Statistical Analysis of 2021 Enhanced Scorecard Field Study

Dave Pierce and Joe Siwinski, Green Section, August 22, 2023

1. Summary

A statistical analysis of data gathered in 2021 by volunteers recruited to record their shots while playing golf using an enhanced scorecard was conducted by the USGA. Limited demographic information about the golfers was available. Only gender, age and Handicap Index® were captured and used to examine golfer performance for putting, driving distance and greens in regulation.

2. Introduction

During the 2021 golf season, the USGA recruited volunteers to capture their shots, hole by hole, on an enhanced scorecard. Our goal was to supplement the golfer performance data gathered on the range with TrackMan with on-course data.

3. The Sample

5275 valid shots were captured during 76 rounds. Multiple rounds were recorded by some golfers, so the number of golfers is less than 76 and unknown since names were not requested. Also, not all rounds were a full 18 holes. Because of the limited sample size, this study should be used to evaluate the potential of the methodology for gathering this type of data and only for directional trends, even though specific numerical results will be provided for these data through the analysis.

Handicap Index was a continuous variable in the models and ranged from 0 to 43. The median Handicap Index was 9.5. Some records did not have a Handicap Index. Age of the golfers ranged from 14 to 80 years old and was a continuous variable. The median golfer age was 67. All the records had an age associated with them. 23.5% of the shots were by female golfers – roughly the percentage of U.S. female golfers. All records had a gender associated with them.

4. Methodology

Exhibit 1 shows the front and back of the data recording template or “enhanced scorecard”. Golfers participating were provided the following instructions:

This scorecard is intended to capture more detailed information than a normal scorecard. For each hole, 8 rows are provided to record the current distance from the hole, and the current lie. As each hole begins from the teeing area, you will see a pre-filled "T" for the lie. The corresponding distance is the length of the hole from whichever set of tees is being played. After teeing off please mark down the club used for your initial tee shot. The club only needs to be recorded for the initial tee shot.

After reaching the location of your tee shot, record the lie of the golf ball according to the legend provided. Estimate the remaining distance to the hole from marked yardages on the course, or with a personal rangefinder/GPS, and record the value next to the lie that was just marked down. An exact distance isn't necessary, a 10–20-yard estimate is sufficient.

Continue recording shot lies and distances in this manner until reaching the putting green. Once on the green please mark your distances in FEET from the hole. If you prefer to continue marking distances in yards, please indicate you are doing so on the scorecard (e.g. 5 YDS, instead of only the number 5).

BAD LIES If you find yourself in a lie that is especially difficult to advance such as a buried bunker lie, nestled down in deep rough, or your intended line of play is blocked by a tree or other obstacle, please CIRCLE the lie you mark down. The intent of circling your lie is to indicate that a particular lie significantly affected the ability of the shot to be advanced towards the flag.

PENALTY STROKES If your ball is lost or out of bounds, or comes to rest in a penalty area, please mark the next stroke's lie as either "OB" or "PA" corresponding to out of bounds or penalty. The distance entry for a shot with a lie of "OB" or "PA" may be left blank.

Exhibit 1: Enhanced Scorecard Template (front & back)

HOLE PAR																		
HOLE	1		2		3		4		5		6		7		8		9	
SHOT	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE
1		T		T		T		T		T		T		T		T		T
TEE CLUB																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
HOLE SCORE																		

HOLE PAR																		
HOLE	10		11		12		13		14		15		16		17		18	
SHOT	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE	DIST	LIE
1		T		T		T		T		T		T		T		T		T
TEE CLUB																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
HOLE SCORE																		

INSTRUCTIONS

This scorecard is intended to capture more detailed information than a normal scorecard. For each hole, 8 rows are provided to record the current distance from the hole, and the current lie. As each hole begins from the teeing area, you will see a pre-filled "T" for the lie. The corresponding distance is the length of the hole from whichever set of tees is being played. After teeing off please mark down the club used for your initial tee shot. The club only needs to be recorded for the initial tee shot.

After reaching the location of your tee shot record the lie of the golf ball according to the legend provided. Estimate the remaining distance to the hole from marked yardages on the course, or with a personal rangefinder/GPS, and record the value next to the lie that was just marked down. An exact distance isn't necessary, a 10-20 yard estimate is sufficient.

Continue recording shot lies and distances in this manner until reaching the putting green. Once on the green please mark your distances in FEET from the hole. If you prefer to continue marking distances in yards, please indicate you are doing so on the scorecard (e.g. 5 YDS, instead of only the number 5).

LIE LEGEND	
T	TEE
F	FAIRWAY/FRINGE
R	ROUGH
B	BUNKER
G	PUTTING GREEN
OB	OUT OF BOUNDS
PA	PENALTY AREA
U	UNPLAYABLE

BAD LIES

If you find yourself in a lie that is especially difficult to advance such as a buried bunker lie, nestled down in deep rough, or your intended line of play is blocked by a tree or other obstacle, please CIRCLE the lie you mark down. The intent of circling your lie is to indicate that a particular lie significantly affected the ability of the shot to be advanced towards the flag.

PENALTY STROKES

If your ball is lost or out of bounds, or comes to rest in a penalty area, please mark the next stroke's lie as either "OB" or "PA" corresponding to out of bounds or penalty. The distance entry for a shot with a lie of "OB" or "PA" may be left blank.

USGA®

ENHANCED SCORECARD

DATE
COURSE
HANDICAP Or avg. 18 hole score
GENDER
AGE

5. Results

5.1. Putting Model

A putt success rate logistic model was trained on the putting portion of the enhanced scorecard data. The parameters in this model were Gender, Age, Handicap Index, and Distance from the hole for each putt. As stated earlier, only gender is a discrete variable. The results of the analysis are shown in Table 1. Distance from the hole was the most statistically significant variable in the model and the most impactful. A negative coefficient for Success Odds Increase indicates that as distance from the hole increases, the odds of successfully making the putt decreases.

Table 1: Putt Success Model Results

Variable Name	Success Odds Increase	P-Value
Distance from Hole [yards]	-36.2%	< 0.001
Age [years]	0.05%	0.904
Gender (Male)	27.1%	0.11
Handicap Index	-0.89%	0.19

Figure 1 shows the model's predicted overall putt percentage as a function of distance from the hole for all golfers. Note that the solid circles do not represent individual data points – they are the model results at 1-yard increments and shown to help read specific success rates at 1-yard intervals. Not shown, but understood, is that putting success at 0 yards is 100%.

Figure 1: Predicted Putt Success Rate vs. Yards from Hole

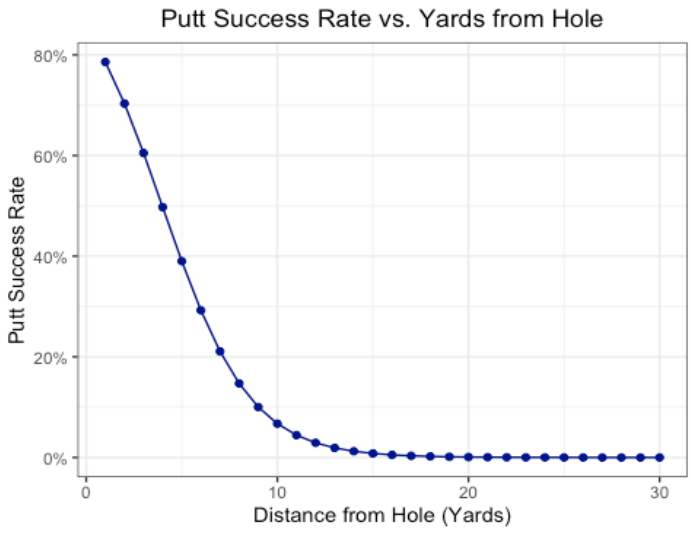


Figure 2 shows the output of the putting success model by gender at varying distances from the hole. The p-value of 0.11 indicates that it is 89% likely that there is a difference in putting success between genders. Male golfers are predicted to have a small increase in putting success rate at closer distances. There is no significant difference in gender putting success rates at distances over about 30 feet or 10 yards.

Figure 2: Predicted Putt Success Rate vs. Yards from Hole (By Gender)

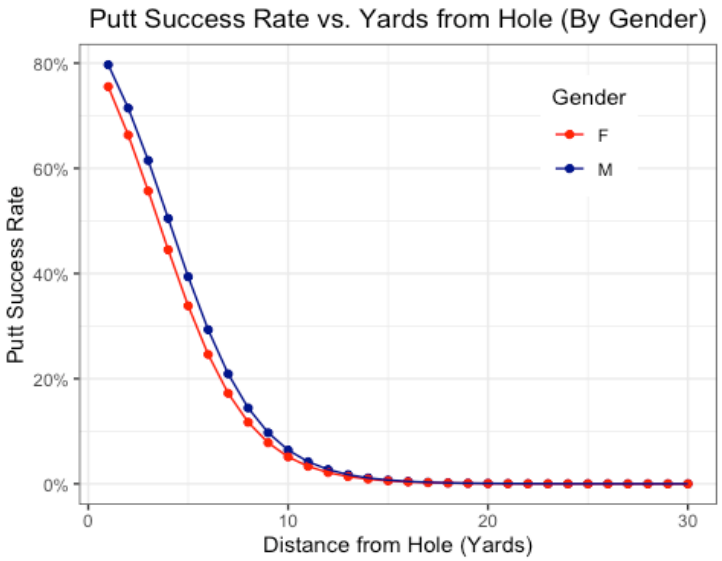
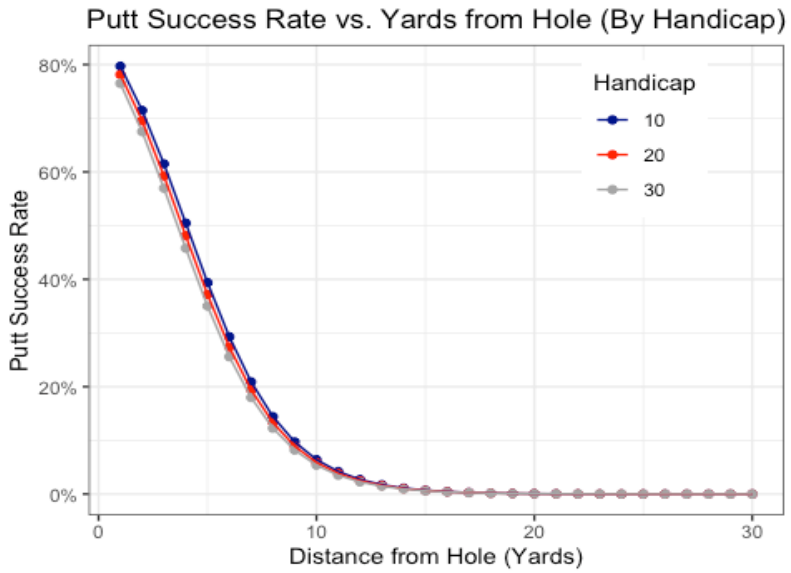


Figure 3 shows that putting success rate differs slightly by Handicap Index. The p-value of 0.19 indicates that it is 81% likely that there is a difference in putting success as Handicap Index changes. On average, golfers with lower Handicap Indices have higher putting success rates at closer distances than higher handicap golfers. At longer distances from the hole, all skill levels experience similar putting success rates. Again, the solid circles do not represent individual data points – they are the model results at 1-yard increments.

Figure 3: Predicted Putt Success Rate vs. Yards from Hole (By Handicap)



To summarize the putting results, the likelihood of making a putt is dominated by the distance from the hole and only marginally affected by skill level (Handicap Index) or gender. Interestingly, age was not a statistically significant indicator of putting performance! The p-value of 0.904 indicates that there is less than a 10% chance that age affects putting performance in these data.

5.2. Driving Distance Model

Drive distance was modeled using the enhanced scorecard data. Table 2 compares the model coefficient estimates for both the Enhanced Scorecard dataset and the TrackMan dataset (documented separately). All variables were statistically significant with a maximum parameter p-value of 0.10 in both models.

The intercept is higher in the enhanced scorecard model compared to the TrackMan model. On average, drives are longer from the enhanced scorecard data for younger, low Handicap Index, male golfers. For young male golfers with low Handicap Index, the models are quite close (313 vs. 310 yds). TrackMan data were collected on a driving range resulting in detailed launch conditions and estimated bounce and roll, whereas the Enhanced Scorecard data consisted of golfers playing actual holes.

Additionally, gender shows an impact on both model's drive distance estimates. However, the TrackMan model shows a bigger disparity between genders. In the TrackMan model, males drive 55 yards more than females on average as compared to 39 yards in the Enhanced Scorecard model.

Handicap Index has a larger impact for driver distance on the course. For every 1-point increase in Handicap Index, drive distance decreases by 2.9 yards on average compared to 1.95 yards on average in the TrackMan model.

Finally, age has similar effects in both models. For every one-year increase in age, total drive distance decreases by about 1 yard on average.

Table 2: Comparison of Drive Distance Models

	Model Type	
Variable	**TrackMan**	**Enhanced Scorecard**
Intercept	257.87	270.83
Gender (Male)	55.55	39.41
Handicap Index	-1.95	-2.91
Age	-1.01	-1.11

5.3. Greens in Regulation (GIR) Model

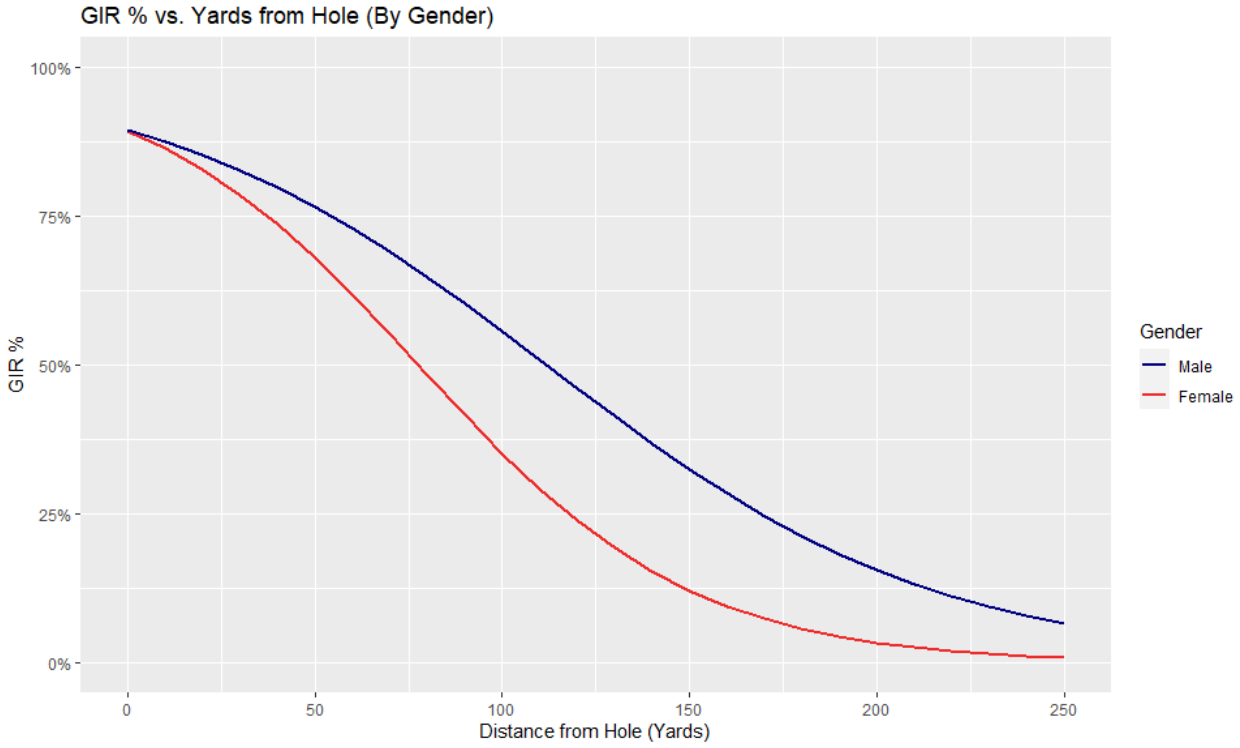
A GIR Success rate logistic regression model was also trained on the data. This model considers golfer demographics (Age, Handicap Index, and Gender) as well as distance from the hole and the lie. The results of this model are shown in Table 3. A tee shot lie contributed most toward GIR success compared to the other possible lies. The odds of a successful GIR essentially doubled when hit from the teeing ground. Shots off the fairway were also somewhat significant (p-value = 0.145) in the model and improved the chance of success in reaching the GIR by about 18%. Finally, Rough and Sand shots were not statistically significant in the model due to the low number of shots and had lower odds increases to GIR. The model also showed that as distance from the hole increases, the odds of GIR decreases and that younger golfers have an increased likelihood of hitting the GIR.

Table 3: GIR Success Model Output

Variable Name	Success Odds Increase	P-Value
Distance from Hole	0.399%	0.0197
Age	-1.19%	0.011
Gender (Male)	41.8%	0.087
Handicap Index	-10.5%	< 0.001
Lie = Tee	209.6%	0.042
Lie = Fairway	117.9%	0.145
Lie = Rough	27.8%	0.65
Lie = Sand	-99.9%	0.976

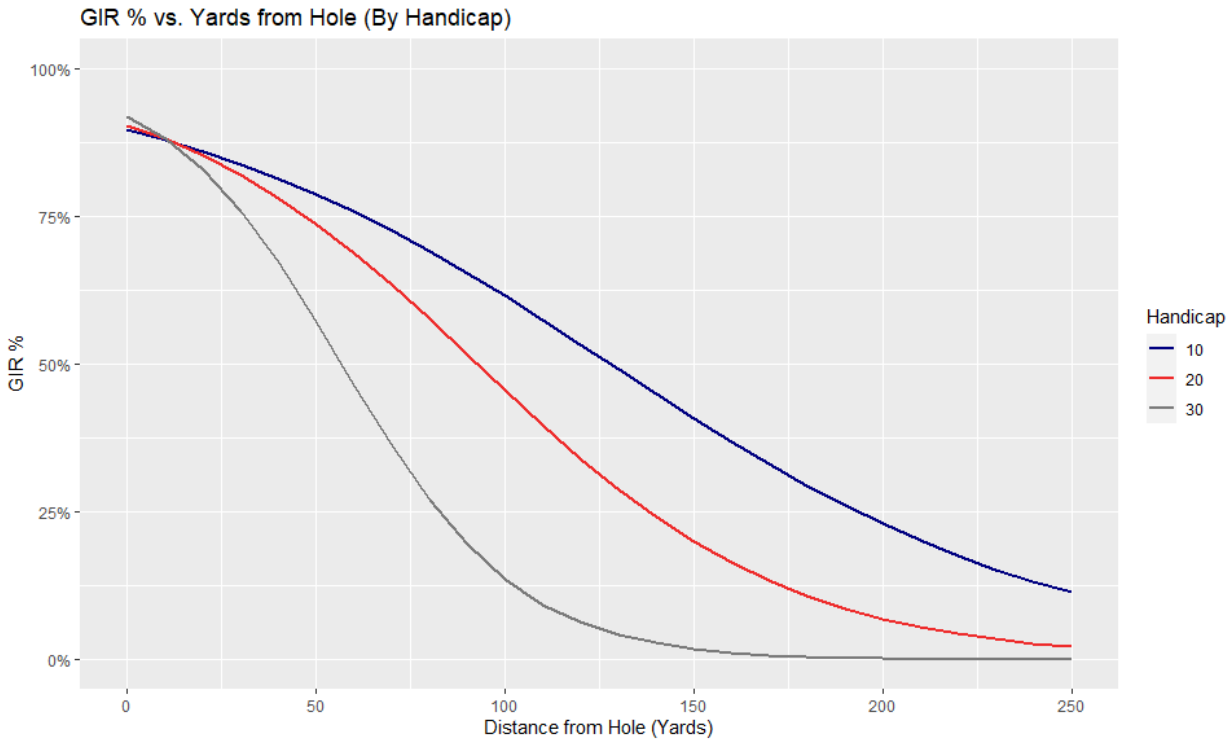
Figures 4 and 5 illustrate visually the results of the data modeling for the predicted GIR percentage as a function of distance from the hole by gender and by Handicap Index. Interpreting the results of the model for these variables using these charts may be easier compared to interpreting the model coefficients in Table 3.

Figure 4: Predicted Greens in Regulation (GIR%) Success Rate vs. Yards from Hole (By Gender)



As shown in Figure 4, male golfers have a greater likelihood of hitting the GIR from any distance. Male golfers have a 50% chance of hitting the green from about 112 yards while female golfers have the same chance from about 80 yards. Figure 5 shows that better golfers, as indicated by lower Handicap Index, also have a better chance at hitting a GIR from all distances. A golfer with a 10.0 Handicap Index would be expected to hit 50% of GIR from about 130 yards, a 20.0 Handicap Index golfer from about 90 yards and a 20.0 Handicap Index golfer from about 57 yards.

Figure 5: Predicted Greens in Regulation (GIR%) Success Rate vs. Yards from Hole (By Handicap Index)



6. Conclusions

Overall, this is a viable method to gather on-course golfer performance data. The relative costs and efforts for this type of data collection are low and the data were able to show statistically significant results. With an increased sample size and additional demographic information, it is likely that additional relationships could be determined from this type of analysis.