

# 逢甲大學學生報告 ePaper

## 個人的理想城市

### My Ideal City

作者：鄭雅云、曾怡華、林子傑、張佳臻、陳威樑、黃姿培、吳東偉

系級：統計學系四甲

學號：D0308071、D0308098、D0333892、D0333730、D0333917、D0335290、D0335021

開課老師：魏裕中 老師

課程名稱：統計專題(一)

開課系所：統計學系

開課學年：106 學年度第一學期

## 摘要

每個人對於衡量理想的居住地區準則不同，但現有的調查僅考慮各地區的主客觀因素提供一個綜合性指標，而並未考量到個人化的差異。本研究欲建構一套個人化決策系統，可供使用者選擇自身關心的衡量項目，經由政府開放平台的客觀數據，並以縝密的統計分析方法，最終以視覺化地圖的呈現個人的理想城市排名。

本研究利用 R 軟體的 Shiny 套件建構出互動式頁面，此程式須包含前端的互動頁面與後端的統計分析兩大部分，首先，使用者於前端勾選欲關注的細項變數後，後端將以收集的資料以勾選的變數進行主成分分析，並計算出各縣市的理想總分與排名，再依設定好的理想區段將縣市分為三等級，最終將各等級以深淺不同的顏色代表，並以 Leaflet 套件將顏色繪製於臺灣地圖上，並傳送至前端的頁面中顯示。

我們藉由本次專題研究，發揮於統計學系學習的知識，建構了一個可供個人需求來進行統計分析的系統，利用簡單的互動式操作，即可讓使用者清楚的解讀複雜的統計分析的結果。

**關鍵字：**資料視覺化、個人化決策系統、主成分分析、R 程式 Leaflet 套件、R 程式 Shiny 套件

## Abstract

The criteria to identify ideal countries are different for each person. However, the reports in general most liveable survey did not evaluate the diversity. In this research, we construct a personalized decision system to find my ideal city.

Shiny package in R program is used to build the interactive web page. Two parts are essential, including frontend visualization and backend data analysis. Based on the items that user selected, data retrieve from the government open data is utilized to estimate the scores for each city via principal components analysis. The ranks for all cities transformed from scores are shown with colors. The Leaflet package is adopted to display the analytic results on the Taiwan map.

In this project, all techniques learned from college are used such as programming and statistical theory. We provide a interactive personalized system to make the statistical results easy to understand.

**Keyword :** Data visualization, personalized decision system, principal components analysis, R Leaflet package, R Shiny package

# 目 次

第一章 緒論.....	1
1.1 研究動機與背景.....	1
1.2 研究目的.....	2
第二章 資料分析.....	3
2.1 資料收集.....	3
2.2 資料彙整.....	3
第三章 研究方法.....	7
3.1 研究流程.....	7
3.2 統計分析.....	8
3.2.1 選擇統計分析方法.....	8
3.2.2 主成分分析.....	8
3.3 視覺化應用.....	11
3.3.1 互動資料視覺化.....	11
3.3.2 空間資料視覺化.....	12
第四章 研究結果.....	14
第五章 結論與討論.....	18
5.1 結論.....	18
5.2 討論.....	19
參考文獻.....	20

## 圖目錄

圖 1 研究流程圖.....	7
圖 2 理想程度分級顏色對應圖.....	12
圖 3 初始頁面.....	13
圖 4 勾選後的頁面.....	14
圖 5 主成分分析數據.....	14
圖 6 陡坡圖.....	15
圖 7 縣市指標分數.....	15
圖 8 縣市排名與分級指標.....	16
圖 9 顯示結果.....	16
圖 10 調整前的版面.....	18

## 表目錄

表 1 原始 63 個細項變數.....	4
表 2 五大面向與刪除細項變數明細.....	5
表 3 細項變數整合成勾選選項一覽.....	6
表 4 勾選選項表.....	6

# 第一章 緒論

## 1.1 研究動機與背景

所得較高就幸福嗎？還是社會福利好才算幸福呢？經濟合作與發展組織(OECD)綜合了客觀的觀察數據與主觀的調查訪談的所有資訊，訂定各國家城市的美好生活指數<sup>[1]</sup>；天下雜誌每年也針對臺灣各個縣市做了幸福城市大調查，共調查了「經濟力」、「環境力」、「施政力」、「文教力」、「社福力」等五大面向，每個面向又包含了十二至十五個細項指標<sup>[2]</sup>，藉由綜合所有面向的指標，以訂定各縣市的幸福程度排名。

一般的幸福指標是從多面向的綜合性評比數據來衡量各縣市的幸福程度，但缺少考慮到每個人對於各項指標的關注差異，舉例來說，有適學年齡孩子的家庭也許會較注重文教方面的表現和治安、環境汙染等因素，即將畢業欲找工作的學生可能會較在意所得支出等經濟上的情況，由於一般的美好生活幸福指標調查並無法反映個人需求上的差異，因此我們欲建構一個依照個人化需求而找到自己心中覺得幸福理想城市的方法。

此外，由於有些幸福指標的計算方式是將每個縣市的各個指標排名後，再加總計算總分予以排序，因此可能會造成一些城市因某項指標分數特別突出，即使在多數指標上略低於其他縣市，最後的總分因那些特別突出的指標而使總分與排名高於其他城市；相反的，有些城市在多數的指標都略高於其他縣市，但可能因為某項指標排名特別低，導致在總分排名上遠遠落後其他城市。為了避免這種「一竿子打翻一船人、一好遮三醜」的現象產生，我們試圖由各指標的原始數據，藉由統計方法設定一套能夠客觀去訂定城市之間排名的方法。

因此本專題研究欲藉由客觀的政府調查數據，並允許主觀的挑選個人關注的面向，以直覺方便的視覺化工具結合適當的統計方法，讓大家能夠更快速有效的找出自己的理想生活城市。

## 1.2 研究目的

由於市面上各機構調查的幸福指標，皆以綜觀性的衡量各城市的美好生活狀態，但每個人認定美好生活在乎的面向或多或少有所不同，所以我們欲由政府豐富多元化的開放資料，允許使用者自行挑選想關注的面向細項，再經由縝密地統計分析，最終提供每個人自己的理想城市排名。

另一方面，為了使一般大眾也能輕鬆易讀地選擇關心的細項並檢視複雜的統計分析結果，因此我們運用了所學的統計軟體程式，並搭配新興的視覺化套件，讓對於不大熟稔統計學的使用者也能藉由我們提供的互動式個人化決策系統，獲得自己心目中的理想城市資訊。

## 第二章 資料分析

### 2.1 資料收集

為了完整的考量各面向的客觀資訊，我們最初由政府資料開放平台(open data)<sup>[3]</sup>，透過網頁中涵蓋分類的十八種資料集，再連結至各政府相關單位網站，以尋找各種面向的細項變數資料。

本研究以臺灣的各縣市做為一個地區單位，因此原先欲納入臺灣所有共 22 個直轄市與縣市納入評比，但由於離島的連江縣、金門縣的資訊較不完整且資料品質有許缺陷，因此我們最終選擇台北市、台南市、新北市、桃園市、高雄市、台中市、新竹市、新竹縣、彰化縣、基隆市、宜蘭縣、花蓮縣、嘉義市、苗栗縣、南投縣、屏東縣、嘉義縣、台東縣、雲林縣、澎湖縣等 20 個直轄市與縣市作為研究對象。

此外，各細項變數資料的年月份單位與年份起迄日並不完全一致，為了使資料具有時間的一致性，因此更加小心的挑選資料變數；以降雨量與氣溫為例，中央氣象局提供了每日的降雨量但僅提供每月的平均氣溫資訊，且兩者的起始日期分別為 2008 年與 1981 年統計時間並不一致，因此無法全數整合納入討論分析；另過於久遠的資料可能與現在的情況差距過大，因此我們最終選擇了最近一年 2016 年以年為單位的資料，使各變數的資訊能最貼近現今的環境。

最終共收集到 63 個具有時間與地點一致性的細項變數，並參考天下雜誌/631 期 2017 幸福城市評分的調查，將 63 個細項變數簡易分為「經濟力」、「環境力」、「施政力」、「文教力」與「社福力」五大面向。各細項的原始變數名稱列於表 1。

### 2.2 資料彙整

原始的 63 個細項變數中，有些因為資料缺漏(如：火災傷亡率)，或是與其他變數相似性太高，因此共計刪除了 7 項變數，僅剩 56 個分析變數(表 2)。

再者，為了於視覺化頁面呈現時，使用者能清晰了解各選項，因此將相似的變數歸為同一個選項予以命名(表 3)，最終共計整合成 26 個勾選選項(表 4)。

表 1、原始 63 個細項變數

15 歲以上民間高等教育比	平均每教保人員負擔幼生數	赤字占歲入比率	104 學年度輟學率
每人出席藝文展演活動次數	兒童及少年安置及教養機構	債務占歲入比率	營利事業營業家數
營利事業銷售額年增率	55 歲以上每百人參與長青學苑人次	每萬機動車肇事數	人均營利事業銷售額
經濟發展支出占歲出比	平均每人每日垃圾清量	酒駕事件死亡數	15 歲以上識字率
每戶可支配所得中位數	接受婦女福利服務占女性人口比率	刑案破獲率	低底盤公車數
平均每人圖書資料費支出	平均每人每日垃圾產量	火災傷亡率	縣市各層級公圖書館借閱量
2015 地方文化支出執行率	每人每日生活用水量	歲入自籌財源比率	每人消費支出
平均每人可分配藏書冊數	身障福利機構安置及日托占身心障礙比率	自籌財源年增率	人口增加率
每年舉辦藝文展演活動次數	每縣市民享有之社福支出	刑案發生率	公立公共圖書館個數
大專以上占就業人口比	志工人數占 15 歲以上人口比率	下水道建設完工率	家戶聯網率
外國專業人員年增率	平均每教師教導學生數-國小	觀光旅館住用率	勞動參與率
文創產業占營利事業家數比	高風險家庭個案平均面訪次數	污水處理率	人均文創產業營收
每十萬人口自殺死亡率	無障礙生活環境業務督導考核總成績	公共運輸使用	失業率
平均每教師教導學生數-高中	平均每鄉鎮市區發展協會與附屬社團數	PSI>100 日數比率	每人每年用電量
教育科學文化支出占總歲出比	安養機構每工作人員服務老人數	垃圾回收率	每人每年油耗量
平均每教師教導學生數-國中	平均每列冊關懷獨居老人被服務次數	平均每人環保經費	

表 2、五大面向與刪除細項變數明細

經濟力	環境力	施政力	文教力	社服力
人均營利事業銷售額	PSI>100 日數比率	赤字占歲入比率	104 學年度輟學率	<u>每十萬人口自殺死亡率</u>
營利事業營業家數	污水處理率	債務占歲入比率	每年舉辦藝文展演活動次數	志工人數占 15 歲以上人口比率
<u>營利事業銷售額年增率</u>	平均每人環保經費	每萬機動車肇事數	每人出席藝文展演活動次數	低底盤公車數
經濟發展支出占歲出比	<u>平均每人每日垃圾清量</u>	酒駕事件死亡數	15 歲以上識字率	兒童及少年安置及教養機構
每戶可支配所得中位數	垃圾回收率	刑案破獲率	15 歲以上民間高等教育比	高風險家庭個案平均面訪次數
每人消費支出	<u>平均每人每日垃圾產量</u>	<u>火災傷亡率</u>	縣市各層級公圖書館借閱量	無障礙生活環境業務督導考核總成績
<u>失業率</u>	每人每日生活用水量	歲入自籌財源比率	平均每人可分配藏書冊數	身障福利機構安置及日托占身心障礙比率
家戶聯網率	每人每年用電量	自籌財源年增率	平均每人圖書資料費支出	接受婦女福利服務占女性人口比率
<u>人口增加率</u>	每人每年油耗量	刑案發生率	公立公共圖書館個數	55 歲以上每百人參與長青學苑人次
大專以上占就業人口比	公共運輸使用率	下水道建設完工率	2015 地方文化支出執行率	安養機構每工作人員服務老人數
外國專業人員年增率		觀光旅館住用率	教育科學文化支出占總歲出比	平均每列冊關懷獨居老人被服務次數
勞動參與率			人均文創產業營收	平均每鄉鎮市區發展協會與附屬社團數
			文創產業占營利事業家數比	每縣市民享有之社福支出
			平均每教保人員負擔幼生數	
			平均每教師教導學生數-國小	
			平均每教師教導學生數-國中	
			平均每教師教導學生數-高中	

紅色底線標示者為刪除的變數

表 3、細項變數整合成勾選選項一覽

經濟力	環境力	施政力	文教力	社服力
人均營利事業銷售額	PSI>100 日數比率	赤字占歲入比率	104 學年度輟學率 1 0	志工人數占 15 歲以上人口比率
營利事業營業家數 1	污水處理率	債務占歲入比率 6	每年舉辦藝文展演活動次數	低底盤公車數
經濟發展支出占歲出比	平均每人環保經費	每萬機動車肇事數	每人出席藝文展演活動次數 1 1	兒童及少年安置及教養機構
每戶可支配所得中位數	垃圾回收率 4	酒駕事件死亡數 7	15 歲以上識字率	高風險家庭個案平均面訪次數 1 6
每人消費支出 2	每人每日生活用水量	刑案破獲率 8	15 歲以上民間高等教育比	無障礙生活環境業務督導考核總成績 1 7
家戶聯網率	每人每年用電量	歲入自籌財源比率	縣市各層級公圖書館借閱量	身障福利機構安置及日托占身心障礙比率
大專以上占就業人口比	每人每年油耗量 5	自籌財源年增率 9	平均每人可分配藏書冊數	接受婦女福利服務占女性人口比率 1 8
外國專業人員年增率	公共運輸使用率	刑案發生率	平均每人圖書資料費支出	55 歲以上每百人參與長青學苑人次
勞動參與率 3		下水道建設完工率	公立公共圖書館個數 1 2	安養機構每工作人員服務老人數 1 9
		觀光旅館住用率	2015 地方文化支出執行率 1 3	平均每列冊關懷獨居老人被服務次數
			教育科學文化支出占總歲出比	平均每鄉鎮市區發展協會與附屬社團數
			人均文創產業營收 1 4	每縣市民享有之社福支出 2 0
			文創產業占營利事業家數比	
			平均每教保人員負擔幼生數	
			平均每教師教導學生數-國小	
			平均每教師教導學生數-國中	
			平均每教師教導學生數-高中 1 5	

1. 企業經營力、2. 家戶收支、3. 勞動力、4. 垃圾汙水處理、5. 天然能源耗量、6. 赤字債務、7. 交通事故、8. 刑案事故、9. 自籌財源、10. 教育程度、11. 藝文展演、12. 圖書資源、13. 文化經費、14. 文創產業、15. 師生比例、16. 兒童福利服務、17. 身心障礙福利服務、18. 婦女福利服務、19. 老人福利服務、20. 社福資源

表 4、勾選選項表

經濟力	環境力	施政力	文教力	社服力
家戶收支	環境汙染情形	赤字債務	教育程	志工服務
家戶聯網率	垃圾汙水處理	交通事故	師生比例	低底盤公車
企業經營力	天然能源耗量	刑案事故	圖書資源	兒童福利服務
勞動力	公共運輸普及	自籌財源	藝文展演	身心障礙福利服務
觀光事業			文創產業	婦女福利服務
			文化經費	老人福利服務
				社福資源

## 第三章 研究方法

### 3.1 研究流程

本研究欲建構一個個人化的系統，使用者可由系統前端的頁面，勾選自身關注的面向，再經由後端的統計程式分析得到各縣市的理想程度，再由前端的頁面以視覺化的地圖呈現分析結果，讓使用者獲得個人的理想城市資訊，研究流程見圖 1。

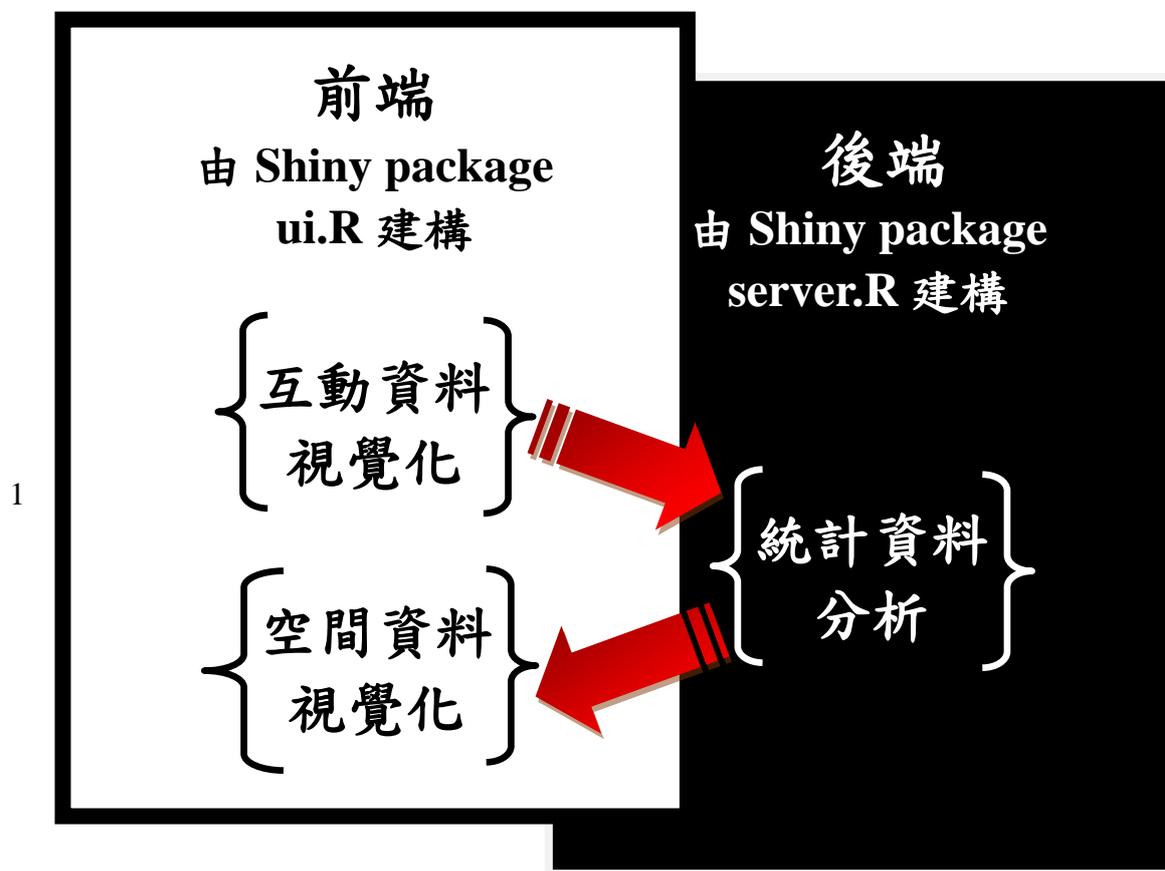


圖 1、研究流程圖

## 3.2 統計分析

### 3.2.1 選擇統計分析方法

本研究的目標是欲從使用者挑選出的細項變數數據，將其整合成一個數值，並由此數值的大小關係，來代表各縣市的理想程度。

迴歸分析是最常見且簡易結合各變數數據的統計模型之一，然而，我們納入分析的變數數量眾多，且部分變數意義可能具有相關性，因此易造成訊息的混雜，此外，因為分析變數個數較探討的縣市數目多，所以可能造成模型的配適和參數的估計不穩定。

主成分分析(Principal components analysis, PCA)這統計方法，可將原先眾多的變數，藉由線性組合將維度縮減成數量較少的新變數，並使新變數間彼此無相關性，且仍保有原先變數所提供的資訊，因此本研究將採用主成分分析將各變數的資訊整合。

### 3.2.2 主成分分析

主成份分析於 1901 年由 Karl Pearson 創建<sup>[4]</sup>，是一種分析、簡化數據集的技術。原有的變數由線性組合成新的變數，以達到資料縮減的目的，但能夠保留住數據本身所提供的重要資訊。其方法主要是通過對共變異數矩陣進行特徵分解，以得到數據的特徵向量(eigenvector)與特徵值(eigenvalue)。由於主成分分析主要依賴數據提供的訊息，所以數據的準確性對分析結果影響很大。

下述以最簡單的二維主成分分析為例，說明此統計方法的理論：

1. 各觀察點原兩變數分別為 $x_1$ 、 $x_2$
2. 將各觀察點投影至新的二維軸線上 $c_1$ 、 $c_2$ ，新產生的變數 $c_1$ 、 $c_2$ 稱為主成分，此觀察點於新軸線上的數值為稱為主成分計分(principal components scores)
3. 新變數 $c_1$ 、 $c_2$ 為原變數 $x_1$ 、 $x_2$ 的線性組合，且均值保持不變為 0
4. 新變數 $c_1$ 、 $c_2$ 與原變數 $x_1$ 、 $x_2$ 的總平方和相同，總變異數也相同
5.  $c_1$ 解釋總變異的百分比一定比任何原始變數 $x_1$ 、 $x_2$ 解釋總變異的百分比大
6. 新變數 $c_1$ 、 $c_2$ 的相關係數為 0，即 $c_1$ 、 $c_2$ 不相關

若將二維推廣至 p 個維度，所有 p 個新軸中能解釋原變數最多的變異，則訂為第一主成分，記為 $c_1$ ；再者，剩下的 p-1 個新軸中，能解釋最多 $c_1$ 未解釋的變

異者，訂為第二主成分，記為 $c_2$ ；以此類推第三、第四至第 $p$ 主成分。

本研究中，我們以相關矩陣來進行主成分分析，其步驟如下：

假設有一組原始變數 $x_1, x_2, \dots, x_p$

1. 原始資料標準化：

$$x^* = \frac{x - \bar{x}}{s}$$

2. 建立相關矩陣：

當變數單位相同時，可用變異數-共變異數矩陣求特徵值與特徵向量，但單位不同時應以相關矩陣求特徵值與特徵向量較為合理。

$$R = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix}$$

3. 將標準化後的變數轉換成主成分，求特徵值與特徵向量，表示如下：

$$c_j = w_{j1}x_1 + w_{j2}x_2 + \cdots + w_{jp}x_p, j = 1, 2, \dots, p$$

其中 $c_1 \dots c_p$ 是 $p$ 個主成分，彼此間無相關。

主成分 $c$ 總變異與原變數 $x$ 總變異相同，即

$$\begin{aligned} \text{Var}(c_1) + \text{Var}(c_2) + \dots + \text{Var}(c_p) &= \text{Var}(x_1) + \text{Var}(x_2) + \dots + \text{Var}(x_p) \\ &= p \end{aligned}$$

若 $c_1 = w_{11}x_1 + w_{12}x_2 + \cdots + w_{1p}x_p$ ，且 $\text{Var}(c_1)$ 最大， $c_1$ 即為第一主成分；

若 $c_2 = w_{21}x_1 + w_{22}x_2 + \cdots + w_{2p}x_p$ ，

且 $(w_{21}, w_{22}, \dots, w_{2p}) \perp (w_{11}, w_{12}, \dots, w_{1p})$ ，

若 $\text{Var}(c_2)$ 最大， $c_2$ 即為第二主成分；以此類推。

$w_{i1}, w_{i2}, \dots, w_{ip}$ 為相關矩陣 $R$ 的第 $i$ 個特徵向量(eigenvector)， $\text{var}(c_i)$ 為相

關矩陣 $R$ 的第 $i$ 個特徵值(eigenvalue)。

4. 保留主成分個數：

決定保留主成分個數的方式可以透過觀察陡坡圖(圖 2)，以觀察主成分解釋比例，根據 Kaiser rule，將僅保留特徵值大於 1 之主成分。

個人的理想城市

5. 建構指標分數：

我們將保留的主成分乘上變異數解釋比例後加總，作為新的指標。

## 3.3 視覺化應用

### 3.3.1 互動資料視覺化

為了讓使用者能簡單依照自己的需求挑選關注的變數，並清楚得到理想的城市排名情況，因此我們建構一個簡易的頁面，可讓使用者能於該頁面操作勾選變數並由地圖的圖示了解分析結果，而不須自行判讀複雜的統計分析數據。

R-studio 程式的 Shiny 套件<sup>[5]</sup>提供了方便的工具，可用於製作資料分析者與使用者互動的頁面，使一般大眾能更簡易的藉由統計的分析了解其關心的議題。Shiny 套件主要分為 ui.R 與 server.R 兩大部分，ui.R 為前端程式碼，用以描述互動頁面的呈現與排版，server.R 為後端程式碼，用以負責分析、計算與繪製圖表，並將結果傳遞給前端；因此當使用者打開瀏覽器時，R-studio 程式的 Shiny 套件會依照中 ui.R 的程式碼，產生使用者看到的頁面內容，並在瀏覽器中顯示。

我們設計的主頁面包含左中右三大部分，左側邊包含兩個版面的切換功能，分別顯示「主程式」和「關於」，點選「主程式」可切換至勾選變數與地圖顯示頁面，點選「關於」可顯示指導老師以及組員的相關資訊；中間部分設計為變數勾選區，將五大面向經濟力、環境力、施政力、文教力、社福力的各細項設定為複選選項，使用者勾選完畢後可按下「點一下，觀看結果」的按鈕後，指令即會套入後台進行資料分析，分析的結果將於頁面右側以地圖呈現；頁面呈現的程式碼皆寫入 ui.R 中。

此外，server.R 城市中包含連結前後端的操作指令與後端統計分析的程式，其中藉由 input 與 output 兩個函式連結前端與後端的相關指令與動作，先將使用者與前端選取的變數條件設定為 input，再讀取本研究蒐集的資料與採用的主成分分析，並將結果數據和地圖圖示設定為 output，並運用 renderPrint 函式生成輸出。

## 3.3.2 空間資料視覺化

### 3.3.2.1 使用空間資料視覺化的原因

使用者挑選的細項變數，經由主成分分析後，則會整合得到各縣市的理想程度的分數，但為了讓使用者更容易辨識排序，因此我們將各縣市的分數由大排到小並可列出總排名。

而相對於文字上的結果顯示，我們想要將結果能夠用更吸引人的方式來呈現。於是我們以「空間資料視覺化」的方式，將結果鑲嵌在臺灣地圖上，透過地圖上顏色深淺的變化，讓使用者更快速的接收結果。

在這個部分我們使用了 R 程式語言中的 Leaflet 套件，Leaflet 源自 JavaScript，用來創建互動式頁面地圖。它除了繪製地圖外，還可以將我們想要顯示的數據嵌入地圖中，這也是我們選擇使用它的重要因素，希望透過這樣的顯示結果，能夠方便使用者在短時間內即可以很輕鬆且清楚地得到答案。

### 3.3.2.2 Leaflet 簡介

Leaflet 是時下廣為運用的開源 JavaScript 庫之一的套件<sup>[6]</sup>，它可以直接使用許多主流的地圖數據作為輔助圖層，如 OpenStreetMap, Mapbox 等，進而將地理信息數據做視覺化的操作。除了在線的地理數據外，也可以很容易的使用 shapefile 格式和 json 格式以及 sp 封包中的空間與經緯度數據。

除了繪製地圖外，還可以在地圖圖層中輕鬆地使用點標記、線條和多邊形等視覺化要素來顯示常用的地理信息。像是我們要定位座標就可以用點標記，要框選某個特定地區則可以用多邊形的方式鎖定範圍。在標示區域還可以再使用當滑鼠移動到該區域時，顯示出相對應的內容。

Leaflet 對於文檔輸出系統也有著良好的支持，可方便地在 knitr / R Markdown 文檔和 Shiny 系統中嵌入地圖。

### 3.3.2.3 Leaflet 運用

#### 1.將數據分級及定義其分級顏色：

首先從資料分析中，得到各縣市的理想程度的分數，但為了讓使用者更容易辨識排序，因此我們將各縣市的分數由大排到小並可列出總排名。之後再依排序劃分為「普通」、「良好」、「優良」三個理想程度，並使用指令來將三個理想程度依序由淺到深對應出指定的顏色。如果是使用原先連續型的資料(程式計算後的原分數)的話，我們可以使用 `colorNumeric` 或 `colorBin` 等的指令來做出顏色淺到深的漸層，但是由於我們是三個不同的分級因此使用 `colorFactor` 函式，圖 2 顯示三個理想程度分級的顏色對應。



圖 2、理想程度分級顏色對應圖

#### 2.使用 Leaflet 函數繪製地圖及顯示結果

先基本建構出我們的網頁地圖，將地圖中心定位在  $120.97^{\circ}\text{E}, 23.98^{\circ}\text{N}$ ，再使用由政府資料開放平台(open data)所提供的「直轄市、縣市界線(TWD97 經緯度)」資料來繪製出臺灣的基本輪廓及各直轄市、縣市之間的界線，並套用在線上的地圖中。在網頁中的設計，還包含地圖右上方製作一個方格，來顯示出我們將分級及顏色對應圖，方便使用者對照。

最後將結果再帶入我們繪製的地圖中，依照結果將各縣市填上分級顏色，並對應顯示在各個縣市中，當滑鼠游標移至直轄市、縣市方塊中，即會顯示出「縣市名稱、等級分級、總排名數」。

#### 3.鑲嵌在 Shiny 系統中

在 Shiny 系統 ui.R 中將頁面右方設定為顯示地圖，並調整其大小以完整顯示出我們使用 Leaflet 函數所繪製出的臺灣地圖。

## 第四章 研究結果

藉由上述的研究方法，我們提供了一個個人化的系統，提供給使用者在主畫面勾選關心的細項變數，並在主畫面右方的方格中顯示繪製出的地圖結果；資料數據的分析和繪製地圖的過程在後端作業，僅提供統計分析者參考，並不提供給一般使用者於頁面前端看見。以下我們以一個實際例子來演示我們的研究結果與使用介面：

介面打開後如圖 3，使用者就可以看到我們彙整的五大面向，總共 26 項的變數。使用者可以自行依照自己對於理想中的縣市標準所考量的因素，去選取符合自己需求的變數。

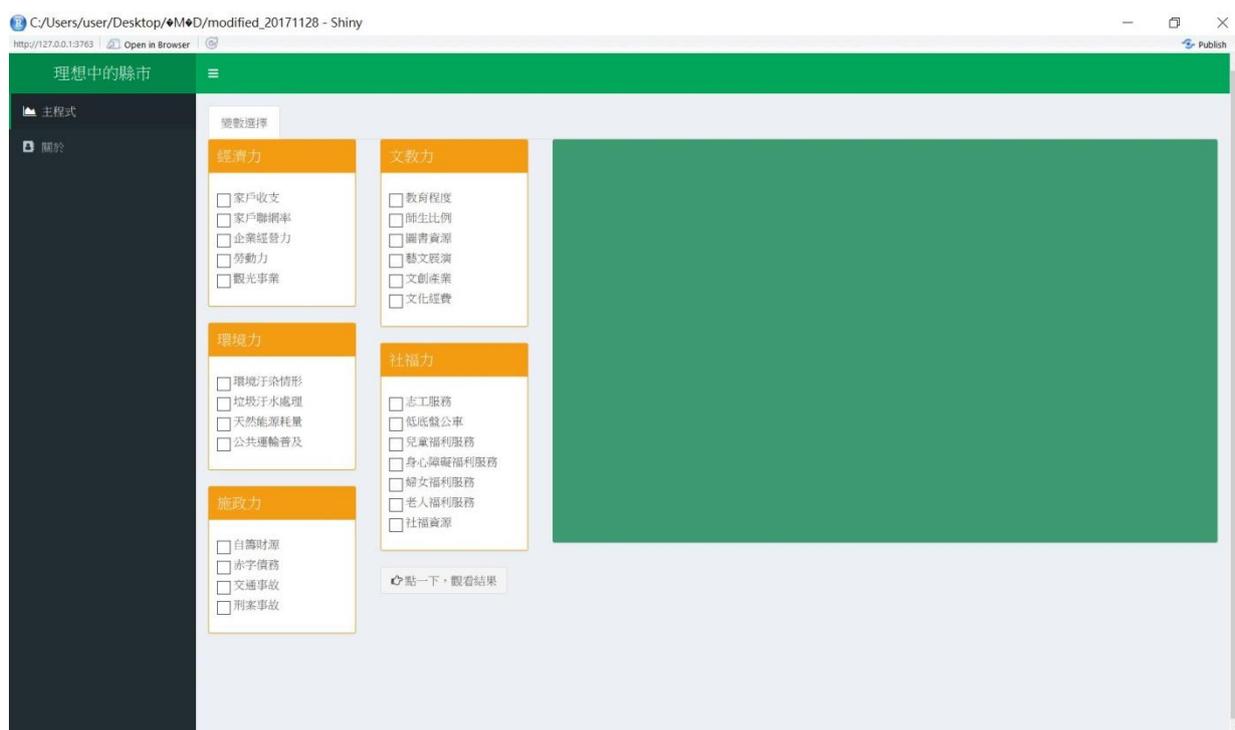


圖 3、初始頁面

假設使用者對於家戶聯網率、觀光事業、垃圾污染處理、天然能源耗量、刑案事故、教育程度、圖書資源、工服務、老人福利服務、社福資源，這 11 個變數感到關心，因此他在「經濟力」、「環境力」、「施政力」、「文教力」、「社福力」五大面向中做勾選，然後按下按鈕來進行統計分析（圖 4）。

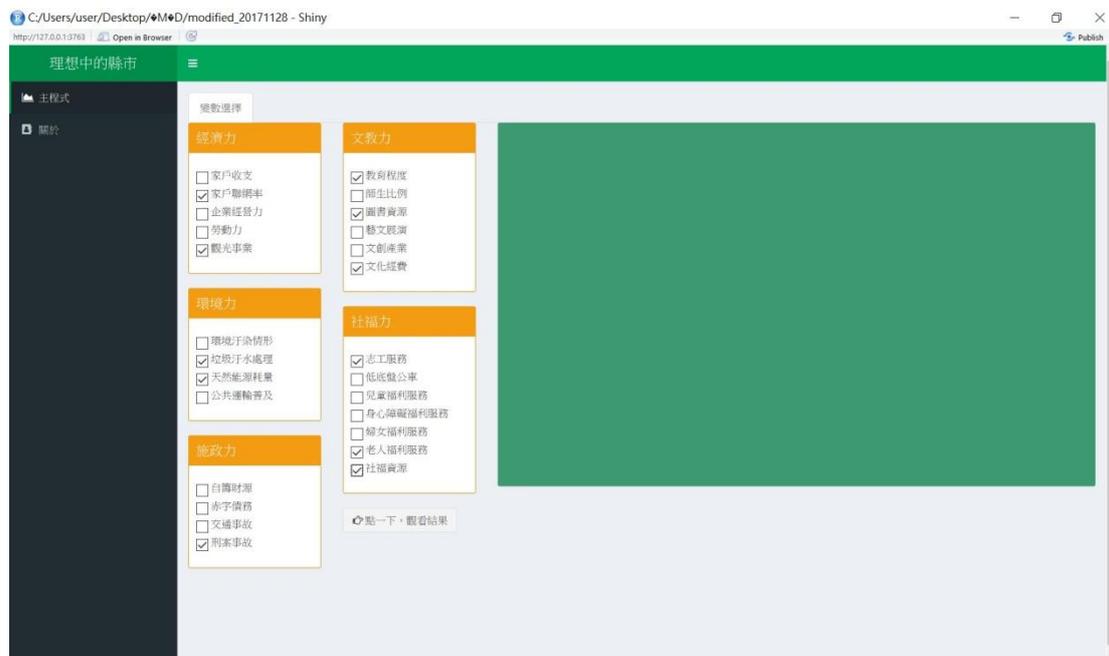


圖 4、勾選後的頁面

選擇好變數之後，Shiny 系統會將得到的指令傳回後端由 R-Studio 程式來進行主成分分析，圖 5 是主成分的各项數據。(註：此結果僅提供給後端統計分析者參考)

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
x8	-0.32366874	-0.048083743	1.033755e-01	0.13285273	-0.11407545	0.076962696	0.2329292767	0.1104561209	0.010328572	-0.10962266
x33	-0.16244995	0.037243828	3.458136e-01	0.08922999	-0.23848628	-0.007549062	0.1247282417	-0.5277032932	0.012515471	0.27233194
x14	-0.27519442	0.043485366	5.728354e-02	0.19356851	0.11729086	0.273071225	0.2283832388	0.2221963674	-0.215889317	0.03740855
x19	-0.25766893	0.184816336	2.242857e-01	-0.05026737	0.19742998	0.007203060	0.0057667231	-0.1327950536	-0.087327171	0.23327782
x27	0.01158896	0.214727551	2.520167e-05	-0.32102234	0.35914588	0.041637726	0.3694415356	0.0009253429	-0.250400219	0.10583096
x34	0.14641489	0.184211825	3.961179e-01	0.05704754	0.04326528	0.023460582	0.1808183670	0.3422570267	-0.010460657	0.23176214
x39	-0.25725110	0.006888596	-1.577919e-01	0.21809517	0.03003929	-0.348712785	-0.1746468485	-0.1460763783	0.007878954	-0.18699641
x43	0.09701466	0.214239046	1.271824e-01	0.37606423	-0.14228621	0.184164831	0.0795203140	-0.0679132333	0.311962795	-0.19114755
x52	0.16295062	0.138382865	3.154496e-01	0.01737267	0.25415656	-0.344772690	0.0467637568	0.0838518422	0.218135839	-0.31886633
x59	-0.23948533	-0.070278929	1.072080e-01	-0.29317957	-0.13516888	-0.196607139	-0.2398813400	0.0929495476	-0.195427150	0.16890931
x62	-0.22533310	-0.157071956	6.660331e-02	-0.32485695	-0.07931445	0.062832951	0.0599882412	-0.1790751922	0.236248671	-0.17000274
x15	-0.11677616	0.336743865	-2.663869e-02	-0.15564778	-0.13105889	0.156869702	0.1247211099	-0.0633480050	0.327084900	-0.28861488
x20	-0.23349353	-0.17595685	2.001036e-01	-0.07311165	-0.03423901	-0.209917933	0.1510094364	0.2862104813	-0.001038078	-0.22889649
x31	0.03565286	0.219430155	3.132866e-01	-0.07240945	0.19508988	0.135858113	-0.2835750337	-0.1744851593	-0.334904580	-0.34341169
x37	-0.20352222	0.180104968	2.438044e-01	0.04103909	-0.39949037	0.040816370	-0.0235438182	0.1956912063	-0.008599667	0.02772154
x40	0.12216938	0.166513175	-1.079916e-01	0.08198381	-0.15830337	-0.428472322	0.4006204978	-0.3030517186	-0.392972230	-0.15660449
x44	0.18815815	-0.223963008	6.776664e-02	0.31367988	0.05118163	-0.028182290	-0.1210899570	0.0402537315	-0.224708489	-0.18691600
x60	0.17514146	0.186313606	-5.170305e-02	-0.04759199	-0.54980770	-0.031077807	0.0257495157	0.0846359859	-0.214416867	0.04455368
x63	-0.01433957	0.335755746	-2.726356e-01	-0.07238512	-0.04201307	-0.174632657	0.0006851194	0.0253355600	0.118716164	0.15191445
x17	-0.26184000	0.188340275	-5.489622e-02	-0.01410738	0.11416620	0.030433527	-0.2241870899	-0.1478372757	0.188313748	0.29797903
x21	0.18906982	-0.282136989	2.379404e-01	0.01490976	-0.18814833	-0.009383327	-0.1912052340	-0.0045034954	-0.115148246	0.08643540
x38	-0.32752203	0.129425092	-4.412459e-02	-0.03748649	0.00734897	0.030293420	-0.1187588206	-0.1285491937	-0.144011778	-0.18135167
x41	-0.01315654	0.337548211	1.052535e-01	0.04626237	0.01919309	-0.386704250	-0.3426889070	0.2525125540	0.073365659	0.09313880
x61	-0.23141083	-0.187241928	-7.680601e-02	-0.22048052	-0.07262578	-0.265885855	0.2551529727	0.1699517095	0.167948629	0.02072102
x32	-0.14207616	0.232305051	-3.443662e-01	0.03976299	-0.08920233	0.223866178	-0.1061767585	0.2537132541	-0.233546866	-0.13897359
x42	-0.13166358	-0.020999276	-1.182376e-01	0.49398229	0.16245913	-0.152323181	0.1370119217	0.0118324884	0.043790685	0.25688057

圖 5、主成分分析數據 (僅截取第一至第十主成分)

接著，根據 Kaiser's rule 與陡坡圖，選擇特徵值大於 1 的主成分，從圖 6 中可以看出特徵值大於 1 的主成分有七個，因此我們保留了前面七個主成分，來做後續的計算。(註：此陡坡圖僅提供給後端統計分析者參考)

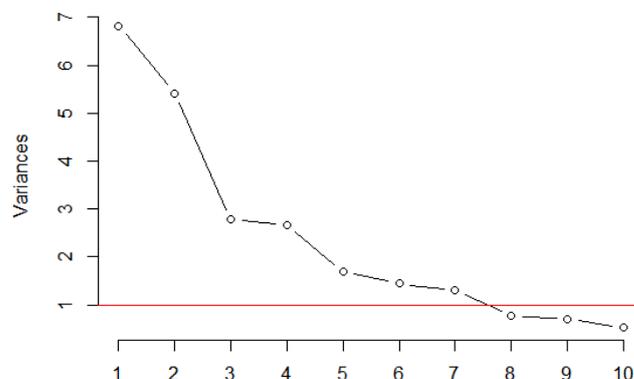


圖 6、陡坡圖

利用新的主成分計算出各個縣市的總分(圖 7)，並依總分高低排序，再依照「優良」、「良好」、「普通」分為三級(圖 8)，用以定義成由深到淺的顏色，方便後續於各縣市中填入色塊。

"New Taipei City"	"-0.14274473081642"
"Taipei City"	"-1.02955123083256"
"Taoyuan City"	"-0.620912402088643"
"Taichung City"	"-0.56382264356257"
"Tainan City"	"-0.615230799080983"
"Kaohsiung City"	"-0.532113289808271"
"Yilan County"	"0.538526798318628"
"Hsinchu County"	"-0.415935121451773"
"Miaoli County"	"-0.017617925267683"
"Changhua County"	"0.140672501265795"
"Nantou County"	"0.453790906311263"
"Yunlin County"	"0.0298954561933492"
"Chiayi County"	"0.479629091807023"
"Pingtung County"	"0.775501275775294"
"Taitung County"	"1.8892436067106"
"Hualien County"	"1.27714745674234"
"Penghu County"	"0.541363438097916"
"Keelung City"	"0.799760343882588"
"Hsinchu City"	"-2.29087456899028"
"Chiayi City"	"-0.696728163205621"

圖 7、縣市指標分數

Yilan County	15	普通	宜蘭縣
Changhua County	12	良好	彰化縣
Nantou County	13	良好	南投縣
Yunlin County	11	良好	雲林縣
Pingtung County	17	普通	屏東縣
Taitung County	20	普通	臺東縣
Hualien County	19	普通	花蓮縣
Penghu County	16	普通	澎湖縣
Keelung City	18	普通	基隆市
Hsinchu City	1	優良	新竹市
Taipei City	2	優良	臺北市
New Taipei City	9	良好	新北市
Taichung City	6	優良	臺中市
Tainan City	5	優良	臺南市
Taoyuan City	4	優良	桃園市
Miaoli County	10	良好	苗栗縣
Hsinchu County	8	良好	新竹縣
Chiayi City	3	優良	嘉義市
Chiayi County	14	普通	嘉義縣
Kaohsiung City	7	良好	高雄市

圖 8、縣市排名與分級指標

最後，系統將繪製好的地圖傳回 Shiny 前端，顯示出來的畫面如圖 9，使用者即可根據各縣市的深淺判別理想程度。

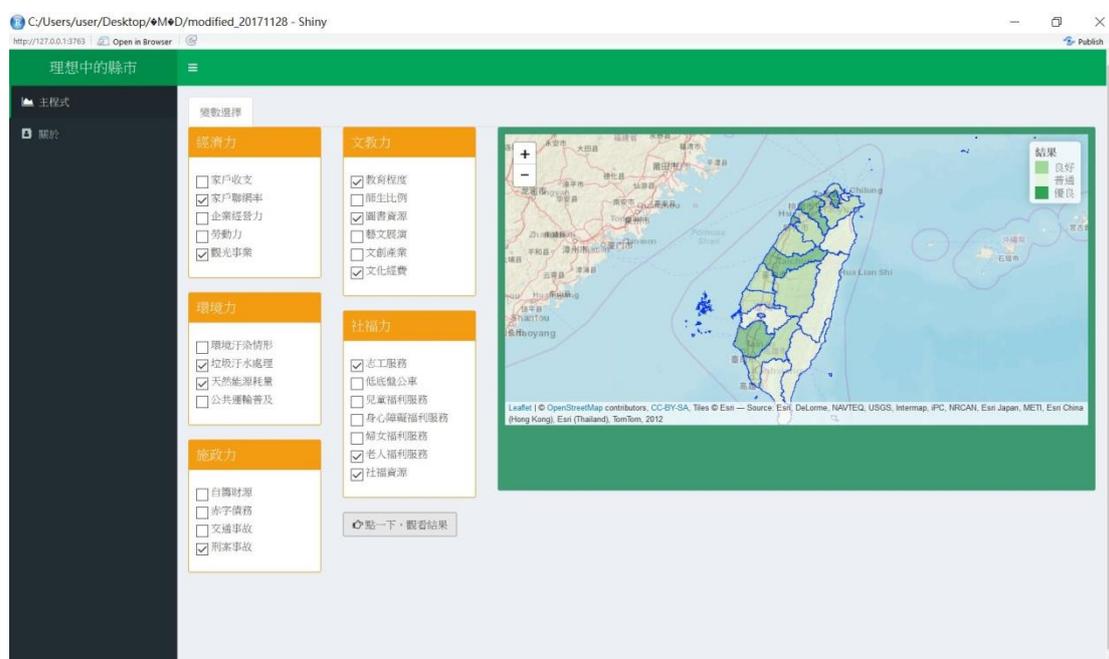


圖 9、顯示結果

## 第五章 結論與討論

### 5.1 結論

本次的專題研究，我們由 R 程式語言建構了一套簡單明瞭、容易操作的個人化決策系統，使用者可以選取自己關心的細項變數，經由後端程式的統計分析，再藉由頁面的地圖顏色，迅速清楚地得到自己的理想縣市的程度。

我們也用完成後的成品試著做出了多組不同變數的分析結果，從中我們可以在選擇不同變數得出的結果來跟天下雜誌做一些比較。根據我們的資料與分析方法，當我們把所有 26 項變數全部勾選，結果顯示六都的顏色都比其他縣市來得深，也就是說明他們的排名等級是比較高的，所以我們的研究結果可以說明，六都在綜合性評比仍是優於其他縣市。而「天下雜誌/第 631 期」依照五大面向去統計分析所得出的結果，顯示臺北市排名第一、其他五都(台南市、新北市、桃園市、高雄市、台中市)分別排在第五、六、七、八、九名，也都是位居前面，兩者相互比較，有一定程度的相似，但是因為使用的變數不同，所以結果並不盡相同。由此可知我們所做的分析結果與市面上其他報導和調查可得到相似的結果，這也說明我們的研究資料與方法是具有可信度的。

而我們也一直在強調，使用者可以依照自己的需求去選擇關心的變數，所以當在單向、或是單一選擇變數時，就不一定會是六都的排名會在前端，反而是沒有這麼都市化的縣市可能會名列前茅，像是當使用者在乎環境更勝於經濟發展時，六都的等級就會相對較低，但是這也是符合一般的常理判斷。比起市面上其他報導和調查都是已經做好排名，且固定變數的分析結果，我們的研究結果更貼近使用者，提供更方便的管道。

總而言之，我們這次的研究建構了一套個人化決策系統，比起市面上其他報導和調查所分析的內容，更加貼近每個使用者的需求與理想，不論是綜合性的評比，或是某些特定變數作評比中，可以知道每個縣市其實都有自己的優、缺點，而在於每個人如何去做選擇，我們提供了更方便、有效的系統提供給使用者判斷自己心目中的理想地區。

## 5.2 討論

本研究提供了全臺灣各縣市的理想程度，以縣市做為分析單位可能仍過於粗略，若能更精細地分析各鄉鎮或鄰里的理想情況，可能會更利於使用者選擇居住區域考量；然而，現有的政府開放平台僅有少部分的變數提供鄉鎮的資訊，即使臺中市政府的網站也僅提供極少數的鄰里訊息，若之後政府的開放平台能提供更多資料，則我們可再更新資料與地圖以提供更精準的訊息。

此外，本研究將各縣市依照排名分為優良、良好、普通三個程度，並未詳列各縣市的排名；我們嘗試過使用經主成分分析得到的縣市總分，由漸層色將結果繪製在地圖中，但可能因總分間過於接近時，漸層的效果反而讓地圖看起來很難辨識；後來也嘗試改用將排名做漸層色的繪製，卻發現排名間隔的顏色區隔太小也不太容易辨識，所以，最後我們決定將排名分組歸類，把 20 個縣市依排名分成 3 等份，讓使用者可以更容易讀取結果。

再者，原本設計的版面將勾選變數的方格都放在上排，顯示地圖放在下方(圖 10)，但這會造成使用者須將頁面往下滑才能看見整個地圖，因此我們又更新了程式重新編排版面，最終使用的介面將顯示結果地圖置於右方，我們希望經由使用者嘗試後能給予我們些許回饋，將來可再根據這些意見編修更清楚的頁面。



圖 10、調整前的版面

## 參考文獻

- [1] OECD Better Life Index. <http://www.oecdbetterlifeindex.org>
- [2] 天下雜誌第 631 期 (2017)
- [3] 中華民國統計資訊網 [http://statdb.dgbas.gov.tw/pxweb/Dialog/CityItemlist\\_n.asp](http://statdb.dgbas.gov.tw/pxweb/Dialog/CityItemlist_n.asp)
- [4] Subhash Sharma (1996). Applied Multivariate Techniques, Wiley, New York.
- [5] Shiny.rstudio <https://shiny.rstudio.com/>
- [6] RPubs. <https://rpubs.com/>