

# 分析 AJAX 传递的 JSON 获取数据

## 获取动态数据两种的思路

1. 分析页面请求
2. 利用 selenium 模拟浏览器行为或其他抓包工具直接获取

效率最高的就是分析出请求数据的 URL 一般都可以。

## 案例：获取豆瓣电影的信息

### 需求

[https://movie.douban.com/explore#!type=movie&tag=%E7%83%AD%E9%97%A8&sort=recommend&page\\_limit=20&page\\_start=0](https://movie.douban.com/explore#!type=movie&tag=%E7%83%AD%E9%97%A8&sort=recommend&page_limit=20&page_start=0)

爬取电影的名称, url, 评分, 封面图片

### 审查元素

打 开  
[https://movie.douban.com/explore#!type=movie&tag=%E7%83%AD%E9%97%A8&sort=recommend&page\\_limit=20&page\\_start=0](https://movie.douban.com/explore#!type=movie&tag=%E7%83%AD%E9%97%A8&sort=recommend&page_limit=20&page_start=0)



选择电影标签

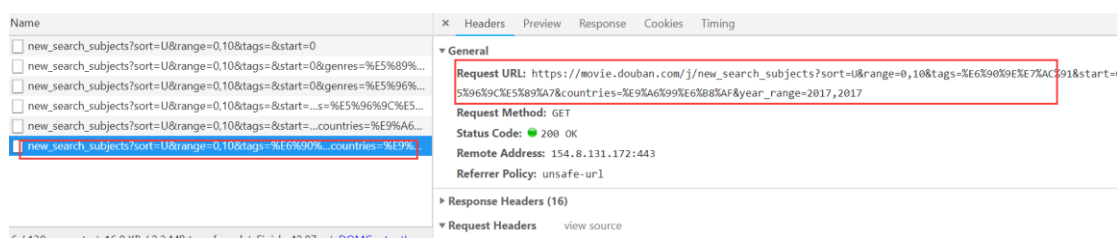
全部形式 电影 电视剧 综艺 动漫 纪录片 短片

全部类型 剧情 喜剧 动作 爱情 科幻 动画 悬疑 惊悚 恐怖 犯罪 同性  
音乐 歌舞 传记 历史 战争 西部 奇幻 冒险 灾难 武侠 情色

全部地区 中国大陆 美国 香港 台湾 日本 韩国 英国 法国 德国 意大利  
西班牙 印度 泰国 俄罗斯 伊朗 加拿大 澳大利亚 爱尔兰 瑞典 巴西 丹麦

全部年代 2018 2017 2010年代 2000年代 90年代 80年代 70年代 60年代 更早

全部特色 经典 青春 文艺 搞笑 励志 魔幻 感人 女性 黑帮 + 自定义标签



### Request URL:

[https://movie.douban.com/j/new\\_search\\_subjects?sort=U&range=0,10&tags=%E6%90%9E%E7%AC%91&start=0&genres=%E5%96%9C%E5%89%A7&countries=%E9%A6%99%E6%B8%AF&year\\_range=2017,2017](https://movie.douban.com/j/new_search_subjects?sort=U&range=0,10&tags=%E6%90%9E%E7%AC%91&start=0&genres=%E5%96%9C%E5%89%A7&countries=%E9%A6%99%E6%B8%AF&year_range=2017,2017)

### 查看 Query string

#### ▼ Query String Parameters view

sort: U  
range: 0,10  
tags: 搞笑  
start: 0  
genres: 喜剧  
countries: 香港  
year\_range: 2017,2017

# sort:T 排序

# range:0,10 评分范围

# tags:电影,剧情 标签 (电影类型)

#genres:类型

#countries: 国家和地区

# start:0 数据的起始位置 从 0 开始 每页 20 条数据, 取值 0, 20, 40...

# year\_range:年份

## 代码实现

```
from urllib import request
from urllib import parse
import os
import json

def load_page(url,page):
    """发送请求加载页面"""
    headers = {
        'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/68.0.3440.75 Safari/537.36'}
    # 构造请求
    req = request.Request(url, headers=headers)
    try:
        # 发送请求
        response = request.urlopen(req)
        if response.getcode() == 200:
            json_data = response.read().decode()
            data = json.loads(json_data)['data']
            print(type(data))
            print(len(data))
            for item in data:
                print('directors:',''.join(item['directors']))
                print('rate:',item['rate'])
                print('star:',item['star'])
                print('title:', item['title'])
                print('url:', item['url'])
                print('casts:',''.join(item['casts']))
                print('cover:', item['cover'])
                print('='*60)

            if not os.path.exists('./data'):
                os.mkdir('./data')
```

```
        filename = 'douban_'+str(page)+'.json'
        file_path = os.path.join('./data', filename)
        with open(file_path, 'w', encoding='utf-8') as fp:
            fp.write(json_data)
    else:
        print('请求出错')
except Exception as err:
    print(err)

def spider(page):
    """爬虫程序"""
    base_url = 'https://movie.douban.com/j/new_search_subjects?'
    # 构造关键字 请求需要带上的参数
    key_words = {
        "sort": "T", # 排序的方式
        "range": "0,10", # 电影评分的范围
        "tags": "电影,剧情,美国", # 检索的标签
        #"playable": "1", # 是否可以播放
        "start": (page-1)*20, # 检索的开始位置 (这里可以去改变的 从 0 开始 一个电
影代表的是一条数据)
        "genres": "喜剧",#类型
        "countries": "中国大陆",# 国家地区
    }
    # 编码
    key_words = parse.urlencode(key_words)
    url = base_url + key_words
    # 调用函数
    load_page(url,page)

if __name__ == "__main__":
    for i in range(3):
        print('page:',i+1)
        spider(i+1)
```

## 案例：获取 A 股上市公司信息

### 需求

爬取上海证券交易所网站，获取 A 股上市公司信息，包括公司代码，公司简称，A 股代码，

A 股简称以及 A 股总资本和 A 股流通资本这几项

## 审查元素

打开 <http://www.sse.com.cn/assortment/stock/list/share/>

查看 A 股上市公司信息

The screenshot shows the SSE website with a QR code on the left and a table of A-share companies on the right. The table has columns for company code, A-share code, A-share name, A-share listing date, A-share total capital, and A-share circulating capital. The first row is for 浦发银行 (Pudong Bank).

公司代码	A股代码	A股简称	A股上市日期	A股总股本	A股流通
600000	600000	浦发银行	1999-11-10	2935208.04	2935208.04
600004	600004	白云机场	2003-04-28	206932.05	206932.05
600006	600006	东风汽车	1999-07-27	200000.00	200000.00
600007	600007	中国国贸	1999-03-12	100728.25	100728.25
600008	600008	首创股份	2000-04-27	482061.41	482061.41
600009	600009	上海机场	1998-02-18	192695.84	192695.84
600010	600010	包钢股份	2001-03-09	4558503.26	4558503.26
600011	600011	华能国际	2001-12-06	1520038.34	1520038.34
600012	600012	皖通高速	2003-01-07	165861.00	165861.00
600015	600015	华夏银行	2003-09-12	1282268.67	1282268.67
600016	600016	民生银行	2000-12-19	4378241.85	4378241.85

The developer tools show the HTML structure of the table. The selected element is a table with class 'table search\_'. The table has a tbody with a tr containing the first row of data.

```

<div class="tab-pane js_tableT01 active" id="tableData">
  <div class="sse_table title2" style="display: none;">
    <div class="table-responsive sse_table_T01 tdclickable">
      <table class="table search_">
        <tbody>
          <tr>
            <td>浦发银行</td>
            <td>600000</td>
            <td>浦发银行</td>
            <td>1999-11-10</td>
            <td>2935208.04</td>
            <td>2935208.04</td>
          </tr>
        </tbody>
      </table>
    </div>
  </div>
</div>

```

发现上市公司条目在 class 为 table-responsive sse\_table\_T01 tdclickable 的 div 下的 table 中

查看网页源码，发现该 div 为空，我们用普通的爬取方法是获取不到数据的

```

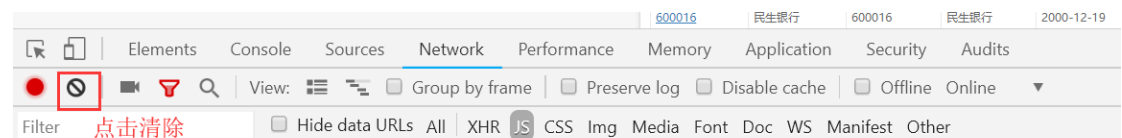
<div class="sse_table title2" style="display: none;">
  <div class="table-responsive sse_table_T01 tdclickable">
    <table class="table search_">
      <script type="text/javascript">
        </script>
    </table>
  </div>
</div>

```

这是因为这个网页是使用了 AJAX 技术的**动态网页**。它可以通过与服务器进行少量数据交换而在不重新加载整个网页的情况下对部分网页进行异步更新。

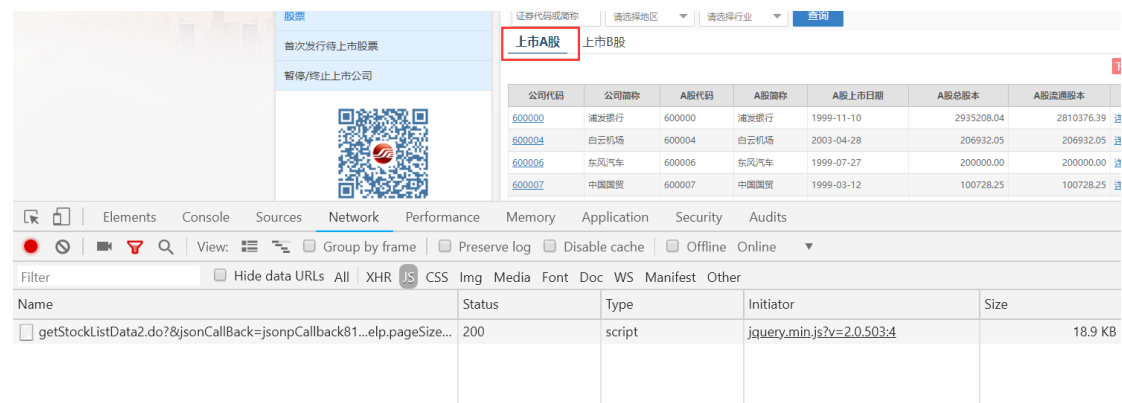
## 分析请求数据的 URL

打开开发者工具，选 Network 进行 JS 分析，点击清除，更方便定位

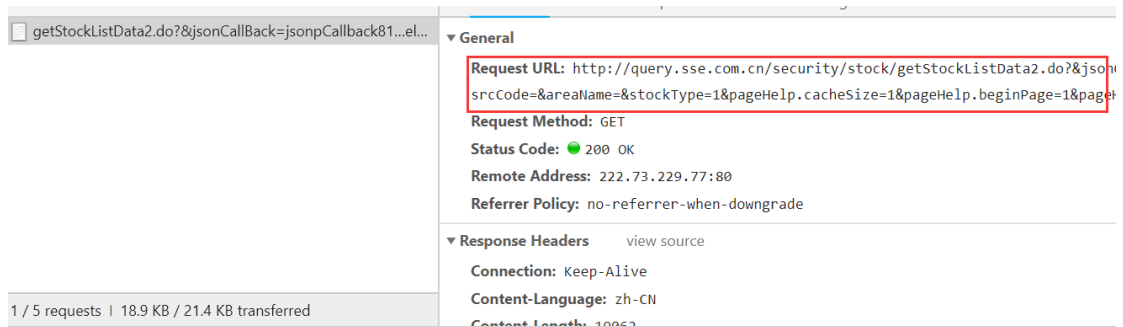


Recording network activity...  
Perform a request or hit **F5** to record the reload

点击网页中上市 A 股，进行加载刷新



出现了一个条目，就是我们要找的数据的 url，（有的网站会出现一堆的 js 条目，就需要去分析一下了）



浏览器里打开目标验证一下有没有我们需要的数据

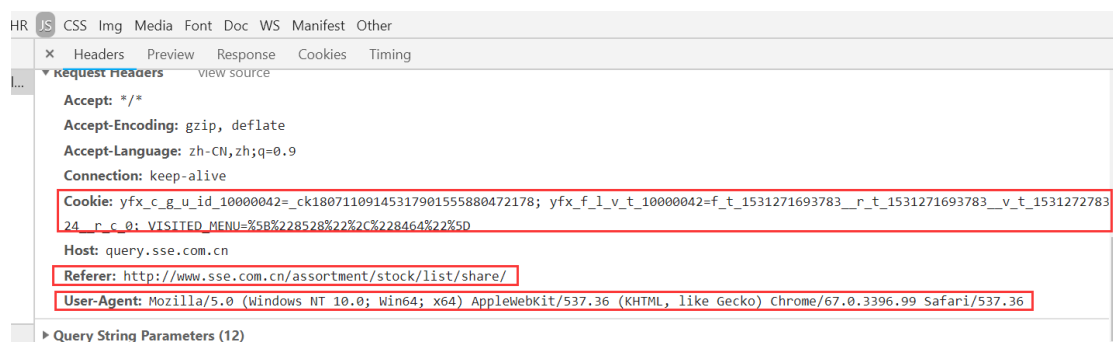
Error 403: SRVE0190E: 找不到文件: /error/error\_cn.jsp

403 码表示我们没有权限浏览目标地址。这是网站的自我保护行为。

## JSON 数据的获取

思路：修改 Request-Headers 中 Cookie, User-Agent, Referer 等信息，让爬虫模拟人的操作

需要修改的内容可以在 Headers 中查看：



## 代码实现

```
import requests
```

```
url='http://query.sse.com.cn/security/stock/getStockListData2.do?&jso
```

```

nCallBack=jsonpCallback81793&isPagination=true&stockCode=&csrcCode=&a
reaName=&stockType=1&pageHelp.cacheSize=1&pageHelp.beginPage=1&pageHe
lp.pageSize=25&pageHelp.pageNo=1&_=1531271694633'
headers={'Cookie':'yfx_c_g_u_id_10000042=_ck1807110914531790155588047
2178;
yfx_f_l_v_t_10000042=f_t_1531271693783__r_t_1531271693783__v_t_153127
2783524__r_c_0; VISITED_MENU=%5B%228528%22%2C%228464%22%5D',
'User-Agent':'Mozilla/5.0 (Windows NT 10.0; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/67.0.3396.99
Safari/537.36',
'Referer':'http://www.sse.com.cn/assortment/stock/list/share/'
}
response=requests.get(url,headers=headers)
print(response.text)
a='{"content":'+response.text[19:-1]+'}' #去除开都的
jsonpCallback81793(和结尾的)
b=json.loads(a)
print(b)

```

## 结果

```

jsonpCallback81793({"areaName":"","csrcCode":"","downloadFileName":null,"execLS
tream":null,"jsonCallBack":"jsonpCallback81793","pageHelp":{"beginPage":1,"cache
Size":1,"data":[{"NUM":"1","totalFlowShares":"2810376.39","endDate":"2018-07-10"
,"LISTING_DATE":"1999-11-10","SECURITY_CODE_A":"600000","COMPANY_CODE":"
600000","SECURITY_CODE_B":"-","END_SHARE_CODE":"-","SECURITY_ABBR_A":" 浦
发      银      行      ","COMPANY_ABBR":"      浦      发      银      行
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"2935208.
04","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"2","totalFlowSh
ares":"206932.05","endDate":"2018-07-10","LISTING_DATE":"2003-04-28","SECURIT
Y_CODE_A":"600004","COMPANY_CODE":"600004","SECURITY_CODE_B":"-","END_
SHARE_CODE":"-","SECURITY_ABBR_A":" 白云机场 ","COMPANY_ABBR":" 白云机场

```



```
,"SECURITY_ABBR_B": "-", "END_SHARE_MAIN_DEPART": "-", "totalShares": "206932.05", "END_SHARE_VICE_DEPART": "-", "CHANGE_DATE": "-"}, {"NUM": "3", "totalFlowShares": "200000.00", "endDate": "2018-07-10", "LISTING_DATE": "1999-07-27", "SECURITY_CODE_A": "600006", "COMPANY_CODE": "600006", "SECURITY_CODE_B": "-", "END_SHARE_CODE": "-", "SECURITY_ABBR_A": "东风汽车", "COMPANY_ABBR": "东风汽车"}, {"SECURITY_ABBR_B": "-", "END_SHARE_MAIN_DEPART": "-", "totalShares": "200000.00", "END_SHARE_VICE_DEPART": "-", "CHANGE_DATE": "-"}, {"NUM": "4", "totalFlowShares": "100728.25", "endDate": "2018-07-10", "LISTING_DATE": "1999-03-12", "SECURITY_CODE_A": "600007", "COMPANY_CODE": "600007", "SECURITY_CODE_B": "-", "END_SHARE_CODE": "-", "SECURITY_ABBR_A": "中国国贸", "COMPANY_ABBR": "中国国贸"}, {"SECURITY_ABBR_B": "-", "END_SHARE_MAIN_DEPART": "-", "totalShares": "100728.25", "END_SHARE_VICE_DEPART": "-", "CHANGE_DATE": "-"}, {"NUM": "5", "totalFlowShares": "482061.41", "endDate": "2018-07-10", "LISTING_DATE": "2000-04-27", "SECURITY_CODE_A": "600008", "COMPANY_CODE": "600008", "SECURITY_CODE_B": "-", "END_SHARE_CODE": "-", "SECURITY_ABBR_A": "首创股份", "COMPANY_ABBR": "首创股份"}, {"SECURITY_ABBR_B": "-", "END_SHARE_MAIN_DEPART": "-", "totalShares": "482061.41", "END_SHARE_VICE_DEPART": "-", "CHANGE_DATE": "-"}, {"NUM": "6", "totalFlowShares": "109347.64", "endDate": "2018-07-10", "LISTING_DATE": "1998-02-18", "SECURITY_CODE_A": "600009", "COMPANY_CODE": "600009", "SECURITY_CODE_B": "-", "END_SHARE_CODE": "-", "SECURITY_ABBR_A": "上海机场", "COMPANY_ABBR": "上海机场"}, {"SECURITY_ABBR_B": "-", "END_SHARE_MAIN_DEPART": "-", "totalShares": "192695.84", "END_SHARE_VICE_DEPART": "-", "CHANGE_DATE": "-"}, {"NUM": "7", "totalFlowShar
```

```
es":"3167721.16","endDate":"2018-07-10","LISTING_DATE":"2001-03-09","SECURIT
Y_CODE_A":"600010","COMPANY_CODE":"600010","SECURITY_CODE_B":"-","END_
SHARE_CODE":"-","SECURITY_ABBR_A":" 包 钢 股 份 ","COMPANY_ABBR":" 包 钢 股 份
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"4558503.
26","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"8","totalFlowSh
ares":"1520038.34","endDate":"2018-07-10","LISTING_DATE":"2001-12-06","SECURI
TY_CODE_A":"600011","COMPANY_CODE":"600011","SECURITY_CODE_B":"-","END
_SHARE_CODE":"-","SECURITY_ABBR_A":" 华 能 国 际 ","COMPANY_ABBR":" 华 能 国 际
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"1520038.
34","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"9","totalFlowSh
ares":"165861.00","endDate":"2018-07-10","LISTING_DATE":"2003-01-07","SECURIT
Y_CODE_A":"600012","COMPANY_CODE":"600012","SECURITY_CODE_B":"-","END_
SHARE_CODE":"-","SECURITY_ABBR_A":" 皖 通 高 速 ","COMPANY_ABBR":" 皖 通 高 速
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"165861.0
0","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"10","totalFlowSh
ares":"1282268.67","endDate":"2018-07-10","LISTING_DATE":"2003-09-12","SECURI
TY_CODE_A":"600015","COMPANY_CODE":"600015","SECURITY_CODE_B":"-","END
_SHARE_CODE":"-","SECURITY_ABBR_A":" 华 夏 银 行 ","COMPANY_ABBR":" 华 夏 银 行
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"1282268.
67","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"11","totalFlowS
hares":"4378241.85","endDate":"2018-07-10","LISTING_DATE":"2000-12-19","SECU
RITY_CODE_A":"600016","COMPANY_CODE":"600016","SECURITY_CODE_B":"-","EN
```

## python 之——动态数据的获取

```
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"224737.18","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"16","totalFlowShares":"213973.93","endDate":"2018-07-10","LISTING_DATE":"2003-10-29","SECURITY_CODE_A":"600021","COMPANY_CODE":"600021","SECURITY_CODE_B":"-","END_SHARE_CODE":"-","SECURITY_ABBR_A":"上海电力","COMPANY_ABBR":"上海电力"},{"SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"240965.71","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"17","totalFlowShares":"1068861.31","endDate":"2018-07-10","LISTING_DATE":"2004-06-29","SECURITY_CODE_A":"600022","COMPANY_CODE":"600022","SECURITY_CODE_B":"-","END_SHARE_CODE":"-","SECURITY_ABBR_A":"山东钢铁","COMPANY_ABBR":"山东钢铁"},{"SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"1094654.96","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"18","totalFlowShares":"1360069.00","endDate":"2018-07-10","LISTING_DATE":"2013-12-19","SECURITY_CODE_A":"600023","COMPANY_CODE":"600023","SECURITY_CODE_B":"-","END_SHARE_CODE":"-","SECURITY_ABBR_A":"浙能电力","COMPANY_ABBR":"浙能电力"},{"SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"1360069.00","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"19","totalFlowShares":"180000.00","endDate":"2018-07-10","LISTING_DATE":"2017-12-15","SECURITY_CODE_A":"600025","COMPANY_CODE":"600025","SECURITY_CODE_B":"-","END_SHARE_CODE":"-","SECURITY_ABBR_A":"华能水电","COMPANY_ABBR":"华能水电"},{"SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"180000.00","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"20","totalFlowS
```

hares": "403203.29", "endDate": "2018-07-10", "LISTING\_DATE": "2002-05-23", "SECURITY\_CODE\_A": "600026", "COMPANY\_CODE": "600026", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "中远海能", "COMPANY\_ABBR": "中远海能", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "403203.29", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-"}, {"NUM": "21", "totalFlowShares": "850183.93", "endDate": "2018-07-10", "LISTING\_DATE": "2005-02-03", "SECURITY\_CODE\_A": "600027", "COMPANY\_CODE": "600027", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "华电国际", "COMPANY\_ABBR": "华电国际", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "986297.67", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-"}, {"NUM": "22", "totalFlowShares": "12107120.96", "endDate": "2018-07-10", "LISTING\_DATE": "2001-08-08", "SECURITY\_CODE\_A": "600028", "COMPANY\_CODE": "600028", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "中国石化", "COMPANY\_ABBR": "中国石化", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "12107120.96", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-"}, {"NUM": "23", "totalFlowShares": "1008817.33", "endDate": "2018-07-10", "LISTING\_DATE": "2003-07-25", "SECURITY\_CODE\_A": "600029", "COMPANY\_CODE": "600029", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "南方航空", "COMPANY\_ABBR": "南方航空", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "1008817.33", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-"}, {"NUM": "24", "totalFlowShares": "1209298.94", "endDate": "2018-07-10", "LISTING\_DATE": "2003-01-06", "SECURITY\_CODE\_A": "600030", "COMPANY\_CODE": "600030", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "浦发银行", "COMPANY\_ABBR": "浦发银行", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "1209298.94", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-"}]

D\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"中信证券","COMPANY\_ABBR":"中信证券  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"1211690.  
84","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"25","totalFlowS  
hares":"770011.75","endDate":"2018-07-10","LISTING\_DATE":"2003-07-03","SECURI  
TY\_CODE\_A":"600031","COMPANY\_CODE":"600031","SECURITY\_CODE\_B":"-","END  
\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"三一重工","COMPANY\_ABBR":"三一重工  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"773862.1  
6","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"}], "endDate":null, "endPage":  
null, "objectResult":null, "pageCount":58, "pageNo":1, "pageSize":25, "searchDate":null  
, "sort":null, "startDate":null, "total":1426}, "result":[{"NUM":"1", "totalFlowShares":"281  
0376.39", "endDate":"2018-07-10", "LISTING\_DATE":"1999-11-10", "SECURITY\_CODE  
\_A":"600000", "COMPANY\_CODE":"600000", "SECURITY\_CODE\_B":"-","END\_SHARE\_C  
ODE":"-","SECURITY\_ABBR\_A":"浦发银行","COMPANY\_ABBR":"浦发银行  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"2935208.  
04","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"}, {"NUM":"2", "totalFlowSh  
ares":"206932.05", "endDate":"2018-07-10", "LISTING\_DATE":"2003-04-28", "SECURIT  
Y\_CODE\_A":"600004", "COMPANY\_CODE":"600004", "SECURITY\_CODE\_B":"-","END\_  
SHARE\_CODE":"-","SECURITY\_ABBR\_A":"白云机场","COMPANY\_ABBR":"白云机场  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"206932.0  
5", "END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"}, {"NUM":"3", "totalFlowShar  
es":"200000.00", "endDate":"2018-07-10", "LISTING\_DATE":"1999-07-27", "SECURITY  
\_CODE\_A":"600006", "COMPANY\_CODE":"600006", "SECURITY\_CODE\_B":"-","END\_S

HARE\_CODE": "-", "SECURITY\_ABBR\_A": "东风汽车", "COMPANY\_ABBR": "东风汽车", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "200000.0", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-", {"NUM": "4", "totalFlowShares": "100728.25", "endDate": "2018-07-10", "LISTING\_DATE": "1999-03-12", "SECURITY\_CODE\_A": "600007", "COMPANY\_CODE": "600007", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "中国国贸", "COMPANY\_ABBR": "中国国贸", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "100728.25", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-", {"NUM": "5", "totalFlowShares": "482061.41", "endDate": "2018-07-10", "LISTING\_DATE": "2000-04-27", "SECURITY\_CODE\_A": "600008", "COMPANY\_CODE": "600008", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "首创股份", "COMPANY\_ABBR": "首创股份", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "482061.41", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-", {"NUM": "6", "totalFlowShares": "109347.64", "endDate": "2018-07-10", "LISTING\_DATE": "1998-02-18", "SECURITY\_CODE\_A": "600009", "COMPANY\_CODE": "600009", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "上海机场", "COMPANY\_ABBR": "上海机场", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "192695.84", "END\_SHARE\_VICE\_DEPART": "-", "CHANGE\_DATE": "-", {"NUM": "7", "totalFlowShares": "3167721.16", "endDate": "2018-07-10", "LISTING\_DATE": "2001-03-09", "SECURITY\_CODE\_A": "600010", "COMPANY\_CODE": "600010", "SECURITY\_CODE\_B": "-", "END\_SHARE\_CODE": "-", "SECURITY\_ABBR\_A": "宝钢股份", "COMPANY\_ABBR": "宝钢股份", "SECURITY\_ABBR\_B": "-", "END\_SHARE\_MAIN\_DEPART": "-", "totalShares": "4558503.

26","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"8","totalFlowShares":"1520038.34","endDate":"2018-07-10","LISTING\_DATE":"2001-12-06","SECURITY\_CODE\_A":"600011","COMPANY\_CODE":"600011","SECURITY\_CODE\_B":"-","END\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"华能国际","COMPANY\_ABBR":"华能国际","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"1520038.34","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"9","totalFlowShares":"165861.00","endDate":"2018-07-10","LISTING\_DATE":"2003-01-07","SECURITY\_CODE\_A":"600012","COMPANY\_CODE":"600012","SECURITY\_CODE\_B":"-","END\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"皖通高速","COMPANY\_ABBR":"皖通高速","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"165861.00","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"10","totalFlowShares":"1282268.67","endDate":"2018-07-10","LISTING\_DATE":"2003-09-12","SECURITY\_CODE\_A":"600015","COMPANY\_CODE":"600015","SECURITY\_CODE\_B":"-","END\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"华夏银行","COMPANY\_ABBR":"华夏银行","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"1282268.67","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"11","totalFlowShares":"4378241.85","endDate":"2018-07-10","LISTING\_DATE":"2000-12-19","SECURITY\_CODE\_A":"600016","COMPANY\_CODE":"600016","SECURITY\_CODE\_B":"-","END\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"民生银行","COMPANY\_ABBR":"民生银行","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"4378241.85","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"12","totalFlowShares":"307565.39","endDate":"2018-07-10","LISTING\_DATE":"2006-10-17","SECURITY\_CODE\_A":"600017","COMPANY\_CODE":"600017","SECURITY\_CODE\_B":"-","END\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"民生银行","COMPANY\_ABBR":"民生银行","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"307565.39","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"}]



## python 之——动态数据的获取

SHARE\_CODE":"-","SECURITY\_ABBR\_A":"上海电力","COMPANY\_ABBR":"上海电力  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"240965.7  
1","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"17","totalFlowSh  
ares":"1068861.31","endDate":"2018-07-10","LISTING\_DATE":"2004-06-29","SECURI  
TY\_CODE\_A":"600022","COMPANY\_CODE":"600022","SECURITY\_CODE\_B":"-","END  
\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"山东钢铁","COMPANY\_ABBR":"山东钢铁  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"1094654.  
96","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"18","totalFlowS  
hares":"1360069.00","endDate":"2018-07-10","LISTING\_DATE":"2013-12-19","SECU  
RITY\_CODE\_A":"600023","COMPANY\_CODE":"600023","SECURITY\_CODE\_B":"-","EN  
D\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"浙能电力","COMPANY\_ABBR":"浙能电力  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"1360069.  
00","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"19","totalFlowS  
hares":"180000.00","endDate":"2018-07-10","LISTING\_DATE":"2017-12-15","SECURI  
TY\_CODE\_A":"600025","COMPANY\_CODE":"600025","SECURITY\_CODE\_B":"-","END  
\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"华能水电","COMPANY\_ABBR":"华能水电  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"1800000.  
00","END\_SHARE\_VICE\_DEPART":"-","CHANGE\_DATE":"-"},{"NUM":"20","totalFlowS  
hares":"403203.29","endDate":"2018-07-10","LISTING\_DATE":"2002-05-23","SECURI  
TY\_CODE\_A":"600026","COMPANY\_CODE":"600026","SECURITY\_CODE\_B":"-","END  
\_SHARE\_CODE":"-","SECURITY\_ABBR\_A":"中远海能","COMPANY\_ABBR":"中远海能  
","SECURITY\_ABBR\_B":"-","END\_SHARE\_MAIN\_DEPART":"-","totalShares":"403203.2

```
9","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"21","totalFlowSh
ares":"850183.93","endDate":"2018-07-10","LISTING_DATE":"2005-02-03","SECURIT
Y_CODE_A":"600027","COMPANY_CODE":"600027","SECURITY_CODE_B":"-","END_
SHARE_CODE":"-","SECURITY_ABBR_A":"华电国际","COMPANY_ABBR":"华电国际
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"986297.6
7","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"22","totalFlowSh
ares":"12107120.96","endDate":"2018-07-10","LISTING_DATE":"2001-08-08","SECU
RITY_CODE_A":"600028","COMPANY_CODE":"600028","SECURITY_CODE_B":"-","EN
D_SHARE_CODE":"-","SECURITY_ABBR_A":"中国石化","COMPANY_ABBR":"中国石化
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"1210712
0.96","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"23","totalFlow
Shares":"1008817.33","endDate":"2018-07-10","LISTING_DATE":"2003-07-25","SEC
URITY_CODE_A":"600029","COMPANY_CODE":"600029","SECURITY_CODE_B":"-","E
ND_SHARE_CODE":"-","SECURITY_ABBR_A":"南方航空","COMPANY_ABBR":"南方航空
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"1008817.
33","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"24","totalFlowS
hares":"1209298.94","endDate":"2018-07-10","LISTING_DATE":"2003-01-06","SECU
RITY_CODE_A":"600030","COMPANY_CODE":"600030","SECURITY_CODE_B":"-","EN
D_SHARE_CODE":"-","SECURITY_ABBR_A":"中信证券","COMPANY_ABBR":"中信证券
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"1211690.
84","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"},{"NUM":"25","totalFlowS
hares":"770011.75","endDate":"2018-07-10","LISTING_DATE":"2003-07-03","SECURI
```

```
TY_CODE_A":"600031","COMPANY_CODE":"600031","SECURITY_CODE_B":"-","END
_SHARE_CODE":"-","SECURITY_ABBR_A":"三一重工","COMPANY_ABBR":"三一重工
","SECURITY_ABBR_B":"-","END_SHARE_MAIN_DEPART":"-","totalShares":"773862.1
6","END_SHARE_VICE_DEPART":"-","CHANGE_DATE":"-"}],stockCode":"","stockType
":"1"))
```

## 扩展作业

问题：要抓 30 页该怎么办？



分析 url:

第二页 url:

[http://query.sse.com.cn/security/stock/getStockListData2.do?&jsonCallBack=jsonpCallback96322&isPagination=true&stockCode=&csrcCode=&areaName=&stockType=1&pageHelp.cacheSize=1&pageHelp.beginPage=2&pageHelp.pageSize=25&pageHelp.pageNo=2&pageHelp.endPage=21&\\_1531271694634](http://query.sse.com.cn/security/stock/getStockListData2.do?&jsonCallBack=jsonpCallback96322&isPagination=true&stockCode=&csrcCode=&areaName=&stockType=1&pageHelp.cacheSize=1&pageHelp.beginPage=2&pageHelp.pageSize=25&pageHelp.pageNo=2&pageHelp.endPage=21&_1531271694634)

第三页 url:

[http://query.sse.com.cn/security/stock/getStockListData2.do?&jsonCallBack=jsonpCallback4267&isPagination=true&stockCode=&csrcCode=&areaName=&stockType=1&pageHelp.cacheSize=1&pageHelp.beginPage=3&pageHelp.pageSize=25&pageHelp.pageNo=3&pageHelp.endPage=31&\\_1531271694635](http://query.sse.com.cn/security/stock/getStockListData2.do?&jsonCallBack=jsonpCallback4267&isPagination=true&stockCode=&csrcCode=&areaName=&stockType=1&pageHelp.cacheSize=1&pageHelp.beginPage=3&pageHelp.pageSize=25&pageHelp.pageNo=3&pageHelp.endPage=31&_1531271694635)

不同之处仅仅在个别关键字部分进行了修改。

def find\_pageA(c): #根据传递参数 c（提取的页数）来选择目标 url 地址

return

```
'http://query.sse.com.cn/security/stock/getStockListData2.do?&jsonCallBack=jsonpCallback4267&isPagination=true&stockCode=&csrcCode=&areaName=&stockType=1&pageHelp.cacheSize=1&pageHelp.beginPage=='+str(c)+'&pageHelp.pageSize=25&pageHelp.pageNo=='+str(c)+'&pageHelp.endPage=='+str(c)+'1&_153127169463'+str(c)'
```

作业：

动态的获取豆瓣电影的 json 数据

<https://movie.douban.com/tag/#/>

[https://movie.douban.com/j/new\\_search\\_subjects?sort=T&range=0,10&tags=&start=0](https://movie.douban.com/j/new_search_subjects?sort=T&range=0,10&tags=&start=0)

中国商标网电子公告爬虫 (json)

<http://sbj.saic.gov.cn/>

<http://sbgg.saic.gov.cn:9080/tmann/annInfoView/homePage.html>