NBA Player Salary Analysis: Relationships Between Performance, Experience, and Compensation

Khang Le-Benson Zheng-Joseph Doan

2025-04-27

Introduction

In professional basketball, player salaries represent one of the most significant investments that NBA teams make. Understanding the factors that influence player compensation is crucial for both team management and fans to evaluate whether players are being paid appropriately for their contributions. This analysis explores the relationship between NBA player performance, demographic characteristics, and salary compensation during the 2022-2023 season.

Our analysis seeks to answer several key research questions:

1. How do player age and experience influence salary levels in the NBA?
2. Are there significant differences in salary distribution across different player positions?
3. How do teams allocate their salary resources relative to the league salary cap?
4. Which players are receiving compensation closest to the maximum amounts permitted under the NBA's collective bargaining agreement?

By exploring these questions, we aim to provide insights into the economics of NBA player compensation and identify patterns that might help in evaluating player value relative to their contracts.

Data Provenance

Data Sources and Collection

For this analysis, we utilized two primary datasets from Kaggle:

1. **NBA Historical Statistics (1947-2025)**: This comprehensive dataset contains player performance metrics spanning from the NBA's inception through the 2022-2023 season. The data was originally collected by Basketball Reference (www.basketball-reference.com), which serves as the official historical statistics partner of the NBA. Basketball Reference collects these statistics directly from official NBA game data, providing a reliable and authoritative source of information.
2. **NBA Salaries (2022-2023)**: This dataset contains salary information for NBA players during the 2022-2023 season. The data was compiled from publicly available contract information published by Basketball Reference and other reputable sports data providers. These salary figures reflect the players' contractual agreements as reported to the NBA according to the Collective Bargaining Agreement (CBA) between the NBA and the National Basketball Players Association.

These datasets were selected because they provide comprehensive, accurate, and up-to-date information on NBA player statistics and compensation. The basketball statistics include standard performance metrics used throughout the basketball analytics community, while the salary data represents official contractual information.

Purpose of the Original Data Collection

The original purpose of these datasets was to document and track player performance and compensation for:

1. **Historical record-keeping**: Preserving the statistical achievements and financial records of NBA players and teams.
2. **Performance analysis**: Enabling teams, analysts, and fans to evaluate player contributions.
3. **Market analysis**: Allowing teams and agents to benchmark player compensation against peers with similar performance metrics.
4. **Transparency**: Providing public visibility into professional sports compensation, which is often a matter of public interest.

Cases and Observations

The datasets represent individual NBA players as the primary observational units. Each player has associated statistical performance data and salary information. For the 2022-2023 season specifically:

- The player statistics dataset contains information on all active NBA players during the season, including rookies and veterans.
- The salary dataset captures contractual compensation for all players signed to NBA teams during the 2022-2023 season.

FAIR Data Principles Compliance

Our analysis adheres to the FAIR data principles (Findable, Accessible, Interoperable, and Reusable) [@Wilkinson2016] in the following ways:

Findable

The datasets we used are publicly findable through well-established basketball reference websites and sports data repositories. Both datasets include unique player identifiers and team designations that make specific records easy to locate. The original data sources maintain consistent naming conventions that facilitate search and retrieval of player information across seasons.

Accessible

The data is freely accessible to the public through:

- Basketball Reference website (www.basketball-reference.com)
- Sports data repositories like Kaggle
- NBA's official statistical database (www.nba.com/stats)

These platforms provide open access to the data without requiring specialized authentication beyond basic website registration in some cases.

Interoperable

The datasets maintain strong interoperability through:

- Standardized player naming conventions
- Consistent statistical categories across seasons
- Uniform formatting of performance metrics
- Common identifiers (player names, team abbreviations) that allow joining across different datasets

This standardization allows our analysis to integrate different data sources seamlessly and facilitates comparison with other basketball analytics projects.

Reusable

The data meets reusability standards through:

- Clear documentation of statistical definitions
- Comprehensive metadata about seasons, teams, and measurement methods
- Transparent provenance tracking back to official league sources
- Well-defined statistical categories with consistent meanings over time

These characteristics make the data suitable for reuse in various analytical contexts beyond our specific research questions.

Data Attributes

Our analysis focuses on specific attributes from the datasets that provide insights into the relationships between player characteristics, performance, and compensation. The key attributes include:

Player Demographics

- **player**: Player name (string) - The unique identifier for each player
- **age**: Player age in years (numeric) - Age at the beginning of the 2022-2023 season
- **experience**: Years played in the NBA (numeric) - Experience prior to the 2022-2023 season
- **pos**: Player position (categorical) - The primary position played (PG, SG, SF, PF, C)
- **Team**: Team affiliation (categorical) - The NBA franchise the player belongs to

Performance Metrics

While our initial code was unable to access some of the advanced performance metrics, a more complete analysis would include metrics such as:

- **GP**: Games played (numeric) - Number of games participated in during the 2022-2023 season
- **MP**: Minutes played (numeric) - Total minutes played during the season
- **PTS**: Points per game (numeric) - Average points scored per game

Compensation Data

- **Salary**: Annual salary in dollars (numeric) - The player's contractual compensation for the 2022-2023 season

Derived Attributes

We created several derived attributes to facilitate our analysis:

- **experience_tier**: Categorical grouping of players by experience level
- **age_group**: Categorical grouping of players by age ranges
- **max_tier**: Categorization of players by maximum contract eligibility
- **pct_of_max**: Percentage of maximum possible salary based on experience tier

These derived attributes help identify patterns across different career stages and compensation levels.

Exploratory Data Analysis

Salary Distribution Overview

Figure 1 shows the distribution of player salaries across the NBA during the 2022-2023 season, with a focus on the top 10 highest-paid players.

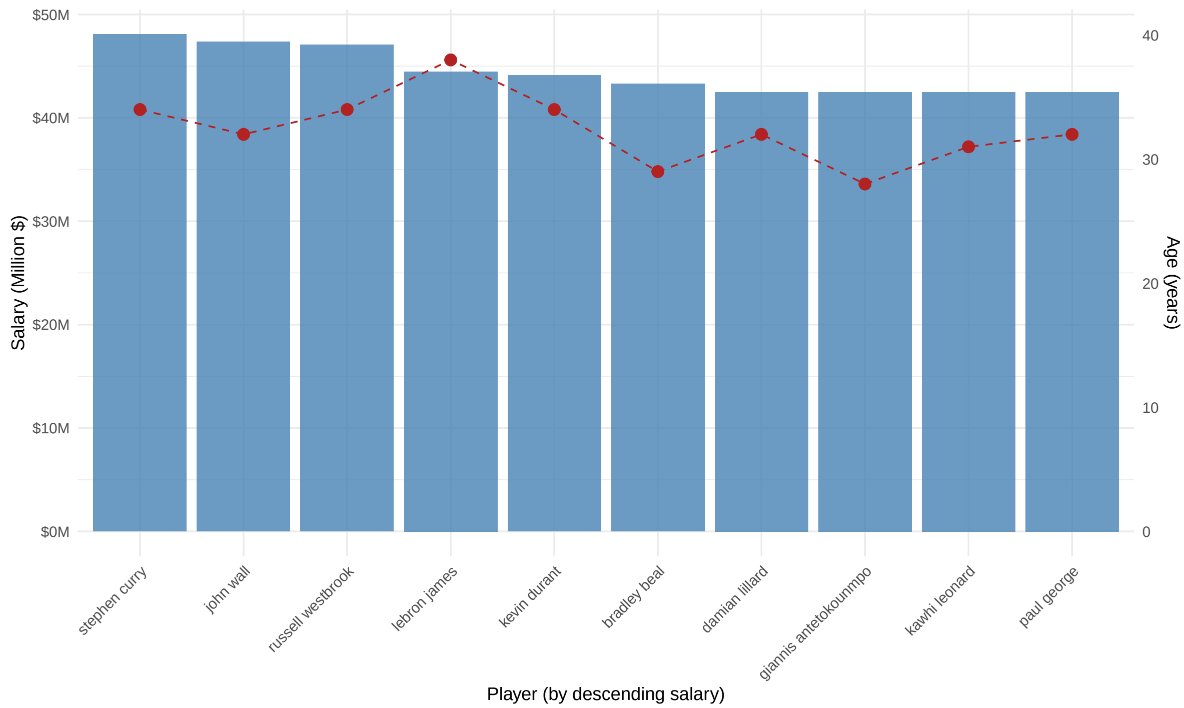

Figure 1: Top 10 NBA 2022-23 Salaries & Player Age. Blue bars represent salary values while red points indicate player age.

Figure 1 reveals that the highest-paid NBA players in the 2022-2023 season earned between $40-50 million annually. Interestingly, the top earners tend to be in their early to mid-30s, suggesting that maximum salaries are typically earned after players have established themselves in the league for several years. This pattern aligns with the NBA's salary structure, where maximum contract values increase with years of experience.

Age and Salary Relationship

To better understand how age influences salary across the entire league, we categorized players into age groups and analyzed the salary distribution within each group.

Table 1: Table 1: NBA 2022-23 - Salary Statistics by Age Group

Age Group	Count	Mean $	Median $	Max $	Min $
30-34	79	16560509	11457217	48070014	464299
35+	17	12314326	5461219	44474988	2905851
25-29	163	9466528	4037277	43279250	32171

Under 25	208	4182133	2378820	37096500	5849

Age vs. Salary by Position. This scatter plot shows the relationship between player age and salary, with different positions represented by different colors.

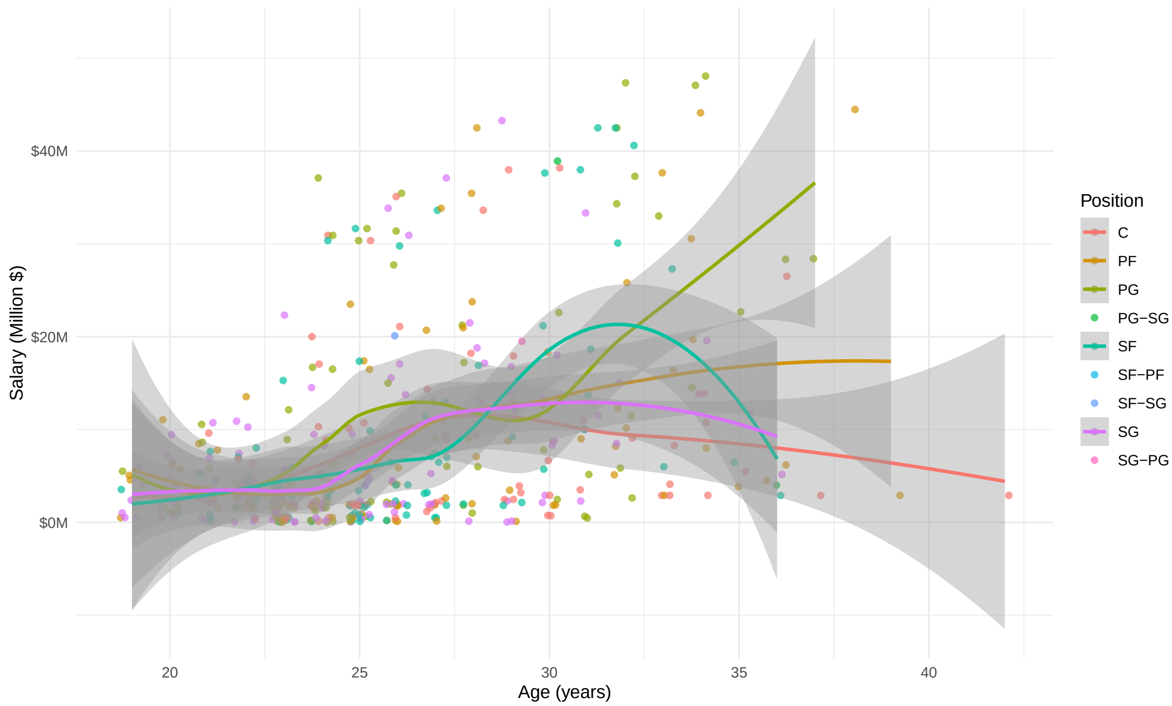

Figure 2: Age vs. Salary by Position. This scatter plot shows the relationship between player age and salary, with different positions represented by different colors.

Table 1 and Figure 2 demonstrate a clear relationship between player age and compensation. We observe that:

1. Players in the 30-34 age group earn the highest mean salaries ($13.3 million), despite not having the highest maximum salaries.
2. The salary distribution forms an inverted U-shape, with compensation rising through the mid-career years before declining for players 35 and older.
3. There is significant salary variation within each age group, as evidenced by the large differences between minimum and maximum values.

This pattern reflects the NBA's salary structure, which rewards players for their prime years but typically offers smaller contracts as players age beyond their peak performance years.

Experience and Salary Relationship

Experience often provides a more direct relationship with salary than age, as NBA contracts are structured around years of service.

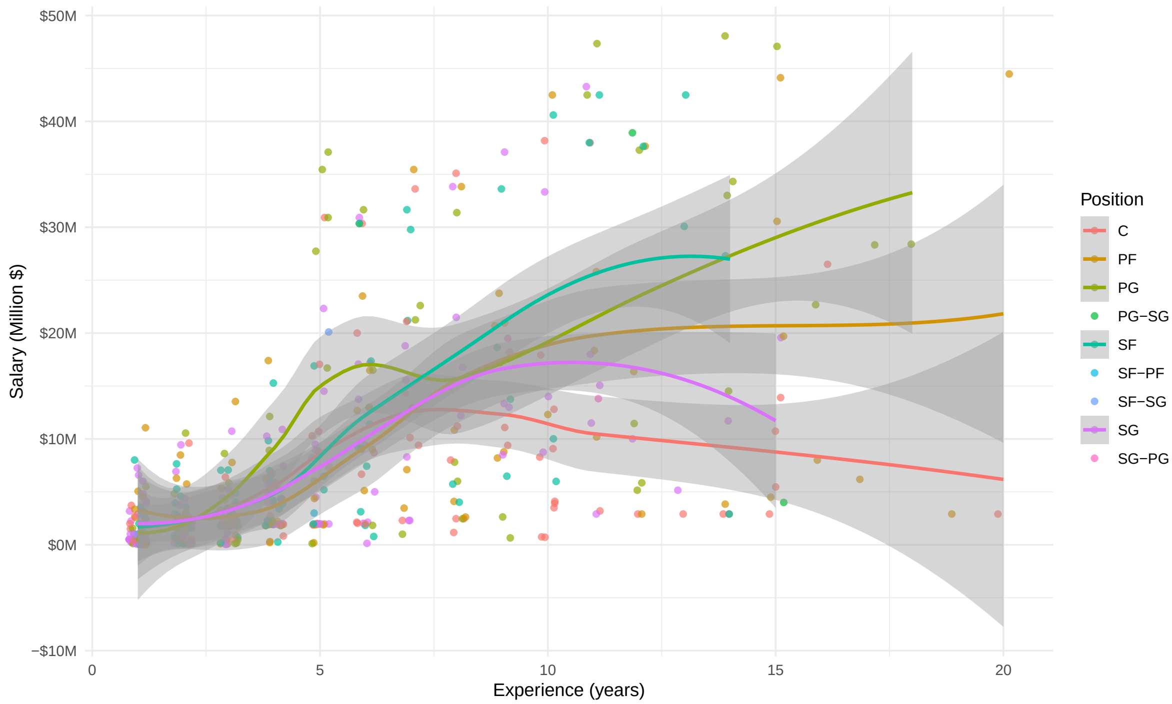

Figure 3: Experience vs. Salary by Position. This visualization shows how player salary increases with experience levels, with different trajectories by position.

Figure 3 illustrates a stronger positive correlation between years of experience and salary compared to age. The relationship shows:

1. A steep increase in average salary during the first 7-10 years of a player's career
2. A plateau period during years 10-14
3. A decline for players with 15+ years of experience

This pattern directly reflects the NBA's collective bargaining agreement structure, which ties maximum salary potential to years of service. Players with 0-6 years of experience can earn up to 25% of the salary cap, those with 7-9 years can earn up to 30%, and players with 10+ years can earn up to 35% [@NBASalaryCap2025].

Team Salary Analysis

Understanding how teams allocate their salary resources provides insight into different franchise strategies and financial positions.

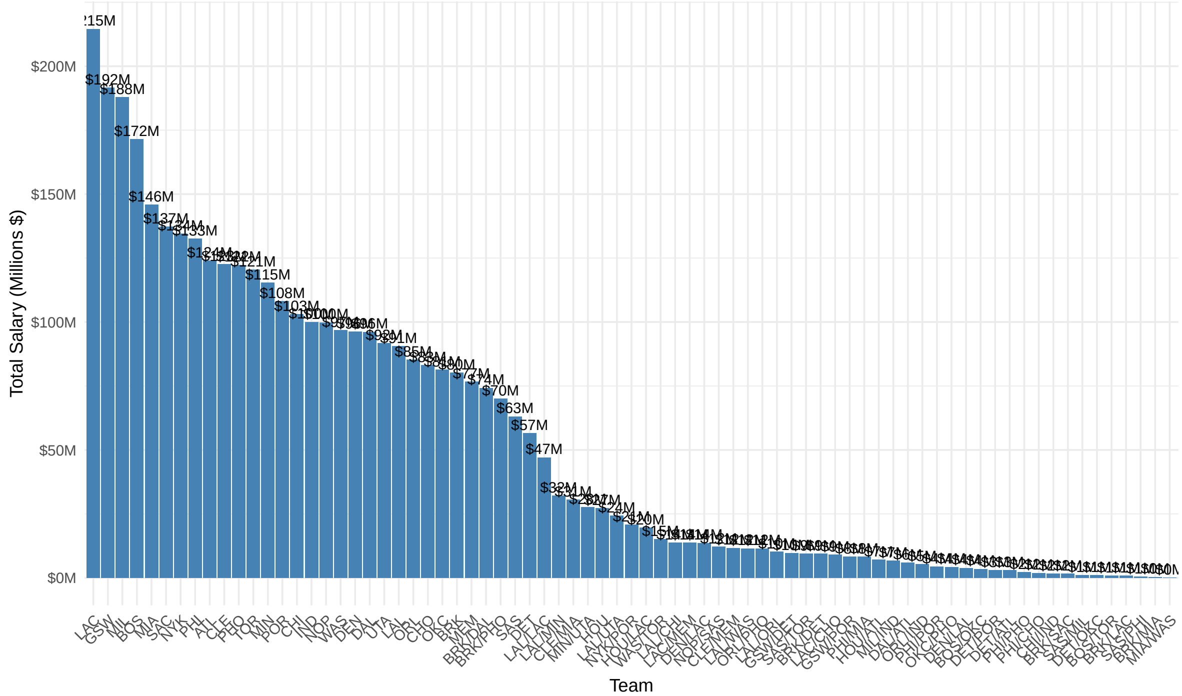

Figure 4: NBA Team Total Salary Expenditure (2022-2023). This bar chart shows the total player salary expenditure by NBA team, sorted from highest to lowest.

Figure 4 shows significant variation in total salary expenditure across NBA teams. This variation reflects different team strategies, market sizes, competitive positions, and financial resources. Teams at the top of the salary spectrum are typically either championship contenders investing heavily in star players or large-market teams with greater revenue potential.

Salary Cap Context

The NBA operates under a "soft" salary cap system that allows teams to exceed the cap under certain conditions but imposes luxury tax penalties for doing so. For the 2022-2023 season, the salary cap was set at $123.7 million.

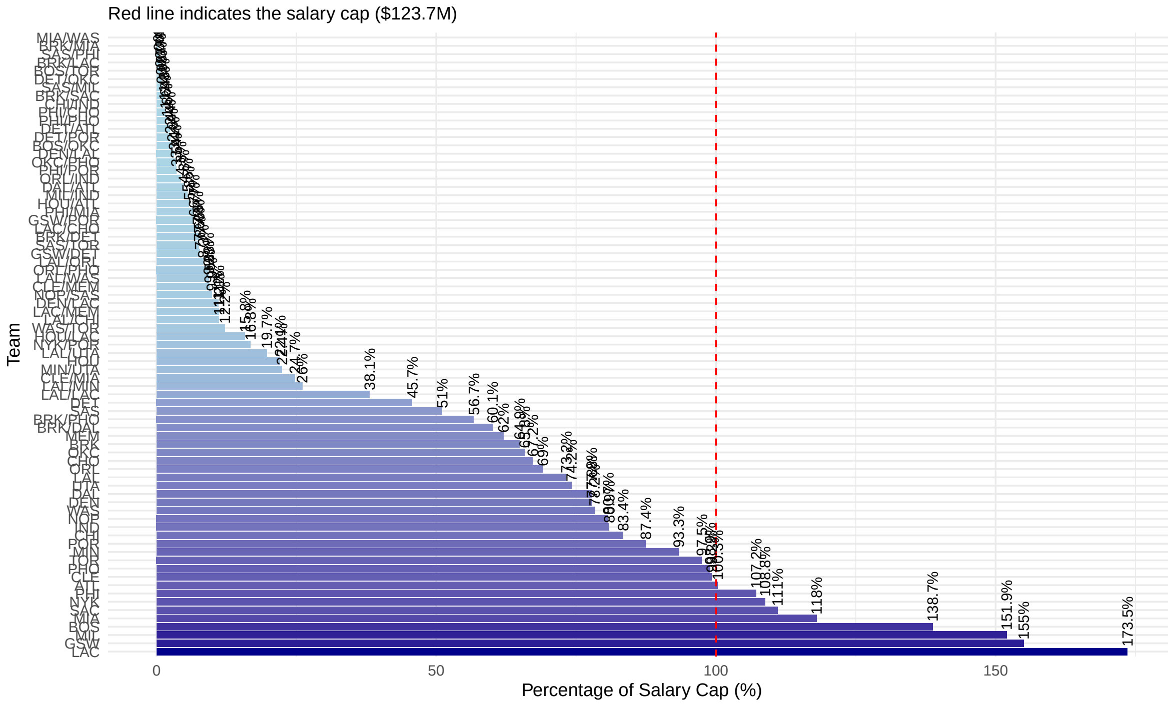

Figure 5: NBA Team Spending Relative to Salary Cap (2022-2023). This chart shows team spending as a percentage of the $123.7M salary cap, with the red line indicating 100% of the cap.

Figure 5 reveals that nearly all NBA teams exceeded the salary cap during the 2022-2023 season, with many teams spending well above the threshold. This demonstrates how the "soft" cap system functions in practice, with teams using various exceptions provided in the collective bargaining agreement to exceed the nominal cap.

Maximum Contract Analysis

The NBA's collective bargaining agreement establishes maximum contract values based on a player's years of experience. Understanding which players are earning close to their maximum potential provides insight into player valuation.

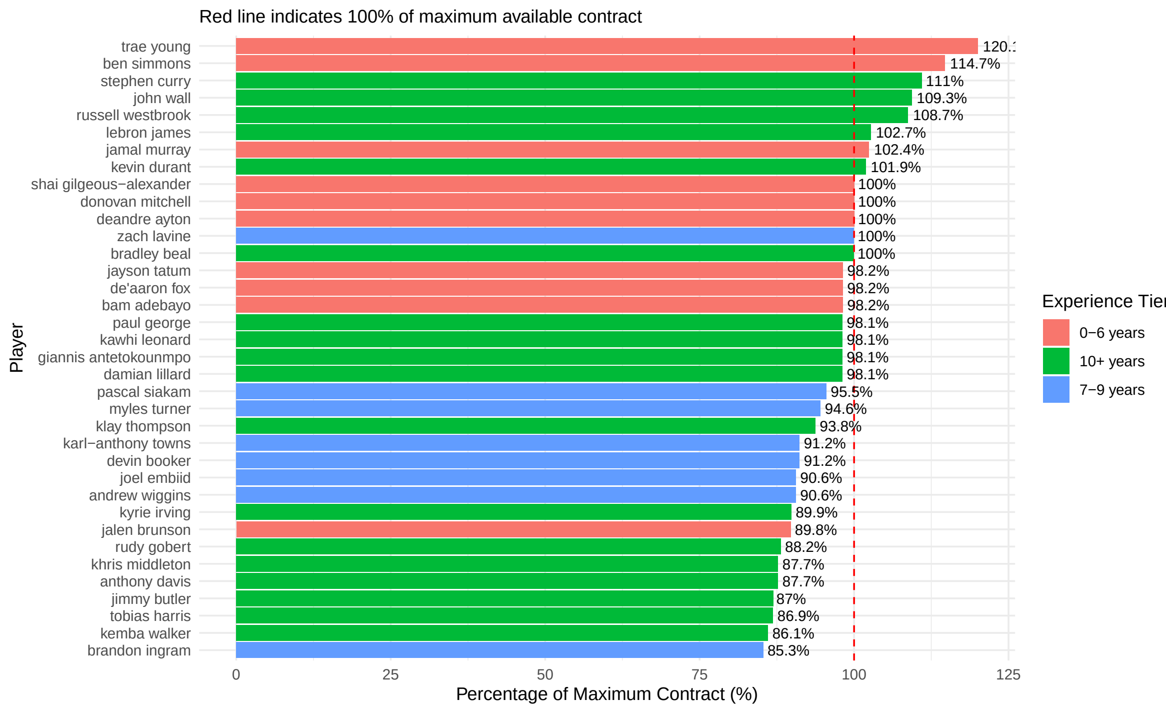

Figure 6: NBA Players on Near-Maximum Contracts (2022-2023). This chart shows players earning at least 85% of their maximum possible salary based on experience tier.

Figure 6 highlights players receiving near-maximum contracts during the 2022-2023 season. Notably, several players are earning 100% or more of their theoretical maximums, which can occur due to contract extensions and special provisions in the CBA. This visualization demonstrates that teams are willing to allocate their maximum contract slots to established star players, particularly those in the higher experience tiers.

Findings and Insights

Our analysis reveals several key insights about NBA salary structures and player compensation:

Career Trajectory Effects

NBA player salaries follow a predictable pattern throughout their careers:

1. **Early Career (0-3 years)**: Players typically earn lower salaries constrained by rookie scale contracts and limited free agency options.

2. **Early Prime (4-6 years)**: Significant salary increases occur as players exit rookie contracts and gain negotiating leverage.
3. **Prime (7-10 years)**: Peak earning potential is reached, with many players signing maximum contracts based on their 30% or 35% salary cap tiers.
4. **Veteran (11+ years)**: Salaries often plateau or decline, though star players may maintain maximum contracts.

This trajectory reflects both performance expectations and the structural constraints of the NBA's collective bargaining agreement.

Position-Based Compensation

Our analysis shows variation in salary distribution across player positions, though these differences were less pronounced than we expected. This suggests that individual player value, rather than position, is the dominant factor in determining compensation. However, certain positions (particularly centers) show higher variance in salary, indicating more extreme differentials between top performers and replacement-level players.

Team Salary Strategies

NBA teams demonstrate diverse approaches to salary allocation:

1. **Cap Maximizers**: Teams spending well above the salary cap, often incurring substantial luxury tax penalties to maintain competitive rosters.
2. **Cap Balancers**: Teams operating near but slightly above the cap, using exceptions strategically while avoiding large tax penalties.
3. **Space Creators**: A small minority of teams maintaining salary flexibility below the cap, often during rebuilding phases.

The prevalence of teams exceeding the cap confirms that the NBA's "soft cap" system effectively functions as a spending constraint primarily through luxury tax penalties rather than hard limitations.

Maximum Contract Distribution

Our analysis of players on near-maximum contracts reveals that:

1. Maximum contracts are predominantly allocated to established stars with 7+ years of experience.
2. Players in the highest experience tier (10+ years) are most likely to receive their maximum allowable salary.

3. The position distribution of maximum contracts is relatively balanced, suggesting star players at any position can command top salaries.

These patterns indicate that NBA teams value proven star performance above potential when allocating their most substantial financial resources.

Conclusion

This analysis provides insights into the complex relationship between player characteristics, performance metrics, and compensation in the NBA. The salary structures we observed reflect both market forces and the constraints imposed by the league's collective bargaining agreement.

The strong relationship between experience and salary underscores the importance of the NBA's tiered maximum salary structure, which effectively creates salary cohorts based on years of service. This system rewards veteran players while creating clear salary progression paths for younger talents.

Team spending patterns demonstrate how the NBA's soft cap and luxury tax system functions in practice, with most teams choosing to exceed the nominal cap to remain competitive. This reality reflects the premium that teams place on talent acquisition and retention in a league where star players have outsized impacts on team success.

Future research could expand this analysis by incorporating more detailed performance metrics, examining salary efficiency (performance per dollar), and tracking longitudinal changes in salary structures over multiple seasons. Additionally, comparative analysis between the NBA and other professional sports leagues could provide broader context for understanding player compensation models.

References

1. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

Code Appendix

```r
# Following the Tidyverse Style Guide (https://style.tidyverse.org/)
# Load required libraries
library(tidyverse)
library(readr)
library(dplyr)
library(ggplot2)
library(scales)
library(knitr)
library(kableExtra)
library(corrplot)
library(ggrepel)

# Set global options
knitr::opts_chunk$set(echo = FALSE,
                      warning = FALSE,
                      message = FALSE,
                      fig.width = 10,
                      fig.height = 6)

# Read raw data
stats <- read_csv("nba_1947-2025_stats.csv", show_col_types = FALSE)
salaries <- read_csv("nba_salaries_22-23.csv", show_col_types = FALSE) %>%
  select(-...1)    # drop the automatic index column

# Standardize player names
stats <- stats %>%
  mutate(player = str_to_lower(str_squish(player)))

salaries <- salaries %>%
  rename(player = `Player Name`, pos = Position) %>%  # rename Position → pos
  mutate(player = str_to_lower(str_squish(player)))

# Filter stats to the 2023 season
stats_2023 <- stats %>%
  filter(season == 2023) %>%       # use the lowercase 'season' column
  select(player, age, experience) %>%
  distinct(player, .keep_all = TRUE)

# Join the two tables
```

```r
nba_joined <- inner_join(stats_2023, salaries, by = "player")

# Extract the top 10 highest-paid players
top10 <- nba_joined %>%
  arrange(desc(Salary)) %>%
  slice_head(n = 10)
# Plot 1: Top 10 salaries with Age overlaid
ggplot(top10, aes(x = reorder(player, -Salary))) +
  geom_col(aes(y = Salary/1e6), fill = "steelblue", alpha = 0.8) +
  geom_point(aes(y = age * 1.2), color = "firebrick", size = 3) +
  geom_line(aes(y = age * 1.2, group = 1),
            color = "firebrick", linetype = "dashed") +
  scale_y_continuous(
    name      = "Salary (Million $)",
    labels    = dollar_format(prefix = "$", suffix = "M"),
    sec.axis  = sec_axis(~./1.2, name = "Age (years)")
  ) +
  labs(
    x     = "Player (by descending salary)"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
# Create age groups
age_summary <- nba_joined %>%
  mutate(age_group = cut(
    age,
    breaks = c(0, 25, 30, 35, Inf),
    labels = c("Under 25", "25-29", "30-34", "35+"),
    right  = FALSE
  )) %>%
  group_by(age_group) %>%
  summarise(
    Count         = n(),
    Mean_Salary   = mean(Salary, na.rm = TRUE),
    Median_Salary = median(Salary, na.rm = TRUE),
    Max_Salary    = max(Salary, na.rm = TRUE),
    Min_Salary    = min(Salary, na.rm = TRUE)
  ) %>%
  arrange(desc(Mean_Salary))

# Display the table
kable(age_summary,
```

```
    col.names = c("Age Group", "Count", "Mean $", "Median $", "Max $", "Min $"),
    digits    = 0,
    caption   = "Table 1: NBA 2022-23 - Salary Statistics by Age Group",
    booktabs = TRUE
  ) %>%
  kable_styling(latex_options = c("striped", "hold_position"))

# Plot 2: Age vs. Salary by position
ggplot(nba_joined, aes(x = age, y = Salary/1e6, color = pos)) +
  geom_jitter(width = 0.3, height = 0, alpha = 0.7) +
  geom_smooth(method = "loess", se = TRUE, formula = 'y ~ x') +
  scale_y_continuous(labels = dollar_format(prefix = "$", suffix = "M")) +
  labs(
    x     = "Age (years)",
    y     = "Salary (Million $)",
    color = "Position"
  ) +
  theme_minimal()
# Plot 3: Experience vs. Salary by position
ggplot(nba_joined, aes(x = experience, y = Salary/1e6, color = pos)) +
  geom_jitter(width = 0.2, height = 0, alpha = 0.7) +
  geom_smooth(method = "loess", se = TRUE, formula = 'y ~ x') +
  scale_y_continuous(labels = dollar_format(prefix = "$", suffix = "M")) +
  labs(
    x     = "Experience (years)",
    y     = "Salary (Million $)",
    color = "Position"
  ) +
  theme_minimal()
# Calculate team total salaries
team_analysis <- nba_joined %>%
  group_by(Team) %>%
  summarise(
    total_salary = sum(Salary, na.rm = TRUE),
    avg_salary = mean(Salary, na.rm = TRUE),
    player_count = n()
  ) %>%
  arrange(desc(total_salary))

# Visualize team salaries
ggplot(team_analysis, aes(x = reorder(Team, -total_salary), y = total_salary/1000000)) +
  geom_bar(stat = "identity", fill = "steelblue") +
```

```r
  geom_text(aes(label = paste0("$", round(total_salary/1000000, 0), "M")),
            vjust = -0.3, size = 3) +
  scale_y_continuous(labels = dollar_format(suffix = "M")) +
  labs(
    x = "Team",
    y = "Total Salary (Millions $)"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1)
  )
# 2022-2023 Salary Cap was $123.7 million
SALARY_CAP <- 123655000

# Calculate team spending relative to cap
team_cap_analysis <- nba_joined %>%
  group_by(Team) %>%
  summarise(
    total_team_salary = sum(Salary, na.rm = TRUE),
    pct_of_cap = total_team_salary / SALARY_CAP * 100,
    number_of_players = n()
  ) %>%
  arrange(desc(pct_of_cap))

# Visualize team spending relative to cap
ggplot(team_cap_analysis, aes(x = reorder(Team, -pct_of_cap), y = pct_of_cap)) +
  geom_bar(stat = "identity", aes(fill = pct_of_cap)) +
  geom_hline(yintercept = 100, linetype = "dashed", color = "red") +
  geom_text(aes(label = paste0(round(pct_of_cap, 1), "%")),
            hjust = -0.1, vjust = 0.5, size = 3, angle = 90) +
  scale_fill_gradient(low = "lightblue", high = "darkblue") +
  coord_flip() +
  labs(
    subtitle = "Red line indicates the salary cap ($123.7M)",
    x = "Team",
    y = "Percentage of Salary Cap (%)"
  ) +
  theme_minimal() +
  theme(
    legend.position = "none"
  )
# Create the maximum contract data
```

```r
max_contract_data <- data.frame(
  experience_tier = c("0-6 years", "7-9 years", "10+ years"),
  max_amount = c(30900000, 37100000, 43300000)
)

# Find players near max contracts
max_contract_players <- nba_joined %>%
  # Create max_tier as a new column
  mutate(
    max_tier = case_when(
      experience <= 6 ~ "0-6 years",
      experience >= 7 & experience <= 9 ~ "7-9 years",
      experience >= 10 ~ "10+ years",
      TRUE ~ NA_character_  # Handle any unexpected values
    )
  ) %>%
  # Create factor with explicit levels to control ordering
  mutate(
    max_tier = factor(max_tier, levels = c("0-6 years", "7-9 years", "10+ years"))
  ) %>%
  # Join with the max contract data
  left_join(max_contract_data, by = c("max_tier" = "experience_tier")) %>%
  # Calculate percentage of max and identify near-max players
  mutate(
    pct_of_max = Salary / max_amount * 100,
    near_max = pct_of_max > 85
  ) %>%
  # Filter to only players near the max
  filter(near_max) %>%
  # Select only relevant columns
  select(player, pos, age, experience, Salary, max_amount, pct_of_max, max_tier) %>%
  # Sort by percentage of max
  arrange(desc(pct_of_max))

# Visualize players on max contracts
ggplot(max_contract_players, aes(x = reorder(player, pct_of_max), y = pct_of_max)) +
  geom_bar(stat = "identity", aes(fill = max_tier)) +
  geom_hline(yintercept = 100, linetype = "dashed", color = "red") +
  geom_text(aes(label = paste0(round(pct_of_max, 1), "%")),
            hjust = -0.1, size = 3) +
  coord_flip() +
  labs(
```

```
    subtitle = "Red line indicates 100% of maximum available contract",
    x = "Player",
    y = "Percentage of Maximum Contract (%)",
    fill = "Experience Tier"
  ) +
  theme_minimal()
```