Predicting Annual Earnings on the PGA



Tour

Rob Zhang

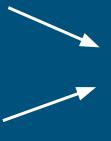








Newsday





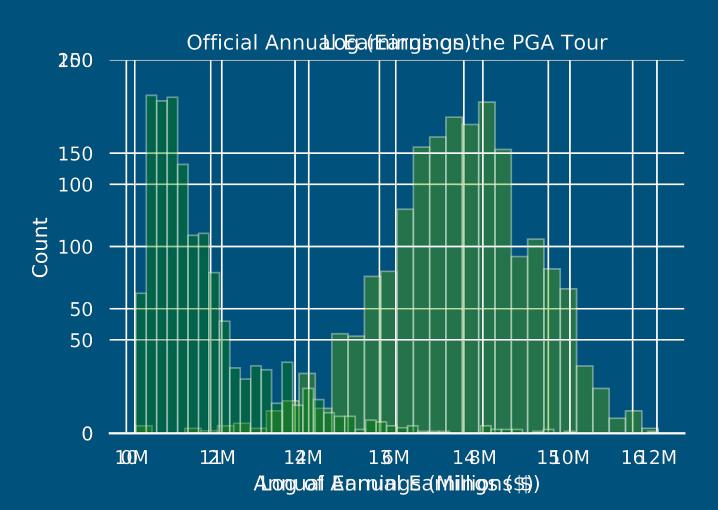








Cleaned Data



Traditional

- Fairway Percentage
- Average Drive Distance
- Greens in Regulation
- Average Scrambling
- Average Putts

Strokes Gained

- SG Off-the-Tee
- SG Approach-the-Green
- SG Around-the-Green
- SG Putting

Model Results

 R^2 : 0.765

rmse: ~ \$740k



	Model	Actual	Difference
2018	2.1	2.7	-0.6
2017	8.5	9.4	-0.9
2016	4.5	5.5	-1.0
2015	11.1	12.0	-0.9
2014	4.1	4.3	-0.2
2013	4.8	3.9	+1.0 (\$ Millions)

Which Stats Actually Matter?



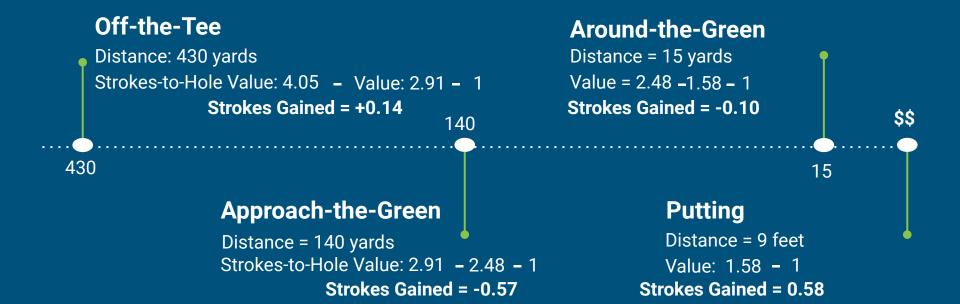
Traditional

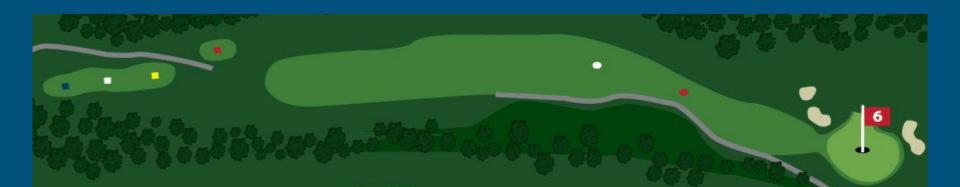
- Fairway Percentage
- Average Drive Distance
- Greens in Regulation
- Average Scrambling
- Average Putts
- Weight, Height, Age

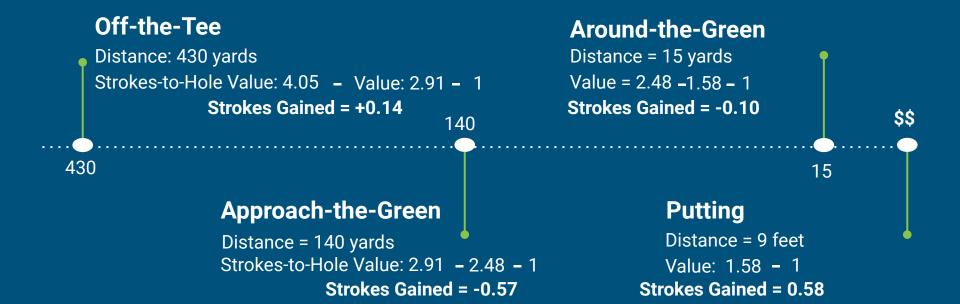
Strokes Gained

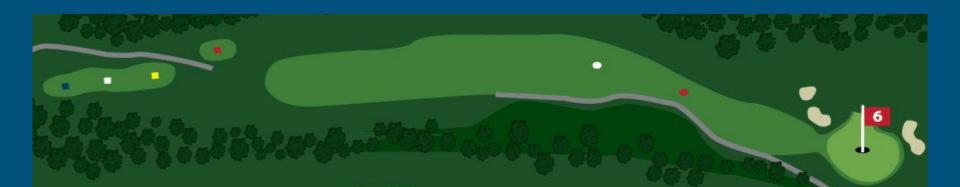
- SG Off-the-Tee
- SG Approach-the-Green
- SG Around-the-Green
- SG Putting

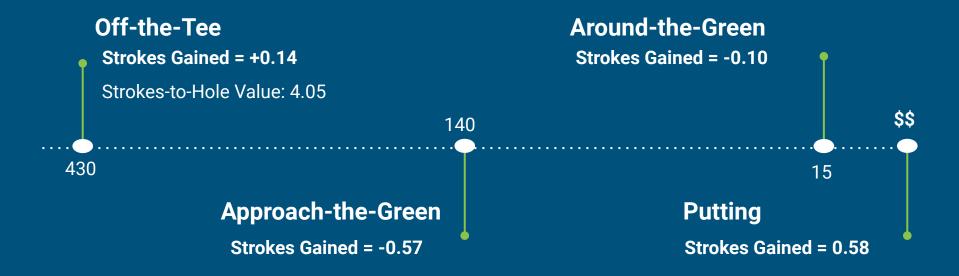
Appendix











Strokes Gained Total = 0.14 - 0.57 - 0.10 + 0.58 = 0.05 = 4.05 - 4



Questions?

