Maximizing Play Success in the NFL: The Impact of Down, Distance, Formation and Field Position

Skylar Furey

***Abstract*—NFL and College Football teams constantly seek to maximize play effectiveness by optimizing offensive strategies against defensive formations. This study analyzes 2022 NFL play data to determine the success rates of various offensive formations against different defensive alignments and coverages based on field position. We visualized offensive and defensive formations, mapped pass attempts using color-coded indicators, and used box plots to display running play distributions. We defined the success rate of each situation as the number of plays that gained yards for the offense. We found plays ran on later downs and with less yards to gain for a first down were much more successful. Additionally, we found that plays ran from the I-Form offensive formation have the highest success rate. For defensive schemes, the Cover 1 scheme allowed the highest success rate against it. We also found that the closer a team was to the end zone, the higher the success rate was for their plays. These insights provide valuable decision-making tools for play selection in real-game scenarios.**

1 INTRODUCTION

Football is a strategic game where offensive and defensive play calling significantly impacts success. Offensive coordinators must evaluate whether running or passing plays will yield the yards needed to continue their drive, and if passing, determine the optimal target areas based on defensive coverage. Whereas defensive coordinators must use offensive player personnel to determine which coverage scheme to play. Understanding these factors allows coaches to refine game plans and improve decision-making. This study uses data-driven insights to explore which offensive formations and play types are most successful against varying defensive schemes and field positions.

2 RELATED WORK

Data analytics has become a crucial tool in improving decision-making and increasing play success rates in the NFL. Several studies have explored how statistical models, machine learning algorithms, and predictive analytics can optimize strategies for play calling, fourth-down decisions, and overall game flow.

The most researched topic focuses on the application of predictive modeling to decision-making during high-leverage situations, such as fourth-down conversions. These decisions are focused on due to their large impact on the game, many games are won or lost on a few fourth-down decisions. Blinkoff et al. [1] analyzed fourth-down decision-making in college football focusing on the Power 5 conferences using predictive models to enhance the likelihood of success. Their study demonstrates the utility of historical data in determining whether to attempt a fourth-down conversion based on opponent tendencies and game context. Similarly, Weller [2] compares fourth-down decision-making in college and NFL football, emphasizing how different field conditions and game scenarios influence the success rates of these high-risk plays.

In the context of offensive play calling, Lee, Chen, and Lakshman [3] developed a model to predict offensive play types in the NFL, showcasing how predictive techniques can help coaches select plays based on specific game situations. This approach allows for a data-driven strategy to increase the chances of success. Moreover, New [4] highlights how NFL teams have increasingly turned to data analytics to assist coaches in evaluating the effectiveness of offensive formations and making better decisions during games. By analyzing metrics like expected points added (EPA), teams can assess play outcomes and adjust their strategies in real-time.

Krishna [5] focused on the role of game flow analysis, examining how momentum shifts during a game can influence the success of subsequent plays. Their research shows that by tracking game flow using data, teams can better predict which plays are most likely to succeed based on the state of the game. In a similar vein, Aarhus, Simikic and Zuber [6] explore optimizing play calling based on factors such as time remaining and distance to the endzone, helping teams make informed decisions under pressure.

Another study by Husowitz, Mixer, and Morrow [7] further investigates the predictive power of data by exploring the use of machine learning in NFL data analytics, applying regression models to predict probabilities and player performance, classification models to play and game outcome predictions, and deep learning to play type and strategy assessments. They also introduce a novel passing statistic derived from machine learning algorithms, showcasing how advanced analytics can enhance quarterback evaluation and overall team strategy.

In addition, a recent study by Zhang et al. [8] presents an advanced machine learning model to analyze player performance and predict the success rate of individual plays. This research highlights the growing reliance on machine learning techniques to enhance play selection and game strategy.

In conclusion, data analytics plays a pivotal role in improving play success rates in the NFL by providing insights into play prediction, situational awareness, and

decision-making. By continuously refining these models and integrating diverse data sources, teams can further enhance their strategies and increase their chances of success on the field.

3 METHODOLOGY

We obtained data for this study from the 2022 NFL season via Kaggle [9]. The dataset was filtered to remove penalty plays and special teams plays, and some offensive formations were combined for simplification. The absolute yard line number was calculated to group plays based on their distance from the end zone regardless of if they were going towards the home or away end zone.

To streamline the representation, we excluded wildcat offensive formations, as well as certain defensive formations (Goal Line, Redzone, Bracket, Miscellaneous, and Prevent). These formations were removed due to the high variability in player alignment, which made it challenging to create a consistent representation.

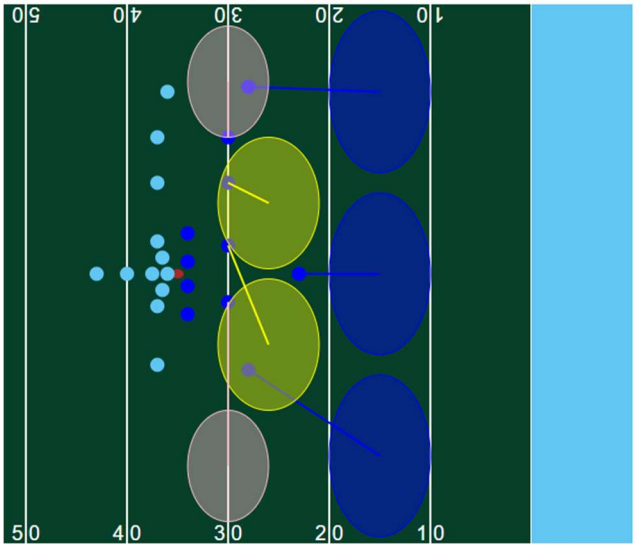

Fig.#1 Main Field Visualization

Figure 1 illustrates the main visualization displaying a football field with Carolina blue players representing the offensive team and Duke blue players representing the defensive team. It showcases the user-selected offensive formation (I-Form), receiver alignment (3x1), defensive scheme (Cover 3), and defensive coverage type (Zone).

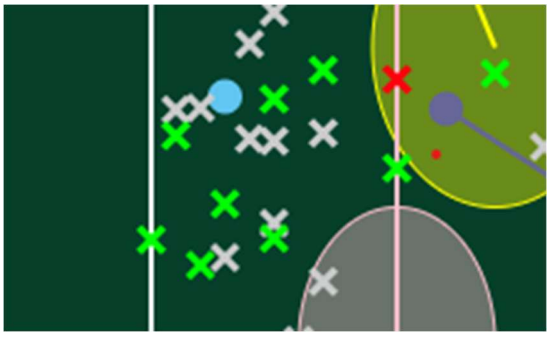

Fig.#2 Pass Attempts

Figure 2 shows how each pass type is represented. Completed passes are indicated by light green crosses, incomplete passes by gray crosses, and interceptions by red crosses. Each cross represents the location of the targeted receiver when the pass arrived.

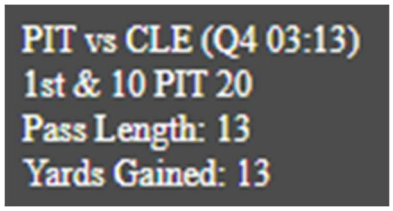

Fig.#3 Pass Attempts Tooltip

Each pass attempt is accompanied by a tooltip, as shown in Figure 3, that provides detailed information, including the teams involved, the quarter and time on the clock, down and distance, the true ball position, the pass length (calculated as the distance from the line of scrimmage to the receiver's location), and the yards gained (determined by the final location of the receiver after being tackled).

For offensive formations, we created simplified representations of player alignments, ensuring compliance with NFL rules on ineligible receivers. While these representations do not capture every real-world formation variant, similar formations were categorized under the same label. The most common variations, such as spacing between wide receivers or flipping the formation (e.g., a 3x1 formation with the majority of the receivers either on the left or right side of the ball), were combined.

Defensive formations were similarly simplified, with a standard alignment of four defensive linemen used for all formations. The number and placement of safeties were determined based on the defensive scheme. Each receiver was assigned a cornerback, and the remaining defenders were spread out as linebackers, positioned five yards past the line of scrimmage.

Ball position was represented in 10-yard increments. For instance, selecting **Own 25** displayed all plays run from the 20 to 29-yard line. To maintain accuracy, we adjusted pass attempt locations to account for the difference between the true line of scrimmage and the visualized line of scrimmage.

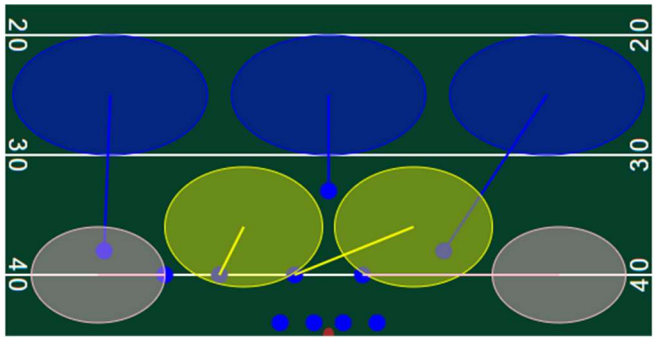

Fig.#4 Zone Coverage Illustration

Figure 4 illustrates how each zone defensive coverages are color-coded to indicate the type of coverage being employed. **Pink** zones represent flat coverage, designed to defend short throws to the sideline and stop outside run

plays. **Yellow** zones represent hook/curl coverage, aimed at defending short to medium passes over the middle of the field and supporting inside run defense. **Blue** zones represent deep coverage, intended to prevent passes beyond 15 yards; these defenders primarily focus on the pass and only engage in run defense when necessary.

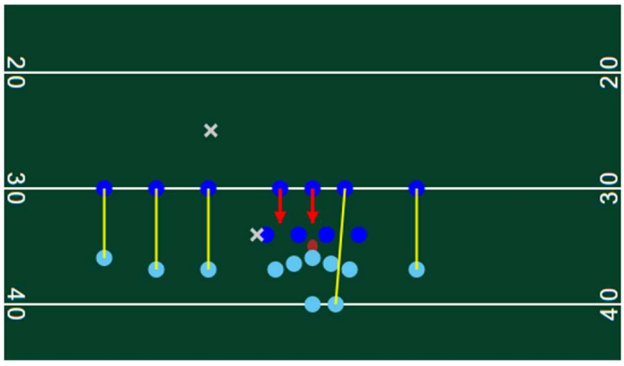

Fig.#5 Man Coverage Illustration

Figure 5 shows how the visualization appears when man coverage is selected, yellow lines indicate which defenders are assigned to cover each receiver. Receivers are guarded by their closest off-ball defenders (non-defensive lineman). This representation simplifies the real-world complexity of man coverage, where defenses often will have defenders from the outside blitz and interior defenders will cover their assumed man coverage responsibility. Employing blitzes from various players can confuse the offensive line and allow the defense to generate quick pressure on the quarterback. This leads to more incomplete passes, interceptions, or sacks, all of which are unsuccessful plays. In running situations, blitzing defenders often outnumber blockers, resulting in more tackles behind the line of scrimmage, also considered unsuccessful plays.

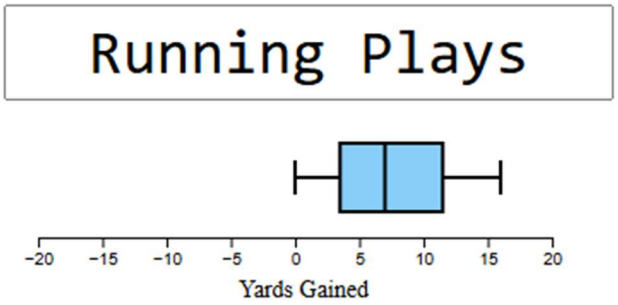

Fig.6 Running Plays Distribution

Figure 6 depicts the distribution of running plays presented as a box and whisker plot. The plot displays the range of yards gained and lost, with limits set to 20 yards gained and 20 yards lost. This design unfortunately removes extreme run plays, but by ensuring a consistent scale, comparison is easier across different situations.

Additionally, the number of total plays in each situation is displayed, allowing users to see the frequency with which each situation occurs. The success rate is also shown, representing the percentage of plays that result in the offensive team gaining yards. The success rate is calculated based on whether the play gains enough yards to improve the team's position for the next play.

Fig.7 Interactivity Options

Figure 7 shows the options that the user has to explore the data including selections for ball position, offensive formations and defensive coverages. These options allow the user to explore play success rates based on different selections.

4 RESULTS AND DISCUSSION

4.1 Play Success Rate by Down

An analysis of play success rates by down reveals a consistent decline across successive downs:

- 1st Down: 72.1%
- 2nd Down: 69.5%
- 3rd Down: 61.0%
- 4th Down: 47.3%

This trend shows how teams are able to be more balanced in early down situations but are often forced to pass in late down situations. On first down, teams choose to run the ball 56% of the time compared to 33.5% and 36% on third and fourth down respectively. The balanced attack keeps the defense guessing and allows the offense to be more successful.

4.2 Impact of Yards to First Down on Success

The success rate is comparable for all yards to go for a first down except for between eleven and twenty yards:

- 0-3 Yards: 66.7%
- 4-7 Yards: 67.0%
- 8-10 Yards: 70.2%
- 11-20 Yards: 63.4%
- 21+ Yards: 71.6%

At first glance, the declining success rates across downs might suggest that losing yards on early downs could somehow be advantageous. However, this interpretation is counterintuitive and misleading. The likelihood of achieving a first down on third or fourth down is significantly higher when the offense needs 7 yards or fewer—53.0% on third down and 51.2% on fourth down—compared to just 23.6% and 12.0%, respectively, when 8 or more yards are needed. Offenses also tend to run the ball more frequently in short-yardage situations, opting to run approximately 40% of the time when 7 or fewer yards are needed, versus only 20% when facing 8 or more yards. In longer-yardage scenarios, defenses are better positioned to anticipate passing plays, allowing them to allocate more coverage to receivers and limit deep passing opportunities.

4.3 Effectiveness of Offensive Formations

The effectiveness of different offensive formations was evaluated by analyzing their success rates. The findings indicate that balanced formations that are structured to allow for both successful run and pass plays are superior:

- Most Effective Formation:
  - I-Form: 75.8%
- Less Effective Formations:
  - Singleback: 73.6%
  - Pistol: 72.2%
  - Shotgun: 66.1%
  - Empty: 61.2%

The Shotgun, Pistol and Singleback formations all have similar personnels allowing for a variety of alignments. We thought that these formations would excel compared to the more specialized formations, but the contrary is true. With the I-Form formation the offensive players are more condensed, allowing for more room to run routes as well as more blockers at the point of attack when running. The Empty formation is the complete opposite with no running back and receivers split out wide across the field to allow for passing options in each part of the field. Although the defense is required to guard the whole field for passes, by eliminating the threat of a run the defense is able to focus on a sole responsibility.

4.4 Defensive Schemes Effectiveness

Defensive coverage schemes were analyzed for their impact on offensive play success. The results indicate that certain coverage types are more effective in limiting offensive success:

- Most Effective Schemes (Lowest Success Rate Allowed):
  - Cover 0: 56.7%, a blitz-heavy man coverage that forces quicker throws.
  - Cover 1: 64.8%, a man coverage with one deep safety, and usually one extra blitzers.
- Less Effective Schemes:
  - Cover 2: 68.4%, a man coverage with two deep safeties that is balanced.
  - Cover 6: 69.2%, a zone coverage with three deep safeties, one covering half of the field and the other two splitting the other
  - Cover 3: 70.6%, a zone coverage with three deep safeties that mitigates deep passing threats.
  - Quarters: 71.8%, a zone coverage with four deep defenders preventing explosive plays.

These results indicate that defensive strategies emphasizing tighter coverage and more defenders near the line of scrimmage lead to lower offensive success rates. Conversely, dropping three or more safeties deep to prevent explosive plays appears to increase vulnerability to consistent yardage gains. Cover 3 was the most frequently used coverage scheme, employed on 38% of plays, despite yielding the highest offensive success rate and the third-highest explosive play rate allowed. Explosive plays, defined as gains of 10 or more yards, occurred at similar rates across most schemes, ranging from 18.8% to 21.1%, with one notable exception: Cover 0, which allowed explosive plays on only 7.7% of snaps. Cover 0 also had the lowest overall offensive success rate against it, yet it was the least utilized scheme, called on just 4.2% of plays. Additionally, Cover 2 was the scheme most frequently targeted by passing plays, with teams opting to pass on 65% of plays against it.

4.5 Success Rate by Play Type

Running plays and Passing plays had differing success rates:

- Run Plays: 71.9%
- Pass Plays: 65.3%

Given that success rate is defined as the percentage of plays that gain any yardage, it is unsurprising that running plays exhibit a higher success rate due to their inherently low variance. Most run plays result in modest gains or minimal losses, aligning with their strategic intent—to move the offense incrementally closer to the first down marker while minimizing risk. In contrast, passing plays, while offering greater potential for explosive gains, carry a higher risk of being incomplete and thus counted as unsuccessful. This tradeoff is evident in the data: 30.4% of passing plays were classified as explosive, compared to only 12.7% of running plays. Despite the higher failure rate, NFL offenses have increasingly favored the passing game, recognizing the challenge of sustaining long drives—often requiring seven or more consecutive successful plays—without a wasted down or penalty pushing the offense into a third-and-long situation. As previously noted, such scenarios significantly reduce the likelihood of conversion.

4.6. Success Rate by Field Position

Field position appears to have a modest impact on play success rates, as illustrated in the following breakdown:

- Own 1-20 Yard Line: 72.1%
- Own 21-50 Yard Line: 70.5%
- Opponent 49-21 Yard Line: 68.6%
- Red Zone (Opponent 20-1 Yard Line): 64.6%

Offensive success rates declined progressively as teams advanced closer to the opponent’s end zone. This trend was primarily driven by a decrease in passing success rate, which fell from 69.6% to 67.4%, then 66.3%, and ultimately dropped to 59.9% in the red zone. In contrast, the success rate of running plays remained relatively stable, ranging from 74.0% to 69.5%. The decline in passing success can be attributed to the shrinking field, which limits route spacing, allowing defenses to better cover receivers. Additionally, the compressed field enables defenders to align closer to the line of scrimmage, improving their ability to contain the run.

5. CONCLUSION

This study has provided a comprehensive analysis of offensive and defensive strategies in the NFL by utilizing data from the 2022 season. By exploring the effectiveness of different offensive formations, defensive coverage schemes, and play types, we have identified key insights that can inform coaching decisions and game planning. Our analysis highlights the importance of formation balance, the influence of defensive coverages, and most importantly field position and down situation in determining play success rates.

Our findings suggest that balanced offensive formations—such as the I-Form—are associated with higher success rates, while pass-heavy formations like the Empty set are more susceptible to defensive pressure. Defensively, schemes that place more players closer to the line of scrimmage, such as Cover 0, proved more effective at limiting both incremental and explosive gains. We found that running plays were more consistent in gaining positive yardage but were significantly less likely to result in explosive plays. We also found that third and fourth down conversion rates are strongly dependent on the distance needed to convert, reinforcing the critical importance of early-down efficiency. Furthermore, offensive effectiveness declined as teams approached the opponent's end zone, particularly in the passing game, likely due to the reduced field space and increased defensive concentration.

Ultimately, this research emphasizes the value of data-driven decision-making in the NFL, and how leveraging historical data and advanced analytics can provide teams with a competitive edge. Future work can build on these findings by continuing to develop machine and deep learning models that assist in decision-making and predict what the opponent will try to do. Additionally, data from multiple years can be compared to see league-wide trends in usage rates for offensive formations and defensive schemes.

REFERENCES

[1] J. Blinkoff, M. Voeller, S. Graham, and J.Wilson, "Predictive modeling of 4th down conversion in Power 5 conferences: Football data analytics," *The Sport Journal*, 2022. https://thesportjournal.org/article/predictive-modeling-of-4th-down-conversion-in-power-5-conferences-football-data-analytics/.

[2] N. Weller, "What College Football Can Teach Us About Fourth Down Decision-Making in the NFL," 2019. https://www.sharpfootballanalysis.com/analysis/what-college-football-can-teach-us-about-fourth-down-decision-making-in-the-nfl/.

[3] P. Lee, R. Chen, and V. Lakshman, "Predicting Offensive Play Types in the National Football League," *Stanford University*, 2016. https://cs229.stanford.edu/proj2016/report/LeeChenLakshman-PredictingOffensivePlayTypesIntheNFL-report.pdf.

[4] C. New, "The Application of Data Analytics to Key Decision Making by NFL Coaches," *Medium*, 2020. https://medium.com/@connornew/the-application-of-data-analytics-to-key-decision-making-by-nfl-coaches-7079f5f29b47.

[5] D. Krishna, "Analysing the Flow of NFL Games with One Simple Statistic," *Towards Data Science*, 2021. https://towardsdatascience.com/analysing-the-flow-of-nfl-games-with-one-simple-statistic-5bfca4e17dd4/.

[6] D. Aarhus, M. Simikic and M.Zuber, " Leveraging Machine Learning to Predict Football Play Calling," *Medium* 2024. https://medium.com/@monikasimikic/leveraging-machine-learning-to-predict-football-play-calling-2b1506f6521b

[7] B. Husowitz, M. Mixer, and S. Morrow, " NFL Data Analytics and Predictions Using Machine Learning," *Springer*, 2024. https://link.springer.com/referenceworkentry/10.1007/978-3-030-93954-0_22-1

[8] C. Joash Fernandes, R. Yakubov, Y. Li, A.K. Prasad, and T. C. Y. Chan, "Predicting plays in the National Football League," *Journal of Sports Analytics*, 2019. https://doi.org/10.3233/JSA-190348

[9] Kaggle, "NFL Big Data Bowl 2022," Kaggle, 2022. [Online]. Available: https://www.kaggle.com/competitions/nfl-big-data-bowl-2022