

Data Visualization

by Bidish Basu

Introduction

In this project I will be exploring and tabulating my findings of the baseball players dataset. I will be using Tableau to communicate my findings and relationships that I have observed in the data.

First Visualization

https://public.tableau.com/profile/bidish.basu#!/vizhome/First_Visualization/Story1

Final Visualization

https://public.tableau.com/profile/bidish.basu#!/vizhome/Final_Visualisation_0/Story1

Summary

I have used many different techniques to represent the Baseball data. I have explored the average, Height and Weight variables for tabulating my findings. I have also created a new variable (Height/Weight) to help me with my findings. In my initial visualization, I found out some stats based on the original variables but in my final visualization, I record my findings on the basis of the variable that I derived (Height/Weight).

From my exploration, I have concluded that the players with a height and weight ratio of 0.4929 have the highest Batting Average. Also, most of the players have a batting average between 0.21 and 0.30. Handedness was not conclusive as there were not equal numbers of players who were left, right and played with both hands.

Design

Initial Design

Initially the chart type and format settings for all the axis and plots were set to default.

Each plot contained gridlines and labels.

X and Y axis of all plots were shown with no color encoding being done as it wasn't felt needed.

I have also used a scatterplot as that helps us to see the scattering of the Batting Averages of the players.

Final Design

I have now removed the grid lines except for the scatterplot as they were needed for aesthetic purposes.

There is a "Calculated Field" introduced i.e. the "Height/Weight" as they allowed me to analyze the data in depth.

I added a filter so as to remove most of the empty data i.e., some records had 0 Homeruns so I removed those rows as it was not helpful for my analysis.

I have used colours for enhancing my analysis.

In my initial design, the sheet titles were not seen in the final visualization. I have made sure to include them.

I have also added a few actions in some of the dashboards.

I chose to use bar charts and scatterplots as my tools to visualise as they provide a great overview of the distribution between the different types of variables I have used.

Feedback

Following are some of the feedbacks which I received based on which I made my final visualizations.

1. The stories should have a title so as to make it clear what the visualization is for.
2. The labelling of the X and the Y axes should be clearer (Avg and HR for example were not clear).

3. Description of the visualization are not clear.
4. A ratio of Height and Weight will be easier for analysis.
5. Include titles for all the pages.

Resources

1. <https://onlinehelp.tableau.com>
2. <https://www.tableau.com/learn/training>
3. <https://community.tableau.com>
4. <https://www.youtube.com>