

data_pull

August 11, 2022

```
[ ]: # !pip install nflfastpy
      # !pip install pandas
      # !pip install numpy
      # !pip install matplotlib
      # !pip install requests
```

1. Section Title.
Blah blah blah.

```
[ ]: #using nfl-env conda python 3.10.4 environment
import pandas as pd
import numpy as np
import nflfastpy as nfl
from matplotlib import pyplot as plt
from matplotlib import style
import requests
from io import BytesIO

pd.options.display.max_rows = 100
pd.options.display.max_columns = 1000
style.use("ggplot")
style.use("seaborn-notebook")
style.use("fivethirtyeight")
```

```
[ ]: df2018 = nfl.load_pbp_data(2018)
      df2018['year'] = 2018
      df2019 = nfl.load_pbp_data(2019)
      df2019['year'] = 2019
      df2020 = nfl.load_pbp_data(2020)
      df2020['year'] = 2020
      df2021 = nfl.load_pbp_data(2021)
      df2021['year'] = 2021
```

```
[ ]: dfTotal = pd.concat([df2018, df2019, df2020, df2021])

####columns i don't know
# sp
#
```

```

fieldsIcareAbout = ['play_id', 'game_id', 'home_team', 'away_team',␣
    ↳'season_type', 'week',
    ↳'posteam', 'defteam', 'side_of_field', 'yardline_100',␣
    ↳'game_date',
    ↳'quarter_seconds_remaining', 'half_seconds_remaining',␣
    ↳'game_seconds_remaining'
    ↳'game_half', 'quarter_end', 'drive', 'sp', 'qtr', 'down',␣
    ↳'goal_to_go',
    ↳'time', 'yrdln', 'ydstogo', 'ydsnet', 'desc', 'play_type',␣
    ↳'yards_gained',
    ↳'shotgun', 'no_huddle', 'qb_dropback', 'qb_kneel',␣
    ↳'qb_spike', 'qb_scramble',
    ↳'pass_length', 'pass_location', 'air_yards',␣
    ↳'yards_after_catch', 'run_location',
    ↳'run_gap', 'field_goal_result', 'kick_distance',␣
    ↳'extra_point_result', 'two_poin_conv_result',
    ↳'home_timeouts_remaining', 'away_timeouts_remaining',␣
    ↳'timeout', 'timeout_team',
    ↳'td_team', 'td_player_name', 'td_player_id',␣
    ↳'total_home_score', 'total_away_score',
    ↳'posteam_score', 'defteam_score', 'score_differential',␣
    ↳'posteam_score_post', 'defteam_score_post',
    ↳'score_differential_post', 'punt_blocked',␣
    ↳'first_down_rush', 'first_down_pass', 'first_down_penalty',
    ↳'thirty_down_converted', 'third_down_failed',␣
    ↳'fourth_down_converted', 'fourth_down_failed', 'incomplete_pass',
    ↳'touchback', 'interception', 'punt_inside_twenty',␣
    ↳'punt_in_endzone', 'punt_out_of_bounds', 'punt_downed',
    ↳'punt_fair_catch', 'kickoff_inside_twenty',␣
    ↳'kickoff_in_endzone', 'kickoff_out_of_bounds', 'kickoff_downed',
    ↳'kickoff_fair_catch', 'fumble_forced', 'fumble_not_forced',␣
    ↳'fumble_out_of_bounds', 'solo_tackle', 'safety',
    ↳'penalty', 'tackled_for_loss', 'fumble_lost',␣
    ↳'own_kickoff_recovery', 'own_kickoff_recovery_td', 'qb_hit',
    ↳'rush_attempt', 'pass_attempt', 'sack', 'touchdown',␣
    ↳'pass_touchdown', 'rush_touchdown', 'return_touchdown',
    ↳'extra_point_attempt', 'two_point_attempt',␣
    ↳'field_goal_attempt', 'kickoff_attempt', 'punt_attempt', 'fumble',
    ↳'complete_pass', 'assist_tackle', 'lateral_reception',␣
    ↳'lateral_rush', 'lateral_return', 'lateral_recovery',
    ↳'passer_player_id', 'passer_player_name', 'passing_yards',␣
    ↳'receiver_player_id', 'receiver_player_name',

```

```

        'receiving_yards', 'rusher_player_id',␣
↪'rusher_player_name', 'rushing_yards', 'later_receiver_player_name',
        'lateral_receiving_yards', 'later_rusher_player_id',␣
↪'lateral_rusher_player_name', 'lateral_rushing_yards',
        'lateral_sack_player_id', 'lateral_sack_player_name',␣
↪'interception_player_id', 'interception_player_name',
        'lateral_interception_player_id',␣
↪'lateral_interception_player_name', 'tackle_for_loss_1_player_id',␣
↪'tackle_for_loss_1_player_name',
        'tackle_for_loss_2_player_id',␣
↪'tackle_for_loss_2_player_name', 'roof', 'surface', 'temp', 'wind',␣
↪'home_coach', 'away_coach', 'special_teams_play',
        'out_of_bounds',
    ]

    ###Special Teams
    # 'punt_returner_player_id', 'punt_returner_player_name',
    # 'lateral' ]

```

```

[ ]: rushing_stats = ['year', 'yardline_100', 'down', 'ydstogo', 'run_location',␣
↪'defteam', 'fantasy_player_id', 'fantasy_player_name', 'roof', 'surface',␣
↪'temp', 'wind', 'score_differential', 'goal_to_go', 'rushing_yards']
rush_filter = dfTotal['play_type'] == 'run'
goalFilter = dfTotal['goal_to_go'] == 0
dfRushing = dfTotal[rush_filter&goalFilter][rushing_stats].copy(deep=True)
dfRushing['rush_attempts'] = 1

```

```

[ ]: rushing_grouped = dfRushing.groupby(['fantasy_player_name', 'year'],␣
↪as_index=False).agg({'rushing_yards' : sum, 'rush_attempts' : sum })

rushing_grouped['ypc'] = rushing_grouped['rushing_yards']/
↪rushing_grouped['rush_attempts']

```

```

[ ]: #Calculate ADOT by receiver and QB
df_adot = dfTotal.loc[(dfTotal['year'] == 2021) &␣
↪(dfTotal['play_type']=='pass'), ['receiver_player_name', 'air_yards',␣
↪'year']]

```

```

[ ]: df_adot['attempts'] = 1
df_adot2 = df_adot.groupby(['receiver_player_name', 'year'], as_index=False).
↪agg({'air_yards' : sum, 'attempts' : sum})
df_adot2['adot'] = df_adot2['air_yards']/df_adot2['attempts']

```

```

[ ]: df_adot2.loc[(df_adot2['attempts']>=100)&(df_adot2['year'] == 2021)].
↪sort_values(by='adot', ascending=False).head(25)

```

```
[ ]: receiver_player_name year air_yards attempts adot
103 C.Sutton 2021 1509.0 100 15.090000
488 T.Lockett 2021 1562.0 108 14.462963
508 V.Jefferson 2021 1469.0 106 13.858491
170 D.Smith 2021 1593.0 118 13.500000
491 T.McLaurin 2021 1697.0 130 13.053846
147 D.Metcalf 2021 1636.0 129 12.682171
351 M.Evans 2021 1754.0 140 12.528571
399 O.Beckham 2021 1358.0 109 12.458716
3 A.Brown 2021 2131.0 176 12.107955
232 J.Chase 2021 1952.0 163 11.975460
363 M.Jones 2021 1422.0 120 11.850000
254 J.Jefferson 2021 2072.0 175 11.840000
474 T.Higgins 2021 1658.0 141 11.758865
85 C.Kirk 2021 1326.0 114 11.631579
68 C.Claypool 2021 1292.0 112 11.535714
447 S.Diggs 2021 1931.0 175 11.034286
151 D.Mooney 2021 1552.0 141 11.007092
318 K.Pitts 2021 1188.0 110 10.800000
5 A.Cooper 2021 1231.0 114 10.798246
382 M.Williams 2021 1587.0 147 10.795918
346 M.Brown 2021 1620.0 152 10.657895
181 Dj.Moore 2021 1730.0 164 10.548780
46 B.Cooks 2021 1410.0 134 10.522388
419 R.Anderson 2021 1156.0 110 10.509091
430 R.Gronkowski 2021 1086.0 106 10.245283
```

```
[ ]: df_adot2
```

```
[ ]: receiver_player_name year air_yards attempts adot
0 A.Abdullah 2021 43.0 53 0.811321
1 A.Armah 2021 0.0 1 0.000000
2 A.Auclair 2021 22.0 5 4.400000
3 A.Brown 2021 2131.0 176 12.107955
4 A.Collins 2021 2.0 10 0.200000
.. ... ..
516 Z.Gentry 2021 89.0 29 3.068966
517 Z.Jones 2021 1068.0 78 13.692308
518 Z.Moss 2021 26.0 32 0.812500
519 Z.Pascal 2021 608.0 69 8.811594
520 Z.Wilson 2021 8.0 1 8.000000
```

```
[521 rows x 5 columns]
```

```
[ ]: from subprocess import call
call(['python', '-m', 'nbconvert', 'data_pull.ipynb'])
```

[]: 1

[]: