

I, Task

Từ Việt Thảo • 18 thg 12, 2023 (Đã chỉnh sửa 11 thg 1)

100 điểm

Đến hạn 20 thg 12, 2023

Nghiên cứu BeautifulSoup4: <https://pypi.org/project/beautifulsoup4/>
Docs: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

Bài tập:

- Sử dụng thư viện python **requests** và **BeautifulSoup4**, hãy crawl dữ liệu giá trong **30 ngày gần nhất** (không tính cuối tuần) của **Top 10 cổ phiếu** trên website <https://finance.vietstock.vn/>.

- Hiển thị dữ liệu giá này theo dạng như sau:

<Mã cổ phiếu>

<Ngày/tháng/năm>: <Giá>

- VD:

HPG

18/12/2023: 26,750

15/12/2023: 26,800

14/12/2023: 27,050

...

- Lưu dữ liệu vào file CSV, các cột phân tách bằng dấu phẩy ","

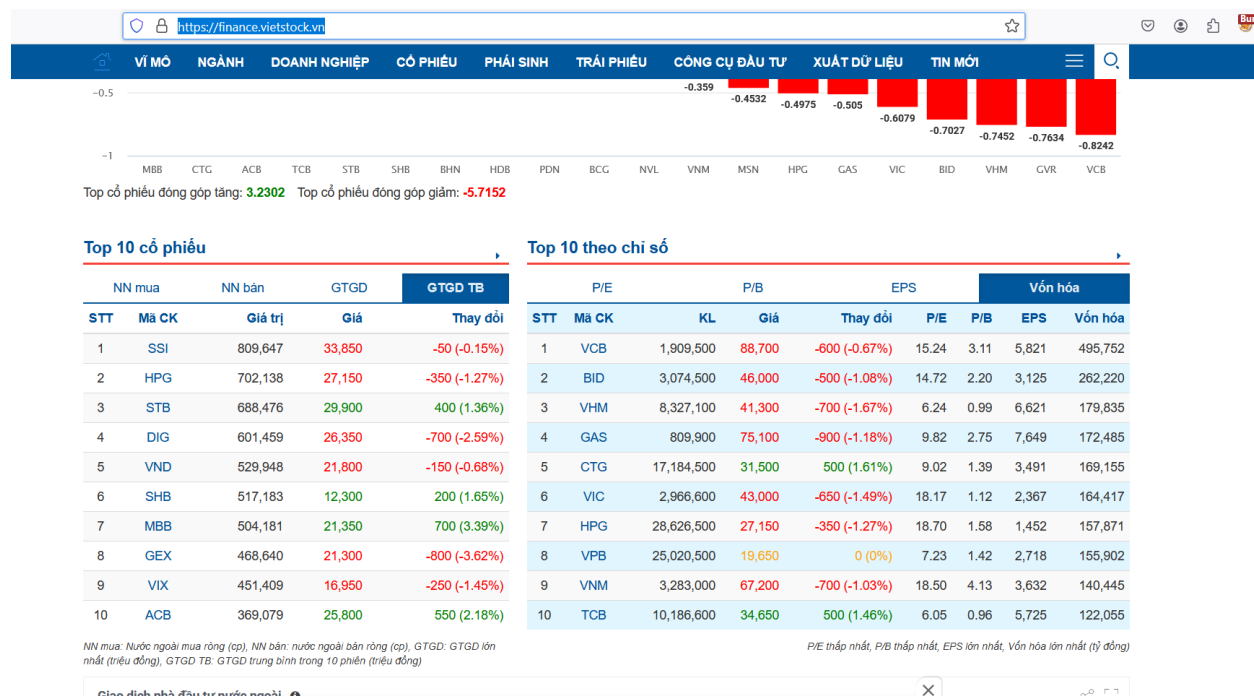
Nâng cao: Sử dụng thư viện matplotlib để vẽ biểu đồ giá của cổ phiếu top 1 sử dụng dữ liệu giá 30 ngày vừa crawl

II, solve

1, Điều tra khảo sát, tìm manh mối

Sau khi nhận được task, em bắt tay vào khảo sát, xem xét các nghiệp vụ của trang web

<https://finance.vietstock.vn/>. Khi vào trang web, em có thể thấy được top 10 cổ phiếu xếp theo **GTGD**



Em tiếp tục click vào những mã cổ phiếu đó để xem lịch sử giao dịch nằm ở đâu và tìm thấy 1 mục là thống kê giao dịch.

https://finance.vietstock.vn/SSI-ctcp-chung-khoan-ssi.htm

VÍ MÔ

NGÀNH

DOANH NGHIỆP

CỔ PHIẾU

PHÁI SINH

TRÁI PHIẾU

CÔNG CỤ ĐẦU TƯ

XUẤT DỮ LIỆU

TIN MỚI

Tiếp tục truy cập vào mục đó em thấy được lịch sử giao dịch mã cổ phiếu theo từng ngày. Giá đóng cửa là giá của cổ phiếu chốt theo ngày đó. Em tiếp tục set thời gian và chọn hiển thị 30 dòng theo đúng 30 ngày để xem và nhận được kết quả trả về là dữ liệu mà chúng ta cần crawling. Trường mà chúng ta cần crawling ở đây sẽ là “Ngày” và “Giá đóng cửa”.

SSI: CTCP Chứng khoán SSI - SSI X

https://finance.vietstock.vn/SSI/thong-ke-giao-dich.htm

90% ☆

VÍ MÔ NGÀNH DOANH NGHIỆP CỔ PHIẾU PHẢI SINH TRÁI PHIẾU CÔNG CỤ ĐẦU TƯ XUẤT DỮ LIỆU TIN MỚI

Trang 1/187 dòng/trang

Bảng Biểu đồ

Từ ngày 13/12/2023 Đến ngày 13/01/2024 Tùy chỉnh cột hiển thị trên lưới-

Khớp lệnh & Thỏa thuận Đầu tư nước ngoài

Hiển thị 30 dòng/trang

Xuất dữ liệu: MetaStock Excel Word

Ngày	Tổng KLGD	Tổng GTGD	Vốn hóa TT	Đóng cửa	+/- giá	+/- % giá	Khớp lệnh	
							KL	GT
12/01/2024	25,278,801	851,824	50,745,844	33,850	-50	-0.15	25,258,800	851,146
11/01/2024	25,571,700	867,148	50,820,801	33,900	200	0.59	25,471,700	863,725
10/01/2024	23,858,800	803,685	50,520,973	33,700	-300	-0.88	23,828,800	802,665
09/01/2024	19,185,200	654,168	50,970,715	34,000	-100	-0.29	18,985,200	647,348
08/01/2024	20,809,000	712,269	51,120,629	34,100	-50	-0.15	20,729,000	709,474
05/01/2024	25,137,200	851,906	51,195,586	34,150	550	1.64	25,137,200	851,906
04/01/2024	58,745,301	1,985,982	50,371,059	33,600	650	1.97	58,713,300	1,984,926
03/01/2024	11,408,702	372,451	49,396,619	32,950	450	1.38	11,408,700	372,451
02/01/2024	18,052,301	592,207	48,722,007	32,500	-300	-0.91	18,052,300	592,207
29/12/2023	12,790,302	420,618	49,171,748	32,800	-200	-0.61	12,790,300	420,618
28/12/2023	11,105,901	365,066	49,471,576	33,000	200	0.61	10,895,900	358,168
27/12/2023	16,311,500	537,400	49,171,748	32,800	50	0.15	16,101,500	530,489
26/12/2023	14,005,700	457,989	49,096,791	32,750	200	0.61	14,005,700	457,989
25/12/2023	14,208,301	461,329	48,796,964	32,550	200	0.62	13,953,300	453,087
22/12/2023	17,547,600	571,995	48,497,136	32,350	150	0.47	17,457,600	569,097
21/12/2023	8,246,601	265,292	48,272,265	32,200	-100	-0.31	8,083,600	260,027
20/12/2023	13,576,600	437,923	48,422,179	32,300	150	0.47	13,506,600	435,823
19/12/2023	15,757,700	500,225	48,197,308	32,150	450	1.42	15,687,700	497,974
18/12/2023	11,377,400	362,100	47,522,696	31,700	-100	-0.31	11,323,400	360,502

Oke vậy khảo sát các tính năng cùng với tìm thấy được dữ liệu chúng ta cần tìm. Giờ đến mục bắt request bằng Burpsuite. Phàm là những gì hiển thị được trên màn hình đều là từ response trả về từ những request. Em mất khoảng gần 1 ngày để tìm ra những request chính trả về những kết quả mà chúng ta cần tìm.

Em đã tìm thấy 1 request trở đến api <https://api.vietstock.vn/finance/toptrading?type=7&catID=1> trả về response đúng là top 10 cổ phiếu mà chúng ta đang tìm kiếm. Chúng ta có thể lấy mã cổ phiếu từ trường "StockCode"

DashboardTargetProxyIntruderRepeaterCollaboratorSequencerDecoderComparerLoggerExtensionsLearn

InterceptHTTP historyWebSockets historyProxy settings

Filter: Hiding CSS, image and general binary content

#	Host	Method	URL	Params	Edited	Status code	Length	MIME type	Extension	Title	Comment	TLS	IP	Cookies	Time	Listener port
1629	https://finance.vietstock.vn	POST	/uaa/heatmapInfo_v2	✓	200	200	1108	JSON				✓	221.132.38.197	language=vi-VN	23:53:40 13 ... 8080	
1628	https://finance.vietstock.vn	POST	/data/GetStockBySector_V3	✓	200	200	17914...	JSON				✓	221.132.38.197		23:53:39 13 ... 8080	
1627	https://api.vietstock.vn	GET	/finance/topstockfin?type=4&cat...	✓	200	200	2610	JSON				✓	221.132.38.195		23:53:39 13 ... 8080	
1626	https://finance.vietstock.vn	POST	/data/getmarketprice	✓	200	200	1308	JSON				✓	221.132.38.197		23:53:39 13 ... 8080	
1625	https://api.vietstock.vn	GET	/finance/toptrading?type=7&cat...	✓	200	200	1771	JSON				✓	221.132.38.195		23:53:39 13 ... 8080	
1624	https://finance.vietstock.vn	POST	/Data/GeneralMarket_GetMarket...	✓	200	200	4048	JSON				✓	221.132.38.197		23:53:38 13 ... 8080	
1623	https://finance.vietstock.vn	POST	/data/GetCommodities	✓	200	200	623	JSON				✓	221.132.38.197		23:53:37 13 ... 8080	
1622	https://finance.vietstock.vn	POST	/data/GetCurrency	✓	200	200	822	JSON				✓	221.132.38.197		23:53:37 13 ... 8080	
1621	https://finance.vietstock.vn	POST	/Data/GeneralMarket_GetTicker...	✓	200	200	611	JSON				✓	221.132.38.197		23:53:30 13 ... 8080	
1620	https://finance.vietstock.vn	POST	/data/heatmapInfo_v2	✓	200	200	1108	JSON				✓	221.132.38.197		23:53:25 13 ... 8080	
1619	https://finance.vietstock.vn	POST	/data/GetStockBySector_V3	✓	200	200	17914...	JSON				✓	221.132.38.197	language=vi-VN	23:53:25 13 ... 8080	
1618	https://finance.vietstock.vn	POST	/Data/GeneralMarket_GetMarket...	✓	200	200	4048	JSON				✓	221.132.38.197		23:53:23 13 ... 8080	
1617	https://finance.vietstock.vn	POST	/Data/Liquidity_GetTickerForChart...	✓	200	200	352	JSON				✓	221.132.38.197		23:53:22 13 ... 8080	
1613	https://finance.vietstock.vn	POST	/Data/GeneralMarket_GetTicker...	✓	200	200	611	JSON				✓	221.132.38.197		23:53:13 13 ... 8080	
1608	https://finance.vietstock.vn	POST	/data/heatmapInfo_v2	✓	200	200	1108	JSON				✓	221.132.38.197		23:53:09 13 ... 8080	

Request

RawHex

1 GET /finance/toptrading?type=7&catID=1 HTTP/1.1

2 Host: api.vietstock.vn

3 User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0

4 Accept: */*

5 Accept-Language: en-US,en;q=0.5

6 Accept-Encoding: gzip, deflate

7 Origin: https://finance.vietstock.vn

8 Referer: https://finance.vietstock.vn/

9 Sec-Fetch-Dest: empty

10 Sec-Fetch-Mode: cors

11 Sec-Fetch-Site: same-site

12 Te: trailers

13 Connection: close

14

15

Response

PrettyRawHexRender

"TradingDate": "2024-01-12T00:00:00",

"StockCode": "SSI",

"ClosePrice": 13850.0,

"Change": -50.0,

"PerChange": -0.15,

"TotalVol": 809647.0,

"Row": 1

},

{

"TradingDate": "2024-01-12T00:00:00",

"StockCode": "SPG",

"ClosePrice": 17150.0,

"Change": -350.0,

"PerChange": -1.27,

"TotalVol": 702138.0,

"Row": 2

},

{

"TradingDate": "2024-01-12T00:00:00",

"StockCode": "BIO",

"ClosePrice": 15600.0,

"Change": 400.0,

"PerChange": 1.36,

"TotalVol": 688476.0,

"Row": 3

},

{

"TradingDate": "2024-01-12T00:00:00",

"StockCode": "BIG",

"ClosePrice": 16150.0,

"Change": -700.0,

Inspector

Request attributes

Request query parameter

Request headers

Response headers

Tiếp tục lần sâu hơn, em truy cập vào địa chỉ <https://finance.vietstock.vn/SSI/thong-ke-giao-dich.htm> để bắt những request trả về giá của cổ phiếu sau 30 ngày. Em đã bắt được request trở tới địa chỉ api <https://finance.vietstock.vn/data/gettradingresult> trả về 1 file json gồm các dữ liệu của cổ phiếu trong vòng 30 ngày.

DashboardTargetProxyIntruderRepeaterCollaboratorSequencerDecoderComarperLoggerExtensionsLearn

InterceptHTTP historyWebSockets historyProxy settings

Filter: Hiding CSS, image and general binary content

#	✓	Host	Method	URL	Params	Edited	Status code	Length	MIME type	Extension	Title	Comment	TLS	IP	C
2120		https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197	
2113		https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197	
2112		https://finance.vietstock.vn	POST	/Notifications/GetTotalNotification...	✓		200	378	JSON				✓	221.132.38.197	
2105		https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197	
2104		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	40303	JSON				✓	221.132.38.197	
2103		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	39539	JSON				✓	221.132.38.197	
2102		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	2916	JSON				✓	221.132.38.197	
2101		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	2948	JSON				✓	221.132.38.197	
2100		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	18652	JSON				✓	221.132.38.197	
2098		https://csm.as.criteo.net	POST	/all?cppv=3&cpp=mefvwkxPeT...	✓		200	221					✓	182.161.73.142	
2097		https://csm.as.criteo.net	POST	/all?cppv=3&cpp=u3fu6kxPeTHL...	✓		200	221					✓	182.161.73.142	
2091		https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197	
2090		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	18652	JSON				✓	221.132.38.197	
2089		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	18652	JSON				✓	221.132.38.197	
2088		https://finance.vietstock.vn	POST	/data/gettradingresult	✓		200	18652	JSON				✓	221.132.38.197	

Request

PrettyRawHex

16JN7V9PFXp1eThibUlleTVxdkYoRlpUNtADuK5hUvg1UTVjYjhcTkdydyYvRjZSeWfjaORlVfZkUXhMTnpZJTJGHW10a21ITCuyQnHVHVBcjAaHozUWkSPUwubBpThT1SRFVLQ0poMBExV0tkeKvabVh5MSUaPA;AnonymousNotification=; dable_uid=27137602.1705163808340; lsShowLogin=true; _gid=GA1.2.221192376.1705163812; _gads=ID=6d3d0a74b50407ce;T=1705163814;RT=1705165065;S=ALNI_M200JfttR7ppsDm_rDp7doCsZANpmw; __gpl=UID=00000c85142e4eed;T=1705163814;RT=1705165065;S=ALNI_MbxvCwxiJtjb8hPcD5FYWj6kTW6A;cto_dna_bundle=Myqyff8zUTJGvjhwzThPU1L2cnFFDWhzPMNLNllnV3JTTWZJcWkTSVpYXKhQbkl1S0R3PMNEZMWrbkZ0Y1Rvb3NLZks3NPF8Q04lQnd0eWkLV3F2cU2LznNYZmhnJTNEJTNE; finance_viewedstock=SSI,User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0Accept: */*Accept-Language: en-US,en;q=0.5Accept-Encoding: gzip, deflateContent-Type: application/x-www-form-urlencoded; charset=UTF-8X-Requested-With: XMLHttpRequestContent-Length: 330Origin: https://finance.vietstock.vnReferer: https://finance.vietstock.vn/SSI/thong-ke-giao-dich.htmSec-Fetch-Dest: emptySec-Fetch-Mode: corsSec-Fetch-Site: same-originTe: trailersConnection: closeCode=SSI&OrderBy=4&OrderDirection=desc&PageIndex=1&PageSize=30&FromDate=2023-12-13&ToDate=2024-01-13&ExportType=default&Cols=TKLGD%2CTGTGD%2CVHHT%2CTGG%2CDC%2CTGPTG%2CKLGDKL%2CTGTGDKL&ExchangeID=1&__RequestVerificationToken=AmpLTSPhLRNU04lyf9ubQzVgMe3uR-UCu_aZFVmh8j_S_3xbpZ8s1D2VyWfoWbZEWZVYH10y5AxjRdoOltbpcXZ_D-YPDhVYKXlgBJLT_Ao1

0 matches

Response

PrettyRawHexRender

"ColSpan":0,"RowSpan":0},],,"Data":[{ "TradingDate":"/Date(1704992400000)\/", "Time":"/Date(1704992400000)\/", "StockCode":"SSI", "KLCPH":1499138669, "KLCPW":1491130137, "ClosePrice":33850, "PriorClosePrice":33900, "CeilingPrice":36250, "FloorPrice":31550, "TotalVol":25278801, "TotalVal":851824031550, "MarketCapital":50745843945650, "HighestPrice":34000, "LowestPrice":33350, "DiffHighLow":650, "OpenPrice":33600, "LastPrice":33850, "Price1":33600, "TotalVol1":336800, "TotalVal1":7956000000, "Price2":33850, "AvrPrice2":33698, "TotalVol2":25022000, "TotalVal2":843190000000,

Search...

0 matches

Search...

Xác định được luôn trường chúng ta cần là **“TradingDate”** và **“ClosePrice”**.

Oke vậy là trong tay em đã có những request và thông tin cần thiết để lấy được những dữ liệu mà chúng ta cần.

2, phân tích, kiểm thử

Trước tiên đến với headers của request. Để biết được những headers nào cần thiết và không cần thiết trong việc định nghĩa lại headers trong code. Em ném các request quan trọng đã bắt được sang repeater để test. Sau khi thử xóa thì em thấy được chỉ cần những headers sau là vẫn có thể nhận được response mình cần

1 x 3 x +

Send Cancel < >

Target: https/

Request

Pretty Raw Hex

```
1 GET /finance/toptrading?type=7&catID=1 HTTP/1.1
2 Host: api.vietstock.vn
3 User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0
4 Referer: https://finance.vietstock.vn/
5 Connection: close
6
7
```

Response

Pretty Raw Hex Render

```
1 HTTP/1.1 200 OK
2 Server: nginx
3 Date: Sat, 13 Jan 2024 17:15:24 GMT
4 Content-Type: application/json; charset=utf-8
5 Content-Length: 1383
6 Connection: close
7 Cache-Control: no-cache
8 Pragma: no-cache
9 Expires: -1
10 X-XSS-Protection: 1; mode=block
11 X-Content-Type-Options: nosniff
12 Strict-Transport-Security: max-age=7776000; includeSubDomains
13
14 [
  {
    "TradingDate": "2024-01-12T00:00:00" ,
    "StockCode": "S91",
    "ClosePrice": 33850.0 ,
    "Change": -50.0 ,
    "PerChange": -0.15 ,
    "TotalVol": 809647.0 ,
    "Row": 1
  },
  {
    "TradingDate": "2024-01-12T00:00:00" ,
    "StockCode": "H96",
    "ClosePrice": 37150.0 ,
    "Change": -350.0 ,
    "PerChange": -1.27 ,
    "TotalVol": 702138.0 ,
    "Row": 2
  },
  {
    "TradingDate": "2024-01-12T00:00:00" ,
    "StockCode": "S75",
    "ClosePrice": 29900.0 ,
    "Change": 400.0 ,
    "PerChange": 1.36 ,
    "TotalVol": 688476.0 ,
    "Row": 3
  },
  {
    "TradingDate": "2024-01-12T00:00:00" ,
    "StockCode": "D16",
    "ClosePrice": 26350.0 ,
    "Change": -700.0 ,
    "PerChange": -2.59 ,
    "TotalVol": 100000.0 ,
    "Row": 4
  }
]
```

0 matches 0 matches

Tiếp tục đến với request tới api lấy dữ liệu 30 ngày trong tháng. Em xóa các headers không quan trọng và thử gửi lại. Trong lúc xóa em có nhận thấy phía bên sever dùng 1 cookie duy nhất để verify request đó là cookie **__RequestVerificationToken**

1 x 3 x +

Send Cancel < >

Request

Pretty Raw Hex

```
1 POST /data/gettradingresult HTTP/1.1
2 Host: finance.vietstock.vn
3 Cookie: __RequestVerificationToken =
  oOc4lmm2k_ow0mCixlpQICHMyb0WqfsIrV9Cy4Lo967GxwlgYQ2cAYfCH2IDLrI912p6XfBXQCfTxyCIRfYPLG
  HFULU-VeR-AvcYN300BMol ;
4 User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101
  Firefox/121.0
5 Content-Type: application/x-www-form-urlencoded;
6 X-Requested-With: XMLHttpRequest
7 Content-Length: 330
8 Referer: https://finance.vietstock.vn/SSI/thong-ke-giao-dich.htm
9 Connection: close
10
11 Code=SSI&OrderBy=&OrderDirection =desc&PageIndex =1&PageSize =30&FromDate =2023-12-13 &
  ToDate =2024-01-13 &ExportType =default &Cols=
  TKGd%2CTGTGD%2CVHTT%2CTGG%2CDC%2CTGPTG%2CKLGDKL%2CGTGDKL &ExchangeID =1&
  __RequestVerificationToken =
  Amp1TSPhLRNUd4lyFsJbQzVgMe3uR-UCu_a2FVmh8j_S_G3xhpZ8sID2VyWfowb2EWZVYH10ySAXjRdoOltDpc
  XZ_D-YPbhVFXLg2ULT_Ao1
```

Response

Pretty Raw Hex Render

```
{
  "Text": "4- $ gia",
  "Css": "text-right",
  "ColSpan": 0,
  "RowSpan": 2
},
{
  "Text": "Khôp lệnh",
  "Css": "text-center",
  "ColSpan": 2,
  "RowSpan": 0
},
{
  "Text": "XL",
  "Css": "text-right",
  "ColSpan": 0,
  "RowSpan": 0
},
{
  "Text": "GT",
  "Css": "text-right",
  "ColSpan": 0,
  "RowSpan": 0
}
],
"Data": [
  {
    "TradingDate": "\Date(1704992400000)\\" ,
    "Time": "\Date(1704992400000)\\" ,
    "StockCode": "SSI",
    "KLCPHL": 1499138669 ,
    "KLCPHY": 1491130137 ,
    "ClosePrice": 33850 ,
    "PriorClosePrice": 33900 ,
    "CeilingPrice": 36250 ,
    "FloorPrice": 31550 ,
    "TotalVol": 25278801 ,
    "TotalVal": 851824031550 ,
    "MarketCapital": 50745843945650 ,
    "HighestPrice": 34000 ,
    "LowestPrice": 33350 ,
    "DiffHighLow": 650 ,
    "OpenPrice": 33600 ,
    "LastPrice": 33850 ,
    "Price1": 33600 ,
    "TotalVoll": 236800 .
  }
]
```

Nhìn vào body của request, em có thấy mục **Code=SSI** và response trả về là thông tin 30 ngày của cổ phiếu **SSI**, mục **FromDate** và mục **ToDate** để trả về dữ liệu mà mình nhập vào từ ngày nào đến ngày nào. Vậy nên em thử thay mã **SSI** bằng mã cổ phiếu khác (**HPG**) thử xem nó có trả về data của cổ phiếu này hay không.

Và bùm, em đã nhận được data dữ liệu của cổ phiếu **HPG** mà em vừa thay


```

import requests
from datetime import *
import re
# redefine Headers
headers = {
    "Host": "api.vietstock.vn",
    "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0",
    "Referer": "https://finance.vietstock.vn/",
    "Connection": "close"
}
# api
url = "https://api.vietstock.vn/finance/toptrading?type=7&catID=1"

proxies = { 'http': 'http://127.0.0.1:8080', 'https': 'http://127.0.0.1:8080' }

# send request and use burpsuite as proxy to debug
r = requests.get(url, headers=headers, proxies=proxies, verify=False)

data = r.json()
stockCode = []
# use only StockCode
for i in data:
    print(i)

```

Data nhận về

```

{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'SSI', 'ClosePrice': 33850.0, 'Change': -50.0,
 'PerChange': -0.15, 'TotalVol': 809647.0, 'Row': 1}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'HPG', 'ClosePrice': 27150.0, 'Change': -350.0,
 'PerChange': -1.27, 'TotalVol': 702138.0, 'Row': 2}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'STB', 'ClosePrice': 29900.0, 'Change': 400.0,
 'PerChange': 1.36, 'TotalVol': 688476.0, 'Row': 3}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'DIG', 'ClosePrice': 26350.0, 'Change': -700.0,
 'PerChange': -2.59, 'TotalVol': 601459.0, 'Row': 4}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'VND', 'ClosePrice': 21800.0, 'Change': -150.0,
 'PerChange': -0.68, 'TotalVol': 529948.0, 'Row': 5}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'SHB', 'ClosePrice': 12300.0, 'Change': 200.0,
 'PerChange': 1.65, 'TotalVol': 517183.0, 'Row': 6}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'MBB', 'ClosePrice': 21350.0, 'Change': 700.0,
 'PerChange': 3.39, 'TotalVol': 504181.0, 'Row': 7}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'GEX', 'ClosePrice': 21300.0, 'Change': -800.0,
 'PerChange': -3.62, 'TotalVol': 468640.0, 'Row': 8}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'VIX', 'ClosePrice': 16950.0, 'Change': -250.0,
 'PerChange': -1.45, 'TotalVol': 451409.0, 'Row': 9}
{'TradingDate': '2024-01-12T00:00:00', 'StockCode': 'ACB', 'ClosePrice': 25800.0, 'Change': 550.0,
 'PerChange': 2.18, 'TotalVol': 369079.0, 'Row': 10}

```

Lấy ra chỉ mã cổ phiếu



```
data = r.json()  
stockCode = []  
# use only StockCode  
for i in data:  
    stockCode.append(i['StockCode'])
```

Debugger bằng burpsuite cũng đã nhận được request.

Filter: Hiding CSS, image and general binary content

#	Host	Method	URL	Params	Edited	Status code	Length	MIME type	Extension	Title	Comment	TLS	IP
3108	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3107	https://finance.vietstock.vn	POST	/Notifications/GetTotalNotification...	✓		200	378	JSON				✓	221.132.38.197
3100	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3093	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3086	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3085	https://api.vietstock.vn	GET	/finance/toptrading?type=7&catID=1	✓		200	1712	JSON				✓	221.132.38.195
3078	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3077	https://api.vietstock.vn	GET	/finance/toptrading?type=7&catID=1	✓		200	1712	JSON				✓	221.132.38.195
3076	https://finance.vietstock.vn	POST	/Notifications/GetTotalNotification...	✓		200	378	JSON				✓	221.132.38.197
3069	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3062	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3055	https://finance.vietstock.vn	POST	/data/getmarketprice	✓		200	1308	JSON				✓	221.132.38.197
3054	http://detectportal.firefox.com	GET	/success.txt?ipv4	✓		200	235	text	txt				34.107.221.82
3053	http://detectportal.firefox.com	GET	/success.txt?ipv6	✓		200	235	text	txt				34.107.221.82
3052	http://detectportal.firefox.com	GET	/canonical.html			200	317	XML	html				34.107.221.82

Request

PrettyRawHex

in

1

GET /finance/toptrading?type=7&catID=1 HTTP/1.1

2

User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0

3

Accept-Encoding: gzip, deflate

4

Accept: */*

5

Connection: close

6

Host: api.vietstock.vn

7

Referer: https://finance.vietstock.vn/

8

9

Response

PrettyRawHexRender

1

HTTP/1.1 200 OK

2

Server: nginx

3

Date: Sat, 13 Jan 2024 17:49:16 GMT

4

Content-Type: application/json

5

Content-Length: 1383

6

Connection: close

7

Cache-Control: no-cache

8

Pragma: no-cache

9

Expires: -1

10

X-XSS-Protection: 1; mode=block

11

X-Content-Type-Options: nosniff

12

Strict-Transport-Security: max-age=7776000; includeSubDomains

13

14

[

{

"TradingDate": "2024-01-12T00:00:00",

"StockCode": "SSI",

"ClosePrice": 33850.0,

"Change": -50.0,

"PerChange": -0.15,

"TotalVol": 809647.0,

"Row": 1

}

{

"TradingDate": "2024-01-12T00:00:00",

"StockCode": "HPG",

"ClosePrice": 27150.0,

"Change": -350.0,

"PerChange": -1.27,

"TotalVol": 702138.0,

"Row": 2

}

]

Tiếp tục gửi request Post với các tham số thay đổi như em đã nói ở trên. Vì dữ liệu trả về là json và ngày trong json đã được parse nên em phải covert nó lại theo đúng định dạng. Sau khi chỉnh sửa lại các tham số và hoàn thiện quá trình tự động. Em đã hoàn thiện code và nhận được data như yêu cầu

```

define cookie to verify request
cookie = {"_RequestVerificationToken": "o0c41mm2k_ow0mCix1pQICHMyb0WqfsIrV9Cy4L0967Gxw1gYq2CAYfCH2IDLrI91ZpGxkBCfTxyCirfYPLGHUfViU-Ver-AvcYW3008Mo1"}

# redefine headers
headers = {"Host": "finance.vietstock.vn",
"User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0",
"Content-Type": "application/x-www-form-urlencoded",
"X-Requested-With": "XMLHttpRequest",
"Content-Length": "330",
"Referer": "https://finance.vietstock.vn/SSI/thong-ke-giao-dich.htm",
"Connection": "close"}

}

# api which to send request
url = "https://finance.vietstock.vn/data/gettradingresult"

# define date
today = date.today()
before = today - timedelta(days=30)

# format Date
def dateFormat(d):
    day = re.sub(r'\D', '', d)
    day = int(day)/1000
    day = datetime.fromtimestamp(day)
    return day

for i in stockCode:
    print("Stock Code: ", i)
    data = "Code={code}&OrderBy=&OrderDirection=desc&PageIndex=1&PageSize=30&FromDate={past}&ToDate={now}&ExportType=default&Cols=TKLGD%2CTGTGD%2CVHTT%2CTGG%2CDC%2CTGPTG%2CKLGDKL%2CGTGDKL&ExchangeID=1&__RequestVerificationToken=AmplTSPhLRNUD4lyFsJbQzVgMe3uR-UCu_aZvFmhj_S_G3xbpZ8s1D2VyWfowbZEWZYH10ySAxjRdo01tdpcX_D-YPDHVKXLgEJLT_Ao1".format(code = i, past = before, now = today)
    r = requests.post(url, headers=headers, cookies= cookie, data=data, proxies=proxies, verify=False)
    # parse json date to normal date
    codeInfor = r.json()
    for i in codeInfor["Data"]:
        print(dateFormat(i["TradingDate"]), i["ClosePrice"], sep= ' : ')

```

Data receive:

```
warnings.warn(  
Stock Code: SSI  
C:\Users\ADMIN\AppData\Local\Programs\Python\Python3  
dding certificate verification is strongly advised.  
warnings.warn(  
2024-01-12 00:00:00 : 33850  
2024-01-11 00:00:00 : 33900  
2024-01-10 00:00:00 : 33700  
2024-01-09 00:00:00 : 34000  
2024-01-08 00:00:00 : 34100  
2024-01-05 00:00:00 : 34150  
2024-01-04 00:00:00 : 33600  
2024-01-03 00:00:00 : 32950  
2024-01-02 00:00:00 : 32500  
2023-12-29 00:00:00 : 32800  
2023-12-28 00:00:00 : 33000  
2023-12-27 00:00:00 : 32800  
2023-12-26 00:00:00 : 32750  
2023-12-25 00:00:00 : 32550  
2023-12-22 00:00:00 : 32350  
2023-12-21 00:00:00 : 32200  
2023-12-20 00:00:00 : 32300  
2023-12-19 00:00:00 : 32150  
2023-12-18 00:00:00 : 31700  
2023-12-15 00:00:00 : 31800  
Stock Code: HPG  
C:\Users\ADMIN\AppData\Local\Programs\Python\Python3  
dding certificate verification is strongly advised.  
warnings.warn(  
2024-01-12 00:00:00 : 27150  
2024-01-11 00:00:00 : 27500  
2024-01-10 00:00:00 : 27350  
2024-01-09 00:00:00 : 27700  
2024-01-08 00:00:00 : 27850
```

Yeah task done ! Và đây là toàn bộ source code của em

```

import requests
from datetime import *
import re
# redefine Headers
headers = {
    "Host": "api.vietstock.vn",
    "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0",
    "Referer": "https://finance.vietstock.vn/",
    "Connection": "close"
}
# api
url = "https://api.vietstock.vn/finance/toptrading?type=7&catID=1"

proxies = { 'http': 'http://127.0.0.1:8080', 'https': 'http://127.0.0.1:8080' }

# send request and use burpsuite as proxy to debug
r = requests.get(url, headers=headers, proxies=proxies, verify=False)

data = r.json()
stockCode = []
# use only StockCode
for i in data:
    stockCode.append(i['StockCode'])

# define cookie to verify request
cookie = {"__RequestVerificationToken": "o0c41mm2k_ow0mCIX1pQICHMyb0WqfsIrV9Cy4Lo967Gxw1gyQ2cAYfCH2IDLrI91ZpGXfBXQCfTxyCIRfYP1GHFUu-Ver-AvcYW300BMol"}
# redefine headers
headers = {"Host": "finance.vietstock.vn",
    "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:121.0) Gecko/20100101 Firefox/121.0",
    "Content-Type": "application/x-www-form-urlencoded",
    "X-Requested-With": "XMLHttpRequest",
    "Content-Length": "330",
    "Referer": "https://finance.vietstock.vn/SSI/thong-ke-giao-dich.htm",
    "Connection": "close"
}
# api which to send request
url = "https://finance.vietstock.vn/data/gettradingresult"
# define date
today = date.today()
before = today - timedelta(days=30)
# format Date
def dateFomat(d):
    day = re.sub(r'\D','', d)
    day = int(day)/1000
    day = datetime.fromtimestamp(day)
    return day

for i in stockCode:
    print("Stock Code: ", i)
    data = "Code={code}&OrderBy=&OrderDirection=desc&PageIndex=1&PageSize=30&FromDate={past}&ToDate={now}&ExportType=default&Cols=TKLGD%2CTGTGD%2CVHT%2CTGG%2CDC%2CTGPTG%2CKLGD%2CKLGD%2CKLGD&ExchangeID=1&__RequestVerificationToken=AmplTSPH LRNUd4lyFsJbQzVgMe3uR-UCu_aZFVmh8j_S_G3xbpZ8s1D2VyWfowbZEWZYH10ySAxjRdo01tdpcXZ_D-YPDhVKXLgEJLT_Aol".format(code = i, past = before, now = today)
    r = requests.post(url, headers=headers, cookies= cookie, data=data, proxies=proxies, verify=False)
    # parse json date to normal date
    codeInfor = r.json()
    for i in codeInfor["Data"]:
        print(dateFomat(i["TradingDate"]), i["ClosePrice"], sep=' : ')

```

Additional: bởi vì em đọc thiếu đề là còn phải lưu dữ liệu dưới file .csv nên em sẽ update lại code. Khi quan sát dữ liệu em nhận thấy trường TradingDate của từng mã cổ phiếu đều giống nhau. Nên em định nghĩa riêng cho nó 1 phần ở từ điển (key Date). Và sử dụng mã cổ phiếu làm key và value của nó là giá cổ phiếu theo ngày được lưu theo mảng. Sau đó sử dụng thư viện csv để tạo file theo dictionary. Với các row là key của dictionary và giá trị của các cột là value.

```

csvData = {"Date": []}

for i in stockCode:
    csvData[i] = []
    print("Stock Code: ", i)
    data = "Code={code}&OrderBy=&OrderDirection=desc&PageIndex=1&PageSize=30&FromDate={past}&ToDate={now}&ExportType=default&Cols=TKLGD%2CTGTGD%2CVHTT%2CTGG%2CDC%2CTGPTG%2CKLGDKL%2CGTGDKL&ExchangeID=1&__RequestVerificationToken=AmplTSPH
LRNUD4lyFsJbQzVgMe3uR-UCu_aZFVmH8j_S_G3xbpZ8s1D2VyWfowbZEWZVYH10ySAXjRdo01tdpcXZ_D-YPDhVKXLgEJLT_Ao1".format(code = i, past = before,
now = today)
    r = requests.post(url, headers=headers, cookies= cookie, data=data, proxies=proxies, verify=False)
    # parse json date to normal date
    codeInfor = r.json()

    for j in codeInfor["Data"]:
        # print(dateFomat(j["TradingDate"]), j["ClosePrice"], sep=' : ')
        if len(csvData["Date"]) < 30:
            # advoid identical data
            csvData["Date"].append(dateFomat(j["TradingDate"]))
            csvData[i].append(j["ClosePrice"])

with open("csvData.csv", "w") as outfile:
    writer = csv.writer(outfile)
    # use dictionary key for row
    writer.writerow(csvData.keys())
    writer.writerows(zip(*csvData.values()))

```

Và output em nhận được

A1													
Date													
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Date	SSI	HPG	STB	DIG	VND	SHB	MBB	GEX	VIX	ACB		
2													
3	1/12/2024 0:00	33850	27150	29900	26350	21800	12300	21350	21300	16950	25800		
4													
5	1/11/2024 0:00	33900	27500	29500	27050	21950	12100	20650	22100	17200	25250		
6													
7	1/10/2024 0:00	33700	27350	29450	26800	21750	12000	20900	22150	16700	25250		
8													
9	1/9/2024 0:00	34000	27700	29200	27400	22150	11700	20850	22450	16950	25000		
10													
11	1/8/2024 0:00	34100	27850	29200	27750	22250	11700	20700	22250	16900	25350		
12													
13	1/5/2024 0:00	34150	27800	29300	27150	22300	11550	20500	22150	16750	25450		
14													
15	1/4/2024 0:00	33600	27750	28600	26500	22300	11450	20100	22250	16750	25300		
16													
17	1/3/2024 0:00	32950	27800	28500	26300	22350	11250	19150	22550	16650	25000		
18													
19	1/2/2024 0:00	32500	27450	27750	26000	22150	10950	18850	22150	16400	24450		
20													
21	12/29/2023 0:00	32800	27950	27950	26750	22250	10800	18650	23600	17100	23900		
22													
23	12/28/2023 0:00	33000	27950	27600	26700	22150	10800	18550	23400	17100	23750		
24													
25	12/27/2023 0:00	32800	27750	27400	26550	22000	10750	18400	23350	17100	23300		
26													
27	12/26/2023 0:00	32750	27800	27500	26750	21850	10700	18450	23450	17100	23250		
28													
29	12/25/2023 0:00	32550	27450	27500	26600	21700	10750	18400	23200	16900	23350		
30													
31	12/22/2023 0:00	32350	27050	27200	25800	21500	10700	18300	22700	16700	23200		
32													
33	12/21/2023 0:00	32200	27200	26600	26150	21750	10650	18150	22600	16850	23250		
34													
35	12/20/2023 0:00	32300	27200	26950	25800	21950	10700	18150	22600	16950	23200		
36													
37	12/19/2023 0:00	32150	27200	26600	25850	21800	10700	18150	22400	16800	23050		
38													
39	12/18/2023 0:00	31700	26600	26300	25450	21550	10700	18000	21950	16400	22800		
40													
41	12/15/2023 0:00	31800	26800	27100	25200	21800	10800	18050	22250	16500	22750		
42													
43													

Yeah task finally done ! và đây là toàn bộ source code của em.

