Homework #2

IR and GenAI

Due Oct./30/2024

Design web crawlers to fetch only the data of “**臺股期貨未平倉餘額(口數)**” from the following links:

- 臺灣期貨交易所-交易資訊-三大法人-查詢-區分各期貨契約-依日期 (taifex.com.tw)

			交易口數與契約金額						未平倉餘額					
			多方		空方		多空淨額		多方		空方		多空淨額	
序號	商品名稱	身份別	口數	契約金額	口數	契約金額	口數	契約金額	口數	契約金額	口數	契約金額	口數	契
1	臺股期貨	自營商	9,853	45,265,577	10,599	48,723,615	-746	-3,458,038	4,249	19,568,423	5,600	25,797,407	-1,351	-(
		投信	19,546	90,072,510	19,576	90,061,703	-30	10,807	31,574	145,464,986	6,042	27,837,777	25,532	11
		外資	84,195	386,873,226	81,417	374,271,491	2,778	12,601,735	24,855	114,527,707	63,926	294,429,398	-39,071	-17

Figure 1

- 臺灣期貨交易所-交易資訊-三大法人-查詢-區分各期貨契約-夜盤 (taifex.com.tw)

期貨契約

單位：口數；千元　　日期2024/10/14

序號	商品名稱	身份別	交易口數與契約金額					
			多方		空方		多空淨額	
			口數	契約金額	口數	契約金額	口數	契約金額
1	臺股期貨	自營商	2,178	9,993,213	2,072	9,513,946	106	479,267
		投信	56	256,672	0	0	56	256,672
		外資	19,258	88,345,745	18,028	82,710,179	1,230	5,635,566

Figure 2

For Figures 1 and 2, the input parameters are:

日期**: set the date on which you watch to fetch the data.**

契約**: 臺股期貨**

The task is to design a crawler to fetch the data (as many days as possible) and arrange the final results to be displayed in a .xlsx file as follows:

	A	B	C	D	E	F	G	H	I	J
1	DATE	自營商多	自營商空	自營商多	投信多	投信空	投信多空	外資多	外資空	外資多空淨額
2	20241004	4371	5502	-1,131	31397	6766	24,631	22503	59503	-37,000
3	20241007			-1,083			24,573			-38,882
4	20241007	5420	5773	-353	31762	6487	25,275	21579	63725	-42,146
5	20241008			-183			25,783			-40,805
6	20241008	5228	5220	8	31884	6476	25,408	19156	60246	-41,090
7	20241009			-705			25,351			-41,498
8	20241009	4737	5357	-620	31862	6340	25,522	20073	60523	-40,450
9	20241011			-704			25,522			-39,601
10	20241011	4949	5642	-693	31732	6170	25,562	23024	64873	-41,849
11	20241014			-587			25,618			-40,619
12	20241014	4249	5600	-1,351	31574	6042	25,532	24855	63926	-39,071

Figure 3

Notice that for the night market (夜盤), the data is only available as 交易口數; so for 未平倉口數 **(open interests)**, you have to derive it from the previous row to compute 未平倉口數 by yourself. E.g., the data of D11 in Figure 3 is computed by D10+106, where 106 is the number circled by red in Figure 2. This rule applies to the calculation for the data in G11 and J11, etc.

Also notice that your script has to be able to fetch data for every trading day, so that the xlsx file can be updated accordingly.

In this homework, you are encouraged to use generative AI for assistance. Describe in a word file regarding how you use GenAI and your own effort to get the job done, including what you have learned. Then compress your source code, the xlsx file, and the word file using your student ID, e.g., hw2_A1105501.zip, and upload to Moodle.