TITLE PAGE

Forecasting NFL Wide Receiver Touchdowns with a Temporal Linear Regression Model

Author: Ruben Chung

Author affiliation: Brown University, Providence, Rhode Island

Contact information author:

Ruben Chung

Ruben_chung@brown.edu

Forecasting NFL Wide Receiver Touchdowns with a Temporal Linear Regression Model

Ruben Chung

Brown University, Providence, Rhode Island

1. Abstract

Forecasting touchdowns for NFL wide receivers is a challenging but valuable problem in football analytics and player evaluation. Touchdowns are notoriously volatile, influenced by red zone usage, quarterback play, and situational variance, making year-to-year outcomes difficult to predict. This study develops a temporal linear regression model to project wide receiver touchdown totals using a feature-rich dataset spanning 1990–2024. The dataset incorporates lagged statistics, two-year rolling averages, player age and experience, team offensive strength, and efficiency metrics such as catch rate and touchdowns per target. The model was trained on 1990–2010 player-seasons and tested on 2011–2024 data, with strict time-bounded feature engineering to prevent data leakage.

Results show strong predictive accuracy ($R^2 = 0.803$, MAE = 0.82 TDs), demonstrating that systematic patterns can be identified even within a highly volatile statistic. Feature importance analysis indicates that efficiency and usage metrics are more reliable predictors than raw prior-year touchdown totals, aligning with football intuition and highlighting regression-to-the-mean

effects. The model generates 2025 projections that identify both elite scorers and likely regression candidates, providing insight into the stability of touchdown production.

This work demonstrates that with careful feature engineering, a transparent and interpretable linear model can yield valuable insights in sports analytics. Beyond forecasting, the results underscore the importance of efficiency and opportunity metrics in understanding touchdown outcomes, offering a framework that can inform research on statistical predictability in professional sports.

2. Introduction

Touchdowns are a critical driver of value in football analytics and player evaluation. However, among all wide receiver (WR) statistics, touchdowns are the most volatile and thus among the hardest to predict. This volatility stems from their dependence on situational factors such as red zone usage, play-calling tendencies, and defensive matchups, which often vary widely from game to game and season to season. As a result, analysts and bettors often find touchdowns to be the hardest statistic to predict due to its inherent volatility. Early football analytics research emphasized team-level scoring and play-calling (Burke, 2009), with less attention given to forecasting individual player outcomes. More recent work has examined broader statistical properties of NFL performance (Lopez, Matthews, & Baumer, 2018) and player-level evaluation metrics such as nflWAR (Yurko, Ventura, Horowitz, & Balasubramanian, 2019). However, few studies have addressed the predictability of touchdowns specifically, particularly using temporally valid models.

This paper introduces a linear regression model designed to project wide receiver touchdowns using a multi-season dataset with both player-level and team-level features. Unlike models that rely solely on prior-year performance, we incorporate lag features, two-year rolling averages, age and experience indicators, and contextual team metrics such as offensive scoring and pace. This richer feature set allows the model to better account for nonlinear career arcs, regression to the mean, role changes within an offense, and changes in structure of the offense itself.

The model is trained on NFL player-seasons from 1990 to 2010 and tested on data from 2011 through 2024. Our approach emphasizes temporal integrity to avoid lookahead bias: all engineered features respect chronological boundaries. Despite the interpretability and transparency of a linear model, our results show it captures meaningful signal in a noisy domain and yields actionable projections. We conclude by applying the model to 2024 data to forecast 2025 WR touchdown outcomes.

For clarity, the following abbreviations are used throughout this paper: TD = touchdowns; RZ = red zone; MAE = mean absolute error; RMSE = root mean squared error; R^2 = coefficient of determination.

2.1 What is Linear Regression

Linear regression is one of the most widely used statistical techniques for modeling the relationship between a dependent variable and one or more independent variables. In our case, the dependent variable y is the number of receiving touchdowns, and the independent variables $x_1, x_2, \dots, x_n$ are the engineered features describing a player's past performance, efficiency, and context (things such as yards per game or catch rate for example)

The core idea of linear regression is to fit a straight-line (or hyperplane, in higher dimensions) relationship between inputs and output, such that the predicted values are as close as possible to the observed values. Linear Regression thus assumes:

1. The relationship between predictors and the target is approximately linear.
2. The residuals (errors) are independent and have constant variance.
3. No predictor is an exact linear combination of others (no perfect multicollinearity).

In practical terms, the model estimates how much the target variable changes (touchdowns), on average, when each feature changes by one unit, holding all other features constant.

2.2 Mathematical Formula of Linear Regression

For a dataset with n observations and p predictors, linear regression models the target y as:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i$$

where:

- $\boldsymbol{y_i}$= the actual value of the dependent variable for observation i
- $\boldsymbol{x_{ij}}$= the value of predictor j for observation i
- $\boldsymbol{\beta_0}$ = intercept term
- $\boldsymbol{\beta_j}$ = coefficient for predictor j
- $\boldsymbol{\varepsilon_i}$ = residual error for observation i

The goal is to find the coefficient vector $\overrightarrow{\beta} = (\beta_0, \beta_1, \dots, \beta_p$ that minimizes the Residual Sum of Squares (RSS)**:**

$$RSS(\beta) = \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2$$

This minimization yields the ordinary least squares (OLS) solution:

$$\beta = (X^T X)^{-1} X^T y$$

where:

- $X = n * (p + 1)$ matrix of features (including a column of ones for the intercept)
- $y = n * 1$ vector of target values

In the context of this paper each β_j represents the expected change in touchdowns for a given player for a change of one-unit for the feature x_j while assuming all other variables remain constant. Positive coefficients would imply a positive correlation between that feature and expected touchdowns scored – and vice-versa.

3. Materials & Methods

3.1 Data Sources

To build a model capable of projecting wide receiver touchdowns, a comprehensive dataset of NFL WR performance spanning over three decades was assembled. The dataset pulls together information from multiple sources and levels — including individual player statistics, rushing contributions, and some context about their team and situation.

The core data were scraped from Pro-Football-Reference.com (Pro-Football-Reference.com, 2025), including:

- Receiving stats (e.g., targets, receptions, yards, touchdowns, yards per reception)
- Rushing stats (for WRs with occasional carries)
- Team stats (e.g., points scored, total offensive plays, yards per play)

These were compiled into season-level records for every wide receiver from 1990 through 2024.

3.2 Feature Engineering

To support accurate preseason forecasting, a variety of derived features were constructed from player and team data. These included both static and historical variables, with all temporal metrics carefully bounded by year to prevent data leakage.

Static Features:

- Age and Experience Indicators: Variables such as player age, total seasons active, and binary flags for career stages (e.g., prime years at ages 25–29 or rookie/sophomore seasons at age ≤ 23).

- Individual Performance: Core metrics including Catch Rate (completions per target), targets per game, and yards per game.

- Scoring Efficiency: Touchdowns per reception **(**TD_Per_Reception**)** and per target **(**TD_Per_Target**),** capturing how effectively players convert opportunities into scores.

- Team Context: Offensive strength measures, such as Team_Offense_Strength (points per game).

Historical Features:

- Prior Season Performance: Key statistics from the previous year, including touchdowns, targets, receptions, and receiving yards (e.g., TD_Prev, Yds_Prev).

- Long-Term Trends: Two-year rolling averages for touchdowns, receptions, and yards (e.g., TD_Avg2, Rec_Avg2, Yds_Avg2) to capture sustained performance over time.

- Career Experience: Number of years a player has appeared in the dataset.

All features were designed to mirror the type of information available during the preseason, ensuring the model's evaluation aligns with real-world forecasting constraints.

3.3 Problem Framing

This section describes the process used to model and project WR touchdown totals. The approach involves preparing the dataset, training a linear regression model and testing it on a different dataset.

We frame WR touchdown prediction as a supervised regression task. For each player-season, the goal is to predict the number of receiving touchdowns (TD) a player will score based on a variety of contextual and historical features. The modeling target is continuous ($TD \in \mathbb{R}^{+}$), and we use a linear regression framework for its transparency and interpretability.

An inherent characteristic of this modeling framework is *regression to the mean*: extreme touchdown totals in one season are statistically likely to move closer to league-average levels in subsequent seasons. In a linear regression context, unless a predictor perfectly explains the target, the model's fitted values tend to be "pulled" toward the overall mean. This property helps prevent overestimation of repeat peak seasons and avoids overreacting to one-off scoring spikes, which is particularly important for touchdowns given their volatility.

3.4 Feature Set and Target Variable

The features used in the model fall into the following categories:

- Lag Features: Stats from the previous season (e.g., TD_Prev, Yds_Prev, Tgt_Prev, Y/R_Prev)
- Rolling Averages: Two-year rolling means for TDs, targets, receptions, and yards (TD_Avg2, etc.)
- Player Traits: Age, experience (Age, Age_Squared, Experience, Prime_Age, Rookie_Sophomore)
- Efficiency Metrics: TD_Per_Target, Catch_Rate, Target_Share, Yards_Per_Game
- Red-Zone Data (RZ): Rz_Targets, RZ_receptions, RZ_yards, RZ_TDs, RZ_INTs, RZ_Catch_Rate
- Team Context: Offensive output per game (Team_Offense_Strength), and scoring environment
- Durability: Games played

The target variable is the actual number of receiving touchdowns scored by a player in that season.

3.5 Temporal Validation and Data Leakage Prevention

To ensure robust backtesting, we implement a chronologically consistent train/ test split of the data set:

- **Training Set**: Player-seasons from 1990–2010
- **Test Set**: Player-seasons from 2011–2024

All temporal features (e.g., lag stats and rolling averages) are computed using only prior seasons up to the year being predicted. This simulates a true forward-looking forecast and ensures the prevention of data-leakage.

3.6 Model Choice and Evaluation

We trained the Linear Regression model using scikit-learn (Pedregosa et al., 2011), with the following characteristics:

- No regularization (ordinary least squares)
- Standardization applied to numeric features via StandardScaler

This choice was made to prioritize interpretability, making it easier to identify which features positively or negatively influence TD outcomes.

We apply 5-fold cross-validation on the training set to assess in-sample error and variance in the model. In this procedure, the training data is split into five equal parts – called folds. The model is trained on four folds and evaluated on the remaining fold, repeating so that each fold serves once as the evaluation set.

Five folds were chosen as a balance between:

- Bias and variance estimation – Too few folds can yield noisy estimates of model performance, while too many folds (e.g., leave-one-out) can produce low-bias but high-variance estimates.
- Computational efficiency – Five folds provide stable performance metrics without making training time excessive.

Metrics reported include:

- MAE (Mean Absolute Error)
- RMSE (Root Mean Squared Error)
- R^2 Score

The final model is evaluated on the full test set (2011–2024) to measure out-of-sample generalization.

4. Results and Data Visualization

The linear regression model was trained on NFL WR data from 1990 to 2010 and evaluated on a temporally isolated test set spanning 2011 through 2024. Results from the out-of-sample backtest show that the model captures meaningful signal in a high-variance target variable: receiving touchdowns.

4.1 Top Features

The figure below shows the features used in training and testing the model, a positive coefficient value implies a positive correlation with touchdowns (and vice-versa). A larger magnitude implies a stronger correlation.

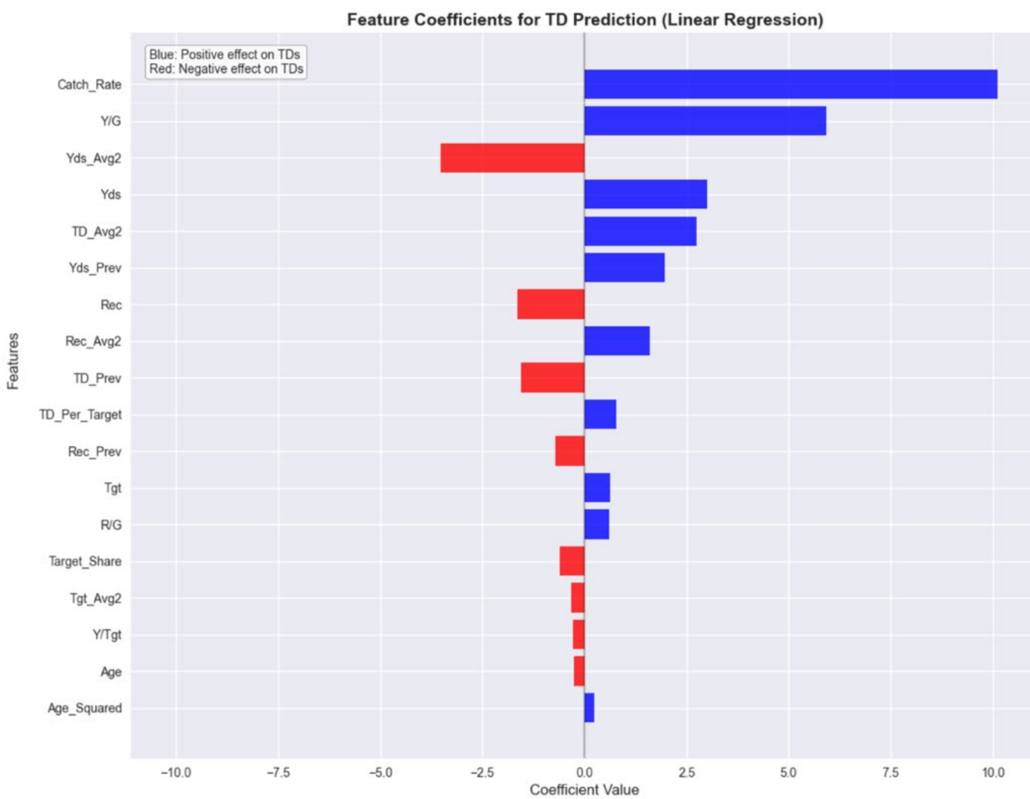

Figure 1. Feature coefficients for touchdown prediction using linear regression. Positive coefficients (blue) indicate features that increase projected touchdowns; negative coefficients (red) decrease them.

The most influential features in the model, ranked by coefficient magnitude, were:

1. Catch Rate – A higher catch rate reflects a receiver's reliability and ability to convert targets into completions. Consistently catching passes increases red-zone efficiency and scoring opportunities.

2. Yards Per Game – Sustained yardage production per game signals a player's central role in the offense and strong target volume, both of which correlate strongly with touchdown opportunities.
3. Two-Year Average Receiving Yards (Yds_Avg2) — In this model, the coefficient is negative conditional on other features (e.g., catch rate, usage). This can occur if high-yardage profiles come from between-the-20s usage while lower-yardage receivers see proportionally more red-zone targets.

Notably, the model emphasizes efficiency and sustained production metrics over raw prior-year touchdown totals. This aligns with established football intuition: touchdowns are more often the product of consistent usage and high-value opportunities than the simple repetition of past high-scoring seasons. By identifying players with stable efficiency profiles, the model highlights candidates most likely to sustain or improve their touchdown output.

4.2 Test Set Performance

The model achieves strong performance on the 2011–2024 test set, as shown in Table 1.

Metric	Value
Mean Absolute Error (MAE)	0.82 TDs
Root Mean Squared Error (RMSE)	1.29 TDs
R^2 Score	0.803

Table 1: Performance metrics for the linear regression model predicting WR touchdowns (2011–2024 test set).

The low mean absolute error (MAE) of 0.82 touchdowns indicates that most predictions are within ±1 TD of actual outcomes. The root mean squared error (RMSE) of 1.29 TDs reflects a modest penalty for larger errors, while the R^2 of 0.803 suggests that the model explains over 80% of the variance in wide receiver touchdown totals. These results demonstrate strong predictive accuracy for a linear model, particularly given the high volatility and perceived non-linearity of touchdowns.

4.3 Cross-Validation (Train Set)

During training (1990–2010), the model achieved a cross-validation MAE of 0.45 ± 0.08, suggesting low variance and good generalization. The consistency between train and test scores supports the model's robustness and lack of overfitting.

4.4 Data Visualization

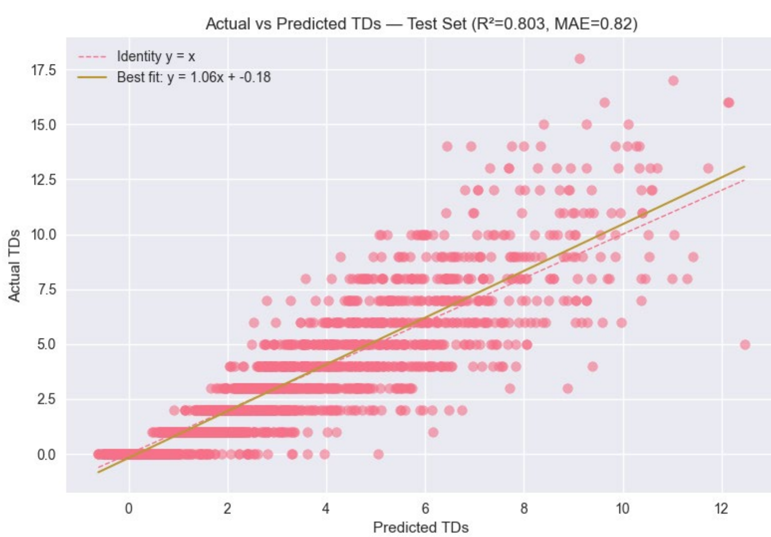

Figure 2. Predicted vs actual WR touchdowns (2011-2024 test set) with identity and best-fit lines; alignment indicates low systematic bias.

The best fit regression line is given by the equation: $y = 1.06x - 0.18$ where y represents actual touchdowns scored, and x represents our models projected touchdowns. Our slope being 1.06, greater than 1, implies a slight tendency to underpredict at the high end of the distribution, while the intercept of –0.18 is close to zero, indicating minimal systematic bias.

A comparison of the model's best-fit line to the identity line ($y =$ x, perfect predictions) reveals systematic patterns. When touchdowns are low (< 3) the identity line is above the model's best fit line. This means that in this range of touchdowns the model slightly overpredicts touchdowns. On the contrary, when touchdowns are higher (> 3) the identity line is below the best fit line. This means the model slightly underpredicts touchdowns. In fact, the distance between the two lines increases as touchdowns increase, so this under projection becomes worse as touchdowns increase (reflecting the increased variance in elite level touchdown scoring). This reflects both the increased variance among elite scorers and the model's natural regression toward the mean.

4.5 Projections for Upcoming Season

After evaluation, the trained model was applied to WRs from the 2024 season to generate touchdown projections for 2025.

The model's projections for 2025's top 10 touchdown scorers are presented in Table 2.

Player	**Tm**	**Age**	**Games**	**TD Projection**	**TD / Game**
Ja'Marr Chase	CIN	24	17	11.03	0.649

Justin Jefferson	MIN	25	17	9.85	0.579
Brian Thomas	JAX	22	17	8.52	0.501
Drake London	ATL	23	17	7.89	0.464
Terry McLaurin	WAS	29	17	7.69	0.452
Amon-Ra St. Brown	DET	25	17	7.4	0.435
A.J. Brown	PHI	27	13	7.26	0.558
Ladd McConkey	LAC	23	16	7.26	0.454
Jerry Jeudy	CLE	25	17	7.14	0.42
Jameson Williams	DET	23	15	7.09	0.473

Table 2. Model projections for the top 10 wide receivers by touchdown total in the 2025 NFL season (based on 2024 data).

These results align closely with expectations for elite WRs, with Ja'Marr Chase topping the list with 11.03 projected touchdowns**.** While this is still an elite projection, it represents a decline from his 2024 production due to reversion to the mean — the statistical tendency for extreme performances to move closer to league-average levels in subsequent seasons.

The list includes many established stars (e.g., Justin Jefferson, Amon-Ra St. Brown, A.J. Brown), however, the model does not incorporate external situational factors for 2025 that could materially influence these projections. For example:

- Justin Jefferson and Drake London will be playing with new quarterbacks.

- Brian Thomas may see target share competition from 2nd overall pick, rookie teammate Travis Hunter.

These contextual elements, while important for interpretation, fall outside the model's current feature set. Future work could incorporate roster changes, quarterback stability, and offensive scheme adjustments to better capture these external influences and provide perhaps more accurate projections.

5. Discussion

The results of our linear regression model provide several important insights into the predictability of WR touchdowns and the broader dynamics of NFL scoring.

5.1 Interpretable and Stable Performance

The model's R^2 of 0.803 and MAE under 1 TD demonstrate that touchdowns, while high variance, are not entirely random. By relying on a broad, temporally valid feature set, the model captures stable indicators of future scoring — particularly catch efficiency, receiving volume, and team context. This shows that touchdown regression and breakout candidates can be detected systematically using historical data.

Moreover, the model's strong cross-validation performance (MAE = 0.45 ± 0.08) further validates its generalizability and lack of overfitting, despite the additive, linear structure of the model.

5.2 Feature Importance Reveals Underlying Mechanics

The most predictive features — such as Catch_Rate, Yards_Per_Game — reflect sustainable opportunity and role in an offense rather than raw scoring alone. This aligns with the football intuition that touchdowns are often a function of consistent usage and efficiency rather than isolated high-touchdown seasons.

Interestingly, past touchdown totals (TD_Prev) did not rank among the top linear features, suggesting that surface-level regression models relying on "he scored X touchdowns last year" may be overly simplistic; guidelines based on prior-year scoring are insufficient for reliable forecasting in such a volatile statistic.

5.3 Limitations

Limitations

While this project provides a strong foundation for touchdown prediction, the model has several notable limitations:

- **Linear form:** Cannot capture nonlinear effects, such as diminishing returns or interactions between variables (e.g., age and usage).
- **No injury or situational awareness:** Lacks inputs for offseason and in-season changes, including depth chart shifts, quarterback changes, and play-calling adjustments.
- **No adversarial defense data:** Matchup strength and opposing defensive rankings were not incorporated.
- **Limited positional scope:** Focuses exclusively on wide receivers, excluding other offensive positions.

Future Work

Future research could address these gaps by:

- Applying the methodology to predict other wide receiver statistics—such as receiving yards or receptions—to build a more complete player profile.
- Expanding the framework to include additional positions, such as tight ends or running backs.
- Incorporating richer data sources, including play-by-play logs, tracking metrics, defensive matchup data.
- Adding offseason and situational information, such as coaching staff changes, injury history, and team roster moves.
- Exploring more advanced modeling techniques—such as tree-based ensembles or neural networks—to capture nonlinear relationships and complex feature interactions.

5.4 Practical Applications

Despite its simplicity, the model offers clear, actionable value across multiple areas. Its linear regression framework makes it highly interpretable, allowing outputs to be analyzed and trusted in real-world decision-making. By ranking players based on projected touchdowns and grouping them into tiers, the framework can inform evaluation. In professional contexts, the model highlights the attributes most strongly associated with scoring, assisting scouts, analysts, and coaching staff in talent assessment and development. In broader analytics applications, the projections and feature importance provide a benchmark for understanding which statistics most reliably translate into future scoring outcomes.

6. Conclusion

This study developed and validated a linear regression model to forecast NFL WR touchdown totals using feature-rich, temporally consistent dataset spanning 1990–2024. By incorporating lagged performance statistics, rolling averages, efficiency metrics, and team-level context, the model achieved strong predictive accuracy in one of football's most volatile metrics, touchdowns.

Evaluation on an out-of-sample test set (2011–2024) yielded an R^2 of 0.803 and an MAE of 0.82 touchdowns, demonstrating that even in the presence of perceived randomness in scoring touchdowns, systematic patterns can be identified and exploited. The model's interpretability allowed for clear insights into which factors most influence touchdown outcomes — with efficiency and volume metrics outperforming raw prior-year touchdown counts as predictors.

The resulting projections for 2025 offer actionable guidance for sports analytics and player evaluation by identifying likely regression and breakout candidates. While the linear framework is interpretable and robust, future work could explore nonlinear modeling, additional contextual variables such as red zone usage or quarterback efficiency, and integration of play-by-play tracking data.

In sum, this work shows that with careful feature engineering and strict prevention of data leakage, a transparent linear regression model can provide valuable and interpretable forecasts for one of football's most volatile statistics.

Acknowledgements

The author is an undergraduate student at Brown University, Providence, RI. This research was conducted independently and did not receive formal guidance or funding from a specific course, program or research grant.

References

Burke, B. (2009). The hidden game of football: A quantitative analysis of scoring and play calling in the NFL. *Journal of Quantitative Analysis in Sports, 5*(3), Article 6. https://doi.org/10.2202/1559-6458.1102

Lopez, M. J., Matthews, G. J., & Baumer, B. S. (2018). How often does the best team win? A unified approach to understanding randomness in North American sport. *The Annals of Applied Statistics, 12*(4), 2483–2516.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12,* 2825–2830.

Pro-Football-Reference.com. (2025). *NFL player statistics.* Sports Reference LLC. https://www.pro-football-reference.com/

Yurko, R., Ventura, S. L., Horowitz, M., & Balasubramanian, V. (2019). nflWAR: A reproducible method for offensive player evaluation in football. *Journal of Quantitative Analysis in Sports, 15*(3), 163–183.

Statements and Declarations

Ethical considerations

Not applicable. This research used publicly available secondary data (NFL statistics) and did not involve human participants, human tissue, or human data requiring ethical approval.

Consent to participate

Not applicable

Consent for publication

Not applicable

Declaration of conflicting interest

The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding statement

The author received no financial support for the research, authorship, and/or publication of this article.

Data availability

All raw data used in this study are publicly available from Pro-Football-Reference (https://www.pro-football-reference.com/). Processed datasets and model predictions generated during the current study are available from the author on reasonable request.