Intro to webscraping

Due - 2022/1/30 - 11:59PM

Deliverables

1. python code file
2. document containing text and the plot

Included

1. Assignment PDF (this)
2. Boilerplate code file
3. Yield curve dataset

Submission

Submit your document and code file on canvas before January 31st.

Scraping

You'll often find data web pages that's not in the format you want. Luckily for us, we can grab that data in a process called web scraping. We can manipulate then manipulate the data as we like. This assignment is a demonstration of how you might decide to do this. We're using pandas on a very well formatted website. For more complex use cases it might make sense to use a library called beautiful soup.

Code

There are 4 functions that you'll implement to create a plot which will contain 12 lines of rates vs time. There are comments in the boilerplate code for how to implement these functions. Note the overall style and comments in the code.

Questions

What do you notice about the graph and how it connects to significant events in the past 20 years?