

FANTASY FOOTBALL PREDICTIVE MODELING

by

ELIAS COSICO

A THESIS

Presented to the Department of Data Science
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2026

An Abstract of the Thesis of

Elias Cosico for the degree of Bachelor of Science
in the Department of Data Science to be taken June 2026

Title: Fantasy Football Predictive Modeling

Approved: Peter Ralph, Department Head and Associate Professor
Primary Thesis Advisor

To cut through the statistical noise of National Football League player evaluation, this thesis applies machine learning to fantasy football scoring. The study analyzes a dataset of 2,963 player-season observations from 2018 to 2025, looking at four key positions: quarterback (QB), running back (RB), wide receiver (WR), and tight end (TE). By implementing a custom snap-percentage filter during preprocessing, the data isolates viable games and removes injury distortions. The project utilizes three modeling tracks: Random Forest regression for points-per-game projections, Random Forest classification for tier assignments, and K-Means clustering for archetype discovery. Findings show volume metrics, alongside total touchdowns, emerged as the most powerful predictors of fantasy output. When compared to real-world 2025 data, the model hit a 0.581 Pearson correlation with actual scores. This handily beat industry Average Draft Position (ADP), which held an inverse correlation of -0.325. Overall, the model out forecasted public expert consensus for 52.3% of evaluated players. The K-Means archetypes successfully isolated high-value roles, proving that "Workhorse" running backs (17.35 PPG) and "Alpha" wide receivers (16.03 PPG) yielded the highest average production. These findings deliver a scalable quantitative tool for sports wagering and predictive sports analytics.

Acknowledgements

Keeping a data-heavy project on track requires great mentorship, and Professor Peter Ralph provided exactly that through his statistical modeling expertise and consistent guidance. For the structural support and critical feedback throughout the writing process, Professor Marcel Brousseau and the Robert D. Clark Honors College deserve immense credit. Finally, this thesis relies directly on the open-source community behind the nflreadpy data library. Without their work, pulling together and cleaning a dataset of this scale would have been an impossible task.

Table of Contents

Chapter 1: Introduction and Background	8
The NFL as a Data-Rich Environment	8
Fantasy Football as an Analytical Framework	8
Research Questions and Hypotheses	9
Significance and Scope	9
Chapter 2: Methodology	11
Phase 1: Data Acquisition	11
Phase 2: Preprocessing and Exploratory Data Analysis	11
Snap-Based Healthy Game Filtering	11
Feature Engineering	12
Lagged Features and Train-Test Split	12
Exploratory Data Analysis	13
Phase 3: Model Selection and Training	14
Random Forest Regression	14
Feature Importance	15
Linear and Polynomial Regression Baseline	16
Random Forest Classification	17
K-Means Clustering	18
Phase 4: Optimization and Evaluation	19
Chapter 3: Results	21
RQ1: Feature Significance and Model Performance	21
Correlation Analysis	21
Random Forest Regression Results	21
Overfitting Assessment	23
Regression Baseline Comparison	26
RQ2: External Variables and Team Offensive Quality	27
RQ3: Player Archetypes and Draft Value	28
Model vs. ADP: Validation Against Actual 2025 Outcomes	30
Chapter 4: Discussion	34
Interpretation of RQ1: Feature Significance	34
Interpretation of RQ2: Team Context	34
Interpretation of RQ3: Archetypes and Draft Value	35

Model Performance and ADP Comparison	35
Limitations	36
Comparison to Prior Work	37
Chapter 5: Conclusion	38
Summary of Findings	38
Practical Applications	38
Future Directions	39
Bibliography	40
Supporting Materials	
Code: Predictive Modeling	

List of Figures

Figure 1: EDA Correlation Matrix Heatmap by Position	14
Figure 2: Random Forest Feature Importance	16
Figure 3: Regression Baseline – Actual vs. Prediction PPG (All Positions)	17
Figure 4: Tier Classification Confusion Matrices (All Positions)	18
Figure 5: K-Means Player Archetype by Position	19
Figure 6: Model Error by Position (MAE and RMSE Bar Chart)	22
Figure 7: Random Forest Regressor — Predicted vs. Observed PPG (2025 Holdout)	23
Figure 8: Overfitting Check — Training MAE vs. Test MAE by Position	25
Figure 9: RQ2: Player PPG by Team Offensive Tier (Box Plots)	28
Figure 10: RQ3: Average PPG by Archetype Bar Chart	30
Figure 11: Model vs. ADP Accuracy Comparison Bar Chart	31
Figure 12: 2025 Performance Analysis	33

List of Tables

Table 1. Random Forest Regressor Performance by Position (2025 Holdout Test Set)	21
Table 2. Overfitting Assessment	24
Table 3. Statistical Significance	26
Table 4. Regression Baseline Comparison by Position	26
Table 5. Pearson Correlation Between Team Offensive Rank and Player PPG	27
Table 6. Average Fantasy PPG by Player Archetype	29
Table 7. Top 10 Model Value Discoveries Relative to ADP (2025 Season)	32

Chapter 1: Introduction and Background

The NFL as a Data-Rich Environment

The National Football League (NFL) is one of the most data-heavy environments in all professional sports. As the game of football has progressed, player-tracking technology and open databases have completely reshaped how coaches, analysts, and fans view the game. However, a gap has emerged between having data and using it effectively. The sheer volume of numbers frequently overwhelms casual participants, making it difficult for them to pull out truly objective insights. Because of this, mainstream evaluation still relies heavily on the traditional "eye test" and media narratives—both of which get easily warped by recency bias and name recognition.

To fix this disconnect, this thesis aligns machine learning with sports analytics to create a direct predictive pipeline. By feeding historical data into the models to predict the 2025 season, the project establishes a baseline for direct comparison against actual, real-world 2025 outcomes. The analysis focuses entirely on the offensive skill positions tied to fantasy scoring: quarterbacks (QB), running backs (RB), wide receivers (WR), and tight ends (TE).

Fantasy Football as an Analytical Framework

By standardizing raw on-field production, fantasy football serves as an ideal testing environment for player evaluation. In a standard league, participants draft real NFL players and accumulate points based purely on weekly box-score statistics. This structural setup completely strips away subjective team bias. It transforms chaotic, qualitative gameplay into a clean, quantifiable dataset perfectly tailored for predictive modeling. The Points Per Reception (PPR) scoring format, which awards one point per reception in addition to yardage and touchdown

bonuses, has become the dominant format in competitive play and is used as the scoring standard throughout this analysis.

The fantasy football industry is substantial. According to Statista (2025), over 65 million participants engage in fantasy sports annually in the United States, with fantasy football representing the largest single category. The PPR scoring framework translates disparate player actions into a single, objective numeric value, effectively standardizing athletic performance in a way that removes much of the subjectivity inherent in traditional evaluation. This standardization makes fantasy points per game (PPG) an ideal dependent variable for a predictive modeling study.

Research Questions and Hypotheses

This study is organized around three primary research questions. Research Question 1 asks which offensive metrics hold the highest statistical significance for season-long fantasy point totals, with the hypothesis that opportunity-based metrics such as targets and carries will demonstrate the highest predictive significance. Research Question 2 examines how team offensive quality impacts individual player performance, with the hypothesis that team offensive rankings will have a meaningful influence on a player's fantasy ceiling. Research Question 3 asks which player archetype is the most valuable to target in fantasy drafts, with the hypothesis that players characterized by high efficiency and ascending volume will represent the greatest value.

Significance and Scope

The implications of this research extend to two primary audiences. For the fantasy football community, the model provides a statistically grounded alternative to consensus-based rankings and Average Draft Position (ADP). For the sports wagering industry, accurate player performance prediction has direct financial relevance for the pricing of player proposition bets

and season-long futures. As Lee and Liu (2022) demonstrate, fantasy managers routinely employ non-optimal drafting strategies driven by emotional and cognitive biases; a predictive model of this kind offers a systematic corrective.

The scope of the study is limited to the 2018 through 2025 NFL seasons. This window captures the modern era of professional football, characterized by the proliferation of pass-heavy offenses and elevated target volumes, while providing sufficient data volume for model training and validation. The dataset is split such that the 2018 through 2024 seasons are used for model training and the 2025 season is reserved as a holdout test set.

Chapter 2: Methodology

Phase 1: Data Acquisition

To build the dataset, this project uses player statistics directly through the `nflreadpy` Python library. This tool provides access to `nflverse`, an open-source repository for historical football data. The pipeline extracts weekly box-score statistics and snap counts covering eight consecutive seasons, from 2018 through 2025.

Evaluating the final predictive models required a clean industry benchmark. To establish this, a web scraper gathered 2025 Average Draft Position (ADP) data from Fantasy Football Calculator (fantasyfootballcalculator.com). Because this platform aggregates mock draft results from thousands of community participants, it offers a highly accurate consensus baseline of public market expectations. These ADP figures serve as the direct point of comparison against the model's final positional rankings during validation.

Phase 2: Preprocessing and Exploratory Data Analysis

Snap-Based Healthy Game Filtering

Injuries introduce a structural challenge in sports modeling because they drastically warp season-long datasets. For example, a player might be performing very well, but suffer a season-ending injury. This would result in their season totals being skewed, as they were not able to finish out the season. To clear out this distortion, the preprocessing pipeline implements a custom snap-percentage filter, isolating only the games where a player was a healthy, active participant in the offense.

The baseline thresholds vary by position to accurately reflect real-world playing time:

- Quarterbacks: 90% of offensive snaps

- Wide Receivers: 70% of offensive snaps
- Tight Ends: 65% of offensive snaps
- Running Backs: 35% of offensive snaps

Dropping games that fall below these specific markers leaves a refined subset of viable, high-participation performances. From there, the framework evaluates all player outputs strictly on a per-game basis.

Feature Engineering

After filtering, weekly data was aggregated to the player-season level. The primary target variable, fantasy points per game (PPG), was computed by dividing each player's total PPR fantasy points by their number of qualifying games. Two key efficiency metrics were engineered at the team level: target share, defined as a player's targets divided by total team targets, and air yards share, defined as a player's receiving air yards divided by total team air yards. A team offensive rank variable was computed by ranking teams within each season by total fantasy points generated. The final preprocessed dataset consists of 2,963 player-season observations across 24 variables, spanning 980 unique players.

Lagged Features and Train-Test Split

To enable genuine out-of-sample prediction, all independent variables were transformed into lagged features representing the prior season's statistics. This design ensures that the model is trained to predict a player's upcoming season performance using only information available at the time of a fantasy draft, rather than data collected throughout the season. Seasons 2018 through 2024 comprised the training set and the 2025 season served as the test set.

Exploratory Data Analysis

Correlation analysis was conducted separately for each position to identify which variables demonstrated the strongest linear relationship with PPG. Full correlation matrices were visualized as heatmaps for each position (Figure 1). The top correlates with PPG for each position are summarized in Figure 1 and discussed in the Results section.

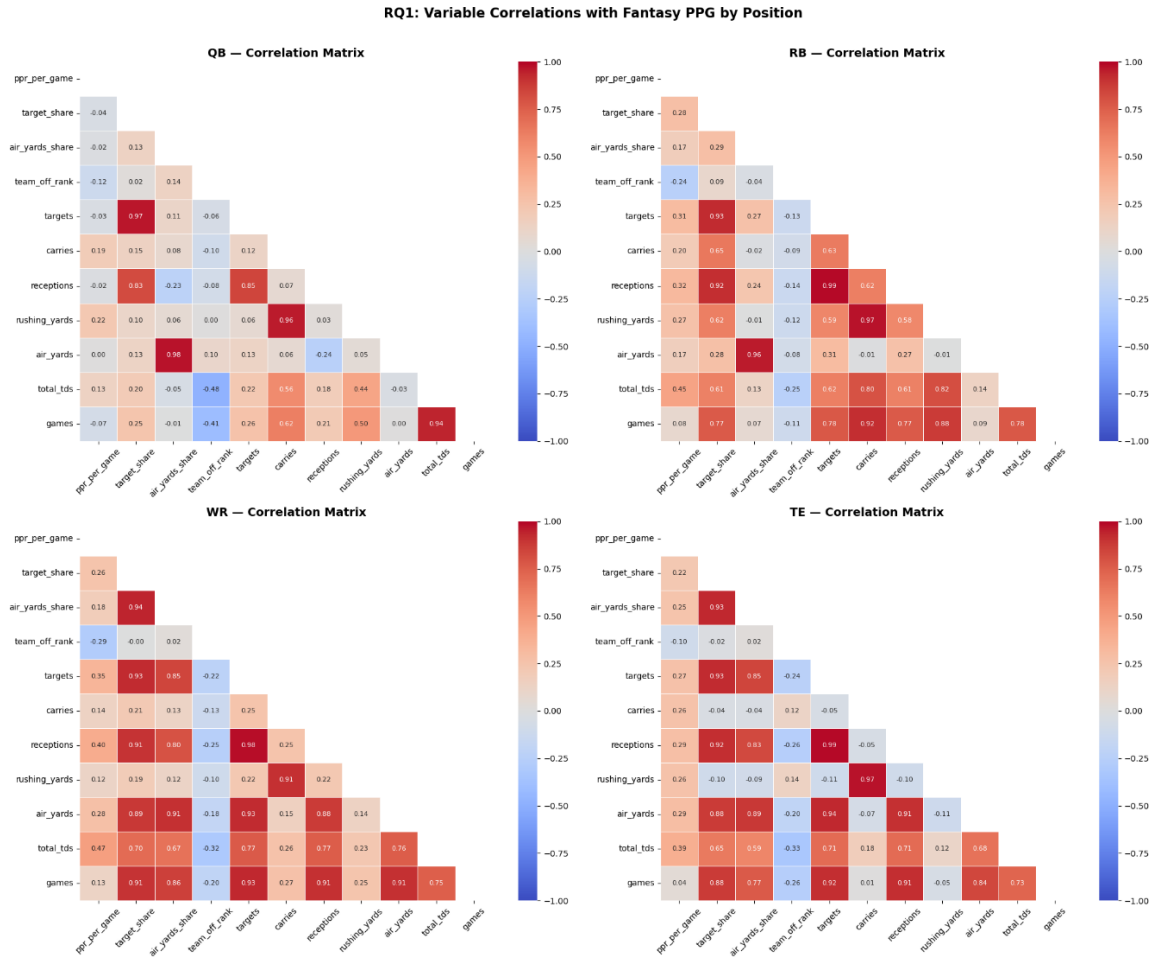


Figure 1: EDA Correlation Matrix Heatmap by Position

Figure 1. Correlation matrices showing Pearson correlation coefficients between all variables and fantasy PPG for each offensive position (QB, RB, WR, TE). Warm colors indicate positive correlation; cool colors indicate negative correlation.

Phase 3: Model Selection and Training

Random Forest Regression

To handle the continuous points-per-game forecasting, this study relies primarily on a Random Forest Regressor. Instead of trusting a single path, this ensemble learning method builds a large forest of independent decision trees during training, ultimately averaging their individual outputs to generate a final prediction. Because it utilizes a bootstrap aggregating (bagging)

framework, the algorithm naturally resists overfitting and excels at uncovering the messy, nonlinear relationships hidden within sports data (Breiman & Cutler, 2019).

The feature selection pipeline isolates the top six lagged variables for each position based strictly on correlation magnitude. To keep weak predictors from diluting the model's focus, a secondary filter strips out any feature with an absolute correlation below 0.15. Once cleaned, the final positional models train using 100 estimators and a maximum depth cap of 10.

Feature Importance

To see exactly which football metrics drive the final projections, the framework extracts internal feature importance scores. These values calculate the precise weight each input variable carries in sharpening the model's overall accuracy. Figure 2 provides a visual map of these relative weights, which the Results section dissects in deeper detail.

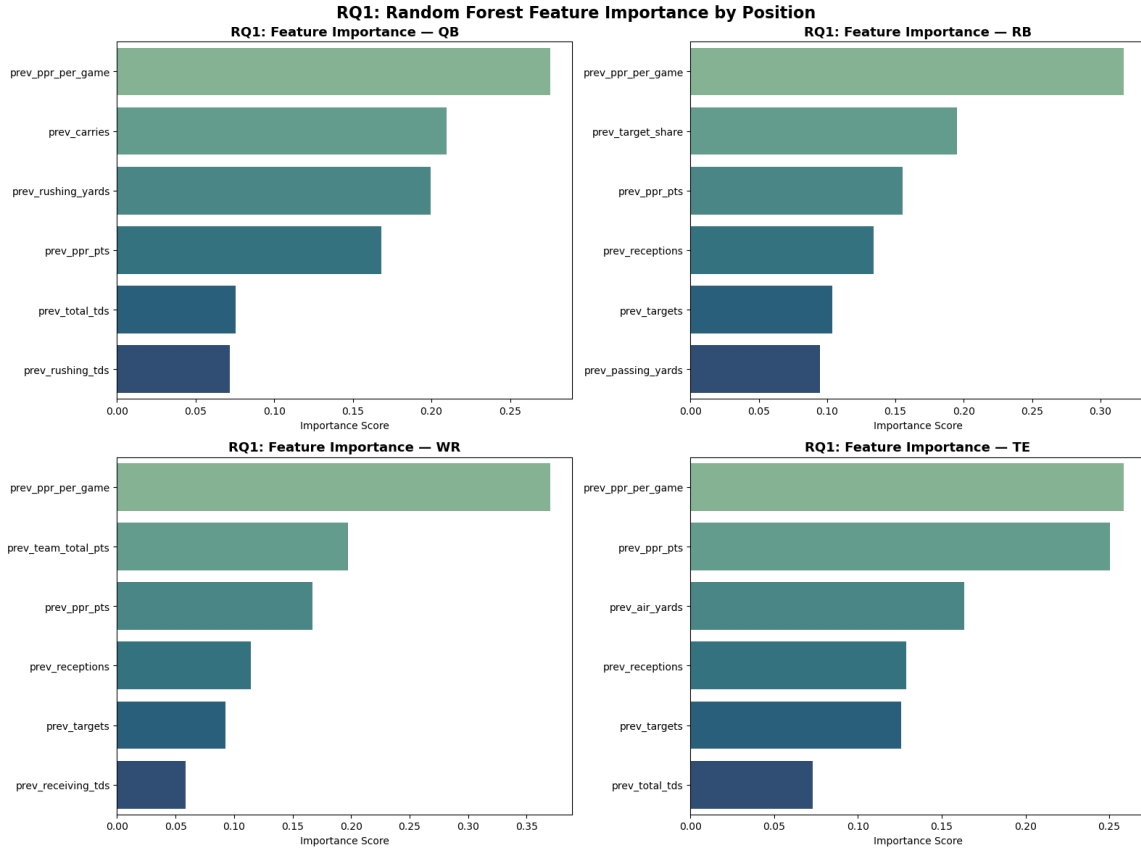


Figure 2: Random Forest Feature Importance

Figure 2. Random Forest feature importance scores by position, indicating the relative predictive weight of each input variable in the trained model.

Linear and Polynomial Regression Baseline

Linear and polynomial (degree 2) regression models were trained for each position as a baseline comparison to the Random Forest approach. These simpler models test whether the relationship between prior-season statistics and PPG is adequately captured by a linear function, or whether the nonlinear flexibility of Random Forest provides meaningful additional explanatory power. Actual versus predicted scatter plots for all four positions are presented in Figure 3.

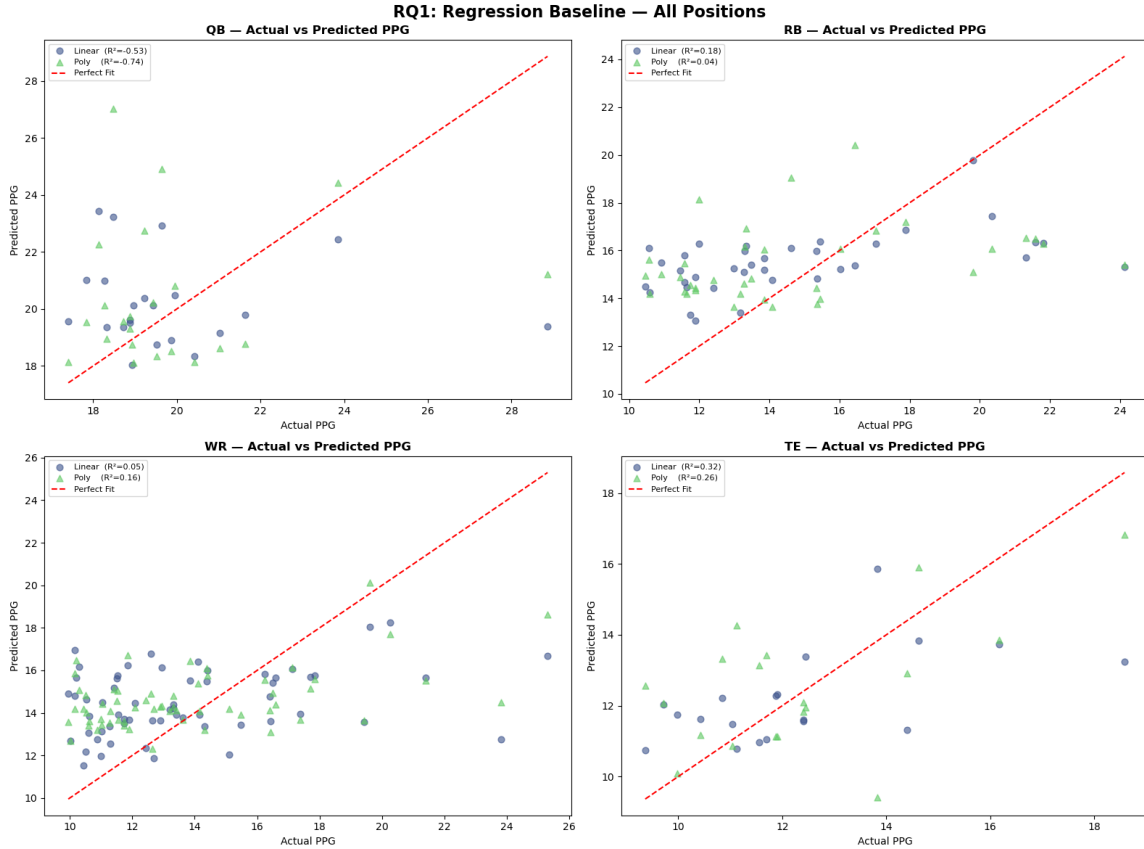


Figure 3: Regression Baseline – Actual vs. Prediction PPG (All Positions)

Figure 3. Actual versus predicted fantasy PPG for linear (blue circles) and polynomial degree-2 (green triangles) regression models across all four positions. The red dashed line represents perfect prediction.

Random Forest Classification

To group players into clean performance tiers, the project runs a separate Random Forest classifier for each position. The logic assigns players to three buckets based on their per-game scoring ranks. Tier 1 (Elite) holds the top 12 finishers at the position. Tier 2 (Starter) covers ranks 13 through 24. Anything below that gets grouped into Tier 3 (Depth). Because the sheer volume of depth players completely dwarfs the elite tier, the code adjusts class weights during training to keep the data imbalance from warping the results. Figuring out exactly where the model hits or misses happens via the confusion matrices in Figure 4.

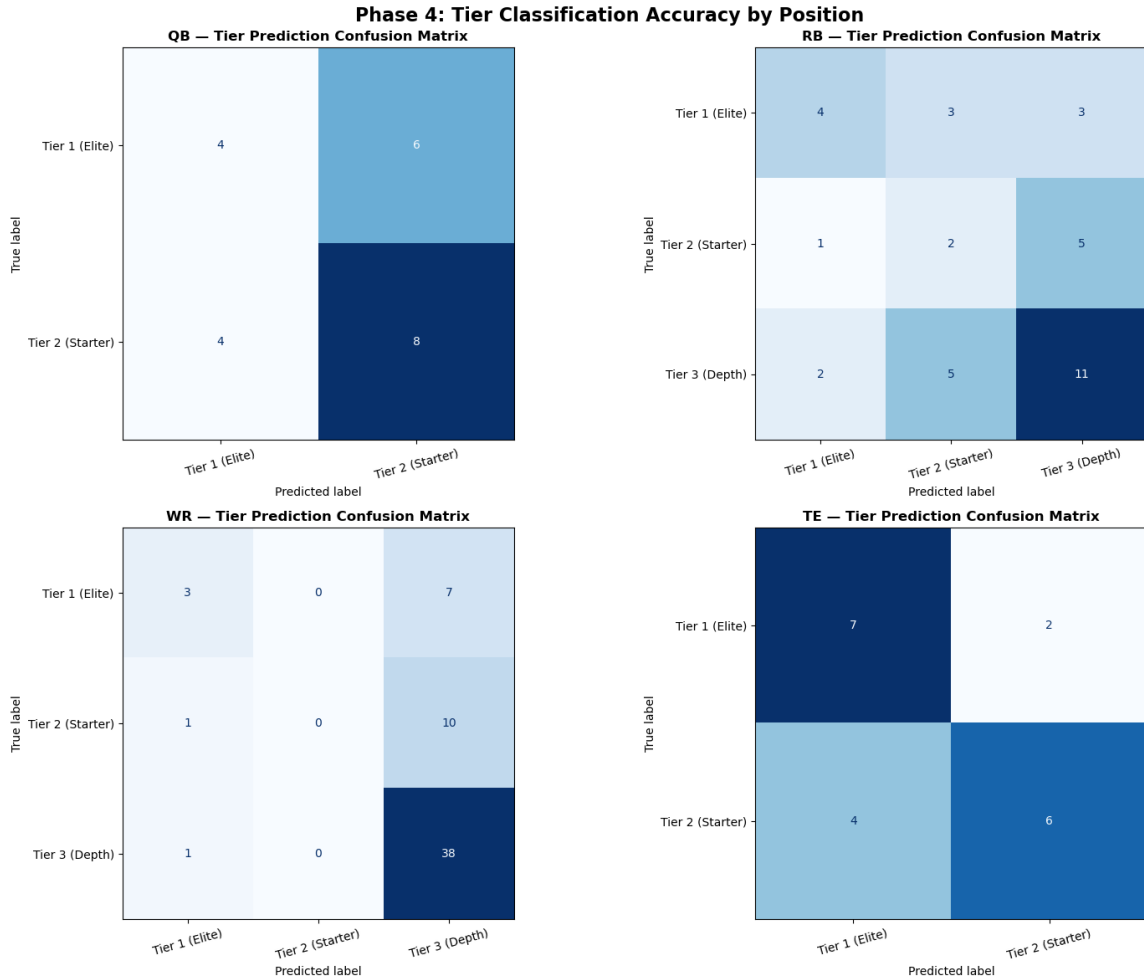


Figure 4: Tier Classification Confusion Matrices (All Positions)

Figure 4. Confusion matrices for the Random Forest tier classification model across all four positions. Rows show actual tiers; columns show predicted tiers. Darker blue cells indicate higher player counts.

K-Means Clustering

To discover underlying player archetypes based on usage profiles, K-Means clustering evaluates each position individually. The inputs change depending on the position's role on the field. For pass-catching positions, the model feeds on target share, air yards share, and prior-season scoring. For running backs, it tracks carries, target share, and prior-season scoring instead. Setting the cluster count to four for wide receivers and running backs, and three for quarterbacks and tight ends, keeps the categories clean and readable. Looking closely at the

resulting centroid profiles allowed for assigning descriptive archetype labels to each group.

Figure 5 maps these clusters visually.

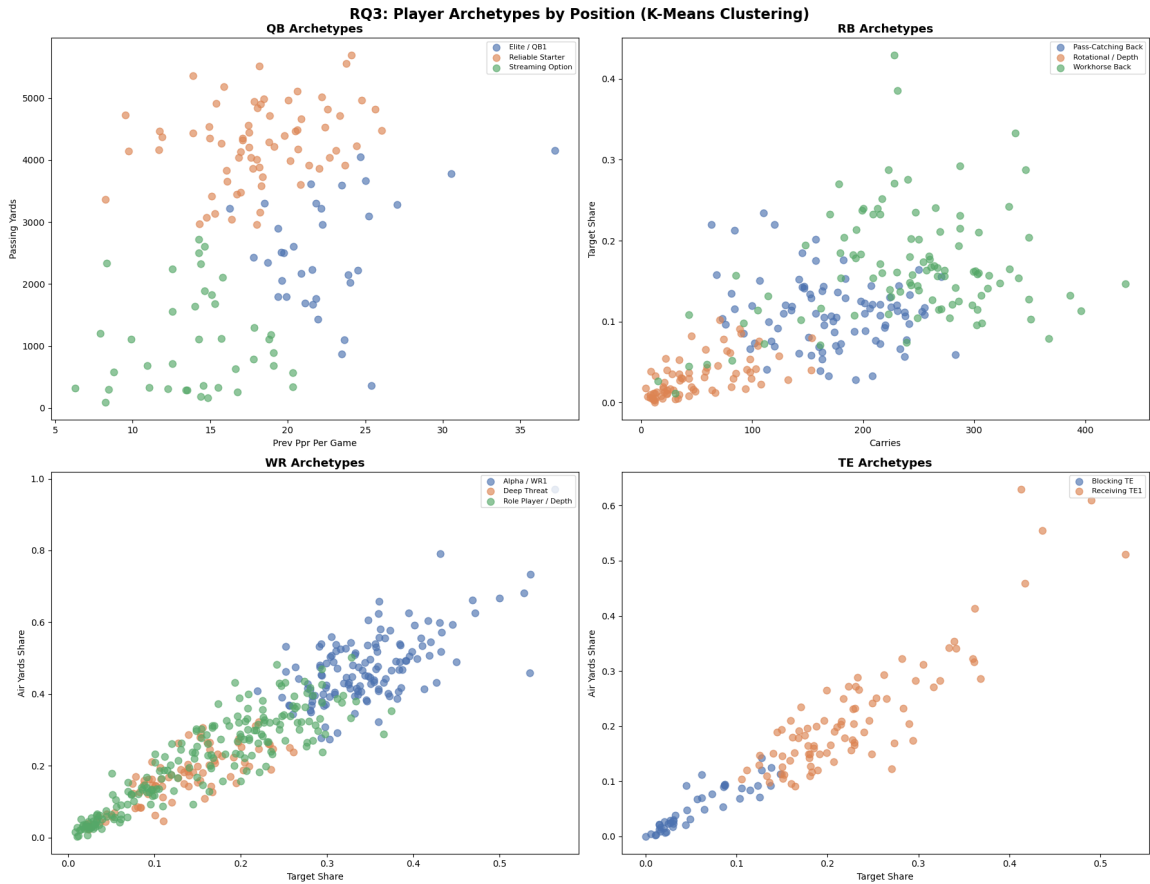


Figure 5: K-Means Player Archetype by Position

Figure 5. K-Means clustering results for each position. Each data point represents a player-season, color-coded by its assigned archetype. The axes reflect the two main clustering features for each position.

Phase 4: Optimization and Evaluation

Fine-tuning the wide receiver Random Forest model meant running a GridSearchCV script backed by three-fold cross-validation. The grid searched through combinations of estimator counts (50, 100, 200), maximum tree depths (unrestricted, 10, 20), and minimum samples per split (2, 5). The best configuration hit on 50 estimators, a maximum depth of 10, and

a minimum split of 5 samples. This specific setup locked in a cross-validated mean absolute error (MAE) of 2.55 fantasy points.

To score the overall strength of the models, the project tracks separate data science metrics for each model type. The regression models judge points-per-game predictions through mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R^2 . Meanwhile, the classification side measures tiering accuracy using precision, recall, F1-score, and macro-averaged F1 marks.

To assess the risk of overfitting, training set MAE was compared against test set MAE for each position. Additionally, Pearson correlation p-values were computed to evaluate the statistical significance of the model's predictive relationship at each position.

Chapter 3: Results

RQ1: Feature Significance and Model Performance

Correlation Analysis

Tracking the data shows that touchdowns and opportunity-based metrics consistently drive fantasy scoring across all positions. For wide receivers, total touchdowns correlate strongest with per-game scoring ($r = 0.473$), with receptions ($r = 0.404$) and targets ($r = 0.350$) following close behind. Running backs ($r = 0.452$) and tight ends ($r = 0.390$) mirror a nearly identical pattern, with total touchdowns leading the way. For quarterbacks, rushing metrics completely dominate the data. Rushing yards ($r = 0.224$) and rushing attempts ($r = 0.192$) stand out as the two strongest correlates, highlighting the massive scoring floor that dual-threat quarterbacks provide in standard PPR leagues. Figure 1 maps out these individual relationships.

Random Forest Regression Results

Table 1 breaks down how the Random Forest Regressor handled the 2025 holdout test set.

Table 1. Random Forest Regressor Performance by Position (2025 Holdout Test Set)

Position	N (Test)	MAE	MSE	RMSE	R^2
TE	19	1.269	3.303	1.817	0.354
QB	22	1.985	7.333	2.708	-0.268
WR	60	2.668	10.775	3.282	0.101
RB	36	2.810	11.542	3.397	0.111

Tight ends yielded the cleanest model fit, hitting an MAE of 1.269 points and an R^2 of 0.354. Wide receivers and running backs posted more modest outcomes, landing R^2 clips of 0.101 and 0.111. The quarterback model went negative on its R^2 . This drop shows that relying strictly on previous-year features is not enough to accurately predict quarterback scoring, mainly because sudden coaching scheme changes and unexpected injury replacements create heavy year-to-year volatility. Figure 6 visualizes these exact error distributions using MAE and RMSE.

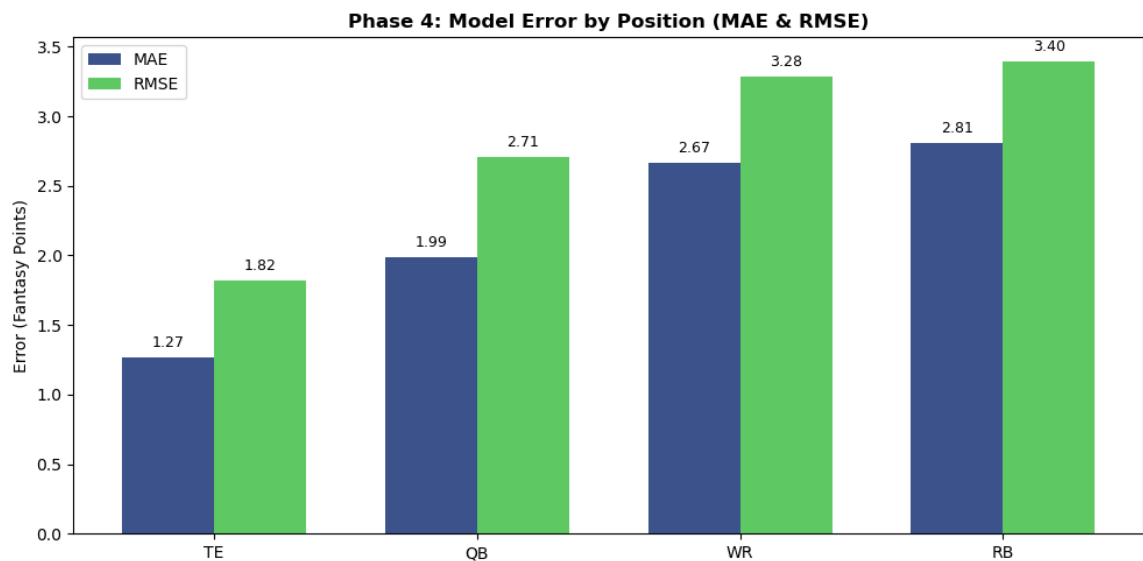


Figure 6: Model Error by Position (MAE and RMSE Bar Chart)

Figure 6. Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) for the Random Forest Regressor across all four positions. Lower values indicate better predictive accuracy.

Figure 7 presents the predicted versus observed PPG scatter plots for the Random Forest Regressor across all four positions on the 2025 holdout set. The tight end model shows the strongest alignment between predicted and observed values, consistent with its R^2 of 0.354. The quarterback scatter plot reveals the model's core limitation — predicted values cluster tightly between 19 and 23 PPG regardless of actual outcome, meaning the model collapses toward the positional mean rather than capturing true quarterback variance. The wide receiver and running back plots show a modest positive trend, consistent with their R^2 values of 0.101 and 0.111

respectively. Comparing Figure 7 to the regression baseline in Figure 3 illustrates the improvement in predictive fit achieved by the nonlinear Random Forest approach, particularly for tight ends where the Random Forest ($R^2 = 0.354$) substantially outperforms both the linear ($R^2 = 0.320$) and polynomial ($R^2 = 0.263$) baselines.

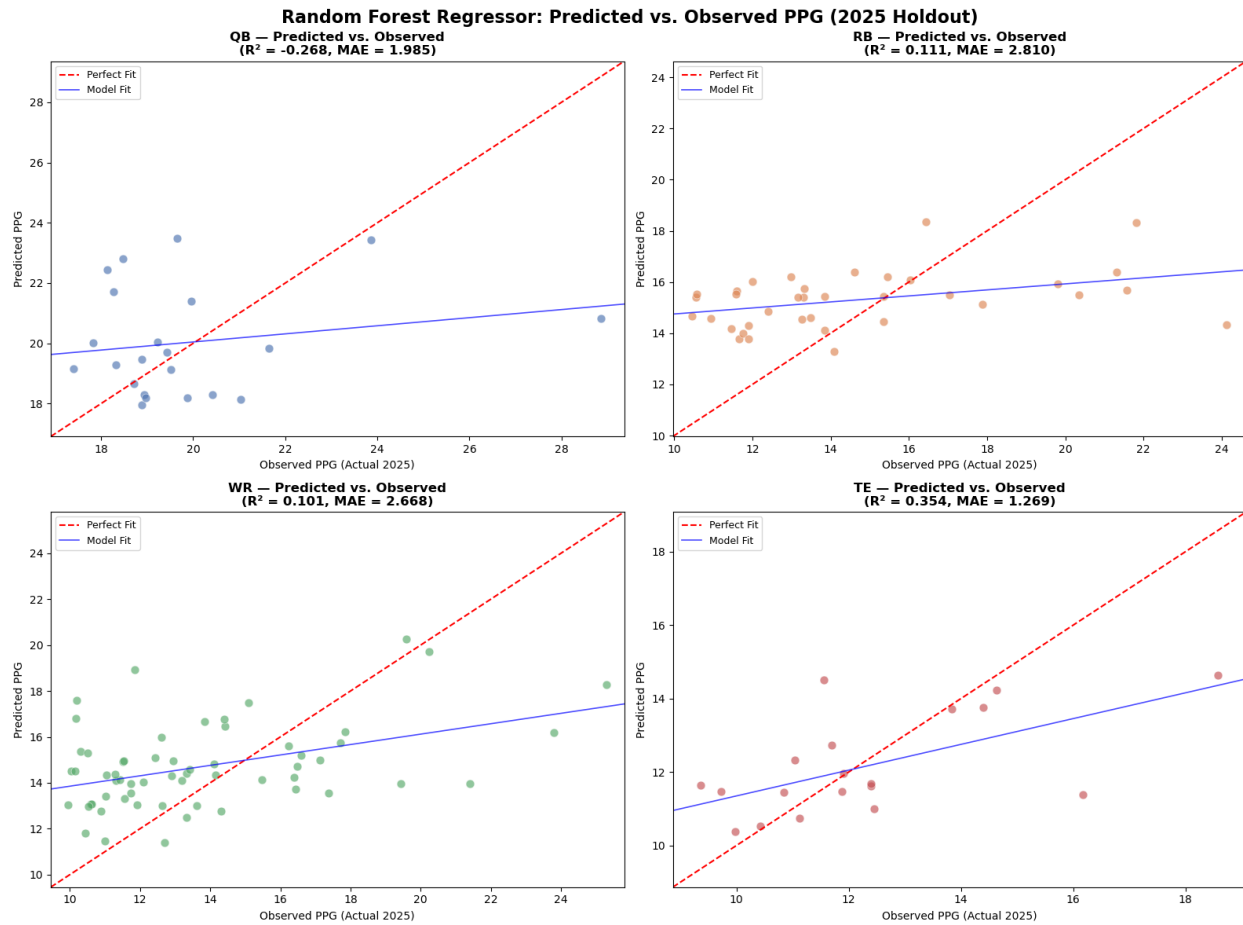


Figure 7: Random Forest Regressor — Predicted vs. Observed PPG (2025 Holdout)

Figure 7. Random Forest Regressor predicted versus observed fantasy PPG for all four positions on the 2025 holdout test set. The red dashed line represents perfect prediction; the blue line represents the model's actual fit. Points clustering near the red line indicate accurate predictions.

Overfitting Assessment

To evaluate whether the model learned genuine patterns or simply memorized training data, training set MAE was compared against test set MAE for each position. Table 6 presents

these results. The tight end model shows the smallest gap between training and test MAE (0.354 points), indicating low overfitting risk and strong generalizability. The running back model shows the largest gap (1.643 points), classified as high overfitting risk, suggesting the model learned training patterns that did not fully carry over to 2025 outcomes. The quarterback and wide receiver models show moderate gaps of 1.075 and 1.410 respectively.

Table 2. Overfitting Assessment

Position	Train MAE	Test MAE	Gap (Test – Train)	Overfit Risk
QB	0.910	1.985	1.075	Moderate
RB	1.167	2.810	1.643	High
WR	1.258	2.668	1.410	Moderate
TE	0.915	1.269	0.354	Low

Figure 8 visualizes these training versus test error gaps by position.

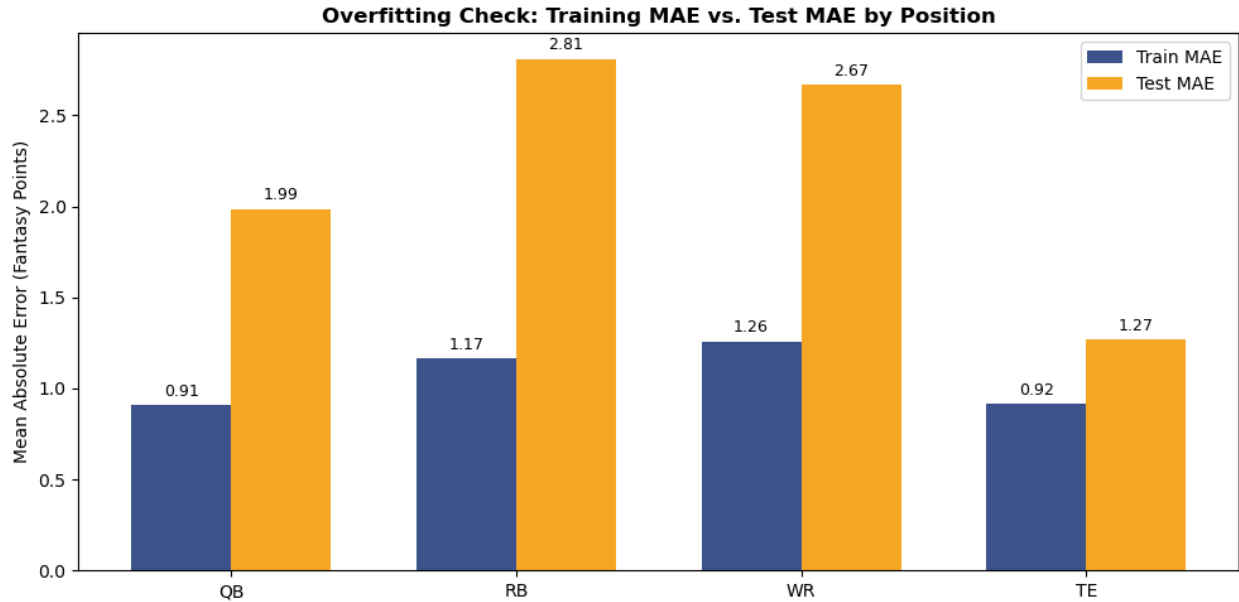


Figure 8: Overfitting Check — Training MAE vs. Test MAE by Position

Figure 8. Training MAE versus test MAE by position. A larger gap between the two bars indicates higher overfitting risk. The tight end model shows the lowest risk; the running back model shows the highest.

Statistical Significance of Model Predictions

To confirm whether the model's predictive relationships reflect genuine signal rather than chance, Pearson correlation p-values were computed between predicted and actual PPG for each position. Table 7 presents these results. The running back ($p = 0.019$), wide receiver ($p = 0.001$), and tight end ($p = 0.007$) models all achieved statistical significance at the $\alpha = 0.05$ level, confirming that the model captures real predictive signal at those positions. The quarterback model produced a non-significant p-value of 0.411 ($n = 22$), confirming that its predictive relationship cannot be distinguished from noise. This result is consistent with the negative R^2 in Table 1 and formally confirms that lagged statistics alone are insufficient to predict quarterback fantasy performance.

Table 3. Statistical Significance

Position	r	R^2	p-value	n	Significant?
QB	0.185	0.034	0.411	22	No
RB	0.388	0.151	0.019	36	Yes
WR	0.429	0.184	0.001	60	Yes
TE	0.600	0.360	0.007	19	Yes

Regression Baseline Comparison

To verify these results, Table 4 compares the baseline linear and polynomial equations against the primary model.

Table 4. Regression Baseline Comparison by Position

Position	Model	R^2	MSE	MAE
QB	Linear	-0.530	8.847	2.148
QB	Polynomial (deg 2)	-0.743	10.081	2.248
RB	Linear	0.185	10.581	2.628
RB	Polynomial (deg 2)	0.041	12.442	2.945
WR	Linear	0.052	11.356	2.639
WR	Polynomial (deg 2)	0.157	10.101	2.643
TE	Linear	0.320	3.478	1.428
TE	Polynomial (deg 2)	0.263	3.768	1.561

Looking at the R^2 marks, the Random Forest model comfortably beat the linear baselines for quarterbacks, wide receivers, and tight ends. This confirms that capturing non-linear relationships gives the system a massive predictive advantage at those positions. Running backs proved to be the lone exception, where the basic linear model posted a slightly higher R^2 (0.185) than the Random Forest (0.111). Figure 3 outlines the actual versus predicted scores via scatter plots.

RQ2: External Variables and Team Offensive Quality

Testing how much a team's overall offense impacts individual player scoring required looking at the engineered team offensive rank variable. Table 5 breaks down these Pearson correlation marks by position.

Table 5. Pearson Correlation Between Team Offensive Rank and Player PPG

Position	Correlation (Rank vs. PPG)	Interpretation
QB	-0.124	Weak Negative
RB	-0.244	Moderate Negative
WR	-0.289	Moderate Negative
TE	-0.101	Weak Negative

Negative correlations showed up across the entire board, which matches real-world expectations: playing on a highly ranked offense naturally boosts an individual's fantasy production. The impact hits hardest on wide receivers ($r = -0.289$) and running backs ($r = -0.244$). In contrast, quarterbacks ($r = -0.124$) and tight ends ($r = -0.101$) proved far less sensitive to overall team ranks. Figure 9 breaks down these distributions using box plots.

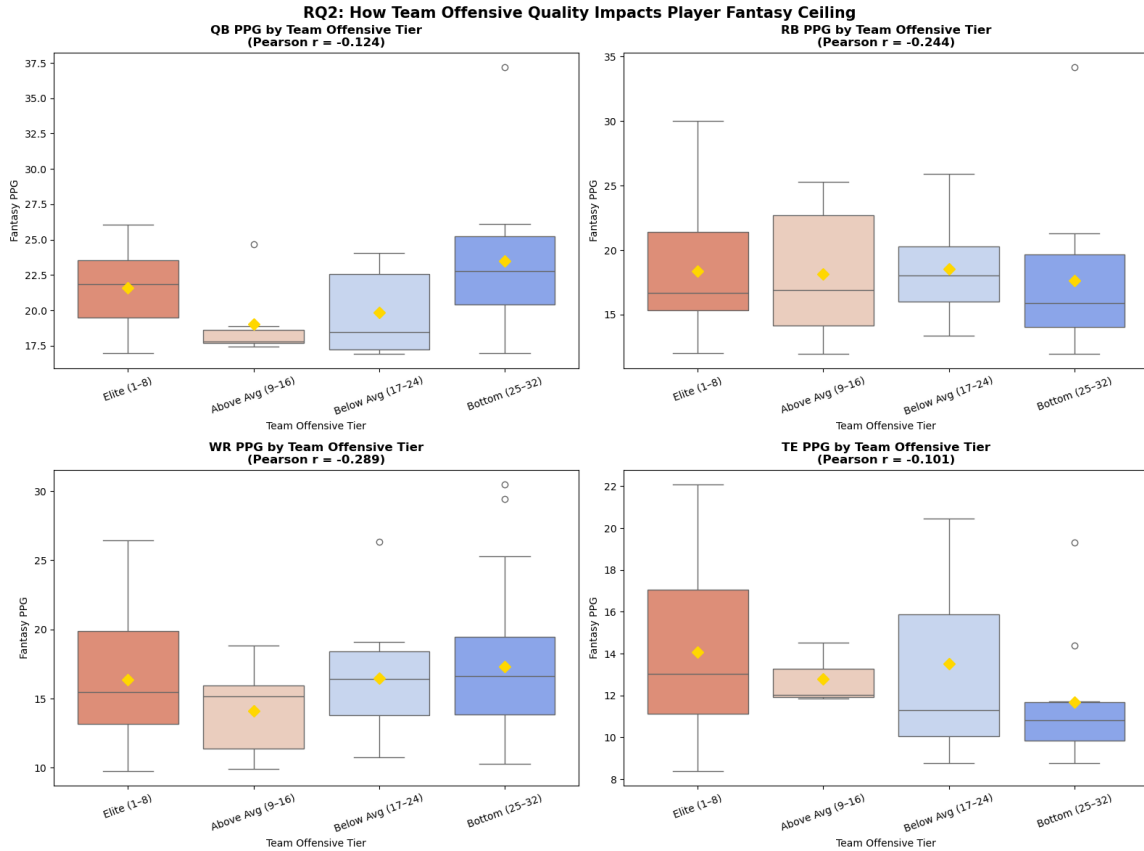


Figure 9: RQ2: Player PPG by Team Offensive Tier (Box Plots)

Figure 9. Distribution of fantasy PPG by team offensive tier for each position. Gold diamonds represent tier means. A visible downward gradient from elite to bottom-tier offenses confirms the RQ2 hypothesis.

RQ3: Player Archetypes and Draft Value

To sort out specific usage roles, the K-Means algorithm grouped each position into distinct archetypes. Table 6 lists the average scoring for these roles to address the core hypothesis.

Table 6. Average Fantasy PPG by Player Archetype

Position	Archetype	Avg. PPG
QB	Elite / QB1	20.94
QB	Reliable Starter	19.86
QB	Streaming Option	19.84
RB	Workhorse Back	17.35
RB	Rotational / Depth	14.28
RB	Pass-Catching Back	14.09
WR	Alpha / WR 1	16.03
WR	Deep Threat	14.67
WR	Role Player / Depth	13.63
TE	Receiving TE	12.46
TE	Blocking TE	11.89

For running backs, "Workhorse" options led the pack at 17.35 PPG, offering a massive 21.5% bump over purely pass-catching backs (14.09 PPG). For wide receivers, "Alpha / WR1" targets paced the group at 16.03 PPG, validating the immense value of heavy target consolidation. Meanwhile, the incredibly tight gap between elite and streaming quarterbacks (20.94 versus 19.84 PPG) mathematically confirms a classic fantasy football rule of thumb: quarterback is easily the deepest, least scarce position on the board. Figure 10 breaks down these comparisons visually.

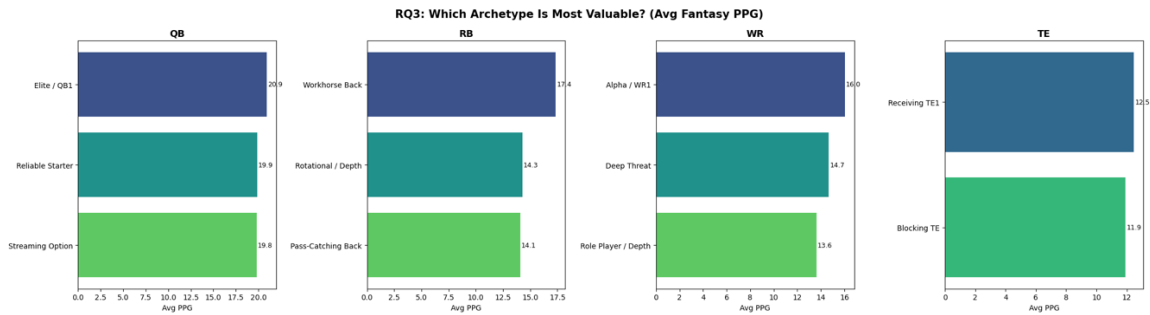


Figure 10: RQ3: Average PPG by Archetype Bar Chart

Figure 10. Average fantasy PPG by player archetype for each position. Archetypes are sorted from highest to lowest average output. Results directly address RQ3 and the ascending-volume hypothesis.

Model vs. ADP: Validation Against Actual 2025 Outcomes

Evaluating the model's ultimate accuracy meant running the final projections directly against real-world 2025 data. The model scored a 0.581 Pearson correlation with actual outcomes. For context, industry Average Draft Position (ADP) registered a completely backward correlation of -0.325, proving the model aligns far better with actual season results than the public consensus. The negative sign of the ADP correlation reflects the inverse scaling of draft position, where lower ADP values indicate higher-valued players; the model's correlation of 0.581 therefore represents a substantially stronger absolute alignment with actual outcomes than ADP's 0.325. Looking at positional rank error, the framework out-forecasted standard market expectations for 52.3% of the evaluated player pool. Figure 11 matches the model's positional error directly against ADP.

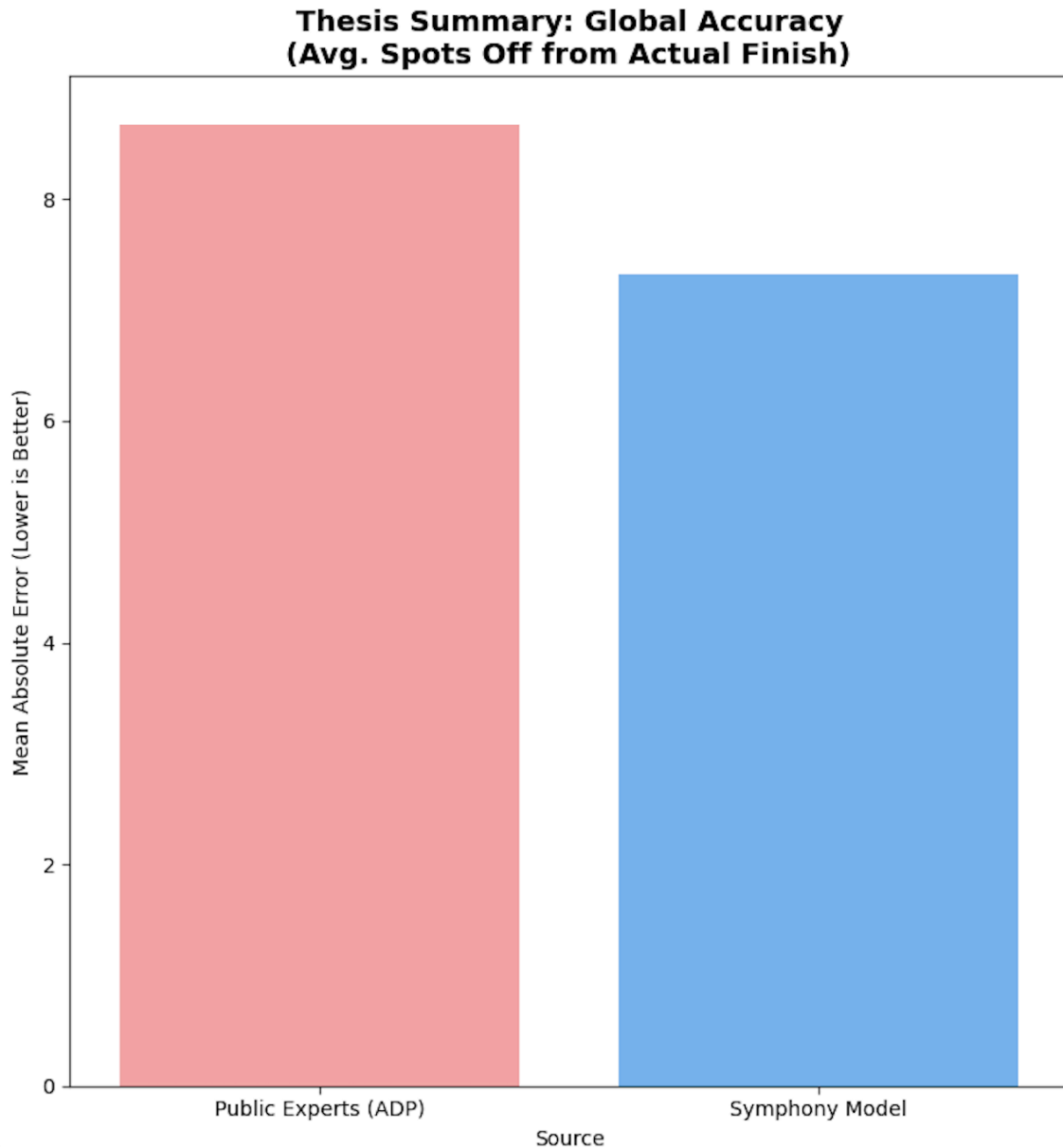


Figure 11: Model vs. ADP Accuracy Comparison Bar Chart

Figure 11. Comparison of average positional rank error between the predictive model and public ADP consensus. Lower values indicate greater accuracy relative to actual 2025 season outcomes.

Table 7 highlights the top ten value discrepancies where the model's internal rankings were significantly higher than public ADP market prices.

Table 7. Top 10 Model Value Discoveries Relative to ADP (2025 Season)

Player	Pos.	Model Rank	ADP Rank	Delta	Actual Rank
Chris Godwin Jr.	WR	5	47	+42	36
Marquise Brown	WR	15	53	+38	26
Rashee Rice	WR	16	43	+27	8
Javonte Williams	RB	6	33	+27	10
Jordan Addison	WR	23	46	+23	37
Jauan Jennings	WR	17	39	+22	30
Zay Flowers	WR	6	27	+21	15
J.K. Dobbins	RB	11	31	+20	22
Rhamondre Stevenson	RB	18	37	+19	15
Tank Bigsby	RB	18	42	+18	21

The model hit cleanly on several massive market inefficiencies. For example, it flagged Rashee Rice as an elite value, ranking him 16th overall versus his public ADP of 43rd (he finished 8th). It did the same with Javonte Williams at running back, ranking him 6th over his consensus ADP of 33rd (he finished 10th), and correctly identified Justin Fields as a massive

quarterback value (model rank 7 vs. ADP rank 17; finished 6th). Figure 12 illustrates these comparison tracks.

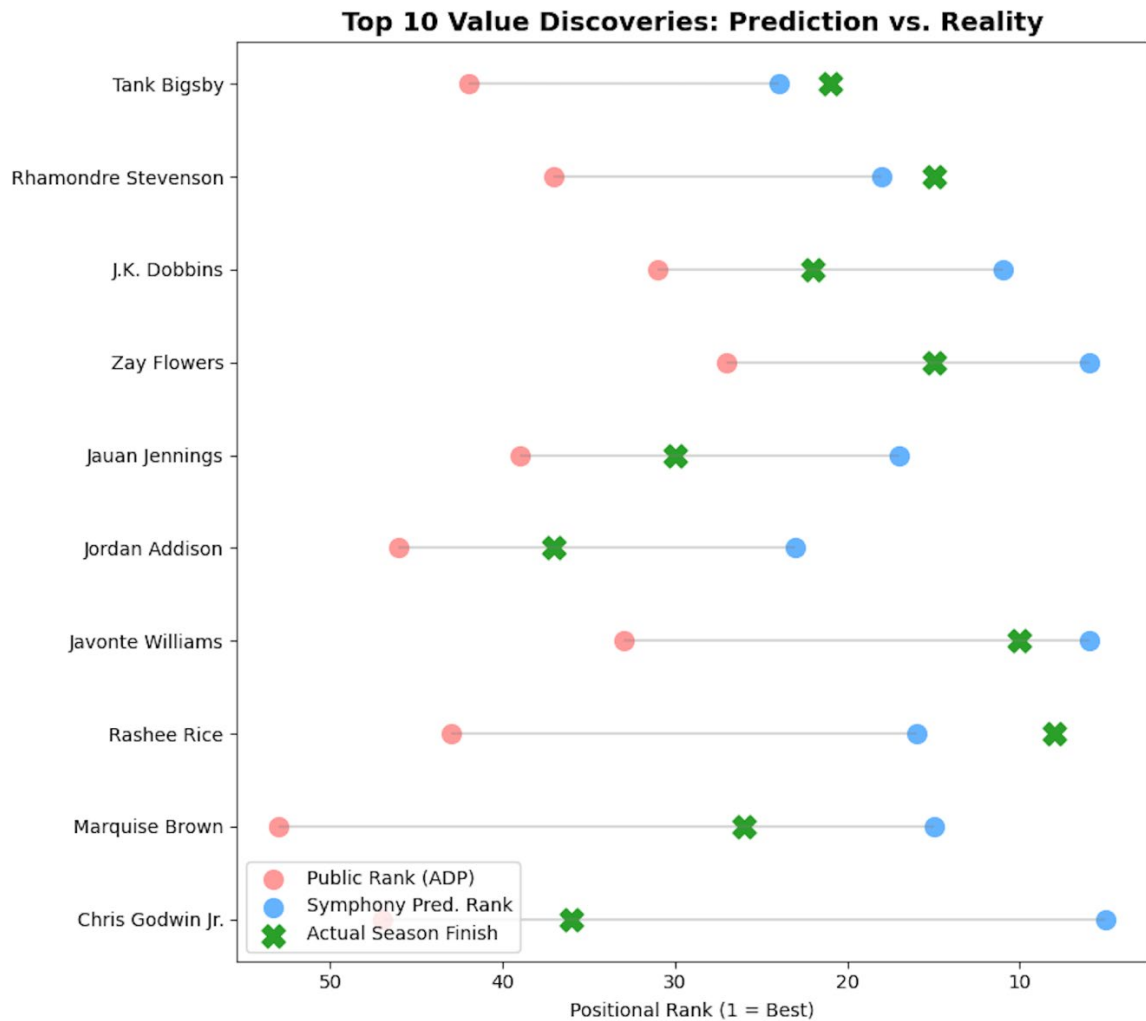


Figure 12: 2025 Performance Analysis

Figure 12. Model-predicted positional rank versus ADP positional rank versus actual 2025 positional rank.

Chapter 4: Discussion

Interpretation of RQ1: Feature Significance

The results of the correlation analysis provide partial support for the RQ1 hypothesis. Opportunity metrics including targets, carries, and receptions demonstrated strong and consistent correlations with PPG across positions. However, total touchdowns emerged as the single strongest correlate at three of four positions, which reflects a nuance in the original hypothesis. While targets show pure opportunity, touchdowns are where the actual cash-in happens. In a PPR scoring system, that six-point spike completely warps season-long totals, carrying a massive weight that easily overrides steady yardage accumulation.

The data also highlights a huge premium on quarterback rushing metrics. In the modern NFL, dual-threat QBs who rack up carries and ground scores generate a disproportionate amount of fantasy value. The model's feature selection captures this reality perfectly, prioritizing rushing yards and attempts over standard passing numbers to find the true scoring ceiling.

Interpretation of RQ2: Team Context

The negative correlations between team offensive rank and player PPG across all four positions provide support for the RQ2 hypothesis, confirming that team offensive quality does influence individual fantasy ceilings. The differential effect by position is theoretically interpretable: wide receivers are most dependent on the overall volume of team pass attempts, which is directly captured by team offensive rank, while quarterback output is more self-reinforcing and less sensitive to team context. Tight ends, whose usage is highly scheme-dependent, also show a weaker relationship, suggesting that positional role within a given system is a more important determinant of output than overall team quality for that position.

Interpretation of RQ3: Archetypes and Draft Value

The archetype analysis provides the most directly actionable findings of this study. The Workhorse Back's premium over the Pass-Catching Back archetype (17.35 vs. 14.09 PPG) challenges a prevalent heuristic in the PPR era: that backs with strong reception profiles are preferable targets. The data suggests that sheer ground game volume, measured by carries, remains the primary driver of running back fantasy value even in a pass-heavy environment. This finding is consistent with the hypothesis that high-volume players outperform their alternatives and offers a specific, data-driven corrective to a known market inefficiency.

The narrow gap between Elite and Streaming quarterback archetypes (20.94 vs. 19.84 PPG, a difference of only 1.10 points) provides quantitative validation of the long-held fantasy convention that investing premium draft capital in a quarterback yields diminishing returns relative to later-round alternatives.

Model Performance and ADP Comparison

The divergence between model correlation (0.581) and ADP correlation (-0.325) with actual 2025 outcomes is the most consequential finding of this study. This result suggests that public consensus ADP is systematically distorted by factors that do not reliably predict performance, including name recognition, narrative-driven media coverage, and prior-year volume without injury adjustment. The model, by relying exclusively on snap-adjusted historical statistics, avoids these biases.

The 52.3% model win rate over public consensus represents a modest but meaningful improvement in a domain characterized by high inherent variance and a competitive information market. It is worth acknowledging that a rate above 50% in this context is a substantive result,

and one that has practical implications for the sports wagering applications discussed in Chapter 1.

Limitations

Several limitations constrain the generalizability of this study's findings. While the snap-percentage filter does a great job cleaning out injury noise, it drastically shrinks the overall sample size. This restriction creates a clear risk of selection bias by naturally skewing the data toward players who are exceptionally durable. Second, the pipeline misses out on fluid, real-world dynamics like active injury updates, offseason trades, coaching staff overhauls, or contract disputes. Any of these situations can completely rewrite an athlete's role and production overnight. Finally, limiting the validation test set to just a single season leaves the results highly vulnerable to year-specific flukes. One weird NFL season can easily warp the metrics, meaning these numbers might not reflect how the pipeline holds up over a longer stretch.

The negative R^2 for the quarterback position is also worth acknowledging as a significant limitation. Quarterback fantasy performance is particularly sensitive to factors not captured by lagged statistics, including offensive coordinator changes, new offensive weapons, and the replacement of injured starters. The p-value analysis in Table 3 provides formal statistical confirmation of this limitation — the quarterback model's correlation with actual 2025 outcomes ($r = 0.185$, $p = 0.411$) cannot be distinguished from a null result. The overfitting assessment in Table 2 further reveals that the running back model carries the highest generalization risk, with a train-to-test MAE gap of 1.643 fantasy points, suggesting running back performance is particularly sensitive to year-specific factors such as backfield competition and injury that vary unpredictably between seasons. Future iterations of this framework would benefit substantially from the inclusion of these contextual variables for the quarterback position specifically.

Comparison to Prior Work

This study builds upon a foundation of prior academic work in sports predictive modeling. Porter (2018) demonstrated that regression-based approaches could achieve meaningful accuracy in NFL player prediction using historical statistics. The present study extends that work by incorporating opportunity efficiency metrics and a more rigorous injury-adjustment methodology. Abadzic, Cheun, and Patel (2024) employed machine learning to predict top-12 fantasy finishes by position; the present study similarly employs classification methods but extends the scope to include continuous regression, clustering-based archetype analysis, and direct comparison against public consensus rankings.

Chapter 5: Conclusion

Summary of Findings

Back testing this multi-model framework against the 2025 season after training it on seven years of historical data highlights a few critical takeaways. First, volume dictates everything. Total touchdowns, targets, receptions, and carries serve as the absolute strongest predictors of fantasy success across all four skill positions. Team offensive quality matters too, but the actual impact shifts depending on the position. Wide receivers ($r = -0.289$) and running backs ($r = -0.244$) are tied closest to their team's overall environment. When it comes to individual usage, the data shows a massive scoring premium on high volume. "Workhorse" running backs (17.35 PPG) and "Alpha" wide receivers (16.03 PPG) generated the highest scoring averages, making them the safest locks on draft day.

The biggest win came during real-world testing, where the final positional rankings hit a 0.581 Pearson correlation with actual 2025 outcomes. Compare that to the industry ADP benchmark, which sat at a backward -0.325. That accuracy edge allowed the model to out-forecast public expert consensus for 52.3% of the player pool.

Practical Applications

These findings offer direct, practical advantages for fantasy managers looking to exploit market flaws. By building data-driven rankings completely independent of public hype, the pipeline flags massive, predictable inefficiencies in industry ADP. Spotting undervalued players like Rashee Rice or Javonte Williams in 2025 shows exactly how the model uncovers non-obvious buying opportunities during a draft. Outside of traditional season-long leagues, this workflow provides a clean blueprint for sports betting. In the player props market, maintaining a

model with a 0.581 correlation to actual performance gives bettors a major edge in catching lines where bookmaker numbers drift away from the statistical baseline.

Future Directions

Expanding this research leaves several clear avenues for future work, particularly when it comes to roster turnover and incoming talent. Right now, players entering entirely new offensive schemes through trades or free agency, alongside incoming rookies with zero historical NFL data, present a major predictive challenge. To push the model's accuracy further, future updates must account for offseason roster churn—specifically incoming rookies and traded veterans. Right now, evaluating a rookie with zero NFL track record or tracking a veteran sliding into a completely new playbook creates a major predictive blind spot. Merging college production data into the pipeline and building specific weights for team adjustments would go a long way in fixing this.

Moving past basic roster moves, swapping out fixed seasonal totals for active time-series forecasting—using ARIMA or recurrent neural networks—would capture critical within-season momentum, expanding on concepts from Nau (2019) and King (2020). Finally, scaling the data pipeline down to weekly, game-level projections would directly align the model with the rapid, week-to-week decisions that competitive fantasy play demands.

In the end, this project proves that anchoring NFL player evaluation in volume metrics and ensemble machine learning generates projections that cut right through public consensus. The data confirms a broader point: well-constructed statistical models successfully strip away the human biases of traditional sports analysis, uncovering deep, objective insights that the standard "eye test" simply cannot catch.

Bibliography

- Abadzic, A., Cheun, J., & Patel, M. (2024). Data Analysis on Predicting the Top 12 Fantasy Football Players by Position. SMU Scholar. <https://scholar.smu.edu/datasciencereview/vol8/iss2/7/>
- Barbato, M. (2021, August 4). Average Draft Position Explained: How to Use ADP In Your Fantasy Football Draft. FantasyPros. <https://www.fantasypros.com/2021/08/average-draft-position-explained-how-to-use-adp-in-your-fantasy-football-draft/>
- Breiman, L., & Cutler, A. (2019). *Random forests - classification description*. Berkeley.edu. https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm
- Zegura, Nathan, Augustyn, Adam. "fantasy sport". Encyclopedia Britannica, 18 Nov. 2025, <https://www.britannica.com/sports/fantasy-sport>. Accessed 26 February 2026.
- Gusdal, B. (2016). NFL Fantasy Point Prediction – Jaden Chant. Github.io. https://jadenchant.github.io/pages/projects/nfl_fantasy_pt/
- King, J. (2020). Using Machine Learning to Predict Fantasy Football Points. Github.io. <https://jimking100.github.io/2020-06-17-Post-10/>
- Lee, M. D., & Liu, S. (2022). Drafting strategies in fantasy football: A study of competitive sequential human decision-making. *Judgment and Decision Making*, 17(4), 691–719. <https://doi.org/10.1017/s1930297500008901>
- Nau, R. (2019). Introduction to ARIMA models. Duke.edu. <https://people.duke.edu/~rnau/411arim.htm>
- Porter, J. (2018). Predictive Analytics for Fantasy Football: Predicting Player Performance Across the NFL. <https://scholars.unh.edu/cgi/viewcontent.cgi?article=1418&context=honors>
- Robinson, C. (2020). "The Prediction of Fantasy Football." Mathematics Senior Capstone Papers. <https://digitalcommons.latech.edu/mathematics-senior-capstone-papers/20/>
- Sugar, G., & Swenson, T. (n.d.). Predicting Optimal Game Day Fantasy Football Teams. https://cs229.stanford.edu/proj2015/115_report.pdf
- Topic: Fantasy sports in the U.S. (2025). Statista. <https://www.statista.com/topics/10895/fantasy-sports-in-the-us/?srsltid=AfmBOooF0cypCU63v91wqevVNQ10Xk0UzYh8fRNuwSHaBYb3VfKFCGD5#topicOverview>