

Task

For each 2018 free agent, you will predict the **Average Per Year (APY)** of their 2018 contract. We define APY as:

$$\text{APY} = \text{Total Contract Value} / \# \text{ of Years}$$

Where “Total Contract Value” includes guaranteed salary as well as all bonuses, and “# of Years” is the length of the contract.

For the first round on Kaggle, you will predict the value of 2017 free agent contracts. For a concrete example of how to set up your prediction task, I recommend you take a look at our [starter code on github](#).

Submission Instructions

Create a submission file in the same format as sample_submission.csv. Create an account on Kaggle.com, if you don't have one already, and join the first round of the competition via [this link](#). Follow instructions on kaggle to submit your submission file.

File descriptions

- NFL Contract Data_v1.xlsx - OTC contract data 2012-2017
 - Contracts sheet - free agent contracts
 - Players sheet - player records
 - Salary Cap sheet - salary cap records
- PFF 0-100 Grades NFLSeason2006to2015-1.xlsx - Pro Football Focus nfl stats 2006-2015
- train.csv - joined contracts table 2006-2015
- val.csv - joined contracts table 2016
- test.csv - joined contracts table 2017
- sample_submission.csv - a sample submission file in the correct format for Kaggle

Data fields - OTC: Contracts

The contracts table contains all the current active contracts in the NFL and the data relevant to the contracts. The data represents the components of the contract when originally signed, so things like bonuses for cap relief aren't included in the fields.

- **Name**- player name
- **Player_id**- players unique identifier
- **Team**- team's name
- **Team_id**- team's id code
- **Position**- current position that the player plays.
- **Role**- Whether the position is an offense, defense, or special teams position
- **Years**- the length (in terms of new years) of the contract
- **Date_signed**- this is something new we are doing so it's not complete.
- **Year_signed**- the year in which the player signed his new contract
- **Start_year**- the first "new year" of a contract. For extensions this will be different than the year signed
- **Total**- the total new money value
- **APY**- the contracts annual value
- **Guarantee**- the amount fully guaranteed at signing
- **GPY**- the full guarantee per year of the contract
- **GP**- the % of the contract fully guaranteed.
- **Total_guarantees**- the total amount of guarantees in a contract (this is the amount fully guaranteed and injury guaranteed)
- **Vested_guarantees**- how much of the injury guarantees become fully guaranteed at some date
- **Signing_bonus**- the bonus money paid as a signing bonus
- **Option_bonus**- any option bonus or secondary signing bonuses negotiated in the original contract
- **Roster_bonus**- The total amount of roster and reporting bonus money in a contract
- **Pergame_bonus**- The total amount of per game bonuses in a contract
- **Workout_bonus**- The total amount of workout money in a contract. I also generally include weight bonuses as workouts
- **Incentives**- The amount of known incentives in a contract (this is money that can be earned in excess of the total new money value via some performance threshold)
- **FA**- the year the contract expires and the type of free agency. If it has a R after the year it means restricted, E means exclusive, O means an option year, V means voidable
- **Contract_type**- the type of contract signed.
 - *UFA* means his last contract expired and the player had enough accrued years to be unrestricted free agent. I consider any player signing as a UFA once a regular season ends, even if re-signing with his old team. The NFL may consider that technically an extension.
 - *RFA* means the player was a restricted free agent at the time of signing
 - *ERFA* means the player was an exclusive rights player at the time of signing
 - *Extension* means a player signed a new contract with years still remaining on his prior contract
 - *Drafted* means a player is on a rookie draft pick contract
 - *UDFA* means a player is on an undrafted rookie contract

- *SFA* means a player is a street free agent and had been released by his prior team
- *Franchise* means any player who signed a contract as a franchise player
- *Transition* means any player who signed a contract as a transition player
- *Other* means a player who doesn't fit into any of the above categories. Generally "other" means a player who renegotiates his contract and takes a pay cut in the process

Data fields - OTC: Players

The players table has a brief bio of all the players we currently have in our databases.

- **Id** – player's unique identification code
- **Name**- players name
- **Accrued**- this helps figure out how many years of 6+ games a player has but its not complete so Id just delete it or bypass it
- **Height, weight, DOB, college**- all self explanatory
- **Original team**- this is the team number of the original draft team
- **Draft year**- the year in which the player was first eligible for the draft
- **Draft round and draft pick**- the round and number of their draft pick if drafted (will read 0 for a UDFA)
- **Notes**- the generic contract notes you see on the site. This probably serves no purpose for you and could be deleted
- **NFL player id**- If you have any stats that can be pulled from nfl.com or any other sources that use their data services this is the id that links with the nfl.com id.

Data fields - OTC: Salary Cap

The cap table contains all the salary cap and cash breakdowns of each contract.

- **Name**- players name
- **Player_id**- players unique identifier
- **Team, team id**- the team name and their id number
- **Position**- the players position
- **Role**- offense, defense, or special teamer
- **Year**- the year of the contract
- **BS**- the base salary of the contract. This is what the player earns in the regular season
- **PB**- the sum of all prorated (signing, option, etc...) bonuses that impact that cap in a given year

- **RB**- the sum of all roster (including per game) and reporting bonuses.
- **Wo**- the workout money promised to a player in a year
- **OB**- LTBE incentives which count on the cap
- **Cp**- cap charge for a player. Generally it's a sum of the prior 5 columns, but there are adjustments for per game bonuses, suspensions, etc...
- **DM**- the amount it would cost to release a player prior to June 1 in any season
- **CS**- the amount of cap room created or lost when cutting a player
- **GBS**- the amount of P5 or other salary that is not a signing bonus that is guaranteed in that contract year
- **Notes**- just some notes on the contract
- **CshBns**- This is the amount of money paid as a signing bonus, option bonus, or other prorated bonus to a player.
- **Cshpay**- This is the amount of money the player earns in a year. Generally it is the sum of the BS, RB, WO, and Cshbns data
- **Cshtocp**- This is the cash to cap charge ratio for the player. I don't fill this out for everyone anymore so no need to even pay attention to it and it's easy enough to just calculate on the fly
- **Tb**- This is a ratio I use to help identify the time a player loses dead money protection. Like the cshtocp I don't always update it so no need to really pay attention. In general a number over 1 means a team would likely cut if the players performance slips and under 1 makes it more likely a player stays.
- **Initial_cash**- This is the amount promised to the player when the contract is signed. So it assumes all per gamers are earned, players are not suspended, etc...When evaluating contract structure and cash flows this is the best data to really use.
- **Old_money**- money existing in the prior contract at the time an extension is signed. To get the new money in a contract you subtract this amount out of the initial_cash
- **Cash_dead**- You can just disregard it.
- **Status**- A player's roster status at the end of a given year. Active, IR, PUP, etc...
- **Contract_type**- same stuff as the in the contracts database.

Data fields - PFF

This dataset is provided by Pro Football Focus, which develops a comprehensive hand-grading system to rank each player. It contains data on how well a player has performed each season in the NFL.

- **Season** - The year for which the grades were accumulated
- **Player_id** - PFF player id. Not the same as OTC player id.
- **Player** - Player Name
- **Pos** - Position played for that season
- **Overall** - PFF grade
- **Overall_rk** - PFF rank

- Pass - Grade for ability to pass the ball on Pass Plays
- Run - Grade for ability to run the ball on Running Plays
- Receiving - Grade for ability to catch the ball and run routes on Pass Plays
- Passblock - Grade for ability to block on Pass Plays
- Runblock - Grade for ability to block on Run Plays
- Passrush - Grade for ability to rush the Quarterback on Pass Plays
- Rundefense - Grade for ability to defend on Run Plays
- Coverage -
- Discipline -
- Pass_snaps - Total number of plays in which the player was asked to pass the ball.
- Receiving_snaps - Total number of plays in which the player was asked to be a receiver on a pass play
- Passblock_snaps - Total number of plays in which the player was asked to block on a pass play
- Run_snaps - Total number of plays in which the player was asked to run the ball on a run play
- Runblock_snaps - Total number of plays in which the player was asked to block on a run play
- Rundefense_snaps - Total number of plays in which the player defended the run
- Passrush_snaps - Total number of plays in which the player rushed the passer
- Coverage_snaps - Total number of plays in which the player covered receivers
- Total_snaps - Total number of plays played
- Games - Total number of games played