

Data Visualisation Project in Tableau

Sanjeev Yadav

January 3, 2018

1 Summary

I have attached the initial and final version of my workbook. They contain my analysis and visualisations on baseball players dataset. First I have shown general information about data and then compared some properties based on handedness of players. Finally, I have created some visualisations for top players. It also shows clearly how the feedback helped me improve my work.

2 Design

1. When looking at the data source, I saw there was only one categorical variable, i.e., handedness. Name has many categories so it is difficult to categorise on the basis of name. The number of records gave me an idea that right handed players are more so they will contribute for large value for every feature which depends on number of observations. Like in sum of homeruns or sum of batting averages or sum of weight or sum of heights. I also saw outliers when taking the sum of height and weight. That is because the outliers were repeated entries of some players. I corrected them by taking average. I have shown in my visualisation when taking sum can provide biased results and taking average can provide unbiased results. I had tried to categorize my plots based on handedness, so there are many bar charts in my visualisation. In order to see the trends of height and weight, I created a scatterplot and came to know about the duplicates because their height and weight were almost double than that of normal players.
2. Based on Feedback 1: Use of black shades as color pattern to make the view as basic as possible. I have used this in the entire workbook.
3. Based on Feedback 2: In most of the plots, I have removed the headers from both the axis and tried to include the information as label, the label for some number looks crowded in **sheet** view, but it is optimised in **dashboard** view, which looks nicer.
4. Based on Feedback 3, I have tried to compare and produce some conclusion in every part of my story.

5. Based on Feedback 4: I have followed the pattern of general to specific before this feedback. First I have provided general visualisations for the data and then information for specific observations. So, you will not see any change based on this feedback.

3 Feedback

1. To use same color pattern throughout because categorical variable have not more than 3 categories.
2. Title to be as subjective and short as possible. Use less text and show the values on bars in bar charts because there are only 3 values to display in each bar chart.
3. Every part of story should end with a new result for the reader.
4. Maintain workflow.

4 Resources

None.