

網路擷取證券交易

```
In [1]: import pandas as pd
import requests
import io
import datetime
import time

In [2]: def crawler(date_time):
page_url = 'https://www.twse.com.tw/exchangeReport/MI_INDEX?response=csv&date=' + date_time + '&type=ALLBUT0999'
page = requests.get(page_url)
use_text = page.text.splitlines()
for i, text in enumerate(use_text):
    if text == '"證券代號","證券名稱","成交股數","成交筆數","成交金額","開盤價","最高價","最低價","收盤價","漲跌(+/-)","漲跌價差","最後揭示買價","最後揭示買量","最後揭示賣價","最後揭示賣量"'
        initial_point = i
        break

df = pd.read_csv(io.StringIO(''.join([text[:-1] + '\n' for text in use_text[initial_point:]]))
df['證券代號'] = df['證券代號'].apply(lambda x:x.replace('=',' '))
df['證券代號'] = df['證券代號'].apply(lambda x:x.replace('"',' '))
return df

def trans_date(date_time):
return ''.join(str(date_time).split(' ')[0].split('-'))

def parse_n_days(start_date,n):
df_dict = {}
now_date = start_date
for i in range(n):
    time.sleep(3)
    now_date = now_date - datetime.timedelta(days=1)
    try:
        df = crawler(trans_date(now_date))
        print("Current date" + ' ' +trans_date(now_date))
        df_dict.update({trans_date(now_date):df})
        print('Successful!!')
    except:
        print('Fails at'+ ' ' + trans_date(now_date))
return df_dict

In [3]: result_dict = parse_n_days(datetime.datetime.now(), 5)
result_dict
```

Fails at 20210221
Fails at 20210220
Current date 20210219
Successful!!
Current date 20210218
Successful!!
Current date 20210217
Successful!!

```
Out[3]: {'20210219':      證券代號      證券名稱      成交股數      成交筆數      成交金額      開盤價      最高價      最低價      \
0      0050      元大台灣50      9,327,689      14,087      1,291,374,488      139.55      139.55      137.60
1      0051      元大中型100      23,852      38      1,107,115      46.32      46.65      46.32
2      0052      富邦科技      2,463,157      1,850      331,450,168      135.10      135.20      133.50
3      0053      元大電子      28,050      24      1,918,151      68.65      68.65      67.85
4      0054      元大台商50      5,000      3      151,440      30.22      30.54      30.22
...      ...      ...      ...      ...      ...      ...      ...
1120      9944      新麗      147,149      141      3,043,068      20.65      20.80      20.60
1121      9945      潤泰新      3,843,711      1,959      156,352,371      40.85      40.90      40.50
1122      9946      三發地產      304,786      212      4,759,228      15.65      15.70      15.55
1123      9955      佳龍      383,235      254      7,058,168      18.40      18.65      18.25
1124      9958      世紀鋼      2,919,153      1,866      322,013,262      109.50      111.50      108.00
```

	收盤價	漲跌(+/-)	漲跌價差	最後揭示買價	最後揭示買量	最後揭示賣價	最後揭示賣量	本益比
0	138.85	-	1.50	138.80	21	138.85	93	0.00
1	46.65	+	0.30	46.46	1	46.65	10	0.00
2	134.60	-	1.55	134.55	13	134.60	58	0.00
3	68.25	-	0.40	68.25	2	68.65	2	0.00
4	30.54	-	0.06	30.32	51	30.54	18	0.00
...
1120	20.80	+	0.10	20.70	3	20.80	37	61.18
1121	40.90		0.00	40.85	35	40.90	2	8.85
1122	15.65	+	0.05	15.60	13	15.65	5	8.94
1123	18.35	+	0.10	18.30	9	18.35	15	0.00
1124	111.50	+	2.00	111.00	15	111.50	135	31.41

```
[1125 rows x 16 columns],
'20210218':      證券代號      證券名稱      成交股數      成交筆數      成交金額      開盤價      最高價      \
0      0050      元大台灣50      10,805,214      11,664      1,515,128,298      140.55      141.25
1      0051      元大中型100      39,353      59      1,822,473      45.83      46.49
2      0052      富邦科技      1,436,221      1,450      195,177,031      136.30      137.00
3      0053      元大電子      55,273      47      3,808,365      68.80      69.20
4      0054      元大台商50      15,307      14      470,032      30.74      30.80
...      ...      ...      ...      ...      ...      ...
1120      9944      新麗      164,538      117      3,396,109      20.45      20.70
1121      9945      潤泰新      6,963,254      3,491      283,746,242      40.25      41.00
1122      9946      三發地產      412,487      341      6,413,494      15.40      15.70
1123      9955      佳龍      202,042      154      3,694,094      18.50      18.50
1124      9958      世紀鋼      4,497,485      2,801      495,217,910      107.00      112.00
```

	最低價	收盤價	漲跌(+/-)	漲跌價差	最後揭示買價	最後揭示買量	最後揭示賣價	最後揭示賣量	本益比
0	139.60	140.35	-	0.20	140.30	10	140.35	94	0.00
1	45.83	46.35	+	0.52	46.35	1	46.44	4	0.00
2	135.00	136.15	-	0.15	136.15	2	136.20	7	0.00
3	68.65	68.65	+	0.05	68.65	24	68.90	1	0.00
4	30.60	30.60	+	0.10	30.60	1	30.62	1	0.00
...
1120	20.45	20.70	+	0.05	20.65	2	20.70	33	60.88
1121	40.25	40.90	+	0.70	40.85	88	40.90	30	8.85
1122	15.40	15.60	+	0.25	15.55	2	15.60	14	8.91
1123	18.10	18.25	+	0.20	18.20	7	18.25	29	0.00
1124	107.00	109.50	+	1.50	109.50	26	110.00	87	30.85

```
[1125 rows x 16 columns],
'20210217':      證券代號      證券名稱      成交股數      成交筆數      成交金額      開盤價      最高價      \
0      0050      元大台灣50      19,308,197      18,365      2,703,193,279      139.00      140.85
1      0051      元大中型100      93,121      69      4,255,124      45.65      45.93
2      0052      富邦科技      2,305,128      2,149      314,678,942      137.00      137.50
3      0053      元大電子      66,303      68      4,545,706      68.30      69.00
4      0054      元大台商50      20,295      18      617,644      30.18      30.53
...      ...      ...      ...      ...      ...      ...
1120      9944      新麗      122,628      100      2,528,551      20.60      20.70
1121      9945      潤泰新      5,650,302      3,095      226,884,304      39.90      40.30
1122      9946      三發地產      475,073      309      7,235,674      15.15      15.55
1123      9955      佳龍      178,221      140      3,231,140      17.95      18.50
1124      9958      世紀鋼      3,837,774      2,538      409,276,755      104.50      108.50
```

	最低價	收盤價	漲跌(+/-)	漲跌價差	最後揭示買價	最後揭示買量	最後揭示賣價	最後揭示賣量	本益比
0	139.00	140.55	+	5.75	140.50	12	140.55	222	0.00
1	45.65	45.83	+	0.93	45.83	17	45.85	6	0.00
2	135.85	136.30	+	6.10	136.30	2	136.35	18	0.00
3	68.10	68.60	+	2.60	68.50	2	68.60	1	0.00
4	30.18	30.50	+	0.52	30.42	52	30.57	1	0.00
...
1120	20.55	20.65	-	0.05	20.55	4	20.65	1	60.74
1121	39.85	40.20	+	0.70	40.15	88	40.25	72	8.70
1122	15.00	15.35	+	0.35	15.30	5	15.35	5	8.77
1123	17.95	18.05	+	0.35	18.05	12	18.10	30	0.00
1124	102.00	108.00	+	4.50	107.50	59	108.00	129	30.42

```
[1125 rows x 16 columns]}
```

```
In [4]: for key in result_dict.keys():
        result_dict[key].to_csv(str(key)+'_csv', index=False)
        print('SAVED')
```

```
SAVED
SAVED
SAVED
```

```
In [5]: df_long = pd.read_csv('20210219.csv')
df_long = df_long[['證券名稱', '收盤價']]
df_long = df_long[df_long['證券名稱'].str.contains('台積電')]
df_long = df_long.reset_index(drop=True)
df_long
```

```
Out[5]:      證券名稱      收盤價
0      台積電      652.00
```