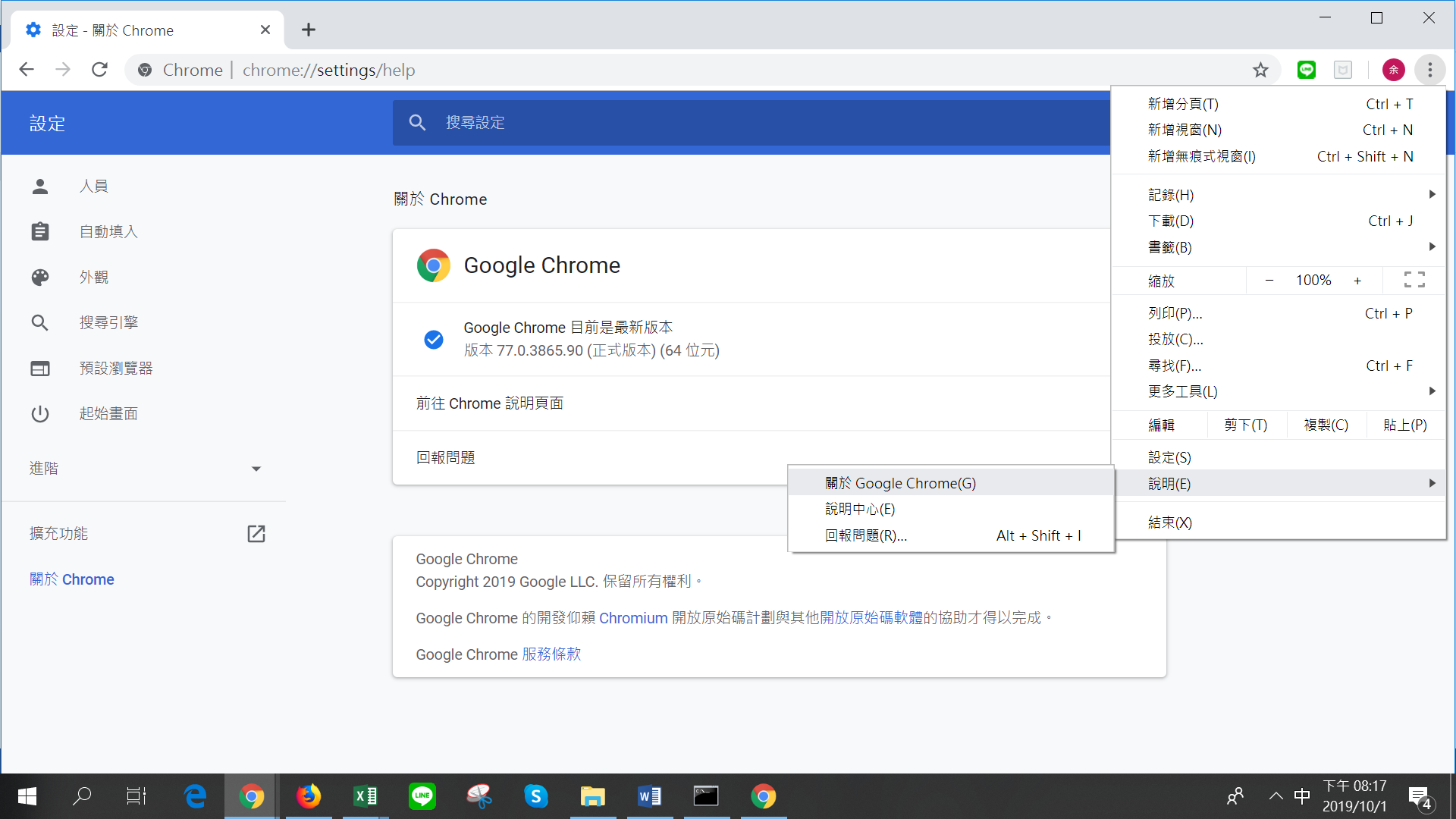
**一、跑一下範例程式: 從鉅亨網頭條新聞撈出最近一些新聞文字資料**

1. 看本機chrome版本, 下載該版本的chrome driver



1. 取代modules裡面的chromedriver.exe
2. 文字資料取回後, 直接開啟會看到亂碼, 解法如下:

從開始功能表點選EXCEL, 開啟空白工作表,

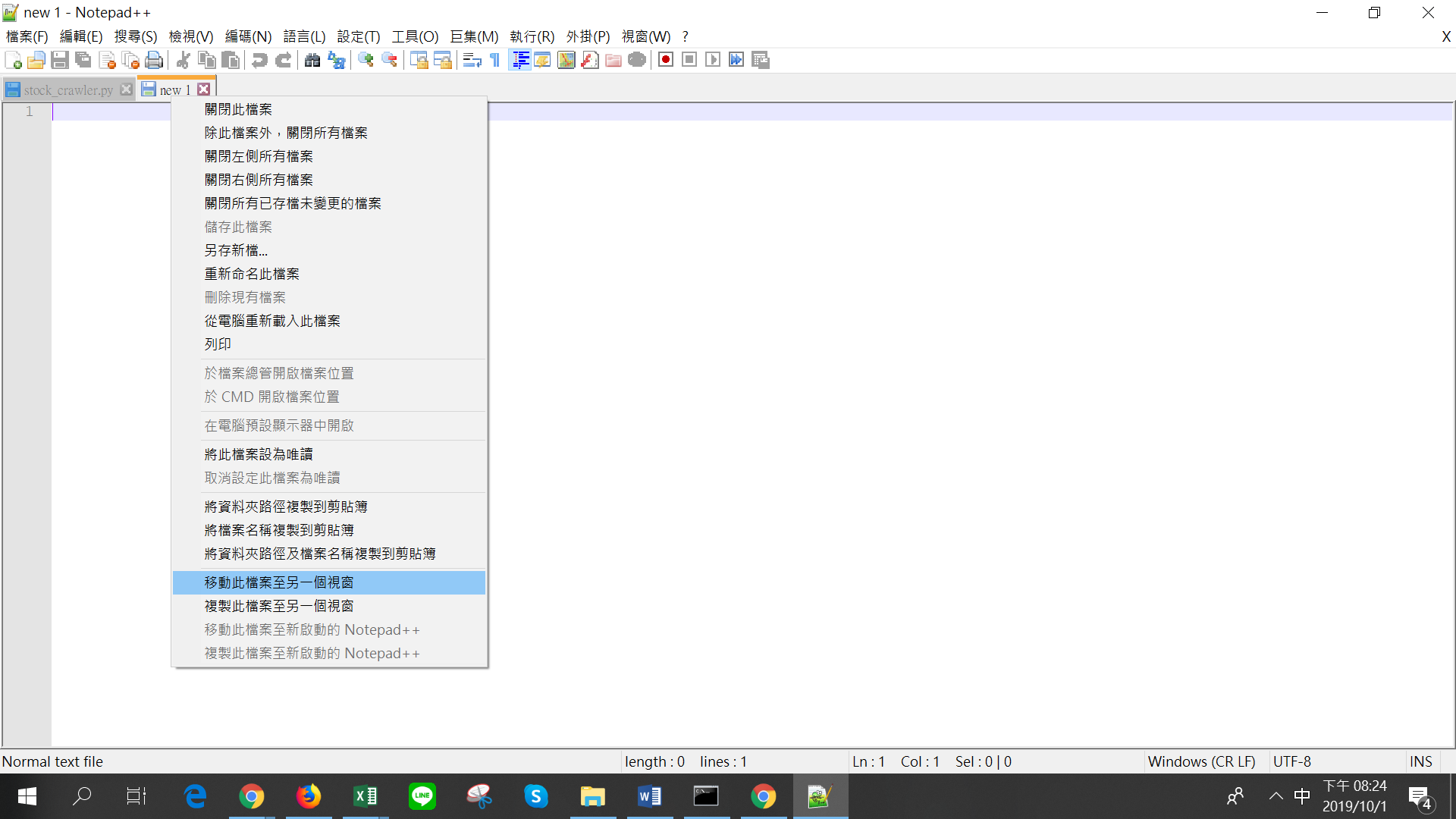
點選資料->取得外部資料->從文字檔->點選result.csv -> 匯入

->點選UTF-8, 分隔符號 -> 下一步, 分隔符號改為逗點

->下一步, 保留一般 -> 完成 -> 確認

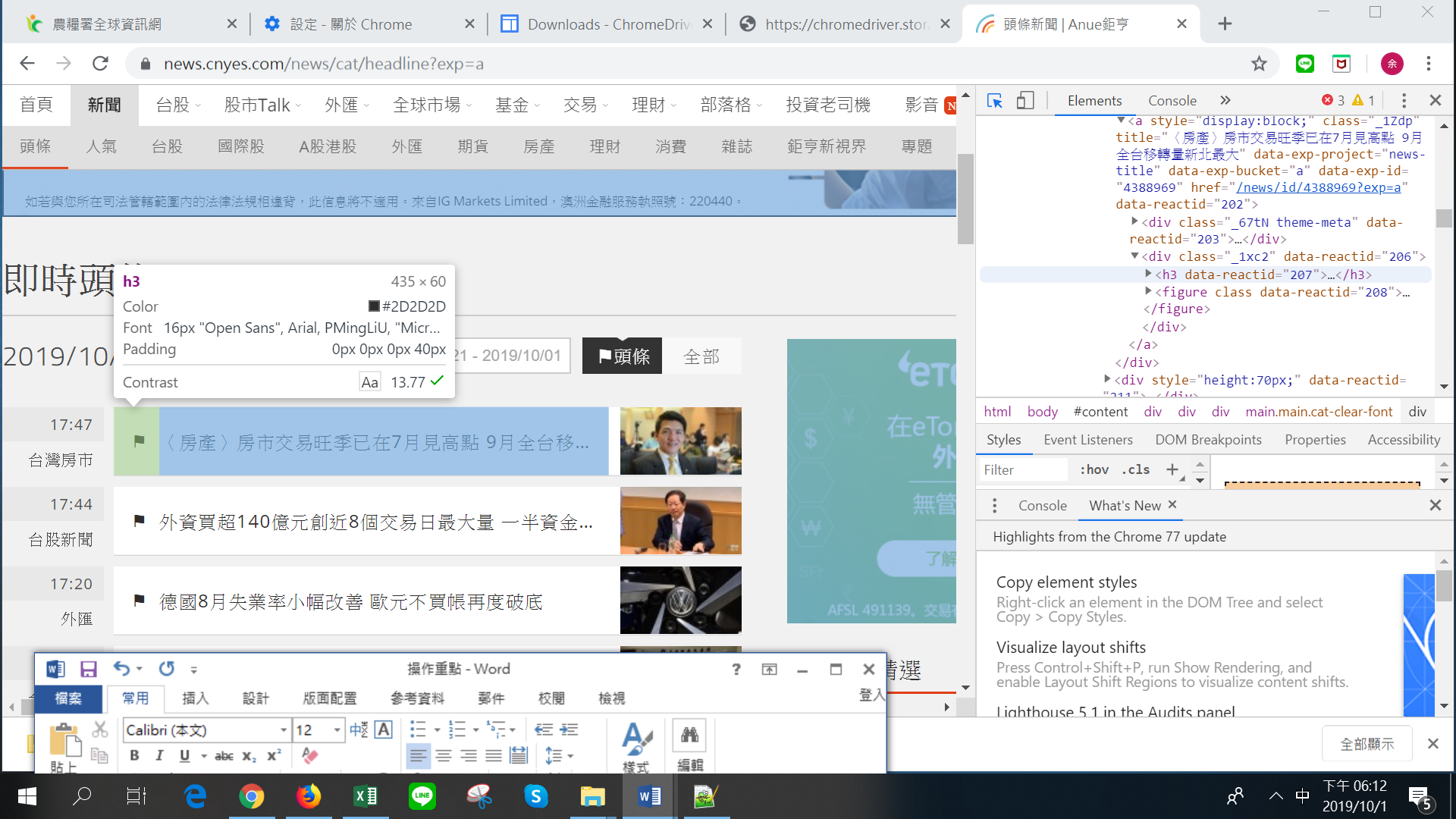
**二、了解範例程式:　可應用到各種網頁爬蟲**

1. 下載安裝Notepad++, 之後可在stock\_crawler.py檔案上按右鍵以Nodepad++開啟。另點選檔案->新增, 另新開空白頁, 配置到新視窗, 以便比對檢視。

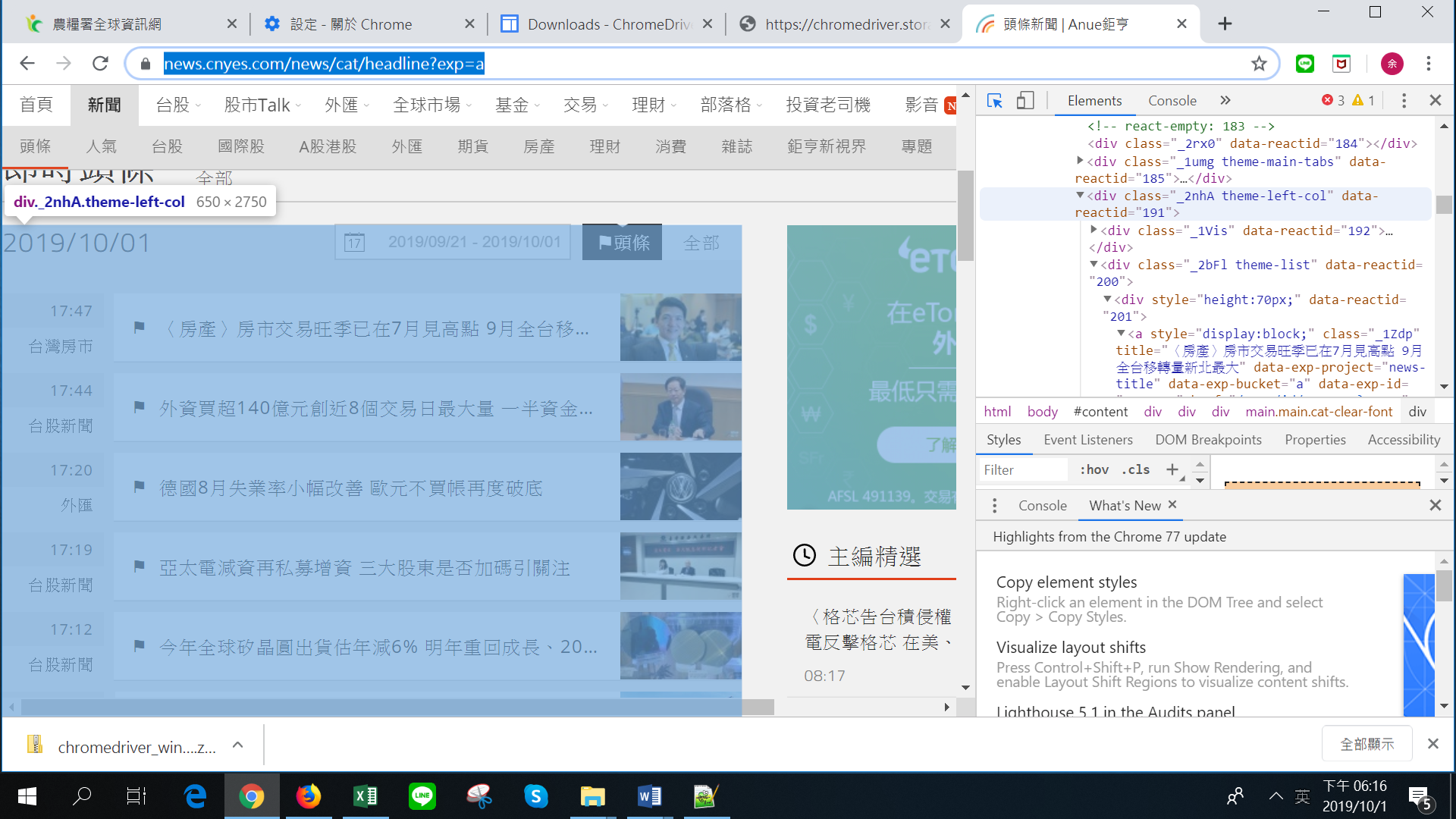


1. 到鉅亨網首頁->新聞頁, 可直接從程式中copy網址出來。

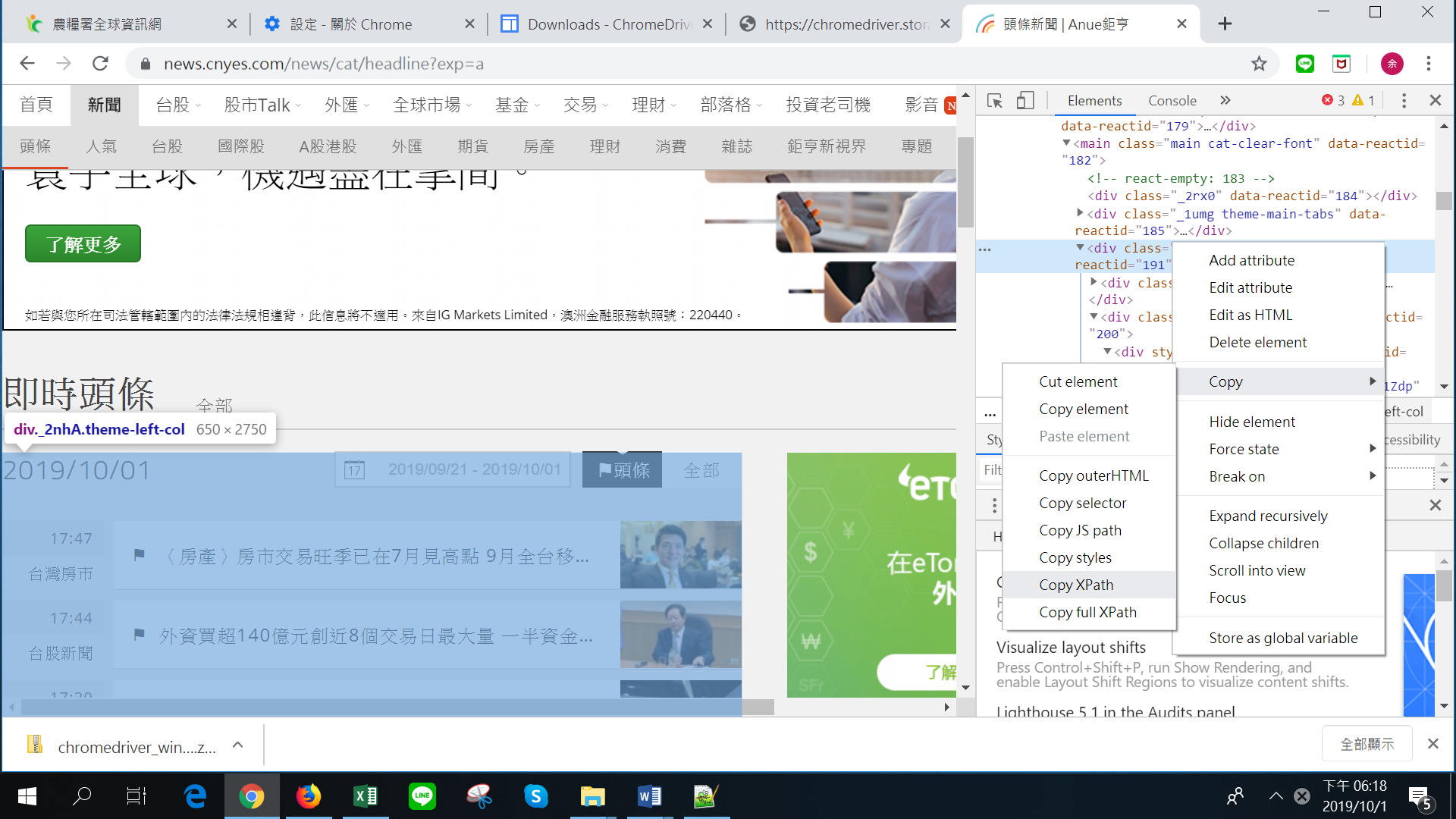
6. 到了新聞頁, 在瀏覽器上按F12, 按箭頭處, 可建立網頁與HTML程式的動態對應(跳走了要重新點選, 呈現藍色才會動態對應)。



7. 找到所需範圍的對應程式片段, 如下:



8. 在所需範圍上點一下可選定右邊程式片段, 不再動態移動, 在選定程式片段上按右鍵, 可以複製XPath, 如下圖。



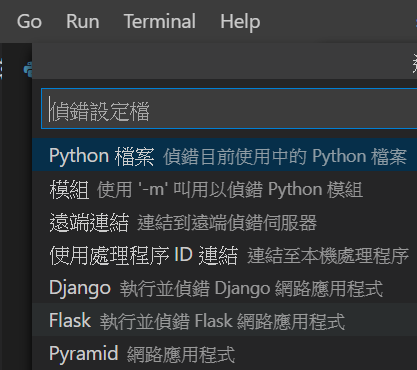
9. 核對stock\_crawler.py程式中#取得新聞資料部分的driver.find\_elements\_by\_xpath(…)參數是否與網站現況, 若不一致則修改之。

10. 可以下載安裝Visual Studio Code檢視程式

10-1. Open Folder, 設定程式所在資料夾

10-2. 設定程式執行斷點

10-3. F5(Run): 下拉選第2個-python, 跑到斷點停止



10-4. F10(Step): 逐步執行可以看已執行過的變數內容。

提醒: Debug時無法編輯(加註解等)

上面有小指引, 可以看一下

