Baseball Story – What Affects Performance

Summary:

This is a dataset that was provided from Udacity to explore and tell a story using visualizations. The dataset had 1,157 baseball players, along with their handedness, batting average, homeruns, height and weight. I wanted to see what affects performance, some the questions asked where; is being left handed better, does it matter how tall or how much the player weighs, will you have more homeruns with a higher batter average?

Initial Version:

https://public.tableau.com/profile/andy.dusanowsky#!/vizhome/BaseballDataInitialVersion/Story

Final Version:

https://public.tableau.com/profile/andy.dusanowsky#!/vizhome/BaseballDataFinalVersion/Story?publish=yes

Design:

In the dataset, I created a new column as "Weight Groups". I created this due to the amount of weights that different players had, to better understand the data and tell a story. This is the formula: IF [Weight] <= 149 THEN 'UNDER 150' ELSEIF [Weight] <=174 THEN '150-174' ELSEIF [Weight] <=199 THEN '175-199' ELSEIF [Weight] <=224 THEN '200-224' ELSE '225+'END

The first tab in the story is the introduction, demonstrating the questions that were going to be asked during the analysis. This was to demonstrate to the reader what I am trying to do.

The second tab breaks the data out to a high level, for a quick, easy view of the data. A filter was added for handedness to see where each handedness fell with weight, height regarding home runs and batting average.

The third tab, I created a plot graph, color coded by handedness to show the relationship batting averages and homeruns. I created this to see if players that high batting average also have a good number of homeruns.

The fourth tab is a review tab that demonstrates the data that is found, trying what attributes in a player, height, weight, handedness, to see if it makes them a better hitting player or a power hitter for homeruns.

The fifth tab looks at the top 10 players for both batting average and homeruns. I created this to see if the previous data found is valid in these players.

The sixth tab is a conclusion tab that reviews all the data that was found and determines if the data that was found holds true.

Items Changed from Feedback:

First feedback received:

Introduction on 1st slide is squeezed out of frame

- Changed to not be squeezed

Slide 2, top portion is squeezed together where can't see all 3 handedness

- Changed to not reflect any squeezing

Line plot change to bar chart

- Changed line chart to bar chart

Missing a caption on your conclusion slide

- Added caption

<u>Second feedback received:</u>
Color scales on bar charts

- Removed color scales on bar charts to reflect only 1 color

Short cuts within captions

- Updated all HR's with the word homeruns

Giving users sliders to "play" with the data for slide 3.

- Added 2 sliders for weight and height

Overplotting on slide 3.

- Made the plot bigger to fill the whole screen and view data better.

Typo on slide 1 "reviewed"

- Corrected typo

Feedback:

Clint V: UDACITY Student: "Hey Andy, your introduction on 1st slide is squeezed out of frame. Wondering if it's the size of your MLB logo that's pushing it up. On slide 2, your top portion is squeezed together where I can't see all 3 handedness. Also, on slide 2, on my first submission yesterday the reviewer sent it back saying we shouldn't use line plots on non-time related data. I had to changed mine to bar charts. I know that's what defines line plots, but I still feel line plots visually represent the data better. You're missing a caption on your conclusion slide. Otherwise looks great!! good luck!"

Jan S: UDACITY Student: "as mentioned before, there are some layout issues ... maybe a screen resolution related problem. slide 2: I would not consider the color scales a very good choice ... have you tried w/o color? (except for handedness - that one needs color) ... try to avoid shortcuts within captions / labels (HR --> home runs / human resources ??) Also there are some issues with overplotting (slide 2 top-center and slide 3) ... try to minimize it? pick boxplots instead of scatter plots or use a log-scale for the Home-run-axis on slide 3slide 3: why not give the user some sliders to "play" with the data you're always looking at weight or height ... have you tried calculating some ratio (e.g. BMI)? typo on slide 1 "reviewed""

References:

Udacity Lesson Videos
https://onlinehelp.tableau.com/current/pro/desktop/en-us/functions_all_alphabetical.html