UDACITY

Data Analyst Nanodegree

Project 10

Create a Tableau Story

Name: Ibrahim Riyad Aldekhiyl

Introduction:

Story link 1:
https://public.tableau.com/profile/ibrahim1170#!/vizhome/Baseballdatastory_0/Baseballdatastory?publish=yes

Story link 2:

https://public.tableau.com/profile/ibrahim1170#!/vizhome/Baseballdatastory_0/Baseballdatastory2?publish=yes

This data visualization created using Tableau, the visualization tells a story or highlights trends or patterns in the Basketball data set. The Baseball data set containing 1,157 baseball players including their handedness (right or left handed), height (in inches), weight (in pounds), batting average, and home runs.

Summary:

This data visualization or story is been created using different visualization methods to give insights on the data set that I'm working on. Some of the visuals are comparing between two variables to know the effect of each among other (Handedness, homeruns, height, ..), you will see this and even more in the data story.

Design:

Initially I was thinking of use handedness as a categorization variables because I thought it's the best fit to be my categorization variables, but when I got to know that the right handed players are representing around 64% of overall players I stopped thinking to link the handedness to all my thoughts.

I tried to know the relationship between Weight, Height and Handedness with the Batting average and Homeruns and see if there any impact of the Height in the Batting average or Homeruns and so on, after analyzing the visuals I got to know that the players with Weight between 180-189 pounds and height around 73 inches are most likely to have higher Home runs than the others, and the players with weight between 180-189 inches and height around 72.5 pounds are most likely to perform better than other players. The players who play in both right and left hands are most likely to have high Batting average than the others by a little bit different with left handed players.

I used scatterplot, bar chart, packed circle chart, line chart to visualize the data set. Based on some feedback I've change the type of some charts (e.g. box plot).

I've added two calculated field for Height to change it from inches to meter and Weight to change it from pounds to kg. There are some errors or duplicate records that cause some outliers in the analysis.

I used bar chart because it’s one of the best plots when you are comparing multiple variables. I used scatter plot, because it’s one of the best when measuring the relationship between Few (2-4) variables, in the scatter plot I used a various color (Red, Green and yellow) to differentiate between the variables. Circle chart is good to summarize a set of categorical data or displaying the different values of a given variable.

Feedback:

One of the feedbacks is about the abbreviation on the legends and the labels, I need to make sure to use full words like "Right" instead of "R" all over the story. Also, I've change some of sheets and dashboards title based on the feedback, because it's not explanatory titles, I did the same for the Story name. Colors in data visualization is very important, because colors just make the shapes visible, sometimes they encode data or categories themselves.