

# Statistics in Sports: (American) Football Overview

---

ZACHARY BINNEY, PHD MPH

OXFORD COLLEGE OF EMORY UNIVERSITY

FALL 2021

# Roadmap

---

- 1. Football is Complex and Chaotic
- 2. Predicting Probabilities of Events
  - Field Goal Model
  - Expected Points Added (EPA)
  - Win Probability Added (WPA)
- 3. Fourth Down Decisions
- 4. Draft Analytics
- 5. NFL Player Valuation Analytics
  - Overview of Passing Stats
    - Box Score Stats, ANY/A
    - Passer Rating, (Total) QBR
    - Air Yards/aDOT vs. YAC
    - EPA/WPA and Translation to WAR
  - 6. Tracking/Next Gen Stats (NGS) Data
    - Play Expectations
    - Player Valuation vs. Expectation
    - Big Data Bowl
  - 7. Lineman Analytics: Pass Blocking and Rushing
    - Survival Analysis

# Intro

---

# “Teams That Run Win More”



- We start with a puzzle: why are these bullshit?

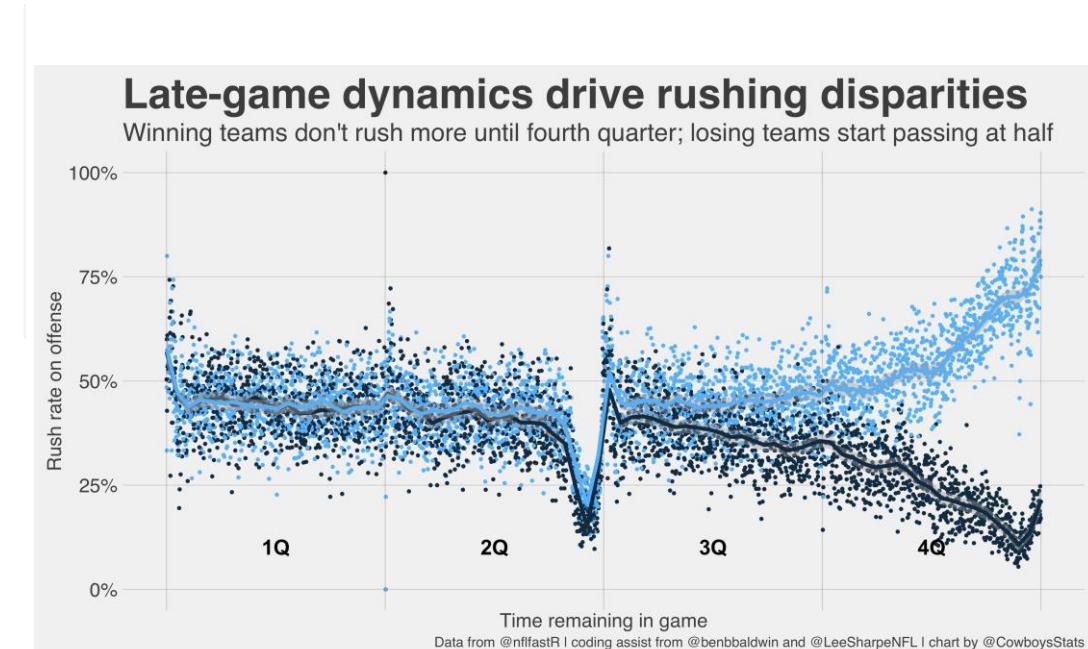


Cowboys Stats & Graphics  
@CowboysStats

You've heard arguments that go something like this:  
"When Ezekiel Elliott gets 25+ carries, his team is 19-4."

This is explained by the incentives teams face late in games. Teams run more when they're already likely to win.

Light blue = winning teams  
Dark blue = losing teams



# Football is Complex and Chaotic

---

- Let's break down a play: Seahawks @ Vikings, 9/26/21
  - MIN 21-SEA 17; **3rd & 4 at SEA 34**

(10:02 - 3rd) (Shotgun) K.Cousins pass short left to J.Jefferson ran ob at SEA 26 for 8 yards.



# Data Contained in a Football Play

---

- Let's break down a play: Seahawks @ Vikings, 9/26/21
  - MIN 21-SEA 17

**3rd & 4 at SEA 34**

(10:02 - 3rd) (Shotgun) K.Cousins pass short left to J.Jefferson ran ob at SEA 26 for 8 yards.

- How could we break down this film and/or play-by-play text into analytical “nuggets” for a computer? In other words, how can we “encode” something like this for analysis?
- **“Tokenization”** (text analysis, not data security)

# Data Contained in a Football Play

- Let's break down a football play:
  - OK, go!

- MIN

1st & 10 at SEA

(12:53 - 4th) (No  
at SEA 44. The R  
T.Lockett to SEA

e:



# Data Contained in a Football Play

- Possible fields for one play from nflfastR play-by-play analysis

```

names(data)
#> [1] "play_id"
#> [2] "game_id"
#> [3] "old_game_id"
#> [4] "home_team"
#> [5] "away_team"
#> [6] "season_type"
#> [7] "week"
#> [8] "postteam"
#> [9] "postteam_type"
#> [10] "defteam"
#> [11] "side_of_field"
#> [12] "yardline_100"
#> [13] "game_date"
#> [14] "quarter_seconds_remaining"
#> [15] "half_seconds_remaining"
#> [16] "game_seconds_remaining"
#> [17] "game_half"
#> [18] "quarter_end"
#> [19] "drive"
#> [20] "sp"
#> [21] "qtr"
#> [22] "down"
#> [23] "goal_to_go"
#> [24] "time"
#> [25] "yrdln"
#> [26] "ydstop"
#> [27] "ydsnet"
#> [28] "desc"
#> [29] "play_type"
#> [30] "yards_gained"
#> [31] "shotgun"
#> [32] "no_huddle"
#> [33] "qb_dropout"
#> [34] "qb_kneel"
#> [35] "qb_spike"
#> [36] "qb_scramble"
#> [37] "pass_length"
#> [38] "pass_location"
#> [39] "air_yards"
#> [40] "yards_after_catch"
#> [41] "on_location"
#> [42] "run_gap"
#> [43] "field_goal_result"
#> [44] "kick_distance"
#> [45] "extra_point_result"
#> [46] "two_point_conv_result"
#> [47] "home_timeouts_remaining"
#> [48] "away_timeouts_remaining"
#> [49] "timeout"
#> [50] "timeout_team"
#> [51] "td_team"
#> [52] "td_player_name"
#> [53] "td_player_id"
#> [54] "postteam_timeouts_remaining"
#> [55] "defteam_timeouts_remaining"
#> [56] "total_home_score"
#> [57] "total_away_score"
#> [58] "postteam_score"
#> [59] "defteam_score"
#> [60] "score_differential"
#> [61] "postteam_score_post"
#> [62] "defteam_score_post"

#> [63] "score_differential_post"
#> [64] "no_score_prob"
#> [65] "opp_fg_prob"
#> [66] "opp_safety_prob"
#> [67] "opp_td_prob"
#> [68] "fg_prob"
#> [69] "safety_prob"
#> [70] "td_prob"
#> [71] "extra_point_prob"
#> [72] "two_point_conversion_prob"
#> [73] "ep"
#> [74] "epa"
#> [75] "total_home_epa"
#> [76] "total_away_epa"
#> [77] "total_home_rush_epa"
#> [78] "total_away_rush_epa"
#> [79] "total_home_pass_epa"
#> [80] "total_away_pass_epa"
#> [81] "air_epa"
#> [82] "yac_epa"
#> [83] "comp_air_epa"
#> [84] "comp_yac_epa"
#> [85] "total_home_comp_air_epa"
#> [86] "total_away_comp_air_epa"
#> [87] "total_home_comp_yac_epa"
#> [88] "total_away_comp_yac_epa"
#> [89] "total_home_raw_air_epa"
#> [90] "total_away_raw_air_epa"
#> [91] "total_home_raw_yac_epa"
#> [92] "total_away_raw_yac_epa"
#> [93] "wp"
#> [94] "def_wp"
#> [95] "home_wp"
#> [96] "away_wp"
#> [97] "wpa"
#> [98] "vegas_wpa"
#> [99] "vegas_home_wpa"
#> [100] "home_wp_post"
#> [101] "away_wp_post"
#> [102] "vegas_wp"
#> [103] "vegas_home_wp"
#> [104] "total_home_rush_wpa"
#> [105] "total_away_rush_wpa"
#> [106] "total_home_pass_wpa"
#> [107] "total_away_pass_wpa"
#> [108] "air_wpa"
#> [109] "yac_wpa"
#> [110] "comp_air_wpa"
#> [111] "comp_yac_wpa"
#> [112] "total_home_comp_air_wpa"
#> [113] "total_away_comp_air_wpa"
#> [114] "total_home_comp_yac_wpa"
#> [115] "total_away_comp_yac_wpa"
#> [116] "total_home_raw_air_wpa"
#> [117] "total_away_raw_air_wpa"
#> [118] "total_home_raw_yac_wpa"
#> [119] "total_away_raw_yac_wpa"
#> [120] "punt_blocked"
#> [121] "first_down_rush"
#> [122] "first_down_pass"
#> [123] "first_down_penalty"
#> [124] "third_down_converted"
#> [125] "third_down_failed"
#> [126] "fourth_down_converted"
#> [127] "fourth_down_failed"
#> [128] "incomplete_pass"

#> [129] "incomplete_pass"
#> [130] "interception"
#> [131] "punt_inside_twenty"
#> [132] "punt_in_endzone"
#> [133] "punt_out_of_bounds"
#> [134] "punt_dowed"
#> [135] "punt_fair_catch"
#> [136] "kickoff_inside_twenty"
#> [137] "kickoff_in_endzone"
#> [138] "kickoff_out_of_bounds"
#> [139] "kickoff_dowed"
#> [140] "kickoff_fair_catch"
#> [141] "fumble_forced"
#> [142] "fumble_not_forced"
#> [143] "fumble_out_of_bounds"
#> [144] "solo_tackle"
#> [145] "safety"
#> [146] "penalty"
#> [147] "tackled_for_loss"
#> [148] "fumble_lost"
#> [149] "own_kickoff_recovery"
#> [150] "own_kickoff_recovery_td"
#> [151] "qb_hit"
#> [152] "rush_attempt"
#> [153] "pass_attempt"
#> [154] "sack"
#> [155] "touchdown"
#> [156] "pass_touchdown"
#> [157] "rush_touchdown"
#> [158] "return_touchdown"
#> [159] "extra_point_attempt"
#> [160] "two_point_attempt"
#> [161] "field_goal_attempt"
#> [162] "kickoff_attempt"
#> [163] "punt_attempt"
#> [164] "fumble"
#> [165] "complete_pass"
#> [166] "assist_tackle"
#> [167] "lateral_reception"
#> [168] "lateral_rush"
#> [169] "lateral_return"
#> [170] "lateral_recovery"
#> [171] "passer_player_id"
#> [172] "passer_player_name"
#> [173] "passing_yards"
#> [174] "receiver_player_id"
#> [175] "receiver_player_name"
#> [176] "receiving_yards"
#> [177] "rusher_player_id"
#> [178] "rusher_player_name"
#> [179] "rushing_yards"
#> [180] "lateral_receiver_player_id"
#> [181] "lateral_receiver_player_name"
#> [182] "lateral_receiving_yards"
#> [183] "lateral_rusher_player_id"
#> [184] "lateral_rusher_player_name"
#> [185] "lateral_rushing_yards"
#> [186] "lateral_sack_player_id"
#> [187] "lateral_sack_player_name"
#> [188] "interception_player_id"
#> [189] "interception_player_name"

#> [190] "lateral_interception_player_id"
#> [191] "lateral_interception_player_name"
#> [192] "punt_returner_player_id"
#> [193] "punt_returner_player_name"
#> [194] "lateral_punt_returner_player_id"
#> [195] "lateral_punt_returner_player_name"
#> [196] "kickoff_returner_player_name"
#> [197] "kickoff_returner_player_id"
#> [198] "lateral_kickoff_returner_player_id"
#> [199] "lateral_kickoff_returner_player_name"
#> [200] "punter_player_id"
#> [201] "punter_player_name"
#> [202] "kicker_player_name"
#> [203] "kicker_player_id"
#> [204] "own_kickoff_recovery_player_id"
#> [205] "own_kickoff_recovery_player_name"
#> [206] "blocked_player_id"
#> [207] "blocked_player_name"
#> [208] "tackle_for_loss_1_player_id"
#> [209] "tackle_for_loss_1_player_name"
#> [210] "tackle_for_loss_2_player_id"
#> [211] "tackle_for_loss_2_player_name"
#> [212] "qb_hit_1_player_id"
#> [213] "qb_hit_1_player_name"
#> [214] "qb_hit_2_player_id"
#> [215] "qb_hit_2_player_name"
#> [216] "forced_fumble_player_1_team"
#> [217] "forced_fumble_player_1_player_id"
#> [218] "forced_fumble_player_1_player_name"
#> [219] "forced_fumble_player_2_team"
#> [220] "forced_fumble_player_2_player_id"
#> [221] "forced_fumble_player_2_player_name"
#> [222] "solo_tackle_1_team"
#> [223] "solo_tackle_2_team"
#> [224] "solo_tackle_1_player_id"
#> [225] "solo_tackle_2_player_id"
#> [226] "solo_tackle_1_player_name"
#> [227] "solo_tackle_2_player_name"
#> [228] "assist_tackle_1_player_id"
#> [229] "assist_tackle_1_player_name"
#> [230] "assist_tackle_1_team"
#> [231] "assist_tackle_2_player_id"
#> [232] "assist_tackle_2_player_name"
#> [233] "assist_tackle_2_team"
#> [234] "assist_tackle_3_player_id"
#> [235] "assist_tackle_3_player_name"
#> [236] "assist_tackle_3_team"
#> [237] "assist_tackle_4_player_id"
#> [238] "assist_tackle_4_player_name"
#> [239] "assist_tackle_4_team"
#> [240] "tackle_with_assist"
#> [241] "tackle_with_assist_1_player_id"
#> [242] "tackle_with_assist_1_player_name"
#> [243] "tackle_with_assist_1_team"
#> [244] "tackle_with_assist_2_player_id"
#> [245] "tackle_with_assist_2_player_name"
#> [246] "tackle_with_assist_2_team"
#> [247] "pass_defense_1_player_id"
#> [248] "pass_defense_1_player_name"
#> [249] "pass_defense_2_player_id"
#> [250] "pass_defense_2_player_name"

#> [251] "fumbled_1_team"
#> [252] "fumbled_1_player_id"
#> [253] "fumbled_1_player_name"
#> [254] "fumbled_2_player_id"
#> [255] "fumbled_2_player_name"
#> [256] "fumbled_2_team"
#> [257] "fumble_recovery_1_team"
#> [258] "fumble_recovery_1_yards"
#> [259] "fumble_recovery_1_player_id"
#> [260] "fumble_recovery_1_player_name"
#> [261] "fumble_recovery_2_team"
#> [262] "fumble_recovery_2_yards"
#> [263] "fumble_recovery_2_player_id"
#> [264] "fumble_recovery_2_player_name"
#> [265] "sack_player_id"
#> [266] "sack_player_name"
#> [267] "half_sack_1_player_id"
#> [268] "half_sack_1_player_name"
#> [269] "half_sack_2_player_id"
#> [270] "half_sack_2_player_name"
#> [271] "return_team"
#> [272] "return_yards"
#> [273] "penalty_team"
#> [274] "penalty_player_id"
#> [275] "penalty_player_name"
#> [276] "penalty_yards"
#> [277] "replay_on_challenge"
#> [278] "replay_on_challenge_result"
#> [279] "penalty_type"
#> [280] "defensive_two_point_attempt"
#> [281] "defensive_two_point_conv"
#> [282] "defensive_extra_point_attempt"
#> [283] "defensive_extra_point_conv"
#> [284] "safety_player_name"
#> [285] "safety_player_id"
#> [286] "season"
#> [287] "cp"
#> [288] "cpoe"
#> [289] "series"
#> [290] "series_success"
#> [291] "series_result"
#> [292] "order_sequence"
#> [293] "start_time"
#> [294] "time_of_day"
#> [295] "stadium"
#> [296] "weather"
#> [297] "nfl_api_id"
#> [298] "play_clock"
#> [299] "play_deleted"
#> [300] "play_type_nfl"
#> [301] "special_teams_play"
#> [302] "st_play_type"
#> [303] "end_clock_time"
#> [304] "end_yard_line"
#> [305] "fixed_drive"
#> [306] "fixed_drive_result"
#> [307] "drive_real_start_time"
#> [308] "drive_play_count"
#> [309] "drive_time_of_possession"
#> [310] "drive_first_downs"
#> [311] "drive_inside20"

#> [312] "drive_ended_with_score"
#> [313] "drive_quarter_start"
#> [314] "drive_quarter_end"
#> [315] "drive_yards_penalized"
#> [316] "drive_start_transition"
#> [317] "drive_end_transition"
#> [318] "drive_game_clock_start"
#> [319] "drive_game_clock_end"
#> [320] "drive_start_yard_line"
#> [321] "drive_end_yard_line"
#> [322] "drive_play_id_started"
#> [323] "drive_play_id_ended"
#> [324] "away_score"
#> [325] "home_score"
#> [326] "location"
#> [327] "result"
#> [328] "total"
#> [329] "spread_line"
#> [330] "total_line"
#> [331] "div_game"
#> [332] "roof"
#> [333] "surface"
#> [334] "temp"
#> [335] "wind"
#> [336] "home_coach"
#> [337] "away_coach"
#> [338] "stadium_id"
#> [339] "game_stadium"
#> [340] "aborted_play"
#> [341] "success"
#> [342] "passer"
#> [343] "passer_jersey_number"
#> [344] "rusher"
#> [345] "rusher_jersey_number"
#> [346] "receiver"
#> [347] "receiver_jersey_number"
#> [348] "pass"
#> [349] "rush"
#> [350] "first_down"
#> [351] "special"
#> [352] "play"
#> [353] "passer_id"
#> [354] "rusher_id"
#> [355] "receiver_id"
#> [356] "name"
#> [357] "jersey_number"
#> [358] "id"
#> [359] "fantasy_player_name"
#> [360] "fantasy_player_id"
#> [361] "fantasy"
#> [362] "fantasy_id"
#> [363] "out_of_bounds"
#> [364] "home_opening_kickoff"
#> [365] "qb_epa"
#> [366] "yac_epa"
#> [367] "yac_mean_yardage"
#> [368] "yac_median_yardage"
#> [369] "xyc_success"
#> [370] "xyc_fd"
#> [371] "xpass"
#> [372] "pass_ue"
```

# Data Contained in a Football Play

- Data for one play in nflfastR

▲ play_id ▾	game_id	old_game_id	home_team	away_team	season_type	week	posteam	posteam_type	defteam	side_of_field	yardline_100	game_date	quarter_seconds_remaining	
1	2254	2021_03_SEA_MIN	2021092611	MIN	SEA	REG	3	MIN	home	SEA	SEA	34	2021-09-26	602

...

qtr	down	goal_to_go	time	yrdln	ydstogo	ydsnet	desc	play_type	yards_gained	shotgun	no_huddle	qb_dropback	qb_kneel	qb_spike
3	3	0	10:02	SEA 34	4	50	(10:02) (Shotgun) 8-K.Cousins pass short left to 18-J.Jefferso...	pass	8	1	0	1	0	0

...

# Predicting Probabilities of Events

---

FIELD GOALS

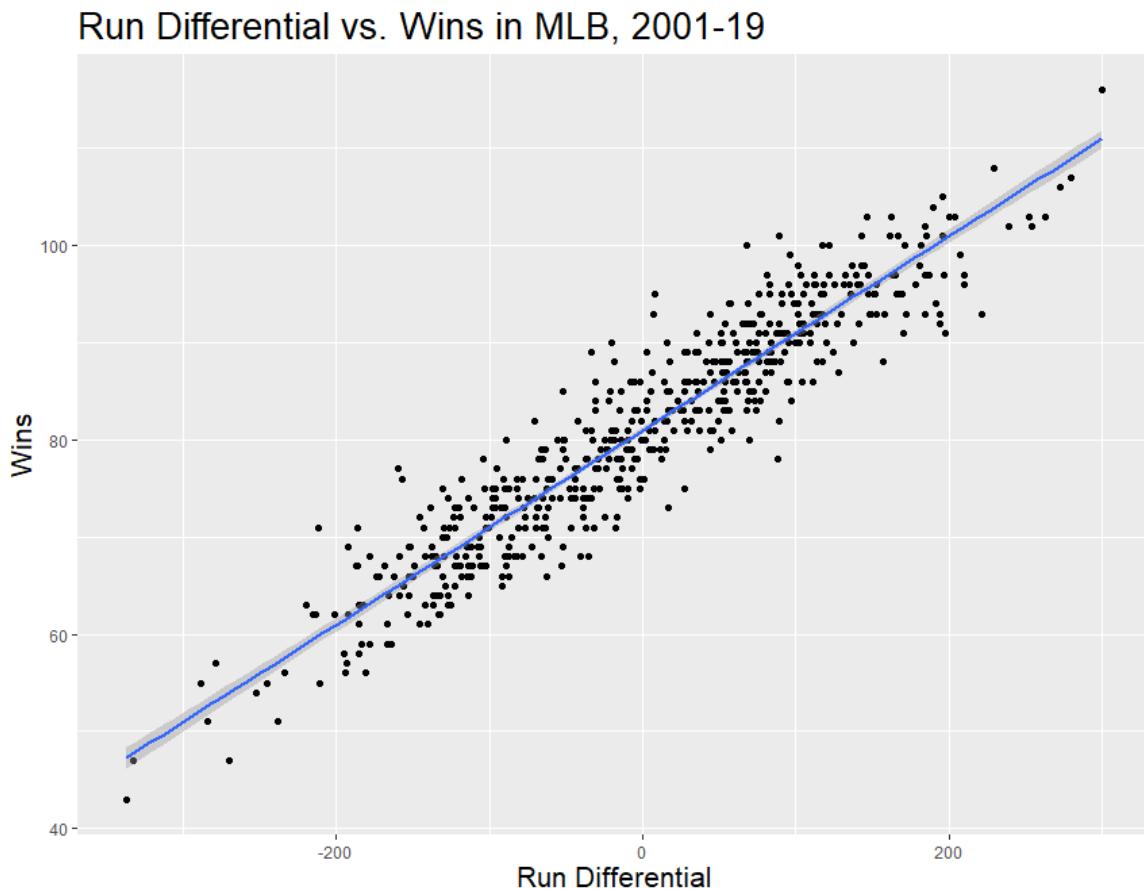
EXPECTED POINTS

WIN PROBABILITY

# Simple Linear Regression

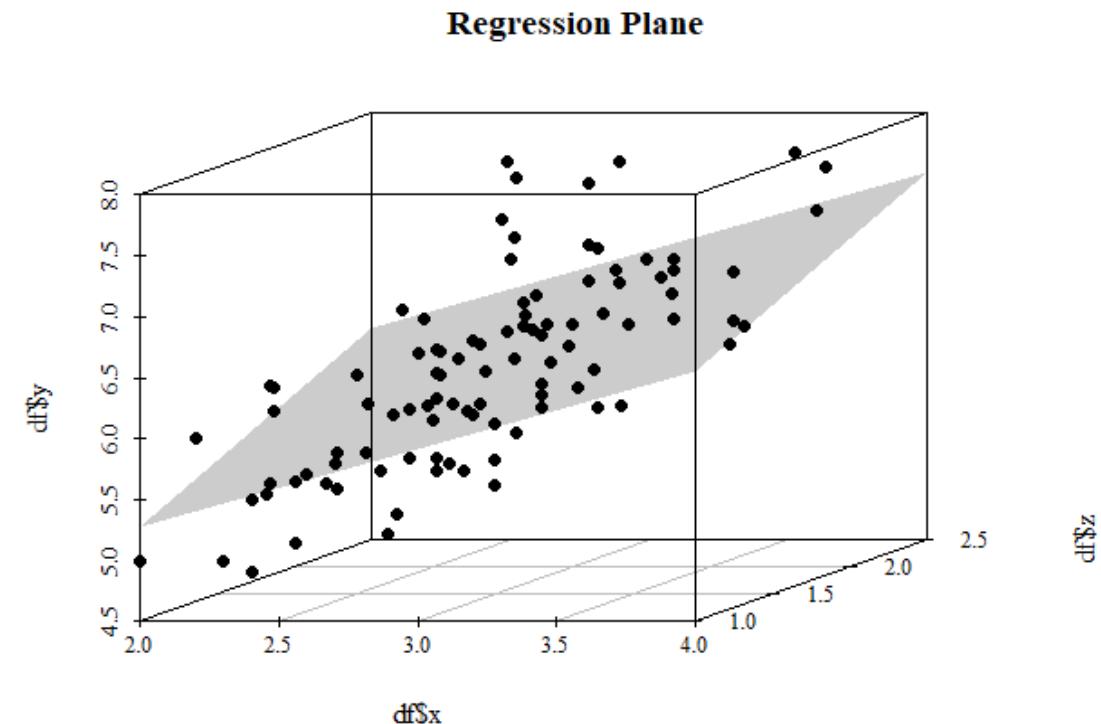
---

- Recall our **simple linear regression model**, where we use one variable (rundiff) to predict the *value* of another (W)
  - $Wins = Intercept + \beta_1 * RunDiff$



# Multiple Linear Regression

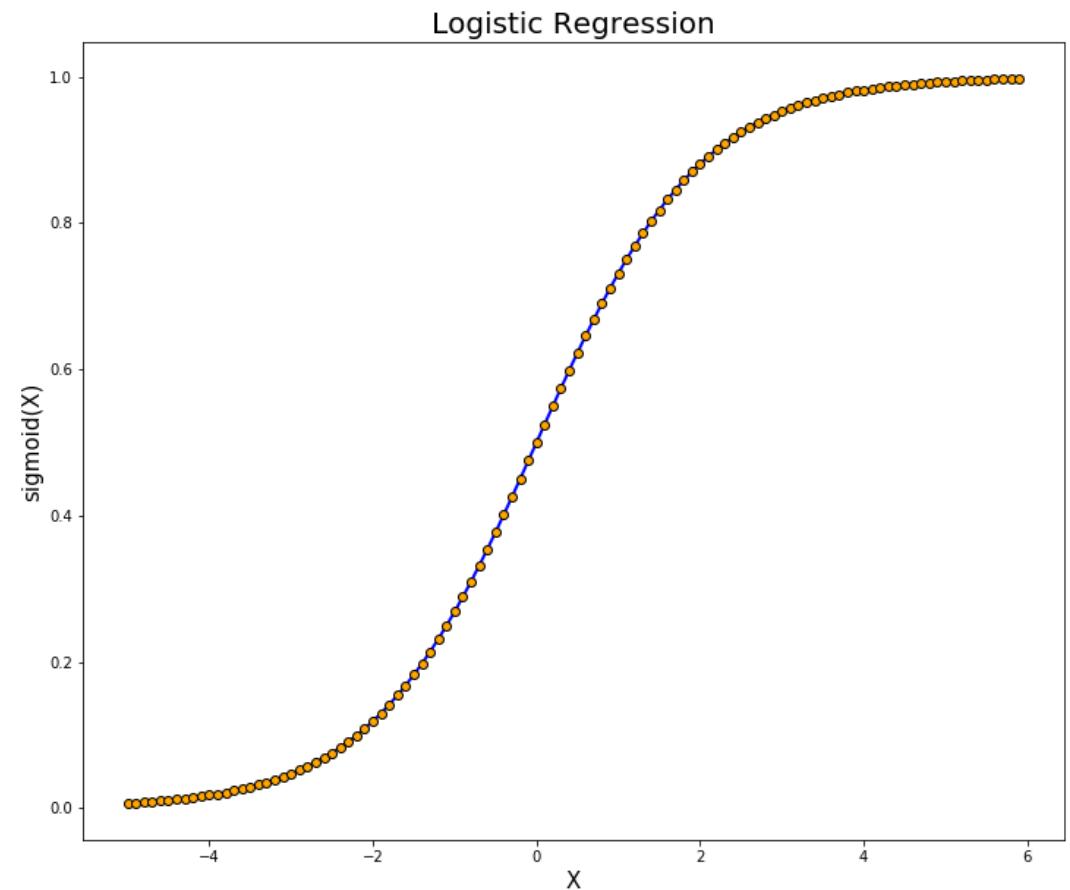
- You can also have a **multiple regression model**, where you use multiple variables to predict the value of another
  - $Wins = Intercept + \beta_1 * RunDiff + \beta_2 * WinsPriorYear$



# Predicting Categorical Variables

---

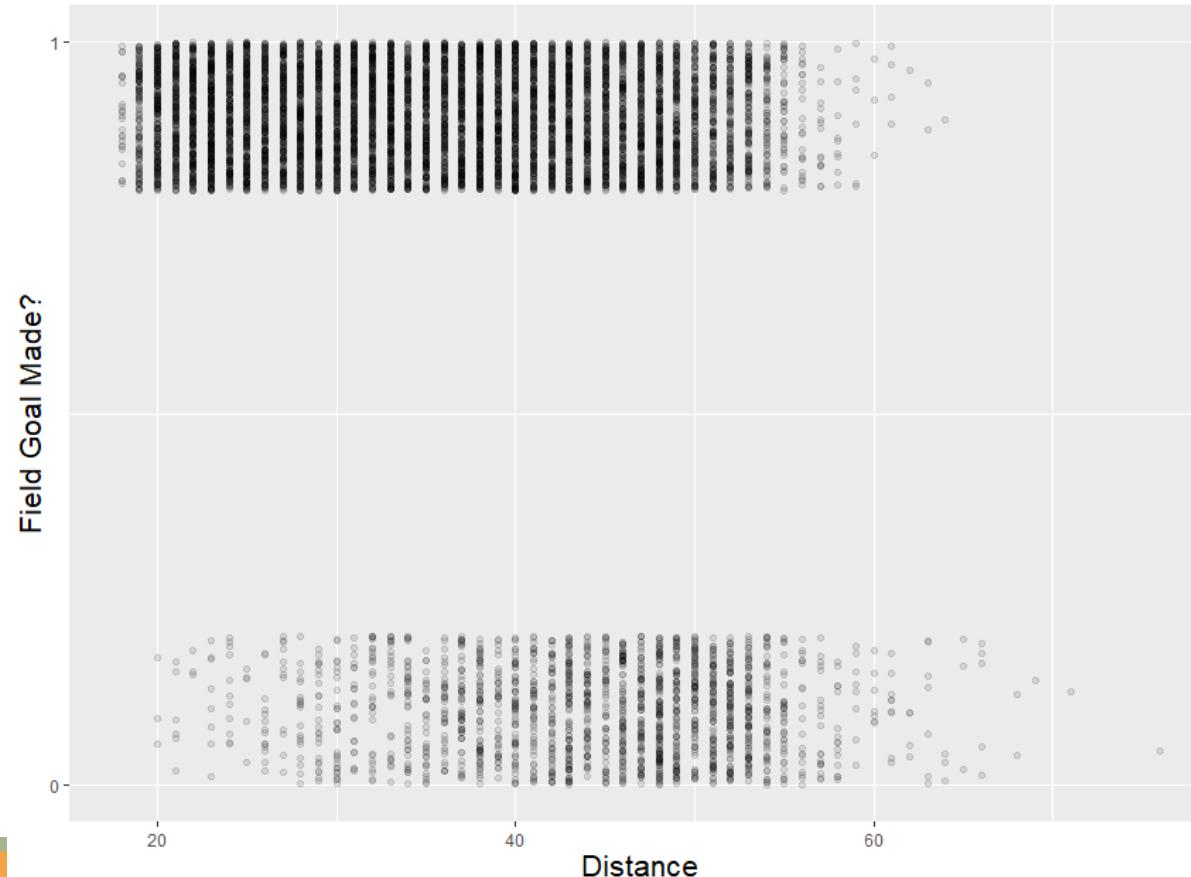
- How do we predict the “value” of a categorical (let’s say binary) variable?
  - One method: predict its *probability*
  - Could we use linear regression for this?
    - To keep limited to {0,1} we’ll predict probability using a slightly different technique called **logistic regression**
    - Skipping underlying math



# Predicting Field Goal Probability

## RAW DATA

NFL Field Goal Makes by Distance, 2005-15



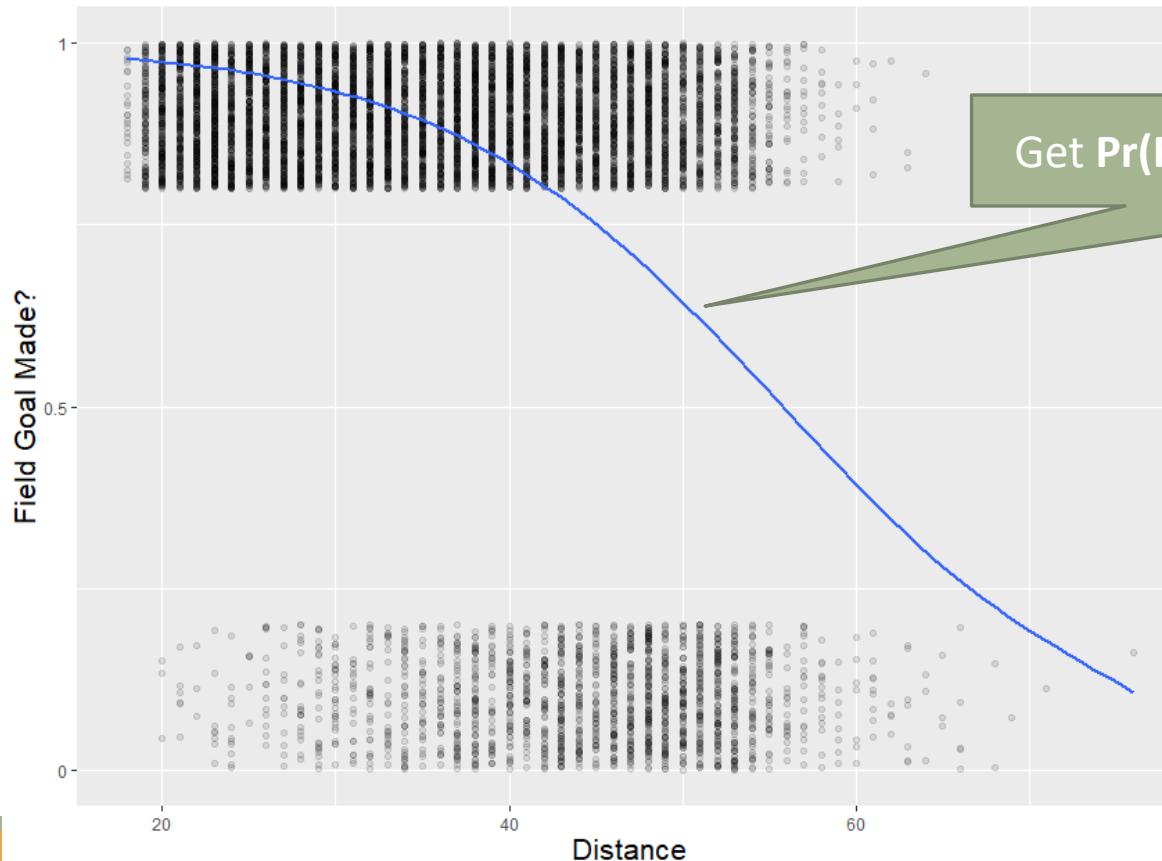
Source:

[https://raw.githubusercontent.com/statsbylopez/StatsSports/master/Data/nfl\\_fg.csv](https://raw.githubusercontent.com/statsbylopez/StatsSports/master/Data/nfl_fg.csv)

# Predicting Field Goal Probability

## RAW DATA + LOGISTIC REGRESSION MODEL PREDICTIONS

NFL Field Goal Makes by Distance, 2005-15



Get  $\text{Pr}(\text{Field Goal Made} = 1 \mid \text{Distance} = X)$

What is the approximate probability of nailing a 40-yd field goal?

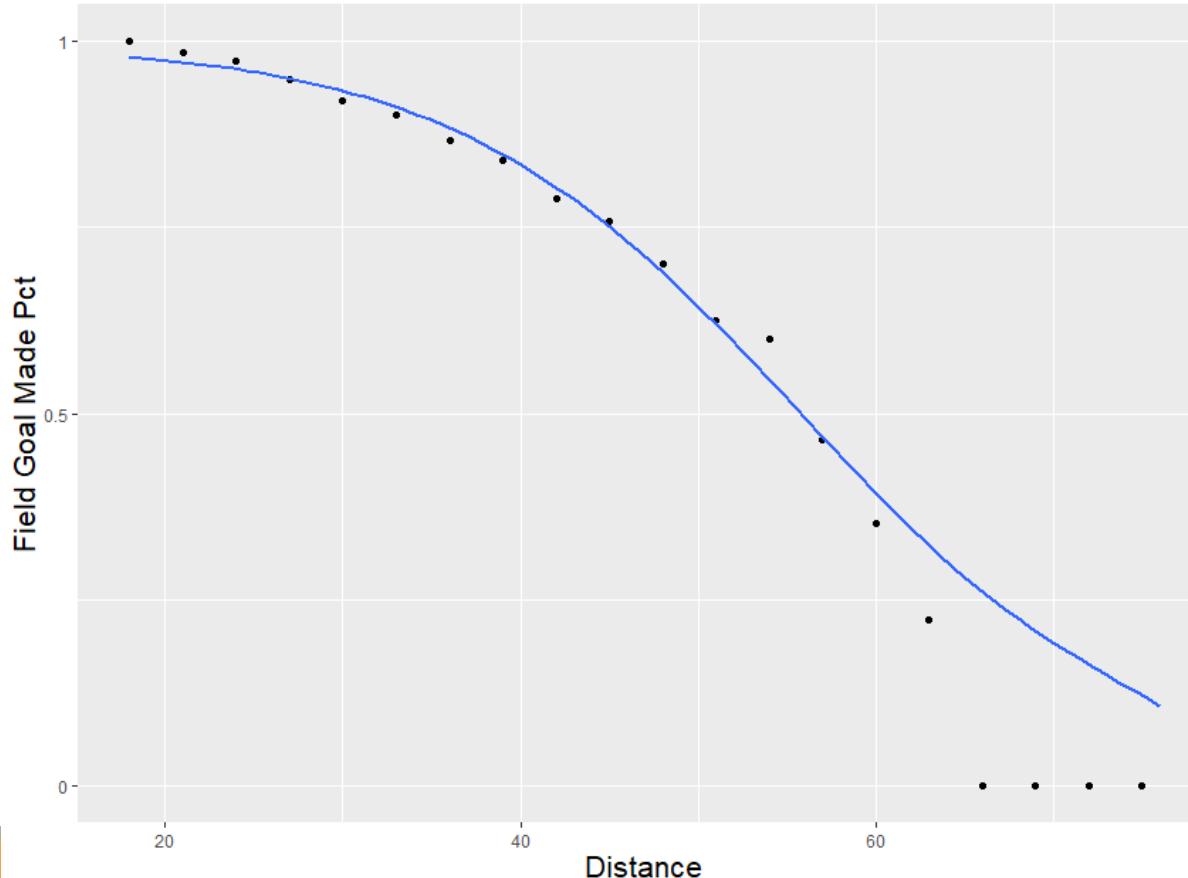
$\text{Pr}(\text{Field Goal Made} = 1 \mid \text{Distance} = 40) = ???$

# Predicting Field Goal Probability

Note we DON'T bin the data to run the model – this is just for a plot that makes more sense

## BINNED RAW DATA + LOGISTIC REGRESSION MODEL PREDICTIONS

NFL Field Goal Makes by Distance, 2005-15



Just like linear regression, could add more variables (e.g. year, grass vs. artificial turf) to our model to improve accuracy.

But we won't bother here.

# Expected Points

---

- Now that we know how to predict the probability of a binary event like a field goal, let's do something more useful
- What's the point of a football team? How do you win?



# Expected Points

---

- Predict value of current game state for offense by calculating **expected** value of next **points** scored
  - From offense's perspective, next score could be: +8, +7, +6, +3, +2, 0, -2, -3, -6, -7, -8
    - Simplify to +7, +3, +2, 0, -2, -3, -7
  - **Expected Points** =

$$\text{Pr}(\text{Next Score } +7)*7 + \text{Pr}(\text{Next Score } +3)*3 + \dots + \text{Pr}(\text{Next Score } -7)*-7$$

- Where do the probabilities come from?

# Expected Points

---

- Fairly simple **logistic regression\*** model based on
  - Down
  - Distance to go for first down
  - Field position (yard line or distance from opponent end zone)
  - Time left (some models)
- Intuition: how would each of these impact **expected points** for the offense?

\* Note model is more complex because predicting probability of several mutually-exclusive events, but not fundamentally different from field goals.

# Expected Points

---

- Calculation example: Vikings, 3<sup>rd</sup> & 4 from SEA 34, 10:66 left in 3<sup>rd</sup> quarter

Event Value	Probability
+7	0.40
+3	0.40
+2	<0.01 (treat as negligible)
0	0.08
-2	0.02
-3	0.10
-7	0.10

- What are the expected points for the Vikings?

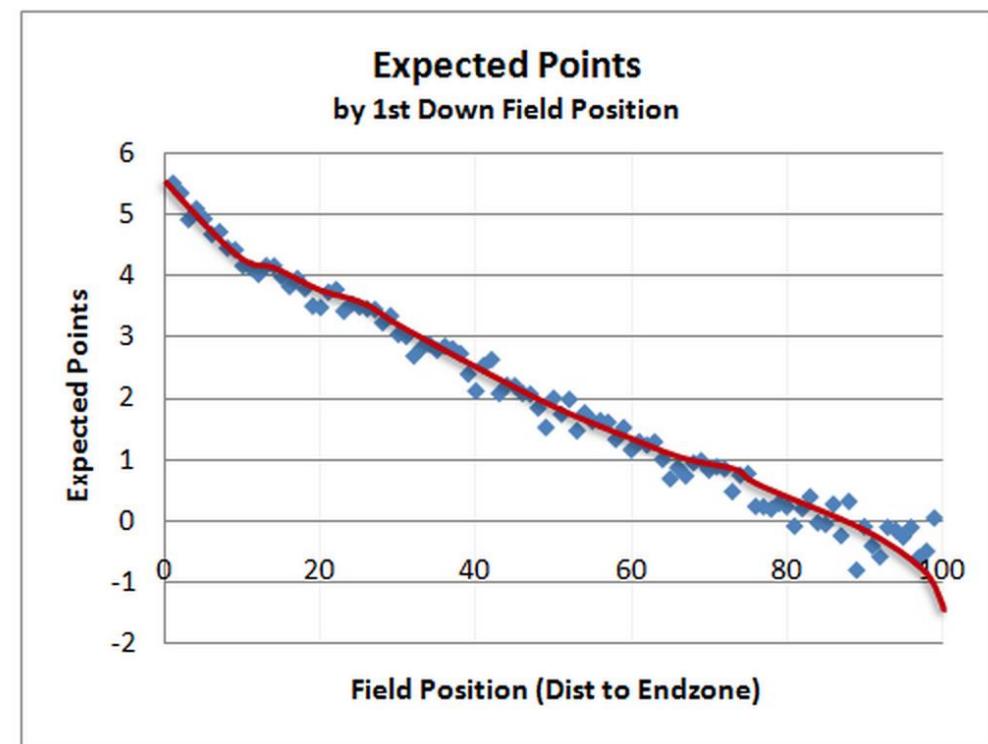
# Expected Points

- Various **expected points** models

TABLE I

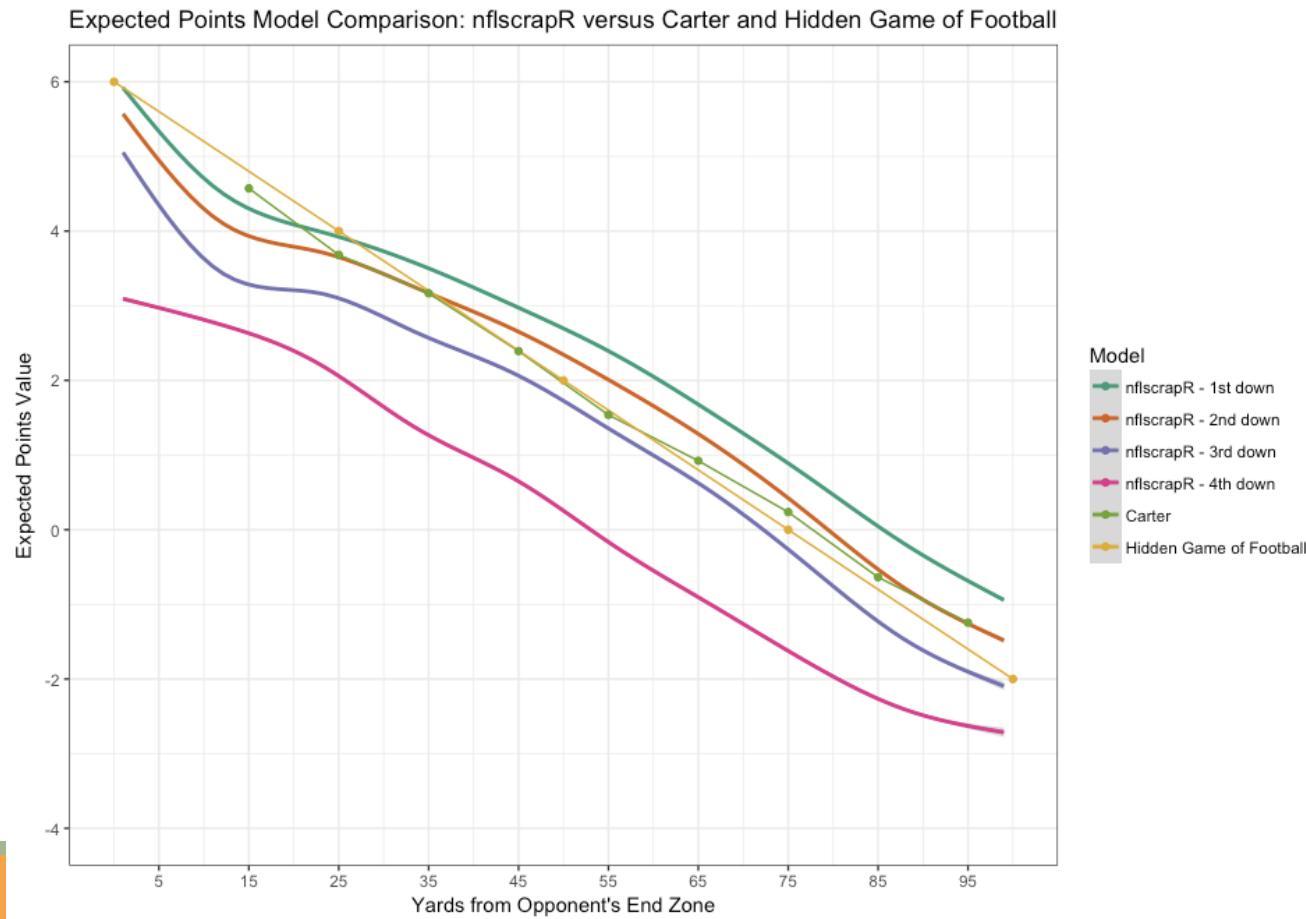
THE EXPECTED POINT VALUES OF POSSESSION OF THE FOOTBALL WITH FIRST DOWN AND TEN YARDS TO GO FOR VARIOUS TEN-YARD STRIPS

Center of the ten-yard strip (yards from the target goal line): $X$	Expected point value: $E(X)$
95	-1.245
85	-0.637
75	+0.236
65	0.923
55	1.538
45	2.392
35	3.167
25	3.681
15	4.572
5	6.041



# Expected Points

- Various **expected points** models



# Expected Points Added

---

- Expected points are nice in their own right, I guess
- But real value comes from being able to compare two game states – like before and after a play
  - **Expected Points Added (EPA)**
  - Sound like anything you remember from baseball...?

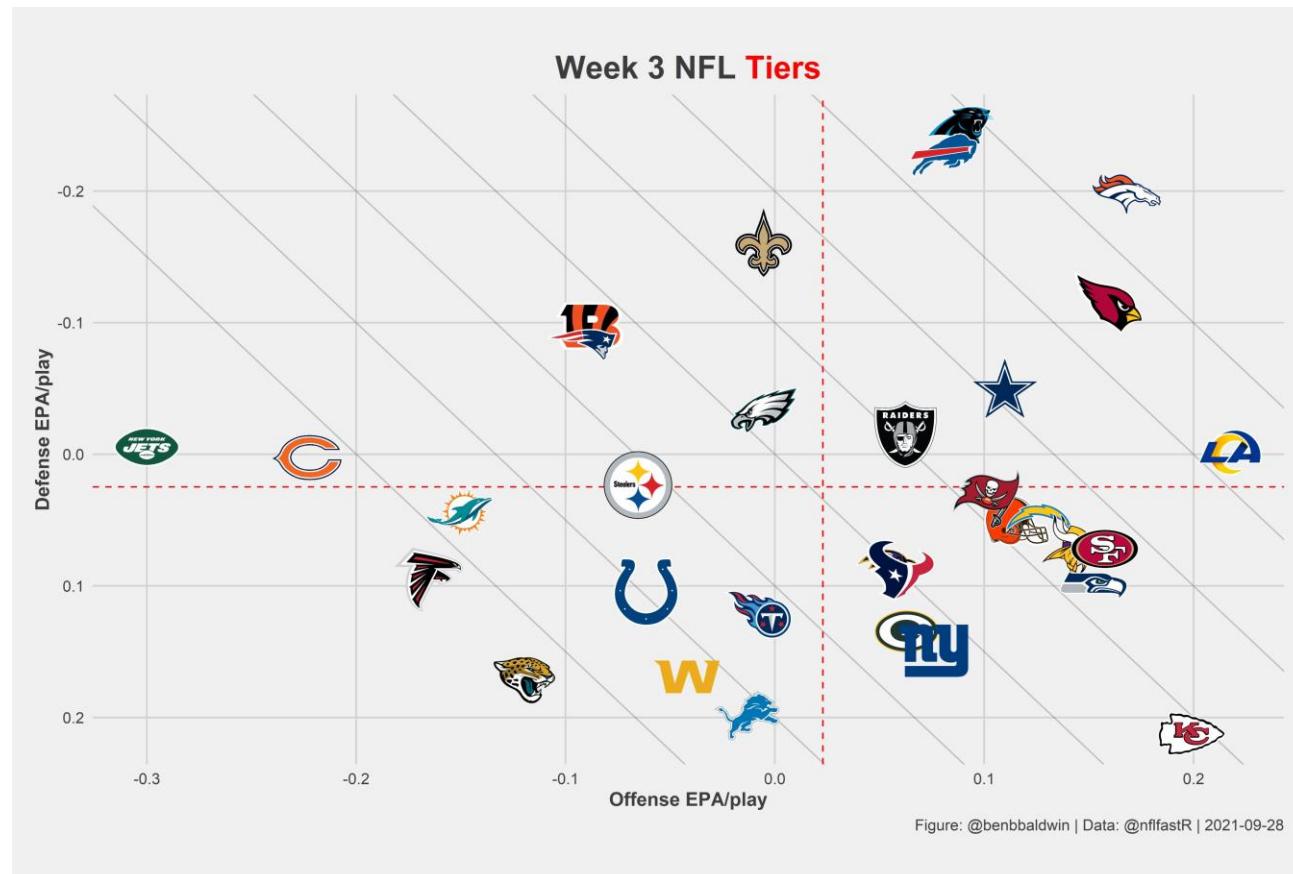
# Expected Points Added

---

- What do you think EPA would say about each of these plays?
  - A completed 3-yard pass on 1<sup>st</sup> & 10?
  - A completed 3-yard pass on 3<sup>rd</sup> & 2?

# Expected Points Added

- EPA can be used to value:
  - Plays
  - Coach decisions (go for it vs. kick on 4<sup>th</sup> down)
  - Team offensive and defensive units (by summing up EPA of all their plays)
  - (With a lot more work to divide credit) individual player value



# Win Probability

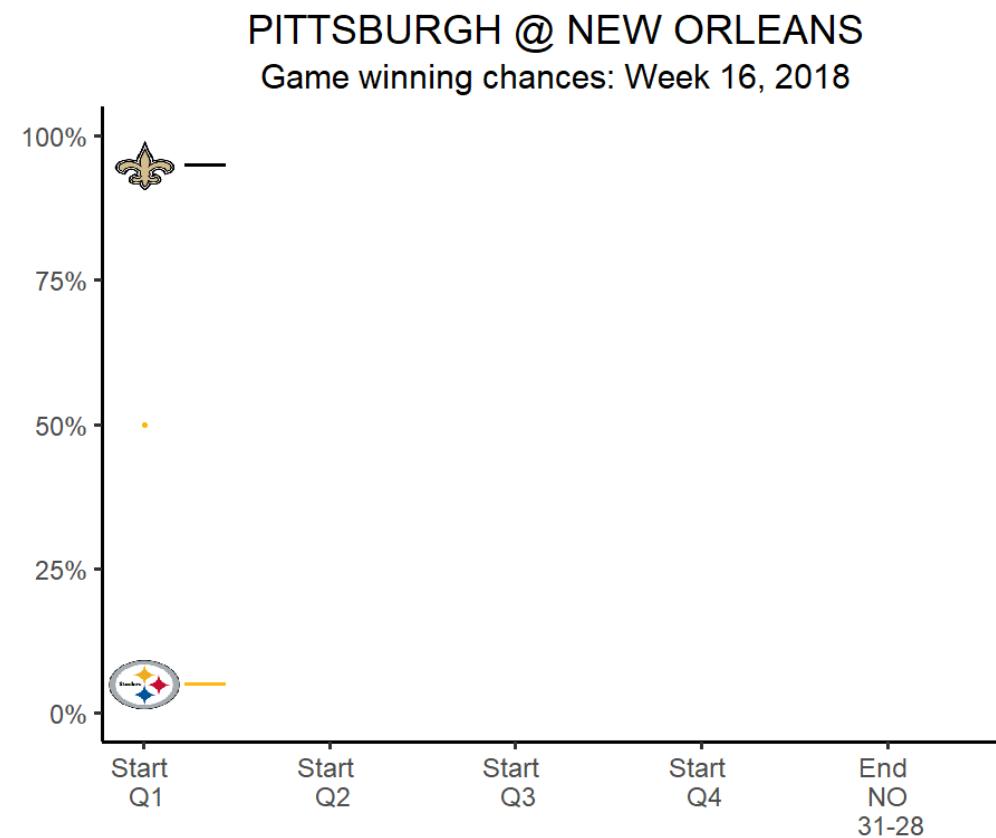
---

- But it's not really about points, is it?
- Fair enough, coach, so what should we-



# Win Probability (Added)

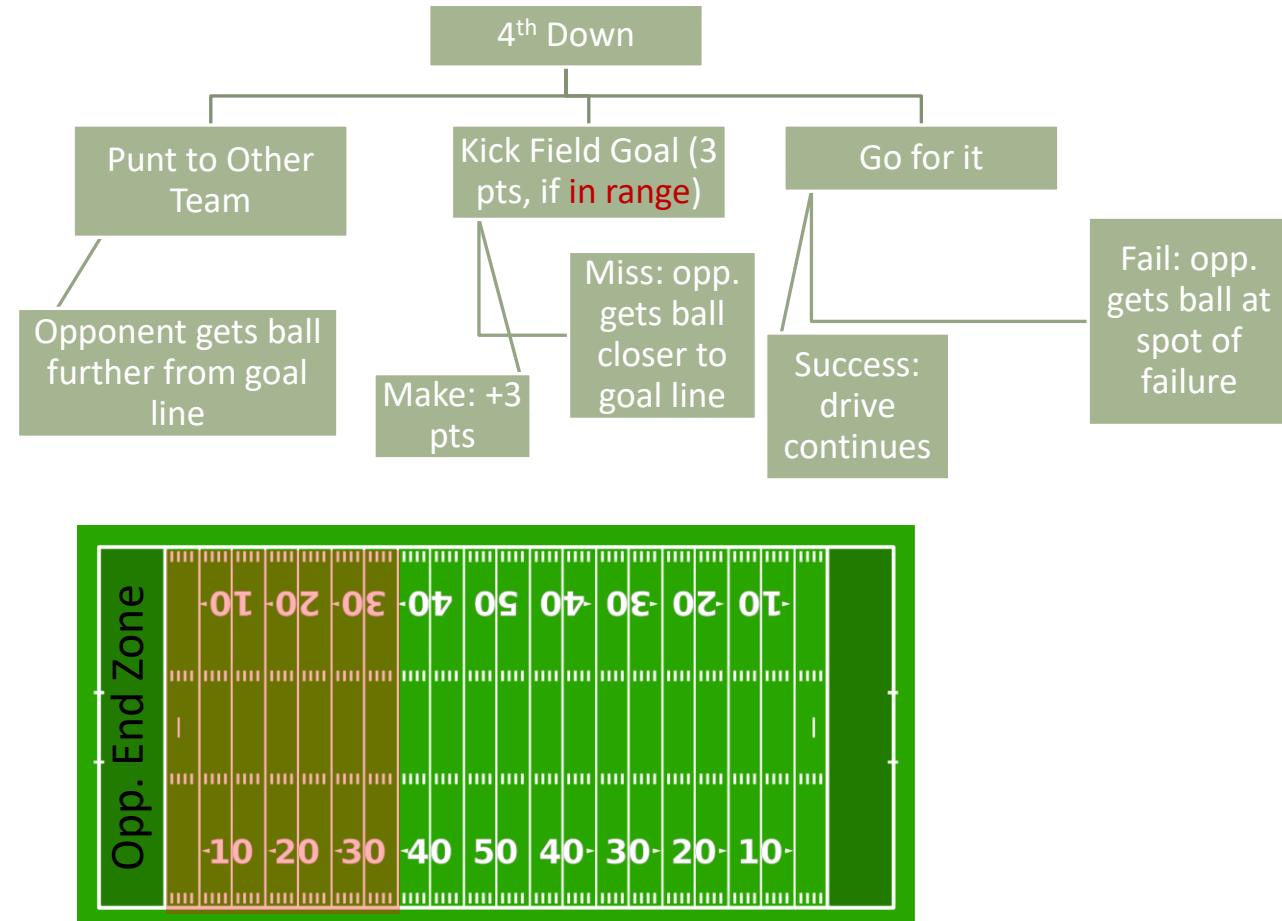
- We could just model the probability of winning directly
  - **Win Probability** models
- Just like going from expected points → EPA, we can calculate **win probability added (WPA)** for any play
- Use EPA or WPA? Case-by-case



# Fourth Down Analyses

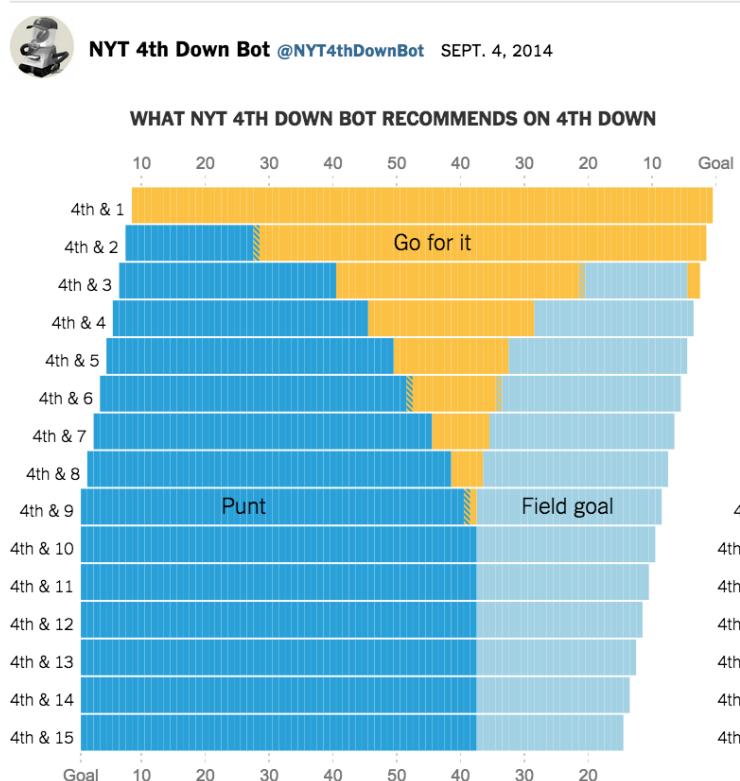
---

# Fourth-Down Decisionmaking



# Fourth-Down Decisionmaking

- What should a coach do? Recommendations (as of 2014):



# Fourth-Down Decisionmaking

---

- Where do these recommendations come from?
  - You'll be shocked to hear, but... **EPA** or **WPA**!
  - Need:
    - 1.  $\text{Pr}(\text{success})$  of your choice
      - 100% for punt, high for field goal (if in range), maybe lower for go-for-it
    - 2. *Weighted EP* or **WP** for each choice:
      - Example: say you're at 4<sup>th</sup> & 1 at opponents' 23 yd line and have an 80% chance of success (converting it).
        - Then  $\text{EP}(\text{Go for it}) = 0.8 * \text{EP}(\text{Go for it and succeed}) + 0.2 * \text{EP}(\text{Go for it and fail})$
        - Compare to  $\text{EP}(\text{FG}) = 0.9 * 3 + 0.1 * \text{EP}(\text{FG Miss})$
      - Then (in many but not all cases) choose highest EP or WP option
        - Time when you shouldn't do this?

# Fourth-Down Decisionmaking

---

- Example:
  - 4<sup>th</sup> - 2 from opponents 23-yard line

```
> round(tab.Fourth.2, 2)
      Decision
    pts.next Go for it Kick
      -8      0.00 0.00
      -7      0.09 0.06
      -6      0.00 0.00
      -3      0.07 0.05
       2      0.01 0.00
       3      0.23 0.81
       6      0.04 0.00
       7      0.37 0.04
       8      0.01 0.00
     <NA>      0.19 0.02
```

Go for it:

- $E(X) = 2.81$

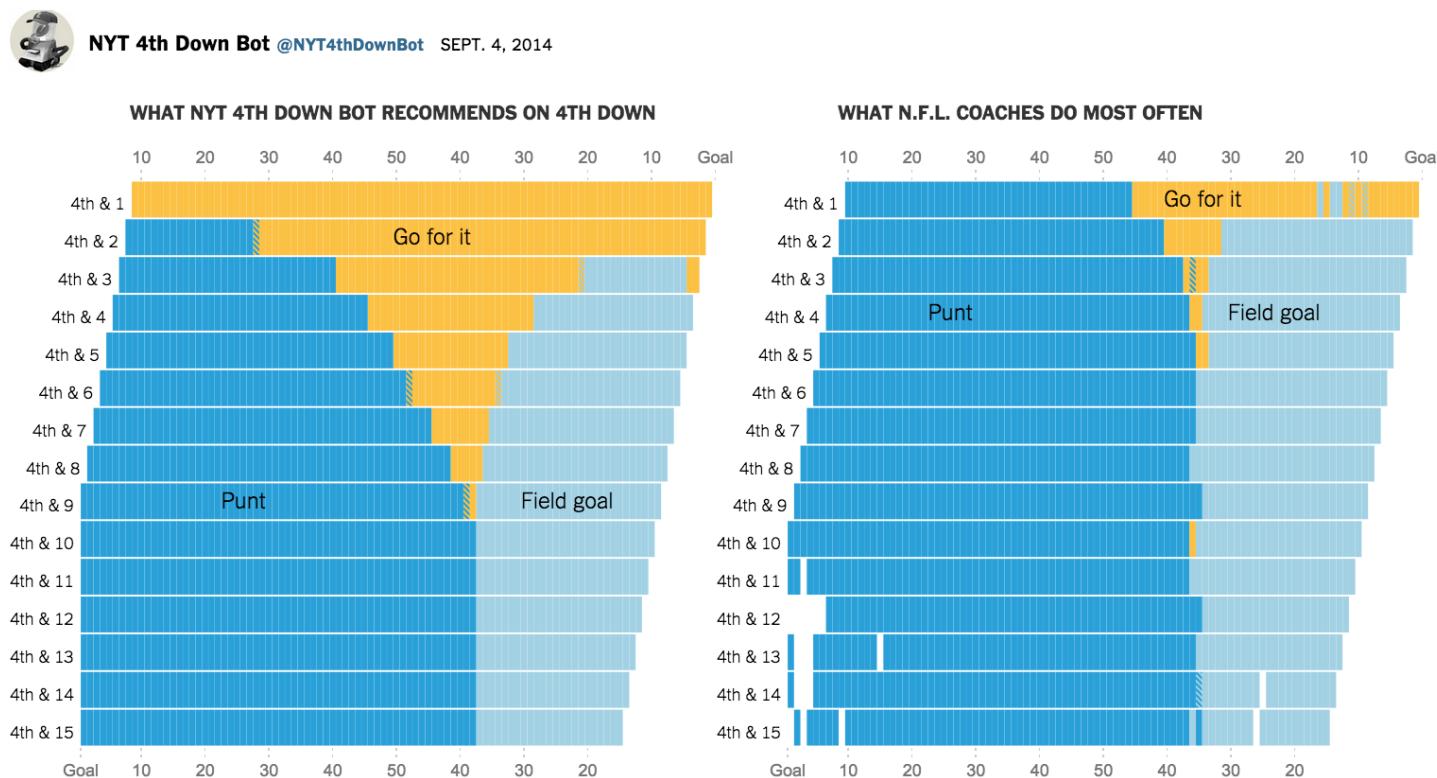
Kick it:

- $E(X) = 2.14$

What do coaches do?

# Fourth-Down Decisionmaking

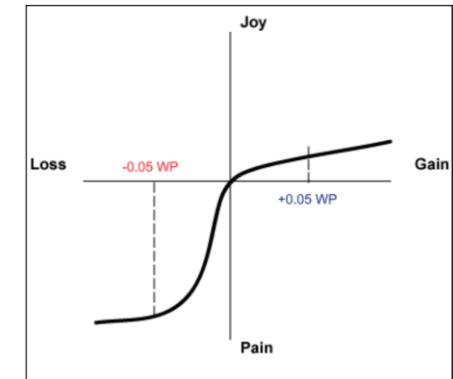
- What do coaches *actually* do?



# Fourth-Down Decisionmaking

- Why would coaches act this way?

- Loss Aversion
- Risk Aversion/Minimax

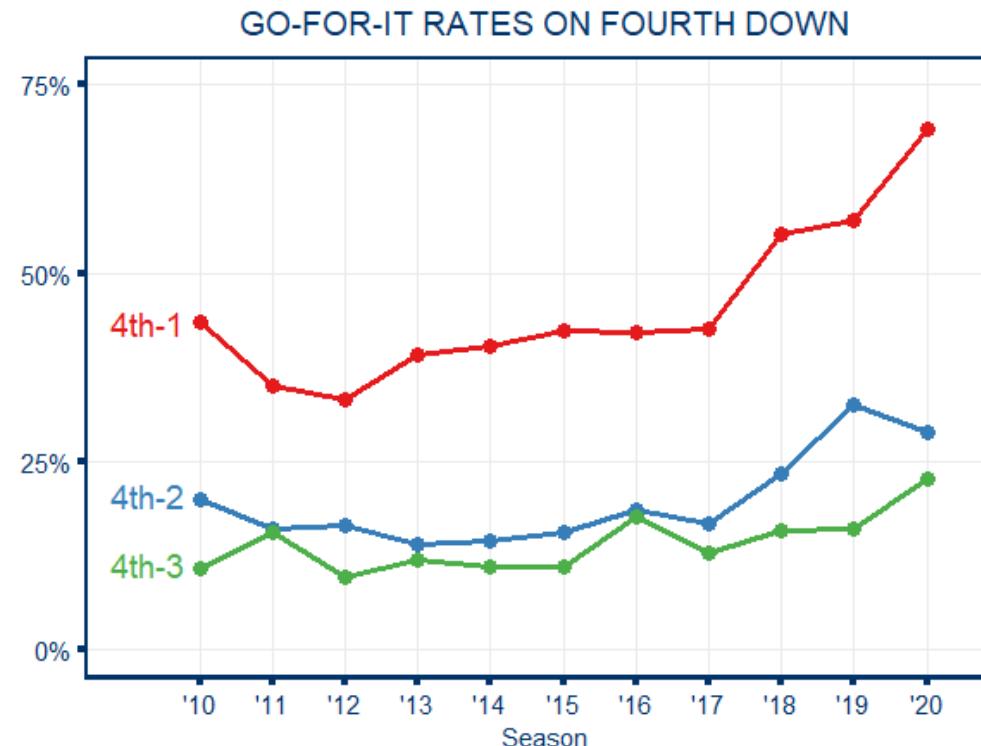


*"Had we done that [gone for it] after what we had done to get down there and [not scored a touchdown], I can imagine what the critique would have been today about the play call." – Brian Billick*

*"You guys might very well be right that we're calling something too conservative in that situation. But what you guys don't understand is that if I make a call that's viewed to be controversial by the fans and the owner, and I fall, I lose my job" – Marvin Lewis*

# Fourth-Down Decisionmaking

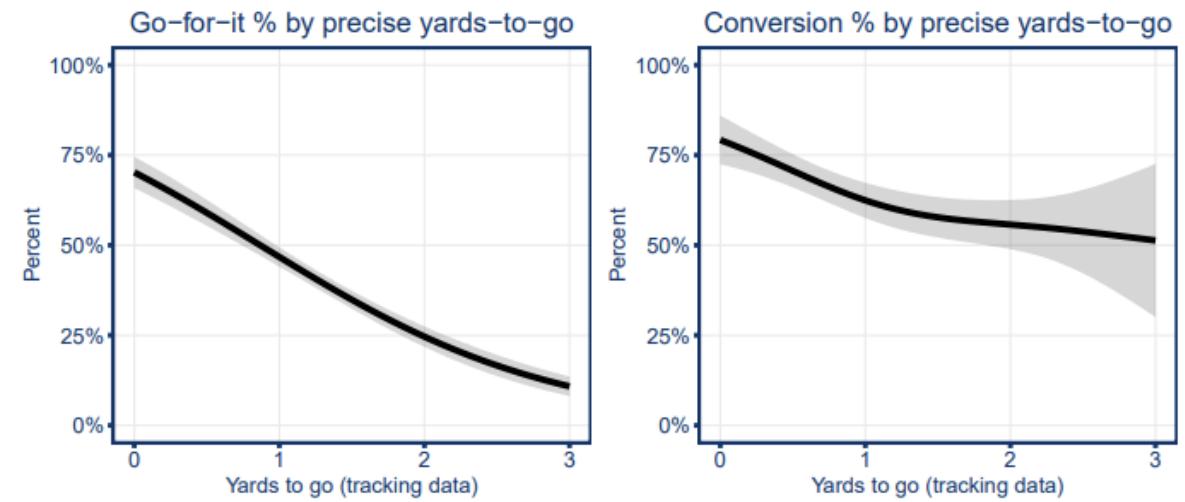
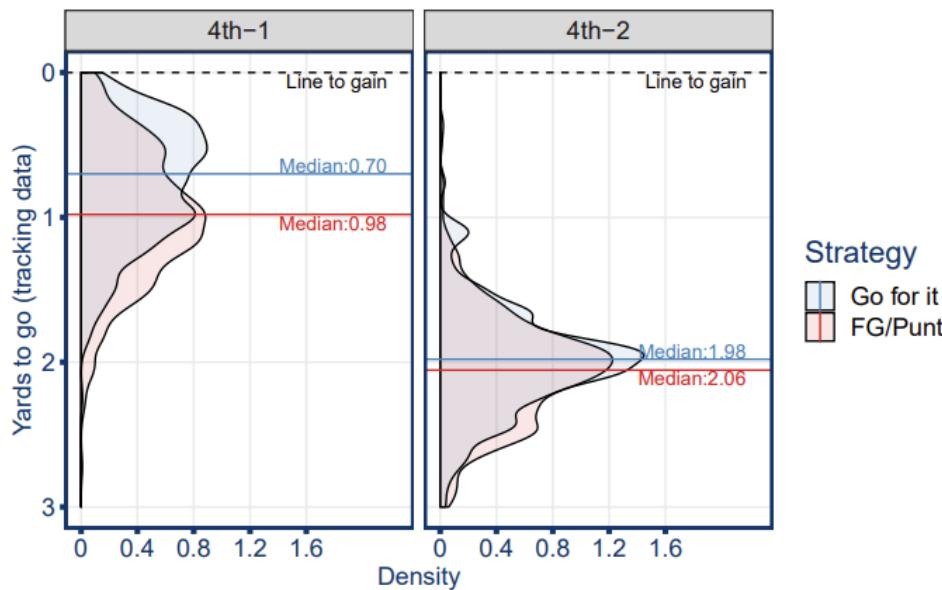
- Well, more like what *did* coaches actually do. We're making inroads!
  - Of course, not perfect. John Harbaugh (BAL) vs. Joe Judge (NYG)
  - Follow [4<sup>th</sup> down decision bot Twitter account](#) for more



# Fourth Down Decisionmaking

- Did coaches know more about what they were doing than we gave them credit for?
  - Once we got better data (ball tracking chip), some very interesting data emerged that changed what we thought we knew
  - Importance of humility and open-mindedness in analytics

Precise distance needed on 4th-down plays, 2017–2019



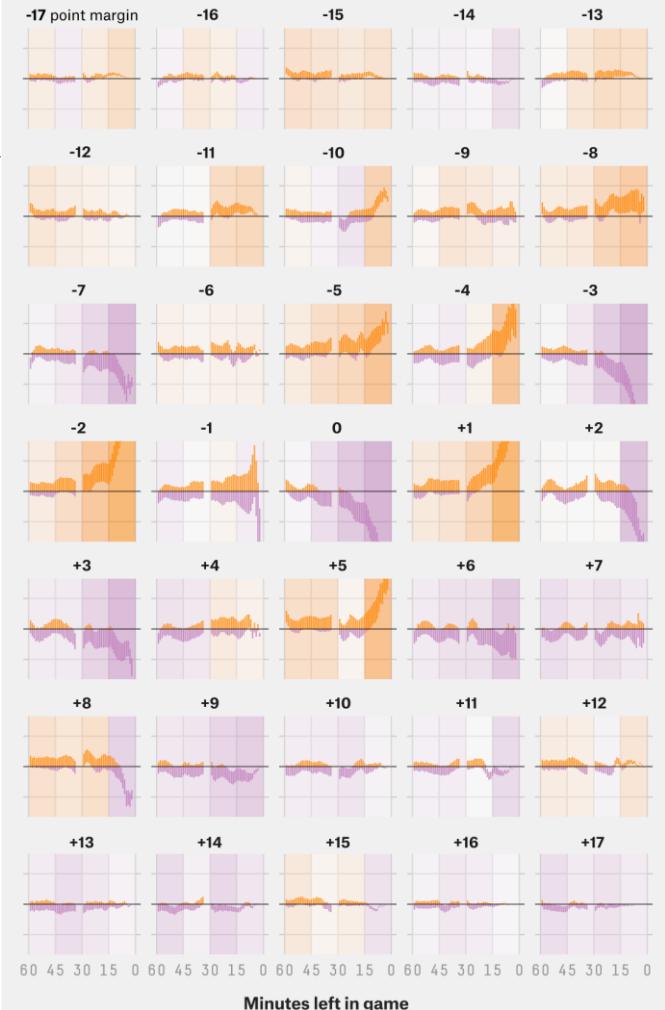
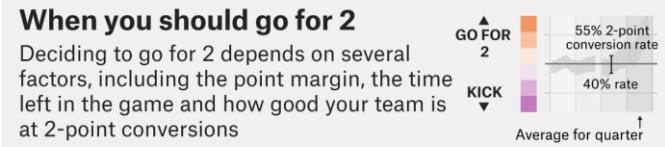
# Two-Point Conversions

- Extremely similar situation

The most common 2-point decision mistakes, 2015 and 2016

SITUATION		WHAT COACHES DID			
MARGIN	QUARTER	BETTER TO GO FOR 2?	CORRECT DECISIONS	CLEARLY INCORRECT DECISIONS	AVG. WIN PROB IMPACT OF MISTAKES
-8	4th	✓	0	27	-0.90
-4	4th	✓	0	19	-1.92
-11	4th	✓	0	16	-0.38
-8	3rd	✓	0	13	-0.79
-15	4th	✓	0	9	-0.16
-11	3rd	✓	0	6	-0.51
-5	3rd	✓	1	5	-0.71
-2	3rd	✓	4	5	-1.02
-15	3rd	✓	0	4	-0.21
-6	4th	✓	0	2	-0.33
-13	3rd	✓	0	2	-0.31
-17	3rd	✓	1	2	-0.22
+1	3rd	✓	4	2	-0.93
+11	4th		5	2	-0.28
+12	4th	✓	7	2	-0.24

Among scenarios in which the correct decision applies to virtually any team, regardless of conversion ability, and coaches made the wrong decision at least twice (minimum -0.1 impact).



FiveThirtyEight

SOURCE: ESPN STATS & INFORMATION GROUP

# Caution on 4<sup>th</sup> Down Analytics Projects

---

- It's been done to death publicly
- Not an “interesting” project to teams, fellow analysts (become a bit of a joke)
- Unless you have new data or a *completely* different approach, leave it be
- (Same thing for 2-pt conversions)

# Draft Analytics

---

# Draft Analytics

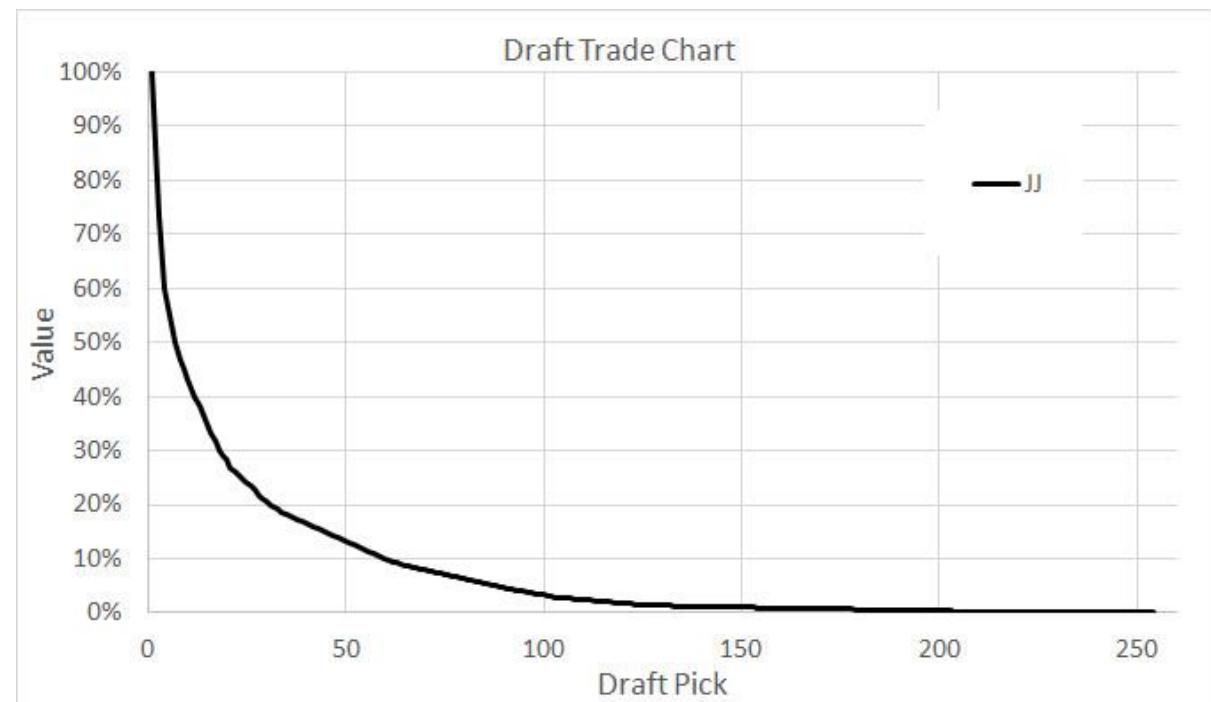
---

- Say you're a GM on Draft Day thinking about trading picks with another team
  - How do you decide when to make a trade?

# Draft Pick Values

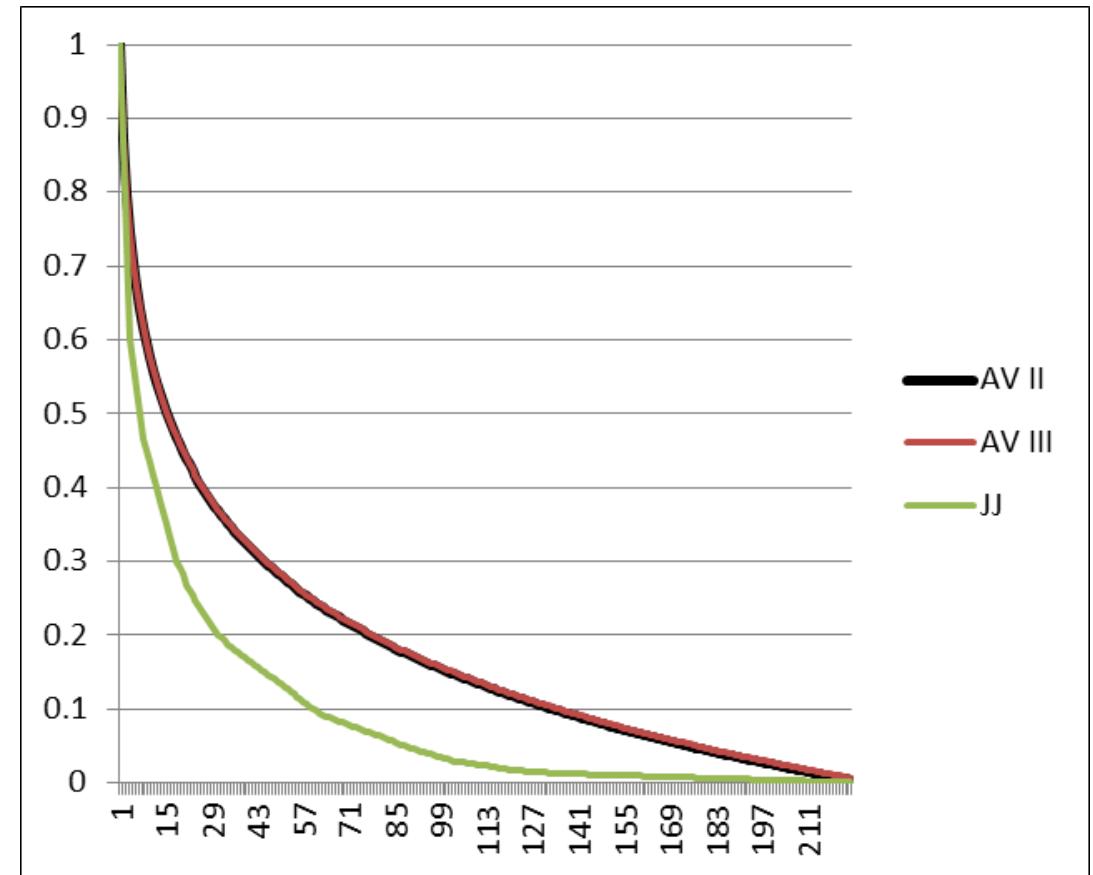
---

- Need to know the value of your picks
  - Attempt #1: Jimmy Johnson, “The Chart” – from 1990s Cowboys.
  - Based on actual trades made in 1980s. Problem?



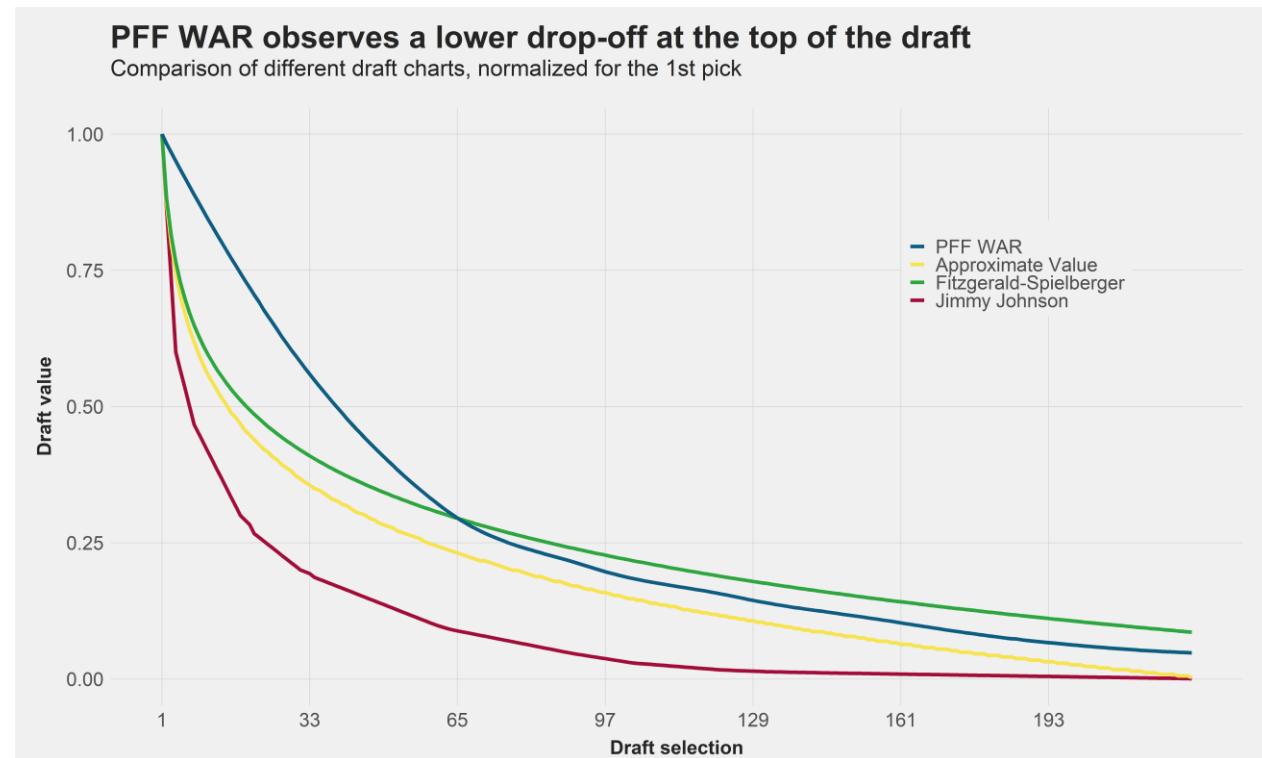
# Draft Pick Values

- Need to know the value of your picks
  - Attempt #2: [Massey and Thaler](#); Chase Stuart at Football Perspective
  - NFL GMs massively overvalued high picks  
→ curve was too steep
  - Based on first 4 years' actual Career AV metric from Pro-Football-Reference



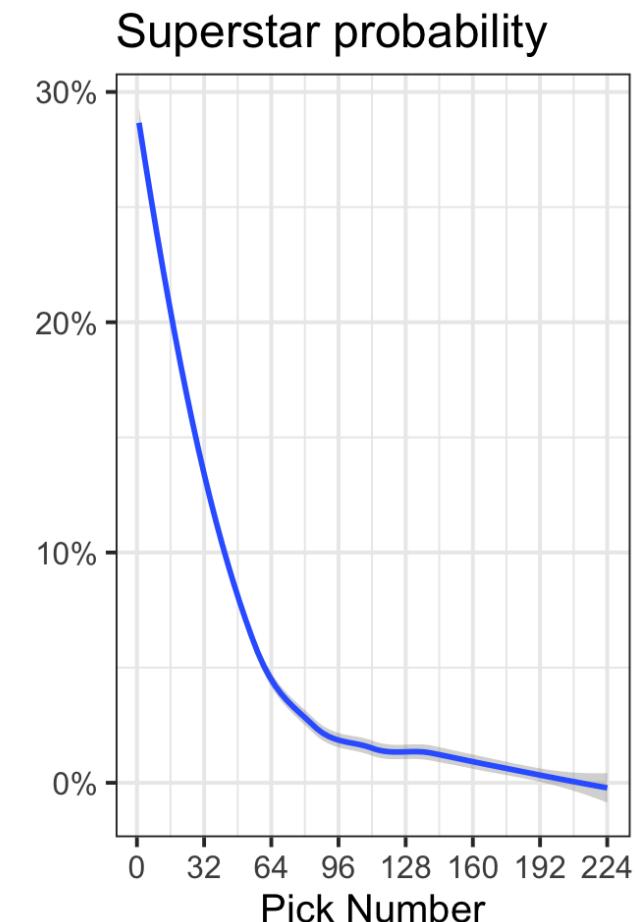
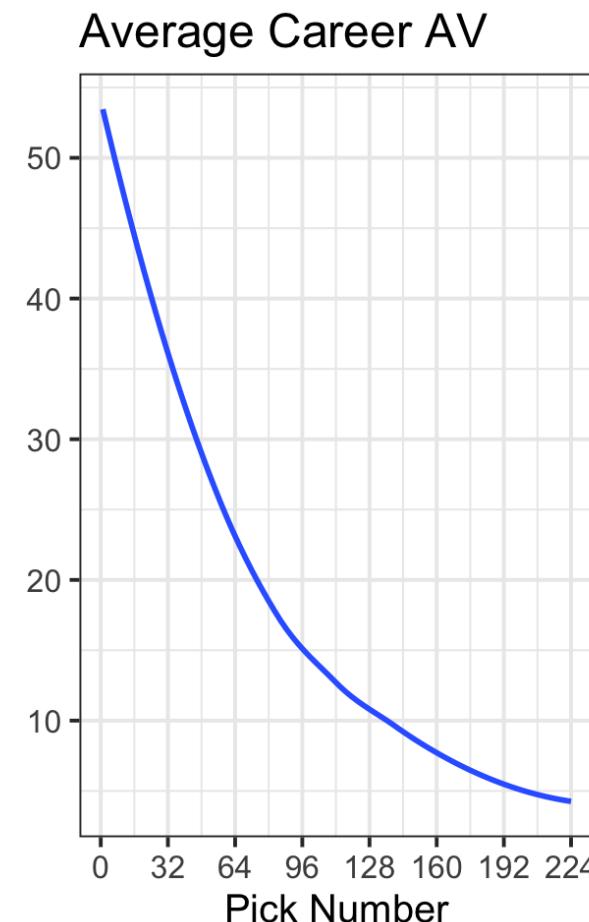
# Draft Pick Values

- Need to know the value of your picks
  - Later attempts: Fitzgerald and Spielberger (based off actual post-rookie contract values); Pro Football Focus (based off PFF WAR); many others
  - All with same conclusion vs. Jimmy Johnson Chart



# Draft Pick Values

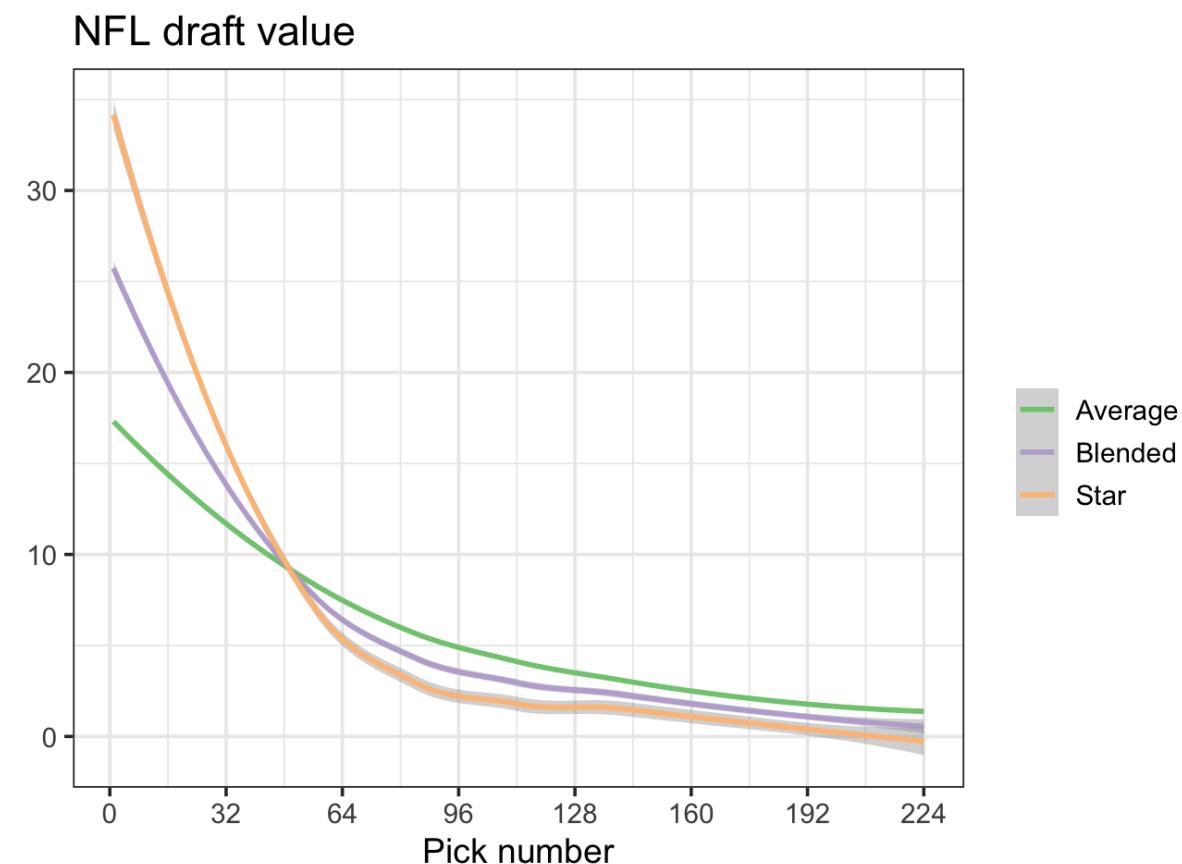
- But...is it possible teams are behaving rationally (or more rationally than we give them credit for) even if they over-value top picks?
  - Instead of making best decisions on “average” player, maybe teams are shooting for a “superstar?”
    - “High-variance” strategy
    - “**Superstar**” = career AV in roughly top 5<sup>th</sup> percentile



# Draft Pick Values

---

- But...is it possible teams are behaving rationally (or more rationally than we give them credit for) even if they over-value top picks?



# Draft Pick Values

- New York Football Jets, 2018 NFL Draft: give Colts #6, 37, 49 (Quenton Nelson, Braden Smith, Dallas Goedert) in exchange for #3 (Sam Darnold)

Avg.  
Curve

Pick Number	Value	Team
3	52.5	to the Jets
6	50.6	to the Colts
37	33.5	to the Colts
49	28.3	to the Colts

Star  
Curve

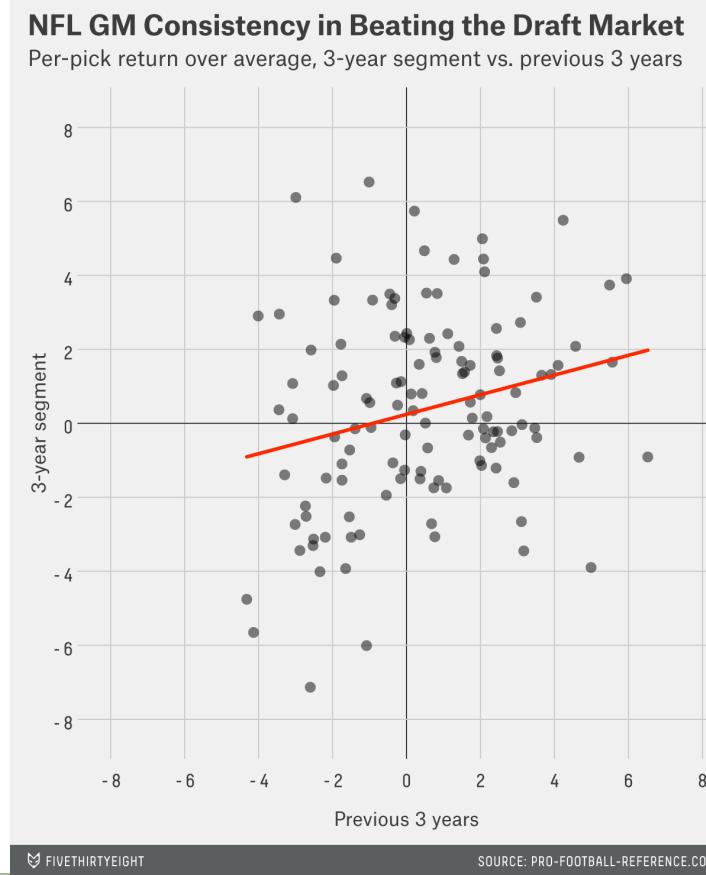
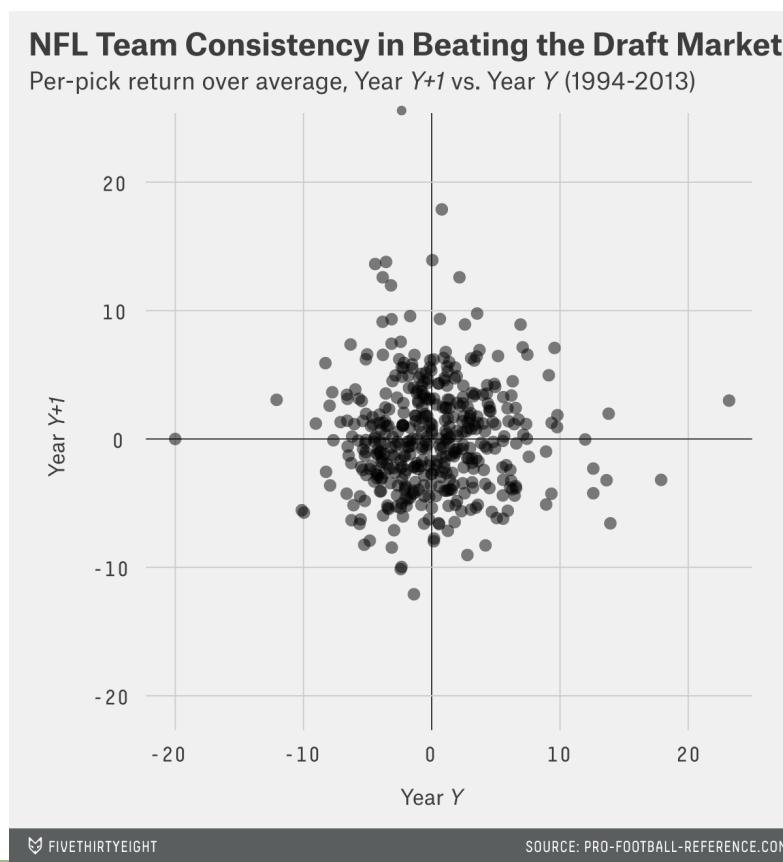
Pick Number	Value	Team
3	0.32	to the Jets
6	0.29	to the Colts
37	0.09	to the Colts
49	0.07	to the Colts

Blended  
Curve

Pick Number	Value	Team
3	27.8	to the Jets
6	25.4	to the Colts
37	10.8	to the Colts
49	8.7	to the Colts

# Drafting: Skill or Luck

- Are there GMs that are especially good or bad at drafting?



# Draft Analytics Conclusions

---

- Analytics current wisdom:
  - Trading up is dumb
  - Trade back as much as possible with dumb trade partners
  - Chances > specific players



# NFL Player Valuation Analytics

---

# Player Valuation – DVOA

- **Context. Is. King.**
- Early modern attempt: Football Outsiders, Aaron Schatz, and Defense-Adjusted Value over Average (DVOA, Teams and Players)

THE ULTRA-SHORT VERSION:

**FOOTBALL OUTSIDERS**™

DVOA measures a team's efficiency by comparing success on every single play to a league average based on situation and opponent.

THE SHORT VERSION:

DVOA is a method of evaluating teams, units, or players. It takes every single play during the NFL season and compares each one to a league-average baseline based on situation. DVOA measures not just yardage, but yardage towards a first down: Five yards on third-and-4 are worth more than five yards on first-and-10 and much more than five yards on third-and-12. Red zone plays are worth more than other plays. Performance is also adjusted for the quality of the opponent. DVOA is a percentage, so a team with a DVOA of 10.0% is 10 percent better than the average team, and a quarterback with a DVOA of -20.0% is 20 percent worse than the average quarterback. Because DVOA measures scoring, defenses are better when they are negative. For more detail, read below.

Please feel free to contact us with questions and comments about our original statistics [using the contact form](#).

2021 Team DVOA Ratings: Overall  
Updated 09/28/2021 03:53 PM EDT

Team	Total DVOA	▲ Prev. Week Rank	W-L	Weighted DVOA	DAVE	Offense DVOA	Defense DVOA	Special Teams DVOA
			W-L		DAVE			
CAR	1 40.1%	1	3-0	1 40.1%	16 0.3%	12 6.6%	1 -38.8%	31 -5.3%
CLE	2 38.8%	9	2-1	2 38.8%	8 10.5%	4 26.3%	10 -7.7%	6 4.9%
LAR	3 37.8%	2	3-0	3 37.8%	4 14.5%	1 37.0%	13 -4.7%	29 -3.9%
ARI	4 37.5%	5	3-0	4 37.5%	13 7.5%	8 14.8%	6 -17.8%	5 4.9%
DEN	5 34.1%	4	3-0	5 34.1%	9 9.8%	7 15.5%	5 -21.7%	26 -3.1%
BUF	6 29.8%	13	2-1	6 29.8%	3 14.9%	16 -0.1%	2 -31.1%	19 -1.2%
TB	7 24.2%	3	2-1	7 24.2%	1 21.5%	5 23.8%	17 0.8%	10 1.2%
NO	8 22.7%	12	2-1	8 22.7%	11 8.6%	19 -3.8%	3 -25.3%	9 1.2%
DAL	9 21.6%	14	2-1	9 21.6%	10 9.1%	6 23.1%	15 -0.5%	23 -2.0%

# Player Valuation – QBs and Passing

- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - Start with “traditional” box score stats (and things we can derive from it, like Completion % and TD/INT ratio)
  - Chargers @ Chiefs, 9/26/21; Final LAC 30-KC 24

Los Angeles Passing

	C/ATT	YDS	Avg	TD	INT	SACKS	
Justin Herbert	26/38	281	7.4	4	0	1-6	
TEAM	<b>26/38</b>	<b>275</b>	<b>7.4</b>	<b>4</b>	<b>0</b>	<b>1-6</b>	

Kansas City Passing

	C/ATT	YDS	Avg	TD	INT	SACKS	
Patrick Mahomes	27/44	260	5.9	3	2	2-9	
TEAM	<b>27/44</b>	<b>251</b>	<b>5.9</b>	<b>3</b>	<b>2</b>	<b>2-9</b>	

# Player Valuation – QBs and Passing

- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - **Adjusted Yards per Attempt (AY/A) – Hidden Game of Football (1988)**
    - $$\frac{\text{pass yards} + 20 * (\text{pass TD}) - 45 * (\text{interceptions thrown})}{\text{Passing Attempts}}$$
  - **Adjusted Net Yards per Attempt (ANY/A) – Chase Stuart at Football Perspective popularized in 2010s:**
    - $$\frac{\text{pass yards} + 20 * (\text{pass TD}) - 45 * (\text{interceptions thrown}) - \text{sack yards}}{\text{Sacks+Passing Attempts}}$$

Los Angeles Passing						
	C/ATT	YDS	AVG	TD	INT	SACKS
Justin Herbert	26/38	281	7.4	4	0	1-6
TEAM	26/38	275	7.4	4	0	1-6

Kansas City Passing						
	C/ATT	YDS	AVG	TD	INT	SACKS
Patrick Mahomes	27/44	260	5.9	3	2	2-9
TEAM	27/44	251	5.9	3	2	2-9

# Player Valuation – QBs and Passing

---

- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - **ANY/A leaders, 2019**

Rk	◆	Quarterback	◆	ANY	◆	DB	◆	ANY/A	▼	RANY/A	◆	SOS	◆	Adj RANY/A	◆
1		Ryan Tannehill		2700		317		8.52		2.36		0.02		2.37	
3		Patrick Mahomes		4199		501		8.38		2.22		-0.08		2.15	
4		Drew Brees		3250		390		8.33		2.17		-0.06		2.11	
2		Lamar Jackson		3471		424		8.19		2.03		0.16		2.19	
5		Matthew Stafford		2517		309		8.15		1.99		-0.16		1.83	
7		Dak Prescott		4856		619		7.84		1.69		-0.10		1.58	
8		Kirk Cousins		3646		472		7.72		1.57		-0.20		1.37	
6		Russell Wilson		4186		564		7.42		1.26		0.32		1.59	
9		Derek Carr		3930		542		7.25		1.09		0.12		1.21	
10		Jimmy Garoppolo		3696		512		7.22		1.06		-0.13		0.93	
12		Aaron Rodgers		4058		605		6.71		0.55		-0.01		0.54	
13		Teddy Bridgewater		1385		208		6.66		0.50		-0.02		0.48	
11		Deshawn Watson		3575		539		6.63		0.47		0.16		0.64	
14		Jared Goff		4188		648		6.46		0.30		0.10		0.40	
16		Gardner Minshew		3237		503		6.44		0.28		-0.14		0.14	

# Player Valuation – QBs and Passing

- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - RTG = Passer Rating
  - Range 0-158.3 (???)
  - Seems great/comprehensive...actually stinks

$$\text{Passer Rating} = \left( \frac{\left( \frac{\text{COMP}}{\text{ATT}} - 0.3 \right) * 5 + \left( \frac{\text{YARDS}}{\text{ATT}} - 3 \right) * 0.25 + \left( \frac{\text{TD}}{\text{ATT}} \right) * 20 + 2.375 - \left( \frac{\text{INT}}{\text{ATT}} * 25 \right)}{6} \right) * 100$$

If Quarterback A completes three passes for three yards in a row, then they would have a passer rating of 79.17.

Meanwhile, Quarterback B throws three straight passes, with the first two falling incomplete, while the third is caught for a 30-yard gain. The quarterback's passer rating in this circumstance is 43.75.<sup>[11]</sup>

Los Angeles Passing

	C/ATT	YDS	Avg	TD	INT	SACKS	RTG
Justin Herbert	26/38	281	7.4	4	0	1-6	125.0
TEAM	26/38	275	7.4	4	0	1-6	125.0

Kansas City Passing

	C/ATT	YDS	Avg	TD	INT	SACKS	RTG
Patrick Mahomes	27/44	260	5.9	3	2	2-9	81.6
TEAM	27/44	251	5.9	3	2	2-9	81.6

# Player Valuation – QBs and Passing

- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - **QBR = (Total) Quarterback Rating**
  - Not *bad* at all. Accounts for rushes, uses EPA framework, attempts to divide credit, scales to 0-100.
  - Rough interpretation: chance of QB winning with an average supporting cast
  - Drawbacks: proprietary (must trust ESPN, can't interrogate it ourselves); maybe outdated vs. modern methods?

 Los Angeles Passing

	C/ATT	YDS	Avg	TD	INT	SACKS	QBR	RTG
Justin Herbert	26/38	281	7.4	4	0	1-6	75.4	125.0
TEAM	<b>26/38</b>	<b>275</b>	<b>7.4</b>	<b>4</b>	<b>0</b>	<b>1-6</b>	<b>--</b>	<b>125.0</b>

 Kansas City Passing

	C/ATT	YDS	Avg	TD	INT	SACKS	QBR	RTG
Patrick Mahomes	27/44	260	5.9	3	2	2-9	64.6	81.6
TEAM	<b>27/44</b>	<b>251</b>	<b>5.9</b>	<b>3</b>	<b>2</b>	<b>2-9</b>	<b>--</b>	<b>81.6</b>

# Player Valuation – QBs and Passing

---

- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - **QBR = (Total) Quarterback Rating leaders, 2019**

Total QBR - All NFL

RK	NAME	QBR	PAA	PLAYS	EPA
1	Lamar Jackson BAL	83.0	66.7	613	103.7
2	Patrick Mahomes KC	77.7	55.8	585	97.3
3	Drew Brees NO	73.3	33.7	419	62.6
4	Dak Prescott DAL	71.9	48.1	690	93.1
5	Russell Wilson SEA	71.5	43.0	674	90.9
6	Matthew Stafford DET	71.3	27.0	353	56.1
7	Deshawn Watson HOU	70.5	42.3	662	91.6
8	Ryan Fitzpatrick MIA	68.3	30.2	620	70.7
9	Ryan Tannehill TEN	64.2	19.0	374	45.8
10	Derek Carr OAK	64.1	27.3	598	70.3
11	Carson Wentz PHI	62.8	32.2	745	90.5
12	Jimmy Garoppolo SF	60.8	20.8	593	69.3
13	Kirk Cousins MIN	60.4	18.1	537	59.0
14	Matt Ryan ATL	59.6	22.3	748	75.1
15	Kyler Murray ARI	57.7	16.2	709	71.6

Source: [https://www.espn.com/nfl/qbr/\\_/season/2019/seasontype/2](https://www.espn.com/nfl/qbr/_/season/2019/seasontype/2)

# Player Valuation – QBs and Passing

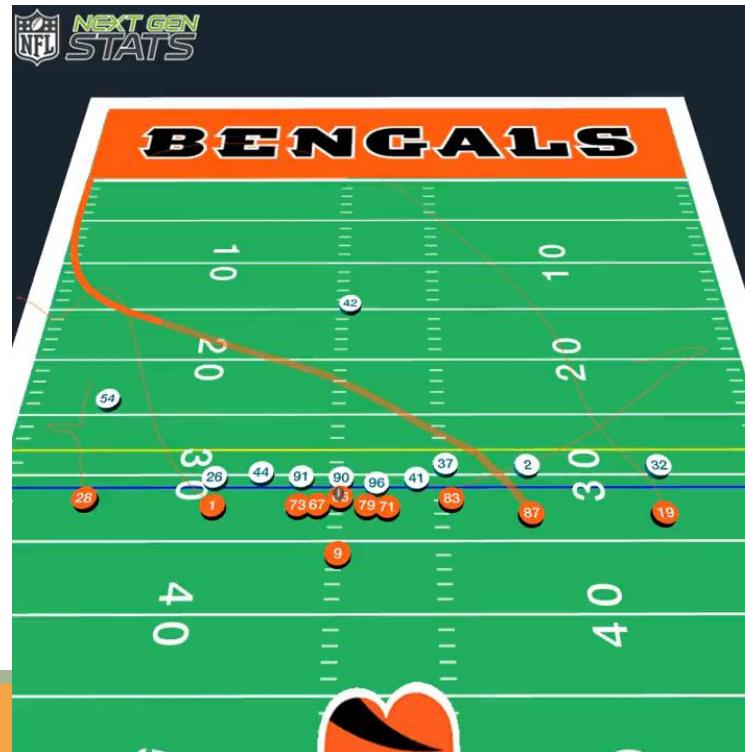
- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - Split passes into two pieces: **air yards** (a/k/a **average depth of target, aDOT**) and **yards after the catch (YAC)**

 Next Gen Stats  @NextGenStats · 14h  
Joe Burrow & C.J. Uzomah (31-yard TD)

Burrow held for 6.21 seconds on the play, enough time for Uzomah to find space in the intermediate on a hitch route turned crosser.

- ◆ Yards After Catch: 17
- ◆ YAC Over Expected: +5
- ◆ TD Probability: 27.5%

#JAXvsCIN | #RuleTheJungle 



Source:

<https://twitter.com/NextGenStats/status/1443773324186685441>

# Player Valuation – QBs and Passing

- Context. Is. King.
- When considering **air yards** vs. **YAC**: credit split for QBs vs. WR/TEs? Implications for evaluating completion percentage, game play styles?

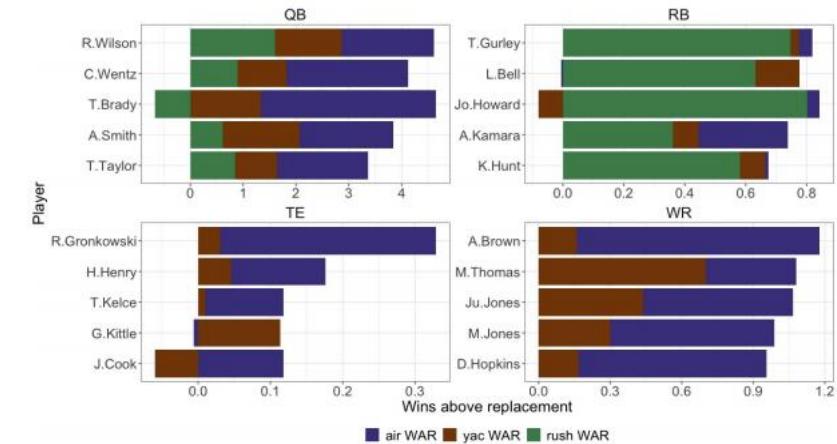
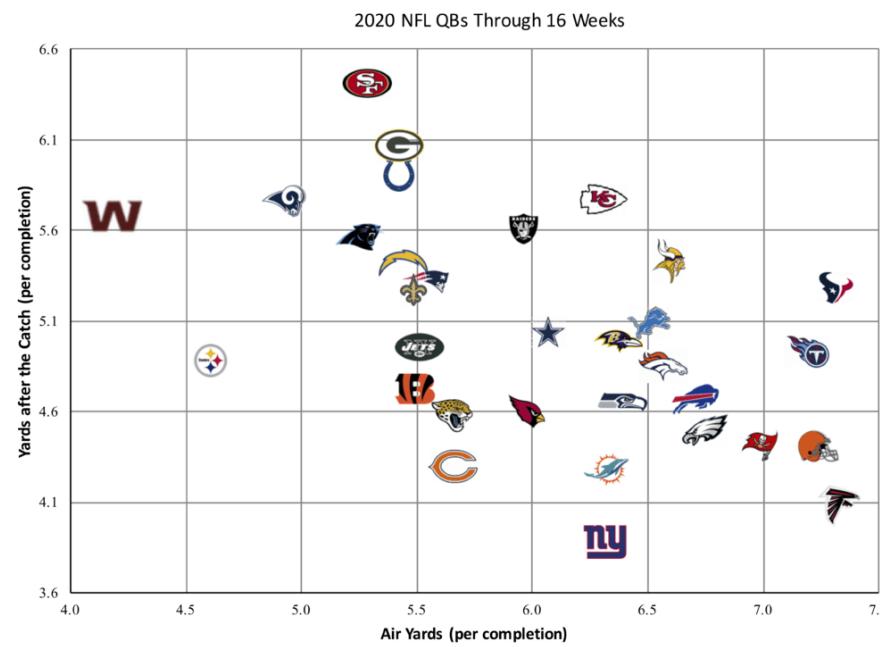
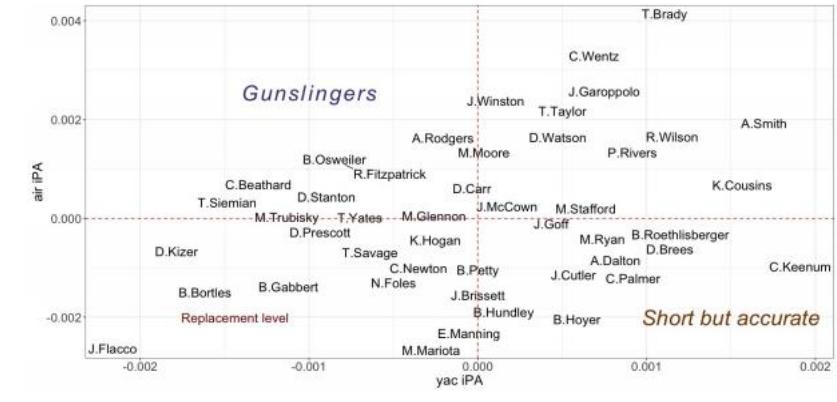
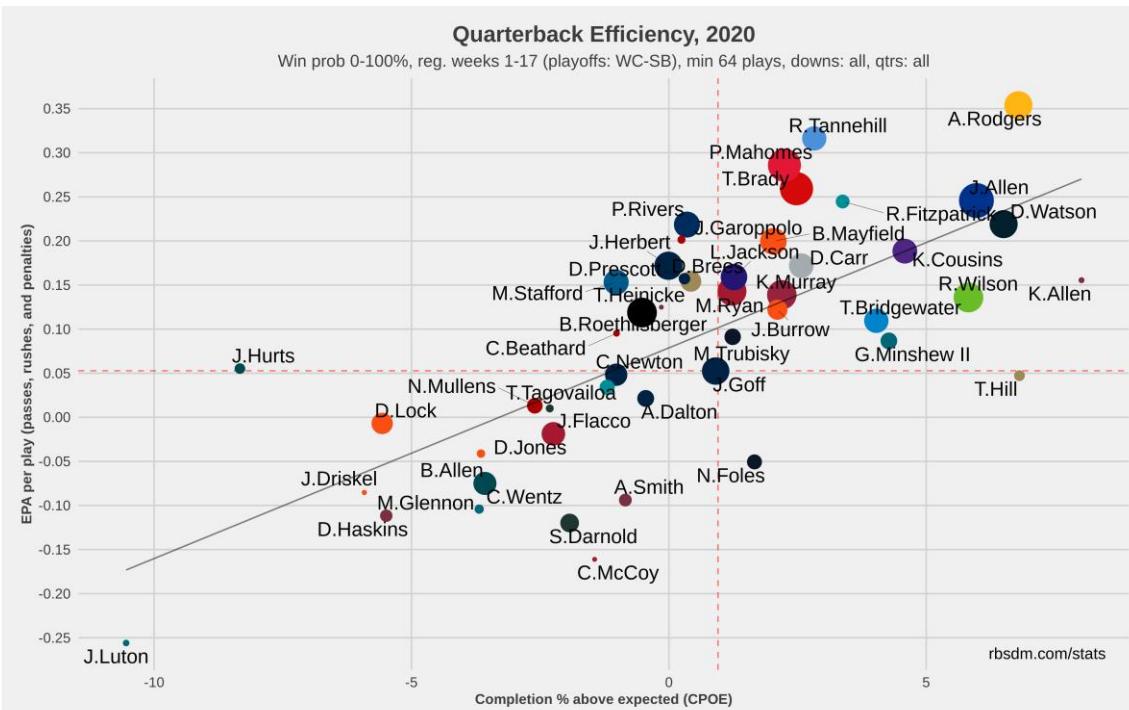


Figure 13: Top five players in WAR by position for the 2017 season.



# Player Valuation – QBs and Passing

- Context. Is. King.
- Let's focus on the most important position on the field: QBs
  - (Direct) EPA and WPA frameworks



Player	Team	Plays	EPA+CPOE composite	EPA/play
A.Rodgers	(Green Bay Packers logo)	700	0.212	0.361
R.Tannehill	(Tennessee Titans logo)	597	0.180	0.327
P.Mahomes	(Kansas City Chiefs logo)	830	0.167	0.296
T.Brady	(New England Patriots logo)	834	0.163	0.280
J.Allen	(Buffalo Bills logo)	879	0.173	0.256
R.Fitzpatrick	(Miami Dolphins logo)	318	0.155	0.247
P.Rivers	(Indianapolis Colts logo)	642	0.133	0.231
D.Watson	(Houston Texans logo)	698	0.165	0.228
K.Cousins	(Minnesota Vikings logo)	610	0.150	0.216
D.Carr	(Oakland Raiders logo)	607	0.138	0.213
B.Mayfield	(Cleveland Browns logo)	658	0.135	0.212

Source: <https://rbsdm.com/stats/stats/>

# Player Valuation – QBs and Passing

- **Context. Is. King.**
- Let's focus on the most important position on the field: QBs
  - Translate EPA and WPA to **WAR**

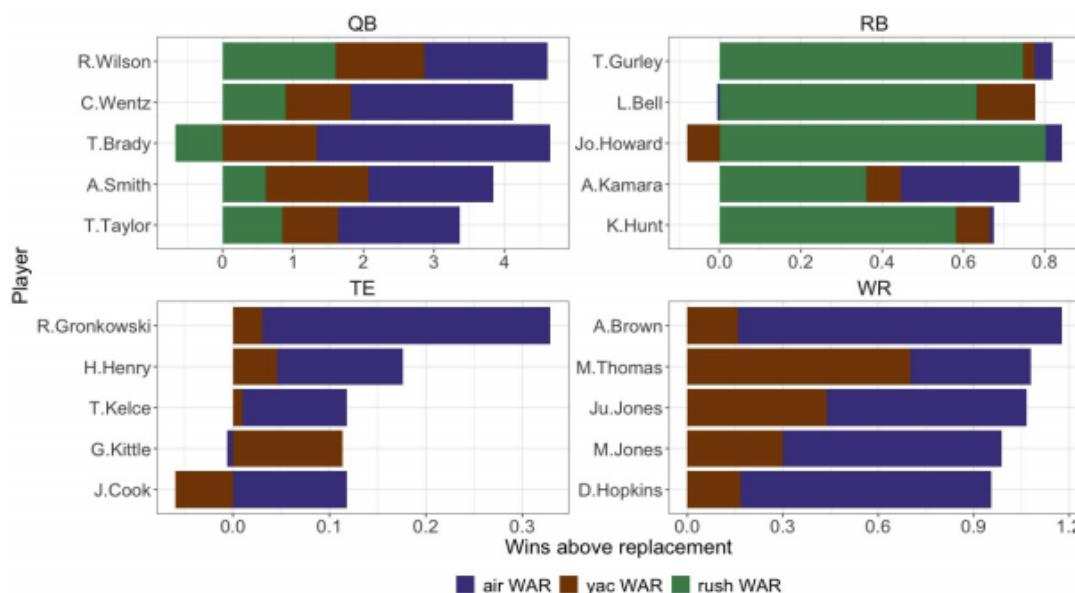
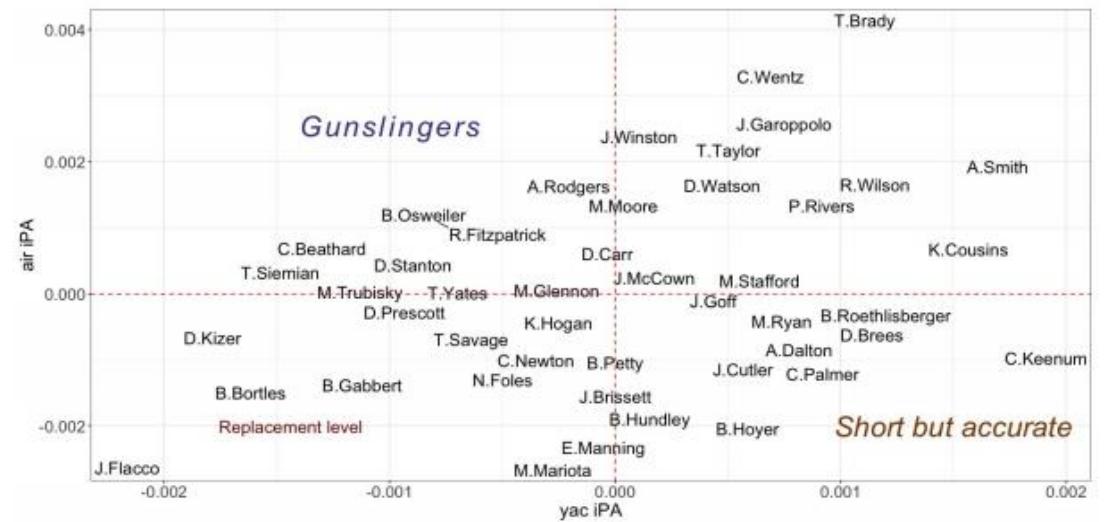


Figure 13: Top five players in **WAR** by position for the 2017 season.



Source: <https://arxiv.org/abs/1802.00998>

# Player Valuation – Other Positions

---

- **Context. Is. King.**
- Can apply many of these same ideas to other (offensive skill) positions
  - Volume vs. Rate stats (Rushing yards vs. YPC)
  - EPA, WPA, and WAR for all players
  - Wide Receivers (WRs) and Tight Ends (TEs): YAC vs. aDOT
  - Running Backs (RBs): yards before contact vs. yards after first contact, etc.

# Tracking/Next Gen Stats (NGS) Data

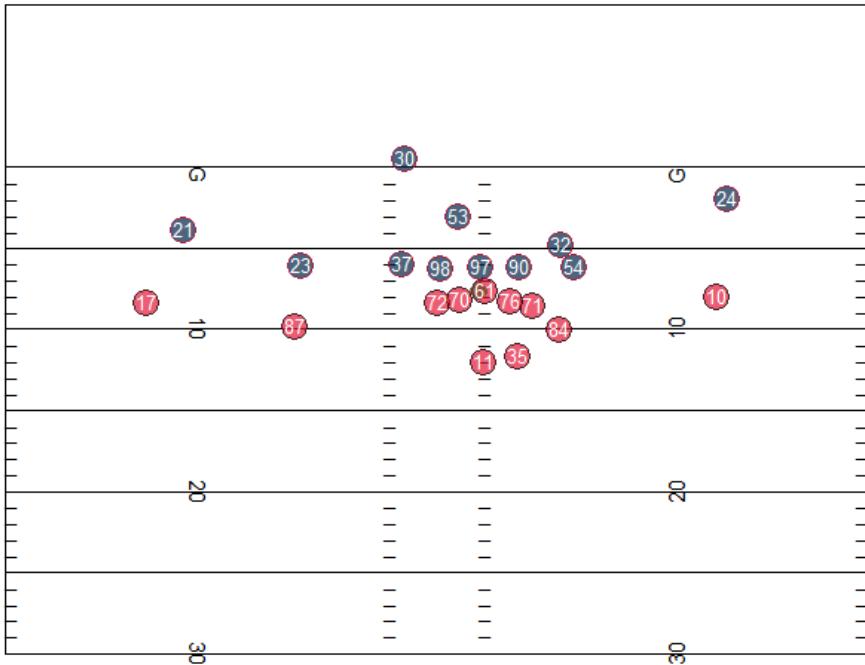
---

APPLICATIONS (PLAY AND PLAYER VALUATION)

BIG DATA BOWL

# What is NFL Tracking Data?

- Also called **Next Gen Stats (NGS)**
- Chips in shoulder pads and ball
- Record and transmit position (and by extension speed, distance, accel) every 0.1s



Performance	Classification	Advanced
Distance Traveled	Formations (offense/defense)	Completion Probability
Max Speed	Coverage	Expected Rushing Yards
Time on Field	Route Detection	Win Probability

# What is NFL Tracking Data?

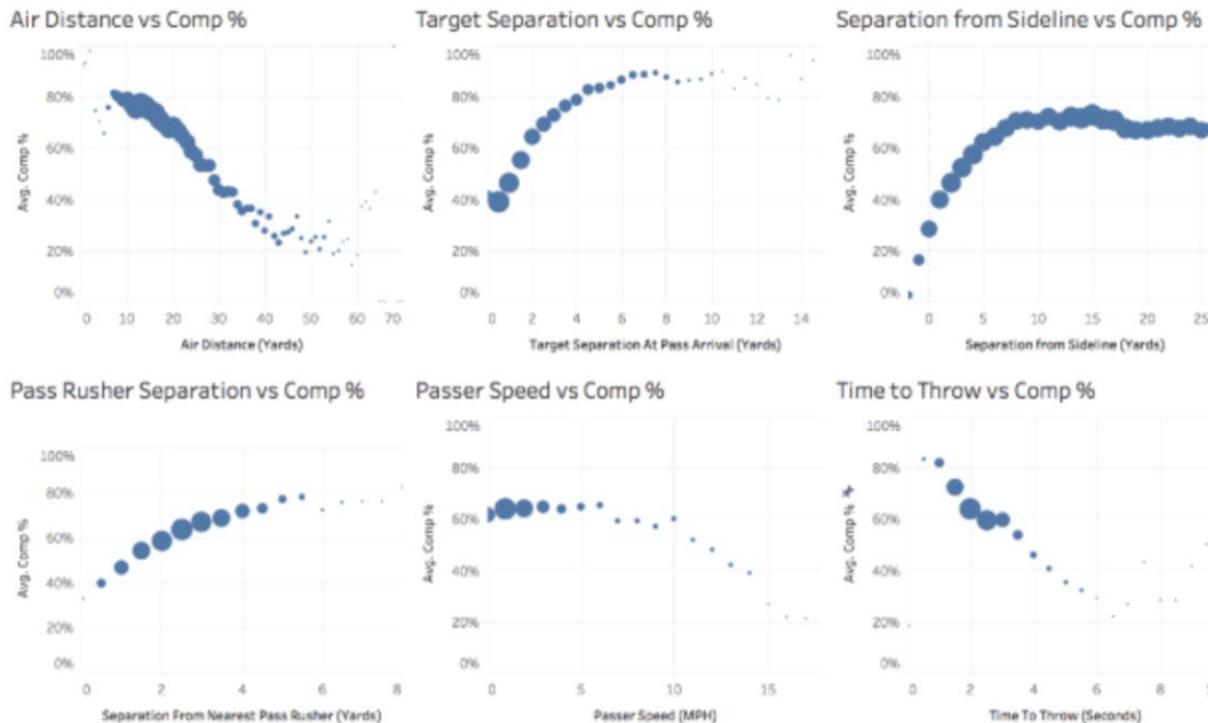
---

- Chips in shoulder pads and ball
- Record and transmit position (and by extension speed, distance, accel) every 0.1s

frame.id	x	y	s	dir	event	displayName
24	60.64	29.70	7.55	175.34	handoff	Cordarrelle Patterson
25	60.77	28.94	7.61	177.10	NA	Cordarrelle Patterson
:	:	:	:	:	:	:
44	55.20	14.62	8.92	226.45	first_contact	Cordarrelle Patterson
:	:	:	:	:	:	:

# Uses of Tracking Data

- Even more advanced play valuation (passes):
- **Expected completion percent** for any pass

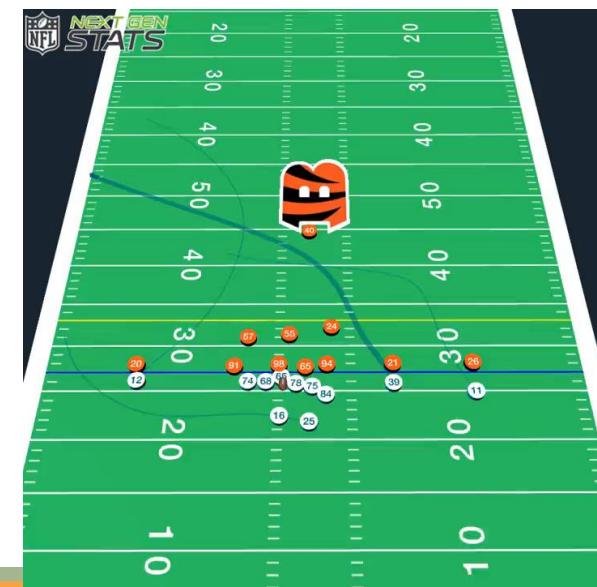


Trevor Lawrence & Jamal Agnew (27 yards)

- ◆ Air Distance: 42.3 yards
- ◆ Sideline Distance: 1.0 yards
- ◆ Target Separation: 2.1 yards

Completion Probability: 29.2%

#JAXvsCIN | #DUUUVAL 🐾

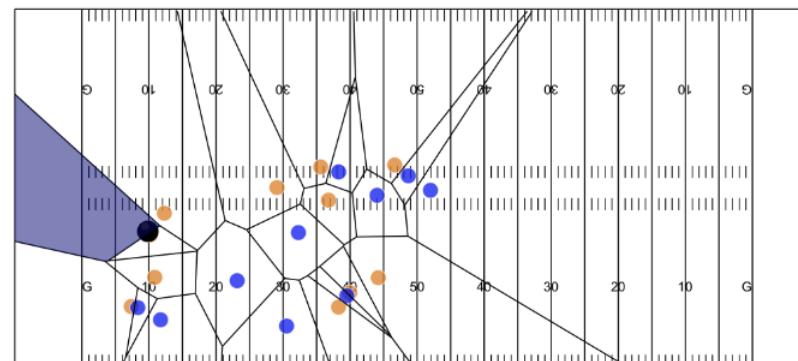


# Uses of Tracking Data

- Even more advanced play valuation (runs):

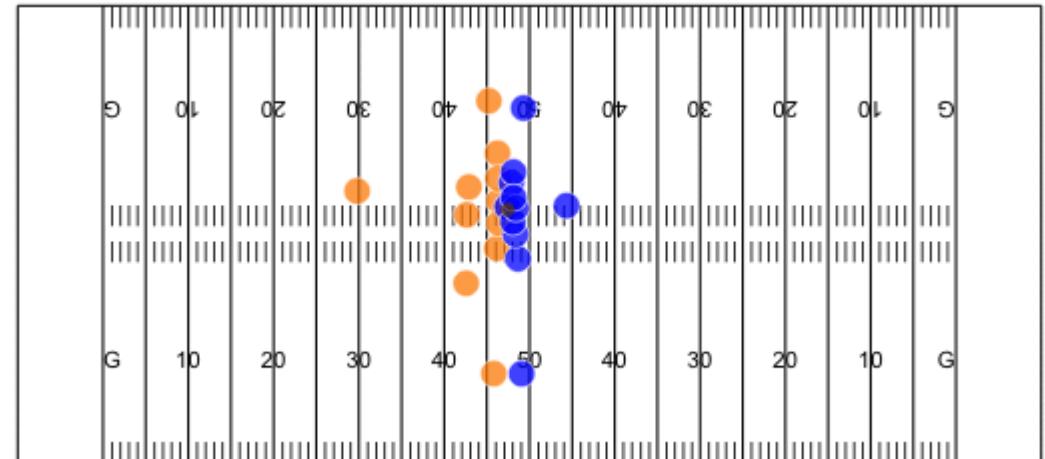
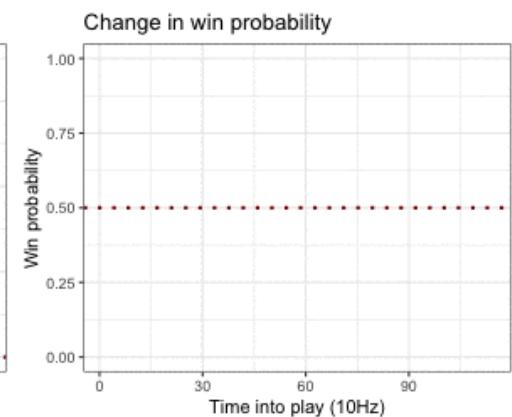
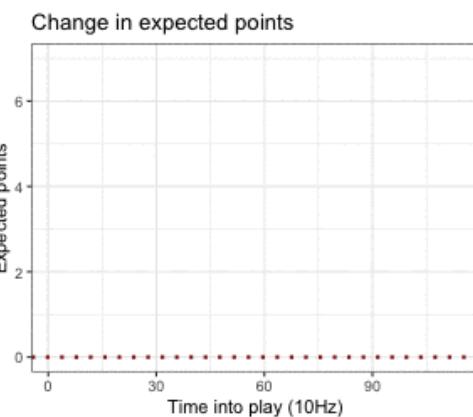
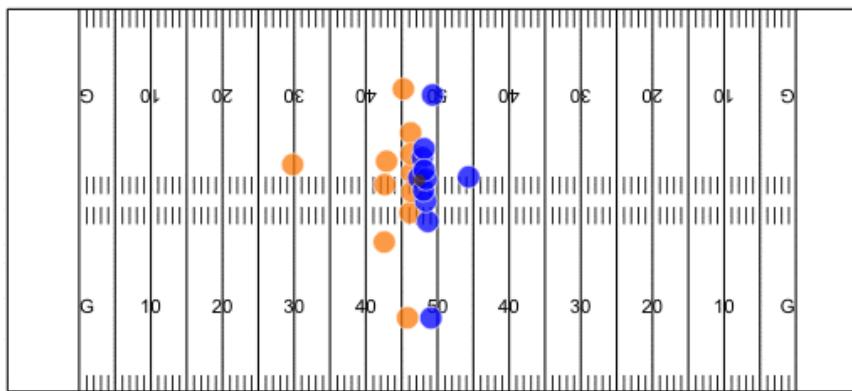
## Ball-carrier model features

- Split players into 3 groups: ball-carrier, offense, and defense:
  - Use (x, y), speed, direction, and distance traveled from previous frame
- Order offense/defense players using distance to ball-carrier
  - e.g. defense2\_x gives x coordinate for second closest defender
- Voronoi tessellations summary:
  - Ball carrier's area
  - x-coordinate of closest and farthest points from target endzone
  - Indicator if surrounded by teammates



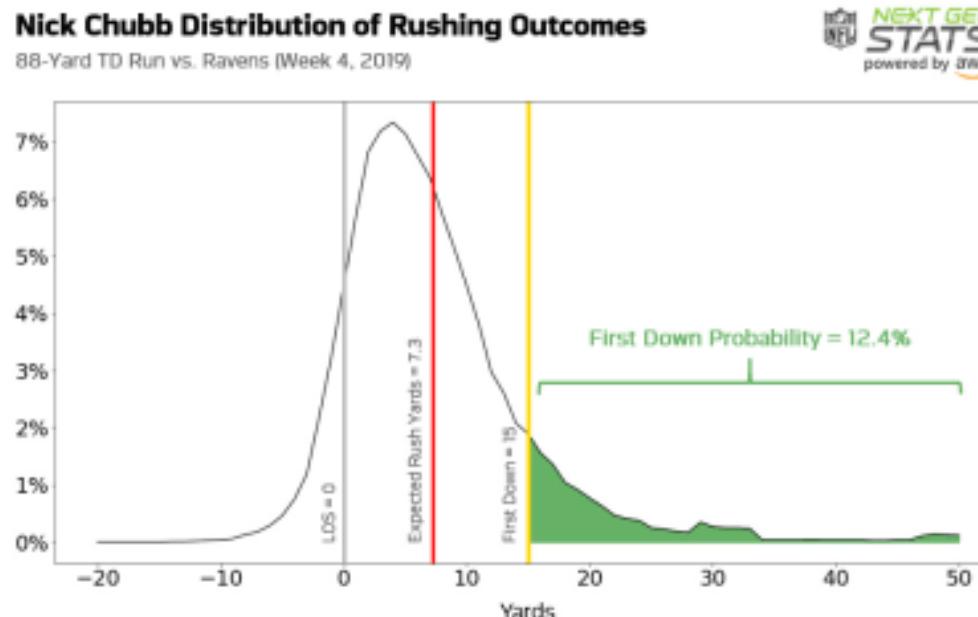
# Uses of Tracking Data

- Even more advanced play valuation (runs):
- **Expected rushing yards** vs. actual rushing yards



# Uses of Tracking Data

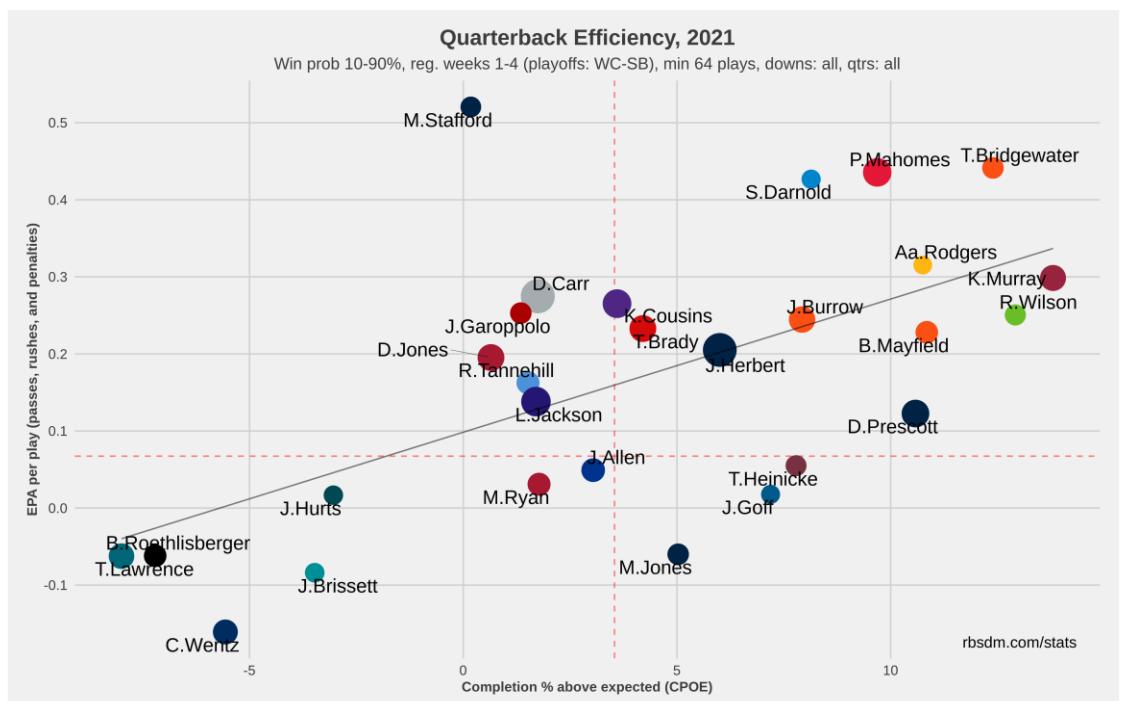
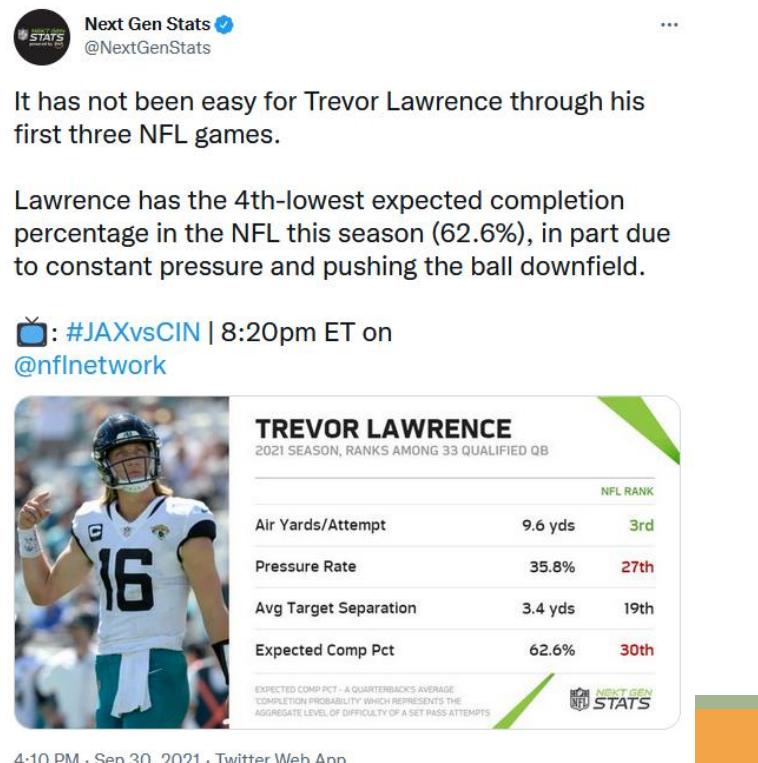
- Even more advanced play valuation (runs):
  - **Expected rushing yards** vs. actual rushing yards
  - “The Zoo” 2019 Big Data Bowl
  - Deep learning model for runs just letting a neural net sort out predictions based on location, speed, and direction of all 22 players relative to rusher and opposing players. Produces:



Source: <https://www.nfl.com/news/next-gen-stats-intro-to-expected-rushing-yards>

# Uses of Tracking Data

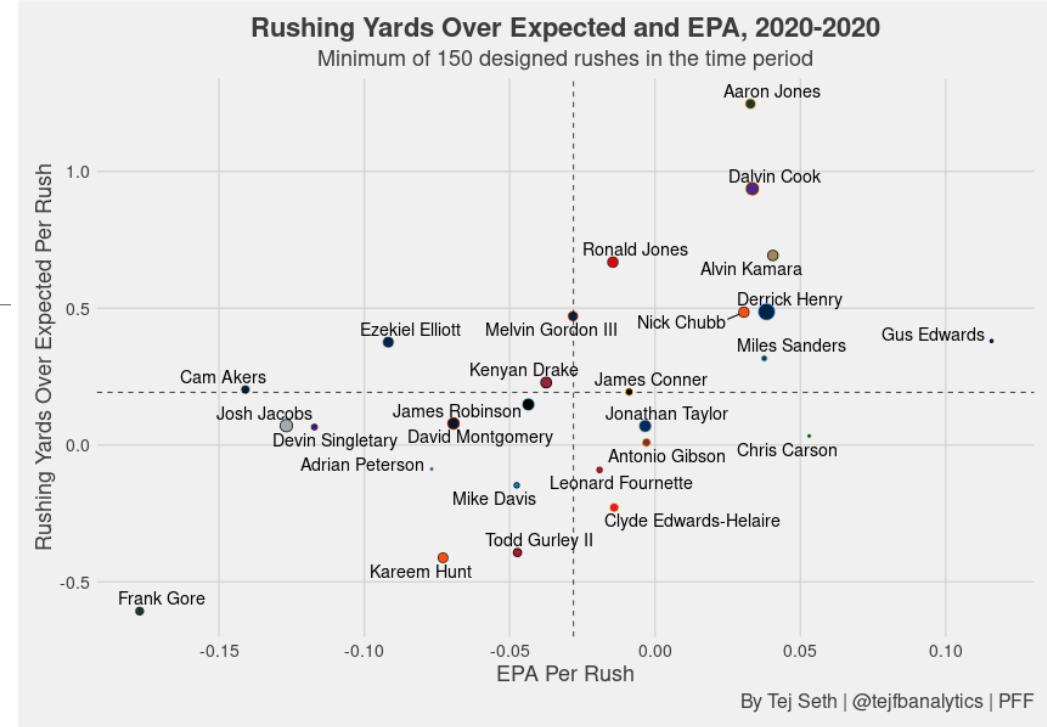
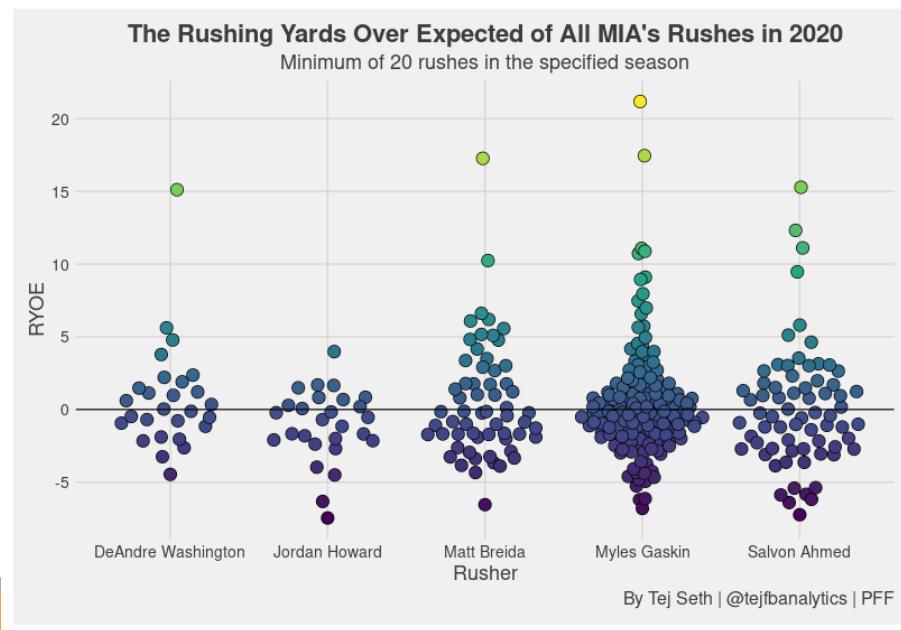
- More advanced player valuation:
  - Context is king.** Key improvement: how do players actually perform **relative to similar situations (that is, vs. how we would expect them to do?)**



Source: <https://twitter.com/NextGenStats/status/1443669550805880846?s=20>; <https://rbsdm.com/stats/stats/>

# Uses of Tracking Data

- More advanced player valuation:
  - Context is king.** Key improvement: how do players actually perform **relative to similar situations (that is, vs. how we would expect them to do?)**



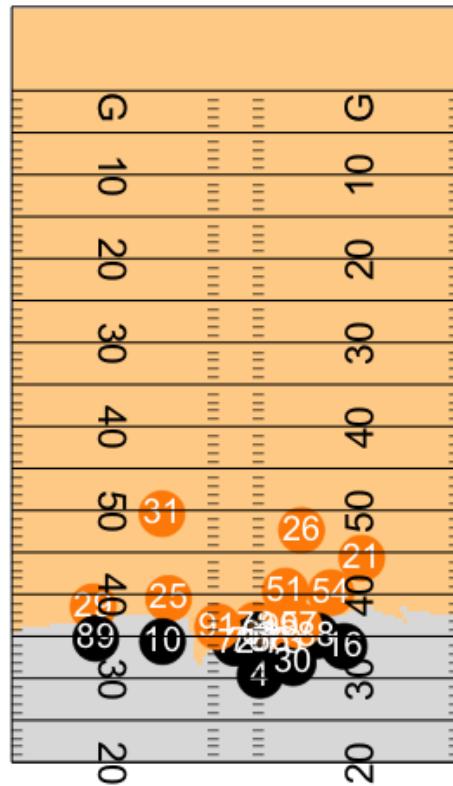
RANK	PLAYER	TEAM	RUSHES	EPA/RUSH	YARDS PER CARRY	EXPECTED YARDS	RUSHING YARDS OVER EXPECTED
1	Aaron Jones	G	220	0.03	5.58	4.33	1.25
2	Dalvin Cook	V	309	0.03	4.95	4.01	0.94
3	Alvin Kamara	S	228	0.04	4.83	4.14	0.69
4	Ronald Jones	TB	230	-0.01	5.16	4.49	0.67
5	Derrick Henry	T	396	0.04	5.09	4.60	0.49
6	Nick Chubb	C	225	0.03	5.43	4.94	0.49
7	Melvin Gordon III	B	213	-0.03	4.68	4.21	0.47
8	Ezekiel Elliott	D	241	-0.09	4.08	3.71	0.38
9	Gus Edwards	R	157	0.12	4.90	4.52	0.38

Source: <https://mfbanalytics.shinyapps.io/RYOE/>

# Uses of Tracking Data

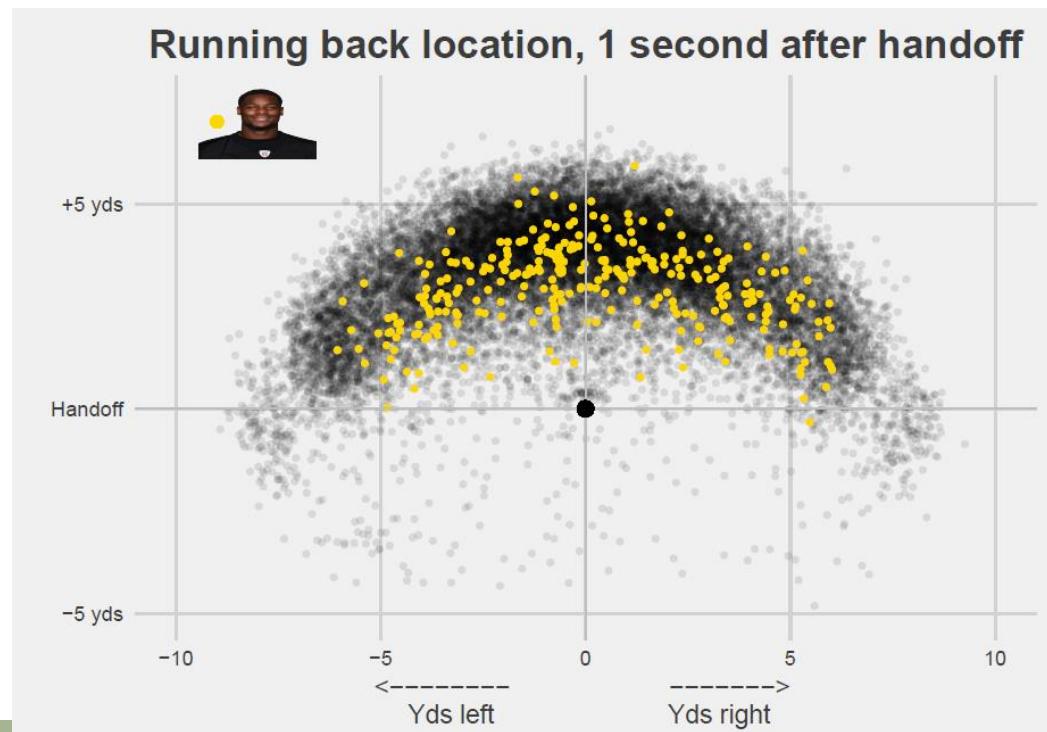
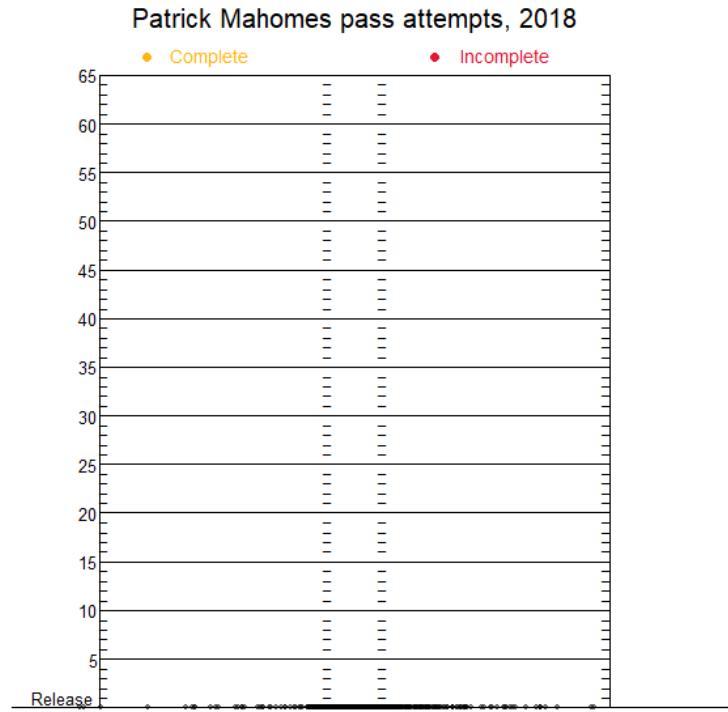
---

- Field ownership/zone of control:



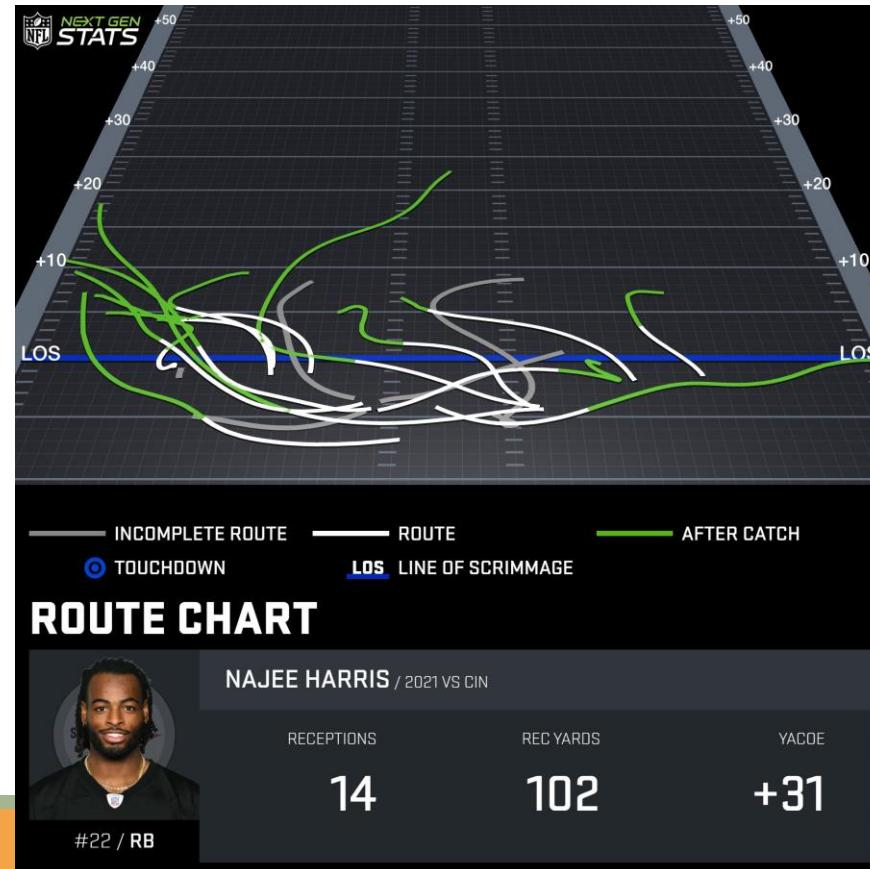
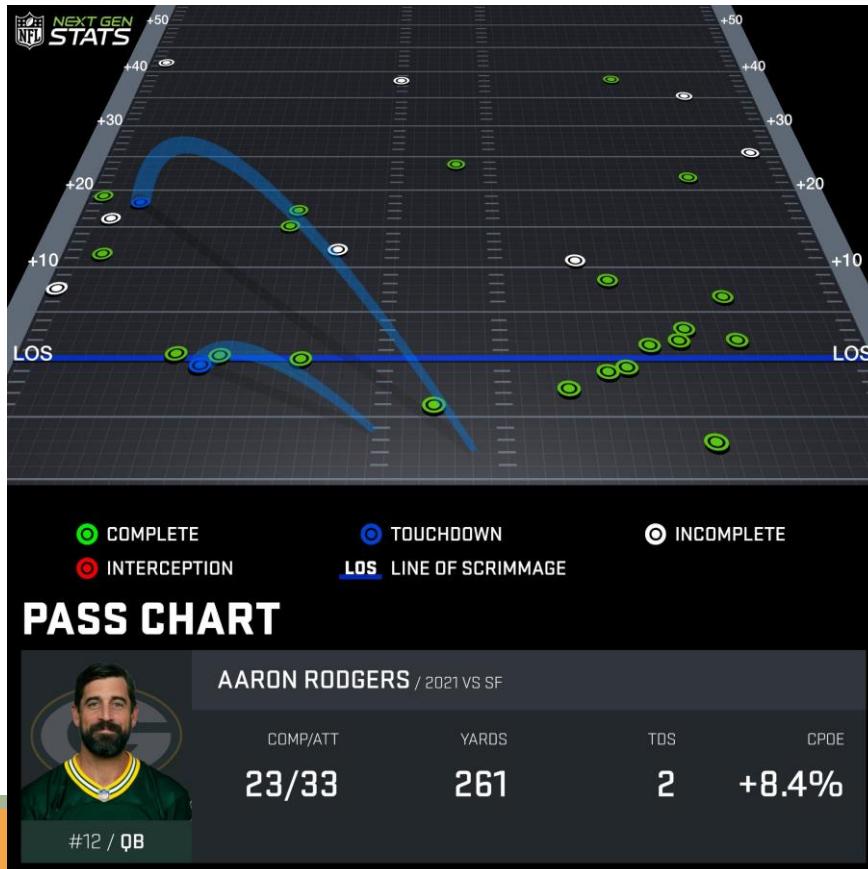
# Uses of Tracking Data

- New ways to look at player trends, do automated opponent scouting



# Uses of Tracking Data

- NFL's own NGS team produces novel public charts and stats for fans and broadcasts (check out [NGS Twitter feed](#) and [website](#) for more)



Source: <https://twitter.com/NextGenStats>

# Uses of Tracking Data

---

- **Big Data Bowl**
  - Since 2018, NFL has released a selection of tracking data for public analysis in a competition
  - Themes:
    - 2018 – WR Route and Passing Analytics
    - 2019 – Predict Rush Yards
    - 2020 – Pass Defense Analytics
    - 2021 – Punt Analytics
  - Other contests on Kaggle from NFL Health & Safety, too (punt play injury prevention, helmet video tagging for head injuries)



# Uses of Tracking Data

---

- **Big Data Bowl**
  - Interested?
    - 2021 contest ongoing now on Kaggle
    - Talk with others, form a team; Oxford Sports Analytics Club?
    - Check the websites:  
<https://operations.nfl.com/gameday/analytics/big-data-bowl/> and <https://www.kaggle.com/c/nfl-big-data-bowl-2021>
    - Follow @statsbylopez and #BigDataBowl on Twitter (and GET TWITTER if you don't already have an account!)
    - If you like streamers, follow [my man Nick Wan \(formerly Reds analytics, now – I'm not kidding – KFC\)](#) analyzing the data on Twitch



# Lineman Analytics: Pass Blocking and Rushing

---

SURVIVAL ANALYSIS

# Lineman Analytics

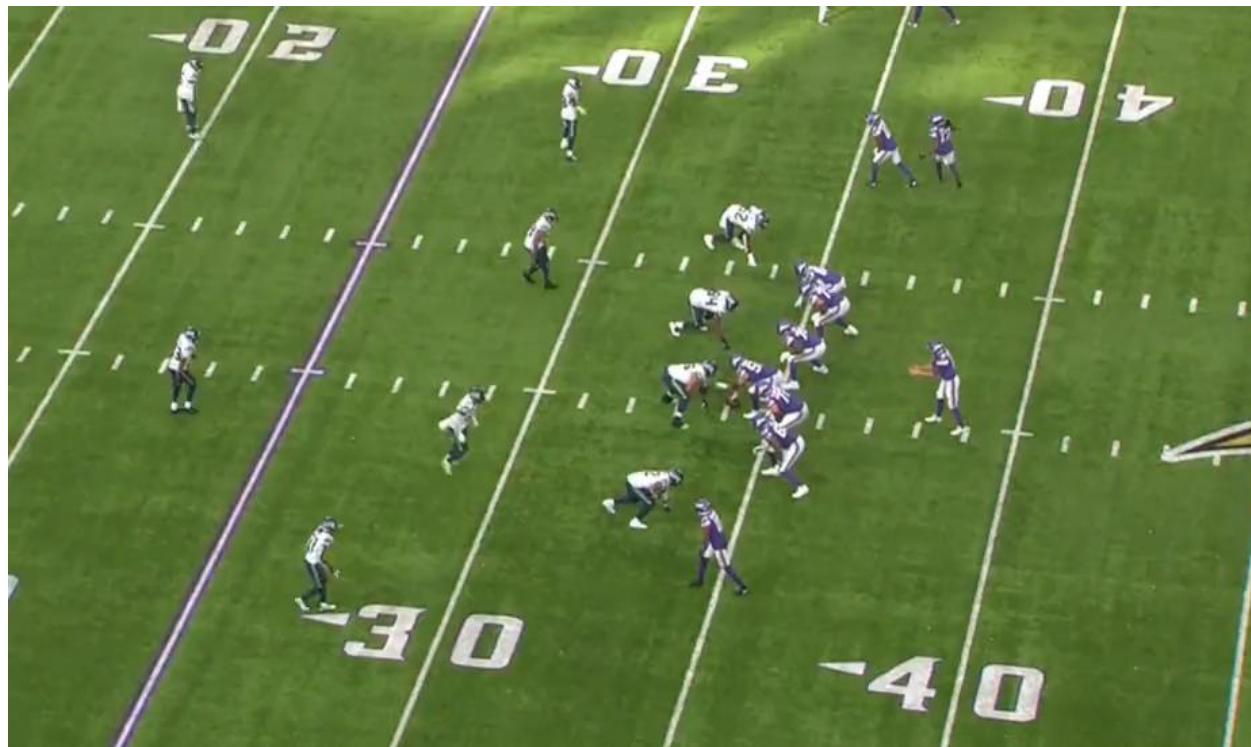
---

- Evaluating offensive skill players (QB, RB, WR, TE) *fairly* advanced
  - “See” their contributions easily in a throw, run, or catch
- Offensive line (OL) and defense lag behind due to few or unstable (INTs, Sacks) individual traditional stats
  - (Public) charting of defense, in particular, has become more accessible in recent years (pass break-ups, QB pressures, etc.) – but still missing contributions on plays where player doesn’t “do” something distinct

# OL Pass Blocking Analytics

---

- How can you evaluate an offensive or defensive lineman?
  - What are their jobs on this play?



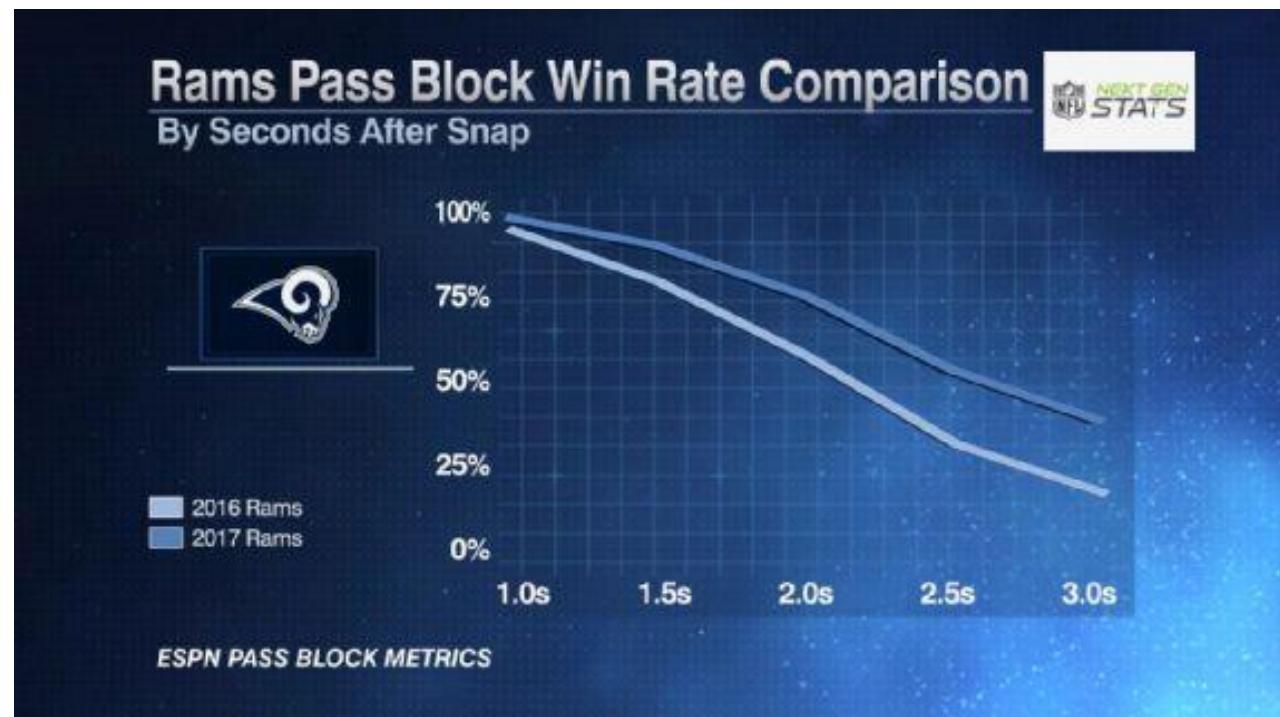
# Linemen Pass Blocking Analytics

---

- We already know we can use statistics to predict:
  - 1. The value of something (**linear regression**)
  - 2. The probability of something (**logistic regression**)
- But we can also predict:
  - 3. The time it takes for something to happen (**survival analysis**)

# Linemen Pass Blocking Analytics

- **Survival analysis:** model the “time-to-event”
  - Pass blocking/rushing: time to “beaten” block
  - Requires tracking data
- Is longer or shorter better here for a.) OLs and b.) DLs?



# Linemen Pass Blocking Analytics

- **Survival analysis:** model the “time-to-event”
  - ESPN also calculates individual **pass rush win rate** (PRWR) and **pass block win rate** (PBWR) for DLs and OLs, respectively
    - % of passing plays in which you beat/sustain block for 2.5s



# Thanks!

---

- Questions? [zbinney@emory.edu](mailto:zbinney@emory.edu), @binney\_z on Twitter

