



# Predicting the Winner of The Masters

---

Kelsey Schroeder and  
Roberto Arguello



# Why The Masters?

---

- The Masters is the highest profile event in the golf world
- The tradition of The Masters is unparalleled by any other golf tournament
- This winner receives not only a payout in excess of one million dollars, but also the green jacket



# Methodology

---

- We used a lasso regression to predict which golfer will win The Masters
- Compared a golfer's statistics with final rank in tournament
- Used data from 2010-2016, predicting following year's results
- Trained on 2010-2014, tested on 2015 (predicting 2016)
- Omitted golfers who did not participate in The Masters or have adequate PGA tour data

# Data

---

We collected our data on the various statistics via <http://www.pgatour.com/stats.html>

Plan to make our data set available online: not much golf data out there in a useful format



# Variables

---

Rank

Age

Yards/Drive

Driving Accuracy

Greens in Regulation

Putting Average

Birdie Conversion

Bounce Back

Club Head Speed: Average, Fastest, Slowest

Going for the Green

GIR Rank

Greens Hit

Hit Fairway Percentage

One Putt Percentage

Par 3, 4, 5 Scoring Average

Proximity to Hole

Rough Tendency

Sand Save Percentage

Scoring Average

Scrambling

Strokes Gained Putting

Strokes Gained Tee to Green

Total Putting

Total Putts Gained

# Variables That Don't Matter

---



- Rank
- Age
- Putting Average
- Total Putting
- Scoring Average on Par 3s and 4s
- GIR Rank
- Total Putts Gained
- Greens in Regulation

# Variables That Do Matter

---

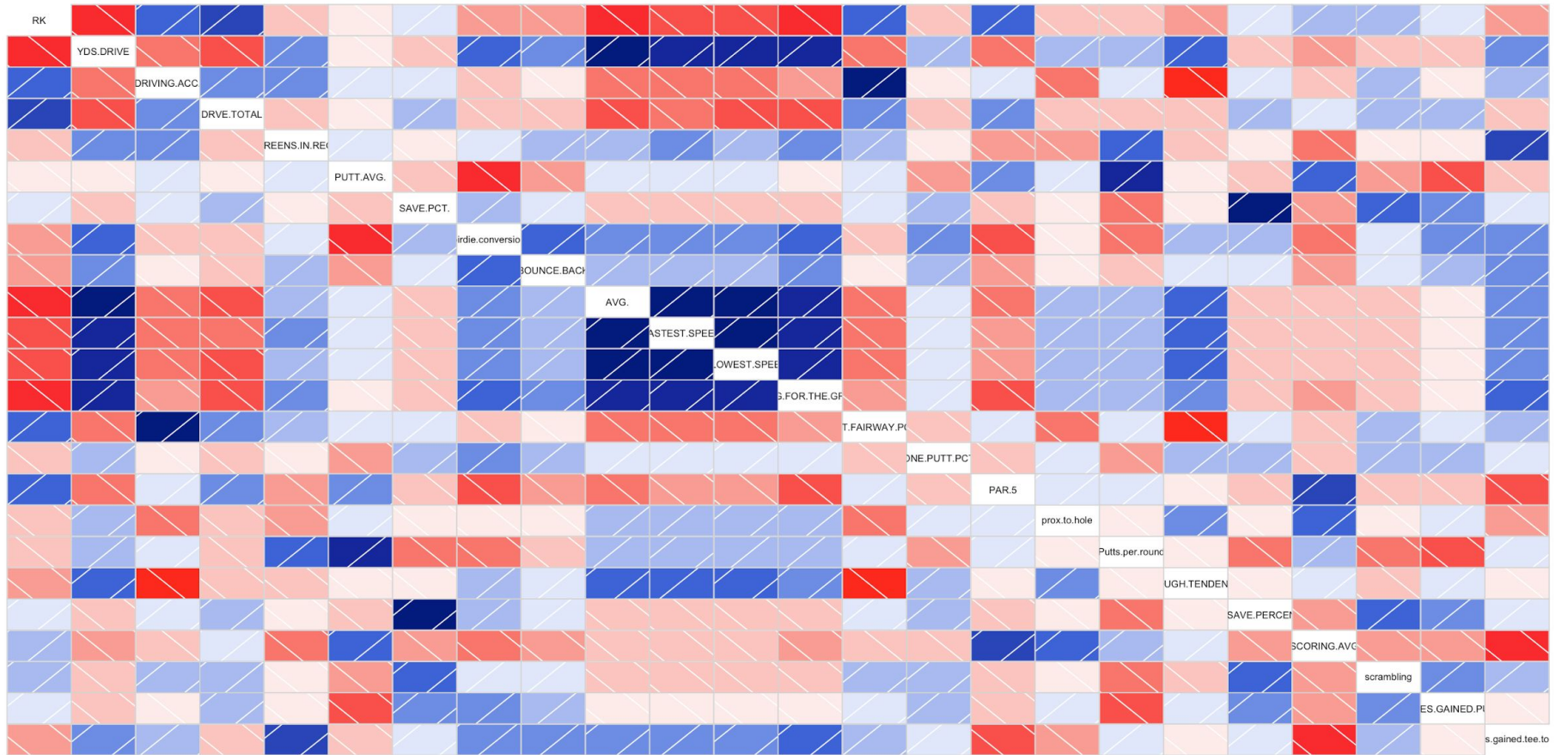
- Strokes Gained Putting
- Strokes Gained Tee to Green
- Proximity to Hole
- Birdie Conversion
- Bounce Back
- Putts Per Round
- Scoring Average
- Hit Fairway Percentage



How do our variables relate  
to each other?

---





# Regression Predictions for the 2016 Masters

---

## Our Top Ten Predictions:

1. Bubba Watson
2. Justin Rose
3. Jason Day
4. Henrik Stenson
5. Brooks Koepka
6. Adam Scott
7. Dustin Johnson
8. Sergio Garcia
9. Hideki Matsuyama
10. Justin Thomas

# Predictions vs Results: 2016 Masters

---

## Our Top Ten Predictions vs Their Results:

|    |                |     |     |                  |     |
|----|----------------|-----|-----|------------------|-----|
| 1. | Bubba Watson   | T37 | 6.  | Adam Scott       | T21 |
| 2. | Justin Rose    | T10 | 7.  | Dustin Johnson   | T42 |
| 3. | Jason Day      | T10 | 8.  | Sergio Garcia    | T34 |
| 4. | Henrik Stenson | T24 | 9.  | Hideki Matsuyama | T7  |
| 5. | Brooks Koepka  | T21 | 10. | Justin Thomas    | T39 |

# Surprises

---



The winner of this year's Masters, Danny Willett, was not included in our data set because he did not accumulate enough tournament rounds in 2015 despite his 11th World Golf Ranking at the time of his victory.

# Surprises

---

- All 10 members of our top 10 predicted managed to make the cut. Furthermore, all 10 finished in the top 42.



# Surprises

---

- Jordan Spieth was not predicted to finish in our top ten despite winning The Masters last year (he finished T2 this year).



# More Predictions: 2017 Masters

---

1. Rory McIlroy
2. Adam Scott
3. Justin Rose
4. Jason Day
5. Dustin Johnson



6. Branden Grace
7. Jordan Spieth
8. Brooks Koepka
9. Phil Mickelson
10. Matt Kuchar



# Surprises

---

- Jason Day was not our prediction to win The Masters in 2017 based off of this year's small sample size.





# Limitations of Golf Predictions

---

- Close leaderboards
- Small differences between tour players' statistics
- Course variations, weather, etc.
- Skill variability year-to-year
- Unpredictability of golf tournaments



# How To Improve Our Predictions

---

- Full year's worth of data (from one Masters to the next Masters)
- Remove highly correlated variables
- Add variables (weather, round, etc)
- More specific variables (e.g. left rough percentage)
- Stroke by stroke or hole by hole (not currently available)
- Other tournament results (Euro Tour)

# Questions?

