PGA Project

Tyler Gorecki

7/5/2022

What is the most important component of success on the PGA Tour?

Background and Data Description

Golf is an incredibly difficult sport both mentally and physically. With the analytics boom impacting all sports, golf has followed suit with new approaches to golf ball/ club technology and in playing strategy. I collected the PGA Tour data by transferring the statistics from both the ESPN (1) and PGA Tour (2) websites into an excel spreadsheet. This data was then converted into csv files and merged in R by the name of each golfer in the dataset. The ESPN data included traditional statistics for the golfers - statistics like final ranking, number of rounds played, number of wins, average driving distance, and 13 others. The PGA Tour data contained more advanced and modern analytics such as club head speed, ball speed, and strokes gained. Strokes gained is a modern way of comparing golfers directly with each other (3). For every round in each tournament, the net number of strokes gained by all players that day will be zero. If a golfer played better than the rest of the field, their strokes gained will be positive and if they didn't play well compared to the field, it will be negative. My goal is to determine what the best approach is to perform well and win on the PGA Tour.

More Specifically. . .

Does driving distance or fairway accuracy matter more? Is greens in regulation or putting the most important? Some players main focus is to make sure they always try to hit the fairway to give them a better lie to hit the green on their following shots. Others attempt to hit the ball as far as they can to give themselves a shorter distance to the green. I want to research which strategy is better correlated with success among PGA Tour golfers on the PGA Tour. I also want to analyze how much impact greens in regulation (the percentage of shots hitting the green in no more than two shots under par - ex: by the second shot on a par four) and putting have in the difference between the top golfers and the rest.

Potential Issues

When comparing the data between both sites, I realized that there was a slight difference in the number of rounds for each golfer in their 2020-21 seasons. The number of rounds was slightly higher from the PGA Tour website (2) and I believe this is due to the fact that they included their postseason tournament in their final statistics, while ESPN (1) did not. The numbers aren't significantly different, but because of this, I will not be comparing any 'total' statistics. Regardless, there is enough data for each golfer that the 'average' statistics can be trusted for appropriate analysis for this dataset.

Scoring

Typical scores for golfers with vs without win(s)
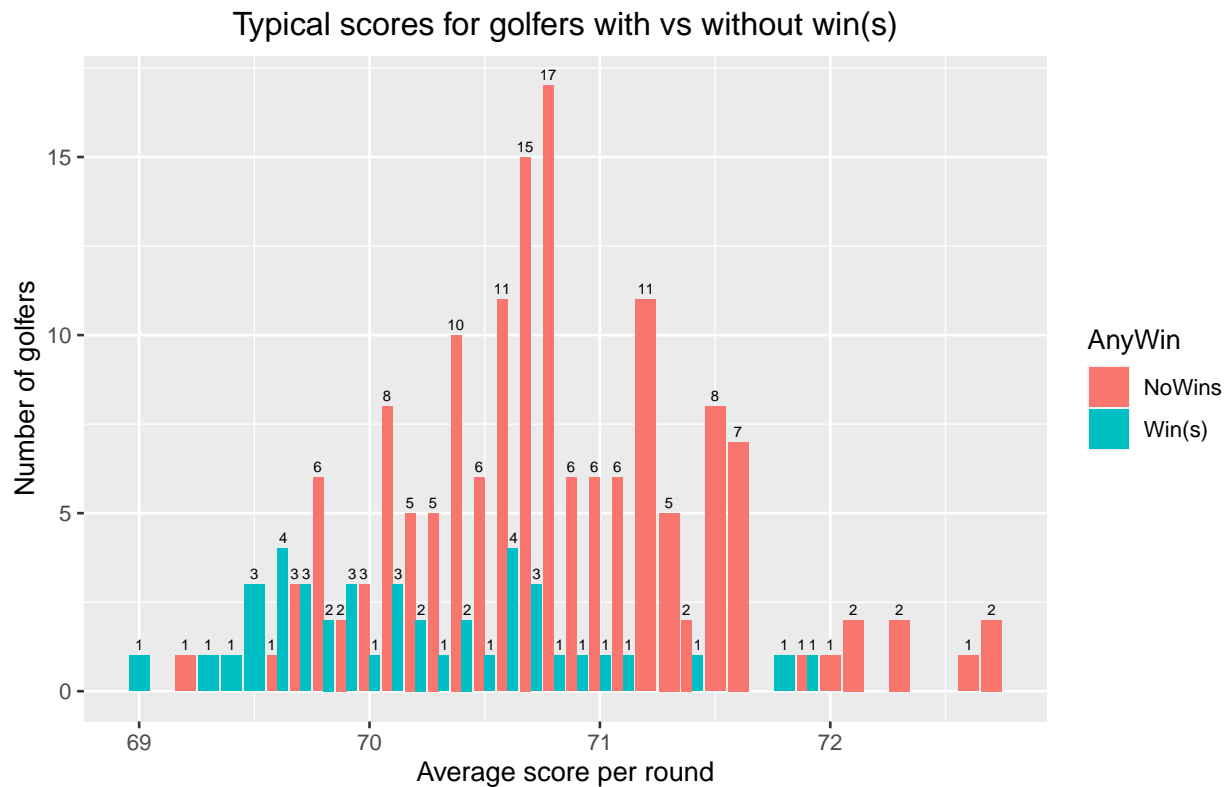
This graph shows what the typical scores are of golfers that won a tournament this season and that of the golfers who didn't. It is expected, but from this numerical summary the golfers who had at least one win on the PGA Tour during the 2020-21 season tended to average a lower average score over every round they played during the season. From this visualization, I know that 2/3, or 28/42 of the winners in this dataset averaged a score of 70.5 or less per round over the season. Looking further at this average score, about 36% of the golfers that averaged this low of a score won at least one tournament, while only about 12% who averaged above this threshold won a tournament. This shows that winning on tour typically isn't a fluke and the most consistent golfers win tournaments more times than not. This is a good baseline when first looking at this data to understand that there is a consistent trend that can be seen among PGA Tour winners.

Driving

Strokes gained vs. driving accuracy
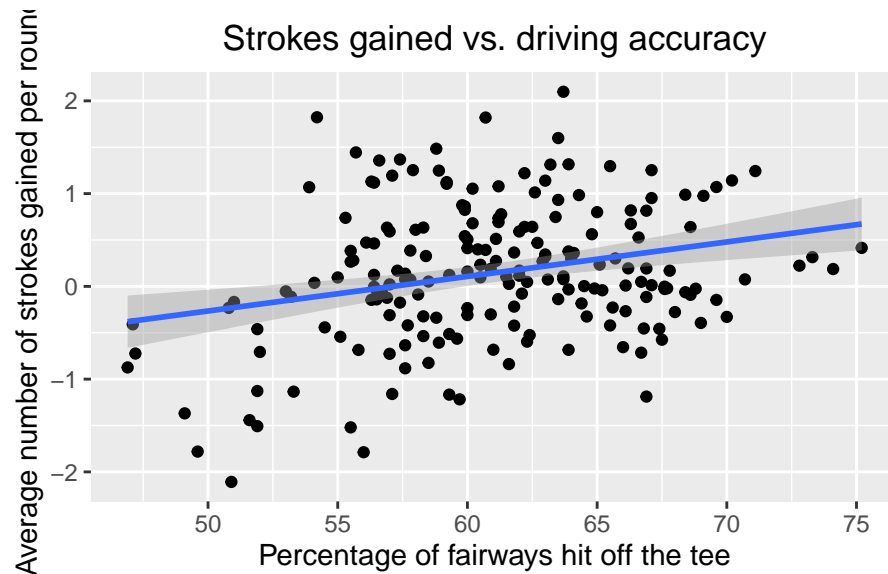
While there is a slight positive correlation in this relationship, this plot shows that driving accuracy isn't a true indicator of success on the PGA Tour unless the percentage of fairways hit is above 70% or below 55%, which is rare. For the data points within the 55-70% range, the average number of strokes gained per round can almost be considered random between -1 and 1.5 strokes gained.

Strokes gained vs. driving distance
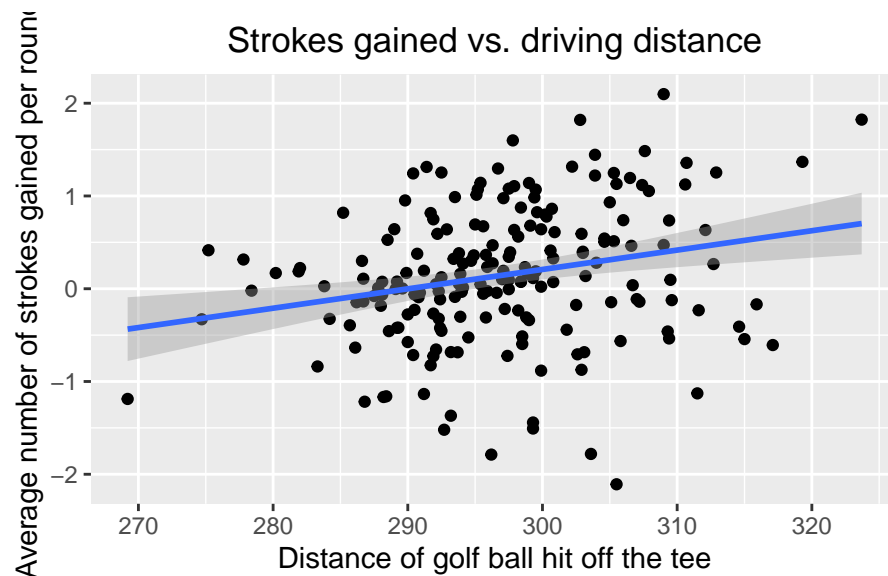
When looking at the relationship between strokes gained and driving distance, while there is a slightly higher positive correlation between the variables, a similar sense of randomness can be seen regardless of the average driving distance. From these plots, I feel comfortable saying that as long as the golfer drives the ball in play, they should just focus on giving themselves the best opportunity to hit the green on their next shot. For some, hitting out of the fairway is a big priority, but others believe that regardless of lie, it is always better to be as close to the green as possible for hitting an approach shot.

Greens in regulation and putting
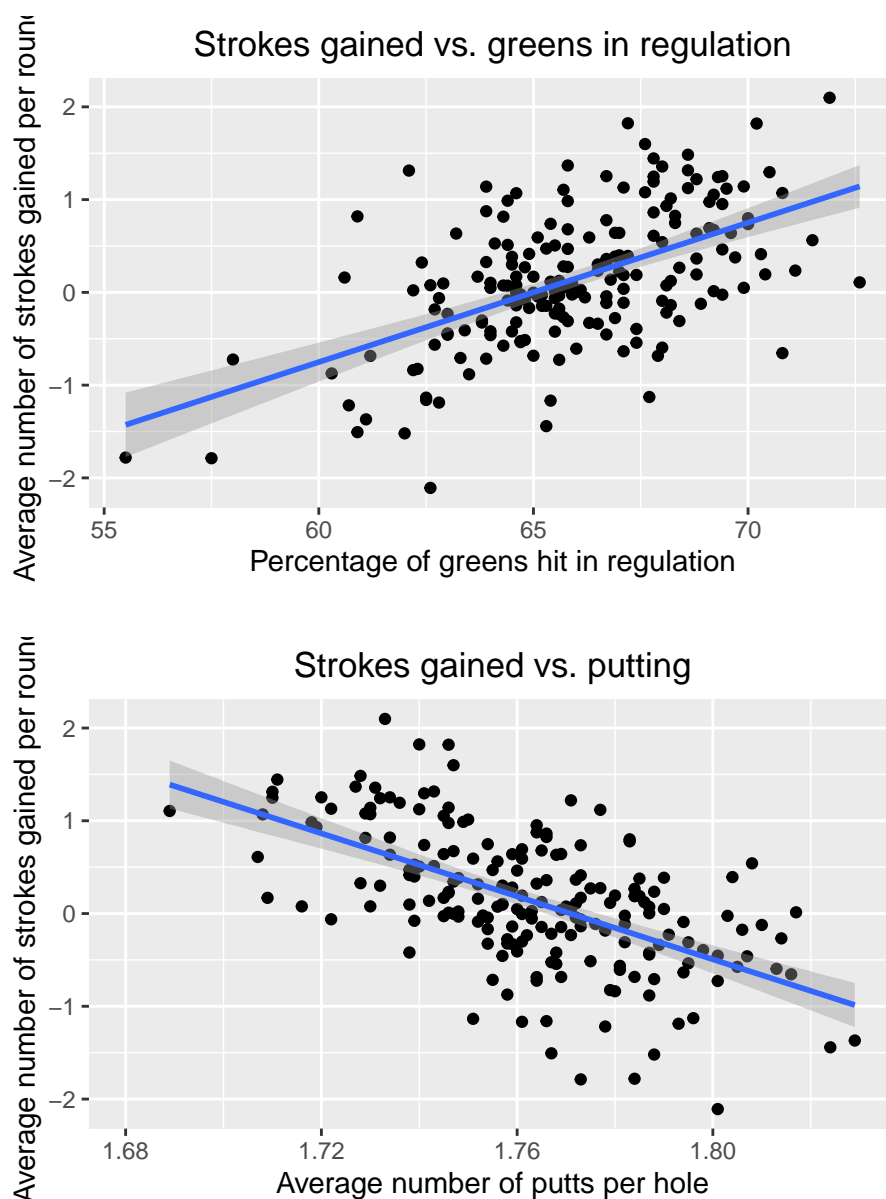
These two plots both have seemingly strong correlations showing relationships between strokes gained and greens in regulation percentage and strokes gained and average number of putts per hole.

For players with at least 67% of greens hit in regulation, there are still multiple who have an average number of strokes gained per round below -0.5, meaning they tend to score lower than the rest of the players on tour. For putting though, of the golfers averaging less than 1.75 putts per hole, there is only one that has an average strokes gained significantly less than 0. While we're talking about an average of .08 average putts per hole and about 1.44 per round, this could be the difference in a golfer winning or losing a tournament. I'm not saying that golfers aren't trying to set themselves up in the best situations for putting, but if there's a most important aspect to a golfer's game I could identify, it would be approach shots. Better approach shots leads to less putts and less putts leads to more success on the PGA Tour.

References

(1) https://www.espn.com/golf/stats/player/_/season/2021
(2) https://www.pgatour.com/stats.html
(3) https://www.pgatour.com/news/2016/05/31/strokes-gained-defined.html