## 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.

## Data

Our first dataset is the Kaggle 1926-2024 NFL Scores[[1]](#footnote-2) Dataset, which includes game results for the 1926-2024 seasons. For our analysis, we are focusing on the data from the 2020-2024 seasons. This dataset includes game outcomes such as the date, teams involved, scores, and the type of game (regular season, playoffs). We will use this data to analyze team success, such as wins and playoff appearances.

The second data source we will use is Spotrac[[2]](#footnote-3), which provides detailed breakdowns of player salaries, team salary caps, and cap space for NFL teams. Spotrac allows us to access up-to-date salary information for individual players and teams, which will be crucial for analyzing the relationship between team payroll and performance. This data will enable us to examine player compensation and overall team spending.

By combining the Spotrac salary data with the Kaggle game scores data, we can perform a detailed analysis to examine how player salaries might correlate with team success, measured by metrics such as wins and playoff appearances, across the 2020-2024 seasons.

*Data Dictionary, Integrated Data*

|  |  |  |
| --- | --- | --- |
| **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.) |
| Team Name | String | The name of the team whose cash values are being analyzed (also the winning team, was used to merge the data) |
| Rank | Integer | The team salary cash ranking (1 being the highest total cash salary) by each season |
| Team | String | Abbreviated version of the team’s name |
| Record | String | The specified team’s win-loss-tie record |
| Signed Players | Integer | Total number of signed players on the team |
| Avg Age | Float | Average age of the signed players on the team |
| Active Cash | Integer | The money that is currently allocated to players and counts against the team’s salary cap for the season |
| Dead Cash | Integer | Money that is still owed to players from a previous contract but is no longer counted against the team’s active salary cap |
| Total Cash | Integer | The total amount of money players will earn in a given season, including both active and dead cash. (Base salary, signing bonuses, roster bonuses, performance incentives, and any dead money) |

## Proposed Analysis

Potential research questions:

* How do the salary caps of the top 5 performing teams (top 5 win-loss records) compare each season?
* What is the overall win-loss record of each team, considering both regular season and postseason performances?
* What are the summary statistics of salaries for teams that won the Super Bowl for the 2020-2024 seasons?
* How do the records of the teams with the highest salary caps compare each season?
* Do teams with higher salaries perform better in the regular season or postseason?
* How do teams with relatively low salary caps perform compared to those with relatively high salary caps in the 2020-2024 seasons?

Why the data analysis is important:

* Managers/coaches of teams may want to see if there is a correlation between pay and player performance when evaluating how to better succeed.
* Players could want to know how their team is comparable to other teams in the league and how that has changed across multiple years to see if they want to negotiate with other teams once free agents.
* Fans might also like to analyze whether the players on their team are being “overpaid” for bad performance which can affect their support for the team. It could be part of the decision of if they buy merchandise and tickets for games or not.

1. Flynn, J. (2024). *NFL scores dataset 1926-2024*. Kaggle. Retrieved March 30, 2025, from <https://www.kaggle.com/datasets/flynn28/1926-2024-nfl-scores> [↑](#footnote-ref-2)
2. Spotrac. (n.d.). *NFL salary cap information*. Spotrac. Retrieved March 30, 2025, from <https://www.spotrac.com/nfl/teams> [↑](#footnote-ref-3)