NBA Salary Cap

2025-11-07

Importing Data

```
library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(janitor)
```

```
##
## Attaching package: 'janitor'

## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

```
library(scales)
```

```
##
## Attaching package: 'scales'

## The following object is masked from 'package:readr':
##
##     col_factor
```

```
library(broom)
library(ggthemes)

nba <- read_csv("NBA Salary Cap Data - 24-25 Salaries.csv") |>
  clean_names()
```

```
## Rows: 30 Columns: 14

## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (10): Team Name, Total Cap, #1 Name, #1 Salary, #2 Name, #2 Salary, #3 N...
## dbl  (4): Preseason Power Rankings (nba.com), Final Power Ranking (The Athle...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

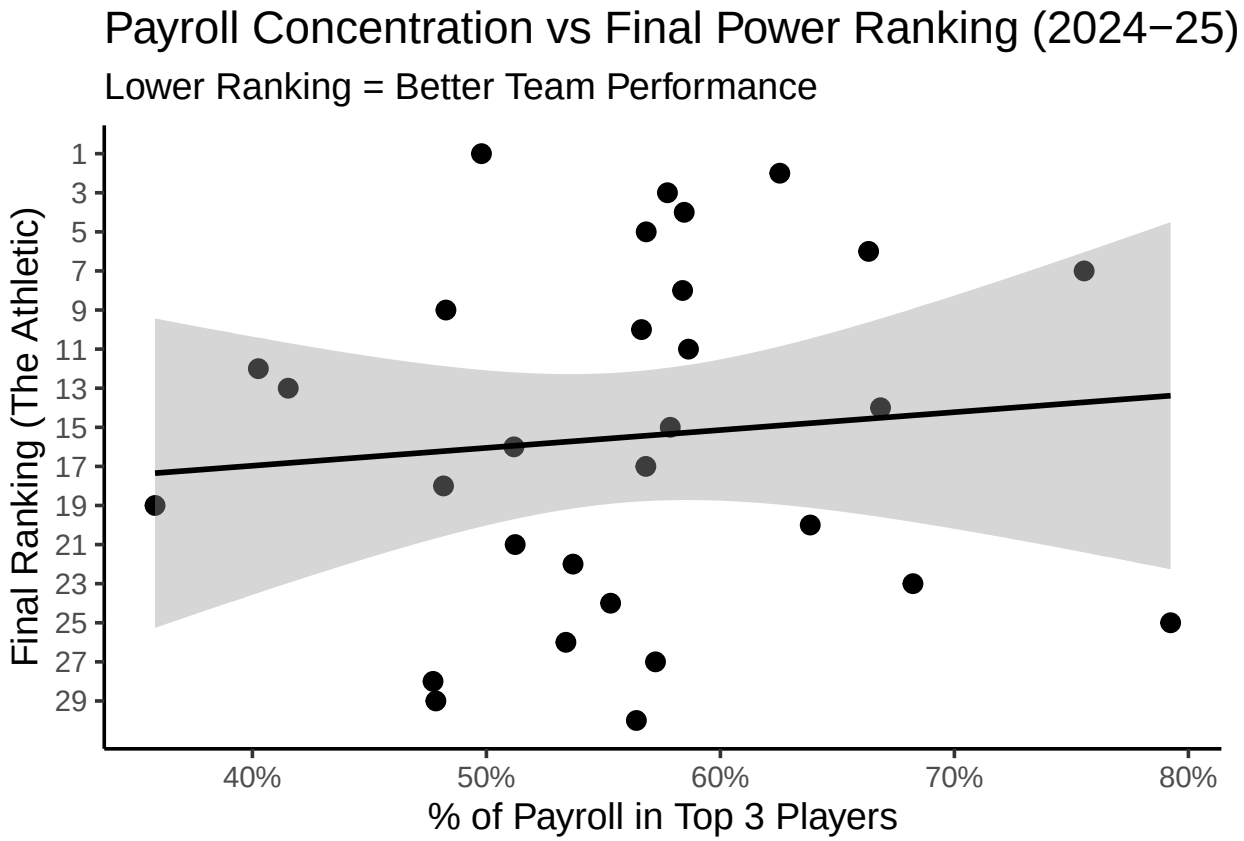
Data Manipulation

```
nba <- nba |>
  mutate(
    total_cap = parse_number(total_cap),
    cap_m = total_cap/1000000,
    salary_1 = parse_number(`number_1_salary`),
    salary_2 = parse_number(`number_2_salary`),
    salary_3 = parse_number(`number_3_salary`),
    rest_of_team = parse_number(rest_of_team),
    top3_total = salary_1 + salary_2 + salary_3,
    top3_share = parse_number(`percent_of_payroll_on_top_3`) / 100,
    rank_change_calc = `final_power_ranking_the_athletic` - `preseason_power_rankings_nba_com`,
    balance_index = rest_of_team / top3_total
  )
```

Data Visualization - Top-3 payroll share vs Final Ranking

```
p_final <- ggplot(nba, aes(x = top3_share, y = final_power_ranking_the_athletic)) +
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = TRUE, color = "black") +
  scale_x_continuous(labels = percent) +
  scale_y_reverse(breaks = seq(1, 30, by = 2)) +  # 1 = best, 30 = worst
  labs(
    title = "Payroll Concentration vs Final Power Ranking (2024-25)",
    subtitle = "Lower Ranking = Better Team Performance",
    x = "% of Payroll in Top 3 Players",
    y = "Final Ranking (The Athletic)"
  ) +
  theme_classic(base_size = 14)
p_final
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
ggsave("NBA_Top3_Allocation_vs_Final_Ranking.png", p_final, width = 9, height = 6, dpi = 300, units = "
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
m_rank_final <- lm(final_power_ranking_the_athletic ~ top3_share + cap_m, data = nba)
summary(m_rank_final)
```

```
##
## Call:
## lm(formula = final_power_ranking_the_athletic ~ top3_share +
##     cap_m, data = nba)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.2671  -7.4053   0.1072   7.6446  13.3428
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 32.29312   18.57150   1.739   0.0935 .
## top3_share  -2.44271   19.65012  -0.124   0.9020
## cap_m       -0.08899    0.11963  -0.744   0.4634
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 8.987 on 27 degrees of freedom
## Multiple R-squared:  0.0297, Adjusted R-squared:  -0.04217
## F-statistic: 0.4133 on 2 and 27 DF,  p-value: 0.6656
```

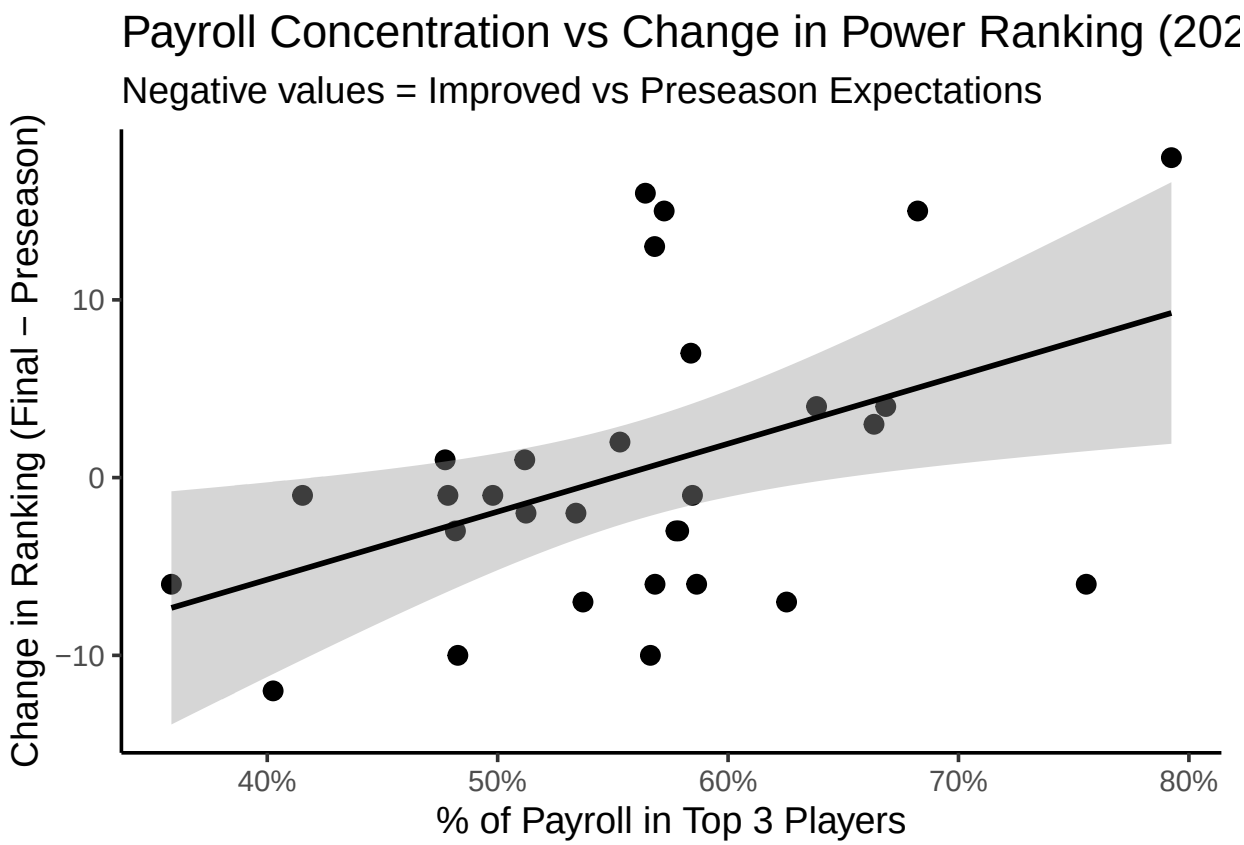
$R^2 = 0.03$, Adjusted $R^2 = -0.04$, $p = 0.67$, (top3_share) = –2.44 ($p = 0.90$), (cap_m) = –0.09 ($p = 0.46$)

Takeaways: There was no statistically significant relationship between payroll concentration and team performance. Teams that allocated a larger share of their payroll to their top three players did not perform meaningfully better or worse in final power rankings, suggesting that overall spending structure had little impact on success in the 2024–25 season.

Data Visualization - Payroll Concentration vs Ranking Change

```
p_change <- ggplot(nba, aes(x = top3_share, y = rank_change_calc)) +
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = TRUE, color = "black") +
  scale_x_continuous(labels = percent) +
  labs(
    title = "Payroll Concentration vs Change in Power Ranking (2024-25)",
    subtitle = "Negative values = Improved vs Preseason Expectations",
    x = "% of Payroll in Top 3 Players",
    y = "Change in Ranking (Final - Preseason)"
  ) +
  theme_classic(base_size = 14)
p_change
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



```
ggsave("NBA_Top3_Allocation_vs_Ranking_Change.png", p_change, width = 9, height = 6, dpi = 300, units =
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
m_rank_change <- lm(rank_change_calc ~ top3_share + cap_m, data = nba)
summary(m_rank_change)
```

```
##
## Call:
## lm(formula = rank_change_calc ~ top3_share + cap_m, data = nba)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.8347  -5.1245  -0.3353   3.2991  16.5889
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -28.57869   15.45271  -1.849   0.0754 .
## top3_share   33.90818   16.35020   2.074   0.0478 *
## cap_m         0.05752    0.09954   0.578   0.5681
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.478 on 27 degrees of freedom
## Multiple R-squared:  0.2117, Adjusted R-squared:  0.1533
## F-statistic: 3.625 on 2 and 27 DF,  p-value: 0.04032
```

$R^2 = 0.21$, Adjusted $R^2 = 0.15$, F-test p = 0.040 (top3_share) = +33.91 (p = 0.048) (cap_m) = +0.058 (p = 0.57)

Takeaways: Higher payroll concentration in the top three players is positively associated with worse ranking change (more positive = underperformed preseason), while total cap is not predictive. In 2024–25, top-heavy teams tended to underperform expectations modestly, though the effect size is moderate and results should be read with small-N caution.