

統計諮詢期末報告書

台南 T-bike 使用量分析

統計 111
H24071215
陳柔漪

統計 111
H24076029
劉米婷

統計 111
E44076178
王思蘋

統計 111
H24071312
劉子維

目錄

一、摘要.....	3
二、背景動機及資料來源：.....	4
2.1 背景動機.....	4
2.2 資料來源.....	4
三、分析步驟流程圖：整個研究架構.....	5
四、使用量研究及視覺化分析.....	5
4.1 各站點使用量現況.....	5
4.1.1 平均每月使用量較低的站點.....	5
4.1.2 建議刪除站點.....	8
4.2 時間vs. 使用量.....	9
4.2.1 2016~2021 整體使用量.....	9
4.2.2 2019年每日總使用量(11月開始).....	10
4.2.3 工作日 vs. 非工作日 每日總使用量.....	11
4.2.4 12/18~12/27站點即時T-bike數.....	12
4.2.5 T-bike設置數量建議.....	14
4.3 天氣vs. 使用量.....	14
4.3.1 氣溫.....	14
4.3.2 相對溼度.....	14
4.3.3 風速.....	15
4.3.4 降雨量.....	15
4.3.5 相關係數.....	15
4.4 疫情vs. 使用量.....	16
4.5 使用量預測.....	18
4.5.1 無疫情、病例數資料.....	18

4.5.2 加入疫情、病例數資料.....	19
五、 結論.....	20
5.1 建議刪除站點.....	20
5.2 T-bike供應車數調整建議.....	20
5.3 每日使用量預測.....	20
六、附錄：可放圖表或方法學介紹.....	20
七、參考文獻.....	20
八、 心得.....	21
九、小組成員貢獻表.....	22

圖目錄

▲ (圖) 1 每月總使用量前5高站點.....	5
▲ (圖) 2 歸仁區T-bike站點分布.....	6
▲ (圖) 3 南科區T-bike站點分布.....	6
▲ (圖) 4 東區T-bike站點分布.....	7
▲ (圖) 5 永康區T-bike站點分布.....	7
▲ (圖) 6 安南區T-bike站點分布.....	8
▲ (圖) 7 台江文化中心附近公車站點.....	8
▲ (圖) 8 2016~2018每月總使用量.....	9
▲ (圖) 9 2019年十一、十二月每日總使用量.....	10
▲ (圖) 10 2019~2021工作日與非工作日每日總使用量盒形圖.....	11
▲ (圖) 11 學校站點可使用T-bike變化量.....	12
▲ (圖) 12 火車站站點可使用T-bike變化量.....	13
▲ (圖) 13 公園站點可使用T-bike變化量.....	13
▲ (圖) 14 氣溫(黃)與使用量(黑)折線圖.....	14
▲ (圖) 15 相對濕度(藍)與使用量(黑)折線圖.....	14
▲ (圖) 16 風速(綠)與使用量(黑)折線圖.....	15
▲ (圖) 17 降雨量(藍)與使用量(黑)折線圖.....	15
▲ (圖) 18天氣資料與使用量的相關係數熱圖.....	15
▲ (圖) 19 2020年(左)與2021年(右)天氣資料與使用量的相關係數熱圖.....	16
▲ (圖) 20 2020 - 2021年疫情資料.....	16
▲ (圖) 21 疫情指數(紅)與使用量(黑)相關折線圖.....	17
▲ (圖) 22 疫情資料與使用量相關係數熱圖.....	17
▲ (圖) 23 隨機森林模型(左)與 XGBoost模型(右)的變數重要性繪圖.....	18
▲ (圖) 24 隨機森林模型(左)與 XGBoost模型(右)的變數重要性繪圖.....	19

一、摘要

此報告針對台南T-bike的使用量進行研究。T-bike為台南市24小時可借可還之公共自行車租借系統，自2016年8月開始正式啟用，迄今總共設有78個站點，租借費率以30分鐘為一個計費單位，使用臺南市市民卡有特別優惠(5元/30分)，其他卡種為10元/30分。

此報告的分析時間範圍是從2019年11月到2021年11月。首先進行各站點的使用量分析，對於T-bike平均使用量較低之站點——歸仁區公所、南科健康生活館、和平公園、永康車站、台江文化中心、裕文圖書館這五個站點我們提出移除或遷移之建議。進而從以下角度對使用量進行分析：

1. 時間與使用量的關係

繪製每年每月T-bike總使用量折線圖與每年工作日vs. 非工作日T-bike使用量盒形圖進行比較，可發現T-bike在不同月份的使用量有較大的落差，以及非工作日的平均使用量高於工作日的平均使用量。

2. 天氣與使用量的關係

觀察氣溫、相對濕度、風速、降雨量與使用量的關係，並分別計算相關係數，發現降雨量與使用量較有相關性，風速與使用量較無相關性。

3. 疫情與使用量的關係

從2016年8月到2021年11月之每年每月T-bike總使用量折線圖可看出，在新冠疫情首次爆發(2020年2月)時使用量明顯有所下降，在三級警戒期間(2021年5月~7月)的使用量也是非常低，因此我們認為疫情也是會影響使用量之因素，並依照中央疫情指揮中心發佈之疫情警戒級數編碼成為類別變數，作為後續預測使用量之變數。

4. 使用量預測

最後我們將以上會影響使用量之因素作為預測變數，建立模型進行每站每日的總使用量預測。我們使用Random Forest與XGBoost模型分別就有疫情變數與無疫情變數資料做訓練，結果發現變數站點名稱在兩個模型中皆為重要變數，整體來說，XGBoost的表現較Random Forest要好，且在加入疫情變數後兩個模型的分數皆有所上升。

二、背景動機及資料來源：

2.1 背景動機

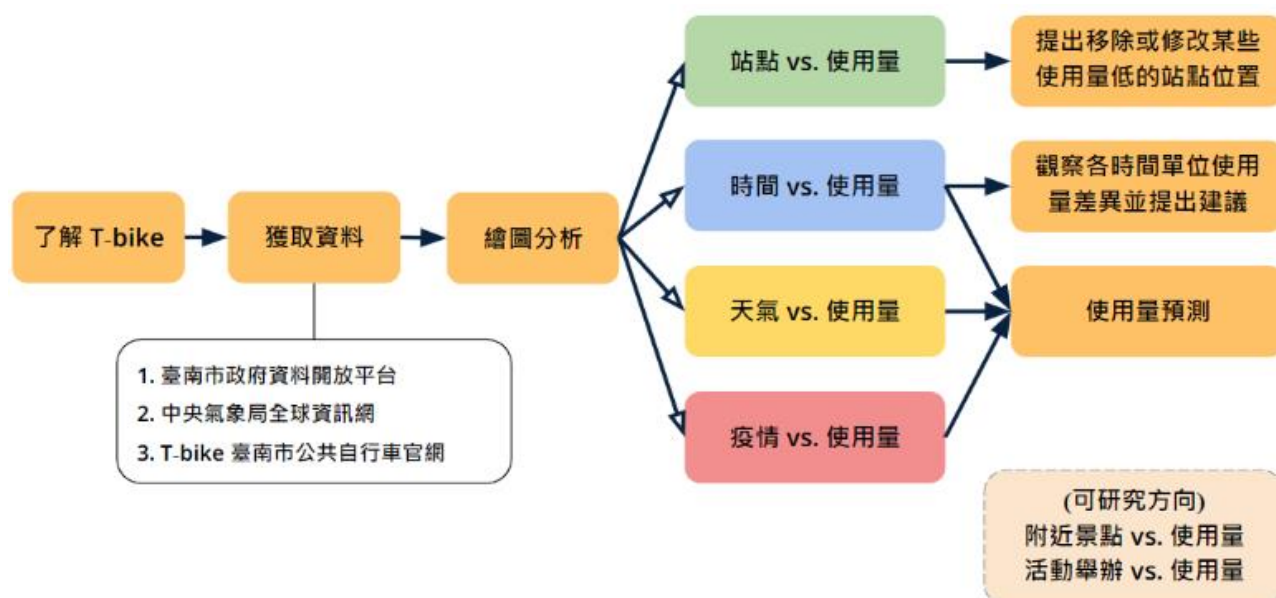
台南T-bike從2016年開始營運至今，與台中Youbike的900多站、高雄YouBike 1000多站相比，T-bike的使用率低，發展狀況極度不佳。因此，我們想從使用量的方向研究T-bike，並以「幫助T-bike提高收益」為目標，且針對以下題目進行討論。

- 增加使用量
 1. 廢除使用量低的站點、在需求量高的地區增加維修頻率
 2. 預測每日每區使用量
- 調整價格
 1. 在需求量高的時段提高費率

2.2 資料來源

- 臺南市政府資料開放平台
 1. T-bike每月使用次數
 - 蒐集日期：2019/11-2021/11
 - 變數：年/月、編號、場站、1-31 日
 2. T-bike租賃站資訊
 - 變數：ID編號、站名、地址、格位數可借車輛數、可停空位數、更新時間、緯度、經度
 3. T-bike累計使用量
 - 變數：ID編號、年、月、當月次數
- 中央氣象局全球資訊網
 1. 台南氣象資料
 - 蒐集日期：2019/11 - 2021/11
 - 變數：氣溫(°C)、相對濕度(%)、風速(m/s)、降水量(mm)
- T-bike臺南市公共自行車官網
 1. 及時租賃站資訊(爬蟲)
 - 蒐集日期：2021/12/18 - 2021/12/27(每日9:00~23:00)
 - 變數：站點區域、站點名稱、可借車輛數、可停車位數

三、分析步驟流程圖：整個研究架構

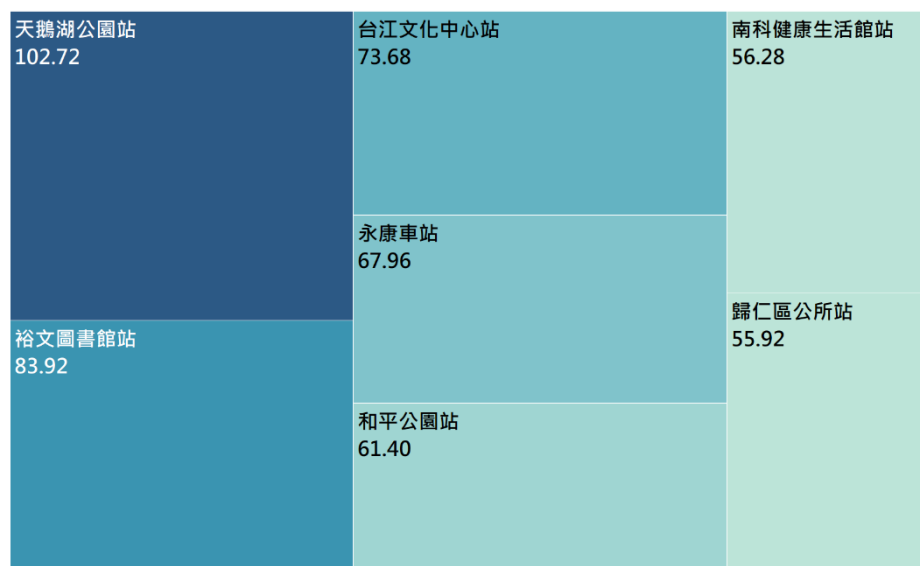


四、使用量研究及視覺化分析

4.1 各站點使用量現況

4.1.1 平均每月使用量較低的站點

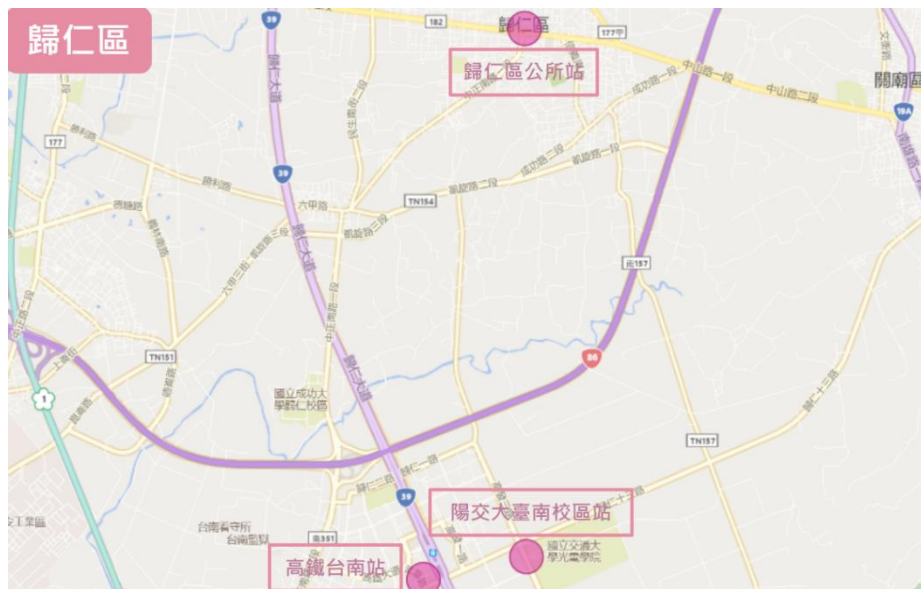
透過heatmap圖得知使用量低的前六名為歸仁區公所站、南科健康生活館站、和平公園站、永康車站、台江文化中心站、裕文圖書館站。除了使用量後續也考量了附近景點，全面分析該站點是否廢除。



▲ (圖) 1 每月總使用量前5高站點

a. 歸仁區公所站

該站點位居小鄉鎮，居民多騎機車行動；外加上與陽交大台南校區站、高鐵台南站距離遙遠(如下圖)，因此可考慮將此站點刪除。



▲ (圖) 2 歸仁區T-bike站點分布

b. 南科健康生活館站

根據地圖可觀察到，該站點非位於南部科學工業園區附近，且附近無著名景點觀光，因此也能考慮將此站點刪除。



▲ (圖) 3 南科區T-bike站點分布

c. 和平公園站

該站點附近皆為住宅大樓，且位於和平公園旁邊；居民僅需走路即能抵達該公園，使用t-bike的頻率極低。外加上該站點離火車站、學校、圖書館等景點皆有段距離，因此可考慮將此站點刪除。

d. 裕文圖書館站

雖然該站點位於復興國小旁，但國小生並不會騎t-bike上下學通勤，所以可考慮刪除此站點(目前已被刪除)。



▲ (圖) 4 東區T-bike站點分布

e. 永康車站

若規劃於永康區的車站設置T-bike，可考量將永康車站移至「大橋火車站」(如下圖)，大橋火車站附近有南台科技大學以及聖功女中，可使使用量提高，進而增加T-bike整體收益。



▲ (圖) 5 永康區T-bike站點分布

f. 台江文化中心站

該站點附近無景點且離臺史博物館距離遙遠，居民使用T-bike頻率低；外加上附近有大量公車站點的設置(如圖x)，因此可考量移至「台江國家公園」附近。國家公園有自行車步道、賞鳥步道等，較多人前往觀光能增加整體T-bike收益。



▲ (圖) 6 安南區T-bike站點分布



▲ (圖) 7 台江文化中心附近公車站點

4.1.2 建議刪除站點

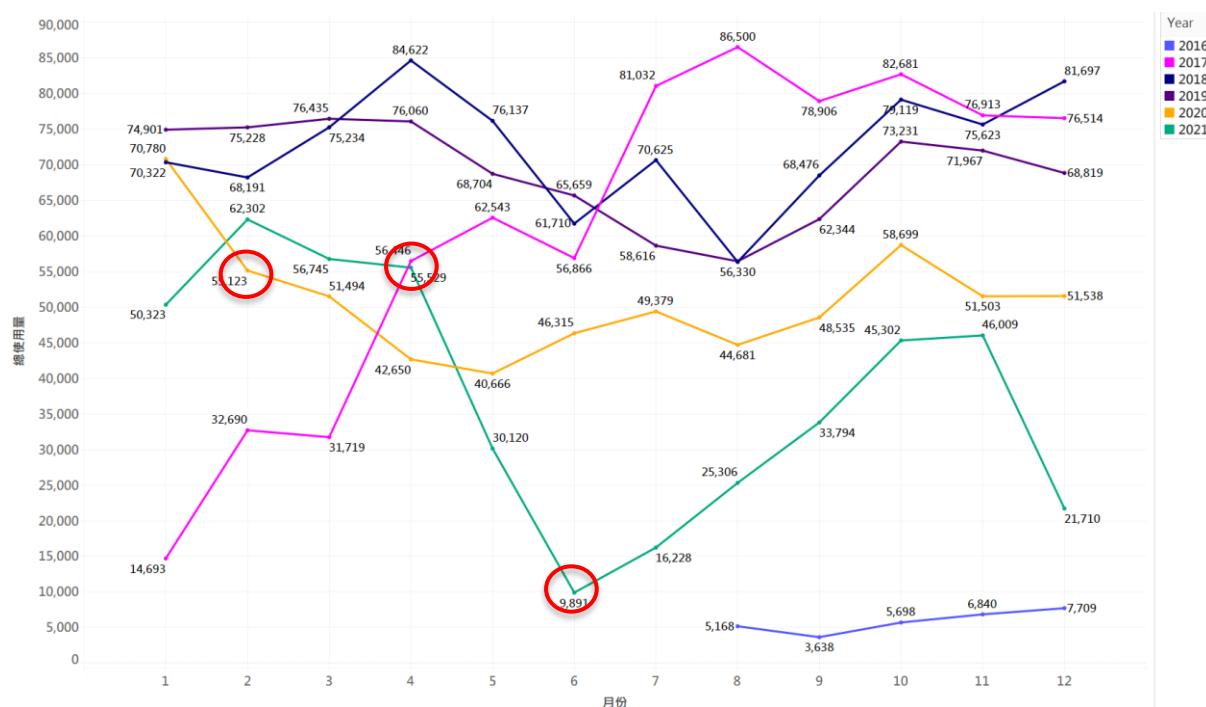
- 最低使用量—歸仁區公所站（歸仁區）→ 刪除
- 第二低使用量—南科健康生活館站（南科）→ 刪除
- 第三低使用量—和平公園站（東區）→ 刪除
- 第四低使用量—永康車站（東區）→ 移至大橋火車站
- 第五低使用量—台江文化中心站（安南區）→ 移至台江國家公園
- 第六低使用量—裕文圖書館（東區）→ 刪除(目前已刪除)

4.2 時間vs. 使用量

4.2.1 2016~2021 整體使用量

(圖)8的資料為2016年八月至2021年12月每個月T-bike總使用量的資料，其中2021年12月的總使用量資料並非完整的。

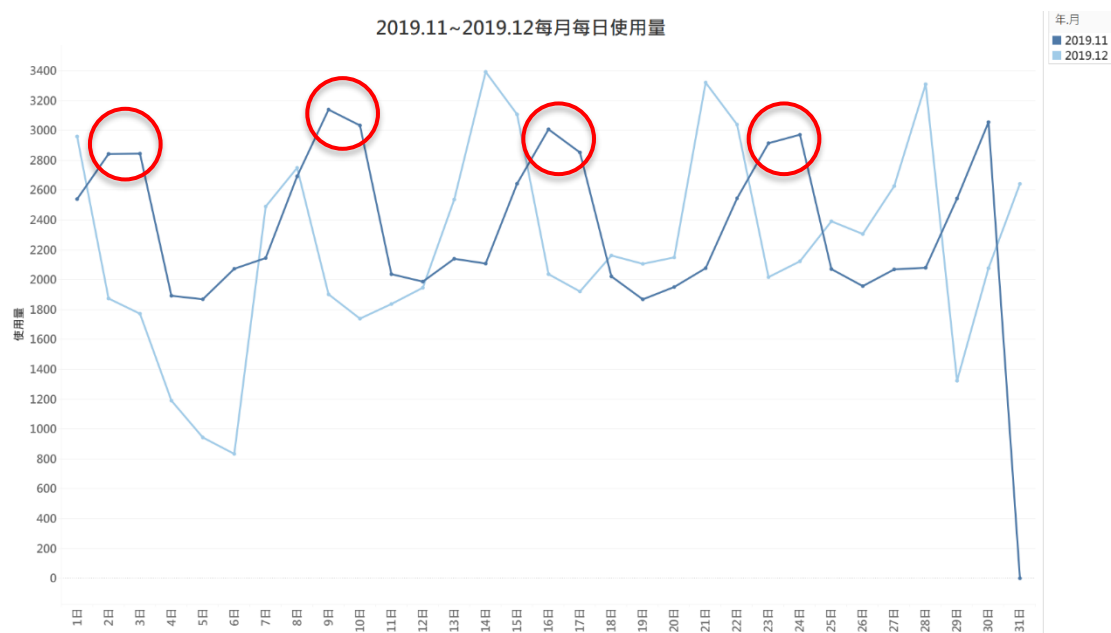
我們先大方向的觀察T-bike的使用趨勢：首先2016年八月開始設置T-bike，到了2018、2019年有較穩定的使用趨勢。並且可以觀察出在五月到九月夏天的季節使用量較少，尤其是七、八月，推測可能的原因有二：一是因為天氣熱所以使用量低、二是因為放暑假，所以學生的使用量降低許多。2020年的寒假開始疫情所以使用量有明顯的下滑，可以從圖上看出三個紅色圓圈的位置應為疫情爆發而下降的點，而目前整體的使用狀況而言為使用量持續回升的階段。



▲ (圖) 8 2016~2018每月總使用量

4.2.2 2019年每日總使用量(11月開始)

接下來，我們想要觀察單個月裡的使用量趨勢，因此，我們將已較具有代表性2019年十一、十二月拿出來觀察。可以發現使用量有周期性的趨勢，可以推測大概是在假日兩天的使用量會較高，並可以猜測大家在非工作日外出如出遊等形成對T-bike的使用量用高於工作日大家通勤的使用量。

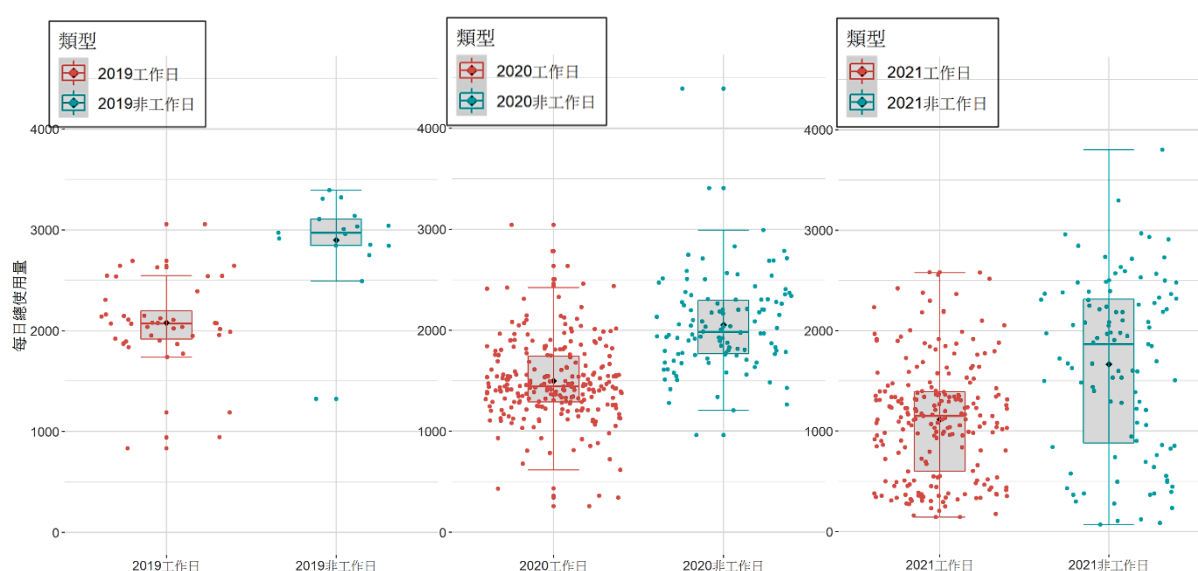


▲ (圖) 9 2019年十一、十二月每日總使用量

4.2.3 工作日 vs. 非工作日 每日總使用量

(圖)10的資料為2019年十一月至2021年十二月每日每站點的總使用量，由於該筆資料最早只到2019年十一月，因此只做2019~2021三年的觀察。為了更仔細的觀察並驗證工作日與非工作日使用量上的差異，我們將每年工作日與非工作日的使用量以盒形圖來表示並在(表)00列出每年各站點屬用量的平均數及中位數。

可以觀察出在2019~2022年非工作日的使用量皆明顯的高於工作日的使用量，而2021年因為疫情的關係整體的使用量都下降了。其中2021年非工作日的使用量更加分散，推測可能是因為非工作日大家外出的機會又更少了，所以許多站點的使用量再非工作日呈現驟降的趨勢。



▲ (圖) 10 2019~2021工作日與非工作日每日總使用量盒形圖

年份	2019		2020		2021	
	工作日	非工作日	工作日	非工作日	工作日	非工作日
mean	2079.45	2899.41	1497.9	2051.7	1116	1665.6
median	2072	2971	1445	1983	1151	1864

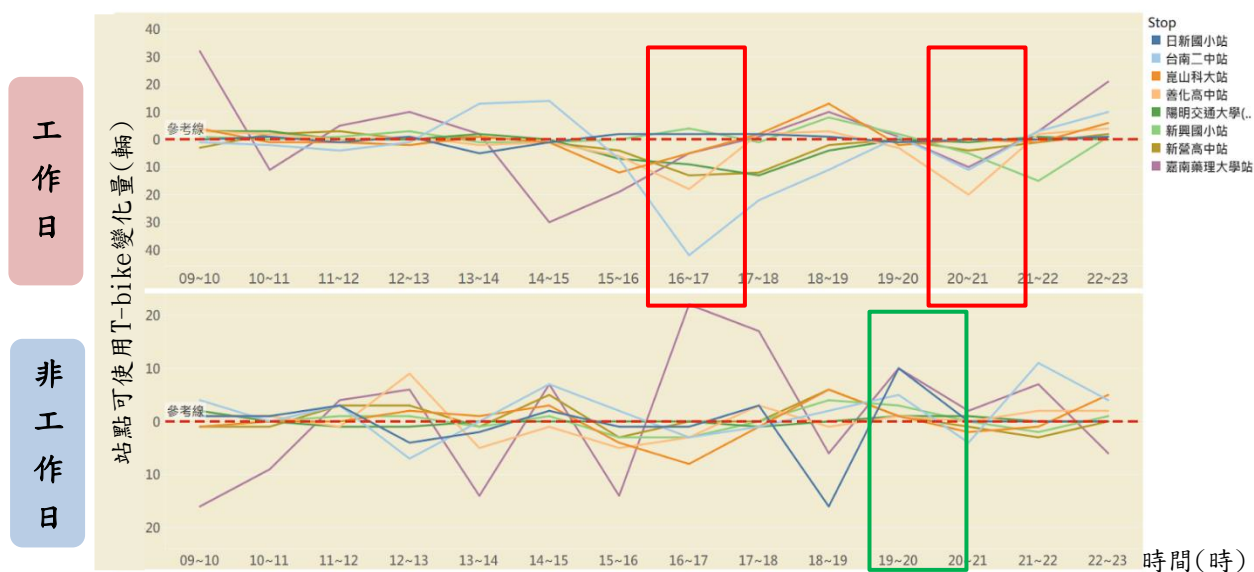
▲ (表) 1 2019~2021工作日與非工作日每日總使用量敘述統計

4.2.4 12/18~12/27站點即時T-bike數

為了可以得到更詳細的T-bike使用狀況，我們用爬蟲的方式至台南T-bike的即時租賃站資訊抓取，抓取的數值為各站點每個時間的可使用車數，抓取的時間為12/18~12/27的早上8點至晚上11點。我們以每個站點在每個時間區段可使用的T-bike車數變化量作為衡量的數據，在參考線以上(變化量 >0)代表該站點在該時段的可使用車輛增加，即歸還車輛的人多於借出的人；在參考線以下(變化量 <0)代表該站點在該時段的可使用車輛減少，即歸還車輛的人少於借出的人。

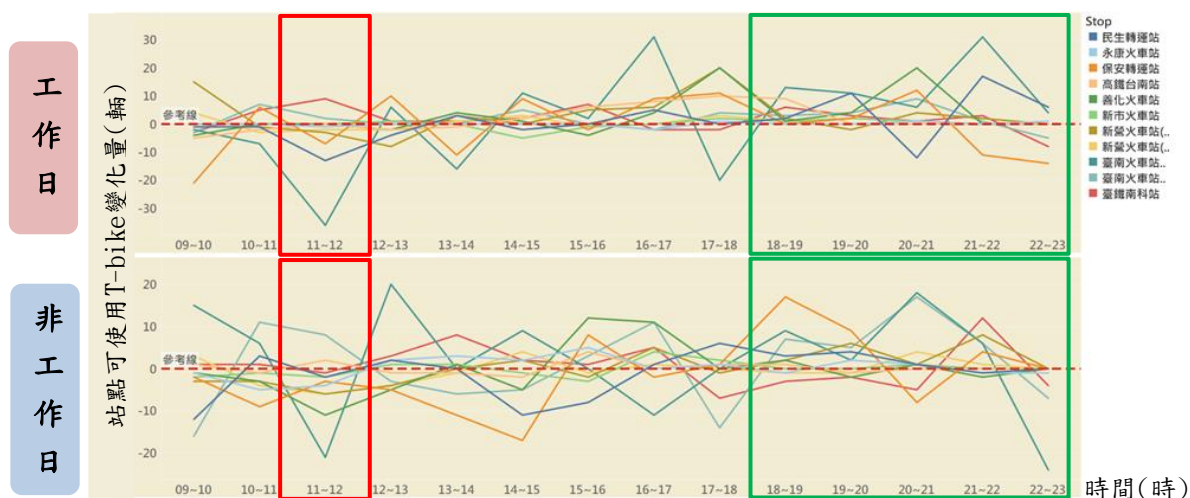
因為站點太多，而我們想要觀察相似地理條件下的T-bike站在使用量上是否也會有相似的趨勢，所以我們選出學校、火車站、公園附近的站點來觀察。

(圖)11為學校站點可使用T-bike變化量，紅色框為各站點可使用T-bike的變化量皆有下降的趨勢；綠色框則為各站點可使用T-bike的變化量皆有上升的趨勢。可以發現工作日在16點~17點之間有較大的借出量，因為學生放學會使用；在20點~21點又有一波使用量。而非工作日時，則是在19點~20點有較大的歸還量。



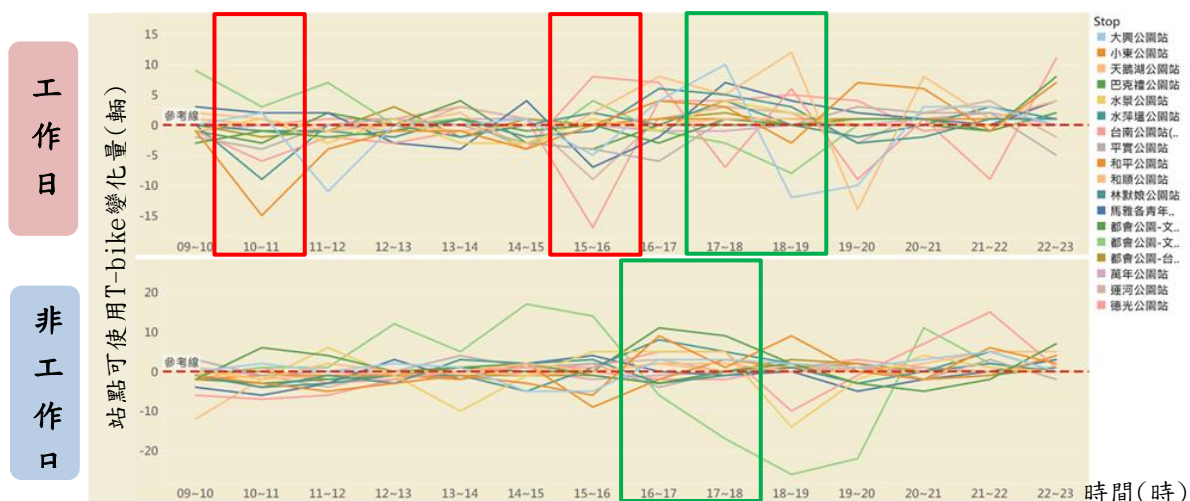
▲ (圖) 11 學校站點可使用T-bike變化量

(圖)12為火車站站點可使用T-bike變化量，火車站附近變化量起伏大，因為配合公車的時間T-bike的使用量變化較頻繁。不論是工作日與非工作日，所有的火車站站點普遍在11點~12點有一波較大的借出量；而在18點~23點之間，多數火車站站點也有大量的歸還量，可以推測那個時間點的火車站有較頻繁的發車車次，且搭乘的人數較多。



▲ (圖) 12 火車站站點可使用T-bike變化量

(圖)13為公園站點可使用T-bike變化量，公園站點在工作日10點~11點間以及15點~16點，有較大的借出量；在17點~19點之間又有較大的歸還量；而非工作日，在16點~18點有稍微比較多的歸還量，其他趨勢則不明顯。



▲ (圖) 13 公園站點可使用T-bike變化量

4.2.5 T-bike設置數量建議

學校附近於工作日16點~17點為學生放學時間，會有瞬間較高的使用量，可以建議T-bike公司在這個時間點前多補車並增加維護T-bike的頻率。而火車站站點使用較頻繁，建議可以在這些站點增加維修的頻率。

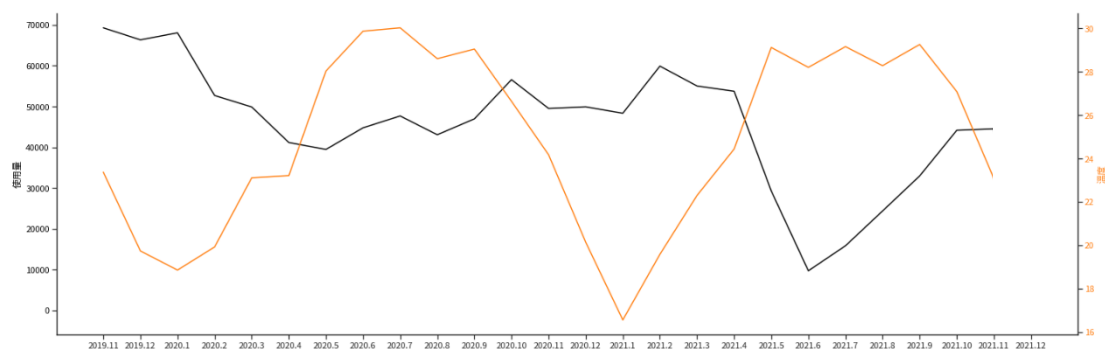
此外，綜合觀察以上地理條件相似的T-bike站，可以發現這些T-bike站在一天中每個時段的使用趨勢並沒有直接的關係，因此無法從各站點的借出及歸還趨勢判斷多數使用者的移動路線。

4.3 天氣vs. 使用量

在我們的大學生活中多少都有以腳踏車代步的經驗，通常最不想騎車上課的時刻不外乎就是下雨天了，因此我們認為天氣可能是影響T-bike使用量的因素之一。我們透過中央氣象局找到所有台南市的天氣觀測站點，並從每一個T-bike站點直線距離最近的觀測站點取得2019/11~2021/11氣溫、相對溼度、風速、降雨量等四個天氣資料，往後進行使用量相關分析。

4.3.1 氣溫

從下方折線圖可以看到在氣溫升高的5-7月，使用量大多有下降的趨勢，整體看起來是呈現負相關的。



▲ (圖) 14 氣溫(黃)與使用量(黑)折線圖

4.3.2 相對溼度

從下方折線圖可以看到在左半邊的2020年當中，4-9月之間的相對濕度與使用量有小小的負相關，而在右半邊的2021年裡兩者就有較明顯的負相關。



▲ (圖) 15 相對濕度(藍)與使用量(黑)折線圖

4.3.3 風速

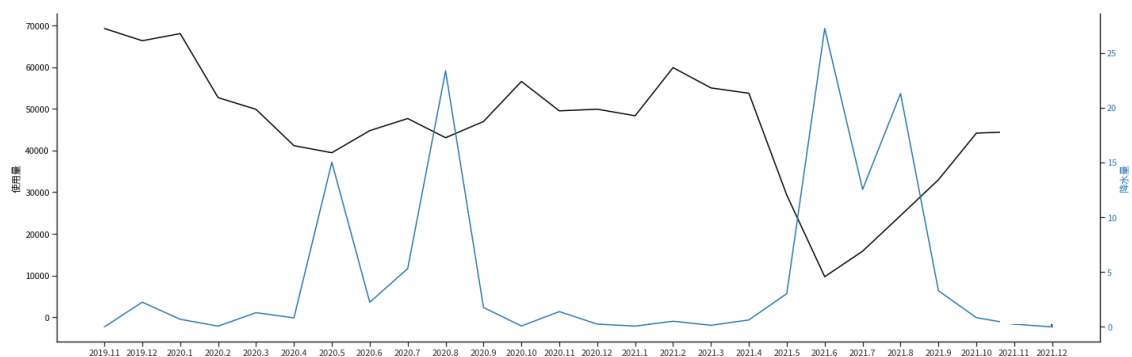
在左半邊的2020年當中，風速與使用向稍微呈現正向關係，但在右半邊的2021年裡，兩者看起來沒有什麼關係，初步認為風速與使用量是較無關的。



▲ (圖) 16 風速(綠)與使用量(黑)折線圖

4.3.4 降雨量

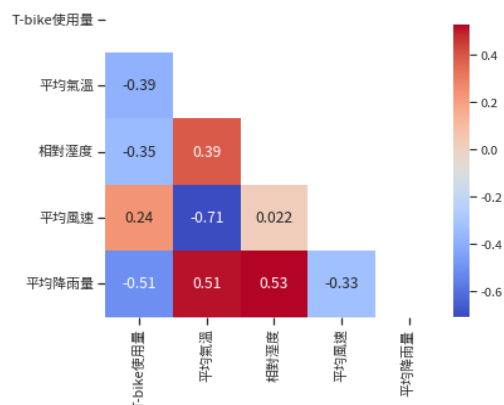
從下方折線圖可以看到在降雨量較多的5-8月，使用量皆是一年當中使用量最低點的時間附近，因此我們認為降雨量與使用量呈現負向關係，而且是有明顯影響的因素。



▲ (圖) 17 降雨量(藍)與使用量(黑)折線圖

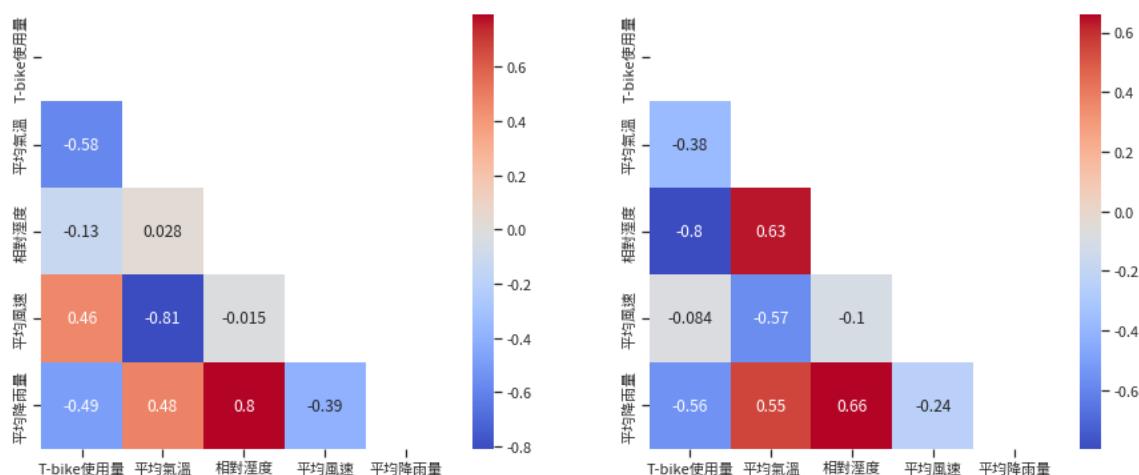
4.3.5 相關係數

下圖為四種天氣資料與使用量的相關係數，可以看到最左邊那欄呈現的相關係數皆與上方的初步討論相同，風速是四種天氣資料裡最不相關的，其他種類則與使用量呈現負相關。



▲ (圖) 18 天氣資料與使用量的相關係數熱圖

若是把2020年與2021年的資料拆開來看，也能驗證我們在上方的描述，較吸引我們思考的部分是在氣溫和降雨量上，這兩個與使用量的相關性不如我們想像中的高相關，可能是因為我們判斷的高相關僅出現在夏冬季的對比，春秋兩季的低相關影響到了整體的數值。

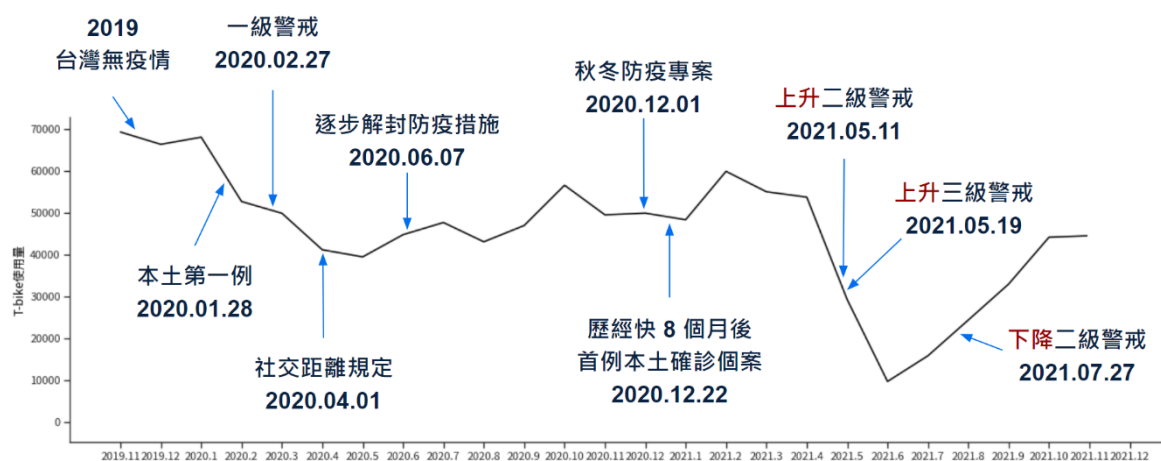


▲ (圖) 19 2020年(左)與2021年(右)天氣資料與使用量的相關係數熱圖

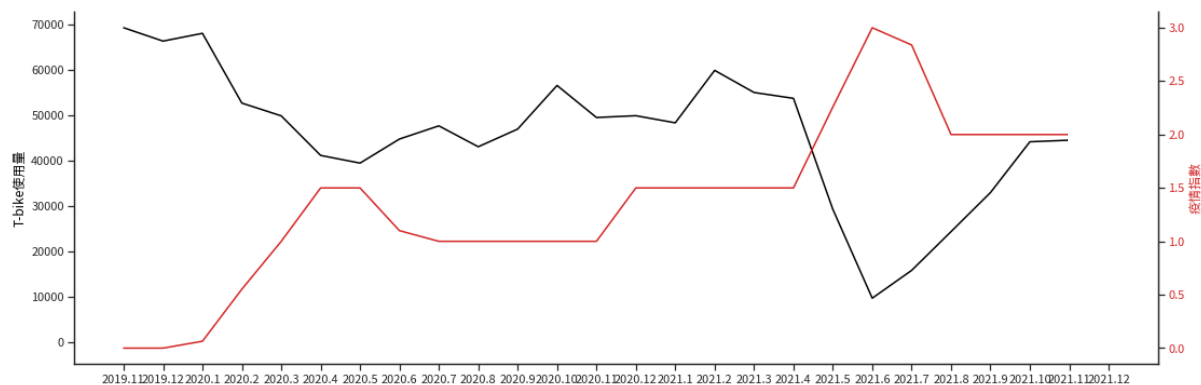
4.4 疫情vs. 使用量

在2020~2021年裡，衝擊我們生活最大的是新冠肺炎的出現，當疫情變的嚴重時，人們會盡量減少出門的頻率，出遠門、來台南旅遊的機會也大幅降低，因此我們也認為疫情會是影響使用量的因素之一。

接著，我們參考維基百科上對於台灣這兩年內有關疫情的詳細記錄(圖20)，包括個案確診、政府頒布命令、措施、警戒等的時間點，編制0~3的疫情指數，如圖21所示，疫情指數與使用量明顯呈現負相關。



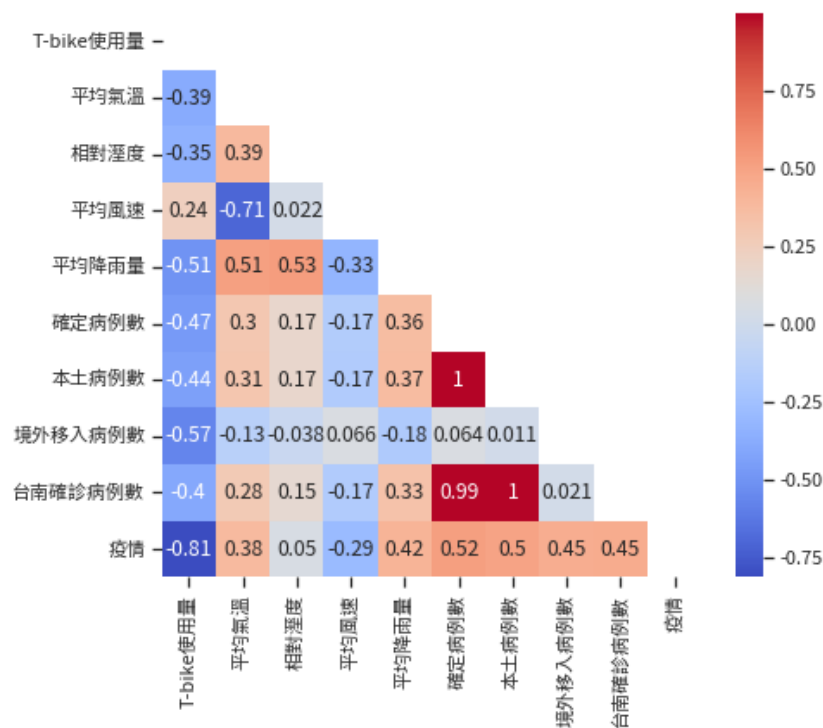
▲ (圖) 20 2020 - 2021年疫情資料



▲ (圖) 21 疫情指數(紅)與使用量(黑)相關折線圖

為了能找到更實際的疫情資料，我們也到政府資料公開平台尋找新冠肺炎的所有確診資料(含有日期、確診地、是否為境外等變數)，最終建構了近兩年的「確定病例數」、「本土病例數」、「境外移入病例數」、「台南確診病例數」四種資料。

將以上五個變數與天氣資料、T-bike使用量進行相關係數繪圖於圖22，可以看到除了風速，其他變數與使用量皆呈負相關，特別是疫情指數對使用量有高度負相關。另外，可以看到台南病例數、本土病例數兩者與確定病例數幾乎為相同資料，故在下一部份進行使用量預測時無須加入。



▲ (圖) 22 疫情資料與使用量相關係數熱圖

4.5 使用量預測

根據4.2至4.4從不同角度切入來研究T-bike使用量，我們希望能在最後使用機器學習模型進行每日的T-bike使用量預測，我們選定以下兩種機器學習模型來進行使用量預測：

1. Random Forests Regression
2. XGBoost Regressor

資料前處理將會移除上方討論過相關性過高的變數，最後，總資料筆數為53728筆，我們將資料以9:1切分訓練集與測試集，即訓練集42982筆、測試集10746筆，以便進行模型評分。

4.5.1 無疫情、病例數資料

在此小節當中，我們先將疫情資料排除，僅使用天氣資料來預測T-bike每日的使用量。使用訓練資料範例如下：

	年	月	站點名稱	平均氣溫	相對溼度	平均風速	平均降雨量	星期	是否為工作日	疫情
38695	2021	3	72	21.5	81.0	1.4	0.0	4	1	1.5
16398	2020	6	14	29.3	85.0	1.7	0.0	2	1	1.5
51991	2021	9	58	29.2	78.0	2.0	12.0	7	0	2.0

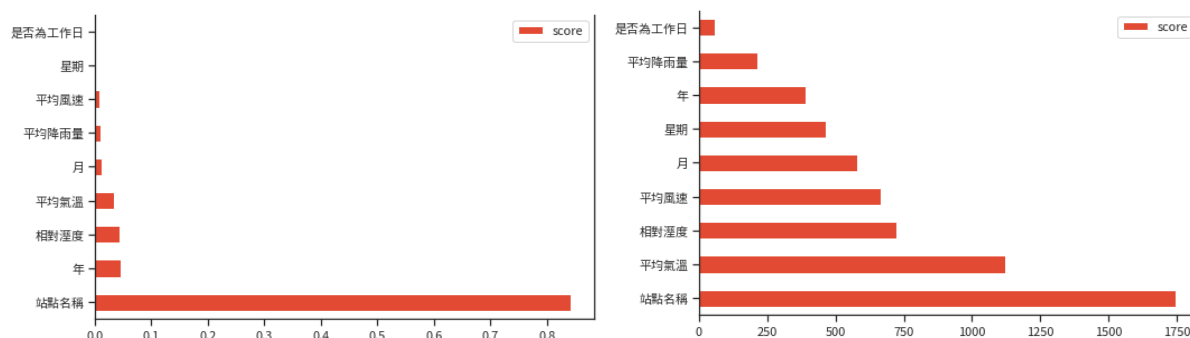
- 模型評估

我們採用RMSE、MAE以及R-square進行評估，從表中可以看到在預測每日使用量上XGBoost勝過隨機森林非常多，尤其兩者的R-square相差約1.6倍。

模型	RMSE	MAE	R ²
Random Forests Regression	22.3196	14.3225	0.495
XGBoost Regressor	13.1778	8.0856	0.8239

- 模型對變數的重要性估計

從下方兩張圖中可以看到，兩個模型都認為站點名稱是非常重要的，而是否為工作日則對使用量的預測沒有影響。另外，可以看出兩個模型看待時間變數（年、月）的重要性也有較大的差異。



▲ (圖) 23 隨機森林模型(左)與 XGBoost模型(右)的變數重要性繪圖

4.5.2 加入疫情、病例數資料

在此小節當中，我們加入疫情、病例數等資料來預測T-bike每日的使用量，並預期此模型在變數增加的情況下，各評分會有所成長、誤差降低。使用訓練資料範例如下：

	年	月	站點名稱	平均氣溫	相對溼度	平均風速	平均降雨量	星期	是否為工作日	疫情	境外移入病例數	確定病例數
35565	2021	3	72	21.5	81.0	1.4	0.0	4	1	1.5	1	1
13906	2020	6	14	29.3	85.0	1.7	0.0	2	1	1.5	0	0
48770	2021	9	58	29.2	78.0	2.0	12.0	7	0	2.0	4	6

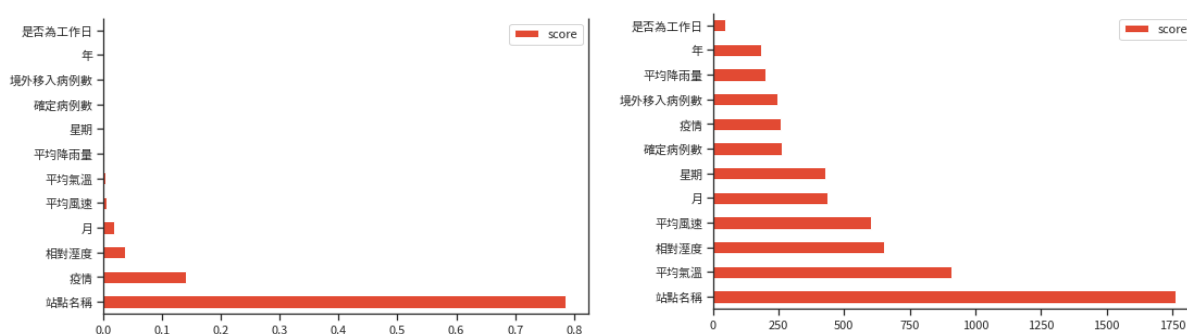
- 模型評估

這次從表中可以看到仍然是XGBoost比隨機森林在預測上準確許多，但在增加變數後兩個模型在各分數上皆有成長，XGBoost的R-square也提升到83%。

模型	RMSE	MAE	R ²
Random Forests Regression	21.7908	14.2258	0.5186
XGBoost Regressor	12.9468	7.8834	0.8301

- 模型對變數的重要性估計

在加入新的變數後，站點名稱仍然是最重要的變數，而疫情指數在隨機森林的模型中是第二重要的變數，但對於XGBoost卻相對不是那麼重要，反而天氣資料在XGBoost模型中仍保持著較高的分數。



▲ (圖) 24 隨機森林模型(左) 與 XGBoost模型(右)的變數重要性繪圖

五、結論

5.1 建議刪除站點

- a. 最低使用量-歸仁區公所站（歸仁區）→ 刪除
- b. 第二低使用量-南科健康生活館站（南科）→ 刪除
- c. 第三低使用量-和平公園站（東區）→ 刪除
- d. 第四低使用量-永康車站（東區）→ 移至大橋火車站
- e. 第五低使用量-台江文化中心站（安南區）→ 移至台江國家公園
- f. 第六低使用量-裕文圖書館（東區）→ 刪除(目前已刪除)

5.2 T-bike供應車數調整建議

學校附近於工作日16點~17點為學生放學時間，會有瞬間較高的使用量，可以建議T-bike公司在這個時間點前多補車並增加維護T-bike的頻率。而火車站站點使用較頻繁，建議可以在這些站點增加維修的頻率。

此外，綜合觀察以上地理條件相似的T-bike站，可以發現這些T-bike站在一天中每個時段的使用趨勢並沒有直接的關係，因此無法從各站點的借出及歸還趨勢判斷多數使用者的移動路線。

5.3 每日使用量預測

我們目前取得的使用量資料僅有兩年的時間，因此無法以每日以外的角度進行資料預測。從4.5節中，我們對於每日使用量預測的結果並不是很差，也看到增加變數能讓模型更進步。

此報告粗淺地從站點、時間、天氣、疫情的角度分析T-bike的使用量，我們認為若要让模型更加進步還能夠討論到台南市活動與市集的舉辦、站點附近的景點與設施、促進觀光之政策、人口居住密度、人口年齡組成、交通工具的使用習慣等等，當討論的範圍能包含到更多面向時，模型預測也會更加穩固、更加有意義。

六、附錄：可放圖表或方法學介紹

無

七、參考文獻

1. T-bike台南公共自行車網站
2. 交通部中央氣象局觀測資料查詢系統
3. 嚴重特殊傳染性肺炎臺灣疫情時間線（2020年） - 維基百科，自由的百科全書
4. 嚴重特殊傳染性肺炎臺灣疫情時間線（2021年） - 維基百科，自由的百科全書

八、心得

• 陳柔漪

大二的時候在個體經濟學上研究過台南市公車的經濟價值，那時候也是從台南市政府資料開放平台取得資料集，但從那之後就覺得單純只有描述載運量、收益的政府資料好像沒辦法做什麼可以發展的報告。剛開始選擇這份題目時，我還是以差不多的角度，而且不曉得T-bike能做出什麼有內容的報告。但是在老師提供很多想法與指導後，讓我發現我的思想過於限制在「使用量的數值」上，在結合了其他領域像是天氣、地域、人口等的資料內容，報告就能夠更加完善與有價值。和組員們一起發想能夠從哪些角度出發描述T-bike使用量、想達成什麼樣的目的等等，過程中我們做過很多嘗試，PPT最大情況有到60頁的圖片吧，但許多結果不如我們所假設的情況而選擇刪去，這段時間很辛苦，討論的時候甚至會覺得好像沒有進度，有幸我們想了許多可結合的資料，在大家的努力下，最後我們也順利完成了報告！

統計諮詢幫助我複習很多以前學的知識，也鍛鍊了很多跟報告相關的口條、資料整理與PPT製作等技能，謝謝老師也總是給予正面的回應，尤其在期末報告上自由發揮的題目，雖然一開始傷透了腦筋，但能看到所有人各自完成的多面向題目非常有趣。製作報告的過程中，從發現問題到解決問題，我得到許多不同層面的思考，最後想謝謝大家在我Loading很重的時候都願意等我、幫助我😊，給大家愛心❤️。

• 劉米婷

這次的統諮報告我們換過蠻多次主題，最後決定做台南T-bike這個題目，但剛開始拿到資料時我們卻發現如果單純從資料去分析的話可以做的事情太少了，尤其在定義問題的部分總覺得我們對於解決問題的論點有點過於單薄了，因為我們只能單純從使用量進行分析而已。後來跟老師討論後才發現我們太過於以使用者的角度切入這個分析的議題了，只考慮到應該怎麼讓使用者有更多更穩定的T-bike，但卻沒有考慮到對企業或對政府而言，他們考慮的是收益。經過老師的提點後我們又重新思考了我們想要分析的方向，雖然最後礙於沒辦法取得T-bike公司的收益，以及每個使用者的花費，所以只能回到我們原本有的資料進行分析，但這讓我知道分析一個議題要思考為什麼要分析這件事、分析後我們能做出什麼貢獻，甚至是製作出的圖表或是檢定等結果是想要觀察什麼。總結而言，從這次的統諮報告我覺得學到最多的是當今天拿到一筆資料我應該從資料去發現問題、定義問題，並且簡潔有力的用圖表或檢定等方式表達自己的論述。

另外，在期中的個人報告真的讓我覺得其實統計也沒有那麼難，尤其是自己準備一個主題報告真的非常有成就感，準備的過程中雖然花了很多時間，但這之後我學會了該用自己的話解釋一個統計方法給大家聽，這整個過程都讓我覺得很有挑戰性但也很有趣。

最後，我真的很謝謝我的組員們！大家真的都有努力的在這次報告上貢獻，也很願意互相幫忙，有這麼棒的組員我真的覺得太幸運了！！

- **王思蘋**

起初我們是做跳繩運動，很有趣的題目但可惜後續無法完成，原本還想說順便減肥一下XD 在這學期統計諮詢課中，感覺到老師都用好奇的角度鼓勵同學嘗試去做，題目也沒有侷限，是我目前修過滿意度前三的課程。

至於這份報告，雖然僅是一小部分的T-bike分析，但最有印象的是，起初找老師討論時才發現我們把T-bike想得太侷限，只將數據所呈現的表面分析，並沒有聯想到從政府、企業的角度來分析使用量，能夠影響到後續的收益，果然我們還是資料分析新生兒哈哈。另外也在這次報告中回憶了修統計繪圖使用的Tableau，跟之前比起來上手很多；但也學到教訓，之後會稍加思考要探討的問題後再著手畫圖，不然會浪費一些時間。

最後這次團體報告最開心是遇到很棒的組員，過程很順利還約了幾次小聚餐。也透過討論得到很多想法、能深入分析的問題等等，比起之前的個人報告，團體合作把報告做到更有邏輯、完整性。

- **劉子維**

這次的報告在選題目的時候就一波三折，原本想要做的內容因為儀器問題沒有辦法繼續進行下去，後來千挑萬選就選擇了T-bike的使用量研究，雖然說現在很少騎T-bike了，但是整個報告做下來還是感覺蠻貼近生活的，也可以看到使用量會因為什麼而變化、怎麼變化的。

一開始是看到有關紐約city bike使用量的分析覺得很有趣—原來公共自行車的使用量有那麼多的面向可以研究分析（使用者年齡、騎行速度、天氣情況等等），所以想說可以來看看台南的T-bike是不是也可以往相似的方向來分析。很可惜的是我們能獲得的資訊蠻有限的，只能單純看到使用量的變化以及結合天氣來分析，不過透過各種視覺化的呈現可以看出使用量變化的規律、找出原因，甚至通過模型可以來預測未來的使用量，這些都還蠻有趣的。作報告的時候參與最多的就是視覺化的圖形繪製，過程中其實並沒有想象中的容易，如何呈現、呈現出來有何意義、可以透露什麼訊息這些都是值得思考的。很感謝組員的互相幫忙、交流想法，以及老師不時提供的建議，讓我們最後能夠呈現這一份完整的報告。

九、小組成員貢獻表

成員	完成內容
陳柔漪	使用量預測模型建立，圖表製作，報告撰寫
劉米婷	資料抓取與清理，圖表分析，報告撰寫
王思蘋	圖表製作，圖表分析，報告撰寫
劉子維	圖表製作，圖表分析，報告撰寫