NFL Red Zone Play Sheet

Summary

The Red Zone is the most critical (and most challenging) area of the field, where the outcome of plays often determines the difference between a win and a loss. Many teams struggle to consistently convert these opportunities into points due to the compressed space and tighter defensive coverage. To address this, we developed a framework that combines statistical and kinematic analysis of player positions with a powerful, user-interactive dashboard. This tool enables coaches and players to make split-second decisions by identifying optimal Red Zone play patterns from 5 to 20 yards from the end zone, including recommended alignments, routes, and acceleration based on predicted defensive coverage. Additionally, this dashboard allows coaches to identify projected defender positions around the targeted receiver for catch probability insights. By leveraging this approach, teams can increase Red Zone efficiency and improve the likelihood of success in the most crucial moments of the game.

Data Description, Exploration, and Preparation

The input, output, and supplementary data from every week were merged using game and play identifiers as join keys. To analyze catch behavior, we created catch-position columns by identifying the targeted receiver's location within a 1-yard radius of the ball's landing point, selecting the moment when the receiver was closest to the ball. We also calculated a defensive-distance metric, defined as the minimum Euclidean distance from the nearest defender to the receiver at the moment of the catch. For analysis purposes, defender distances were binned into intervals, allowing us to track catch rates across discrete distance ranges. Throws with a catch-time interval of less than 0.2 seconds were excluded, as such data is considered unreliable.
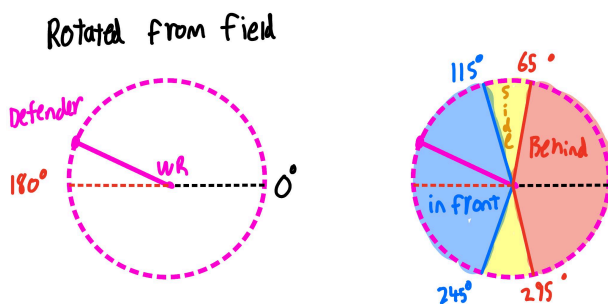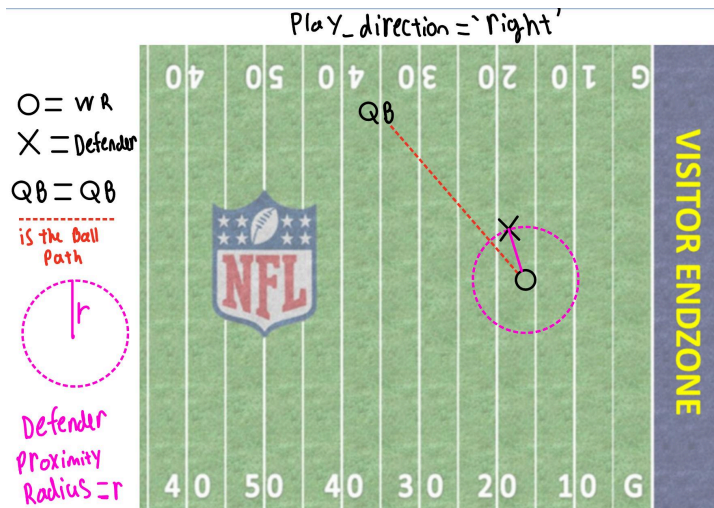
Red Zone plays were stratified into three intervals based on yardline distance from the end zone: 5–10 yards, 10–15 yards, and 15–20 yards. This grouping balances context and sample size: plays within the 5 are often run plays. A custom function mapped each play to its corresponding interval, yielding a filtered dataset of 1,982 unique Red Zone plays for analysis. For each unique play, we grouped the receiver formation and alignment, route type, quarterback snap position, and defensive coverage into distinct categories to perform statistical analysis. We also created flexibility with each receiver position by binning receiver alignment within the standard deviation across successful plays using the same formation and route combination.

Touchdown identification was performed using text pattern matching on play descriptions, searching for the keyword "TOUCHDOWN". Across 2,692 red zone plays analyzed, 318 resulted in touchdowns, yielding an overall baseline success rate of 11.8%. This baseline rate served as the prior probability to reference in later Bayesian simulation models.

To calculate receiver acceleration patterns we computed frame-by-frame displacement on consecutive x-y positions, then derived instantaneous speed by dividing displacement by the 0.1 second frame interval. Acceleration was calculated by computing the rate of change of speed. To create an acceleration effort percentage, we divided by a realistic red zone top acceleration of 2.5 yards / sec².

Finally, we wanted to classify the defender position relative to the targeted receiver and the ball path at the catch. Using the ball path as a reference axis, we classified defender position by transforming the coordinate system so that the receiver was centered at the origin. Defender locations were expressed in this receiver-centric frame and mapped onto a unit-circle–like representation. We then computed the defender's angular position relative to the ball path using the `arctan2(y, x)` function in Python, where x and y denote the defender's

coordinates relative to the receiver. This angle was used to categorize the defender as being 'in front', 'behind', or to the 'side' of the receiver. These categories were added as a new column, 'd_pos', to the dataframe. An abstraction of the field is shown below with an example of 'd_pos' categorization below.
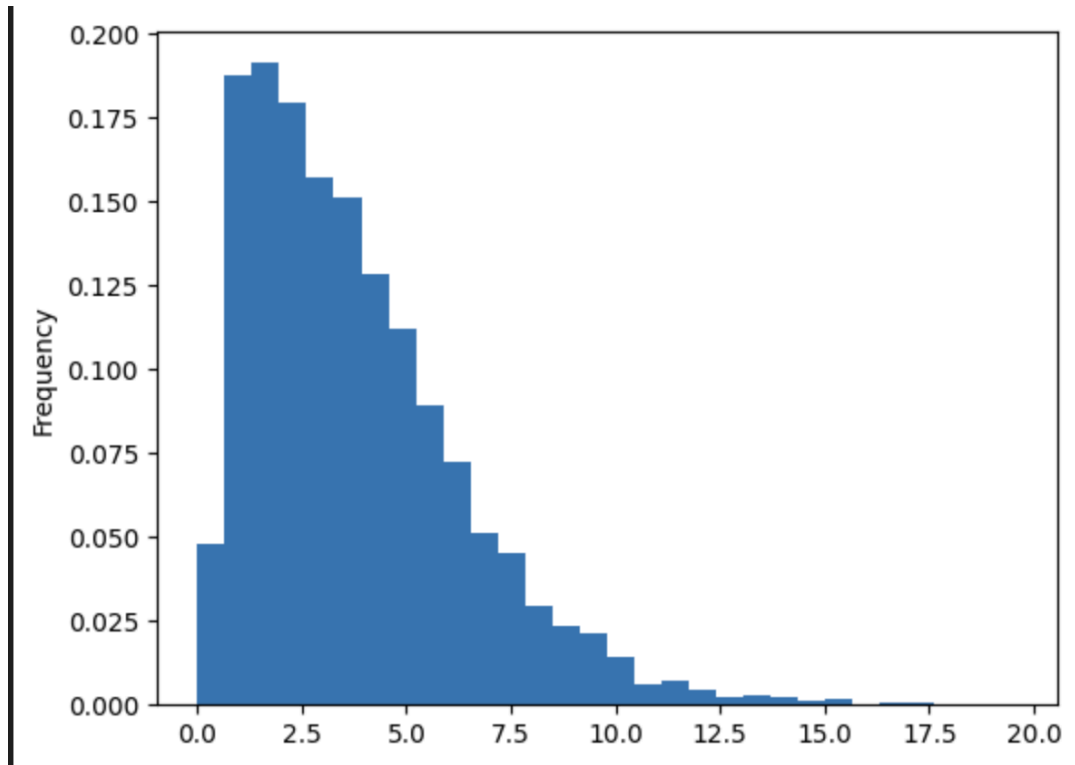

In this example the defender would be labeled as 'in front' in the 'd_pos' column because the defender falls within the bounds of 115 to 245 degrees relative to the receiver and ball.

Key Statistical Findings at (Micro Receiver-Level & Macro Play-Level)
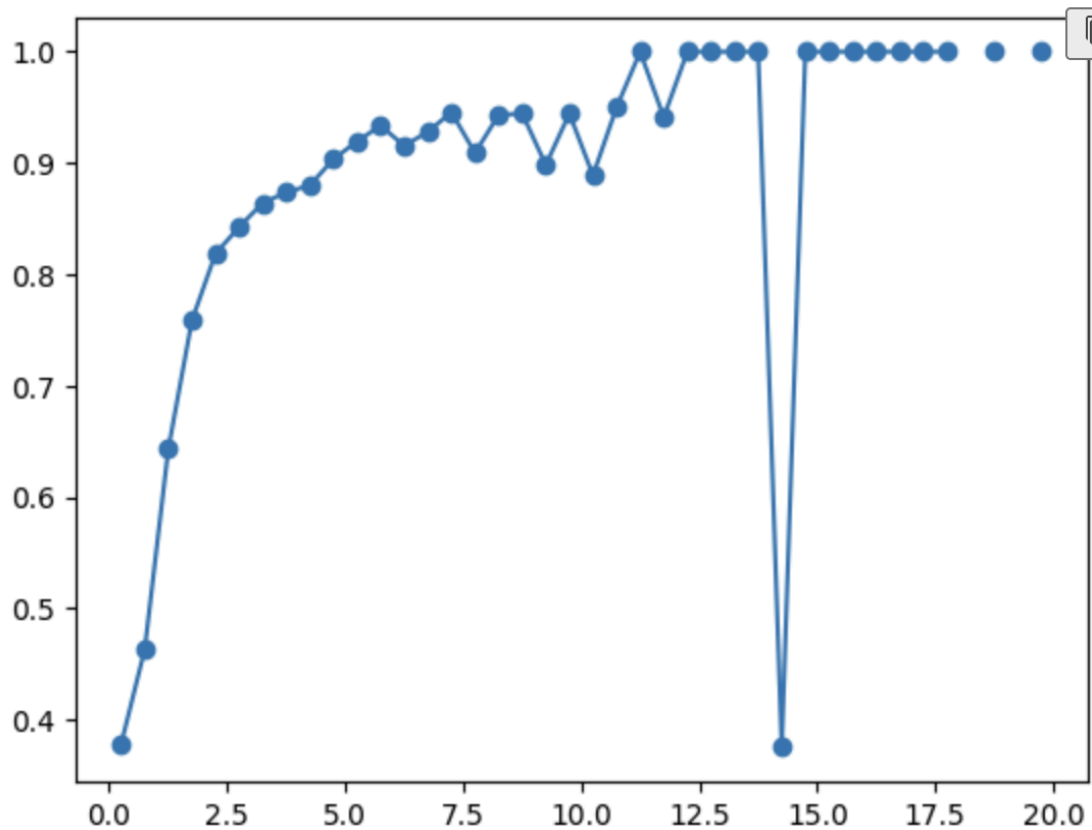
Micro Level

We first plotted the distribution of defender-to-receiver distance at the moment of the catch for every target.

This visualization is informative in illustrating that catches are frequently made with minimal separation; however, it does not quantify the relationship between defender proximity and the likelihood of a completed catch.

To address this, we discretized defender distance into 0.5-yard bins (e.g., [0.0, 0.5), [0.5, 1.0), …). For each bin, we then calculated the proportion of targets that resulted in completed catches. In the resulting plot, defender distance is shown on the x-axis, while the completion probability (catch rate) is shown on the y-axis.
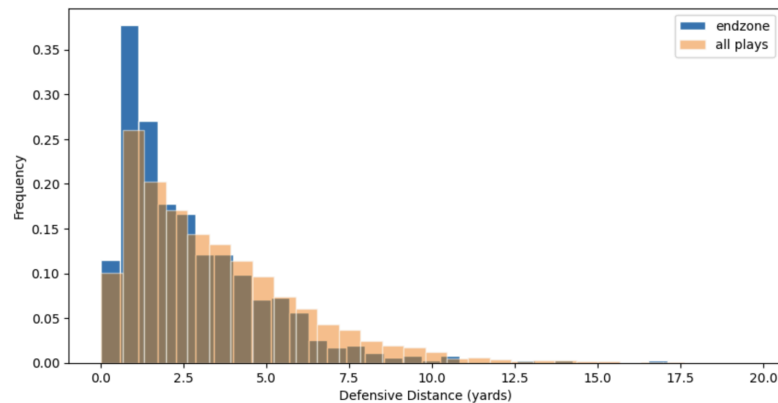
As expected, the likelihood of a completed catch increases as the distance from the nearest defender grows. However, the defender-distance distribution is heavily right-skewed. As a result, observations beyond about 6 yards—such as the outlier near 14.0 yards—are rare and therefore unreliable for statistical inference.

Despite this limitation at larger separations, the data reveals a clear logistic relationship between defender proximity and catch rate, where sample sizes are substantial (around 0.5 - 6.0 yds in separation). In this range, marginal increases in separation produce large gains in catch probability, whereas additional separation beyond this range yields imperceptible differences.

We next aimed to test if the average defender distance is lower in the Red Zone. If significantly less, this would provide a quantitative explanation for why offenses can move the

ball effectively between the 20s but struggle near the goal line. Reduced separation in the Red Zone would push plays into the **steepest portion of the logistic curve**, where even small losses in defender separation result in disproportionately large declines in catch probability. (Red Zone v.s All-Plays distribution for defensive distance is shown below.)

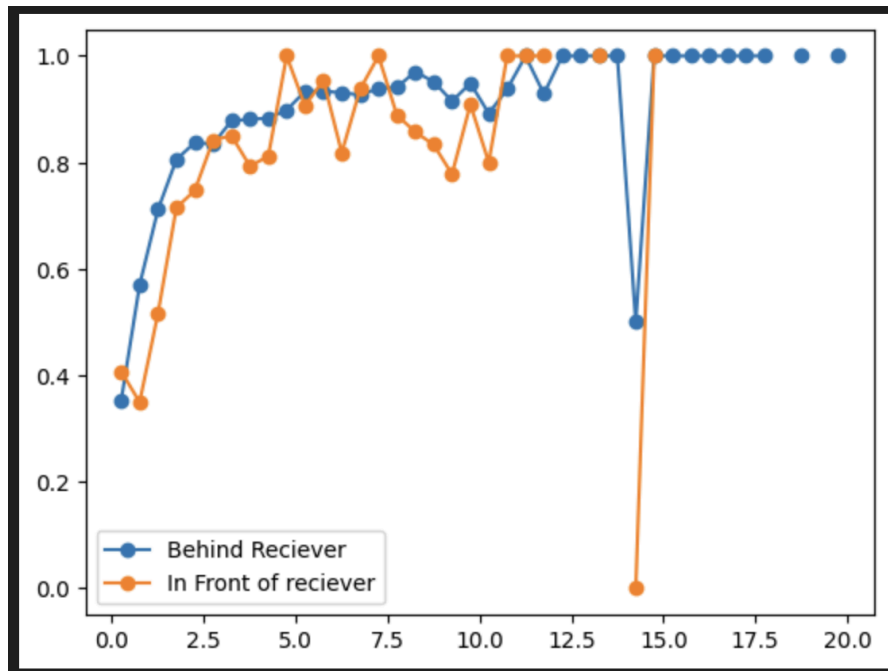There is a pronounced difference in skewness between the two distributions: the Red Zone ("endzone") distribution is substantially more right-skewed, indicating that defenders are typically closer to the targeted receiver near the goal line. To determine whether this shift could be attributed to random variation, we conducted a hypothesis test using 1,000 trials, comparing median defender distance in the Red Zone to all plays.
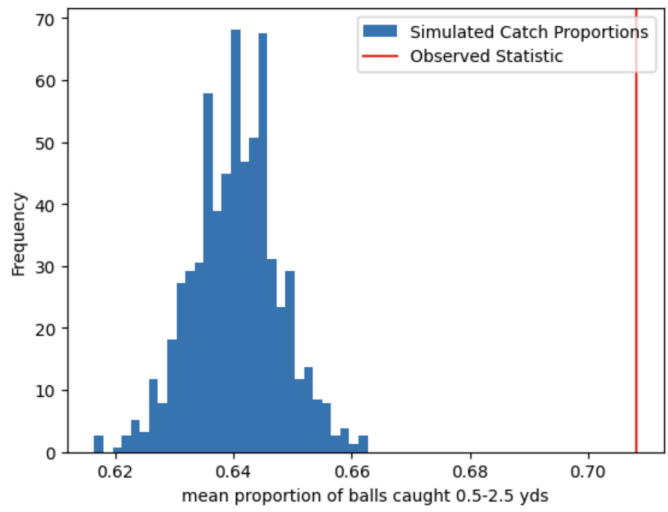
The resulting p-value was zero, indicating that the observed difference in separation is extraordinarily unlikely to be due to random chance. In practical terms, defender separation is consistently smaller in the Red Zone, with a greater concentration of defenders positioned closer to the targeted receiver compared to plays in the rest of the field. Crucially, the observed median defender distance in the Red Zone falls within the 0.5–2.5 yard range—the steepest region of the catch-rate curve. As a result, a disproportionately large share of Red Zone targets occur under conditions that inherently suppress catch rates due to extreme defensive proximity.

Having established that end-zone targets occur at substantially lower separation, we next examined whether the defender's position relative to the ball and receiver (in front of vs behind) further influences catch rate under these constrained conditions, using the previously defined `d_pos` variable.
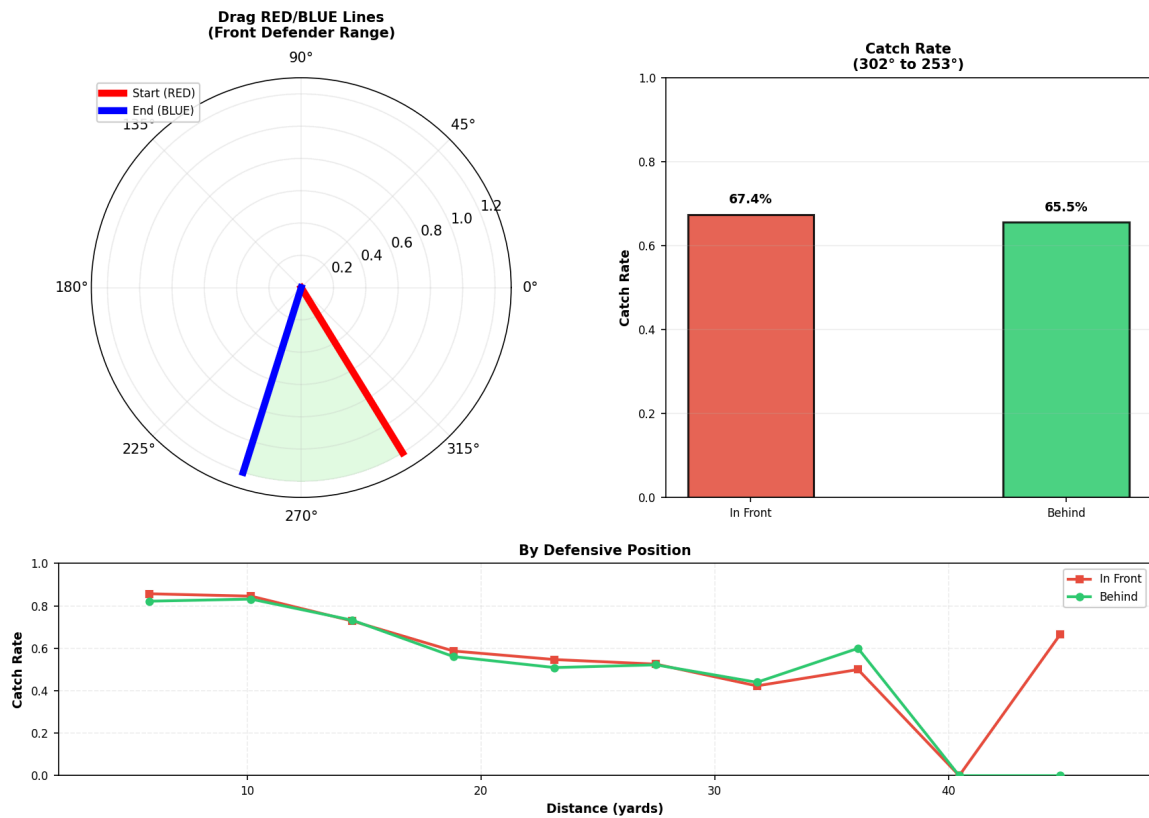
In the range where data is most abundant and statistical inference is most reliable (0.5 -- 2.5 yards), a clear pattern emerges: targets with the defender positioned **behind the receiver** exhibit perceptively higher catch rates. If statistically significant, this distinction has direct implications for Red Zone play design, where average defender proximity is smallest. Because catch probability is lowest at minimal separation (most plays occur in this regime) any advantage in this range affects the greatest number of plays in the most consequential area of the field.

To formally test this effect, we calculated the mean catch rate for targets where the defender was positioned behind the receiver and located between 0.50 and 2.50 yards away at the moment of the catch. This catch rate was 70.8% on average. We then conducted a permutation test to evaluate whether this observed advantage could plausibly arise by random chance, thereby assessing the statistical significance of defender position at low separation distances.

Across the 1,000 permutation simulations, none produced a mean catch rate that approached the observed value when the defender was positioned behind the receiver. Instead, the null distribution was centered at approximately 0.64 (64%) for the same separation interval. This result provides strong evidence that defender positioning is not random in its effect: when a receiver successfully places the defender behind him at low separation distances (0.5 -- 2.5 yards), the expected catch rate increases by roughly 10 percentage points on average.

To give use to our findings for coaches and players in a game time situation we developed a defender position vs catch rate probability interactive dashboard.

Drag RED/BLUE Lines (Front Defender Range) / **Catch Rate (302° to 253°)** / **By Defensive Position**

This dashboard enables coaches to determine optimal player routes based on defender proximity and the predicted defender angle. Using a unit-circle–style representation, the coach sets the defender angle within the dashboard, and the system instantly displays the corresponding catch probability in the adjacent bar chart. Additionally, if the coach is uncertain about exact defender distance but knows the defender's relative position to the ball and targeted receiver, the dashboard provides a guide for receiver positioning along the route.
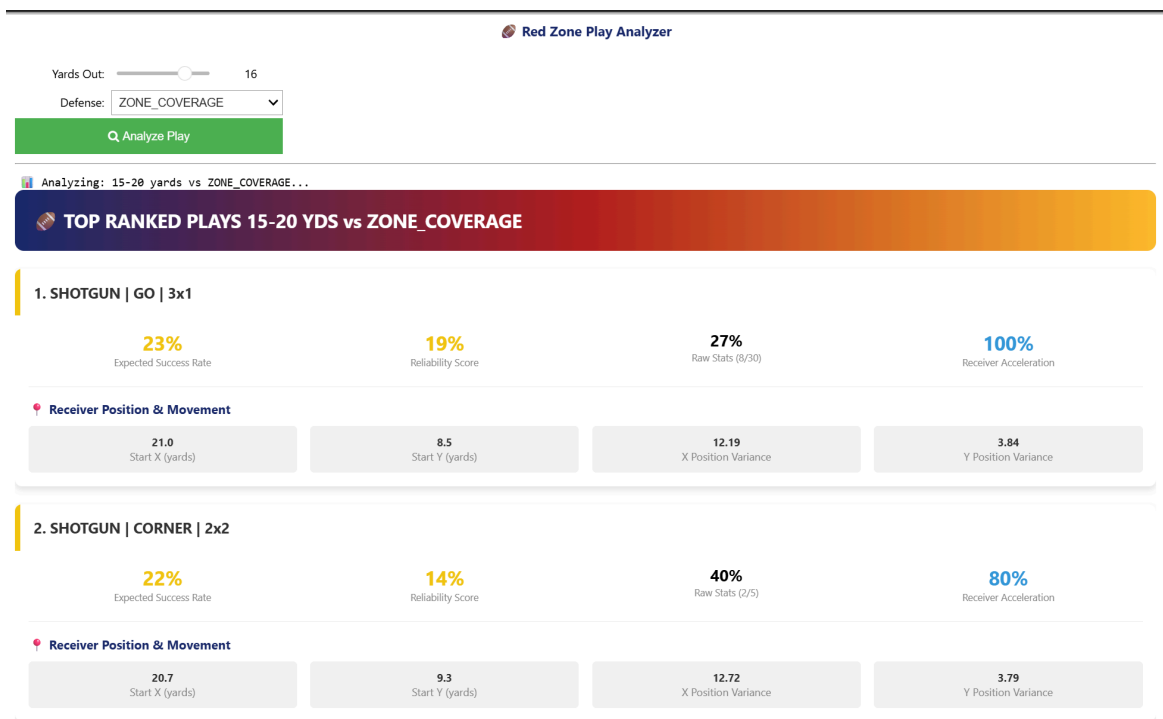
Macro Level

Since using only the data's success rates for each play gives a small sample size and uncertainty in our results, our analysis implemented a Bayesian simulation approach using Beta distributions to simulate more trials and gain a better understanding of the reliability of these

plays in the red zone. For each play formation, the model combined observed data (touchdowns and attempts) with a prior distribution reflecting overall red zone performance.

In Bayesian Statistics, a prior represents an initial estimate before statistically analyzing the data, the prior was constructed with alpha (Significance Level / Type I Error / false positive) and beta (Type II error / false negative) parameters derived from the global average touchdown rate. We then augmented it with the observed successes and failures to create a posterior distribution, which is an updated distribution after analysis. We utilized Monte Carlo simulation to draw 10,000 samples from this posterior Beta distribution, using the 25th percentile of simulated rates (lower bound to account for overconfidence).

Pattern Recognition and Recommendation Engine

Our special recommendation engine works by filtering plays by red zone interval and defensive coverage type, then grouping plays by the combination of offensive formation, receiver route, and alignment. For patterns with at least one attempt, the system calculated both raw and Bayesian-adjusted success rates.

🏈 **Red Zone Play Analyzer**

Yards Out: ————◯———— 16

Defense: [ZONE_COVERAGE ▼]

[🔍 Analyze Play]

📊 Analyzing: 15-20 yards vs ZONE_COVERAGE...

🏈 TOP RANKED PLAYS 15-20 YDS vs ZONE_COVERAGE

1. SHOTGUN | GO | 3x1

23%	**19%**	**27%**	**100%**
Expected Success Rate	Reliability Score	Raw Stats (8/30)	Receiver Acceleration

📍 **Receiver Position & Movement**

21.0	8.5	12.19	3.84
Start X (yards)	Start Y (yards)	X Position Variance	Y Position Variance

2. SHOTGUN | CORNER | 2x2

22%	**14%**	**40%**	**80%**
Expected Success Rate	Reliability Score	Raw Stats (2/5)	Receiver Acceleration

📍 **Receiver Position & Movement**

20.7	9.3	12.72	3.79
Start X (yards)	Start Y (yards)	X Position Variance	Y Position Variance

The analysis revealed that red zone success rates varied substantially based on tactical combinations. Certain formation-route-alignment patterns showed reliability scores exceeding 30%, nearly triple the baseline rate, while others fell below 10%. Defensive coverage type dramatically influenced optimal play selection, with different patterns excelling against man versus zone coverages.

Acceleration effort metrics indicated that successful red zone plays required moderate but not full acceleration once the ball was snapped. This implies that a red zone touchdown is more likely when receivers reach full effort in speed later in the route rather than at the start. Starting position analysis showed that successful plays featured relatively low positional flexibility, suggesting that precise alignment is key.

The Bayesian approach successfully balanced historical performance with statistical uncertainty, preventing overconfidence in plays that weren't as commonly run while still suggesting formations that were proven to work with high reliability.

Impact

This analysis successfully transformed raw NFL tracking data into actionable coaching intelligence for red zone play calling. By combining spatial catch probability modeling with Bayesian statistical techniques and interactive visualization, the system provided coaches with reliable, data-driven play recommendations tailored to specific game situations.

Future enhancements could integrate quarterback performance metrics to account for passer accuracy, develop play sequencing recommendations to exploit defensive adjustments across drives, and expand beyond touchdown optimization to include success probabilities at any part of the field. The framework established here provides a robust foundation for ongoing NFL analytics innovation in a time where the power of data continues to increase.