

網頁爬蟲、數據處理及視覺化：全台灣七天 氣溫預報

本報告展示了使用數據分析和網頁爬蟲技術，對台灣未來一週天氣預報進行深入分析和視覺化的過程。我們將探討如何從官方資料來源獲取、處理和呈現全台灣的氣溫數據。

E by Eve Lee



報告大綱

1

摘要

概述專案目標、數據來源和使用的工具。

2

流程

詳細說明數據獲取、處理和視覺化的步驟。

3

圖像處理

展示數據視覺化的結果和解釋。

4

結論、應用與未來發展

總結專案成果和潛在應用。

5°C	 22 - 24°C	 22 - 25°C	23
8°C	 22 - 26°C	 21 - 28°C	22
4°C	 21 - 23°C	 22 - 24°C	22
7°C	 21 - 26°C	 21 - 29°C	21
4°C	 21 - 23°C	 21 - 25°C	21
8°C	 22 - 28°C	 21 - 31°C	22
5°C	 21 - 25°C	 22 - 27°C	23
8°C	 22 - 28°C	 22 - 30°C	22
5°C	 22 - 25°C	 22 - 26°C	23
7°C	 21 - 28°C	 20 - 31°C	22

摘要

目標

通過數據處理和視覺化，提供易於理解的全台灣七天氣溫預報，展示數據擷取及分析能力。

數據來源

從【氣象資料開放平臺】爬取【鄉鎮天氣預報-台灣未來1週天氣預報】的資訊。

- [1]<https://opendata.cwa.gov.tw/index>
- [2]<https://opendata.cwa.gov.tw/dataset/forecast/F-D0047-091>

技術工具

運用**Requests**、**JSON**、**API Key**等進行數據獲取和處理。
。使用**Matplotlib**和**datetime**模組實現數據視覺化。

流程

數據擷取

使用**Requests**模組和**API Key**從氣象資料開放平臺獲取**JSON**格式的天氣預報數據。

1

2

數據處理

運用**JSON**模組解析和清理原始數據，提取關鍵天氣信息。

3

數據視覺化

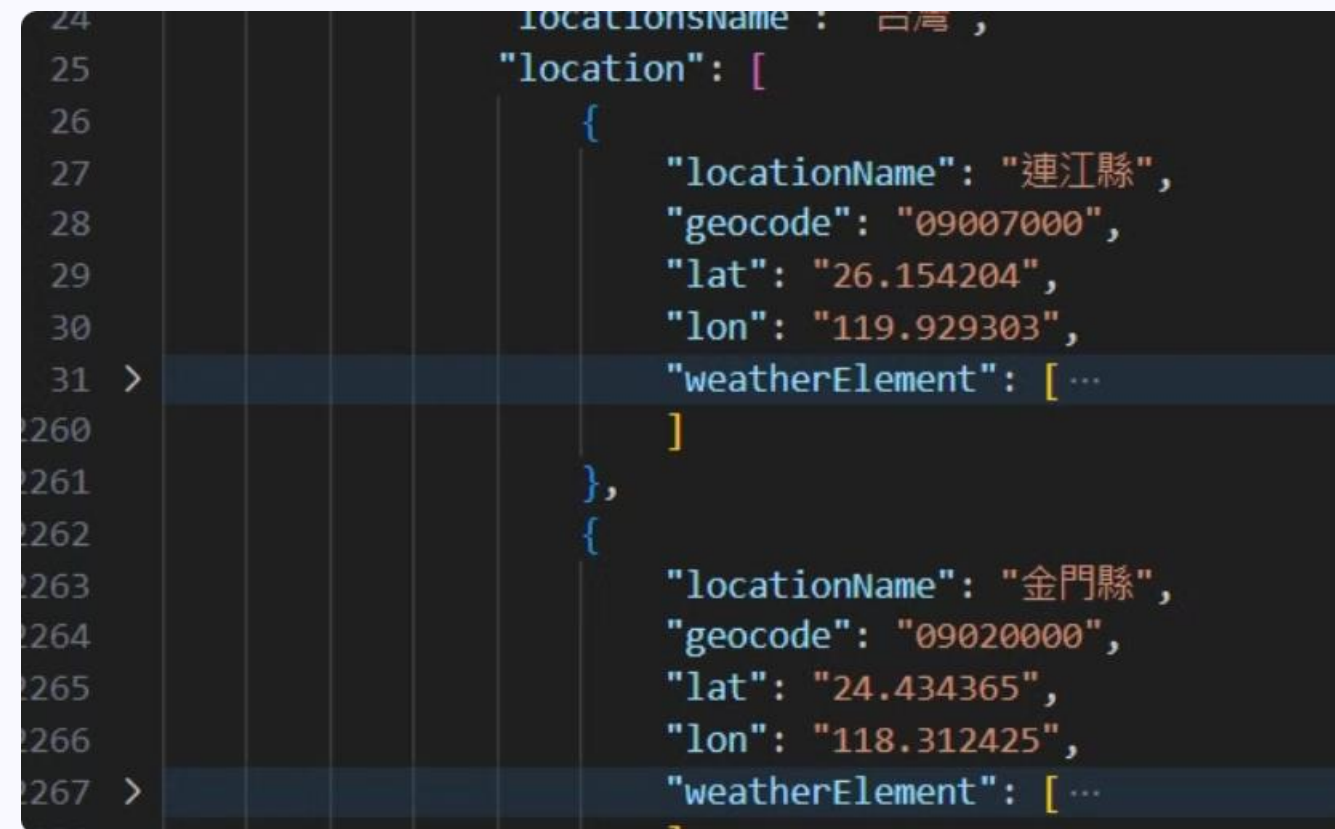
利用**Matplotlib**模組創建圖表，展示溫度和降雨機率等天氣參數的變化趨勢。應用**datetime**模組處理日期和時間，生成準確的七天天氣預測時間軸。

數據擷取



Requests 模組 & API Key

【氣象資料開放平臺】過URL下載檔案以及 **RESTful API** 資料擷取方法取用資料，申請會員後可得到**授權碼(API Key)**並利用 **Requests** 模組從此網頁取得 **JSON** 格式的天氣預報數據。



JSON 格式數據清理

取得數據後，利用 **JSON** 模組解析和清理原始數據，取得關鍵天氣資訊，如平均溫度、溼度等。此專案將專注在**平均溫度**的資料擷取。

數據擷取結果

- 【鄉鎮天氣預報-台灣未來1週天氣預報】提供各個縣市的每 6 個小時的平均溫度
- 結果如右圖

```
=== 縣市名稱：臺北市, latitude: 25.035095, longitude: 121.558742 ===  
平均溫度  
開始時間: 2024-09-17T18:00:00+08:00, 結束時間: 2024-09-18T06:00:00+08:00  
溫度: 攝氏度 29  
開始時間: 2024-09-18T06:00:00+08:00, 結束時間: 2024-09-18T18:00:00+08:00  
溫度: 攝氏度 30  
開始時間: 2024-09-18T18:00:00+08:00, 結束時間: 2024-09-19T06:00:00+08:00  
溫度: 攝氏度 28  
開始時間: 2024-09-19T06:00:00+08:00, 結束時間: 2024-09-19T18:00:00+08:00  
溫度: 攝氏度 31  
開始時間: 2024-09-19T18:00:00+08:00, 結束時間: 2024-09-20T06:00:00+08:00  
溫度: 攝氏度 29  
開始時間: 2024-09-20T06:00:00+08:00, 結束時間: 2024-09-20T18:00:00+08:00  
溫度: 攝氏度 31  
開始時間: 2024-09-20T18:00:00+08:00, 結束時間: 2024-09-21T06:00:00+08:00  
溫度: 攝氏度 29  
開始時間: 2024-09-21T06:00:00+08:00, 結束時間: 2024-09-21T18:00:00+08:00  
溫度: 攝氏度 31  
開始時間: 2024-09-21T18:00:00+08:00, 結束時間: 2024-09-22T06:00:00+08:00
```

視覺化處理準備-程式整理

數據分類

為了方便後續的圖像資訊整理，我們將擷取的數據依照縣市分類，存入各別的 **Dictionary** 及安排將縣市依照地區呈現資訊（北部、中部、南部、東部、離島）

函式整理

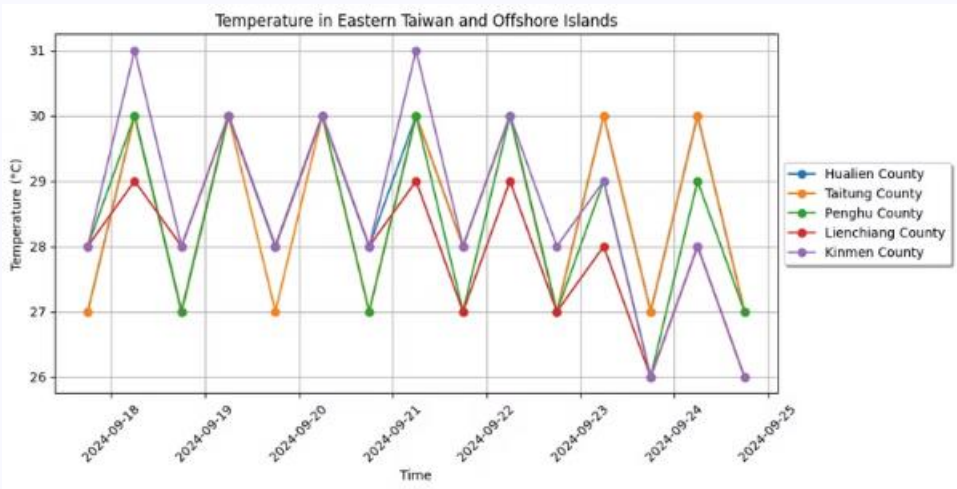
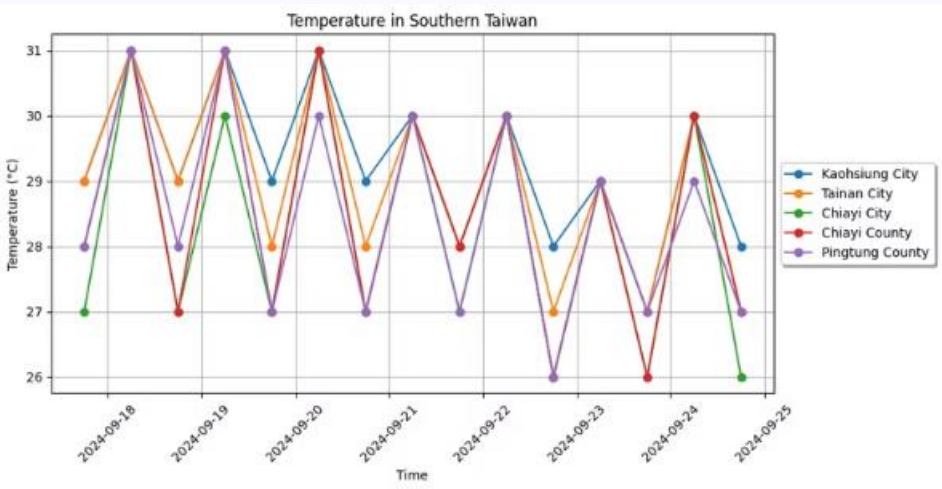
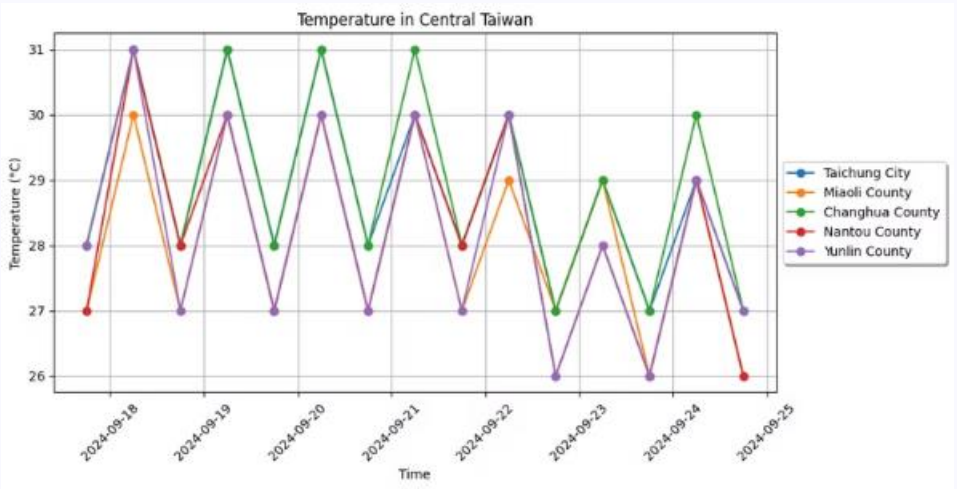
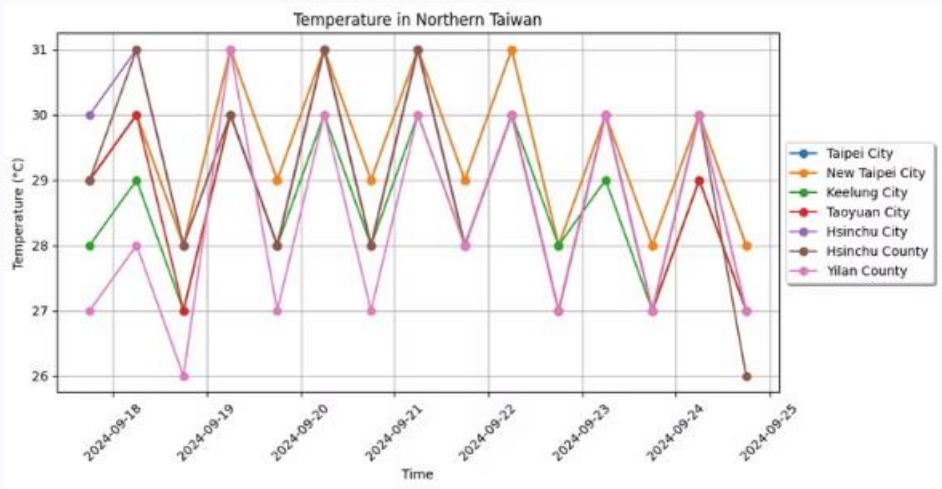
由於這是同一個程式，為了提升程式碼的可讀性，使用函式來組織程式邏輯。

```
# 取得 JSON 資料 (dict 格式)
obj = res.json()
with open('weather_data.json', 'w', encoding='utf-8') as f:
    json.dump(obj, f, ensure_ascii=False, indent=4)
...
make it into a function so cleaner look
...
# 印出每一個縣市的一週天氣
> def weather(city, weather_dict): ...

# 北部
weather('臺北市', tp_weather)
weather('新北市', nt_weather)
weather('基隆市', kl_weather)
weather('桃園市', ty_weather)
weather('新竹市', hsinCi_weather)
weather('新竹縣', hsinCo_weather)
weather('宜蘭縣', yil_weather)
```

數據視覺化結果

利用matplotlib，將擷取的【鄉鎮天氣預報-台灣未來1週天氣預報】平均溫度數據整理成北部、中部、南部及東部與離島資訊。



結論、應用與未來發展

實用價值

生成的天氣預報圖表直觀明了，可為大眾提供實用的氣象信息。

技術應用：數據擷取、分析及數據化處理

展示了在實際項目中運用**Python**進行網頁爬蟲和數據分析的能力。

未來發展

- 可進一步擴展，加入更多氣象參數，創作**Dashboard**，提供更全面的天氣分析服務。
- 可利用**Seaborn**等其他數據視覺化模組，提供更美化的圖表及色調的調配



參考資料與感謝

[1] 【氣象資料開放平臺】 <https://opendata.cwa.gov.tw/index>

[2] 【鄉鎮天氣預報-台灣未來1週天氣預報】（資料來源）
<https://opendata.cwa.gov.tw/dataset/forecast/F-D0047-091>

[3] 圖1 <https://www.cwa.gov.tw/V8/C/W/County/index.html>

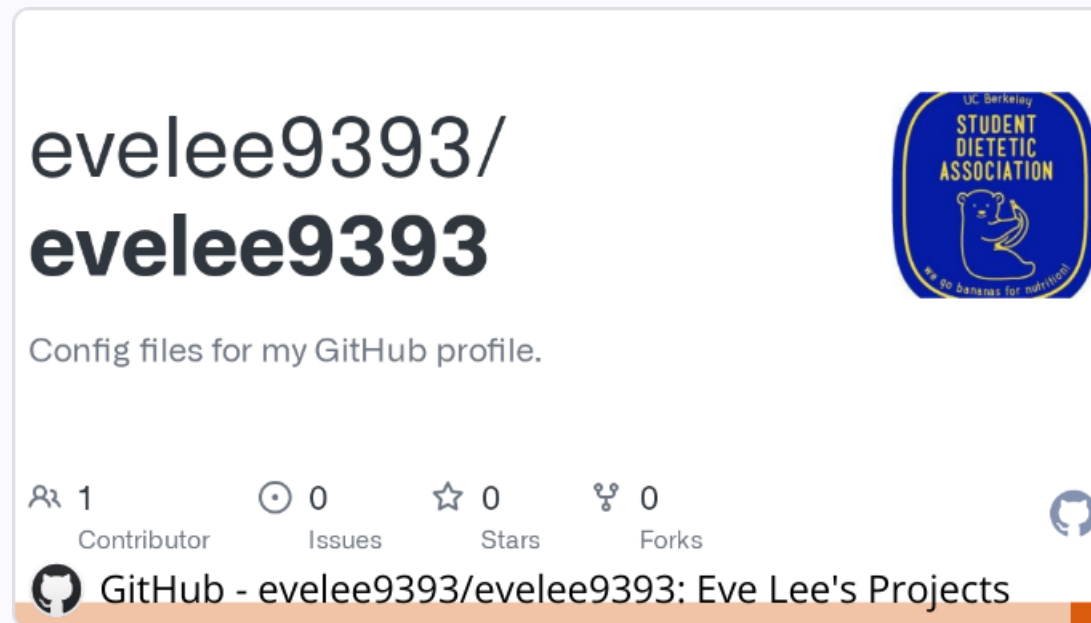
[4] 圖2 <https://www.cwa.gov.tw/V8/E/W/week.html>

[5] 圖3 <https://consoleflare.com/blog/wp-content/uploads/2023/05/Matplotlib-vs-Seaborn.png>

最後，也要感謝各位的耐心聆聽。

李敬怡的 GitHub 頁面

歡迎到我的 **GitHub** 頁面，查看這個專案的原始碼和更多有趣的作品。



<https://github.com/evelee9393/evelee9393>

Thank you!