Looking for a Hitter

Arturo Guerrero

Udacity Data Analysis NanoDegree

Summary

This analysis is focused on how 1.157 baseball players perform when batting.

I will focus on their handedness, their height, their weights, the batting average and how many homeruns they have achieved. The purpose of this analysis is to find out the features of a good baseball hitter.

Design

The final report is a story design on Tableau, but I have created several worksheets and dashboards to create the story “Looking for a Hitter”.

“Looking for a Hitter” starts with a description of the players profile. The analysis in this first step shows that the average player is: right handed, weights approximately 180 pounds and is 72 inches tall. This section includes three bar-charts for displaying the distribution of the player concerning their handedness, their weight and their height for a categorical comparison.

The next step is to have a look on the players scores. Most player have a batting average of 24%. This part has two bar-charts grouped in categories (batting average percentage and number of homeruns) were the number of players are displayed.

Then we look on the specific features of the players: handedness, weight and height and the impact of these features when batting. In a first insight we see that left and both handed players are better batters than right handed and left handed are also homerunners, while right and both handed are not so good hitting homeruns. But if we have a deeper insight on this subject, we will find out that the better a player perform while batting the greater the possibilities to achieve a homerun. And in this case the better the average of batting the better the right-handed players perform in comparison to left and both handed. This part of the story has two types of visualizations: two bar-chart for a categorical approach and a scatterplot to see if there is a correlation between two variables: number of homeruns and batting average percentage and their relation to the handedness included with different colours and different trend lines. I included to simple interaction elements for a better visualization of the handedness and the batting average.

If we look to the weight, we will see a double effect on batting. The average while batting is better on thinner players than on heavier players, but heavier players scores higher number of homeruns. This section is based in bar-charts for have a look to the distribution of the weight between different categories.

In one side we will find out, that smaller players are better hitters than taller players, if we have a look on the height of the players, but on the other side we will see that the height doesn´t have a great impact for hitting homeruns. This last section is also made up with to bar-chart for analysing the height distribution of the player grouped by heights.

Conclusion

It could be possible to change the approach and add other type of features to have a different perspective in this analysis (speed ball throwing, type of ball throwing, handness of the thrower that should have also effects on the batting response), but attending to this dataset it can be conclude that:

If you are looking for a hitter it is recommended to look for a right handed player, not very tall (66 inches - 168cm) and not very heavy (140 pounds - 64kg).

If you are looking for a homerunner you should look for heavier player, while the height is not so important. In any case the player should have demonstrate a good batting average over 22%.

Feedback

I´ve shared my project on the Udacity Student Hub. I received some recommendation and suggestions from the mentors, that I added to a second version of this project.

I´ve uploaded both versions in Tableau Public:

First version:

https://public.tableau.com/profile/arturo3591#!/vizhome/Lookingforahitter/Lookingforahitter?publish=yes

The second and reviewed version:

https://public.tableau.com/profile/arturo3591#!/vizhome/Lookingforahitterv2/Lookingforahitter?publish=yes

The first change they suggested me was to review my text boxes, because they were overlapping the plots. I change the width and position of the overlapping plots and check it out on Tableau Public in 100% zoom on the screen.

The second subject they mentioned was to change labels of the axes with suitable title to all the plots. In this case I changed the titles and labels. Many were with an automatic Spanish description, due to my Tableau version. I changed it manually to English looking for proper descriptions.

And the third issue I should address was to review the caption areas to make it clearer. I changed each caption area to have better description.

Resources

Baseball Data

https://s3.amazonaws.com/udacity-hosted-downloads/ud507/baseball_data.csv

Tableau 2018.3