

國立勤益科技大學資訊工程系  
Scripting程式語言期末專題計畫書

[illegible]

**專題參與人員：**

班級：丁班      學號：3B117131      姓名：利巽予  
班級：丁班      學號：3B117146      姓名：李振瑋  
班級：丁班      學號：3B117131      姓名：吳旻哲

**關鍵字：**地址轉郵遞區號、Python、SQLite

指導老師：朱孝國 教授

中華民國 113 年 01 月 03 日

# 目錄

目錄 .....	1
圖目錄 .....	2
一、作業摘要 .....	3
二、背景資料 .....	3
(一)、使用 SQLite 資料庫 .....	3
(二)、使用 Pandas 處理 CSV 檔案 .....	3
(三)、使用 requests 模組進行網路請求 .....	4
(四)、基本的資料清理和處理技巧 .....	4
三、環境介紹 .....	4
四、實作方法 .....	5
(一)、實作流程 .....	5
(二)、基本環境安裝(請點擊) .....	6
(三)、to_Zipcode(搜尋郵遞區號) .....	8
(四)、結論 .....	9
(五)、結果展示 .....	10
五、實作心得 .....	11
(一)、實作的心得 .....	11
(二)、遇到的困難 .....	11
(三)、分組的感想 .....	12
(四)、組員的互動 .....	12
六、附註 .....	12
(一)、專題參與人員 .....	12
七、參考文獻 .....	13

# 圖目錄

圖1 SQLite.....	3
圖2 Pandas.....	3
圖3 Requests Python.....	4
圖4 to_Zipcode程式流程圖 .....	5
圖5 to_Zipcode程式流程心智圖 .....	6
圖6 切換至指定專案目錄.....	6
圖7 下載SINOPAC_AIGO_Competition專案 .....	6
圖8 Python環境安裝腳本.....	7
圖9 cmd執行畫面.....	10
圖10 抓取的request[Zipcode] .....	10
圖11 CSV正規化後.....	10
圖12 資料呈現網站.....	11

## 一、作業摘要

本專案旨在透過 Python 學習處理 CSV 資料集及使用 SQLite 資料庫儲存郵遞區號資訊。核心任務是將地址轉換成相應郵遞區號，以增進資料結構性和可用性。專案涵蓋主要技能包括運用 SQLite 資料庫、Pandas 處理 CSV 檔案、網路請求以及基本資料清理。環境設定包括 Python、SQLite、Pandas、requests，並假設在支援 Python 的作業系統中運行。

實作過程涉及讀取 CSV 檔案、創建並操作 SQLite 資料庫、利用網路請求查詢郵遞區號及更新資料集，最終將清理後資料儲存為新 CSV 檔案。此過程不僅增進了對資料處理和清理的理解，也鞏固了使用資料庫和網路請求的技巧。通過此項目，我們深入理解數據處理技巧，尤其是在地址與郵政編碼轉換的地方。儘管過程中遇到挑戰，但團隊合作與互相學習幫助成功完成任務。

## 二、背景資料

本專案使用 Python，涉及到以下主要知識點：

### （一）、使用 SQLite 資料庫

首先，我們使用 SQLite 資料庫，這是一種輕量級的資料庫管理系統，因其簡單和高效能而廣受歡迎。在此專案中，SQLite 的使用使我們能夠有效地儲存和管理大量的郵遞區號資訊。這不僅提高了數據的可訪問性，還簡化了數據管理流程。



圖1 SQLite

### （二）、使用 Pandas 處理 CSV 檔案

接著，我們運用了 Pandas 這一強大的數據分析庫來處理 CSV 檔案。Pandas 提供了豐富的數據結構和分析工具，使得從 CSV 檔案中讀取、清理、轉換和分析數據變得更加直接和高效。在這個專案中，Pandas 的應用不僅增強了我們對數據的控制能力，還提高了數據處理的效率和準確性。



圖2 Pandas

### (三)、使用 requests 模組進行網路請求

此外，我們還利用了 requests 模組來進行網路請求。這個強大的 HTTP 庫讓我們能夠輕鬆地發送 HTTP 請求，從而從互聯網上獲取或發送數據。在本專案中，通過 requests 模組，我們能夠有效地從網上獲取郵遞區號相關資訊，這對於數據集的完整性和準確性至關重要。



圖3 Requests Python

### (四)、基本的資料清理和處理技巧

最後，本專案還包括了基本的資料清理和處理技巧。資料清理是數據分析過程中不可或缺的一部分，它幫助我們去除或糾正數據集中的錯誤和不一致性。透過有效的資料清理，我們確保了數據的質量和可靠性，從而使最終分析的結果更加精準可靠。

## 三、環境介紹

在本專案中，我們建立了一個專門針對數據處理和分析的高效工作環境。這個環境的核心是 Python 程式語言，一種廣受歡迎且功能強大的語言，特別適合於數據科學和網路開發。Python 的易讀性和簡潔的語法使得它成為學習和使用最廣泛的編程語言之一，特別適合處理複雜的數據集和進行快速開發。

專案中還積極利用了 SQLite 資料庫，這是一種嵌入式的輕量級數據庫系統。由於其自包含、高可靠性和簡單的設置，SQLite 成為了儲存和管理數據的理想選擇。它特別適合於需要輕量化、無需大型數據庫系統支持的應用程序，並且可以無縫整合到 Python 環境中。

除此之外，Pandas 和 requests 這兩個 Python 模組在我們的專案中發揮了關鍵作用。Pandas 是一個數據分析和處理的強大工具，提供了快速、靈活且直觀的數據結構，用於處理和分析大量數據。而 requests 模組則為我們提供了一種簡單的方法來發送 HTTP 請求，這對於從互聯網獲取數據至關重要。

最後，我們的專案設定在一個支援 Python 開發的作業系統環境中。這種環境提供了必要的支持，使得 Python 及其相關模組可以高效運行。無論是在 Windows、Linux 還是 macOS 上，Python 的靈活性和跨平台特性確保了各種作業系統都能提供穩定和強大的開發環境。

透過這些工具和技術的綜合應用，我們的專案環境不僅提供了處理複雜數據的能力，還保證了開發過程的高效和靈活性。這樣的環境對於任何專注於數據處理和分析的項目來說，都是不可或缺的。

## 四、實作方法

### (一)、實作流程

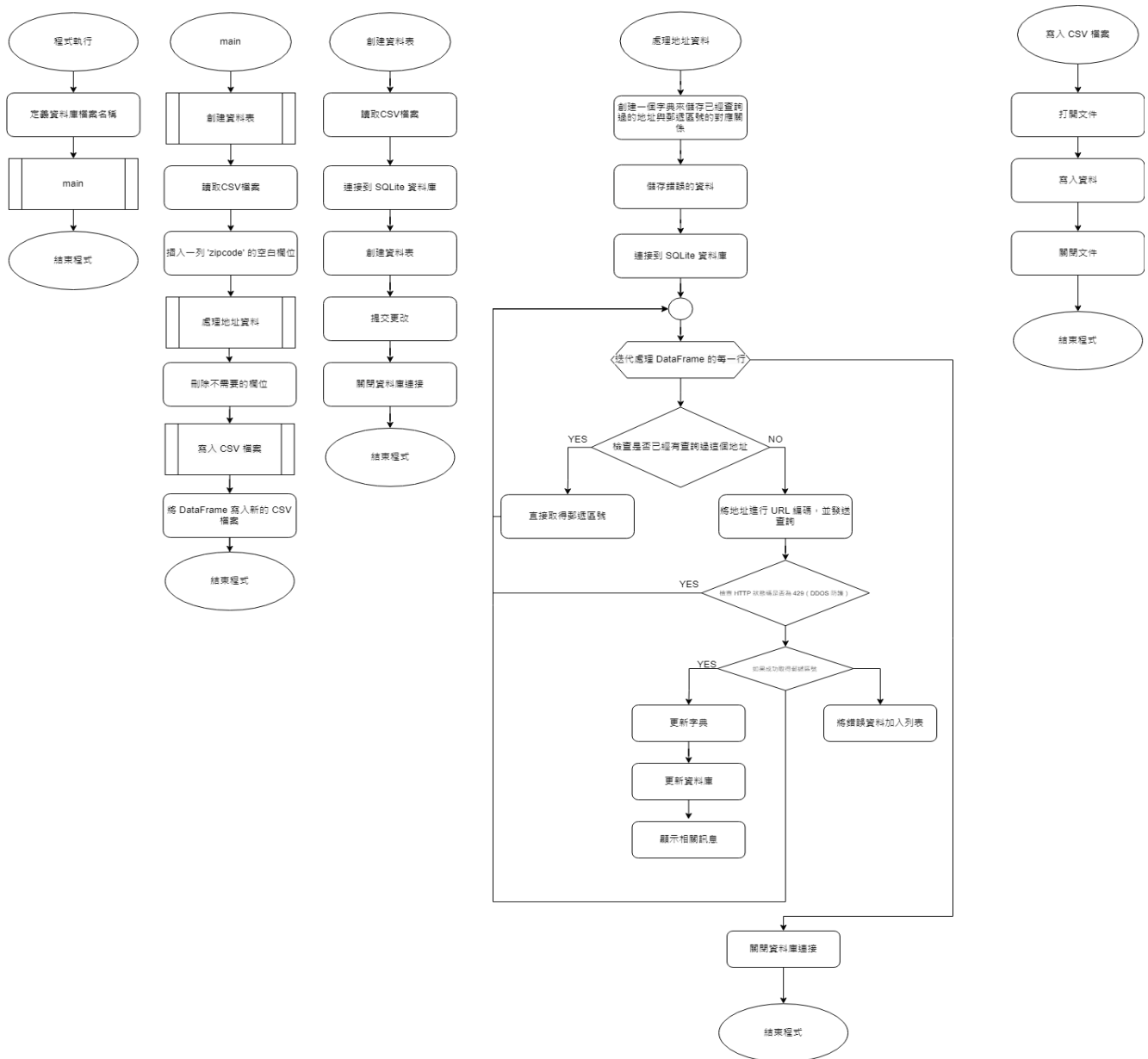


圖4 to\_Zipcode程式流程圖



圖5 to\_Zipcode 程式流程心智圖

## (二)、基本環境安裝(請點擊)

步驟 1 :切換至指定專案目錄(請在git bash執行 註:以切換至桌面為例)

```

bash

User@DESKTOP-G1BPJGQ MINGW64 ~/Desktop/SINOPAC_AIGO_Competition (main)
$ cd /Desktop

```

圖6 切換至指定專案目錄

步驟 2 :下載 \*[SINOPAC\\_AIGO\\_Competition](https://github.com/TW-blast77/SINOPAC_AIGO_Competition)\* 專案(請在git bash執行)

```

git clone https://github.com/TW-blast77/SINOPAC_AIGO_Competition.git

```

```

bash

User@DESKTOP-G1BPJGQ MINGW64 ~/Desktop/SINOPAC_AIGO_Competition (main)
$ git clone https://github.com/TW-blast77/SINOPAC_AIGO_Competition.git

```

圖7 下載 SINOPAC\_AIGO\_Competition專案

### 步驟 3：Python環境

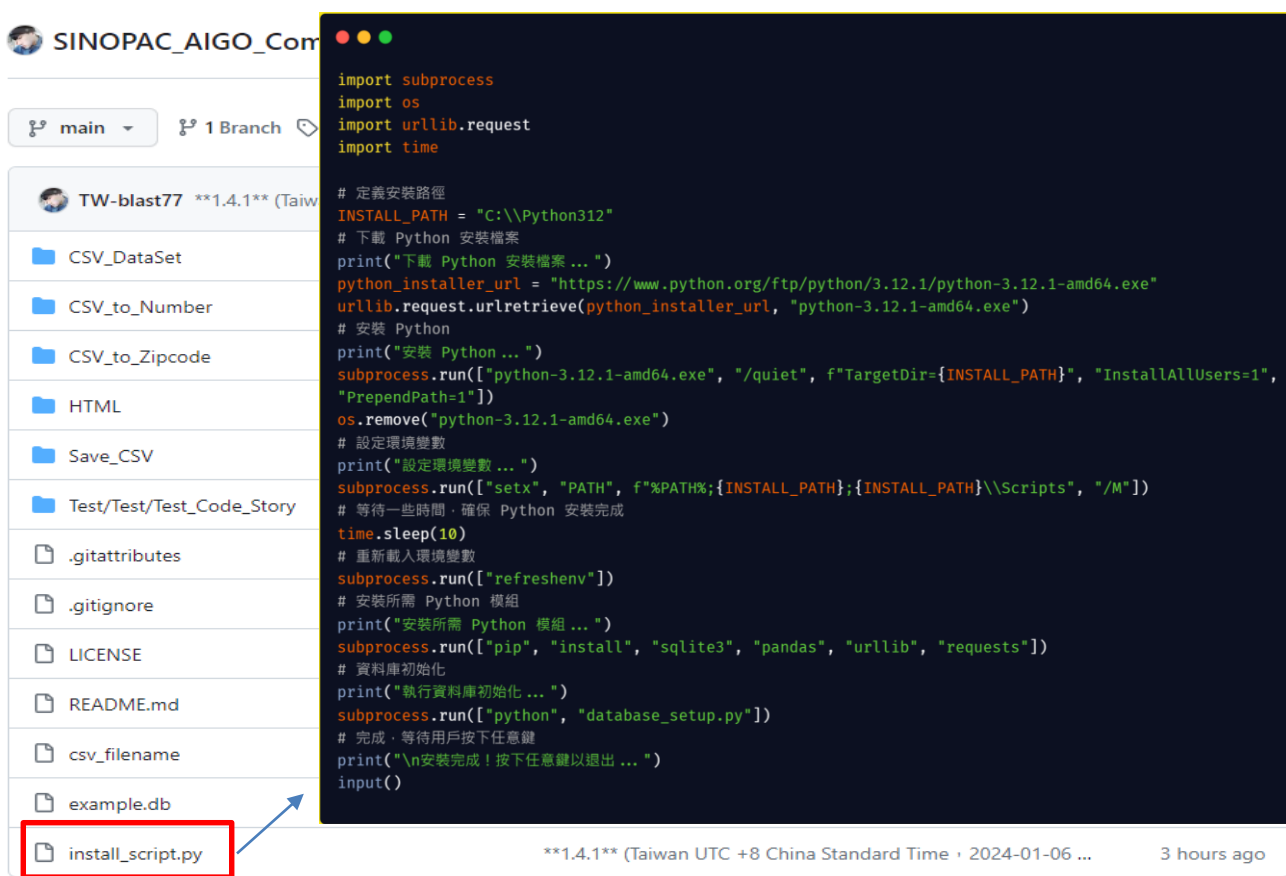


圖8 Python環境安裝腳本

請在 **\*\*cmd\*\*** 中執行，安裝python基本環境以及python模組。

[\\*python install\\_script.py\\*\(請點擊\)](#)

- ◆ 下載 Python 安裝檔案：使用 **urllib.request.urlretrieve** 函式從指定的 URL Python 安裝檔案。
- ◆ 安裝 Python：使用 **subprocess.run** 函式執行命令行指令，安裝 Python。使用 **/quiet** 參數以無聲方式進行安裝，**TargetDir** 參數指定安裝路徑，**InstallAllUsers** 參數表示將 Python 安裝為所有使用者，**PrependPath** 參數表示將 Python 的路徑添加到系統環境變數中。
- ◆ 設定環境變數：使用 **subprocess.run** 設定環境變數，將 Python 的路徑添加到 **PATH** 中。
- ◆ 等待安裝完成：使用 **time.sleep** 函式等待一段時間，確保 Python 安裝完成。



- ◆ 重新載入環境變數：使用 `subprocess.run(["refreshenv"])` 重新載入環境變數。
- ◆ 安裝所需 Python 模組：使用 `subprocess.run(["pip", "install", ...])` 安裝所需的 Python 模組，這裡安裝了 `sqlite3`、`pandas`、`urllib`、`requests`。
- ◆ 資料庫初始化：使用 `subprocess.run(["python", "database_setup.py"])` 執行資料庫初始化腳本。
- ◆ 等待使用者輸入：使用 `input()` 函式等待使用者按下任意鍵，以確保使用者有足夠的時間檢查安裝狀態。

### (三)、to\_Zipcode(搜尋郵遞區號)

#### 1、[測試資料\(請點擊\)](#)(資料集Sinopac所擁有)

標頭：

ID, 縣市, 鄉鎮市區, 路名, 土地面積, 使用分區, 移轉層次, 總樓層數, 主要用途, 主要建材, 建物型態, 屋齡, 建物面積, 車位面積, 車位個數, 橫坐標, 縱坐標, 備註, 主建物面積, 陽台面積, 附屬建物面積。

測試資料型態(CSV)：

CSV指的是透過逗號（或其他分隔符號）來分隔不同欄位的文本格式。CSV文件通常是純文本文件，其中包含一個表格，每行代表表格中的一條記錄，而每個欄位則由分隔符號進行區分。

#### 步驟1：匯入所需的模組環境

Python 資料庫操作的第一步是匯入必要的模組。這一步是至關重要的，因為它允許程式訪問與資料庫互動所需的功能。匯入的模組可能包括 `sqlite3`、`pandas`、`urllib`、`requests`、`time`和`csv`等。

#### 步驟2：定義資料庫檔案名稱

定義資料庫檔案名稱是Python 資料庫操作的重要部分。這一步涉及指定資料庫檔案的名稱和位置，程式將在其中存儲和訪問數據。

#### 步驟3：創建資料表的函式

創建創建資料表的函式是Python 資料庫操作的重要部分。這函式負責建立資料庫表的結構，包括用於存儲信息的字段和數據類型。

#### 步驟4：處理地址資料的函式

處理地址資料的函式是Python 資料庫操作的關鍵步驟。該函式涉及處理地址資料，查詢相關信息，並使用獲得的詳細信息更新資料庫。

#### 步驟5：將資料寫入CSV檔案的函式

將資料寫入CSV檔案的函式是Python 資料庫操作的重要部分。這函式使程式能夠從資料庫中導出數據並將其保存在CSV檔案中，以進行進一步分析或共享。

#### 步驟6：主程式

主程式執行包括將上述步驟整合到一個統一的工作流程中。本部分展示了如何使用各種函式和模組來實現Python 資料庫操作中所需的結果。

### (四)、結論

Python 資料庫操作是現代程式設計中不可或缺的一環，它提供了豐富的工具和函式庫，使得開發者能夠輕鬆地與資料庫進行互動。在本文中，我們深入探討了Python 資料庫操作的步驟和相關函式，並突顯了每個步驟的重要性。首先，我們學習了匯入所需模組的重要性。

這些模組包括 `sqlite3` 用於處理 `SQLite` 資料庫、`pandas` 用於處理資料框架、`urllib` 和 `requests` 用於處理網路請求，以及 `csv` 用於處理 CSV 檔案。這一步為後續的操作提供了必要的基礎。接著，我們定義了資料庫檔案名稱，確保了程式能夠正確地連接到 `SQLite` 資料庫。這一步是建立資料庫連接的關鍵，確保了後續操作能夠順利進行。

在 `創建資料表` 的函式中，我們實際上建立了一個名為 `"zipcodes"` 的資料表，並指定了各種欄位，包括 `id`、`city`、`district`、`road`、`zipcode` 等。這使得資料庫具備了結構化的能力，以更有效地存儲和管理數據。在處理地址資料的函式中，我們進行了與外部 API 的互動，查詢郵遞區號，同時更新了資料庫中的相應欄位。

這一步驟展示了 Python 資料庫操作的實際應用，涉及到網路請求、資料處理和資料庫更新等多個方面。將資料寫入 CSV 檔案的函式使得我們能夠方便地將處理後的資料導出，進行進一步的分析或與他人分享。這擴大了 Python 資料庫操作的應用範疇，使得資料更具有可移植性。

## (五)、結果展示



```
問題 1 輸出 偵錯主控台 終端機 連接埠 GITLENS

& C:/Users/User/AppData/Lo

-----
項目: 1 號
request : http://zip5.5432.tw/zip5json.py?adrs=新北市樹林區大有路
目前地區為: 新北市樹林區大有路
找尋到郵遞區號為: 23854
-----

項目: 2 號
request : http://zip5.5432.tw/zip5json.py?adrs=新北市三重區力行路一段
目前地區為: 新北市三重區力行路一段
找尋到郵遞區號為: 24162
-----

項目: 3 號
request : http://zip5.5432.tw/zip5json.py?adrs=新北市板橋區民生路三段
目前地區為: 新北市板橋區民生路三段
找尋到郵遞區號為: 22047
-----
```

圖9 cmd執行畫面

```
{"zipcode": "", "dataver6": "11207(beta)", "adrs": "\u65b0\u5317\u5e02\u4e09\u91cd\u5340\u529b\u884c\u8def\u4e00\u6bb5", "new_adrs6_2":  
"\u65b0\u5317\u5e02\u4e09\u91cd\u5340\u529b\u884c\u8def\u4e00\u6bb5", "new_adrs6":  
"\u65b0\u5317\u5e02\u4e09\u91cd\u5340\u529b\u884c\u8def\u4e00\u6bb5", "new_adrs2":  
"24162\u65b0\u5317\u5e02\u4e09\u91cd\u5340\u529b\u884c\u8def\u4e00\u6bb5", "new_adrs":  
"24162\u65b0\u5317\u5e02\u4e09\u91cd\u5340\u529b\u884c\u8def\u4e00\u6bb5", "dataver": "11112", "zipcode": "24162"}
```

圖10 抓取的request[Zipcode]



ID	Zipcode	鄉鎮市區	路名	土地面積	使用分區	移轉層次	總樓層數	
PU-1	23854	1	1	2.0393784158445825	1	3	7	1,1,華廈(10層含以下有)
PU-2	24162	2	2	-0.07777861276769849	1	5	5	1,1,公寓(5樓含以下無)
PU-3	22047	3	3	-0.555765608691798	1	2	20	2,1,住宅大樓(11層含)
PU-4	24349	4	4	-0.7448376409295061	1	6	16	1,1,住宅大樓(11層含)
PU-5	11065	5	5	-0.6109226334770453	1	3	7	1,1,華廈(10層含以下)
PU-6	24154	2	6	0.5490572363530009	1	5	5	1,1,公寓(5樓含以下無)
PU-7	24341	4	4	-0.6959533758257184	1	8	14	2,1,住宅大樓(11層含)
PU-8	33850	4	7	-0.7807197489924143	1	3	13	1,1,住宅大樓(11層含)
PU-9	324247	6	8	-0.2272748833041037	1	14	14	1,1,住宅大樓(11層)
PU-10	23941	7	9	0.2198455041053685	1	4	5	1,1,公寓(5樓含以下無)

圖11 CSV正規化後

Show
entries

Search:

ID	Zipcode	土地面積	使用分區	移轉層次	總樓層數	主要用途	主要建材	建物型態	屋齡	建物面積	車位面積	車位個數	橫坐標	縱坐標	備註	主建物面積	陽台面積	附屬建物面積	Action
1	13657	3.23	2	8	7	4	7	建物型態	14.45	34.97	16.13	5	781.90	668.45	備註	119.35	8.52	37.32	<div>通過</div> <div>不通過</div>
2	19412	1.79	2	7	6	1	3	建物型態	7.20	24.45	13.71	2	208.84	211.02	備註	418.71	22.66	26.17	<div>通過</div> <div>不通過</div>
3	23771	2.71	5	5	6	6	5	建物型態	10.43	15.20	26.14	2	875.77	228.53	備註	244.30	49.11	6.33	<div>通過</div> <div>不通過</div>
4	17604	2.26	10	7	2	2	9	建物型態	2.76	95.82	19.44	5	535.07	708.74	備註	205.87	7.68	59.32	<div>通過</div> <div>不通過</div>
5	23270	1.62	2	4	2	1	4	建物型態	19.92	24.80	8.40	4	992.27	208.85	備註	33.19	35.12	52.70	<div>通過</div> <div>不通過</div>
6	20901	4.87	10	2	8	4	5	建物型態	1.31	8.54	1.85	4	765.84	882.14	備註	84.04	17.82	16.24	<div>通過</div> <div>不通過</div>

圖12 資料呈現網站

## 五、實作心得

### (一)、實作的心得

這學期的期末作業為我提供了一個實踐所學知識的機會，也是一次全面性的挑戰。作為專案開發的一部分，我們需要將理論知識轉化為實際的解決方案，這要求我們不僅要有堅實的理論基礎，還需要具備實際應用的能力。這讓我深感這是一個寶貴的學習機會，讓我更加了解如何應對真實世界中的挑戰。

### (二)、遇到的困難

在整個專案的開發過程中，我覺得最讓我收穫良多的是對專案管理和團隊協作的深入認識。從一開始的需求分析，到設計和實施，每個步驟都需要團隊成員的密切合作。我學到了如何有效地分工和溝通，以確保每個人都在專業領域充分發揮自己的優勢。這也教導我如何在有限的時間內合理分配任務，以確保整個專案按計劃進行。

然而，這次作業並非一帆風順。我們遇到了技術上的一些挑戰，特別是在與外部API整合和資料庫操作方面。這要求我們不僅要快速理解新技術，還需要

找到解決問題的方法。這些困難使我們更深刻地理解了專業領域的複雜性，也增強了我們解決實際問題的能力。

### **(三)、分組的感想**

與團隊的合作讓我有新的視角。每個團隊成員都帶來了獨特的見解和技能，這讓我們能夠以更全面的方式來解決問題。透過分組作業，我學到了如何充分發揮每個成員的潛力，並將各種專業背景融合成一個協同工作的整體。這也讓我更加珍惜團隊合作的價值，因為在合作中我們可以互相學習，共同成長。

### **(四)、組員的互動**

在這次作業中，我與組員1利巽予及組員2吳旻哲的互動讓整個專案更加順利進行。利巽予負責大部份的Word製作，他的專業能力使得我們的文件結構更為完整，內容更為豐富。在與他的討論中，我深刻體會到專業領域的深度和廣度。而吳旻哲則負責PPT簡報的製作，他的創意和設計技巧為我們的簡報增色不少。我們透過密切的溝通，共同解決了一些技術上的難題，並有效地分工協作。這次的合作經歷不僅讓我學到了更多的知識，也建立了彼此之間的信任與友誼，也非常感謝老師透過分組的方式讓我們組員互相配合磨合，共同完成一份專題。

## **六、 附註**

### **(一)、專題參與人員**

人員1：

學號：3B117145

姓名：吳旻哲

工作：（簡報PPT製作100%）

人員2：

學號：3B117131

姓名：利巽予

工作：（Word製作70%）

人員3：

學號：3B117146

姓名：李振瑋

工作：(Word製作30%、本專題製作100%)

## 七、參考文獻

- [1] Python SQLite 資料庫教學與範例 <https://shengyu7697.github.io/python-sqlite-tutorial/>
- [2] Pandas官網 <https://pandas.pydata.org/>
- [3] [Python 資料科學]pandas 基礎介紹 - 進入資料科學的領域  
<https://medium.com/seaniap/pandas%E5%9F%BA%E7%A4%8E%E4%BB%8B%E7%B4%B9-%E9%80%B2%E5%85%A5%E8%B3%87%E6%96%99%E7%A7%91%E5%AD%B8%E7%9A%84%E9%A0%98%E5%9F%9F-be9894b3548>
- [4] Python官網 <https://www.python.org/>
- [5] STEAM 教育學習網  
<https://steam.oxxostudio.tw/category/python/spider/requests.html>
- [6] Os各種作業系統介面<https://docs.python.org/zh-tw/3/library/os.html>
- [7] 子進程管理Subprocess <https://docs.python.org/zh-tw/3/library/subprocess.html>
- [8] 時間處理time-python教學-steam教育學習網  
<https://steam.oxxostudio.tw/category/python/library/time.html>
- [9] Python 中的自動 Script 入門  
<https://www.ibm.com/docs/zh-tw/spss-statistics/29.0.0?topic=scripts-getting-started-autoscripts-in-python>
- [10] 永豐AIGO資料集：<https://tbrain.trendmicro.com.tw/Competitions/Details/30>