



Python 程式教學-chapter 26

- Selenium 套件v2
- 8891二手車網查詢
- 連續頁面抓取方式

■ 為何不使用requests進行抓取？

```
1 import requests
2
3 resp=requests.get('https://auto.8891.com.tw/?key=altis&page=1')
4
5 resp
```

<Response [404]>



反爬蟲機制

■ 增加用戶代理 (useragent)

```
1 import requests
2
3 headers = {'user-agent': 'Mozilla/5.0 (Macintosh Intel Mac OS X 10_13_4)\
4 AppleWebKit/537.36 (KHTML, like Gecko) Chrome/66.0.3359.181 Safari/537.36'}
5
6
7 resp=requests.get('https://auto.8891.com.tw/?key=altis&page=1',headers=headers)
8
9 print(resp)
10
11
12 getPageInfo(resp.text)
```

<Response [200]>

[]



資料還是為空

■ 使用Selenium套件



auto.8891.com.tw

Google Cloud Co... C# 的變數與運算 -... 技術事件簿: C# 如... [C#] 使用 Win32 A... JSON Editor Onlin... 在Android中使用2... 使用Unity3D的設...

<<返回首頁 新車 中古車 中古車APP

登入 註冊

8891中古車

找汽車 找機車 中古車商 賣車 服務保障

透過webdriver 定位跟輸入
車種搜尋

瑪莉

搜尋

廠牌： Mercedes-Benz/賓士 BMW Toyota Mitsubishi/三菱 Honda Ford/福特 Mazda Nissan Volkswagen/福斯 Audi 更多

Lexus Hyundai/現代

車款： C300 Mazda3 Focus Corolla Altis Civic GLC300 Yaris CR-V

車齡： 1年以內 1-3年 3-5年 5-10年 10年以上 請選擇 - 請選擇 年

價格： 10萬以下 10-25萬 25-35萬 35-50萬 50-100萬 100萬以上 萬 確定

其他： 地區 車色 車種 車門數 燃料 乘客數 變速系統 排氣量

Das WeltAuto.
福斯原廠認證中古車

原廠認證 值得信賴
立即前往

Das WeltAuto.

- 資料抓取
- 標題/金額/連結



Maserati Levante 2017款 3.0L

【齊樂車酷】2017年 瑪莎 Levante S 新車6XX萬 配備非常好

358萬

2017年

3.0萬km

台北市 | 23小時內更新 | 2442次瀏覽

車身號已審



Maserati Ghibli 2014款 3.0L

《SAVE認證車》僅跑3萬餘~全車內外如新~免200輕鬆入主瑪莎拉蒂~可全額貸

198萬

2014年

3.0萬km

桃園市 | 2小時內更新 | 4277次瀏覽

車身號已審

■ 儲存成CSV



1	標題	價格	連結
2	Maserati Ghibli 2014款 3.0L	169.8萬	https://auto.8891.com.tw/usedauto-infos-2257556.html
3	Maserati Ghibli 2014款 3.0L	170萬	https://auto.8891.com.tw/usedauto-infos-2236386.html
4	Maserati Ghibli 2014款 3.0L	180萬	https://auto.8891.com.tw/usedauto-infos-2244081.html
5	Maserati Ghibli 2014款 3.0L	198萬	https://auto.8891.com.tw/usedauto-infos-2233572.html
6	Maserati GranTurismo 2010款 4.7	199萬	https://auto.8891.com.tw/usedauto-infos-2259499.html
7	Maserati Ghibli 2016款 3.0L	209.8萬	https://auto.8891.com.tw/usedauto-infos-2264523.html
8	Maserati Ghibli 2016款 3.0L	247.8萬	https://auto.8891.com.tw/usedauto-infos-2246683.html
9	Maserati Ghibli 2016款 3.0L	249.3萬	https://auto.8891.com.tw/usedauto-infos-2263898.html
10	Maserati Ghibli 2016款 3.0L	249萬	https://auto.8891.com.tw/usedauto-infos-2252448.html
11	Maserati Ghibli 2016款 3.0L	258萬	https://auto.8891.com.tw/usedauto-infos-2243476.html
12	Maserati GranTurismo 2015款 4.7	295萬	https://auto.8891.com.tw/usedauto-infos-2256049.html
13	Maserati GranTurismo 2011款 4.7	338萬	https://auto.8891.com.tw/usedauto-infos-2235997.html
14	Maserati Levante 2017款 3.0L	358萬	https://auto.8891.com.tw/usedauto-infos-2238798.html
15	Maserati Levante 2018款 3.0L	388萬	https://auto.8891.com.tw/usedauto-infos-2224594.html
16	Maserati Quattroporte 2007款 4.2	88萬	https://auto.8891.com.tw/usedauto-infos-2217416.html
17	Maserati 228 1990年以前款 2.8L	89.9萬	https://auto.8891.com.tw/usedauto-infos-2247140.html
18	Maserati Quattroporte 2006款 4.2	99萬	https://auto.8891.com.tw/usedauto-infos-2248420.html

■ 資料的錯誤

搜尋 altis→



Toyota Corolla Altis 2013款 1.8L

【上好汽車】原廠保養僅29000km超低里程 原廠選配天窗精裝影音

38.8萬

2013年

2.9萬km

台北市 | 11分鐘內更新 | 170次瀏覽

車身號已審



Toyota Corolla Altis 2013款 1.8L

8891嚴選巨將汽車 國民神車房車首選 僅跑3.9萬里 定速巡航 恆溫 雙安

35.8萬

2013年

3.9萬km

高雄市 | 13分鐘內更新 | 2056次瀏覽

售價真實

100%有車

車況真實

8891嚴選承諾：售價真實、100%到店有車、車況真實，看車如有不實，8891立賠5000元！



Toyota Corolla Altis 2016款 1.8L

【泰洋】ALTIS 影音 Z版大包 一手車況佳 認證車 零頭款 可全貸

47.8萬

2016年

8.0萬km

台中市 | 37分鐘內更新 | 278次瀏覽

車身號已審

全部車

8891嚴選車

該筆連結需要跳離

■ 檢查None



```
for r in results:
    if r.find('span',class='ib-price')!=None:
        titles=r.find('span',class_='ib-it-text').text.strip()
        prices=r.find('span',class_='ib-price').text.strip()
        urls=url+r.get('href')
        cars.append([titles,prices,urls])
```

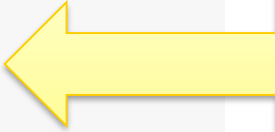
找出關鍵部分進行是否是None的檢查

■ 使用 try except

```
def getPageInfo(html):  
  
    cars=[]  
    soup=BeautifulSoup(html,'lxml')  
    if soup!=None:  
        results=soup.find(id='result-items').find_all('a')  
        #print(results)  
        for r in results:  
            try:  
                #if r.find('span',class_='ib-price')!=None:  
                title=r.find('span',class_='ib-it-text').text.strip()  
                price=r.find('span',class_='ib-price').text.strip()  
                url=r.get('href')  
                cars.append([title,price,url])  
            except:  
                print('該連結無資料..')
```

■ 將單頁分析邏輯宣告成函式

```
def getPageInfo(html):  
  
    cars=[]  
    soup=BeautifulSoup(html,'lxml')  
    if soup!=None:  
        results=soup.find(id='result-items').find_all('a')  
        #print(results)  
        for r in results:  
            try:  
                #if r.find('span',class_='ib-price')!=None:  
                title=r.find('span',class_='ib-it-text').text.strip()  
                price=r.find('span',class_='ib-price').text.strip()  
                url=r.get('href')|  
                cars.append([title,price,url])  
            except:  
                print('該連結無資料..')
```



這邊使用
try except:
來跳開錯誤

■ 連續頁面查找



Toyota Corolla Altis 2015款 1.8L

{小白/免保人/免頭款}[全額貸，超額貸，0元交車]FB:(碩-誠信嚴選中古車)

45萬

2015年

5.5萬km

桃園市 | 46分鐘內更新 | 128次瀏覽

車身號已審

查車在店

<上一頁

1

2

3

4

5

...

25

下一頁>

共 25 頁，到第

頁

確定

查找最大頁面

■ 分析網址連結跟頁數資訊

<https://auto.8891.com.tw/?key=altis&page=1>

<https://auto.8891.com.tw/?key=altis&page=3>

將主網址連結跟頁數分開

[https://auto.8891.com.tw/?key=altis&](https://auto.8891.com.tw/?key=altis&page=2)
[page=2](#)



頁數是關鍵

■ 找出最大頁數區段

▼<ul class="list clearfix">

::before

```
<li class="page prev"><上一頁</li>
<li class="page number" id="page1">1
</li>
<li class="page number at" id=
"page2">2</li>
<li class="page number" id="page3">3
</li>
<li class="page number" id="page4">4
</li>
<li class="page number" id="page5">5
</li>
```

▶<li class="gradient">...

```
<li class="page number" id="page25">
25</li> == $0
```

```
<li class="page next">下一頁></li>
```

▶<li class="total">...

```
<li class="sure">確定</li>
```

::after

```
pages=soup.find('ul',class_='list clearfix').find_all('li')[-4]
if pages!=None:
    pages=int(pages.text.strip())
    print('共%d頁'%pages)
```

■ 進行連續抓取

```
all_cars=[]

for i in range(pages):
    url=main_url+'&page='+str(i+1)
    print(url)
    chrome=getWebDriver(url,path,option='hide')
    if chrome!=None:
        try:
            cars=list(getPageInfo(chrome.page_source))
            #print(cars)
            if cars!=None:
                all_cars+=cars
                print('資料讀取成功!')


        except Exception as e :
            print(e)
        finally:
            chrome.quit()

    time.sleep(3)

all_cars
```



組合成連結



使用迴圈跟函式進行抓取



使用迴圈抓取請用暫停秒數

■ 使用pandas 套件

```
import pandas as pd

df1=pd.DataFrame(cars_sort,columns=['標題','價格','連結'])

#ascending=True 正排序
df1=df1.sort_values(by='價格',ascending=True)

df1.to_html('cars.html')

df1
```

	標題	價格	連結
0	Toyota Corolla Altis 2006款 1.8L	10.8萬	/usedauto-infos-2259507.html
1	Toyota Corolla Altis 2005款 1.8L	10.8萬	/usedauto-infos-2258966.html
2	Toyota Corolla Altis 2001款 1.6L	10.8萬	/usedauto-infos-2218784.html
3	Toyota Corolla Altis 2001款 1.8L	10.8萬	/usedauto-infos-2219571.html
4	Toyota Corolla Altis 2007款 1.8L	10.8萬	/usedauto-infos-2230276.html
5	Toyota Corolla Altis 2008款 1.8L	10.8萬	/usedauto-infos-2244069.html
6	Toyota Corolla Altis 2005款 1.8L	10萬	/usedauto-infos-2225728.html
7	Toyota Corolla Altis 2006款 1.8L	11.3萬	/usedauto-infos-2254671.html
8	Toyota Corolla Altis 2009款 1.8L	11.5萬	/usedauto-infos-2253080.html
9	Toyota Corolla Altis 2009款 1.8L	11.7萬	/usedauto-infos-2220719.html
12	Toyota Corolla Altis 2002款 1.8L	11.8萬	/usedauto-infos-2232765.html

	標題	價格	連結
0	Toyota Corolla Altis 2006款 1.8L	10.8萬	/usedauto-infos-2259507.html
1	Toyota Corolla Altis 2005款 1.8L	10.8萬	/usedauto-infos-2258966.html
2	Toyota Corolla Altis 2001款 1.6L	10.8萬	/usedauto-infos-2218784.html
3	Toyota Corolla Altis 2001款 1.8L	10.8萬	/usedauto-infos-2219571.html
4	Toyota Corolla Altis 2007款 1.8L	10.8萬	/usedauto-infos-2230276.html
5	Toyota Corolla Altis 2008款 1.8L	10.8萬	/usedauto-infos-2244069.html
6	Toyota Corolla Altis 2005款 1.8L	10萬	/usedauto-infos-2225728.html
7	Toyota Corolla Altis 2006款 1.8L	11.3萬	/usedauto-infos-2254671.html
8	Toyota Corolla Altis 2009款 1.8L	11.5萬	/usedauto-infos-2253080.html
9	Toyota Corolla Altis 2009款 1.8L	11.7萬	/usedauto-infos-2220719.html
12	Toyota Corolla Altis 2002款 1.8L	11.8萬	/usedauto-infos-2232765.html
13	Toyota Corolla Altis 2008款 1.8L	11.8萬	/usedauto-infos-2257321.html
10	Toyota Corolla Altis 2006款 1.6L	11.8萬	/usedauto-infos-2262317.html
11	Toyota Corolla Altis 2004款 1.8L	11.8萬	/usedauto-infos-2256299.html
14	Toyota Corolla Altis 2004款 1.8L	11.9萬	/usedauto-infos-2236822.html
15	Toyota Corolla Altis 2004款 1.8L	11萬	/usedauto-infos-2242233.html
16	Toyota Corolla Altis 2010款 1.8L	11萬	/usedauto-infos-2254275.html
17	Toyota AT2EPN 2006款 1.8L	11萬	/usedauto-infos-2256469.html
18	Toyota Corolla Altis 2008款 1.8L	11萬	/usedauto-infos-2248864.html
19	Toyota Corolla Altis 2012款 1.8L	12.3萬	/usedauto-infos-2242619.html
20	Toyota Corolla Altis 2002款 1.8L	12.5萬	/usedauto-infos-2244153.html
21	Toyota Corolla Altis 2005款 1.8L	12.5萬	/usedauto-infos-2235139.html
29	Toyota Corolla Altis 2003款 1.8L	12.8萬	/usedauto-infos-2255599.html
34	Toyota Corolla Altis 2011款 1.8L	12.8萬	/usedauto-infos-2256329.html
33	Toyota Corolla Altis 2007款 1.8L	12.8萬	/usedauto-infos-2236429.html
32	Toyota Corolla Altis 2005款 1.8L	12.8萬	/usedauto-infos-2168279.html
30	Toyota Corolla Altis 2005款 1.8L	12.8萬	/usedauto-infos-2240868.html

輸出成HTML格式

■ 程式練習

- 請整合以上程式碼，寫成函式傳入欲查詢的車種，進行自動最大頁數查詢及資料蒐集回傳

```
url='https://auto.8891.com.tw'  
path=r'C:\webdriver\chromedriver'  
print(getCarList('altis',url,path))
```

```
altis  
查找中...  
https://auto.8891.com.tw/?key=altis  
共30頁  
=====
```

頁數	查詢結果
第1頁查詢..	該連結無資料..
第2頁查詢..	資料讀取成功!
第3頁查詢..	資料讀取成功!
第4頁查詢..	資料讀取成功!
第5頁查詢..	資料讀取成功!
第6頁查詢..	資料讀取成功!
第7頁查詢..	資料讀取成功!

```
=====
```