1. 開發環境：MacOS version==11.2.3 using Python=3.7
2. 程式邏輯：

Google search(可自行輸入關鍵字跟頁數) -> google news button -> 爬蟲標題以及連結做儲存

INSIDE news website -> 爬蟲所有大標題以及連結儲存

1. 函式功能說明：

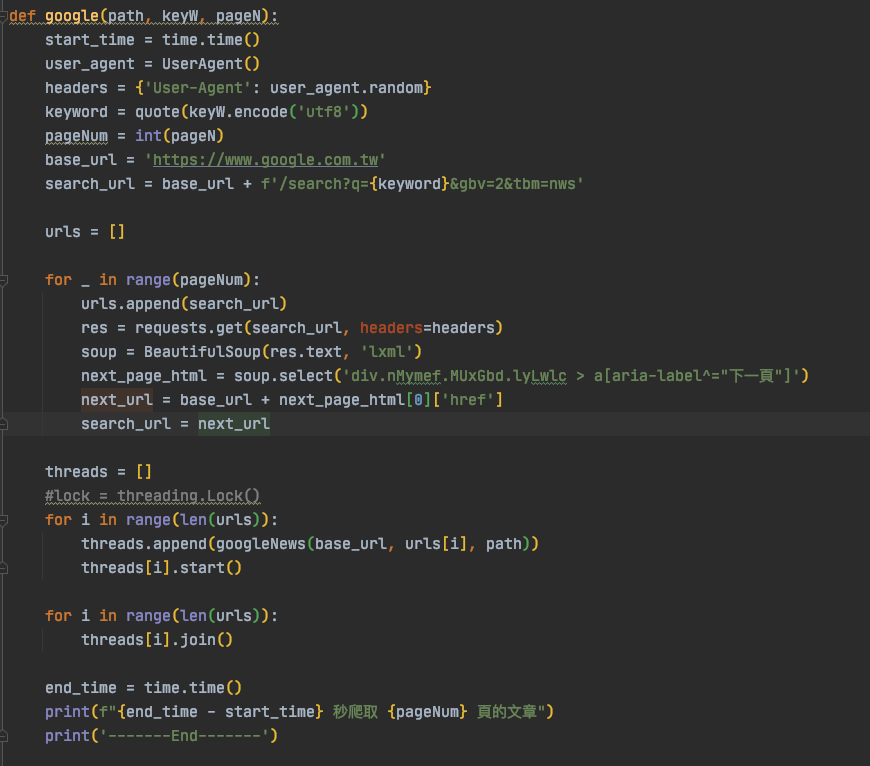
GoogleNews.py:

Google這個function先去偽造user\_agent以避免多次爬蟲被鎖ip,

再來讓使用者可以自行在GUI中輸入關鍵字和所要爬的頁數，

我先將所有每一頁(搜尋頁)的url存在urls這個list當中，為了能讓我丟給threads時能夠平行運行

丟給threads幫我們運行

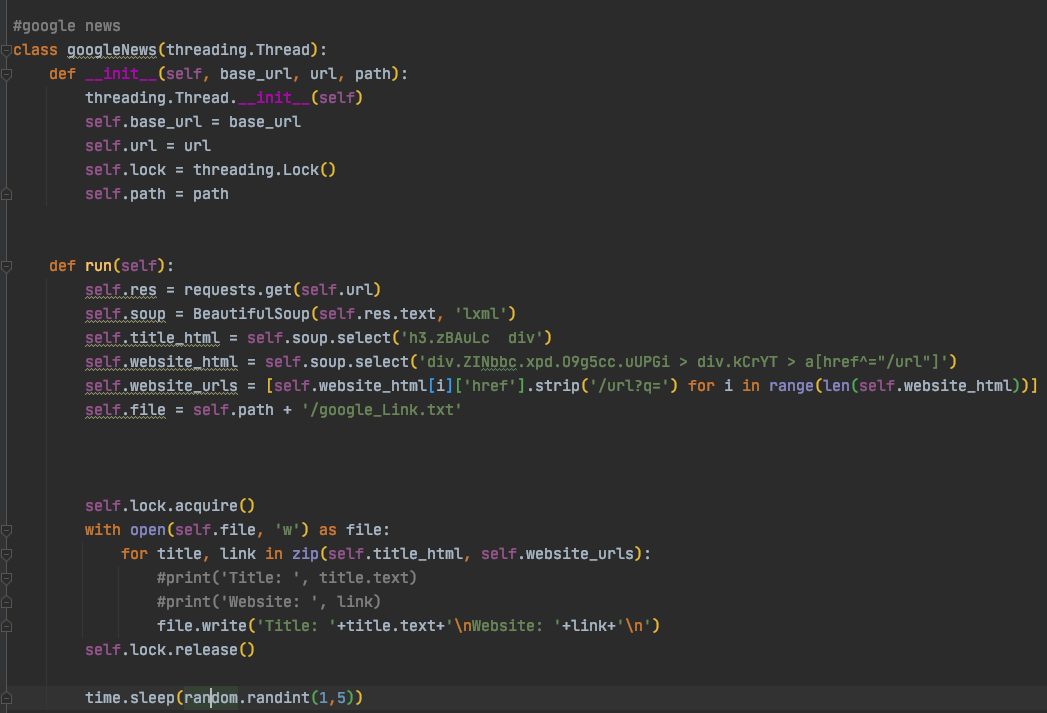


利用物件導向的方式寫threads，

我將我的每一頁的url給傳入，

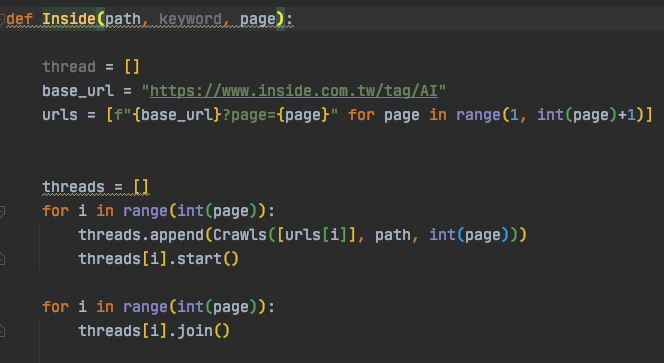
在這里利用beautifulsoup來爬我的標籤，

最後輸出title以及連結並且因為要連續寫入，所以利用lock來能夠一比一比鎖住，最後輸出給使用者想要的路徑．

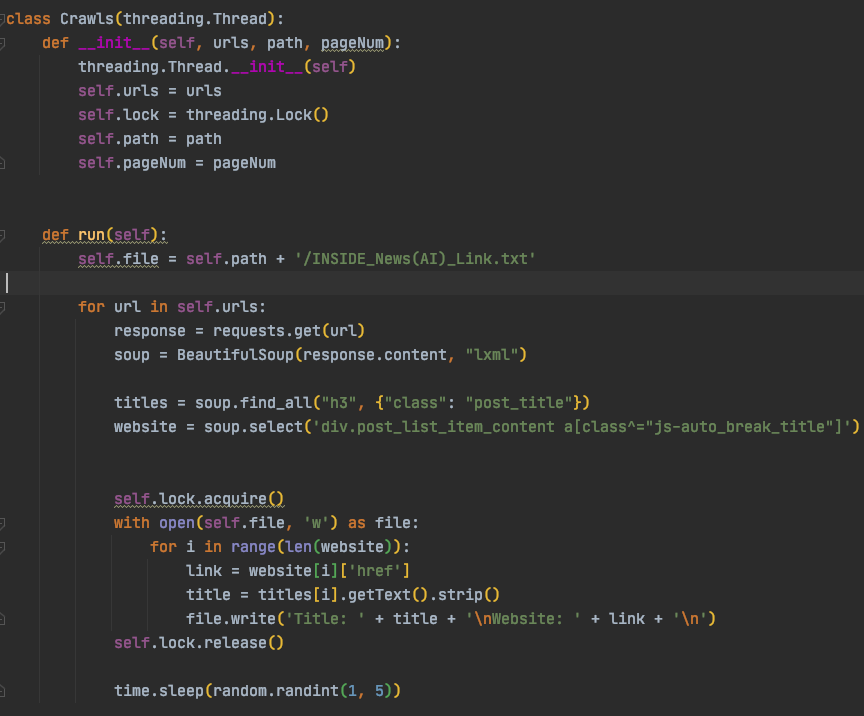


INSIDENews.py:

這裡的function大同小異，一樣是透過url的方式，將多個urls丟給threads去做執行



這裡和上面google news的也幾乎一樣，會去依照使用者選取的路徑儲存檔案，並且利用beautifulsoup來爬我的標籤，最後將我的title以及website給儲存．



GUI.ipynb:

我自行設計的GUI

利用tkinter以及物件導向的概念去寫



在這些functions中，會依照點擊的地方而有反應，

如print\_selection會在上方黃色地方顯示點擊了哪種新聞，

exitWindow則是點擊exit時，會結束GUI

searchPage則會先需要使用者輸入完東西以後，並且跳出對話框，要選擇哪個路徑做儲存，並且依照所選的新聞網站來儲存．

