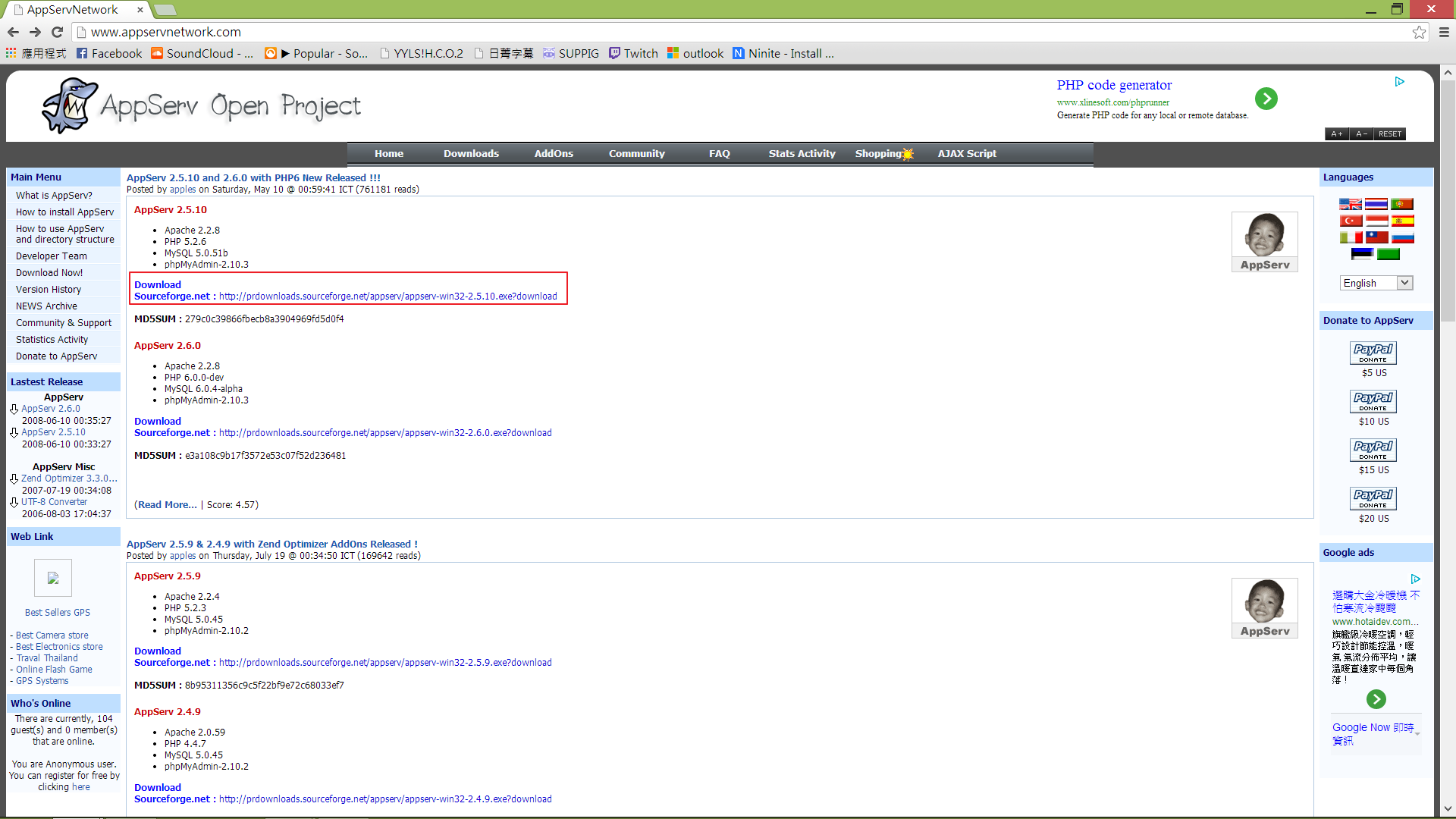
## 環境配置

### 下載Appserv

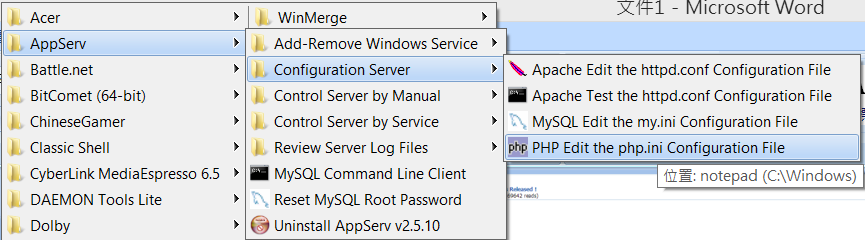
<http://www.appservnetwork.com/>

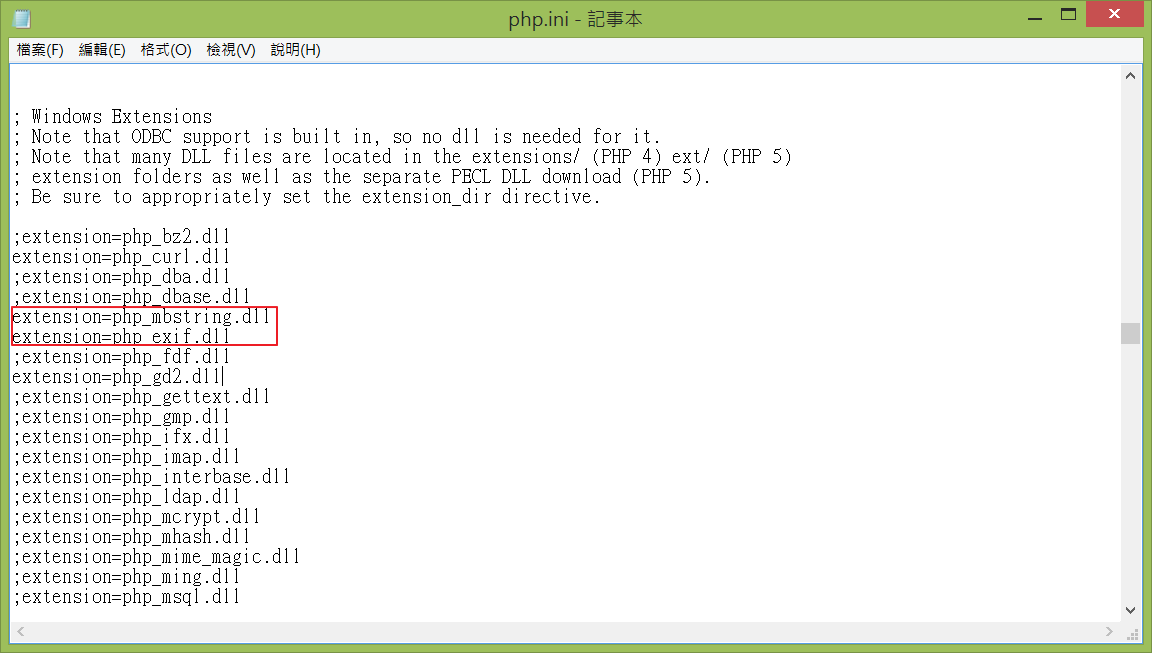


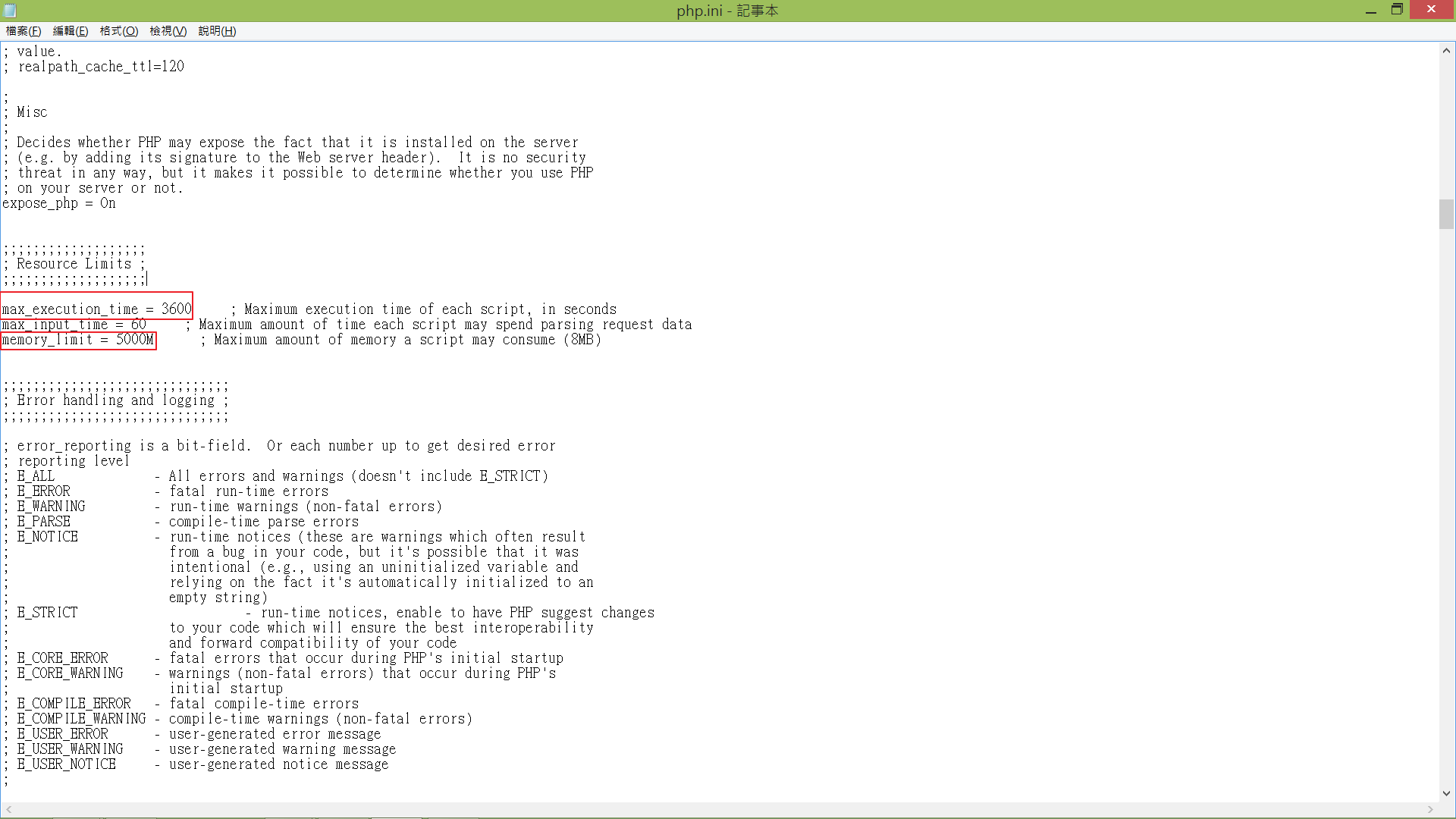
### 設定config

* php.ini:

1. 從附屬應用程式裡找到並開啟php.ini
2. 找到extension=php\_mbstring.dll，移動到extension=php\_exif.dll的前面
3. 找到max\_execution\_time，設定為3600
4. 找到memory\_limit，設定為5000M







### 將程式碼放進去

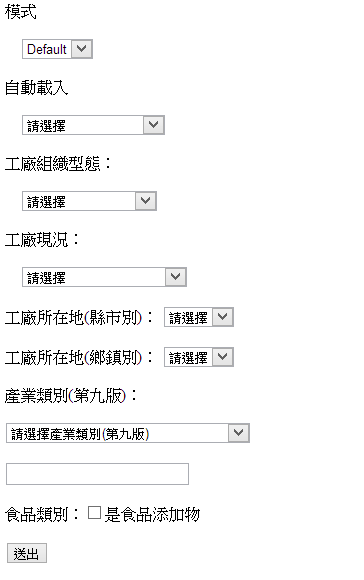
找到%appserv\_home%\www\ 將WebCrawler資料夾放進去

Appserv預設安裝路徑為C:\appserv所以就是會在C:\appserv\www底下

## 網頁版

### 抓取URL

1. 連接<http://localhost/WebCrawler/index.php>



1. 模式[Default]: 抓取url

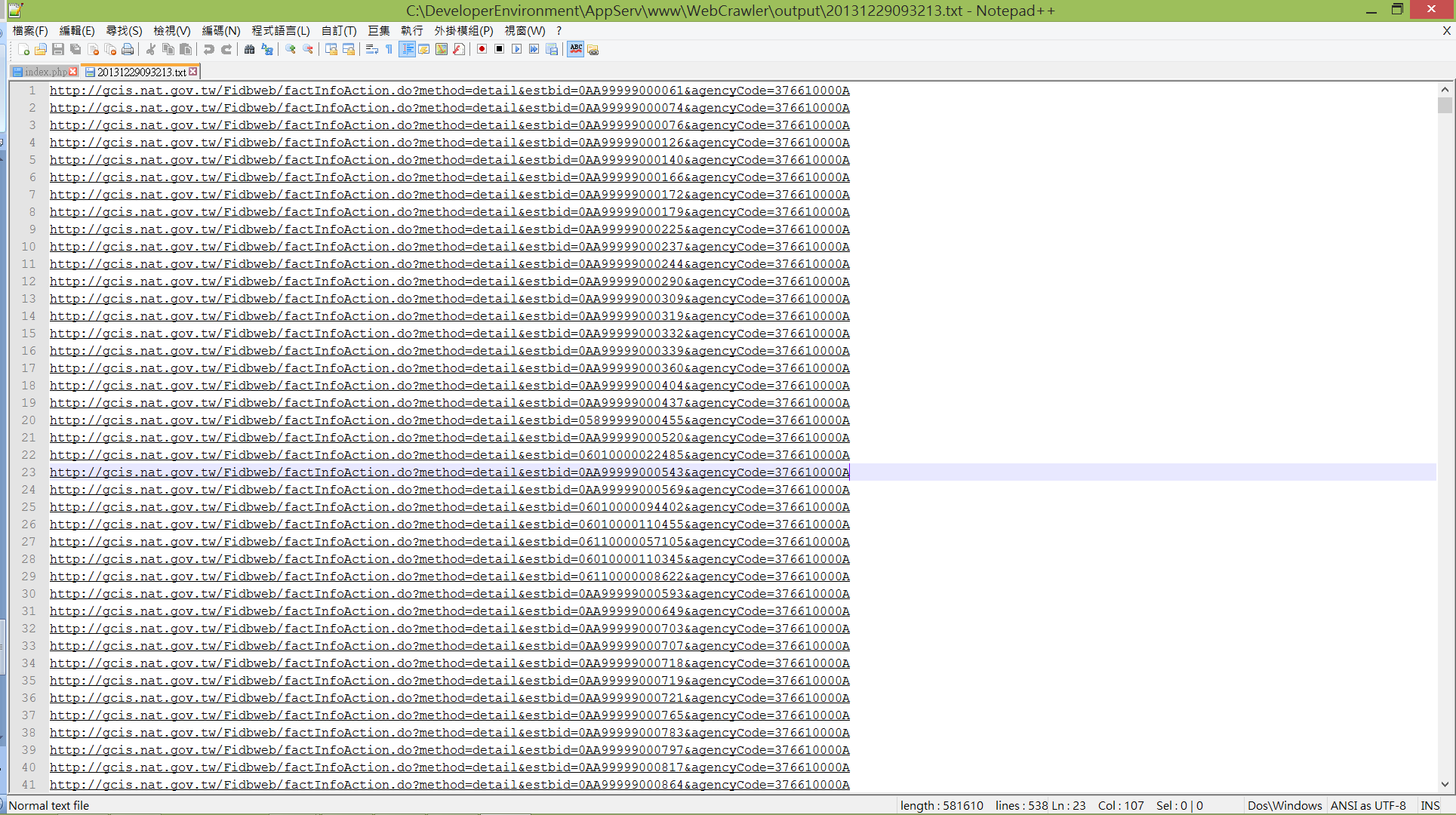
模式[Count]: 選擇工廠所在地(縣市別)，抓取數量

1. 自動載入[請選擇]: default，請從下方選擇三個條件去查詢

自動載入[工廠所在地(鄉鎮別)]:選擇工廠所在地(縣市別)，並從工廠所在地以外其他條件則一個選擇

自動載入[工廠現狀]:從工廠現狀以外選擇兩個條件查詢

1. 按送出產生.txt檔案



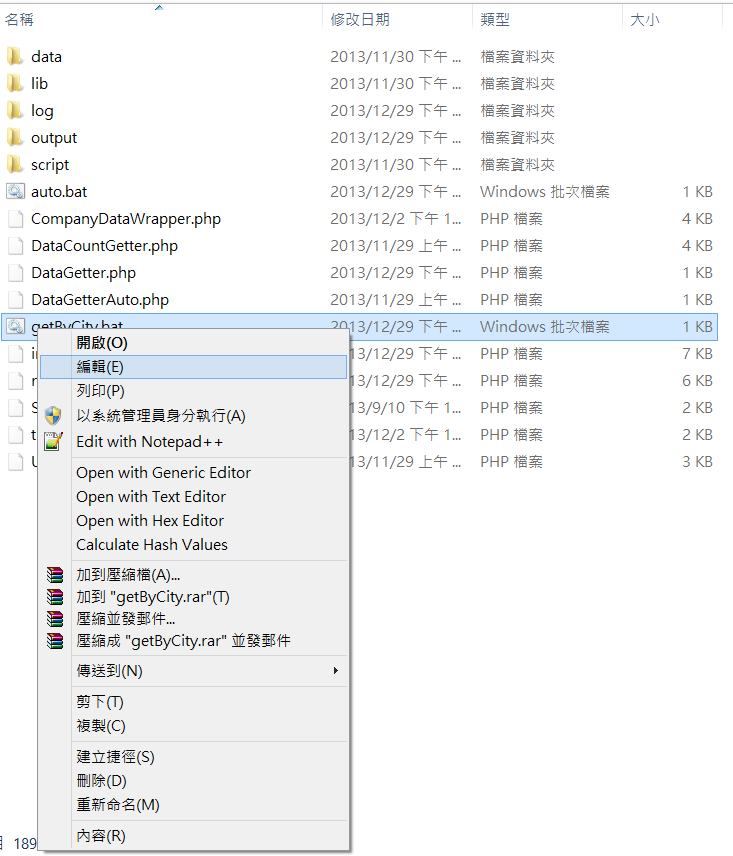
### 轉換URL變成資料

1. 連接<http://localhost/WebCrawler/transform.php>
2. 將剛剛下載之txt檔案上傳

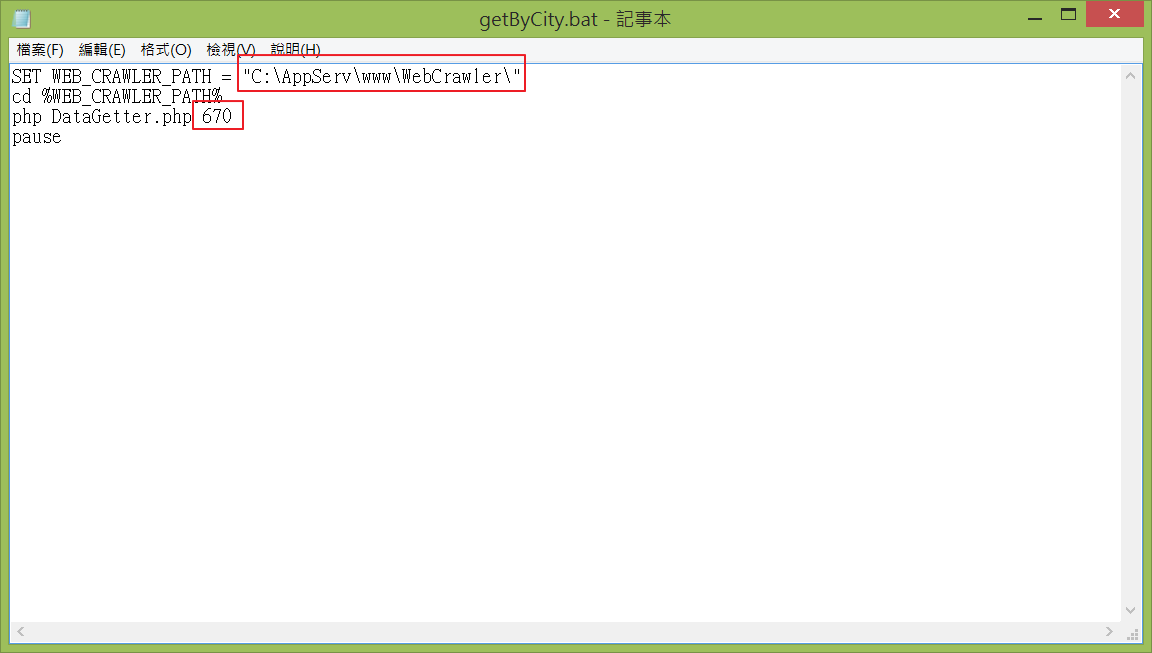
## Bat執行版

### getByCity.bat

1. 在WebCrawler資料夾底下有getByCity.bat
2. 右鍵編輯



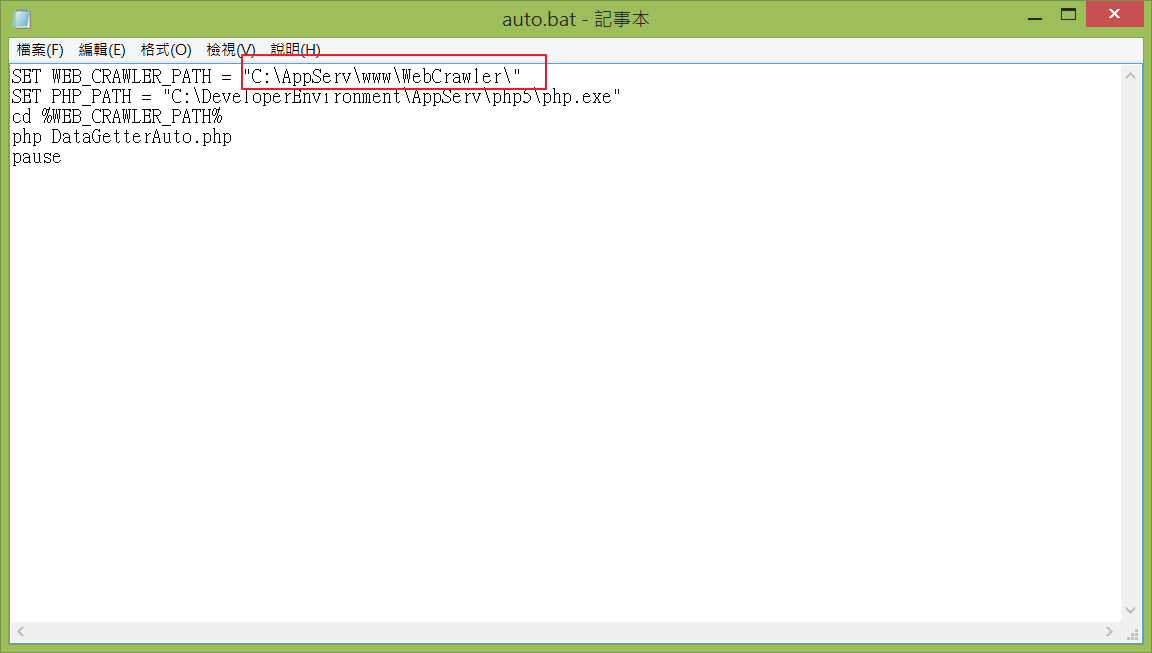
1. WEB\_CRAWLER\_PATH設定為WEB\_CRAWLER資料夾位置
2. 第三行後面數字打要搜尋的城市碼，數字可參考附錄



1. 點兩下執行bat檔案
2. 產生的檔案在output資料夾底下，檔名為city\_${cityCode}.txt

### Auto.bat

1. 在WebCrawler資料夾底下有auto.bat
2. 右鍵編輯
3. WEB\_CRAWLER\_PATH設定為WEB\_CRAWLER資料夾位置



1. 點兩下執行bat檔案，會根據上次執行的城市跑下一個城市
2. 產生的檔案在output資料夾底下，檔名為city\_${cityCode}.txt

## Windows排程

## 附錄

### 城市碼一覽

"630" => "臺北市"

"10017" => "基隆市"

"650" => "新北市"

"10002" => "宜蘭縣"

"10004" => "新竹縣"

"10018" => "新竹市"

"10003" => "桃園縣"

"10005" => "苗栗縣"

"660" => "臺中市"

"10007" => "彰化縣"

"10008" => "南投縣"

"10010" => "嘉義縣"

"10020" => "嘉義市"

"10009" => "雲林縣"

"670" => "臺南市"

"640" => "高雄市"

"10016" => "澎湖縣"

"10013" => "屏東縣"

"10014" => "台東縣"

"10015" => "花蓮縣"

"09020" => "金門縣"

"09007" => "連江縣"