## Motivation

My motivation for writing this paper stems from the fact that fourth-downs are some of the most consequential moments in any football game. They can make or break a game, shift momentum, alter win probabilities, and redefine how it unfolds. Fourth-down decisions are strategic choices that can be measured and refined so that each play, team, and game can be improved based on what is optimal at the moment to increase the chance of winning.

Traditional metrics have focused on conversion rates, raw expected points added, or win probability, but they often fail to capture the interplay between offensive aggression and the defensive strength facing that aggression. Our metric, Adjusted fourth-down Efficiency (A4DE), combines the rate at which fourth-down plays are converted into first downs with the quality of the resulting play (measured by EPA) while adjusting for the opposing defense's ability to stop these plays. A4DE is innovative in that it accounts for defensive quality by using a percentile ranking of opponents' "stop rates" (calculated as one minus the conversion rate). By weighting EPA with a defensive adjustment, our metric provides a clearer picture of how well a team or player executes in fourth-down situations compared to the league average. This provides analysts, coaches, and fans a more nuanced understanding of fourth-down aggressiveness and efficiency.

## History

There has been significant analysis on fourth-down conversions since NFL data packages such as nflreadr and nflverse were released to the public by the NFL. Analysts from football teams, sports TV shows, and sports data companies have used these tools to generate statistics that help fans learn and understand the game. News anchors often discuss these stats to question teams' aggressiveness, evaluate whether fourth-down decisions are warranted, and analyze trends in the football world.

Articles on Medium, for example, examine conversion rates—breaking them down by year, by yards-to-go on fourth-down attempts, and by success rate—driven by an interest in whether these metrics have been increasing over time. In addition, some developers have created Shiny applications that calculate the probability of converting on fourth-down based on the possession team, defensive team, field position, quarter, score differential, and other relevant metrics. This tool aims to project what decision a team should make based on completion and win probabilities.

Although fourth-down analysis is popular due to its game-changing impact, many existing studies do not fully account for defensive context. While many analysts break down success rates in detail, few have robustly integrated both offensive performance and defensive strength into a single measure. These analyses may reveal that a team's raw conversion rate is high or low, but they often do not determine whether this success is attributable to excellent offensive play or a particularly weak opposing defense. Our Adjusted fourth-down Efficiency (A4DE) metric fills this gap by adjusting offensive success for the strength of the opposition's

defense. By emphasizing context, A4DE provides an alternative perspective to traditional measures, making it a potentially more robust predictor of long-term success and situational decision-making.

## Development

Data and Methodology:

Using publicly available NFL play-by-play data from the nflverse package, I isolated fourth-down plays from the 2024 season and a multi-year period of 2021–2024 for additional historical perspective. We first classified fourth-down plays into three main decisions: Go for it (offensive plays such as runs and passes), Punt, and Field Goal. I then computed several statistics for each play group, including: Conversion Rate: The rate at which plays resulted in a first down. EPA: A measure of the play's impact on the game. Count: Total number of plays in each category.

Adjusting for Defensive Quality:

A key innovation in our metric is the adjustment for defensive strength. We computed a defensive "stop rate" for each team, which is defined as 1 – conversionRate. I then calculated the defensive stop percentile for each team, allowing us to compare defenses on a relative scale. The adjustment factor was applied as follows:

1. Adjusted EPA: adj_epa = avg_epa × (1 + (1 − def_stop_percentile)). This means that a play against a weaker defense will see its EPA boosted.
2. A4DE Calculation: Finally, we multiply the conversion rate by the adjusted EPA: A4DE = conversionRate × adj_epa. This composite metric captures both the likelihood of converting on fourth-down and the effectiveness of that conversion in terms of expected points, weighted by the defensive quality.

Considerations and Limitations:

While our metric provides a more holistic view of fourth-down decision-making, some factors remain unaccounted for. Factors such as field position nuances, weather conditions, or player injuries are not directly incorporated. Although win probability is used in some filtering steps, further refinement could adjust for situational pressure, like score differential or time remaining, more explicitly. With access to richer datasets, like player tracking data, future iterations could refine how individual player performance contributes to overall efficiency.

## Results

To begin, I isolated and categorized all Philadelphia Eagles fourth-down plays from the 2024 season, classifying each as a punt, field goal, or go-for-it attempt. From these data, the Eagles emerged with a 70.97% fourth-down conversion rate across 315 attempts, substantially above the league-wide average of 57.11% (7,628 attempts). This gap underscores the Eagles' unusually high propensity to go for it, and to succeed when they do, compared to their peers.

A more detailed examination of "most aggressive" versus "least aggressive" teams revealed a similar pattern. Teams like Detroit and Washington showed large gaps between their actual go-for-it rates and model-projected go-for-it rates (e.g., Detroit's over_expected of 0.178), indicating a coaching philosophy that regularly exceeds conservative thresholds. In contrast, teams such as San Francisco and Minnesota had narrower gaps and ranked among the least aggressive, with smaller over_expected margins that hint at a more conservative decision-making process.

Most Aggressive Teams:

| posteam <chr> | actual_go_for_it_rate <dbl> | expected_go_for_it_rate <dbl> | over_expected <dbl> |
|---|---|---|---|
| DET | 0.3271028 | 0.14953271 | 0.17757009 |
| WAS | 0.2835821 | 0.20149254 | 0.08208955 |
| CLE | 0.2767296 | 0.11320755 | 0.16352201 |
| CHI | 0.2676056 | 0.10563380 | 0.16197183 |
| ATL | 0.2672414 | 0.09482759 | 0.17241379 |
| NYG | 0.2642857 | 0.10000000 | 0.16428571 |

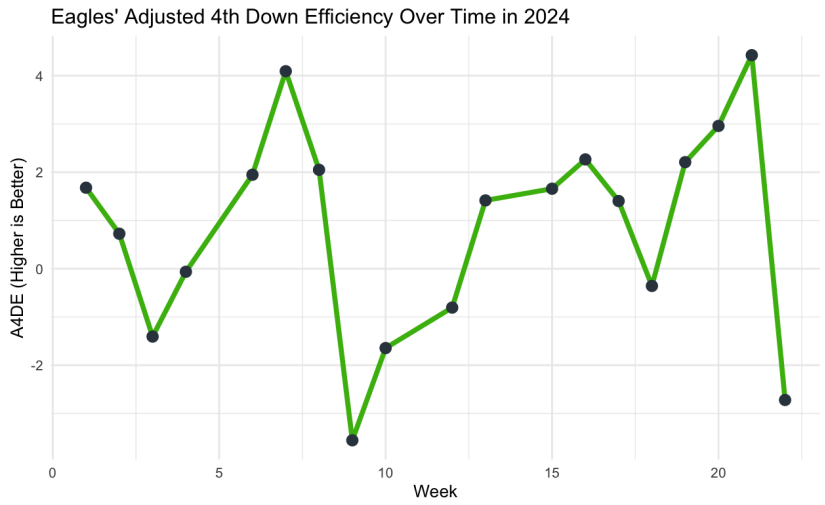
Least Aggressive Teams:

| posteam <chr> | actual_go_for_it_rate <dbl> | expected_go_for_it_rate <dbl> | over_expected <dbl> |
|---|---|---|---|
| SF | 0.1538462 | 0.06730769 | 0.08653846 |
| MIN | 0.1452991 | 0.05128205 | 0.09401709 |
| PIT | 0.1439394 | 0.02272727 | 0.12121212 |
| HOU | 0.1428571 | 0.09523810 | 0.04761905 |
| TB | 0.1428571 | 0.08163265 | 0.06122449 |
| LAC | 0.1268657 | 0.06716418 | 0.05970149 |

**Where Do the Eagles Fit?**

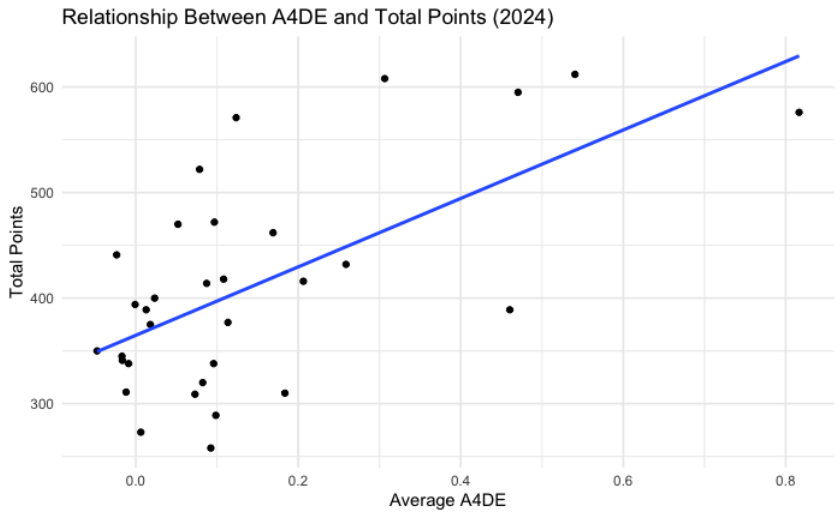
Although the Eagles did not top the charts in sheer aggressiveness, their actual go-for-it rate still surpassed their expected rate by a considerable margin. This confirms that they consistently seek an edge on fourth-downs when conversion odds are reasonable, yet remain mindful of not surrendering field position at inopportune moments.

## Adjusted 4th Down Efficiency (A4DE)

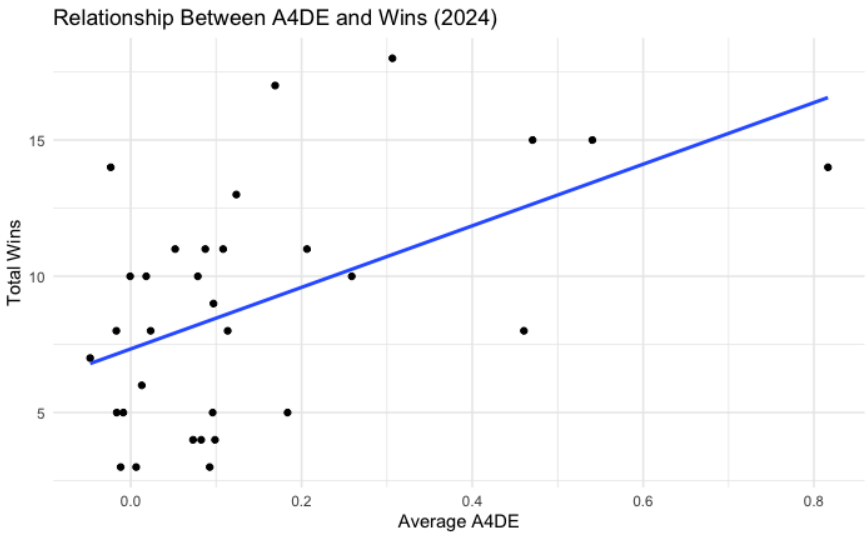



To gain a clearer picture of the Eagles' week-to-week performance, I tracked their Adjusted 4th Down Efficiency throughout the 2024 season. As shown in the graph, the Eagles generally maintained an A4DE above zero, indicating they executed fourth-down plays effectively against a variety of defensive strengths. Notably, there are a few spikes, around Weeks 5 and 6 and again towards the end of the season, where the Eagles' A4DE rose significantly, suggesting that they converted high-leverage fourth-downs against tougher defenses or generated particularly impactful scoring opportunities during those periods. Meanwhile, the occasional dips below the trendline may reflect matchups against elite defenses, key injuries, or shifts in offensive strategy. The one clear exception is the final data point, the Super Bowl, where the Eagles attempted only one fourth-down play late in the game with their backups, resulting in a negative A4DE. Nevertheless, the fact that their A4DE remained positive for most of the season underscores that the Eagles' strong fourth-down execution was not simply a product of facing weaker opponents; rather, it reflects a consistently effective approach that translated into sustained drives, increased scoring opportunities, and ultimately, more wins.

## Predicting Wins and Points



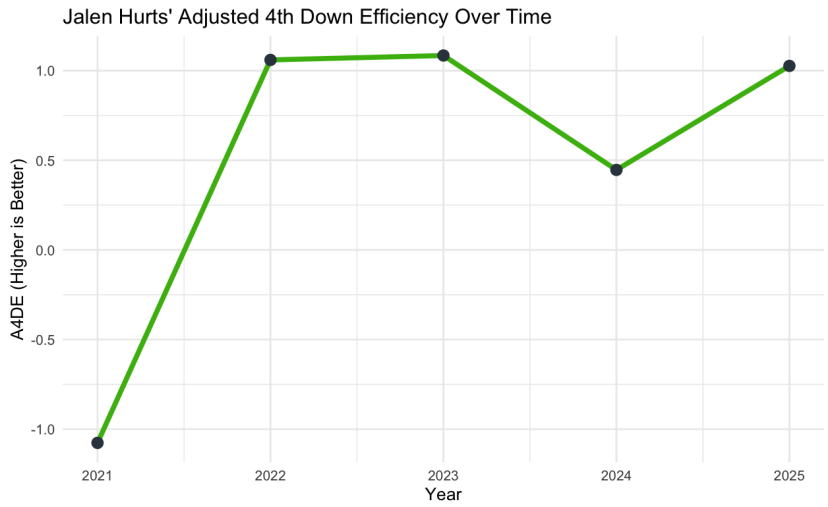

To examine how Adjusted 4th Down Efficiency correlates with overall team success, I plotted each team's average A4DE against both their total points and their total wins for the 2024 season. As shown in the graph above, there is a clear positive relationship between A4DE and total points scored: teams with higher A4DE values often surpassed 500 points on the season, whereas teams with lower A4DE values tended to remain below the 400-point mark. This trend supports the notion that when teams not only choose to go for it on fourth-down but also execute effectively, especially against quality defenses, they generate more sustained drives and thus capitalize on more scoring opportunities. To examine the relationship between A4DE and offensive production, I ran a linear regression using total points as the dependent variable and A4DE of a team as the independent variable. The model output indicates a statistically significant positive relationship between these two measures (Estimate = 324.24, $p < 0.001$). Interpreted practically, this suggests that a one-point increase in A4DE corresponds to roughly 324 additional points over the course of a season, though, in reality, teams' A4DE values tend to cluster within a narrower range. With an R-squared value of about 0.386, A4DE explains a notable portion of the variation in total points, highlighting its potential as a meaningful indicator of a team's offensive success.




A similar pattern emerges in the graph above, which compares average A4DE to total wins. Here, an upward-sloping regression line again indicates that teams with higher A4DE tend to rack up more victories. While no single metric can guarantee success, the consistency of this relationship across the league reinforces that strong fourth-down execution can be a meaningful advantage. Collectively, these findings underscore that A4DE does more than just describe how frequently a team is willing to gamble on fourth-down—it captures how well they take those risks, and how that skill positively influences both scoring output and win totals. To assess how well A4DE predicts team success, I ran a linear regression using total wins as the dependent variable and A4DE of a team as the independent variable. The model's output, shown above, reveals that A4DE has a positive and statistically significant effect on total wins (Estimate = 11.2974, $p = 0.00295$). In practical terms, this suggests that a one-point increase in A4DE is associated with roughly 11 additional wins over a season, an understandably large effect given that most teams' A4DE values cluster within a relatively narrow range. While the R-squared value of 0.259 indicates that many other factors influence a team's total wins, as one would expect in a complex sport like football, the low p-value confirms that A4DE meaningfully captures an element of performance tied to on-field success.

When correlating a team's A4DE with outcomes like total points scored and overall wins, a positive relationship clearly emerged. High A4DE teams often prolong drives by converting critical fourth-downs, thereby improving their field position and scoring opportunities. While no single metric guarantees victory, A4DE proved a meaningful indicator of how risk-taking on high-leverage downs can yield tangible advantages.

**Stability Across Seasons**




Looking at Jalen Hurts' A4DE across multiple seasons provides insight into how stable this metric can be over time. As the chart shows, Hurts' A4DE took a significant leap from 2021 to 2022, suggesting that improvements in offensive scheme, personnel, or Hurts' own development elevated his ability to convert high-leverage fourth downs against varying defensive strengths. The subsequent dip and rise in later years, while remaining firmly positive, demonstrate that, although fourth-down performance can fluctuate due to factors like injuries or coaching adjustments, the overall trend remains strong. This year-to-year consistency indicates that A4DE likely captures a repeatable skill set in decision-making and execution under pressure. In other words, if a quarterback like Hurts excels at converting critical downs one season, he often carries that success forward, barring major external changes. Where we do see notable swings, whether up or down, they tend to coincide with shifts in team context, such as roster turnover or scheme tweaks. Taken together, Hurts' multi-season A4DE trajectory reinforces that the metric is not only meaningful in a single season but also robust enough to track a player's evolving performance over several years.

To further explore whether A4DE reflects a repeatable skill for Jalen Hurts, I examined both the variability in his A4DE and its year-to-year correlation. His standard deviation of roughly 1.47 indicates a moderate amount of fluctuation in A4DE across seasons, while the yearly correlation of –0.14 suggests that changes in Hurts' A4DE may be driven more by external factors, such as injuries, roster changes, or coaching adjustments, than by a strictly stable trend over time. Despite these fluctuations, Hurts has maintained a relatively high A4DE each season, reinforcing that his fourth-down decision-making and execution remain valuable assets to the Eagles on-field success.

**Why A4DE Matters**

When ranking teams using raw EPA, those with impressive scoring plays and aggressive fourth-down attempts might initially rank highly. However, raw EPA fails to distinguish whether

those plays occurred against a weak or strong defense. In contrast, A4DE adjusts each play's EPA by accounting for the quality of the opposing defense. This adjustment yields refined rankings: teams that face stronger defenses and still convert on fourth-downs tend to rank higher with A4DE, while teams benefiting from easier match-ups may drop in the rankings. Moreover, this context provides better performance insights; for example, a team that appears mediocre by raw EPA standards might actually be executing exceptionally well when considering that their opponents are among the toughest defensively, whereas a team that seems dominant with raw EPA might have inflated numbers due to facing weaker defenses. By incorporating defensive context, A4DE offers a more accurate picture of a team's effectiveness in high-pressure situations, highlighting not only offensive success but also the resilience and skill required to perform well against strong defenses. This comprehensive view makes A4DE a more robust and informative metric than raw EPA alone.

Taken together, these findings confirm that A4DE is both analytically meaningful and relatively stable over time. By integrating defensive context with expected points added, A4DE goes beyond simply tracking "go-for-it" frequency. It quantifies how effectively teams execute in these high-leverage moments, and connects directly to real-world outcomes such as wins, points, and year-to-year performance trends. This blend of descriptive power and predictive value makes A4DE a compelling lens for understanding how and why aggressive fourth-down strategies translate into sustained success on the field.

## Conclusion

Adjusted fourth-down Efficiency is a novel metric that synthesizes traditional measures, such as conversion rate and EPA, with an adjustment for the opposing defense's strength. This combined approach offers a more nuanced picture of fourth-down decision-making by capturing both the aggressiveness of play-calling and the situational effectiveness of those decisions. A4DE is important because it acknowledges that not all fourth-down successes are created equal; rather, they depend heavily on the quality of the opponent. Furthermore, the metric highlights the contributions of key players, such as Jalen Hurts, and identifies how their performance influences overall team efficiency. For coaches and analysts, understanding A4DE could lead to refined in-game decision-making and provide a strategic edge, while its broader applications extend beyond the NFL to other sports where situational decisions are critical. By addressing both offensive and defensive dimensions, A4DE offers an analytically robust alternative to existing measures, making it a compelling candidate for inclusion on major sports analysis platforms such as ESPN. If ESPN adopted A4DE, fans and coaches could track and compare how effectively teams execute high-leverage plays against tough defenses each week, gleaning deeper strategic insights and making more informed decisions than raw conversion rates alone.