遇到的問題:

標題:網頁設計是"太優良"or"太沒有邏輯" ?

Agoda:

1.網頁讀取需要滾動視窗:

(1)因訂房網設計成需要滾動頁面讀取下半部資料:

解決辦法:使用Selinium

(2)用滑鼠滾動太慢了:

解決辦法: driver.find\_element\_by\_xpath('//body').send\_keys(Keys.END)

需要from selenium.webdriver.common.keys import Keys

(3)需要更注意time.sleep的設置:

滾動式網頁，因頁面讀取的時間比較長，若沒有設置time.sleep，容易導致element找不到、無法點擊的問題，例如按鈕幽靈化

解決辦法:time.sleep()秒數長一點

還是不行就使用，driver.execute\_script("arguments[0].click(),button")

(4)頁面讀取常常卡住:

Selenium讀取頁面預設為normal，需要整頁loading完成才開始執行

解決辦法:

selenium.webdriver.common.desired\_capabilities import DesiredCapabilities

desired\_capabilities = DesiredCapabilities.CHROME

desired\_capabilities["pageLoadStrategy"] = "eager"

driver=webdriver.Chrome("C:\\Users\\Yuet\\Downloads\\chromedriver.exe")

可以設定成eager（互動式）或 none:

2.彈跳式廣告

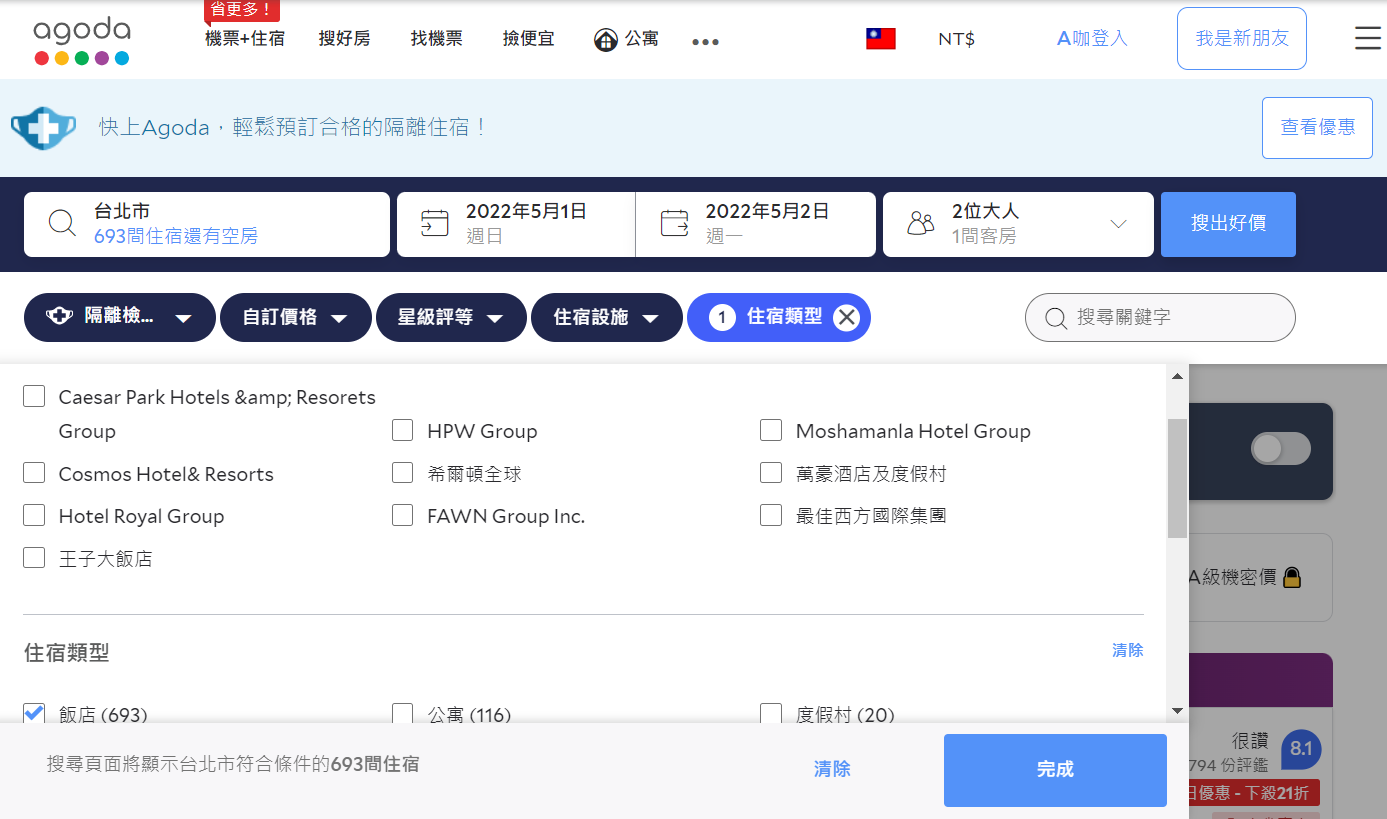
(1)防疫需求諮詢:

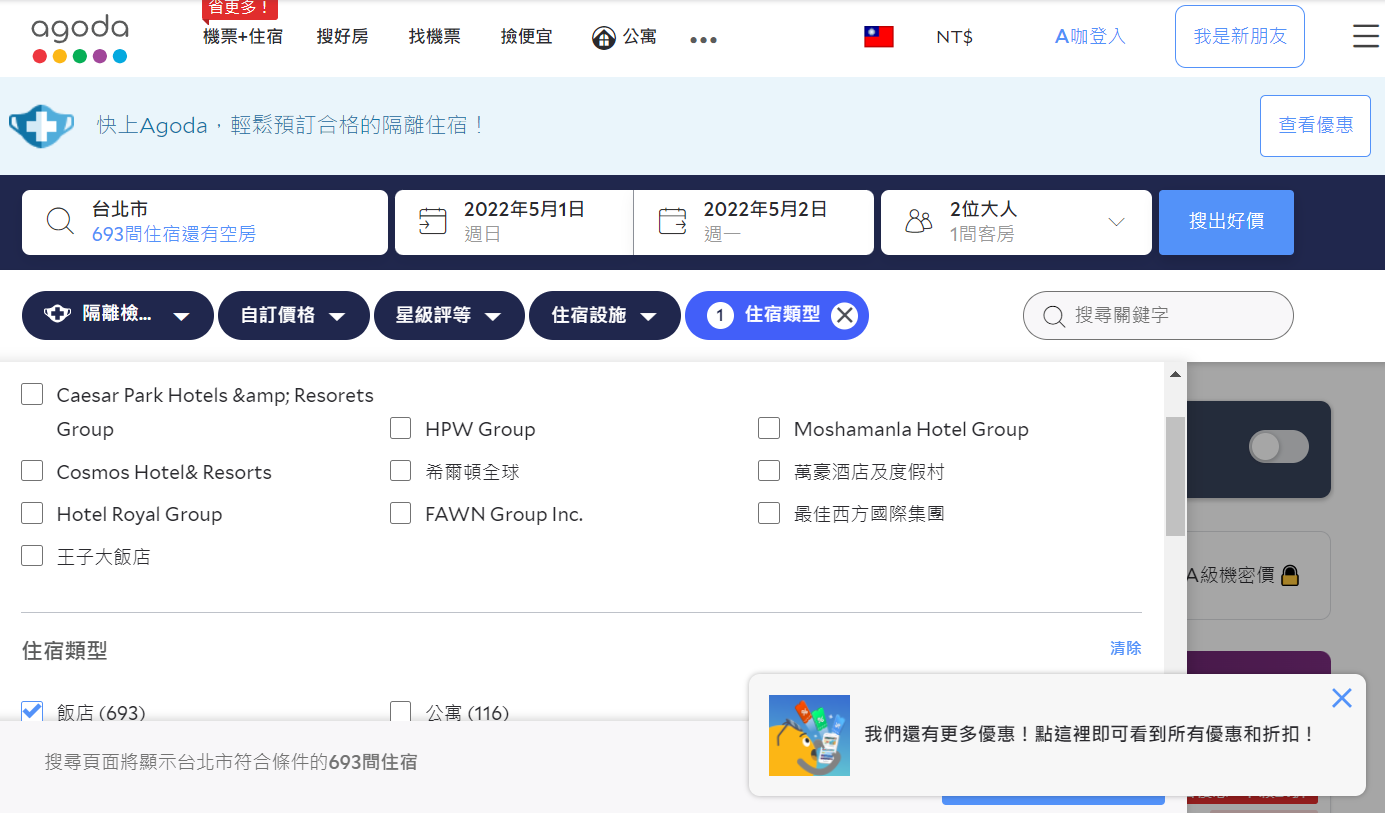


(2)廣告擋住按鈕:

WebDriver會拒絕點擊看不到的element，廣告擋住就無法點，

解決辦法: driver.execute\_script("arguments[0].click(),button")





3.查詢空房有誤:

查詢日期並沒有空房，但是飯店資訊一樣會出現



解決辦法: try except

但出現NameError: name "NoSuchElementException" is not defined

需要特別 from selenium.common.exceptions import NoSuchElementException

4.地址:

(1)格式不統一:

a.多數的形式:"秀山街4號, 西門町, 台北市, 台灣, 10042"

b.中式地址:"臺北市中正區忠孝東路二段89號3樓(入住櫃台請至３樓), 中正區, 台北市, 台灣, 100013"

c.中文+英文地址:"No. 7, Lane 83, Section 1, Zhongshan North Road, 中山區, 台北市, 台灣, 1044"

(2)地址有誤:

a. 位於西門町的飯店，多數地址沒有顯示區，只能用replace



b.縣市有誤，如位於新北市，但地址顯示台北市，data clean用篩選處理