Analysis Plan

Data sources

* <https://www.advancedsportsanalytics.com/nfl-raw-data>
* 2019-2022
* Data from all positions on 3 separate CSVs
  + Pass/Rush/Receive
  + D/ST
  + Kicking

Questions to answer

* Can we determine the most consistent players by points from the 2022 year?
  + Are there any consistent players?
  + Does age correlate with points?
* Who are the streakiest players?
  + Are they worth picking up?

League rules

* Points
  + Quarterbacks
    - Passing touchdowns: 4 points
    - Passing yards: 1 point/25 yards
    - Interceptions: -1 point
    - 2pt conversion: 2 points
  + Rushing/receiving
    - Rushing/receiving yards: 1 point/10 yards
    - Rushing/receiving touchdowns: 6 points
    - Receptions: 1 point
  + Defense
    - Return touchdown: 6 points
    - Sack: 1 point
    - Interception: 2 points
    - Fumble recovery: 2 points
    - Safety: 2 points
    - Blocked kick: 2 points
    - 0 Points allowed: 10 points
    - 1-6 points allowed: 7 points
    - 7-13 points allowed: 4 points
    - 14-20 points allowed: 1 points
    - 21-27 points allowed: 0 points
    - 28-34points allowed: -1 points
    - +35 points allowed: -4 points
    - Extra point returned: 2 points
  + Kickers
    - 0-19 yd field goal: 3 points
    - 20-29 yd field goal: 3 points
    - 30-39 yd field goal: 3 points
    - 40-49 yd field goal: 4 points
    - 50+ yd field goal: 5 points
    - Point after attempt made: 1 point
  + All
    - Fumble: -2 points

Season dates

* 2019-09-05 to 2019--12-29
* 2020-09-10 to 2021-01-03
* 2021-09-09 to 2022-01-08

Cleaning the data

* Remove duplicate data
* Update opponents score to scored\_against
  + Subtract other teams defensive scores
* **Combine datasets?**
  + Separate offense from defense because offense is individual and defense isnt.
* Convert to the correct data type?
  + Probably stick to CAST, but maybe for dates
* Add column for season name based off of dates above