## Research Question

Is there a relationship between anthropometric data with baseball performance?

Basebll performance here is defined by two variables.

1) Batting Average (%)

2) Home Runs (Absolute)

First Visualization :- https://public.tableau.com/profile/sofyan.sahrom#!/vizhome/DAND-FirstVisualisation/DAND-FirstVisualisation

Final Visualization :- <https://public.tableau.com/profile/sofyan.sahrom#!/vizhome/DAND-FinalVisualisation/DAND-FirstVisualisation>

## Summary

### Introduction

Anthropometric Data has been used to profile athletes for various reasons such as performance monitoring and talent identification. For this data set, several anthropometric data was made availible to help correlate with performance. Body Mass Index (BMI) was created as an indirect (very) reflection of the athletes stature. As mentioned above, performance is defined by two variables:

1) Batting Average (%)

2) Home Runs (Absolute)

### Conclusion

Therefore based on the observations made, we can conclude that:

1) Left handedness tend to be at least consistent with the performance (both batting average and home runs) and "possibly" even better (relative) wise.

2) Athletes who are vertically challenged and heavier seems to have a slightly better batting average

3) For Home runs, it is a bit more open, both vertically challenged and vertically gifted seems to be better at it. , however the difference may not be significant when compared in relative terms.

## Methods / Design - First Visualisation

To answer the research question, we adopted 6 stages:

#### Stage 1: Surveying the Battlefield

Before we can do any EDA, we must look at the distribution of the samples size. From the histogram below, we can see we have a relatively normal distribution in terms of 1) height, 2) weight, and 3) Body Mass Index (BMI), with a wide spread. This would allow a better reliability in terms of detecting pattern and coorelation etc.

Stage 2: Handness of Athletes.

Having surveyed the distribution, we now proceed to other details such as the handness of the athletes and its relationship with performance. Performance here is defined as both batting average and home runs. From the bar chart below we can see that left handed athletes tend to have better batting average and home runs. However, it is important to "not jump the gun" and consider the relative percentage of left handed athletes as shown in the pie chart below. Since there are almost double the number of right handed (as compared to left handed) and 7 times (as compared to ambidextrous), it is logical that the average will be spread out and "diluted" which could partially explain the lower average. Neverthless it does seem that left-handed athletes in general have better performance (taking into consideration BOTH batting average and home runs).

Stage 3: Performance vs. BMI

Having looked at the handedness, we will now look at overall stature. To do so, we will use BMI which integrates both height and weight. From the bar graph below, there does not seem to be much difference in batting average, possibly a very slightly higher average for bigger stature. However for the home runs, it seems that the bigger athletes have a higher average home runs. This makes sense, as batting average is more dependent on the "batting skill" so stature will have a smaller part to play. Home runs on the other hand is dependent on the ability to run or cover the distance around the diamond, thus a bigger athelte will possible longer limb length might be able to cover the distance better. However it can be also argued otherwise, thus, we may need to look deeper such as the actual height.

Stage 4: Performance vs. Height

As observed previously, there is a possibility that the stature or anthropometric data might be a factor in the performance. Thus we will now delve further and investigate height. From the line chart below, we can see the opposite of what we saw earlier, that the the vertical challenged (shorter) athletes are the ones with the better batting average. For the home runs we observe a similiar pattern, however the vertical gifter also seems to have decent home runs.

Stage 5: Performance vs. Weight

For weight, we see the same pattern for batting average, the lighter athletes seems to have a higher average batting average. However for the home runs, it is the opposite, it seems the heavier athletes seems to have a better average home runs.

Stage 6: Summary

See Conclusion above

**Feedback**

After seeking feedback. The feedback can be summarised into the 3 following points.

1) Color Scheme

a) Boring - Your colour scheme is very boring. Try using differennt colors for the 2 performance variables. Try using other colors for other graphs.

b) Consistency - You should use the same color scheme for home runs, so that it easier to associate. Right now, different page has different color for the home runs.

2) Your pie is for "anorexic"/malnourished.

Overall Size is small. So is the font size also. Make it bigger.

3) Label Values.

While the visualisation is cretive, you should still try to provide a quantitative number.

**Final Visualisation**

After the feedback both written and visual, we addressed them and produced the final visualisation.

**Resources**

https://onlinehelp.tableau.com/current/pro/desktop/en-us/publish\_workbooks\_tableaupublic.html

https://www.tableau.com/learn/training

https://community.tableau.com/thread/137413

http://onlinehelp.tableau.com/v10.4/pro/desktop/en-us/help.htm#extracting\_data.html

https://www.youtube.com/watch?v=znV12jYDyB0