

Project Report

NFL Team's Player Salaries

Megan Tetrick, Brendan Sullivan, Anne Oddan

1. Introduction

The National Football League (NFL) is a professional American football league comprising 32 teams divided into two conferences: the American Football Conference (AFC) and the National Football Conference (NFC). Each conference is further segmented into four divisions, each consisting of four teams. Each of these teams plays a 17-game regular season (they previously played 16-game regular seasons prior to 2021). Across these 17 games, teams play for a spot in the playoffs, where the top teams from each conference compete to reach the Super Bowl, the NFL's championship game.

Our analysis focuses on examining the relationship between player salaries and team success. Player salaries in the NFL vary significantly based on a player's position, skill level, experience, and contract negotiations. Conversely, team success is typically assessed by win-loss records, playoff appearances, and ultimately, their performance in the postseason.

The issue we are seeking to explore is whether a high payroll, which generally signifies higher salaries for players, correlates with enhanced team performance. While it may seem logical that more expensive players would produce better results, various factors might influence the outcome. Through our analysis, we aim to uncover patterns or insights into how player compensation relates to overall team success.

2. Data

This project utilizes two primary datasets to examine the intricate relationship between player salaries and team success in the NFL: the NFL Scores Dataset and the Player Salaries Dataset. By delving into these datasets, we aim to uncover patterns and insights that can inform decisions for managers, players, and fans alike.

2.1 NFL Scores Dataset

The first dataset is the Kaggle 1926-2024 NFL Scores Dataset, which is a comprehensive collection of game results spanning nearly a century of NFL history. For the purpose of our analysis, we will focus on data from the 2020 through 2024 seasons. This dataset is rich with detailed game outcome data, including the date and day of the week the game was played, the names of the winning and losing teams, their respective scores, the type of game (regular season or playoff), and the season year. This dataset will be instrumental in exploring various measures of team success, such as wins, losses, and playoff appearances, providing a robust foundation for our analysis.

2.2 Player Salaries Dataset

The second dataset is sourced from Spotrac, a leading provider of detailed breakdowns of player salaries, team salary caps, and cap space for NFL teams. This dataset offers up-to-date salary information on players and teams, which is essential for analyzing the financial aspects of team success. Key fields in this dataset include the salary cash ranking of both winning and losing teams, their win-loss-tie records, the number of signed players, the average age of signed players, and detailed financial data such as active cash, dead cash, and total cash for both winning and losing teams. This dataset provides a comprehensive view of player compensation and financial strategy, which is crucial for understanding the economic factors influencing team performance.

2.3 Merging the Datasets

By merging the Spotrac salary data with the Kaggle game scores data, we can perform a detailed and nuanced analysis to examine how player salaries might be linked to team success. The merged dataset will enable us to address several compelling research questions, such as:

- How do the salary caps of the top 5 performing teams (top 5 win-loss records) compare each year? This question will help us understand if financial investment correlates with top-tier performance.
- What is the overall win-loss record of each team, including both regular and postseason performance? This will provide a holistic view of team success across different phases of competition.

- What are the summary statistics of salaries for Super Bowl-winning teams for the 2020-2024 seasons? This analysis will reveal the financial characteristics of championship teams.
- How do the records of the top salary cap teams compare each season? This will help us understand if higher financial investment consistently leads to better performance.
- Do teams with higher salaries perform better during the regular season or postseason? This question will explore the impact of financial investment on different stages of competition.
- How do relatively low salary cap teams perform compared to relatively high salary cap teams in the 2020-2024 seasons? This comparative analysis will provide insights into the performance of teams with varying financial strategies.

This combined dataset will provide important insights into the relationship between player compensation and team performance, educating managers, players, and fans on the determinants of team success in the NFL.

Table 1 Data Dictionary

Field Name	Data Type	Description
Date	Date	The date the game was played
Day of the Week	String	The day of the week the game was played (e.g., “Thu”, “Sun”, “Mon”)
Winning Team Name	String	The winning team
Losing Team Name	String	The losing team
Winning Team Score	Integer	The number of points scored by the winning team (WT)
Losing Team Score	Integer	The number of points scored by the losing team (LT)
Type	String	The type of game (i.e., “Regular Season” or “Playoff”)
Year	Integer	The season year (e.g., 2020, 2021, etc.)
Winning Rank	Integer	The winning team's salary cash ranking (1 being the highest total cash salary) by each season
Winning Team Record	String	The winning team’s win-loss-tie record

Winning Signed Players	Integer	Total number of signed players on the winning team
Winning Avg Age	Float	Average age of the signed players on the winning team
Winning Active Cash	Integer	The money that is currently allocated to players and counts against the winning team's salary cap for the season
Winning Dead Cash	Integer	Money that is still owed to players from a previous contract but is no longer counted against the winning team's active salary cap
Winning Total Cash	Integer	The total amount of money players will earn in a given season, including both active and dead cash, of the winning team.(Base salary, signing bonuses, roster bonuses, performance incentives, and any dead money)
Losing Rank	Integer	The losing team's salary cash ranking (1 being the highest total cash salary) by each season
Losing Team Record	String	The losing team's win-loss-tie record
Losing Signed Players	Integer	Total number of signed players on the losing team
Losing Avg Age	Float	Average age of the signed players on the losing team
Losing Active Cash	Integer	The money that is currently allocated to players and counts against the losing team's salary cap for the season
Losing Dead Cash	Integer	Money that is still owed to players from a previous contract but is no longer counted against the losing team's active salary cap
Losing Total Cash	Integer	The total amount of money players will earn in a given season, including both active and dead cash, of the losing team.(Base salary, signing bonuses, roster bonuses, performance incentives, and any dead money)

3. Analysis

3.1 *Salary in relation to team success*

To analyze the relationship between team performance and financial expenditure, we undertook a series of data processing and analysis steps. The goal was to identify the five worst-performing teams each year based on their win-loss records and total cash spent, and to compare these with the top five winning teams.

First, we converted the salary data from strings to numerical values. The columns Winning Total Cash and Losing Total Cash contained salary information formatted as strings with dollar signs and commas. These non-numeric characters were removed, and the values were converted to floating-point numbers to enable accurate numerical analysis.

Next, we extracted the number of wins from the Losing Record column, which contained win-loss records in a string format (e.g., "10-20"). Using a regular expression, we extracted the number of wins and converted these values to integers.

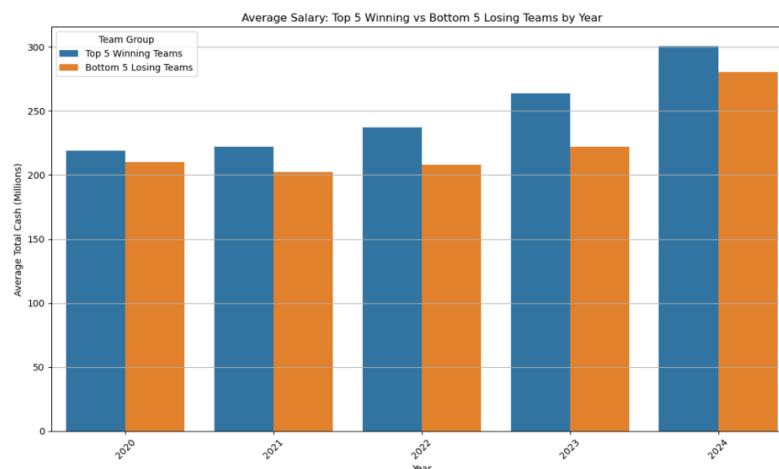
Following this, we grouped the data by Year and Losing Team. For each group, we calculated the maximum values of Losing Wins and Losing Total Cash. This aggregation provided a summary of the worst-performing teams each year based on their win-loss records and total cash spent.

Within each year, we ranked the teams based on their Losing Wins in ascending order. The five teams with the highest ranks, representing the worst-performing teams, were identified.

Similarly, we grouped the data by Year and Winning Team Name, calculated the maximum values of Wins and Winning Total Cash, and ranked the teams based on their Wins in descending order to identify the top five winning teams each year.

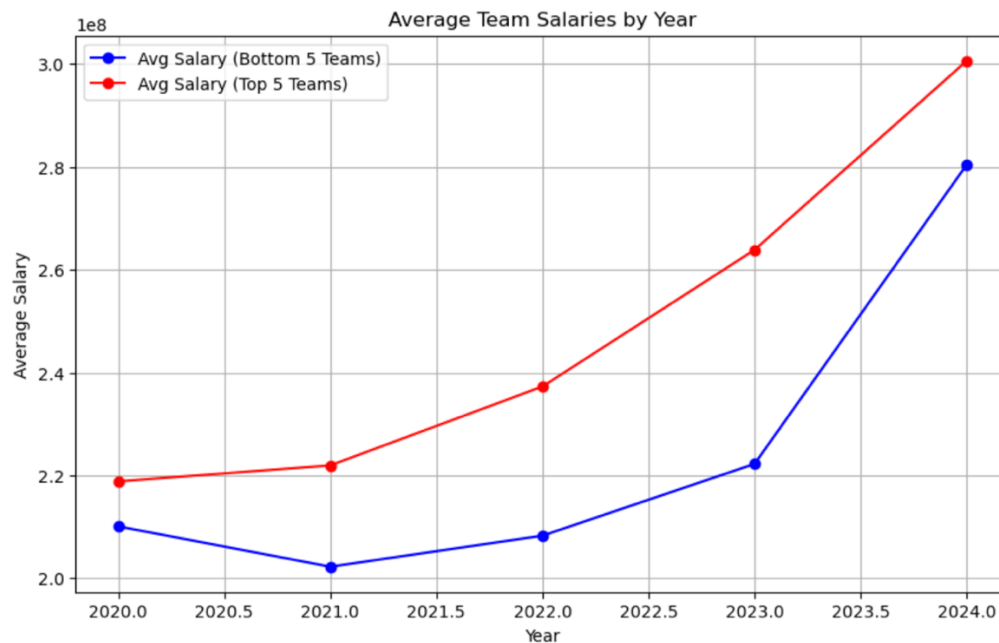
To visualize the financial expenditure of the worst and best-performing teams, we created a bar graph (Figure 1) showing the average salaries of the bottom 5 losing teams and the top 5 winning teams by year. Each year has two bars in the graph, one representing the bottom 5 teams and the other representing the top 5 teams.

Figure 1 Bar Graph of Average Team Salaries by Year for Top 5 and Bottom 5 Teams



The bar chart allowed us to see that there might be some degree of a pattern between the two groups of teams. Because of this, we decided to create a line graph using the same data and find the r-squared value to see if the correlation was actually significant. The blue line represents the average salary of the bottom 5 losing teams, while the red line represents the average salary of the top 5 winning teams.

Figure 2 Line Graph of Average Team Salaries by Year for Top 5 and Bottom 5 Teams



The line graph (Figure 2) allowed us to see that the average salary of the bottom 5 teams appeared to be lower but interestingly, the analysis revealed a very weak correlation between the average salaries of the bottom 5 losing teams and the top 5 winning teams, with an R-squared value of 0.001199. This indicates that there is almost no linear relationship between the financial expenditure of the worst and best-performing teams. This finding suggests that factors other than financial expenditure may play a more significant role in determining team performance.

Overall, this analysis provides valuable insights into the financial strategies of teams and highlights the need for a more comprehensive approach to improving team performance beyond just financial investment.

3.2 Impact of Team Salaries on Regular Season and Postseason Performance

To determine if higher-paid teams are better in the regular season or playoffs, we conducted an analysis of how team salaries correlate with their Super Bowl performance compared to the average league salaries.

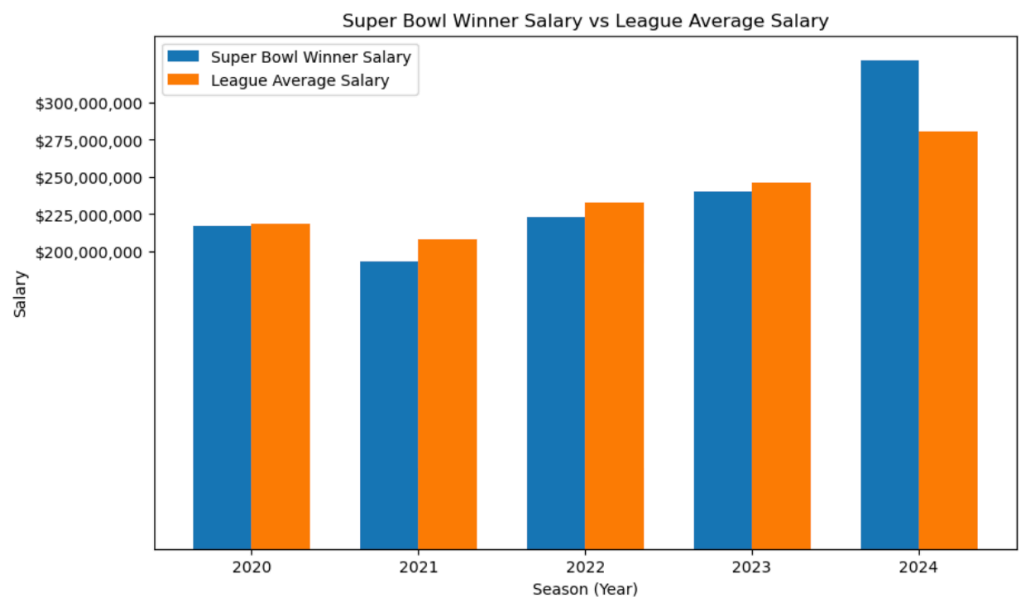
We first converted the date data to a datetime format and cleaned the dataset to have only playoff games. From the playoff games, we pulled the last game of each year, which is the Super Bowl. we then pulled the salary data of the winning teams of those Super Bowl games.

Next, we combined the salary data for winning and losing teams for all years without duplicates for each team per year. This allowed us to calculate the average salary per year across the entire league.

To compare Super Bowl winner salaries with the league average salaries, we merged Super Bowl winner salary data and the league average salary data. The comparison provided an easy view of how Super Bowl-winning teams' salaries relate to the league average.

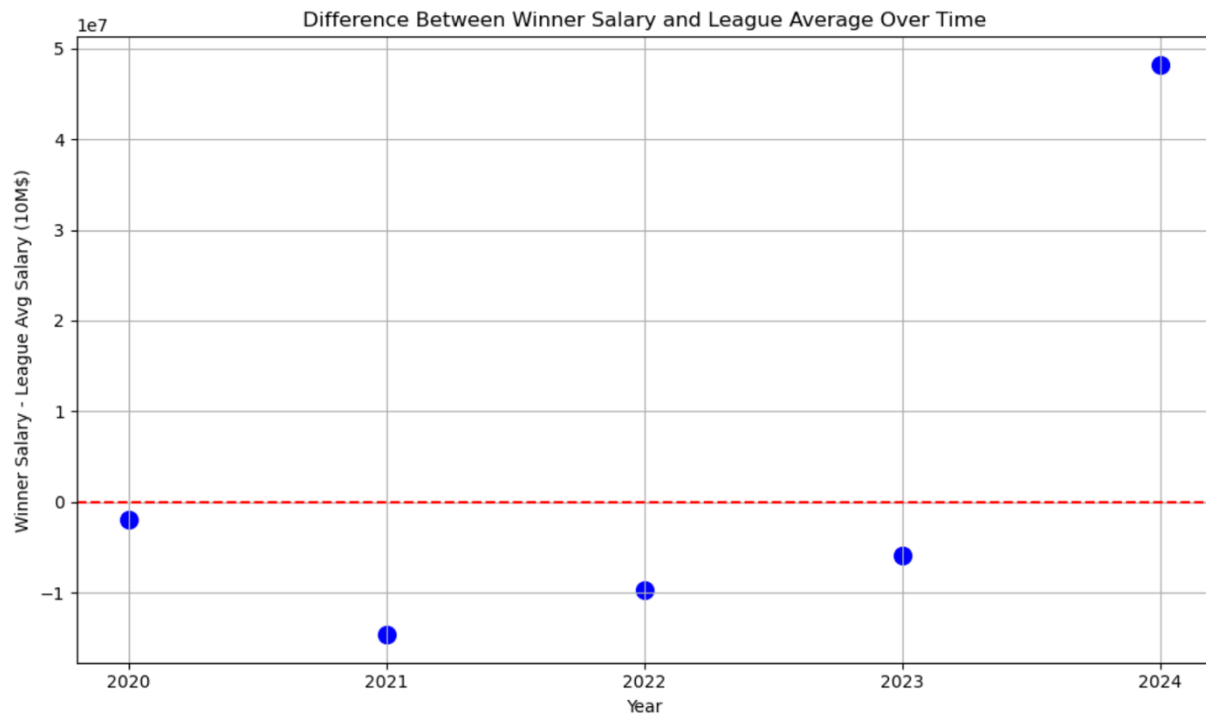
To make this comparison clearer, we created a bar graph comparing the Super Bowl winner salaries and the league average salaries from year to year. In Figure 3, you can see that the graph indicated that the Super Bowl-winning team had a higher total salary than the league average in only the most recent year.

Figure 3 Bar Chart Comparing Salaries with Super Bowl Winner and League Average



To further visualize this inconsistent-looking data we created a scatterplot. The scatterplot shown in Figure 4 keeps the league average salary as a constant (red line) and has Super Bowl winner points plotted in relation to the league average salary. If the winner is below the

Figure 4 Scatterplot Comparing Difference in Salaries between Super Bowl Winner and League Average each Year



The above scatterplot allows us to more easily see that the Super Bowl winner's salary does not appear to be consistently above the league average salary. In fact, for four seasons before the most recent, the Super Bowl winner had a lower salary than the average. The most recent year's winner is significantly above the league-average salary, but it can be inferred that this is just an instance where player salary coincided with success and not a trend.

To statistically test the significance of this difference and verify our belief that there is not a visible trend, we performed a paired t-test between the Super Bowl winner salaries and the league average salaries. The t-test resulted in a t-statistic of 0.2827 and a p-value of 0.7915. Given the alpha level of 0.05, the p-value is greater than 0.05, leading us to fail to reject the null hypothesis. This indicates that there is no significant difference between Super Bowl winner salaries and league-average salaries.

In conclusion, the analysis suggests that teams with higher salaries do not necessarily perform better in the postseason, as evidenced by the lack of a significant difference between the salaries of Super Bowl-winning teams and the league average. This finding highlights the complexity of factors influencing team performance and suggests that financial investment alone may not be a decisive factor in achieving postseason success.

4. Conclusion

For this project, we analyzed the relationship between player salaries and team performance in the NFL using two primary data sets: the NFL Scores Dataset and the Player Salaries Dataset. Our research aimed to identify patterns and findings that would be beneficial to managers, players, and fans when making decisions. Below are the primary findings of our research questions:

1. Is there an intense correlation between team spending and performance?

Our analysis revealed that there was an extremely poor correlation between the average expenditure of the worst 5 losing teams and the top 5 winning teams, with an R-squared of 0.001199. This indicates that there is effectively no linear correlation between the worst and top teams' spending. It is possible there are other variables beyond spending that have a more significant impact on team performance.

2. Are higher-paid teams better in the regular season or playoff?

We compared the salaries of Super Bowl-winning teams to league-average salaries. The data showed the salaries of Super Bowl-winning teams were slightly lower than the league average for four years but were significantly higher in another year. A paired t-test produced a t-statistic value of 0.2827 and a p-value of 0.7915, indicating no significant difference between Super Bowl winner salaries and league average salaries. This suggests that more lucrative teams are not necessarily better performers during the postseason.

Overall, this study provides us with valuable information about the financial strategy of NFL teams and the multidimensionality of the determinants of team performance. Financial investment may not be the key driver that leads to success, and more comprehensive strategies are needed to enhance team performance. There may be future research including examination of other determinants like team management, player development, coaching effectiveness, team cohesion, and the impact of injuries. Additionally, exploring the role of advanced analytics and

technology in optimizing team performance could provide further insights. Understanding these factors holistically can help teams develop more effective strategies to achieve sustained success in the NFL.