



國立臺東大學 理工學院

NATIONAL TAITUNG UNIVERSITY COLLEGE of SCIENCE and ENGINEERING

104 學年度第二學期 教學增能計畫

資料處理與分析 成果報告書

課程名稱：資料處理與分析

開課系所：應用數學系

執行人：高嘉宏

中華民國 105 年 6 月 23 日



一、 課程主旨與目標

- **課程主旨：**資料科學或資料工程領域，在近年來有著相當蓬勃的發展；其中，巨量資料、開放資料、Open API、資料探勘、機器學習等主題，更受到學界與業界的大量關注，相對應的業界職缺更是大幅度的成長。對於應用數學系的同學來說，本身已具備了數學與統計的理論基礎，在資料處理與分析的領域上已有一定的優勢。然而，在資訊科技大幅度被應用的今日，若同學亦可掌握相關電腦技術與工具，更可成為資料科學或資料工程領域的專業人才。藉此，本學期開立的資料處理與分析課程，希望由電腦與工程的領域切入，讓同學了解在資料處理與分析的工作上，有哪些資訊科技可供使用，並協助特定工作的進行。另一方面，課程亦大幅使用了政府開放資料與 Open API 作為練習的基礎，希望同學可透過實作進行學習與實務應用上的創意發想。最後，亦鼓勵與協助同學參與學校、組織或企業的開放資料或創新應用競賽，累積同學未來研究或職場上的能力與優勢。
- **課程目標**
 - 資料處理與分析流程介紹
 - 巨量資料與開放資料的認識
 - 開放資料的認識、練習與應用
 - 地理資訊呈現（Google Map）的實際練習
 - 數據與統計圖表呈現（Google Charts）的實際練習
 - Open API 的認識與使用
 - Python 資料處理與分析工具（SciPy、NumPy、Matplotlib）的使用



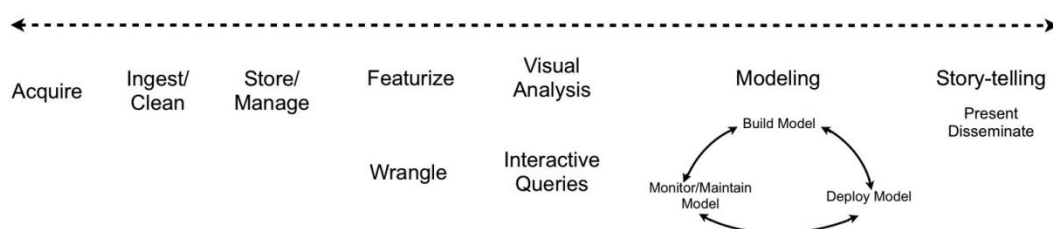
二、課程執行成果

(一)課程執行量化成果

項目(參考項目)	數量	內容說明
參與競賽或論文發表	1	參與臺東大學理工學院學習競賽，獲得佳作
增加學生動手實作或課程互動	16	每次上課均規劃學生實作時間，透過開放資料，撰寫程式，實際進行資料處理與分析工作
增加實務或業界實習參訪	1	邀請業界人士演講
強化學生學習動機及提升學習風氣	16	每次上課均補充最新社群活動、業界訊息、TED演講、資料處理與分析專案、競賽訊息等等，透過實際的案例激發學生的學習動機與風氣
開發數位教材或新教材	1	所有的教材均由教師自行製作，並公開於網路 (https://sites.google.com/site/nttudatacourse/)，學生皆可取得教材並練習
跨領域學習	1	基於數學與統計理論，結合電腦資訊領域的技術與工具，進行資料處理與分析的學習與實務操作
具體成果或作品	11	課堂期末專案以分組方式進行，共產出 11 組作品，而其中一組同學亦在 5 月參與理工學院學習競賽，獲得佳作

(二)課程整體執行質化成果

- 對於資料處理與分析的全盤認識：介紹資料處理與分析流程，使得同學對於資料處理與分析工作有一完整的認識。另一方面，流程中的任一階段，所需使用到的知識與相關工具均作介紹，使得同學有更為深入的了解，對於日後研究或職場工作奠定全盤的知識基礎。



Big Data Now: 2013 Edition, O'Reilly, Feb. 2014



Becoming a Data Scientist – Curriculum via Metromap

- 對於開放資料與 Open API 的認識與實際應用：在課堂中大幅使用開放資料與 Open API 作為實務的資料處理與分析操作。透過實際資料的練習，可讓同學處理常見的資料格式（如 CSV、XML、JSON 等等），並輔以程式語言的撰寫，練習後續的處理與分析工作。另一方面，也透過 Open API 的練習，讓同學了解如何介接外部服務資源。藉此，同學亦可以熟知基於特定資料格式與 Open API 的應用開發工作，對於日後研究或職場工作能有相當的幫助。

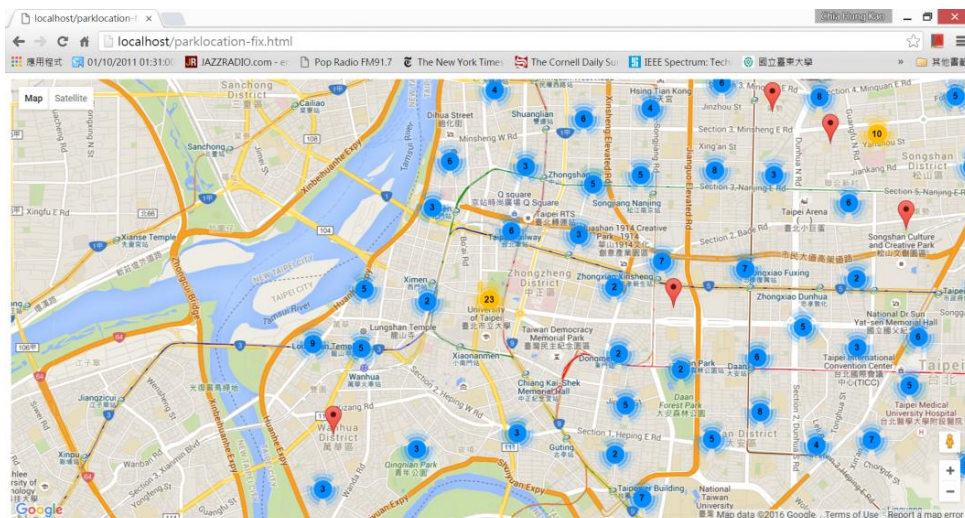


政府開放資料平台

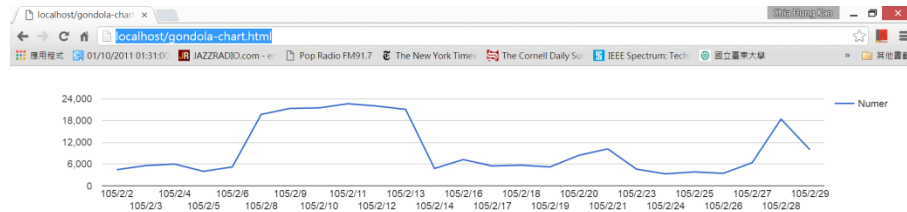


Open API 介接練習與使用

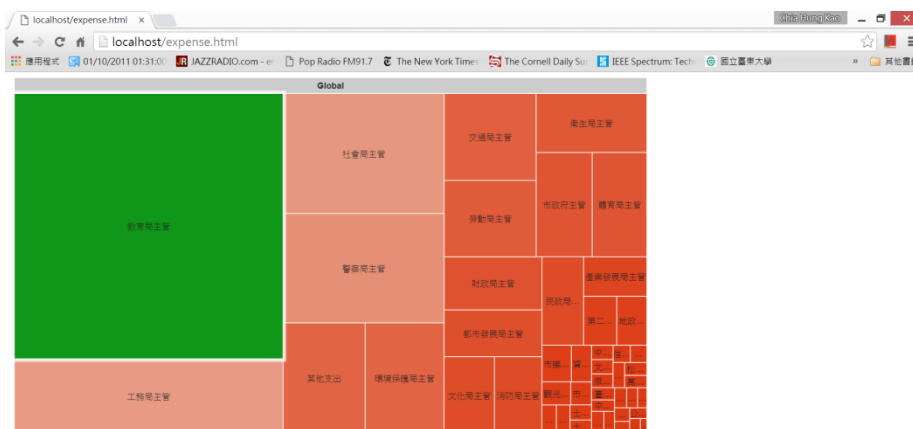
- 資料分析與呈現工具（Google Map 與 Google Charts）的使用：基於開放資料或 Open API 介接，再以 Google Map、Google Charts 等工具進行地理資訊與統計圖表的呈現，可結合同學在數學與統計的理論基礎，進行進一步的資料處理與分析工作。而此跨領域的整合，亦可讓同學了解與實際練習數學與資訊領域的工作。



同學課堂練習：基於開放資料透過 Google Map 呈現地理資訊

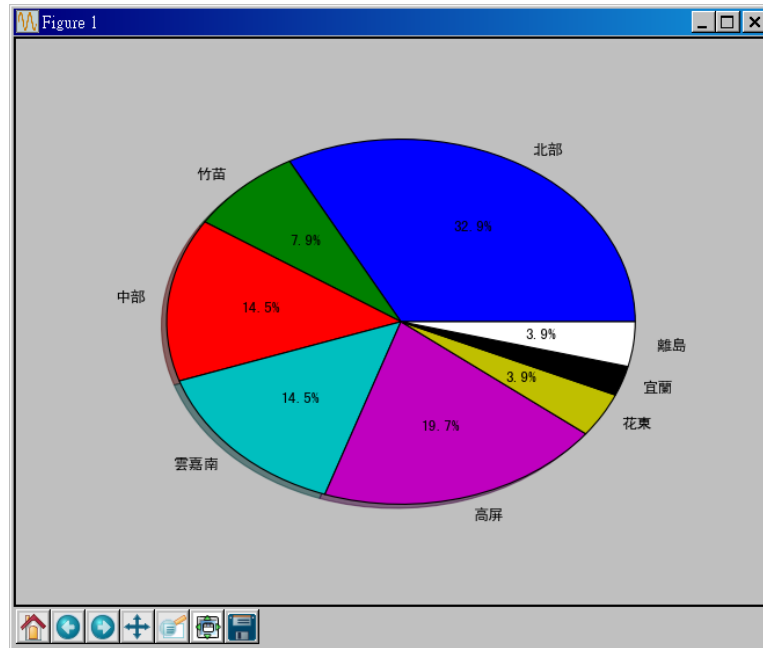


同學課堂練習：基於開放資料透過 Line Chart 呈現數值紀錄

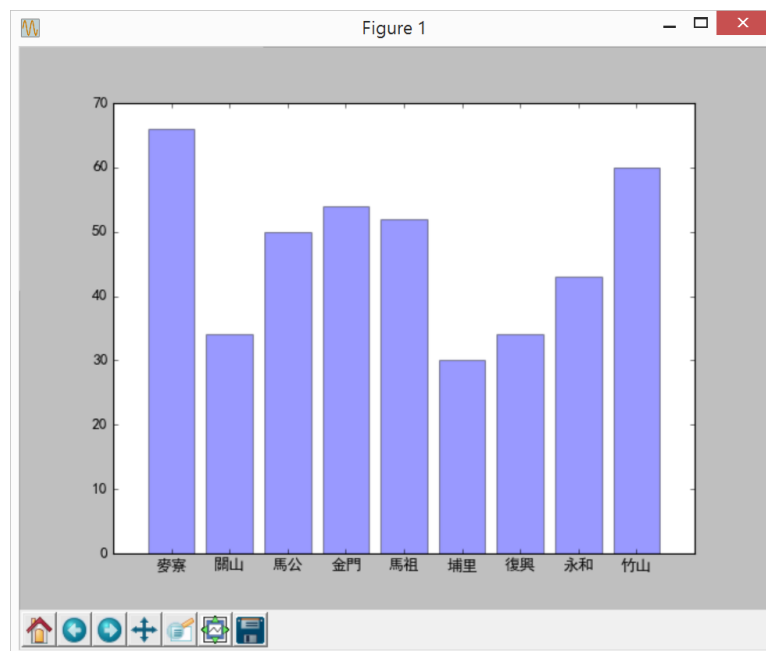


同學課堂練習：基於開放資料透過 Treemap 呈現數值比例資訊

- Python 資料處理與分析工具（SciPy、NumPy、Matplotlib）的使用：課堂前半段以介紹開放資料、地理資訊繪製、統計資訊繪製，以及網頁前端呈現方法為主。後半段則因應資料處理與分析工作，介紹 Python 程式語言與相關套件，如 SciPy、NumPy、Matplotlib 等等。同學在物件導向程式設計課程中，已有學過 Python 程式語言，因此，搭配著 Python 的基礎，相當適合深入學習關於資料處理與分析的方法。同樣基於實際的開放資料，同學以 Python 程式語言進行資料的讀取，進行進一步的分析，並透過 Matplotlib 套件繪製圖形。對於應用數學系的同學來說，日後的研究或職場工作，很有可能需要用到程式語言（如 Python 或 R）進行資料處理與分析，透過課程的介紹，累積同學對於程式語言與相關套件的知識與經驗，對於日後的工作勢必能產生效益。



同學課堂練習：基於開放資料透過 Python 與 Matplotlib 進行處理



同學課堂練習：基於開放資料透過 Python 與 Matplotlib 進行處理

三、學生具體學習成效

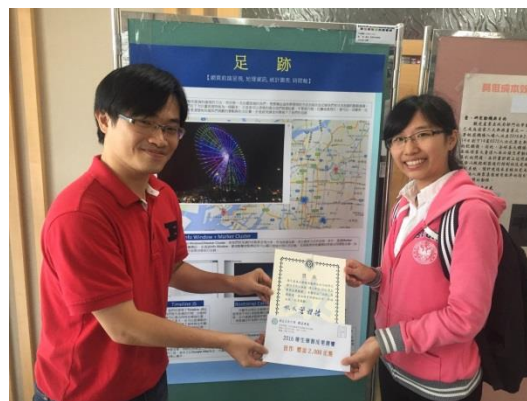
(一)課程學習成效說明

項目	本課程執行前相關作法	獲得經費補助後本課程有所創新改革之作法及成效	學生學習成績、成果與說明
激發學生學習動機與興趣	不定期補充業界訊息	每次上課均補充最新社群活動、業界訊息、TED演講、資料處理與分析專案、競賽訊息等等	透過實際案例激發學生的學習動機與風氣，同學亦逐步關注資料科學的發展與應用
提升教學品質、強化學用合一	方法介紹與程式碼範例練習	實際透過開放資料、Open API 作為練習的基礎，撰寫程式並做地理資訊、統計圖表的呈現	每次上課均規劃實作時間，並搭配作業，讓同學透過實際開放資料，撰寫程式，實際進行資料處理與分析工作
具體成果或作品	課堂作業和期末專案開發	除了課堂作業與期末專案開發之外，鼓勵並協助同學產出作品並參與競賽	獲得理工學院學習競賽佳作

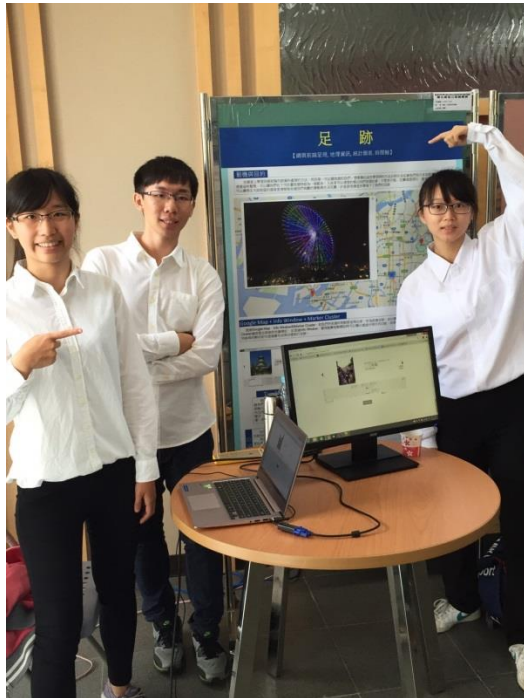
(二) 成果照片



課堂練習



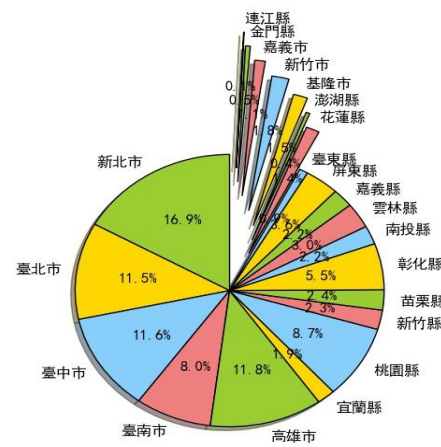
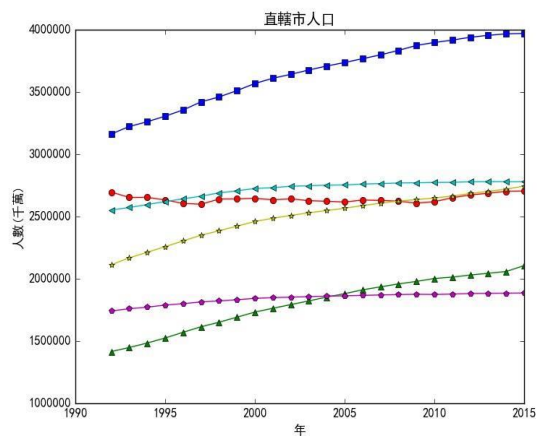
理工學院學習競賽



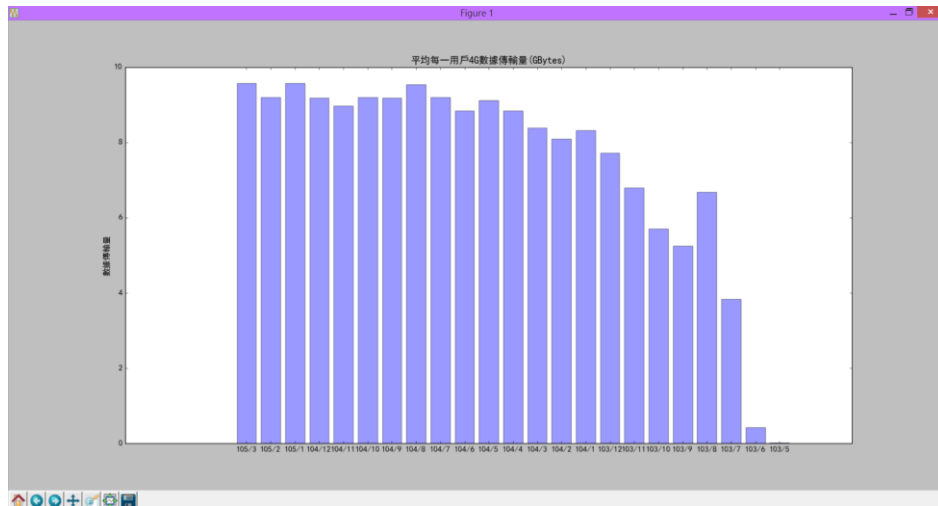
理工學院學習競賽



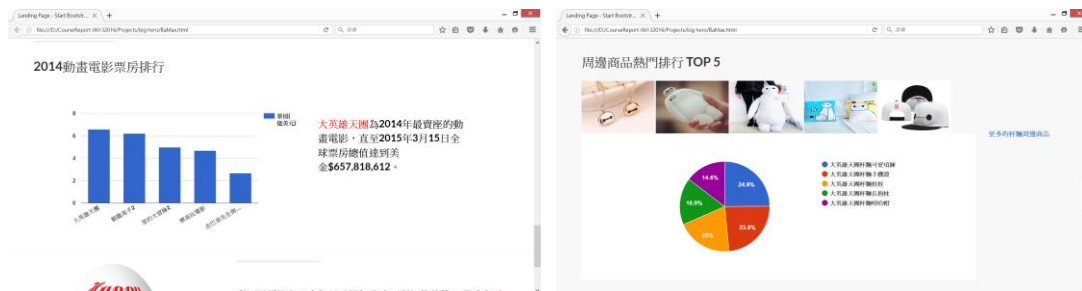
理工學院學習競賽



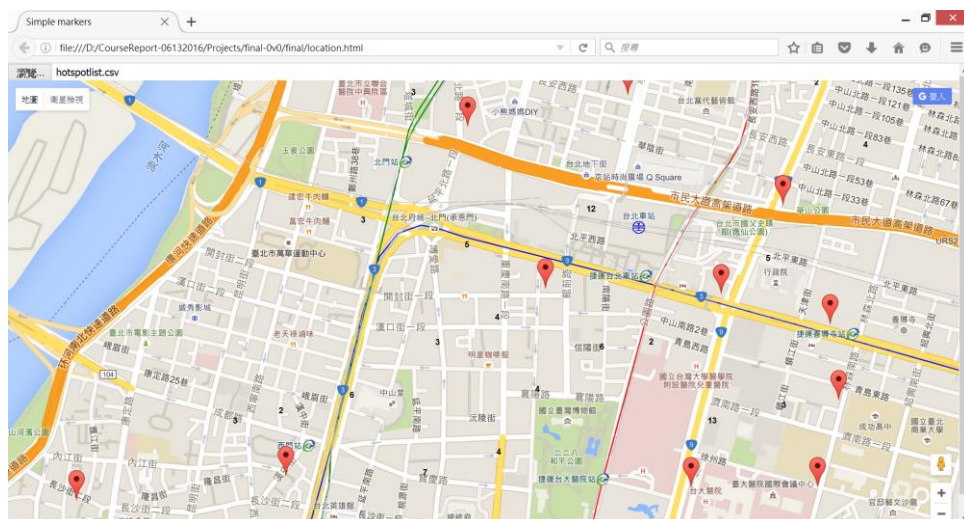
期末專案：直轄市人口變動分析 (透過 Python 與 Matplotlib)



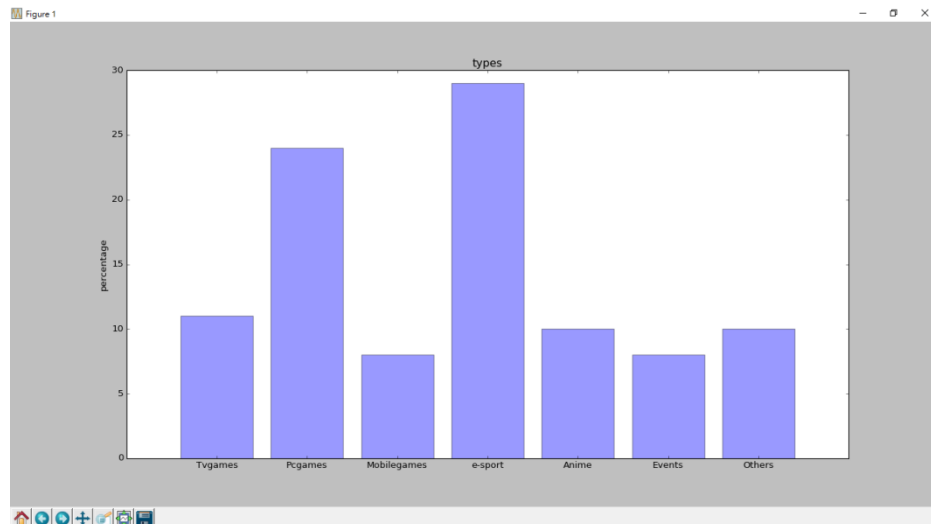
期末專案：網路數據用量分析（透過 Excel、Python 與 Matplotlib）



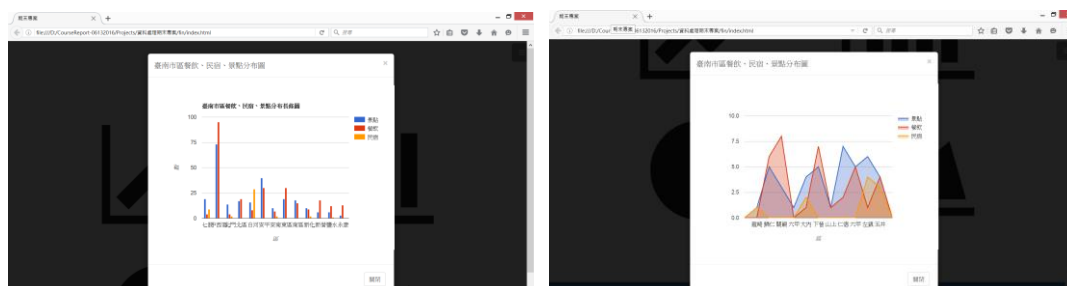
期末專案：大英雄天團介紹（透過 Google Charts 與 Bootstrap 以前端呈現）

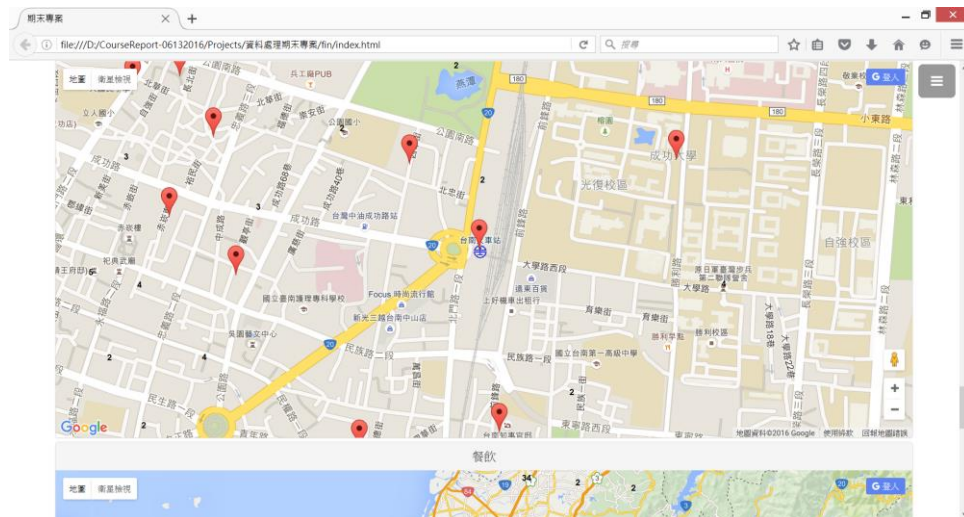


期末專案：全台 iTaiwan 熱點地圖（透過 Google Map 以前端呈現）

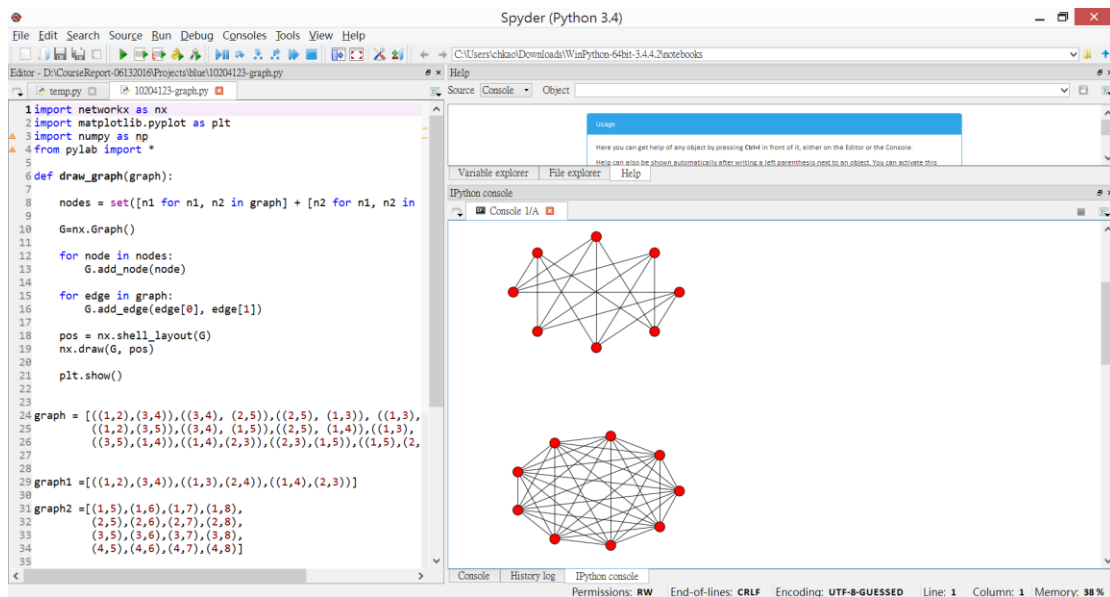


期末專案：遊戲社群討論主題分析（透過 Python 與 Matplotlib）





期末專案：台南旅遊地圖（透過 Google Map、Google Chart、Bootstrap 呈現）



期末專案：圖論圖形繪製（透過 Python、Matplotlib、numpy、networkx）



四、課程導向

- **在地結合與創新應用：**在本次課堂中，大幅使用了開放資料做為同學練習的資料基礎。舉例來說，台北市政府開放資料、政府開放資料、台南市政府開放資料等皆為使用的對象。未來課堂的設計與進行上，將鼓勵同學以台東的開放資料或 Open API 作為基礎，進行資料處理與分析，以及創新應用的開發。藉此，學習與開發的成果亦可回饋給台東或當地社群，促成更加深入的在地參與及合作。
- **教學與實務操作的精進：**在課堂的教學與實務操作上，每一主題介紹後，可以透過小型專案或實作的方式，由同學分組完成。透過分組進行的方式，團隊合作完成實作，一方面可避免單人練習的無趣感，二方面亦可讓同學在合作之中應用與分享所學，藉此提升學習效率。並且，以資料科學或資料工程的工作來說，絕對需要不同領域專業的協同合作，在課堂上引入合作專案與實作的進行，亦可讓同學了解未來工作的型態。
- **業界參訪與經驗分享：**學校的教學往往缺乏業界實際的動態，因此，在未來課堂的規劃上，希望可以安排業界參訪或業師經驗分享的活動。讓同學深刻了解當前所學，在業界的實際應用。藉此，除了能夠提高同學學習的品質之外，亦可即早了解業界樣貌，對於未來職場進行準備。
- **社群活動參與：**在資料科學或資料工程領域上，有相當多的社群可以參與，如台灣資料科學愛好者、資料視覺化、Data Science 社群等等，相關的活動也相當蓬勃，如台灣資料科學年會等。然而，對於同學來說，受限於學校的地理位置，參與社群活動較為困難。若時間與經費許可，可鼓勵與協助同學多多參與相關社群，獲取最新的知識、方法與工具，更可提升同學的學習成果。