This dataset contains [Next Gen Stats](https://nextgenstats.nfl.com/) tracking data for running plays. You must use features known at the time when the ball is handed off (TimeHandoff) to forecast the yardage gained on that play (PlayId).

Because this is a time-series code competition that will be evaluated on future data, you will receive data and make predictions with a time-series API. This API provides plays in the time order in which they occurred in a game. Refer to the starter notebook [here](https://www.kaggle.com/dster/nfl-big-data-bowl-official-starter-notebook) for an example of how to complete a submission.

Note: Before the evaluation period begins, we will be updating the train.csv file to include current season games. Before Stage 2 begins, Kaggle will update the train.csv file to include current-season games through Stage 1. Please take note should you want to retraining to be a part of your model submission.

Columns

Each row in the file corresponds to a single player's involvement in a single play. The dataset was intentionally joined (i.e. denormalized) to make the API simple. All the columns are contained in one large dataframe which is grouped and provided by PlayId.

* GameId - a unique game identifier
* PlayId - a unique play identifier
* Team - home or away
* X - player position along the long axis of the field. See figure below.
* Y - player position along the short axis of the field. See figure below.
* S - speed in yards/second
* A - acceleration in yards/second^2
* Dis - distance traveled from prior time point, in yards
* Orientation - orientation of player (deg)
* Dir - angle of player motion (deg)
* NflId - a unique identifier of the player
* DisplayName - player's name
* JerseyNumber - jersey number
* Season - year of the season
* YardLine - the yard line of the line of scrimmage
* Quarter - game quarter (1-5, 5 == overtime)
* GameClock - time on the game clock
* PossessionTeam - team with possession
* Down - the down (1-4)
* Distance - yards needed for a first down
* FieldPosition - which side of the field the play is happening on
* HomeScoreBeforePlay - home team score before play started
* VisitorScoreBeforePlay - visitor team score before play started
* NflIdRusher - the NflId of the rushing player
* OffenseFormation - offense formation
* OffensePersonnel - offensive team positional grouping
* DefendersInTheBox - number of defenders lined up near the line of scrimmage, spanning the width of the offensive line
* DefensePersonnel - defensive team positional grouping
* PlayDirection - direction the play is headed
* TimeHandoff - UTC time of the handoff
* TimeSnap - UTC time of the snap
* Yards - the yardage gained on the play (you are predicting this)
* PlayerHeight - player height (ft-in)
* PlayerWeight - player weight (lbs)
* PlayerBirthDate - birth date (mm/dd/yyyy)
* PlayerCollegeName - where the player attended college
* HomeTeamAbbr - home team abbreviation
* VisitorTeamAbbr - visitor team abbreviation
* Week - week into the season
* Stadium - stadium where the game is being played
* Location - city where the game is being player
* StadiumType - description of the stadium environment
* Turf - description of the field surface
* GameWeather - description of the game weather
* Temperature - temperature (deg F)
* Humidity - humidity
* WindSpeed - wind speed in miles/hour
* WindDirection - wind direction

