ISGB 79AA - Advanced Python for Financial Programming
Assignment 3 – Access Web Page Data with Pandas and Beautiful Soup – Yield Curves

This assignment has 2 parts. Each part involves using Python to get data from a web page with US Treasury interest rates, available at:
https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yield

You should submit 2 notebooks for this assignment., one for each part. You can use one or more cells for each part. Name your notebook like: LastnameFirstnameAsn3<part>, for example, SmithJohnAsn3A. Include in each cell of your notebook a Python comment at the start of the cell, describing the cell's statements.

Part A – Accessing a Web Page Table Using pandas

Use pandas read_html to access the Treasury yields from the page indicated above. In particular,

1) Call the read_html function to get all HTML tables from the above web page. How many tables are returned?

2) Determine which of the returned tables in the list is the table of interest rates over time (e.g., the table at index 0, index 1, etc.).

3) From the list of returned DataFrames, set the DataFrame with the interest rates into its own variable.

4) For the DataFrame in step #3, set the Date column to be the row names (index). Display this DataFrame.

5) From the table of yields (one row for each date), extract the latest yields into a Python list. This may be the first or last row of the table. (The particular date will depend on when you run your program.)

6) Plot the latest yield curve. This is a scatter chart with yields (in percent) on the Y axis, and time (in years) on the X axis. Time should be represented as a list, e.g., time = [1/12, 2/12, …, 30]

7) Plot the time series of the 2 year and 10 year yields. These are the values in the columns labels '2 yr' and '10 yr'.

Part B – Accessing a Web Page Table Using Beautiful Soup

Use Beautiful Soup to access the Treasury yields from the page indicated above. In particular,

1) Locate the <table> element in the page corresponding to the Treasury yields.

2) Access the first row in the table.

3) Similar to part A step 6, plot the latest yield curve.

Appendix – Plotting notes

Scatter plots

There are several options for creating a simple scatter plot in Python. Here is a simple example using matplotlib, assuming x and y are lists of numbers, with the same number of elements:

```
import matplotlib.pyplot as plt
plt.scatter(x, y)
plt.show()
```

Line charts

matplotlib can plot 1 or more time series, taking the data from a DataFrame. Here is an example, assuming df is a DataFrame with columns 'A' and 'B':

```
fig, ax = plt.subplots()
plt.plot("A", data=df)
plt.plot("B", data=df)
plt.legend()
ax.tick_params(labelbottom=False) # supress x-axis labels
plt.show()
```