

eupoly

Day 22 REGULAR EXPRESSION

空氣污染網站爬蟲實作練習



出題教練:張維元





本日知識點目標

● 能夠利用 selenium + BeautifulSoup 撰寫動態爬蟲

空氣污染網站爬蟲實作練習

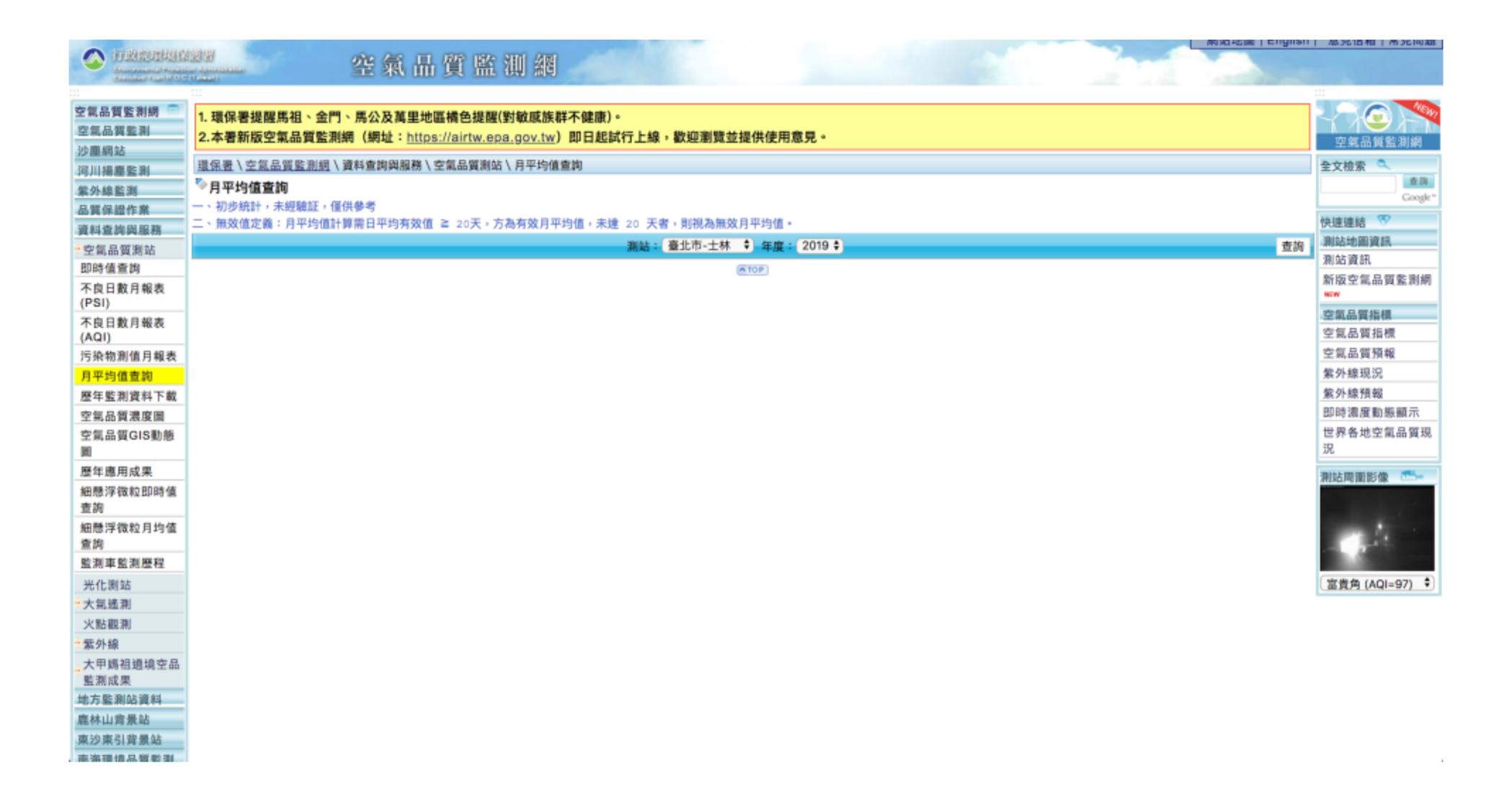


● 目標網站:空氣品質監測網

目標網址: https://taqm.epa.gov.tw/taqm/tw/MonthlyAverage.aspx

● 目標內容:抓取「監測項目」的表格

●本日為實作練習,建議搭配範例程式閱讀

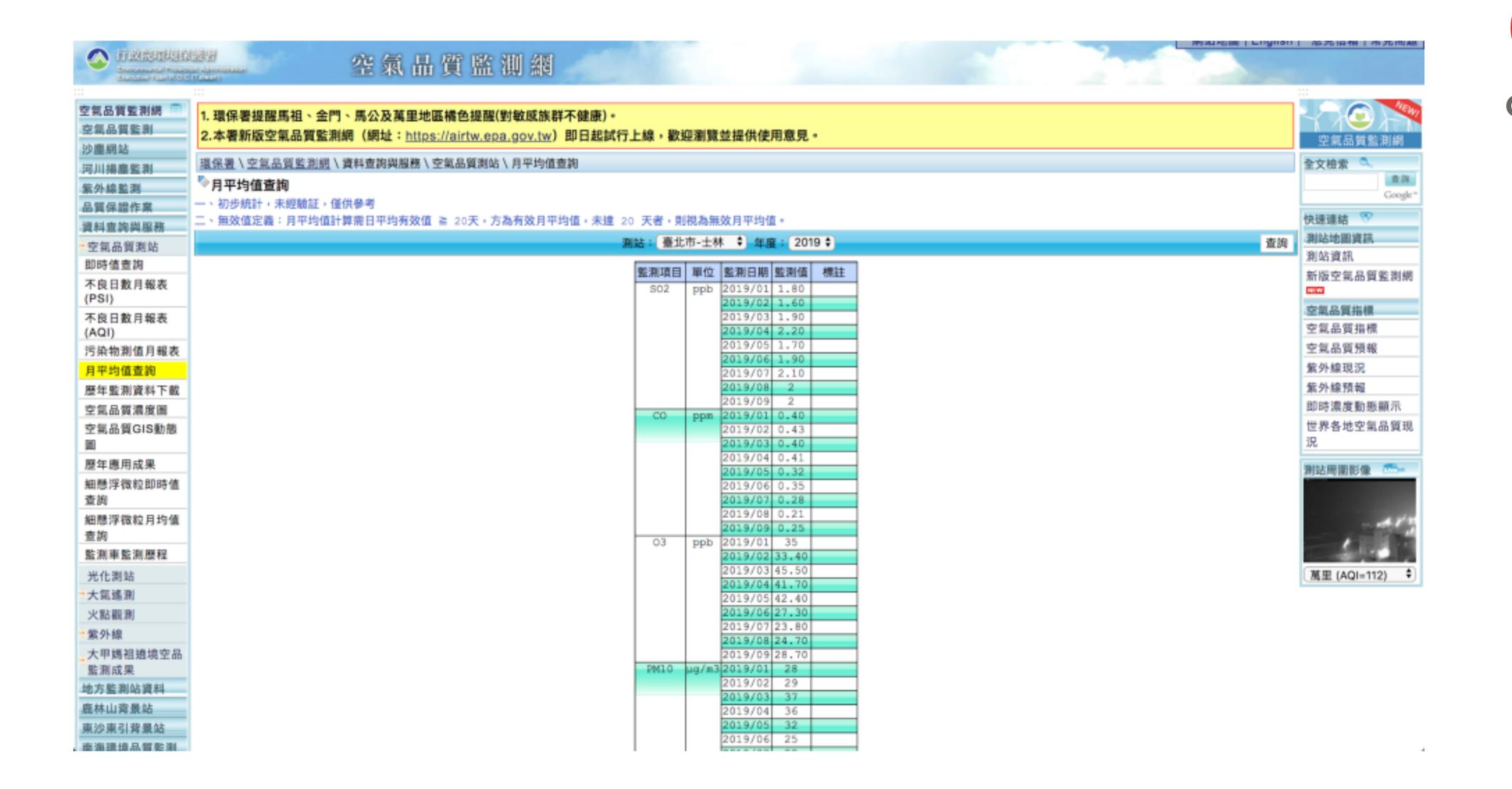


這是我們想要爬取的網頁,我們發現他一開始是沒有資料的

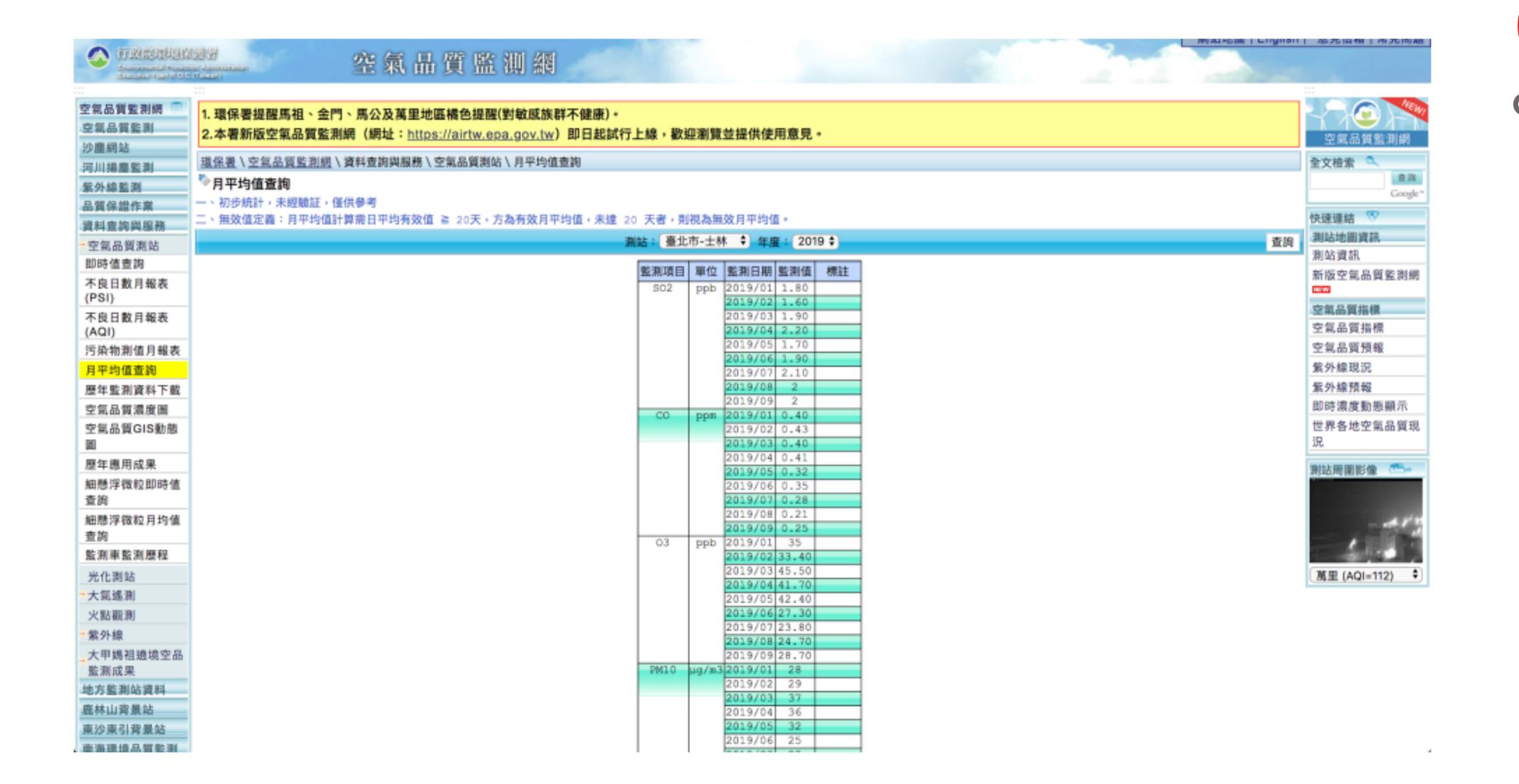




點擊查詢後,才會載入資料



而且載入的過程中網址也沒有改變,因此屬於動態網頁

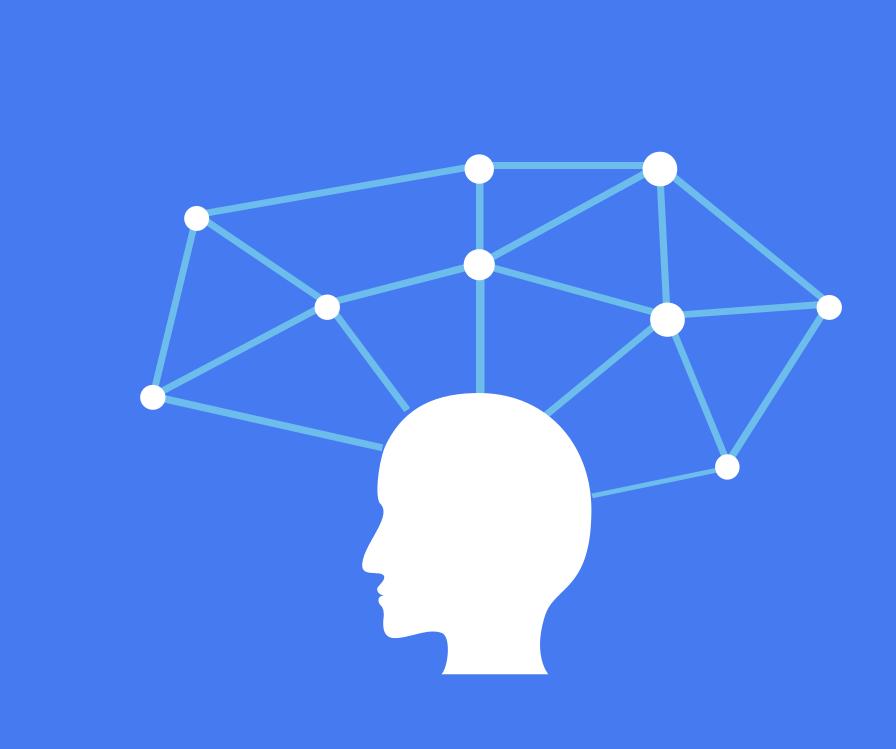


那我們一起跟著範例看這種需要使用者互動的動態爬蟲該怎麼處理?

重要知識點複習



● 能夠利用 selenium + BeautifulSoup 撰寫動態爬蟲



解題時間 LET'S CRACK IT

請跳出 PDF 至官網 Sample Code &作業開始解題

