Laila Hussain Alqawain

Data Analyst Nanodegree

Project: Create a Tableau Story

Baseball Data Story

- **Introduction**

  To create a tableau story, I choose Baseball. It contains 1,157 observations about baseball players and 5 variables which are players' names, handedness (right (R), left (L), and both (B)), height (in inches), weight (in pounds), batting average, and home runs. After cleaning data. The dataset contains 871 observations and the same variables in addition height/weight ratio. I calculated by height divided on weight.

  1st Story

  https://public.tableau.com/profile/laila7356#!/vizhome/Story1_282/Story1?publish=yes

  2nd Story

  https://public.tableau.com/profile/laila7356#!/vizhome/Story2_54/ThePerformanceofBaseballPlayers?publish=yes

- **Summary**

  In this project, I made different relationships between the variables like, home run vs. height/weight ratio, batting average vs. height/weight ratio, home run vs. batting average, and batting average/home run/number of records vs. handedness. To explanatory the performance of baseball players.

- **Design**

1st version

I made 4 plots to analyze the data.

- Home Run vs. Height/Weight Ratio, I made scatter plot and filter handedness by color, to explanatory the relationship.
- Batting Average vs. Handedness, I made a box plot to explanatory the distribution.
- Batting Average vs. Height/Weight Ratio, I made scatter plot and filter handedness by color, to explanatory the relationship.
- Batting Average/Home Run/Number of Records vs. Handedness, I made a bar plot to explanatory the distribution.

2nd version

I added Home Run vs. Batting Average plot instead of Batting Average vs. Handedness plot. For All plots, I changed features of variables to viewing clear insight, and filtered by handedness.

- Home Run vs. Batting Average, I made scatter plot and changed features of variables to viewing clear insight. I found a weak positive relationship between home run and batting average. The highest peak is on 563 in home run and 0.262 in batting average.
- Home Run vs. Height/Weight Ratio, I made scatter plot and I found there is no relationship between home run and height/weight ratio. The ratio with the highest home run (563) is 0.369 and the home run with the highest ratio (0.492) is 13.
- Batting Average/Home Run/Number of Records vs. Handedness, I made a scatter plot instead of bar chart. I found the number of players who use the right hand greater than who use the left hand and both. But the players who use left-hand greater than other in batting average and home run.
- Batting Average vs. Height/Weight Ratio, I made scatter plot, I found there is a weak relationship between batting average and height/weight ratio. The ratio with the highest batting average (0.328) is 0.423. The batting average with the highest ratio (0.492) is 0.245.

- **Feedback**
  - Change the features of variables to viewing data more clearly.
  - Remove Batting Average vs. Handedness plot to avoid riddance.
  - Filter by handedness without unacceptable color.
  - Rename the story name to make the reader understand what your story about.
  - Change Batting Average/Home Run/Number of Records vs. Handedness plot to scatter to see the distribution more clearly.

- **Conclusion**

  Even though the number of players who use the right hand is greater, the highest performance was for the players who use the left hand.
  In my opinion, this conclusion maybe change due to there are no relationships between some variables.

- **Resources**
  - https://kb.tableau.com/articles/HowTo/dual-axis-bar-chart-multiple-measures
  - https://www.tableau.com/about/blog/2017/2/top-10-tableau-table-calculations-65417
  - https://techgraphs.fangraphs.com/how-to-use-tableau-for-baseball-data/