

# 2018 PGA Tour Golf Data

Manda Bucklin, Nate Wagner

## Statistics For Data Science II

1. The data comes from [Kaggle](#).

2. Size of Data (excluding players who didn't play in 2018):

- 20 variables, 193 observations

### Player Name

- Player Name

### Numeric Variables:

#### Rounds

- Number of rounds played in that year (4 per tournament)

#### Fairway Percentage

- The percentage of time a tee shot comes to rest on the fairway

#### Year

- Stats from the calendar gold year 2018

#### Avg Distance

- The average number of yards per measured drive

#### Gir

- The percent of time a player was able to hit the green in regulation (greens hit in regulation/holes played). Note: A green is considered hit in regulation if any portion of the ball is touching the putting surface after the GIR stroke has been taken.

#### Average Putts

- The average number of putts per green in regulation

#### Average Scrambling

- The percent of time a player misses the green in regulation, but still makes par or better.

#### Average Score

- The weighted scoring average which takes the stroke average of the field into account. It is computed by adding a player's total strokes to an adjustment and dividing by the total rounds played. The adjustment is computed by determining the stroke average of the field for each round played. This average is subtracted from par to create an adjustment for each round. A player accumulates these adjustments for each round played.

**Points**

- The cumulative points for the year that the player has earned in the regular season of the FedExCup points race.

**Wins**

- Winning a PGA Tour event.

**Top 10**

- The total number of times the player has finished in the top 10 at official events for the year.

**Average SG Putts**

- The sum of the values for all holes played in a round minus the field average strokes gained/lost for the round is the player's Strokes gained/lost for that round. The sum of strokes gained for each round are divided by total rounds played.

**Average SG Total**

- The per round average of the number of strokes the player was better or worse than the field average on the same course & event.

**SG:OTT**

- The per round average of the number of strokes the player was better or worse than the field average on the same course & event minus the Players Strokes Gained putting value.

**SG:APR**

- The sum of the values for all holes played in a round minus the field average strokes gained/lost for the round is the player's Strokes gained/lost for that round. The sum of strokes gained for each round are divided by total rounds played.

**SG:ARG**

- The number of Around the Green strokes a player takes from specific locations and distances are measured against a statistical baseline to determine the player's strokes gained or lost on a hole.

**Money**

- The total official money a player has earned year-to-date.

**Categorical Variables****Driving Distance**

- 3 Levels
  - Long: a drive greater than 302.0 yards
  - Average: a drive 290.0-302
  - Short: a drive less than 290

**Won That Year Dummy**

- Dummy variable indicating whether a player won that year
  - 1: Player won a PGA Tour event
  - 0: Player did not win a PGA Tour event