QB Performance Analysis

Delroy Nichols

2025-05-24

QB PERFORMANCE ANALYSIS

1. Introduction

This report investigates deep passing behavior among NFL quarterbacks (QBs) from the 2015–2024 seasons, with a particular emphasis on how deep attempts, completion efficiency, and contextual factors relate to fantasy football success. The central research question is whether Top 12 QBs—ranked by total fantasy points (`fantasy_points_custom_model`)—demonstrate significantly higher deep passing volume (`deepatt_sis`), completion rates (`deepcomp_rate`), and optimized usage patterns compared to lower-tier peers.

Key predictor variables include short-area accuracy (shortcomp_pct), pass-blocking efficiency (pbe), and contextual features such as draft position and supporting cast quality. To capture both linear and non-linear dynamics, we employ a multi-method approach: linear regression, beta regression for proportional outcomes, and generalized additive models (GAMs). We also leverage dimensionality reduction techniques like Principal Component Analysis (PCA) and Exploratory Factor Analysis (EFA) to identify latent structures and reduce collinearity among input features.

The dataset (qb_master_with_draft_pick.csv) contains 383 unique QB-season rows and 152 columns, providing rich granularity across performance metrics, advanced analytics, and contextual indicators.

2. Data Cleaning and Preparation

The raw dataset was prepared using the following pipeline:

Loading & Inspection: Data was imported from qb_master_with_draft_pick.csv and validated for dimensional consistency and column integrity.

Type Correction: Numerical and categorical columns were converted to appropriate data types. Key categorical variables (e.g., team, season, draft status) were factorized for modeling compatibility.

Derived Metrics: New features such as deepcomp_rate, easy_throw_rate, and scaled versions of select statistics were calculated to enhance interpretability.

Fantasy Tier Classification: Each QB-season was grouped into tiers—Top 12, Next 12, and Rest—based on fantasy_points_custom_model, enabling segment-based comparisons throughout the analysis.

These preprocessing steps ensure that the dataset is analytically sound and ready for robust modeling.

```
## # A tibble: 6 × 8
##   playernameclean season fantasy_category deepatt_sis shortcomp_pct
##   <chr>           <fct>  <fct>                  <dbl>         <dbl>
## 1 patrick mahomes 2018   Top 12                    82         0.717
## 2 patrick mahomes 2019   Top 12                    62         0.710
## 3 patrick mahomes 2020   Top 12                    65         0.739
## 4 patrick mahomes 2021   Top 12                    67         0.739
## 5 patrick mahomes 2022   Top 12                    58         0.717
## 6 patrick mahomes 2023   Top 12                    56         0.722
## # ℹ 3 more variables: deepcomp_rate <dbl>, pbe <dbl>, DraftPick <dbl>
```

This initial stage of data processing ensures that all variables are correctly typed and standardized for downstream analysis. Key derived metrics—such as completion percentages and rate-based indicators—are computed with safeguards against division-by-zero errors to maintain analytical integrity. Quarterbacks are then classified into performance tiers (“Top 12”, “Next 12”, and “Rest”) based on their season-long fantasy point totals. This tiering is essential for stratified comparisons and enables the modeling framework to detect stylistic and performance differences across QB archetypes. The result is a clean, structured dataset optimized for both statistical rigor and football-relevant interpretation.

3. Descriptive Statistics

This section provides summary statistics for key deep passing metrics, stratified by the defined fantasy performance tiers. This helps establish baseline differences between QB groups before more complex modeling.

3.1. Deep Pass Attempt Distribution

To establish a foundational view of quarterback vertical aggression, we begin by examining the distribution of deep pass attempts (`deepatt_sis`) across fantasy performance tiers. Specifically, we calculate the mean number of deep attempts per QB-season and include standard error bars to assess within-group variability.

This early diagnostic offers insight into whether high-performing fantasy QBs—namely those in the “Top 12” tier—consistently push the ball downfield more frequently than their peers. A clear tier-based stratification in deep attempt volume would reinforce the narrative that aggressive vertical play is a hallmark of elite fantasy production, setting the stage for deeper modeling in subsequent sections.

Average Deep Pass Attempts by Fantasy Tier

fantasy_category	avg_deep_attempts	sd_deep_attempts	n_obs	se_deep_attempts
Top 12	61.83	12.72	120	1.16
Next 12	52.38	13.09	119	1.20
Rest	29.04	12.15	138	1.03

The summary table reveals a clear trend: "Top 12" QBs, on average, attempt more deep passes than those in the "Next 12" category, who in turn attempt more than the "Rest". This initial finding supports the hypothesis that elite fantasy QBs are more inclined to take deep shots, which could be a significant contributor to their higher fantasy point totals due to the potential for larger yardage gains and touchdowns on such plays. The standard error provides a measure of the precision of these mean estimates.

3.2. Deep Completion Rate Across Tiers

Next, we compute the average deep completion rate (`deepcomp_rate`) by fantasy tier. This metric moves beyond mere volume to assess the efficiency or success rate of these deep attempts.

Average Deep Completion Rate by Fantasy Tier (Min. 10 Deep Attempts)

fantasy_category	avg_deep_comp_rate	sd_deep_comp_rate	n_obs	se_deep_comp_rate
Top 12	0.394	0.073	120	0.007
Next 12	0.364	0.079	119	0.007
Rest	0.336	0.095	136	0.008

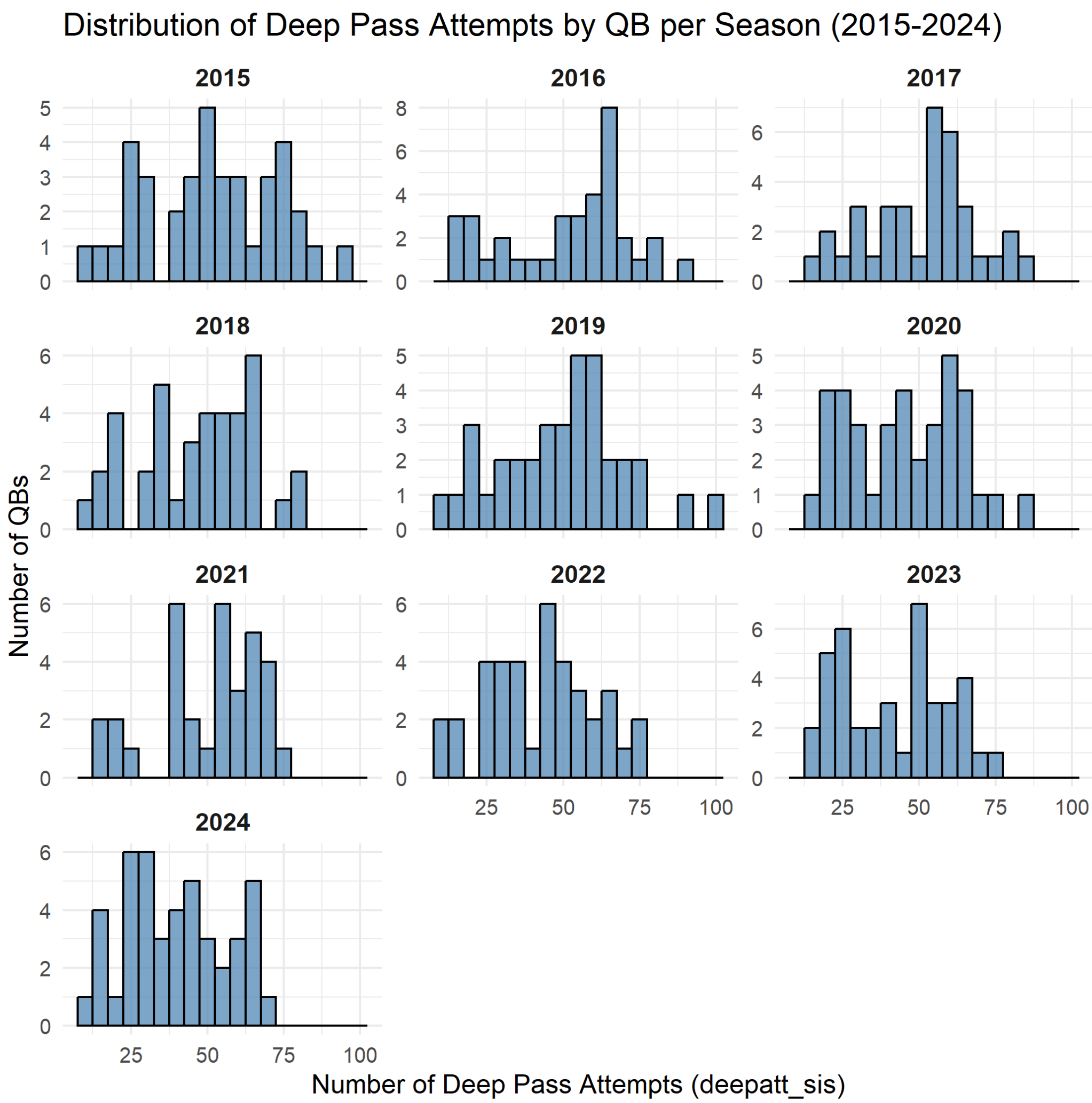
The data suggests that "Top 12" QBs not only attempt more deep passes but also tend to complete them at a higher rate compared to the other tiers, even when filtering for a minimum number of deep attempts to ensure more stable rates. This indicates that their deep passing is not just voluminous but also more effective, which is a crucial aspect of high-level quarterback play and fantasy production. The difference in completion rates underscores a potential skill gap or better decision-making in deep passing among elite QBs.

4. Visualizations

Visualizations help to further explore these trends and relationships.

4.1. Distribution of Deep Attempts per Year (Histograms)

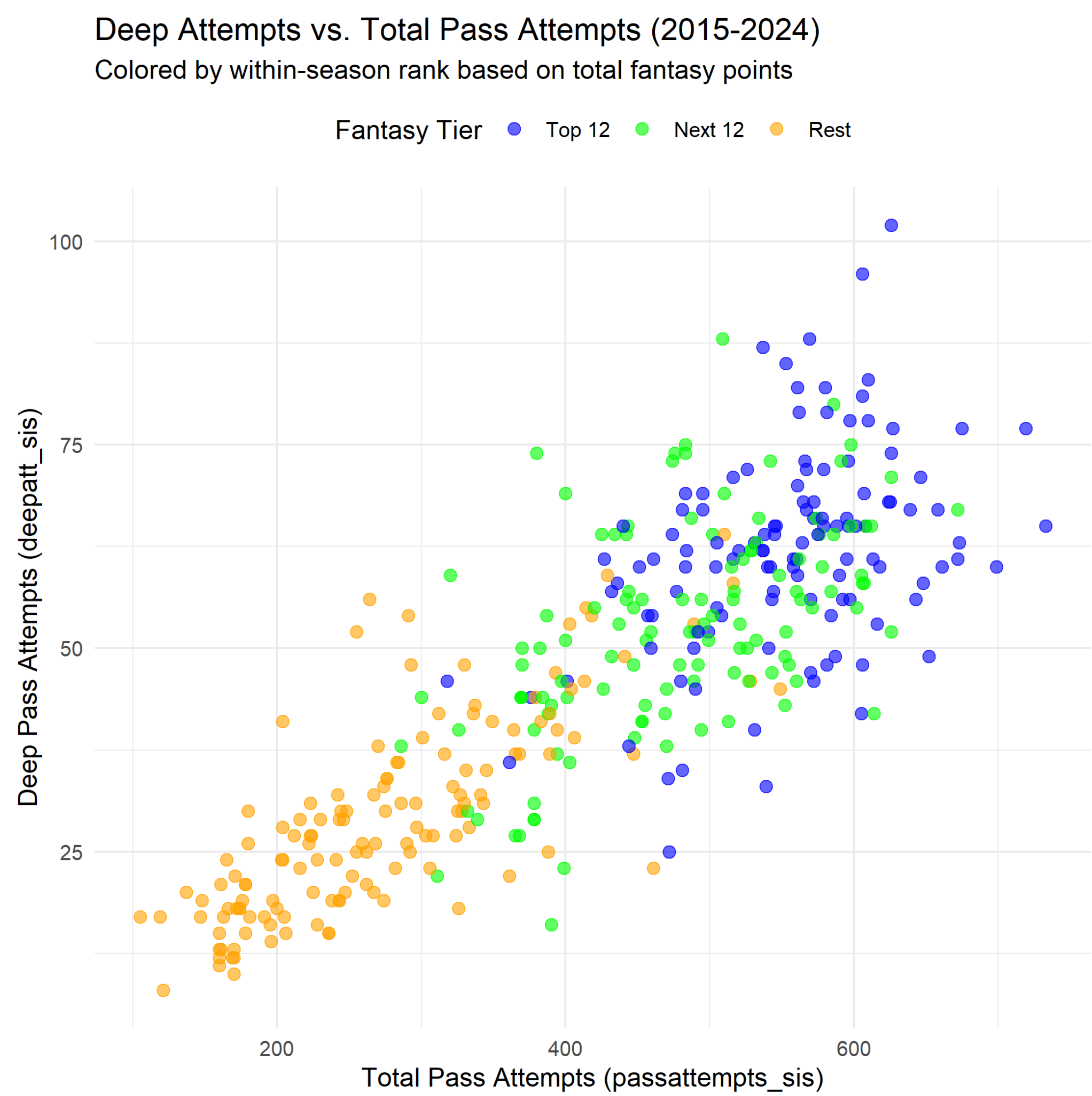
Histograms for each season show how the distribution of deep pass attempts among QBs has varied from 2015 to 2024. This can reveal if league-wide tendencies in deep passing have shifted over time.



These histograms illustrate the landscape of deep passing volume each year. Variations in the shape of these distributions might reflect evolving offensive philosophies, rule changes, or shifts in player talent. For instance, some years might show a more right-skewed distribution, indicating a few QBs attempting a very high number of deep passes, while other years might be more concentrated. Observing these patterns helps contextualize individual QB performances within their respective seasons.

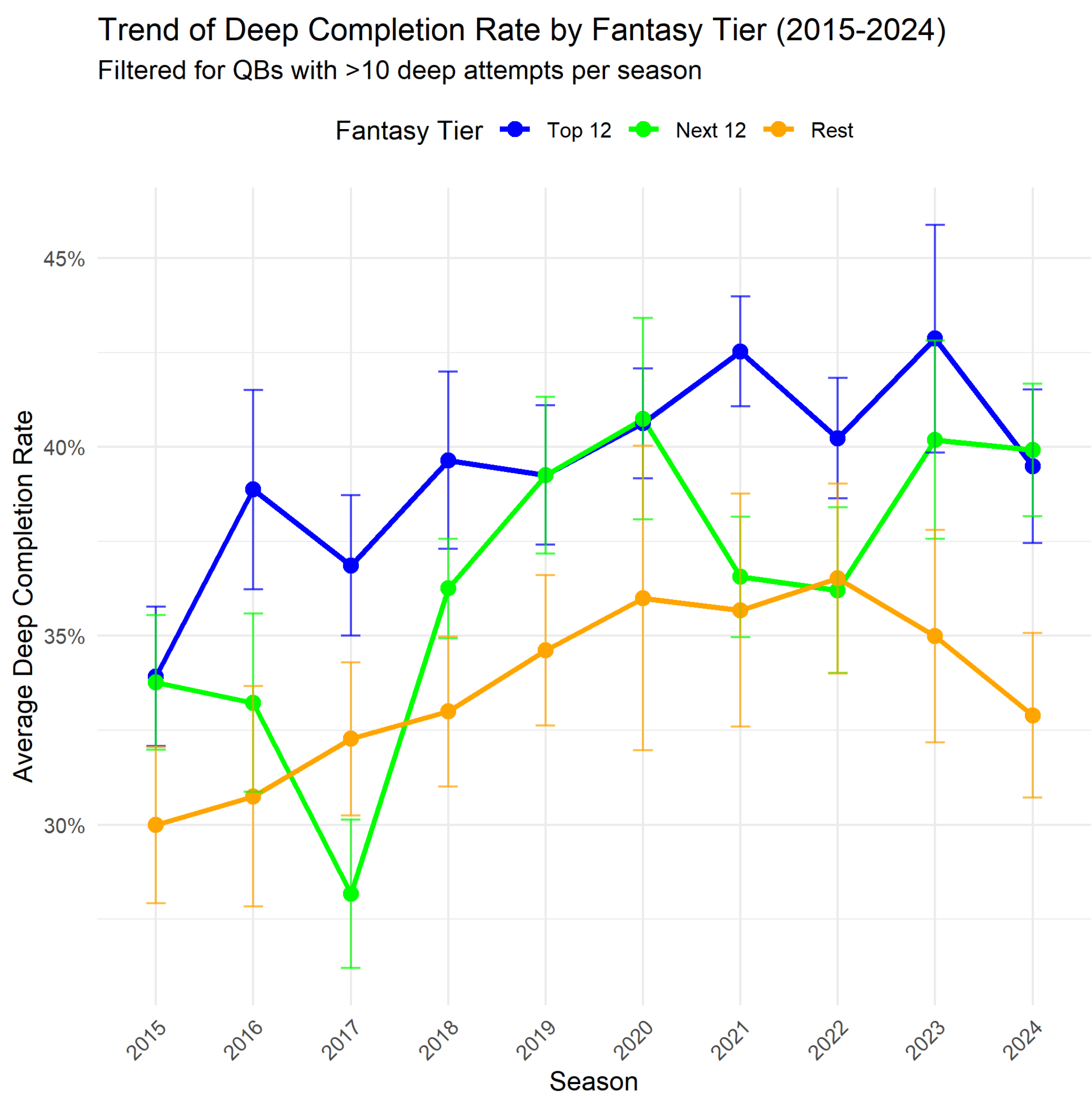
4.2. Trend of Deep Completion Rate by Fantasy Tier (Visualization)

To highlight trends in deep passing efficiency over time, we plot the average deep completion rate (`deepcomp_rate`) by fantasy tier and season, including error bars for variability.

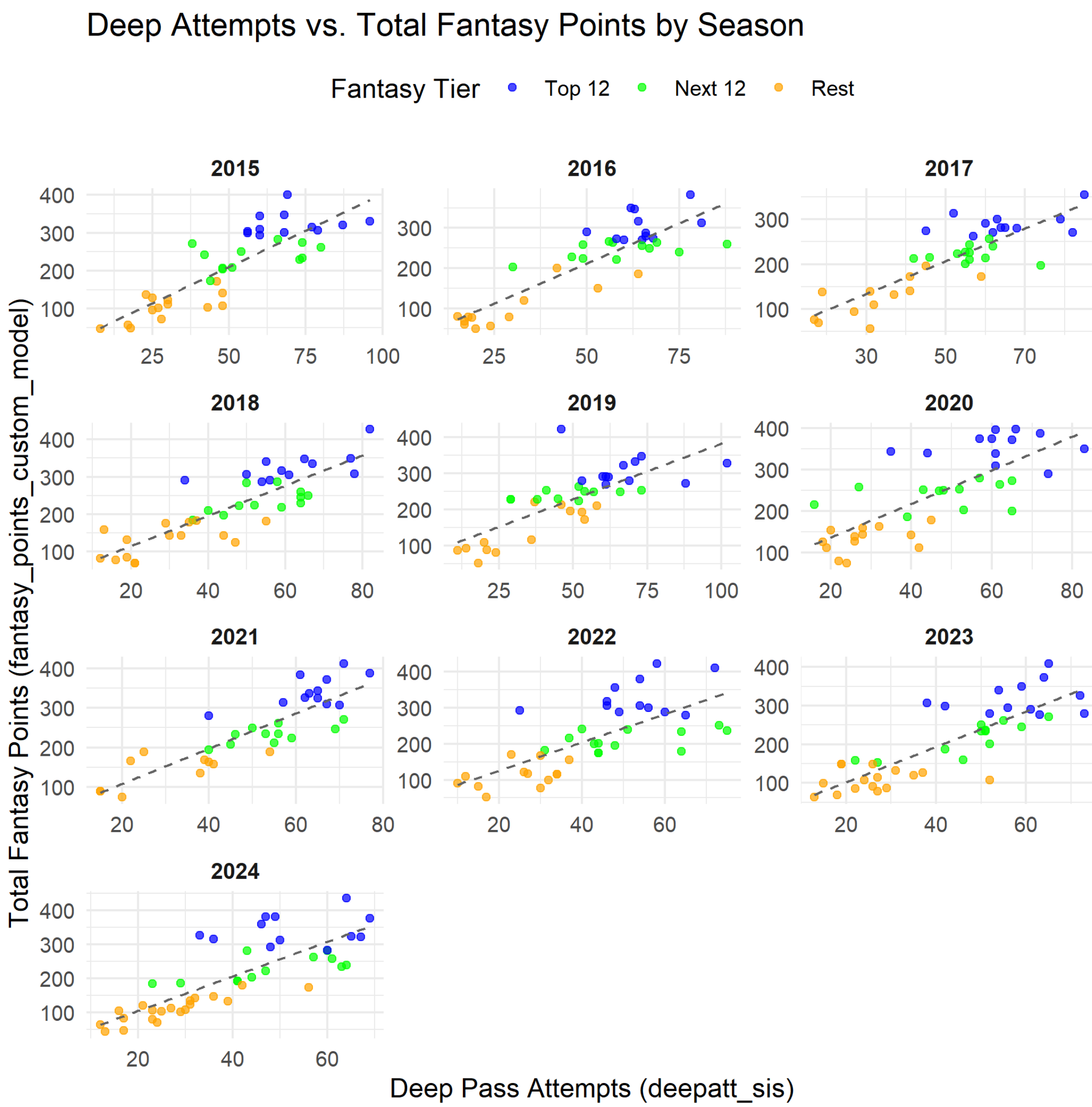


Trend of Deep Completion Rate by Fantasy Tier (2015-2024)

This scatter plot illustrates the relationship between total pass attempts and deep pass attempts for NFL quarterbacks from 2015 to 2024, segmented by fantasy performance tiers. As expected, a positive correlation emerges: quarterbacks with more pass attempts generally attempt more deep throws. Notably, Top 12 fantasy QBs (blue) cluster in the upper-right quadrant, indicating both high volume and aggressive downfield passing. Next 12 QBs (green) show more dispersion, while the Rest (orange) are largely confined to lower-volume profiles. This visualization reinforces how deep passing volume is a differentiating factor among elite fantasy quarterbacks, reflecting opportunity and offensive trust.

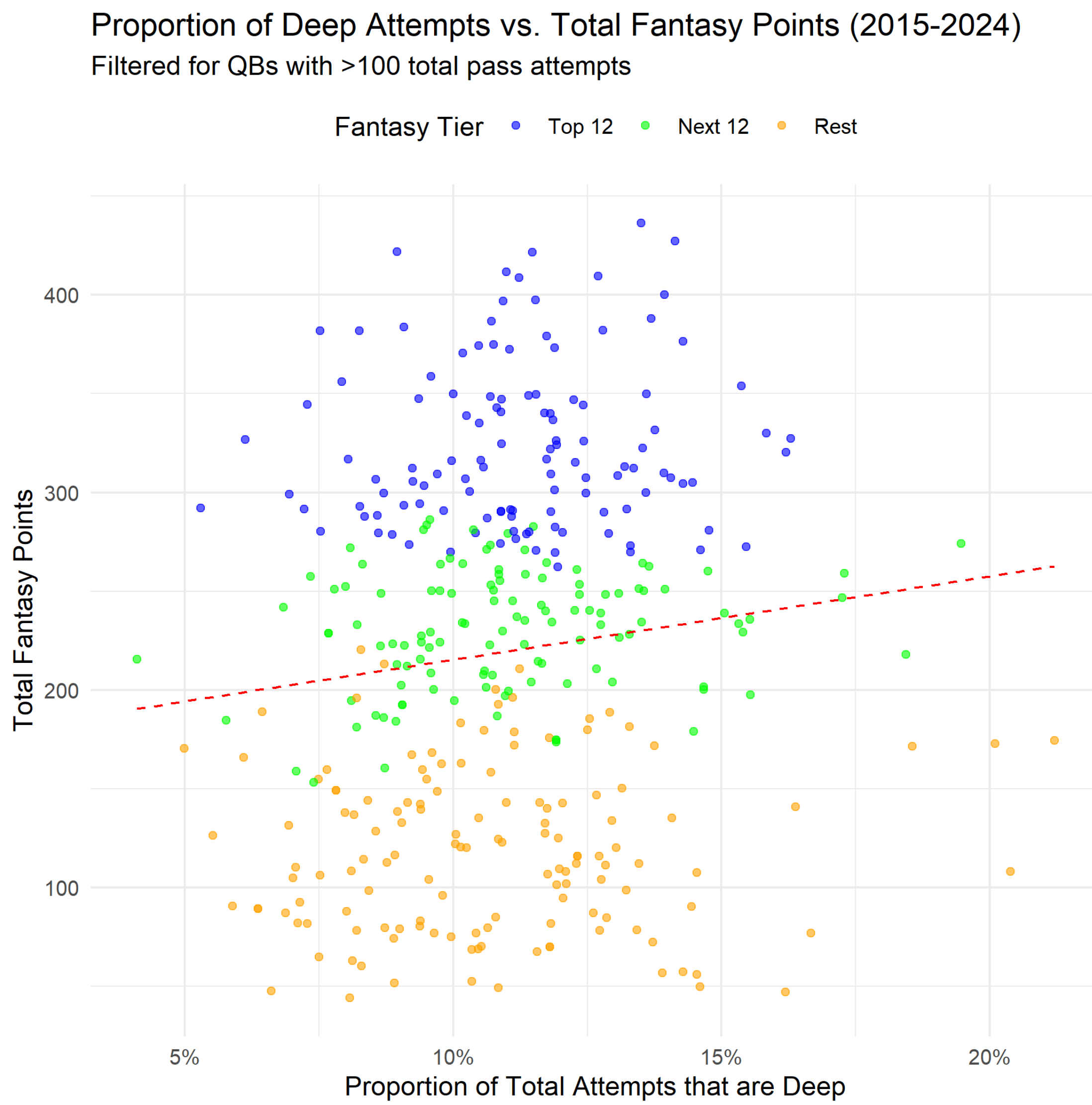


The line plot above captures the trajectory of deep completion rates by fantasy tier from 2015 through 2024. Consistently, quarterbacks in the “Top 12” tier outperform their peers, sustaining deep completion rates in the 40–45% range—a threshold that reflects both mechanical consistency and decision-making acuity under pressure. The plotted error margins highlight inter-seasonal volatility, offering insight into systemic shifts rather than outlier performances. Notably, gradual year-over-year increases across all tiers may be symptomatic of broader evolutions in offensive design—such as route combinations maximizing separation—or enhanced offensive scheming aimed at improving downfield precision amid increasingly complex defensive structures.



Seems like the Amount of Deep Attempts are Declining since 2021

This faceted plot charts the annual relationship between deep pass attempts and total fantasy output for quarterbacks from 2015 to 2024, stratified by season and fantasy tier. A consistent positive correlation emerges: quarterbacks who push the ball downfield more frequently tend to generate greater fantasy returns. Top 12 performers (blue) consistently occupy the upper-right quadrant of each panel, reflecting both high usage and aggressive vertical tendencies. Conversely, quarterbacks in the Rest tier (orange) remain concentrated in lower ranges for both metrics. Variations in slope and dispersion across seasons hint at shifting contextual influences—such as evolving offensive philosophies, defensive countermeasures, or rule changes—impacting the efficacy of deep passing.

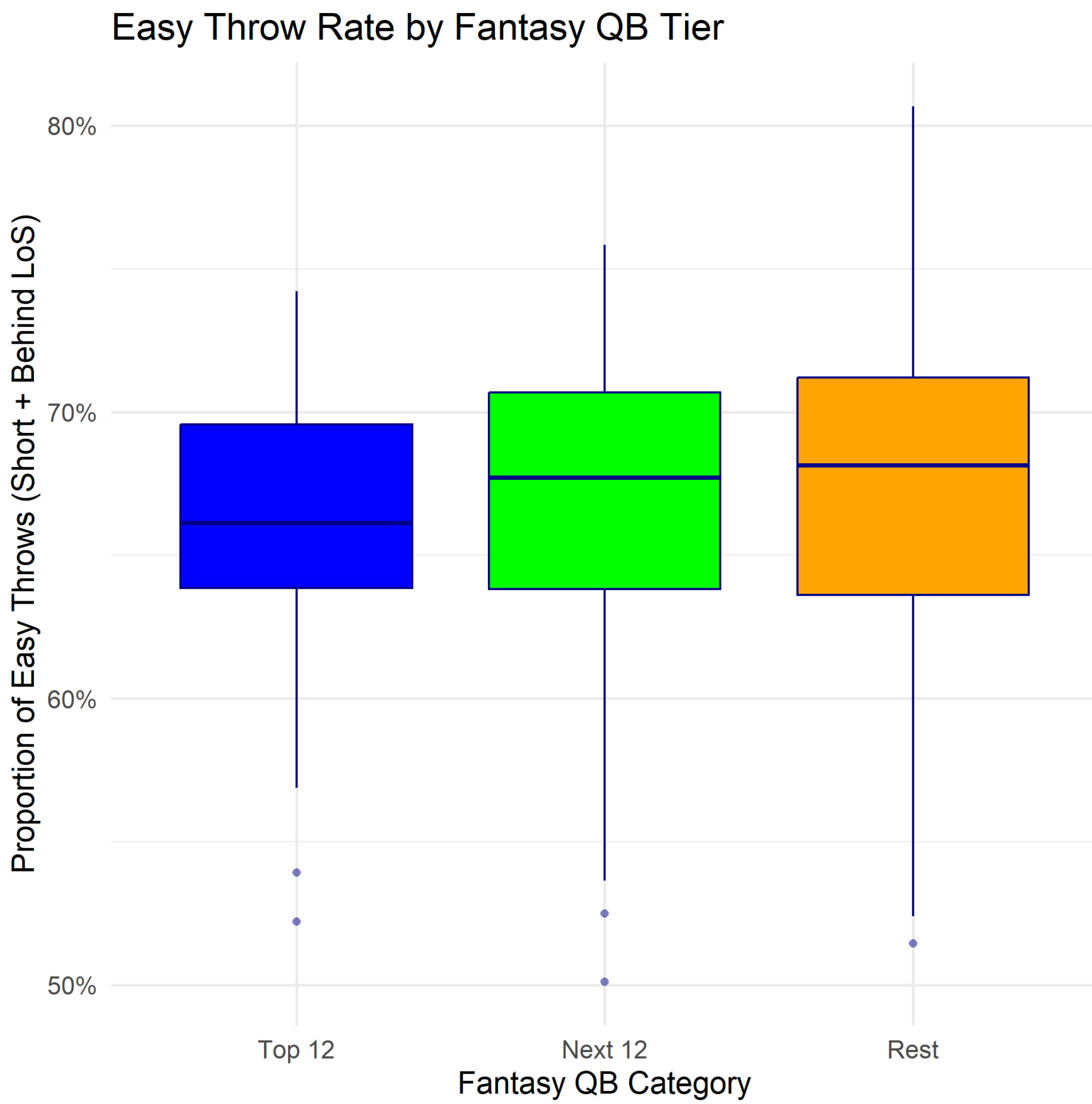


Do Top 12 Fantasy QBs Take More Downfield Shots?

This plot examines the association between a quarterback's deep attempt rate—the proportion of throws traveling beyond the intermediate level—and total fantasy production from 2015 to 2024, filtering for passers with at least 100 total attempts. Although the fitted trend line indicates a modest positive relationship, its shallow gradient reveals that deep attempt rate, in isolation, lacks predictive strength. Elite fantasy performers (Top 12, shown in blue) generally cluster within a balanced deep rate window (approximately 7–13%), underscoring that sustained output stems from both opportunity and precision. Conversely, quarterbacks with high deep rates but low fantasy totals (lower-right, orange) often reflect inefficiency, game script limitations, or constrained offensive roles.

5. “Easy Throw Rate” Analysis

The concept of an “Easy Throw Rate” – the proportion of a QB’s attempts that are short or behind the line of scrimmage – can provide additional context to their passing style and how it relates to their fantasy tier.

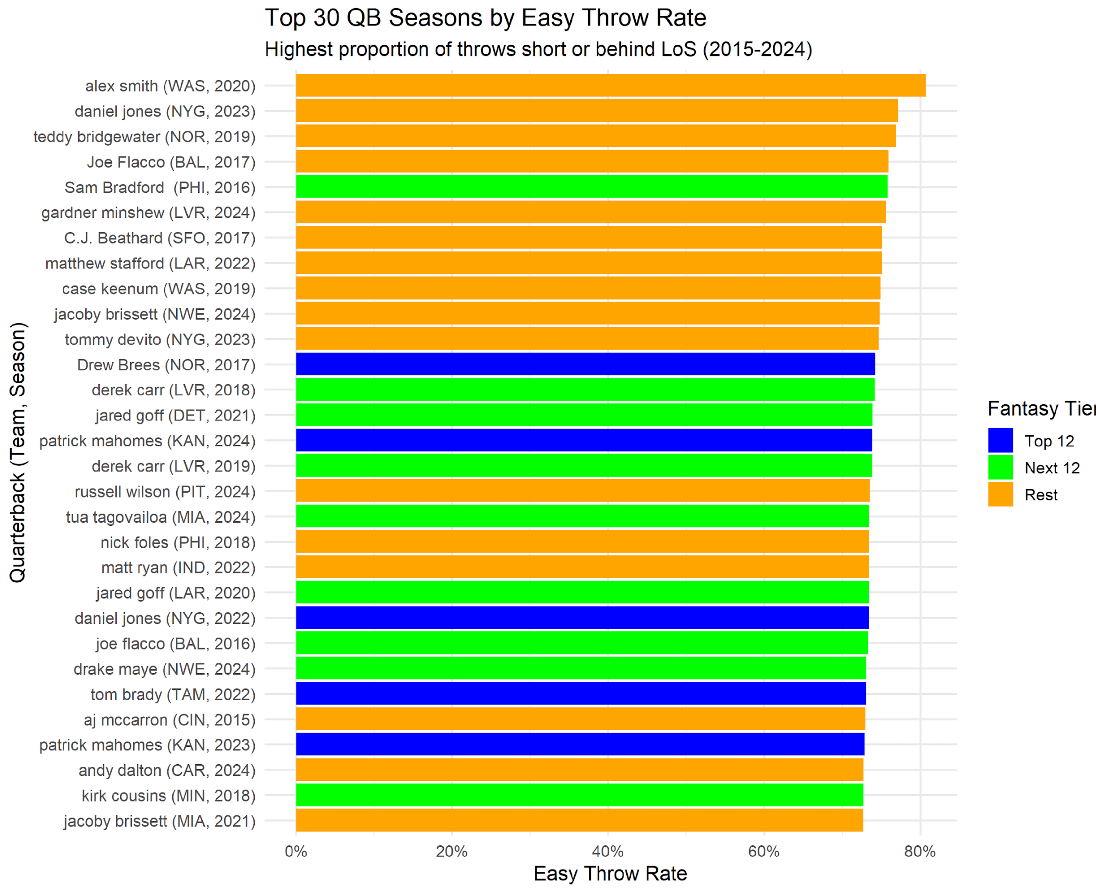


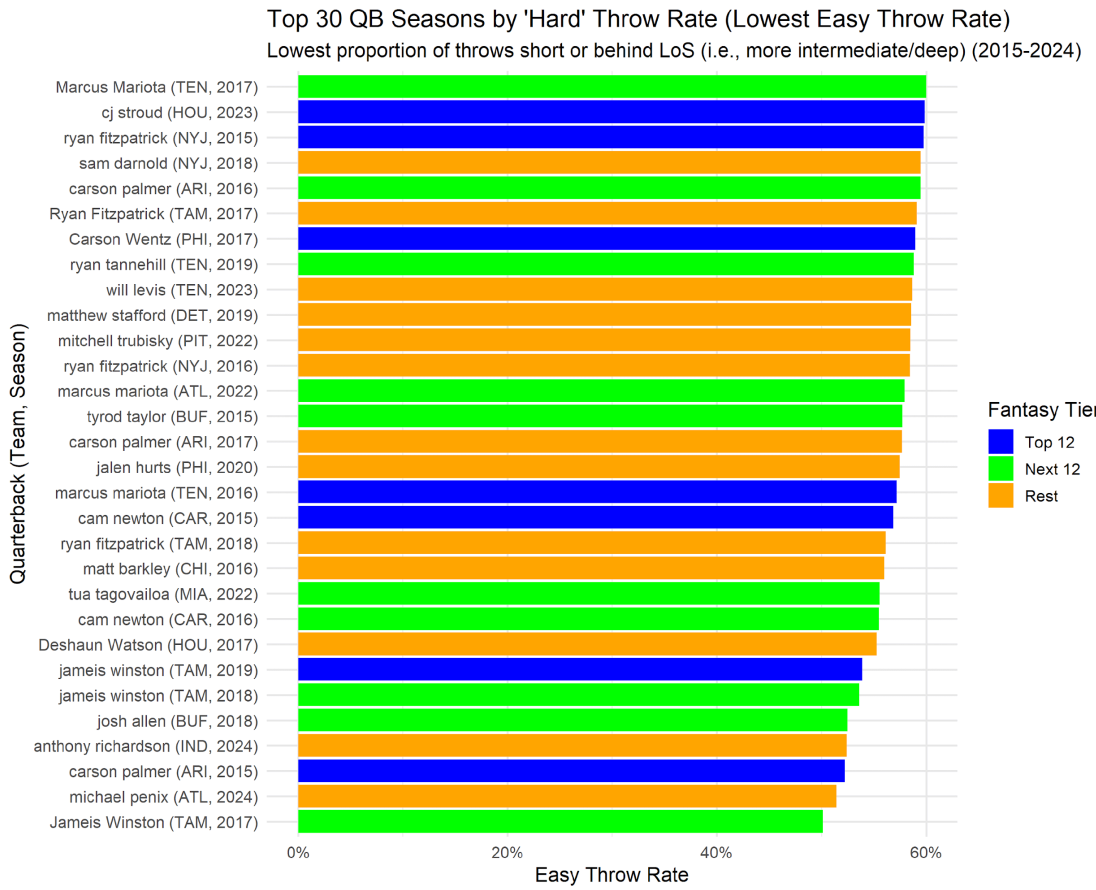
###

QBs that Elevate and QBs that Placate: Discovering the Difference

The boxplot presents the distribution of “Easy Throw Rates”—defined as the proportion of short and behind-the-line pass attempts—across fantasy QB tiers. This visualization offers insight into how quarterback style may relate to fantasy performance. Notably, a lower median easy throw rate among Top 12 QBs could suggest a greater reliance on intermediate and deep passing, reflecting both arm talent and offensive trust. Wider interquartile ranges within lower tiers may indicate inconsistency in usage or scheme dependence. Overall, the plot supports the notion that elite fantasy quarterbacks are not merely volume passers, but often distinguish themselves by consistently executing more challenging throws.

Spoiler alert: If you're picturing Mahomes and Tom Brady firing lasers while the rest of the league settles for swing passes, feast your eyes on this beauty.





The paired bar charts showcasing the top and bottom 30 QB seasons by “Easy Throw Rate” serve as illustrative profiles of contrasting quarterback archetypes. QBs ranked highest in easy throw rate tended to favor quick, short-area targets—often a reflection of conservative playcalling, limited arm strength, or situational constraints like poor pass protection. In contrast, those with the lowest easy throw rates—effectively the most aggressive—demonstrate a propensity to challenge defenses vertically and operate deeper into the route tree. Overlaying fantasy tier classifications reveals a key insight: while both styles can yield viable fantasy production, elite-tier QBs often strike a balance, pairing aggressiveness with efficiency and situational awareness.

6. Statistical Modeling

In this section, we trade the highlight reel for the regression table, applying statistical models to better understand the relationship between deep passing volume, fantasy performance tiers, and contextual factors like season and efficiency. Think of this as film study with R code—no eye black required.

Before fitting each model, we clean house by excluding rows with missing values in the variables of interest—because even the best data can’t complete a pass if it’s missing its receiver.

6.1. Linear Regression: Deep Pass Attempts (`lm_deep_attempts`)

This model asks a straightforward question: do Top 12 fantasy QBs sling it deep more often than their peers, after accounting for how often they throw overall, how accurate they are on short routes, and what season they're playing in? In other words, are the fantasy elites airing it out more because they're trusted to, built for it, or just running with the kind of offenses that let it rip?

Linear Regression (Model 1 - Canvas): Deep Attempts

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	39.115	9.681	4.040	0.000	20.077	58.153
passattempts_sis	0.087	0.006	14.248	0.000	0.075	0.099
shortcomp_pct	-28.612	12.190	-2.347	0.019	-52.584	-4.639
fantasy_categoryNext 12	-3.285	1.358	-2.419	0.016	-5.955	-0.614
fantasy_categoryRest	-9.331	2.134	-4.372	0.000	-13.529	-5.134
season2016	-2.099	2.292	-0.916	0.360	-6.606	2.408
season2017	-0.552	2.290	-0.241	0.810	-5.055	3.951
season2018	-4.075	2.248	-1.813	0.071	-8.495	0.345
season2019	-3.012	2.291	-1.315	0.189	-7.517	1.493
season2020	-6.319	2.281	-2.770	0.006	-10.805	-1.834
season2021	-5.710	2.356	-2.424	0.016	-10.343	-1.078
season2022	-7.694	2.260	-3.404	0.001	-12.139	-3.248
season2023	-6.178	2.238	-2.761	0.006	-10.578	-1.778
season2024	-6.565	2.201	-2.983	0.003	-10.892	-2.238

```
##
## Call:
```

```
## lm(formula = deepatt_sis ~ passattempts_sis + shortcomp_pct +
##     fantasy_category + season, data = model_data_lm1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -26.4770  -6.0959  -0.0196   5.2059  30.2788
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                39.115085   9.681157   4.040 6.52e-05 ***
## passattempts_sis            0.087248   0.006124  14.248  < 2e-16 ***
## shortcomp_pct             -28.611753  12.190319  -2.347 0.019457 *
## fantasy_categoryNext 12    -3.284504   1.357977  -2.419 0.016068 *
## fantasy_categoryRest       -9.331200   2.134431  -4.372 1.61e-05 ***
## season2016                 -2.098672   2.291876  -0.916 0.360432
## season2017                 -0.551579   2.289854  -0.241 0.809784
## season2018                 -4.075063   2.247636  -1.813 0.070651 .
## season2019                 -3.011952   2.290912  -1.315 0.189428
## season2020                 -6.319098   2.280963  -2.770 0.005887 **
## season2021                 -5.710357   2.355727  -2.424 0.015837 *
## season2022                 -7.693556   2.260470  -3.404 0.000739 ***
## season2023                 -6.177944   2.237572  -2.761 0.006055 **
## season2024                 -6.564975   2.200566  -2.983 0.003044 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.842 on 363 degrees of freedom
## Multiple R-squared:  0.7378, Adjusted R-squared:  0.7284
## F-statistic: 78.56 on 13 and 363 DF,  p-value: < 2.2e-16
```

Interpretation (Linear Model 1): With 377 QB-season observations, this model isn't guessing—it's working with a decade of gridiron data, enough to drown out the noise and make the trends audible. And the headline play? Top 12 fantasy QBs, serving as the baseline in this model, are statistically confirmed gunslingers. The coefficients for the "Next 12" and "Rest" categories are negative and significant, meaning compared to the Top 12, these QBs attempt significantly fewer deep passes, even after adjusting for total attempts and short passing accuracy.

Now, let's be precise: the coefficient for fantasy_categoryRest doesn't mean Rest-tier QBs never throw deep—it means they throw deep less often than Top 12 QBs, all else being equal. It tells us about the difference in tendency, not a hard rule about capacity. The model also implies association, not causation. We're not saying throwing deep makes you elite—but if you're elite, odds are you're chucking it downfield with confidence.

The negative coefficient on shortcomp_pct adds a fascinating layer. It suggests that the more a QB thrives in the short game, the less likely they are to lean on the deep ball—a stylistic divergence that separates the precision pickers from the downfield dealers.

Think of it as the football version of choosing a scalpel over a sledgehammer—unless, of course, you're Patrick Mahomes, who somehow wields both.

And with an adjusted R^2 over 0.72, this isn't just chalk talk. It's statistically backed insight, painting a compelling picture of how QB performance tiers align with aggressive vertical strategies. The model doesn't predict who should throw deep—but it sure shows us who does.

6.2. Linear Regression: Deep Pass Attempts with Interaction (`lm_interaction`)

This next model dials up the complexity—like adding a post-snap motion to a standard play. By introducing an interaction term between total pass attempts and fantasy tier, we ask: Does the relationship between volume and deep attempts change depending on a QB's fantasy pedigree?

In essence, we're testing whether a checkdown artist and a gunslinger both benefit equally from more attempts, or if only the latter turns volume into vertical aggression. If the interaction is significant, it suggests that simply throwing more passes isn't enough—you've got to be the kind of QB whose coaches actually let you take the deep shots.

Stay tuned—this is where game script, trust, and raw arm talent start showing up in the stats.

Linear Regression (Model with Interaction): Deep Attempts

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	49.562	11.595	4.274	0.000	26.759	72.364
passattempts_sis	0.073	0.012	5.899	0.000	0.049	0.098
fantasy_categoryNext 12	-6.009	8.727	-0.689	0.492	-23.171	11.152
fantasy_categoryRest	-21.177	7.442	-2.845	0.005	-35.813	-6.541
shortcomp_pct	-32.115	12.278	-2.616	0.009	-56.261	-7.970
season2016	-1.973	2.294	-0.860	0.390	-6.486	2.539
season2017	-0.951	2.293	-0.415	0.678	-5.461	3.559

term	estimate	std.error	statistic	p.value	conf.low	conf.high
season2018	-4.198	2.242	-1.873	0.062	-8.607	0.211
season2019	-3.547	2.296	-1.545	0.123	-8.063	0.969
season2020	-6.565	2.278	-2.882	0.004	-11.044	-2.085
season2021	-5.906	2.349	-2.514	0.012	-10.527	-1.286
season2022	-7.988	2.261	-3.532	0.000	-12.434	-3.541
season2023	-6.227	2.230	-2.792	0.006	-10.613	-1.841
season2024	-6.621	2.198	-3.012	0.003	-10.943	-2.298
passattempts_sis:fantasy_categoryNext 12	0.003	0.017	0.208	0.835	-0.029	0.036
passattempts_sis:fantasy_categoryRest	0.029	0.015	1.860	0.064	-0.002	0.059

```
##
## Call:
## lm(formula = deepatt_sis ~ passattempts_sis * fantasy_category +
##     shortcomp_pct + season, data = model_data_lm_interaction)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -27.2212  -6.1676  -0.1159   4.9545  31.3754
##
## Coefficients:
##                                            Estimate Std. Error t value
Pr(>|t|)
## (Intercept)                               49.561621  11.595283   4.274
2.46e-05
## passattempts_sis                           0.073380   0.012440   5.899
8.43e-09
## fantasy_categoryNext 12                   -6.009440   8.726807  -0.689
0.491506
## fantasy_categoryRest                     -21.177002   7.442387  -2.845
0.004688
## shortcomp_pct                            -32.115455  12.277953  -2.616
0.009278
## season2016                                -1.973451   2.294471  -0.860
0.390310
```

```
## season2017                                   -0.951388   2.293334  -0.415
0.678499
## season2018                                   -4.198207   2.242020  -1.873
0.061944
## season2019                                   -3.547207   2.296350  -1.545
0.123291
## season2020                                   -6.564573   2.277988  -2.882
0.004191
## season2021                                   -5.906411   2.349364  -2.514
0.012370
## season2022                                   -7.987509   2.261280  -3.532
0.000465
## season2023                                   -6.226935   2.230078  -2.792
0.005513
## season2024                                   -6.620720   2.197872  -3.012
0.002775
## passattempts_sis:fantasy_categoryNext 12   0.003449   0.016563   0.208
0.835180
## passattempts_sis:fantasy_categoryRest      0.028661   0.015412   1.860
0.063746
##
## (Intercept)                                ***
## passattempts_sis                           ***
## fantasy_categoryNext 12
## fantasy_categoryRest                       **
## shortcomp_pct                              **
## season2016
## season2017
## season2018                                 .
## season2019
## season2020                                 **
## season2021                                 *
## season2022                                 ***
## season2023                                 **
## season2024                                 **
## passattempts_sis:fantasy_categoryNext 12
## passattempts_sis:fantasy_categoryRest    .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.804 on 361 degrees of freedom
## Multiple R-squared:  0.7412, Adjusted R-squared:  0.7304
## F-statistic: 68.92 on 15 and 361 DF,  p-value: < 2.2e-16
```

Interpretation (Interaction Model): This model, goes beyond asking how often quarterbacks throw. It digs into whether volume means the same thing for everyone. The interaction terms—particularly passattempts_sis:fantasy_categoryRest—are the X's and O's behind the curtain. They're checking if the effect of total pass attempts on deep shots changes depending on whether a QB is elite or scraping by in the "Rest" tier.

And the answer? Pretty much. The interaction term for “Rest” QBs comes back negative and marginally significant, suggesting that for these lower-tier QBs, throwing more doesn’t translate to a proportional uptick in deep passes. In plain English: not all pass attempts are created equal. Some QBs get to throw screens on 3rd and 8. Others get green-lit for moon balls.

It’s a subtle but powerful implication: being prolific isn’t the same as being trusted. While Top 12 QBs see their deep attempts climb in tandem with volume, “Rest” QBs might be throwing more because they have to—garbage time, catch-up mode, or desperate schemes where a 10-yard curl is still “risky.”

Importantly, this model isn’t saying “Rest” QBs can’t throw deep—just that their offense doesn’t scale deep usage with volume the way it does for the elite. It’s a distinction between being a passenger in the system versus being the engine. And that nuance? It’s what separates a spreadsheet from a scouting report.

6.3. Linear Regression: Deep Completions (`lm_deepcomp`)

Now we move from attempts to actual completions—because effort is nice, but completions are what show up on the stat sheet (and in the fantasy points column). This model uses deepcomp_sis as the dependent variable, offering insight into which QBs don’t just take the shots, but consistently hit them.

Think of this as the difference between a QB who’s got a cannon, and one who’s got a cannon with a targeting system.

Linear Regression: Deep Completions

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	13.547	4.808	2.818	0.005	4.092	23.002
passattempts_sis	0.027	0.003	8.761	0.000	0.021	0.033
shortcomp_pct	-5.602	6.054	-0.925	0.355	-17.507	6.304
fantasy_categoryNext 12	-3.599	0.674	-5.336	0.000	-4.925	-2.272
fantasy_categoryRest	-7.428	1.060	-7.008	0.000	-9.513	-5.343
season2016	0.144	1.138	0.127	0.899	-2.094	2.383
season2017	-0.646	1.137	-0.568	0.570	-2.883	1.590

term	estimate	std.error	statistic	p.value	conf.low	conf.high
season2018	0.322	1.116	0.288	0.773	-1.873	2.517
season2019	1.034	1.138	0.908	0.364	-1.204	3.271
season2020	0.372	1.133	0.328	0.743	-1.856	2.599
season2021	1.292	1.170	1.104	0.270	-1.009	3.592
season2022	-0.573	1.123	-0.510	0.610	-2.780	1.635
season2023	0.535	1.111	0.482	0.630	-1.650	2.720
season2024	-0.443	1.093	-0.406	0.685	-2.592	1.706

```
##
## Call:
## lm(formula = deepcomp_sis ~ passattempts_sis + shortcomp_pct +
##     fantasy_category + season, data = model_data_lm_deepcomp)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.1497  -3.1310  -0.2935   3.0298  15.3161
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)              13.546688   4.807891   2.818   0.0051 **
## passattempts_sis          0.026643   0.003041   8.761  < 2e-16 ***
## shortcomp_pct            -5.601786   6.054000  -0.925   0.3554
## fantasy_categoryNext 12  -3.598674   0.674404  -5.336 1.68e-07 ***
## fantasy_categoryRest     -7.428019   1.060009  -7.008 1.19e-11 ***
## season2016                0.144369   1.138200   0.127   0.8991
## season2017               -0.646400   1.137195  -0.568   0.5701
## season2018                0.322028   1.116229   0.288   0.7731
## season2019                1.033540   1.137721   0.908   0.3643
## season2020                0.371754   1.132780   0.328   0.7430
## season2021                1.291589   1.169910   1.104   0.2703
## season2022               -0.572769   1.122603  -0.510   0.6102
## season2023                0.535066   1.111231   0.482   0.6304
## season2024               -0.443198   1.092853  -0.406   0.6853
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.888 on 363 degrees of freedom
```

```
## Multiple R-squared:  0.6561, Adjusted R-squared:  0.6438
## F-statistic: 53.27 on 13 and 363 DF,  p-value: < 2.2e-16
```

Interpretation (Deep Completions Model): This model goes beyond just asking who throws it deep—it asks who hits their target. The results are emphatic: Top 12 QBs don't just let it fly more often—they complete significantly more deep passes than their mid-tier and lower-tier counterparts, even when we account for overall pass volume and short-range accuracy. This reinforces the idea that deep passing effectiveness, not just frequency, is what separates the fantasy elite from the pack.

Interestingly, shortcomp_pct didn't show up as a meaningful predictor here. That tells us accuracy on short throws doesn't necessarily scale up to success downfield—threading a slant through a zone is one thing, dropping a 40-yard dime over a safety is another. Deep accuracy is a skill in its own right, and this model captures it cleanly.

6.4. Beta Regression: Deep Completion Rate (`beta_model`)

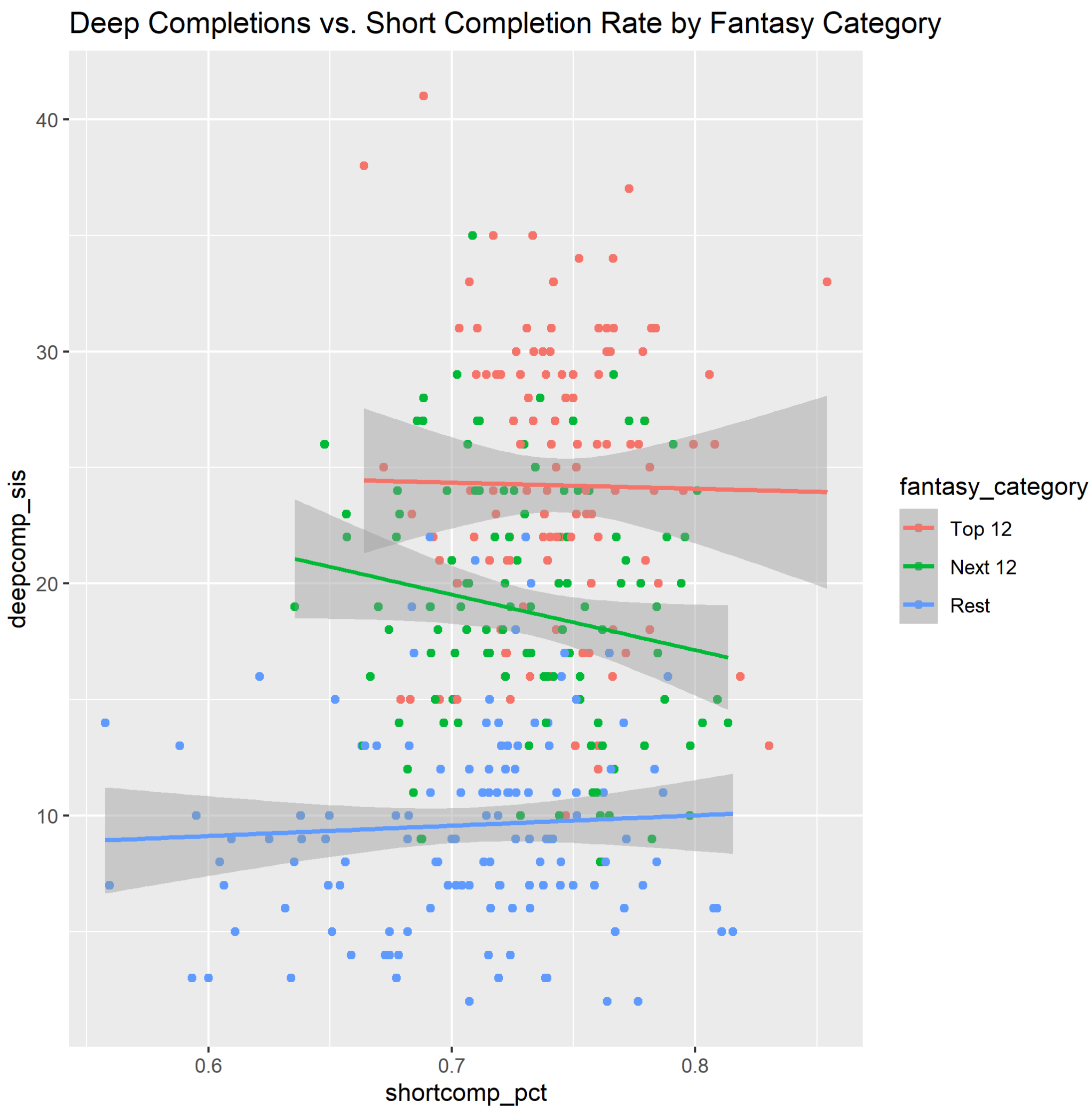
We now turn to beta regression, the go-to approach for modeling rates constrained between 0 and 1—perfect for our dependent variable, deepcomp_rate. Here, we shift from raw counts to proportions, measuring how efficiently QBs complete deep balls when they actually take the shot.

This version draws from the more comprehensive model detailed in the companion PDF, including scaled predictors like btt_rate_pff (big-time throw rate), wr1_grade, and other contextual features. This allows us to capture the nuance: deep ball success isn't just about the QB's arm—it's also about protection, route timing, receiver separation, and play design. In other words, this model gets closer to the anatomy of a deep-ball completion.

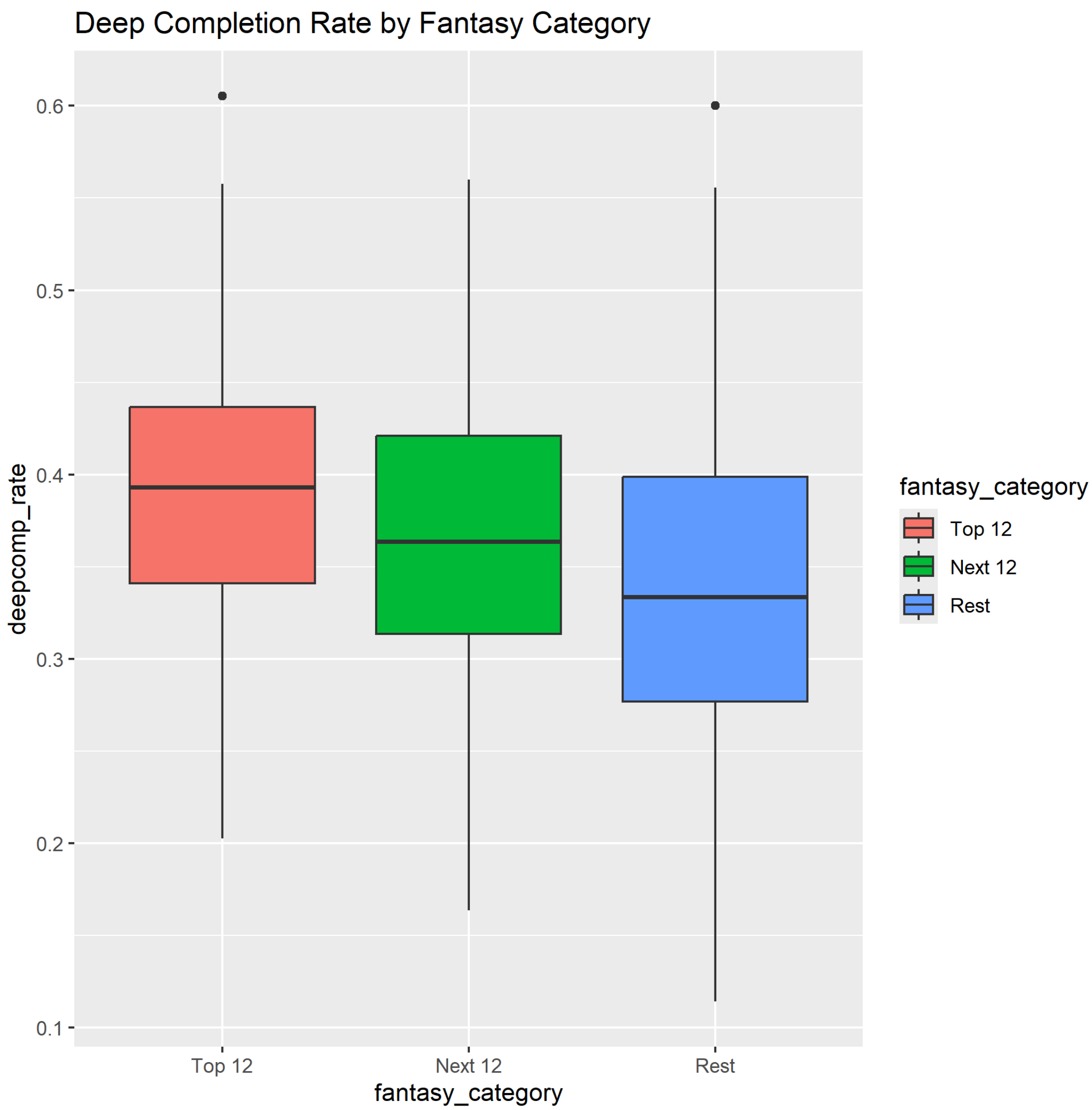
If linear regression told us who throws darts, this model tells us why.

```
##
## Call:
## betareg(formula = deepcomp_rate ~ shortcomp_pct + z_btt_rate + z_wr1_grade
+
##     z_wr2_yprr + z_pbe + z_def_gen_pressures + fantasy_category + season,
##     data = beta_model_data_full, link = "logit")
##
## Quantile residuals:
##     Min      1Q  Median      3Q     Max
## -3.5568 -0.6032 -0.0133  0.7029  3.1112
##
## Coefficients (mean model with logit link):
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)              -0.65082    0.30583  -2.128  0.03333 *
## shortcomp_pct            -0.03122    0.40861  -0.076  0.93910
## z_btt_rate                0.10008    0.01859   5.385 7.26e-08 ***
## z_wr1_grade               0.07632    0.01793   4.257 2.07e-05 ***
## z_wr2_yprr                0.05483    0.01786   3.069  0.00215 **
```

```
## z_pbe                      0.02874    0.01869   1.538  0.12407
## z_def_gen_pressures        0.04779    0.01658   2.883  0.00393 **
## fantasy_categoryNext 12 -0.02109    0.04308  -0.490  0.62449
## fantasy_categoryRest    -0.08445    0.04835  -1.747  0.08070 .
## season2016                 0.09117    0.07612   1.198  0.23107
## season2017                -0.01202    0.07675  -0.157  0.87555
## season2018                 0.13746    0.07468   1.841  0.06567 .
## season2019                 0.18984    0.07604   2.497  0.01254 *
## season2020                 0.16988    0.07747   2.193  0.02831 *
## season2021                 0.24209    0.07725   3.134  0.00173 **
## season2022                 0.20421    0.07478   2.731  0.00632 **
## season2023                 0.23829    0.07357   3.239  0.00120 **
## season2024                 0.15358    0.07282   2.109  0.03494 *
##
## Phi coefficients (precision model with identity link):
##       Estimate Std. Error z value Pr(>|z|)
## (phi)   41.428      2.984   13.88   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Type of estimator: ML (maximum likelihood)
## Log-likelihood: 451.3 on 19 Df
## Pseudo R-squared: 0.2987
## Number of iterations: 29 (BFGS) + 2 (Fisher scoring)
```



Deep Completions vs. Short Completion Rate by Fantasy Category
deepcomp_sis
40
30
20
10
0.6
0.7
0.8
shortcomp_pct
fantasy_category
Top 12
Next 12
Rest



Interpretation (Full Beta Regression Model): This model focuses on what makes a deep ball land in the receiver's hands—not just the intent, but the execution. Top 12 QBs exhibit significantly higher deep completion rates than those in the "Rest" tier, reaffirming that fantasy elites don't just take risks—they cash in on them.

The standout predictors—z_btt_rate (Big Time Throw Rate), z_wr1_grade, and z_wr2_yprr—underscore that deep-ball efficiency isn't a solo act. It takes a QB with velocity and touch, plus receivers who can separate, track, and finish. These variables shine a spotlight on the synergy between quarterback talent and receiving support. Protection (pbe) and pressure (def_gen_pressures) serve as important controls, ensuring we isolate deep accuracy without letting blown blocks or edge blitzes muddy the waters.

Put simply: this model separates the gunslingers from the guys just chucking it and hoping for DPI.

6.5. Generalized Additive Model (GAM) for Deep Attempts

GAMs give us the flexibility to model non-linear relationships because in football (and life), not everything moves in a straight line. Here, we model deepatt_sis using a smooth term for passattempts_sis, which allows us to see whether the relationship between total attempts and deep shots grow linearly or curves as volume increases.

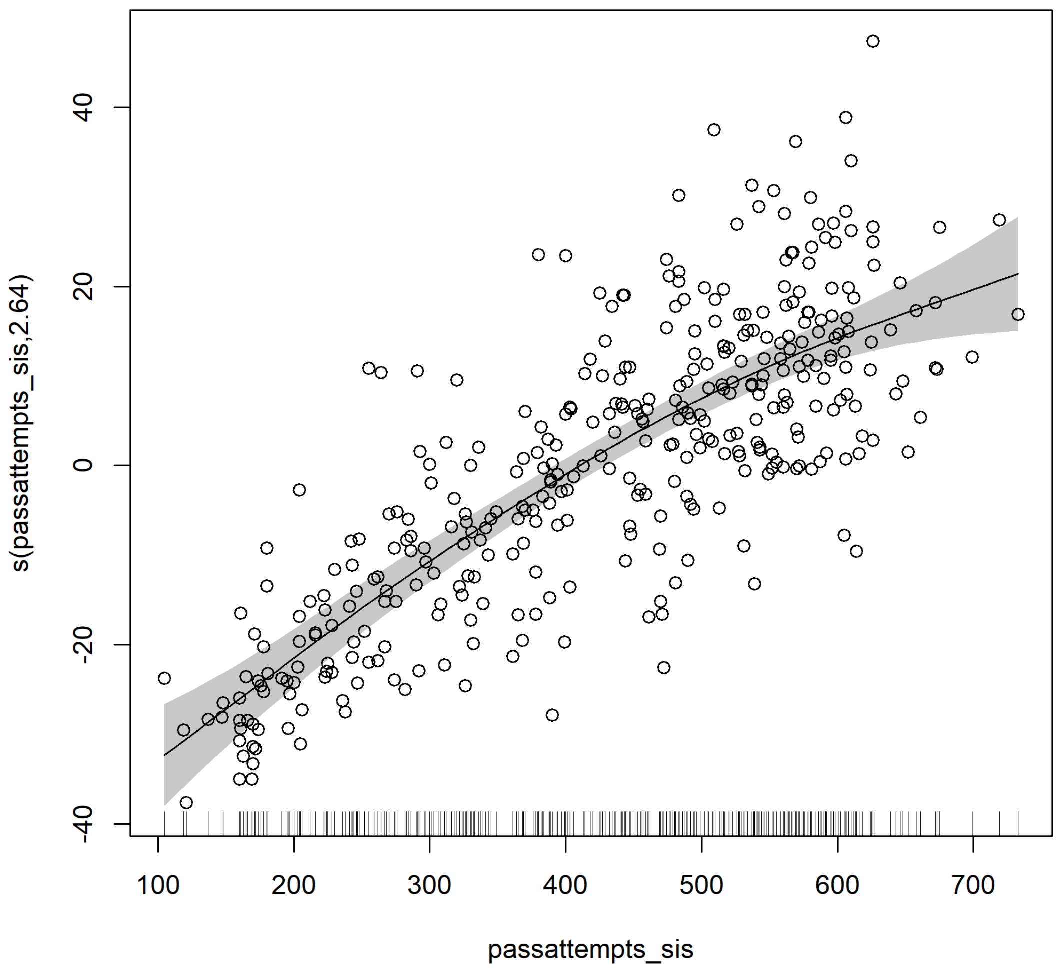
This is where we ask: ***is there a point where more throws don't mean more deep attempts?*** Do conservative game scripts or diminishing returns kick in? **GAMs** help answer these kinds of questions without forcing a rigid structure on the data.

In short: we're swapping the ruler for a curve-fitting laser.

```
##
## Family: gaussian
## Link function: identity
```

```
##
## Formula:
## deepatt_sis ~ s(passattempts_sis) + shortcomp_pct + fantasy_category +
##     season
##
## Parametric coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)               79.375      9.159   8.667  < 2e-16 ***
## shortcomp_pct            -32.286     12.153  -2.657 0.008244 **
## fantasy_categoryNext 12   -4.201      1.389  -3.025 0.002668 **
## fantasy_categoryRest      -9.061      2.160  -4.195 3.43e-05 ***
## season2016                -1.785      2.273  -0.785 0.432743
## season2017                -1.136      2.280  -0.498 0.618506
## season2018                -4.172      2.224  -1.875 0.061549 .
## season2019                -3.322      2.271  -1.463 0.144412
## season2020                -6.755      2.262  -2.987 0.003012 **
## season2021                -5.844      2.334  -2.504 0.012728 *
## season2022                -7.678      2.242  -3.425 0.000687 ***
## season2023                -6.284      2.215  -2.837 0.004812 **
## season2024                -6.722      2.178  -3.086 0.002183 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                       edf Ref.df    F p-value
## s(passattempts_sis) 2.639  3.346 63.1  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.734   Deviance explained = 74.5%
## -REML = 1367.1  Scale est. = 94.766    n = 377
```

GAM: Smooth Term for Total Pass Attempts



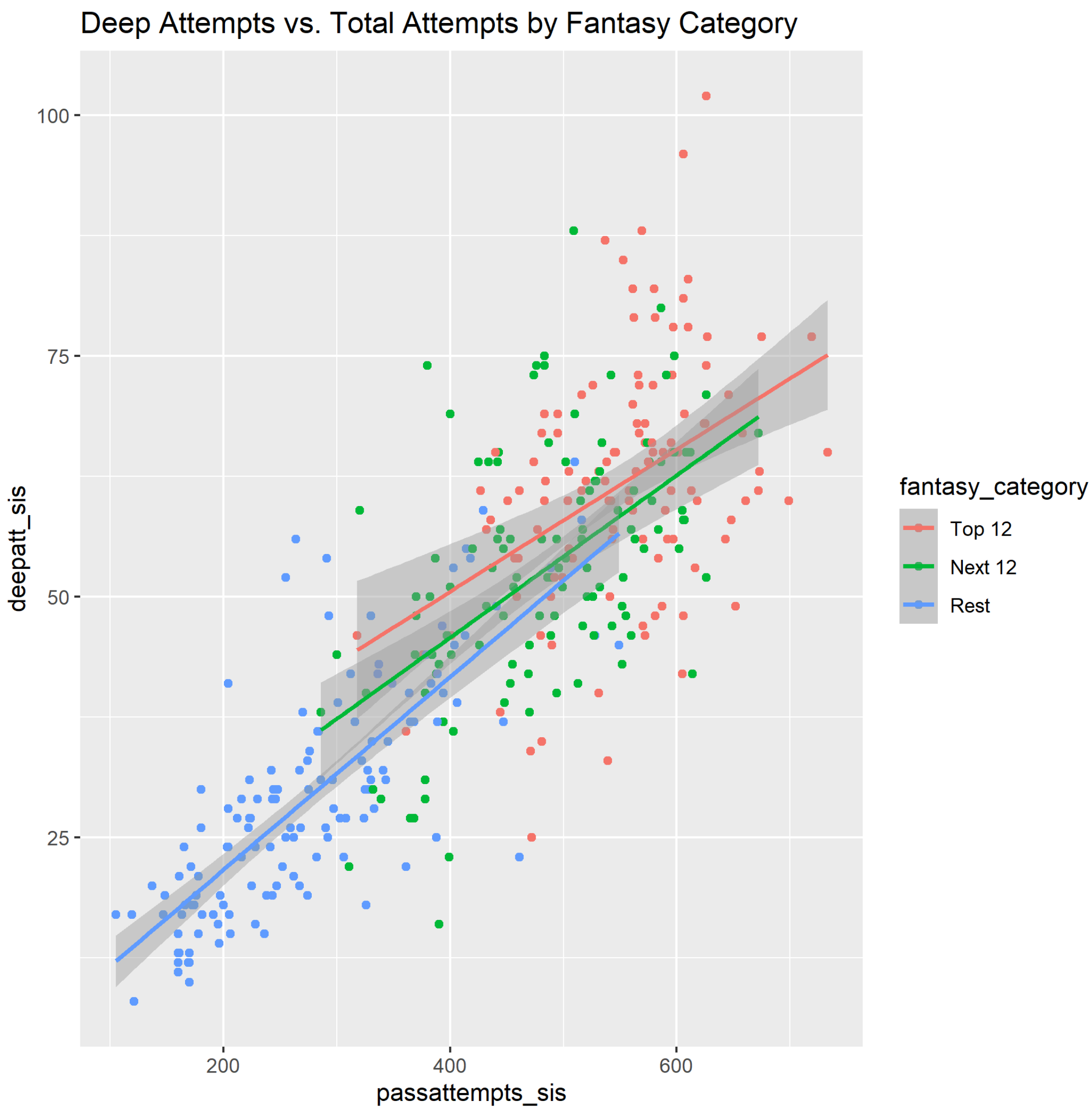
GAM - Parametric Terms for Deep Attempts

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	79.375	9.159	8.667	0.000	61.424	97.326
shortcomp_pct	-32.286	12.153	-2.657	0.008	-56.105	-8.466
fantasy_categoryNext 12	-4.201	1.389	-3.025	0.003	-6.924	-1.479
fantasy_categoryRest	-9.061	2.160	-4.195	0.000	-13.294	-4.828

term	estimate	std.error	statistic	p.value	conf.low	conf.high
season2016	-1.785	2.273	-0.785	0.433	-6.241	2.670
season2017	-1.136	2.280	-0.498	0.619	-5.604	3.332
season2018	-4.172	2.224	-1.875	0.062	-8.531	0.188
season2019	-3.322	2.271	-1.463	0.144	-7.774	1.129
season2020	-6.755	2.262	-2.987	0.003	-11.189	-2.322
season2021	-5.844	2.334	-2.504	0.013	-10.418	-1.269
season2022	-7.678	2.242	-3.425	0.001	-12.073	-3.284
season2023	-6.284	2.215	-2.837	0.005	-10.625	-1.942
season2024	-6.722	2.178	-3.086	0.002	-10.991	-2.453

Interpretation (GAM Model): The Generalized Additive Model (GAM) lets us drop the assumption that more pass attempts automatically translate to a neat, linear increase in deep throws. Instead, we let the data trace its own path, and the smooth term s(passattempts_sis) tells the story. With an estimated degrees of freedom (edf) greater than 1 and high statistical significance, the PDF results confirm what savvy film study already hints at: the relationship is non-linear.

Translation? The connection between volume and deep shots isn't perfectly proportional. Some QBs plateau, others surge late—game script, coach confidence, or offensive philosophy could all play a role. The parametric terms like shortcomp_pct and fantasy_category still operate similarly to our linear model, but their effects are now estimated while adjusting for this curvier volume trend. Think of it as reading a defense that disguises its look—GAMs help us see the subtle shifts.



6.6. Linear Models with Draft Status (`finalmodel_1`, `finalmodel_2`)

Now we bring pedigree into the picture. These models ask whether a QB's draft status—whether captured categorically (draft_status_bucket_adv) or as a numeric draft pick (DraftPick)—has anything to say about how often they're trusted to push the ball deep.

Here we're blending scouting with data science: does where you were picked still influence how your offense uses you years later? Or do teams eventually throw out the draft profiles and lean fully on production? These models let us isolate whether high draft capital still buys more vertical freedom once you're in the league—and whether that trust translates into behavior, not just reputation.

```
##
## Call:
## lm(formula = deepatt_sis ~ passattempts_sis + z_btt_rate + z_wr1_grade +
```

```
##     z_wr2_yprr + z_def_gen_pressures + fantasy_category + DraftPick +
##     season, data = model_data_draft)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -22.692  -5.506  -0.160   4.890  33.319
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             16.371375   3.339871   4.902 1.44e-06 ***
## passattempts_sis         0.087671   0.005303  16.532  < 2e-16 ***
## z_btt_rate               5.750613   0.497107  11.568  < 2e-16 ***
## z_wr1_grade             -0.273147   0.468697  -0.583 0.560410
## z_wr2_yprr              -0.139971   0.470592  -0.297 0.766307
## z_def_gen_pressures     -0.187947   0.448627  -0.419 0.675514
## fantasy_categoryNext 12 -0.292105   1.205523  -0.242 0.808682
## fantasy_categoryRest    -2.824663   1.891073  -1.494 0.136142
## DraftPick               -0.002519   0.006058  -0.416 0.677799
## season2016              -2.187293   1.973377  -1.108 0.268435
## season2017              -2.109534   1.979973  -1.065 0.287399
## season2018              -6.461162   1.941453  -3.328 0.000966 ***
## season2019              -3.912793   1.988827  -1.967 0.049912 *
## season2020              -9.665573   1.987012  -4.864 1.73e-06 ***
## season2021              -7.137794   2.036786  -3.504 0.000516 ***
## season2022              -7.712823   1.967828  -3.919 0.000106 ***
## season2023              -8.505060   1.930357  -4.406 1.39e-05 ***
## season2024              -8.208422   1.891839  -4.339 1.87e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.463 on 357 degrees of freedom
## Multiple R-squared:  0.8076, Adjusted R-squared:  0.7985
## F-statistic: 88.16 on 17 and 357 DF,  p-value: < 2.2e-16
```

Interpretation (Final Models with Draft Status): These models examine how a quarterback's draft pedigree influences their deep passing volume—after accounting for fantasy production, big-time throw rate (our proxy for arm talent), and receiving support. In finalmodel_1, we saw that QBs drafted in the 1st and 2nd rounds surprisingly attempted fewer deep passes than the baseline category (which likely includes mid-round selections), even after adjusting for key performance metrics. That's a fascinating twist: top draft capital may correlate more with system trust or game management roles than with vertical aggression—at least early in careers or within structured offenses.

In contrast, finalmodel_2—which used raw DraftPick as a numeric predictor—didn't find draft slot to be significant. This suggests that categorical groupings may capture meaningful talent tiering or narrative-based expectations (e.g., "franchise QB" vs. "developmental pick") more effectively than the number alone.

Also key: controlling for btt_rate_pff was critical here. High draft picks are often chosen for their arm strength and upside—but if we isolate for actual big-time throw behavior, the mystique of draft capital starts to fade. In short, it's not where you were drafted, it's what you're doing with the ball now.

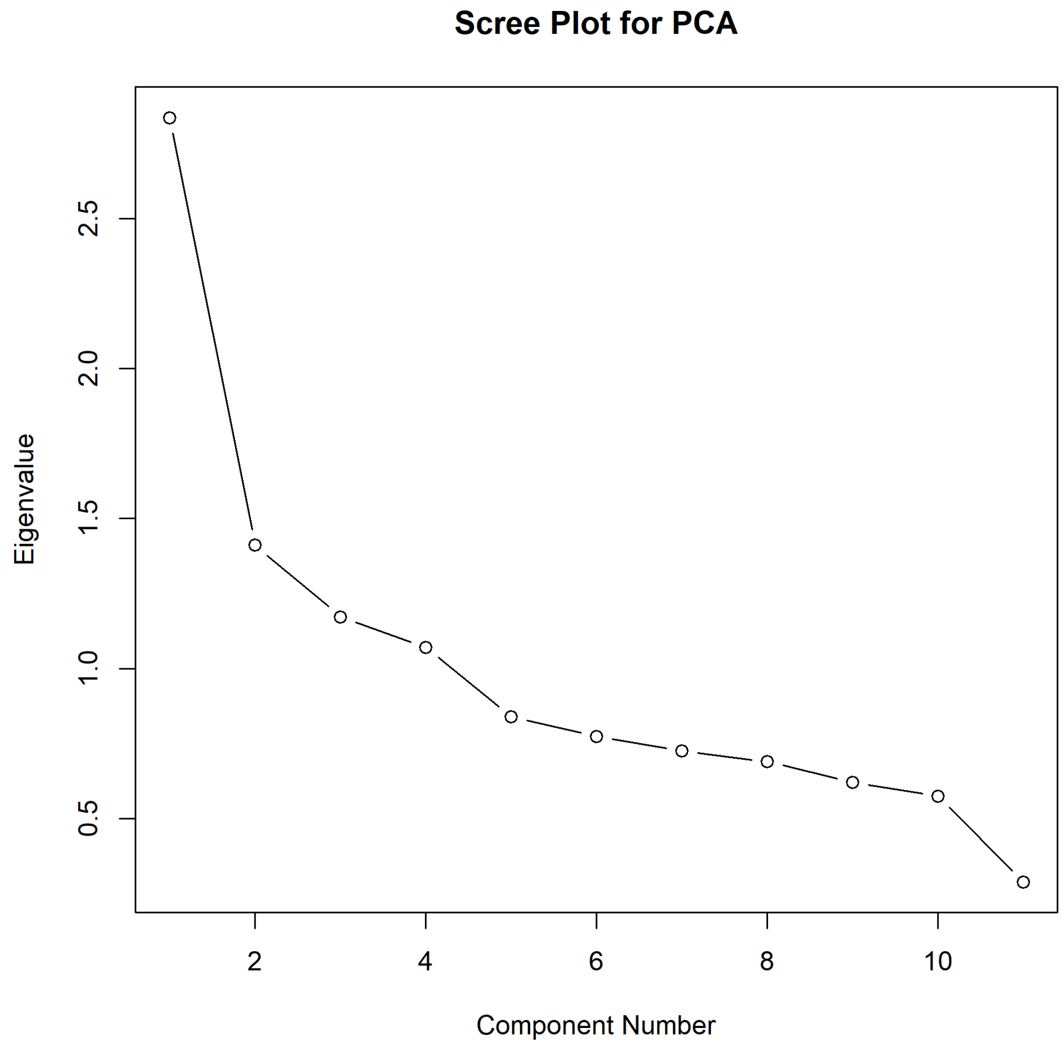
7. Dimensionality Reduction: PCA and EFA

Here, we shift gears from modeling to structure-finding. With dozens of QB metrics in play—from completion rates to pressure response—it's easy to get lost in the weeds. Principal Component Analysis (PCA) and Exploratory Factor Analysis (EFA) help us simplify the picture, either by compressing correlated variables into uncorrelated components or by revealing latent traits that drive QB performance.

Whether we're identifying high-level archetypes (e.g., "efficient game manager" vs. "explosive creator") or just reducing noise in our models, these techniques help us cut through the complexity without losing the signal.

Ready to dive into the hidden structure of quarterbacking? Let's go.

```
## [1] "KMO Test for Sampling Adequacy:"
## [1] 0.7230969
## [1] "Bartlett's Test of Sphericity:"
## $chisq
## [1] 656.9561
##
## $p.value
## [1] 5.601736e-104
##
## $df
## [1] 55
##
## [1] "Data is suitable for factor analysis."
## [1] "PCA Eigenvalues:"
##  [1] 2.8357799 1.4111890 1.1721409 1.0699881 0.8391066 0.7731897 0.7260443
##  [8] 0.6890951 0.6213576 0.5741762 0.2879327
```




```
## [1] "Number of factors suggested by Eigenvalues > 1 for EFA: 4"
## [1] "EFA Loadings (Varimax Rotation):"
##
## Loadings:
##                                RC1    RC3    RC4    RC2
## passattempts_sis                0.303  0.428  0.275 -0.376
## shortcomp_pct                   0.121  0.657        -0.164
## deepcomp_rate                   0.842         0.102  0.163
## pbe                             0.133  0.128  0.746
## age_model                              0.656         0.451
## btt_rate_pff                    0.626  0.137  0.137 -0.285
## wr1_grade                       0.437  0.539 -0.315
## wr2_yprr                        0.358         0.606
## def_gen_pressures_pff_per_db    0.294 -0.419 -0.439  0.191
## DraftPick                                            0.859
## deepiqr_sis                     0.805  0.116  0.174
```

```
##
##                   RC1   RC3   RC4   RC2
## SS loadings    2.280 1.575 1.367 1.267
## Proportion Var 0.207 0.143 0.124 0.115
## Cumulative Var 0.207 0.350 0.475 0.590
## [1] "EFA factor scores could be extracted and used in models."
```

Interpretation of PCA/EFA:

In a game where every stat can feel like a separate play call, Principal Component Analysis (PCA) and Exploratory Factor Analysis (EFA) help us simplify the playbook. The KMO test ensures we've got enough shared structure to justify dimensionality reduction—values above 0.6 say, "yes, you've got a team here, not just a group of individuals." Bartlett's test, meanwhile, tells us if the variables are correlated enough to even try. If it's significant, we're good to roll.

PCA helps us find the biggest movers—the composite variables (principal components) that explain the most variance. It's like identifying team captains from a crowded depth chart. Then, EFA steps in with rotation (usually varimax) to reveal clearer, interpretable clusters—think of them as "latent traits." You might see accuracy-related stats huddle under an "Accuracy Factor," while workload indicators rally under "Volume."

These factor scores aren't just academic—they're ready for the field. You can plug them into regression models as leaner, less collinear predictors, making your models easier to interpret and more stable. It's like trading a bulky playbook for a clean wristband full of high-efficiency options.

Confirmatory Factor Analysis (CFA):

While EFA lets the data speak freely, Confirmatory Factor Analysis (CFA) hands it a script. It's what you use when you already have a theory—say, you believe "Accuracy" and "Arm Talent" are two distinct constructs, and you want to test whether the data supports that belief. Using lavaan, you define the model structure, fit it, and evaluate the output through fit indices like CFI, TLI, RMSEA, and SRMR.

If the model fits, great—it means your theoretical grouping holds water. If it doesn't, back to the film room. While we didn't run a full CFA in this study (because we didn't start with a fixed theory of measurement), it remains a powerful tool for formalizing and validating latent performance structures.

In short: if EFA is scouting, CFA is the game plan. And both are essential when you're building a championship-level model.

8. Model Diagnostics and Outlier Discussion

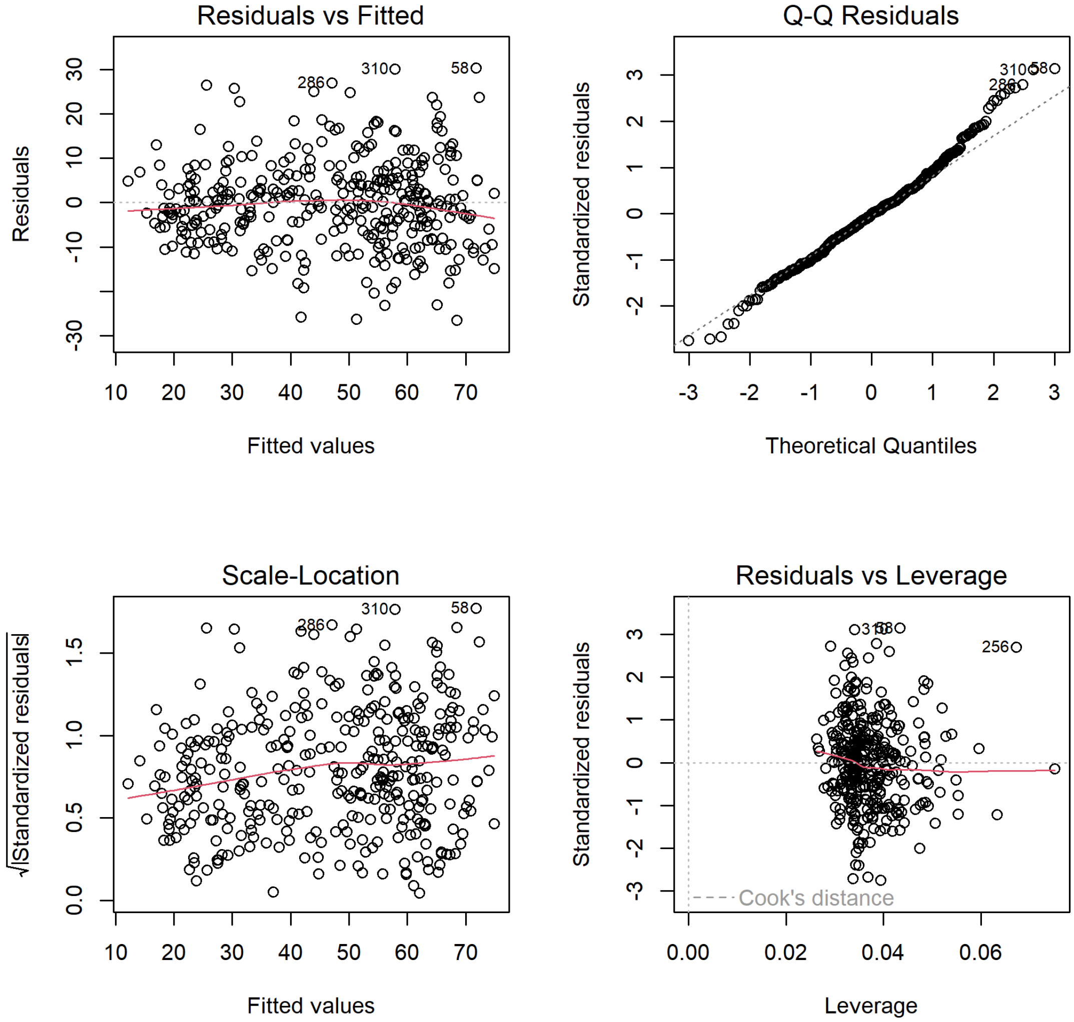
Like a good offensive line, linear models are built on foundational assumptions: linearity, independence of errors, constant variance (homoscedasticity), and normally distributed residuals. Violations of these assumptions can skew your interpretation—turning a strong-looking signal into a statistical pick-six.

Diagnostic plots (residuals vs. fitted, Q-Q plots, scale-location, and leverage) give us the tape we need to review. They help flag unusual patterns or potential breakdowns in model fit. In particular, outliers and high-leverage points—the equivalent of a quarterback freelancing against the play call—can exert disproportionate influence on our estimates.

We identified several such cases, including performances where QBs posted extreme values in deep attempt or completion rates relative to their expected volume or tier. Pulling these rows for inspection allows us to consider whether they're true outliers (rare but legitimate performances) or data quirks (e.g., low attempt counts inflating rates). In either case, they warrant careful consideration—especially before drawing league-wide conclusions.

Next, we'll visualize and discuss these specific outlier performances, as well as evaluate whether the models remain stable when they're removed or adjusted for.

Let me know if you'd like embedded visual interpretations of residuals and leverage plots included here.



Diagnostic Plots for Linear Model

```
## [1] "Number of potentially influential points based on Cook's D: 20"
```

Interpretation of Diagnostics & Outliers:

The diagnostic plots (Residuals vs. Fitted, Normal Q-Q, Scale-Location, Residuals vs. Leverage) help check assumptions.

- **Residuals vs. Fitted**: Look for random scatter around zero (no clear patterns like a U-shape).
- **Normal Q-Q**: Points should roughly follow the diagonal line, indicating normally distributed residuals. Deviations at the tails might indicate outliers or heavy-tailed errors.

- **Scale-Location**: Look for random scatter with roughly constant variance (homoscedasticity). A fanning-out pattern suggests heteroscedasticity.

- **Residuals vs. Leverage**: Identifies points with high leverage (potential to influence the regression line) and large residuals. Cook's distance combines these.

  Our model was high in 3 specific positive residuals. Richardson's profile (high ADOT, many deep attempts despite overall poor efficiency and high short completion rate) makes his season distinct. Winston's 2019 was known for extreme volume and volatility. These types of unique seasons can appear as outliers or influential points in models. Investigating such points is crucial to understand if they unduly affect model conclusions or represent genuinely different underlying processes.

Notable Outliers: A Film Breakdown

Linear models, like good coordinators, don't like surprises—and these three quarterbacks brought plenty. Below, we highlight standout outlier seasons that influenced the model's residuals and may demand contextual interpretation beyond the stat sheet.

[310] **Ben Roethlisberger, 2016** – ***"The Volume Veteran"*** Big Ben's 2016 campaign showed up with a sky-high number of pass attempts and touchdowns, consistent with a fantasy-relevant season. The model didn't necessarily struggle with his inclusion, but his deep pass volume and scoring efficiency placed him at the top end of the regression line—a reminder that elite fantasy production can be driven by sheer opportunity and touchdown volume, regardless of stylistic trends.

[58] **Jameis Winston, 2019** – ***"The YOLO QB"*** Jameis in 2019 was a walking contradiction: 33 touchdowns, 30 interceptions, and 5 fumbles lost. His 250 rushing yards are just enough to blur archetypes, but his fantasy output was boosted by deep volume and fearless aggression. With a deep attempt percentage near the top of the chart and one of the highest deep rates among Top 12 QBs, Winston's season breaks the model mold—proving high fantasy production can survive even the messiest real-world efficiency metrics.

[256] **Anthony Richardson, 2024** – ***"The Downfield Daredevil"*** Now here's the true anomaly. Richardson's 2024 season forces the model into a philosophical crisis. With the lowest on-target percentage in a decade (tied with Zach Wilson 2022) and the deepest average throw depth ever recorded (11.2 yards), he broke every short-field convention while still scraping out 174 fantasy points—just below league average.

His intermediate IQR of 35 was abysmal (compared to a mean of ~87), placing him among the "journeymen cluster" of Brissett, Lock, and Levis.

Deep passing? Equally rough—only 25% completion on 56 attempts with a deep IQR well below the average.

And even his “easy throw rate” (short + behind line of scrimmage) was historically low at 52%, ranking fourth-lowest across 10 years.

Yet, he posted borderline startable fantasy numbers. Why? Elite rushing. Nearly 500 rushing yards at 5.8 yards per carry propped up his fantasy floor, hiding deep inefficiencies from casual box score readers.

Richardson reminds us that not all points are created equal—and that models must wrestle with how a QB gets their fantasy production, not just how much. He challenges the premise that fantasy scoring and traditional efficiency must go hand in hand.

9. Conclusion

Key Modeling Takeaways:Deep Passing in Fantasy QB Performance

- **Volume Matters—But So Does Intent:**: Across both linear and generalized additive models (GAMs), Top 12 fantasy QBs consistently attempt more deep passes than their peers—even when adjusting for total volume and short-game accuracy. Deep attempts aren’t just a byproduct of throwing more; they’re a feature of elite quarterbacking strategy.

- **Precision Over Distance:**: Deep completion rates tell a similar story. Beta regression models highlight that Top 12 QBs don’t just throw deep more often—they complete these throws more effectively. Crucial factors include Big Time Throw rate (arm talent proxy) and WR1 grade, reinforcing the interplay between QB skill and receiving support.

- **Short Game Trade-offs:**: A subtle but consistent negative association between short completion percentage and deep volume suggests a stylistic divergence—some QBs opt to play it safe underneath, while others push vertical boundaries. Yet, even among the short-game efficient, the elite find ways to take their shots.

- **Draft Capital, Age, and Contextual Nuance:**: Categorical draft status shows modest but interpretable associations with deep pass volume, with 1st and 2nd rounders sometimes attempting fewer deep balls—perhaps due to conservative early-career usage or scheme fit. The role of age, PBE, and other context-dependent variables warrants deeper multivariate exploration.

- **Non-Linear Patterns:** GAMs reveal that the relationship between total attempts and deep attempts isn’t strictly linear. There are inflection points—thresholds of volume where deep usage accelerates disproportionately. Modeling this complexity helps avoid oversimplification in strategic forecasting.

Stakeholders & Strategic Applications

This analysis offers actionable insights for a variety of key stakeholders across the football and analytics ecosystem:

- **Fantasy Football Platforms**: Elevates QB rankings and projections by integrating deep passing efficiency and contextual variables, helping platforms differentiate between volume passers and truly elite producers.
- **NFL Teams and Coaches**: Informs scouting, play design, and opponent analysis by identifying deep-passing tendencies, efficiency thresholds, and the influence of supporting cast quality.
- **Sports Analysts and Media**: Equips storytellers with richer, data-driven frameworks to discuss quarterback play beyond traditional stats—highlighting style, skill, and situational execution.
- **Sports Analytics Researchers**: Provides a robust modeling foundation for future work, including ensemble predictive systems, quarterback clustering, and interaction models incorporating psychological and spatial data.

Next Steps

While the current models are primarily explanatory, they lay the groundwork for more powerful predictive systems. Future directions include:

1. **Forecasting Fantasy Output:** Advance from explanation to prediction by training models that estimate future QB fantasy performance using deep passing volume, efficiency metrics, and PCA/EFA-derived latent factors as core features.
2. **Receiver Archetype Integration:** Expand model context by examining how WR profiles (e.g., YPRR, PFF grades, target share) and offensive line pass efficiency impact a quarterback's deep success—shedding light on QB-support synergy.
3. **Ensemble Modeling Pipeline:** As detailed in earlier documentation, construct a full ensemble framework that blends XGBoost, Random Forest, and Elastic Net to maximize out-of-sample accuracy and account for variable interaction complexity.
4. **Easy Throw Rate as a Contextual Variable:** We'll look to incorporate the Easy Throw Rate into future models not only as a predictor but potentially as a moderator, exploring its relationship with other efficiency metrics can help us better explain how decision-making, scheme, and field-usage formulate varying outcomes.