



Python 程式教學 Chapter23

- 網頁爬蟲v3
- 使用api方式進行資料抓取(學習使用post功能)
- 全台縣市ibon資料(儲存成excel格式)

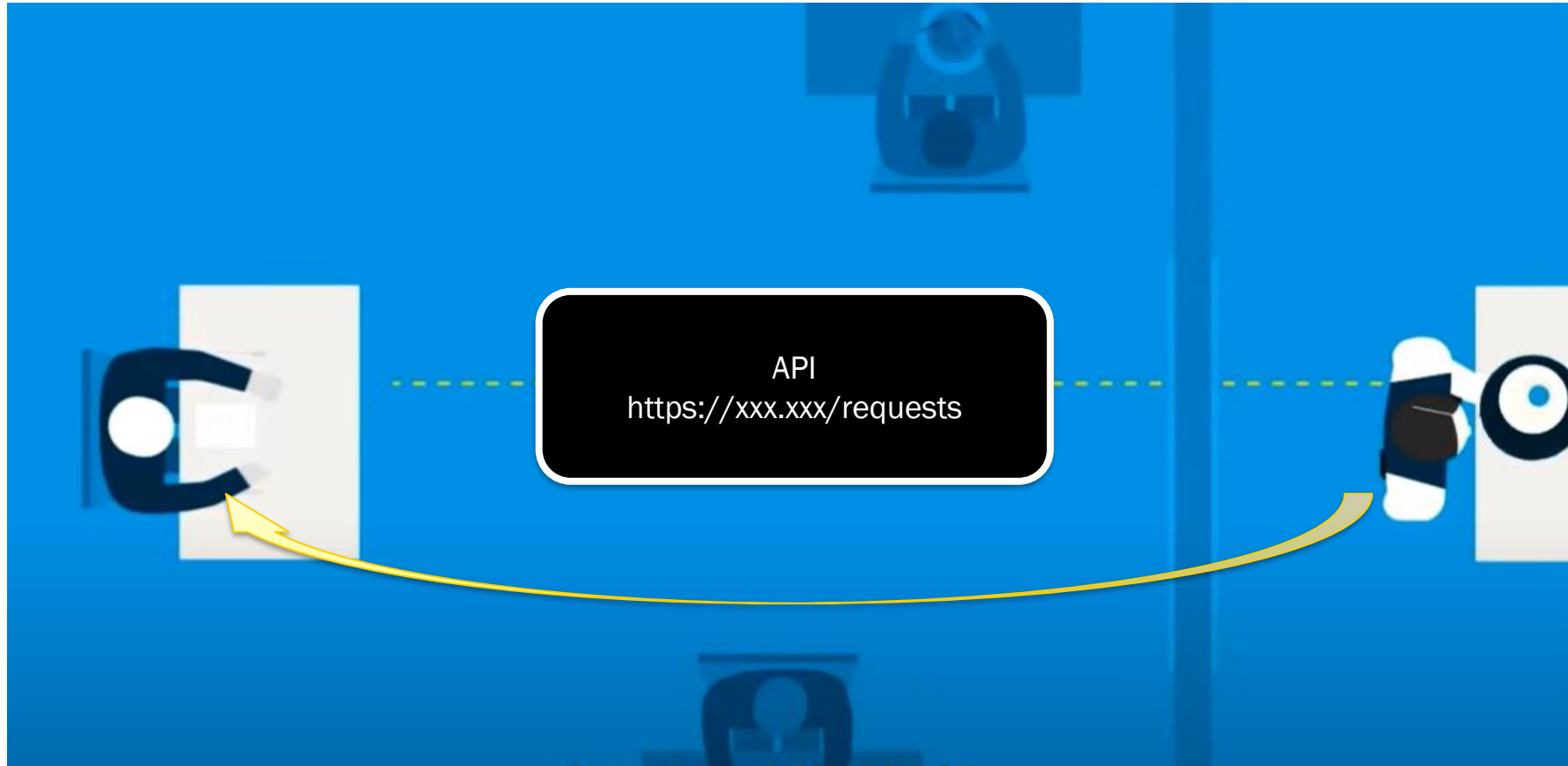


使用api進行資料抓取

- 觀察網頁資料傳輸過程
- 檢查XHR發送找出api
- 模擬FormData(表單)
- 使用post發出回傳
- 收到資料後進行分析

■ 什麼是API?

- 應用程式介面(Application Programming Interface)



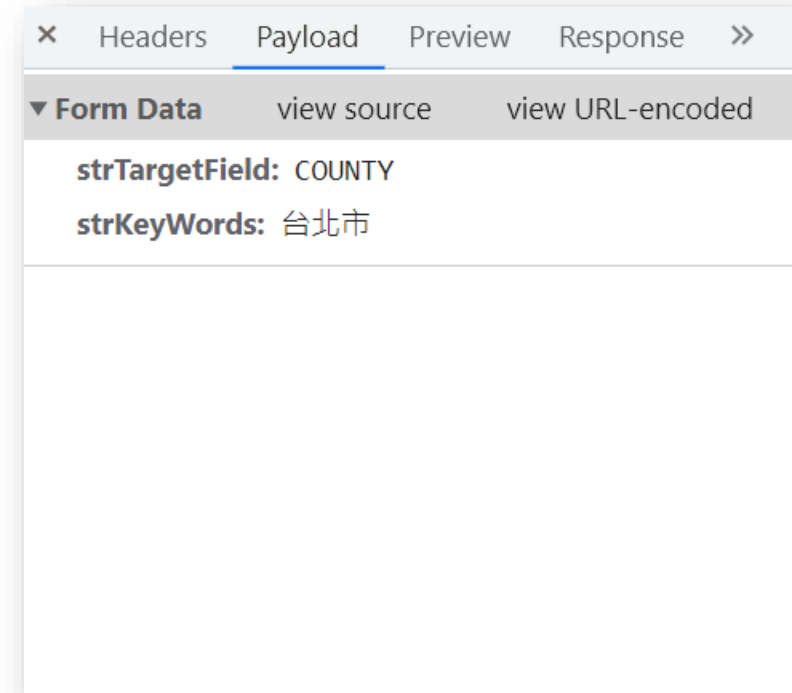
- <https://www.youtube.com/watch?v=zvKadd9Cf1c>

■ GET 跟 POST 差別

■ GET將表單資訊(參數)都顯示在URL之上傳送(不安全)

- https://www.ibon.com.tw/retail_inquiry.aspx?strTargetField=COUNTY&strKeyWords=台北市

■ POST將表單資訊放在REQUEST. Body一併傳送(較為安全)




■ 全台 ibon 地點查詢

http://www.ibon.com.tw/retail_inquiry.aspx#gsc.tab=0



能夠直接爬所有縣市資料嗎?



請選擇欲查詢門市：

地址：

店號	店名	地址
110242	敦維	台北市大安區東豐街43號45號1樓
110334	安信	台北市大安區信義路四段74號1樓
110448	古亭	台北市中正區汀州路二段210號
110817	千翔	台北市中正區許昌街17號1樓
110828	港華	台北市內湖區港華街40號1樓
111278	前港	台北市士林區後港街98之1號2號1樓
111441	合維	台北市大安區四維路170巷8號1樓
111636	江東	台北市中山區長安東路二段43號45-1號1樓
112374	敬群	台北市中山區樂群二路154號1樓
112684	塔優	台北市松山區撫遠街197號199號
113089	秀湖	台北市內湖區成功路四段359號1樓

- 門市地點是使用選取縣市後發送資料給後端server後進行呈現，故無法一次直接爬到所有資料



請選擇欲查詢門市：

地址：可查地址、店號

全部行政區

查詢 ▶

店號	店	地址
110242	敦	東豐街43號45號1樓
110334	安	信義路四段74號1樓
110448	古	州路二段210號
110817	千	昌街17號1樓
110828	港	華街40號1樓
111278	前	港街98之1號2號1樓
111441	合	維路170巷8號1樓
111636	江	安東路二段43號45-1號1樓
112374	敬	群二路154號1樓
112684	塔	遠街197號199號
113089	秀	成功路四段359號1樓



關鍵：使用api+post

- 找出api連結

```
api_url="http://www.ibon.com.tw/retail_inquiry_ajax.aspx"
```

- 使用requests.post進行資料回傳

```
resp=requests.post(request_url, data=formData)
```



第一步驟：查詢ibon網頁

- http://www.ibon.com.tw/retail_inquiry.aspx#gsc.tab=0
- 切換開發者工具頁面
- 爬目前所有縣市地點
- 儲存成list資料格式

■ 抓取目前縣市地點

地圖上包含所有縣
市地點資料



請選擇欲查詢門市：

台北市

全部行政區

地址：可查地址、店號/店名、縣市/鄉鎮市區/郵遞區號

查詢

店號	店名	地址
110242	敦維	台北市大安區東豐街43號45號1樓
110334	安信	台北市大安區信義路四段74號1樓
110448	古亭	台北市中正區汀州路二段210號
110817	千翔	台北市中正區許昌街17號1樓
110828	港華	台北市內湖區港華街40號1樓
111278	前港	台北市士林區後港街98之1號2號1樓
111441	合維	台北市大安區四維路170巷8號1樓
111636	江東	台北市中山區長安東路二段43號45-1
112374	敬群	台北市中山區樂群二路154號1樓
112684	塔優	台北市松山區撫遠街197號199號
113089	秀湖	台北市內湖區成功路四段359號1樓
113160	指南	台北市文山區指南路一段34號36號1樓

```

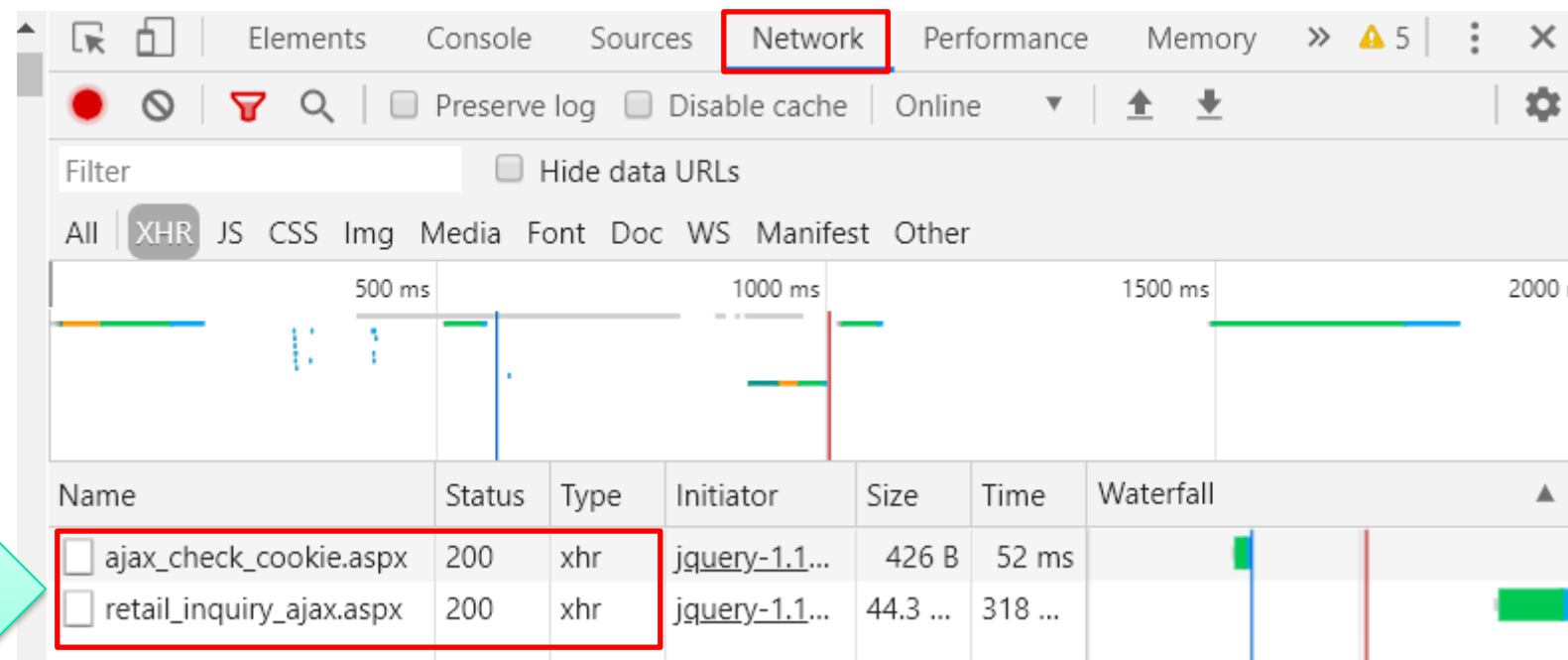
<div>...</div>
<div>...</div>
<div id="googlesearchbar" style="width:10px;height:16
-100px;left:-100px;z-index:-100;">...</div>
<!--wrapper Start-->
<div id="wrapper">
  <!--header Start-->
  <header>...</header>
  <!--header-->
  <!--contentContainer-->
  <div class="contentContainer" style="min-height: 54
  <!--content-->
    <div class="content">
      <div class="mainArea" id="B011">
        <!--notabcontent-->
        <div class="mainContainer" style=" min-height
        <div class="searchmap" style="float:left; m
        relative; background:url(image/map.gif) no-re
        555px; margin:15px 0 0 0;">
...
<p style="left: 320px; top: 25px;" onclick=
基隆市</p>
<p style="left: 265px; top: 28px;" onclick=
台北市</p>
<p style="left: 270px; top: 64px;" onclick=
新北市</p>
<p style="left: 210px; top: 45px;" onclick=
桃園市</p>
<p style="left: 159px; top: 80px;" onclick=
新竹市</p>
<p style="left: 210px; top: 87px;" onclick=
新竹縣</p>
<p style="left: 175px; top: 126px;" onclie
">苗栗縣</p>
<p style="left: 155px; top: 170px;" onclie
">台中市</p>
<p style="left: 110px; top: 213px;" onclie
">彰化縣</p>
<p style="left: 106px; top: 255px;" onclie
">雲林縣</p>
<p style="left: 190px; top: 225px;" onclie
">南投縣</p>

```

```
# =====  
# 找出ibon地點  
# =====  
import requests  
from bs4 import BeautifulSoup  
  
url='http://www.ibon.com.tw/retail_inquiry.aspx#gsc.tab=0'  
resp=requests.get(url)  
  
#citys=[]  
if resp.status_code==200:  
    soup=BeautifulSoup(resp.text,'lxml')  
  
    ps=soup.find('div',class_='searchmap').find_all('p')  
  
    citys=[p.text for p in ps]  
#     for p in ps:  
#         print(p.text)  
  
else:  
    print('網路資料要求有錯誤')  
  
print(citys)
```

['基隆市', '台北市', '新北市', '桃園市', '新竹市', '新竹縣',
'苗栗縣', '台中市', '彰化縣', '雲林縣', '南投縣', '嘉義縣',
'嘉義市', '台南市', '高雄市', '屏東縣', '台東縣', '花蓮縣',
'宜蘭縣', '連江縣', '金門縣', '澎湖縣']

- 切換觀察視窗頁面，重新刷新F5
- 查找Network標籤
- 觀察找出xhr (XMLHttpRequest (XHR) 物件與伺服器互動)



找出XHR

■ 查找主要傳送資料的API

Name	×	Headers	Preview	Response	Cookies	Timing
<input type="checkbox"/> ajax_check_cookie.aspx	1			<response><passed>True</passed><message>0</message></response>		
<input type="checkbox"/> retail_inquiry_ajax.aspx						

Name	×	Headers	Preview	Response	Cookies	Timing
<input type="checkbox"/> ajax_check_cookie.aspx	1	7631</td>		<td align="center">安居 </td>	<td>台北市大安區安居街33號	
<input type="checkbox"/> retail_inquiry_ajax.aspx	2					
	3					
	4					
	5					
	6					
	7					
	8					
	9					
	10					
	11					
	12					

確認為該API

■ 檢查回傳的格式(觀察Response)



```
x Headers Preview Response Cookies Timing
1 <table class="font16" width="100%" cellpadding="5" cellspacing="0" bgcolor="ffffff"> <tr> <td>
2
3 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-t
4
5 <html xmlns="http://www.w3.org/1999/xhtml">
6 <head><title>
7
8 </title></head>
9 <body>
10 <form name="form1" method="post" action="retail_inquiry_ajax.aspx" id="form1">
11 <div>
12 <input type="hidden" name="__VIEWSTATE" id="__VIEWSTATE" value="/wEPDwULLTE2MTY2ODcyMjlkZAuyeJSXcs66wN8m
13 </div>
14
15 <div>
16
17 <input type="hidden" name="__VIEWSTATEGENERATOR" id="__VIEWSTATEGENERATOR" value="4B292396" />
18 </div>
19 <div>
20
21 </div>
22 </form>
23 </body>
24 </html>
25
```

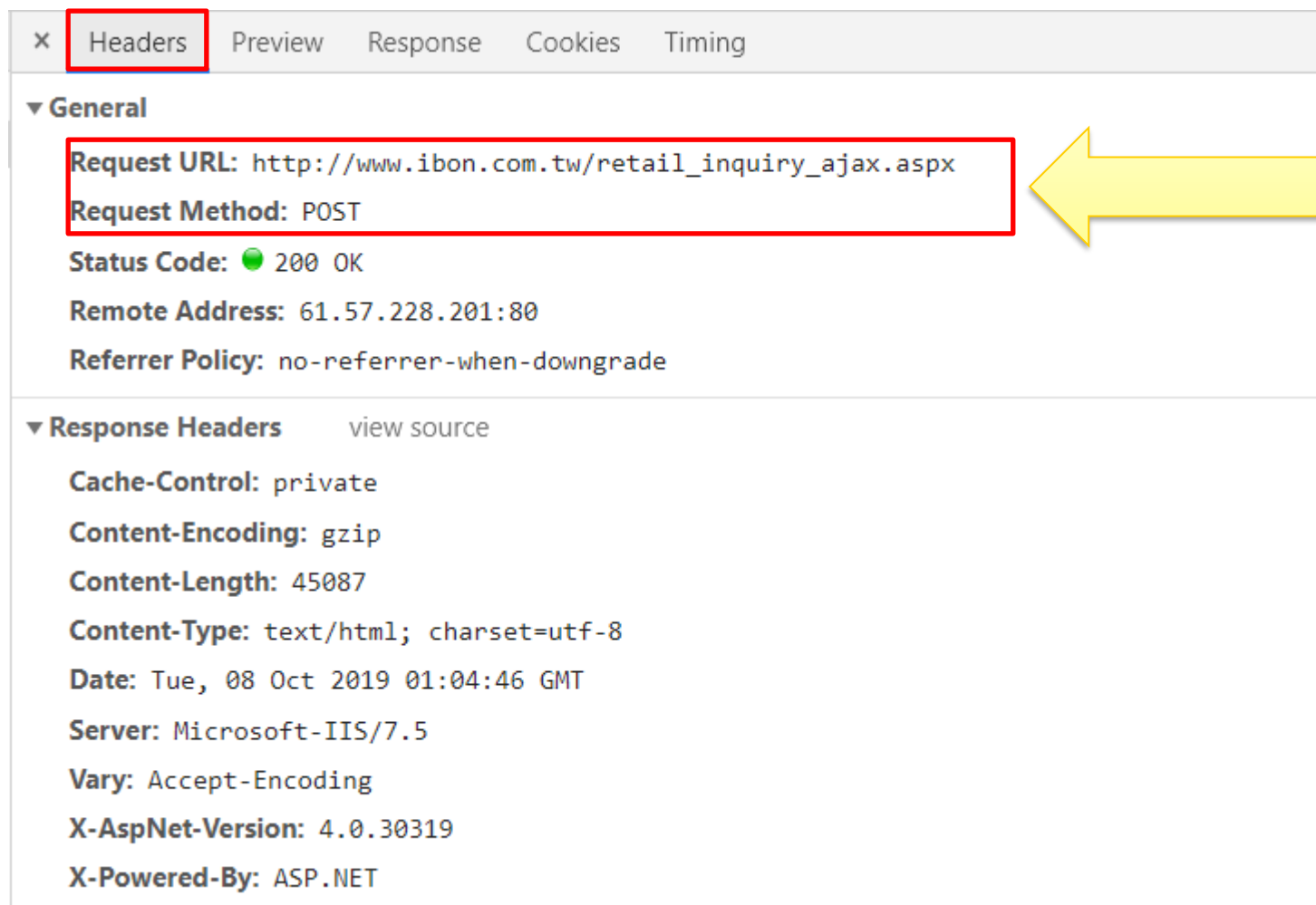
■ 觀察Preview視窗



× Headers Preview Response Cookies Timing			
店號	店名	地址	
110242	敦維	台北市大安區東豐街43號45號1樓	
110334	安信	台北市大安區信義路四段74號1樓	
110448	古亭	台北市中正區汀州路二段210號	
110817	千翔	台北市中正區許昌街17號1樓	
110828	港華	台北市內湖區港華街40號1樓	
111278	前港	台北市士林區後港街98之1號2號1樓	
111441	合維	台北市大安區四維路170巷8號1樓	
111636	江東	台北市中山區長安東路二段43號45-1號1樓	
112374	敬群	台北市中山區樂群二路154號1樓	
112684	塔優	台北市松山區撫遠街197號199號	
113089	秀湖	台北市內湖區成功路四段359號1樓	
113160	指南	台北市文山區指南路一段34號36號1樓	
113551	忠聯	台北市中正區忠孝東路二段100號1樓	
113676	統家	台北市大安區忠孝東路四段216巷27弄1號1樓	
113908	延龍	台北市大同區延平北路四段16號18號1樓	

要抓取的資料

- 切到Headers 標籤
- 複製api



× Headers Preview Response Cookies Timing

▼ General

Request URL: `http://www.ibon.com.tw/retail_inquiry_ajax.aspx`

Request Method: POST

Status Code: 200 OK

Remote Address: 61.57.228.201:80

Referrer Policy: no-referrer-when-downgrade

▼ Response Headers view source

Cache-Control: private

Content-Encoding: gzip

Content-Length: 45087

Content-Type: text/html; charset=utf-8

Date: Tue, 08 Oct 2019 01:04:46 GMT

Server: Microsoft-IIS/7.5

Vary: Accept-Encoding

X-AspNet-Version: 4.0.30319

X-Powered-By: ASP.NET

確定api連結(複製)
及requests method 為post

- Headers標籤
- 確定FormData(表單回傳資料)

▼ Request Headers [view source](#)

Accept: text/html, */*; q=0.01
Accept-Encoding: gzip, deflate
Accept-Language: zh-TW,zh;q=0.9,en-US;q=0.8,en;q=0.7,ja;q=0.6
Connection: keep-alive
Content-Length: 61
Content-Type: application/x-www-form-urlencoded; charset=UTF-8
Cookie: _ga=GA1.3.1553351690.1570440187; _gid=GA1.3.1211956634.1570440187; _gat=1
Host: www.ibon.com.tw
Origin: http://www.ibon.com.tw
Referer: http://www.ibon.com.tw/retail_inquiry.aspx
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/77.0.3865.90 Safari/537.36
X-Requested-With: XMLHttpRequest

需要回傳的
資料

▼ Form Data [view source](#) [view URL encoded](#)

strTargetField: COUNTY
strKeyWords: 台北市

整理需要的資料

api_url=http://www.ibon.com.tw/retail_inquiry_ajax.aspx

```
from_data={  
    "strTargetField":"COUNTY"  
    "strKeyWords":"台北市"  
}
```

準備代入之前查詢到的
全省ibon縣市串列
citys[]

▼ Form Data [view source](#)

strTargetField: COUNTY

strKeyWords: 台北市

■ 第一次api測試(預設為台北市)

```
import requests
from bs4 import BeautifulSoup
```

```
api_url='http://www.ibon.com.tw/retail_inquiry_ajax.aspx'
```

```
form_data={
    'strTargetField':'COUNTY',
    'strKeyWords': '台北市'
}
```

```
#改成post並傳入form_data
```

```
resp=requests.post(api_url,data=form_data)
```

```
if resp.status_code==200:
    soup=BeautifulSoup(resp.text,'lxml')
    print(soup)
```

```
else:
    print('網路資料要求有錯誤')
```

確定有資料回傳

```
SendProcess_Retail_Inquiry_ShowRetailInformation('997571',
;" style="cursor:pointer;font-size:large; "><a
href="javascript:void(0)">997571</a></td> <td
align="center">聚鑫 </td> <td>台北市士林區延平北路六段350號
</td></tr><tr style="background-color:#E9E9E9;"> <td
align="center" class="banner" onclick="return
SendProcess_Retail_Inquiry_ShowRetailInformation('999083')
;" style="cursor:pointer;font-size:large; "><a
href="javascript:void(0)">999083</a></td> <td
align="center">匯陽 </td> <td>台北市內湖區陽光街383號
</td></tr><tr style="background-color:#FFFFFF;"> <td
align="center" class="banner" onclick="return
SendProcess_Retail_Inquiry_ShowRetailInformation('999751')
;" style="cursor:pointer;font-size:large; "><a
```

程式練習

- 觀察Response視窗回傳html格式進行資料爬取
- 儲存至ibon_data[]二維串列

```
[['店號', '店名', '地址'], ['110242', '敦維', '台北市大安區東豐街43號45號1樓'],  
['110334', '安信', '台北市大安區信義路四段74號1樓'], ['110448', '古亭', '台北市  
中正區汀州路二段210號'], ['110817', '千翔', '台北市中正區許昌街17號1樓'],  
['110828', '港華', '台北市內湖區港華街40號1樓'], ['111278', '前港', '台北市士林  
區後港街98之1號2號1樓'], ['111441', '合維', '台北市大安區四維路170巷8號1樓'],  
['111636', '江東', '台北市中山區長安東路二段43號45-1號1樓'], ['112374', '敬群',  
'台北市中山區樂群二路154號1樓'], ['112684', '塔優', '台北市松山區撫遠街197號199  
號'], ['113089', '秀湖', '台北市內湖區成功路四段359號1樓'], ['113160', '指南',  
'台北市文山區指南路一段34號36號1樓'], ['113551', '忠聯', '台北市中正區忠孝東路二  
段100號1樓'], ['113676', '統家', '台北市大安區忠孝東路四段216巷27弄1號1樓'],  
['113908', '延龍', '台北市大同區延平北路四段16號18號1樓'], ['114510', '松錦',  
'台北市中山區松江路295號295-1號1樓'], ['114554', '敦北', '台北市松山區敦化北路  
149號1樓'], ['114668', '敦禾', '台北市大安區敦化南路二段265巷6號1樓'],
```

```
import requests
from bs4 import BeautifulSoup

api_url='http://www.ibon.com.tw/retail_inquiry_ajax.aspx'


form_data={
    'strTargetField':'COUNTY',
    'strKeywords':'台北市'
}

#改成post並傳入form_data
resp=requests.post(api_url,data=form_data)

ibon_data=[]
if resp.status_code==200:
    soup=BeautifulSoup(resp.text,'lxml')
    trs=soup.find('table').find_all('tr')
    #找出每列
    for tr in trs:
        #將每次資料進行儲存(店號/店名/地址)
        temp_data=[]
        #找出每列表格內資料
        tds=tr.find_all('td')
        for td in tds:
            temp_data.append(td.text.strip())

        ibon_data.append(temp_data)

print(ibon_data)
```



進行資料抓取

合併程式(ch23_1.py+ch23_5.py)

- 將citys當作參數進行自動循環抓取

```
['基隆市', '台北市', '新北市', '桃園市', '新竹市', '新竹縣',  
'苗栗縣', '台中市', '彰化縣', '雲林縣', '南投縣', '嘉義縣',  
'嘉義市', '台南市', '高雄市', '屏東縣', '台東縣', '花蓮縣',  
'宜蘭縣', '連江縣', '金門縣', '澎湖縣']
```

```
form_data={  
    'strTargetField': 'COUNTY',  
    'strKeywords': '台北市'  
}
```

將form_data內的
"strKeywords" 修改value
進行自動資料抓取

■ 思考邏輯

```
from_data={  
    "strTargetField":' COUNTY',  
    "strKeyWords":' 台北市'  
}  
  
for city in citys:  
    from_data["strKeyWords"] = city  
    resp=requests.post(api_url,data..  
    .....  
    .....  
    .....
```



```
for city in citys:  
    #改成post並傳入form_data  
    form_data['strKeyWords']=city  
    resp=requests.post(api_url,data=form_data)  
  
    ibon_data=[]  
    if resp.status_code==200:  
        soup=BeautifulSoup(resp.text,'lxml')  
        trs=soup.find('table').find_all('tr')  
        #找出每列  
        for tr in trs:  
            #將每次資料進行儲存(店號/店名/地址)  
            temp_data=[]  
            #找出每列表格內資料  
            tds=tr.find_all('td')  
            for td in tds:  
                temp_data.append(td.text.strip())  
  
            ibon_data.append(temp_data)
```

- 去除每一次抓取資料前頭的欄位資料
- 將每縣市資料再合併成一個dict進行儲存 `ibon={}`

```
[['店號', '店名', '地址'], ['110242', '敦維', '台北市大安區東豐街43號45號1樓'],  
['110334', '安信', '台北市大安區信義路四段74號1樓'], ['110448', '古亭', '台北市  
中正區汀州路二段210號'], ['110817', '千翔', '台北市中正區許昌街17號1樓'],  
['110828', '港華', '台北市內湖區港華街40號1樓'], ['111278', '前港', '台北市士林  
區後港街98之1號2號1樓'], ['111441', '合維', '台北市大安區四維路170巷8號1樓'],  
['111636', '江東', '台北市中山區長安東路二段43號45-1號1樓'], ['112374', '敬群',  
'台北市中山區樂群二路154號1樓'], ['112684', '塔優', '台北市松山區撫遠街197號199
```

`ibon["台北"]`=台北市的二維串列資料

進行資料整理

```
ibon_data=ibon_data[1:]  
ibon[city]=ibon_data
```



```

import requests
from bs4 import BeautifulSoup

url='http://www.ibon.com.tw/retail_inquiry.aspx#gsc.tab=0'
resp=requests.get(url)
#citys=[]
if resp.status_code==200:
    soup=BeautifulSoup(resp.text,'lxml')
    ps=soup.find('div',class_='searchmap').find_all('p')
    citys=[p.text for p in ps]
    # for p in ps:
    #     print(p.text)

else:
    print('網路資料要求有錯誤')

print(citys)

api_url='http://www.ibon.com.tw/retail_inquiry_ajax.aspx'

form_data={
    'strTargetField':'COUNTY',
    'strKeywords':'台北市'
}

```

使用測試
 if ibon!={}:
 print(ibon['台北市'][0])

```
ibon={}
```

```

for city in citys:
    #改成post並傳入form_data
    form_data['strKeywords']=city
    resp=requests.post(api_url,data=form_data)

    ibon_data=[]
    if resp.status_code==200:
        soup=BeautifulSoup(resp.text,'lxml')
        trs=soup.find('table').find_all('tr')
        #找出每列
        for tr in trs:
            #將每次資料進行儲存(店號/店名/地址)
            temp_data=[]
            #找出每列表格內資料
            tds=tr.find_all('td')
            for td in tds:
                temp_data.append(td.text.strip())
            ibon_data.append(temp_data)

            #進行資料整理
            ibon_data=ibon_data[1:]
            ibon[city]=ibon_data

    else:
        print('網路資料要求有錯誤')

```


儲存成excel格式

	A	B	C	D	E	F	G	H	I	J	K
1	店號	店名	地址								
2	112477	極緻	基隆市仁愛區仁二路197號1樓								
3	112879	碇內	基隆市暖暖區源遠路158號160號								
4	113746	德欣	基隆市中山區復興路328號之6號之7號1樓								
5	117140	德信	基隆市信義區東信路50號52號								
6	117896	深澳坑	基隆市信義區深澳坑路2-6號2-7號								
7	118419	聖心	基隆市中山區西定路38號40號								
8	118866	新壯觀	基隆市安樂區麥金路64號1樓								
9	119216	篤鑫	基隆市七堵區堵南街22-1號1樓								
10	119788	富華	基隆市中山區中華路9號								
11	123525	滿福	基隆市信義區深澳坑路166之30號1樓								
12	127103	暖鑫	基隆市暖暖區暖暖街554號								
13	127114	新五福	基隆市七堵區福五街80號82號1樓								
14	128900	港西	基隆市仁愛區港西街6-2號								
15	132415	正光	基隆市七堵區自治街9號11號								
16	138697	新福樂	基隆市信義區深溪路43號45號1樓								
		基隆市	台北市	新北市	桃園市	新竹市	新竹縣	苗栗縣	台中市	彰化縣	...

頁籤使用縣市地點進行分類

- 安裝 xlsxwriter 套件
- `pip install xlsxwriter`

```
系統管理員: Anaconda Prompt (Anaconda3)

(base) C:\WINDOWS\system32>pip install xlsxwriter
Collecting xlsxwriter
  Downloading XlsxWriter-1.2.9-py2.py3-none-any.whl (141 kB)
    |██████████████████████████████████████████████████████████████████████████████| 141 kB 595 kB/s
Installing collected packages: xlsxwriter
Successfully installed xlsxwriter-1.2.9

(base) C:\WINDOWS\system32>
```

使用pandas套件

```
writer = pd.ExcelWriter('ibon便利生活站地點_test.xls', engine = 'xlsxwriter')
count=0
for city in citys:
    count+=1
    if count==1:
        df = pd.DataFrame.from_dict(ibon[city])
        df.columns=['店號','店名','地址']
        df.index=df['店號']
        df=df.drop('店號',axis=1)

        df.to_excel(writer,sheet_name=city)

writer.save()
writer.close()
```

自行新增columns
將index設定
成[店號]
將多餘的[店號]
進行移除

傳入每個縣市的
二維資料

將每個縣市的儲
存到獨立的sheet



程式練習

使用get+params
進行傳遞跟回傳

■ 台灣觀光年曆(使用get+form_data進行資料傳遞)



Browser address bar: taiwan.net.tw/ml.aspx?sNo=0001019

Taiwan THE HEART OF ASIA

網站導覽 | 聯絡我們 | 行政資訊網 | 兒童網

旅遊情報 | 觀光 | 節慶活動 | 美食 | 購物 | 住宿 | 旅遊指南 | 行程推薦 | 多媒體影音

:: 家 > 節慶活動 > 臺灣觀光年曆

臺灣觀光年曆

資料總筆數: 92 每頁筆數: 10 總頁數: 10



國際級

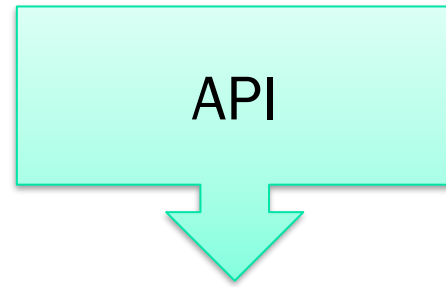
2019臺北馬拉松

2019-12-15

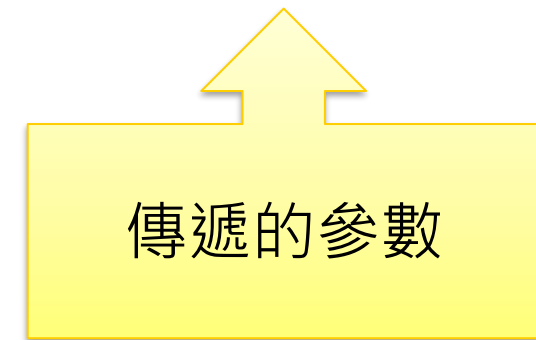
臺北馬拉松是我國規模最大國際城市馬拉松賽事，路線納入承恩門(北門)、麗正門(南門)和景福門(東門)等歷史韻味濃厚的古城門，並行經101大樓、仁愛路林蔭大道、臺北美術知名景點，充分呈現臺北市古今融合的城市風貌。

<https://www.taiwan.net.tw/ml.aspx?sNo=0001019>

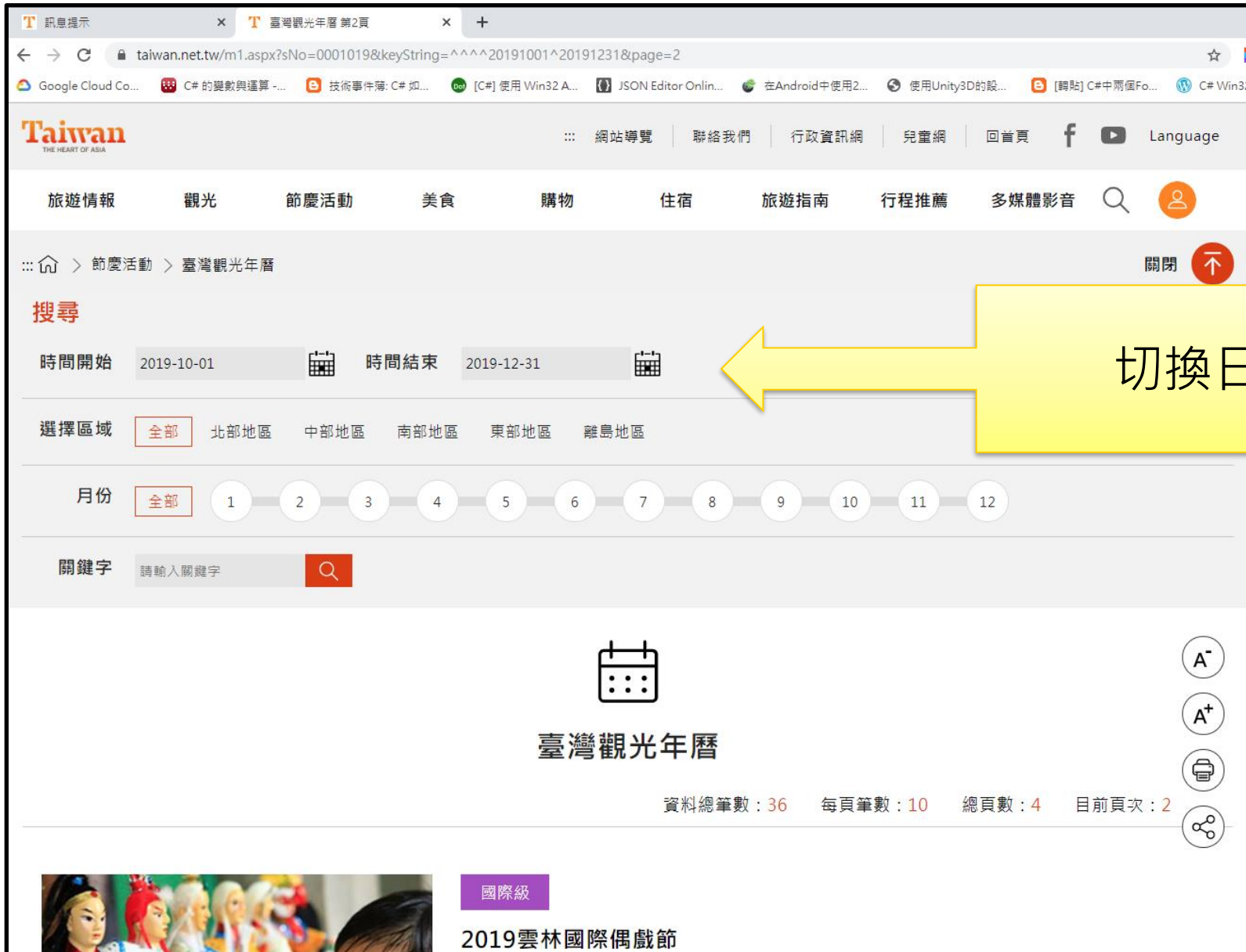
- 使用get傳遞資訊並取得回傳



<https://www.taiwan.net.tw/ml.aspx?sNo=0001019>



■ 檢視url組成格式

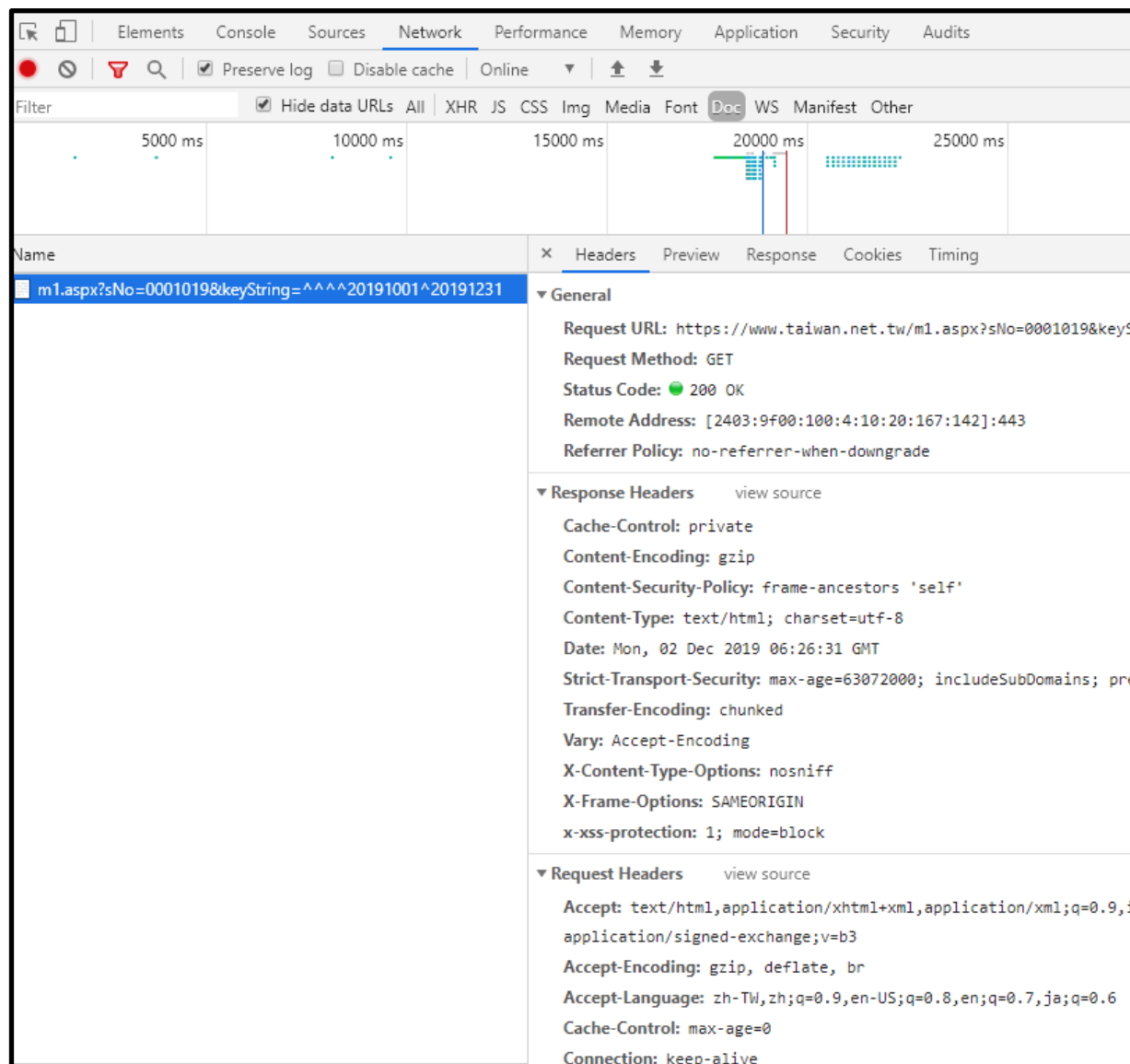


The screenshot shows the Taiwan Tourism Year Calendar website. The URL in the browser is `taiwan.net.tw/m1.aspx?sNo=0001019&keyString=^^^^20191001^20191231&page=2`. The page features a search bar with the following fields:

- 時間開始: 2019-10-01
- 時間結束: 2019-12-31
- 選擇區域: 全部, 北部地區, 中部地區, 南部地區, 東部地區, 離島地區
- 月份: 全部, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
- 關鍵字: 請輸入關鍵字

A yellow callout box with the text "切換日期" (Switch Date) and a large yellow arrow points to the date selection fields. Below the search bar, the page title is "臺灣觀光年曆" (Taiwan Tourism Year Calendar). The page also displays statistics: 資料總筆數: 36, 每頁筆數: 10, 總頁數: 4, 目前頁次: 2. At the bottom, there is a section for "2019雲林國際偶戲節" (2019 Yunlin International Puppet Festival).

■ 檢視Network→Doc標籤



The screenshot shows the Chrome DevTools Network tab. The 'Doc' tag is selected in the filter. The request is to `m1.aspx?sNo=0001019&keyString=^^^^20191001^20191231`. The 'General' tab is expanded, showing the following details:

- Request URL:** `https://www.taiwan.net.tw/m1.aspx?sNo=0001019&keyString=^^^^20191001^20191231&page=2`
- Request Method:** GET
- Status Code:** 200 OK
- Remote Address:** [2403:9f00:100:4:10:20:167:142]:443
- Referrer Policy:** no-referrer-when-downgrade

The 'Response Headers' section shows:

- Cache-Control: private
- Content-Encoding: gzip
- Content-Security-Policy: frame-ancestors 'self'
- Content-Type: text/html; charset=utf-8
- Date: Mon, 02 Dec 2019 06:26:31 GMT
- Strict-Transport-Security: max-age=63072000; includeSubDomains; preconnect
- Transfer-Encoding: chunked
- Vary: Accept-Encoding
- X-Content-Type-Options: nosniff
- X-Frame-Options: SAMEORIGIN
- x-xss-protection: 1; mode=block

The 'Request Headers' section shows:

- Accept: text/html,application/xhtml+xml,application/xml;q=0.9,application/signed-exchange;v=b3
- Accept-Encoding: gzip, deflate, br
- Accept-Language: zh-TW,zh;q=0.9,en-US;q=0.8,en;q=0.7,ja;q=0.6
- Cache-Control: max-age=0
- Connection: keep-alive

▼ General

Request URL: `https://www.taiwan.net.tw/m1.aspx?sNo=0001019&keyString=^^^^20191001^20191231&page=2`

Request Method: GET

Status Code: 200 OK

Remote Address: [2403:9f00:100:4:10:20:167:142]:443

Referrer Policy: no-referrer-when-downgrade

▼ Query String Parameters

[view source](#)
[view URL encoded](#)

sNo: 0001019

keyString: ^^^^^20191001^20191231

page: 2

■ url連結

<https://www.taiwan.net.tw/m1.aspx?sNo=0001019&keyString=^^^20191001^20191231&page=2>

■ api連結

<https://www.Taiwan.net.tw/m1.aspx>

■ 傳遞參數:

[sNo=0001019&keyString=^^^20191001^20191231&page=2](https://www.taiwan.net.tw/m1.aspx?sNo=0001019&keyString=^^^20191001^20191231&page=2)



三個參數

使用get&form_data格式

- 網址
 - <https://www.taiwan.net.tw/m1.aspx?sNo=0001019>
- 使用get+傳遞參數
 - <https://www.taiwan.net.tw/m1.aspx?sNo=0001019&keyString=^^10^20191024^20191031>
- 查找頁數，並自動下一頁查詢

```
import requests
from bs4 import BeautifulSoup

url='https://www.taiwan.net.tw/m1.aspx?sNo=0001019&keyString=^^^^20191001^'

main_url='https://www.taiwan.net.tw/m1.aspx'

from_data={
    "sNo": "0001019",
    "keyString": "^1^^^^",
    "page": 1
}

resp=requests.get(main_url,params=from_data)

if resp.status_code==200:
    soup=soup=BeautifulSoup(resp.text,'lxml')
    #print(soup)
    pages=soup.find('div',class_='title news-title').find_all('span')[2]

    pages=pages.b.text

    print('共%s頁'%pages)

    for page in range(1,int(pages)+1):
        from_data={
            "sNo": "0001019",
            "keyString": "^1^^^^",
            "page": str(page)
        }

        input('分析第%d頁，按任一鍵開始'%page)

        resp=requests.get(main_url,params=from_data)

        if resp.status_code==200:
            soup=soup=BeautifulSoup(resp.text,'lxml')
            #print(soup)

            lis=soup.find('ul',class_='columnBlock-list').find_all('li')

            print(len(lis))

            for li in lis:
                print(li.find('a')['href'])
                print(li.text.strip())
```

2019年暫停辦理

玻璃是新竹市的在地文化，也是當地積極推動的創意產業，每二年舉辦一屆的「國際玻璃藝術節」更是年度盛事！「國際玻璃藝術節」秉持著振興地方玻璃產業、結合地方文化內涵、活化在地觀光旅遊等理念，透過國際交流與全球接軌，讓玻璃產業與文化得以在世界發光發熱。並將臺灣玻璃工藝的獨特性、文化性與藝術性推展至國際舞臺，讓世界看見臺灣層出不窮的玻璃創意。

「國際玻璃藝術節」活動包括「玻璃藝術街」，讓民眾瀏覽各式玻璃藝品；「示範表演」讓民眾瞭解玻璃藝品細緻的創作過程，進行玻璃工藝的推廣教育；「家庭日」提供親子同樂的環境，讓民眾創造屬於自己的玻璃紀念品；「戶外裝置藝術」讓民眾貼近大型玻璃藝品，見識不同於室內展品的風華；「藝文表演活動」則可讓大眾感受嘉年華般的歡樂氣氛。

m1.aspx?sNo=0001019&lid=080064
全國級

臺北溫泉季

2019-10-31~2019-11-04

北投一向以泡湯聞名，北投溫泉歷史悠久，從日治時期開發至今，遊客始終絡繹不絕。位於臺北市近郊的北投，擁有豐富的自然資源、傳統的溫泉鄉文化、酒家菜的美食饗宴、便捷的大眾交通運輸，為旅客們假日休閒放鬆的好去處。近幾年經過臺北市政府與北市溫泉發展協會積極的推廣，旅客來到北投已有更多元的選