## **About Dataset**

***Update 2023-10-13:*** The data now includes 2022 season.

***Update 2022-08-06:*** The data now includes 2021 season.

***Update 2021-08-02:*** The data now includes 2020 season and metrics for 2019 have been updated.

***Update 2020-08-03:*** The data now includes 2017, 2018 and 2019 seasons. Keep in mind that metrics like gp, pts, reb, etc. are not complete for 2019 season, as it is ongoing at the time of upload.

### **Context**

As a life-long fan of basketball, I always wanted to combine my enthusiasm for the sport with passion for analytics 🏀📊. So, I utilized the NBA Stats API to pull together this data set. I hope it will prove to be as interesting to work with for you as it has been for me!

### **Content**

The data set contains over two decades of data on each player who has been part of an NBA teams' roster. It captures demographic variables such as age, height, weight and place of birth, biographical details like the team played for, draft year and round. In addition, it has basic box score statistics such as games played, average number of points, rebounds, assists, etc.

The pull initially contained 52 rows of missing data. The gaps have been manually filled using data from [Basketball Reference](https://www.basketball-reference.com/). I am not aware of any other data quality issues.

### **Analysis Ideas**

The data set can be used to explore how age/height/weight tendencies have changed over time due to changes in game philosophy and player development strategies. Also, it could be interesting to see how geographically diverse the NBA is and how oversees talents have influenced it. A longitudinal study on players' career arches can also be performed.