at?
8. Handedness caption on the last viz needs to move so it isn't on top of the data table. Can you use icons to show handedness instead of color? Or color the background instead of the text?

Resources

Wikipedia

https://en.wikipedia.org/wiki/List_of_Major_League_Baseball_career_home_run_leaders

[https://en.wikipedia.org/wiki/Batting_average_\(baseball\)](https://en.wikipedia.org/wiki/Batting_average_(baseball))

MLB

<https://www.mlb.com/glossary/standard-stats/batting-average>

<https://www.mlb.com/glossary/standard-stats/home-run>