Baseball Story

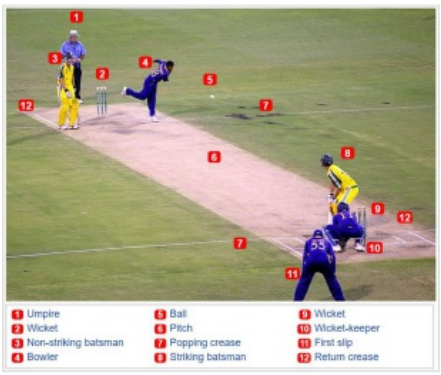
Once Upon A Time there lived a guy named Padmaraj (Me) who thought baseball game is a variation of Cricket game. Unlike in Cricket, you have bat to hit the ball, in baseball you use wicket to hit the ball. Since, you use wicket to hit the ball, you do not have wicket placed on the ground.

Fortunately, one day he found a task to Analyze Baseball Score Dataset which opened his eyes and could see that baseball as a entire different game. On the same day he also came to know that HR does not only stands for Human Resource but also stands for 'Home Run'. Having experience in Bengaluru Traffic, first he thought Home Run is all about running back home as early as possible to skip the traffic. He thought, he knows the game completely and can kick start analysis on the data set. And then the miracle happened, he saw the MLB videos to get the domain knowledge and he then got know the actual game.

He then happily carried out the data analysis and visualization on the data set and his analysis are shown in the next story slides. Please review and give feedback.

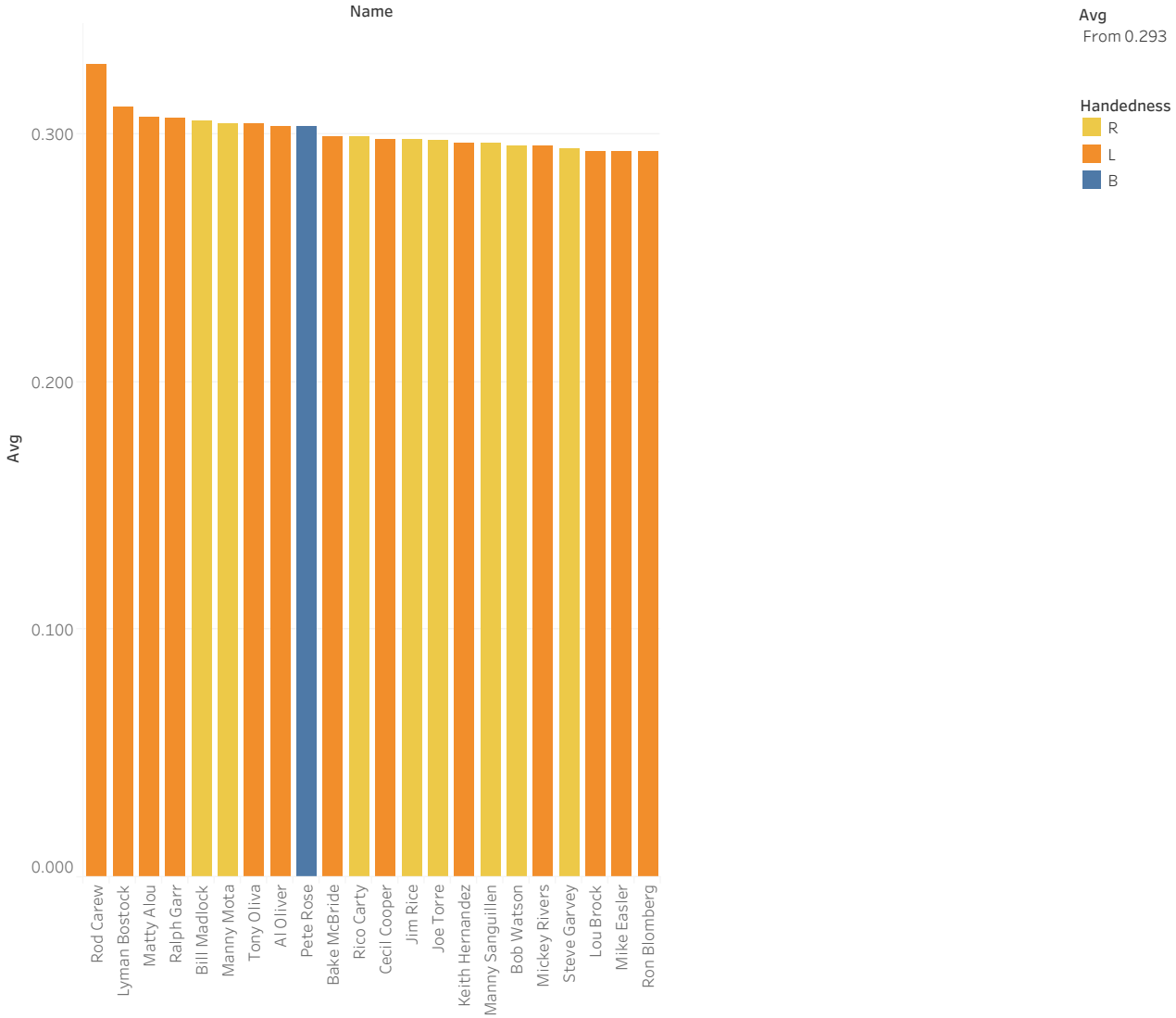
Baseball Picture credit: By SecondPrint Productions on Flickr - From Flickr; description page is (was) here, CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=3083155

Cricket Image Source: Wikipedia



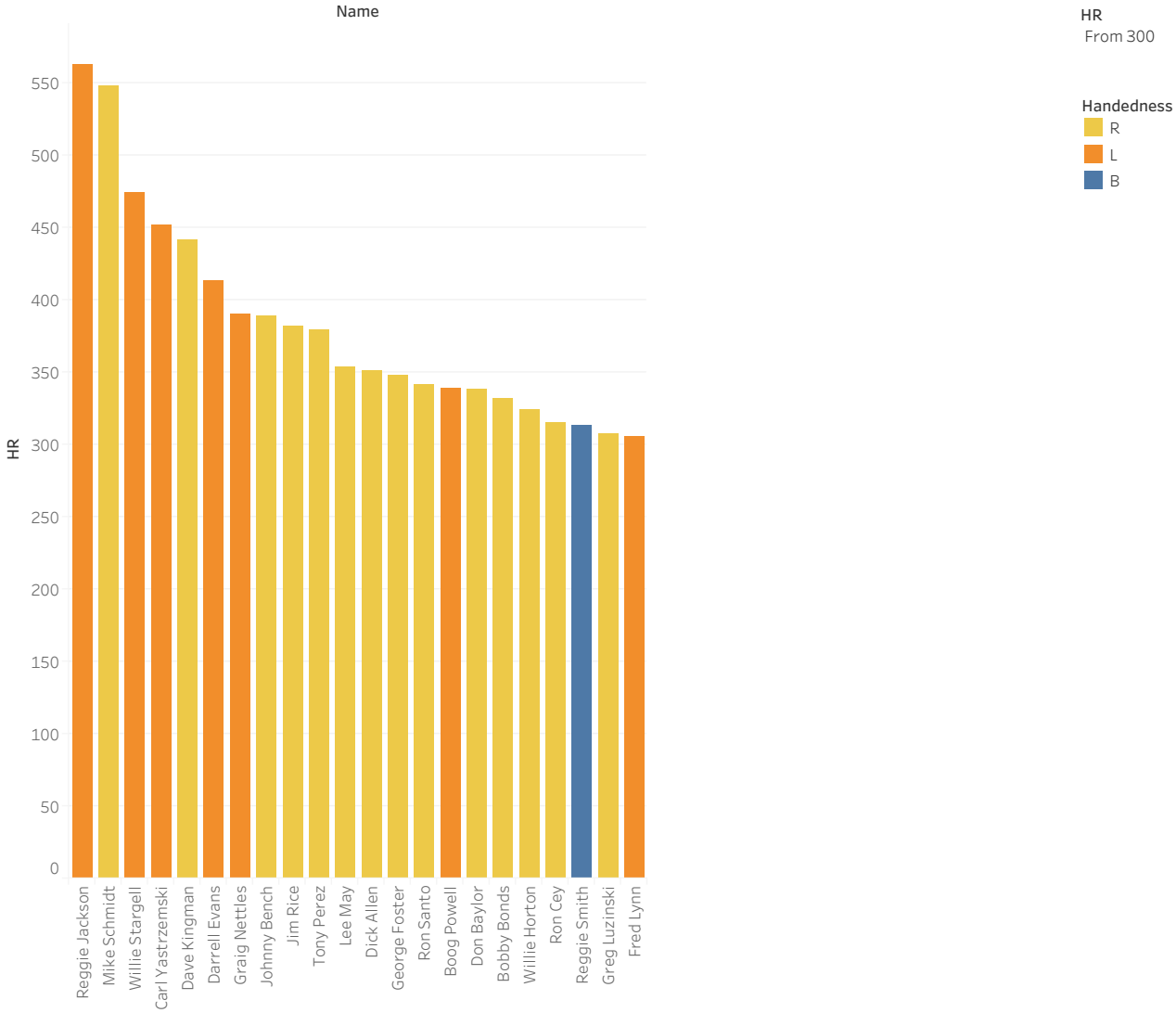
Baseball Story

Once Upon A Time | Top Career Batting Averages | Top Most Career Home Runs | Ouch !!! | Left or Right | Initial Plot | Pitchers and Batters



Baseball Story

Once Upon A Time | Top Career Batting Averages | Top Most Career Home Runs | Ouch !!! | Left or Right | Initial Plot | Pitchers and Batters



Baseball Story

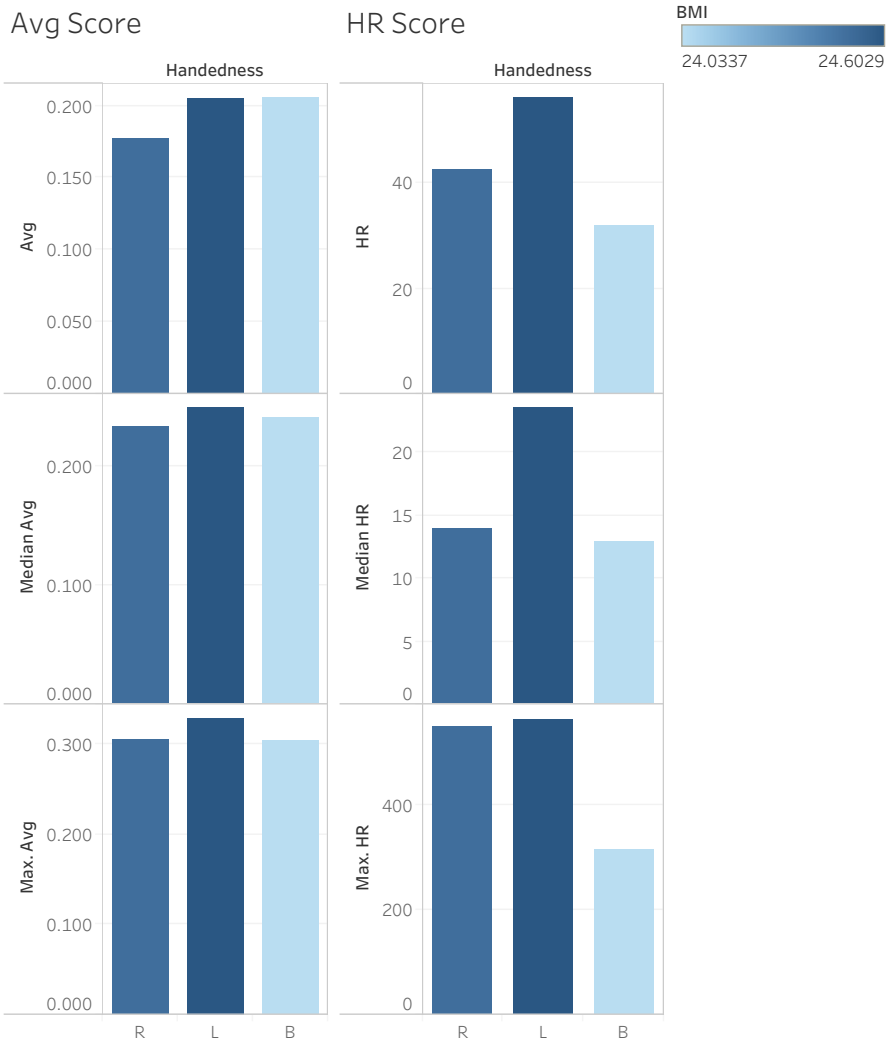
If you had an eagle eye, you might have already said "ouch" by now.
1. In both the previous visualizations, we see that topper is left handed batsman.
2. Most of the Top 20 Avg scorer are left handed batsmen.
3. Most of the top 20 HR scorer are right handed batsmen.

The 'ouch' factor is that the above values tend to bias towards the left handed batsmen because they have the high score. High score is going to pull average towards left handed batsman.

So....Who is going to rescue right handed batsmen .... yes... The Quartiles...especially IQR.

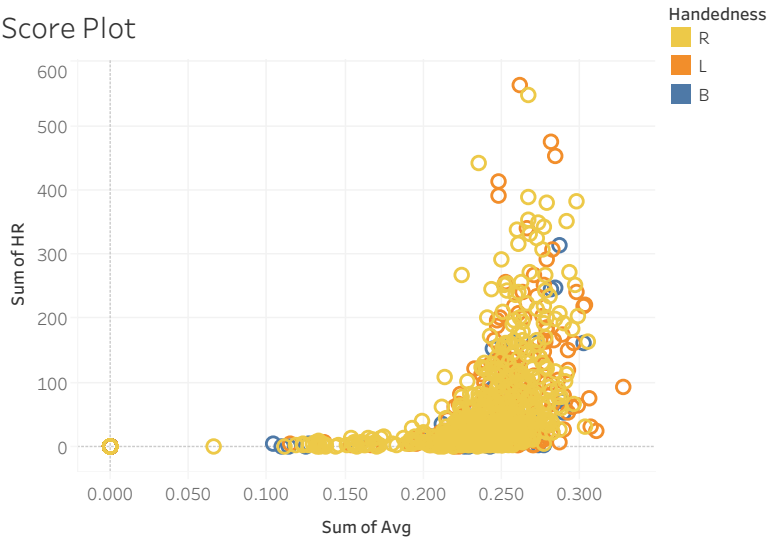
Baseball Story

Wow.... Is Baseball meant for left handed batter ?

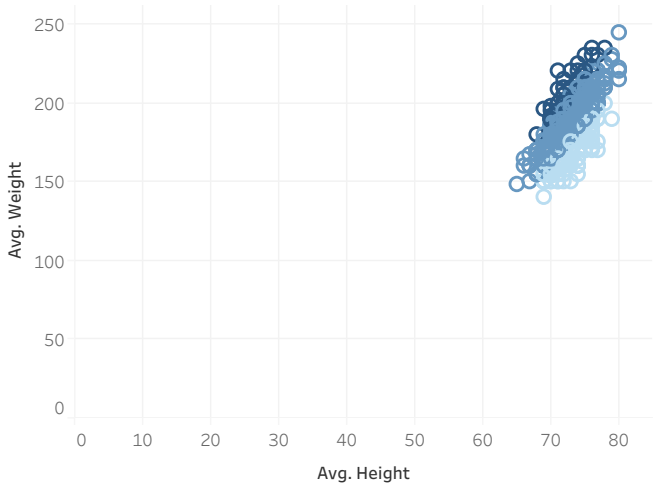


Baseball Story

Score Plot

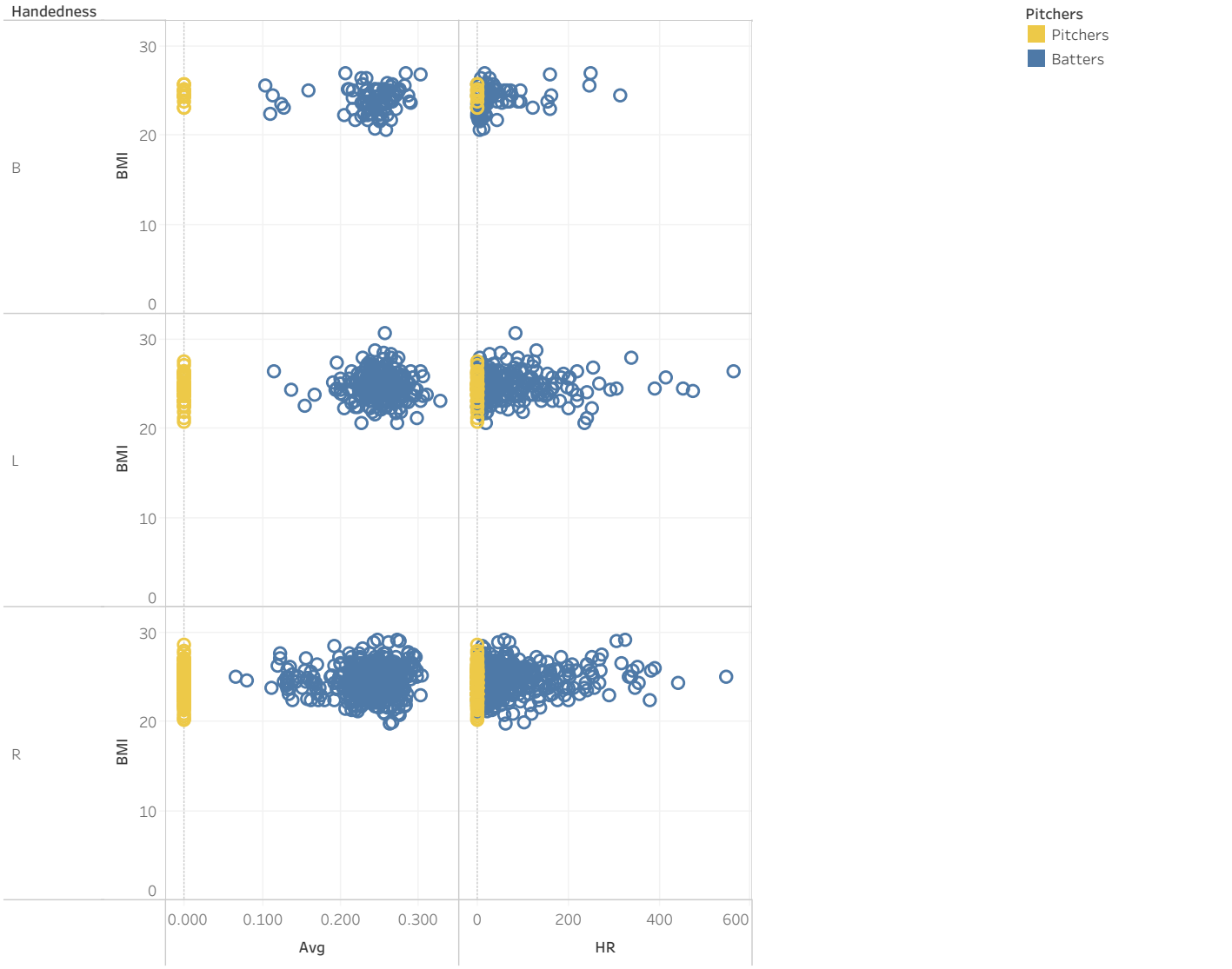


Body Mass Index (BMI) Plot



Baseball Story

Ouch !!! | Left or Right | Initial Plot | Pitchers and Batters | Pitcher Intro | Pitcher per Handedness Counts | IQR in Rescue operation

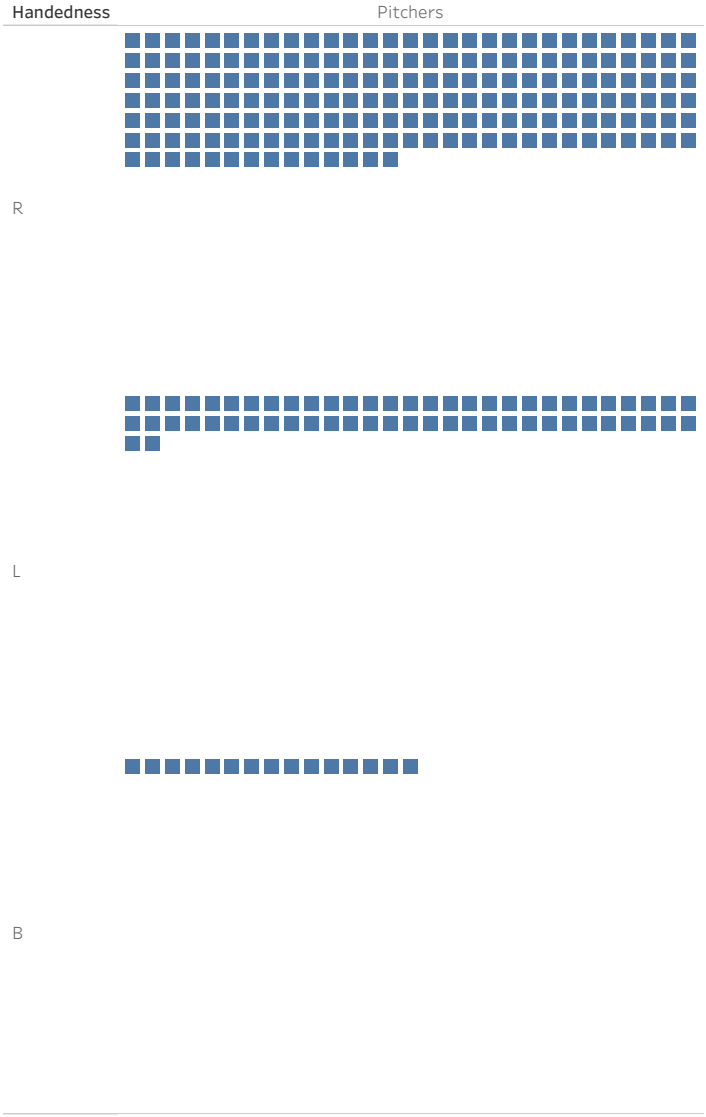


Baseball Story

People plotted in yellow are the Pitchers. They are the ones who try to break the Catchers Nose but batters are the ones who comes in between and spoils thier plan. Clearly, they do not have room in the analysis of the for batting score. Since they lied that they are batsmen and thus got into our dataset; we are going to classify them as Outliers. There are 266 of them out of 1157 players and we will clean/take them out.

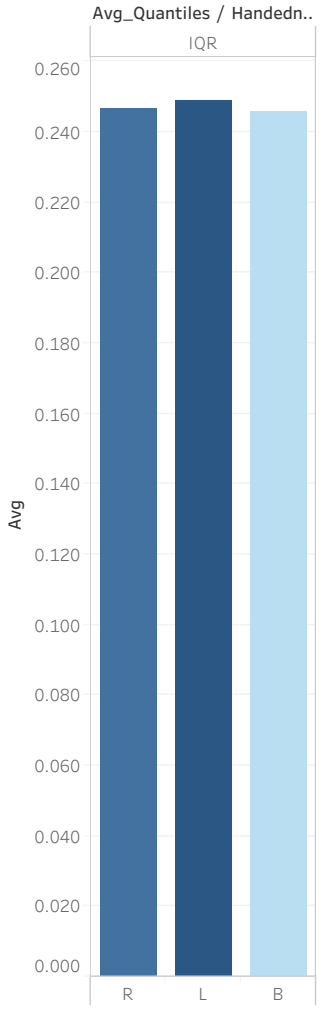
Now, with the outlier removed, Inter Quartile Range is 0.135 - 0.258

Baseball Story

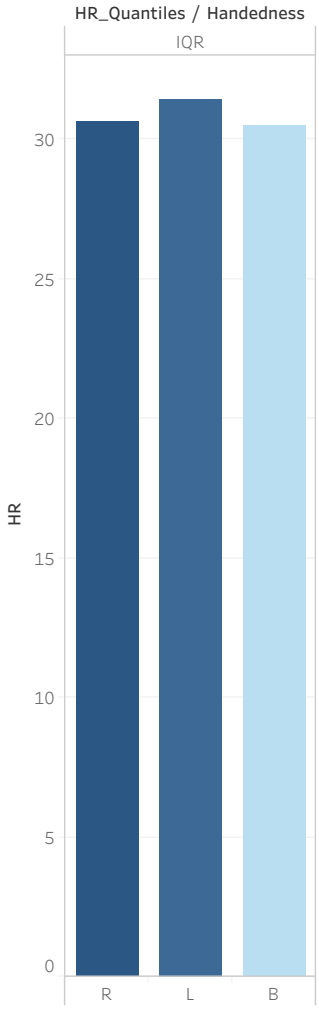


Baseball Story

AVG Quantile Distribution

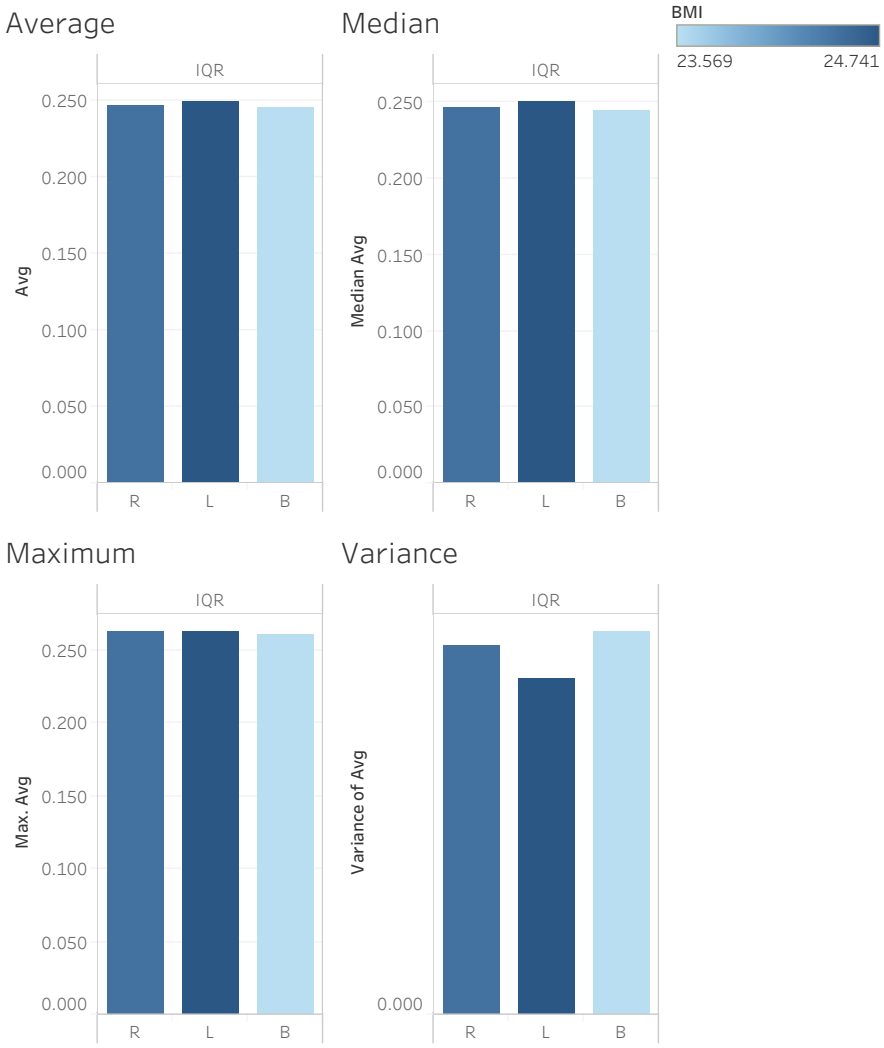


HR Quantile Distribution



Baseball Story

Avg Score in IQR



Baseball Story

Pitcher per Handedness Counts	IQR in Rescue operation	Detailed look on Avg Score in IQR	Avg IQR Study	Detailed look on Avg Score in Q4	Avg Q4 Study	HR in IQR

Mean, Median and Maximum score comparison between left, right and both handed batsment indicates that Avg Score is only marginally different from each other.

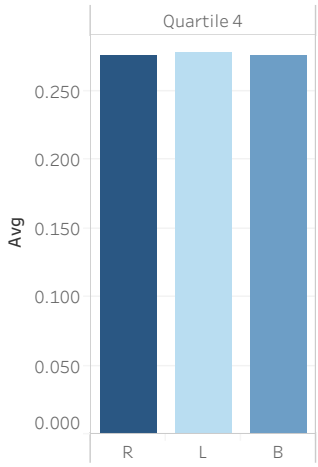
Here dip in the variance for Left handed batsmen indicates that left handed batsmen have (slightly) higher average when compared to other batsmen. Hence the difference between the max and the average reduces.

Baseball Story

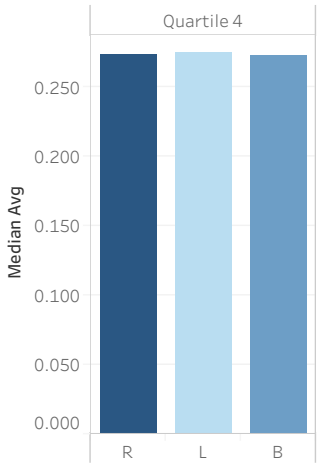
Avg Score in Quartile 4

BMI
24.5309 24.6569

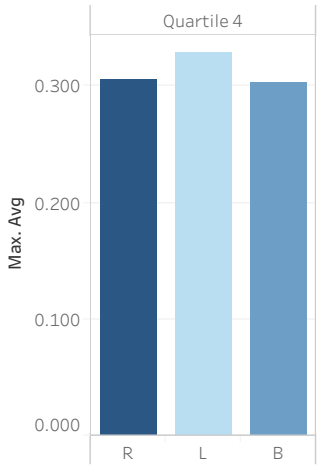
Average



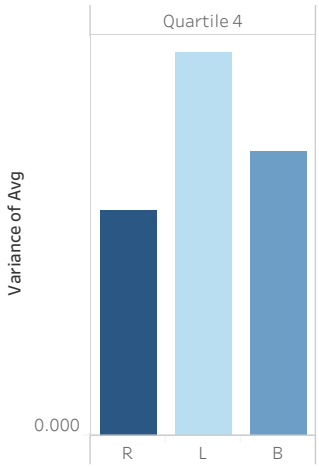
Median



Maximum



Variance



Baseball Story

Quartile 4 contains all the top scorer in the Dataset. Again in Quartile 4, we see that mean and median are almost the same.

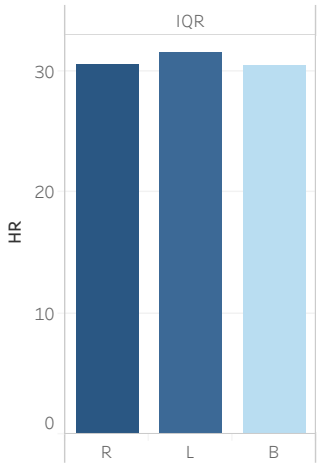
Here, Max and Variance chart complement each other. Since average is comparable(almost the same) as that of other handed batsmen and we have highest score from left handed batsman ; we have variance high for left handed batsmen.

Similar conclusion can be taken from the next two slides which talk about Home Run comparison between right, left and both handed batsmen in both IQR region and Quartile 4 region. Unlike we saw in case of Avg score, HR score comparison indicates right handed or both handed player have higher score than that of left handed batsmen.

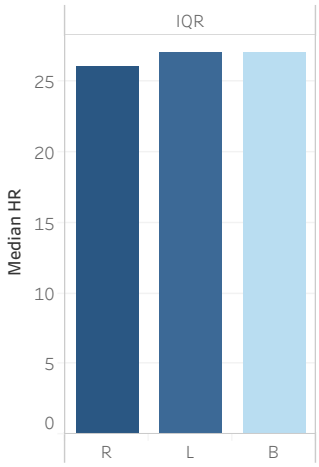
Baseball Story

HR Score in IQR

Average

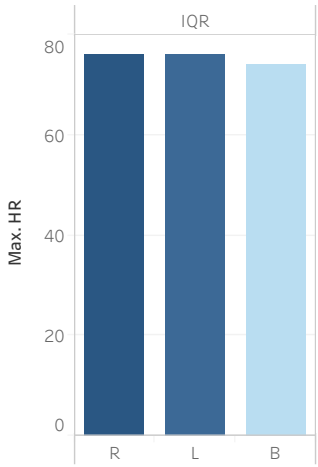


Median

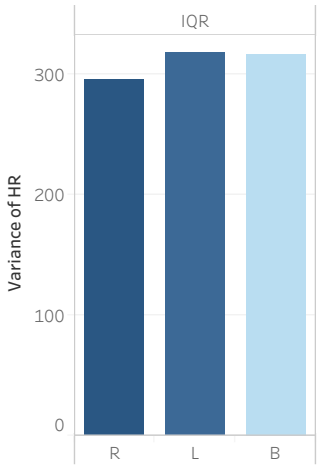


BMI
24.2296 24.6843

Maximum



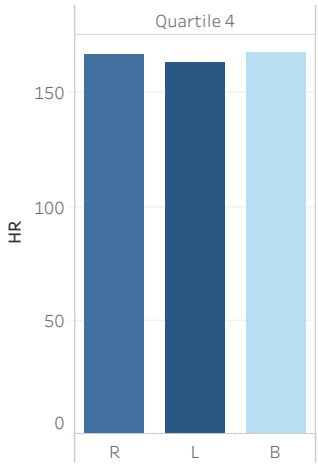
Variance



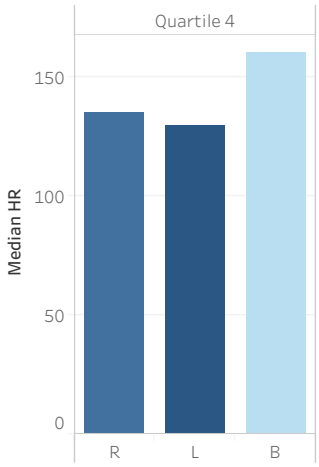
Baseball Story

HR Score in Quartile 4

Average



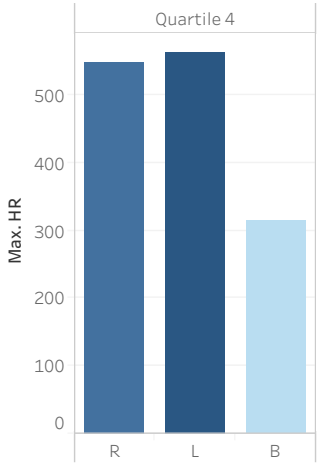
Median



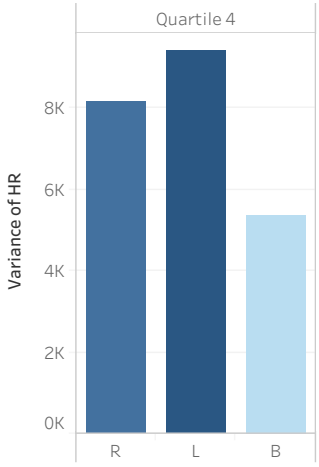
BMI

24.5842 24.8954

Maximum



Variance



Baseball Story

Conclusion

Both (left and right) handed battting score analysis indicates that there are no valid evidences to prove any one of them is better. The only thing we can derive out of the analysis is the BMI range. All the players in IQR and Q4 fall under the range of 23.5 - 24.89. And around 24.5 units being the top performer's BMI.

It seemed like a left handed person has created this data set to win over right handed folks. Well, at least post this analysis, we agree that it not true. So, we shake hand !!! (of course, with right hand :P)

Baseball Story

Best Pitcher :

Over the course of this project, I have watched lot of baseball game snippets and kind of gained respect for it too. I have googled about couple of players from the dataset. The Dataset talks about batting score but includes Pitcher batting score too. Pitcher stat info was not there in the Dataset, so I took the all time best pitcher (as indicated by Google search result) snap here.Picture credit goes to Wikipedia and Google. As a minor tribute to the game and great players, I am putting snaps of key players I have Googled. A Grand Salute !!! :)

Baseball Story

Best Left handed Batsmen (from the Dataset)

Rod Carew

Baseball first baseman

Rodney Cline Carew is a former Major League Baseball first baseman, second baseman and coach of Panamanian descent. Wikipedia

Born: 1 October 1945 (age 72 years), Gatún, Panama

Hits: 3,053

Batting average: .328

Home runs: 92

Hall of fame induction: 1991

Spouse: Rhonda Carew (m. 2001), Marilynn Levy (m. 1970–2000)

Reggie Jackson

American baseball athlete

Reginald Martinez Jackson is an American former professional baseball right fielder who played 21 seasons for the Kansas City / Oakland Athletics, Baltimore Orioles, New York Yankees, and California Angels of Major League Baseball. Wikipedia

Born: 18 May 1946 (age 72 years), Wyncote, Pennsylvania, United States

Height: 1.83 m

Home runs: 563

Spouse: Jennie Campos (m. 1968–1973)

Number: 44 (New York Yankees), 9 (Oakland Athletics / Outfielder), 9 (Baltimore Orioles), 44 (Los Angeles Angels)

Baseball Story

Best Right Handed Batsmen (from the Dataset)

Bill Madlock

Baseball player

Bill "Mad Dog" Madlock, Jr. is a former Major League Baseball player. From 1973 to 1987, Madlock was a right-handed hitter who won four National League batting titles. Wikipedia

Born: 2 January 1951 (age 67 years), Memphis, Tennessee, United States

Height: 1.8 m

Bats: Right-handed

Weight: 86 kg

Runs batted in: 860

Batting average: .305

Mike Schmidt

Baseball third baseman

Michael Jack Schmidt is an American former professional baseball third baseman who played 17 seasons in Major League Baseball for the Philadelphia Phillies. Wikipedia

Born: 27 September 1949 (age 68 years), Dayton, Ohio, United States

Home runs: 548

Spouse: Donna Wightman (m. 1974)

Batting average: .267

Number: 20 (Philadelphia Phillies / Infielder)

Hall of fame induction: 1995

Baseball Story

Thank You :)