The Most Popular/Best Performing Route Combinations in 2020

By: Nicolas Zacharis

Tools/Data Used For This Analysis

- I utilized Python for the entirety of this project
  - All tables and calculated numbers were generated through the many capabilities of Python
- I only used the provided data in the project description in order to come up with my conclusions
  - This data was more than sufficient for my analysis of the prompt, however it did come with some limitations that I will discuss

Part 1: What Were the Most Popular Route Combinations in the NFL in 2020?

- Assumptions Made
  - How did i consolidate/clean up the data prior to analysis?
    - I created route groupings which allowed for easier analysis and less route combinations to be working with
    - Below are the groupings I made:

Original Route Name	**New Route Group**
Any route with ‘Flat’	Flat
Any route with ‘Swing’	Swing
Chip Routes: ‘Curl’, ‘Drag’, ‘Seam’	Curl, Drag, Seam
Screens: ‘Bubble’, ‘Beneath’, ‘Quick’, ‘Tunnel’, ‘Drag’	WR Screen

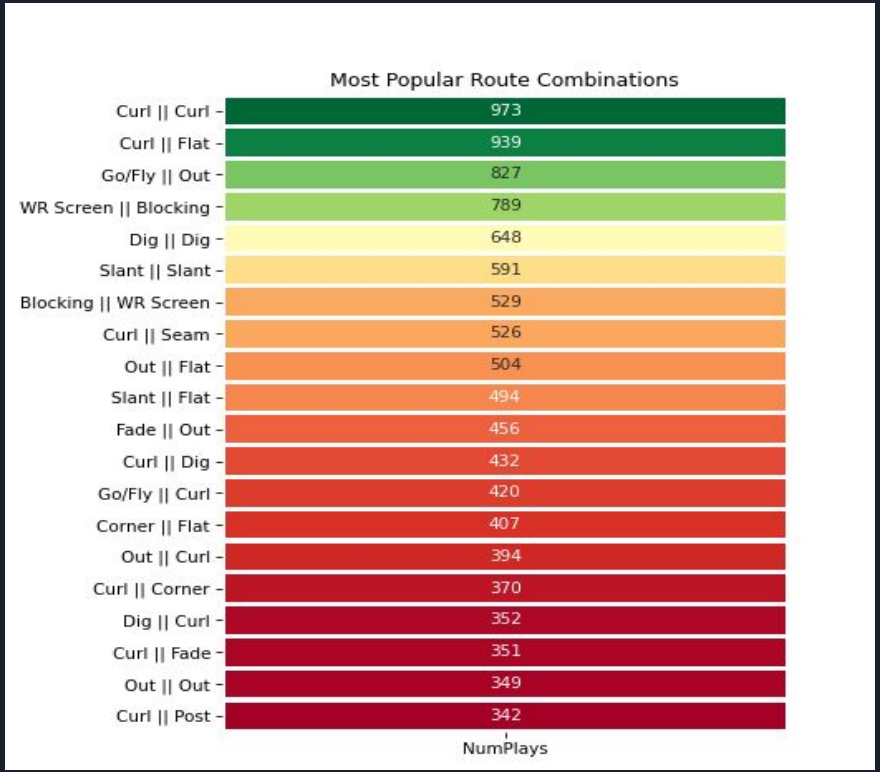
  - How did I define what a route combination was?
    - I defined a route combination as any combination of an outside receiver and the receiver lined up directly inside of them (WR1 and WR2) , ignoring side of the field

Part 1: What Were the Most Popular Route Combinations in the NFL in 2020?

- Evaluation Method
  - In order to determine the most popular route combos, I believe a pure count of the number of times each combination is used accurately depicts the popularity of a route combination
  - In order to have play by play data alongside route data, I first joined those two tables together
  - I then went through the joined table play by play, creating a counter for each route combo I encountered that incremented if a route combo appears in a play

Part 1: What Were the Most Popular Route Combinations in the NFL in 2020?

- Final Results
  - I filtered out the massive outlier of both WRs running a 'Flat' (about 16000 plays) and here are the results
  - The table on the right displays the number of plays by route combination for the 20 most popular combos in the format of :
    - **'WR1's route' || 'WR2's route'**



Part 2: Which Route Combinations Perform Best Against Each Coverage Scheme?

- Assumptions Made
  - For the next part of this project I examined key metrics to compare route combos performances
    - Due to a limited amount of data I had to come up with cutoffs/qualifying numbers so that each route combo at least **qualified for analysis** vs a particular coverage scheme
    - These cutoffs will make more sense in a bit:
      - ***Minimum of 5 Targets***
      - ***Minimum of 5 Receptions***
      - ***Minimum of 30 Routes Run***
    - Each of these are important in limiting any outliers (**ex. 1 target vs Cover 1**) yet at the same time including as much data as possible in my analysis
      - I found these numbers to be the sweet spots

Part 2: Which Route Combinations Perform Best Against Each Coverage Scheme?

- Evaluation Method
  - To come up with the top performing route combinations, I decided to analyze ***7 Key Metrics***
    - The following metrics were computed and ranked on a ***route combination vs cov scheme*** basis
      - *Yards/Target* - **Depicts all around performance of a route combo**
      - *Yards/Route Run* - **Adds in a level of increased detail of performance on a play basis**
      - *EPA/Target* - **More of an advance metric to see how much avg EPA a route generates**
      - *TD/Reception* - **Depicts success in the redzone more accurately**
      - *1st Down % / Reception* - **Great for depicting success on 3rd downs**
      - *INT % / Target* - **Good metric for analyzing the "safety" of a route combo**
      - *Success Rate (Positive EPA % / Play)* - **Analyzes Positive EPA generation more on a per play basis rather than just relying on big plays**
    - The assumptions made stated on the previous slide are utilized for each of these metrics in order to get the most accurate yet impactful analysis possible\
    - I felt like each of these metrics played a role in coming up with a route combinations overall performance against a given coverage scheme

Part 2: Which Route Combinations Perform Best Against Each Coverage Scheme?

- Evaluation Method
  - For each of the previously described metrics, I computed a qualifying route combo's value for that metric and ranked them compared to all other qualifying route combos
  - This generated the **top 3** route combinations against each coverage scheme for that given metric
  - For example, the following depicts my results for **Success Rate**

Top Routes By Success Rate

	1st	1st SuccessRate	2nd	2nd SuccessRate	3rd	3rd SuccessRate
Cover 0	WR1: Slant WR2: Slant	71.4%	WR1: Slant WR2: Flat	60%	WR1: Out WR2: Flat	50%
Cover 1	WR1: Curl WR2: Post	75%	WR1: Out WR2: Out	73%	WR1: Curl WR2: Curl	63.2%
Cover 2	WR1: Go/Fly WR2: Curl	81.8%	WR1: Curl WR2: Dig	70%	WR1: Go/Fly WR2: Out	69%
Cover 3	WR1: Go/Fly WR2: Curl	74.5%	WR1: Curl WR2: Corner	74.1%	WR1: Fade WR2: Out	73.7%
Cover 4	WR1: Curl WR2: Fade	85.7%	WR1: Fade WR2: Out	68.2%	WR1: Curl WR2: Post	66.7%
Cover 6	WR1: Curl WR2: Flat	81.8%	WR1: Out WR2: Flat	71.4%	WR1: Go/Fly WR2: Curl	66.7%
Man Cover 2	WR1: Curl WR2: Curl	75%	WR1: Slant WR2: Slant	72.7%	WR1: Out WR2: Curl	66.7%
Screen	WR1: WR Screen WR2: Blocking	45.7%	WR1: Blocking WR2: WR Screen	43.3%		
Tampa 2	WR1: Dig WR2: Dig	66.7%	WR1: Curl WR2: Curl	57.1%	WR1: Curl WR2: Flat	50%

Part 2: Which Route Combinations Perform Best Against Each Coverage Scheme?

- Final Results
  - For each of the metrics I generated a similar table and gave each **1st**, **2nd,** and **3rd** place route combination an associated "**score**" of **3**, **2,** and **1**
  - I then summed up these scores for each route combo vs each coverage and divided that value by 21 (7 key metrics multiplied by the max score of 3 for each) to get a ***final confidence score***
  - The following depicts this final score and thus what I came up with as the best performing route combinations vs each coverage scheme!
    - The **Gold,Silver, and Bronze** columns represent the **1st, 2nd, and 3rd** best performing routes against each coverage scheme

Top Routes By Coverage

	1st	1st Score	2nd	2nd Score	3rd	3rd Score
Cover 0	WR1: Slant WR2: Flat	61.9	WR1: Slant WR2: Slant	38.1	WR1: Out WR2: Flat	28.6
Cover 1	WR1: Curl WR2: Fade	33.3	WR1: Curl WR2: Post	33.3	WR1: Out WR2: Flat	14.3
Cover 2	WR1: Go/Fly WR2: Out	33.3	WR1: Fade WR2: Out	28.6	WR1: Go/Fly WR2: Curl	23.8
Cover 3	WR1: Corner WR2: Flat	33.3	WR1: Go/Fly WR2: Curl	28.6	WR1: Curl WR2: Corner	19.0
Cover 4	WR1: Curl WR2: Fade	42.9	WR1: Fade WR2: Out	23.8	WR1: Curl WR2: Seam	19.0
Cover 6	WR1: Go/Fly WR2: Curl	42.9	WR1: Out WR2: Flat	19.0	WR1: Go/Fly WR2: Out	19.0
Man Cover 2	WR1: Slant WR2: Slant	57.1	WR1: Curl WR2: Curl	19.0	WR1: Slant WR2: Flat	19.0
Screen	WR1: WR Screen WR2: Blocking	90.5	WR1: Blocking WR2: WR Screen	38.1		
Tampa 2	WR1: Dig WR2: Dig	57.1	WR1: Go/Fly WR2: Out	33.3	WR1: Curl WR2: Curl	9.5

Limitations and Future Analysis

- With more expansive resources a more in-depth analysis could have been conducted
  - Firstly, I would not have had to limit my data to just the thresholds i stated initially if I had more route data
    - This resource would just include having more play by play route data
  - Secondly, having the total points gained/lost on a play by play basis would have helped incorporate that into my analysis
  - Finally, having data on who the coverage player on a route is would have sparked an analysis into how route combinations performed against players in certain coverage schemes
    - This would be substantial in having another layer of knowledge on route combination performances