# Factors Determining NBA Salaries

Info Crunchers

## Background

In an article published by Dario AS, a sports newspaper company, it was revealed that during the 2022-2023 NBA season, "the lowest-paid player in the league is held by Ishmail Wainright of the Phoenix Suns, who earns a total of $633,891" while "the highest-paid player in the league, Steph Curry earns $48,070,014 a year" (Rudder, 2023). What causes this considerable discrepancy between NBA players and their respective salaries? Fortunately, many possible predictors within the NBA can be analyzed to determine whether correlations exist between an NBA player's salary and their statistics. In this report, the three main research question being explored to answer these ongoing disparities are:

- Does a player with a higher true shooting percentage have a higher salary on average?
- Does a player that plays more games per season have a higher salary on average?
- Do older players have a high salary than younger players on average?

Firstly, in a past study, it was discovered that "a higher field goal attempt rate represents the players' offensive importance in the team, [and thus] these players with high field goals will usually be paid more" (Yin, 2022). This proposes a potential relationship between salary and a player's true shooting percentage since that statistic is heavily dependent on field goals. Therefore, it is hypothesized that if the true shooting percentage increases, then player's salary will increase accordingly.

Moreover, in a study published by The Sports Journal, it was brought into attention that many believe that being in the league for a long period of time, indicating more firsthand experiences in play and thus possessing increasingly vital skills, justifies the demands for a higher salary (Sigler, 2018); these firsthand experiences are accumulated through a player participating in more games played per season. This sparks the question whether playing more games per season, developing crucial hands-on skills for their profession, correlates to a higher salary.

Lastly, in a study published by the Atlanta Press, it was uncovered that some evidence points to the conclusion that "for every 1 year increase in the NBA player's age, the player's salary will increase by $0.414 million" (Zhao, 2022). This suggests that NBA players who are older tend to make more money than NBA players who are younger. In order to examine this, this report will continue onto analyzing the trend between explanatory predictors, such as age, with the response variable of player salaries.

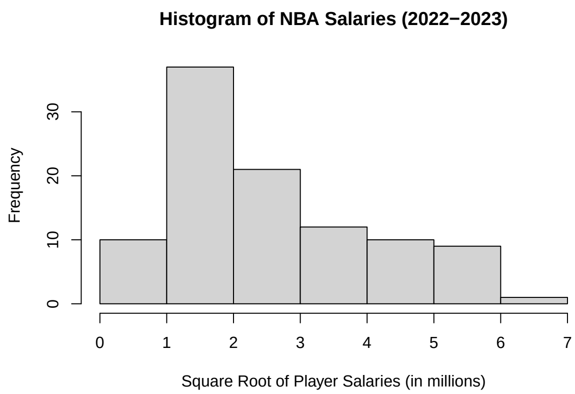
## Data Description

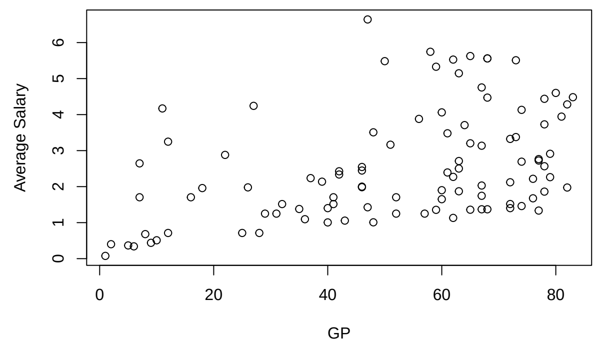
In this report, the data collected consists of NBA players in the 2022-2023 NBA season gathered to reveal possible indicators of NBA annual salaries. The data combines three different datasets from Basketball Reference, for the majority of the advanced statistics; Hoopshype, for NBA player salaries; and the official NBA website, for previously attended colleges of each player. These datasets were utilized because they include valuable statistics such as a player's position, their NBA team, individual shooting percentages, etc., which are all variables that can contribute to NBA salaries. In the original data, the dataset featured over 400 observations, accounting for all NBA players in the 2022-23 season. In this study, a simple random sample of 100 players is extracted from the original observations.

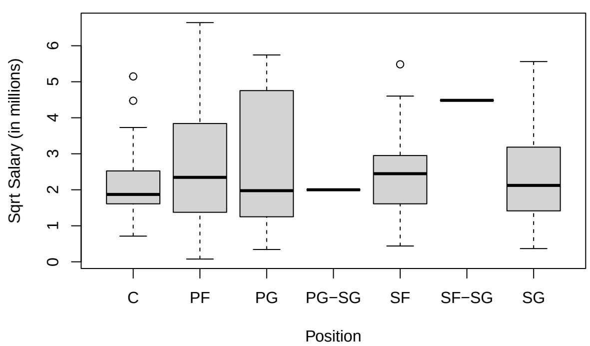
In addition, manipulations made to the dataset include removing repeated predicators throughout the three data sources. Secondly, explanatory variables that measure uninterested values to the research and/or suggest clear multicollinearity were also removed. Thirdly, this report deemed it necessary to transform the response variable from the normal Salary in millions into the square root of those values to better account for the relationships between the response and the predictors. The only potential issue in the finalized dataset is that the sampling method could not be completely representative of the population. Specifically, a few players without complete statistics were removed from the sample consideration. This report also sought to remove any duplicate players that transferred teams from consideration to avoid any confusion within the sample population, however, the examined statistics for a duplicate observation were not necessarily the same. The final concern involves the method of sampling; a different sampling method may have been more representative of the population of NBA players by accounting for more of the variability within the players. Lastly, the sources of the data are all reputable sports news websites, so they are reasonably trustworthy sources.

# Exploratory Data Analysis

```
     [,1]
Age 0.315
TS% 0.304
```




# Conclusion

In the above graphics, the first summary is the histogram for the response variable, player salary in millions. The original graphical representation of a player's annual salary appears to be unimodal and continuous but extremely skewed to the right. In order to make the response variable more suitable for regression analysis, the variable in consideration was transformed by taking the square root of the annual salary results to "pull in" extreme values. Although a right skew is still visibly present, this modification allowed the histogram to appear slightly more symmetric with values more evenly represented in the x-direction.

For the relationships with the explanatory variables, the scatterplot features games played on the x-axis versus the square root of the salary in millions. While there does seem to be a slightly positive linear relationship between the two variables, the observations are displaying a “fanning out” trend as x increases; this indicates that the relationship between the two is not necessarily strong. Similarly, in the numerical summaries of the correlation table, the correlations for the variables of interest, age and true shooting percentage, are 0.315 and 0.304 respectively. These values suggest weak positive relationships with salaries. In terms of the age variable, there is likely not enough data sampled on different ages to fully understand whether a linear relationship exists with salaries. A majority of the current players range from late 20s to early 30s, and since it is a simple random sample of 100 observations, the few older players may not have been fully accounted for when compared to the potential results if a cluster sampling method was preformed. For the true shooting percentage, the weak relationship with salaries may encourage further evaluations of a different variable, but the correlations are not determined to be insignificant until the model-building step where its relationships with multiple variables are considered, so they are still kept.

Lastly, among the graphics, there is a box plot between the qualitative variable, player positions, and their squared root salary. Although the few players with multiple positions, PG-SG and SF-SG, have a lack of data available, the results from the other individual playing positions are noteworthy. The salaries for the center, small forward, and shooting guard exhibit a pretty similar distribution for salaries with 50% of the players earning around two million to six million yearly. Two outliers exist in the center position earning around 20 and 25 million each, and one outlier exists in the small forward position earning slightly above 25 million dollars annually. In addition, while the median salary for all positions is around four million, the point guard position is notable for its large disparity within the upper quartile. This variation in position salaries stipulates a potential importance that can be further analyzed in the multiple regression model.

Ultimately, since the explanatory variables of interest are not extremely strong compared to the other variables available, to continue, the report should consider including some stronger predictors, such as games started (r=0.621), win share (r=0.659), and value over replacement player (r=0.659), to ensure a more meaningful statistic summary from the regression.

## Appendix A: Data Dictionary

| Variable Name | Abbreviated Name | Description |
|---|---|---|
| Player | Player | Individual players in the NBA 2022-23 season |
| Salary | sqrtSalary | Square root of the player's salary in millions of U.S. dollars |
| College | College | U.S. university the player last attended |
| Position | Position | Player's NBA team position; C=center, PF=power forward, PG=point guard, SG=shooting gaurd, SF=small forward |
| Team | Team | Player's NBA team |
| Age | Age | Player's age as of February 1st of the 2022-23 season |
| Games Played | GP | Number of games the player played in |
| Games Started | GS | Number of games which the player started in |
| Minutes Played | MP | Number of total minutes the player played |
| Minutes Per Game | MPG | Average minutes per game played |
| Player Efficiency Rating | PER | Measure of per-minute production standardized such that the league average is 1 |
| True Shooting Percentage | TS% | Measure of shooting efficiency accounting 2-point field goals, 3-point field goals, and free throws. |

| Variable Name | Abbreviated Name | Description |
|---|---|---|
| Free Throw Percentage | FT% | Percent of free throws made to attempted |
| 3 Pointer Percentage | 3P% | Percent of field goal attempts from 3-point range |
| Usage Rate Percentage | UR% | Estimate of percentage of team plays used by a player while they were on the floor |
| Offensive Win Share | OWS | Estimate of the number of wins contributed by a player on offense |
| Defensive Win Share | DWS | Estimate of the number of wins contributed by a player on defense |
| Win Share | WS | Estimate of the number of wins contributed by a player |
| Value Over Replacement Player | VORP | Box score 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 |
| Field Goal Made | FGM | Number of average field goals made per game |
| Field Goal Attempted | FGA | Number of average field goals attempted per game |

## Appendix B: Data Rows

| | Player | sqrtSalary | College | Position | Team | Age | GP | GS | MP | MPG |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Cameron Johnson | 2.426499 | North Carolina | PF | TOT | 26 | 42 | 41 | 1199 | 28.5 |
| 2 | Pat Connaughton | 2.393406 | Notre Dame | SF | MIL | 30 | 61 | 33 | 1443 | 23.7 |
| 3 | Kevin Love | 5.527836 | UCLA | PF | TOT | 34 | 62 | 20 | 1240 | 20.0 |
| 4 | Anthony Gill | 1.355024 | Virginia | PF | WAS | 30 | 59 | 8 | 624 | 10.6 |
| 5 | Keita Bates-Diop | 1.370664 | Ohio State | PF | SAS | 27 | 67 | 42 | 1452 | 21.7 |
| 6 | KZ Okpala | 1.379178 | Stanford | PF | SAC | 23 | 35 | 3 | 248 | 7.1 |

| | PER | TS% | FT% | 3P% | UR% | OWS | DWS | WS | VORP | FGM | FGA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 17.1 | 0.617 | 0.842 | 0.404 | 20.6 | 2.1 | 1.4 | 3.5 | 1.4 | 5.3 | 11.3 |
| 2 | 10.0 | 0.531 | 0.659 | 0.339 | 13.8 | 0.7 | 1.8 | 2.5 | 0.3 | 2.7 | 6.9 |
| 3 | 13.1 | 0.548 | 0.879 | 0.334 | 19.2 | 0.6 | 2.0 | 2.6 | 0.8 | 2.7 | 6.8 |
| 4 | 11.1 | 0.604 | 0.731 | 0.138 | 12.3 | 1.0 | 0.4 | 1.3 | -0.2 | 1.2 | 2.2 |
| 5 | 14.9 | 0.609 | 0.793 | 0.394 | 16.7 | 2.1 | 0.6 | 2.7 | 0.5 | 3.5 | 6.9 |
| 6 | 6.2 | 0.554 | 0.875 | 0.333 | 8.4 | 0.1 | 0.2 | 0.3 | -0.1 | 0.5 | 1.1 |

## Appendix C: References

1. Your resources go here in APA styling. Hanging indent is not required. Links should be surrounded by <>.

HoopsHype. (2023). NBA Player Salaries. HoopsHype. Retrieved October 15, 2023, from https://hoopshype.com/salaries/players/

NBA. (2023). NBA Advanced Stats. NBA.com. Retrieved October 15, 2023, from https://www.nba.com/stats/players/traditional?SeasonType=Playoffs

Rudder, P. (2023, May 26). Who is the lowest paid NBA player in the 2022-2023 season?. Diario AS. https://en.as.com/nba/who-is-the-lowest-paid-nba-player-in-the-2022-2023-season-n/

Sigler, K. (2018, June 19). NBA players' pay and performance: What counts? The Sport Journal. https://thesportjournal.org/article/nba-players-pay-and-performance-what-counts/

Sports Reference. (2023, April 10). 2022-23 NBA Player Stats: Advanced. Basketball Reference. Retrieved October 15, 2023, from https://www.basketball-reference.com/leagues/NBA_2023_advanced.html

Yin, C., & Jamin, W. (2022). How do different factors affect the NBA players' salary. https://rstudio-pubs-static.s3.amazonaws.com/928945_ef3cd002eb2140e0a0875b74f805518f.html

Zhao, Y. (2022). Model prediction of factors influencing NBA players' salaries based on multiple linear regression. Proceedings of the 2022 2nd International Conference on Economic Development and Business Culture (ICEDBC 2022), 1439–1445. https://doi.org/10.2991/978-94-6463-036-7_213