Raymond Atherley
Project 4: Create a Tableau Story

SUMMARY

Baseball, known as American's pastime, has been played since the 1840's. A season usually has about 162 games. Two common measures of player performance is the batting average and the number of home runs. We were given a dataset of 1157 players consisting of their performance scores, height, and weight. We will determine the physical characteristics of the top performers.

DESIGN

The initial draft (of worksheets) included a Top 20 home run scorers horizontal chart and scatterplot comparing top 20 to the other players. In addition, a set for heights greater than 73 inches and equal to and less than 73 inches was created to use as a filter. These worksheet were removed to focus on just the top 10 players and the height filter was not used for the analysis.

The initial story comprises five slides in total. The first slide is an overview of player statistics. It consists of charts showing the median player height and weight, and the median batting average and amount of home runs. It has showed the total amount of home runs, total amount of players as well as their handedness (right-hand, left-hand, or both).
The first slide showed mostly bar charts as the data of handedness was categorical. The information of median height, weight, number of home runs, and batting average also showed multiples of a simplified single bar to easily display the information. The sum of home runs and number of players were shown as single point as they were both quantitative variables.

The second slide relayed statistics by height using a mixture of three multiple vertical charts (two bar charts and a histogram) with height on the horizontal axis. Height ranged from 65 inches to 80 inches. The top chart displayed the percentage of a specific height, and the middle chart showed the percentage of total home runs for players that height. The bottom chart showed the median average for that specific height. By doing it way with the three stacked charts, we can easily see that the taller players tend to have a much greater percentage of the homeruns. In addition, we see that the shorter players tend to have higher batting averages. A ranging color of blue in accordance with the number of players under each height was used to help the audience focus on the higher number areas.

The third slide showed the statistics for the top ten home run hitters with an dynamic slider that could change the number of top players. There was also an interactive option to choose the handedness of the players. The slide consists of a listing of the top ten players with a horizontal bar chart of their total homeruns, and a stacked chart of their total home runs as a percentage of the total home runs. Most notably, there was a table which showed the median height and weight of the top ten compared to the same of the other players.

The fourth slide was similar to third slide, but was relaying the batting average information. In addition, it did not have a stacked bar with a percentage of the top ten's batting average as I did not think it would translate well or make sense to do so.

The fifth slide presented two dual-axis line charts with the top ten players. The top one showed the number of home runs and batting average for each of the top ten home run hitters. The bottom one showed the batting average and number of home runs of the players with the top ten batting average. Line charts where used to the clearly any overlapping of the trends as the top home run hitters where different from the players with the highest batting average.

Design Changes

Slide 1
Changed question from "What physical characteristics result in better performances?" to "Do top players have certain physical characteristics?"
Removed Total Homeruns vs Total Number of Players Scatterplot summation graph.
Added labels to Number of Players vs Handedness plot, and Median Batting Average and Home Runs
Added a baseball picture.
Add question at top of slide, Do top players have certain physical characteristics?
Add intro text: Here's a snapshot of statistics for 1157 Major League players.
Updated/corrected axis labels.

Slide 2
Changed Stats by Height plot to a side by side bar chart for only Percentage of All Players and Percentage of Total Home Runs using orange and blue coloring. Now you can easily see which heights scored the greater percentage of homeruns.
Stack separate plot below of Median Batting Average vs Height
Removed color mark for Number of Players
Updated Aliases/Row labels for the legend

Slide 3
Updated Aliases/Row labels for Handedness filter, Top N filter and Text table.

Slide 4
Updated Aliases/Row labels for Handedness filter, Top N filter and Text table.

Slide 5
Updated Aliases/Row labels for the legend.
Updated axis labels to 'Batting Average'.

FEEDBACK

Slide 1

1. Why is this analysis being done?
2. The number of players were presented as the first visual in the slide. I wondered if this was total players altogether. However another slide showed the total players, but this might have made more sense initially if totals slide was first.
3. Showing totals may be redundant when looking at one side handedness. Possibly use a description vs a table? Possibly say three characteristics.
4. Add median batting average label to vertical axis.

Slide 2

1. 4. Make word home runs in slide 2.
2. % of players heights in slide heading
3. say height percentage
4. Do we need count of numbers
5. Consider inches on all slides

good point showing 74 or 16% score 23% of homeruns!

Slide 3

Labeling to better explain if possible.

Slide 4

Height on graph to be removed.

slide 5

Remove name of top of slide if possible

RESOURCES

Tableau.com
youtube.com
interworks.com

VERSIONS

Initial Draft of Worksheets:
https://public.tableau.com/profile/ray5513#!/vizhome/Udacity_Project4BaseballStory1pre/Top20_HR

Initial Story:
https://public.tableau.com/profile/ray5513#!/vizhome/Udacity_Project4BaseballStory1/Story1?publish=yes

Final Story:
https://public.tableau.com/profile/ray5513#!/vizhome/Udacity_Project4BaseballStory1a/Story1a?publish=yes