

# Golf Performance

Marc Edwards, Yedin Lui, Xinzhe Qi  
36-493: Sports Analytics Capstone

## The Client

Carnegie Mellon University Golf  
Head Coach Daniel Rodgers

## Project Goals

With the Carnegie Mellon Men's and Women's Golf programs evolving into front-runners in NCAA Division III, Coach Daniel Rodgers is looking to leverage round data collected from each player to identify individual and team strengths and weaknesses.

Utilize R Shiny to create an user-friendly interactive dashboard

## Data Collection

Each player on the men's and women's team completes a Google Form after each round recording statistics on hole-by-hole score, fairways/greens in regulation, number of putts, approach distance to green, length of first putt, etc.

## Data Mutation

The Google Form outputs a spreadsheet that we then load into R. Before performing analysis and creating any visualizations, we modify the format of the raw data to allow for easy manipulation.

## Exploratory Data Analysis

After formatting the data, we analyze the information provided through tabulations and visualizations to extract high-level summaries of the data

## R Shiny Application Development

To ensure the data visualization is beneficial to our client, we develop the base code for a user-friendly application mirroring an interactive dashboard.

## Application & Data Integration

Upon completion of the shiny app development, we integrate existing analysis and incorporate golf statistics into the application.

## Next Season Preparation

Our next steps include establishing standard procedures for data entry, and creating an application that will run with the updated Google Form provided by Coach Rodgers.

# Mutating the Data

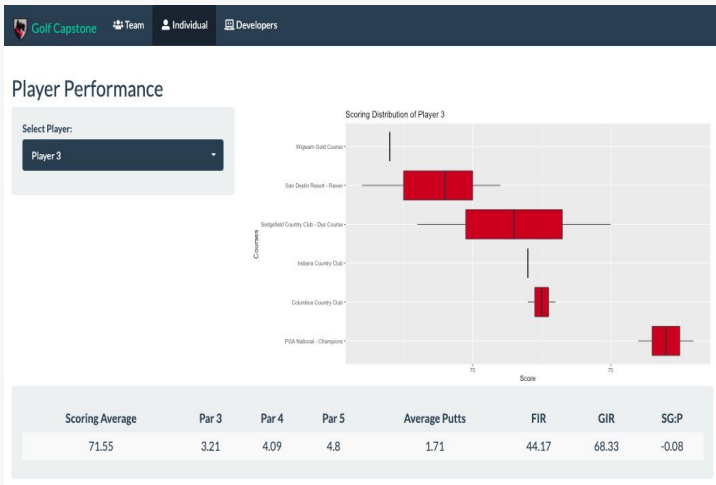
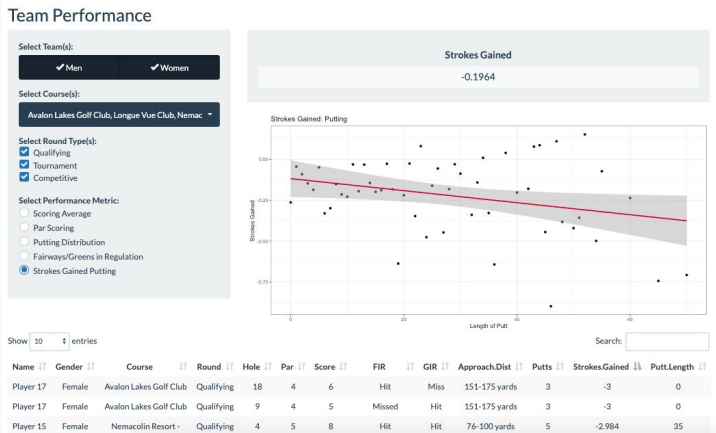
In our original dataset, each row corresponded to one player who played one round of golf. This meant that all of the data including the holes, pars, scores, approach distance, putts, putt length, and greens and fairways in regulation were all in one row. In order to solve this problem, we utilized a ‘for’ loop to transpose the data and make it easier to read for further use. After creating the new data frame, the data went from 168 observations of 151 variables to 3192 observations of 14 variables. By elongating the data, it made it easier for us to filter and subset the data by using consistent variable names across the 3192 observations.

Original Format

Name	Date	Course	Score	Round	Par	Score	Par	Fairway Hit	Green Hit	Approach Di	Putts	First P	Penalty	Shot	Score	Par	Fairway Hit	Green Hit
Player 1	8/28/19	Longue Vue	83	Qualifying	70	5	4	Missed	Hit	126-150 yarc	3	40	No Penalty	5	4	4	Hit	Hit
Player 2	8/28/19	Longue Vue	81	Qualifying	70	5	4	Hit	Miss	151-175 yarc	2	10	No Penalty	5	4	4	Hit	Hit
Player 3	8/30/19	Longue Vue	70	Qualifying	70	4	4	Hit	Hit	126-150 yarc	2	24	No Penalty	5	4	4	Hit	Hit
Player 3	8/31/19	Avalon Lakes	71	Qualifying	72	5	4	Hit	Miss	176-200 yarc	2	7	No Penalty	5	4	4	Hit	Hit
Player 4	8/31/19	Avalon Lakes	73	Qualifying	72	4	4	Missed	Miss	126-150 yarc	1	4	No Penalty	5	4	4	Missed	Hit
Player 5	8/30/19	Longue Vue	79	Qualifying	70	4	4	Hit	Hit	151-175 yarc	2	25	No Penalty	5	4	4	Hit	Hit
Player 6	8/31/19	Avalon Lakes	75	Qualifying	72	5	4	Missed	Miss	126-150 yarc	2	10	No Penalty	5	4	4	Missed	Miss
Player 1	8/31/19	Avalon Lakes	83	Qualifying	72	5	4	Missed	Hit	151-175 yarc	3	25	No Penalty	5	5	4	Missed	Miss
Player 7	8/31/19	Avalon Lakes	81	Qualifying	72	5	4	Hit	Miss	101-125 yarc	2	6	No Penalty	5	6	4	Hit	Miss
Player 8	8/31/19	Avalon Lakes	77	Qualifying	72	4	4	Missed	Hit	151-175 yarc	2	7	No Penalty	5	5	4	Missed	Miss
Player 9	8/28/19	Longue Vue	80	Qualifying	70	5	4	Hit	Hit	151-175 yarc	3	30	No Penalty	5	4	4	Hit	Hit
Player 10	8/31/19	Avalon Lakes	76	Qualifying	72	7	4	Missed	Miss	176-200 yarc	3	38	No Penalty	5	5	4	Missed	Miss
Player 11	8/31/19	Avalon Lakes	83	Qualifying	72	4	4	Hit	Hit	101-125 yarc	2	20	No Penalty	5	6	4	Missed	Miss
Player 9	8/31/19	Avalon Lakes	78	Qualifying	72	5	4	Hit	Miss	126-150 yarc	2	15	No Penalty	5	5	4	Missed	Miss
Player 2	8/30/19	Longue Vue	81	Qualifying	70	5	4	Missed	Miss	51-75 yards	2	18	No Penalty	5	5	4	Hit	Hit
Player 2	8/31/19	Avalon Lakes	86	Qualifying	72	6	4	Missed	Miss	176-200 yarc	2	10	Red Penalty	4	4	4	Hit	Hit
Player 15	8/31/19	Avalon Lakes	82	Qualifying	72	5	4	Hit	Hit	126-150 yarc	3	20	No Penalty	5	6	4	Hit	Hit
Player 16	8/31/19	Avalon Lakes	75	Qualifying	72	6	4	Missed	Miss	176-200 yarc	2	15	Red Penalty	5	4	4	Hit	Hit
Player 12	8/31/19	Avalon Lakes	77	Qualifying	72	4	4	Hit	Hit	151-175 yarc	2	20	No Penalty	5	6	4	Missed	Miss
Player 15	9/1/19	Nemacolin R	81	Qualifying	72	6	5	Hit	Miss	101-125 yarc	2	45	No Penalty	5	5	4	Missed	Hit
Player 15	9/1/19	Nemacolin R	87	Qualifying	72	5	5	Hit	Hit	51-75 yards	2	30	No Penalty	5	5	4	Hit	Miss
Player 1	9/1/19	Nemacolin R	81	Qualifying	72	5	5	Hit	Miss	76-100 yard	1	10	No Penalty	5	5	4	Missed	Miss
Player 10	9/1/19	Nemacolin R	73	Qualifying	72	5	5	Hit	Hit	101-125 yarc	2	28	No Penalty	5	6	4	Missed	Miss
Player 12	9/1/19	Nemacolin R	74	Qualifying	72	5	5	Hit	Hit	Less than 50	2	25	No Penalty	5	3	4	Hit	Hit
Player 11	9/1/19	Nemacolin R	72	Qualifying	72	6	5	Hit	Hit	101-125 yarc	3	30	No Penalty	5	3	4	Hit	Hit
Player 11	9/1/19	Nemacolin R	71	Qualifying	72	4	5	Hit	Hit	Less than 50	1	8	No Penalty	5	4	4	Hit	Hit
Player 16	9/1/19	Nemacolin R	93	Qualifying	72	6	5	Hit	Miss	101-125 yarc	2	18	No Penalty	5	7	4	Hit	Miss
Player 16	9/1/19	Nemacolin R	86	Qualifying	72	5	5	Hit	Hit	76-100 yard	2	30	No Penalty	5	7	4	Hit	Miss
Player 13	9/1/19	Nemacolin R	75	Qualifying	72	5	5	Missed	Hit	76-100 yard	2	40	No Penalty	5	4	4	Missed	Hit
Player 13	8/30/19	Avalon Lakes	85	Qualifying	72	3	4	Missed	Hit	151-175 yarc	1	40	No Penalty	5	4	4	Missed	Hit
Player 17	8/31/19	Avalon Lakes	81	Qualifying	72	5	4	Hit	Miss	176-200 yarc	2	0	No Penalty	5	4	4	Missed	Hit

Revised Format

Name	Course	Round	Hole	Par	Score	FIR	GIR	Approach Di	Putts	PuttLength	Gender
Player 1	Longue Vue	(Qualifying		1	4	5 Missed	Hit	126-150 yarc	3	40	Male
Player 2	Longue Vue	(Qualifying		1	4	5 Hit	Miss	151-175 yarc	2	10	Male
Player 3	Longue Vue	(Qualifying		1	4	4 Hit	Hit	126-150 yarc	2	24	Male
Player 3	Avalon Lakes	Qualifying		1	4	5 Hit	Miss	176-200 yarc	2	7	Male
Player 4	Avalon Lakes	Qualifying		1	4	4 Missed	Miss	126-150 yarc	1	4	Male
Player 5	Longue Vue	(Qualifying		1	4	4 Hit	Hit	151-175 yarc	2	25	Male
Player 6	Avalon Lakes	Qualifying		1	4	5 Missed	Miss	126-150 yarc	2	10	Male
Player 1	Avalon Lakes	Qualifying		1	4	5 Missed	Hit	151-175 yarc	3	25	Male
Player 7	Avalon Lakes	Qualifying		1	4	5 Hit	Miss	101-125 yarc	2	6	Male
Player 8	Avalon Lakes	Qualifying		1	4	4 Missed	Hit	151-175 yarc	2	7	Male
Player 9	Longue Vue	(Qualifying		1	4	5 Hit	Hit	151-175 yarc	3	30	Male
Player 10	Avalon Lakes	Qualifying		1	4	7 Missed	Miss	176-200 yarc	3	38	Male
Player 11	Avalon Lakes	Qualifying		1	4	4 Hit	Hit	101-125 yarc	2	20	Male
Player 9	Avalon Lakes	Qualifying		1	4	5 Hit	Miss	126-150 yarc	2	15	Male
Player 2	Longue Vue	(Qualifying		1	4	5 Missed	Miss	51-75 yards	2	18	Male
Player 2	Avalon Lakes	Qualifying		1	4	6 Missed	Miss	176-200 yarc	2	10	Male
Player 12	Avalon Lakes	Qualifying		1	4	4 Hit	Hit	151-175 yarc	2	20	Male
Player 1	Nemacolin R	Qualifying		1	5	5 Hit	Miss	76-100 yards	1	10	Male
Player 10	Nemacolin R	Qualifying		1	5	5 Hit	Hit	101-125 yarc	2	28	Male
Player 12	Nemacolin R	Qualifying		1	5	5 Hit	Hit	Less than 50	2	25	Male
Player 11	Nemacolin R	Qualifying		1	5	6 Hit	Hit	101-125 yarc	3	30	Male
Player 11	Nemacolin R	Qualifying		1	5	4 Hit	Hit	Less than 50	1	8	Male



# Individual and Team Performance Metrics

## Scoring Average

Average score across all golf rounds  
Goal: 72 strokes per round (regulation par)

## Par Scoring

Average score recorded across all holes, given par  
Goal: Average at or below respective par (3, 4, 5)

## Putting Distribution

Average number of putts across all holes  
Goal: 2 or less putts per hole

## Fairways/Greens in Regulation

Percent of fairways hit on the drive / Percent of greens hit in regulation  
Goal: Greater than 50% of fairways / Greater than 60% of greens

## Strokes Gained: Putting

Measure of putting performance compared to average PGA Tour player, taking into account length of putt  
Goal: Positive strokes gained in putting