

# Data Story Telling

By Ryan Reardon

---

Hitting Performance Based on BMI and Batting Stance

## **Introduction**

I used the baseball dataset for the data story telling visualization project. The dataset is limited terms of hitting statistics, just homeruns and batting average. I focused on the height and weight (BMI) of players and their batting stance.

## **Summary**

I found the majority players tend to be in the high normal category to low overweight categories (.23-.26) Overweight players do hit slightly more homeruns on average but it's not a huge difference and batting averages are consistent across BMI classifications.

Lefties do tend perform slightly better; they hit more homeruns on average and their batting average is .20 pts higher than righties or switch hitters.

Your best hitters on average are lefties with a BMI between 23 and 26.

## **Design**

I added a BMI category and BMI classification to the original dataset; in addition, I deleted the players without hitting statistics. It made no sense to look at pitchers from the American League after the designated hitter rule.

Since, we were limited to two hitting categories the only way to do any analysis is to focus height and weight of players and their batting stance.

I initially used the histograms to look at the distribution of homeruns and batting average and plotted them in a scatterplot to show their relationship.

Then I looked at the height and weight relationship to homeruns and batting average using histograms, line charts, scatter plots, box plots, and bar chart. The histograms, I plotted homeruns and batting average against BMI. Homeruns and batting average are also plotted against height, weight, and BMI bins. Then, I used the scatter plots to show the relationship with BMI, bar charts for counts, and box plot to show their distributions.

For batting stance, I used the bubble chart to show the average, box plot for distributions and bar chart for counts.

I used a combination of dashboards with a filter feature, to filter data as the user wishes and color-blind palette to distinguish categorical variables but make it easier for those have trouble identifying reds, greens, etc.

## **Feedback**

I initially posted my first visual on the Udacity forum but didn't receive feedback, so I asked my wife for feedback. I used the reviewer's comments as additional feedback, I did post on forum, student hub etc. and waited for

**Initial Feedback:** She suggested titles separating the graphs on the dashboards and the ending slide to close out the presentation story.

**Reviewers Feedback:** Reviewer suggested I update the first couple graphs with a better description on the row and add a line chart to help the viewer visual the differences in height, weight, and BMI etc. Changes were made in the third visual.

First Visual: [https://public.tableau.com/profile/ryan.reardon#!/vizhome/Baseball\\_Data\\_Viz1/Story1](https://public.tableau.com/profile/ryan.reardon#!/vizhome/Baseball_Data_Viz1/Story1)

Second Visual: [https://public.tableau.com/profile/ryan.reardon#!/vizhome/Baseball\\_Data\\_Viz2/Story2?publish=yes](https://public.tableau.com/profile/ryan.reardon#!/vizhome/Baseball_Data_Viz2/Story2?publish=yes)

Third Visual: [https://public.tableau.com/profile/ryan.reardon#!/vizhome/Baseball\\_Data\\_Viz3/Story2?publish=yes](https://public.tableau.com/profile/ryan.reardon#!/vizhome/Baseball_Data_Viz3/Story2?publish=yes)