

爬蟲臺灣銀行牌告匯率資料

臺灣銀行牌告匯率為臺灣官方正式匯率宣告單位，因為匯率變動對於投資風險管控必須要非常即時，因此快速爬蟲是本程式一大重點，臺灣銀行牌告匯率網址如下：

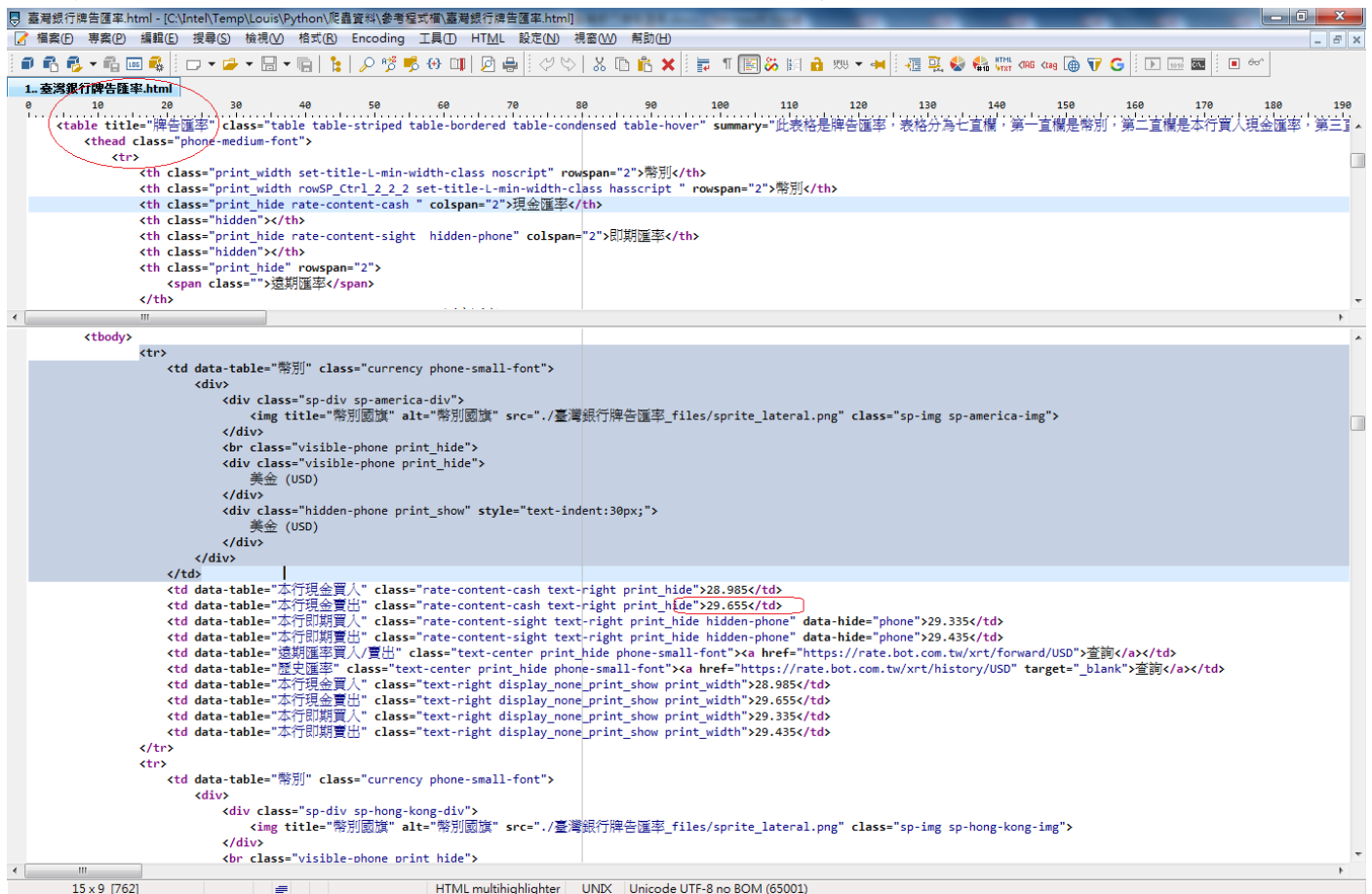
幣別	現金匯率		即期匯率		遠期匯率	歷史匯率
	本行買入	本行賣出	本行買入	本行賣出		
美金 (USD)	29.09	29.76	29.44	29.54	查詢	查詢
港幣 (HKD)	3.649	3.853	3.775	3.835	查詢	查詢
英鎊 (GBP)	37.3	39.42	38.3	38.72	查詢	查詢
澳幣 (AUD)	20.71	21.49	20.98	21.21	查詢	查詢
加拿大幣 (CAD)	21.53	22.44	21.92	22.14	查詢	查詢
新加坡幣 (SGD)	20.89	21.8	21.38	21.56	查詢	查詢
瑞士法郎 (CHF)	31.44	32.64	32.1	32.39	查詢	查詢
日圓 (JPY)	0.2689	0.2817	0.2762	0.2802	查詢	查詢
南非幣 (ZAR)	-	-	1.623	1.703	查詢	查詢
瑞典幣 (SEK)	2.98	3.5	3.32	3.42	查詢	查詢
紐元 (NZD)	18.95	19.8	19.33	19.53	查詢	查詢
泰幣 (THB)	0.8191	1.0091	0.9315	0.9715	查詢	查詢
菲律賓比索 (PHP)	0.5278	0.6608	-	-	查詢	查詢
印尼幣 (IDR)	0.00168	0.00238	-	-	查詢	查詢
歐元 (EUR)	33.88	35.22	34.5	34.9	查詢	查詢
韓元 (KRW)	0.02312	0.02702	-	-	查詢	查詢
越南盾 (VND)	0.0009	0.0014	-	-	查詢	查詢
馬來幣 (MYR)	5.915	7.54	-	-	查詢	查詢
人民幣 (CNY)	4.135	4.297	4.207	4.257	查詢	查詢

因為要快速抓取，因此用 `thread` 的方式，詳見 "臺灣銀行牌告匯率.py"，如下圖所示：

```
1 import threading, csv, ast, requests, time
2 from bs4 import BeautifulSoup
3
4 def job(single_tr):
5     cell = single_tr.findAll("td")
6     currency_name = cell[0].find("div", {"class": "visible-phone"}).contents[0]
7     currency_name = currency_name.replace("\r\n", "")
8     currency_name = currency_name.replace(" ", "")
9     currency_rate = cell[2].contents[0]
10    if Mydict[currency_name] == currency_rate:
11        return
12    else:
13        lock.acquire()
14        print(currency_name, currency_rate, Mydict[currency_name])
15        Mydict[currency_name] = currency_rate
16        file = open('MyCurrDict.csv', 'w', encoding='utf8')
17        file.write(str(Mydict))
18        file.close()
19        lock.release()
20
21 file = open('MyCurrDict.csv', 'r', encoding='utf8')
22 contents = file.read()
23 Mydict = ast.literal_eval(contents)
24 file.close()
25 lock = threading.Lock()
26
27 while True:
28     html = requests.get("https://rate.bot.com.tw/xrt?Lang=zh-TW")
29     bsObj = BeautifulSoup(html.content, "lxml")
30     All_tr = bsObj.find("table", {"title": "牌告匯率"}).find("tbody").findAll("tr")
31     Th = []
32     for single_tr in All_tr:
33         Th.append(threading.Thread(target = job, args = (single_tr,)))
34         Th[i].start()
35         i+=1
36     for j in range(i):
37         Th[j].join()
38     time.sleep(1)
39     print(".",end='')
40
41 Python file length:1,343 lines:40 Ln:13 Col:67 Sel:0|0 Windows (CRLF) UTF-8-BOM INS
```

其中我將匯率資訊存在 `MyCurrDict.csv`，因此只要觀察該檔即可了解最後牌價，程式中

```
bsObj.find("table", {"title": "牌告匯率"}).find("tbody").findAll("tr")
```



```
<table title="牌告匯率" class="table table-striped table-bordered table-condensed table-hover" summary="此表格是牌告匯率，表格分為七直欄，第一直欄是幣別，第二直欄是本行賣入美金匯率，第三直欄是幣別，第四直欄是本行買入美金匯率，第五直欄是本行即期匯率，第六直欄是本行即期匯率，第七直欄是本行即期匯率">
  <thead class="phone-medium-font">
    <tr>
      <th class="print_width set-title-L-min-width-class noscript" rowspan="2">幣別</th>
      <th class="print_width rowSP_Ctrl_2_2_2 set-title-L-min-width-class hasscript" rowspan="2">幣別</th>
      <th class="print_hide rate-content-cash" colspan="2">現金匯率</th>
      <th class="hidden"></th>
      <th class="print_hide rate-content-sight hidden-phone" colspan="2">即期匯率</th>
      <th class="hidden"></th>
      <th class="print_hide" rowspan="2">遠期匯率</th>
    </tr>
    <tr>
      <th class="print_hide rate-content-cash" colspan="2">現金匯率</th>
      <th class="hidden"></th>
      <th class="print_hide rate-content-sight hidden-phone" colspan="2">即期匯率</th>
      <th class="hidden"></th>
    </tr>
  </thead>
  <tbody>
    <tr>
      <td data-table="幣別" class="currency phone-small-font">
        <div class="sp-div sp-america-div">
          
        </div>
        <br class="visible-phone print_hide">
        <div class="visible-phone print_hide">美金 (USD)</div>
        <div class="hidden-phone print_show" style="text-indent:30px;>美金 (USD)</div>
      </td>
      <td data-table="本行現金賣入" class="rate-content-cash text-right print_hide">28.985</td>
      <td data-table="本行現金賣出" class="rate-content-cash text-right print_hide">29.655</td>
      <td data-table="本行即期賣入" class="rate-content-sight text-right print_hide hidden-phone" data-hide="phone">29.335</td>
      <td data-table="本行即期賣出" class="rate-content-sight text-right print_hide hidden-phone" data-hide="phone">29.435</td>
      <td data-table="遠期匯率賣入/賣出" class="text-center print_hide phone-small-font"><a href="https://rate.bot.com.tw/xrt/forward/USD">查詢</a></td>
      <td data-table="歷史匯率" class="text-center print_hide phone-small-font"><a href="https://rate.bot.com.tw/xrt/history/USD" target="_blank">查詢</a></td>
      <td data-table="本行現金賣入" class="text-right display_none_print_show print_width">28.985</td>
      <td data-table="本行現金賣出" class="text-right display_none_print_show print_width">29.655</td>
      <td data-table="本行即期賣入" class="text-right display_none_print_show print_width">29.335</td>
      <td data-table="本行即期賣出" class="text-right display_none_print_show print_width">29.435</td>
    </tr>
    <tr>
      <td data-table="幣別" class="currency phone-small-font">
        <div class="sp-div sp-hong-kong-div">
          
        </div>
        <br class="visible-phone print_hide">
      </td>
      <td data-table="本行現金賣入" class="rate-content-cash text-right print_hide">28.985</td>
      <td data-table="本行現金賣出" class="rate-content-cash text-right print_hide">29.655</td>
      <td data-table="本行即期賣入" class="rate-content-sight text-right print_hide hidden-phone" data-hide="phone">29.335</td>
      <td data-table="本行即期賣出" class="rate-content-sight text-right print_hide hidden-phone" data-hide="phone">29.435</td>
      <td data-table="遠期匯率賣入/賣出" class="text-center print_hide phone-small-font"><a href="https://rate.bot.com.tw/xrt/forward/USD">查詢</a></td>
      <td data-table="歷史匯率" class="text-center print_hide phone-small-font"><a href="https://rate.bot.com.tw/xrt/history/USD" target="_blank">查詢</a></td>
      <td data-table="本行現金賣入" class="text-right display_none_print_show print_width">28.985</td>
      <td data-table="本行現金賣出" class="text-right display_none_print_show print_width">29.655</td>
      <td data-table="本行即期賣入" class="text-right display_none_print_show print_width">29.335</td>
      <td data-table="本行即期賣出" class="text-right display_none_print_show print_width">29.435</td>
    </tr>
  </tbody>
</table>
```

```
cell = single_tr.findAll("td")
```

```
currency_name = cell[0].find("div", {"class": "visible-phone"}).contents[0]
```

```
currency_rate = cell[2].contents[0]
```

利用以上三行指令將[美金(USD)]及[29.655]抓取然後利用[美金(USD)]當作 key 把抓取來的資料與 MyCurrDict.csv 值比較，如果匯率沒改變，我就不做任何處理結束執行緒，如果有改變我就更新 MyCurrDict.csv 資料庫再結束執行緒。

```
for single_tr in All_tr:
    Th.append(threading.Thread(target = job, args = (single_tr)))
    Th[i].start()
    i+=1
for j in range(i):
    Th[j].join()
```

這段程式就是將匯率所有幣別全部轉成執行緒，然後將參數用 args=()方式建立，最後全部 join 才繼續下一步。最後執行結果如下圖所示：

```
In [1]: runfile('C:/Intel/Temp/Louis/Python/爬蟲資料/爬蟲資料/匯率.py', wdir='C:/Intel/Temp/Louis/Python/爬蟲資料/爬蟲資料')
美金(USD) 29.655 29.688
港幣(HKD) 3.839 3.822
英鎊(GBP) 39.36 39.33
澳幣(AUD) 21.46 21.55
人民幣(CNY) 4.293 4.299
.....美金(USD) 29.66 29.655
港幣(HKD) 3.84 3.839
英鎊(GBP) 39.38 39.36
澳幣(AUD) 21.47 21.46
加拿大幣(CAD) 22.46 22.45
新加坡幣(SGD) 21.74 21.73
瑞士法郎(CHF) 32.49 32.48
泰幣(THB) 1.0071 1.0068
菲律賓比索(PHP) 0.6596 0.6598
印尼幣(IDR) 0.00233 0.00238
歐元(EUR) 35.05 35.04
馬來幣(MYR) 7.51 7.509
人民幣(CNY) 4.294 4.293
.....
```