

SkyScanner

1. 網頁分析:

在開始撰寫程式之前要先對整個網站作一些分析處理包括訂票流程、一些規則、網址分析...之類。

訂票流程:

1. 先選擇要往返、單程和多個城市模式。

(往返):



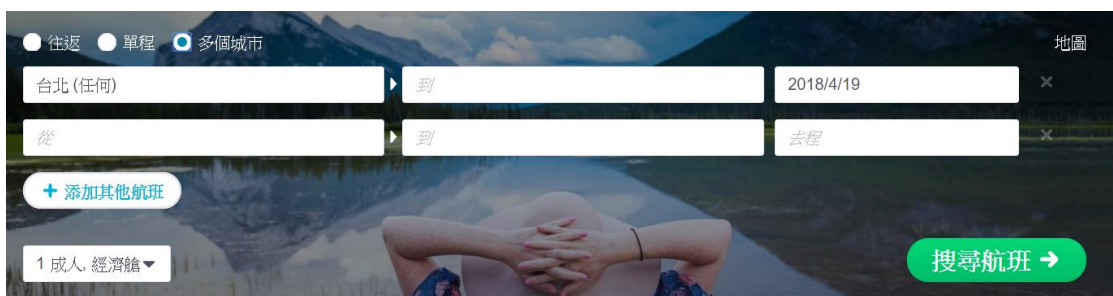
The image shows the SkyScanner flight search interface for a round trip. At the top, there are three radio buttons: "往返" (Round trip) which is selected, "單程" (One way), and "多個城市" (Multiple cities). To the right is a "地圖" (Map) link. Below this, the search form has fields for "從" (From) with "台北 (任何)" (Taipei (any)), "到" (To) with "國家/地區、城市或機場" (Country/region, city or airport), "去程" (Outbound) with "2018/4/19", "回程" (Return) with "2018/4/26", and "客艙等級及旅客" (Cabin class and passengers) with "1 成人, 經濟艙" (1 adult, Economy). There are also checkboxes for "新增鄰近機場" (Add nearby airports) and "僅限直達航班" (Direct flights only). A green "搜尋航班" (Search flights) button with a right arrow is on the right.

(單程):



The image shows the SkyScanner flight search interface for a one way trip. The "單程" (One way) radio button is selected. The "回程" (Return) field now shows "(單程)" (One way). The "客艙等級及旅客" (Cabin class and passengers) field shows "1 成人, 經濟艙" (1 adult, Economy). The "搜尋航班" (Search flights) button is still present.

(多個城市):



The image shows the SkyScanner flight search interface for multiple cities. The "多個城市" (Multiple cities) radio button is selected. The search form is more complex, with multiple input fields for "從" (From) and "到" (To) destinations, and "去程" (Outbound) and "回程" (Return) dates. There is a "+ 添加其他航班" (Add other flights) button. The "客艙等級及旅客" (Cabin class and passengers) field shows "1 成人, 經濟艙" (1 adult, Economy). The "搜尋航班" (Search flights) button is still present.

2. 選擇目的地，這邊目的地有三種模式分別是世界各地、國家、飛機場三種。

(世界各地):



The screenshot shows a flight search form with a scenic background of a person relaxing by a lake. At the top, there are three radio buttons for '往返' (Round Trip), '單程' (One Way), and '多個城市' (Multiple Cities), with '往返' selected. To the right is a '地圖' (Map) link. The form fields are: '從' (From) set to '台北 (任何)' (Taipei (Any)), '到' (To) set to '世界各地' (Worldwide), '去程' (Outbound) set to '2018/4/19', '回程' (Return) set to '2018/4/26', and '客艙等級及旅客' (Cabin Class & Passengers) set to '1 成人, 經濟艙' (1 Adult, Economy). Below the '到' field, there are two checkboxes: '新增鄰近機場' (Add nearby airports) and '僅限直達航班' (Direct flights only), both currently unchecked. A green '搜尋航班' (Search flights) button with a right arrow is on the right side.

(國家:日本(JP)):



This screenshot is identical to the previous one, but the '到' (To) field is set to '日本 (JP)' (Japan (JP)).

(機場:東京 羽田(HND)):



This screenshot is identical to the previous ones, but the '到' (To) field is set to '東京 羽田 (HND)' (Tokyo Haneda (HND)).

3. 去程跟回程時間設定

(去程):

The screenshot shows a flight booking interface. At the top, there are radio buttons for "往返" (Round Trip), "單程" (One Way), and "多個城市" (Multiple Cities). Below these, there are input fields for "從" (From) and "到" (To). The "從" field is set to "台北 (任何)" and the "到" field is set to "東京 羽田 (HND)". There are also checkboxes for "新增鄰近機場" and "僅限直達航班". The "去程" (Departure) date is set to "2018/4/19" and the "回程" (Return) date is set to "2018/4/26". The "客艙等級及旅客" (Cabin Class and Passengers) section is set to "1 成人, 經濟艙". A green button labeled "搜尋航班" (Search Flights) is on the right. A calendar pop-up is open, showing the month of April 2018. The date "19" is selected, and the "特定日期" (Specific Date) tab is active.

週日	週一	週二	週三	週四	週五	週六
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	1	2	3	4	5
6	7	8	9	10	11	12

(回程):

The screenshot shows the same flight booking interface as above, but with the "回程" (Return) date set to "2018/4/20". The calendar pop-up is open, showing the month of April 2018. The date "20" is selected, and the "特定日期" (Specific Date) tab is active.

週日	週一	週二	週三	週四	週五	週六
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	1	2	3	4	5
6	7	8	9	10	11	12

在這邊日期選擇部分理所當然會注意到去的時間跟回來的時間不能問題，像是當出發日期選擇 4/16 而回程日期選擇 4/14 時，出發時間也會跟著一起跳

到 4/14，以防止回程比去程還早的情況發生。

4. 選擇客艙等級跟大人小孩數。

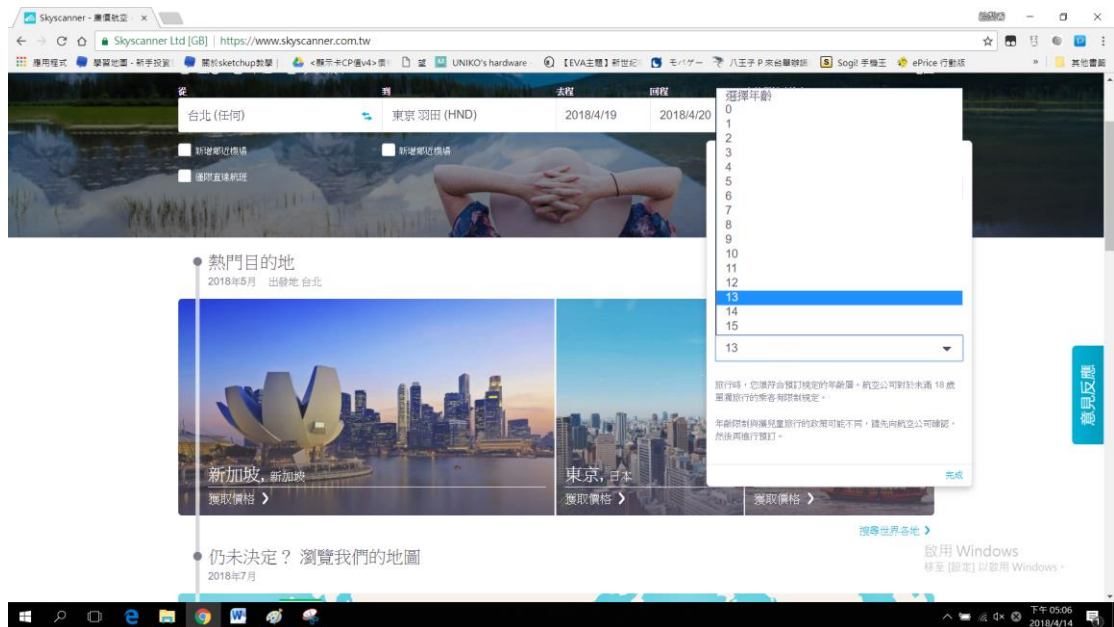
(選擇艙等):

The screenshot shows a flight booking interface. At the top, there are radio buttons for "往返" (Round Trip), "單程" (One Way), and "多個城市" (Multiple Cities). Below this, there are fields for "從" (From) and "到" (To). The "從" field is set to "台北 (任何)" and the "到" field is set to "東京 羽田 (HND)". There are also fields for "去程" (Outbound) and "回程" (Return) dates, both set to "2018/4/19" and "2018/4/20" respectively. A dropdown menu for "客艙等級及旅客" (Cabin Class & Passengers) is open, showing "1 成人, 經濟艙" (1 Adult, Economy Class). Below this, there are checkboxes for "新增鄰近機場" (Add nearby airports) and "僅限直達航班" (Direct flights only). A "熱門目的地" (Popular destinations) section is visible, showing "2018年5月 出發地 台北" (May 2018, Departure Taipei) with a list of destinations. On the right, a "客艙等級" (Cabin Class) dropdown menu is open, showing options: "經濟艙" (Economy Class), "經濟艙" (Economy Class), "高端經濟艙" (Premium Economy Class), "商務艙" (Business Class), and "頭等艙" (First Class). Below this, a "孩童" (Children) section shows "0" children, with a range of "0-15 歲" (0-15 years old). A note states: "旅行時，您須符合預訂規定的年齡層。航空公司對於未滿 18 歲單獨旅行的乘客有限制規定。" (During travel, you must meet the age requirements specified in the booking. Airlines have restrictions for passengers traveling alone who are under 18 years old.) and "年齡限制與攜兒童旅行的政策可能不同，請先向航空公司確認，然後再進行預訂。" (Age restrictions and policies for traveling with children may vary, please confirm with the airline first, then proceed with booking.) A "完成" (Complete) button is at the bottom right.

(填寫大人小孩數):

The screenshot shows the same flight booking interface as above, but with the "客艙等級及旅客" (Cabin Class & Passengers) dropdown menu open. The "客艙等級" (Cabin Class) dropdown menu is still open, showing the same options. Below it, the "大人" (Adults) section shows "1" adult, with a range of "年滿 16 歲" (16 years and older). The "孩童" (Children) section shows "0" children, with a range of "0-15 歲" (0-15 years old). The same note about age restrictions is present. A "完成" (Complete) button is at the bottom right.

(填寫小孩年齡(0~15))



這邊要注意的是當小孩數大於兩位且年齡都介於(0~2 歲)時，需要一名大人

陪同，因此如果有三位介於(0~2 歲)的兒童時，需要有兩名大人做陪同。



5. 選擇機票價錢，藉於剛剛選擇目的地時有分三種方式查詢機票價錢，因此會出現三種不同頁面。

(搜尋世界各地):

The screenshot shows the Skyscanner website interface. At the top, there's a search bar with '台北 (任何)' (Taipei (any)) as the origin and '世界各地' (Worldwide) as the destination. The dates are set to 2018/4/19 and 2018/4/20. Below the search bar, a list of destinations is shown with their starting prices:

目的地	最低機票價錢 (起)
台灣	NT\$2,782
澳門	NT\$4,358
新加坡	NT\$4,714
香港	NT\$5,517
馬來西亞	NT\$5,525
中國	NT\$5,704

On the right side, there are two yellow promotional banners for Scoot flights, both offering a 9% discount with the code MUSTGO. The banners mention a starting price of TWD 1,688* for flights to 29 destinations.

(點開國家之後，或是直接搜尋台灣也會出現這些):

This screenshot shows the Skyscanner website after selecting specific destinations from the previous search. The destinations listed are:

- 台東 (Tainan): NT\$2,782 起
- 金門 (Kinmen): NT\$2,927 起
- 花蓮 (Hualien): NT\$3,229 起
- 馬公 (Makou): NT\$3,503 起
- 馬祖 (Mazu): NT\$3,503 起

The right side of the page features the same Scoot promotional banners as in the previous screenshot, along with a section for hotel recommendations in Taipei.

(點開地區之後，或搜尋地區機場後會出現機票價錢):

Skyscanner

機票 酒店 租車

台北 (任何) - 台東 (TTT)
1 成人 | 經濟艙

顯示整個月份

獲取價格通知

你最愛排序的 25 個結果

排序方式 最優

轉機次數

- ☒ 直飛
NTS\$2,776
- ☐ 轉機1次
無
- ☐ 轉機2次以上
無

出發時間

去程
00:00 - 23:59

回程
00:00 - 23:59

最優	最便宜	最快
NTS\$2,776 0 小時 50 分鐘 (平均)	NTS\$2,776 0 小時 50 分鐘 (平均)	NTS\$2,776 0 小時 50 分鐘 (平均)
17:30 TSA	17:30 TSA	17:30 TSA
0 小時 50 分鐘	0 小時 50 分鐘	0 小時 50 分鐘
18:20 TTT	18:20 TTT	18:20 TTT
08:40 TTT	08:40 TTT	08:40 TTT
0 小時 50 分鐘	0 小時 50 分鐘	0 小時 50 分鐘
09:30 TSA	09:30 TSA	09:30 TSA
17:30 TSA	17:30 TSA	17:30 TSA
0 小時 50 分鐘	0 小時 50 分鐘	0 小時 50 分鐘
18:20 TTT	18:20 TTT	18:20 TTT
18:50 TTT	18:50 TTT	18:50 TTT
0 小時 50 分鐘	0 小時 50 分鐘	0 小時 50 分鐘
19:40 TSA	19:40 TSA	19:40 TSA

10 個優惠，價格低至 NTS\$2,776

選擇

10 個優惠，價格低至 NTS\$3,076

選擇

訂房每晚
港幣\$1,200起
香港迪士尼難忘親子假期
輕鬆啟程
立即預訂

PREDATOR TRITON 700
立即購買

意見反應

機票價錢分三種方式排列，分別是最優、最便宜、最快。

(最優:結合價錢跟速度的最佳組合，會考慮到中途停留次數和麻煩程度等因素):

依最優排序的 1046 個結果

排序方式 最優

最優

NT\$5,273

4 小時 03 分鐘 (平均)

最便宜

NT\$5,273

4 小時 03 分鐘 (平均)

最快

NT\$14,944

4 小時 (平均)

scoot

12:05

TPE

3 小時 50 分鐘

直飛

16:55

CTS

peach

16:45

CTS

4 小時 15 分鐘

直飛

20:00

TPE

3 個優惠，價格低至

NT\$5,273

選擇 →

scoot

12:05

TPE

3 小時 50 分鐘

直飛

16:55

CTS

EVA AIR

16:00

CTS

4 小時 15 分鐘

直飛

19:15

TPE

5 個優惠，價格低至

NT\$12,599

選擇 →

(最便宜):

依最便宜排序的 1046 個結果

排序方式 最便宜優先

最優

NT\$5,273

4 小時 03 分鐘 (平均)

最便宜

NT\$5,273

4 小時 03 分鐘 (平均)

最快

NT\$14,944

4 小時 (平均)

scoot

12:05

TPE

3 小時 50 分鐘

直飛

16:55

CTS

peach

16:45

CTS

4 小時 15 分鐘

直飛

20:00

TPE

3 個優惠，價格低至

NT\$5,273

選擇 →

捷星亞洲航空 + 捷星日本

07:00

TPE

8 小時 55 分鐘

轉機1次 KIX

16:55

CTS

peach

16:45

CTS

4 小時 15 分鐘

直飛

20:00

TPE

2 個優惠，價格低至





NT\$8,914

選擇 →

(最快):

依最快排序的 1046 個結果

排序方式 最快優先

最優 NT\$5,273 4 小時 03 分鐘 (平均)	最便宜 NT\$5,273 4 小時 03 分鐘 (平均)	最快 NT\$14,944 4 小時 (平均)
<div><div></div><div>08:40 TPE</div><div>3 小時 50 分鐘 直飛</div><div>13:30 CTS</div><div>由 中華航空 營運</div></div> <div><div></div><div>15:00 CTS</div><div>4 小時 10 分鐘 直飛</div><div>18:10 TPE</div><div>由 中華航空 營運</div></div>	<div>18 個優惠，價格低至 NT\$14,944</div> <div>選擇 →</div>	
<div><div></div><div>12:05 TPE</div><div>3 小時 50 分鐘 直飛</div><div>16:55 CTS</div><div></div></div> <div><div></div><div>15:00</div><div>4 小時 10 分鐘</div><div>18:10</div><div></div></div>	<div>4 個優惠，價格低至 NT\$15,388</div> <div>選擇 →</div>	

選擇好想要的機票點下去之後就會顯示航班資料和 Skyscanner 所蒐尋的訂票網頁。

去程 2018年4月19日 週四

所有時間都是當地的



12:05

TPE

3 小時 50 分鐘

16:55

CTS

直飛

↑

 酷航 TR892

3 小時 50 分鐘

12:05

TPE 台北桃園

16:55

CTS 札幌千歲

抵達：2018年4月19日 週四 | 旅行總時間：3 小時 50 分鐘

回程 2018年4月20日 週五



16:45

CTS

4 小時 15 分鐘

20:00

TPE

直飛

↑

 樂桃 MM725

4 小時 15 分鐘

16:45

CTS 札幌千歲

20:00

TPE 台北桃園

抵達：2018年4月20日 週五 | 旅行總時間：4 小時 15 分鐘

預定您的機票

基本經濟艙, 1 位成人

訂票前需知

GotoGate

★★★★★ 3191

NT\$5,273

選擇 →

Trip.com

★★★★★ 24084

✈ 透過 Skyscanner 預訂

NT\$5,711

訂購 →

Kiwi.com

★★★★★ 1990

✈ 透過 Skyscanner 預訂

NT\$6,329

訂購 →

選擇要訂票的網站頁面之後就會跳轉到那家網頁的訂票頁面

謝謝您，即將完成...



已找到您所需機票



即將帶您進入GotoGate



現在通過GotoGate完成
預訂

機票詳細資料

TPE - CTS 12:05 2018年4月19日 (星期四)
CTS - TPE 16:45 2018年4月20日 (星期五)



NT\$5,272.35

經濟艙

The screenshot shows the GotoGate website interface. At the top, there's a navigation bar with the GOTO GATE logo and links for '我的預訂', '會員登入', and '您的行程: NT\$ 0'. Below the navigation bar, there's a search bar with '台北' and '台中' selected, and dates '2018年4月19日' and '2018年4月20日'. The search results show a round trip from Taipei (TPE) to Taichung (SPK) with a total price of NT\$ 5,272. The itinerary includes a departure on April 19th at 12:05 and a return on April 20th at 16:45. The website also displays a sidebar with '信賴之選' (Trusted Choice) and '永遠包含' (Always Included) sections. The bottom of the page shows a summary of the flight details and a button to '預訂' (Book).

這邊這個部分通常會有個問題存在，就是 skyscanner 搜尋的票價會跟實際

跳轉網頁之後所顯示的價錢不一樣。

(價錢不一樣部分，像搜尋台北到台東)

去程 2018年4月19日 週四

所有時間都是當地的



17:30
TSA

0 小時 50 分鐘
直飛

18:20
TTT

▼

回程 2018年4月20日 週五



08:40
TTT

0 小時 50 分鐘
直飛

09:30
TSA

▼

預定您的機票

基本經濟艙, 1 位成人

訂票前需知

華信航空 航空公司

★★★★★ 421

NT\$2,776

選擇 →

(但實際按下去之後，價錢也不一樣)

4/16 Monday

4/17 Tuesday


4/18 Wednesday

4/19 Thursday

4/20 Friday

4/21 Saturday

4/22 Sunday

 Outbound: Sungshan → Taitung

2018/4/19 Thursday

Flight No.	Departure time	Arrival time	Aircraft type	Full Fare	Promotion	Select Ticket Fare and Type
AE391	07:20	08:10	E90	\$2303	—	Buy tickets
AE393	17:30	18:20	E90	\$2303	\$1388Up	Buy tickets

※Please be reminded.All passengers can buy the promotion tickets, please depending on the actual needs, choose the best one.

[Select again](#)

網頁分析:

採用普通的 php + JavaScript 來撰寫

(首頁)

```
1
2
3 <!DOCTYPE html>
4 <html class="skyscanner legacy-button " lang="zh-TW" >
5 <head>
6
7 <meta charset="utf-8">
8 <meta name="viewport" content="width=device-width, initial-scale = 1.0, maximum-scale=1.0, user-scalable=no" />
9 <meta http-equiv="X-UA-Compatible" content="IE=10" />
10 <meta http-equiv="Content-Language" content="zh-TW"/>
11
12 <meta property="fb:pages" content="390474687991471" /><meta name="google-site-verification" content="1GBxHSCeJbxqqC8u-wVGQ8mWtGzQzVity16HyyCe0ao" /><meta name="msvalidate.01"
13 content="0E3C71340522093E540BA582021C05F8" />
14
15 <meta name="theme-color" content="#ffffff" />
16 <meta name="msapplication-TileColor" content="#21c4d9" />
17 <meta name="msapplication-TileImage" content="//logos.skyscnr.com/framework/images/icons/mstile-144x144.png" />
18 <link rel="manifest" href="/android-chrome-manifest.json">
19 <link rel="apple-touch-icon" sizes="57x57" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-57x57.png" />
20 <link rel="apple-touch-icon" sizes="60x60" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-60x60.png" />
21 <link rel="apple-touch-icon" sizes="72x72" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-72x72.png" />
22 <link rel="apple-touch-icon" sizes="76x76" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-76x76.png" />
23 <link rel="apple-touch-icon" sizes="114x114" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-114x114.png" />
24 <link rel="apple-touch-icon" sizes="120x120" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-120x120.png" />
25 <link rel="apple-touch-icon" sizes="144x144" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-144x144.png" />
26 <link rel="apple-touch-icon" sizes="152x152" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-152x152.png" />
27 <link rel="apple-touch-icon" sizes="180x180" href="//logos.skyscnr.com/framework/images/icons/apple-touch-icon-180x180.png" />
28 <link rel="mask-icon" href="//logos.skyscnr.com/framework/images/icons/website-icon.svg" color="#454758">
29 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/favicon-16x16.png" sizes="16x16" />
30 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/favicon-32x32.png" sizes="32x32" />
31 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/favicon-96x96.png" sizes="96x96" />
32 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/android-chrome-192x192.png" sizes="192x192" />
33
34 <script>window['start-time'] = new Date().getTime();</script>
35
36 <script type="text/javascript">
37 (function() {
38     var executeImmediately = false;
39
40     window.addLoadEventHandler = function (fn) {
41         executeImmediately ? setTimeout(function () {
42             fn(); }, 0) : window.addEventListener('load', fn, false) : window.attachEvent ? window.attachEvent('onload', fn) : void 0;
43         }
44         addLoadEventHandler(function () { executeImmediately = true; });
45     }
46
47     window.injectScript = function (src, callback, type, crossOrigin) {
48         var an = document.createElement('script');
49         an.type = type || 'text/javascript';
50         an.async = true;
51         if (['anonymous', 'use-credentials'].indexOf(crossOrigin) > -1) {
52             an.crossOrigin = crossOrigin;
53         }
54         if (callback) an.onload = callback;
55         an.src = src.indexOf('http') == 0 ? src : src.indexOf('https') == 0 ? src : ('https:' == document.location.protocol ? 'https://' : 'http://') + document.location.host + src;
56         s.parentNode.insertBefore(an, s);
57     }
58
59     window._collectedErrors = [];
60     var existingHandler = window.onerror;
61     window.onerror = function() { window._collectedErrors.push(arguments) && existingHandler && existingHandler.apply(this, arguments); }
62 })();
63 </script>
64
65 <title>
66 Skyscanner 從台北到東京的廉價航班
67 </title>
68
69
70 <link rel="dns-prefetch" href="//api.mixpanel.com">
71 <link rel="dns-prefetch" href="https://secure.skyscanner.net">
72 <link rel="dns-prefetch" href="https://www.google-analytics.com">
73 <link rel="dns-prefetch" href="//cdn.moxel.com">
74 <link rel="dns-prefetch" href="https://www.googleletsgoservices.com">
```

啟用
移至

(選擇票的頁面)

```
27 <link rel="mask-icon" href="//logos.skyscnr.com/framework/images/icons/website-icon.svg" color="#454758">
28 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/favicon-16x16.png" sizes="16x16" />
29 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/favicon-32x32.png" sizes="32x32" />
30 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/favicon-96x96.png" sizes="96x96" />
31 <link rel="icon" type="image/png" href="//logos.skyscnr.com/framework/images/icons/android-chrome-192x192.png" sizes="192x192" />
32
33 <script>window['start-time'] = new Date().getTime();</script>
34
35
36 <script type="text/javascript">
37 (function() {
38     var executeImmediately = false;
39
40     window.addLoadEventHandler = function (fn) {
41         executeImmediately ? setTimeout(function () {
42             fn(); }, 0) : window.addEventListener('load', fn, false) : window.attachEvent ? window.attachEvent('onload', fn) : void 0;
43         }
44         addLoadEventHandler(function () { executeImmediately = true; });
45     }
46
47     window.injectScript = function (src, callback, type, crossOrigin) {
48         var an = document.createElement('script');
49         an.type = type || 'text/javascript';
50         an.async = true;
51         if (['anonymous', 'use-credentials'].indexOf(crossOrigin) > -1) {
52             an.crossOrigin = crossOrigin;
53         }
54         if (callback) an.onload = callback;
55         an.src = src.indexOf('http') == 0 ? src : src.indexOf('https') == 0 ? src : ('https:' == document.location.protocol ? 'https://' : 'http://') + document.location.host + src;
56         s.parentNode.insertBefore(an, s);
57     }
58
59     window._collectedErrors = [];
60     var existingHandler = window.onerror;
61     window.onerror = function() { window._collectedErrors.push(arguments) && existingHandler && existingHandler.apply(this, arguments); }
62 })();
63 </script>
64
65 <title>
66 Skyscanner 從台北到東京的廉價航班
67 </title>
68
69
70 <link rel="dns-prefetch" href="//api.mixpanel.com">
71 <link rel="dns-prefetch" href="https://secure.skyscanner.net">
72 <link rel="dns-prefetch" href="https://www.google-analytics.com">
73 <link rel="dns-prefetch" href="//cdn.moxel.com">
74 <link rel="dns-prefetch" href="https://www.googleletsgoservices.com">
```

啟用 Wi
移至 (設定)

網址分析:

首頁: <https://www.skyscanner.com.tw/>

我們先將網址簡單分成三個部份，分別是目的地、單程回程、人數種類這三格，而艙等部分因為題目是需要便宜機票，如果選擇其他部分就會有違原意，因此就固定在經濟艙就好。

1. 目的地:

搜尋飛往世界各地(往返):

<https://www.skyscanner.com.tw/transport/flights-from/tpet/180419/180420?adults=1&children=0&adultsv2=1&childrenv2=&infants=0&abinclass=economy&rtn=1&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home>

搜尋國家(日本 JP 成人:1 經濟艙 往返):

<https://www.skyscanner.com.tw/transport/flights/tpet/jp/180419/180420?adults=1&children=0&adultsv2=1&childrenv2=&infants=0&abinclass=economy&rtn=1&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home>

搜尋地區機場(成田 NRT 成人:1 經濟艙往返):

<https://www.skyscanner.com.tw/transport/flights/tpet/nrt/180419/180420?adults=1&children=0&adultsv2=1&childrenv2=&infants=0&cabinclass=economy&rtn=1&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home#results>

將這三條網址刪除同樣部分之後會顯示:

世界各地: <https://www.skyscanner.com.tw/transport/flights-from/tpet/>

國家: <https://www.skyscanner.com.tw/transport/flights/tpet/jp/>

地區機場: <https://www.skyscanner.com.tw/transport/flights/tpet/nrt/>

我們可以注意到當選擇飛往世界各地時，會出現 **flights-form** 而其他的話就只會出現 **flights**，而後面出現的 **jp** 跟 **nrt** 就分別是國家名稱跟機場名稱的簡寫，因此在使用者在輸入時一定要知道想要去的地方的名稱的簡寫。

2. 單程跟往返分別

單程(成田 NRT 成人:1 經濟艙 單程):

<https://www.skyscanner.com.tw/transport/flights/tpet/nrt/180419?adults=1&children=0&adultsv2=1&childrenv2=&infants=0&cabinclass=economy&rtn=0&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home#results>

往返(成田 NRT 成人:1 經濟艙 往返):

<https://www.skyscanner.com.tw/transport/flights/tpet/nrt/180419/180420?adults=1&children=0&adultsv2=1&childrenv2=&infants=0&cabinclass=economy&rtn=1&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home#results>

將這兩條網址刪除同樣部分之後會顯示：

單程：

<https://www.skyscanner.com.tw/transport/flights/tpet/nrt/180419?>

往返：

<https://www.skyscanner.com.tw/transport/flights/tpet/nrt/180419/180420?>

[420?](#)

這邊可以看到他將 2018-04-19 改寫成 180419 而且單程跟往返理所當然時間就分別為一個跟兩個，需要注意的是單程時間完之後接?而往返是第一個接

/第二階?。

3. 人數種類

這邊就相對麻煩許多，在這邊先將網址(成田 NRT 成人:1 經濟艙 往返)做分

段解析:

<https://www.skyscanner.com.tw/transport/flights/tpet/nrt/180419/180420?adults=1&children=0&adultsv2=1&childrenv2=&infants=0&cabinclass=economy&rtn=1&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home#results>

基本: <https://www.skyscanner.com.tw/transport/flights/>

出發地: tpet

目的地: nrt

往返日期: 180419/180420

大人數:1

小孩數:0

大人數 v2 : 1

小孩數 v2 :

幼兒數 : 0

剩下:

&cabinclass=economy&rtn=1&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home#results

當初看到這邊時就有對這邊有疑問，想說都已經有大人跟小孩的類別，為甚

麼還要分出 v2 版，後來在多方實驗才知道規則在哪。

The screenshot shows a flight booking interface. At the top, there's a section titled "客艙等級" (Cabin Class) with a dropdown menu currently set to "經濟艙" (Economy). Below this, there are two sections for passenger selection: "大人" (Adults) and "孩童" (Children). The "大人" section shows a minus button, the number "1", a plus button, and the text "年滿 16 歲" (16 years and older). The "孩童" section shows a minus button, the number "0", a plus button, and the text "0-15 歲" (0-15 years). Below these sections, there is a note in Chinese: "旅行時，您須符合預訂規定的年齡層。航空公司對於未滿 18 歲單獨旅行的乘客有限制規定。" (During travel, you must comply with the age requirements specified in the booking. Airlines have restrictions for passengers traveling alone who are under 18 years old.) Another note follows: "年齡限制與攜兒童旅行的政策可能不同，請先向航空公司確認，然後再進行預訂。" (Age restrictions and policies for traveling with children may differ, please confirm with the airline first before making a booking.) At the bottom right, there is a button labeled "啟用 Wi-Fi" (Enable Wi-Fi) and a link "移至「設定」" (Go to 'Settings') with a "完成" (Done) button below it.

在首頁一開始填寫的大人跟孩童數分別對應 adults 跟 children 數但是後面的 adultsv2 跟 childrenv2 還有 infants 對應的是小孩的年齡區別。

Adults = 首頁所填寫的大人數

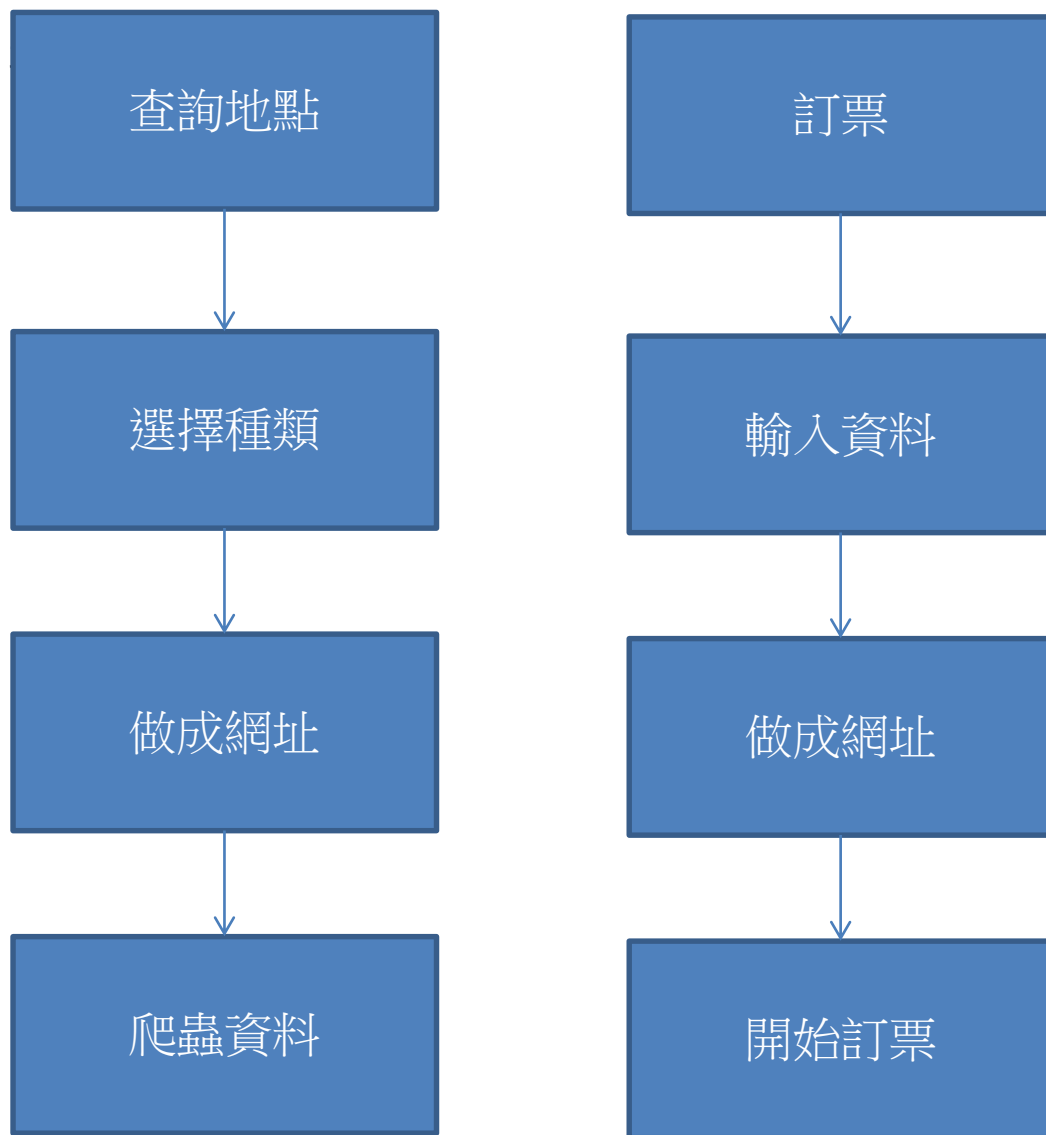
Children = 介於 2 ~ 13 歲的兒童

Adultsv2 = 超過 13 歲的兒童

childrenv2 = 印出所填的兒童年齡，中間以%7c 連接

infants = 介於 0~1 的兒童數

2. 程式架構



查詢地點:

這部分其實只是想要試試看網路爬蟲這部分的功用，因為原本想說可以讓使用者透過這一項功能得知有哪些地方可以去，可是後來發現網頁本身並不會提供要去哪裡地方的簡寫，因此這部分只是拿來用來顯示能通往世界各地哪裡。

訂票:

在開始定票之前，會先讓使用者輸入要訂的機票的相關資料，分別是單程或往返、日期、人數、年齡，輸入完之後進行一連串的規則判斷，都做好之後就會開始製作網址，讓頁面成功到達使用者所輸入的內容，之後開始訂票。

3. 程式

失敗

4. 遇到的困難:

網路爬蟲:

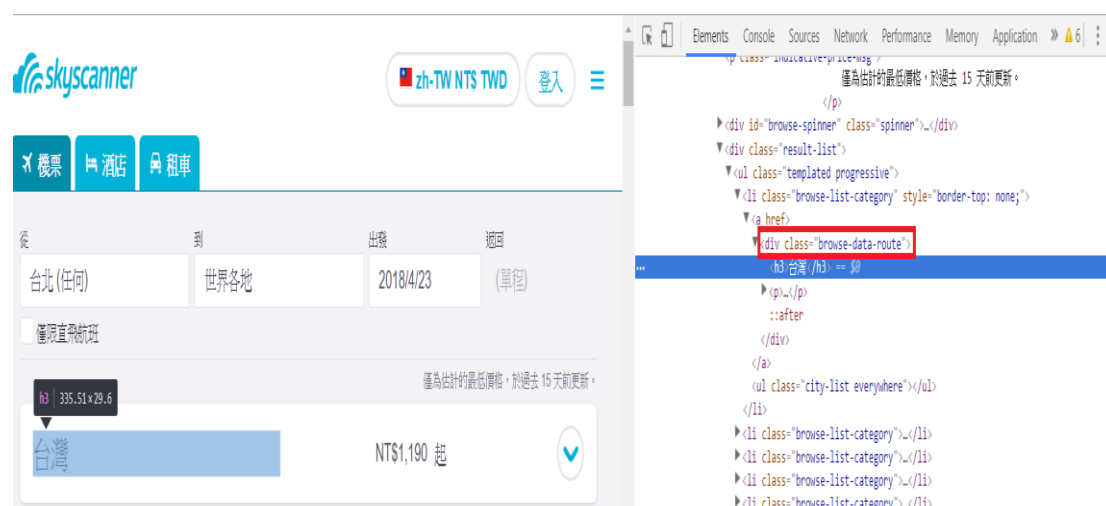
在一開始設定的目標是希望能讓使用者可以查查看今天可以從台灣到世界的哪裡，才做這一部分的，但不做還好一做就知道這網頁是很高竿的。

1. urllib+BeautifulSoup:

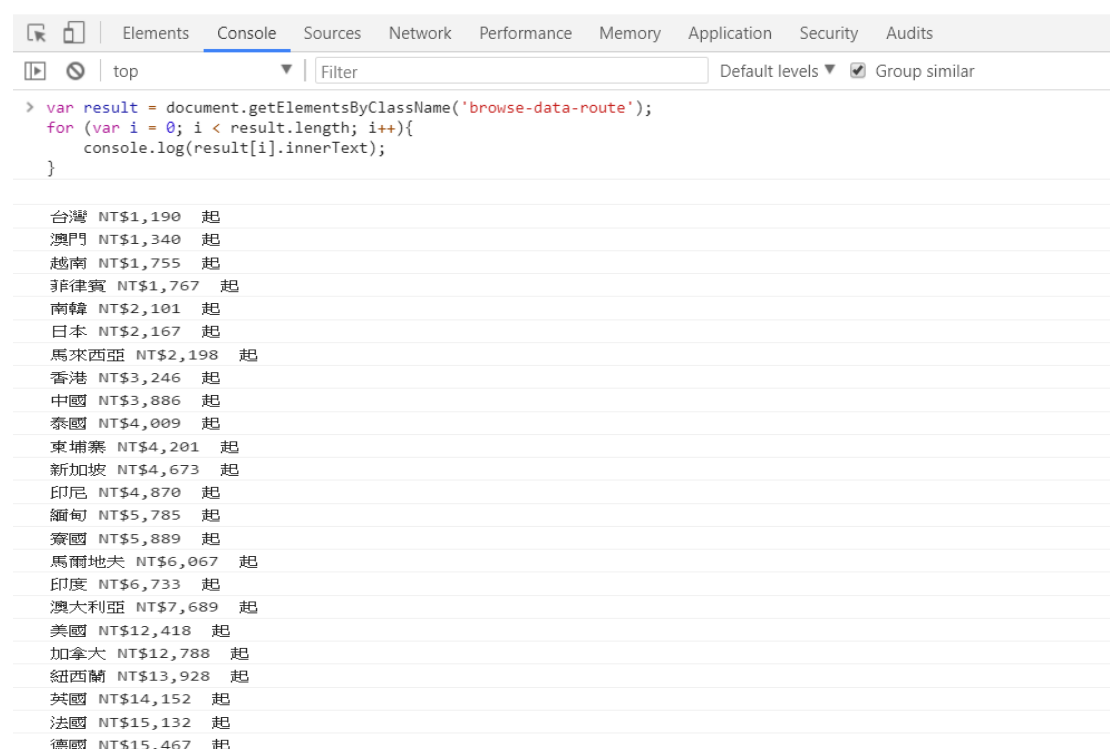
一開始是想要 urllib + BeautifulSoup 這一個組合去做網頁爬蟲，urllib 負責將網頁存取下來而 BeautifulSoup 則是將網頁做解析，在一開始用 urllib 配合上述所說到的製作網址，想要試圖直接抓取搜尋世界各地結果的網頁內容時，發現程式會回報 http error 406 Not Acceptable，那時查了資料發現，是因為網頁訪問時缺少 user-agent 才會被系統擋住，因此找了找 user-agent 並加上去之後發現還是不行，因此又找了其發方法。

2. requests + BeautifulSoup:

這一次使用 requests 後加上 user-agent 之後就可以成功抓取 skyscanner 的網頁資料，但此時又有其他問題了，在模擬操作時，有先試試看能不能查詢網頁所要內容的元件名稱 `<div class="browse-data-route">`



並且使用 Chrome 內建的工具來確認抓的元件是不是所想要跟內容呈現如何



都確認好之後才開始用 requests 來抓，雖然 requests 有抓下來網頁，但是

BeautifulSoup 怎樣都抓不到剛剛上面所寫的網頁元件的內容，而且不僅如此，

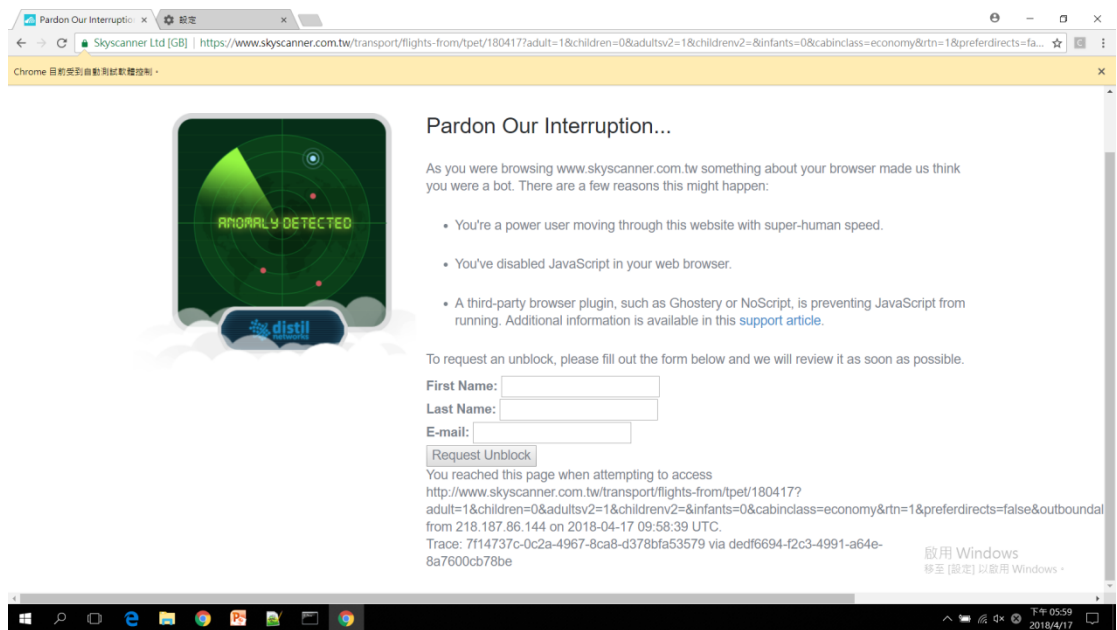
抓了 3、4 次之後，skyscanner 會覺得你是 BOT 因此拒絕你的網頁訪問，回傳的網頁內容會明顯變短，從原本 94KB 變成 5KB 左右，要等幾個小時之後才會回復正常，後來花了花時間查詢資料之後，發現是因為現在網頁的一些搜尋功能都會利用 JavaScript 來去做使用，因此原本內容的網頁叫做靜態網頁，javascript 所呈現的網頁叫做動態網頁，而 requests 只能抓到靜態網頁的內容，故 BeautifulSoup 怎樣都抓不到所需的檔案內容。

3. Selenium + phantomjs

Selenium 原本是用來瀏覽器自動化的一套小工具，可以讓程式直接操控瀏覽器進行各種網路操作，原本是想用 selenium+phantomjs 讓他直接在命令提示字元裡面跑網頁然後再擷取網頁內容，但無奈新版 selenium 已經不支援 phantomjs 了，因此就不能使用。

4. Selenium + chromedriver(一)

Chromedriver 是 chrome 的自動化工具，這個跟 phantomjs 的差別是在直接用 chrome 來進行網頁操作，理論上因該可以跟正常使用者操作一樣，只是變成是程式來做使用，但是 skyscanner 似乎直接擋掉這一類自動化程式的操作，我一直接瀏覽上面所述的製作網址的網址之後，他會直接出現下圖所示，直接阻擋



因始我開始解析是不是 user-agent 或是 cookie 有問題，所以我就利用 chrome 裡面內建的 network 這項檢測網路封包的工具來查詢從首頁到搜尋結果出來這一段裡面有什麼封包內容，經過搜尋有找到兩個有關於 cookie 內容的封包，一個網址標頭是 google

▼ General

Request URL: https://www.google.com/afz/ads?g=KESBEN9E5N8F308N58C979E9A308N58E80undeFinedE7X9A84E888AAAE78F8AD8oe=utf-8&ie=utf-8&r=m&client=pub-6496241184674058&channel=7004543321&hl=zh-TW&type=0&exp=21404&format=n&ad=n&nocache=9141524112558766&num=0&output=uds_ads_only&v=3&sl=10&u_his=3&u_tz=480&dt=1524112558769&u_w=1536&u_h=864&biu=750&bih=715&psw=750&psw=1240&fmm=0&ui=s10s&of3f2t12sdl1s11-8jsv=18255&url=https%3A%2F%2Fwww.skyscanner.com.tw%2Ftransport%2Fflights-from%2Ftpet%2F180417%2Fadults%3D1&children%3D0&infants%3D0&cabinclass=economy&rt=1&preferdirect=false&outboundal=218.187.86.144 on 2018-04-17 09:58:39 UTC. Trace: 7f14737c-0c2a-4967-8ca8-d378bfa53579 via dedf6694-f2c3-4991-a64e-8a7600cb78be

Request Method: GET

Status Code: 200

Remote Address: 108.177.125.105:443

Referrer Policy: no-referrer-when-downgrade

▼ Response Headers

alt-svc: hq=":443"; ma=2592000; quic=51303432; quic=51303431; quic=51303339; quic=51303335, quic=":443"; ma=2592000; v="42,41,39,35"

cache-control: private, max-age=3600

content-encoding: br

content-length: 6693

content-type: text/html; charset=UTF-8

date: Thu, 19 Apr 2018 04:35:56 GMT

server: gws

status: 200

x-xss-protection: 1; mode=block

▼ Request Headers

:authority: www.google.com

:method: GET

:path: /afz/ads?g=KESBEN9E5N8F308N58C979E9A308N58E80undeFinedE7X9A84E888AAAE78F8AD8oe=utf-8&ie=utf-8&r=m&client=pub-6496241184674058&channel=7004543321&hl=zh-TW&type=0&exp=21404&format=n&ad=n&nocache=9141524112558766&num=0&output=uds_ads_only&v=3&sl=10&u_his=3&u_tz=480&dt=1524112558769&u_w=1536&u_h=864&biu=750&bih=715&psw=750&psw=1240&fmm=0&ui=s10s&of3f2t12sdl1s11-8jsv=18255&url=https%3A%2F%2Fwww.skyscanner.com.tw%2Ftransport%2Fflights-from%2Ftpet%2F180417%2Fadults%3D1&children%3D0&infants%3D0&cabinclass=economy&rt=1&preferdirect=false&outboundal=218.187.86.144 on 2018-04-17 09:58:39 UTC. Trace: 7f14737c-0c2a-4967-8ca8-d378bfa53579 via dedf6694-f2c3-4991-a64e-8a7600cb78be

裡面的 cookie 內容是這樣

Name	Value	Domain	Path	Expires / Ma...	Size
Request Cookies					136
NID	128=WBHhml3gZjGuB5jxNeFko5bA2-77J-s4h4TPvyLpbIVA4adP...	N/A	N/A	N/A	136
Response Cookies					0

另外一個是網址標頭是 skyscanner

▼ General

Request URL: https://www.skyscanner.com.tw/transport/flights-from/tpet/180423?adults=1&children=0&adultsv2=1&childreenv2=0&infants=0&cabinclass=economy&rtn=0&preferdirects=false&outboundaltsenabled=false&inboundaltsenabled=false&ref=home
Request Method: GET
Status Code: 200
Remote Address: 210.203.10.219:443
Referrer Policy: no-referrer-when-downgrade

▼ Response Headers

cache-control: no-cache
content-encoding: gzip
content-language: zh-TW
content-length: 29011
content-type: text/html; charset=utf-8
date: Thu, 19 Apr 2018 04:35:54 GMT
expires: -1
pragma: no-cache
server: Microsoft-IIS/7.5
set-cookie: traveller_context=b7cbe336-593d-46cc-b165-bc5f58248754; expires=Sat, 18-Apr-2020 04:35:54 GMT; path=/
set-cookie: abgroup=23928338; Expires=Thu, 17 May 2018 04:35:16 GMT; path=/; httpOnly
set-cookie: ssab=Ads_UseESIAids_V1::b&fps_route_summary_traffic_shift_V6::b&new_price_alerts_ui_V2::b&fps_lus_qss_automatic_rules_V19::a&fps_mbmd_V11::b&mag Mashups_V6::b&fbw_enable_payment_billing_review_V14::a&UtsidTravellerIdentity_V11::b&Car_AATeet_V4::b&fbw_braintrust_new_cvv_service_V6::b&mapinsp_VES_Use_Browse_Proxy_V4::on&glu_springCleanRollout_V2::a&MPT_Saddlebag_V3_0::b&dbbook_flot_trafficcontrol_V12::a&W_LavelPrice_EnableforPmarker_24_02_2017_19_49_V1::b&MPT_Reacti6_upgrade_V2::b&fbw_remove_footer_V2::a&TripsTopicPage_Hidethetopicpage_27_02_2018_34_33_V1::a&skippy_website_link_service_integration_V23::b&dbbook_cath_trafficcontrol_all_web_V2::a&Hfe_OfficialPartner_It2_V2::b&scalfold_wireup_dont_delete_V1::a&Tts_ConductorDayviewProxy_V5_4::b&FLUX_G0T7291_SendPriceTraceToMixpanel_V6::b&HFE_NewReviewsWithSocial_V1::a&fps_lus_client_quote_service_split_traffic_V22a::b&fbw_auth_V26::b&AAExperiment_V8::a&fps_lus_send_quotes_to_slipstream_V35::noexperiment&fss_NewSearchControls_V6::c&dbbook_drag_trafficcontrol_all_web_V2::a&CAKE_DealsWidget_DayviewCheapest_V1::a&Hsc_MexicanToAS2_V4::b&ATBT_Spring_Clean_Carhire_Sea_rch_Controls_V16::b&FLUX787_QuoteBlacklist_V2::a&Hfe_PricePerNight_V2::b&TurnFeatureTests::on&Web_Migration_home_V4::c&hbe_cr_sorting_V2::a&DEAL_Default_To_Two_Guests_V3::b&fbw_auth_inlin_e_V7::b&RAC_web_travel_rank_timetable_V4::on&Hsc_ChildrenAgeView_V10::b&fss_springclean_datepicker_V5::b&rts_magpie_sow_data_collection_V4::budgetscheduled&dbbook_sune_trafficcontrol_web_V_3::a&TCS_Send_Searching_Email_V4::b; expires=Sat, 19-May-2018 04:35:54 GMT; path=/
set-cookie: scanner=adults::1&originalAdults::1&adultsv2::1&children::0&originalChildren::0&infants::0&charttype::1&rtn::false&preferDirects::false&includeOnePlusStop::true&cabinclass::Economy&tripType::OneWayTrip&nrcr::false&lang::TW¤cy::TWD&outboundAlts::false&inboundAlts::false&from::TPET&toym::1804&today::23&usrplace::TW&wy::08to::&iy m::&iday::&fromCy::&toCy::&legs::&childrenV2::; expires=Sun, 29-Apr-2018 04:35:54 GMT; path=/
set-cookie: utm_source:: expires=Thu, 19-Apr-2018 04:35:54 GMT; path=/
set-cookie: X-Mapping-rsqbjcb=898fb6ea89a3b8cacf2a4879cbcd5e8; path=/; max-age=1200
set-cookie: has_app:: expires=Thu, 19-Apr-2018 04:35:54 GMT; path=/
status: 200
vary: Accept-Encoding

啟用 Windows
移至 (設定) 以啟用 Windows。

裡面封包有這麼多

Name	Value	Domain	Path	Expires / Max-Age	Size
Request Cookies					4922
ASP.NET_SessionId	96bcnigqd	N/A	N/A	N/A	29
DAPROPS	*bjs.webG1 bjs.geoLocation:1 bjs.web5qDatabase:1 bjs.indexedD8:1 bjs.webSocket...	N/A	N/A	N/A	694
D_HID	04A71A8E-2416-361D-8F86-D568113CF246	N/A	N/A	N/A	44
D_HID	AAE986F6-3C78-3812-9168-1842791A1D74	N/A	N/A	N/A	44
D_SUID	218.187.86.144:YXdgRnwIAVofMibNjYfdSH2IUDv/pq9hyS2gNGkHk	N/A	N/A	N/A	66
D_UID	390270A3-5F0D-339F-9EE5-3888034F4C2B	N/A	N/A	N/A	44
D_ZUID	F0D48113-5199-3DCD-98DD-9E1930C9C696	N/A	N/A	N/A	44
D_ZUID	1A434898-29DF-3AB8-92D4-138DCF681E18	N/A	N/A	N/A	45
RT	*sl=0&ss=1524112544667&tt=0&obo=0&zh=&dm=skyscanner.com.tw&isi=02373...	N/A	N/A	N/A	209
X-Mapping-fpkjgdih	1819C40E87D011577866563ED130F118	N/A	N/A	N/A	53
X-Mapping-gfrnkhhl	D4939D98142848265C9DB9A4F11CD396	N/A	N/A	N/A	53
X-Mapping-gfrnhaic	8886FC6A591734A705679F5306110A76	N/A	N/A	N/A	53
X-Mapping-rsqbjcb	898fb6ea89a3b8cacf2a4879cbcd5e8	N/A	N/A	N/A	53
_gads	ID=48024e976991531T=1523960447S=ALNI_MbhU_wpU5fyPRRmmySlmmwOnH...	N/A	N/A	N/A	70
_ga	GA13.250470981.1523960449	N/A	N/A	N/A	32
_gat	1	N/A	N/A	N/A	8
_gat_uatracr	1	N/A	N/A	N/A	18
_gid	GA13.343558848.1523960449	N/A	N/A	N/A	33
abgroup	23928338	N/A	N/A	N/A	18
acq	263356c3-e647-4d14-b142-b62b9ee31c77 263356c3-e647-4d14-b142-b62b9ee31...	N/A	N/A	N/A	79
akaas_flights-home	1526703646--rv=1--id=30c6ae9021256d14286c61d93b59a2b	N/A	N/A	N/A	72
cto_lwid	14a46e9f-cbfd-4c2b-81e5-24229b981c4	N/A	N/A	N/A	47
firstvisit	overlay=show	N/A	N/A	N/A	27
mp_243474895430ccc5017faa456fa3d38_mispanel	%78%22distinct_id%22%3A%20%22162d3107a54f-04aecdf6fa4411-b34356b-144...	N/A	N/A	N/A	727
preferences	q3774037b1914b3f99a275b7f817ab12	N/A	N/A	N/A	46
scanner	adults::1&originalAdults::1&adultsv2::1&children::0&originalChildren::0&infants::...	N/A	N/A	N/A	463
settings	acql:true	N/A	N/A	N/A	22
ssab	new_price_alerts_ui_V2::b&fps_lus_qss_automatic_rules_V19::a&fps_mbmd_V11::b...	N/A	N/A	N/A	1648
ssassociate	N/A	N/A	N/A	N/A	14
ssculture	locale::zh-TW¤cy:TWD&market::TW	N/A	N/A	N/A	53
sspref	authcheckedexpiry:1524112726121	N/A	N/A	N/A	42
traveller_context	b7cbe336-593d-46cc-b165-bc5f58248754	N/A	N/A	N/A	56
var	26	N/A	N/A	N/A	8
Response Cookies					2462
X-Mapping-rsqbjcb	898fb6ea89a3b8cacf2a4879cbcd5e8	/	/	20.0 min	73
abgroup	23928338	/	/	2018-05-17T04:3...	74
has_app		/	/	2018-04-19T04:3...	56
scanner	adults::1&originalAdults::1&adultsv2::1&children::0&originalChildren::0&infants::...	/	/	2018-04-29T04:3...	475
ssab	Ads_UseESIAids_V1::b&fps_route_summary_traffic_shift_V6::b&new_price_alerts_ui...	/	/	2018-05-19T04:3...	1623
traveller_context	b7cbe336-593d-46cc-b165-bc5f58248754	/	/	2020-04-18T04:3...	102
utm_source		/	/	2018-04-19T04:3...	59

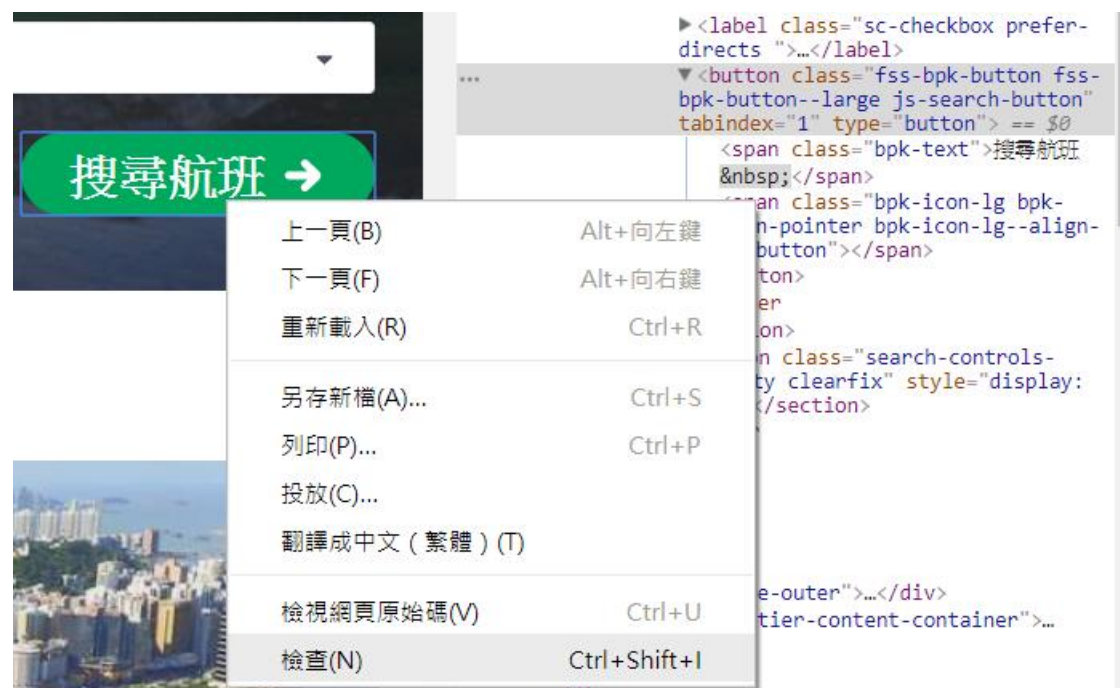
因此我就開始嘗試在開啟網頁內容時添加這些 cookie 到裡面，但是結果還是依樣會出現 pardon our interruption 的網頁畫面，但是還是有偶而會出現成功進入網頁的畫面，但是他會顯示搜尋不到內容，等到重整或是案其他按鈕跳到其他頁面時又會被阻擋，而且在試 cookie 的時候我把原本電腦的 chrome 的 uid 的 cookie 加到裡面也只有成功一次之後就被擋住了，不僅如此還把我原本使用的

chrome 一起擋住。

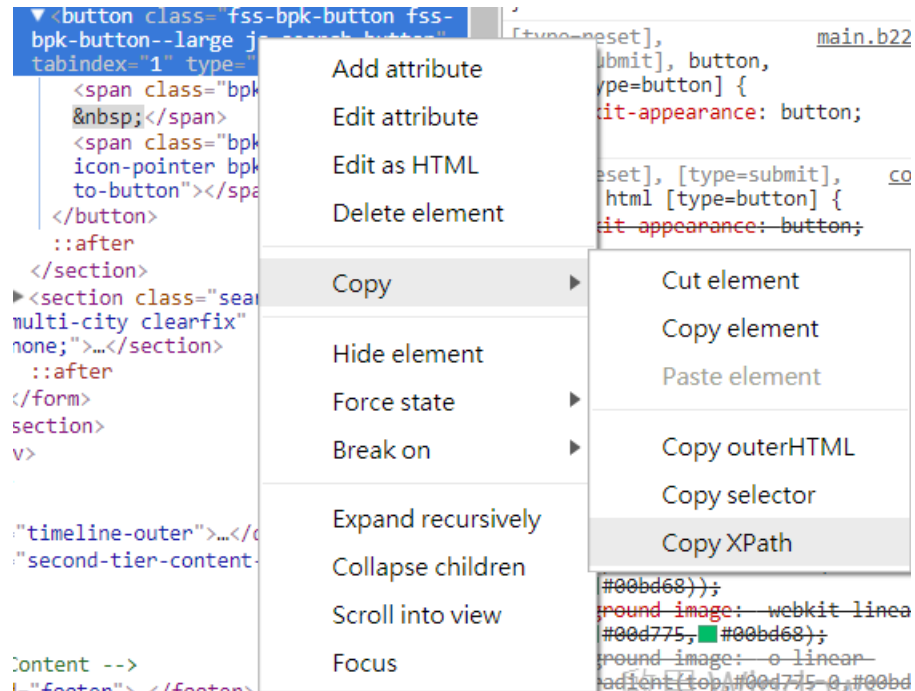
5. Selenium + chromedriver(二)

經過一連串好幾天的測試之後突然想到以往我都是直接製作網址然後讓瀏覽器直接開啟他來進行動作，我就在想會不會因為這是原本要設定的 cookie 跟一些不知道什麼的東西沒設定到，才會造成被發現是 bot 而被阻擋，因此我想說直接開啟官網(“https://www.skyscanner.com.tw/”)來進行自動化操作，這一想法成功奏效了，以往開一次被擋之後就要好幾個小時甚至到一天才能從新訪問網頁，改成開啟網頁之後大概就 55 波不會被擋，接下來開始進行自動化操作，利用 chrome 的工具抓取 xpath 來進行元件定位。

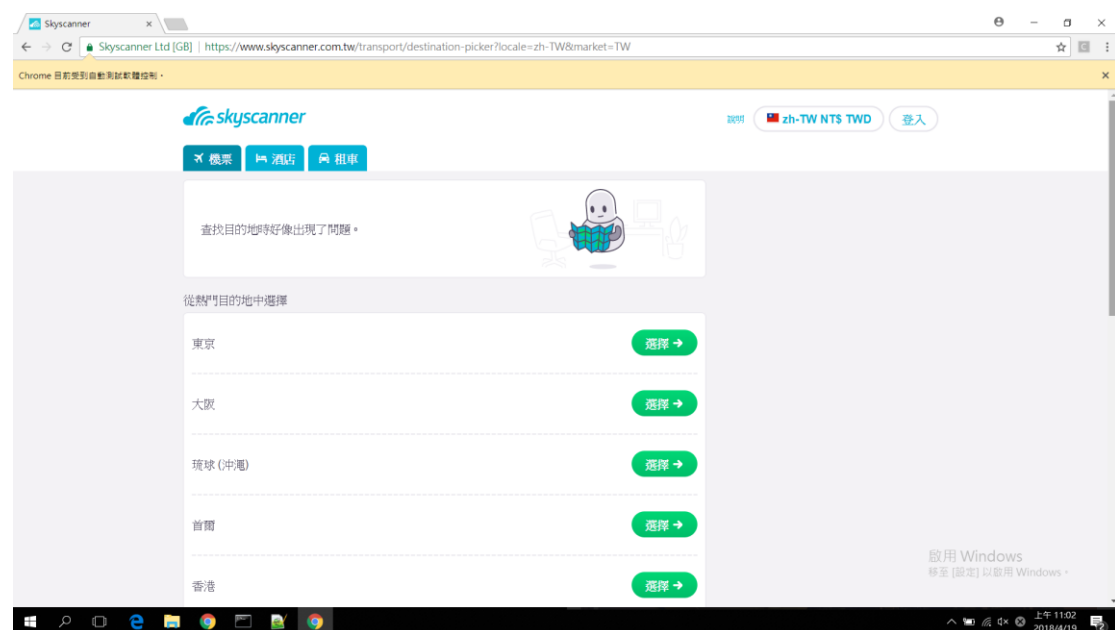
先到想要找尋的元件上案右鍵點選檢查



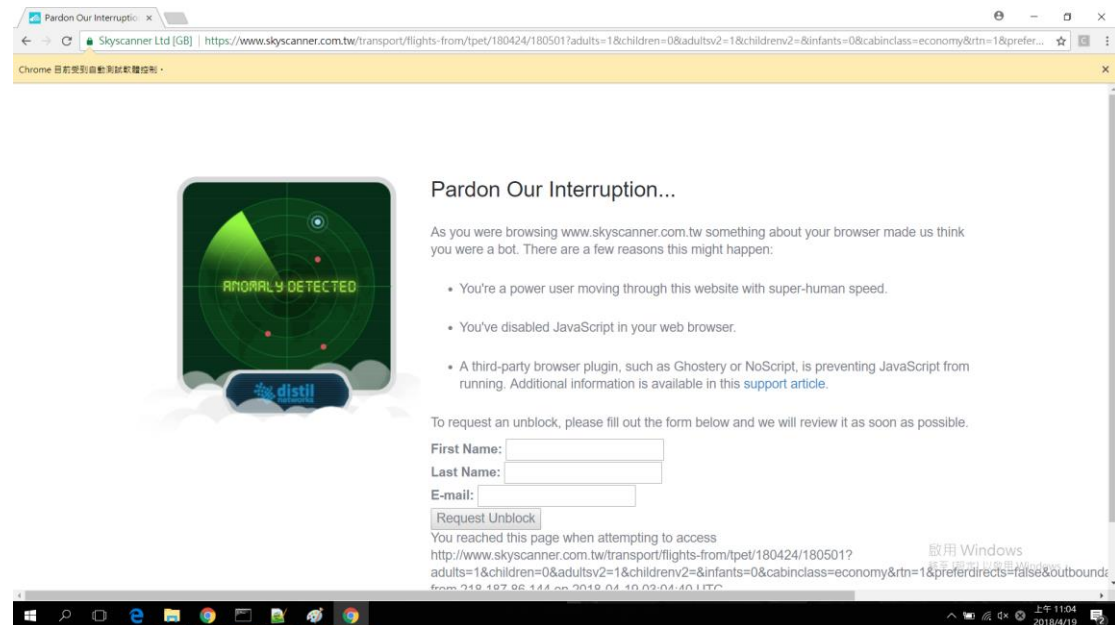
在剛剛的元件上案右鍵選 copy 然後在選 xpath，就可以得到那個元件的絕對位置



原本有成功進到首頁之後想說直接用滑鼠操作看看，但是畫面會顯示這種，會發現他網址有問題沒有顯示該顯示的網址內容



重試了幾次之後還會出現這種網址正常但被擋住的內容



但是除此之外更悲劇的是網址元件的 xpath 會每次都不一樣，原本用 chrome

抓取剛剛圖裡面的 button 的內容是這樣

```
//*[@id="js-search-controls-container"]/form/section[2]/button
```

但是用自動化程式抓的卻有兩個

```
//*[@id="js-destination-input"] 跟 //*[@id="destination-fsc-search"]
```

每次開起都會不一樣之外，就算成功抓到，網頁也不一定會成功進入。

訂票程式:

因為原本要用的方法就是上面網路爬蟲的最後一個方法，但是遇到元件位置不一樣，而且連網頁都進不去，根本就做不了，除此之外 skyscanner 有時候會搜到沒有票的情形還有上面說的票價不一樣的問題，故因此就不做了。

5. 心得

這次作業在選擇時原本以為一點都不難，想說都看過網頁原始碼，看起來就是基本的 html + javascript 的格式，沒想到光是從分析網址這裡就花了許多時間去用，從人數到小孩年齡分布然後跟變數的值的東西就不一樣，沒想到這還不是最難的，網路爬蟲這邊就花了我 2 3 個禮拜在試，因為一定失敗就要等好幾個小時以上，有時到會等到一天，而且不是說可以讓你試很多次而是試一次之後就直接被阻擋，浪費很多時間在用這邊，想到的方法都不成功，該鎖的鎖該檔的檔，以我目前的能力因該沒有能力破解，不過後來想想這種本來就已經算是個網路爬蟲網站，他能避開一些限制抓到其他航空公司的資料本身當然知道漏洞在哪，也當然知道要怎麼阻擋這種想要爬蟲的使用者。