**Intelligent Devices and Cloud Computing**

**Final Project Report**

鄭理文 王以彥

1. **題目 :**

Real-time Parking Lots Assistant

1. **成員 :**

r05942103 王以彥 : 負責資料蒐集、處理、架設spark cluster

r05942080 鄭理文 : 負責後端管理、web ui、app設計及呈現

1. **摘要 :**

我們的project構想，是想幫助在路上行車時，常常會遇到想停車但不知道附近哪裡有停車場，或是開到停車場卻發現大排長龍沒有停車位，因此設計了web和app介面，結合停車場資訊以及google map，在地圖上呈現停車場位置和資訊，讓使用者能即時知道附近停車場狀況。

1. **作法 :**

**4-1 資料來源 :**

資料來源採用政府資料公開平台-新北市路外公共停車場資訊、新北市公有路外停車場即時賸餘車位數兩份公開資料，網址見附錄，原本預計採用新北市及台北市停車場資訊，但因網路上公開資訊中，台北市停車場並沒有隨時更新的開放資料，故最後只採用新北市停車場資訊(每3~5分鐘平台上的停車場資訊會做一次更新)。

**4-2 系統架構 :**

1. 使用python每分鐘為頻率至政府公開平台上抓取最新資料。
2. 於google cloud platform 上，使用google dataproc建立一個spark cluster，包含4 CPU master以及2個各2 CPU worker。
3. 資料經過spark處理後，存入mongodb，並用node.js、Express管理後端部分，本次project mongodb架設在實驗室server上。
4. 結合google map api，使用HTML、CSS、JavaScript建立web ui，使用swift建立ios app。

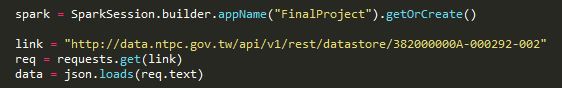
**4-3 實作細節 - 存取靜態資訊 :**

首先先處理的是停車場資訊靜態資料，也就是只須在一開始更新一次即可，其中包含停車場ID、名稱、地區、地址、費率、經緯度、車位總數、營業時間等固定資訊，實作方法為附檔中的save\_static\_info.py，利用python的request，向資料網站請求資料，因原始資料中的座標系統為TWD97資料，故需轉換成常用的經緯度，接著使用pymongo package，將所有資料直接存入mongodb server。

**4-4 實作細節 - spark即時資料處理 :**

實作方法為練習使用pyspark對資料進行處理，再藉由mongo-spark connector將資料存入mongodb，如附檔spark\_final.py。

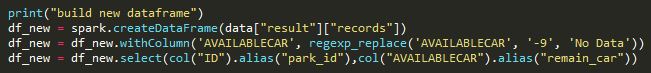
一、首先建立spark session，並至公開平台抓取即時剩餘車位資料



二、從mongodb中讀取靜態停車場資訊，儲存至dataframe

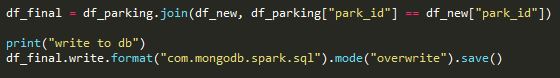


三、處理即時車位資料，資料中包含停車場ID及剩餘車位，有部分停車場並沒有提供即時剩餘車位，在資料中是顯示-9，轉換為”No Data”，並重新命名column name。



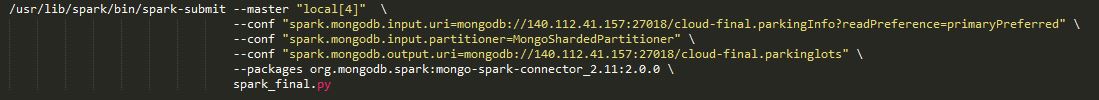
四、將靜態資料和動態資料結合，最後經由mongo-spark connector存

入mongodb中。

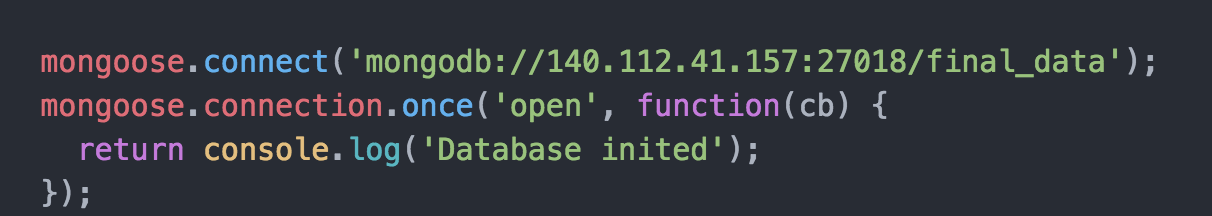


**4-5 實作細節 – 即時更新 :**

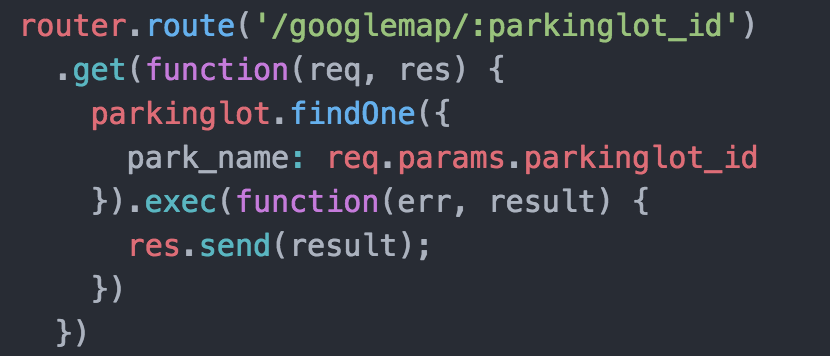
利用linux系統中的crontab工作排程，設定每分鐘執行一次script(附檔中的spark\_script.sh)，在script中設定mongodb的相關資訊，並設定mongo-spark package版本。



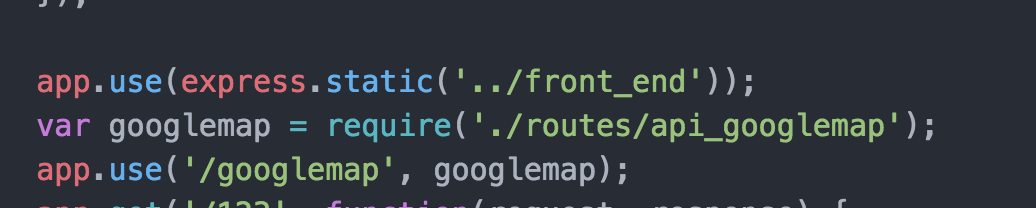
**4-6 實作細節 – 後端 :**

使用mongoose讀取MongoDB內的資料

在server內自行寫API 使Web和APP取得資料



透過Express 呈現Frontend的網頁



**4-7 實作細節 – 前端 web :**

利用JQuery取得Backend寫好的API



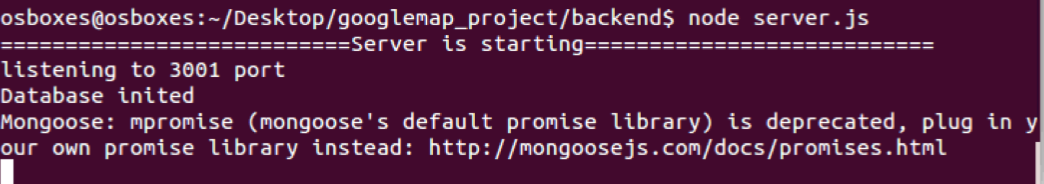
從Google API 申請而得到的token

/Users/liwencheng/Desktop/螢幕快照 2017-06-22 下午8.21.03.png

初始化設定GoogleMap執行環境

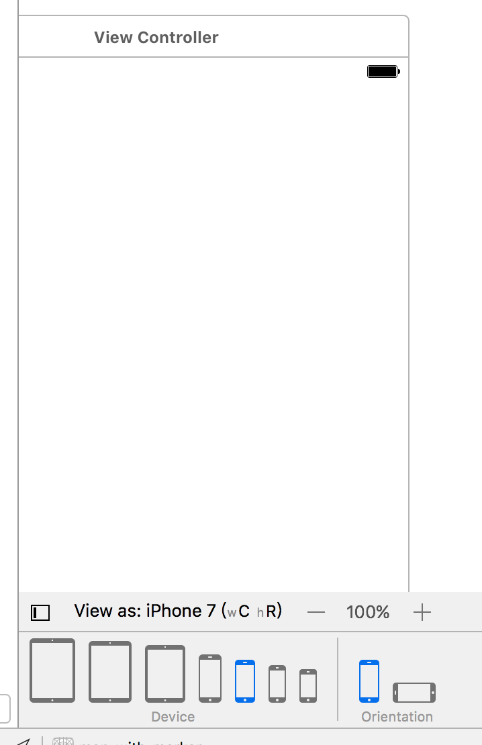


程式成功執行



**4-8 實作細節 – 前端 ios app :**

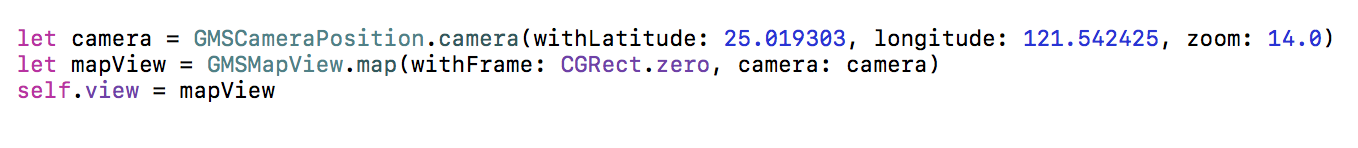
制定此APP運用的介面

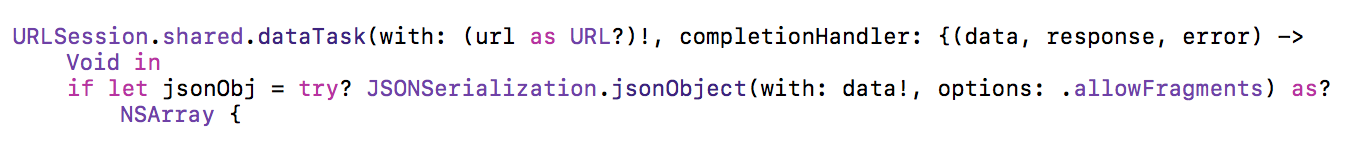


從Google API 申請而得到的token

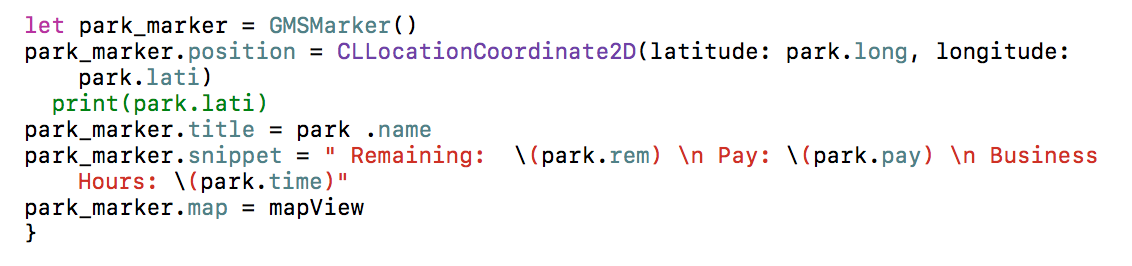


初始化設定GoogleMap執行環境



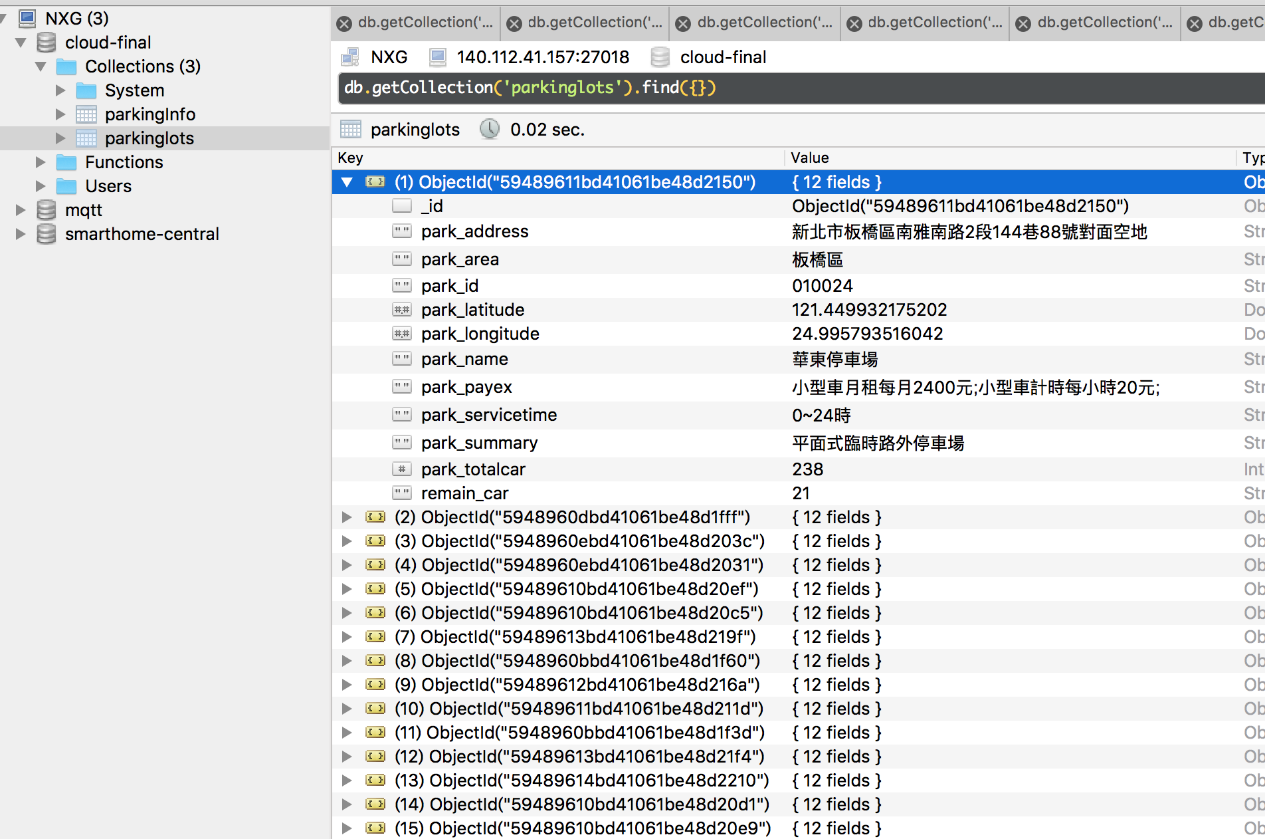
從Backend的API取得資料並確認為JSON格式

標記停車場位於的點



1. **成果**

於MongoDB database中確認所有資料正確無誤



web ui 成果如下圖，於google map上可看到我們標出的各個停車場標點



點擊標記可查看即時剩餘車位等即時資訊



app部份各個停車場地圖資訊如下圖



點擊標記可查看剩餘車位等即時資訊



1. **討論**

在這次的final project中，主要結合了之前作業中的項目，運用到了spark做資料處理，用mongodb做資料庫，並用HTML、CSS、JavaScript呈現網頁，以及swift建立ios app，主要遇到比較多問題是在如何運用google map api，尤其是app部分，因之前並沒有接觸過swift，在使用上遇到蠻多語法上的問題。

總結起來我們的project還有些許地方可以做改進 :

* 目前我們是設定google map起始座標畫面在台大博理館，可以改為使用使用者目前位置來做起始座標。
* 目前畫面上顯示的是所有停車場，可以嘗試只顯示使用者周遭幾公里內的停車場，或許可以讓使用者在顯示全部和顯示周遭之間切換。
* 可以考慮讓使用者對停車場做一些搜尋篩選的動作，例如搜尋目前還有車位或車位數大於一定值的停車場。
* 雖然台北市停車場並沒有動態即時資料，但有靜態停車場資訊，應該可以也將這些資訊放上google map，使用者雖然無法確認這些停車場的即時剩餘車位，但還是能知道台北市周遭哪邊會有停車場，以及其相關資訊。

1. **附錄及相關參考資料**

* 新北市公有路外停車場即時賸餘車位數 : <http://data.gov.tw/node/26701>
* 新北市路外公共停車場資訊 : <http://data.gov.tw/node/26653>
* mongo-spark connector :

<https://docs.mongodb.com/spark-connector/master/python-api/#pyspark-shell>

* run pyspark script with crontab :

<https://stackoverflow.com/questions/39604706/running-pyspark-using-cronjob-crontab>

* Setting up Google Cloud Dataproc with Python 3 :

<https://blog.sourced.tech/post/dataproc_jupyter/>