ADA Data Project #2 Write Up

For my second ADA data project I looped in my Text Mining Data Set Share and Acquire and Analyze’s theme of golf. The golfers referenced in those projects are on this list at #8 and #15 respectively. I pulled the data from Official World Golf Ranking’s website which has an open API. There are other golf data websites but most sit behind a paywall of some sort. I pulled my data set from the main page which has the top 300 golfers in the world.

As I worked on pulling the data, I realized I could do it two different ways: pulling by website element into a table pretty immediately or pulling the text and then cleaning and organizing. I chose to pull by website element but then discovered some challenges. The numbers associated with the rankings hinged on different website element IDs so I couldn’t just pull them like I did with the player names and countries. I would’ve had to do a lot more complicated coding to pull those so I ended up sticking with my simpler pull and coding the rankings myself. This didn’t include all the different player statistics which would’ve been cool.

However, I did some googling around and found this guy who did it the more coding intense way in order to include those statistics. You can check it out here: <https://github.com/bradklassen/Official_World_Golf_Ranking_Scraper>

I think down the line for a larger project it would be worth replicating and player around with Brad Klassen’s code so I could dig more into the statistics of each player. Golf is very data friendly and provides lots of opportunities for analysis. Jon Rahm the top ranked golfer as of this data set pull would be fun to track on specific points levels. Breaking it down by points would also enable broader ranking comparison to past top players.

For my own statistics I created a groupby to see the count of top golfers by country. This is something that could be fun to track over time- especially since the United States leads with 111. Japan comes in second place with 26 which is an extreme difference. I think in some ways that could be traced back to my Acquire and Analyze which focused on the Twitter Activity and Marketing of a golf event a few weeks ago. It’s pretty evident that the U.S. has made men’s professional golf a whole industry.