

PS4_R*ann*

bradleyrann

February 2023

1 Part 1

I'm interested in scraping data from two main sources really. First, I would like to sparse professional tennis data and track how well American players are doing (they've been poor since Andy Roddick but a youngish generation has started to climb the ranks). I would also like to scrape data that would relate to health, whether it's epidemiological or clinical data about certain treatments or something I could apply to health economics.

2 Part 2

The results of `class(df1)` and `class(df2)` are as follows: `class(df1) tbl_df, tbl, data.frame`
`class(df) "tbl_spark" "tbl_sql" "tbl_lazy" "tbl"`

The column names are different because `df1` is created only in tidyverse whereas `df` is a copy of it in spark so it's added different languages and functions like SQL.