The Recipe for Success in the NBA

The Late Round Sleepers: Alex Jackson, Evan Finley, Aliya Kamran, Jack Griffin

Defining the data set

- We investigate what makes NBA players successful in the context of team play
- We explore the NBA-Raptor data set. RAPTOR is a player efficiency rating system. Each observation is an NBA player with variables about their statistics
- The data set has 808 observations and 23 variables with our investigation focusing on the following variables

Key Variables:

player_name	Player name
season_type	Regular season (RS) or playoff (PO)
team	Basketball-Reference ID of team
raptor_offense	Points above average per 100 possessions added by player on offense, using both box and on-off components
raptor_defense	Points above average per 100 possessions added by player on defense, using both box and on-off components
raptor_total	Points above average per 100 possessions added by player on both offense and defense, using both box and on-off components
war_total	Wins Above Replacement between regular season and playoffs
pace_impact	Player impact on team possessions per 48 minutes

What we aim to investigate

1. Does the average offensive and defensive raptor for players differ between the regular season and post season?

2. Does the average raptor differ greatly from the average WAR between members of the All NBA teams?

3. How are the variables pace impact, raptor total, and WAR total impacted by player type?

4. Do better defensive players have a better WAR in the playoffs than the regular season?

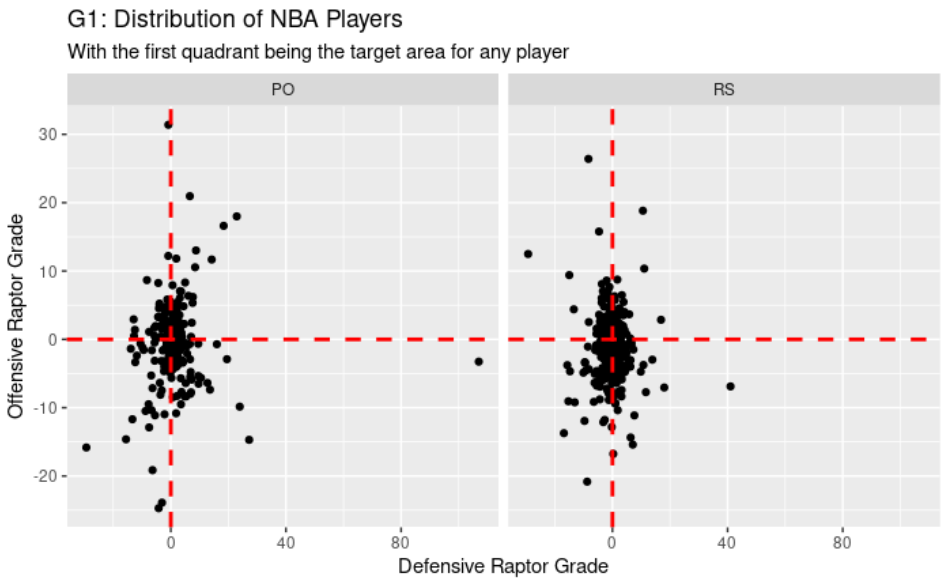
5. Do players with better offensive capabilities have greater impact on team possessions per 48 minutes as compared to players without offensive capabilities?

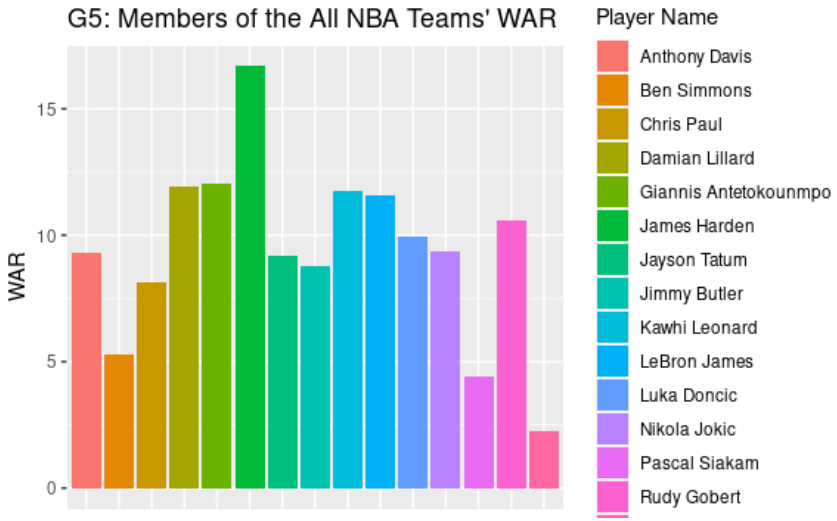
6. Does a team's average raptor correlate to that team's average winning percentage?

7. Does a team's average WAR correlate to that team's average winning percentage?

8. Which statistic, total raptor or total WAR, is better at indicating player success?

Here is what we found #1



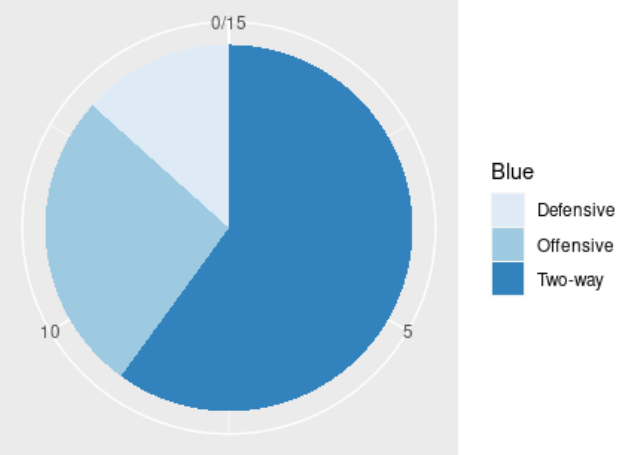


- In both graphs, there seems to be a high density of data point around the axis, meaning that the majority of NBA players have both an offensive and defensive raptor around zero.

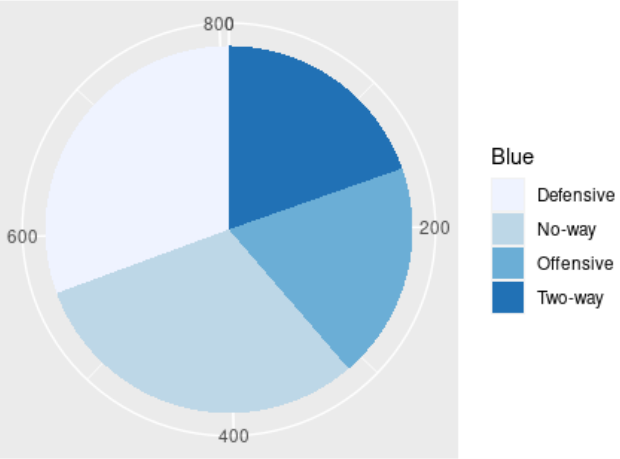
- It is not obvious who the league MVP was (Giannis Antetakounmpo), due to the fact that James Harden is towering above everyone else. Harden finished 3rd in league MVP voting, which is peculiar because based on this metric, he would have been by far the most valuable player. Antetakounmpo, on the other hand, barely reaches the title of second highest WAR and is average among other All-NBA players.[1] *This raises the question of why Antetakounmpo was MVP instead of Harden?*

Here is what we found #2

G2: Two-Way players have the highest representation on the Al



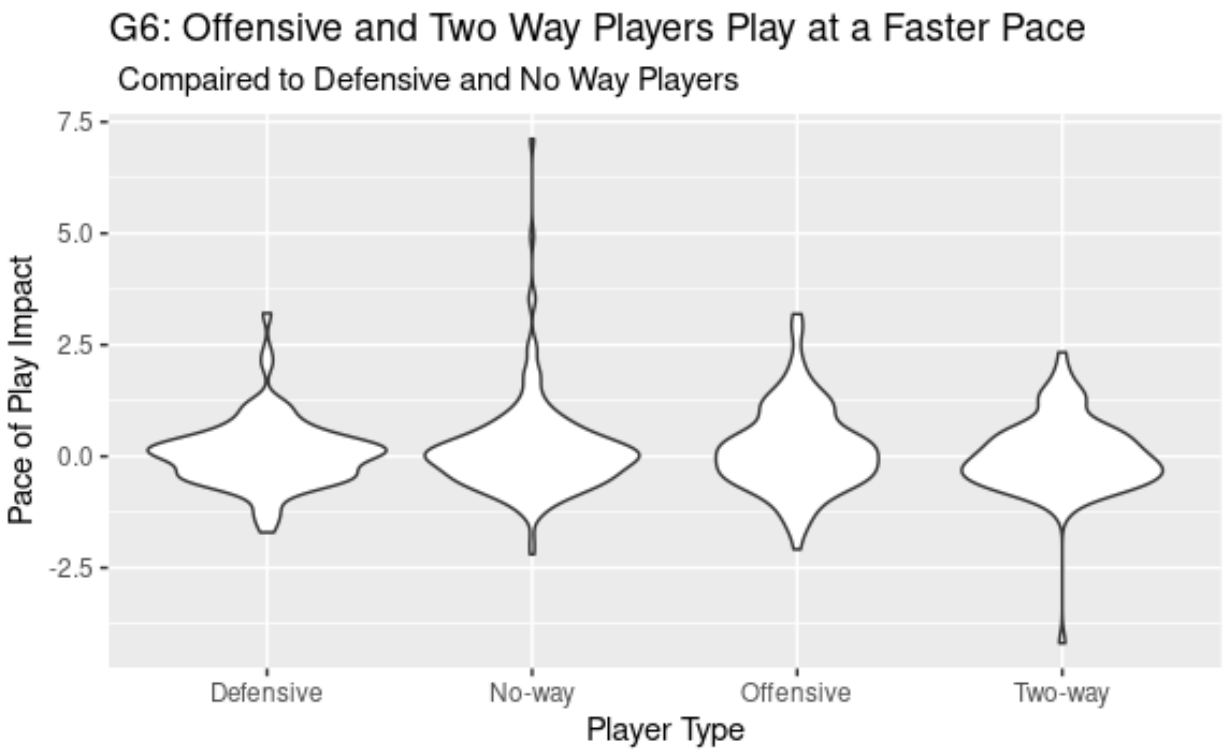
G3: Defensive and No-way Players make up the majority of the



- The first pie chart looks at the 15 All-NBA players. Two-way players dominate (8/15).
- In comparison, the pie chart of the entire NBA shows that two-way players have nearly the smallest proportional representation.

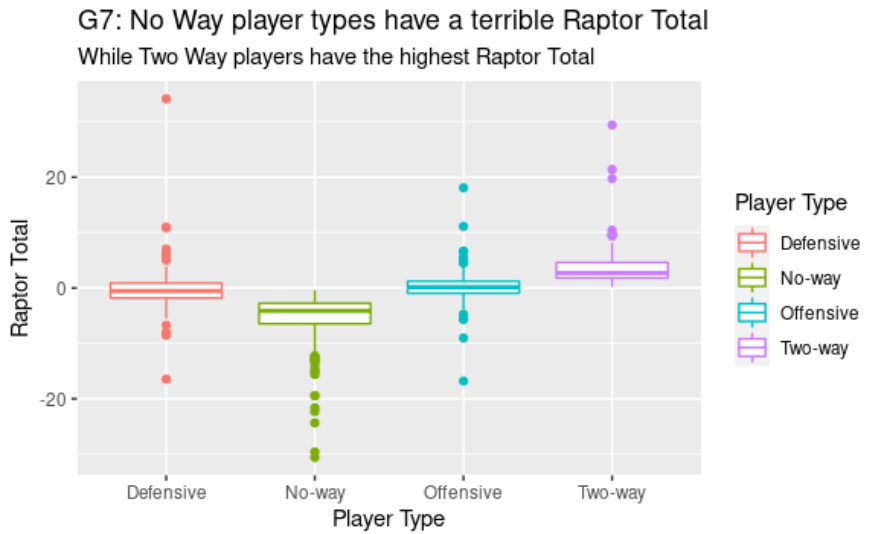
- What this signifies is that two-way players are a more "elite" player type and find more success in the NBA being elite on both ends of the floor, offensive and defensive.
- Defensive players have a small representation of All-NBA players, yet the most of any type in the NBA, so that signifies the difficulty of being the best in the league as a defensive player.

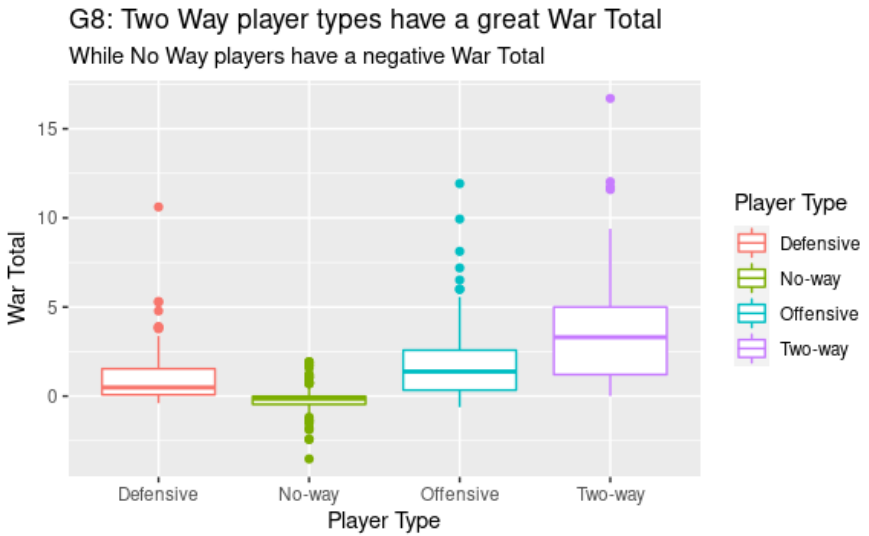
Here is what we found #3



- The offensive and two-way players often look to score and play at a fast pace, which this violin plot clearly shows due to the larger representation of those players above the 0.0 line.
- The defensive players want to slow the game down and force the offensive player to be uncomfortable and play slowly. The no way players are often role players, meaning that they do not affect the pace of the game enough to make it much faster.
- There are a few no way players and two-way players that greatly separate themselves from their respective trends

Here is what we found #4





- No way players have an incredibly low raptor total. This makes sense, but visualizing it makes it much more clear how low their total raptor really is. It also effectively shows how important two-way players are, as they had the highest total raptor. In addition, it is interesting to note that defensive players had a higher total raptor than offensive players

- Two-way players have a wildly high WAR. Two-way players generate higher WARs typically because they excel on both ends of the floor.
- The no way players have a negative WAR total because their impact on the game hurts their teams. Offensive and defensive players have a very close total WAR value, with defensive players slightly ahead.

Here is what we found #5

Effect of Offensive and Defensive Raptor on WAR in the regular season

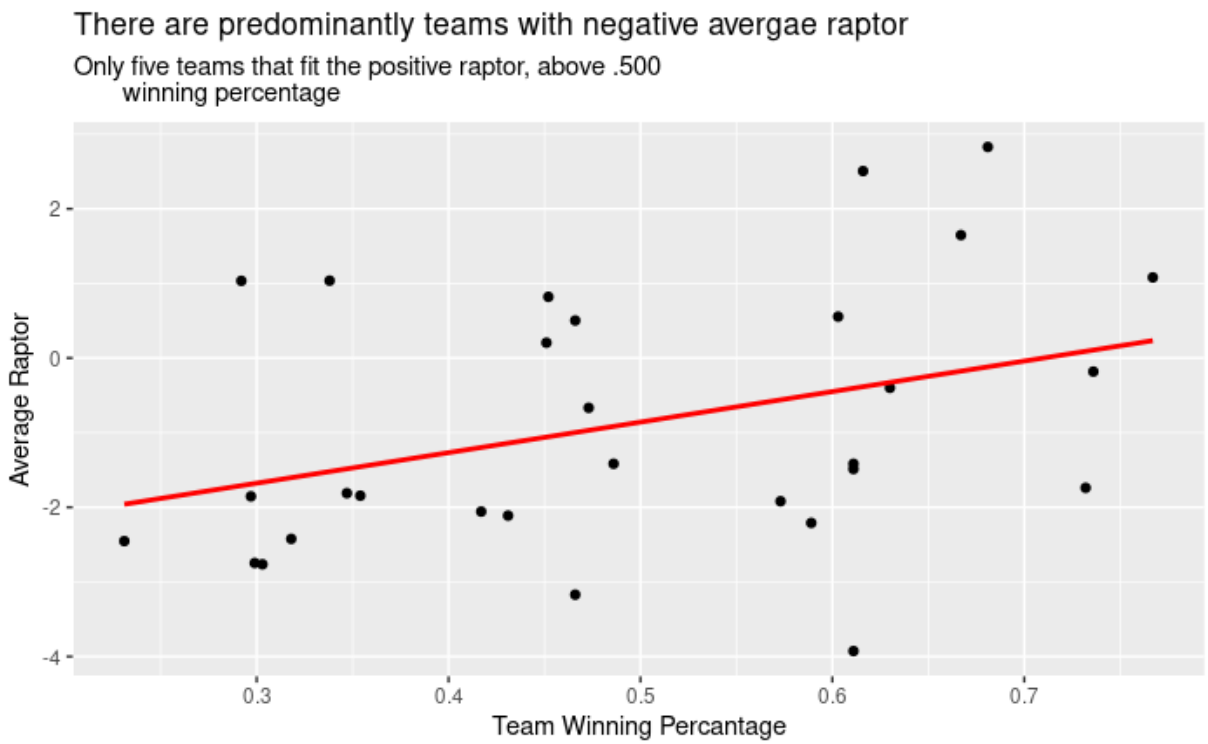
In the linear model test we conducted, the results were as follows:

Expected regular season WAR = 1.02 + 0.026 * Defensive RAPTOR + 0.125 * Offensive Raptor

- This means that offensive efficiency is actually more important to WAR (wins above replacement) than defensive efficiency. The *nature of the league has definitely changed*. Defense used to be the most critical element to winning a game[2], but now with such high scoring games, the importance of efficient offense seems to outweigh the importance of efficient defense.
- After performing two correlation tests, we found that offensive raptor had a higher correlation with WAR than defensive raptor. The issue is that the conditions for linearity, equal variance, and normality all not met. Additionally, the R-squared test showed that only roughly 9.4% of the variability in the WAR of players can be explained by raptor offense and raptor defense for the regular season.

Here is what we found #6

Team Average Raptor and Winning Percentage

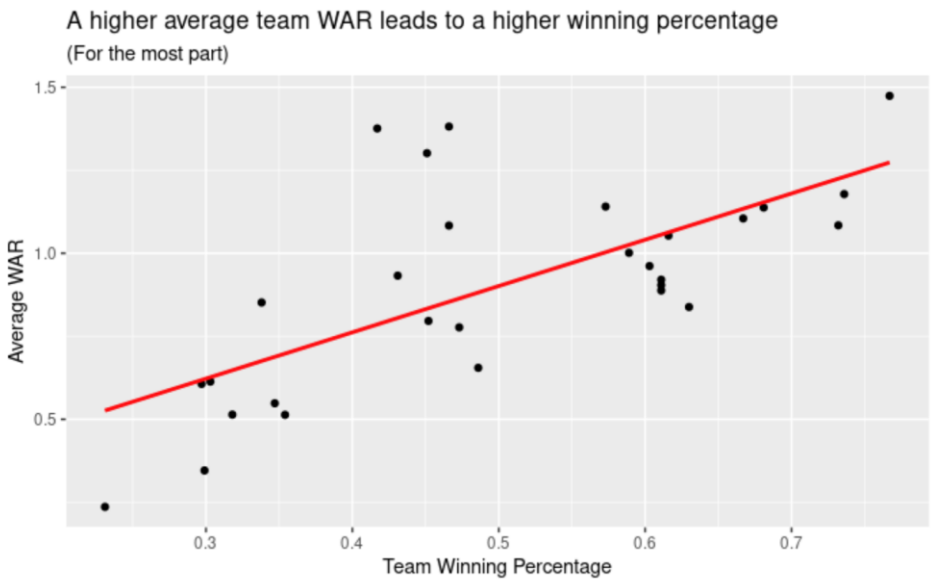


Expected team winning percentage = 0.52338 + 0.03233 (team average raptor)

- For every single unit increase in team average raptor, it is expected there will be an increase in winning percentage of 3.23%.
- We tested for equal variance and linearity first. Overall, we observed a roughly linear and equal variance relationship in our data. Next, we tested for normality using a QQ plot. This came out less fitting, but we realized there are only 30 data points so it is expected that the tests won't exactly fit.

Here is what we found #7

Team Average WAR and Winning Percentage



Expected team winning percentage = 0.2012 + 0.3326(team average raptor)

- For every single unit increase in team average WAR, it is expected there will be an increase in winning percentage of 33.26%. The large increase of win percentage is due to the fact that the team average WARs are very small and very similar, a one point jump would create a tremendous impact, so the 33% increase to the win percentage is understandable.

- We tested for equal variance and linearity first. Overall we observed a roughly linear and equal variance relationship in our data points. Next, we tested for normality using a QQ plot. This came out less fitting than the raptor test, but the normality is close enough to pass the test. With only 30 datapoints, it is hard to tell an exact fit from these tests, so the overall linear model should be used with slight caution.

Here is what we found #8

Top 10 Players in terms of Raptor and WAR

war_reg_season <dbl>	player_name <chr>
16.705593	James Harden
12.031285	Giannis Antetokounmpo
11.926011	Damian Lillard
11.730348	Kawhi Leonard
11.591113	LeBron James
10.614410	Rudy Gobert
9.936519	Luka Doncic
9.380334	Nikola Jokic
9.303934	Anthony Davis
9.158699	Jayson Tatum

raptor_total <dbl>	player_name <chr>
103.72756	Alize Johnson
40.90943	Frank Mason
34.97245	Vincent Poirier
34.13053	Marques Bolden
30.55619	Patrick Patterson
29.41299	Max Strus
27.56142	Vic Law
25.89373	Bol Bol
21.76403	BJ Johnson
21.39367	Tyler Zeller

The top 10 players in terms of WAR are all household names and Allstars. They are some of the best players in the NBA. Conversely, the top 10 players in terms of Raptor are not well-known players. This being the case, I would argue that WAR is the more accurate variable whenever looking for the most valuable players. WAR shows the players that are most important to a team's win above replacement, meaning that the best players would have the highest WAR.

Summary of Findings and Suggestions for Improvement

Findings

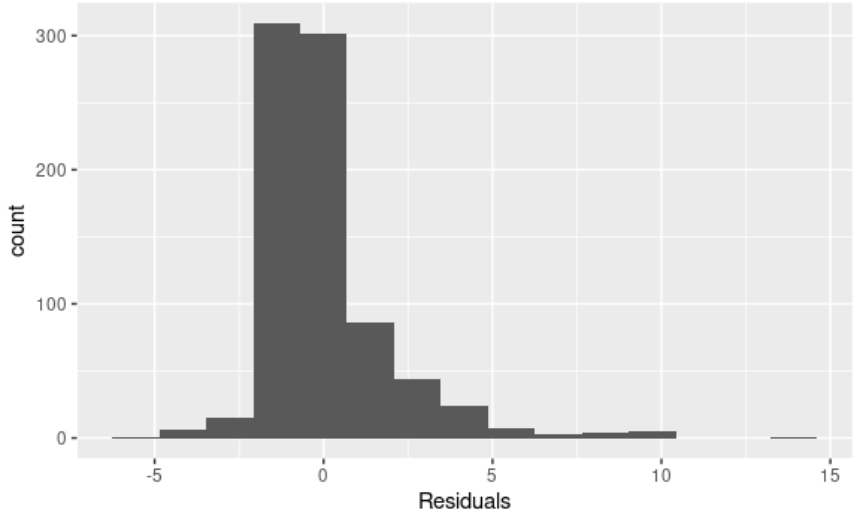
- Before we analyzed this dataset, we hypothesized that defensive players were the most valuable type of player to a team's success.
- However, based on our analysis, we concluded that offensive players are the most important player type to success
- Moreover, two-way players are invaluable to a team, and are key to a team's overall success.
- The majority of NBA players are no way players, but the most valuable players are two-way players.

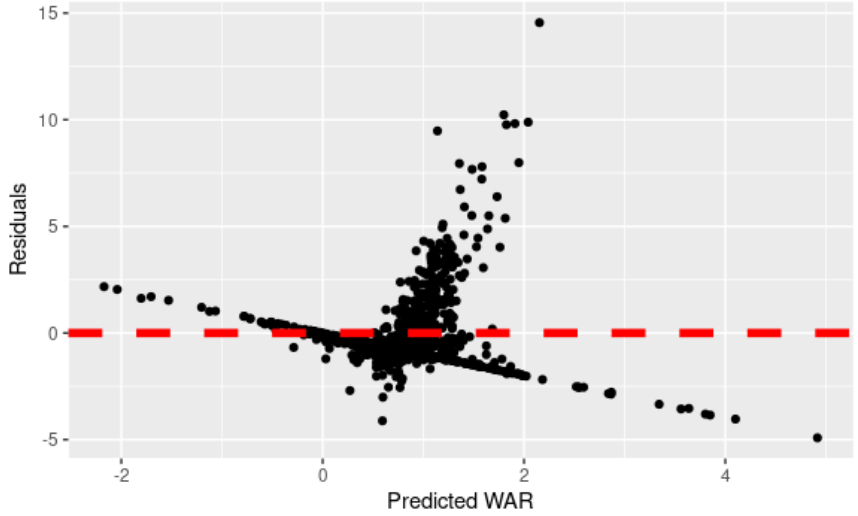
Suggestions for Improvement

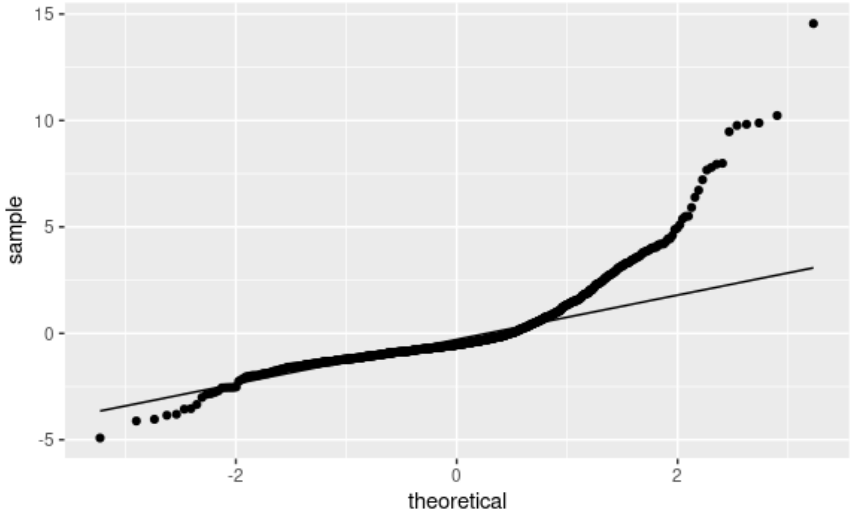
- Our linear models did not always fit their diagnostic test. We tried to augment the data using log augmentation, but that did not work because we had negative values. If we could figure out a way to fit the data better, our results would be more accurate.
- In our first correlation test, the QQ plots showed some non-normality. We could have come up with more accurate conclusions if the condition of normality was met.
- In addition, we performed a T-test on the full population. Due to the fact that NBA roster changes every year, performing a T-test on the full population might suggest accurate predictions about future years.
- Our data is numeric and does not necessarily contain all of the nuances of the NBA, such as coaching, team morale, synergy between teammates, and injuries.
- We would have chosen a similar data set but one with a random population of NBA players throughout history.
- We would have chosen a data set that included simpler statistics as well, such as points per game, rebounds per game, and assists per game,

Appendix

Effect of Offensive and Defensive Raptor on WAR in the regular season
- Diagnostic tests

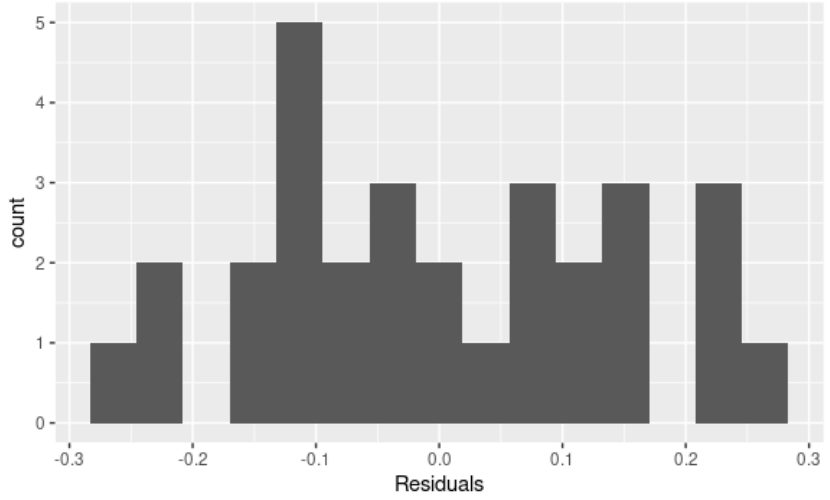


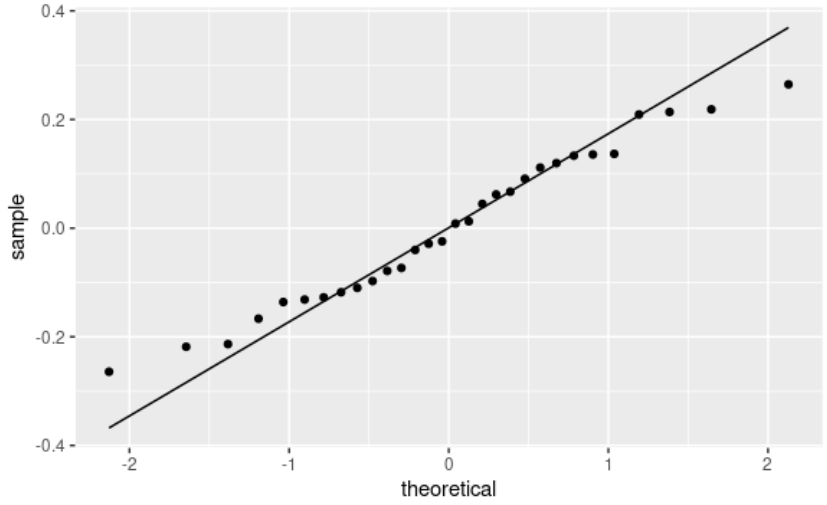


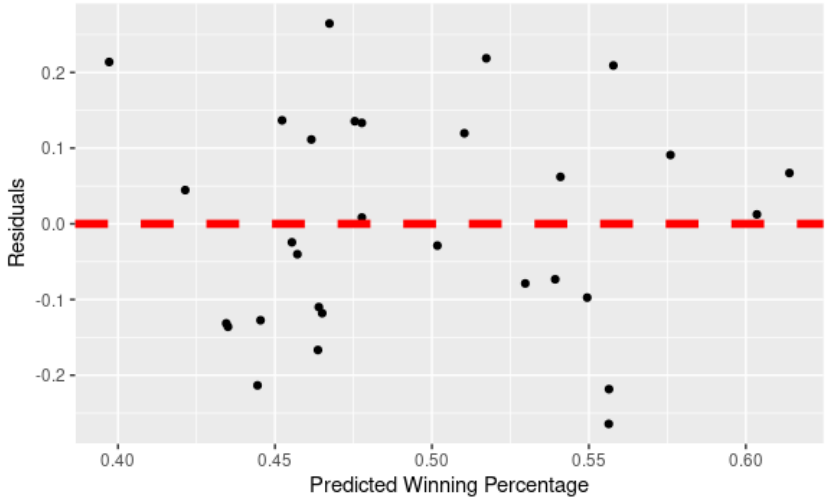


Appendix

Raptor Total and Winning Percentage
- Diagnostic tests

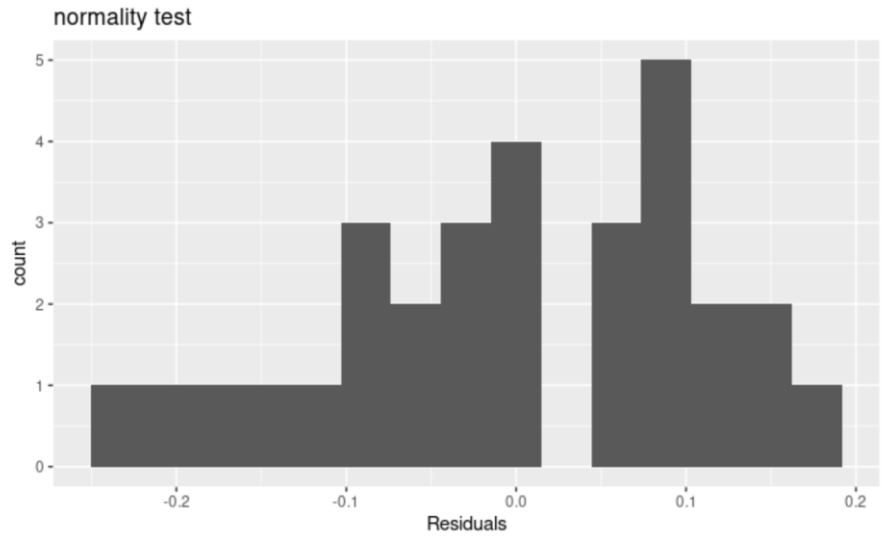


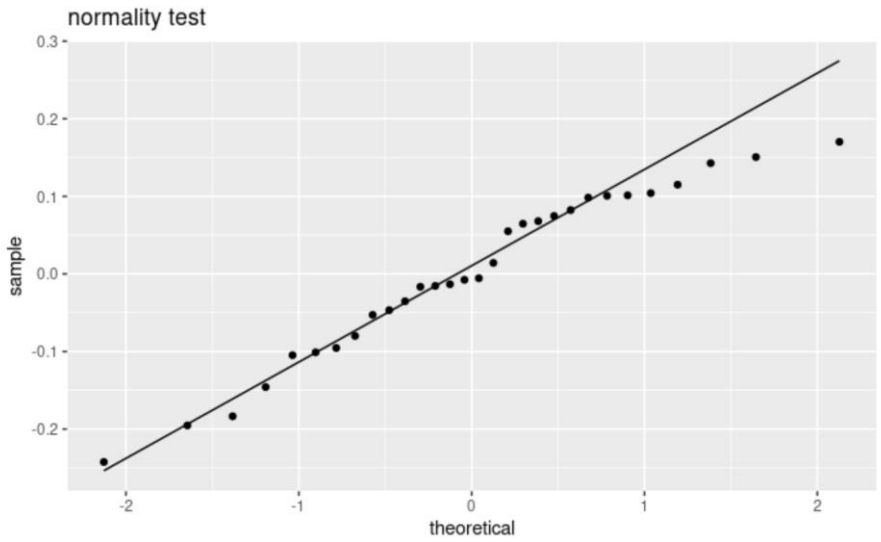


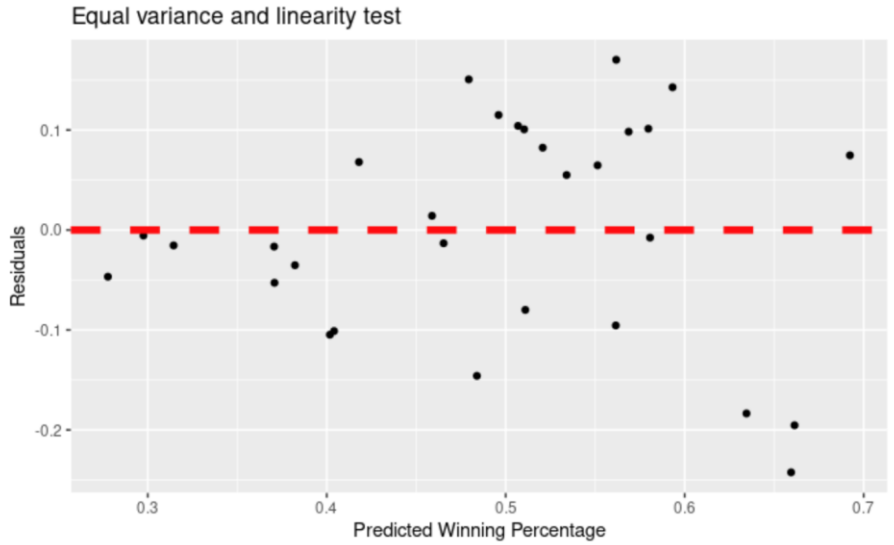


Appendix

WAR Total and Winning Percentage
- Diagnostic tests







Appendix

Sources:

1)

https://www.nba.com/news/nba-2019-20-all-nba-teams-official-release

2)

https://bleacherreport.com/articles/2265297-does-defense-really-win-championships-in-the-nba