What Is the Best Body Type For The NBA?

Zachary Shinnick – A1851035

Abstract

This study aims to determine the optimal body type for an NBA basketballer. The data utilised to address this question was sourced from Kaggle, and has been scraped from the NBA stat API to contain statistical, biological and biographical data on every NBA player from 1996 to 2020. The data set consists of over 11,000 players, each player is represented by an individual row, and there are 22 columns of data about that certain player. This report utilises investigations between players body composition and player performance to discover what the optimal body type is for the NBA.

Introduction and Motivation

The problem studied in this investigation is determining the trends in NBA player's body compositions and performance. Specifically, I analysed correlation between players height and weight to determine the optimal body type for individual player statistics, value to team and efficiency. This investigation will present significant findings as the NBA has a large fanbase all over the world and basketball fans will be able to determine the key factors that make up the best NBA players. Additionally, the findings will be insightful for basketball coaches/management who are seeking to discover future talent by providing the body type most likely to be have a successful NBA career. This topic was chosen as I am an enthusiastic basketball player and fan and wanted to combine this with my passion for analytics. I also want to investigate into the often-presented idea that "basketball is only for big men" and determine if this statement is true or false.

Dataset:

To address the proposed investigation, the data source selected comes from the NBA stat API and contains data since 1996 on each player NBA player (over 11,00 players) and was found on Kaggle, consisting over 11,000 rows and 22 columns[1]. The dataset incorporates demographic variables, including height, weight, nationality, country of birth, age, as well as biographical data including, college attended, position selected in draft and team played for. The data set also includes playing statistics: games played, average points, rebounds, assist for each season, as well as advanced analytical statistics: net rating, true shooting percentage and usage percentage.

Methods

To create visualisations, I utilised Python and Jupyter Notebook. Specifically, Panda's library was employed to read in the Dataset as a .CSV file and transform the data into a form which could be analysed and visualized. Pandas' library also filtered out players who did not play in the season, as it was determined the analysis would benefit from the increased accuracy of filtering out these players as they do not contribute to any statistical breakdown.

Pythons Data Frame feature was utilised in creating subsets of data, grouping players into categories and filtering the data with inbuilt features (such as "*.mean()")* based on the needs for that specific analysis. Data Frame was also used to create new columns for investigation, such as creating a body mass index (BMI) column through calculations of the listed height and weights of NBA players to create another biological analysis perspective.

After the data had been successfully filtered, libraries Seaborn and Matplotlib were used for graphing the data. It was determined the visualisations that best displayed the trends of the data was a combination of box plots, line graphs, scatter plots and column graphs.

Matplotlib was also employed for styling of the graphs, including appending axis titles, fonts, size and colours.

Experimental Setup:

I investigated three different biological body composition factors, height, weight and BMI and compared the effects these elements have on different measures for successful NBA performance. The results from each graph and data investigation were then interpreted, and comparisons were made between different findings, to conclude with an optimal body type for the NBA. This allowed for clear findings and conclusions to be taken and compared to answer the research question, *What is the Optimal Body Type for the NBA?*

---

[1] Justinas Cirtautas 2022, *NBA Players*, Kaggle.com, viewed 9 September 2022, <https://www.kaggle.com/datasets/justinas/nba-players-data/metadata>.

Results:

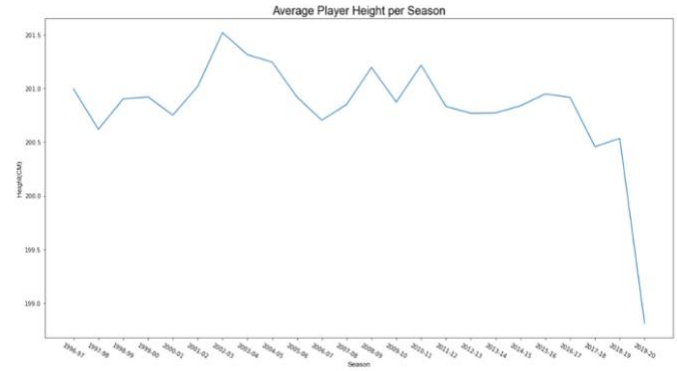


Figure 1: Line plot of average NBA height by Season (made with seaborn)

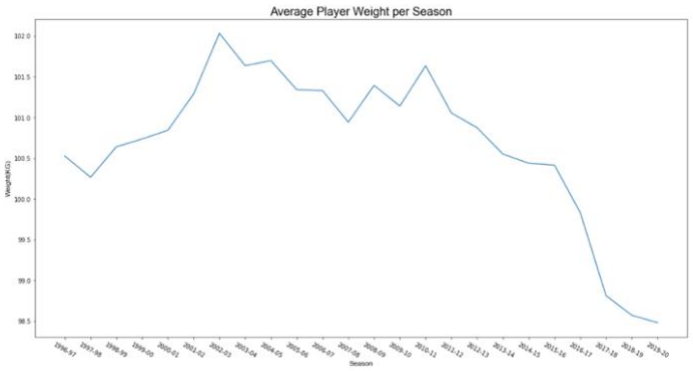


Figure 2: Line plot of average NBA weight by Season (made with seaborn)

This initial investigation provided a clear trend of NBA players getting smaller over time, when analysing the average height (figure 1) and weight (figure 2) over the seasons from 1996 – 2020.

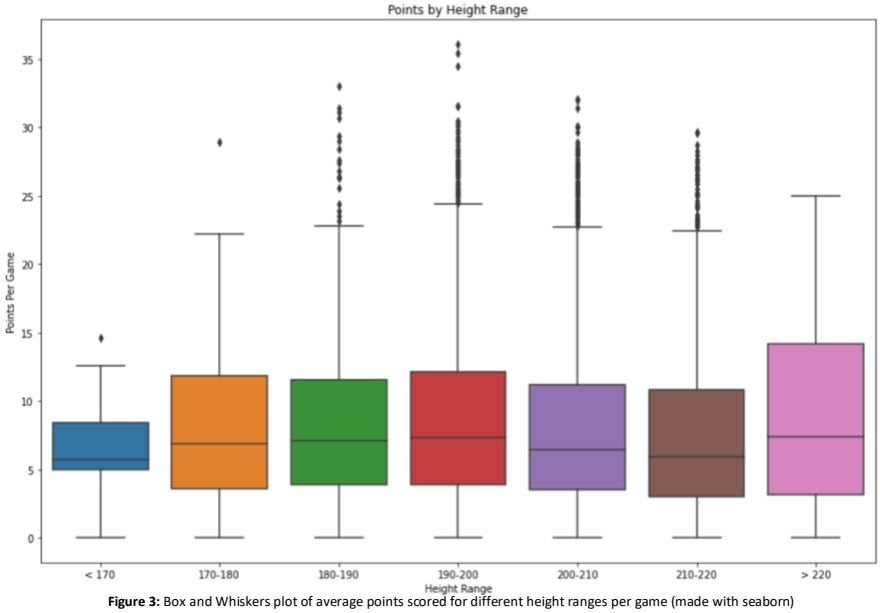


Figure 3: Box and Whiskers plot of average points scored for different height ranges per game (made with seaborn)

Figure 3 demonstrates that 220cm and over players have the greatest median scoring per player, as well as highest upper quartile. Another observation is the 190cm – 200cm category has the highest number of outliers who can be considered elite scorers, however, this can be determined as an irrelevant data observation, as the question proposed aims to address and conclude with the optimal NBA player body type on average.

Average Rebounds and assist's per game were also investigated and visualized in box and whiskers plots (appendix 1 and 2), however these were not included into the findings as it was determined these statistics specifically were heavily influenced by position and not an accurate representation of output by height ranges.

Figure 4: Line plot of average player statistics for 5kg weight increments (made with Matplotlib)

Figure 4 visualises the effect of NBA players weight on individual statistical categories, by grouping players into 5kg increments and graphing the average per game statistics of that weight group. All three of the statistical categories dramatically increase at a player weight of 135kg and over, reaching optimal statistical performance at the largest data category of 155kg. It can be further noted the average points of players remain relatively consistent from the 60kg to 135kg range before the dramatic increase. Rebounds consistently increases as a player's weight increases and this increase is amplified once a player is over 135kg. Assists initially decrease from the smallest weight groups of 60kg and 65kg until 135kg, where the increases follow the trend of players having statistically increased averages past this point. From this investigation it is concluded that over 135kg is the ideal weight range for players based on individual statistics. Note that an outlier was removed from the data set for this visualisation, Sim Bhullar, as he is the only player in the data set above 155kg (10kg heavier than anyone else and only played a total of three minutes in the NBA). With Bhullar in the data set there was a dramatic decline at the end of the graph (due to his

limited playing time and statistics) which was determined to not be an accurate representation of the findings, therefore was removed from the data set.

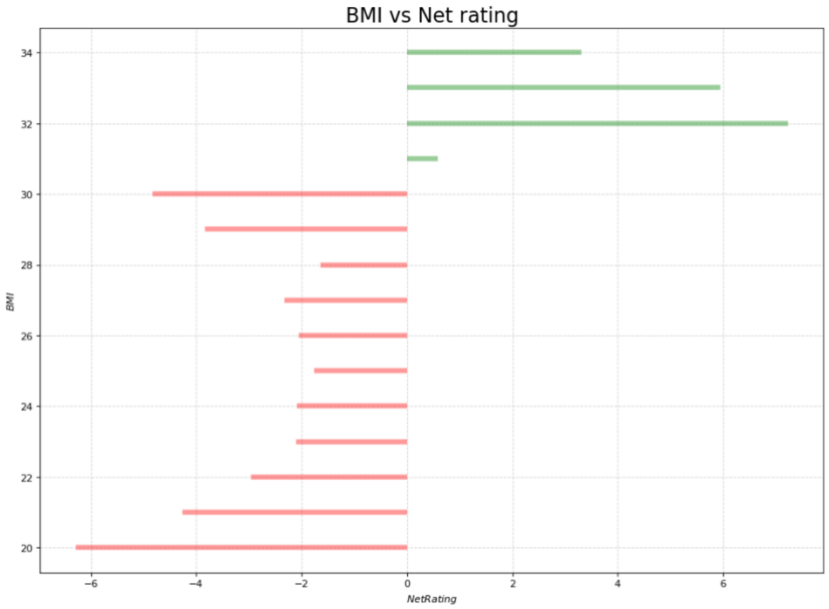


Figure 5: Column Graph, graphing BMI vs Net rating (made with Matplotlib)

The visualisation presented in figure five analyses and compares the impact BMI has on net rating. This investigation builds on the previous as it analyses a player's value to their team, opposed to individual statistics. Net rating is a sports statistic used to measure a player's impact on the game, represented by the difference between their team's total scoring versus their opponent's when the player is in the game[2].

The results highlight the NBA players with the smallest BMI (BMI of 20) have the greatest negative value to their team (net rating of -6.3). It is also observed that players with a BMI of 30 and under, on average have a negative net rating, meaning their team is losing when they are playing. Interestingly, the four largest BMI categories (31, 32, 33 and 34) all have positive impact and are the only categories that do. Specifically, a BMI of 32 is ideal, demonstrated by the highest net rating of +7.2. This investigation highlights the larger massed players (BMI over 30) have the greatest value to an NBA team.

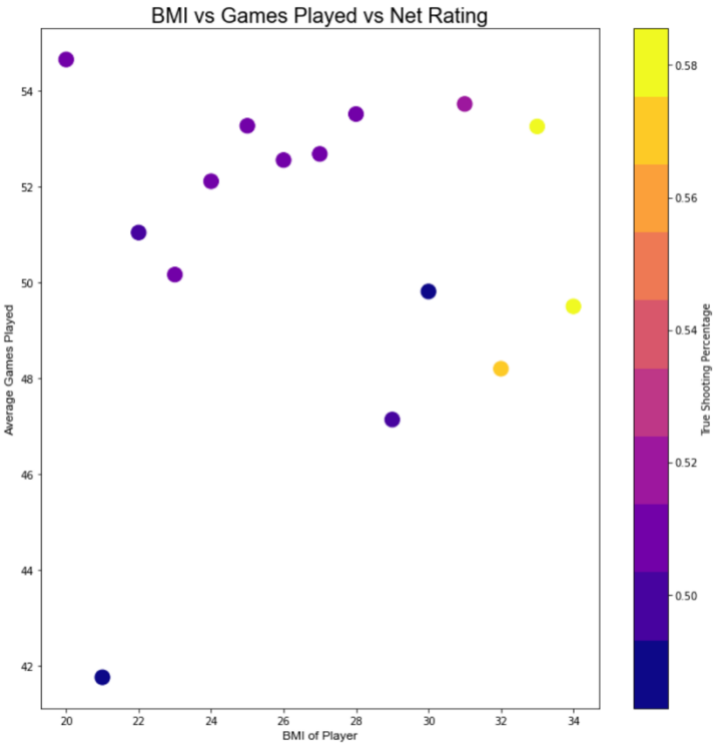


Figure 6: Scatterplot with Hue graph, investigating BMI vs Games Played vs True Shooting Percentage

Figure 6 visualises the impact BMI has on a player's durability throughout a season (games played) and offensive efficiency (true shooting percentage). True shooting percentage measures a player's efficiency at shooting the ball by Comparing how many shot attempts players take against their made shots[3].

The results highlight that there is no correlation between BMI and a player's durability, as the points are scattered at random across the x and y axis.

There is obvious correlation evident in the data when observing the effects of BMI and offensive efficiency. The data presents the four largest BMI index's (31,32,33 and 34) are the four most efficient BMI categories. When looking at the difference between the smallest (20 BMI) players and the largest (34 BMI) there is a 7.8% increase in efficiency For the larger players (20 BMI = 50.1%, 34 BMI = 57.9%). From this investigation, the key takeaway is that players with a BMI over 30 are the most offensively efficient and also there is no trend in the data to suggest BMI has any effect on a player durability throughout a season.

[2]Stat Glossary 2022, NBA Stats, NBA Stats, viewed 9 September 2022, <https://www.nba.com/stats/help/glossary/>.

[3] Ibid

Conclusions and Discussion

Each of these investigations provide unique takeaways which allow for an in-depth conclusion to determine the optimal player body type for the NBA. The investigation into *Points by height range* highlighted that players over 220cm will be the most effective at scoring points, when compared to players of other height categories. The next investigation into *Weight vs Individual Performance Data* presented the findings that players over 135kg dramatically on average have better statistical averages then players who are less then 135kg. When analysing *BMI vs Net Rating* it was found the four largest BMI categories are the most valuable players to a team's success, forming the key takeaway that players with a BMI of 31 and over, on average have the highest positive impact to a team and are the most valuable. The final investigation into *BMI vs Games Played vs Net Rating* provided the insight that there is no correlation in the data to suggest BMI has any influence on a player's durability over a season. This was somewhat surprising, as before the investigation I hypothesised the larger massed NBA players may be more valuable, however may have a decreased durability over a season due to carrying larger amounts of weight which could increase risk of injury. This investigation also highlighted the increased efficiency of the larger players when compared to smaller player, constructing a key takeaway that player with BMI largest four BMIs are the most effective. This adds onto the previous investigation's findings, that not only are the bigger players the most valuable in terms of team performance, they are the most efficient offensively.

These findings contradict the reasoning behind the trend of NBA players getting smaller (figure 1) as clearly the investigation demonstrates the benefits of having players with larger body compositions. However, it must also be noted this investigation did not consider the makeup of a team, as it unreasonable to assume a team of players consisting only of the largest biological factors will be the best team. Another consideration of these findings is that these body types are extremely rare, for example only 0.000038% of the world's population is over 210cm[4] and it was determined the ideal height was over 220cm.

With these factors being considered, this investigation concludes that on average, the best NBA players will be over 220cm and 135kg and have a BMI over 31. These biological makeups of a player are extremely rare, however when there is a player in the NBA draft who meet these requirements, it would be in a team's best interest to heavily consider this type of player.

An example of the ideal NBA player build is NBA hall of fame player Yao Ming who stands at 229cm tall and weighed over 140 kg throughout his career. Ming ranks 21st in NBA history for net rating and 22nd in player efficiency all time[5].

Figure 7: Yao Ming, an example of the ideal NBA player.

References

Justinas Cirtautas 2022, NBA Players, Kaggle.com, viewed 9 September 2022, <https://www.kaggle.com/datasets/justinas/nba-players-data/metadata>.

Percentage Calculator Staff 2013, *How Many 7 Footers Are in the World?*, Percentagecalculator.mes.fm, viewed 9 September 2022, <https://percentagecalculator.mes.fm/interesting-facts/how-many-7-footers-are-in-the-world>.

[1]Stat Glossary 2022, NBA Stats, NBA Stats, viewed 9 September 2022, <https://www.nba.com/stats/help/glossary/>.

Yao Ming | Biography & Facts | Britannica 2022, Encyclopædia Britannica, viewed 9 September 2022, <https://www.britannica.com/biography/Yao-Ming>.

[4] Percentage Calculator Staff 2013, *How Many 7 Footers Are in the World?*, Percentagecalculator.mes.fm, viewed 9 September 2022, <https://percentagecalculator.mes.fm/interesting-facts/how-many-7-footers-are-in-the-world>.

[5] Yao Ming | Biography & Facts | Britannica 2022, Encyclopædia Britannica, viewed 9 September 2022, <https://www.britannica.com/biography/Yao-Ming>.

Appendix

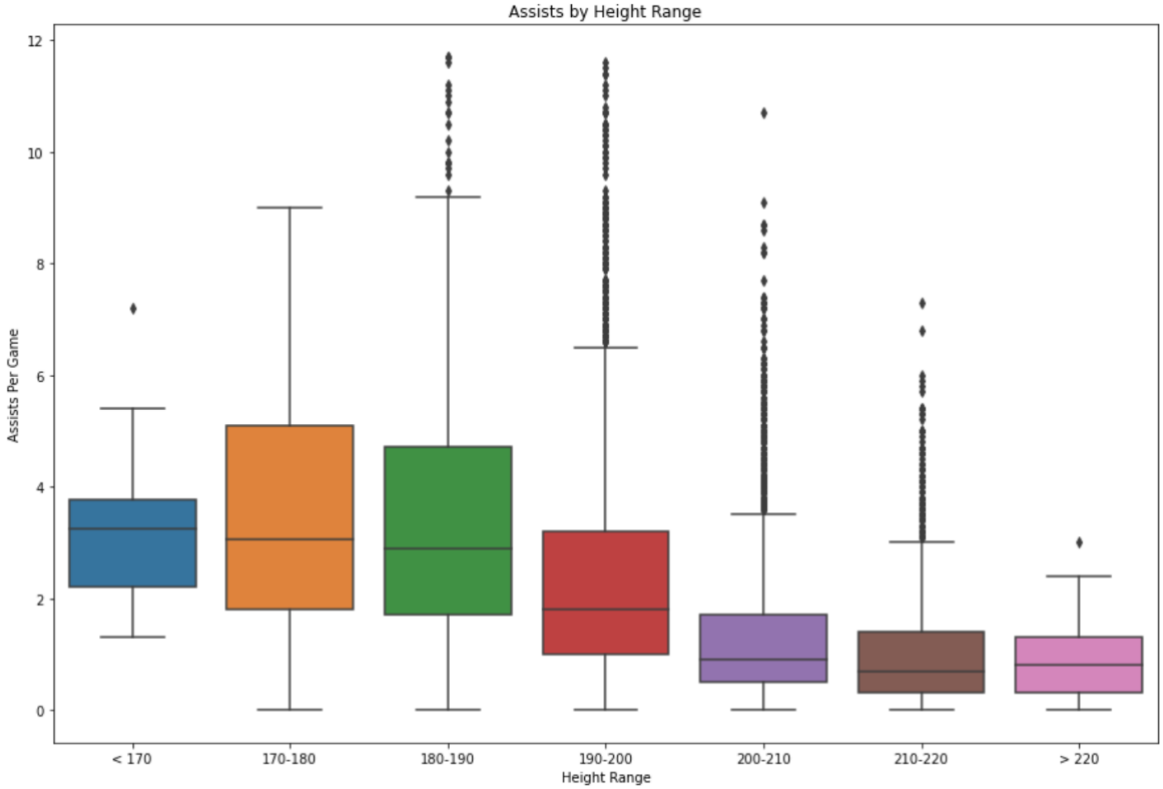
Addressing Feedback:

A piece of feedback received from my original proposal of my project, was a suggestion in a way to analyse my data source for visualisations purposes and evaluation of reliability of conclusions in final report. The suggestion was to investigate the changes in players body compositions from 1996-2020 and investigate whether it is consistent or there have been changes. This suggestion was suggested by three people, including my project tutor, therefore leading me to heavily address this suggestion. To address the feedback, the first investigation I analysed was based on this suggestion: analysing the trends of height and weight from 1996 to 2020. I believe this was a very important part of my final investigation, as initially the trend of players getting smaller made me propose the question "as players are getting smaller, are the smaller players statistically better than larger players?".

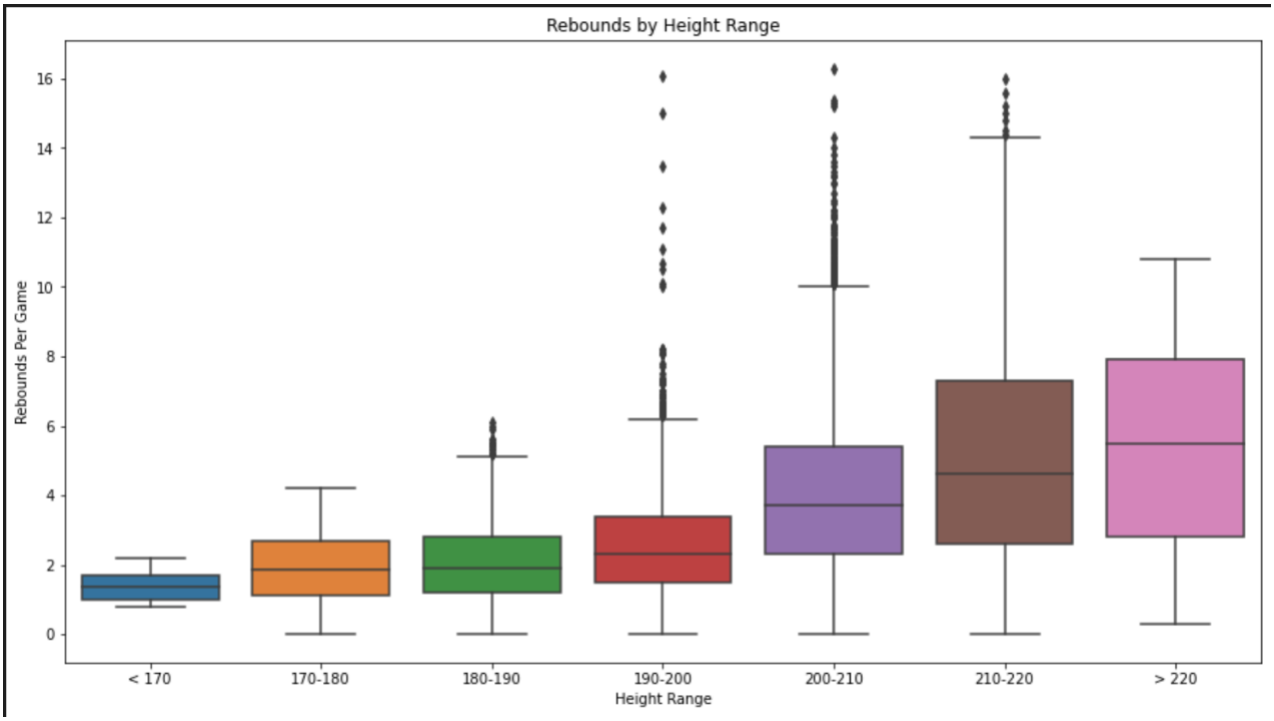
Another piece of feedback received was the concern around my question and data set having to many avenues for analysis and was presented by two of my classmates. I agreed with this feedback and revised my original investigation question which was broader than the final one. Initially, I aimed to investigate all aspects of a player including country of origin, college attended, position in draft to determine if these factors also influence a players NBA career. However, after receiving this feedback, I decided it would be best to narrow my question to just analyse body composition, as I wanted to have more in detail findings in a specific category.

I also received the suggestion to look into whether different body types take longer to reach their peak and play well, and if some body types are good initially, but fall off quickly. This suggestion would definitely be an interesting investigation to conduct and would add another level of analysis to my findings, however due to the time constraints of this assignment, I decided to leave this out of my investigation. In addition to this suggestion, I also received the idea to analyse American players compared with international players. Again, this would be a great investigation, but due to time constraints I did not address this feedback.

Appendix Figures



Appendix 2: Box and Whiskers plot of average assists per game for different height ranges per game (made with seaborn)



Appendix 2: Box and Whiskers plot of average rebounds per game for different height ranges per game (made with seaborn)