Statistical Modeling of NFL Wide Receiver Fantasy Performance

Ethan He
Providence Day School
tiancheng.ethan.he@gmail.com

July 16, 2025

Abstract

This study develops and evaluates statistical models for predicting NFL wide receiver fantasy football performance, a notoriously volatile domain. Using data from the 2022 and 2023 seasons, key predictor variables were engineered—including target share, yards per route run, quarterback quality, depth chart movement, and age to forecast fantasy points for the following season. Both ordinary least squares (OLS) and robust regression models were constructed. Diagnostic tests confirmed the validity of the models, while out-of-sample validation on 2024 data demonstrated moderate predictive accuracy ($R^2 = 0.419$). Though the models captured core trends, they were outperformed by expert consensus rankings in mean absolute rank error (8.581 vs. 6.674), reflecting the enduring value of contextual insights. These findings highlight both the strengths and limitations of quantitative modeling in fantasy sports and suggest that future improvements in predictive accuracy may be achieved by expanding feature sets and adopting hybrid analytical approaches.

1 Introduction

Fantasy football is a game enjoyed by millions of fans, where participants assemble virtual teams of real-life NFL players and compete based on those players' weekly performances. The immense popularity of fantasy football has fueled a demand for accurate player projections, as managers seek every possible edge in drafting and lineup decisions. However, forecasting outcomes—particularly for wide receivers—remains a persistent challenge due to factors such as injuries, evolving offensive schemes, and unpredictable player development.

To provide context for the analysis, here are the standard fantasy point scoring rules for wide receivers:

- 0.1 points per yard gained (10 yards = 1 point)
- 1 point per reception
- 6 points per touchdown

These rules form the basis for evaluating wide receiver performance in fantasy football and are central to the construction and validation of predictive models.

This study aims to evaluate the effectiveness of statistical modeling for predicting NFL wide receiver fantasy performance. By leveraging a comprehensive set of features—including target share, yards per route run, competition change, quarterback skill, and age—this research constructs and validates regression models, compares their predictive accuracy to industry expert projections, and explores the strengths and limitations of quantitative forecasts. The goal is to provide actionable insights for fantasy managers and to advance the integration of analytics and expert judgment in fantasy football strategy.

2 Data and Feature Engineering

The predictive power of any statistical model in fantasy football depends heavily on the quality of its data and the relevance of its features. This section outlines the process of assembling the dataset, cleaning the data, and engineering the key variables used to model wide receiver fantasy performance.

2.1 Data Sources and Preparation

Data were compiled from multiple reputable sources, including official NFL statistics and leading fantasy football platforms. To ensure the analysis focused on meaningful contributors, only wide receivers who played at least 50% of their team's snaps and logged a minimum of 250 total snaps were included (FantasyPros, 2023). Player names were standardized across datasets to ensure accurate merging and consistency.

2.2 Variable Definitions and Coding

Drawing on prior research and domain expertise (Kapania, 2012; Yardbarker, 2024), the following variables were selected to serve as the foundation for the modeling process. Predictor variables were based on data from the previous season, while the dependent variable—fantasy points—was defined on data from the upcoming season.

- **Target Share:** Represents the proportion of team passing attempts directed at a player, directly reflecting their involvement in the offense. Based on prior research and domain knowledge, a *positive* correlation with fantasy points is expected, as increased opportunity typically leads to higher production.

- **Yards Per Route Run (YPRR):** Quantifies a receiver's efficiency by measuring the average yards gained per route run. This metric captures both productivity and usage. A *positive* relationship with fantasy points is anticipated, as more efficient receivers generally accumulate more points.

- **Competition Change:** Captures changes in a player's depth chart position, such as moving up to a starting role or facing new competition for targets (see Appendix A for further details on variable coding). A decrease in competition for targets is expected to be *positively* correlated with fantasy points, while increased competition may have a *negative* association.

- **Quarterback Skill:** Reflects the quality of the quarterback throwing to the receiver, which can significantly influence receiver performance. A better quarterback is expected to be *positively* correlated with fantasy points, as higher quarterback skill should enhance a receiver's scoring opportunities.

- **Age:** Differentiates between younger and older receivers, as age can influence durability, physical ability, and career stage. Drawing on established trends in football analytics, a *negative* correlation with fantasy points is expected, since wide receivers often experience declines in performance as they age past certain thresholds. (The Fantasy Footballers, 2025).

Table 1 summarizes the definition and coding scheme of each variable.

Variable	**Description**	**Coding Scheme**
Target Share	Proportion of team passing targets	Numeric (continuous)
Yards Per Route Run	Yards gained per route run	Numeric (continuous)
Competition Change	Change in depth chart opportunity	1 = less, 0 = no change, -1 = more
Quarterback Skill	Projected QB's prior performance	1 = high, 0 = neutral/rookie, -1 = low
Age	Under or over 30 years old	0 = under 30, -1 = 30 or older

Table 1: Variable definitions and coding schemes used in the analysis.

3 Materials and Methods

3.1 Exploratory Data Analysis

Exploratory data analysis (EDA) was conducted to examine distributions, identify outliers, and highlight relationships among key variables before modeling.

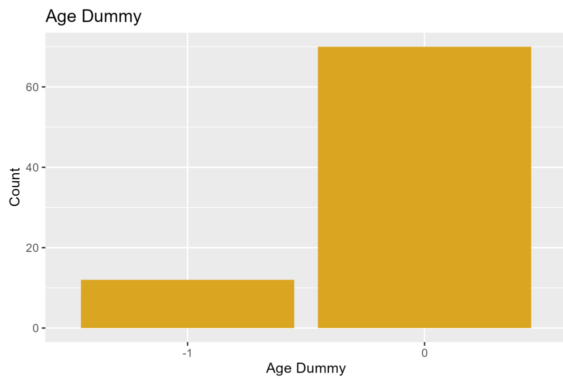
3.2 Univariate Analysis: Categorical Features

Bar plots were used to visualize the following categorical variables:

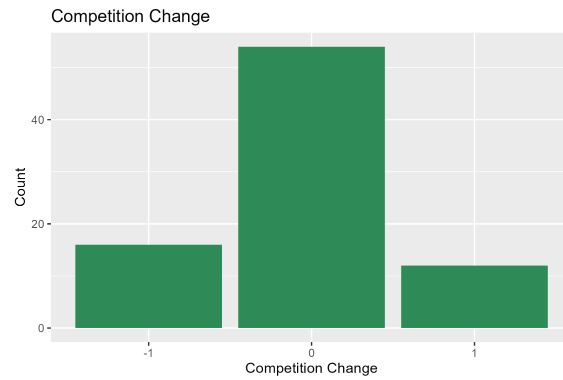
- **Age Group:** Figure 1a shows most receivers were under 30 (0), reflecting both youth-related performance advantages and the dataset's filtering criteria.

- **Competition Change:** Figure 1b indicates most players had no change in competition (0), suggesting stable wide receiver depth charts.

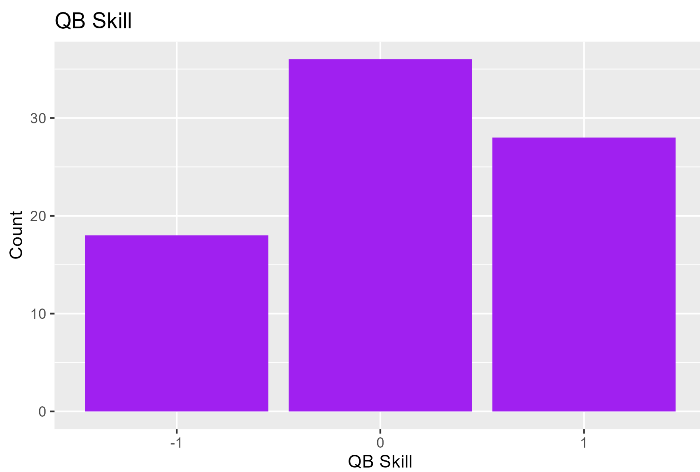
- **Quarterback Skill:** Figure 1c shows most receivers played with neutral or strong quarterbacks, aligning with the NFL's pass-heavy trend.



(a) Bar plot of wide receiver age groups (0 = under 30, −1 = 30 or older).



(b) Bar plot of competition change variable (1 = less competition, 0 = no change, −1 = more competition).



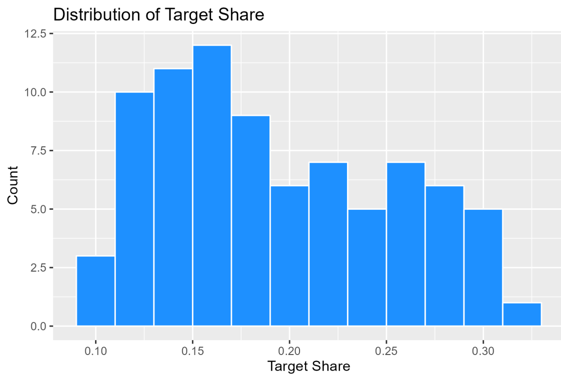
(c) Bar plot of quarterback skill variable (1 = high, 0 = neutral/rookie, −1 = low).

3.3 Univariate Analysis: Continuous Features

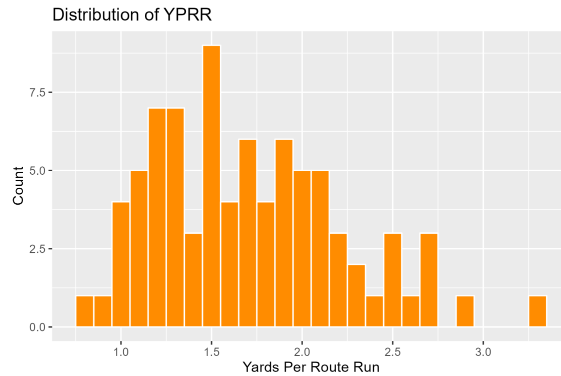
Histograms were used to visualize the distributions of continuous variables:

- **Target Share:** Figure 2a shows most receivers had a target share between 13% and 25%, with a long tail of high-usage players—reflecting overall involvement in the passing game.

- **Yards Per Route Run (YPRR):** Figure 2b shows most players fell between 1.0 and 2.0 YPRR, with a few outliers above 2.5 indicating exceptional efficiency.



(a) Histogram of target share among wide receivers.



(b) Histogram of yards per route run (YPRR) among wide receivers.

3.4 Bivariate Analysis: Correlation Structure

To assess relationships between predictors and the response variable, a correlation matrix was computed and visualized as a heatmap (Figure 3). Target share and Yards Per Route Run (YPRR) exhibited the strongest positive correlations with fantasy points ($r = 0.77$ and $r = 0.72$, respectively), reaffirming their importance in the modeling process. Competition change and quarterback skill showed weaker but directionally consistent associations.

The age variable, coded as -1 for players 30 or older, displayed a negative correlation with fantasy points. This reflects a survivorship bias: older players included in the sample had high snap counts and likely sustained productivity, making them unrepresentative of aging receivers generally. Thus, the result does not indicate a broader trend of improved performance with age but highlights the durability of top-tier veterans.

Finally, the moderately high correlation between target share and YPRR ($r = 0.77$) will be addressed during model diagnostics to evaluate potential multicollinearity.

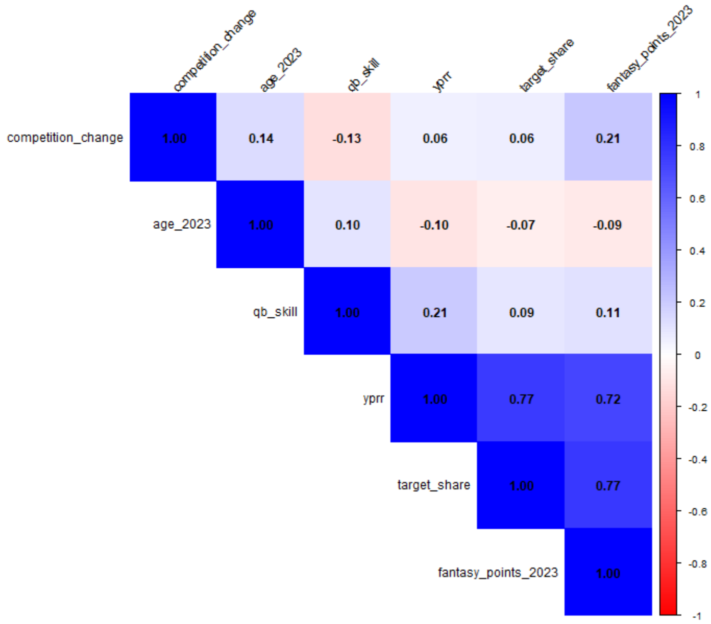


Figure 3: Correlation matrix heatmap for all predictors and fantasy points. Strong positive correlations are observed between target share, YPRR, and fantasy performance.

The key insights revealed by exploratory data analysis provide a solid foundation for the subsequent construction and evaluation of statistical models for wide receiver fantasy performance.

3.5 Model Construction

The objective of this analysis is to construct a statistical model capable of predicting wide receiver fantasy football performance using a set of theoretically and empirically motivated features. Multiple linear regression was selected as the primary modeling approach, allowing for the quantification of relationships between the response variable and several predictors simultaneously.

The response variable, total fantasy points for the season, was modeled as a linear function of the following predictors: target share, yards per route run (YPRR), competition change, quarterback skill, and age. The regression model is specified as:

$$\begin{aligned}\text{FantasyPoints}_{i,t+1} = \beta_0 + \beta_1 \cdot \text{TargetShare}_{i,t} + \beta_2 \cdot \text{YPRR}_{i,t} \\ + \beta_3 \cdot \text{CompetitionChange}_{i,t} + \beta_4 \cdot \text{QB\_Skill}_{i,t} + \beta_5 \cdot \text{Age}_{i,t} + \epsilon_{i,t+1}\end{aligned}$$

where:

- $\text{FantasyPoints}_{i,t+1}$ = Total fantasy points for player i at time t+1

- $\text{TargetShare}_{i,t}$ = Proportion of team passing targets at time t
- $\text{YPRR}_{i,t}$ = Yards per route run at time t
- $\text{CompetitionChange}_{i,t}$ = Change in depth chart opportunity at time t
- $\text{QB\_Skill}_{i,t}$ = Indicator for the quality of the team's starting quarterback at time t
- $\text{Age}_{i,t}$ = Indicator for being 30 years old or older at time t
- $\epsilon_{i,t+1}$ = Random error term

After fitting the model to the 2022 and 2023 season data, the estimated coefficients and their associated statistics are summarized in Table 2.

Variable	**Estimate**	**Std. Error**	**t-value**	**p-value**
Intercept	-43.98	64.10	-0.686	0.495
Target Share	730.17	143.73	5.080	$<$0.001
Yards Per Route Run	47.27	17.20	2.748	0.007
Competition Change	24.93	9.73	2.561	0.012
QB Skill	3.31	7.86	0.421	0.675
Age (30 or older)	-1.57	2.21	-0.711	0.480

Table 2: OLS regression coefficients for the 2023 wide receiver fantasy points model.

All variables were included based on their theoretical relevance and prior evidence of predictive value (Yardbarker, 2024). Automated variable selection was not performed at this stage to maintain interpretability and ensure that all hypothesized factors were assessed.

Of the five predictors included in the model, target share, yards per route run, and competition change were statistically significant at the 5% level. The signs of all estimated coefficients conformed to hypothesized directions, with the exception of the age variable for the reasons outlined in Section 3.3.

Checks for model assumptions such as multicollinearity, normality, and homoscedasticity are presented in the subsequent Model Diagnostics section.

4 Model Diagnostics

This section evaluates the assumptions and overall fit of the model. A comprehensive suite of diagnostic tests and graphical analyses was conducted to assess model validity, identify potential issues such as outliers or heteroscedasticity, and ensure reliable inference.

4.1 Model Fit Statistics

The OLS regression model demonstrated strong explanatory power and statistical significance, as summarized below:

- **Residual Standard Error:** 50.02
- **Adjusted $\mathbf{R^2}$:** 0.645
- **F-statistic:** 30.47 on 5 and 76 degrees of freedom ($p < 2.2e{-}16$)

These metrics indicate that the model explains approximately 64.5% of the variance in wide receiver fantasy points for the 2023 season, with the overall regression highly significant.

4.2 Diagnostic Plot

The diagnostic plot shown in Figure 4 summarizes key regression diagnostics including residuals behavior, normality, homoscedasticity, and leverage.

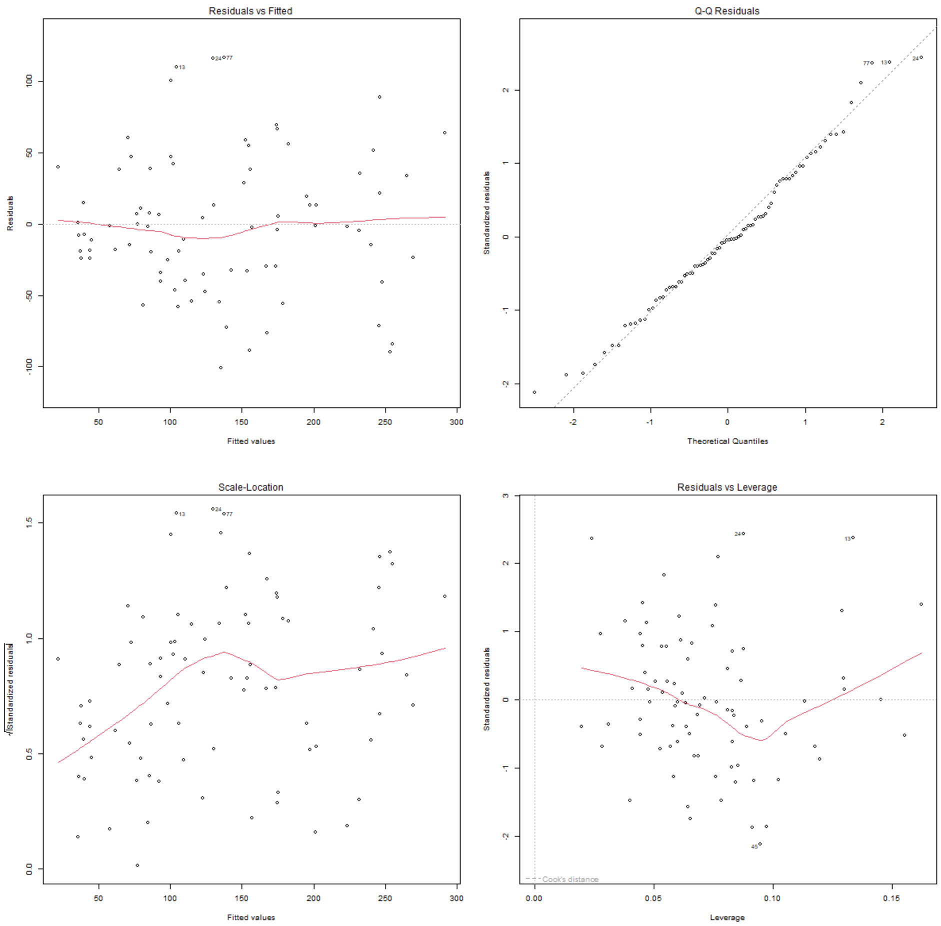


Figure 4: Diagnostic plot for the OLS regression model assessing linearity, normality, homoscedasticity, and influential observations.

The above plots provide a comprehensive check of the regression model's assumptions.

- **Residuals vs. Fitted Plot (Top Left):** This plot assesses linearity and constant variance. The residuals generally scatter randomly around zero, which supports the linearity assumption. However, there is some curvature and slight fanning in the vertical spread of residuals at both ends of the fitted range, suggesting mild heteroskedasticity. A few moderate outliers—such as observations #13, #77 and #24—are visible and labeled for reference.

- **Normal Q-Q Plot (Top Right):** This plot checks whether the residuals follow a normal distribution. Most points align closely with the diagonal line, indicating approximate normality. Deviations at the tails (e.g., points #13, #77, and #24) show some skew, but these are not severe enough to challenge the model's validity.

- **Scale-Location Plot (Bottom Left):** This plot evaluates homoscedasticity more directly by displaying the square root of standardized residuals. The red trend line is relatively flat across the range of fitted values, though a slight upward slope at higher values reinforces the presence of mild heteroskedasticity.

- **Residuals vs. Leverage Plot (Bottom Right):** This plot identifies points that may exert undue influence on the model. While a few observations (e.g., #24, #45, and #13) exhibit elevated leverage, none exceed the Cook's distance threshold, suggesting no single data point is disproportionately driving the model's results.

All significant outliers were manually reviewed and found to reflect real-world player outcomes rather than data entry errors or anomalies. These observations were retained in the dataset to preserve the model's applicability to the full range of wide receiver performances—including cases shaped by injuries, breakout seasons, or unexpected depth chart changes. Removing them would risk sanitizing the dataset and weakening the model's relevance for practical fantasy forecasting.

While some mild heteroskedasticity is present, formal statistical tests (discussed in Section 4.3) did not detect significant non-constant variance. The use of robust regression in later sections serves to complement OLS by offering additional stability, particularly in the presence of the outliers identified here.

4.3 Statistical Tests

To assess model assumptions, several key statistical tests were conducted:

- **Shapiro-Wilk Test for Normality (Table 3):** This test evaluates whether the model residuals are normally distributed. The resulting p-value of 0.378 indicates no significant departure from normality, supporting the use of OLS inference.

- **Breusch-Pagan Test for Heteroscedasticity (Table 4):** This test checks for non-constant variance (heteroscedasticity) in the residuals. The p-value of 0.165 suggests that there is no significant evidence of heteroscedasticity, further validating the regression assumptions.

Table 3: Shapiro-Wilk Test for Normality of Residuals

Test	Statistic	p-value
Shapiro-Wilk	0.984	0.378

Table 4: Breusch-Pagan Test for Heteroscedasticity

Test	Statistic	df	p-value
Breusch-Pagan	7.842	5	0.165

- **Multicollinearity Assessment (Table 5):** Variance Inflation Factors (VIFs) were computed for all predictors to assess multicollinearity. All VIF values are below common thresholds of concern (typically 5 or 10), indicating that multicollinearity is not a major issue in this model. Notably, while target share and yards per route run exhibit a moderately high correlation (r = 0.77), the VIFs suggest this does not substantially inflate variance in coefficient estimates.

Table 5: Variance Inflation Factors (VIF) for Predictors

Predictor	VIF
Target Share	2.28
Yards Per Route Run	1.59
Competition Change	1.06
Quarterback Skill	1.08
Age	1.10

4.4 Endogeneity Considerations

Endogeneity is unlikely in this model, as all predictors—such as target share, YPRR, competition change, quarterback skill, and age—are based on prior-season data or fixed attributes determined before the target season. These variables are not influenced by in-season developments and are unlikely to correlate with unobserved factors affecting outcomes. The model also excludes jointly determined variables, minimizing the risk of bias. Given these conditions, the use of ordinary least squares (OLS) is appropriate.

4.5 Summary

These diagnostic results collectively affirm that the OLS regression model is appropriate for the data, with assumptions reasonably satisfied. Minor concerns related to outliers and slight heteroscedasticity motivate the use of robust regression methods, which are explored in subsequent sections.

5 Validation and Outlier Analysis

This section evaluates the predictive performance of the OLS regression model on out-of-sample data, examines the impact of outliers, and introduces robust regression as a method to address them.

5.1 Validation Procedure and Metrics

The OLS model, trained on 2022 and 2023 data, was validated using 2024 outcomes. Players who recorded zero fantasy points—typically deep depth chart receivers with little playing time—were excluded to prevent distortion of evaluation metrics and better reflect practical fantasy decision-making.

Table 6: Validation Metrics for OLS Model

Metric	Value
Root Mean Squared Error (RMSE)	71.60
Validation R^2	0.419

The model captured general performance trends, as indicated by a moderate R^2 and RMSE. However, prediction accuracy varied, with several significant outliers. This reflects a common challenge in fantasy modeling, where injuries, role changes, and unexpected breakouts often drive outcomes beyond what preseason data can capture (Draft Sharks, 2023; Bleacher Report, 2018). The model is useful for ranking tiers of players but less reliable for forecasting extreme cases.

5.2 Residual Analysis and Outlier Identification

To assess prediction errors, residuals were plotted against predicted values (Figure 5). While most values clustered near zero—indicating accurate forecasts—several players exhibited large residuals.

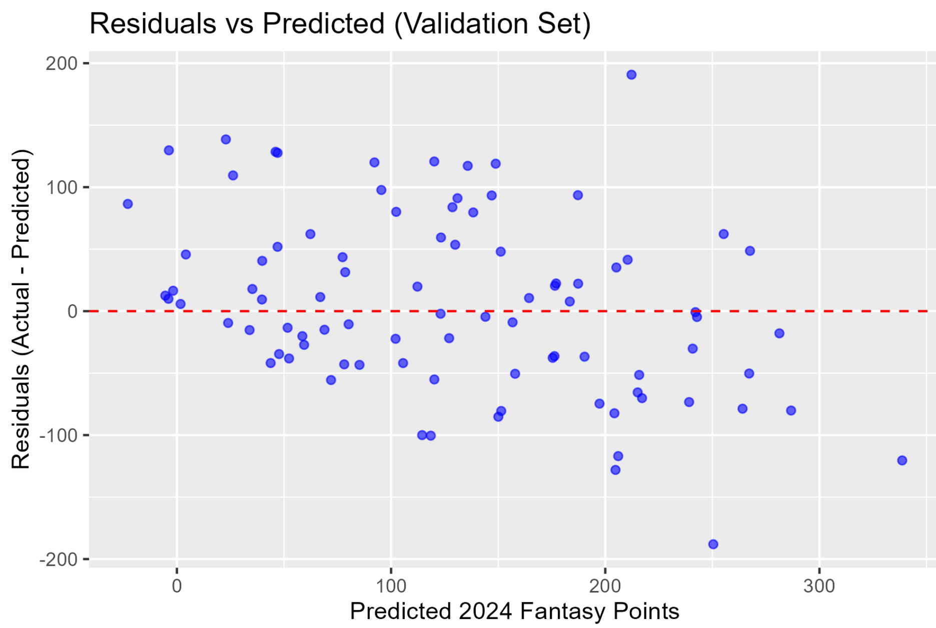


Figure 5: Residuals versus predicted values for the OLS model on the validation set. Large residuals indicate outliers.

The top 10 outliers were identified by absolute residual magnitude (Table 7). Most were tied to injuries or unexpected shifts in usage.

Table 7: Top 10 Outliers in Validation Set

Player Name	Actual FP (2024)	Predicted FP	Residual
Ja'Marr Chase	403.0	212.27	190.73
Brandon Aiyuk	62.4	250.44	-188.04
Alec Pierce	161.4	22.86	138.54
Allen Lazard	126.0	-3.74	129.74
Rashod Bateman	174.6	46.07	128.53
Chris Olave	76.7	204.78	-128.08
Quentin Johnston	174.7	47.01	127.69
Jerry Jeudy	240.9	120.16	120.74
Tyreek Hill	218.2	338.63	-120.43
Jameson Williams	212.2	92.23	119.98

These deviations illustrate the volatility of fantasy outcomes. Ja'Marr Chase greatly outperformed projections, likely due to elite skill and favorable conditions. Others, like Brandon Aiyuk and Chris Olave, underperformed due to injuries. Such cases emphasize the limits of statistical models in capturing in-season dynamics.

5.3 Discussion of Robust Regression

Rather than exclude outliers, robust regression was used to down-weight their influence while retaining the full dataset. This approach preserves the unpredictable and dynamic nature of fantasy football, making the model more applicable to the wide range of outcomes managers actually face. As with the OLS model, only wide receivers with at least 250 snaps and 50% snap share were included (FantasyPros, 2024).

The robust regression produced similar coefficient directions but reduced residual standard error from 50.02 to 43.84, reflecting improved fit and resilience to high-error observations.

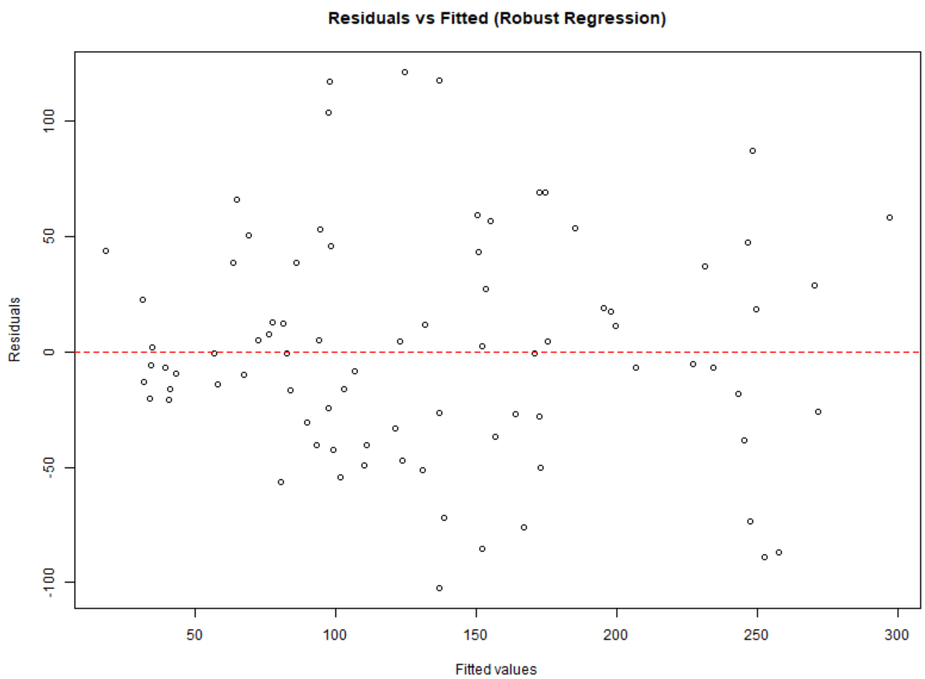


Figure 6: Residuals versus fitted values for the robust regression model. Residuals are more tightly clustered around zero compared to OLS.

5.4 Summary

The OLS model effectively identifies player tiers but is limited by its inability to capture extreme, context-driven outcomes. Robust regression enhances model stability by mitigating the influence of outliers. Together, these results highlight the importance of combining statistical analysis with qualitative football insight to better navigate fantasy football's unpredictability.

6 Comparison with Expert Rankings

A critical benchmark for any statistical model in fantasy football is its ability to compete with, or surpass, the accuracy of industry expert rankings. This section compares the performance of the regression model against expert consensus rankings, highlighting both the strengths and limitations of data-driven projections relative to expert judgment.

6.1 Methodology for Comparison

To provide a fair and relevant evaluation, the top 24 wide receivers predicted by the regression model were compared to the actual top 24 fantasy finishers, as well as to expert consensus rankings published before the 2024 season (sourced from Yahoo Fantasy (Yahoo Sports, 2024)). Players not present in a given list were assigned a rank of 25 (the maximum rank plus one) to avoid artificially inflating accuracy. The primary metric used for comparison was mean absolute rank error (MARE), which quantifies the average difference between predicted and actual ranks.

6.2 Summary Table of Ranking Accuracy

Table 8: Mean Absolute Rank Error (MARE) for Model and Expert Rankings

Ranking Source	MARE
Regression Model	8.581
Expert Consensus (Yahoo)	6.674

As shown in Table 8, the expert consensus rankings achieved a lower mean absolute rank error than the regression model, indicating closer agreement with actual outcomes. This finding aligns with prior analyses showing that expert rankings—which integrate statistical knowledge with qualitative insights—often outperform automated models for high-variance positions like wide receiver (Draft Sharks, 2023; Bleacher Report, 2018).

6.3 Visual Comparison of Top 24 Accuracy

To further illustrate the ranking accuracy of each method, Figure 7 presents a comparison matrix displaying the model, expert, and actual ranks for the top 24 wide receivers. Green highlights indicate cases where the model or expert projection was within five spots of the actual finish, providing a clear visual of where each method succeeded or struggled.

	player_name	model_rank	expert_rank	actual_rank
1	Ja'Marr Chase	16	3	1
2	Justin Jefferson	7	4	2
3	Amon-Ra St. Brown	4	5	3
4	Brian Thomas Jr.	25	25	4
5	Drake London	25	10	5
6	Malik Nabers	25	23	6
7	Terry McLaurin	25	25	7
8	CeeDee Lamb	3	1	8
9	Jaxon Smith-Njigba	25	25	9
10	Garrett Wilson	17	6	10
11	Davante Adams	10	12	11
12	Jerry Jeudy	25	25	12
13	Ladd McConkey	25	25	12
14	Mike Evans	19	11	14
15	Courtland Sutton	25	25	15
16	DJ Moore	9	18	16
17	Tee Higgins	25	25	17
18	Tyreek Hill	1	2	18
19	Jakobi Meyers	25	25	19
20	A.J. Brown	5	7	20
21	Jordan Addison	25	25	21
22	Jameson Williams	25	25	22
23	Nico Collins	11	15	23
24	Jauan Jennings	25	25	24

Showing 1 to 24 of 24 entries.

Figure 7: Comparison of model and expert rankings against actual top-24 wide receiver fantasy finishes. Green highlights indicate predictions within five ranks of the actual outcome.

6.4 Interpretation and Insights

- The expert consensus outperformed the regression model in overall accuracy, achieving a lower mean absolute error and more frequent near-matches to actual rankings. This reflects a unique advantage of expert analysts, who incorporate up-to-date team dynamics, player health, and qualitative assessments that go beyond quantifiable data.

- Despite not outperforming the experts, the regression model accurately identified several top performers, including Amon-Ra St. Brown and CeeDee Lamb. However, it struggled with breakout rookies like Brian Thomas Jr. and Malik Nabers, whose roles were not yet established during the offseason.

- A notable pattern in Figure 7 is that both the model and expert rankings performed significantly better within the top 12 than for WR13–24. This decline in accuracy likely reflects the fact that non-elite receivers are more dependent on situational factors—such as game script, quarterback consistency, and matchup variability—that are harder to account for using preseason projections.

- The comparison highlights the distinct advantages of both statistical models and expert-based approaches. While models deliver consistency and data-driven objectivity, expert rankings incorporate contextual insights and nuanced judgment. Although a hybrid strategy was not used in this study, the results suggest that combining both methods could enhance forecasting accuracy in future applications.

7 Discussion

Interpretation and Broader Implications

The findings of this study reinforce several key points for both researchers and fantasy football managers:

- **Statistical models are powerful for capturing broad trends and ranking tiers of players**, providing transparency, reproducibility, and a systematic approach to projection.
- **Expert judgment remains essential for edge cases and nuanced evaluations**, as experts can incorporate preseason reports and contextual information that models cannot.
- **Outlier handling through robust regression adds stability**, but cannot fully compensate for the unpredictability of injuries and breakout performances.
- **The integration of analytics and domain knowledge is optimal**: The most reliable fantasy football strategies will combine model-driven projections with expert insights and up-to-date news.
- **Fantasy managers and platforms can use these models** to support draft strategy, waiver wire prioritization, and tier-based rankings, particularly in tools that update dynamically as new data becomes available.

Limitations

- The model's feature set is limited to variables that are quantifiable and available prior to the season, omitting injury risk, and qualitative context.
- The analysis is restricted to wide receivers and recent seasons; generalizability to other positions or timeframes remains untested.
- Some unexplained variance is inevitable in a domain characterized by high variance and frequent "black swan" events.
- By retaining all players regardless of injury or volatility, the model reflects realistic conditions but may absorb additional noise.

Future Research

To further advance the accuracy and applicability of statistical models in this space, future research should consider:

- Incorporating additional features such as injury history, advanced efficiency metrics, coaching changes, and preseason depth chart movement.
- Exploring machine learning models capable of capturing nonlinear relationships and complex interactions.
- Developing automated pipelines for real-time data updates and expanding coverage to other positions.
- Investigating hybrid approaches that systematically blend statistical projections with expert or crowd-sourced insights.

In sum, while this work demonstrates that statistical modeling provides a strong, objective foundation for fantasy football projections, its effectiveness may be enhanced when paired with the contextual awareness and adaptability of expert judgment. Until statistical models are able to meaningfully account for contextual and qualitative factors, expert judgment will remain a critical component of accurate fantasy forecasting.

8 Conclusion

This study demonstrates both the potential and the limitations of statistical modeling for predicting NFL wide receiver fantasy performance. By leveraging key quantitative features—target share, yards per route run, competition change, quarterback skill, and age—ordinary least squares (OLS) and robust regression models were constructed and validated. The models captured a substantial portion of the explainable variance in wide receiver outcomes, with target share and YPRR emerging as the most influential predictors.

Despite careful model construction and diagnostics, the statistical models did not surpass expert consensus rankings in predictive accuracy. Experts achieved a lower mean absolute rank error (MARE) than the models, underscoring the enduring value of qualitative insights and contextual knowledge—such as awareness of injuries and team dynamics—that are not easily quantified or captured in automated models.

Outlier analysis and robust regression provided additional stability and transparency, confirming that unpredictable events and player-specific circumstances remain a challenge for purely data-driven approaches. While the models are systematic, reproducible, and grounded in established analytics, their performance is constrained by reliance on prior-season metrics and preseason indicators.

In summary, statistical modeling offers a strong, objective foundation for fantasy football projections and can approach—but not yet match—the accuracy of seasoned experts. Progress will depend on how well future models can capture those qualitative elements experts currently intuit. Future research should explore the incorporation of additional features, alternative modeling techniques, and prior-season micro-trends—such as increasing usage rates, target volume, or snap share trajectories—that may help identify breakout candidates, especially among rookies and emerging players.

Acknowledgments

The author has no acknowledgments to report.

References

Bleacher Report. (2018). *Help or hindrance? the accuracy of fantasy football projections.* Retrieved from `https://bleacherreport.com/articles/68563-help-or-hindrance-the-accuracy-of-fantasy-football-projections` (Accessed 2025-07-08)

Draft Sharks. (2023). *The danger of relying too heavily on projections.* Retrieved from `https://www.draftsharks.com/article/the-danger-of-relying-too-heavily-on-projections` (Accessed 2025-07-08)

FantasyPros. (2023). *Nfl wide receiver snap counts 2023.* Retrieved from `https://www.fantasypros.com/nfl/reports/snap-counts/wr.php?year=2023&show=perc` (Accessed 2025)

FantasyPros. (2024). *Nfl wide receiver snap counts 2024.* Retrieved from `https://www.fantasypros.com/nfl/reports/snap-counts/wr.php?year=2024&show=perc` (Accessed 2025)

Kapania, N. (2012). *Predicting fantasy football performance with machine learning techniques* (CS229 Project Report). Stanford University. Retrieved from `https://cs229.stanford.edu/proj2012/Kapania-FantasyFootballAndMachineLearning.pdf`

The Fantasy Footballers. (2025). *Dynasty lifecycles: Players with age cliff concerns for 2025.* Retrieved from `https://www.thefantasyfootballers.com/dynasty/dynasty-lifecycles-players-with-age-cliff-concerns-for-2025/` (Accessed July 12, 2025)

Yahoo Sports. (2024). *Wide receiver rankings, 2024.* Retrieved from `https://sports.yahoo.com/wide-receiver-rankings-fantasy-football-2024-144656627.html` (Accessed 2025)

Yardbarker. (2024). *Understanding analytics: Wr metrics in fantasy football.* Retrieved from https://www.yardbarker.com/nfl/articles/understanding_analytics_wr_metrics_in_fantasy_football/s1_17198_40650875 (Accessed 2025-07-08)

A Appendix

A.1 Competition_Change Variable Calculation

The `competition_change` dummy variable was calculated based on changes in a player's expected opportunity for targets/touches between the 2022 and 2023 seasons. The assignment rules are as follows:

Value	Definition
1	Competition decreased: Player is expected to face less competition for targets/touches. Assigned if a player moves up the depth chart (e.g., WR3 to WR2).
0	No significant change: Player's competition for targets/touches is largely unchanged; depth chart role remains neutral.
-1	Competition increased: Player is expected to face more competition for targets/touches. Assigned if a player moves down the depth chart (e.g., WR2 to WR3).

Table 9: Calculation and coding of the `competition_change` variable.

Sample Calculation Table

Player Name	Team	Competition_Change
DJ Moore	CHI	1
Cooper Kupp	LAR	0
Davante Adams	LV	0
Mack Hollins	ATL	1
Michael Pittman Jr.	IND	0
Donovan Peoples-Jones	CLE	-1
Christian Kirk	JAX	-1

Table 10: Examples of `competition_change` assignments with reasoning.

Reasoning Examples:

- **DJ Moore (CHI, 1):** Moved from WR1 on CAR to WR1 on CHI, entering a less crowded depth chart and likely increasing opportunity.

- **Mack Hollins (ATL, 1):** Was WR3/WR4 on LV, now projected as WR2 for ATL, moving up the depth chart.
- **Donovan Peoples-Jones (CLE, -1):** Was WR2 for CLE, but with the addition of Elijah Moore and Amari Cooper still present, moves down to WR3.
- **Christian Kirk (JAX, -1):** Moves from clear WR1 to WR2 with the addition of Calvin Ridley.