

NFL Player and Team Performance Database

For my database project, I chose to work with two NFL-related datasets that explore the relationship between individual player performance and overall team success. These datasets offer a rich opportunity to analyze how different aspects of player statistics contribute to team outcomes across multiple seasons.

Dataset 1: NFL Player Stats

- Source: Kaggle
- Link: <https://www.kaggle.com/datasets/zynicide/nfl-football-player-stats>
- Overview: This dataset contains detailed statistics for NFL players across multiple seasons. Each row corresponds to one player's performance in a given year, including information such as player name, position, team, and a wide range of statistics (e.g., passing yards, rushing touchdowns, tackles, field goals made).
- Size: Approximately 30,000 rows and 50+ attributes.
- Structure: Key columns include player name, position, team, year, and various performance metrics.

Dataset 2: NFL Team Stats (2003–2023)

- Source: Kaggle
- Link: <https://www.kaggle.com/datasets/nickcantalupa/nfl-team-data-2003-2023>
- Overview: This dataset includes summary performance metrics for all NFL teams over the past two decades. It includes yearly data on wins, losses, point differentials, offensive and defensive rankings, and playoff appearances.
- Size: Approximately 640 rows (32 teams × 20 years), with 20–30 attributes.
- Structure: Key fields include season, team, wins, losses, points scored/allowed, and playoff indicators.

Potential Insights & Questions

- How strongly do quarterback stats correlate with team win totals?
- Do rushing or receiving stats better predict playoff appearances?
- Which player positions have the highest statistical impact on team success?
- How do player and team performances evolve over time (e.g., by era or rule change)?

These datasets provide a compelling foundation to build a normalized relational database that can support deep analytical queries and visualizations of player-team dynamics across the NFL.