LS88: Sports Analytics

Football

Outline

- ➔ Comparisons to baseball/basketball
- ➔ What are some simple ways to quantify performance?
- ➔ What do they miss or fail to capture about the game?
  How much does the underlying strategic complexity obfuscate performance measurement?
- ➔ What kind of evaluation can we actually do?

Emphasis on an overview of football and (some of) the current state of things

Modeling for Football

A first task is to consider how we want to model football

Let’s look at how we modeled baseball and basketball

Modeling for Baseball

- ➔ Discrete plays
  PAs are sequential and not fluid
- ➔ Discrete Outcomes
  Can succinctly label the outcome without a lot of loss of information
- ➔ We assume the sequencing of events is a matter of luck
  Thus our evaluations have been mostly around removing this component, ie. no RBIs
- ➔ More generally: we aim to be context independent
  Score, who's on base, etc doesn't matter
- ➔ If we do care, then we do it separately
  Analyzing clutchness, context-dependent stats, analyzing matchups

Modeling for Basketball

- ➔ Basketball is far more fluid

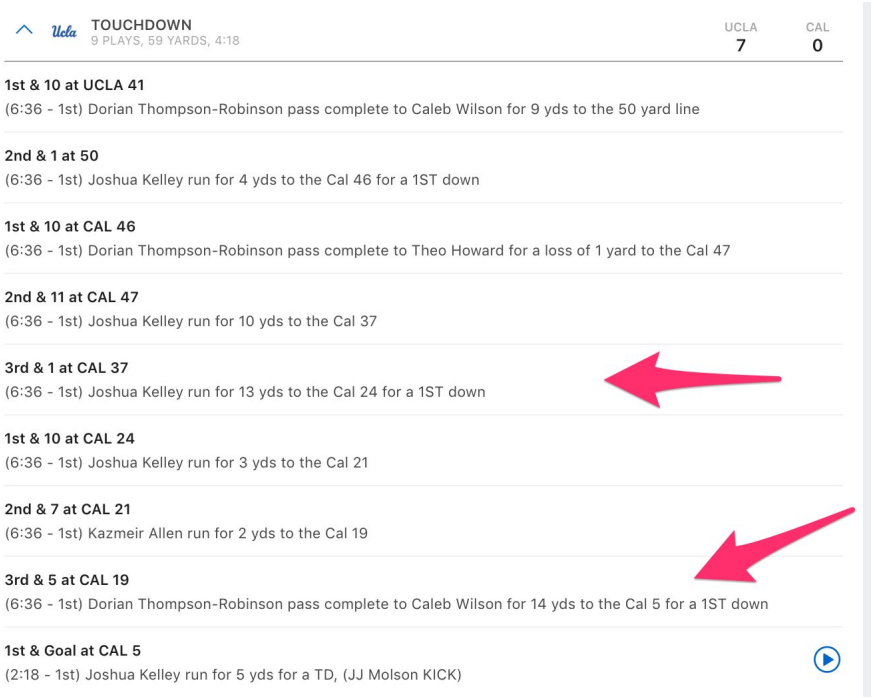
  10 players moving in coordination in a plane: (x, y) coordinates in time
- ➔ Discrete aspects of basketball
  - ◆ Possessions
  - ◆ Shooting
- ➔ We still use context independence
  - ◆ Shooting, PER: when, where, how a contribution is made doesn't matter
- ➔ Our objective for basketball: broaden the context

  Who's on court, coordination through screens/passes, etc

Challenges in Football: Third Downs Matter

DOWNS 6 PLAYS, 27 YARDS, 2:52	UCLA 20	CAL 7
(6:03 - 3rd) JJ Molson kickoff for 65 yds for a touchback		
1st & 10 at CAL 25 (6:03 - 3rd) Brandon McIlwain pass complete to Nikko Remigio for 8 yds to the Cal 33		
2nd & 2 at CAL 33 (6:03 - 3rd) Patrick Laird run for 10 yds to the Cal 43 for a 1ST down		
1st & 10 at CAL 43 (6:03 - 3rd) Patrick Laird run for 4 yds to the Cal 47		
2nd & 6 at CAL 47 (6:03 - 3rd) Brandon McIlwain run for 4 yds to the UCLA 49		
3rd & 2 at UCLA 49 (6:03 - 3rd) Christopher Brown Jr. run for 1 yd to the UCLA 48		
4th & 1 at UCLA 48 (6:03 - 3rd) Christopher Brown Jr. run for no gain to the UCLA 48		

Challenges in Football: Third Downs Matter



TOUCHDOWN
9 PLAYS, 59 YARDS, 4:18

UCLA	CAL
7	0

1st & 10 at UCLA 41
(6:36 - 1st) Dorian Thompson-Robinson pass complete to Caleb Wilson for 9 yds to the 50 yard line

2nd & 1 at 50
(6:36 - 1st) Joshua Kelley run for 4 yds to the Cal 46 for a 1ST down

1st & 10 at CAL 46
(6:36 - 1st) Dorian Thompson-Robinson pass complete to Theo Howard for a loss of 1 yard to the Cal 47

2nd & 11 at CAL 47
(6:36 - 1st) Joshua Kelley run for 10 yds to the Cal 37

3rd & 1 at CAL 37
(6:36 - 1st) Joshua Kelley run for 13 yds to the Cal 24 for a 1ST down

1st & 10 at CAL 24
(6:36 - 1st) Joshua Kelley run for 3 yds to the Cal 21

2nd & 7 at CAL 21
(6:36 - 1st) Kazmeir Allen run for 2 yds to the Cal 19

3rd & 5 at CAL 19
(6:36 - 1st) Dorian Thompson-Robinson pass complete to Caleb Wilson for 14 yds to the Cal 5 for a 1ST down

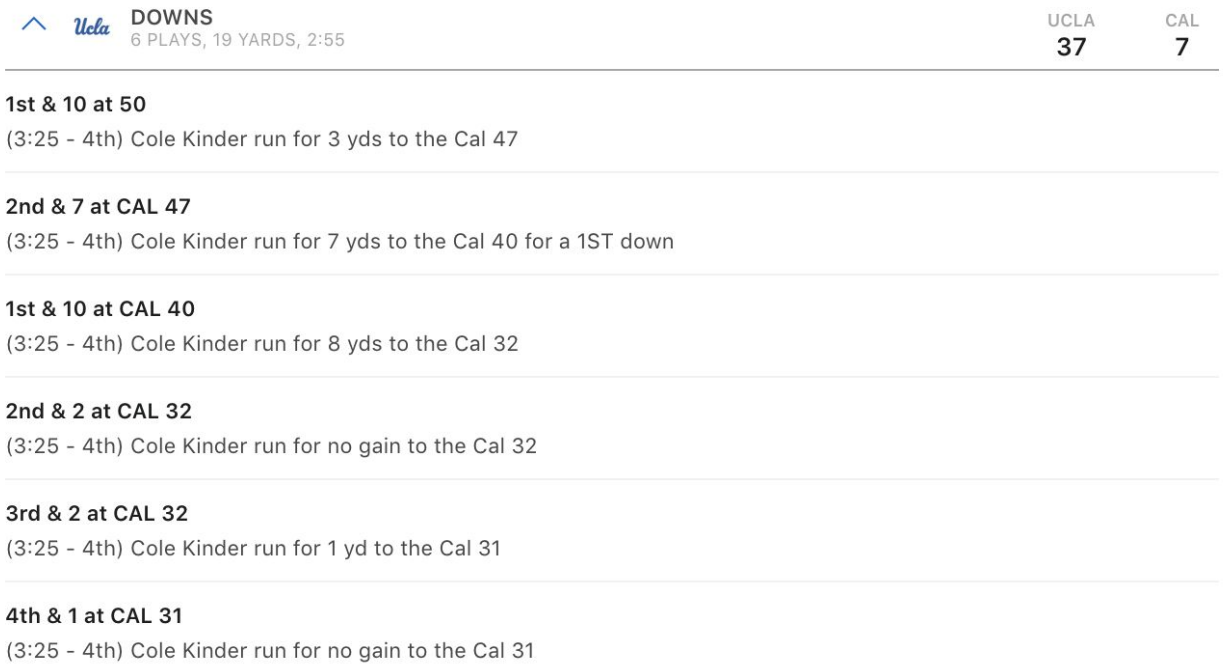
1st & Goal at CAL 5
(2:18 - 1st) Joshua Kelley run for 5 yds for a TD, (JJ Molson KICK)

Challenges in Football: One Yard Runs

DOWNS 6 PLAYS, 27 YARDS, 2:52	UCLA 20	CAL 7
(6:03 - 3rd) JJ Molson kickoff for 65 yds for a touchback		
1st & 10 at CAL 25 (6:03 - 3rd) Brandon McIlwain pass complete to Nikko Remigio for 8 yds to the Cal 33		
2nd & 2 at CAL 33 (6:03 - 3rd) Patrick Laird run for 10 yds to the Cal 43 for a 1ST down		
1st & 10 at CAL 43 (6:03 - 3rd) Patrick Laird run for 4 yds to the Cal 47		
2nd & 6 at CAL 47 (6:03 - 3rd) Brandon McIlwain run for 4 yds to the UCLA 49		
3rd & 2 at UCLA 49 (6:03 - 3rd) Christopher Brown Jr. run for 1 yd to the UCLA 48		
4th & 1 at UCLA 48 (6:03 - 3rd) Christopher Brown Jr. run for no gain to the UCLA 48		

Challenges in Football: One Yard Runs

Challenges in Football: Garbage Time



DOWNS 6 PLAYS, 19 YARDS, 2:55	UCLA 37	CAL 7

1st & 10 at 50
(3:25 - 4th) Cole Kinder run for 3 yds to the Cal 47

2nd & 7 at CAL 47
(3:25 - 4th) Cole Kinder run for 7 yds to the Cal 40 for a 1ST down

1st & 10 at CAL 40
(3:25 - 4th) Cole Kinder run for 8 yds to the Cal 32

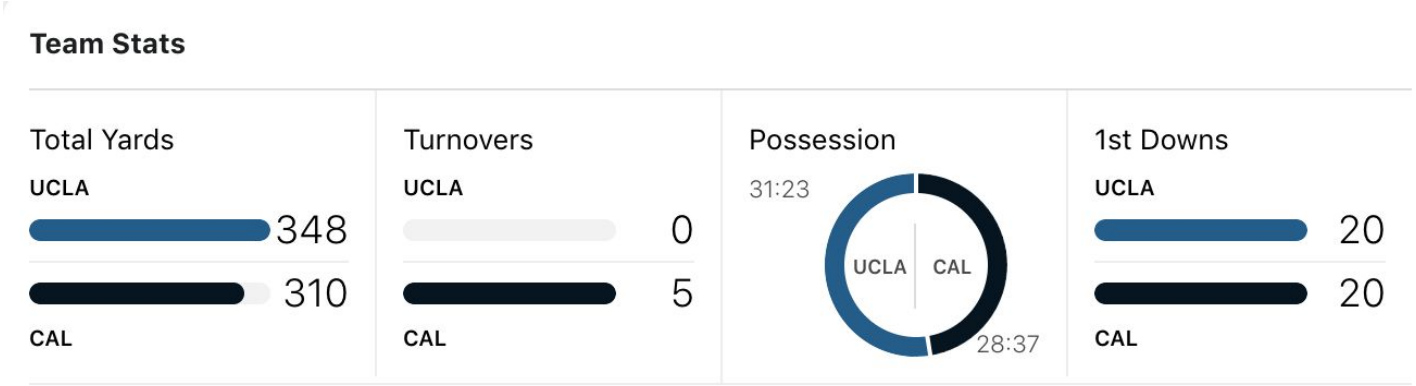
2nd & 2 at CAL 32
(3:25 - 4th) Cole Kinder run for no gain to the Cal 32

3rd & 2 at CAL 32
(3:25 - 4th) Cole Kinder run for 1 yd to the Cal 31

4th & 1 at CAL 31
(3:25 - 4th) Cole Kinder run for no gain to the Cal 31

Simple Performance Measures

➔ Yards, first downs, time of possession



Simple Performance Measures

Matchup	Ucla	Cal
1st Downs	**20**	**20**
3rd down efficiency	5-14	3-14
4th down efficiency	1-2	1-5
Total Yards	**348**	**310**
Passing	**141**	**179**
Comp-Att	13-15	23-42
Yards per pass	9.4	4.3
Interceptions thrown	0	2
Rushing	**207**	**131**
Rushing Attempts	55	33
Yards per rush	3.8	4.0
Penalties	**4-30**	**3-36**
Turnovers	**0**	**5**
Fumbles lost	0	3
Interceptions thrown	0	2
Possession	**31:23**	**28:37**

Simple Performance Measures

- ➔ Individual performance
  - Yards (rushing, passing, receiving)
  - TDs, Ints
  - Yards per Pass Attempt, Yards After Catch (YAC), Yards per Rush
  - Passer Rating
- ➔ Pros and cons
  - Generally speaking, this stuff works: more yards is good, interceptions are bad, etc
  - Yards aren't always everything though
  - Good: 5 yard run on 3rd and 3, Not good: 5 yard run on 3rd and 13

Performance Evaluation: QB Passer Rating

The mythical and mysterious QB Passer Rating

It's a linear performance metric!

$$\text{QB Passer Rating} = 100 \times \frac{1}{6} \times (A + B + C + D),$$

$$A = \left(\frac{\text{COMP}}{\text{ATT}} - .3\right) \times 5, \quad B = \left(\frac{\text{YDS}}{\text{ATT}} - 3\right) \times .25,$$

$$C = \left(\frac{\text{TD}}{\text{ATT}}\right) \times 20, \qquad D = 2.375 - \frac{\text{INT}}{\text{ATT}} \times 25$$

Performance Evaluation: QB Passer Rating

$$\text{QB Passer Rating} = 100 \times \frac{1}{6} \times (A + B + C + D),$$

$$A = \left(\frac{\text{COMP}}{\text{ATT}} - .3 \right) \times 5, \quad B = \left(\frac{\text{YDS}}{\text{ATT}} - 3 \right) \times .25,$$

$$C = \left(\frac{\text{TD}}{\text{ATT}} \right) \times 20, \qquad D = 2.375 - \frac{\text{INT}}{\text{ATT}} \times 25$$

The good: Comp / Att, Yds / Att, TD / Att
The bad: Int / Att

Performance Evaluation: QB Passer Rating

$$\text{QB Passer Rating} = 100 \times \frac{1}{6} \times (A + B + C + D),$$

$$A = \left(\frac{\text{COMP}}{\text{ATT}} - .3\right) \times 5, \quad B = \left(\frac{\text{YDS}}{\text{ATT}} - 3\right) \times .25,$$

$$C = \left(\frac{\text{TD}}{\text{ATT}}\right) \times 20, \qquad D = 2.375 - \frac{\text{INT}}{\text{ATT}} \times 25$$

Devised in the early 1970s, it was used as a ranking system for QBs
Each of the 4 components was calibrated on data from the 60s to average 1
Overall average has grown from 66.7 to 88.6

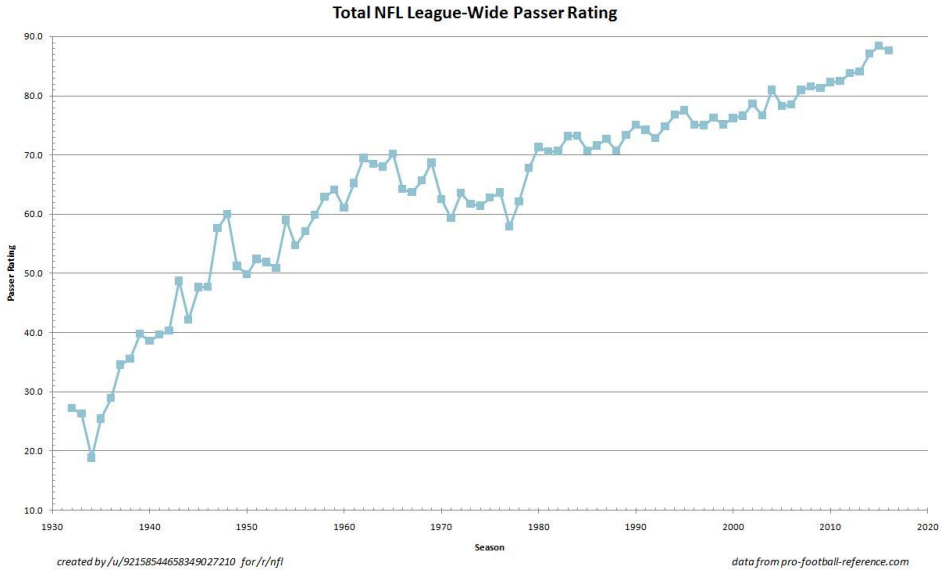
Performance Evaluation: QB Passer Rating

Passer Rating misses a lot of things about QB play

- ➔ Rushing
- ➔ Sacks
- ➔ Fumbles
- ➔ Crucial situations (3rd down or 4th quarter, if this matters to you)
- ➔ Quality of teammates (WRs and O-Line)
- ➔ Attribution of performance (interceptions and yards)

Performance Evaluation: QB Passer Rating

Passer Rating generally captures quality though: it's better than nothing



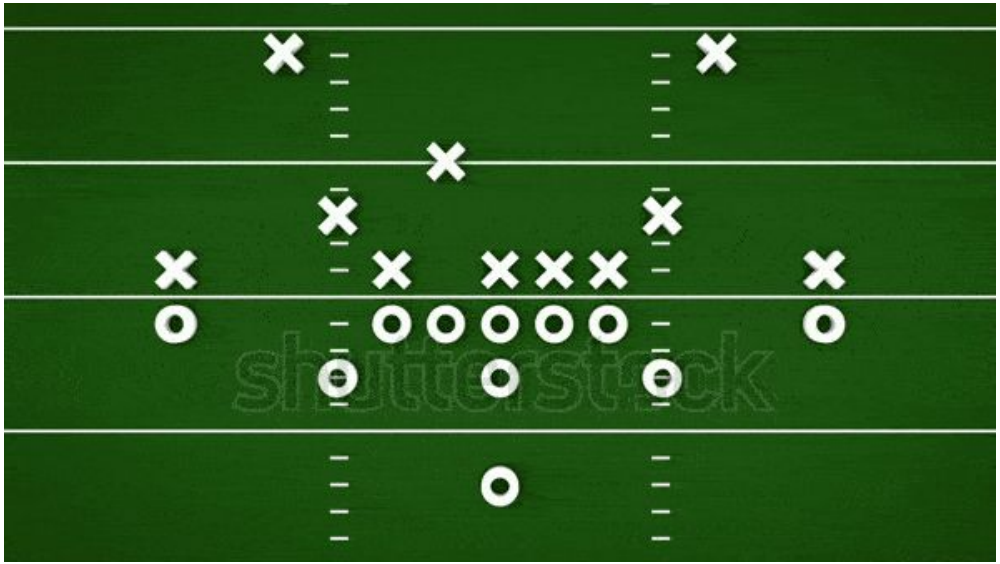
Complexity of Football

- ➔ Ultimately we want to evaluate *players*, not just teams
  - ◆ As we'll see, this is not an easy task
- ➔ 22 players moving in a plane
  - ◆ But each play is an intricate ballet of coordinated movement
  - ◆ Relatively few discrete, isolated actions
- ➔ *HUGE* tactical/strategic considerations
  - ◆ What was the player *supposed* to be doing matters
    - Blocking? Route? Coverage?
  - ◆ Lots of players will not show up in anything we conventionally think about in a play
    - 5+ players blocking on a 5 yard run

Plays and Strategy

Play Diagrams

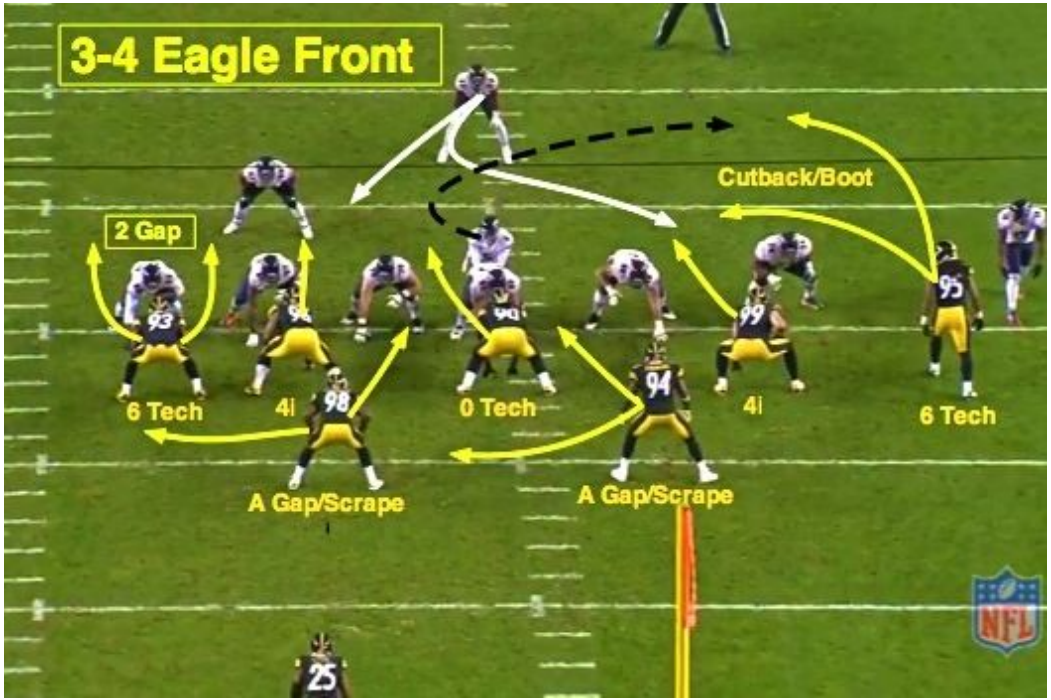
- ➔ Roles, assignments, actions for each player



Plays and Strategy

Play Diagrams

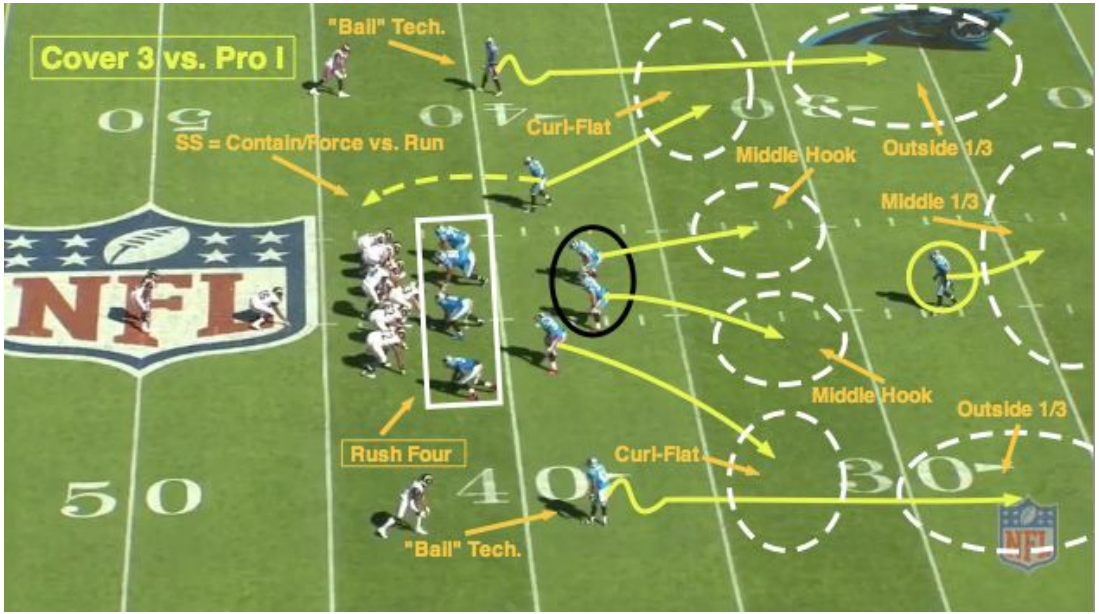
- ➔ Roles, assignments, actions for each player



Plays and Strategy

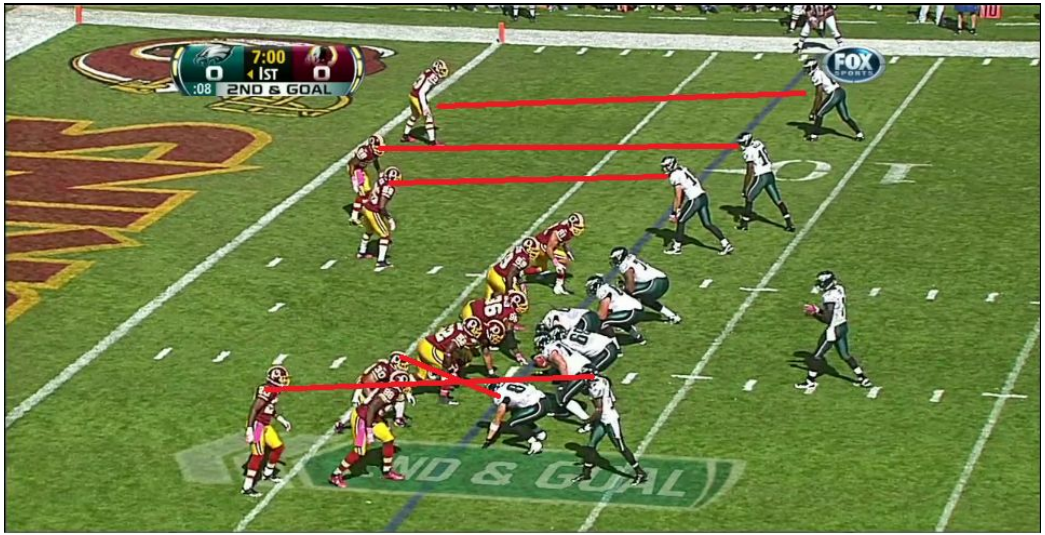
Play Diagrams

- ➔ Roles, assignments, actions for each player



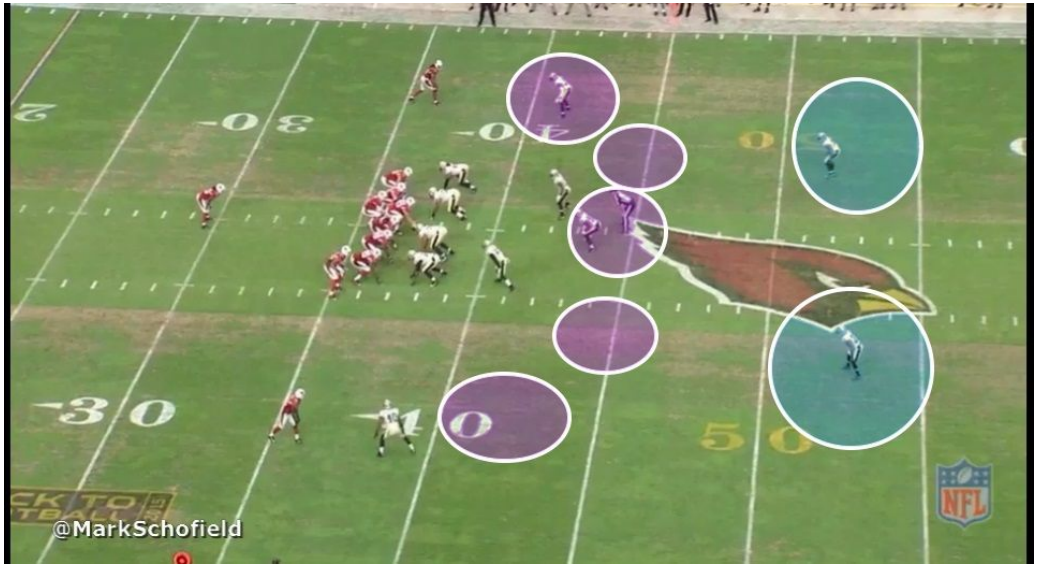
Defensive Coverage: Man

Defensive backs are assigned a player to cover



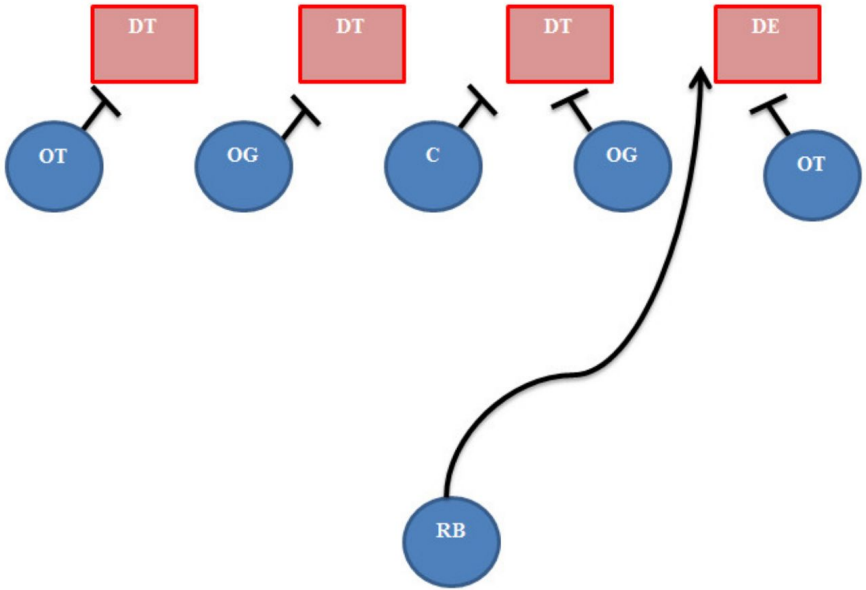
Defensive Coverage: Zone

Defensive backs are assigned a *zone* to cover



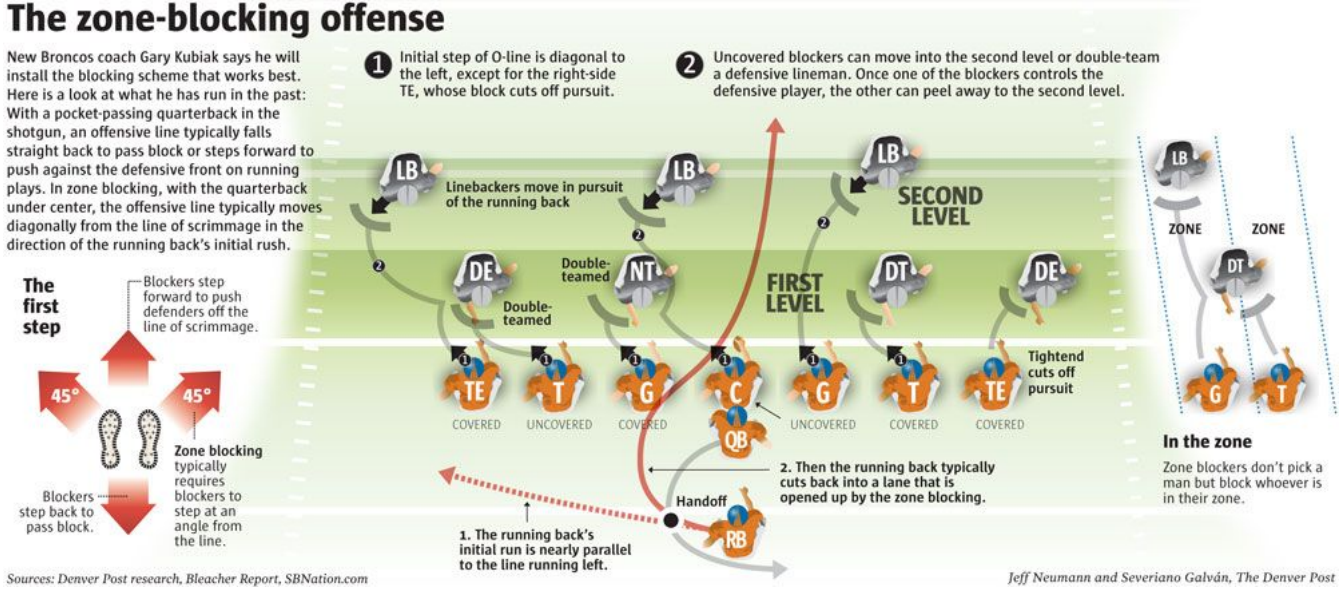
Blocking: Man

Linemen are assigned a player to block



Blocking: Zone

Linemen are assigned a *zone* and to block anyone in the zone



Playcalling & Strategy

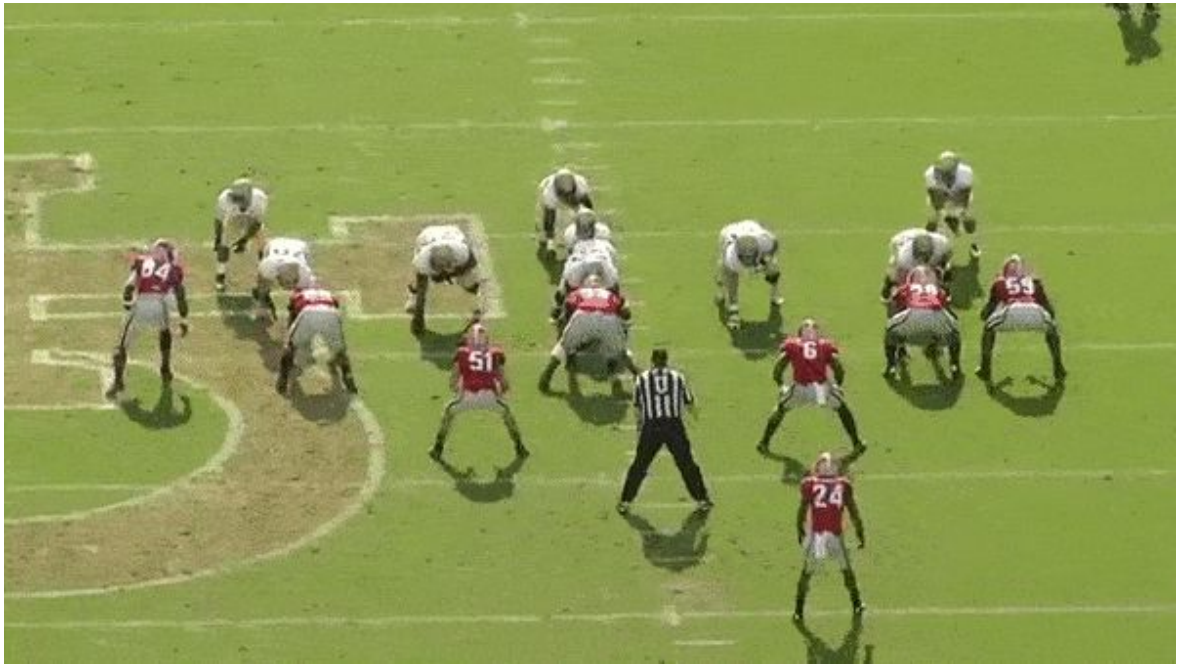
- ➔ Executing offensive plays is part of a broader strategy
- ➔ Multiple playcalls for each play
  - ◆ QBs use audibles through keywords
    Omaha, Kill, Blue 42, etc
- ➔ Sequencing, callbacks, and obfuscation
  - ◆ Use different plays from within a single formation
  - ◆ Attack an overly aggressive defense
    Counters & Reverses
  - ◆ Option plays create a built in attack against an overly aggressive defense

Counters

Reverses

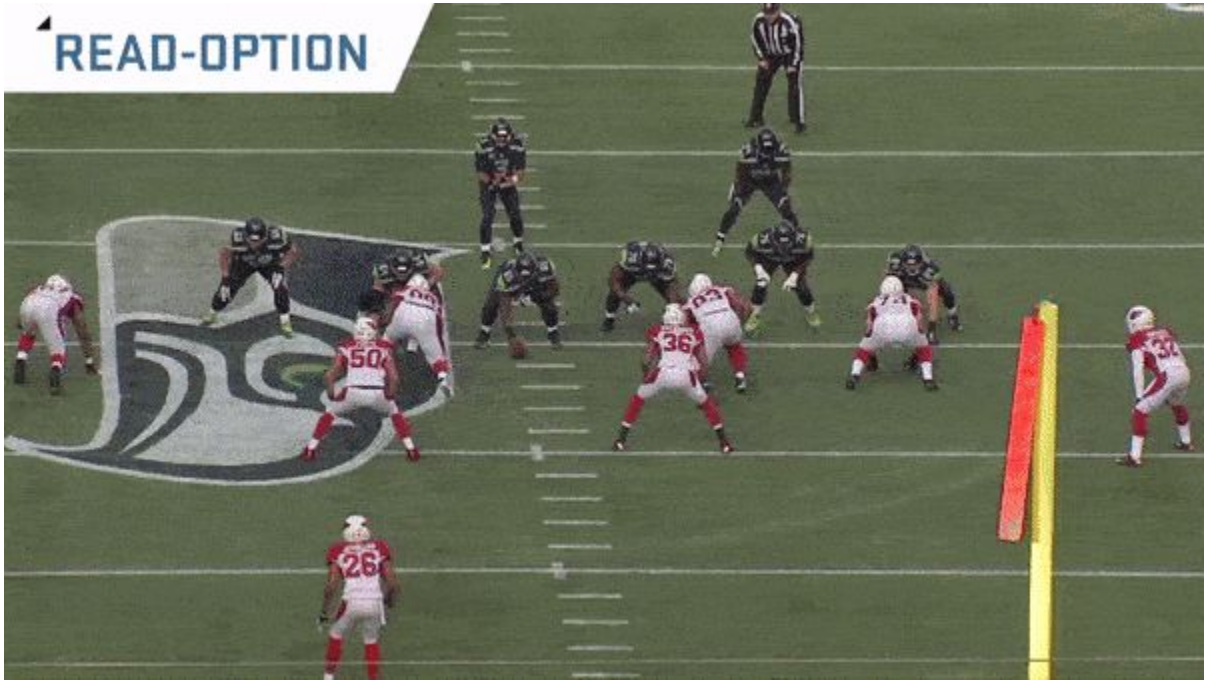

22 WASHINGTON ST 8-2 24
10 COLORADO 8-2 31
4TH 6:24
1ST & 10
FOX CFB
1ST & 10
2 OHIO ST 17
MICHIGAN ST 16
FINAL

Option Plays: Triple Option

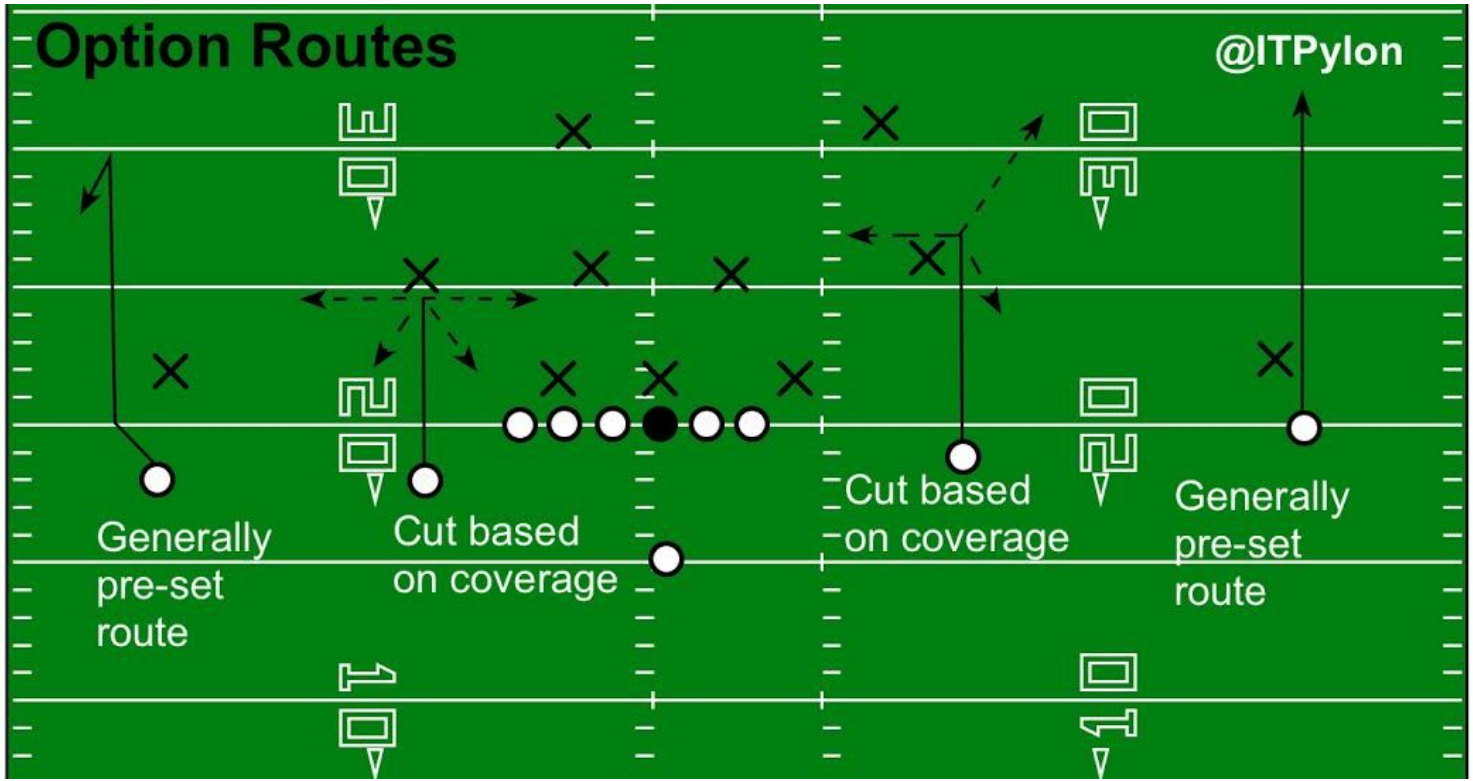


Option Plays: Read Option

Option Plays: Read Option



Option Plays: Option Routes



Option Plays: Run-Pass Option

Why do we care about the strategy?

This is *domain knowledge*

➔ Knowing the game means knowing what to try to do
➔ And what's not really going to work

Now that we know there are so many moving parts, we better account for this if we're going to evaluate players!

But the game is so complex, evaluation is inevitably going to miss a lot.

So what *can* we do?

Performance Evaluation

- ➔ Scouting/grading/subjective evaluation
  - Use film, use “experts”, use domain knowledge
  - Build a rating by inferring (as best you can) what was supposed to happen
- ➔ Quantitative evaluation
  - Collect performance data
  - Use domain knowledge, but for smarter modeling choices
  - Data reveals the evaluation instead of grades/ratings
  - We can also use this to rate overall team performance (coaching+players+scheme). Maybe not the most useful for a coach or team but as a fan or for gambling purposes, it’s something

Pro Football Focus

- ➔ Uses domain knowledge through scouting/grading
  - ◆ Analysts watch the video and grade performance on plays
  - ◆ Positive for good plays, Negative for bad
- ➔ All that complexity in strategy? Conventional data not tracking linemen or other key play features?
- ➔ Using humans, can easily build in more about the play

  Dropped ball but blown coverage, linemen blocking, route running, option decisions, etc

Pro Football Focus

- ➔ There is no free lunch though
- ➔ It most certainly has an element of subjectiveness
  - ◆ Also bias from judges
  - ◆ Proprietary and opaque too (can you even review/challenge their grade??)
- ➔ PFF claims it doesn't blindly/excessively infer the play
  - ◆ Rein in subjectivity by apparently getting analysts to be certain
  - ◆ If they're uncertain, they make no grade (that's a good rule!)
  - ◆ But can we even query any of this?

Evaluation via Film

Remember, there is no free lunch…

Bill Belichick on evaluating performance on film

> *It might even look to us like somebody made a mistake but then we look at it more closely maybe somebody besides him made a mistake and he was trying to compensate. … You think a player did something that he shouldn't have done but maybe he got a call, a line call or a call from a linebacker or he thought the quarterback said something so he did what he thought was the right thing or maybe it was the right thing but that call shouldn't have been made or should have been on the other side.*

Evaluation via Film

More Bill Belichick on evaluating performance on film

> *…There are plenty of plays where I have no idea what went wrong. …These two guys made a mistake but I don't know which guy it was or if it was both of them. … I don't know how you can know that unless you're really part of the team and know exactly what was supposed to happen on that play.* ***I know there are a lot of experts out there that have it all figured out but I definitely don't.***

QBR

- ➔ Proprietary ESPN quantitative metric

  Developed by our friend Dean Oliver!
- ➔ Some notable features
  - ◆ Incorporates all contributions: passes, rushes, turnovers, penalties
  - ◆ Accounts for each play's performance, ie. 5 yards on 3rd and 13 isn't treated the same as 5 yards on 3rd and 3
  - ◆ Accounts for credit attribution, ie. yards after catch only given based on expected outcome
  - ◆ Adjusts for play difficulty
  - ◆ Adjusts for garbage time (no longer provides a boost for "clutch" plays)
  - ◆ Adjusts for opponent

DVOA Rating

- ➔ DVOA stands Defense-adjusted Value Over Average
- ➔ …Value Over Average … sound familiar?
- ➔ DVOA is a proprietary quantitative metric from *Football Outsiders*
- ➔ To understand it, start with the question (or Zen kōan?)
  - ◆ One running back runs for three yards.
  - ◆ Another running back runs for three yards.
  - ◆ Which is the better run?

This will be a terse overview covering the main ideas and how they relate to what we've seen
See more here: http://www.footballoutsiders.com/info/methods

DVOA Rating: What is the context?

- ➔ What is the down and distance?
- ➔ Is it third-and-2 or second-and-15?
- ➔ Where on the field is the ball?
- ➔ Does the player get only three yards because he hits the goal line and scores?
- ➔ Is the player's team running out the clock and facing a tough run defense?
- ➔ Or is the team down big and facing a pass defense?
- ➔ Is the opposing defense porous or really good?

Motivating DVOA Rating

- ➔ Conventional NFL stats measure performance with net yards
- ➔ You really want to score points by getting in the endzone.
- ➔ Yards are a means to an ends but not all yards are created equal
- ➔ 6 yards on 3rd and 10 isn't helpful.
  - ◆ The drive still probably ended and the defense gave it to you to prevent the first down
- ➔ Fantasy doesn't help: 1 yard TD run nets the same as a WR who breaks tackles to go 60 yards but stopped short of the goalline. Why does the RB get the TD credit?

DVOA Rating Goals

- ➔ DVOA distributes credit for scoring points and winning games
- ➔ DVOA assigns every play a value based on total yards and yards towards a first down
- ➔ Based on *The Hidden Game of Football* by Pete Palmer, Bob Carroll, and John Thorn

DVOA Rating: Value Assignments

THGF value assignments

- ➔ First down: play is considered a success if it gains 45 percent of needed yards
- ➔ Second down: play needs to gain 60 percent of needed yards
- ➔ Third or fourth down: only gaining a new first down is considered success

DVOA builds on the THGF value assignments

- ➔ Fractional points and adjustments for big plays
- ➔ Red zone performance
- ➔ Other areas of exceptional importance

The VOA in DVOA Rating

Return to the original question

- ➔ Player A gains 3 yards in a situation when an average RB gains only one yard
  Player A has value above others at his position
- ➔ Player B gains 3 yards in a situation when an average NFL back gains four yards
  Player B has negative value relative to others at his position.

After adjustments, we have something akin to RE24

Cumulative value contribution based on expected outcome

The VOA in DVOA Rating

To get *Value Over Average*

1. Add up every play by a certain team or player
2. Divide by the total of the various baselines for success in all those situations

The D in DVOA Rating

- ➔ We didn't do this with baseball but DVOA is defense-adjusted
- ➔ This is far more important with an unbalanced schedule like in football
- ➔ Each play is adjusted according to the defense's ability to defend that play
- ➔ Defenses are also rated by the quality of the opposing offenses
  - ◆ Apparently they stuck with the name defense-adjusted instead of offense-adjusted
- ➔ Final scale of around +/- 30% for Off and +/- 25% for Def
  - ◆ 0% is average

Expected Points

We can finish off with an insight into the quantitative approach

- ➔ Expected Points is the foundation of QBR and other models
- ➔ Fundamental to evaluation
  - ◆ Expected Points Added is like RE24: how effective was a play in expectation
- ➔ Expected Points and decision making
  - ◆ Try to maximize points (or more generally, win probability)
  - ◆ The NYT 4th Down Bot

Expected Points

- ➔ We have the ball 1st and 10 at our own 42 yard line
  - How many points do we expect the next score *in the game* to be worth to us?
  - Suppose it's +3 points: this possession for us is worth +3 points
- ➔ We get the ball 1st and 10 at our own 1 yard line
  - We have a long way to go to score
  - We're likely to give the ball back to the other team with great position for them
  - The expected points here is negative, ie. in favor of our opponent
- ➔ We get the ball 1st and goal at our the opponent's 1 yard line
  - We're probably going to score a touchdown (we need just 1 yard!)
  - The expected points here is about +6 in our favor

Expected Points

We'll do this to how we did it for Run Expectancy

1. Our State: Kickoff and the start of a drive (1st and 10 from the …)

   We restrict to the first and third quarters only
2. Our value: next score in the game after the drive started

   Similar to runs scored in remainder of inning
3. Average: by yardline, so average the value of all drives starting at the 42 yardline

Summary

- ➔ The basic evaluations from box scores: Yards, TDs, Ints, etc
- ➔ Passer Rating is out-dated but isn't totally worthless
- ➔ The inherent strategy and coordination makes player evaluation comparatively really hard
- ➔ A "scouting" approach can be used ala Pro Football Focus
- ➔ A "quantitative" approach can be used ala Football Outsiders or ESPN
- ➔ The holy grail: automatic tracking and evaluation via advanced modern technology