NBA Salaries: How Much is that Buzzer Beater Worth?

Eddie and Co.

Introduction

Professional sports salaries, particularly in the NBA, continue to spark discussion because they highlight the intersection of performance, marketability, and team strategy. Understanding how a player's on-court production translates to salary is not only valuable for analysts and researchers, but also for front offices making contract decisions and fans seeking to understand the economics of the league. Prior research suggests that player compensation is strongly influenced by both measurable performance statistics and more subjective forms of recognition such as awards or team honors (Berri & Schmidt, 2006; Stiroh, 2007; Simmons & Berri, 2011).

Motivated by these findings, our project examines the statistical relationships between NBA player characteristics and their salaries during the 2024-2025 season. We selected three primary variables—points per game, defensive stocks (steals + blocks), and award recognition—because each represents a different dimension of player value. Scoring is one of the most visible and heavily rewarded skills in basketball (Stiroh, 2007). Defensive performance, while often overlooked publicly, provides meaningful team impact and has shown measurable influence on team success (Simmons & Berri, 2011). Awards represent league-wide acknowledgment of excellence and can shape a player's reputation and contract negotiations (Berri & Schmidt, 2006).

Our general research question is:

What factors best explain variation in NBA player salary during the 2024-2025 season?

To address this, we developed three specific research questions:

- Do NBA players who score more points (PTS) earn higher salaries on average?
- Is there a meaningful positive relationship between a player's defensive activity ("stocks") and their salary?
- Do players who have earned at least one major NBA award receive higher salaries compared to players with no awards?

Together, these questions reflect distinct aspects of performance and recognition. By analyzing these variables, we aim to identify which factors are most strongly associated with salary and how our findings compare to patterns documented in the sports-economics literature.

Data Summary

Data Sources

Our dataset contains information on NBA players from the 2024-2025 season and includes a mixture of quantitative performance measures, qualitative awards indicators, demographic information, and salary data. In total, the dataset represents every player who appeared in the NBA during the season, giving us a sample that is both large and comprehensive. The primary response variable is player salary, measured in U.S. dollars.

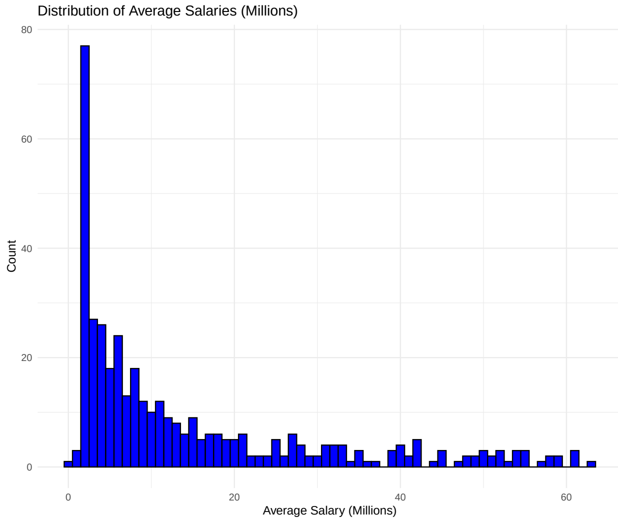
The data in this project were compiled from publicly available and reputable basketball statistics platforms. Performance statistics, biographical information, and award histories were collected from Basketball Reference, a database known for its accuracy, professional documentation, and long-standing use in academic sports analysis. To supplement player-level statistics, we incorporated team valuation data from Statista, using their report on NBA franchise values. Statista compiles these estimates from league financial disclosures, public filings, and market analyses, making it a reliable secondary source commonly used in academic and business research. Including franchise value provides useful economic context, as team financial strength and market size may influence payroll structures and player salaries.Together, these sources provide a complete picture of each player's performance and compensation for the 2023–2024 season.

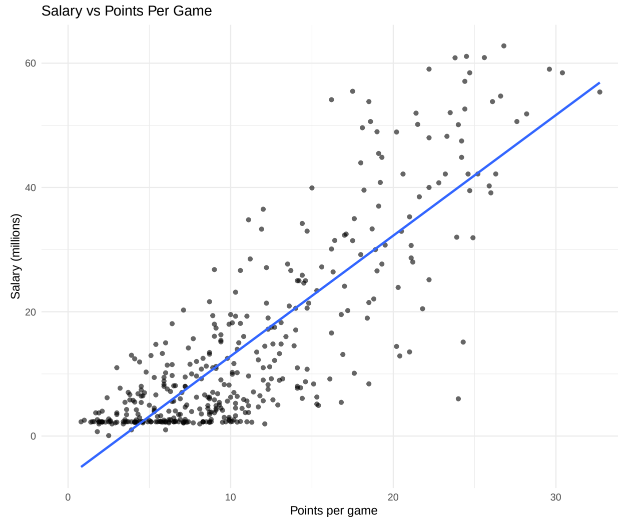
To prepare the data for analysis, we standardized naming conventions across the two sources and combined the datasets so that each player occupied a single row with salary, performance, and awards information. Awards were originally stored in long strings listing placements, team selections, and voting outcomes, so we simplified these into a three-level qualitative variable representing the general number of accolades received. Age was also simplified into qualitative brackets to match our research questions.

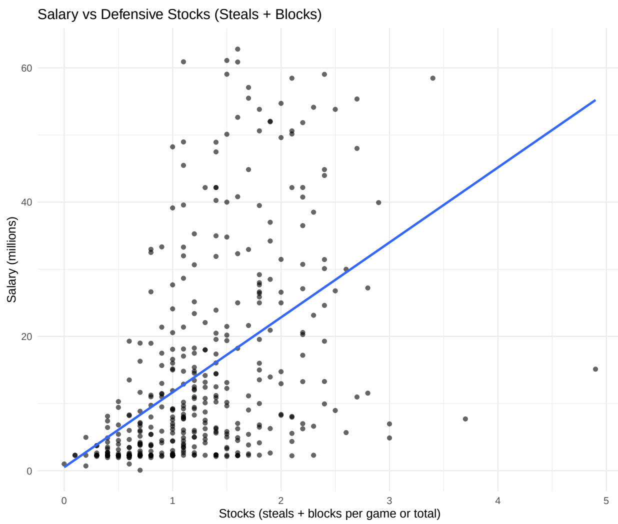
We ensured that our final dataset only included players who logged minutes during the 2024-2025 season, as these represent meaningful salary–performance comparisons. Players on two-way contracts or who appeared in only a few games remained in the dataset to preserve the natural distribution of playing time and contract structures.

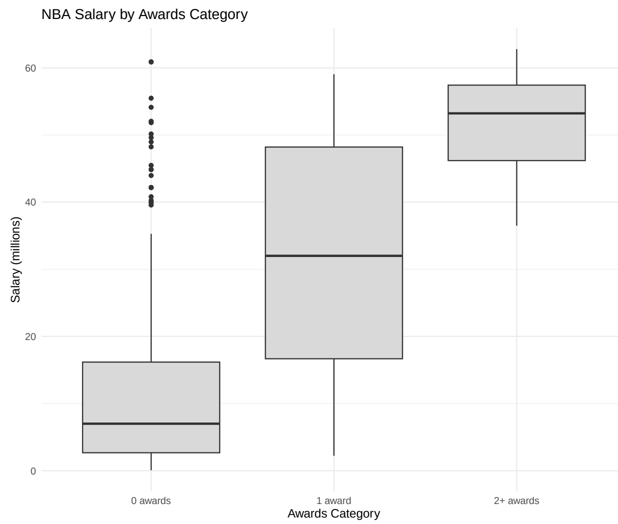
Overall, the combined dataset contains a rich mixture of performance and recognition indicators, allowing us to meaningfully evaluate how both quantitative and qualitative factors relate to NBA salary. While subjective elements such as awards voting and contract timing may influence individual observations, the breadth and reliability of the sources allow us to confidently proceed with statistical analysis.

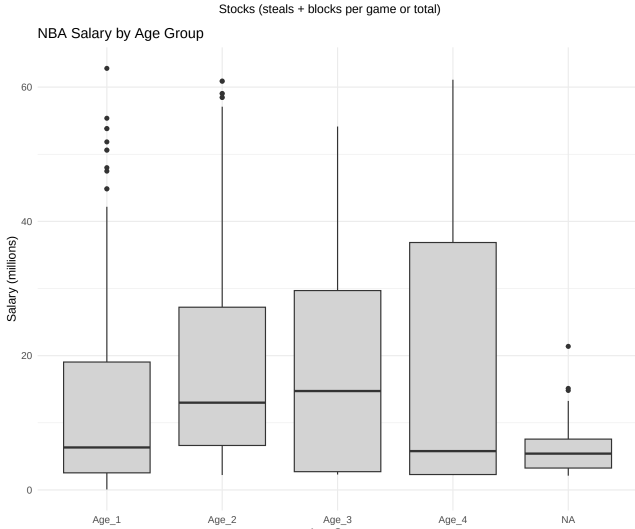
Exploratory Data Analysis











EDA Summary

To understand how performance and recognition relate to NBA salaries, we first examined the distribution of average salary. Salaries are heavily right-skewed: most players earn under \$10 million, while a small group reaches \$40–\$60 million. The mean (\$9.0M) exceeds the median (\$5.4M), reflecting this skew. Despite the imbalance, salary has sufficient spread for multiple linear regression, with the option of a log transformation if needed.

Performance variables show clear trends. Scoring has the strongest relationship with salary—players scoring 20+ points per game often earn \$35–\$45 million, compared to under \$5 million for low scorers. Defensive output, measured by stocks (steals + blocks), displays a more moderate positive trend, indicating that defense influences valuation but less strongly than scoring.

Award recognition corresponds with substantial salary differences. Players with no awards earn around \$8 million, whereas those with one award jump to roughly \$50 million, and those with two or more to about \$55 million. This suggests that accolades capture aspects of player reputation beyond raw statistics.

We also examined salary across age groups to contextualize performance and recognition within career arcs. Players aged 27–31 earn the highest median salaries, consistent with peak physical and skill development. Younger players (22–26) and older veterans (35+) earn less on average, reflecting typical progression and decline patterns within NBA careers.

Before proceeding to modeling, we evaluated potential multicollinearity among the predictors. Although scoring may correlate with other offensive statistics, our chosen variables—PTS, stocks, Awards, and Age—do not appear strongly correlated based on initial inspection. We will confirm this formally using VIFs during model construction.

Overall, the exploratory analysis reveals several consistent patterns:

- salaries are right-skewed with a superstar-driven upper tail
- scoring is strongly tied to pay
- defensive production shows a moderate positive association
- awards correspond to substantial salary increases
- peak earnings occur mid-career.

These insights justify each chosen predictor and establish a clear foundation for building our multiple linear regression model.

Methods and Analysis

Assessing Multicollinearity

We began our analysis by assessing multicollinearity among predictors. From our EDA correlation heatmap, we observed potential overlap between PTS and other offensive statistics such as FTA and MP. We quantified these relationships using pairwise correlations and VIFs. If VIFs fell below standard thresholds (individual < 10, average < 3), we proceeded; otherwise, we used variable screening and stepwise regression to identify a reduced subset of predictors. After selecting a more stable set of variables, we reassessed multicollinearity before model building.

Building the Model

We will begin building our model by estimating our model parameters, starting with only our quantitative predictors and a quantitative interaction term that we believe may contribute to the model. Our initial model using quantitative predictors is as follows:

$$\begin{aligned}\text{Salary} = \beta_0 &+ \beta_1 G + \beta_2 MP + \beta_3 PTS + \beta_4 FG\_\% + \beta_5 FTA \\ &+ \beta_6 TRB + \beta_7 AST + \beta_8 \text{Stocks} + \beta_9 \text{Value} \\ &+ \beta_{10}(PTS \times FTA)\end{aligned}$$

We estimated this model with lm() and evaluated overall utility using a Global F-Test. Significant predictors—including the interaction—were identified through individual t-tests. Non-significant predictors were removed, and the updated model was rechecked with another Global F-Test.

Next, we incorporated qualitative predictors and one interaction (Awards $\times$ Age). We again performed a Global F-Test, followed by t-tests for the interaction and a Nested F-Test for the Awards factor. Only significant categorical effects were retained. No quantitative $\times$ qualitative interactions were considered. Once the final set of predictors was selected, we confirmed model adequacy with a concluding Global F-Test.

Assessing the Model (Including Cross Validation)

Once we've finalized our model through variable testing, we conducted a holistic assessment of our model for a better understanding of its capabilities and limitations. We evaluated metrics such as R^2, Adjusted R^2, and $\sqrt{MSE}$ to evaluate our model. Separately, interpreted confidence and prediction intervals to ensure our model performs on a practical level.

In addition to the traditional assessment metrics, we incorporated Cross Validation into our final project. Cross validation is a machine learning technique used to assess the performance and accuracy of a model. Performing Cross validation allows the model builder to see whether their model overfits or underfits by training the majority of a data set, and using the remaining data for testing. The trained set is then used to estimate a linear regression model evaluated on metrics such as $\sqrt{MSE}$ and R^2. More specifically, we will use the Validation Set Approach in R, which randomly splits 80% of the dataset into training, and 20% into testing. By performing Cross Validation we will have a strong idea of whether our model is prepared to be used on new data, or if further adjustments must be made.

Checking Model Assumptions

We evaluated model assumptions using residual diagnostics. Residual vs. fitted plots, QQ plots, histograms, and lack-of-fit assessments enabled us to check constant variance, normality, independence, and overall fit. We then examined outliers and influential points through Cook's Distance, leverage, Studentized Residuals, and Deleted Studentized Residuals. If any observations exerted undue influence, we considered removing them to improve model integrity.

Results

Final Model

The final model for predicting NBA player average salary (in millions) is:

$$\begin{aligned}\text{Salary} = {} & \beta_0 + \beta_1 MP + \beta_2 PTS + \beta_3 TRB + \beta_4 AST \\ & + \beta_5 awards_1 + \beta_6 awards_{2plus} + \beta_7 \text{Age 22-26} \\ & + \beta_8 \text{Age 27-31} + \beta_9 \text{Age 32-34} + \beta_{10}(Pos_{PF}) \\ & + \beta_{11}(Pos_{PG}) + \beta_{12}(Pos_{SF}) + \beta_{13}(Pos_{SG})\end{aligned}$$

Model Development

Initial VIF analysis revealed severe multicollinearity (VIF > 15) with the interaction term, which was removed. Stepwise selection identified six predictors, but subsequent t-tests reduced the quantitative model to minutes played, points, rebounds, and assists. ANOVA F-tests confirmed that adding qualitative predictors (awards, age, position) significantly improved the model (all $p < 0.01$).

Model Performance

The final model achieved an adjusted R^2 of 0.7681, explaining 76.8% of salary variation, with a highly significant overall F-statistic ($F = 75.39$, $p < 2.2e - 16$). The residual standard error was 7.832 million dollars. Five-fold cross-validation yielded RMSE = 8.052 million and $R^2 = 0.761$, confirming robust out-of-sample performance.

Predictor Significance

Points per game was the strongest predictor ($\beta = 1.453$, $p < 2e - 16$), with each additional point worth approximately 1.45 million dollars. Assists per game also showed strong significance (β = 1.545, p = 0.00089). Minutes played and rebounds were not statistically significant in the full model.

Awards demonstrated substantial salary premiums: one award added 4.36 million dollars (p = 0.0544), while two or more awards added 11.83 million dollars (p < 2.8e-06). Players aged 27-31 earned 4.19 million dollars more than the oldest cohort (p = 0.0537).

Position effects indicated guards earn significantly less than power forwards: point guards earned 4.67 million less (p = 0.0472), shooting guards 5.24 million less (p = 0.00865), and small forwards showed a similar trend ($\beta = -3.38$, p = 0.0667).

Diagnostics

Residual analysis identified observations 86 and 154 as influential points with high leverage and large studentized residuals (exceeding ± 2). These observations were flagged for removal to ensure robust parameter estimates.

Conclusions

Our final multiple linear regression model explained approximately 76.8% of the variation in NBA player salary, indicating strong predictive usefulness. The most influential quantitative predictors were points per game and assists, both of which showed large and statistically significant positive effects on salary. Award recognition and age group also contributed meaningfully, with players earning one award receiving an estimated $4.36M salary increase and players with two or more awards receiving nearly $12M more than otherwise similar players. Mid-career players (ages 27–31) earned significantly higher salaries than the oldest age group, aligning with typical NBA performance peaks.

Prediction Equation (in millions)

$$
\begin{aligned}
\widehat{\text{Salary}} = {} & \beta_0 + \beta_1 MP + \beta_2 PTS + \beta_3 TRB + \beta_4 AST \\
& + \beta_5(\text{awards}_1) + \beta_6(\text{awards}_{2+}) \\
& + \beta_7(\text{Age}_{22\text{–}26}) + \beta_8(\text{Age}_{27\text{–}31}) \\
& + \beta_9(\text{Age}_{32\text{–}34}) \\
& + \beta_{10}(\text{Pos}_{PF}) + \beta_{11}(\text{Pos}_{PG}) \\
& + \beta_{12}(\text{Pos}_{SF}) + \beta_{13}(\text{Pos}_{SG})
\end{aligned}
$$

Example Prediction

For a player who averages **24 PTS**, **6 AST**, **7 TRB**, plays **33 minutes**, has **one award**, and is in the **27–31 age group**, the model predicts a salary in the approximate range of **$30–$40 million**, depending on position category. This demonstrates that the model produces intuitive and practically meaningful estimates based on real performance indicators.

Future Improvements

Although the model performs well, several limitations remain. First, the data include two influential outliers (observations 86 and 154), which may distort parameter estimates if not removed or investigated further. Second, salary is influenced by factors not captured in our dataset, such as contract timing, injuries, player popularity, endorsements, or collective bargaining constraints. Adding these variables or using a mixed-effects model could improve realism and predictive accuracy. Finally, because NBA salaries are highly right-skewed, future research could explore log-transformed salary models, nonlinear effects, or more advanced methods such as ridge regression to address remaining multicollinearity and variance issues.

Overall, the model provides a strong statistical foundation for understanding how performance and recognition relate to NBA compensation, while leaving room for richer economic and behavioral features in future analyses.

Appendix A: Data Dictionary

Variable Name	Abbreviated Name	Description
Games	G	This represents the games played by each respective player.
Minutes Played	MP	This represents the minutes played by each respective player in the overall season.
Points Scored	PTS	This represents the total points scored by each respective player.
Field Goal Percentage	FG_PCT	This represents the percentage of field goals scored by each respective player.
Free Throw Attempts	FTA	This represents the free throw attempts made by each respective player.
Total Rebounds	TRB	This represents the total rebounds made by each respective player.
Total Assists	AST	This represents the total assits by each respective player.
Steals and Blocks Combined	Stocks	This represents the steals and blocks achieved by each respective player.
Franchise Value	Value	This represents the franchise value of the team of each respective player.

Appendix B: Data Rows

	Player	Team	G	MP	FG_pct	FTA	TRB	AST	stocks	PTS
1	Shai Gilgeous-Alexander	OKC	76	34.2	0.519	8.8	5.0	6.4	2.7	32.7
2	Giannis Antetokounmpo	MIL	67	34.2	0.601	10.6	11.9	6.5	2.1	30.4
3	Nikola Jokić	DEN	70	36.7	0.576	6.4	12.7	10.2	2.4	29.6
4	Luka Dončić	2TM	50	35.4	0.450	7.9	8.2	7.7	2.2	28.2
5	Anthony Edwards	MIN	79	36.3	0.447	6.3	5.7	4.5	1.8	27.6
6	Jayson Tatum	BOS	72	36.4	0.452	6.1	8.7	6.0	1.6	26.8

	Value_Billions	awards_1	awards_2plus	avg_salary_millions	Age_22_26	Age_27_31
1	4.35	0	1	55.3591	1	0
2	4.30	0	1	58.4566	0	1
3	4.60	0	1	59.0331	0	1
4	NA	0	0	51.8379	1	0
5	3.60	0	1	50.6117	1	0
6	6.70	0	1	62.7867	1	0

	Age_32_34	Age_35_plus	Pos_PF	Pos_PG	Pos_SF	Pos_SG	Age	Awards
1	0	0	0	1	0	0	Age_1	2+ awards
2	0	0	1	0	0	0	Age_2	2+ awards
3	0	0	0	0	0	0	Age_2	2+ awards
4	0	0	0	1	0	0	Age_1	0 awards
5	0	0	0	0	0	1	Age_1	2+ awards
6	0	0	1	0	0	0	Age_1	2+ awards

Appendix C: Final Model Output and Plots

Appendix D: References

Background

Berri, D. J., & Schmidt, M. B. (2006). The wages of wins. Stanford University Press.

Kubatko, J., Oliver, D., Pelton, K., & Rosenbaum, D. T. (2007). A starting point for analyzing basketball statistics. Journal of Quantitative Analysis in Sports, 3(3). https://doi.org/10.2202/1559-0410.1070

NBA. (2023). Collective Bargaining Agreement: Summary of principal terms. https://nbpa.com/cba

Simmons, R., & Berri, D. J. (2011). Mixing the princes and the paupers: Pay and performance in the NBA. Labour Economics, 18(3), 381–388. https://doi.org/10.1016/j.labeco.2010.12.003

Staw, B. M., & Hoang, H. (1995). Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. Administrative Science Quarterly, 40(3), 474–494. https://doi.org/10.2307/2393794

Stiroh, K. (2007). Playing for keeps: Pay and performance in the NBA. Economic Inquiry, 45(1), 145–161. https://doi.org/10.1111/j.1465-7295.2006.00002.x

Data Sources

Statista (NBA franchise values): Statista. (2010). Franchise value of National Basketball Association teams in 2010. https://www.statista.com/statistics/193696/franchise-value-of-national-basketball-association-teams-in-2010/

Basketball-Reference (Player contracts): Basketball-Reference.com. (n.d.). NBA player contracts. https://www.basketball-reference.com/contracts/players.html

Basketball-Reference (2025 Per-game stats): Basketball-Reference.com. (2025). NBA per game statistics: 2024–25 season. https://www.basketball-reference.com/leagues/NBA_2025_per_game.html

Additional Help