Do the biggest players dominate the sport of Basketball?

Sid Regmi

The Question

Do the biggest players dominate the sport of Basketball?

Most people who are not familiar with the sport of basketball will tell you that height and weight are deciding factors in a player. Taller players might dominate the sport, but that is only true in amateur leagues. In the NBA, players are performing at near-perfect levels of play. There has been a wide range of all-stars.

Methodology

By looking at 20 years of data, we will see the different eras of the NBA. The goal is to see how dominate the bigger players are by looking at the net rating of the players. We can also do a compare and contrast by taking players for each other and comparing them by doing a tall vs short. For this analysis we will only look at the who is dominating the league. To do this we will compare the net rating of a player.

1) Late Transitional Era (2002 - 2009)
2) Early Modern Era (2009 - 2013) #I made this up
3) Modern Era (2013-now)

We can also look at the changes in the sports philosophy by looking at different seasons in each era.

Looking at 2002 vs 2016 and the 2022 season can help us notice the trends and make a prediction on if the biggest players will dominate the game in the future. The first step in all of this is to find how big the biggest players are.

Why these years? 2002 was a know as the start of the end of the Transitional Era. 2016 was when the new style of basketball became popular and every team incorporated tactics creating more three point opportunities. 2022 was the most recent season so we can can see how much the game has changed since 2016.

Explaining net rating

Net rating is what I used to determine if a player is dominating the league. Net rating is the offensive rating minus the defensive rating. Net rating tells us how much better or worse the team is when a specific player is on the court. For this analysis, a player with a net rating above +10 will be considered a dominating player. Net rating is usually in the range of +15 to -15. A +15 player is usually the MVP. A -15 player should not be allowed to play in the league.

Getting the data

The data was downloaded from kaggle and then uploaded to the Notebook.

[https://www.kaggle.com/datasets/justinas/nba-players-data/data]

For now i just need data for 2002, 2016, and 2022 seasons.

In [4]:
```
#I need to get 3 different tables for each season that i want to compare.
#2022 because it was the most recent and some big players dominated
#2016 because it was when most of the bigger players didn't do much
#2002 because it was when the biggest players were the most wanted
season_2022 = data.where("season", "2021-22").drop(1,2,5,6,7,8,9,11,13,12,15,16,17, 18, 19)
season_2016 = data.where("season", "2016-17").drop(1,2,5,6,7,8,9,11,13,12,15,16,17, 18, 19)
season_2002 = data.where("season", "2002-03").drop(1,2,5,6,7,8,9,11,13,12,15,16,17, 18, 19)
season_2002.show(5)
season_2016.show(5)
season_2022.show(5)
```

player_name	player_height	player_weight	gp	net_rating	season
Popeye Jones	203.2	113.398	26	-8.8	2002-03
Michael Doleac	210.82	118.841	75	-5.3	2002-03
Michael Finley	200.66	102.058	69	12.6	2002-03
Michael Jordan	198.12	97.9759	82	0	2002-03
Predrag Savovic	198.12	102.058	27	-9.8	2002-03

... (423 rows omitted)

player_name	player_height	player_weight	gp	net_rating	season
Metta World Peace	200.66	117.934	25	-17.7	2016-17
Matthew Dellavedova	193.04	89.8112	76	-1.1	2016-17
Maurice Harkless	205.74	99.7902	77	3.4	2016-17
Maurice Ndour	205.74	90.7184	32	-1.4	2016-17
Meyers Leonard	215.9	115.666	74	-2.6	2016-17

... (481 rows omitted)

player_name	player_height	player_weight	gp	net_rating	season
Usman Garuba	203.2	103.873	24	7	2021-22
Vlatko Cancar	203.2	107.048	15	0.8	2021-22
Vit Krejci	203.2	88.4504	30	-12.8	2021-22
Victor Oladipo	193.04	96.6151	8	-1.4	2021-22
Vernon Carey Jr.	205.74	122.47	7	-47.6	2021-22

Filtering the data

For now i just need data for 2002, 2016, and 2022 seasons.

I then filtered the data so that a player has played more than the average. If a player only has a few games played they might skew the data. I don't want that .

```
#this is what i used to find the average amount of games
avg_gp_2002 = int(season_2002.column('gp').mean())
avg_gp_2016 = int(season_2016.column('gp').mean())
avg_gp_2022 = int(season_2022.column('gp').mean())


season_2002_filtered = season_2002.where("gp", are.above(avg_gp_2002))
season_2016_filtered = season_2016.where("gp", are.above(avg_gp_2016))
season_2022_filtered = season_2022.where("gp", are.above(avg_gp_2016))

season_2002_filtered.show(5)
season_2016_filtered.show(5)
season_2022_filtered.show(5)
```

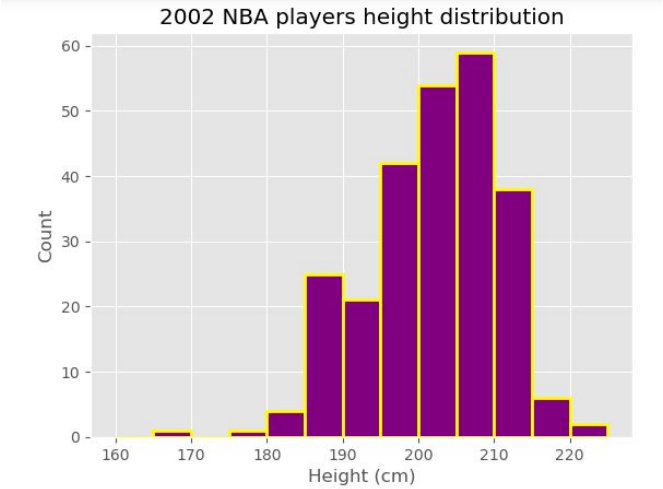
player_name	player_height	player_weight	gp	net_rating	season
Michael Doleac	210.82	118.841	75	-5.3	2002-03
Michael Finley	200.66	102.058	69	12.6	2002-03
Michael Jordan	198.12	97.9759	82	0	2002-03
Michael Redd	198.12	97.5223	82	2.6	2002-03
Predrag Drobnjak	210.82	123.377	82	-3	2002-03

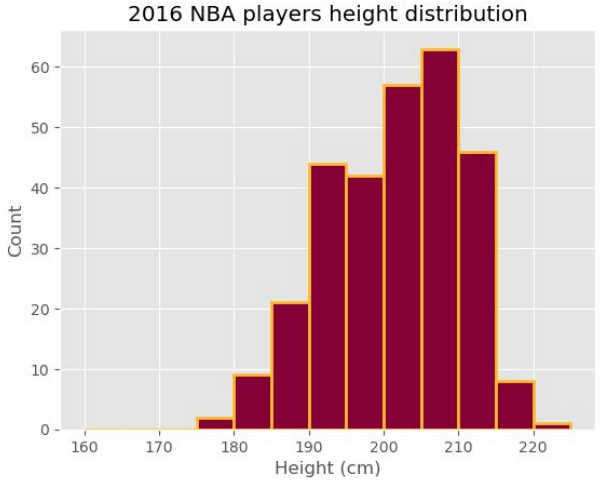
... (250 rows omitted)

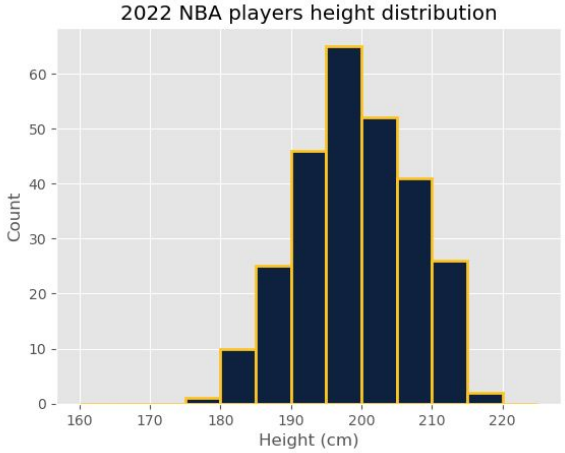
player_name	player_height	player_weight	gp	net_rating	season
Matthew Dellavedova	193.04	89.8112	76	-1.1	2016-17
Maurice Harkless	205.74	99.7902	77	3.4	2016-17
Meyers Leonard	215.9	115.666	74	-2.6	2016-17
Marc Gasol	215.9	115.666	74	1.4	2016-17

Histograms for height

Taking that data. I can create histograms for height. I found that the tallest player are 210 cm (about 6ft 11in).

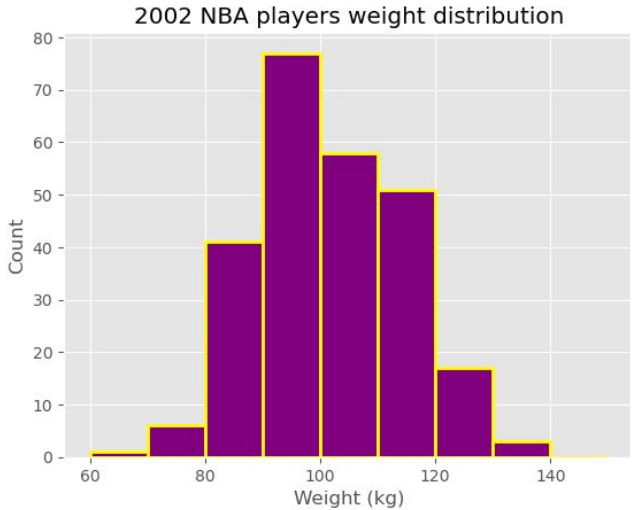


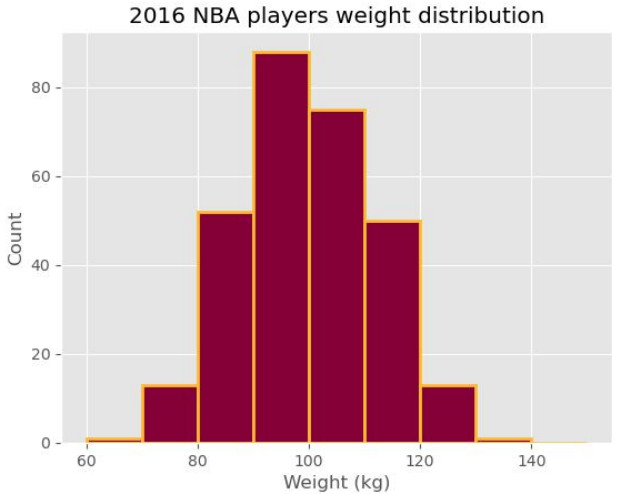


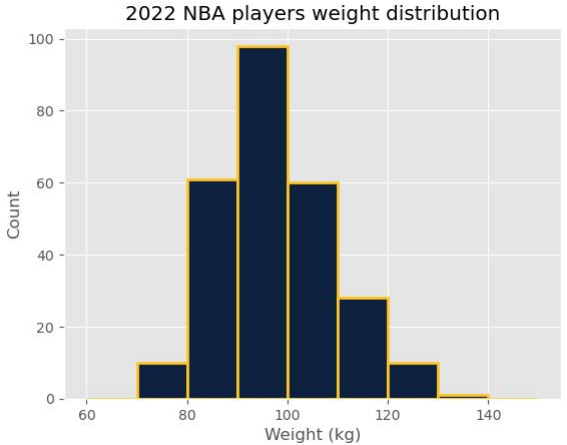


I did the same with weight

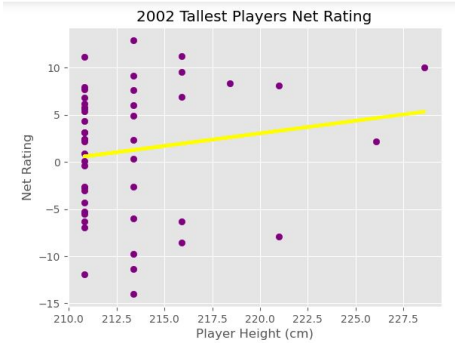
I found that the biggest players are more than 120 kg (264 lbs).

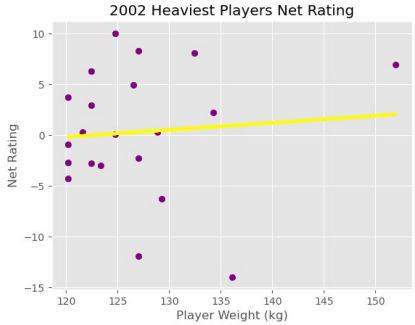


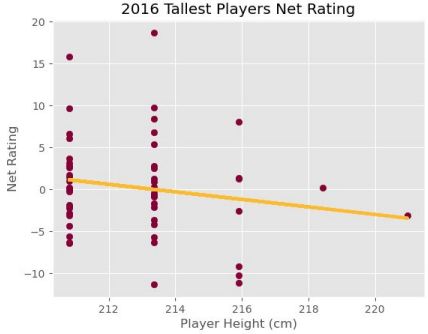


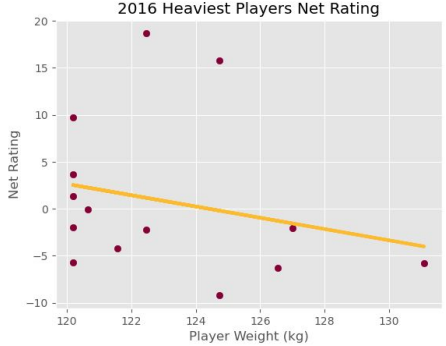


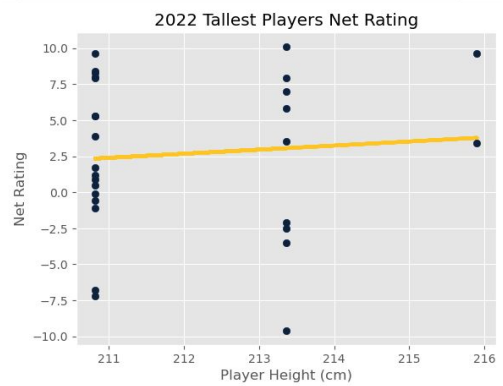
Find the Net Rating of the tallest and heaviest players

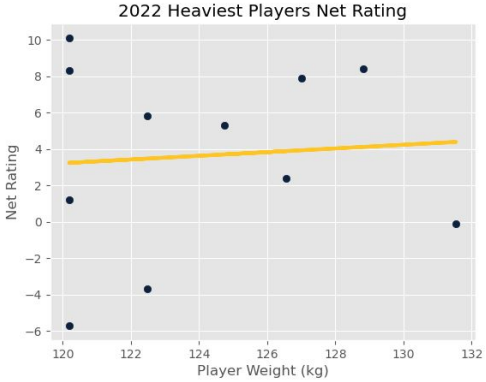












2002: Tallest and Heaviest Players Dominated

- Tallest players had above a +5 net rating, some reaching +10.
- Heaviest players also showed dominance with strong net ratings.

2016: Decline for Tallest and Heaviest Players

- Negative slope in both plots indicates a struggle for the tallest and heaviest players.
- Two outliers identified for further individual player analysis.

2022: Recovery for Tallest and Heaviest Players

- Tallest and heaviest players have regained performance.
- While showing improvement, they no longer dominate the league as in 2002.

What happens when we look at individual players in each Era?

I separated each era and assigned a tall player vs a short player. LeBron James will be the standard good player for each Era as he played the last 20 years and was dominant for most of his career.

Late Transitional Era:

- Tall Player: Yao Ming dominated.
- Short Player: Dwyane Wade made a name for himself.

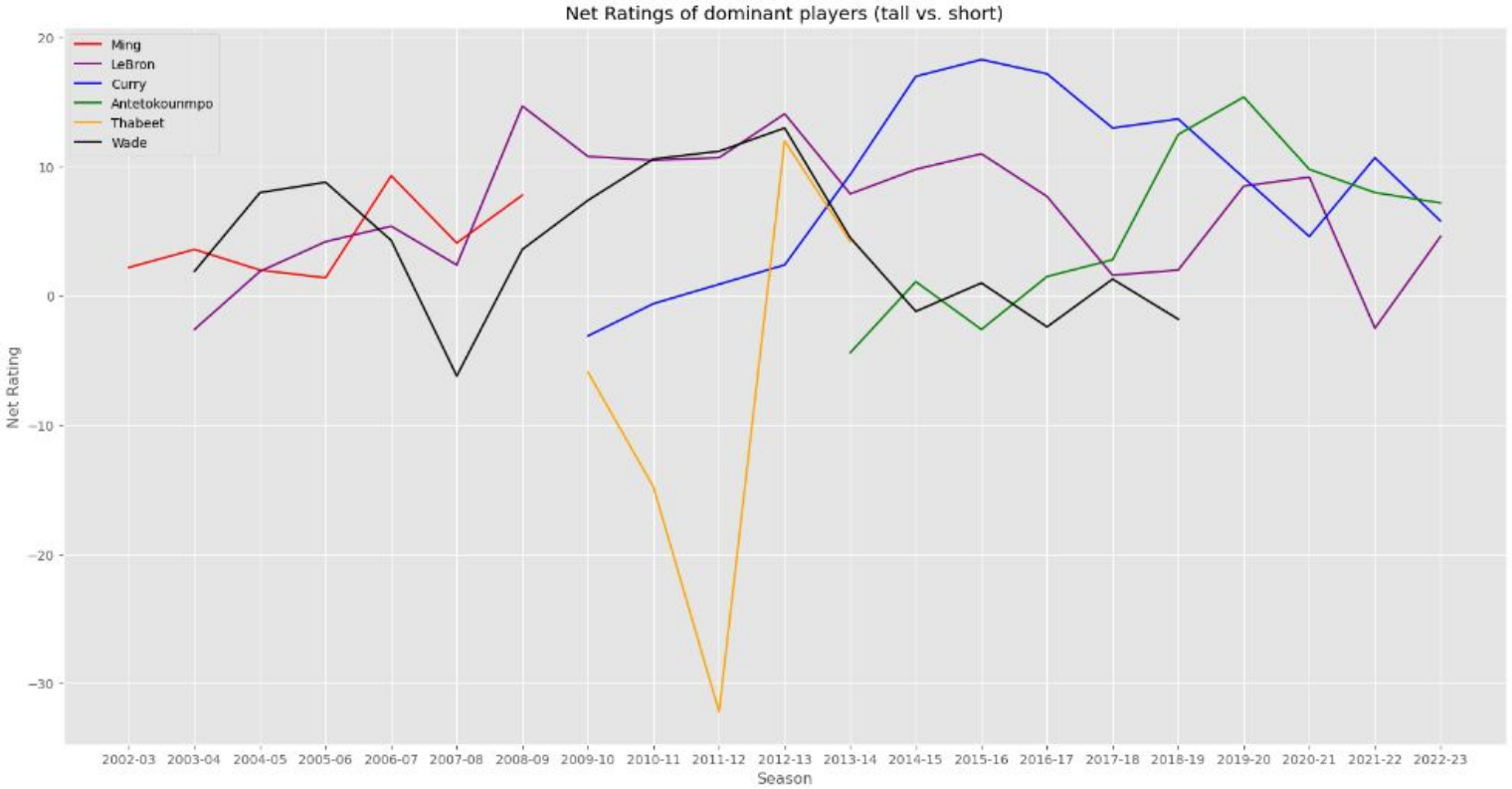
Early Modern Era:

- Few tall players, Hasheem Thabeet notable.
- Overall, tall players saw less court time.
- Short Player: Dwyane Wade dominated during this period.

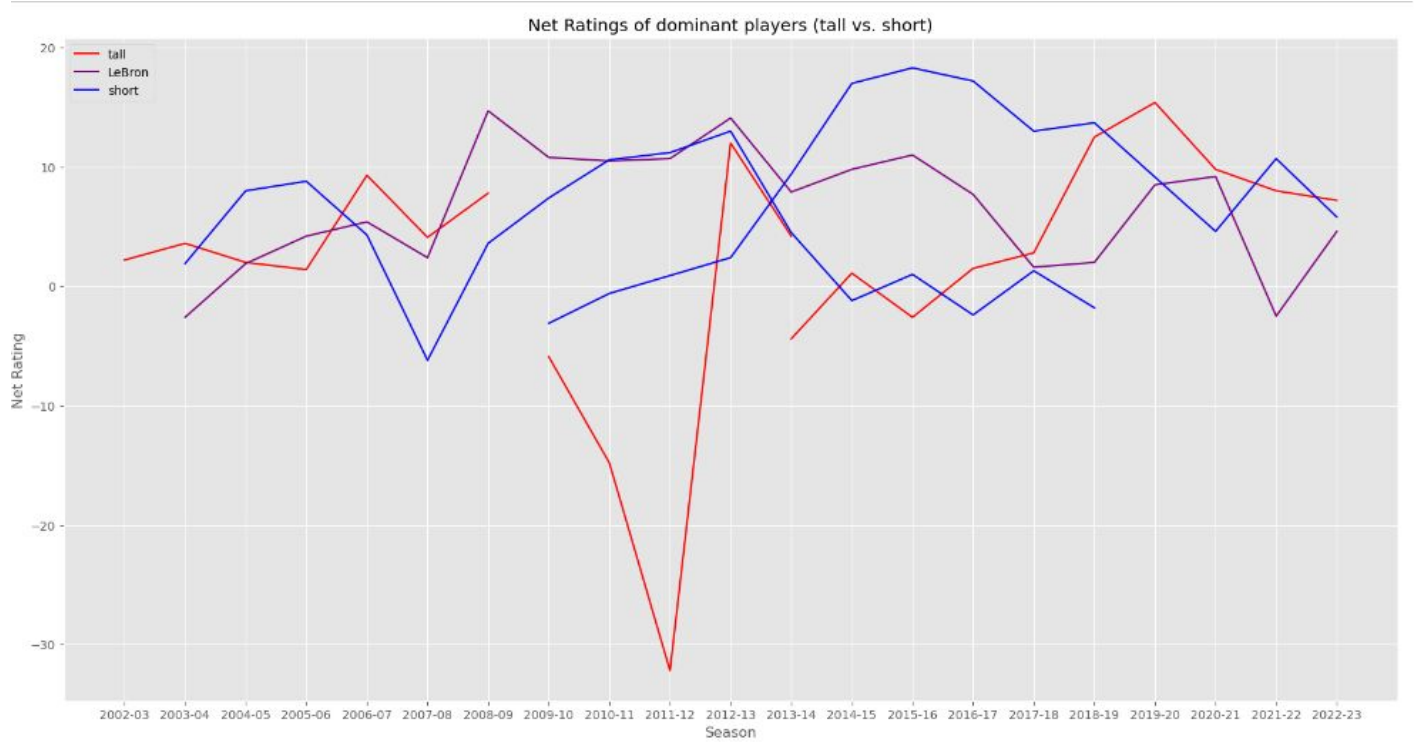
Modern Era:

- Tall Player: Giannis Antetokounmpo.
- Short Player: Stephen Curry.
- Both dominating the sport, representing the "new school" of basketball.
- Overlap with early Modern Era as these players started in that period.

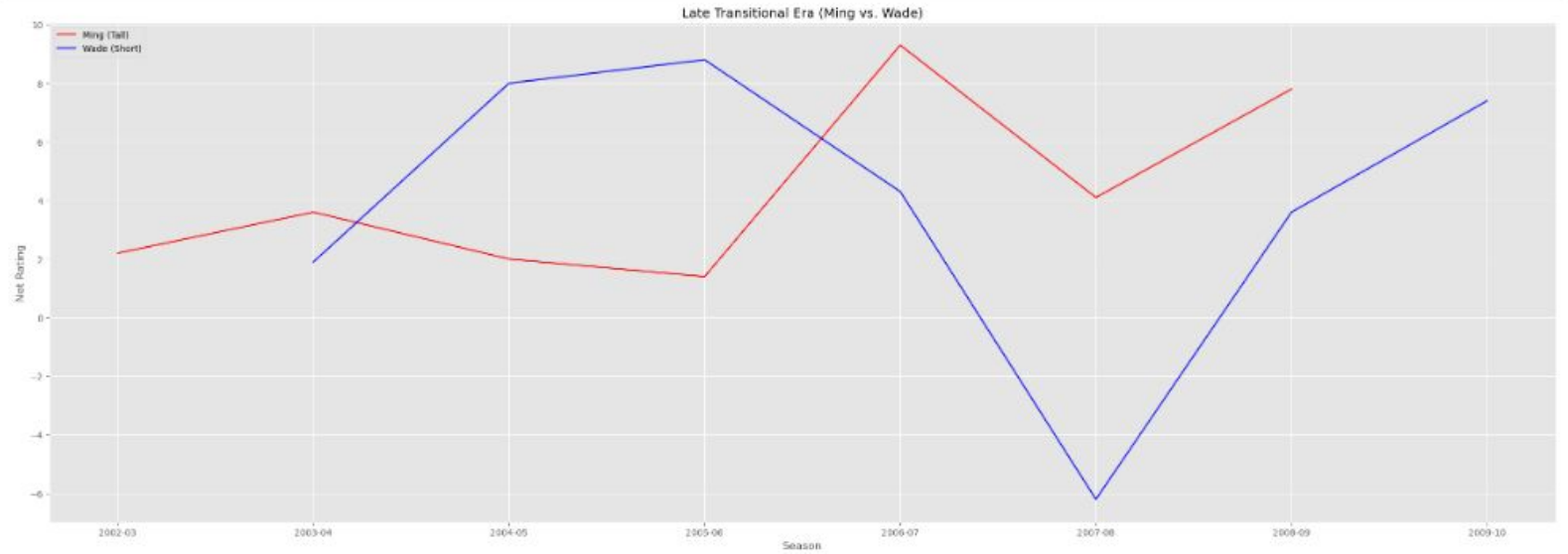
Messy plot. A lot happening. Let's simplify.



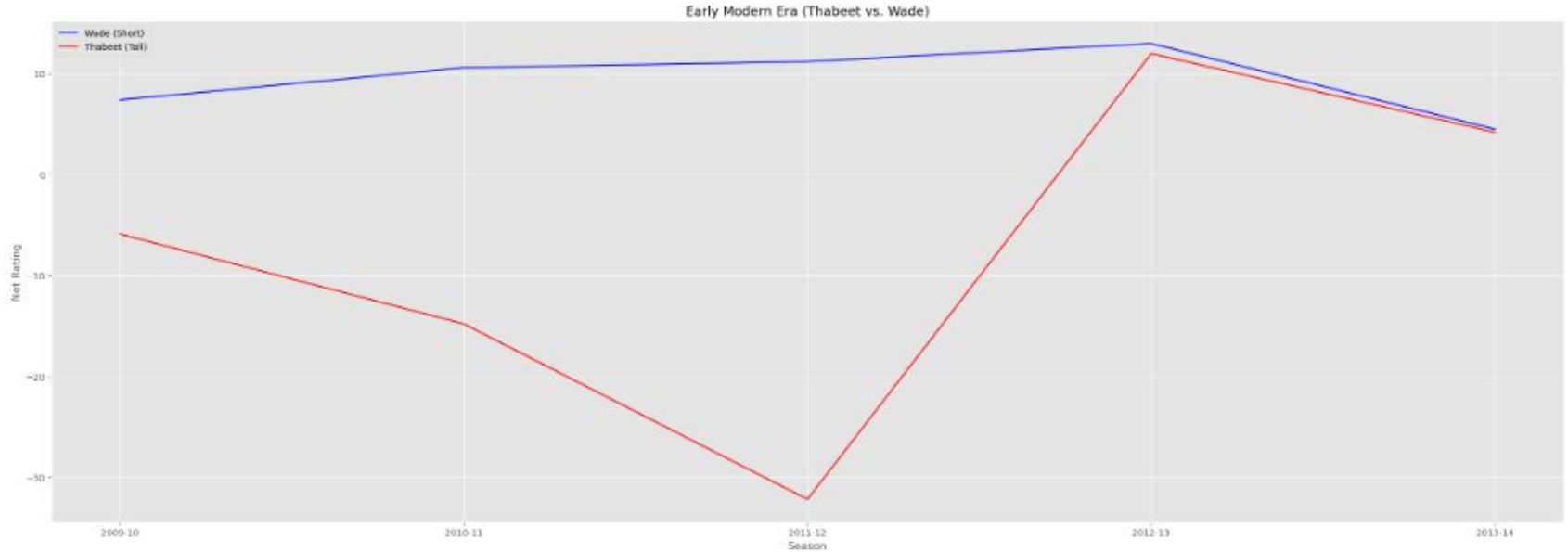
Separating players in Tall and Short. More simplification?



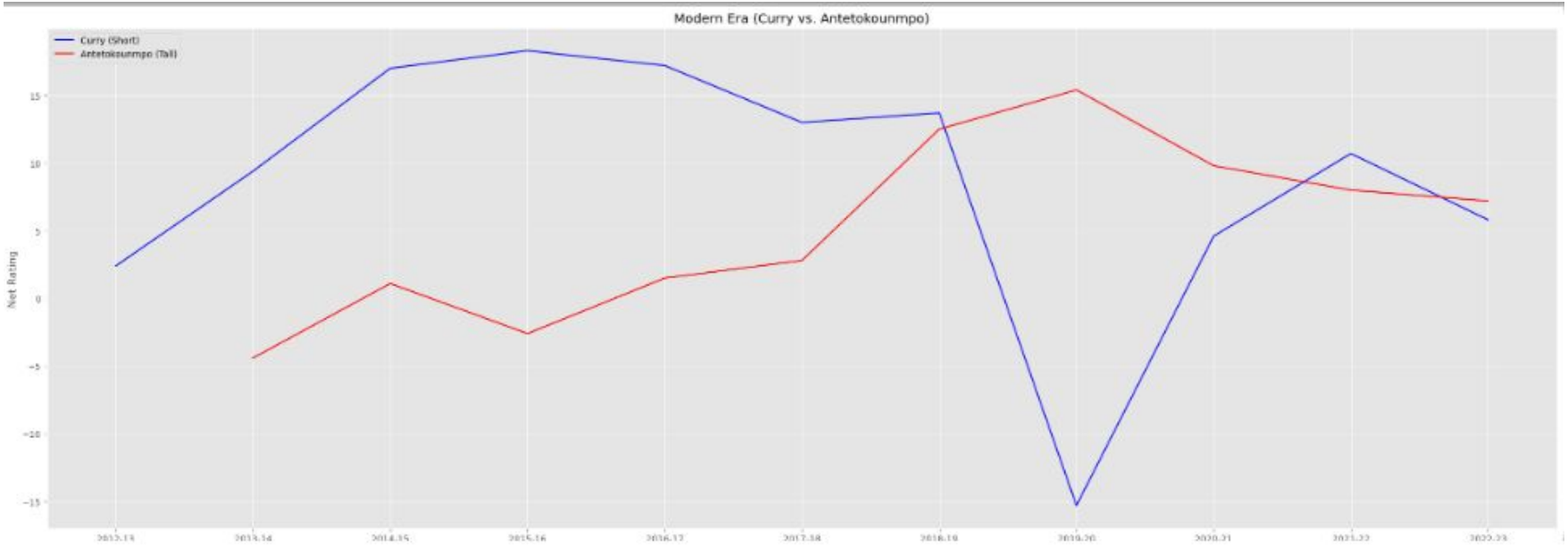
Late Transitional Era (Ming vs. Wade)



Early Modern Era (Thabeet vs. Wade)



Modern Era (Curry vs. Antetokounmpo)



Conclusion

Taller Players' Performance Trends:

- Strong performance for taller players in 2002.
- Decline in performance for 7-foot-tall players in subsequent seasons.
- Slight comeback observed for taller players in 2022.

Comparison of Tall vs Short Players:

- Tall players never fully dominated the sport.
- Challenges for tall players in maintaining long-lasting careers.
- Struggles noted during the Early Modern Era.
- Positive shift observed in 2022, indicating taller players making a more significant impact.

Are we in a new Era?

- Analysis suggests a potential shift in the NBA landscape, but more analysis on this is needed.
- Resemblance to the Transitional Era, with taller players improving their long-range shooting.
- Anticipation of taller players regaining dominance in the near future.

Questions? Concerns?