Create a Story: Baseball Statistics V2 Write-Up

Allison Senden

**Public Tableau Link V2:**

<https://public.tableau.com/profile/allison.senden#!/vizhome/baseball_stats_final_0/BaseballStatsFinal>

**Summary:**

The dataset analyzed for my final project was statistics collected on 1, 157 baseball players. Handedness, batting average, height, and weight are some attributes that were collected during the gathering phase. Let's dig into some of the findings I came across during my analysis.

**Details V2:**

1. First, I dedicated a dashboard to telling my viewers exactly what they are going to expect in this summary of my analysis. Why the data exists and what it contains. So this slide shows a little synopsis of the type of data that’s going to be presented.
2. The second dashboard shows my first glimpse into the data. I wanted to show the distribution of the batting average performance, since after all, that’s what we were going to be ultimately analyzing players for. I made a chart of batting average versus the number of home runs per that batting average. For this chart, I made a calculated variable where the batting average was in bins so that the chart would be more visually appealing and easier to read.
3. That gave me the idea to see how the number of homeruns a player got should be displayed. So next, I show that distribution. Again, here I created a calculated variable where the number of homeruns were in bins so it was more visually appealing and readable. The y-axis had the count of the number of players that achieved that number of homeruns.
4. Now that I had a good idea of how the performance of players was generally distributed, I wanted to see how other features might affect the overall performance of baseball players. I decided to make a dashboard of how height and weight affected both the batting average and number of home runs players got.
5. I found that there was a relation between performance and height, but not weight. I took the same charts I made earlier for tracking height against performance and added in the 1 other variable I haven’t analyzed yet, handedness. Here I was able to decipher whether or not handedness actually played a major role in the player’s performance. I found that there wasn’t enough of a trend to continue with the analysis. It’s not a big enough sample size to decide at this point in time.
6. Not only did I want to check whether handedness impacted the average number of homeruns a player got, but what their batting average was. So, this is the next slide I chose to reflect. For each handedness, wanted to see how the distribution plotted out against batting average. This led me to the conclusion that handedness did not really have an effect on the batting average of a baseball player. Every hand mapped out to about the same batting average.
7. My summary slide came next. I had included all the features that the dataset had to offer in my analysis. This slide, I tried to choose the slides that best reflected the overall findings from the dataset. Meaning, I wanted to portray the ultimate conclusion was that height was the main driver in overall performance.

**Feedback for V1(Version 1):**

This is the feedback I received from my roommate.

* What do you notice in the visualization?  
  Why does the final slide not include the height to # of homerun graphic since the conclusion is around that trait having the largest correlation? Also, the top boxes could be a little larger so you could view more of the text without scrolling.
* What questions do you have about the data?  
  In the first graph – why does the batting average of 0.0 have the highest number of homeruns? Doesn’t seem right that player with no batting average could have hit homeruns only?
* What relationships do you notice?  
  In the height analysis it is easy to notice that the two graphs follow a very similar pattern just with higher spikes in the homerun graphic
* What do you think is the main takeaway from this visualization?  
  Ability to identify characteristics or traits of baseball players that have the highest number of homeruns. Based on the data presented it would seem that height has the largest impact to number of homeruns.
* Is there something you don’t understand in the graphic?  
  Why did the graphics not use a sample size when comparing the handedness of batters? Since the number of right-handed players is much higher than lefties and even higher compared to switch hitters the graph is a little bit confusing visually. While the distribution is fairly easy to see it might be more visually appealing to use the same number of records for each type.

**Feedback:**

Feedback from another individual was short, but I will include a summary of what she had to say here:

I’m in advertising so I’m used to telling a story with the data. I followed your findings very easily through the link. The only thing I got caught up on sometimes is what you were trying to portray exactly because there were labels, but they didn’t specify the dimensions you were talking about. If you change that, I think you’re good to go. I also liked how you ended with specifying the exact takeaway from this analysis.

**Resources:**

<https://onlinehelp.tableau.com/current/pro/desktop/en-us/calculations_bins.html>

<https://onlinehelp.tableau.com/v10.5/pro/desktop/en-us/help.htm#extracting_data.html>

<https://interworks.com/blog/lkrennwallner/2015/03/27/interacting-your-trend-line-tableau-part-1/>