

题目
使用selenium爬取大麦网数据

代码：

```
from selenium import webdriver
import time
import re
import csv
from selenium.webdriver.common.by import By
from selenium.webdriver.common.keys import Keys
from selenium.common.exceptions import TimeoutException, NoSuchElementException
```

```
browser = webdriver.Firefox(executable_path=r"geckodriver.exe")
url= 'https://www.tmall.com'
dic = []
```

```
try:
    browser.get(url)
except TimeoutException:
    print("Time out")
```

```
try:
    input_firse = browser.find_element_by_id('mq')
    input_firse.send_keys('衣服')
    input_firse.send_keys(Keys.ENTER)
```

```
time.sleep(6)
```

```
product = browser.find_elements_by_class_name('product')
for i in product:
    #价格
    price = i.find_elements_by_class_name('productPrice')
    for j in price:
        Price = j.text
        print(Price)

    #标题
    title = i.find_elements_by_class_name('productTitle')
    for j in title:
        Title = j.text
        print(Title)

    #店铺名
    shop = i.find_elements_by_class_name('productShop')
    for j in shop:
        Shop = j.text
        print(Shop)

    #成交量和评价数
```

```

        status = i.find_elements_by_class_name('productStatus')
        for j in status:
            item = re.split('\n',j.text)
            print(item[0])
            print(item[1])
            print('*****')

        info = {'标题':Title,'价格':Price,'店铺':Shop,'月成交量':item[0],'评价数':item[1]}
        dic.append(info)

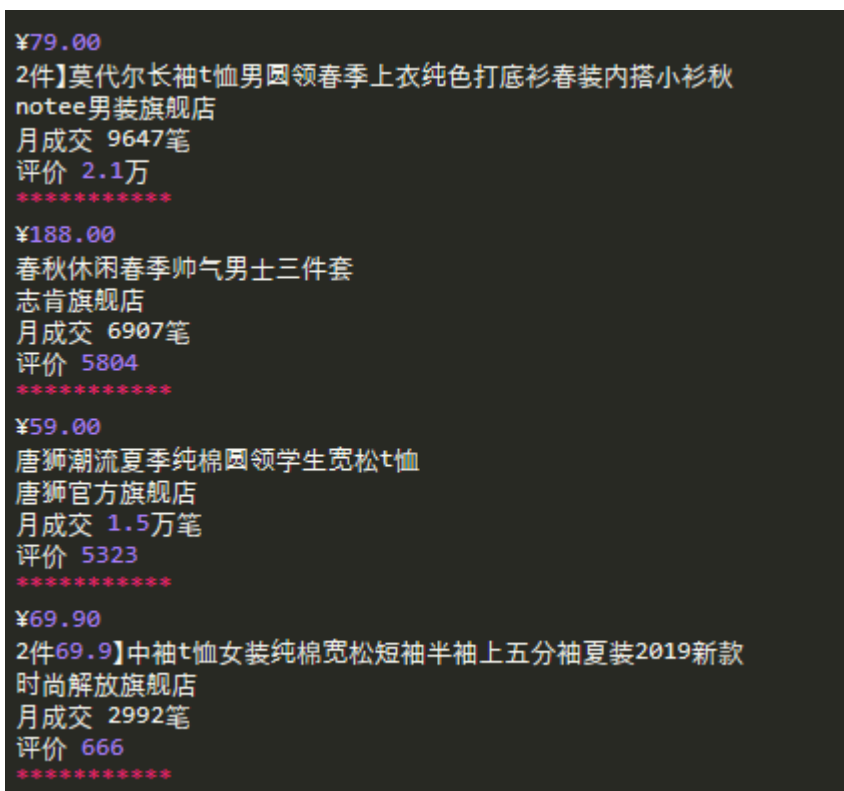
except NoSuchElementException:
    print("No Element")

finally:
    browser.close()
    print("执行完成！ ")

with open('2016110308_何丹阳_实验5.csv','w',newline="",encoding='utf-8-sig') as f:
    writer = csv.DictWriter(f, fieldnames=['标题','价格','店铺','月成交量','评价数'])
    writer.writeheader()
    writer.writerows(dic)

```

截图



微博链接：

<https://chengdu.anjuke.com/sale/jinjiang/m58/>