

PS5_A*deniran*

femi adeniran

March 2023

1 Scraping

To chose what I wanted to scape was a bit of a hassle for me at first. So I googled commonly scraped websites and found out about yellowpages and it instantly clicked. I had just made a brand new meatpie recipe it would be nice to get the phone numbers of all restuarants listed in Norman and ask them if they are willing to taste and maybe buy to resell my favorite delight.

I went the way of the python, as I am thoroughbred pythonista but not so thoroughbred as to be confident to clang it out as fast as it needed to be turned in. So I made my way to ChatGPT and for the first time we worked together on the problem.

It was a pretty interesting experience pasting errors back at it, in many cases i knew what was wrong but waited to see its response , all in all I found It was easier when you guided it to do things in small chunks and tell him what the problem was as opposed to just posting only the error message.

E.g pagination is causing trouble, does href match? I ended up scraping phones and address, I think I can use the map data to see places that are closer to me and I might go back and scrape categories to filter it even better.

2 API CALLS

I knew for APIs I was gonna lean on english football or soccer as it fondly known in these here parts. I had a couple of packages in mind, and thinking of a few different types of viz. I usually scraped fbref but Statsbomb has an interesting R package for football data.

I decided since I used python for the first I could use a little R, I got data on shooting stats for 35000 players across Europe's Major football leagues and their minor subdivisions. Working in R can be pretty straightforward. I enjoy that but I think in the future i might use it for only a part of the analysis , I fear I would feel more at home in a Jupyter environment