

程式設計與資料科學導論-期末報告

臺北市近五年內空氣污染分析

B10208017 地理三 鄭哲愷

前言與研究動機

第一就是有鑑於過往台灣工業化發展，工業活動的增加造成台灣空氣品質的下降，並且這些污染同時造成人體上的危害，例如：空氣污染物中的PM2.5已被世界衛生組織列為一級污染物、其餘空氣污染物也已被證實對人體有各種不同程度的危害。最後一點則是在新冠疫情之後空氣污染是否有因為人類活動的減少，進而改善空氣品質。所以綜合以上三點，希望能夠透過調查台北市內空氣污染變化，調查出台北市內空氣污染的趨勢變化、空氣污染物與天氣狀況的相關聯性與不同污染物對台北市的污染程度。

臺北市各測站相關資料介紹

根據環保署資料，我們在臺北市總共有七個站點，分別是古亭、萬華、中山、松山、士林、大同與陽明，再根據站點的特性可以分為一般測站、交通與工業測站以及背景測站，一般測站主要監測涵蓋系統中全部污染源，測量整體性的空氣品質；交通與工業測站主要監測針對特定污染源，測量其造成的空氣品質；背景測站則是監測排除系統中全部污染源，測量其背景性的空氣品質，在這七個站點中一般測站是古亭、萬華、中山、松山與士林，交通與工業測站是大同測站，背景測站是陽明，測站分別設置位置是古亭(古亭國小)、萬華(福星國小)、中山(新興國中)、松山(松山國小)、士林(中正高中)、大同(重慶北路與民權西路交叉口)與陽明(陽明山鞍部氣象測站)。

其中這次分析所需監測項目共有臭氧 O₃、懸浮微粒 PM_{2.5} PM₁₀、氮氧化物 NO₂、一氧化碳 CO、二氧化硫 SO₂這六種污染物，氣象資料則是包含氣溫、雨量與風向這三種氣象狀況。

資料來源與處理流程

資料來源：從環境部空氣品質監測網下載歷年資料，選取台北市境內「環境部的空氣品質測站」，因為這次使用的並非是台北市設置站點，所以僅有七個測站，而不使用的原因則是因為資料取得上的困難，台北市設置站點的資料在歷年資料上無法搜尋到符合這次報告目的所需，所以選用環境部的測站資料。

資料處理：

第一步、依照測站名稱與年份讀取，並依照測站名稱與年份存進字典。

第二步、去除欄位數值前空格，並將欄位轉為字串型式

第三步、去除非計算所需欄位

第四步、日期轉換，並新增各種日期形式欄位

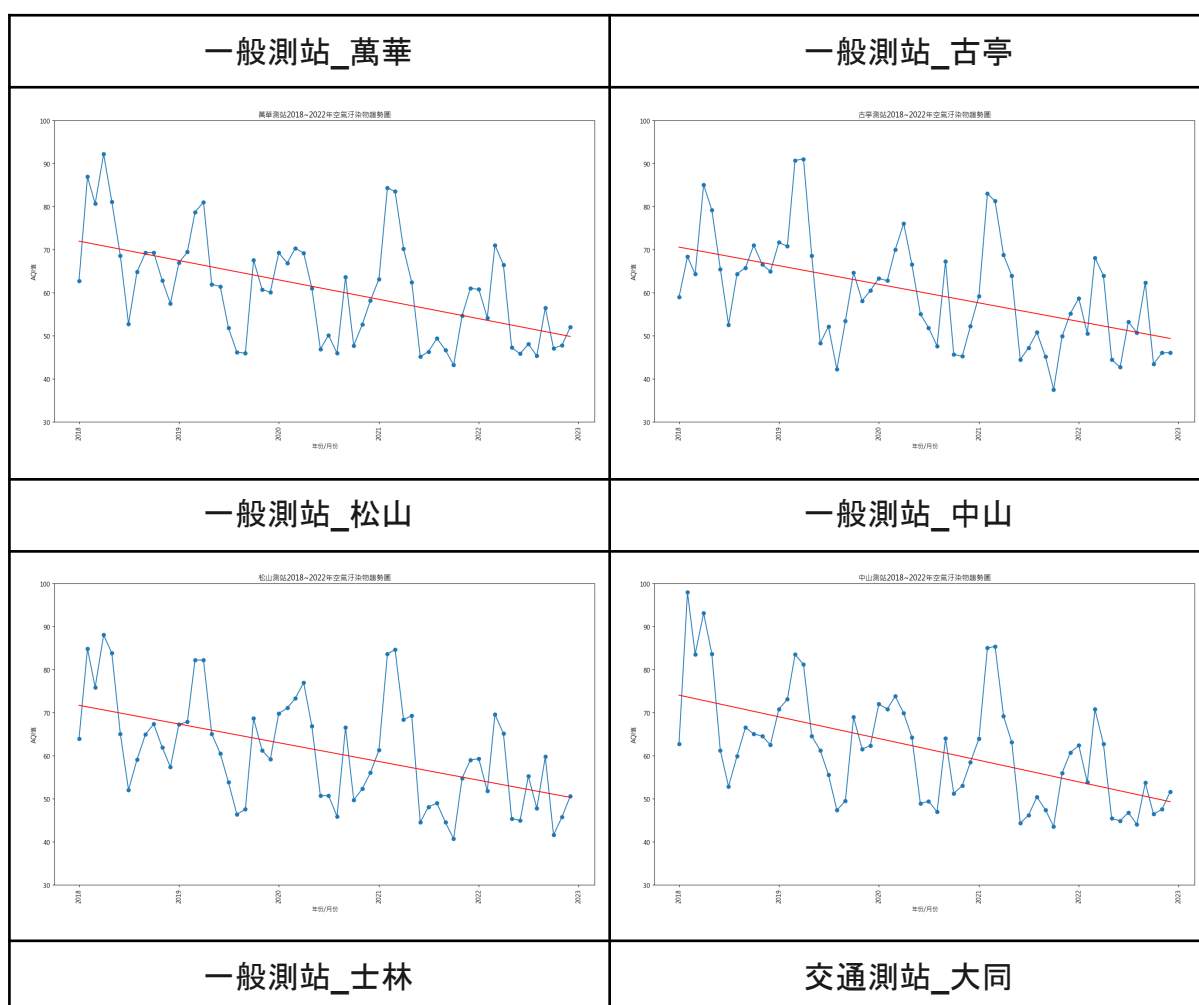
第五步、計算各種污染物濃度的AQI值並存入天氣情況

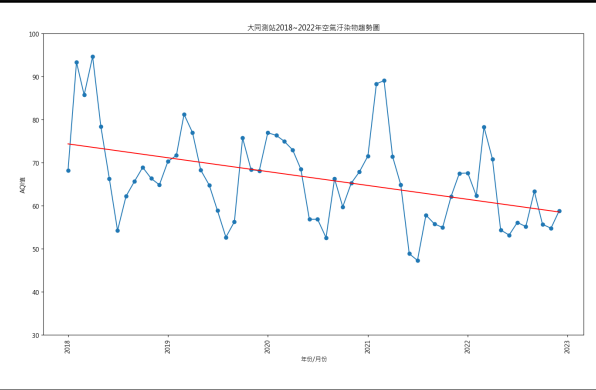
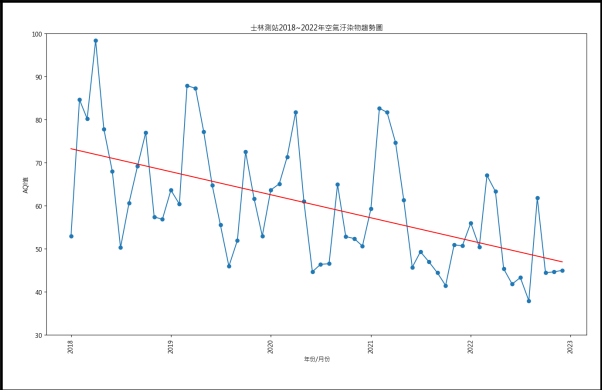
第六步、依照主要三點目標進行統計分析與視覺化輸出

CO	O ₃	SO ₂	NO ₂	PM _{2.5}	PM ₁₀
----	----------------	-----------------	-----------------	-------------------	------------------

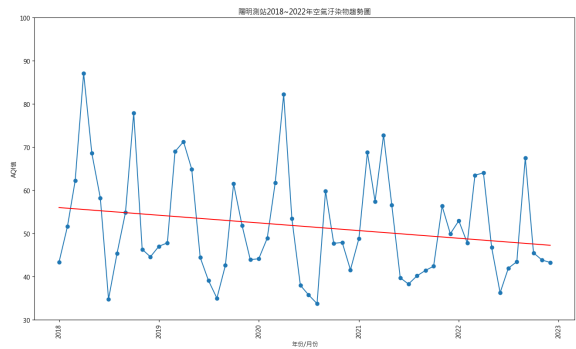
將前8小時濃度 值加總平均	將前8小時濃度 值加總平均， 並四捨五入到 小數點後第一 位	SO2每小時濃 度值轉換AQI， 若計算後AQI 值超過150則 要替換成SO2 的前24小時濃 度平均值	NO2即時濃度 值轉換AQI	0.5 * 前12小時 平均值 + 0.5 * 前4小時平均值 (四捨五入到小 數點第一位) (必須前4小時2 筆有效 & 前12 小時6筆有效)， 並要四捨五入 到小數點後第 二位	0.5 * 前12小時 平均值 + 0.5 * 前4小時平均值 (四捨五入到小 數點第一位) (必須前4小時2 筆有效 & 前12 小時6筆有效)， 並四捨五入到 小數點後第一 位
------------------	--	---	-------------------	---	--

各測站趨勢圖結果分析

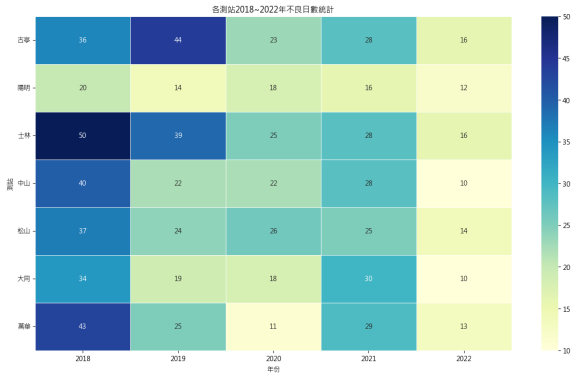




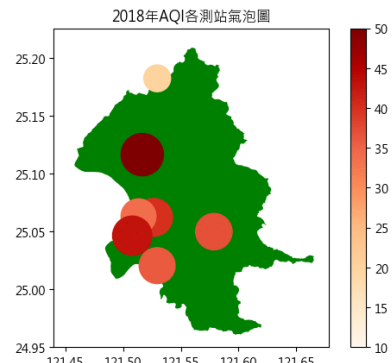
背景測站_陽明



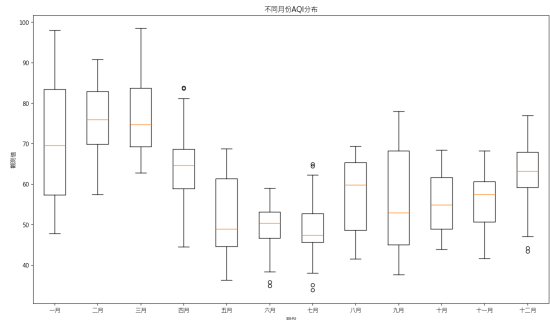
台北市內各測站不良日數熱力圖



台北市內各測站不良日數氣泡圖



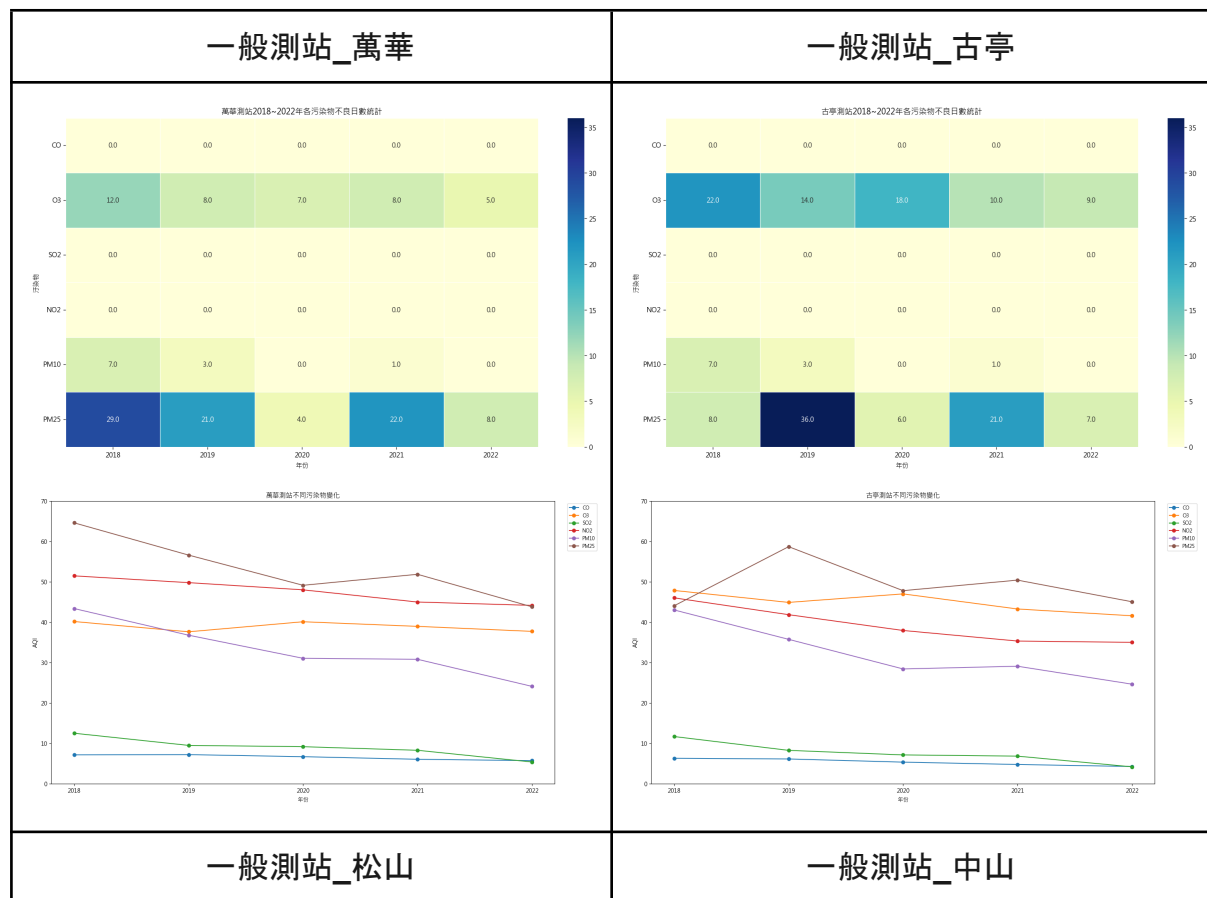
台北市內各月資料箱型圖

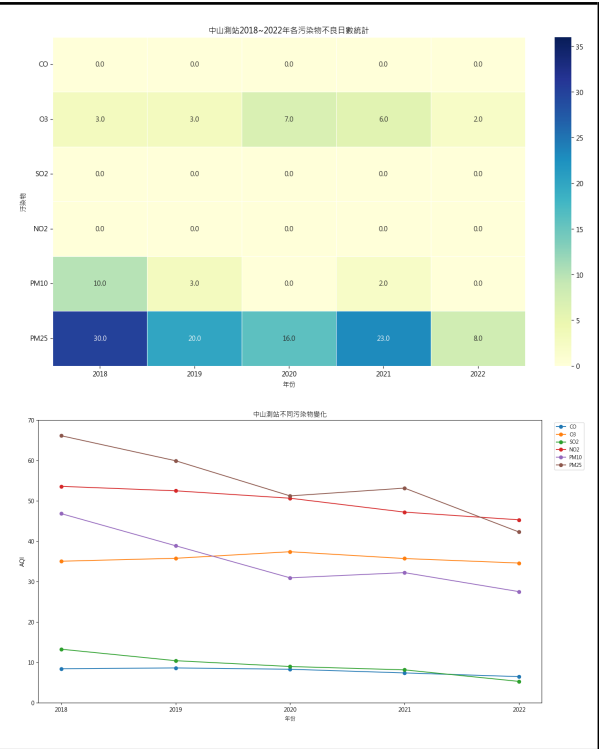
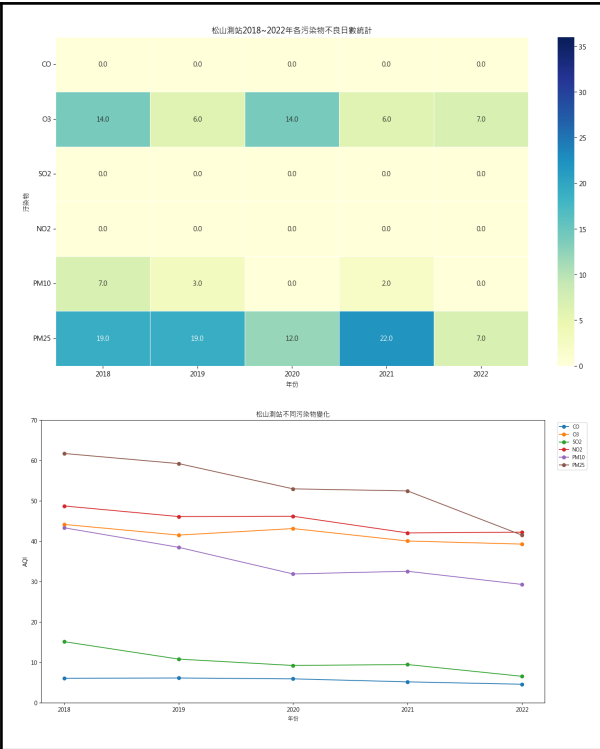


分析

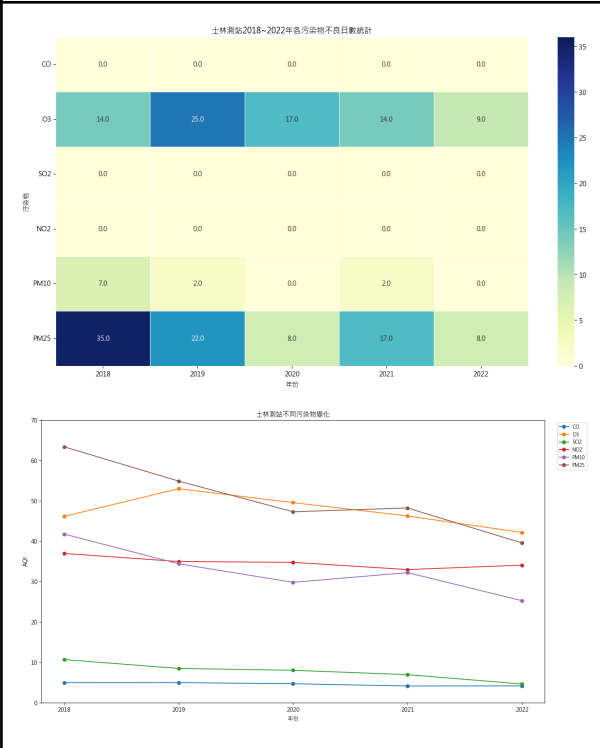
可以從七個測站中的進行簡單回歸後的趨勢線都是下降的趨勢，其中以士林的下降趨勢最為明顯，而陽明測的下降趨勢最為不明顯，而平均值與中位數最高皆是大同測站，最低則是陽明測站，而且可以從一般測站與交通測站看到圖內有約五個明顯的波峰，在依照月分繪製箱型圖之後可以發現台北市境內各月中以夏天(5~7月)的資料中位數與分佈值較低，並且會在冬末春初(1~3月)的資料達到最高中位數與分佈值，最主要原因推測是與盛行風向與氣候有關，因為來自西伯利亞與中國的高壓冷氣團將污染物從中國帶入台灣境內，以及溫度的降低造成污染物的消散難度增加，所以造成空氣污染的惡化。在從各測站每年的不良日數統計來說，從結果我們可以發現從2018年到2022年的不良日數的日數持續減少，但在2022年出現多數測站不良日數的增加，經查證發現這是2018年以後首次出現空氣污染突然惡化的情況，並且也是自2018年以來第一次排名劣於香港、位居東亞第三的情況，具體原因暫時還沒有確切的說法去證明2021年空氣指標的數值惡化。

各測站不同污染物影響結果分析

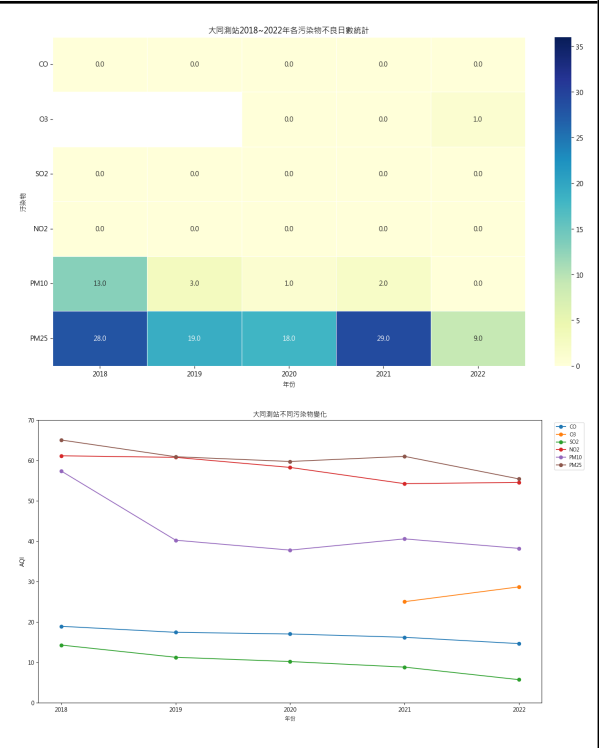




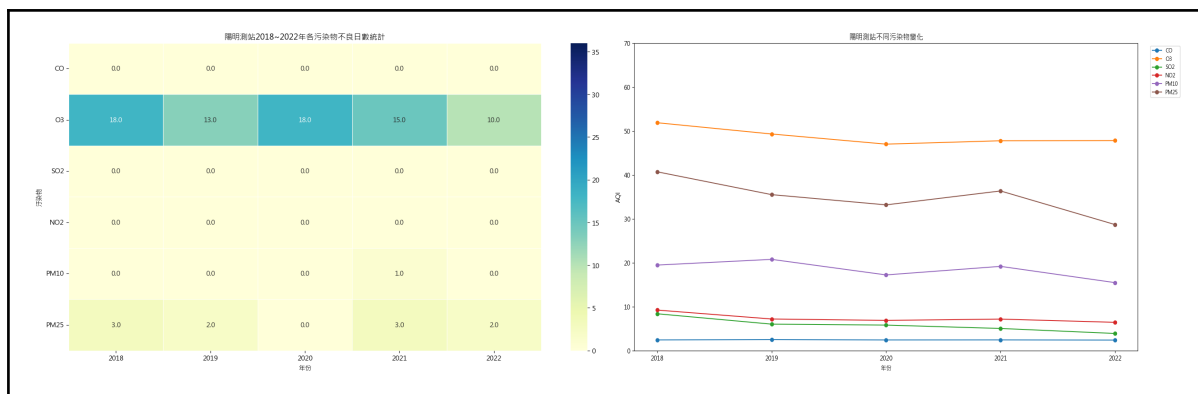
一般測站_士林



交通測站_大同



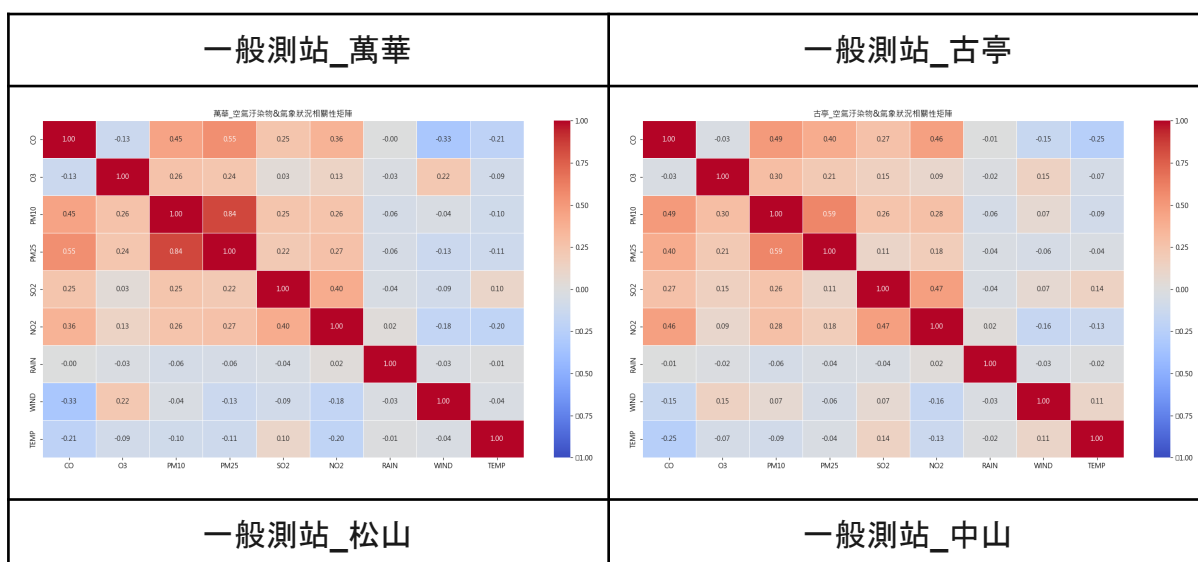
背景測站_陽明

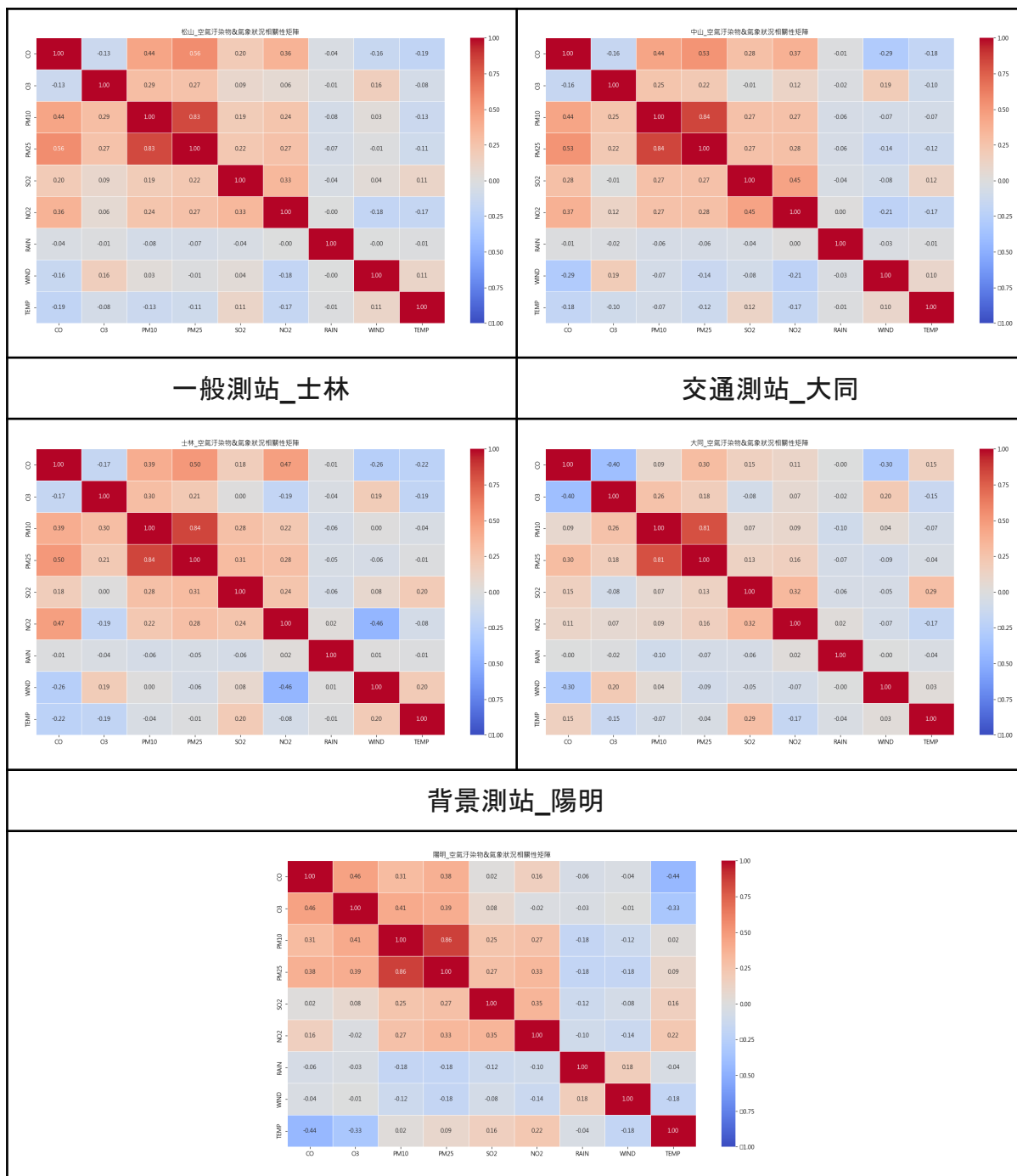


分析

以上是各個測站的各種空氣污染物的年平均折線圖與年不良日數統計圖，可以從每個測站觀察到在台北市境內CO與SO2的影響程度與AQI值都遠低於其他空氣污染物，前面有提到CO與SO2的來源與燃燒與石化產業有關，原因可能是台北市本身無火力發電廠與其他石化產業發展，而一般測站與交通測站的PM2.5與PM10的AQI值與不良日數相對其他污染物都相對較高，所以對交通發達的台北市來說，影響是最嚴重的，最後則是O3，可以發現O3對各個測站都有影響並且AQI值的變動幅度不大，並且可以在陽明測站觀測到影響程度較為嚴重，前面有提到O3是一種衍生性污染物，所以並不是直接排出的污染物，其可能原因是陽明測站的環境本身較容易發生光化學反應，所以造成其污染程度較為嚴重。

各測站不同污染物與氣象情況結果分析





分析

假設: 空氣汙染物與天氣情況有相關性

天氣情況與汙染物互動關係

降雨: 酸沉降(空氣汙染物與雨水結合降落地面, 減少空氣汙染程度)

陽光與溫度: 光化學反應(初級汙染物與光進行反應產生衍生性汙染物)

風速: 擴散與消散(風速的強度會影響汙染物停留在一地的時間與程度)

在分析過程與結果發現, 天氣狀況與不同汙染物之間相關性並不高, 但我的假設是空氣汙染物與天氣情況應該要有關係, 但進一步繪製散佈圖發現資料呈現群聚分佈, 推測可能是因為測量資料多是來自人為汙染物, 或者是汙染物與天氣情況的關

係並不是一對一關係，污染物對整個環境系統的互動所以需要做更加全面性的分析，這是這次分析可能需要改善的部分。

結論

從以上結果來看，台北市境內確實可以發現空氣污染的情況有所改善，但者僅止於這七個測站，從地理分佈來說，這些測站大多集中於台北市人口較為集中的地方，在東部的測站較為少，所以這次研究的缺點是在測站數量的不足，但在大多數台北市人口集中的地方這份研究結果是可行的。並在對人體影響程度的方面，從各個測站的指標污染物與各種污染物的不良日數統計來看，大多數污染物堆人體有害的日數正在減少，所以可以推論台北市的空氣污染情況對人體有害的情況正在減少。但在其象狀況與空氣污染物的關聯性分析來說，我們無法得到明確的污染物對氣象條件的一對一關係，故推測有兩種可能，一是因為人造的污染物本身與天氣情況無關聯性，二是空氣污染物是與整個環境系統互動，所以沒有絕對的一對一關係，而是一對多或是多對多的互動關係。所以總結來說，台北市的空氣污染物整體來說情況有改善，並且在天氣情況的相關性分析發現，這些污染物可能是與整個環境系統互動/是人為污染物所以與天氣情況的單一相關性分析的相關性低。