Analyzing Performance of Baseball players – by Babatunji Ogunjobi

1. Overview: I analyzed the 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 in a bid to draw up valid insights on how these metrics affect baseball player performance. I have tried to make the report a lot more relatable to by conducting a mini- research on major league baseball and its record holders.

2. Within the analysis, I tried to answer the following questions

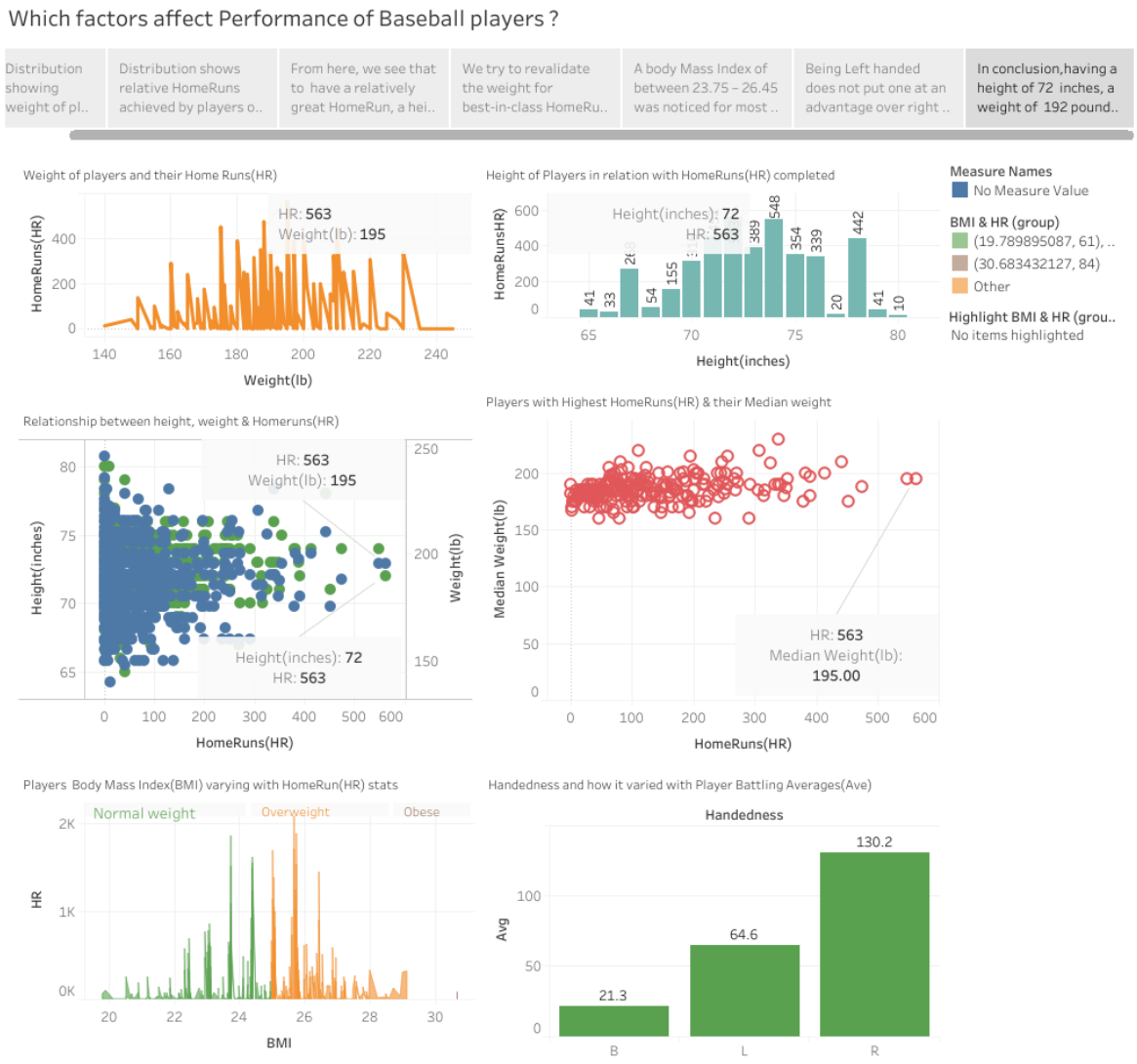
I. Is weight a true determinant for high performance as it relates to Homeruns in Baseball?
II. Is Height a true determinant for high performance as it relates to Homeruns in Baseball?
III. Is a particular weight & height associated with the best players with the highest homeruns?
IV. Is there a relationship between Body Mass Index (BMI) & Homeruns?
V. Are the base ballers within the dataset relatively overweight?
VI. Does a particular handedness affect batting average?
VII. Are left-handed players at an advantage over right-handed players as it relates to batting average?
VIII. Is there a relationship between Weight, height, BMI and batting average?
IX. What is the ideal body configuration in terms of weight, height & BMI that guarantees good Homerun & Batting average performance?

3. Design Choices

1. I resorted to doing a bit more research on baseball to add more context to the analysis
2. For the independent weight & height graph against Homeruns I used line-charts & bar-charts respectively to try to show that the variables changed per player
3. I used scatter plots do show the correlation relationship between the weight and height and how they affected Homerun in one graph and how in relation to BMI, they affected the batting averages of players
4. I tried to use as many colors as possible to distinguish different variables from each other
5. I ensured all my charts had legends and their axis was properly labeled to ensure clear information is passed to readers
6. I used the calculation module within tableau to calculate the Body Mass index of players in a bid to add depth to the analysis being done

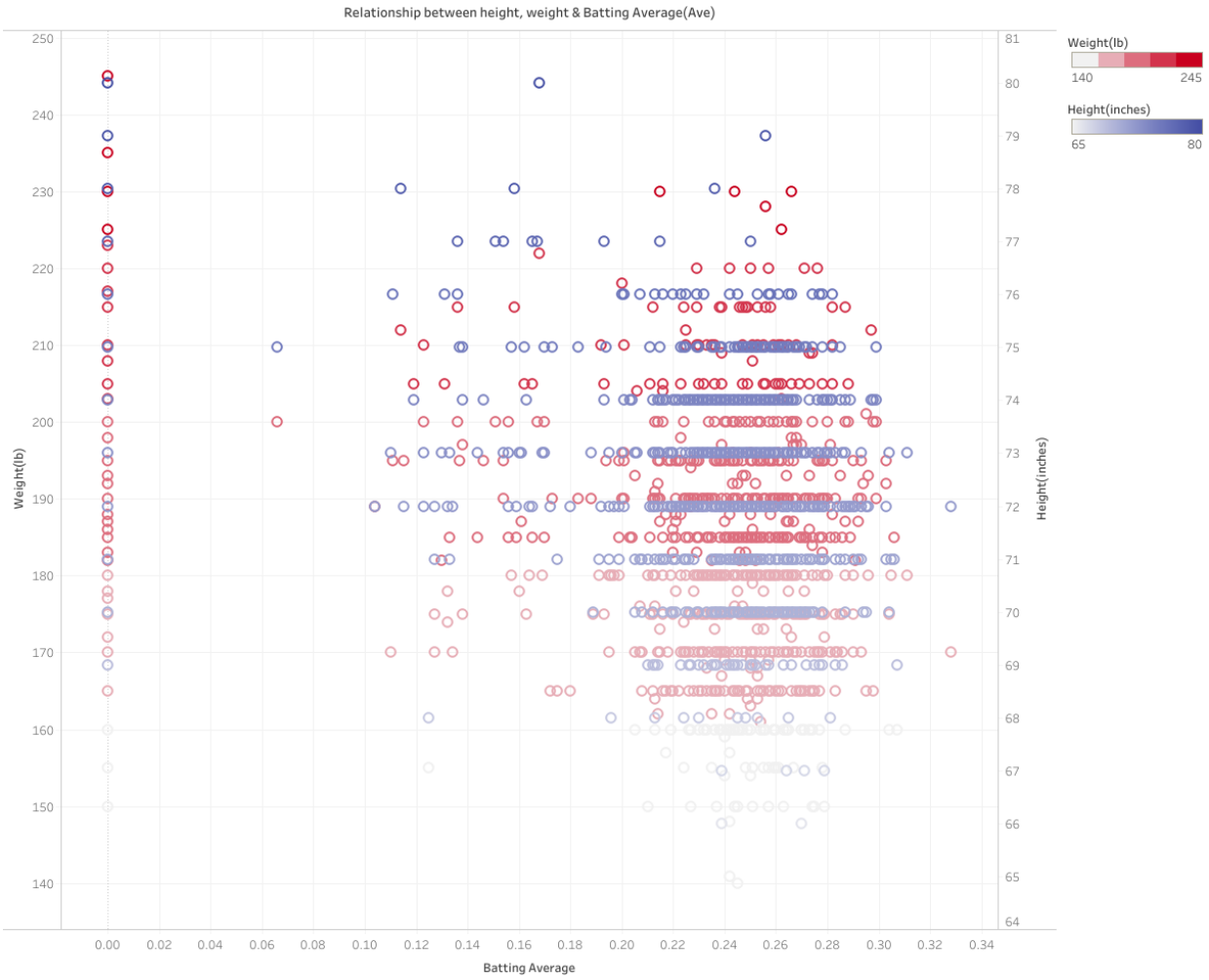
4. First Story:

https://public.tableau.com/profile/babatunji#!/vizhome/mybaseballdata-1/Story1?publish=yes

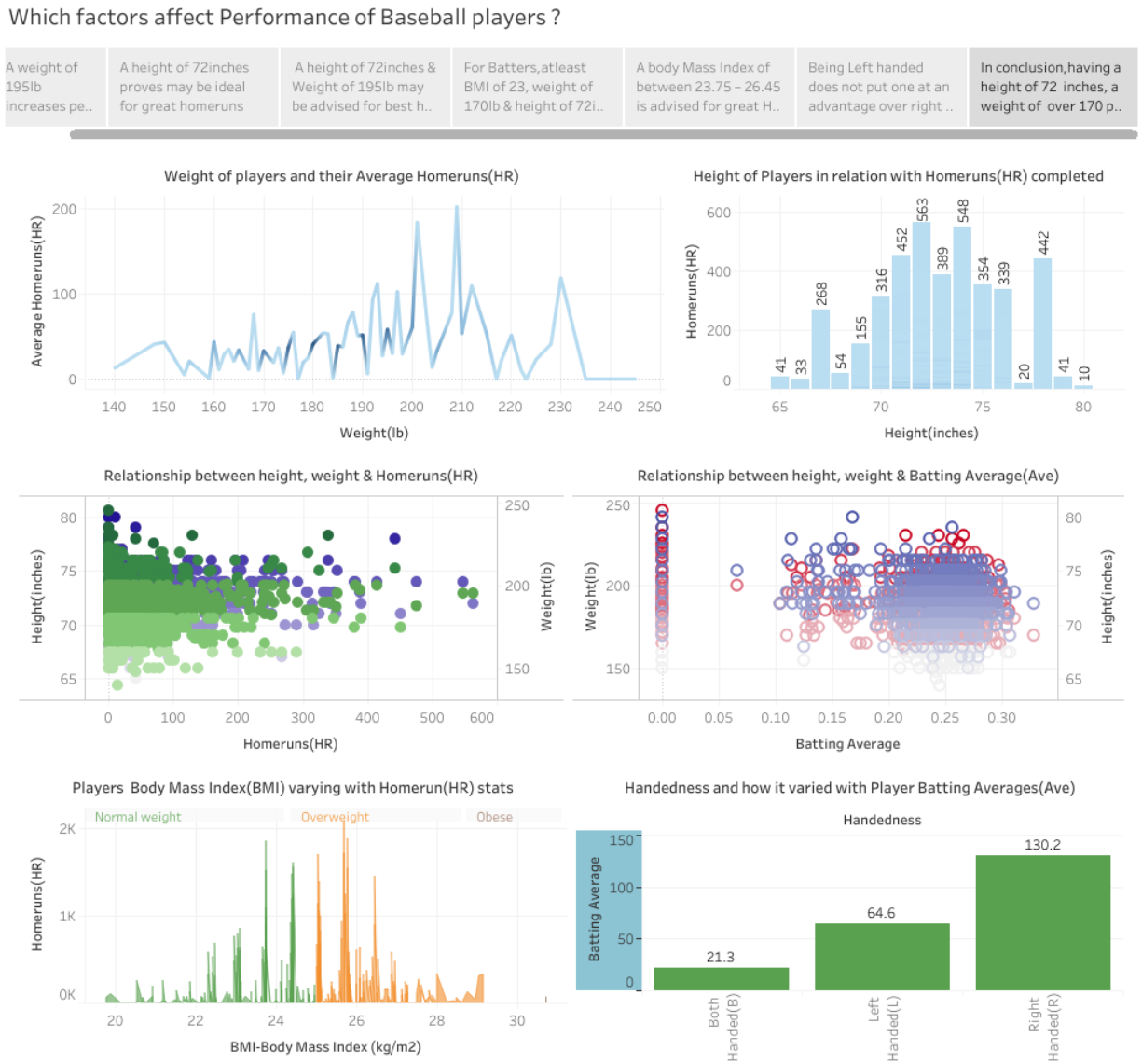


5. Final Story:

https://public.tableau.com/views/mybaseballdata-final/HomerunsvsWeightlbHeightInches?:embed=y&:display_count=yes&publish=yes


Relationship between height, weight & Batting Average(Ave)
Weight(lb)
140
245
Height(inches)
65
80
Weight(lb)
Height(inches)
Batting Average
250
240
230
220
210
200
190
180
170
160
150
140
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
65
64
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0.16
0.18
0.20
0.22
0.24
0.26
0.28
0.30
0.32
0.34

6. Summary



i. Analyzing if body configuration affects homeruns

Even though **Barry Bonds**, who holds the Major league Baseball home run record with **762 home runs** over a **22years career**, had a weight of between **220-240pounds** and a height of **74inches** during his playing years, wearefanatics.com tends to believe that the size or shape of major league baseball players does not determine performance.

Quoting **wearefanatics.com**: “While athletic ability is an essential part of being a Major League Baseball player, there are absolutely no size or shape requirements for being added to a big league roster. If you play hard and tick all the boxes managers are looking for, you can make it as a pro player regardless of how short, tall, thin, or heavy you are.”

http://wearefanatics.com/baseball-players-shapes-sizes/

We are fanatics also states the following

- Average weight of an MLB player: 207pounds
- Average height of a MLB player: 74inches

Assumption: The provided dataset in the file baseball_data.csv is a list of data of college students who are about to be drafted as Major league baseball players.

I answered the following questions with the provided dataset.

a) **Question**: Is **weight** a true determinant for high performance as it relates to **Homeruns** in Baseball?
Answer: Even though one player with a weight of 196pounds had the highest homeruns of 548, other players with relatively lower and higher weights seemed to do pretty well. However I noticed that the graph seemed to form an imaginary bell curve, which can mean that there is a median weight for high performance.

b) **Question**: Is **Height** a true determinant for high performance as it relates to **Homeruns** in Baseball?
Answer: The player with the highest homerun of 563 had a height of 72inches. You can also find players with lower and higher heights who had great homeruns and some who did not have. The imaginary bell curve was also noticeable within this analysis.

c) **Question**: Is a particular **weight** & **height** associated with the best players with the **highest homeruns?**
Answer: By plotting the weight & height against the Homeruns, I was able to determine that even though weight & heights are not currently considered for signing up Major league baseball players, players may be advised to strive to achieve a certain weight & height as the best Homeruns in my data were achieved at a weight of **195pounds** & a height of **72inches**

d) **Question**: Is there a relationship between Body Mass Index (BMI) & Homeruns?

Answer: Since body mass index aims to measure if an individual is overweight by dividing the weight in kilograms by the square of height in meters, I tried to check if this affected the Homeruns of the players. I noticed that a player with a BMI of 25.68(partially overweight) had a homerun of 414 and the trends shows that most players who had the best homeruns had a BMI of between 23.75 – 26.45.

e) **Question**: Are the base ballers within the dataset relatively overweight?
Answer: Generally, the weight chart also had the imaginary bell curve, which depicted that majority of the players, strived to achieve the above BMI level but were not overweight. Only a select few of them were overweight.

ii. Looking further into average Batting & Body configuration

From an analysis done by perceptionaction.com for average overall batting average between 1992 – 2015 in the **Major-league**, it showed that overtime, right handed batters have had a higher batting average than their left handed batters and being left handed does not increase performance.

Within the analysis I tried to check with the provided dataset if perceptionaction.com thesis is correct as Outfielder Ty Cobb, whose career ended in 1928, has the highest batting average in Major League Baseball (MLB) history was left handed. He batted .366 over **24 seasons**, mostly with the Detroit Tigers.

As such within this analysis I answered the following questions.

a) **Question: Does a particular handedness** affect **batting average?**
Answer: I noticed that a larger proportion of players analyzed who had high batting averages were right handed. This could be because in the world, the numbers of right-handed people are more than left handed people. It can also mean that right-handedness leads to increased batting averages but Ty Cobbs who holds the highest record was left handed so I am inclined to say the reason why we had improved batting averages in right handed people is because we have more right-handed people in the world.

b) Question: Are **left-handed** players at an advantage over **right-handed** players as it relates to **batting average?**
Answer: From the data, the left handed players seemed to perform worse than right handed people so we can infer that being left handed does not put one at an advantage for improved batting averages.

c) **Question**: Is there a relationship between Weight, height, BMI and batting average?
Answer: In terms of the relationship between Weight, height, BMI and batting average, the best batting average was achieved by a player who had a BMI of 23, weight of 170lb & height of 72inches.This follows the trend for homeruns in relation to height, weight & BMI even though it appears that the weight requirements for batters is less than that for runners.

Final Conclusion:
In conclusion, having a height of 72inches, a weight of over 170pounds & a BMI of over 23, may aid the performance of a baseball player. Also regardless of the hand used (left, right or both) for batting, a player can always perform. This partially explains Barry Bonds & Ty Cobbs success even though a lot of it may have had to with skill as well.

7. Feedback

I showed my colleagues at work. They do not use tableau or any other data analysis tool but they understand the relevance of data and the need to draw valid insights to make better decisions.
They wanted me to include

a) A view that let's them see player BMI within the sheets were I analyzed **Weight (lb) vs. HR** & **Height (inches) vs. HR** without them having to scroll to the BMI page.
b) They felt the page showing "**HR vs. Median weight**" was irrelevant as it has same information as analysis showing **"Weight (lb) vs. HR"** so I took it out.
c) They also wanted me to show if there was a relationship between **Weight**, **height**, **BMI** and **batting average** so I added one more analysis page.

d) Ensure all legends are properly labeled and the stories are short, direct & concise

The above suggestions were implemented in the final worksheet.

8. References.

Barrybonds
https://www.reuters.com/article/us-baseball-crime-bonds-steroids/bonds-weight-gain-used-as-steroid-evidence-by-many-idUSN1639833020071116
https://en.wikipedia.org/wiki/Barry_Bonds
https://en.wikipedia.org/wiki/List_of_Major_League_Baseball_career_home_run_leaders

TyCobbs
https://en.wikipedia.org/wiki/List_of_Major_League_Baseball_career_batting_average_leaders
https://www.detroitathletic.com/blog/2011/12/22/wrong-handed-the-most-famous-lefties-in-detroit-sports-history/

Batting Average Analysis
http://perceptionaction.com/battinghand/

Handedness
https://en.wikipedia.org/wiki/Handedness

HomeRun
https://en.wikipedia.org/wiki/Home_run

Right handed people vs. left handed people in the world
https://www.livescience.com/19968-study-reveals-lefties-rare.html

Glossary

- **Handedness**: the tendency to use either the right or the left hand more naturally than the other.
- **Homerun**: In baseball, a home run (abbreviated HR) is scored when the ball is hit in such a way that the batter is able to circle the bases and reach home safely in one play without any errors being committed by the defensive team in the process (Wikipedia).