Building the Statistically Optimal Offensive Scheme

Nathan Wright

Overview

This project is the culmination of compiling results from several different research studies and applying them to offensive scheming to illustrate what implementing these strategies looks like in a practical application sense

Data & Considerations

- All evidence in this study stems from results found when examining 2021 to 2023 NFL play-by-play data provided by the nfl_data_py library
- All plays filtered down to plays that:
  - Have no penalties
  - Take place in neutral game scripts
    - |Score Differential| ≤16
  - Are regular plays
    - No spikes, kneels, kickoffs, field goals, two point conversions

Elements of a Scheme (studies involved)

- Defensive Analysis
  - Modeling Expected Offensive Play Concepts and Predicted Defensive Response
  - Sack Expectancy Rates
- Decision-Making
  - Analyzing and Estimating 4th Down Conversion Probability
- Formations & Personnel
  - YPC Across Formations & Personnel Grouping
- Playcalling
  - Early-Down Playcalling Analysis
  - Targeted Route Efficacy
  - Redzone Playcalling Analysis
  - Run Gap Analysis

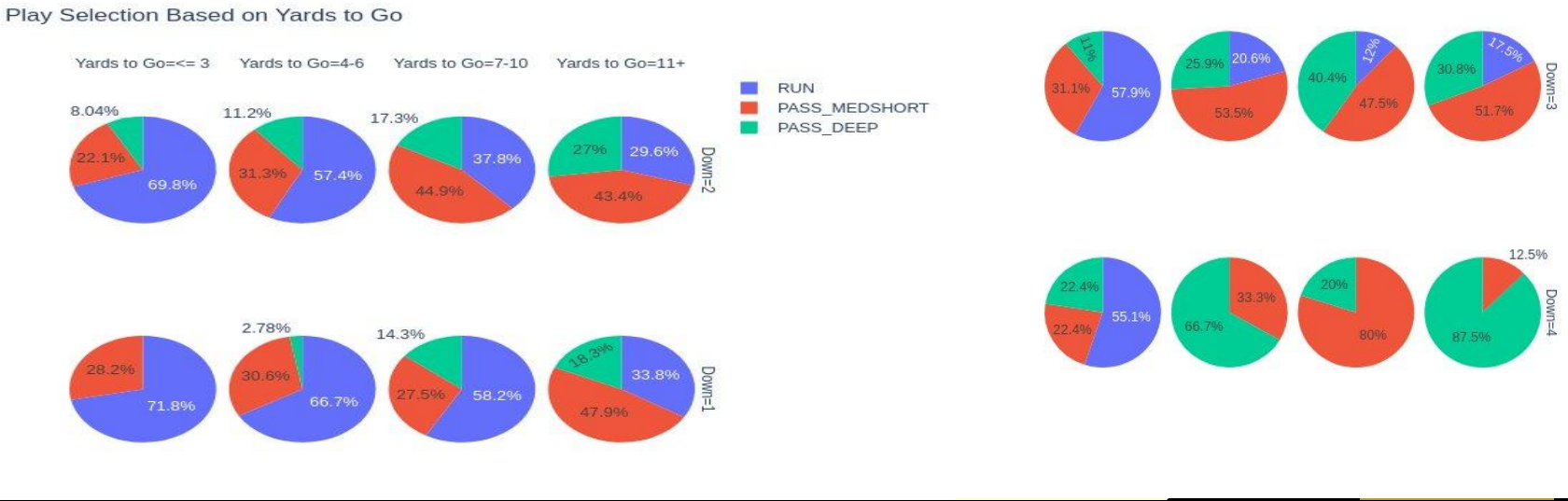
Defensive Analysis

Defensive Analysis

How a defense's decisions affect the offense is largely based upon knowing their expected reaction to a given look being seen from the offense. Reasoning backwards, a defense makes its coverage and personnel calls based on what it sees and expects from the offense. By knowing how a defense affects your plays, what it expects to see from you, and how it will react, you can make effective play calls to exploit what a defense will be deploying.

Defensive Analysis - Offensive Expectations

Show below is a distribution of expected offensive play selection, estimated using a contextual factors and league-wide tendencies

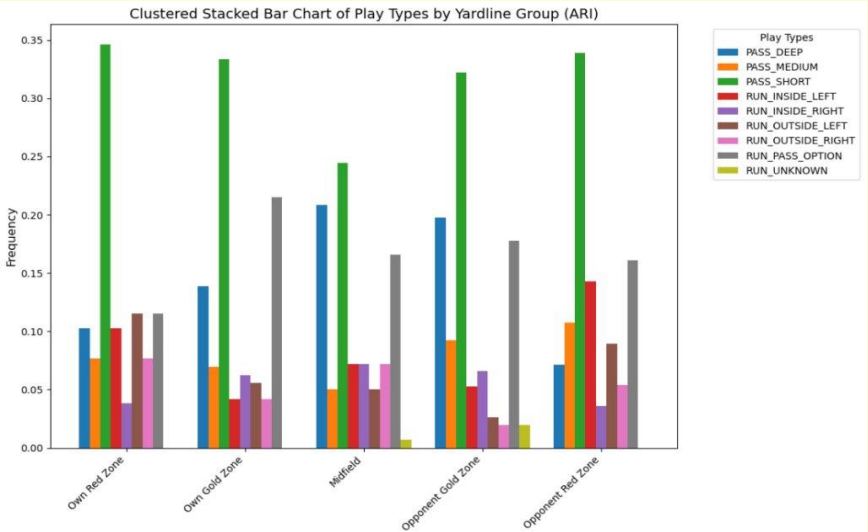


Defensive Analysis - Offensive Expectations (cont.)

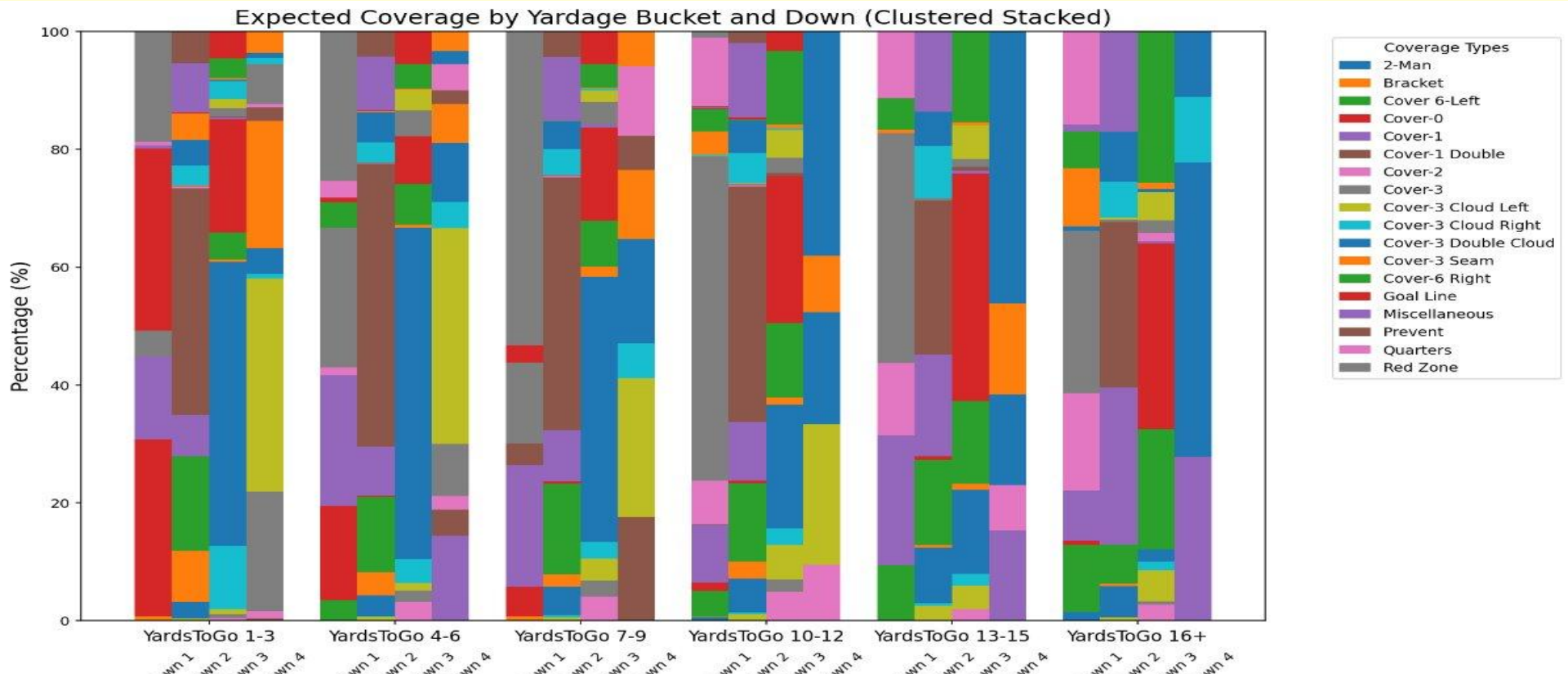
Using features such as:

- ★ Offensive Formation/Personnel
- ★ Down/Distance
- ★ Location on Field
- ★ Score/Time Remaining
- ★ Other Player Stats and Situational Factors

We built a predictive model that can identify the type of play an offense is running with at 75% accuracy.



Defensive Analysis - Coverage Prediction



Defensive Analysis - Coverage Prediction (cont.)

Similarly to our offensive play calling model, we used features such as:

- Defensive Alignment/Personnel
- Down/Distance
- Location on Field
- Score/Time Remaining
- Other Player Stats and Situational Factors

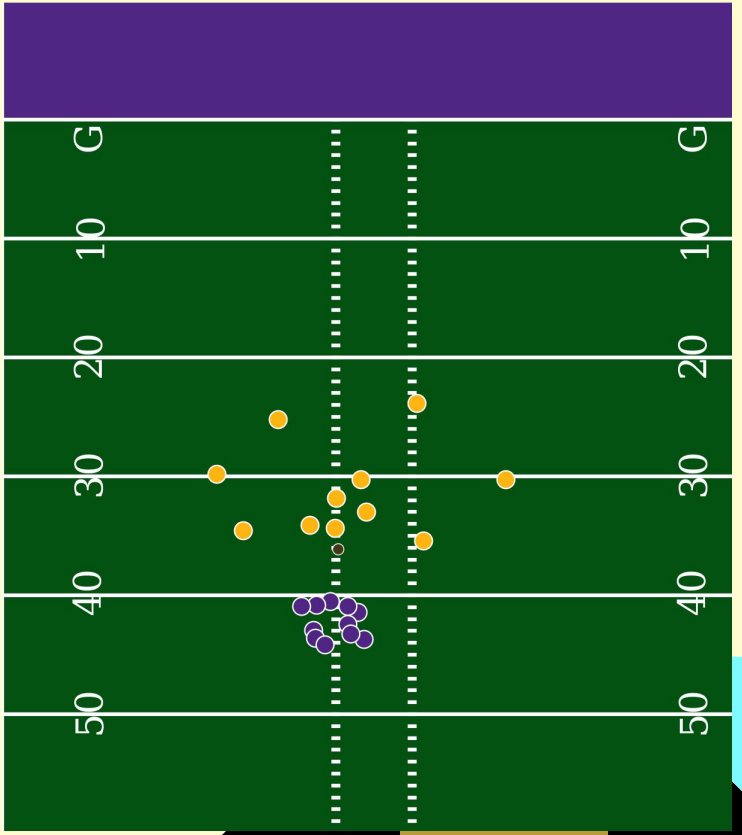
Along with the expected offensive play type derived from our model to build a predictive model that can identify a defense's coverage with an 85% accuracy.

Model Accuracy: 0.8644630034342804

	precision	recall	f1-score	support
2-Man	0.97	0.79	0.87	570
Bracket	1.00	0.88	0.94	236
Cover 6-Left	0.95	0.71	0.82	2020
Cover-0	0.97	0.88	0.92	1996
Cover-1	0.89	0.87	0.88	10263
Cover-1 Double	0.99	0.85	0.92	186
Cover-2	0.90	0.80	0.85	6131
Cover-3	0.78	0.97	0.86	16204
Cover-3 Cloud Left	0.99	0.70	0.82	97
Cover-3 Cloud Right	0.96	0.66	0.78	125
Cover-3 Double Cloud	1.00	0.74	0.85	38
Cover-3 Seam	0.94	0.72	0.82	1967
Cover-6 Right	0.97	0.71	0.82	2128
Goal Line	0.99	0.92	0.95	454
Miscellaneous	0.94	0.85	0.89	34
Prevent	0.99	0.94	0.96	157
Quarters	0.91	0.78	0.84	6737
Red Zone	0.97	0.97	0.97	1905
accuracy			0.86	51248
macro avg	0.95	0.82	0.88	51248
weighted avg	0.88	0.86	0.86	51248

Defensive Analysis - Coverages

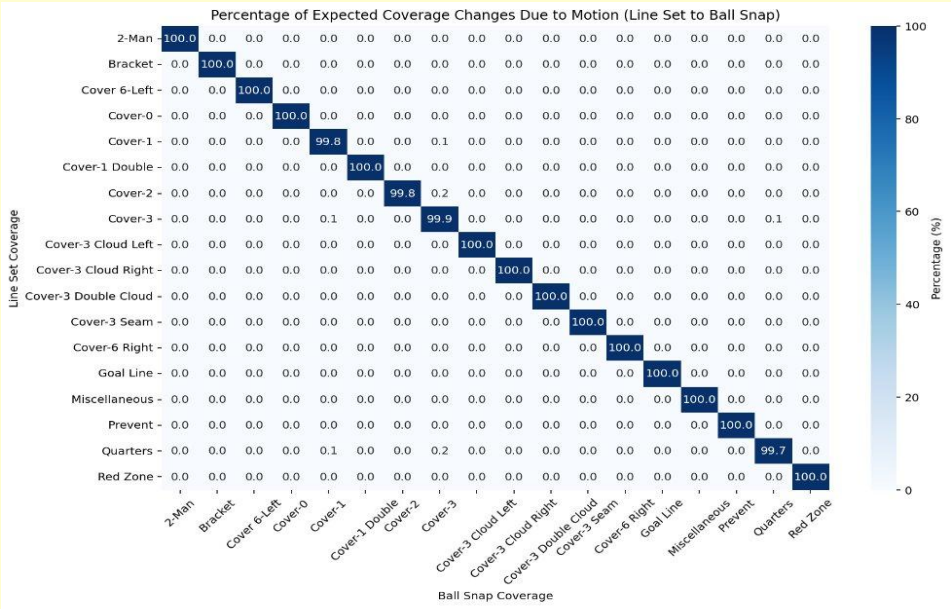
- This example of the play prediction models comes from a 36-yard touchdown pass from Kirk Cousins to Justin Jefferson
- This play occurred on a first down from the opposing 36-yard line with 45 seconds remaining in the half
- When the Vikings initially come set, they are lined up in a 3x1 singleback formation, they then motion the tight end to the left side of the field to create a 2x2 alignment
- Given these factors along with the other variables taken into consideration, our model correctly predicted this play to be a **deep pass**
- Based on the situational factors and the expectation of a deep pass from the offense, our defensive coverage model correctly predicted the coverage to be **Cover-3**



Defensive Analysis - Alignment Effects

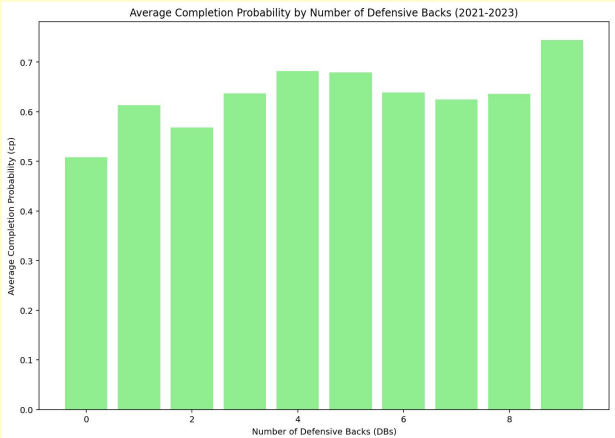
Motion doesn't affect expected coverage, but rather exposes the coverage as teams shift and call out assignment changes in response to motion or shifts from the offense.

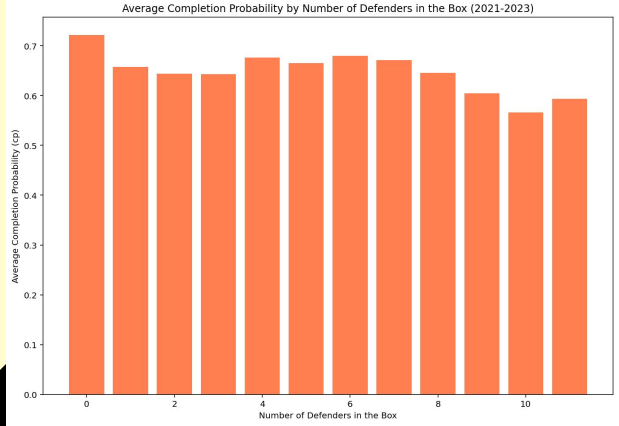
★ Expected Coverage does not change 99% of plays despite pre-snap decoys and shifts from defense



Defensive Analysis - Alignment Effects (cont.)

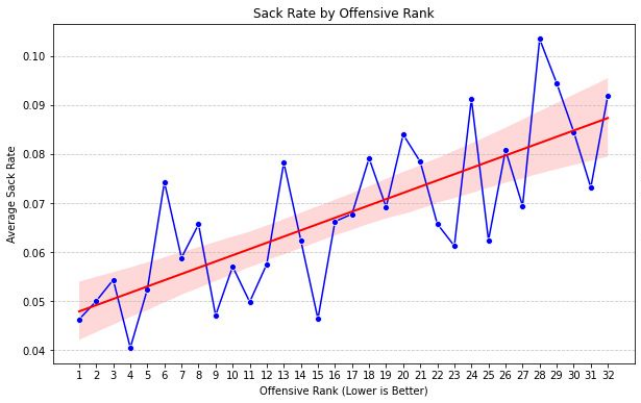
Ultimately, the pre-snap alignment of a defense does not have a significant effect on the chances of a given play's success. Knowing the expected back-end coverage scheme of a defense, training a QB to identify where pressure will come from, and ignoring the "eye candy" will lead to higher success rates rather than reading too much into what a defense is showing pre-play

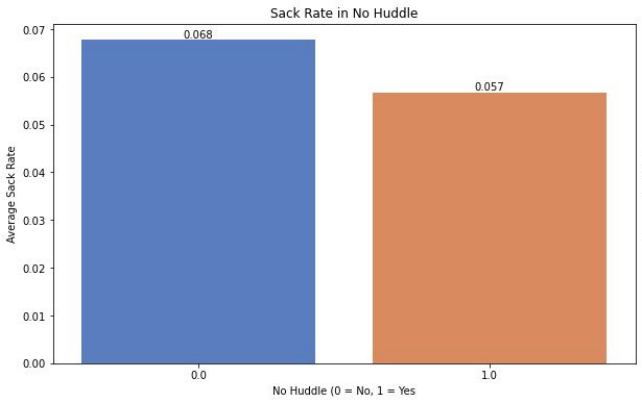


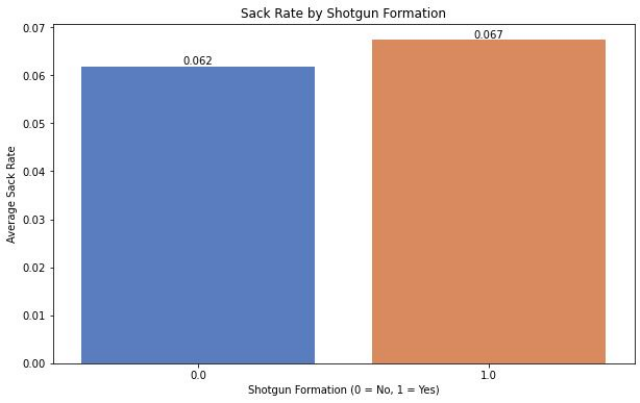


Defensive Analysis - Sack Rates

- ★ Offensive volume affects sack rates the heaviest
- ★ Operating out of no huddle keeps the defense off-guard
- ★ More likely to be sacked while passing out of shotgun



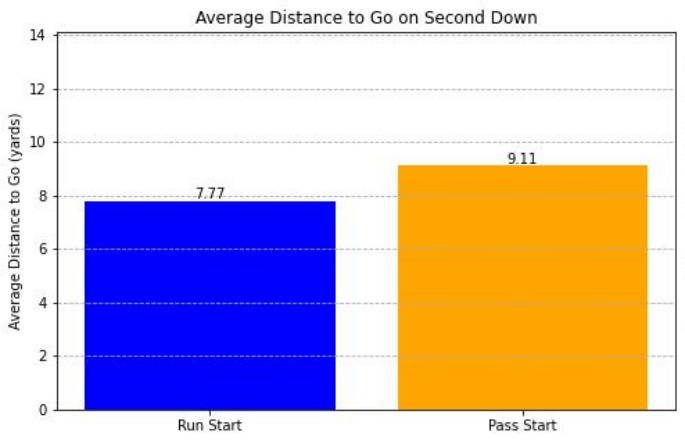


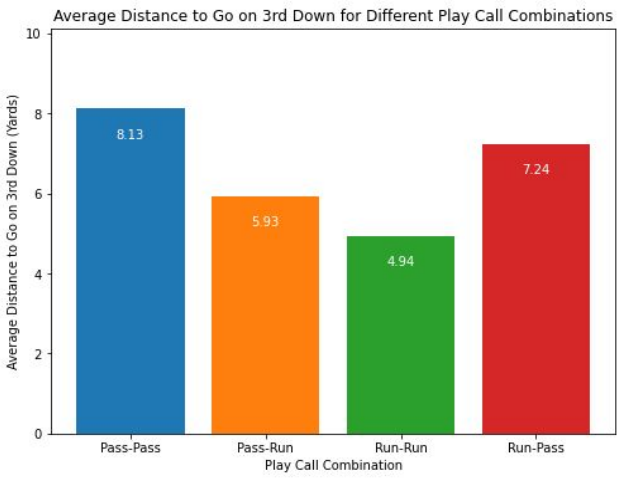


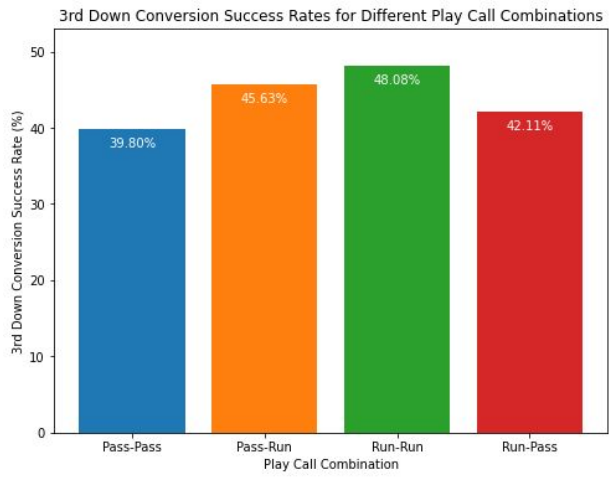
Decision Making

Decision Making - Early Downs

- Shorter distance to go on 3rd down directly correlates to higher conversion rates, evidencing the importance of “staying ahead of the sticks”.
- To do this, incorporating a run on first or second down



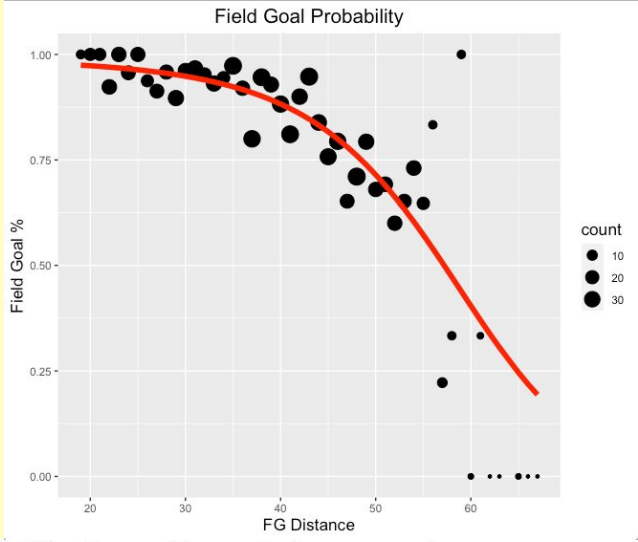


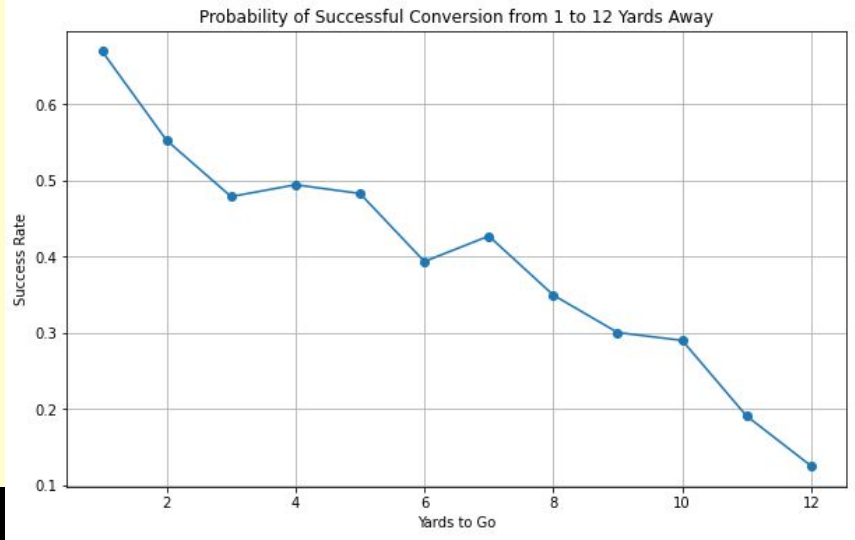


Decision Making - 4th Down

The most influential down on game outcomes & more importantly: Job Security

Despite a rise in conversion attempts over recent seasons, coaches still attempt conversions at a sub-optimal rate when factoring the tradeoff between probabilities and expected gains





Decision Making - 4th Down (cont.)

When deciding whether to go for it, algorithms designed to predict whether a conversion attempt will be successful, further informing the decision rather than relying purely on probabilities.

A random forest model was used to predict whether a team will be successful in a conversion attempt

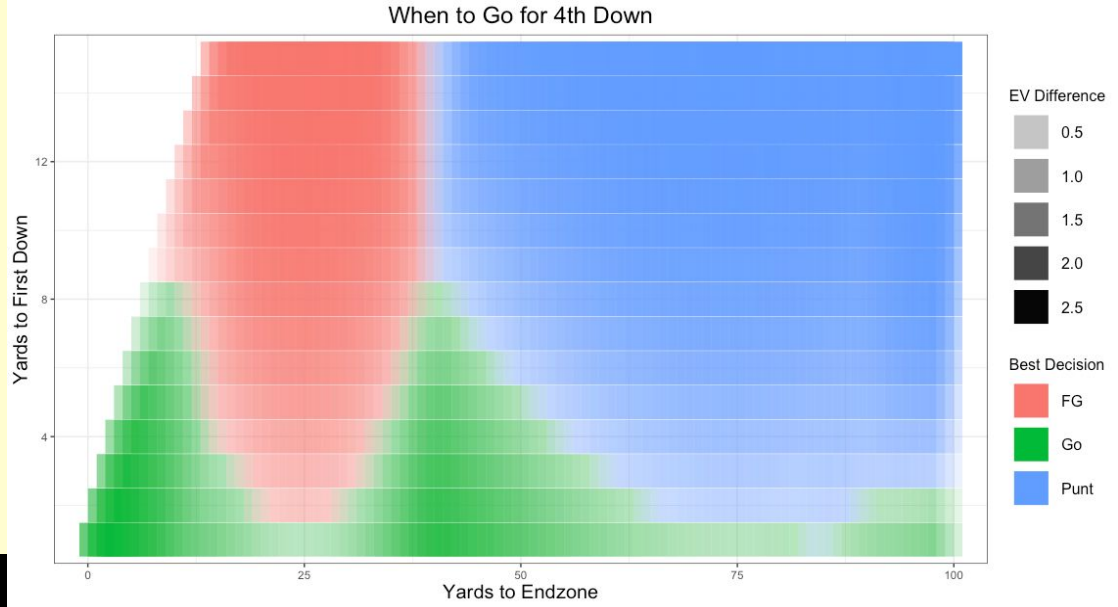
Most Important Features

0	ydstogo	-1.456882
1	yards_to_endzone	0.917689
2	ydsondrive	1.507169
3	cpoe	2.504830

cross_val_scores [0.85089974 0.89460154 0.88174807 0.87371134 0.84536082]
Mean cross-validation accuracy: 0.8692643044549863

Decision Making - 4th Down (cont.)

Essentially, a team will be successful on a conversion a lot more than one may think, but ultimately the factors listed previously affect a team's probability the most. When positioning properly in accordance with those factors, an offense should go for it whenever possible in these situations.



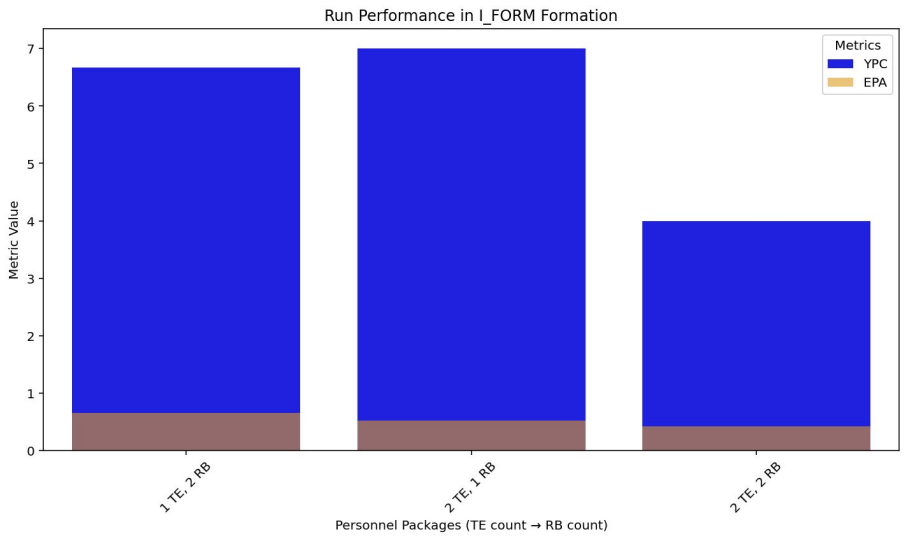
Formations & Personnel

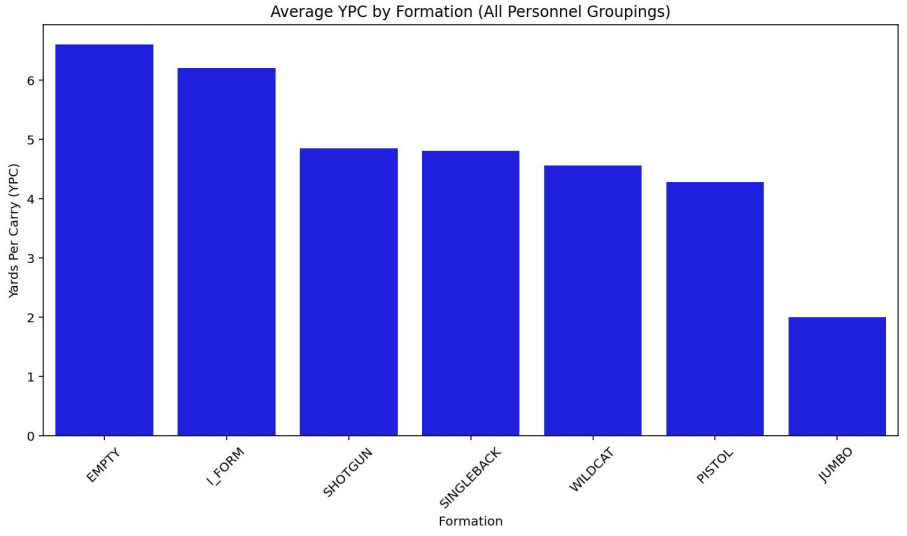
Formations and Personnel

- After understanding the defensive response and analyzing peripheral metrics like drive sustenance and 3rd/4th down conversions, the next step is optimizing the formations and personnel in your playbook.
- Effective deployment of formations and personnel creates mismatches and keeps defenses guessing, enhancing overall offensive efficiency.

Form & Personnel - Run

- Different formations influence run efficiency, with personnel effectiveness varying by formation.
- I-Form and Empty sets (for QB-designed runs or jet sweeps) yield the highest yards per carry (YPC).





Form & Personnel - Pass

- Passing efficiency remains relatively consistent across receiver alignments, personnel, and formations
- To keep defenses off-balance, mix in varied formations while emphasizing passing plays from your most efficient run formations.

Play Calling

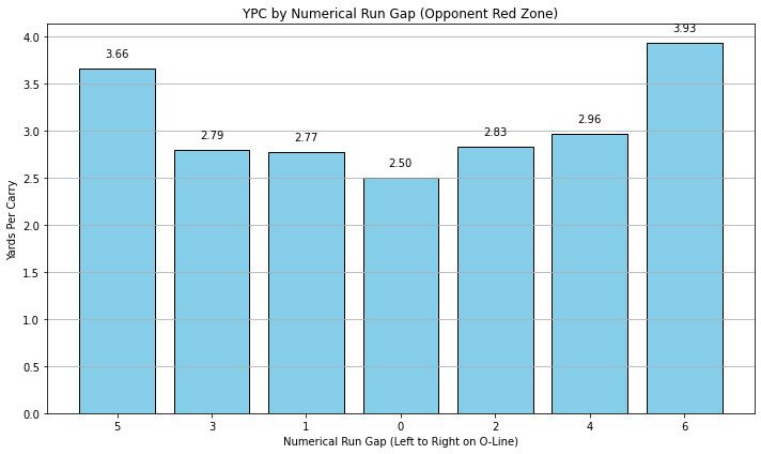
Play Calling - Concepts

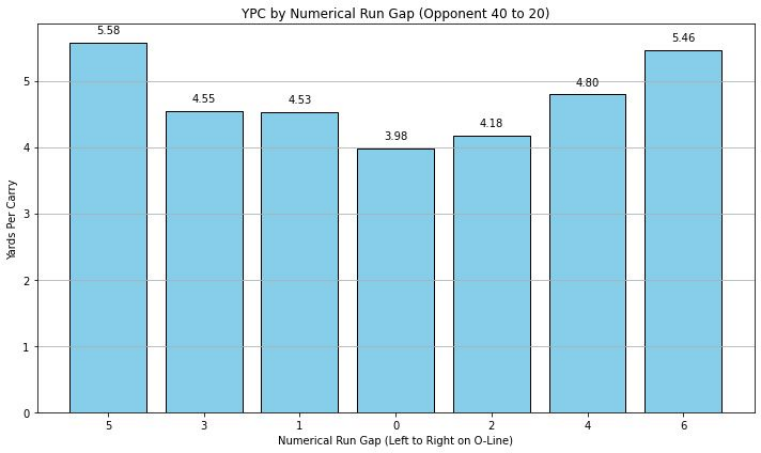
- Two studies on run game characteristics and passing route design reveal the most efficient and effective approaches for each play concept.
- Integrating these findings into your play design ensures you not only have plays for each defense and situation, but also the most effective play designs across the board.

Play Calling - Run Scheme

- Running outside the tackles averages the most yards, but its effectiveness varies by field location.
  - For example, running outside in the Gold Zone (opponent's 40-20) yields a higher yards per carry compared to inside runs vs the Red Zone.

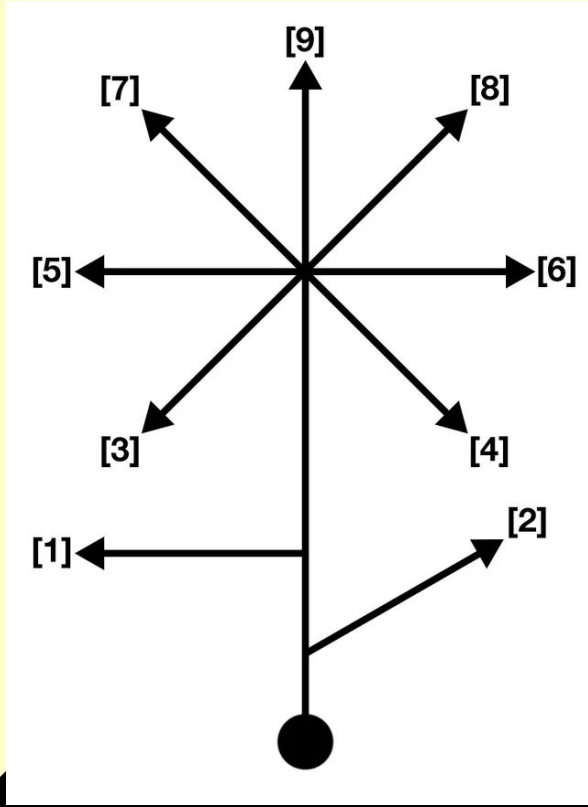
- Prioritize outside runs in areas where they provide the most benefit, and use them sparingly in other parts of the field to minimize the risk of negative plays.





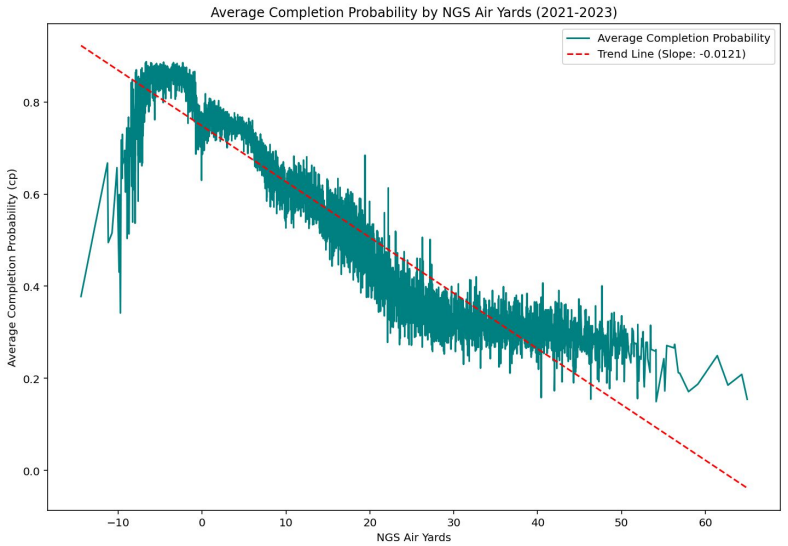
Play Calling - Pass Scheme

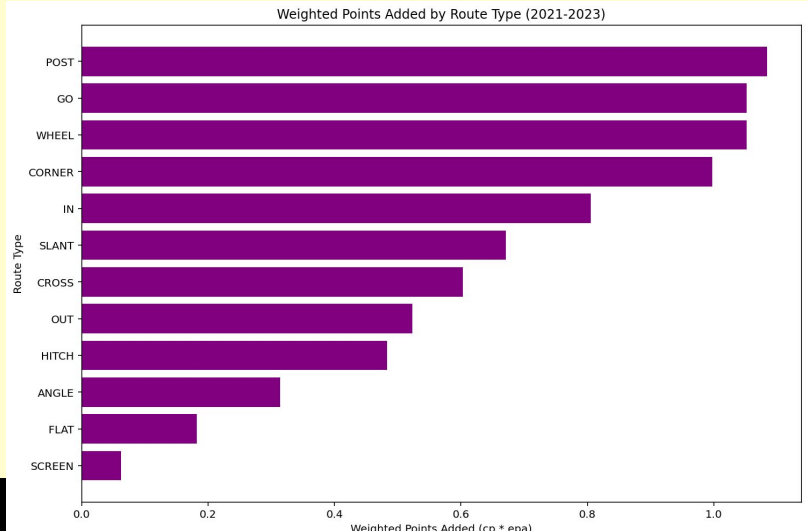
- A study on the efficacy of targeted receiver routes reveals which routes are most effective against various defensive coverages.
- By designing play calls to exploit each coverage type and incorporating these effective routes, you can optimize passing game efficiency and output.

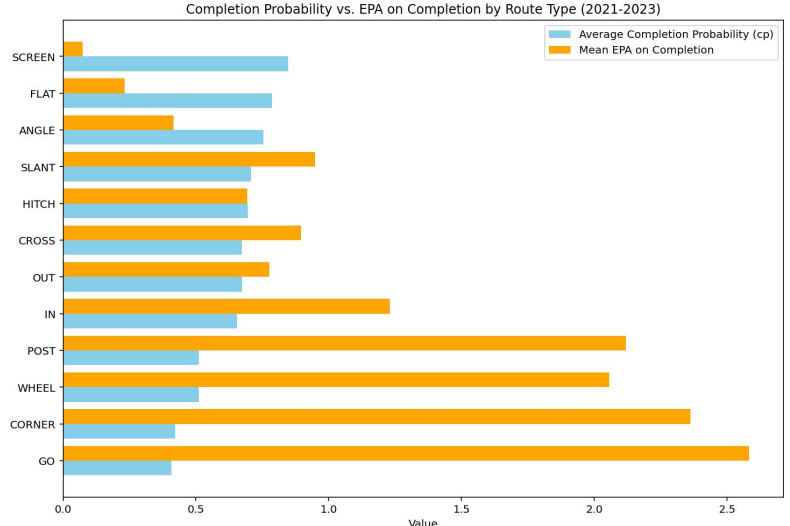


Play Calling - Pass Scheme (cont.)

- Direct and abrupt tradeoff between risk/reward
- Completion Probability (cp) plummets with increased air yards
  - Conversely, Expected Points Added (EPA) upon completion skyrockets

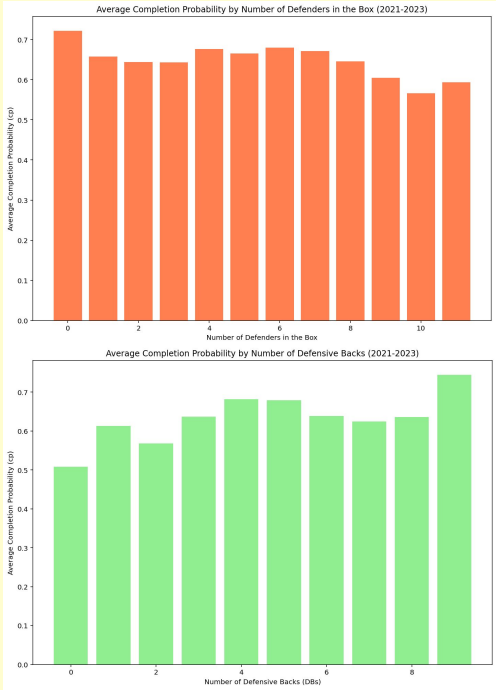


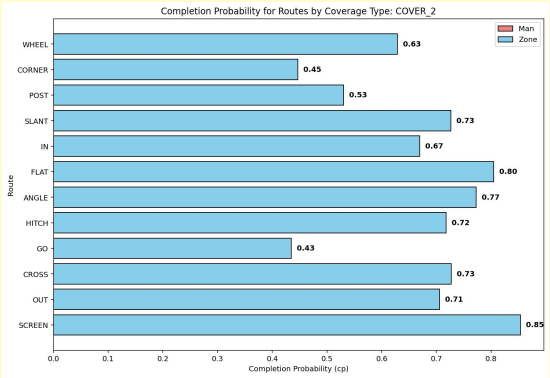


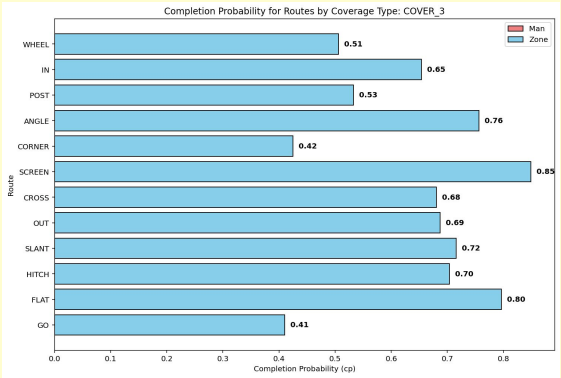


Play Calling - Pass Scheme (cont.)

- Low variance in completion probability (CP) across changes in personnel and pre-snap alignment.
- High variance in CP based on the routes run against varying coverages.

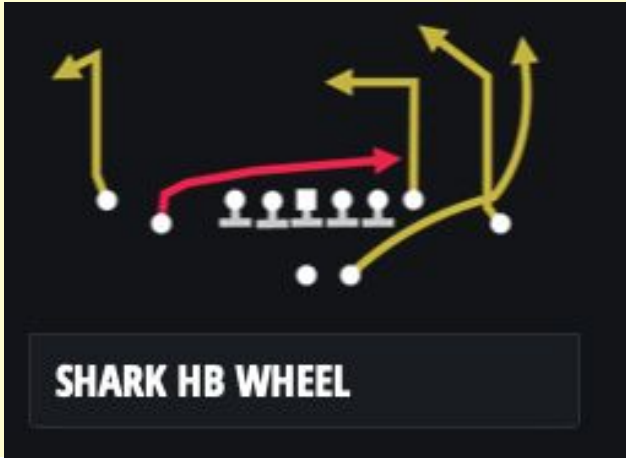


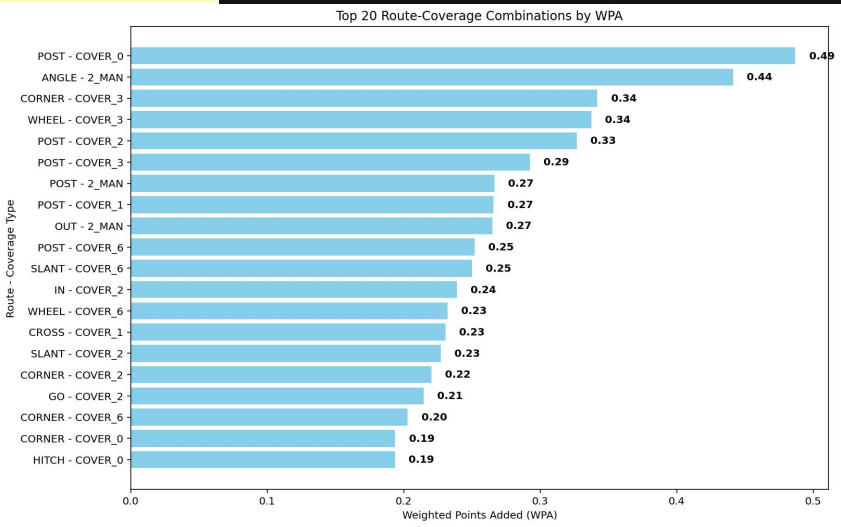




Play Calling - Pass Scheme (cont.)

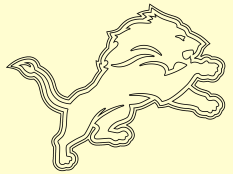
- Recurrence of certain routes vs several different coverages
  - Post
  - Corner
  - Wheel
  - Cross/In/Slant
- Highlights importance of having downfield routes with cuts rather than simple go routes
- Presence of Angle, Wheel routes shows effect a strong receiving RB can have



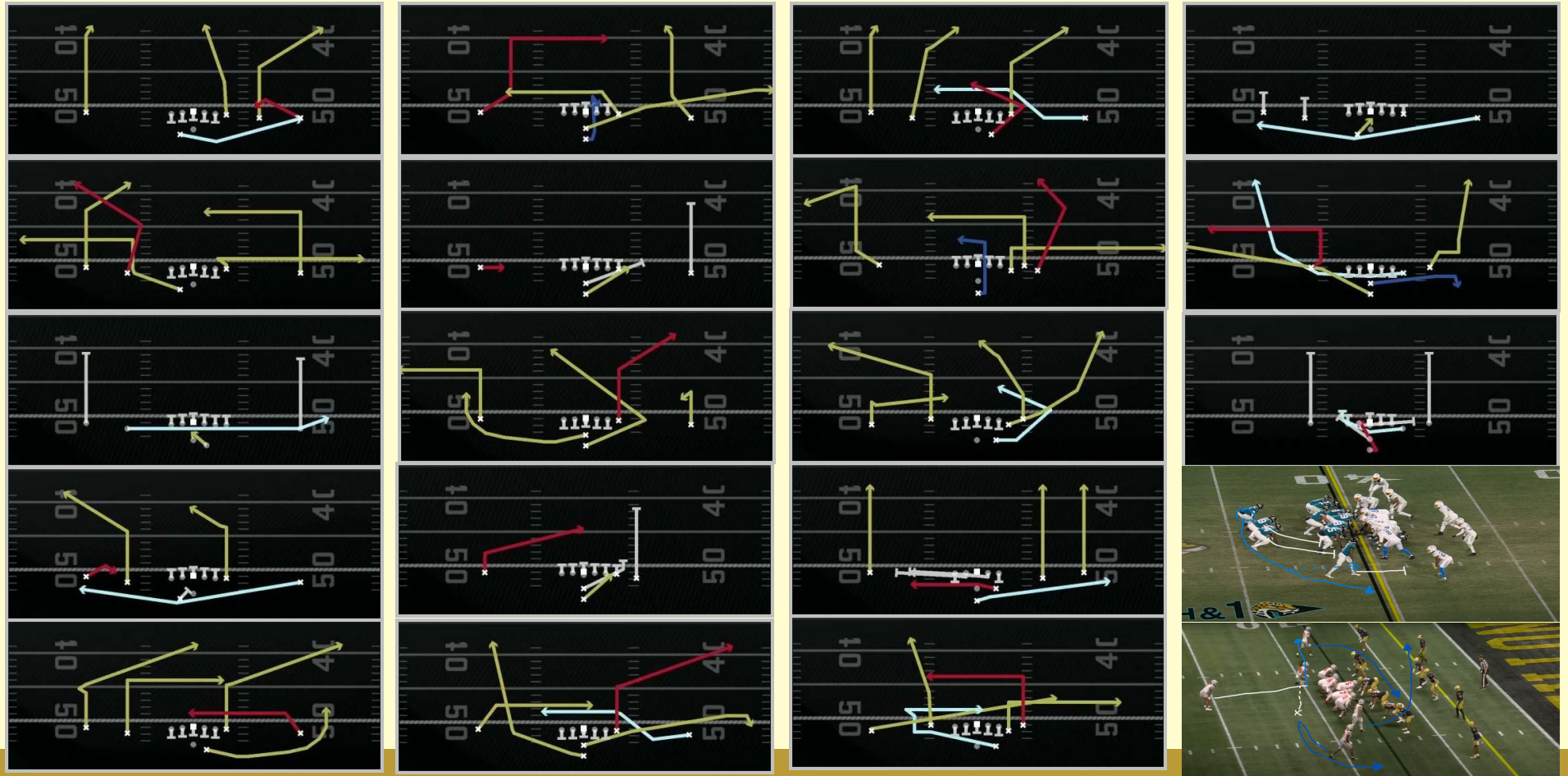


Compiling all of this together

Real Life Application - NFL Examples



- Several top NFL offenses already apply these principles, resulting in elite efficiency and innovation.
- Unsurprisingly, the most effective units are led by play callers regarded as offensive masterminds: **Sean McVay (Rams), Ben Johnson (Lions), Mike McDaniel (Dolphins).**
- These offenses consistently leverage:
  - Heavy pre-snap motion usage
  - Diverse play calls from consistent formations/personnel
  - Smart route combinations and a balanced run scheme (inside vs. outside)
- Merge this with the **aggressiveness** of Dan Campbell (oft criticized by results-based thinkers for attempting too many 4th downs and constantly creating tricks and exotic plays)



Play Design - Example Plays

Real Life Application - Roster Construction

- **Wide Receivers (WR):**
  - **X:** Vertical threat who stretches the field and runs the full deep route tree.
  - **Z:** Route technician with strong spatial awareness to exploit coverage gaps in the intermediate game.
  - **Y (Slot):** YAC-focused speedster who thrives on quick throws and checkdowns, turning short gains into explosive plays.
- **Tight Ends (TE):**
  - Versatile blockers who can hold the edge, run seam routes, and serve as reliable short/intermediate possession targets.
- **Running Backs (RB):**
  - Ideally in a split-backfield or as a modern workhorse: perimeter runner with burst, solid pass protection skills, route-running ability out of the backfield, and the intelligence to diagnose leverage and beat coverage as a checkdown option.
- **Quarterback (QB):**
  - Decision-maker who can read defenses pre- and post-snap, anticipate pressure, identify coverages, and—ideally—offer mobility to keep QB run concepts in the playbook.

Depth & Flexibility: Preparing for In-Game Adaptation

- Depth across position types ensures continuity when faced with injuries or shifting game scripts.
- Prioritize flexibility: WRs who can move inside/outside, RBs who can line up in the slot, and TEs who can block in-line or flex out.
- Roster construction should reflect the same principles of efficiency and versatility that the playbook is built around.
- Pair your most effective play designs with the personnel that brings them to life.
- Be aggressive

Thank You for Your Time!