

大專新進人員訓練工作報告

姓 名： 王翔禾 單位： 台化工務部自控處

報告類別： ☐ 輪班訓練 ☒ 基層實務工作訓練第（二）次報告

訓練部門： 台化工務部自控處智能專案組

起迄日期： 2022/01/01~2022/03/14

工作項目： 太陽能案場發電效率統計自動化及太陽能案場發電效率統計表優化

報告項目： 太陽能案場發電效率統計自動化及統計報表優化

| 評 核 (評 語) | |
|---|--|
| (2) 一 級 主 管 | (1) 二 級 主 管 |
| 本次課程實務訓練，以爬蟲技術及數據統計分析為主，針對python相關套件應用有相當深入瞭解，有研發精神，值得嘉許！ 羅榮達 2022/3/9 | 1. 目前主要工作為太陽能案場發電效率統計表程式設計及統計報表優化。 2. 該員設計工作均能如期完成，工作及學習態度認真表現佳。 職：曹志旭 3/9 |
| (4) 經 營 主 管 | (3) 經 理 室 |
| 王翔禾 3/14 | 林建忠 3/9 |

表號：P0002703 規格：A4

註：一. 報告內容應包括訓練期間之 1. 工作心得感想 2. 所得之工作觀念及精神 3. 自我檢討 4. 建議意見。

二. 訓練部門主管應詳細評核「訓練工作報告」內容，並批註意見。

目錄

| | |
|--|----|
| 一、 前言 | 2 |
| 二、 太陽能案場發電效率統計自動化 | 3 |
| 2-1 系統架構 | 3 |
| 2-2 Python 網路爬蟲介紹 | 5 |
| 2-3 Python API 串接介紹 | 9 |
| 2-4 HyperText Transfer Protocol 傳輸協定介紹 | 11 |
| 2-5 Python Openpyxl 套件介紹 | 14 |
| 三、 太陽能案場發電效率統計表優化 | 16 |
| 3-1 系統架構 | 16 |
| 3-2 Python 統計分析套件介紹 | 18 |
| 3-2-1 Python NumPy 套件介紹 | 18 |
| 3-2-2 Python Pandas 套件介紹 | 19 |
| 3-2-3 NumPy 及 Pandas 交互關係 | 20 |
| 四、 工作心得 | 23 |

一、前言

職目前承辦工作為「太陽能案場發電效率統計自動化」及「太陽能案場發電效率統計表優化」。

「太陽能案場發電效率統計自動化」主要協助經理室專案組人員將每日手動抓取太陽能案場資訊之程序自動化，每日共需抓取新港紡紗、紡紗南側、紡紗北側、複材南/北側、公用保養廠及苛化房、麥寮 PTA 及 AROMA 廠、寧波煤場等共 17 案場，抓取內容為當日發電量(kWh)、容量因素%、PR 值、模組溫度等，每日彙總於 Excel 報表中，提供業主做營運分析與管理。

「太陽能案場發電效率統計表優化」主要將第一案所蒐集之資料，依照不同廠區/部門進行統計分析，廠區分成新港、寧波、麥寮、福懋等區域，而部門則分成工務部、它部、福懋等三部分，各廠區/部門可再細分成各個案場，並依照當月累計「容量因素%」進行排序，可讓業主一眼看出各案場發電效率優劣。

此兩案旨在協助太陽能發電管理人員將每日所需進行的太陽能案場發電效率統計作業，以自動化程式替代人工操作，可為業主節省大量時間，提升工作效率。

本季工作心得，以上述二案為主軸，詳如後續說明。

二、太陽能案場發電效率統計自動化

本案主要以爬蟲及 API 串接形式蒐集各太陽能案場資料，於自動化改善前，使用者須每日於寧波、福懋、台化新港紡紗、盛齊綠電等四個不同網站平台，每日抄寫 17 個案場所需資料，資料蒐集過程繁瑣複雜，因此本案撰寫自動化程式每日早上八點自動抓取上述各案場資料，經統計後自動產出報表，以利業主及時獲取所需資訊並減輕人工負擔，案場列表如下表 2-1 所示。

| 太陽能案廠彙總表 | | | | | | | | | |
|----------|-----------------|---|------------------|----|-------------------|----|-------|----|------|
| 1 | 新港紡紗廠南面 | 5 | 麥寮PTA CCR大樓 | 9 | (舊)原絲六廠屋頂 | 13 | 嘉義新港廠 | 17 | 寧波煤場 |
| 2 | 新港紡紗廠北面 | 6 | 新港PABS複材廠 | 10 | 新港公用廠苛化廠房及MGO倉庫屋頂 | 14 | 福懋S9A | | |
| 3 | 麥寮-工事處和資材辦公室連倉庫 | 7 | 麥寮SM3廠CCR屋頂 | 11 | 麥寮PTA保養廠 | 15 | 福懋S9B | | |
| 4 | 麥寮ARO2廠CCR屋頂 | 8 | 新港公用廠保養工廠及施工場所屋頂 | 12 | 麥寮-預製場及資材倉庫屋頂 | 16 | 福懋S7A | | |

表 2-1、太陽能案場彙總表

2-1 系統架構

以下為本案工程開發時所用到的程式語言、套件及測試工具，如表 2-2 所示。系統架構圖如圖 2-1 所示，使用 Python 語言之第三方套件「Requests」，配合測試工具 Postman 開發網頁爬蟲及 API 串接程式，再以 Python 語言設計下載程式

配合第三方套件「Openpyxl」將下載之數據彙總於每月 Excel 報表中，供使用者查看。

| | |
|---------|--------------------------|
| 作業系統 | Windows 10 |
| 程式語言、套件 | Python、Requests、Openpyxl |
| 測試工具 | Postman |

表 2-2、系統環境

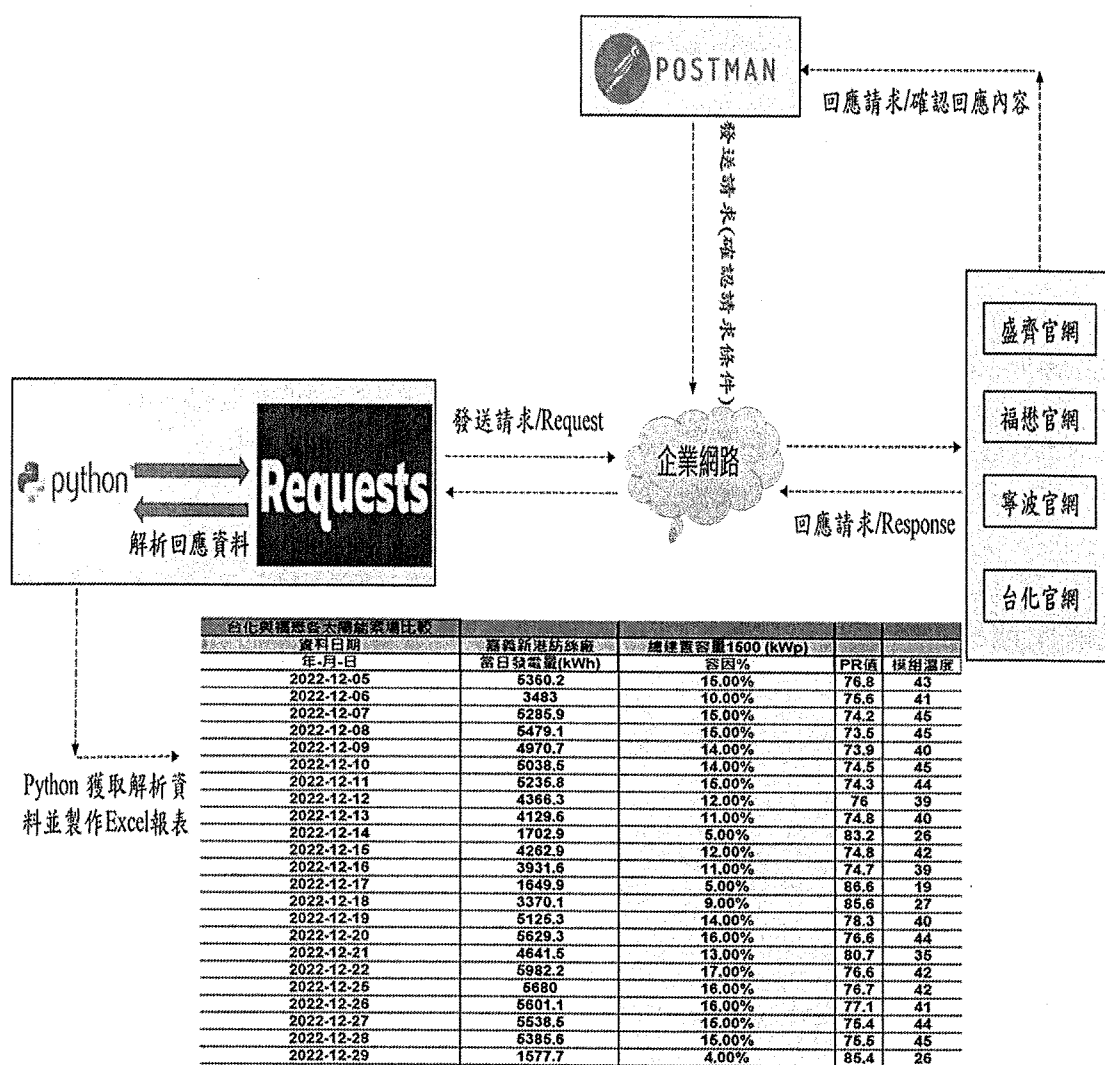


圖 2-1、系統架構圖

2-2 Python 網路爬蟲介紹

網路爬蟲 (spider 或 web crawler)，是一種可以自動瀏覽全球資訊網的網路機器人，許多的搜尋入口網站 (例如 Google)，都會透過網路爬蟲收集網路上的各種資訊，進一步分析後，提供使用者搜尋的資料；許多 AI 程式開發者也會自行開發不同的爬蟲程式，進行大數據收集與分析的動作。

而網頁爬蟲又可大致分為兩種分別為靜態網頁爬蟲及動態網頁爬蟲，靜態網站爬蟲是指網站完成一個請求 (request) 與回應 (response) 後，用戶端即不再與伺服器有任何的交流，所有的互動都只與瀏覽器的網頁互動，資訊不會傳遞到後端伺服器如圖 2-2；動態網站爬蟲是指網站會依照使用者的行為不斷的與伺服器進行交流如圖 2-3，常見的動態爬蟲需要進行「登入驗證」的動作，使用者需告知目標伺服器登入資訊(帳號、密碼等)，當目標伺服器確認使用者身分後，才可繼續進行後續的資料擷取動作。

通常動態網站爬蟲實作比較複雜，撰寫爬蟲程式時必須要知道網站需要使用者提供那些資訊，才能取得所需要的資料，而靜態網站爬蟲比較容易實作，只要爬蟲已經閱讀完整份網頁，就可以取得這個網頁所有的資訊進行分析。

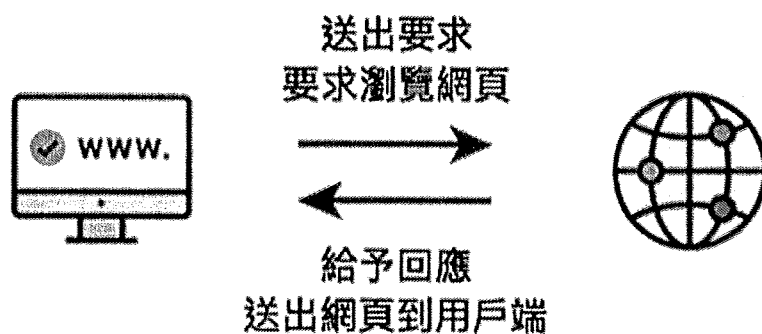


圖 2-2、Python 靜態網路爬蟲示意圖



圖 2-3、Python 動態網路爬蟲示意圖

本案所~~所~~撰寫的爬蟲程式簡單來說是一個可以自動化抓取網頁內容的程式，以瀏覽網頁為例，平常我們瀏覽網站，就是對目標網頁伺服器提出一個請求(Request)，這時候目標伺服器就會收到資訊，並且回傳(Response)給前端 html, css, javascript 等程式資訊，接著再由瀏覽器會把這些資訊渲染成所呈現之網頁，同樣的爬蟲原理就是，藉由你的爬蟲程式去

進行 Request，等待程式接收到 Response 後，解析出需要的資訊，進行更進一步應用。

以下用本案為例，簡易的架構如下，藉由初始網址 (URL)，驗證登錄資訊(目標伺服器收到請求中夾帶的資訊例如帳號、密碼等，經伺服器確認使用者身份後，會返回一組 TOKEN 或建立 SESSION 連線，以作為使用者後續操作時的身分判斷)，接著對目標網址發出 Request 請求，等待目標伺服器回應後，藉由網頁解析器獲取解析(網頁解析器能將伺服器傳送回來之原始碼轉換成能夠被理解閱讀之程式代碼，常用的有 Python 自帶的 html.parser 或 Python 第三方套件 beautifulsoup、Xpath 等)，最後再寫入 Excel 報表中，其流程如下圖 2-4、圖 2-5 及圖 2-6 所示。

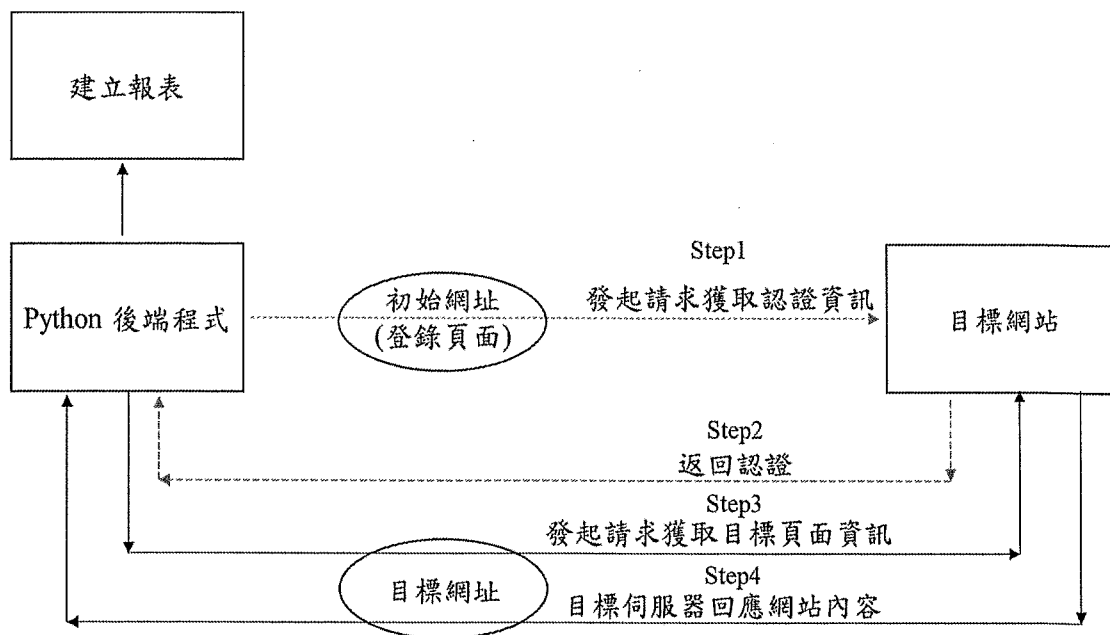


圖 2-4、Python 網路爬蟲流程圖

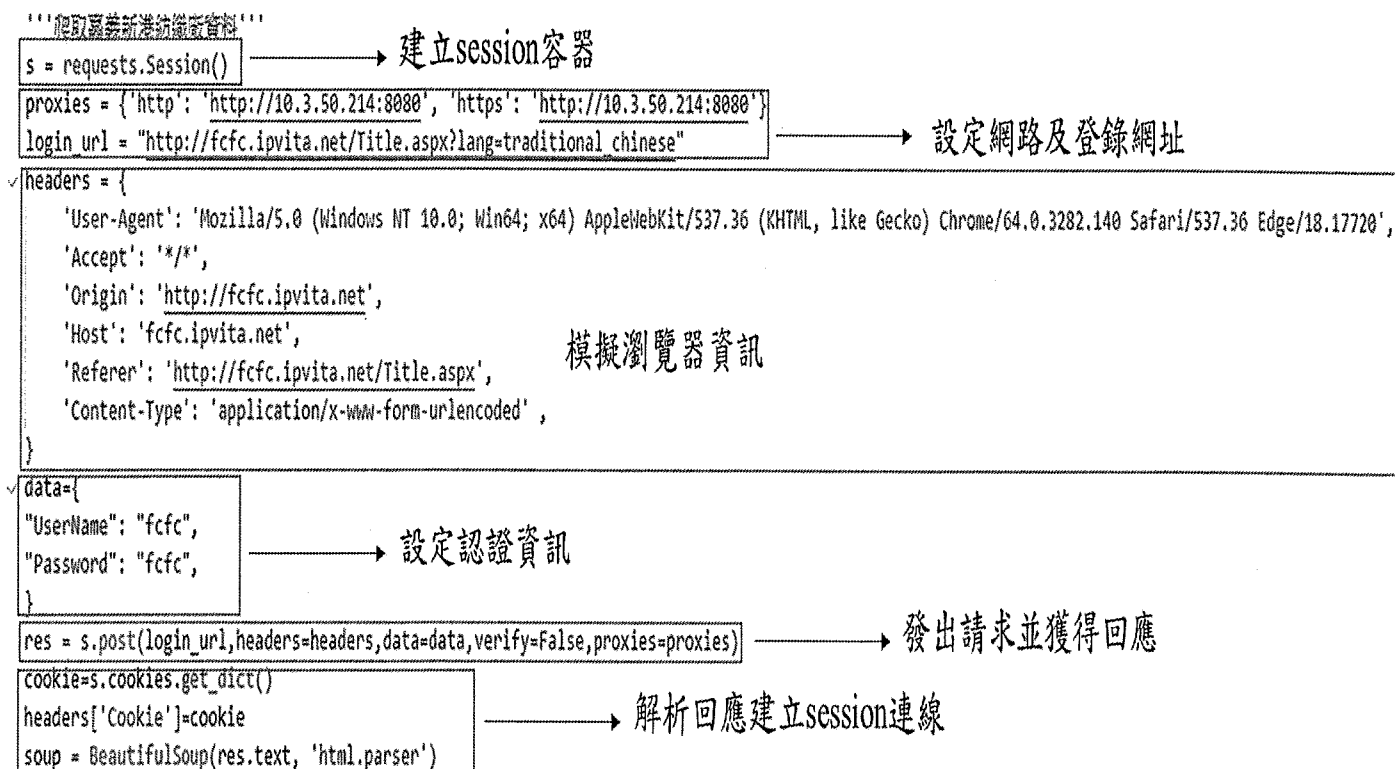


圖 2-5、Python 網路爬蟲程式示意圖(驗證登錄資訊)

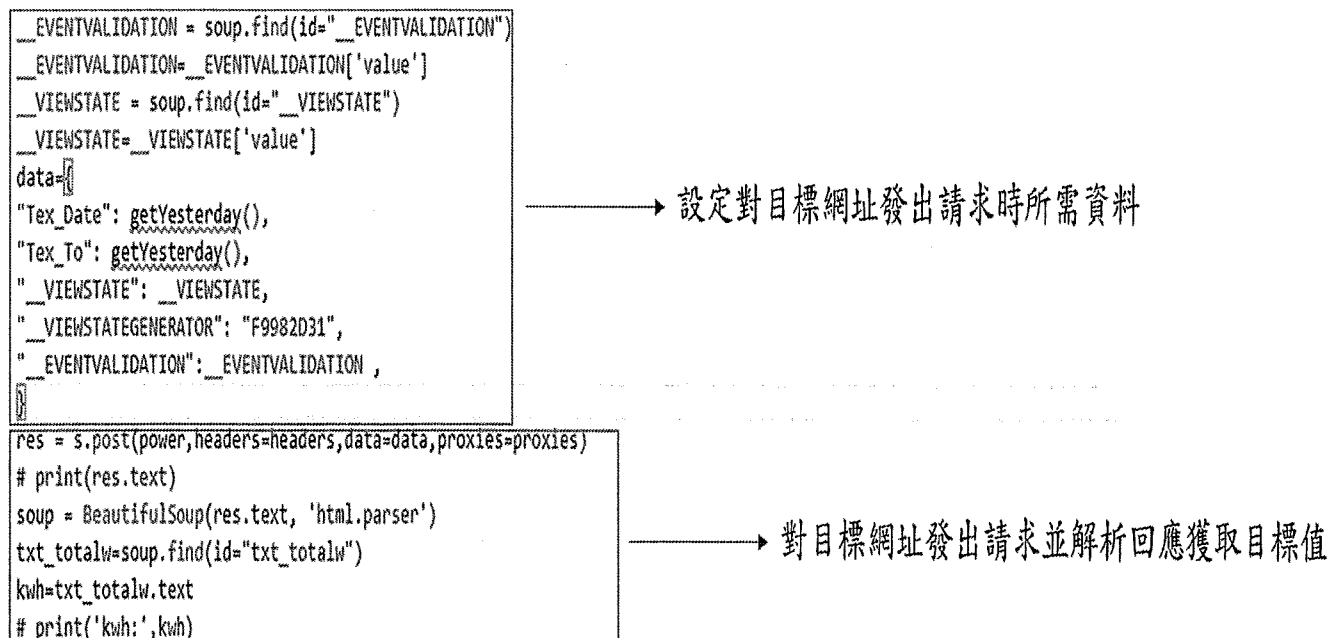


圖 2-6、Python 網路爬蟲程式示意圖(獲取目標值)

2-3 Python API 串接介紹

應用程式介面 Application Programming Interface (API) 是一種程式設計介面，它定義程式與程式之間的互動介面，像是如何進行呼叫或發出請求，應使用的資料格式，及遵循的設計規範等。

簡單來說，API 的工作是讓 A、B 兩端程式進行資料傳遞，一端提供資料輸入、另一端依據資料回傳結果，其目的在於「不必了解對方的技術與邏輯即可加速開發的共識」。

在產品開發或者網頁開發中，資料 API 介面就像後端遞給前端的數據連接線。後端程式需要呈現的資料，只需要通過這根數據線將資料傳輸給前端的程式，前端就可將其按需求進行後續的視覺化，而且這個介面還可供外界使用，如同本案以 API 串接盛齊綠電網站並獲取所需資訊。

與網頁爬蟲不同的是，API 的設計更為簡單高效，此介面已經存儲好使用者所需要的資料，使用者不需要再花費額外的時間去解析網頁，如圖 2-7 及圖 2-8 所示。

<https://solar.billionwatts.com.tw/WsApp/information>

API 網址

Send Json Example

```

{
  "api": "ctActinometerRawData",
  "token": "86878f407120a519be88db88bdb868adD0001",
  "para": {
    "siteNo": "P2017003",
    "dataNo": "00048200",
    "startDatetime": "2019-06-26 13:01:00",
    "endDatetime": "2019-06-26 13:02:00"
  }
}

```

API 文件定義需傳送給目標伺服器之參數資訊

| Code | Name | Data Type | Remark |
|---------------|--------|-----------|---|
| api | API 名稱 | string | |
| token | 驗證碼 | string | 參照 1.2 token 驗證碼 |
| langCode | 傳送語系 | string | zh_TW, en_US |
| para | | | |
| siteNo | 電站代碼 | string | |
| dataNo | 設備識別碼 | string | API 傳遞參數說明 |
| startDatetime | 開始日期時間 | datetime | 當地資料觀測時間 YYYY-MM-DD HH:mm:ss, 限定 3 個月內 |
| endDatetime | 結束日期時間 | datetime | 當地資料觀測時間 YYYY-MM-DD HH:mm:ss, 開始日期時間及結束日期時間限定為同月份, 不可跨月, 查詢區間不可超過 2 天 |

圖 2-7、盛齊綠電 API 文件說明

```
data = {
  "api": "ctActinometerRawData",
  "token": "565798ed630328ca70188d74dd55afd7N000171574",
  "langCode": "zh_TW",
  "para": {
    "siteNo": siteNo,
    "dataNo": dataNo,
    "startDatetime": "2023-01-04 13:00:26",
    "endDatetime": "2023-01-04 14:00:26",
  }
}

res = requests.post('https://solar.billionwatts.com.tw/WsApp/information', proxies=proxies, json = data, verify=False)
res=json.loads(res.text)
data=res['data']
```

設定對API網址發出請求時所需資料

對API網址發出請求並獲取目標值

圖 2-8、Python API 串接程式示意圖

2-4 HyperText Transfer Protocol 傳輸協定介紹

在前面章節中有提到網頁的請求(Request)及回應(Response)，其為 HyperText Transfer Protocol (HTTP) 傳輸協定的組合結構，HTTP 是一種用來傳輸超媒體文件（例如 HTML 文件）的應用層協定，被設計來讓瀏覽器和伺服器進行溝通。

一般來說傳輸資料的兩端會分為 客戶端 (Client) 跟 伺服器端 (Server)，Client 端以網頁來說就是你的瀏覽器、電腦，主要是發送「請求 request」到 Server 端，而 Server 端收到 request 後開始處理資料，完成後會回傳「回應 response」到 Client 端，其組成結構如圖 2-9 所示，而網頁所呈現的資訊，其實就是眾多的 request 跟 response 交互而成，其步驟可拆解為下列幾點：

1. 瀏覽器發送 HTTP request 到 Server
2. Server 回傳 response 到瀏覽器
3. 瀏覽器解析 response 內容，渲染成網頁面

| HTTP 的組成結構 | |
|-----------------|------------------|
| HTTP Request | HTTP Response |
| Request Headers | Response Headers |
| Request Method | Response Status |
| Request URL | Response Body |
| Message Body | |

圖 2-9、HTTP 的組成結構

從客戶端發出 HTTP Request 時，通常會定義以下的資訊，以瀏覽 Google 首頁為例，如圖 2-10 所示。

| HTTP Request - Headers | | |
|------------------------|-----------------------------------|--------------|
| Method | URL | Message Body |
| GET | <u>www.google.com</u> | — — — |
| POST | <u>www.google.com/person</u> | |
| PATCH | <u>www.google.com/person/id/1</u> | |
| PUT | | |
| DELETE | | |

圖 2-10、HTTP Request 所定義的資訊內容

透過日常的郵政系統來比喻，假如我們今天要寄出一封信件，若我們選擇 GET 掛號的寄件方式，而地址 URL 是 Google 首頁(www.google.com)，Message Body 則會是信件內容。上述的內容代表我們要對 www.google.com 做一個 GET 的動作，才得以進入到 Google 首頁，而常見的 Method 方法如下：

- GET：讀取資料
- POST：新增資料（常搭配 form 標籤使用）
- PATCH：修改資料
- PUT：修改資料（若資料不存在，則新增資料）
- DELETE：刪除資料

通常一個 Method 會搭配一個 URL，也會對應伺服器端一組特定的資源。若單純是瀏覽網頁(例如進入 Google 搜尋頁面)則使用 GET 方法與目標伺服器進行溝通，若需要在此頁面進行搜尋，那客戶端就會需要送出資料(例如於搜索框輸入的關鍵字)，這筆資料則使用 POST 方法傳送給目標伺服器，進而使目標伺服器能夠解析客戶端搜尋需求。

從伺服器端回傳 HTTP Response 時，通常會定義以下的資訊，以瀏覽 Google 首頁為例，如圖 2-11 所示。

| HTTP Response - Headers | | |
|-------------------------|--|-----------------------------|
| Status Code | Content-Type | Response Body |
| 200 | text/html、text/plain 或 application/json | <html>....</html> 通常是程式碼 |
| 300 | | |
| 400 | | |
| 500 | | |

圖 2-11、HTTP Response 所定義的資訊內容

狀態碼 (status code)是一種狀態碼的呈現，也是客戶端與伺服器端溝通的狀態情況，常見的狀態代碼如下。

- 1XX 訊息類 (收到請求，請求者繼續執行操作)
- 2XX 成功類 (操作被成功接受並處理)，例如：200 成功回應
- 3XX 重定向類 (需進一步操作才能完成)，例如：301

成功轉向

- 4XX 客戶端錯誤類（請求語法錯誤或無法完成請求），
例如：404 找不到資源
- 5XX 伺服器錯誤類（後端的問題），例如：500 伺服器
錯誤

content-type 這個項目，它則會定義回應格式，例如 text/html、text/plain 或 application/json 等等，客戶端才會知道該如何打開訊息，而 Response Body 通常會是一個一個的程式碼檔案，最終藉由瀏覽器渲染畫面並呈現於使用者。

2-5 Python Openpyxl 套件介紹

在開發 Python 網頁爬蟲或 API 串接時，除了可以將取得的資料儲存在資料庫外，也能匯出以 Excel 檔案製作成報表，進而使用其中的圖形化功能，來製作統計圖表，有效率的進行資料分析。

Openpyxl 是一個開源專案，其是一個讀寫 Excel 文檔的 Python 庫，其不僅能夠同時讀取和修改 Excel 文檔，也可以對 Excel 檔內儲存格進行詳細設置，包括儲存格樣式等內容，甚至還支援圖表插入、列印設置等內容，使用 openpyxl 可以讀寫 xltm、xltx、xlsm 及xlsx 等類型的檔案，也可以處

理資料量較大的 Excel 檔，跨平臺處理大量資料是其它模組沒法相比的，因此，openpyxl 成為處理 Excel 檔案的首選函數庫。

而本案則利用此套件將每日搜集的資料彙總於每月 Excel 報表中，供使用者查看，如圖 2-12 及圖 2-13 所示。

| 台化與福懋各太陽能案場比較 | | | | | |
|---------------|------------|-----------------|------|------|--|
| 資料日期 | 嘉義新港紡絲廠 | 總建置容量1500 (kWp) | | | |
| 年-月-日 | 當日發電量(kWh) | 容因% | PR值 | 模組溫度 | |
| 2022-12-05 | 5360.2 | 15.00% | 76.8 | 43 | |
| 2022-12-06 | 3483 | 10.00% | 75.6 | 41 | |
| 2022-12-07 | 5285.9 | 15.00% | 74.2 | 45 | |
| 2022-12-08 | 5479.1 | 15.00% | 73.5 | 45 | |
| 2022-12-09 | 4970.7 | 14.00% | 73.9 | 40 | |
| 2022-12-10 | 5038.5 | 14.00% | 74.5 | 45 | |
| 2022-12-11 | 5235.8 | 15.00% | 74.3 | 44 | |
| 2022-12-12 | 4366.3 | 12.00% | 76 | 39 | |
| 2022-12-13 | 4129.6 | 11.00% | 74.8 | 40 | |
| 2022-12-14 | 1702.9 | 5.00% | 83.2 | 26 | |
| 2022-12-15 | 4262.9 | 12.00% | 74.8 | 42 | |
| 2022-12-16 | 3931.6 | 11.00% | 74.7 | 39 | |
| 2022-12-17 | 1649.9 | 5.00% | 86.6 | 19 | |
| 2022-12-18 | 3370.1 | 9.00% | 85.6 | 27 | |
| 2022-12-19 | 5125.3 | 14.00% | 78.3 | 40 | |
| 2022-12-20 | 5629.3 | 16.00% | 76.6 | 44 | |
| 2022-12-21 | 4641.5 | 13.00% | 80.7 | 35 | |
| 2022-12-22 | 5982.2 | 17.00% | 76.6 | 42 | |
| 2022-12-25 | 5680 | 16.00% | 76.7 | 42 | |
| 2022-12-26 | 5601.1 | 16.00% | 77.1 | 41 | |
| 2022-12-27 | 5538.5 | 15.00% | 75.4 | 44 | |
| 2022-12-28 | 5385.6 | 15.00% | 75.5 | 45 | |
| 2022-12-29 | 1577.7 | 4.00% | 85.4 | 26 | |

圖 2-12、Excel 每月彙總報表示意圖

(scrapy) D:\0.工程案\03-太陽能案場發電效率統計自動化\new lp 新案廠 (24 and 25)\scrapycase\python scrapytotal.py

爬取嘉義新港紡絲廠資料

{'PM00464': {'date': '2023-01-11', 'PR': '85.9', 'ESH': '1.55', 'temp': '28.86', 'kwh': '2275.7', 'kwp': '1500', 'name': '嘉義新港紡絲廠'}}

爬取福懋案場資料

{'S9A059857A': {'date': '2023-01-11', 'PR': '88.15404577089945', 'ESH': '1.55', 'temp': '29.076000213623047', 'kwh': '4381.1279999999982', 'kwp': '2600.32', 'name': '福懋 S9A/S9B/S7A Total'}}

爬取寧波案場資料

{'NB': {'date': '2023-01-11', 'PR': '11.93', 'ESH': '1.55', 'temp': 'None', 'kwh': '5550.16', 'kwp': '2053.0', 'name': '寧波台塑案場'}}

爬取盛豐案場資料

寫入json檔案

json檔案寫入完畢

資料抓取完畢，json檔案寫入完畢，即將寫入EXCEL報表

檔案存在，新增資料

排程式定時抓取資料

圖 2-13、定時資料抓取式畫面

三、太陽能案場發電效率統計表優化

本案主要將第一案所蒐集之資料，依照不同廠區進行統計分析，計算當月累積發電量及累積容量因素之加權平均，於自動化改善前，使用者須手動將每日報表計算後進行排序比較，因此本案撰寫自動化程式每日自動統計上述報表資料，統計後之報表，可讓使用者清楚看出各案場「容量因素%」大小排序，節省人工作業時間。

3-1 系統架構

以下為本案工程開發時所用到的程式語言及套件如表 3-1 所示。系統架構圖如圖 3-1 所示，使用 Python 語言之第三方統計分析套件「Numpy、Pandas」，配合前述提及之第三方套件「Openpyxl」將統計分析完之資料，呈現於每月統計報表中，以供使用者查看。

| | |
|---------|------------------------------|
| 作業系統 | Windows 10 |
| 程式語言、套件 | Python、Numpy、Pandas、Openpyxl |

表 3-1、系統環境

太陽能案場發電效率月報表

| 台化與福懋營太極能案場比較 | | | | | |
|---------------|--------|-----------------|------|----|--|
| 資料日期 | 新港新港紗廠 | 總裝置容量1600 (kWp) | | | |
| 2022-12-05 | 6300.2 | 16.00% | 76.8 | 43 | |
| 2022-12-06 | 3493 | 16.00% | 75.6 | 41 | |
| 2022-12-07 | 6286.9 | 16.00% | 74.2 | 46 | |
| 2022-12-08 | 6479.1 | 16.00% | 73.6 | 46 | |
| 2022-12-09 | 4970.7 | 14.00% | 73.9 | 40 | |
| 2022-12-10 | 6038.6 | 14.00% | 74.6 | 46 | |
| 2022-12-11 | 6236.8 | 16.00% | 74.3 | 44 | |
| 2022-12-12 | 4366.3 | 12.00% | 76 | 39 | |
| 2022-12-13 | 4129.6 | 11.00% | 74.8 | 40 | |
| 2022-12-14 | 1702.9 | 6.00% | 83.2 | 28 | |
| 2022-12-15 | 4262.9 | 12.00% | 74.8 | 42 | |
| 2022-12-16 | 3931.6 | 11.00% | 74.7 | 39 | |
| 2022-12-17 | 1648.9 | 6.00% | 86.6 | 19 | |
| 2022-12-18 | 3370.1 | 9.00% | 86.6 | 27 | |
| 2022-12-19 | 6126.3 | 14.00% | 78.3 | 40 | |
| 2022-12-20 | 6029.3 | 16.00% | 76.6 | 44 | |
| 2022-12-21 | 4641.6 | 13.00% | 80.7 | 36 | |
| 2022-12-22 | 6982.2 | 17.00% | 76.6 | 42 | |
| 2022-12-23 | 6680 | 16.00% | 76.7 | 42 | |
| 2022-12-24 | 6601.1 | 16.00% | 77.1 | 41 | |
| 2022-12-25 | 6638.6 | 16.00% | 76.4 | 44 | |
| 2022-12-26 | 6386.6 | 16.00% | 76.6 | 46 | |
| 2022-12-27 | 1677.7 | 4.00% | 86.4 | 26 | |



Python 搭配 Pandas 讀取當月
報表資料並進行統計分析

| 台化與福懋營太極能案場比較 | | | | | | | | | | |
|---------------|------------|-----------------|---------------------|-------|----------|--------------|---------------------|-------|----------|--------------|
| 0 | 資料日期 | 新港S43廠 CCR廢排 | 總裝置容量146.7 (kWp) | | | 新港PTA廠 廢排 | 總裝置容量115.2 (kWp) | | | 新港-新製場 廢排 |
| 1 | 年-月-日 | 發電量 (kWh) | 容量% | PR值 | 模組 溫度 | 發電量 (kWh) | 容量% | PR值 | 模組 溫度 | 發電量(kWh) |
| 2 | 2023-01-04 | 205.3 | 0.06 | 76.82 | 42 | 336.8 | 0.12 | 82.85 | 42 | 1458.6 |
| 3 | 2023-01-06 | 0 | 0 | 0 | 37 | 192.5 | 0.07 | 81.71 | 38 | 823.6 |
| 4 | 2023-01-08 | 0 | 0 | 0 | 44 | 350.7 | 0.13 | 75.5 | 41 | 1578.7 |
| 5 | 2023-01-09 | 0 | 0 | 0 | 38 | 242.9 | 0.09 | 82.78 | 34 | 1079.5 |
| 6 | 2023-01-10 | 0 | 0 | 0 | 35 | 134.3 | 0.05 | 88.4 | 33 | 650.7 |
| 7 | 2023-01-11 | 0 | 0 | 0 | 29 | 88.8 | 0.03 | 90.72 | 29 | 426.6 |
| 8 | 2023-01-12 | 0 | 0 | 0 | 40 | 193.3 | 0.07 | 88.77 | 37 | 913.2 |
| 9 | 2023-01-15 | 0 | 0 | 0 | 43 | 32.1 | 0.12 | 79.69 | 41 | 1459.3 |
| 10 | 2023-01-16 | 0 | 0 | 0 | 26 | 196.7 | 0.05 | 91.49 | 25 | 726.3 |
| 11 | 2023-01-17 | 0 | 0 | 0 | 39 | 92.3 | 0.03 | 94.39 | 35 | 473.1 |
| 12 | 2023-01-18 | 233.8 | 0.07 | 46.29 | 39 | 235.9 | 0.09 | 82.82 | 35 | 1104.4 |
| 13 | 2023-01-29 | 616.1 | 0.17 | 85.33 | 39 | 433.7 | 0.16 | 78.75 | 38 | 1956 |
| 14 | 2023-01-30 | 650.5 | 0.18 | 83.09 | 46 | 463.1 | 0.17 | 74.3 | 40 | 2077.2 |



Python 搭配 Openpyxl
匯出分析報表

| 工務部2022年12月份 太陽能發電量統計 | | | | | |
|-----------------------|----|--------------|---------------|----------------|-------------|
| 廠區 | 項次 | 案場 | 裝置容量 (kWp) | 發電量 (kWh/月) | 容量因數 (%) |
| 新港 | 1 | 新港紡紗廠(南側舊案場) | 1,497 | 138,789 | 12.46% |
| | 2 | 新港紡紗+PABS廠 | 5,183 | 476,304 | 12.35% |
| | 3 | 新港公用廠保養/淨化廠房 | 346 | 30,234 | 11.74% |
| 麥寮 | 4 | 麥寮工事資材 | 881.0 | 67,936 | 10.37% |
| 台灣廠區小計 | | | 7,909 | 713,263 | — |
| 寧波 | 5 | 寧波煤場 | 2,053 | 186,293 | 12.20% |
| 工務部合計 | | | 9,962 | 899,556 | — |

圖 3-1、系統架構圖

3-2 Python 統計分析套件介紹

3-2-1 Python NumPy 套件介紹

NumPy 是 Python 的一個用於科學計算的基礎套件。它提供了多維陣列物件，多種衍生的物件（例如隱藏陣列和矩陣）和一個用於陣列快速運算的混合的程式，包括數學，邏輯，排序，選擇，I/O，離散傅立葉轉換，基礎線性代數，基礎統計操作，隨機模擬等等。

NumPy 陣列和標準 Python 序列有以下幾點區別：

1. NumPy 陣列創建的時候有一個固定的大小，不像 Python 清單可以動態地增加要素。改變 array 的大小會創建一個新的陣列並刪除原陣列。
2. NumPy 陣列中的要素必須是同一種資料類型，從而在記憶體中是同樣的大小。唯一例外是可以由 Python（包括 NumPy）物件作為要素組成陣列，因此允許有要素大小不同的陣列的存在。
3. NumPy 陣列更有利於大規模資料的高級數學運算。通常來說，這些運算執行更高效，並且代碼量比用 Python 自帶的序列來實現更少。

目前越來越多的科學和數學計算使用 Python 套件 NumPy 陣列，雖然這些套件通常支援 Python 序列輸入，但

它們通常在處理前把輸入轉化為 NumPy 陣列。因此，想要高效地使用基於 Python 的科學或數學計算軟體，只是瞭解如何使用 Python 內置的序列類型已經無法滿足，更必須知道如何使用 NumPy 陣列。

3-2-2 Python Pandas 套件介紹

Pandas 模組是 Python 語言一個高效能、簡易使用的資料操作和分析工具，最初由 AQR Capital Management 於 2008 年 4 月開發並於 2009 年底開源出來，目前由專注於 Python 資料包開發的 PyData 開發 team 繼續開發和維護，屬於 PyData 專案的一部分。Pandas 最初被作為金融資料分析工具而開發出來，因此 Pandas 為時間序列分析提供了很好的支援，Pandas 的名稱來自於面板資料 (panel data) 和 python 資料分析 (data analysis)，可以將它看作是微軟 Excel 或是 MacOS Numbers 試算表軟體的程式版本，就是透過 Python 來操作像試算表內的資料。

主要特色如下：

1. 方便簡單的異質數據讀取、轉換和處理。
2. 提供兩種主要的資料結構：Series 用來處理時間序列相關的資料，是個擁有標籤的一維陣列；DataFrame

則是用來處理像表格一樣的二維的資料，並且可以隨意更改結構和任意欄位的資料型態。

3. 資料載入 Pandas 的資料結構物件後，可以提供的方法，統一進行資料處理，如合併、補值，過濾或取代等等。
4. 多輸入來源及多輸出整合，例如：可以從資料庫讀取資料進行處理後，可以輸出到資料庫或檔案等

3-2-3 NumPy 及 Pandas 交互關係

目前 NumPy 與 Pandas 均常被使用於資料分析，其彌補了 python 語言本身不足之處，最主要是因為 Python 自身的動態語言特性而帶來的運行速度方面的損失，造成 Python 在大資料處理方面的時候效率不佳。

對於運算速度方面的損失，NumPy 已經做了相當程度的優化，可以對大陣列的資料進行高效處理。

優化包括下列所述

1. NumPy 是在一個連續的區塊中存儲資料，獨立於其他 Python 內置物件，如此便可以加速資料索引的速度。
2. NumPy 調用了大量的用 C 語言編寫的演算法庫，使

得其可以直接操作記憶體，不必進行 Python 動態語言特性所含有的前期類型檢查工作，從而大大提高了運算速度。

3. NumPy 所獨有的可以在整個陣列上執行複雜的計算也能夠大幅提高運算效率（基於 NumPy 的演算法要比純 Python 快 10 到 100 倍）。

雖然 NumPy 有著以上的種種出色的特性，其本身則難以應付所有資料分析需求，一方面由於 NumPy 幾乎僅專注於陣列處理，另一方面則是資料分析牽涉到的資料特性眾多，需要處理各種表格和混雜資料，遠非純粹的陣列可以解決的，而這些就須以 Pandas 進行處理。

Pandas 含有使資料清洗和分析工作變得更快更簡單的資料結構與操作工具。經常是和其他工具一起使用，如數值計算工具 NumPy 和 SciPy，分析庫 statsmodels、機器學習庫 scikit-learn，以及資料視覺化庫 matplotlib。其中 NumPy 則是構建 Pandas 的基礎，後者大量借鑒了 NumPy 編碼風格。

Pandas 功能特性廣泛，其包含的函數類型也眾多，資料結構有 Series 與 DataFrame，函數類型有索引函數、彙總函數、載入以及保存眾多檔案格式函數、與資料庫交互函數、字串處理函數、缺失資料處理函數、簡單的繪圖函數、資料聚合

(groupby) 分組運算 (apply) 函數、透視表交叉表函數以及時間序列處理方面的各種函數。

Python 因為有了 NumPy 與 Pandas 而不同於 Java、C# 等程式語言，成為了數據分析領域的首選，而本案則利用此兩套件將 Excel 報表中資料讀取並進行統計分析並呈現於統計分析報表中，供使用者查看，如圖 3-2、圖 3-3 所示。

| 2023年2月份台化公司各太陽能案場再生能源發電量統計 | | | | | | | | |
|-----------------------------|----|----------------|------------|-----------|----------|---------|----------|---------|
| 廠區 | 項次 | 管理單位 | 案場 | 裝置容量(kWp) | | 本月 | | 2023年累計 |
| | | | | 裝置容量(kWp) | 發電量(kwh) | 容量因數(%) | 發電量(kwh) | 容量因數(%) |
| 新港 | 1 | 新港紡紗廠(南側舊案場) | 新港公用廠 | 1,497 | 3,908 | 3.63% | 3,908 | 3.63% |
| | 2 | 新港紡紗+PABS廠 | 新港公用廠 | 5,185 | 14,704 | 3.94% | 14,704 | 3.94% |
| | 3 | 新港公用廠保養/苛化廠房 | 新港公用廠 | 346 | 899 | 0.49% | 899 | 0.49% |
| | 4 | 新港原絲NF6廠房 | 纖維部-原絲廠 | 209 | 623 | 113.00% | 623 | 113.00% |
| | 5 | 新港PABS辦公室及原料倉庫 | 塑膠部-新港PABS | 277 | 尚未完成 | 尚未完成 | 尚未完成 | 尚未完成 |
| | 小計 | | | 7,237 | 20,134 | 121.06% | 20,134 | 121.06% |
| 麥寮 | 6 | 麥寮資材倉庫 | 麥寮資材倉庫 | 518 | 1,182 | 3.18% | 1,182 | 3.18% |
| | 7 | 麥寮工事處及儲運處辦公室 | 新港公用廠 | 364 | 2,035 | 3.21% | 2,035 | 3.21% |
| | 8 | 麥寮PTA廠CCR及保養廠房 | 化三部-麥寮PTA | 313 | 676 | 5.93% | 676 | 5.93% |
| | 9 | 海豐ARO2廠CCR | 化一部-ARO2 | 197 | 448 | 3.19% | 448 | 3.19% |
| | 10 | 海豐SM3廠CCR辦公室 | 化二部-SM3 | 147 | 259 | 2.45% | 259 | 2.45% |
| | 小計 | | | 1,538 | 4,600 | 17.96% | 4,600 | 17.96% |
| 寧波 | 1 | 寧波熱電煤場南側 | 寧波熱電廠 | 2,053 | 14,964 | 10.12% | 14,964 | 10.12% |
| | 小計 | | | 2,053 | 14,964 | 10.12% | 14,964 | 10.12% |
| 合計 | | | | 10,828 | 39,699 | 0.49% | 39,699 | 0.49% |

圖 3-2、Excel 每月累計廠區統計報表示意圖

| 工務部2023年2月份 太陽能發電量統計 | | | | | | |
|----------------------|--------|----|--------------------|-----------|------------|---------|
| 部別 | 廠區 | 項次 | 案場 | 裝置容量(kWp) | 發電量(kwh/月) | 容量因數(%) |
| 工務部 | 新港 | 1 | 新港紡紗廠(南側舊案場) | 1,497 | 3,908 | 3.63% |
| | | 2 | 新港紡紗+PABS廠 | 5,185 | 14,704 | 3.94% |
| | | 3 | 新港公用廠保養/苛化廠房 | 346 | 899 | 0.49% |
| | 麥寮 | 4 | 麥寮工事資材 | 881 | 2,035 | 3.21% |
| | 台灣廠區小計 | | | 7,909 | 21,546 | -- |
| | 寧波 | 5 | 寧波煤場(南) | 2,053 | 14,964 | 10.12% |
| 小計 | | | | 9,962 | 36,510 | -- |
| 他部 | 新港 | 1 | 原絲六廠NF6(9/19併聯) | 209 | 623 | 4.14% |
| | | 2 | 麥寮PTA保養廠(10/5併聯) | 115 | 234 | 2.83% |
| | 麥寮 | 3 | 麥寮ARO2-CCR(11/2併聯) | 197 | 448 | 3.19% |
| | | 4 | 麥寮PTA-CCR(11/22併聯) | 198 | 442 | 3.10% |
| | | 5 | 麥寮SM3-CCR(12/28併聯) | 147 | 259 | 2.45% |
| | 小計 | | | 866 | 2,006 | -- |
| 福懋 | | 1 | 福懋斗六 | 2,600 | 7,766 | 4.15% |

圖 3-3、Excel 每月累計部門統計報表示意圖

四、工作心得

轉眼間，來到自控處已近半年時間。期間承接「太陽能案場發電效率統計自動化」及「太陽能案場發電效率統計表優化」兩案，第一案主要以爬蟲及 API 串接形式蒐集各太陽能案場資料，撰寫自動化程式每日自動抓取各案場資料，經計算後每日自動產出報表，在執行本案過程中學習到了之前較少接觸之前端語法，需要看懂前端的語法格式才能用 Python 程式將所需資料擷取下來。

在設計期間最大困難點為報表格式，由於業主時常需更改報表格式或新增資料，導致寫好的程式需要時常修正，因此更深刻的了解到在接案時與業主需要清楚地溝通，把所有可能的變化都預估在工程中，以免工時捉襟見肘，或對程式刪改造成難以預知的錯誤。

而第二案「太陽能案場發電效率統計表優化」屬於第一案的延伸，將第一案所蒐集之資料進行統計分析，並排序後呈現於 excel 報表中，這次吸取前一案的經驗加強彼此之間配合與溝通，使專案能夠如期且順利的完成。

目前也接手太陽能案場自行開發監控系統，欲以新港公用廠保養屋頂為試行案場，初步通訊程式碼已撰寫完成，

擬先以廠商提供之小型逆變器做測試，以利後續開發。

在這段時間學習到不一樣的專業與處理事情的角度，發現自己還有許多不足的地方需要更努力，也感謝公司給予資源與環境讓我學習與增進自己的能力。