

# Data\_607\_Assignment\_1

Matthew Tillmawitz

2024-08-29

## Overview of 538 RAPTOR Player Ratings

RAPTOR is a scoring metric developed by 538 to attempt to more accurately rate players and teams based on the criteria valued by modern teams that go beyond the traditional “box” metrics. While including a box component, RAPTOR also tracks on/off court team performance for each player to determine their impact on the game. The accompanying articles have some breakdowns of the “best” players according RAPTOR, an interesting note is the relative undervaluing of LeBron James by the model during a period of his career where he was widely considered to be one of the best players to ever play in the NBA.

RAPTOR articles:

<https://projects.fivethirtyeight.com/nba-player-ratings/>

<https://fivethirtyeight.com/features/introducing-raptor-our-new-metric-for-the-modern-nba/>

<https://fivethirtyeight.com/features/how-our-raptor-metric-works/>

## Filtering the Data

```
library("readr")
library("dplyr")
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
raw_raptor_by_team = read_csv('https://raw.githubusercontent.com/Tillmawitz/data_607_assignment1/main/i')
```

```
## Rows: 7289 Columns: 23
```

```
## -- Column specification -----
```

```
## Delimiter: ","
```

```
## chr (4): player_name, player_id, season_type, team
```

```
## dbl (19): season, poss, mp, raptor_box_offense, raptor_box_defense, raptor_b...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

# Removing predictive and component columns, keeping only raptor composites
columns_to_keep <- c("player_name", "player_id", "season", "season_type", "team", "mp", "raptor_offense")
raptor_data_only = raw_raptor_by_team[columns_to_keep]

# Cleaning up naming
final_raptor_df <- raptor_data_only %>%
  rename("team_id" = "team",
         "minutes_played" = "mp")
```

## Conclusions

A particularly interesting way to extend this data set would be to include the regular season and playoff win/loss records of each team as appropriate and analyze the performance of teams with “star” players to those of teams with more consistent player ratings. Star players can be defined as players who are significantly above the average for their teams, although super teams consisting of multiple players significantly above the league average may also need to be broken out. If extended further to include individual head to head matchups and whether a given player had significant minutes in the game then the impact of having star defenders versus star offensive players or all around stars could also be analyzed in greater detail.