# 🏀Exploratory Analysis of NBA Data🏀

## W200 - Fall 2018 - Section 2 - Group 1

**Team Members:** Anusha Praturu, Tom Goter, Noah Pflaum

**Github Repo:** https://github.com/UCB-INFO-PYTHON/Project2_Section2_Group1

### I. Abstract

Exploratory analysis of an NBA Player of the Week data set, supplemented by data from the NBA API was pursued to explore a hypothesis from Horace Grant that the league has tended toward "small ball" due to rule changes such as the addition of the three point shot and the elimination of forearm checking. Initial results, such as the average weight of an NBA player dropping 10lbs in the past 20 years and the average height of players taking 3pt shots increasing 2 inches, support this hypothesis. Additionally, NBA player and team data was evaluated to investigate how NBA Players of the Week affect overall team performance and even how they affect an individual game. Through the exploratory data analysis discussed herein, it is shown that teams with a high number of Players of the Week awards in a given season tend to perform at a higher level than those with fewer Players of the Week. There are exceptions to this however, and these exceptions tend relate to teams performing at a high level without having a "superstar". Thus, the relationship is that Players of the Week help teams win, but teams winning doesn't necessarily help create Players of the Week. The exploratory data analysis discussed herein just breaks the surface on the NBA API dataset as shown by the variety of analyses performed. Non-exhaustive areas for future data exploration are included for consideration.

### II. Introduction and Questions

In Horace Grant's blog post "How the Game has Changed" he claims that the addition of the three point shot and the elimination of hand/forearm checking has transitioned the NBA from a big man's game to a style referred to as "small ball". Grant defines small ball as the following:

> *"Small ball is played by teams who sacrifice player size and low post emphasis to instead utilize smaller, agile outside scorers. Small ball teams build around a strong forward, such as LeBron James, Kevin Durant or Carmelo Anthony, who are then surrounded by players who can fluidly switch between positions."*

This definition is suggestive that the physical size of NBA players is decreasing, that shots are being scored from further away, and that a dominant forward player still exists. This study examines these claims by Grant primarily utilizing Player of the Week data from a Kaggle dataset and further NBA statistics from NBA.com.

In addition to exploring the changes in the physical attributes of the NBA Players of the Week over the past three decades, the relationship between team performance and the frequency of having Players of the Week was investigated as a means to explore the impact these players have.

NBA team and player data since the inception of the league was investigated to determine whether NBA Players of the Week are representative of the league as a whole. Specifically evaluated were questions of 1) how the NBA player population have physically changed with consideration for how the league (e.g., rules) has changed and 2) correlation of team performance with individual player recognition (i.e., Player of the Week awards). The datasets used in this analysis are discussed in the following section along with a discussion of how these datasets were cleaned and verified.

## III. Data

This section provides details about the pedigree, acquisition and quality/cleanliness of the datasets used. Assumptions are noted, as appropriate, throughout Sections III and IV. The analysis discussed herein relied primarily on two sources of data. These data sources were:

1) NBA Player of the Week Dataset - Maintained by Jacob Baruch on Kaggle
2) NBA.com (API) - Harnessed through Python Package nba-api

### A. Player of the Week Dataset

This data set has 1155 observations and 14 features. These feature columns are given in Table 1 which also includes a short description of each feature. As appropriate, key items of note have been included in **bold maroon**. Feature names were not ideal for data processing as some contained spaces and were inconsistent with other data sources. This was a relatively trivial inconvenience and is addressed in following sections.

**Table 1: Listing and Description of Features in NBA Player of Week Dataset**

| | |
|---|---|
| Active Season | Column of just zeros - **not useful information** |
| Player | Player full name - **not fully consistent with NBA.com names** |
| Team | Team city and name at the time of the award - **some teams have changed city or name during this period of interest** |
| Conference | East/West - for awards 2002 season and beyond - **blank for seasons prior to this** |
| Date | Date of award (Day/Month/Year format) |
| Position | Abbreviated position of player ('PG', 'SG', 'F', 'C', 'SF', 'PF', 'G', 'FC', 'GF', 'F-C', or 'G-F') |
| Height | Height of player in 'FEET-INCHES' format **(not useful format for mathematical operations)** |
| Weight | Weight of player in pounds (lbs) |
| Age | Age of player |
| Draft Year | Year player was drafted into NBA (YYYY format) |
| Seasons in league | Seasons player has been in league |
| Season | Season in format 'Year-Year+1' format since the NBA season spans calendar years |
| Season short | Year the season ends (i.e., Year+1 from above) |
| Real_value | **1.0 for awards given prior to 2002 season (one award per week)**<br>**0.5 for awards given 2002 and beyond (one award per conference per week)** |

The values for each feature were investigated to ensure that they were reasonable; however, a 100% check of all awards and player information was not conducted. Some items of note related to consistency with other data were already highlighted in Table 1. These additional reasonability checks included looking at the distributions of numeric parameters (e.g., height, weight, age) and the sets of unique values for string parameters (e.g., team, player, position). Figure 1 shows histograms of the weights, heights and ages of the NBA Players of the Week. No adjustments were made to the bounds of the axes. While there are some outliers in weight, height and age, these values were checked and determined to be correct (e.g., high weight → Shaq, high age → Michael Jordan returning to play on the Wizards).

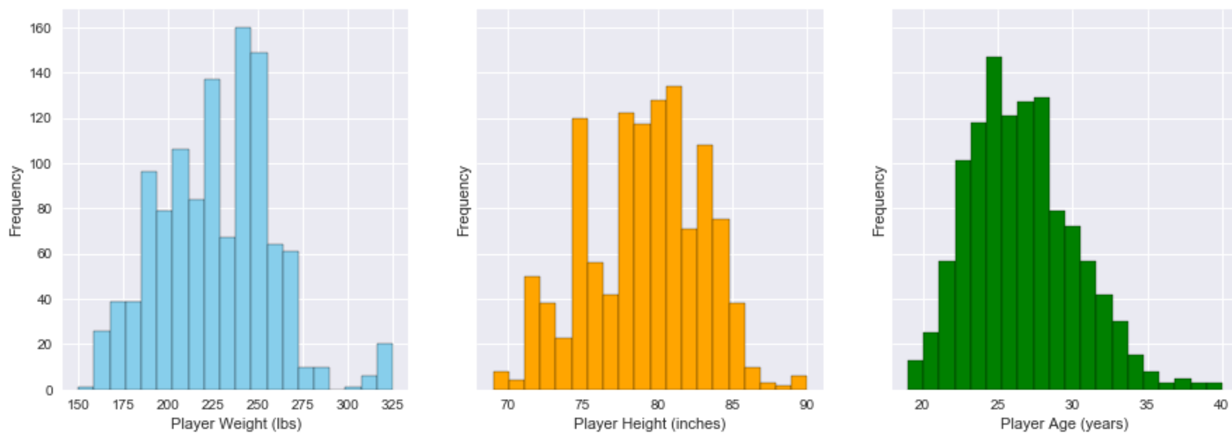


**Figure 1: Distribution of NBA Player Data from Player of the Week Dataset**

<u>Cleaning</u>

Once the dataset above was assessed for veracity and consistency several operations were performed to fix the identified issues. The Player of the Week dataset was cleaned in the following ways:

1) 'Player ID's were added to allow accurate merging with player data. In order to add Player IDs, eight player names were modified due to inconsistencies in naming convention between the Player of the Week data set and NBA API data.
2) 'Franchise' was added as a feature. This maps the specific team name and city to a particular current team nickname (one of the 30 current NBA teams). These team nicknames are consistent with the 'nickname' variable coming out of the NBA API and allows for finding the Team ID which is then used to easily join datasets (item 3). This process also eliminated team naming differences such as 'Sixers' versus '76ers'.
3) 'Team id' was added as a feature to the Player of the Week dataset to allow for accurate merging of Player of the Week data with team data. This addition was performed using the Franchise feature mentioned above.
4) 'Team city' was added as a feature. This was extracted from the existing 'Team' feature and is used to impute missing conference data from seasons prior to 2002 (see item 5).
5) Conference data was imputed for Player of the Week awards occuring before the 2002 season. This was done with a look-up between cities and conferences in the data post-2002.
6) Feature names were converted to lowercase and whitespace was replaced by an underscore

Feature Relationships

In order to understand the nature of the data and assess which variables might be worth investigating further, a correlation matrix of the numerical columns in the Player of the Week dataset was created and is shown as Figure 2.

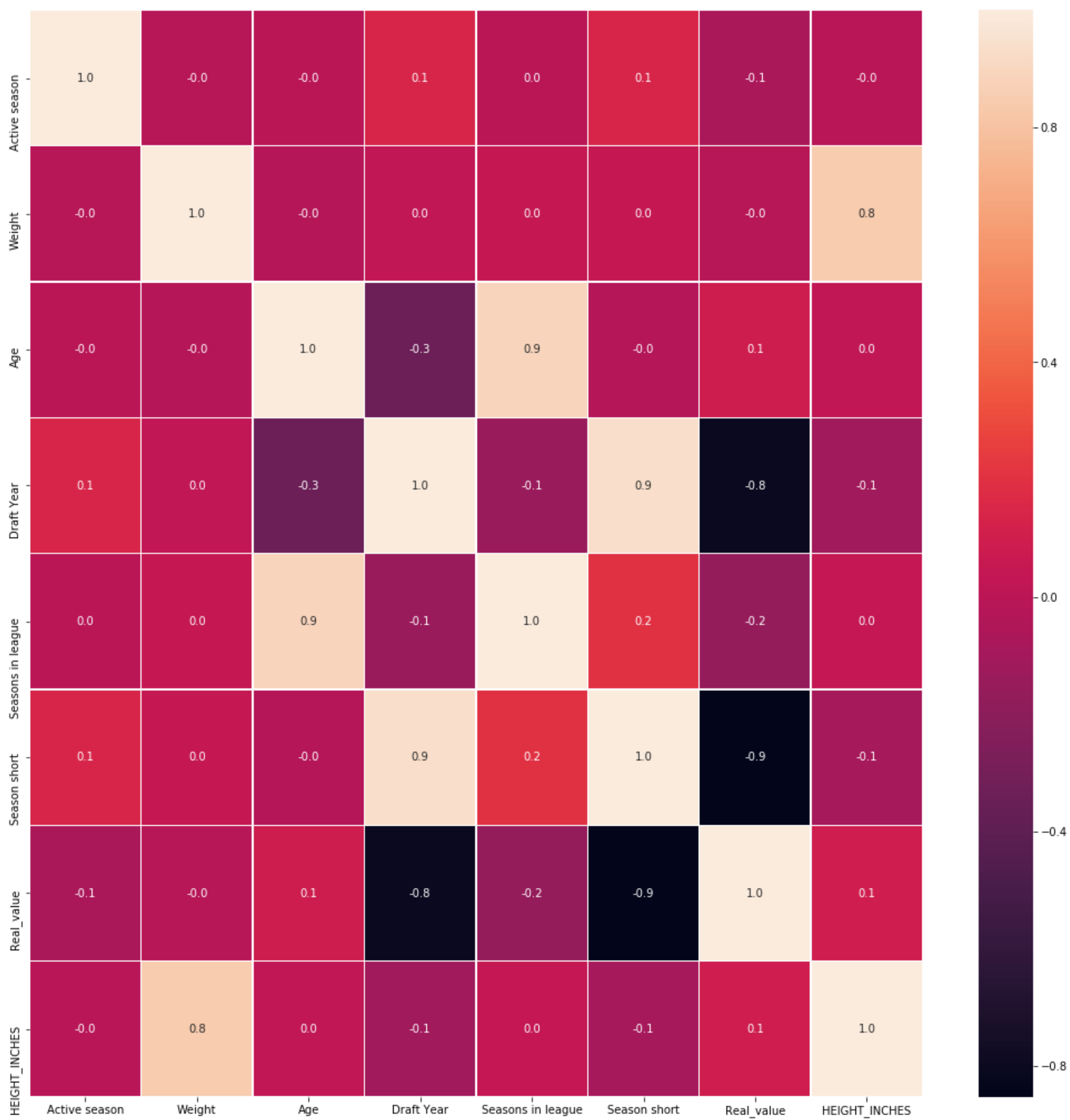


**Figure 2: Correlation Matrix of NBA Player of the Week Data Set**

Certain findings from the above correlation matrix were expected, such as the strong relationship between Draft Year and Award Value (displayed as Real_value). This is because a rule change in the 2001-2002 season enforced that Player of the Week titles be awarded on the conference level, rather than the league. Consequently, a score of 1.0 was given to players who were awarded the title before the rule change and a score of 0.5 was awarded to players awarded the title after the rule change. Other findings from the dataset were unexpected, such as the negative correlation between age and draft year, suggesting that the average age of drafted players has changed over the years. Such findings guided our attempts at further analysis.

### B. NBA API Datasets

The NBA API was utilized to generate new datasets to supplement the Player of the Week data acquired from Kaggle. The NBA API can be harnessed using the nba-api package for Python available through PIP. Using this tool, two primary additional datasets were generated. These are listed below, and the data acquisition and cleaning processes used for each are discussed in more detail in the sub-sections below.

1) Yearly aggregated statistics by NBA player
2) Yearly aggregated statistics by NBA team

***Team Data***

<u>Collection</u>

This dataset was generated in order to investigate the relationship between team performance and the number of Player of Week awards each team received in a given year. It was generated by using the 'teamyearbyyearstats' endpoint of the NBA API and contains data for every *existing* team/franchise from every season from 1950 until now. This dataset, as generated, contains 960 observations and 35 features. The number of observations was checked for reasonability relative to how many teams there were over the history of the NBA (Figure 2) noting that this only includes the teams that are still in existence today (regardless of city or name change). This distinction is made, since in the early years there were some teams such as the Providence Steamrollers and Toronto Huskies which existed for 3 and 1 season, respectively, at the inception of the NBA and thus provide little value to this analysis.

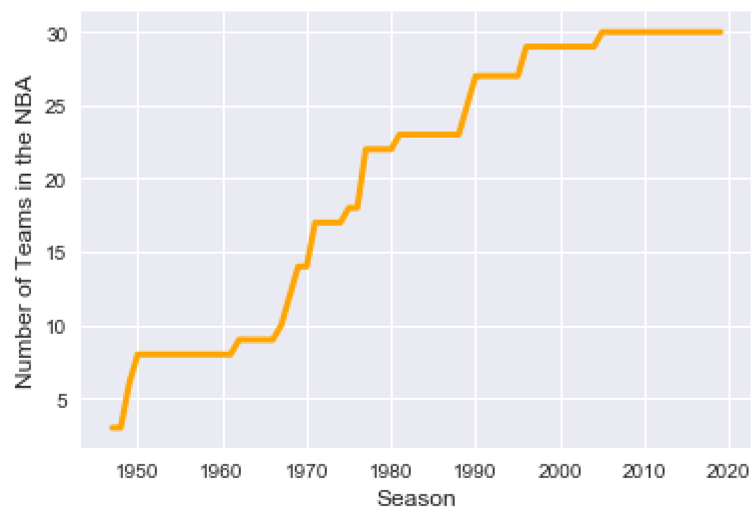


**Figure 3: Number of Teams in the NBA from 1947 until Today**

The features included in the NBA Team Data dataset include a range of variables including items such as wins, losses, winning percentage, team shooting percentage, et cetera. All variables are not listed here for brevity, but as these variables become important to the analysis discussed in the next section they will be explained in more detail. Figure 3 is presented as an example of some of the checks placed on the team data features and includes histograms showing the distribution of team win percentage, shooting percentage, and total number of blocks for all teams from 1984 and beyond. The dataset was subsetted to beyond 1984 for this comparison, as this is the period of overlap with the aforementioned Player of Week dataset. The axes of Figure 3 have not been modified. The distributions appear reasonable (e.g., winning percentage peaked near 50%), and there are no extreme outliers (e.g., 0 or greater than 1.0 for the fractions) which would likely indicate issues with these features.

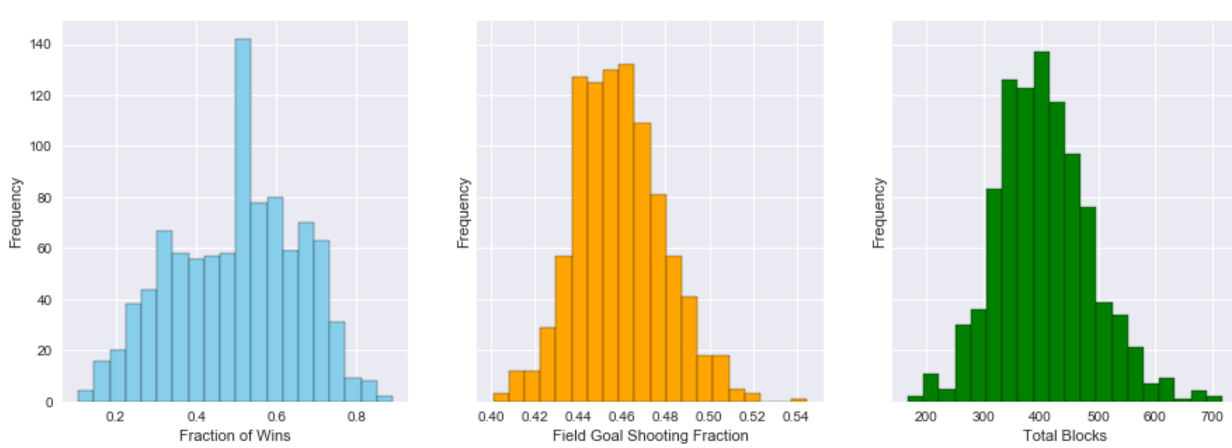


**Figure 4: Example of Reasonability Checks Placed on NBA Team Data**

While the distribution of the data within some of the feature columns appeared reasonable, some of the column names and descriptive variables were inconsistent with those from the NBA Player of the Week dataset. Namely 'TEAM_NAME' and 'TEAM_CITY' were separate columns when data was pulled from the NBA API; whereas, both of these were lumped together in 'Team' in the Player of the Week dataset. Discussion of how this inconsistency was overcome was included in the Data Cleaning section regarding the Player of the Week data.

<u>Cleaning</u>

As discussed in the preceding section, the team data from the NBA API from 1984 and beyond was investigated for accuracy and cleanliness. It was determined that based on the aforementioned changes made to the Player of the Week dataset to enable consistent merges (e.g., adding the team_id feature to the Player of the Week dataset), that limited additional cleaning was needed of the team data.

***Player Data***

<u>Collection</u>

In order to pull the data from NBA.com web API, a python package called nba-api was utilized. The nba-api package works by creating endpoint wrappers for the NBA.com web API and allows users to specify their desired query parameters. Internally, nba-api leverages another python package called requests that enables making HTTP requests from python. Three main endpoints were used and data used from each endpoint were:

1. Player game statistics for each season
    a. It contains 26077 samples with 28 features
    b. Features used: minutes played, field goals made, field goals attempted, three point field goals made and three point field goals attempted
2. Player biography statistics for each season (data available from 1996-97 onwards)
    a. It contains 10563 samples with 35 features
    b. Features used: player height and weight
3. Player career summary statistics
    a. It contains 4371 samples with 31 features
    b. Features used: player height and weight

<u>Cleaning</u>

The intent was to combine player game statistics to their physical features (i.e. dataset 1 with 2 and/or 3) to validate portions of Grant's claims. However, in doing so, several gaps and issues with the data arose.

The first issue encountered was that the NBA.com throttles the number of requests that can be made. Through trial and error it was found that making a requests every 0.4-0.5 seconds allowed for there not to be any interruptions. For this project over

20000 requests were used (some data went unused) resulting in about 3-4 hours worth of data collection. While the total time isn't excessively large, it was somewhat troublesome when it's realized certain data is missing and the planned time to complete work is held up by time taken to pull the data.

When trying to gather data related to physical characteristics of players, initially it was thought that it would be best to use player data for each season to more accurately represent the player for that season. The problem was that this data, dataset 2 described above, is only available from the 1996-97 season onwards and this study looks further back in time than this. That said, several spot check were performed that highlighted that accuracy of this data on a per season basis isn't a huge concern to the NBA and what is listed is a representative number, but with a fair amount of error:

- Kevin Durant has publicly stated he lists himself as 6'9" despite being 6'11" so he's seen as a small rather than power forward.
- Kevin Love has the opposite situation where he's just a tad over 6'7" yet lists himself as 6'10"
- Lebron James was listed as 250 lbs during his time at Miami yet sources from within the locker room saw him weigh in at 278 lbs after a game
- Eric Piatkowski managed to grow and shrink an inch 4 times through his career including growing an apparent inch at age 33.

Therefore it was decided dataset 3, which only has a single data point for each height and weight, would suffice. However, further issues were realized when this dataset showed to have large gaps in data as shown in Figure 5. Even Antawn Jamison, a two time All-Star that played 16 seasons in the league, was missing this data. In the plot below there are two sets of trends shown: the number of played minutes where player physical data (height and weight) is present and missing, and secondly how much as a percentage of all players is that data unavailable. As can be seen, starting from mid 80s this dataset (#3) lacks a lot of player data.

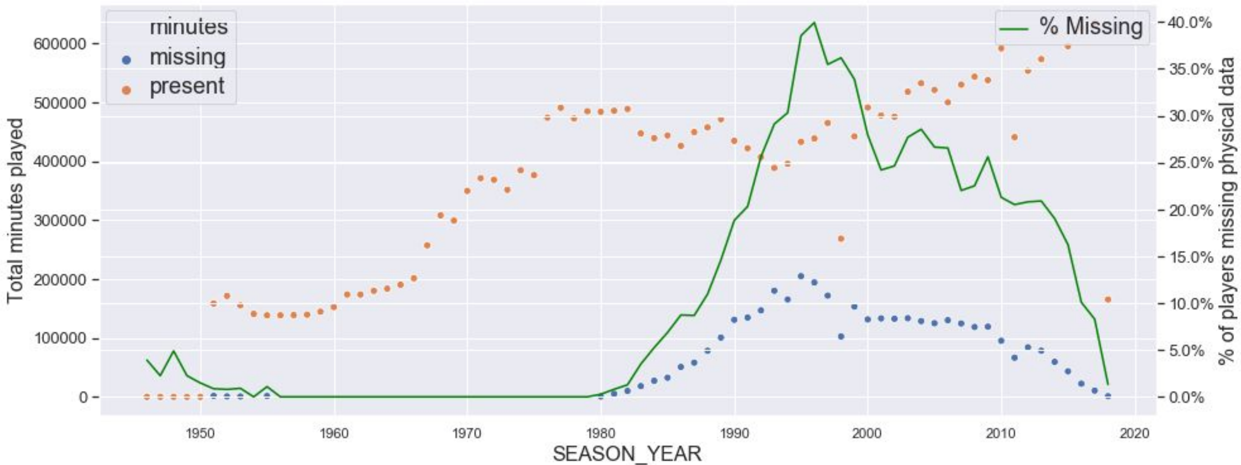


**Figure 5: Missing Physics Data in NBA API Data by Season**

What was then done is that dataset 2, which has the physical characteristics data by season for each player, was averaged for each player throughout each season they played to have a single height and weight number for each player. There was no average done by number of minutes played each season or any other variable. After this averaging, dataset 2 and 3 were combined, and it transformed the data gaps into that shown in Figure 6.

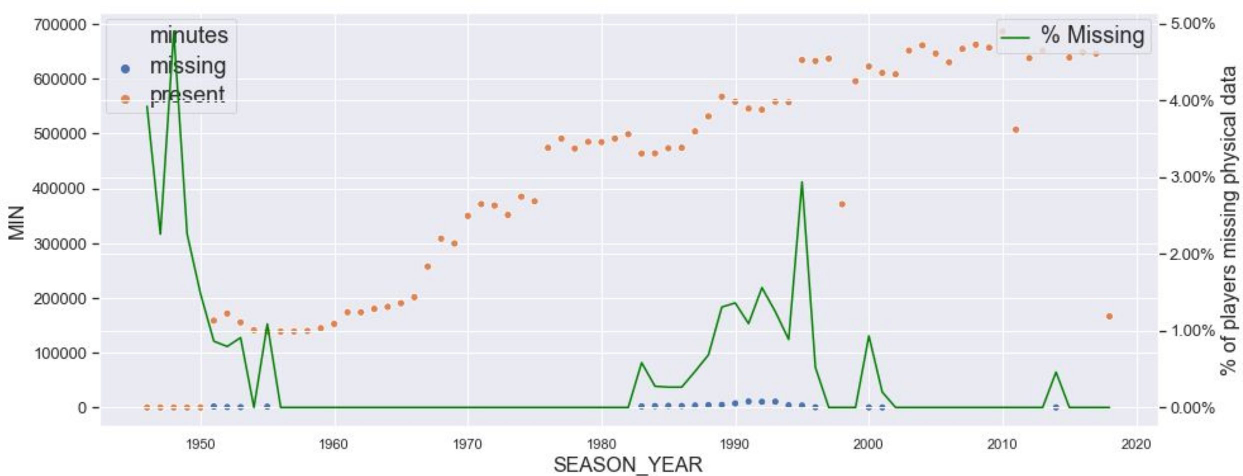


**Figure 6: Missing Physics Data after Season-averaging**

The % of missing physical data for players has dropped from a peak of about 40% down to a peak of 5%, but more realistically a peak of about 3% for the time frame of interest.

*Sanity Checks*

Various sanity checks were performed while cleaning and transforming the data. Most of these checks involved looking at the minimum and maximum of variables and comparing that against known values. For example, Muggsy Bogues is known to be the shortest player in the NBA at 5’3” (63 inches) and the data set aligned with this.

## IV. Discussion of Analysis

To examine the physical characteristics of players over time, a variety of height and weight related metrics were examined. Looking at height the following variables were explored:

- HEIGHT_INCHES = The average height in inches of all players in the NBA
- HEIGHT_MINS_AVG = The average height per player minute on the court
- HEIGHT_PTS_AVG = The average height per player per point scored

What can was found is that the height of players has been relatively stable since the middle of the 1980s at 79 inches (6’7”), but even going back further the average height since the mid 1960s has been 78 inches (6’6”). As can be seen in Figure 7, there is little deviation between each height metric.

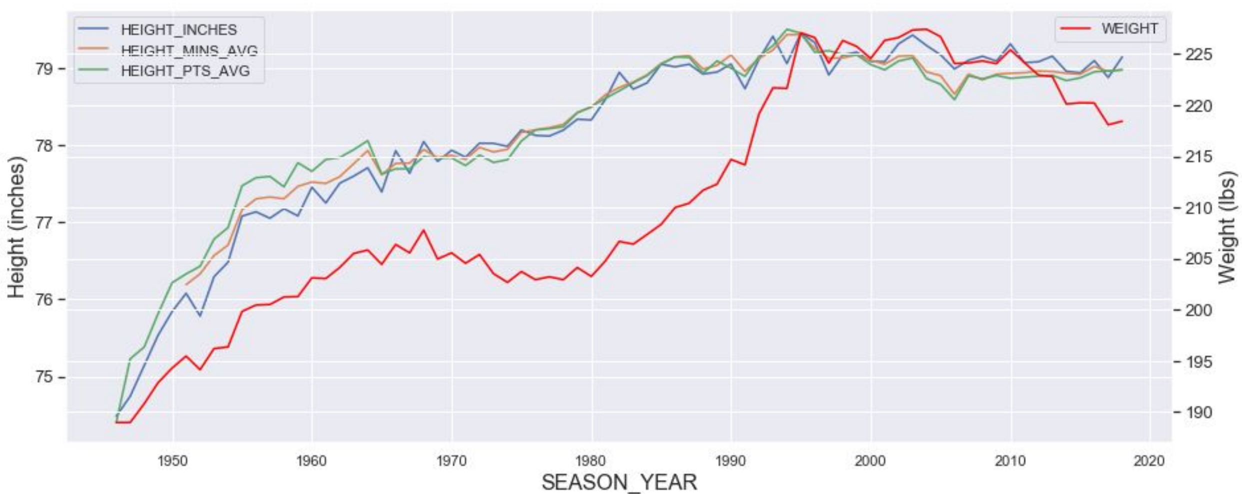


**Figure 7: Height and Weight of NBA Players by Season**

The weight of each player does show a steep increase from the 1980s (although this is a trend noticed in many performance sports) before peaking in the mid 1990s before dropping close to 10lbs to today’s current average. While 10lbs may not seem like a lot (~4% decrease from the peak) when looking at the distribution of player weights who are 78-80 inches in height (Table 2) it turns out that 10 lbs is about 40% of the interquartile range suggestive that player weight has decreased noticeably. This aligns with Grant’s claim of the decreased use of post play where physical mass is beneficial.

**Table 2: Descriptive Statistics of Player Weights Feature**

| | |
|---|---|
| Count | 2230 |
| Mean | 217.15 lbs |
| Std | 17.33 lbs |
| Minimum | 170 lbs |
| 25% | 209 lbs |
| 50% | 215 lbs |
| 75% | 225 lbs |
| Maximum | 270 lbs |

Looking beyond physical data, "small ball" tends to entail shooting from a greater distance away. Unfortunately, the exact distance of each shot is not available on the NBA.com API, but it does breakdown 2pt vs 3pt field goals - note: the 3pt field goal was only introduced in 1979. The following graph shows the field goal shooting percentage overtime (FG%), the 3pt field goal shooting percentage (FG3%) and the proportion of shots that are 3pt attempts (FG3%TOTAL). Looking at 1979 onwards it can be seen that players have become better 3pt shooters, and that players are taking a lot more 3pt attempts relative to the total number of shots taken (Figure 8). This also perhaps accounts for the ~5% decrease in overall field goal shooting percentage peak from 1983 as 3pt shots are harder to make. This also aids Grant's argument that "small ball" from a non-physical perspective is on the rise as 3pt shooting attempts have drastically increased.

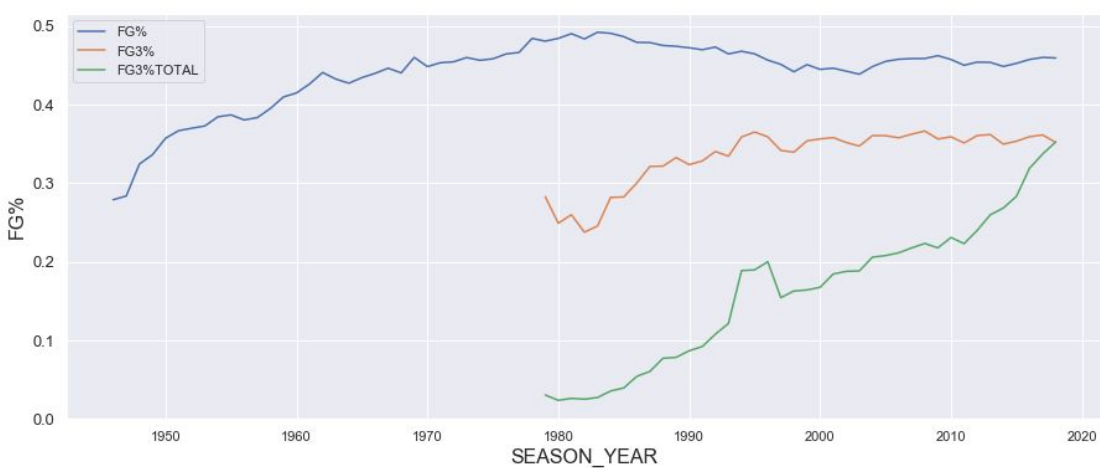


**Figure 8: Shooting Percentage and Percentage of 3pt Shots by Season**

If the idea of players' physical stature and the rise in 3pt attempts are analyzed together, what can be seen is that the average height of 3pt field goal attempts and makes has increased by close to two inches. Moreover, the discrepancy between the height of 3pt field goals made vs attempted has been all but eliminated, suggesting that taller players are shooting the ball better than they used to (or shorter players are shooting it worse) as shown in Figure 9.

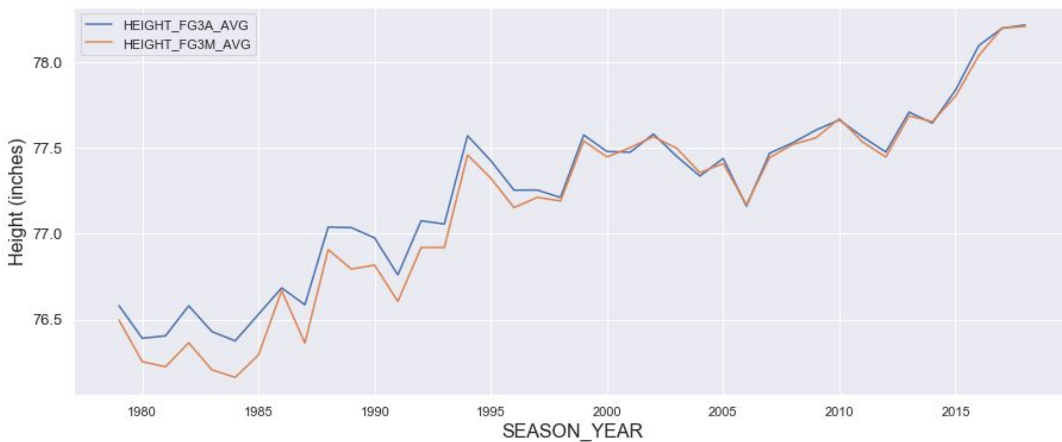


**Figure 9: Average Height of Player Making and Attempting 3pt Shots**

Again, this finding is in alignment with Grant's claims related to the rise of small ball.

Using the clean Player of the Week data, the relationship between team performance and the Player of the Week award was then examined. This was done using a progressive approach. First, the cumulative number of wins by each team was compared to each team's cumulative number of Player of the Week awards. The result of this comparison has been plotted as FIgure 10 and shows what appears to be reasonable correlation between high performing teams (lots of wins) and teams with highly recognized players (high numbers of Player of the Week awards). The two teams with both very few Player of the Week awards and very few wins are explained as the most recent expansion teams in the NBA (Grizzlies - 1995, Pelicans - 2002). The complete distribution of cumulative Player of the Week awards by NBA Franchise is given as Figure 11, with the two aforementioned outliers identified.

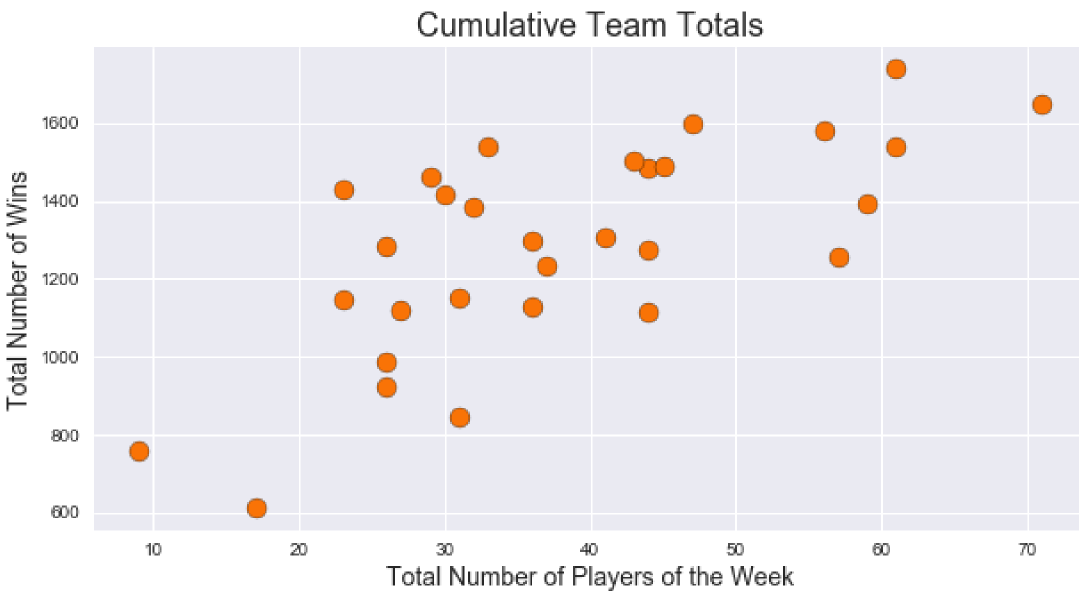


**Figure 10: Relationship between Cumulative Team Wins and Player of the Week Awards - 1984 to 2018**

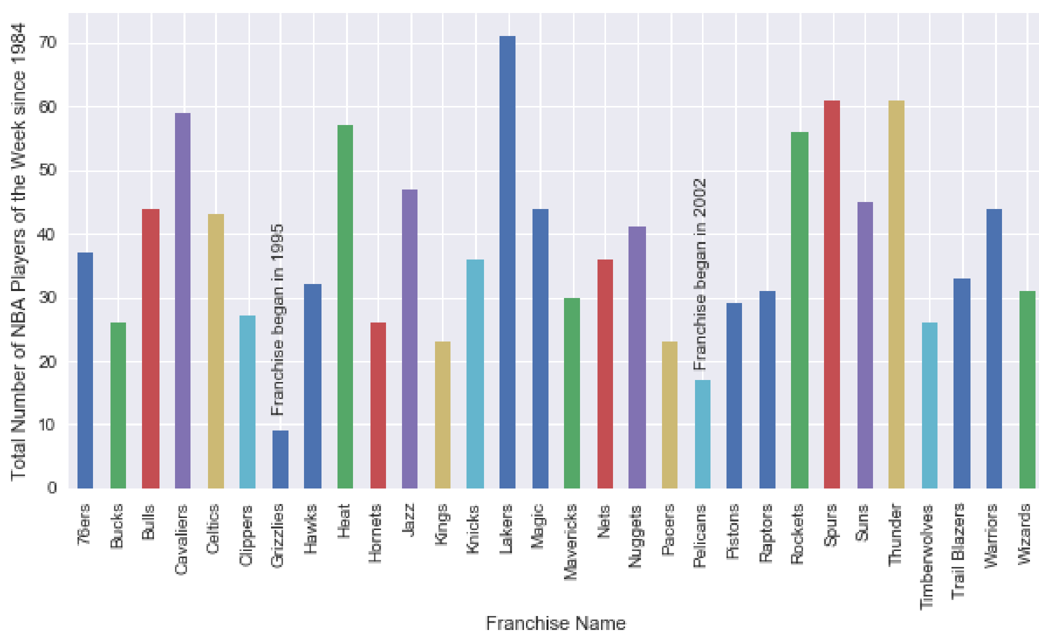


**Figure 11: Cumulative Player of the Week Awards by Franchise**

Noting that there appeared to be some amount of correlation between teams winning and having larger numbers of Players of the Week, the data was further assessed to better understand this. One method of further assessing this data was to bin teams by winning percentage into Bad Teams (i.e., winning percentage < 40%) Great Teams (i.e., winning percentage > 70%) and Fine Teams (i.e., the others) and then compare these bins to the “real value” of their Players of the Week. The “real value” is the weighted sum of the number of Players of the Week in a given season. A value of 1.0 is given for each award given prior to 2002, and a value of 0.5 is given for each award 2002 and beyond. This distinction is made because prior to 2002, only one award per week was awarded, but starting in 2002 one award was given to a member of each conference. The results of this analysis are shown in Figure 12. Again a clear distinction between team caliber (i.e., winning percentage) and receipt of Player of the Week awards is seen.

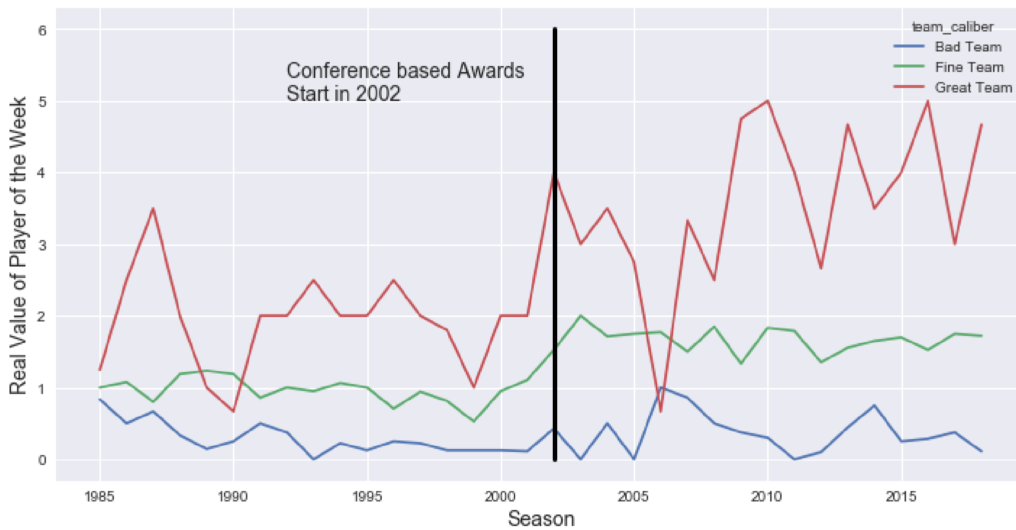


**Figure 12: Real Value of Player of the Week Awards NBA Teams**

Figure 13 shows the distribution of teams, binned by number of Players of the Week per season, versus their winning percentage. This is a bit clearer representation of the data shown in Figure 12. However for Figure 13, each Player of the Week award, regardless of date of receipt, is weighted the same. The relationship shown in Figure 13 clearly indicates that the teams producing five or more Players of the Week in a season are always winning teams (i.e., winning percentage greater than 50%). The figure also shows the great diversity in winning percentage with teams with very few Player of the Week (POW) awards. This was an interesting finding; regardless of number of Player of the Week awards, the maximum winning percentage is quite flat. However, since the median and mean appear to be increasing with an increasing number of Player of the Week awards, a correlation of sorts can be assumed.

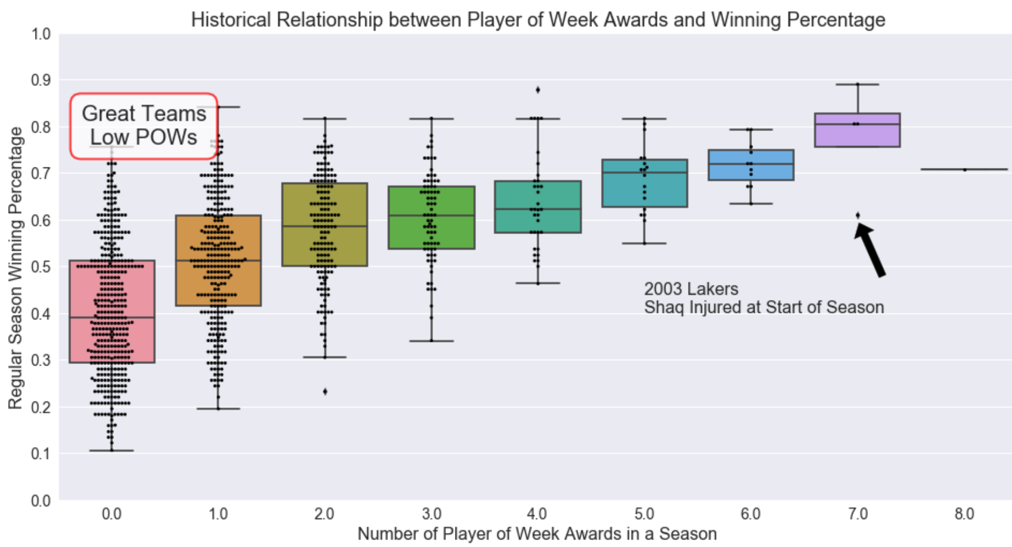


**Figure 13: Team Winning Percentage vs. Number of Players of the Week**

In evaluating the data shown in Figure 13, a couple of follow-up questions are worth asking: 1) What explains the outlier at high number of Player of the Week award with a relatively low (still good) winning percentage and 2) What distinguishes teams with very different numbers of Players of the Week but the same winning percentage. The outlier in the lower right hand corner of Figure 13 was quickly identified as the 2003 Lakers. Upon further investigation it was noted that the Lakers had quite a slow start because Shaq was out for the first 12 games due to foot surgery which led to the slowest start for the Lakers in 10 seasons. However they still had Kobe and eventually Shaq which led to them piling up Player of the Week awards. The damage due to the slow start was done (lost 19 of first 30 games).

Better understanding of the second question above (i.e., What distinguishes teams with very different numbers of Player of the Week awards but similar winning percentages?) is a bit more challenging and requires further analysis. For this portion of the analysis, teams were separated into two distinct datasets. The first dataset was the teams with winning percentages greater than 70% and yet fewer than two (absolute) Player of the Week awards on their team in a given season. The second dataset was the subset of teams with winning percentages greater than 70% and five or more (absolute) Player of the Week awards in a given season.

The aggregated statistics for these teams were compared, and while there were small differences in shooting percentage, blocks and rebounds, no key difference really stood out as explaining the discrepancy. The analysis then went a step deeper by looking at the points per minute of each player on each of the teams in the above datasets (players with fewer than 100 minutes for an entire season were not considered). The median and maximum points per minute for each team in both datasets was then compared. A graphical representation of this comparison is shown as Figure 14. For ease of comparison, the horizontal and vertical axes are consistent across all plots in the Figure. The comparison shows that there tends to be a larger difference between the median and maximum points per minute on teams with higher numbers of Player of the Week awards.  This could indicate that these teams are more reliant on certain players. This seems to follow logically from the Player of the Week award truly being an award to a player and not a team, but it would need additional evaluation to determine the statistical significance of the difference.

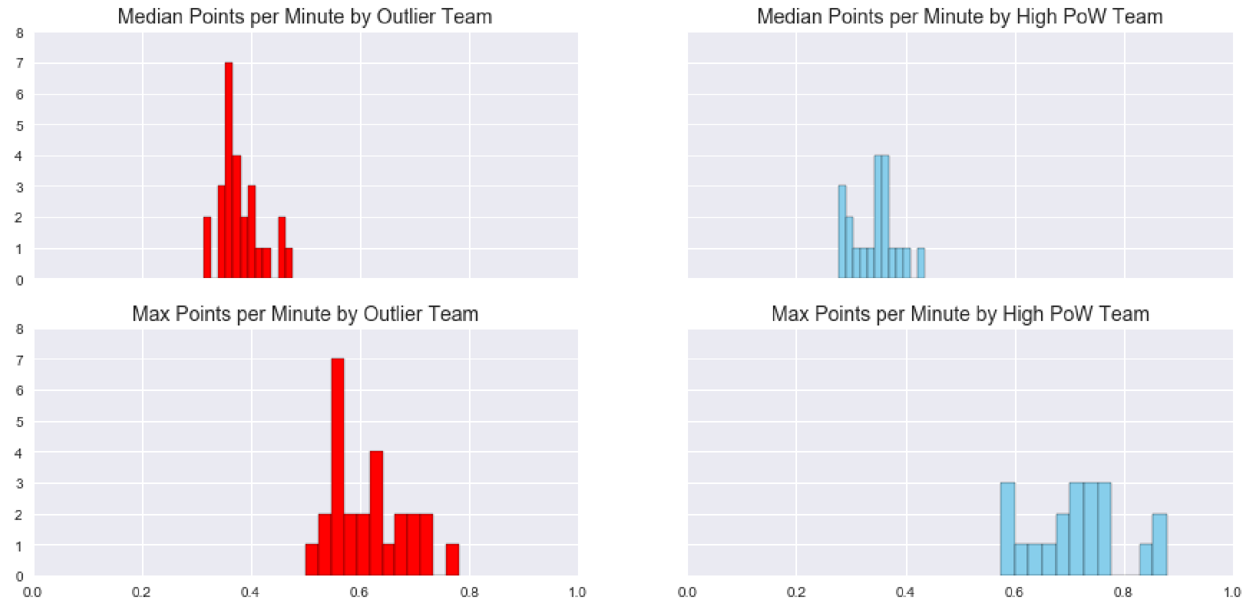


**Figure 14:  Comparison of Median and Maximum Points per Minute by Teams with High Winning Percentages**

For now it seems clear that having consistent Players of the Week on a particular NBA team is correlated to a higher winning percentage. So the next question asked during this analysis is “What physically and statistical attributes make for a Player of the Week?” and “Are these traits distinguishable from the larger population of NBA players.

In an exploration of player age in the Player of the Week dataset, it was found that the average age of an NBA Player of the Week has in general gotten younger since 1969, when data first became available as shown in Figure 15.

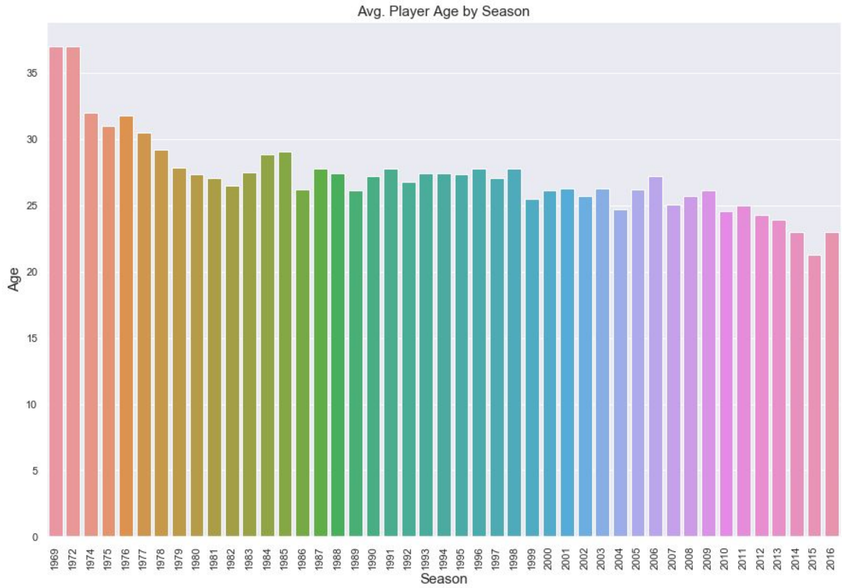


**Figure 15: Average Player Age by Season**

This data is indicative that over time, the demographic of a star player in the NBA has trended younger as players get drafted out of high school or in college. Additionally it was found that player age is correlated with the number of seasons played in the NBA (below), as expected and is shown as Figure 16..

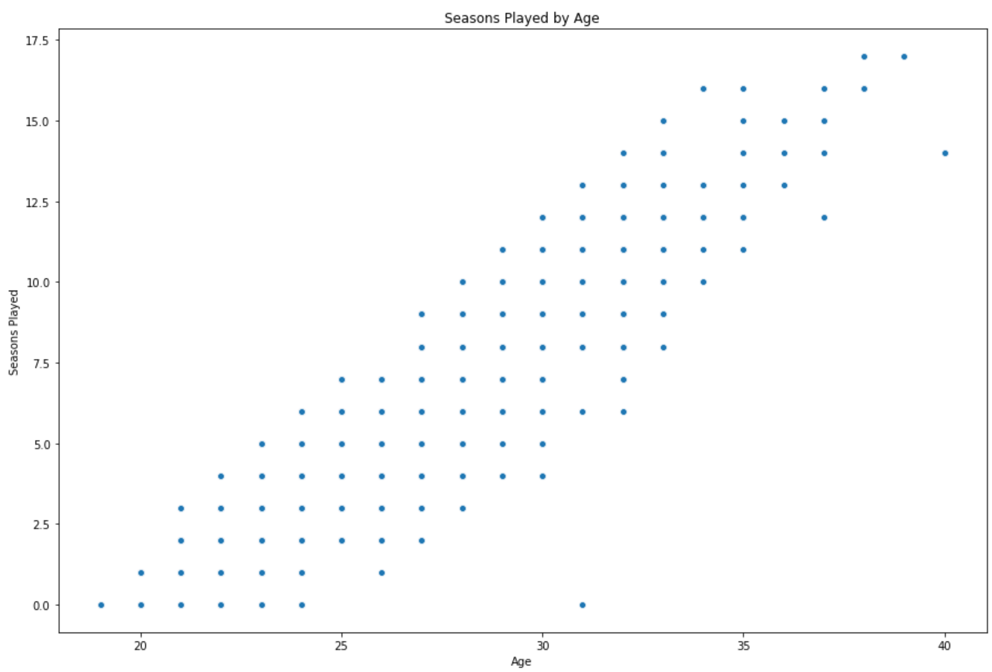


**Figure 16: Player Age by Seasons Played**

A notable outlier from the data above is the player that appears to have joined the NBA at the age of 31. This player was identified as Arvydas Sabonis, a Lithuanian professional basketball player who was drafted in by the Portland Trail Blazers in the first round of the 1986 NBA draft, but did not play his first game until 1995 at the age of 31.

A deeper investigation of the robust NBA API dataset also allowed for an analysis on a season, team, and even game level, even allowing the possibility to research a particular player's performance after being awarded Player of the Week. For context, we can first look at last season and see the Golden State Warriors performance over time the year that they most recently won the NBA championship (Figure 17).


**Figure 17: Warriors 2017 Win/Loss Margin**

Interestingly, you can observe that the Warriors experienced mostly wins over the season, but also saw higher margins of win and loss towards the beginning and end of the season. While they won more games than they lost, it is apparent that their margins of loss were more significant than their margins of victory, especially at the end of the season.

Going one level deeper, the NBA API allows for the analysis of a single game or player by using the Play-by-Play dataset. As an example, the April 3rd, 2018 matchup between the Golden State Warriors and Oklahoma City Thunder was selected. This was Kevin Durant's first season with the Warriors, and a game against his previous team in Oklahoma City immediately after he was named Player of the Week in the western conference. It was a close game that ended with a Warriors win after several lead changes, shown below in Figure 18.

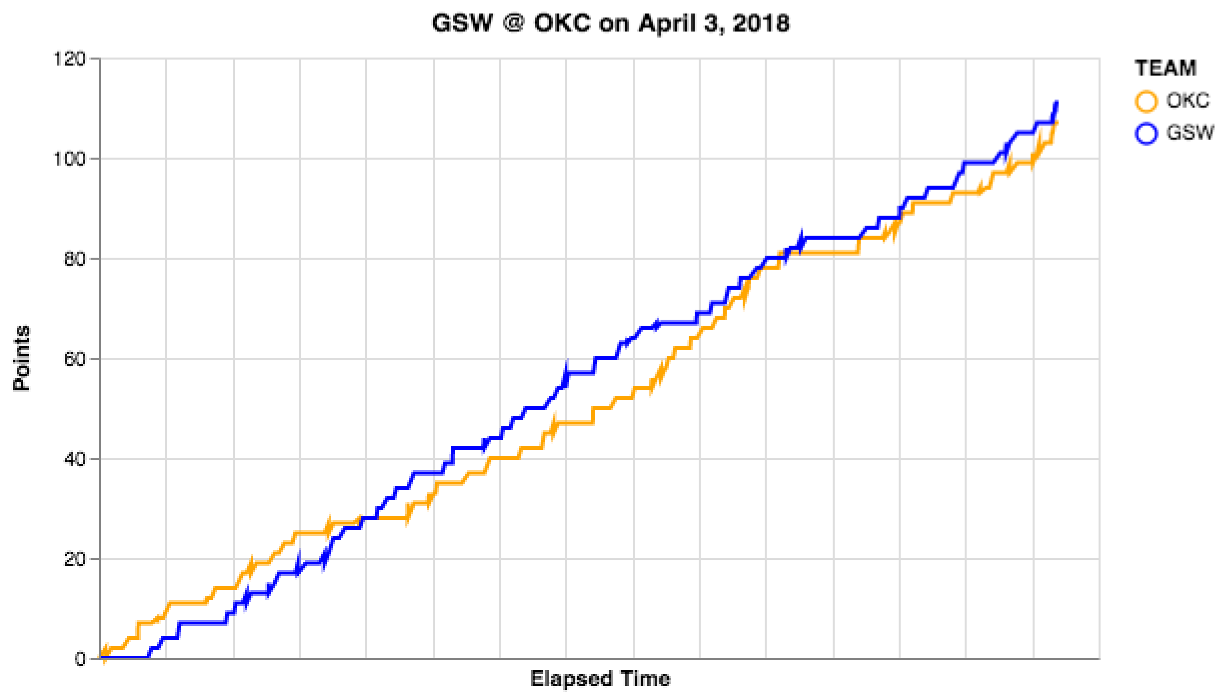


**Figure 18: Warriors at Oklahoma City (April 3, 2018)**

The dataset also allows for a closer look into the performance of the player of the week himself. In order to look into Kevin Durant's performance in this particular game against his previous team, we can look to his "highlight plays" and see where he was able to contribute to the Warriors win (Figure 19).

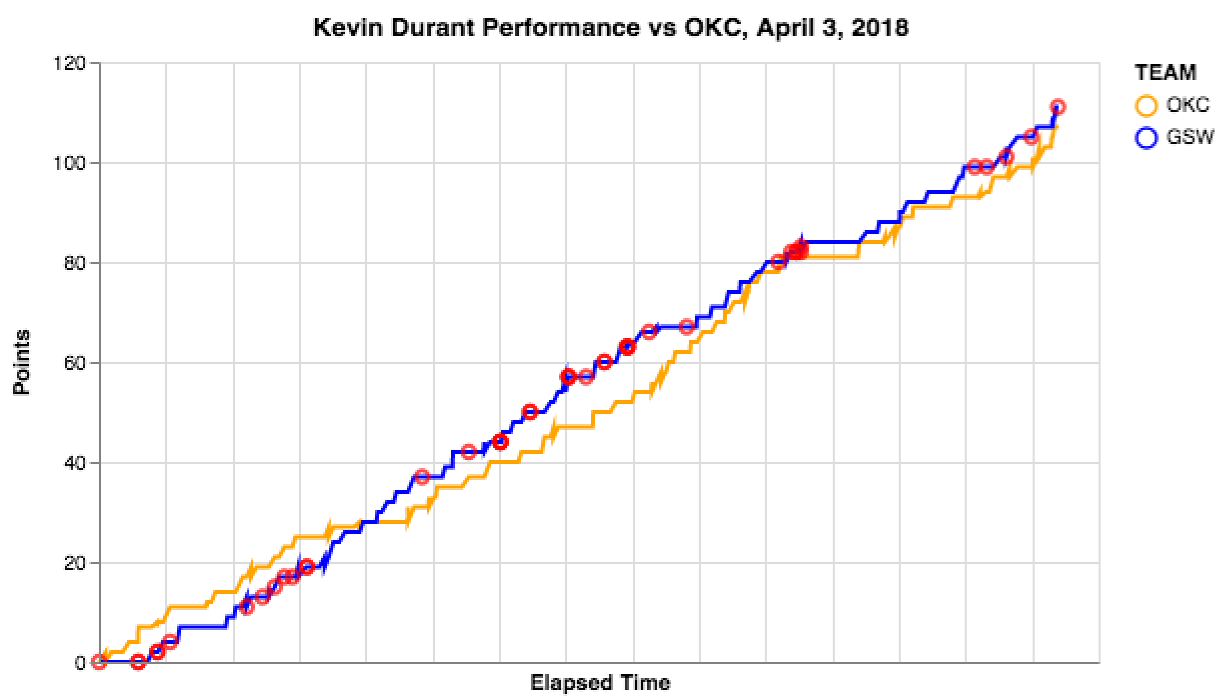


**Figure 19: Kevin Durant Performance Against Oklahoma City (April 3, 2018)**

In the above visual, you can see that Kevin Durant performed well early in the game, taking a break at the end of the first quarter before resuming. He again took a break in the fourth quarter before returning to the game to help deliver the Warriors' win.

## V. Conclusions

The exploratory data analysis discussed herein used a publicly-available dataset from Kaggle containing information on NBA Players of the Week and data extracted from the NBA API to explore two primary questions. The first question addressed was exploring how the physical representation has changed through the history of the NBA. This exploration was guided by supposition from Horace Grant, who claims that changes to the NBA rules have led to more of a "small ball" game, relying more on smaller players. The findings from this analysis are in general agreement with this hypothesis and further studies are proposed in the section below.

The secondary question asked was exploring the relationship between team performance and NBA Players of the Week. This exploration shows that teams with high numbers of Players of the Week tend to have high winning percentages, but teams can also succeed without Players of the Week. Based on some limited assessment, this is posited as being due to the fact that these high performing teams without many Players of the Week (or any) are simply more balanced and get more production from the average player on the team.

## VI. Areas for Continued Investigation

With respect to the investigation between team performance and the number of player of week awards in a season, follow-up work is needed to look in greater detail at team performance throughout a season with the Player(s) of a Week performance. For example, can a Player of the Week rally his team and create a winning momentum through exceptional, consistent play? This evaluation would require a finer detail to the team statistics than is currently in the datasets used, but would be an interesting evaluation. Another question in a similar vein is "Are teams with better records more likely to get their player chosen as Player of the Week?". This question would get at identifying if there may be any bias in the selection process for Player of the Week. To answer this question, one would want to use the week-by-week dataset to bin players by their weekly performance and their team's winning percentage and compare likelihood of being chosen as Player of the Week given similar performance.

## VII. References


1) Player of the Week Data Set - Kaggle - Maintained by J. Baruch
2) NBA API - Hosted on GitHub - S. Patel
3) Horace Grant Blog Post
4) Team Github Repository