Create a Tableau Story: Baseball Data

Initial:
https://public.tableau.com/profile/r007#!/vizhome/Udacity-BaseballStoryInitial2/Story1

Final:
https://public.tableau.com/profile/r007#!/vizhome/Udacity-
BaseballStoryFinalBackup/UdacityBaseballAttributeAnalysis

Summary:
This analysis covers a 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. The goal of the visualization was to show which baseball player attributes (handedness, height, weight) were the best predictor of a higher batting and number of home runs.

Initial Design:
Since there were only three attributes to analyze against batting average and home runs, the design I chose was to show how each correlated to the data. As expected, Height and Weight were strongly correlated, but I chose to still show them separately.

The goal was to see how height, weight, and handiness impacted batting average and home runs, so I created my story showing each of these sequentially, followed by a conclusion.

For height and weight, given that there were 1,157 data points, I went with a scatter plot overlaid with a trend line to show correlation. The dashboard for each of these was done on the same page so that height and weight could be seen next to the each other.

Handiness is a categorical attribute, so I choose to show it in two ways. First, a pie chart showing the proportion of handiness (left, right, both) for the 1,157 data points. Second, I used a bar chart of the median value for batting average and home runs against handedness. Median was a better choice than average for this data to compensate for any extreme outlier data.

Final Design:
Since handiness seemed to have a big impact, I went back and overlaid this onto height and weight. In addition, colors and filters were added to allow for more user interaction and the pie chart was converted to a bar chart to show distribution.

Feedback:
The three points of feedback that I incorporated were 1) overlaying of handiness with the over attributes, 2) adding colors and filters to allow the user to better see what handiness data affected the batting average and home runs, 3) switch from pie chart to bar chart for showing distribution of handiness.

[From Review #1] Initial design decisions such as chart type, visual encodings, layout, legends, or hierarchy are included at the beginning of the Design section in the writeup.
You've described some of the analysis choices you made such as using the median instead of the mean. And you've also documented the addition of handedness as a dimension in your charts and exchanging the pie chart for a bar chart. Good job!

The idea behind this requirement is to develop the habit of considering how visual design choices can help us to communicate more effectively and make a data product more user-friendly. What were some other design choices that went into your project such as chart type, dashboard organization and layout, sequencing, color palettes, etc.? What were the advantages of these choices compared to options that you didn't choose? How did the design elements you settled on help you to communicate your findings more effectively and make your project more user-friendly? See if you can go into a bit more detail.

[From Review #1] The selected finding is clearly communicated. Design choices foster communication between the reader and the visualization.

- Charts are more user-friendly when they include titles. In my experience people almost instinctively look for titles in the upper margin of a chart to help figure out what a chart is plotting. Please add concise but descriptive titles to all of your charts. You may have noticed that Tableau deactivates the titles of charts that are added to a story point as a single worksheet. To get around this, try first placing the single chart onto a dashboard, then into the story point.
- Likewise, your story point needs a descriptive title instead of the default "Story 1".
- Abbreviations facilitate data analysis and storage but in this last stage of the data analysis workflow, when we report the results of the analysis, it's best practice to avoid the ambiguity and obscurity of abbreviations. Not all reader will be familiar with the abbreviations in this dataset, so it's best to change R, L, B, HR, and Avg. to their full-word counterparts. This is easier than you think to do in Tableau. Just right-click on the axis label, legend, or object where you want to edit an abbreviation and choose the related **Edit** option. A window will pop up where you can make the necessary modifications.
- Try grouping charts into dashboards to avoid large empty spaces on the story point slides.
- Sequence your charts logically, beginning with descriptive overview stats like the distribution of handedness before moving to charts that show more complex relationships.

Resources:
https://kb.tableau.com/articles/howto/creating-a-pie-chart-with-percent-of-total-of-variable-sized-bins
https://community.tableau.com/thread/106166
https://onlinehelp.tableau.com/current/pro/desktop/en-us/buildexamples_boxplot.htm
https://www.evolytics.com/blog/tableau-201-make-box-whisker-plot/