Project: Create a Tableau Story

Student: Sherif Shawkat

Baseball players analysis summary report

Introduction:

This report is part of the requirements needed to complete the project “Create a Tableau Story” for Udacity Data Analyst nanodegree. The project aims at analyzing a dataset of our choice using “Tableau”. The dataset I chose is “Baseball players” dataset.

Stories URLs:

V1:

https://public.tableau.com/profile/sherif.shawkat#!/vizhome/Baseballanalysis-SherifShawkat-V1/Story1

Final version:

https://public.tableau.com/profile/sherif.shawkat#!/vizhome/Baseballanalysis-SherifShawkat-V2/Baseballplayersanalysissummarystory

Summary:

- Looking at the Handedness distribution, we will find that over 63% of the players in our dataset are right-handed, over 27% are left-handed and less than 10% use both hands.
- Left-handed and both-handed players have almost similar Batting average of 0.2, while right-handed players have a less Batting average of 0.17.
- Left-handed players are the highest averaging Home Runs players with an average of 56, followed by right-handed players with an average of 42, while the lowest averaging Home Runs players are the both-handed players with an average of 32.
- There’s a weak positive correlation between Weight and Average Home Runs.
- Players who have height = 67 inches or weight = 209 lbs. are averaging the highest Home Runs.
- There is no clear relationship between Home Runs and Batting average.

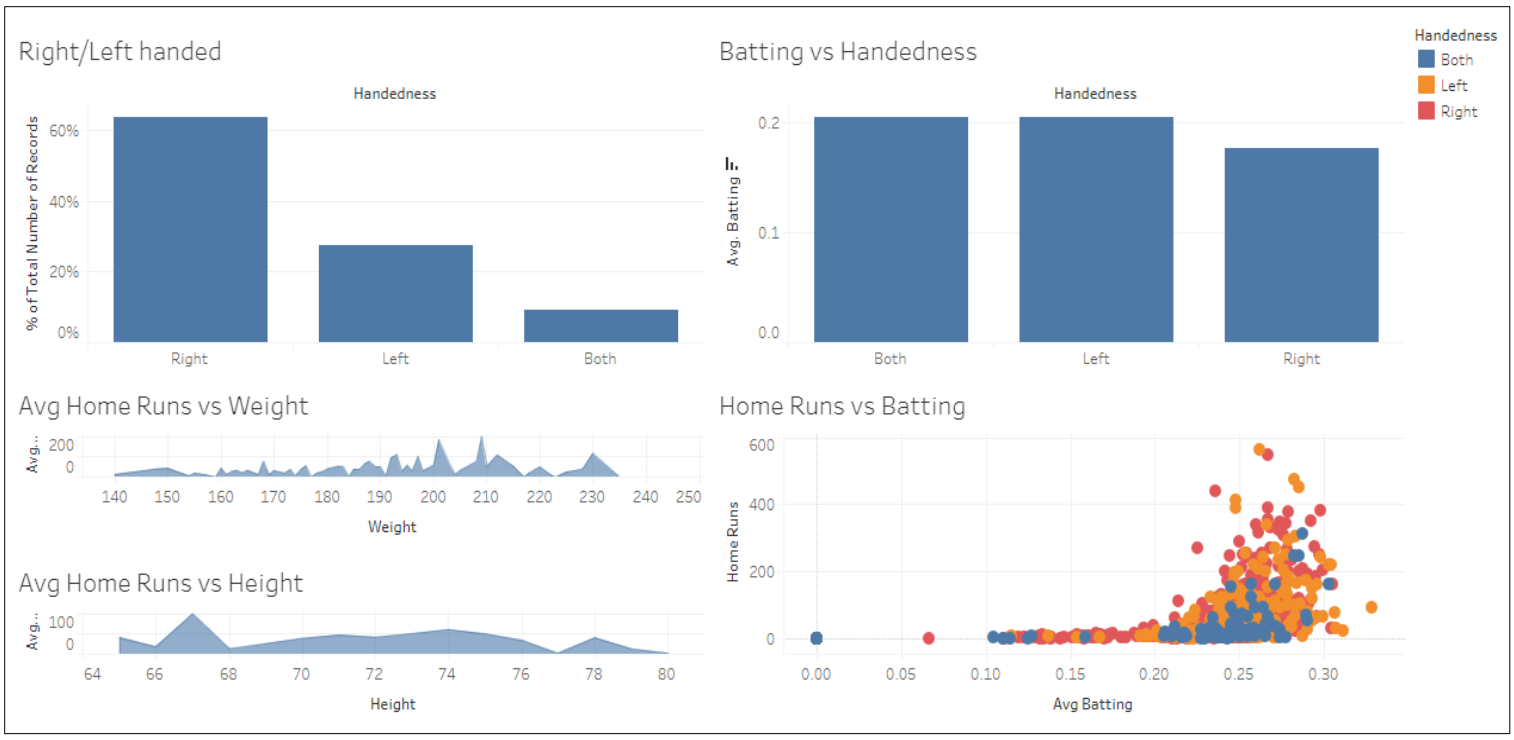

Fig 1: Snap of one of the Baseball players analysis summary story

Design:

- For the Handedness count statistics, I used a bar chart which best describes the distribution per type. I also converted the values to be % of total records so that I can know the portion of each category in the whole dataset. Finally, I sorted the bars in a descending order.
- For Handedness vs Batting and Home Runs (each one independently), I used also the bar chart and sorted it in a descending order.
- For “Avg Home Runs vs Weight”, “Avg Home Runs vs Height”, “Avg Batting vs Weight” and “Avg Batting vs Height”, I used a continuous area chart with weight or height as a dimension on the x-axis, Avg Home Runs or Avg Batting as an aggregated value on the y-axis.
- For the Batting vs Home Runs plot, I used a scatter plot to plot the home runs and batting for each player independently, I used the shape as a filled circle, then I dragged the Handedness column to the Marks-Colour to differentiate between handedness in the plot.
- All charts titles have been modified to describe what the chart means, as well as the story title.

Feedback:

I posted the V1 of the story on Udacity student hub, and received a feedback from Amelia P. (one of Udacity mentors). The feedback included 3 modifications:

1- **Guarantee all charts and the story have an informative title:** all charts titles and the story titles have been modified
2- **Adjust axes titles:** all chart axes have been modified to explain what the axis means.

3- **Change the acronyms in the handedness column:** an alias has been created on the Handedness column to replace “L” with “Left”, “R” with “Right” and “B” with “Both”.

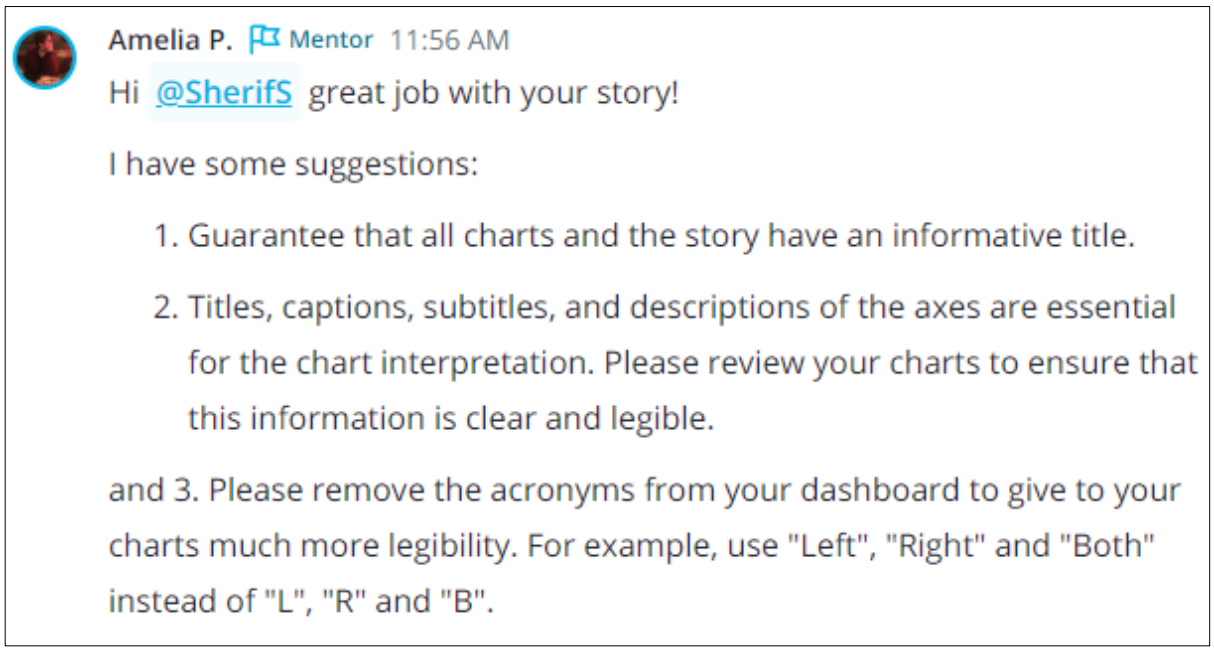

Figure 2: Snap of the feedback received from one of Udacity mentors on the student hub