

Revised Project Notes

Project Name: Height-Weight Relation to Baseball Batting Average and HR

Date: 4/22/2018

My revised Tableau baseball report has included suggestion provided by the reviewer. My response to the reviewer are noted in *purple*.

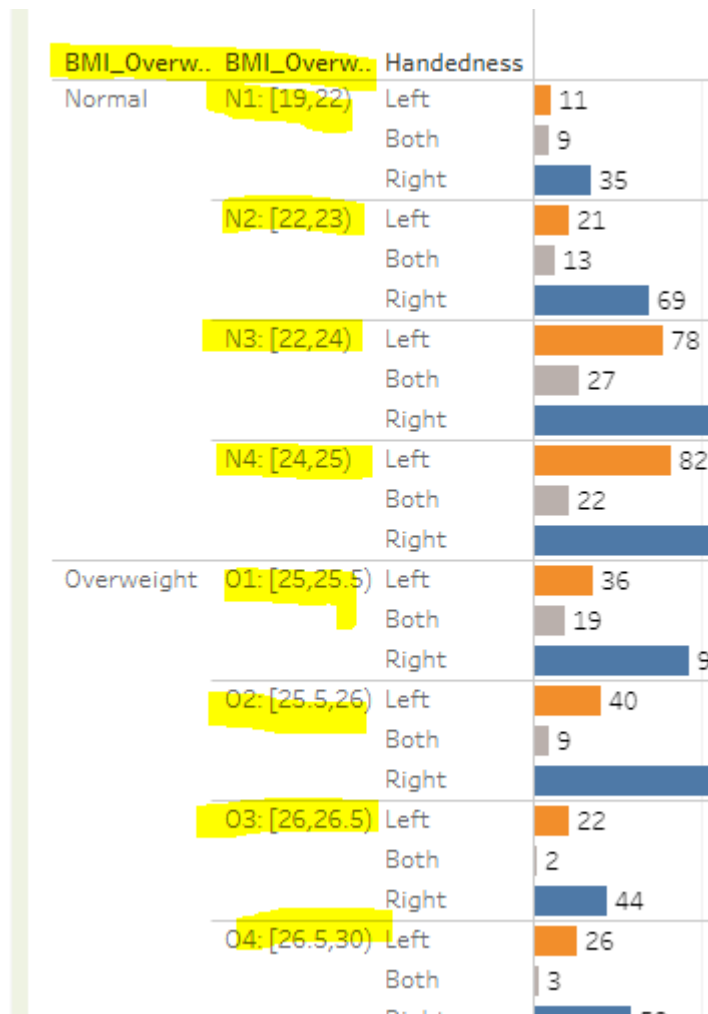
Visualization is Explanatory

The selected finding is clearly communicated. Design choices foster communication between the reader and the visualization.

You've made some excellent design decisions that make your presentation both more interpretable and more user-friendly:

Here are a few features that need improvement:

- Always be sure and include units of measure somewhere on the plot axes (e.g. "Height (in.)" or "Weight (lbs.)").
 - *Axes' units have been included.*
- There is as you pointed out a good deal of variation in HR and batting average that isn't captured by weight and height. There are trends there, though. Not strong ones, but still I think it's work including trend lines in your scatter plots.
 - *I've included the trend lines for meaningful comparison's scatter plots. e.g. height-weight trend was not included.*
 - *I would show the R-squared and regression results if the model was significant but none of them were.*
- The BMI stratification slide is difficult to make sense of primarily because the labels for the first two columns of row labels are truncated. Also, unless you've got some knowledge of BMI and its stratification, it's difficult to make sense of the the values marked in yellow:
 - *I've incorporated these suggestions by:*
 - *Renaming the BMI category and subcategories;*
 - *Using a less technical subcategory naming convention; and*
 - *Including reference on BMI by Centers for Disease Control and Prevention (CDC).*



I would recommend perhaps renaming these stratifications to something that might be more immediately interpretable to a general audience. As you pointed out in your write-up, a high BMI does not in the world of professional sports mean a player is obese. Barry Bonds likely had a very high BMI when he was hitting lots of home runs, but you also made reference to the reason for this in your write-up.

- *Orientation was switched to better point out a trend for a specific subgroup;*
- *Enriched the charts with explanations or call outs were included to better convey the story.*

- I don't know a whole lot about baseball, but I think 0.225 - 0.270 is a wide range for batting average. It would be interesting to know the full range for all MLB players but I bet it doesn't drop much beyond 0.225 and that there aren't too many players above .300. I would recommend that you take this into account in your analysis, especially when looking for relationships between batting average and height (I suspect shorter players are better batters because they are faster).

- *I didn't find any significant relationship between height or weight with Avg or HR, which is why I didn't include them in the story. Yes, shorter may tend to be faster (note: Usain bolt is not short) but shorter may be negatively correlated with strength.*
- *Simple scatter plots analyses, which included plus (linear and non-linear) regression results did not show significant relationships when height and weight were stacked against Avg and HR. The R-squared values for different regression models were less than 5%.*

Design

A reader's summary of the graphic would closely match the written summary in the writeup, or a reader would identify at least one main point or relationship that the graphic attempts to convey.

I found one trend for a specific subgroup (right handed for HR vs. BMI). However, overall, there was no clear trend. I've sliced, diced, and transformed the data in meaningful manners. In the Tableau's 'Conclusion' part, I've noted data weaknesses and possible reasons on why meaningful relationship was not found from the data and its visualization:

Problems with Height and Weight

Height and weight may theoretically affect performance. However, this is not supported by the data and analyses. At the professional level, there are arguably other important variables such as discipline, college experience, eye-hand coordination level, and years in the league.

Height and weight are problematic because HR is career numbers while height and weight would be at one point of their career or their average weights. Although, one may argue that height and weight should not fluctuate greatly because these are before-steroids era (Aw, Snap!!!).

BMI also may not be the best metric for healthy or fit measurement because additional weights for these athletes could be muscle-based weights and not those of fat.

There isn't a clear narrative to accompany and explain the trends illustrated in or missing from your charts. Your second slide is nicely annotated. You've decided against using the caption boxes to explain your findings in words, so the reader doesn't have the advantage of your commentary in guiding them through what are subtle relationships. Please add further commentary or narrative to explain your findings. I realize that you heeded feedback that said the captions were too verbose, but there are other options available, too.

I've tightened up the narrative with story writing, better orientation, explanations or call outs, and other changes mentioned above.