網頁原始碼: 靜態, 為沒有執行JS的結果

python沒法執行JS語法,無法取代瀏覽器只能控制瀏覽器,

替代request套件改用selenium套件

凡request都能改用selenium

browser=webdriver.Chrome(): 建構物件(程序)

browser.get(): 開啟網頁

browser.page_source: 回傳目前開啟網頁之原始碼,

也就是r.text;不用加header,因為瀏覽器是親自上陣

browser.find_element(By.ID, 'id'): 通常用在找單一元素(id唯一)

browser.send_keys(): 對元素傳送文字,

或是Keys.ENTER送enter按鈕

程式速度>瀏覽器載入速度,會抓不到,需等待

無頭模式: 背景模式,

browser.add_argument('--headless')

下拉式選單: <select name="">,回傳<option value="">

(u""): unicode

指向性等待: 等到了結果再執行

時間性等待: 強制等待,在時間內等待, sleep()

隱性等待: 最常等待時間(30秒), 整個程式都結束了

顯性等待: 最多20秒, 每隔0.5秒看一次;

間格時間最小就是0.5秒

selenium可加表頭,控制瀏覽器告知服務端為手機模式

應用程式封裝成執行檔.exe

檢查是否為主程式,是則執行window():

if __name__ == '__main__':
window()