

# 應用視覺化分析 在社區發展資料之可能性初探 ——以Tableau軟體為例

洪婕瑜、鄭夙芬、鄭期緯、吳美鈴

## 壹、前言

現今視覺化分析已然成為學術與實務界的趨勢，筆者在本文中，一方面將描述視覺化分析的意涵，同時說明視覺化在社會福利服務領域的應用，另一方面，將聚焦於筆者與團隊所使用之視覺化分析Tableau軟體之介紹並以其應用於社區發展相關研究為例加以說明，期望提供相關領域專業者在探究視覺化分析之應用上提供參考。

## 貳、視覺化分析的意涵

視覺化是人類最早使用的一種溝通技巧，在遠古時期文字系統尚未出現時，古代人就透過壁畫及雕刻圖像化的方式，將當時的生活及文化記錄下來，也讓後代學者能透過視覺化的圖形了解早期的生活型

態（王筱雯，2021）；隨著電腦科技不斷創新，網際網路的快速發展，視覺化分析軟體應用許多領域，借助人類視覺系統，使人們能夠很直觀、直覺地去對資料數據進行探索，對於資料數據所要傳達的資訊進行最直接的吸收，如此不但加速了人們的學習，還能在學習的同時，對於資料數據有不一樣的洞察與發現（Klerkx et al., 2014）。

視覺化設計最主要的目的，即是將大數據以圖像表示，並且要能將資料縮小到一個螢幕的範圍內呈現出來，同時不扭曲原來數據所要傳達的訊息，如此便能輕鬆地透過一張視覺化圖表來吸收數據所要傳達的訊息（呂偉僑等人，2015；林昱成等人，2021；謝明汗，2019）；好的視覺化設計，可使閱讀者對於數據所要傳達的資訊一目了然，對於不同的資料進行不同顏色、形狀的搭配，讓讀者心曠神怡

(Cruz-Orengo et al., 2014)。

彭其捷(2017)指出資料視覺化的優點，包括：(一)幫助資料降維，去掉無意義的資料，把重要的資訊過濾出來，避免不必要的干擾；(二)有助於同時引導左右腦共同解決問題，左腦擅長邏輯型的工作，右腦擅長視覺型的操作，資料視覺化；(三)使人們更容易理解數據；(四)提升決策信心；(五)提升溝通效率；(六)提升閱讀的吸引力。

視覺化的目的即是為了讓閱讀者能更快速透過單一螢幕的圖表，藉由不同顏色、曲線、形狀等來快速掌握資料所要表達的意義，對於資訊的吸收更為輕鬆與便利，也有利於辨識重要資訊，幫助閱讀者進行溝通與決策，是現代資料呈現的一項趨勢。

## 參、視覺化分析在社會福利服務的應用

視覺化分析應用於社會福利服務領域，林淑娥與許韶芹(2018)指出內政部於1990年代成立「國土資訊推動小組」，開啟我國地理空間資料收集及地理資訊系統(Geographic Information System, 簡稱GIS)開發的濫觴；2005年臺北市少年輔導委員會、2005年臺北市政府社會局分別將業務資訊結合GIS系統形成少年生活地圖、高風險家戶預警系統(張淑慧，

2005；薛承泰，2008)，2016年臺北市政府社會局亦以既有GIS技術基礎上，開發「臺北市福利健康地圖」、「臺北市社會福利資訊系統」等(林淑娥、許韶芹，2018)。

另一方面，由DSP智庫驅動有限公司推動的「資料力做公益」計畫(Data For Social Good, 簡稱D4SG)，運用公益性質組織(非營利、政府或其他)累積的大量營運數據，激發各種創新的公共服務，截至目前已經協助超過40個非營利組織、40個政府機關(廖立文等人，2022)與其合作，其可謂目前國內使用視覺化分析在社會福利服務應用上範圍最廣泛的一項計畫。D4SG團隊協助政府機關及非營利組織進行相關數據分析，包括運用個案記錄、服務人口分布、服務表單等資料，設定分析指標，透過視覺化分析，以儀表板呈現現況、各指標關聯性以及進行各項預測(林祁衡等人，2019；林峰生等人，2017；陳林彰、謝宗翰，2018；廖立文等人，2022)。

此外，Tableau軟體的官方網頁(<https://www.tableau.com/zh-tw>)收錄許多使用Tableau軟體作為分析工具的相關研究及其數據，並開放大眾進行檢索與收藏，其中亦不乏有關社會福利之數據分析，例如，Murugappan(2018)分析美國各城市兒少保護案件之對象、數量、分布、處遇時間、轉介的趨勢等；McKeown

（2013）分析蘇格蘭2011至2012的各項福利支出比例；Schiavone（2021）分析美國各州在2020年使用食品券的家庭數量，以及特別針對南卡羅來納州30年來使用貧困家庭臨時援助金的趨勢。

除Tableau軟體之外，國外學者亦使用其他視覺化軟體進行有關社會工作與社會福利相關研究，例如，Wang等人（2018）使用GIS來分析中國的社會工作碩士班畢業生之專業承諾，探索社會工作畢業生不同程度的工作意向、相關預測因素，研究發現，不同地區社會工作畢業生工作意向的預測因素存在很大的地域差異。

## 肆、Tableau軟體的介紹

有關視覺化分析軟體，目前包括：Google Chart、Tableau軟體、Microsoft Power BI、Matplotlib等，對於大數據的圖像化均有相當的助益，我國不管在中央、直轄市、各縣市政府的網頁多有視覺化統計專區，提供民眾可以一目了然了解相關統計資料，由此顯示，應用各軟體呈現視覺化分析，已為政府行政部門、企業以及學術等領域之趨勢，本文以下將以Tableau軟體為例進行介紹。

Tableau是一套視覺化數據分析軟體，透過大數據資料庫的連結，進行簡易的操作，即可以得到視覺化的圖像，藉由其儀表板，能夠更生動地展現數字間的變

化，讓閱讀者快速理解數據資料所要呈現的重點，並透過快速切換圖像，了解不同數據解釋的方法。

有關Tableau 軟體的特性，筆者整理如下三項。

### 一、數據整合

Tableau軟體是數據處理軟體，透過匯入Excel表來整合不同資料庫的數據，對於不同政府機關的資料，將資料格式統一以Tableau軟體所需格式進行調整，即可將跨數據來源整合於單一分析檔案中。

呂宗憲（2023）提到，Tableau特別針對巨量數據的分析場景進行優化，不需要程式碼就能輕鬆載入千萬級的資料量，而在高效能的電腦上，甚至能載入數億筆的資料量展開分析，讓分析師一次性分析長期性歷史資料，不像過去需要將資料切成不同檔案，並針對每個切開的檔案進行相同操作，可大幅提升產能。

### 二、簡易的儀表板設計

Tableau軟體具有貼近一般使用者的設計，若想調整不同的分析維度，也能夠簡易用拖拉資料表的方式完成。在進行團隊會議時，團隊成員對於資料的呈現與意義歸屬時有不同見解，Tableau軟體的儀表板具有直覺性的操作，能快速呈現不同的圖像，以供團隊成員檢視其使用的適切性（圖1）。

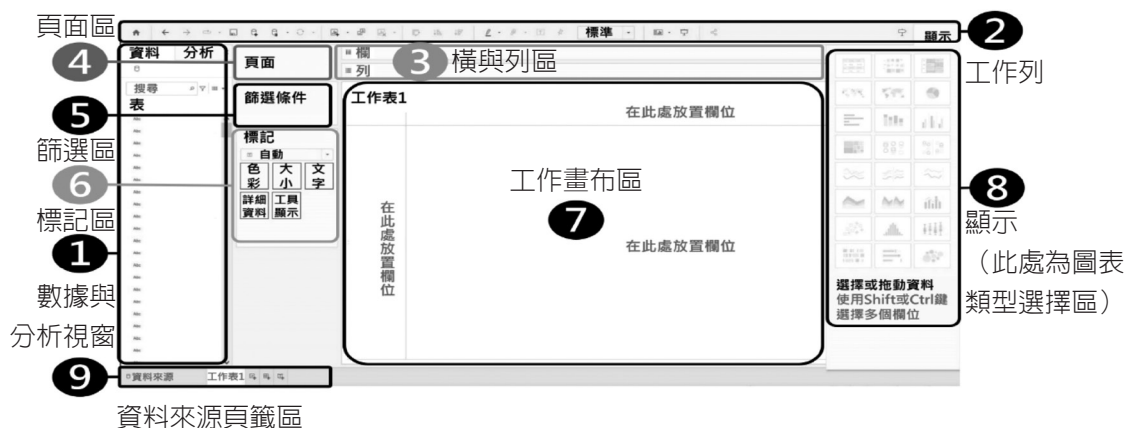


圖 1 Tableau 軟體操作儀表板

資料來源：作者改繪自彭其捷（2017）。

根據Freeman（2000）的視覺化設計原則，動態的介面可讓使用者有豐富的互動感受，進而做到資訊瀏覽的目的；Verbert等人（2013）透過15種不同的儀表板設計讓受測者進行對數據的吸收，實測結果得出儀表板的設計能讓受測者更有效的吸收數據所傳達的訊息，同時對數據所傳達的訊息有更進一步的衍生想法。

### 三、豐富的視覺化圖表

Tableau軟體擁有十足豐富的視覺化圖表，在不同圖像的輔助之下，能將表格呈現統計描述性或推論性統計之數據意義表達更加完整。

有關Tableau軟體內建的各项圖表類型及其適用時機，筆者介紹如下，若仍需更清楚的了解，可參考Tableau官方網站（<https://help.tableau.com/current/>

[pro/desktop/zh-tw/what\\_chart\\_example.htm](https://help.tableau.com/current/pro/desktop/zh-tw/what_chart_example.htm)）；另受限於本文的印刷條件，筆者將色彩改以黑、白、灰色調，或是虛線等線條進行虛擬繪製，以期呈現資料在圖形上的辨別。

#### （一）檢視一量值隨時間的變化

若要顯示某個量值隨時間的變化，例如，此量值在過去一年的變化、此量值變化的時間為何，或此量值變化的速度有多快等，可以選擇折線圖呈現（圖2），筆者以虛擬某地人口不同年度變化相關數據繪製成折線圖。

#### （二）關聯性

若需要找出兩個變數之間的關聯性，例如，這兩個量值是否相關、關聯性有多強、某些量值的關聯性是否強於其他量

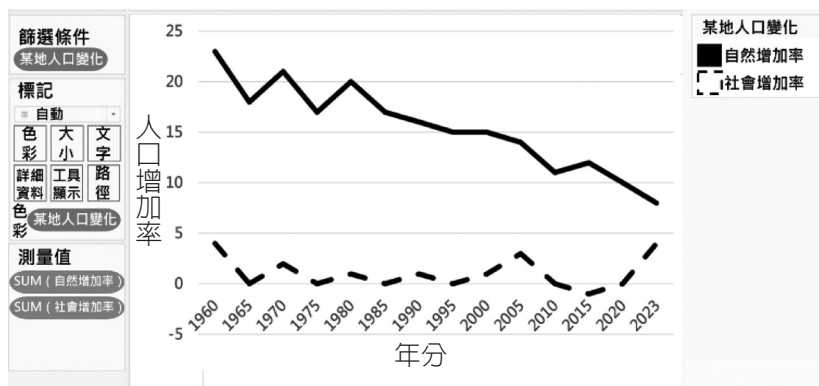


圖 2 折線圖

資料來源：作者繪製。



圖 3 填補泡泡圖

資料來源：作者繪製。

值，或是這些量值的關聯性強度如何等，可透過散點圖呈現，但需注意，關聯性不等於因果關係，在本文的第五部份第四項評估中，筆者即以二個散點圖分別呈現不同行政區與資源總數比例不同分級之關聯，以及不同行政區資源總數分數與不同社區培力輔導分級之關聯。

### (三) 量級

量級是指兩個或以上維度項目的相對大小，例如，了解哪個維度的量值最高，或是在同一維度下，最低和最高量值之間的差距有多大，可使用長條圖、填補泡泡圖或折線圖呈現（圖3），筆者以虛擬某行政區之社區資源種類繪製成填補泡泡圖。



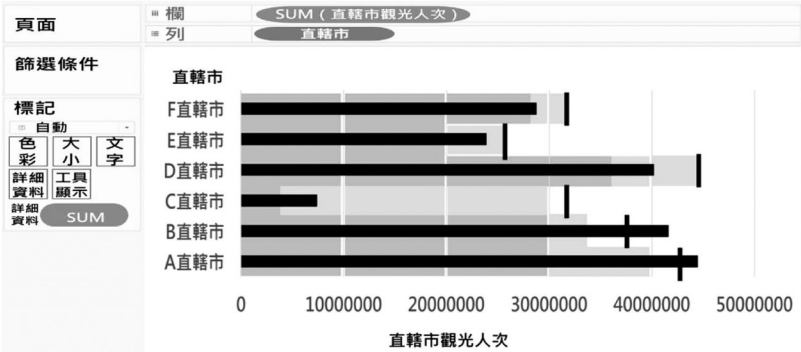


圖 4 標靶圖

資料來源：作者繪製。

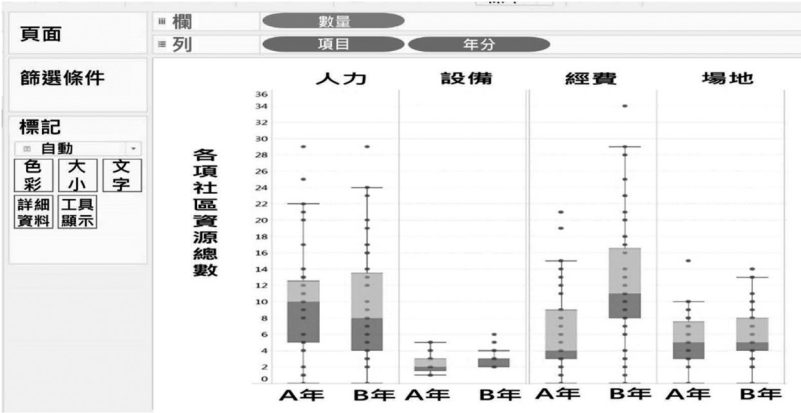


圖 5 盒鬚圖

資料來源：作者繪製。

#### (四) 偏差

偏差是指某個量值距離基準線（如：平均值或中位數），其之間所在的偏差有多大，例如，此量值與標準值相差多遠、此量值的偏差有多重要，或是偏差是否有規律等，可使用標靶圖或長條圖呈現（圖4），筆者以虛擬各直轄市實際與預期觀光人次繪製成標靶圖。

#### (五) 分布

若要顯示某事件在某個群體中的頻率，例如，時間是否圍繞特定概率聚集在一起、哪個群體購買的商品最多，或是我們工作日中最忙的時間是什麼時候等，可選擇長條圖、人口金字塔圖、柏拉圖或盒鬚圖呈現（圖5），筆者以虛擬某地兩年度之各項社區資源總數消長表現與分散集

中情形繪製成盒鬚圖。

#### （六）排名

若要顯示量級，還想描繪所有維度項目的相對排名，例如，企業、公司績效排名、各產業平均收入，或是各產業平均收入最低的值是多少等，可採用長條圖呈現。

#### （七）占比

若要顯示某項目單一部分占總體的比例，例如，此值占總數的多少、企業每年度銷售額如何變化，或是不同商品對銷售額的貢獻是否因地而異，可透過圓餅圖、區域圖、堆疊長條圖或樹狀圖呈現。

#### （八）空間

若要在資料中顯示精確定位並顯示地理格局，例如，企業的分公司在哪個城市的銷售額最高，或是一地的家庭福利服務中心收案人數有多少，可以使用地圖顯示，其中地圖又包括填充地圖、點分布圖、符號地圖和密度地圖，在本文的第伍部份第四項評估中，筆者即以填充地圖呈現某市之各行政區資源總數比例分級分布情形。

#### （九）流程

流程圖是一種隨時間變化的而變動的圖表，顯示同一起點和終點隨時間變化的

路徑，例如，最長的運輸路線，或是一地中交通最壅塞的地段，可用Sankey圖呈現。

#### （十）同時比較多個指標的維度

若要同時比較多個指標下的維度項目，例如，想要了解一地的大眾運輸工具使用率，需比較一地多個大眾運輸工具的使用情形（如：捷運、公車、輕軌等），可採用雷達圖來呈現，在本文的第伍部份第四項評估中，筆者即以二個雷達圖分別呈現不同行政區社區培力輔導分級社區之各類型資源總數比例分布。

筆者基於Tableau軟體眾多的圖表功能及詳細的使用時機指引，加上簡易且直覺的儀表板，同時，國內較缺乏Tableau軟體應用於福利服務之案例，因此透過接受公部門委託研究，嘗試以Tableau軟體的分析方法呈現區域發展觀點下的各行政區社區發展現況與趨勢分析，以下將以該研究案中之資源指標為例，說明Tableau軟體的應用與檢視。

### 伍、我們透過Tableau軟體說了什麼？

筆者與團隊接受公部門委託，運用Tableau軟體的分析方法，以次級資料分析（Secondary Data Analysis）為主來呈現區域發展觀點下的各行政區社區發展現

況與趨勢分析。基於區域發展被廣泛認同的指標大多為社經指標，且為貼近社區發展脈絡，以及次級資料取得性等考量，因此，擬定分析指標包括：「人口」、「資源」、「會務」、「業務」、「財務」與「社區分級」等。以下將以「資源」指標之視覺化分析歷程進行描述與相關舉例，此處所指之資源意指區域內擁有之各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構等資源。

筆者與團隊根據CRISP-DM方法論（Cross Industry Standard Process for Data Mining, CRISPDM）進行分析。Chapman等人（1999）指出CRISP-DM各階段流程與任務如下：

- 一、業務理解（Business Understanding）：  
了解研究問題與需求，進而確認研究變項與屬性，以及可獲取的資料。
- 二、了解數據（Data Understanding）：進行資料蒐集、整理，了解數據所代表的意義，進而確認資料的正確性。
- 三、資料準備（Data Preparation）：進行資料清洗，將數據呈現一致性的結構，以利後續Tableau軟體的執行。
- 四、塑模（Modeling）：將資料依照研究者想要的分析單位進行歸類，有助於進行後續的資料評估與分析，達到研究者的需求與期待。
- 五、評估（Evaluation）：將塑模後的資料透過軟體操作產生成果，進行分析，並

賦予意義，進而找出有價值的資訊。

六、部署（Deployment）：找出有價值的資訊後，將相關建議提供研究單位進行使用，進而對企業或者是政府單位的政策提出改善策略。（Chapman et al., 1999）

透過CRISP-DM流程，可以確保資訊的品質與產出結果的質量、減少發現知識所需要的技能，亦可大大減少成本與時間（Chapman et al., 1999）。筆者進一步參考Chapman提出的CRISP-DM流程圖，依照筆者與團隊的研究經驗進行繪製（圖6），並形成研究流程，分述如下。

## 一、業務理解

### （一）確認分析指標

區域發展指標相當多元，被廣泛認同的指標大多為社經指標，例如，各項公共建設、交通、公共衛生、住宅、環境品質、大眾秩序與安全、教育文化、地方財政、家庭收支與設備、就業以及人口等（國家發展委員會，2017）；從社區發展脈絡中的資源觀點來看，筆者及團隊採用「正式資源」做為指標，再思考指標下的相關細項，以及考慮公部門現有資料庫有哪些次級資料？在本案中，將指標細項含括：「教育－各級學校數」、「宗教－宗教團體數」、「社會－社福與慈善團體機構數」及「醫療－醫療院所與藥局數」。



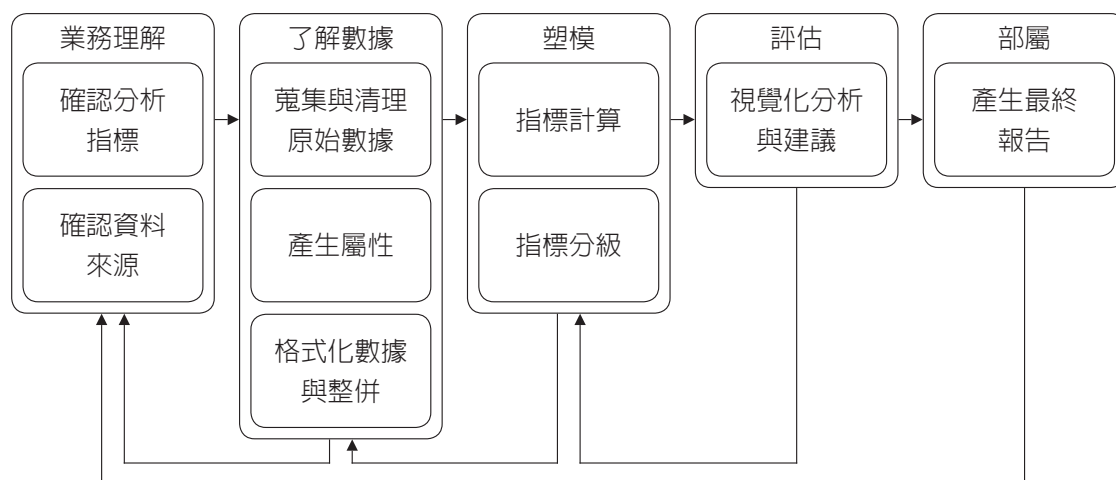


圖 6 本研究流程圖

資料來源：作者參考Chapman等人（1999）之CRISP-DM流程圖後自行繪製。

## （二）確認資料來源

以本案例來說，在確立研究指標後，筆者及團隊開始檢視次級資料如何取得，包括透過教育局全球資訊網蒐集各行政區各級學校名冊；透過民政局網站蒐集各行政區宗教團體名冊；透過健保特約醫事機構查詢系統，蒐集各行政區醫療院所與藥局名冊；透過社會局福利地圖系統，蒐集各行政區社福與慈善團體機構名冊，再進一步將名冊資料進行資料清理。

## 二、了解數據與資料準備

### （一）蒐集與清理原始數據

在蒐集數據的過程中，需要將資料重新以數字彙整成Excel檔，並將無法使用的數據清除，筆者與團隊面臨政府資料跨越至研究所用之間的障礙。

### 1. 政府資料為人工輸入，定義不一致

因政府資料往往由人工輸入，未有明確且嚴謹的定義數字或符號，例如，「0」究竟是「數量為0」還是「無資料」，不同人員輸入時有各自見解，更何況是跨局處之資料，若沒有將定義統整，將導致後續研究上的混淆、誤用。

### 2. 資料闕漏，僅能輸入「Null」

在整理原始數據時，發現因數據的缺漏，但已無從追溯，因此僅能輸入「Null」，代表「無資料」。若「Null」的數據筆數過多，可能會造成分析結果無法反映真實樣貌，甚至無法分析。

### 3. 各局處均為原始資料，未格式化

在進行數據分析前，需要將數據以統一格式輸入Excel，但此階段之數據尚未格式化，均為原始資料，筆者與團隊在資源

的部分共蒐集31個檔案，每個檔案中亦有上百筆數據，將數據格式化可謂一大挑戰。

## （二）產生屬性

使用Tableau軟體進行視覺化分析前，需要先將數據格式化，這個格式分為「維度（Dimension）」與「量值（Measure）」，若Excel欄位中的內容被歸類在「維度」，則Tableau軟體會將其視為一種類別，以字串的方式呈現，另一方面，若Excel欄位中的內容為數字，通常會被Tableau軟體自動放置到「量值」，且通常會自動加總數據（彭其捷，2017）；以本案例來說，筆者及團隊將各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構等資源之數量資料的性質定義為「量值」，代表其數值可以被加總，形成各級學校總數、宗教團體總數、醫療院所與藥局總數、社福與慈善團體機構總數，以及全部相加的資源總數。

## （三）格式化數據與整併

將資源指標下的資料依各指標性質進行分類後，筆者與團隊將各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構等資料類別分類拆分為更細項，例如，各級學校進一步分類為幼兒園、國小、國中、高中職及大專院校，並根據該市各行政區的幼兒園、國小、國中、高中

職及大專院校數量，分別使用Excel進行統計，其他資料處理方式亦同上述，輸入完成後，將各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構等資料予以彙整，整併為一份該市各行政區資源總數的Excel總表檔案。

## 三、塑模

### （一）指標計算

待數據均完成建檔後，在塑模的階段中，筆者與團隊須將指標下的分析類型定義，並根據定義將細項資料的數據進行計算，例如，某年某行政區資源總數比例之定義是指該行政區各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構等數量加總占全市各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構加總之百分比，同時，依據研究所需要呈現的分析單位，例如，不同行政區、不同社區培力輔導分級，便可以進一步呈現，成為最終可用之數據。

### （二）指標分級

以本案例來說，經由計算所獲得的資源指標之數據後，基於該數據並沒有一項標準以判定該數據的意義，例如，資源總數比例屬於高、中或低，因此，筆者與團隊進一步將資源總數比例進行分級，透過計算出最大值及最小值，以及其所分布之區間後，以四分位法進行指標分級，平均

分為四級，藉此賦予各數據意義，例如，資源總數比例在該市各行政區平均得分分級方式為「第一級0-1.27%」、「第二級1.28-1.92%」、「第三級1.93-4.04%」、「第四級4.05%以上」，分級數越高代表資源總數比例越高，並將數據根據指標分級輸入Excel後，匯入Tableau軟體進行分析。

#### 四、評估：視覺化分析

##### （一）填充地圖

以本案例來說，地圖的作用主要用於描述各行政區整體資源總數比例的分級情形（圖7），第一級包含A、D及J行政區，第二級包含F、I及M行政區，第三級包含C、G及K行政區，第四級則包含B、E及H行政區，分級數越高代表資源總數比例越高。

以A區舉例，整體而言，可以發現A區的整體資源總數比例最低是其劣勢，同時透過地理位置的判斷，了解其資源輸送上的困難，因地處中央，與其他行政區之間有山脈阻隔，對於資源的開發較為困難。

另一方面，透過地圖還能以地理位置判斷該行政區可能採取的策略，例如，隔壁的C行政區屬於資源總數比例分級第三級，可試圖連結C行政區的資源作為短期策略，但開發在地資源應還是重要的中長期目標。

##### （二）散點圖

以本案例來說，散點圖能夠呈現各行政區資源總數比例分級的分布狀況，並從圖中按照比例排列，能夠更清楚呈現資源豐富與不足的行政區分別為哪些，同時可

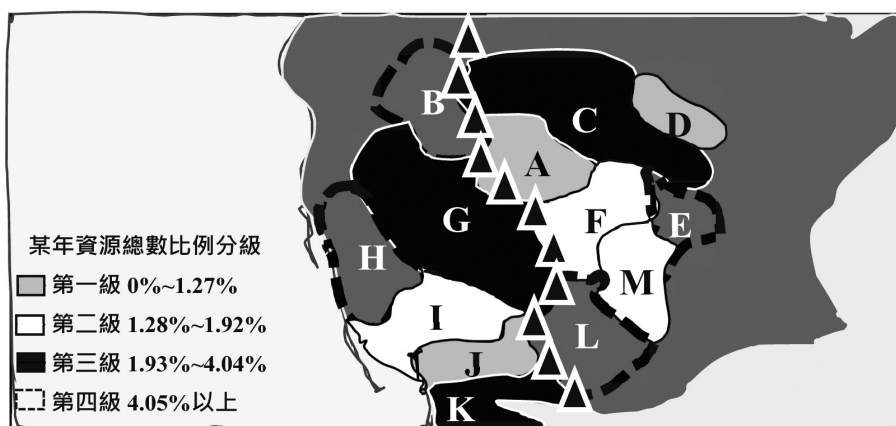


圖 7 以填充地圖呈現某市各行政區資源總數比例分級分布情形

資料來源：作者繪製。

以看到多數行政區所集中的資源總數分級情形，以及少數行政區擁有較高的資源總數分級情形（圖8）。

另一方面，可以將社區培力輔導分級與各行政區資源總數分數交叉呈現，同時將資源總數分數以三分位法平均切割為高、中、低（圖9），可以發現資源總數分數低者多為停滯型、起步型社區，資源總數分數高者則可能為停滯型、潛力型、起步型或穩定型社區，進一步從其行政區分布來看，若資源總數分數與社區輔導分級級數均有一致的高分，意即穩定型社區落在資源總數分數的高分區間，即顯示其發展條件可以最為完整並具有優勢。

然而從圖9之示例中可以發現，有的行政區會出現資源總數分數在高分區間，但社區分級屬於停滯型，顯示該行政區的停滯型社區未必意味缺乏資源，應需要更了解該行政區的社區背景特質與社區輔導

方式，藉以提供更對焦的培力協助。

### （三）雷達圖

雷達圖可以同時比較多個指標下的維度項目，以本案例來說，雷達圖用以呈現各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構之各項資源數比例，同時，加上社區培力輔導分級的類別，可以同時了解停滯、潛力、起步與穩定型社區在該行政區資源數，藉以快速了解該行政區內所擁有的資源現況以及差異，以圖10的A行政區之穩定型、起步型、潛力型與停滯型社區在各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構資源總數比例等之雷達圖形相似，但在資源比例之級數多以停滯型為最高；潛力型與起步型在社福與慈善團體機構資源數比例與停滯型級數相同，宗教團體資源數比例則以潛力型稍高，由此可見該行政區可見停

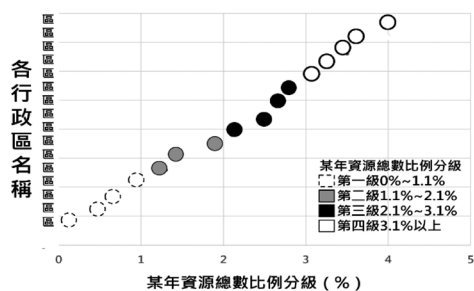


圖 8 以散點圖呈現不同行政區與資源總數比例不同分級之關聯

資料來源：作者繪製。

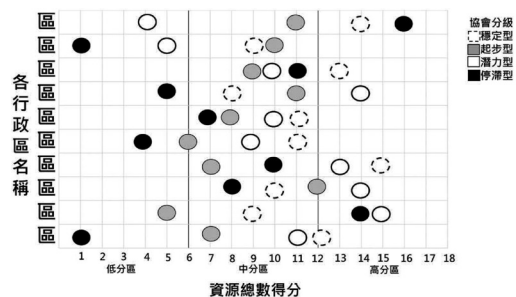


圖 9 以散點圖呈現不同行政區資源總數分數與不同社區培力輔導分級之關聯

資料來源：作者繪製。



圖 10 以雷達圖呈現 A 行政區社區培力輔導分級社區之各類型資源總數比例分布

資料來源：作者繪製。

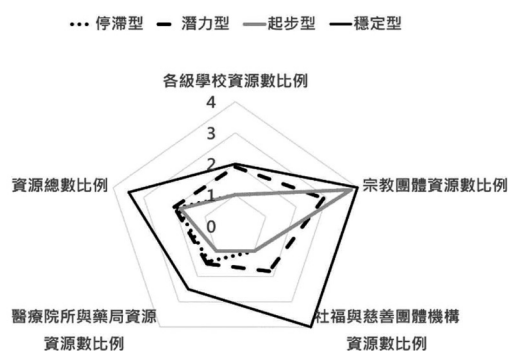


圖 11 以雷達圖呈現 B 行政區社區培力輔導分級社區之各類型資源總數比例分布

資料來源：作者繪製。

滯型的資源比例高有其可以強化的發展潛力，同時，社區培力輔導分級指標不一定在每一行政區均是一體適用。

另一方面，在圖11的B行政區之穩定型社區在各級學校、宗教團體、醫療院所與藥局、社福與慈善團體機構以及資源總數比例等之級數均為四類型之最高，其中，社福與慈善團體機構資源數比例之級數最高，此一樣貌顯示該行政區的穩定型社區符合在資源上的優勢現況，但，該行政區的潛力型雷達圖形雖與穩定型相似，但比例整體縮減，而停滯型擁有宗教團體資源優勢，其他資源數比例均偏低，顯見宗教團體資源為該行政區該類型社區可藉以動員之處。

## 五、部屬：產生最終報告

在本分析中，透過視覺化發現各行政

區的相似與差異樣貌，且各區域所具備的資源數量及種類也有所落差，透過視覺化能夠呈現整體資源最為缺乏的區域，亦能發現相對資源豐沛之處。另一方面，透過視覺化分析也呈現，過去對於社區培力輔導分級的定義與現況有所差距，據此反映出社區培力輔導分級的定義與標準可能需要根據現況更精準的重新劃分，同時也提醒社區發展實務界與學術界，同一標準之社區培力輔導分級不一定適用於每一行政區。

## 陸、結論

### 一、與非視覺化相比，視覺化能幫助什麼？

在過去，對於人文社會科學研究的分析工具往往提及SPSS、SAS等統計軟



體，資料的呈現上，不論推論性統計還是描述性統計，均多以表格為主，以致於在資料的解讀上僅有平面的呈現，例如，數字的高低、分布的多寡等。加入視覺化分析後，數字能夠被「立體」的呈現，其意義不僅停留在高或低，更進一步的能夠觀察極端值及判斷相對優勢與劣勢。

在觀察極端值方面，觀察極端值為何重要？極端值是指一個觀察值與其他觀察值數值呈現很大的差異，所以懷疑該觀察值並不是經由同一機制所產生（Stevens, 1990）；然而極端值的存在，將會嚴重影響到很多統計量之估計值，導致後續詮釋有所偏誤；例如，極端值可能影響資料分布的情況（Wilcox & Keselman, 2003）；極端值在迴歸分析中的斜率估計值容易產生偏差（Pedhazur, 1997）。

過去極端值因為沒有視覺化的呈現，難以被研究人員發現，甚至會忽略，以致讓研究人員可能做出不正確的詮釋（Dunne et al., 2012），另一方面，研究人員也可能不清楚應如何判斷極端值，究竟要使用什麼方法、什麼圖像，才能辨識極端值，對於現今常用的統計軟體而言，往往是一大疑慮。雖然SPSS與SAS兩者均有繪製統計圖之功能，或是利用Excel也可以進行多種圖像的資料視覺化，但其產生之圖像無法得知極端值個案為何，儀表板的操作也較不方便，以致常用的統計圖往往僅停留於簡易的長條圖、圓

餅圖等，忽略數據背後的極端值是常有的事。

現在透過視覺化分析，以散點圖的呈現，較容易發現極端值，或稱離群值的發生，例如，透過各行政區呈現各類資源比例之散點圖，可以發現不同行政區分別有哪些資源有極端值產生？哪幾個行政區有同一類資源的極端值產生？再進一步解讀對於這些行政區共同或差異的意義。另一方面，透過雷達圖的呈現，再加上社區發展協會分級，可以了解目前該行政區不同分級之社區發展協會在各類資源之相對優勢與劣勢，透過雷達圖形判斷該行政區各社區分級具有較多資源的類別？或較少資源的類別？進而能夠針對此現況進行深入探究。

## 二、避免對視覺化理解上的謬誤

視覺化分析畢竟是統計分析的一環，統計解讀上可能發生的偏差在視覺化分析上也會發生，但視覺化分析有著圖像的外衣，比起單純數字的呈現，圖像更容易使人信服，正如謝明汗（2019）所言，視覺化設計借助人類視覺系統，使人們能夠很直觀、直覺地去對資料數據進行探索以及對於資料數據所要傳達的資訊進行最直接的吸收，相較於平面的文字，視覺化的圖像較容易深植於閱讀者的腦海。

有關視覺化分析最常見的謬誤，主要展現在「相關不蘊涵因果（Correlation

Does Not Imply Causation)」，又稱為相關不代表因果，是科學和統計學經常強調的重要觀念，意思是若兩個事物有明顯的相關時，不一定表示兩者之間有因果關係（Pearl, 2009）。當圖像產生趨勢或差異時，在統計的判斷上會稱之為有差異或相關，但統計上的結果不一定就全然等於因果關係，有時只是數理上的湊巧，並非真正的起因與導果，視覺化的呈現，容易讓人誤以為分析所呈現的圖像即為真正的趨勢，而忽略背後可能的成因並非為表面上所看到的，因為這種趨勢可能是毫無邏輯可言，並沒有絕對證據可以證明。

大量的文獻閱讀以及實務脈絡的佐證是很重要的，筆者與團隊在進行視覺化分析的解讀時，往往會透過自身累積的社區實務經驗加以輔助，以促進分析與現況之連結，避免以表面的方式解讀分析結果或成為無用之數據分析。

### 三、資料視覺化是資料治理的驅動力

現今我國政府將資料視覺化已成趨勢，例如，2020年開始在我國盛行的新型冠狀病毒肺炎疫情，衛生福利部疾病管制署COVID-19全球疫情地圖（<https://covid-19.nchc.org.tw/>）即透過資料視覺化，例如，全球地圖呈現病例數、死亡數、目前蔓延至多少國家，並計算死亡率；或是透過長條圖呈現本土、境外移入以及死亡病例趨勢圖。透過視覺化的呈

現，能夠快速反映當局我國現況，以驅動政府施政決策，成為資料治理的推進工具，可能帶來更便民的行政服務，抑或提升政府的施政效能，甚至是將數據當作證據，用來捍衛政府自己的政策。

與臺灣相比，國外運用資料驅動政府決策的起步更早，過去20年來美國知名的資料治理的相關例子，包括紐約警察的犯罪跟蹤統計系統、巴爾的摩市的數據追蹤和管理工具、歐巴馬政府的證據導向政策及預算、紐約市長資料分析辦公室的數據商用方案，以及用數據驅動績效衡量等，這些都是過去政府善用數據管理和分析，來輔助施政決策的不同作法（余至浩，2017）。

然而，若要將資料視覺化，以利長期反映施政狀況，目前在我國政府機關大數據的管理下，還有待改進之處，畢竟政府人員對於數據的繕打格式往往是為了方便操作、好管理，但用於分析與研究之間可能存在許多資料清理的難題，因政府資料為人工輸入，每個輸入的人員定義不一致，以致研究人員在分析時光是資料清理就占去大半時間；另一方面，數據的闕漏情形也是一大問題，闕漏可能會影響分析結果的真實性，但在政府機關保存的數據中卻沒有標示為何缺少該數據，且礙於數據筆數龐大，隨著時間的推移，也日漸難以找回當初的數據。因此，若政府機關可以在資料的管理與保存上，採用有一致性

的格式，以利研究人員或相關數據分析師的後續作業，同時應有資料健檢制度，定期檢查資料的健全性與標準化，以利後續應用上減少發生數據缺漏的疑慮，將有助於資料驅動政府決策往前邁向更大的一步。

（本文作者：洪婕瑜為第一作者，高雄醫學大學醫學社會學與社會工作學系碩士

生；鄭夙芬為第二作者兼通訊作者，高雄醫學大學醫學社會學與社會工作學系副教授；鄭期緯為第三作者，國立東華大學民族社會工作學士學位學程助理教授；吳美鈴為第四作者，高雄醫學大學醫學社會學與社會工作學系助教）

**關鍵詞：**視覺化分析、視覺化分析軟體、Tableau、社區發展

## 參考文獻

- 王筱雯（2021）。《台灣高壓用電趨勢視覺化分析》（碩士論文，淡江大學）。臺灣博碩士論文知識加值系統。<https://hdl.handle.net/11296/7hxc28>
- 余至浩（2017年10月15日）。〈數位政府決策要有新思維，讓數據帶頭說話〉。iThome。<https://www.ithome.com.tw/news/117276>
- 呂宗憲（2023）。《視覺化Tableau互動分析平台之學習成效》（碩士論文，南臺科技大學）。臺灣博碩士論文知識加值系統。<https://hdl.handle.net/11296/539b46>
- 呂偉僑、吳淑娟、蘇希洵、張敬仁、孫天龍、黃建華（2015）。〈以互動式可視化儀表盤做為社區老人跌倒風險評估工具〉。《台灣復健醫學雜誌》，43（2），99-110。[https://doi.org/10.6315/2015.43\(2\)04](https://doi.org/10.6315/2015.43(2)04)
- 林祁衡、蕭玉資、胡文馨、徐紹婷、李佳昇、余佑駿（2019年1月13日）。〈脫貧就業：協助經濟弱勢者重回勞動市場〉。D4SG。<https://d4sg.org/employment-counseling/>
- 林昱成、盧鈺欣、林懿範（2021）。〈會計領域學位論文研究議題視覺化分析——大數據下之實證〉。《中華會計學刊》，17（2），357-402。[https://doi.org/10.6538/TAR.202112\\_17\(2\).0005](https://doi.org/10.6538/TAR.202112_17(2).0005)
- 林峰生、林廷州、林恆宇、胡馨勻、葉憲周、郭泰竹、郭家諭（2017年12月26日）。〈高風險危機警示模組〉。D4SG。<https://d4sg.org/high-risk-family-warning-module/>
- 林淑娥、許韶芹（2018）。〈雲端上的社會福利健康地圖——從臺北市布建社福設施據點經驗談起〉。《社區發展季刊》，161，20-31。
- 國家發展委員會（2017）。《都市及區域發展統計彙編》。
- 張淑慧（2005）。〈地理資訊系統（GIS）在少年犯罪防治工作之應用〉。《社區發展季刊》，111，116-125。
- 陳林昶、謝宗翰（2018年6月6日）。〈未被滿足的法律扶助需求〉。D4SG。<https://d4sg.org/>

legal-aid/

彭其捷（2017）。《大數據時代必學的超吸睛視覺化工具與技術：Excel+Tableau成功晉升資料分析師》。博碩。

廖立文、陳薇亘、黃鼎豪、鄭雅憶（2022年7月22日）。〈讓社工外勤不再危險〉。D4SG。  
<https://d4sg.org/fieldwork-risk/>

薛承泰（2008）。〈社會福利資訊及高風險預警系統：內容與功能簡介〉。《社區發展季刊》，123，282-294。

謝明汗（2019）。《視覺化國際油價：Tableau儀表板之應用》（碩士論文，淡江大學）。臺灣博碩士論文知識加值系統。<https://hdl.handle.net/11296/fr9wpk>

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (1999, March). *The CRISP-DM user guide*. <https://s2.smu.edu/~mhd/8331f03/crisp.pdf>

Cruz-Orengo, L., Daniels, B. P., Dorsey, D., Basak, S. A., Grajales-Reyes, J. G., McCandless, E. E., Piccio, L., Schmidt, R. E., Cross, A. H., Crosby, S. D., & Klein, R. S. (2014). Enhanced sphingosine-1-phosphate receptor 2 expression underlies female CNS autoimmunity susceptibility. *The Journal of Clinical Investigation*, 124(6), 2571-2584. <https://doi.org/10.1172/JCI73408>

Dunne, C., Shneiderman, B., Gove, R., Klavans, J., & Dorr, B. (2012). Rapid understanding of scientific paper collections: Integrating statistics, text analytics, and visualization. *Journal of the American Society for Information Science and Technology*, 63(12), 2351-2369. <https://doi.org/10.1002/asi.22652>

Freeman, L. C. (2000). Visualizing social networks. *Journal of Social Structure*, 1. <https://www.cmu.edu/joss/content/articles/volume1/Freeman.html>

Klerkx, J., Verbert, K., & Duval, E. (2014). Enhancing learning with visualization techniques. In J. Spector, M. Merrill, J. Elen, & M. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 791-807). Springer. [https://doi.org/10.1007/978-1-4614-3185-5\\_64](https://doi.org/10.1007/978-1-4614-3185-5_64)

McKeown, A. (2013, December 18). *Welfare spending in Scotland 2011-12*. <https://public.tableau.com/app/profile/andy.mckeown2231/viz/welfarespendingtreemap/Scotlandonly>

Murugappan, K. (2018, May 30). *Child welfare*. [https://public.tableau.com/app/profile/kumar.murugappan3353/viz/ChildWelfare\\_0/Table](https://public.tableau.com/app/profile/kumar.murugappan3353/viz/ChildWelfare_0/Table)

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, 3, 96-146. <https://doi.org/10.1214/09-SS057>

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: An explanation and prediction*. Holt, Rinehart & Winston.

Schiavone, M. (2021). *Welfare in the US*. [https://public.tableau.com/app/profile/miranda.schiavone/viz/DIY\\_16190352387210/Dashboard1](https://public.tableau.com/app/profile/miranda.schiavone/viz/DIY_16190352387210/Dashboard1)

- Stevens, J. P. (1990). *Intermediate statistics: A modern approach*. Lawrence Erlbaum Associates.
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500-1509. <https://doi.org/10.1177/0002764213479363>
- Wang, Y., Guo, Y. Q., & Zeng, S. H. (2018). Geographical variation and social work students' job intentions in China: A geographic information systems approach. *Social Work*, 63(2), 161-169. <https://doi.org/10.1093/sw/swy004>
- Wilcox, R. R., & Keselman, H. J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods*, 8(3), 254-274. <https://doi.org/10.1037/1082-989X.8.3.254>