# Mining NBA

David Gavrilović

October 27, 2019

# Contents

# 1 Stats 101

## 1.1 Traditional stats

- **Pos** - Position. Traditionaly, position can be one of the following: ***PG*** - point guard, ***SG*** - shooting guard, ***SF*** - small forward, ***PF*** - power forward and ***C*** - center. Nowdays, one player usually plays multiple positions and usually is one of the: ***Point*** - primarly PG, ***Combo guard*** - plays PG and SG, ***Wing*** - SF and SG, ***Forward*** - PF and SF, ***Big*** - usually C but can also be PF.

- **G** - Games. Number of games player played in during a season.

- **GS** - Games started. Number of games player started. Cannot be greater than G.

- **MP** - Minutes played (Per game or total in a season).

- **FG** - Field goals.

- **FGA** - Field goals attempts.

- **FG%** - Field goal percentage. Calculated as FG / FGA.

- **3P** - 3-point field goals.

- **3PA** - 3-point field goal attempts.

- **3P%** - 3-point percentage. Calculated as 3P / 3PA.

- **2P** - 2-point field goals.

- **2PA** - 2-point field goal attempts.

- **2P%** - 2-point percentage. Calculated as 2P / 2PA.

- **FT** - Free throws.

- **FTA** - Free throw attempts.

- **FT%** - Free throws percentage. Calculated as FT / FTA.

- **eFG%** - Field goal percentage that takes into account that a 3-point field goal is, by one point, worth more than a 2-point field goal. Calculated as (FG + 0.5 * 3P) / FGA.

- **ORB** - Offensive rebounds.

- **TRB** - Defensive rebounds.

- **AST** - Assists.

- **STL** - Steals.

- **BLK** - Blocks.

- **TOV** - Turnovers.

- **PF** - Personal fouls.

- **PTS or PPG** - Points or Points per game.

## 1.2  Advanced stats

- **ORtg** - Offensive rating. An estimate of points produced/scored by a player/team per 100 possessions [1].

- **DRtg** - Defensive rating. An estimate of point allowed per 100 possessions [1].

- **PER** - Player efficiency rating. A measure of a per minute production standardized such that the league average is 15 [2].

- **TS%** - True shooting percentage. Points per scoring attempt converted to the 2-point field goal percentage needed to score that many points per attempt. Calculated as PTS / (2 * FGA + 0.44 * FTA).

- **3PAr** - 3-Point attempt rate. Percentage of FGA from 3-point range.

- **FTr** - Free throw attempt rate. Number of FTA per FGA.

- **ORD%** - Offensive rebound percentage. An estimate of the percentage of available offensive rebounds a player grabbed while he was on the floor.

- **DRB%** - Defensive rebound percentage. An estimate of the percentage of available defensive rebounds a player grabbed while he was on the floor

- **TRB%** - Total rebound percentage. An estimate of the percentage of available rebounds a player grabbed while he was on the floor.

- **AST%** - Assist percentage. An estimate of the percentage of teammate field goals a player assisted while he was on the floor.

- **STL%** - Steal Percentage. An estimate of the percentage of opponent possessions that end with a steal by the player while he was on the floor.

- **BLK%** - Block percentage. An estimate of the percentage of opponent two-point field goal attempts blocked by the player while he was on the floor.

- **TOV%** - Turnover percentage. An estimate of turnovers per 100 plays.

- **USG%** - Usage percentage. An estimate of the percentage of team plays used by a player while he was on the floor.

- **OWS** - Offensive win shares. An estimate of the number of wins contributed by a player due to his offense [3].

- **DWS** - Defensive win shares. An estimate of the number of wins contributed by a player due to his defense [3].

- **WS** - Win shares. An estimate number of wins contibuted by a player [3].

- **WS/48** - Win shares per 48 minutes. League average is around 0.100.

- **OBPM** - Offensive Box plus/minus. A box score estimate of the offensive points per 100 possessions that a player contributed above a league-average player, translated to an average team [4].

- **DBPM** - Defensive Box plus/minus. A box score estimate of the defensive points per 100 possessions that a player contributed above a league-average player, translated to an average team [4].

- **BPM** - Box plus/minus. A box score estimate of the points per 100 possessions that a player contributed above a league-average player, translated to an average team [4].

- **VORP** - Value over replacement player. An estimate of the points per 100 team possessions that a player contributed above a replacement-level (-2.0) player, translated to an average team and prorated to an 82-game season [4].

# 2   Calculating prime of an average player

In this section we will try to determine when does an average player hit their *prime*. But first, what is prime? In order to answer that question we must first define something called *peak*. Peak is the best season player had in his career. Knowing that, prime can be defined as player's seasons in which he is close, or somewhat close, to his peak. Knowing length of a player's prime is important, because prime is probably the most important thing when discussing player's career.

In this paper only individual statistics will be used in process of estimating prime years of a player. Because the ultimate stat that can correctly measure when do players hit their prime does not exist (yet), multiple other stats will be used, both traditional and advanced. Those stats are, mostly, PTS, FG and FGA from traditional and PER, WS/48, BPM and VORP from advanced stats. Reasoning behind first three is, well, player in his prime should be best version of himself, so he will probably score and shoot more than in his non-prime seasons. Stats such as PER, BPM and VORP exist so we could determine impact on a game by player, or how good a player is. Prime season should have higher values for this stat than for other seasons. WS per 48 is used to determine players contribution to winning, per 48 minutes, the lenght of a game without overtime. It's also expected for this stat to be higher for seasons in player's prime than for the other seasons.

Of course, different players can hit their prime at different age. Let's check prime years od some of the players inducted, or soon to be inducted, into the Naismith memorial basketball hall of fame. Selected players are chosen without any particular reason. Players used are Bird 1, Shaq 2 and Duncan 3. As we can see from the shown box plots, mentioned players have the best stats when they were around 26 to 30 years old. From that we can conclude that that was the lenght of their prime.

In the next step of determining prime of players, I checked at what age do players win awards, such as *regular season MVP*, *FMVP*, *DPOY* and *Sixth man of the year*. Plots are shown in figure 4. Every plot but plot for Sixth man of the year is somewhat similar. The number of awards players won while in their late twenties and eary thirties is greater than number of awards in other years.
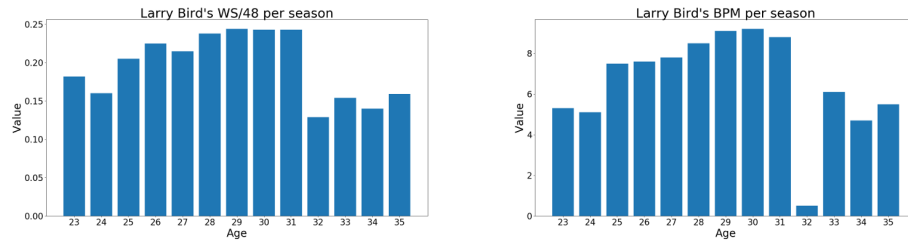
5
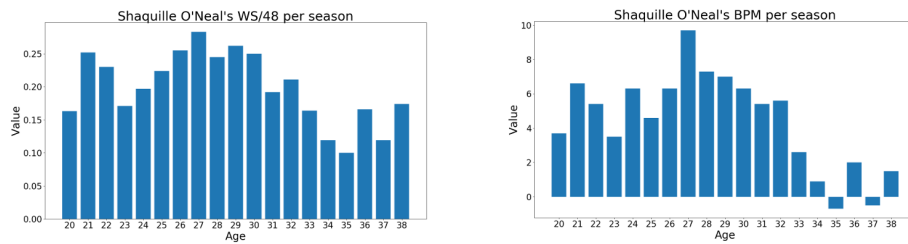
Figure 1: Bird's WS/48 and BPM per season



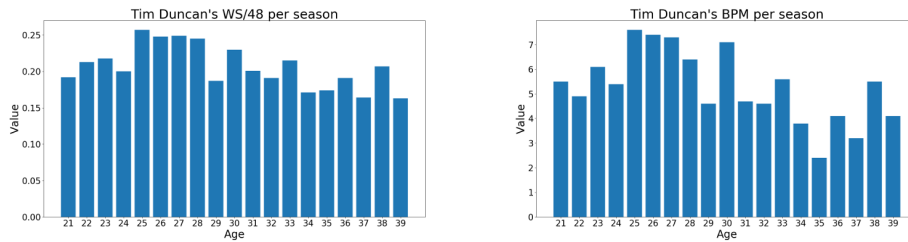Figure 2: Shaq's WS/48 and BPM per season



Figure 3: Duncan's WS/48 and BPM per season

Defensive player of the year award plot is different, where players aged from 23 to 31 won with roughly the same frequency, with an exception of players who were 28 years old when they won, and the ones who were 27. After that there is a significant drop. These plots are showing us that the best players in the season are usually ones between 26 and 31 years old. But those players are stars and superstars. What about an average player?

First, I checked age of players who played in the NBA. Results can be seen in histogram (figure 5). Note that there are two histograms. One represents every player who played in the NBA (in the three point era, from 1979/80 season). The other is filtering out every player (from the same era) that has played in less than 25 games and less then 15 minutes per game. Those players are filtered because I don't consider them regular contributors for the team they
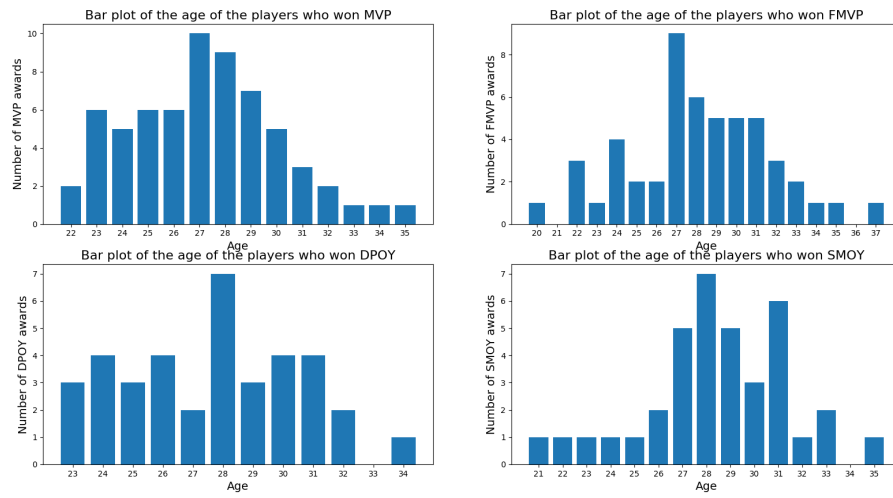
Figure 4: Age of award winners

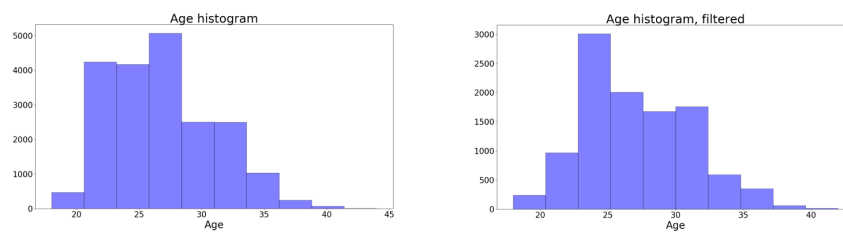are playing for, so they might not be important. Also, I just wanted to check on their histogram as well.



Figure 5: Age histograms

TODO

# References

[1] Basketball-Reference. Calculating individual offensive and defensive ratings. on-line at: https://www.basketball-reference.com/about/ratings.html.

[2] Basketball-Reference. Calculating per. on-line at: https://www.basketball-reference.com/about/per.html.

[3] Basketball-Reference. Nba win shares. on-line at: https://www.basketball-reference.com/about/ws.html.

[4] Basketball-Reference. About box plus/minus (bpm). on-line at: https://www.basketball-reference.com/about/bpm.html.