

chapter 7

數據與地圖資料 視覺化概論

/ 蔡天怡

資料視覺化是資料分析過程中十分重要的環節，透過視覺化圖表呈現分析結果，可以簡潔清晰的表達資料中蘊藏的意義與趨勢。本章介紹資料視覺化的基本概念及意涵，並介紹 **Tableau** 視覺化工具。本章共有四個小節，包括：資料視覺化的基本概念與原則、資料視覺化工具選介與準備工作、以政府開放統計資料談數據資料視覺化，以及地圖資料視覺化。學習完本章後，讀者將了解資料視覺化的重要原則，並能以 **Tableau** 工具繪製長條圖與地理圖表。

7-1 資料視覺化的基本概念與原則

什麼是資料視覺化？

你可曾在閱讀充滿統計數據的文字或欄位複雜的表格時感到暈頭轉向？對於資訊超載的情況感到認知無法負荷，甚至感到焦慮？若你曾有上述感受，資料視覺化或許有助化解上述情形。

「資料視覺化」是用視覺呈現的方式為資料說出它的故事。顧名思義，這是種將資料轉化成視覺形式的歷程，即透過圖像和／或顏色等視覺要素來呈現資料，而其主要目的在透過圖像傳遞資訊，以達成更有效而精準的溝通。

將資料轉化為視覺呈現的過程中，勢必包含資料的處理與分析，方得有效呈現其內涵。因此，資料視覺化與資料分析密不可分，這包含了人文社會科學理解資料的各種分析方式及許多領域皆常使用的統計分析；如何透過適當的工具將資料批次清理與轉換，則與電腦科學息息相關，而其視覺呈現中的圖像規劃更與藝術設計等領域相關。也就是說，資料視覺化的相關領域可能橫跨了人文藝術、社會科學與理工等學門，是相當跨領域的學問。

根據 Stephen Few（2012），資料視覺化的重要性可從視覺是人類非常重要的感官談起，人們非常仰賴視覺感知解讀所見事物。一般說來，人們平時感知週遭事物時，大約 70% 來自視覺；30% 來自其他感官。因此，掌握視覺線索（visual cues）如何影響感知是相當重要的。

圖 7-1 是以《100 Things Every Designer Needs to Know About People》書中實例為基礎所作的。根據該書作者 Susan Weinschenk，在沒有特殊提示的情況下，一般人閱讀圖中左右兩側的文字時，會依照顏色此一視覺線索所引導的方向來理解詞彙的意涵。因此，絕大多數的人會順著顏色指示的方向，將左圖讀作「STOP TALK」、「QUIET NOW」；將右圖讀作「STOP QUIET」、「TALK

NOW」。你是否也受到視覺線索的影響了呢？本節在介紹過資料視覺化的功能與原則後，將探討資料視覺化原則並進一步說明視覺線索的運用。

STOP	TALK	STOP	TALK
QUIET	NOW	QUIET	NOW

圖 7-1 視覺線索可能影響閱讀方向

資料視覺化的主要功能

資料視覺化的核心價值包括：幫助人們思考和溝通、使人得以更有效地了解問題、更清楚而精確地說故事。藉由提供更多資訊，以克服人們認知記憶能力上的限制，以協助我們發掘未曾注意的資訊、協助找出各種關係、協助解決問題。

理想的分析與呈現方式使人得以作更多有意義的比較、更大量的比較（greater quantity）、更多面向的比較（more dimensions），以及更多觀點的比較（multiple perspectives）。

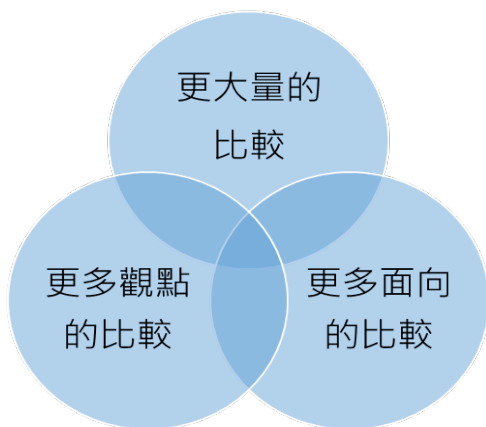


圖 7-2 資料視覺化有助達成理想的資料分析與呈現

其應用層面亦非常廣泛，從各種科學資料視覺化到新聞資料視覺化等等，從不同的學術領域到日常生活，舉凡涉及資料分析與處理的範疇，皆可應用資料視覺化。

資料視覺化的基本原則

視覺呈現的故事邏輯安排與文章的段落安排類似，行文往往須依特定邏輯與架構安排，讓文章容易閱讀；視覺呈現同樣須有其敘事的故事邏輯，讓讀者或受眾容易理解這些圖像所欲傳達的意涵，也才能發揮資料視覺化之功能，達到有效溝通的目標。

至於視覺呈現的故事邏輯如何安排？為了幫助合理地理解與思考，視覺呈現通常會先提供概覽（**overview**），再進一步提供局部和篩選過的資訊（**zoom and filter**）。最後，如果是載體可提供互動設計（如：網站上的視覺呈現），則可依閱聽人之需求提供細節（**details-on-demand**）。

這個安排與報導或學術文章撰寫的段落安排相似，尤其在學術寫作中，研究結果的呈現通常也由研究發現的概況展開（如：整體分布情形、各項描述性統計結果），再進一步依照研究問題所關注的個別議題，針對相關細節作進一步對照與討論等。最後，若有特別值得討論卻未必落在原先預設研究問題之重要發現，也會在此進一步說明。這種從整體概況到具體細節（**from general to specific**）的脈絡安排，是資料視覺化在規劃圖像順序與關係的重要觀念，也是讓圖文架構一致又簡單易懂的好作法。

就視覺呈現的設計而言，Edward Tufte 在《Beautiful Evidence》一書中提及資料視覺化的基本原則，以下簡述各項原則的要點：

1. 原則一：呈現比較關係

將資料視覺化時，記得不斷詢問自己：「這可以和什麼作比較？」由於理想的資料視覺呈現旨在協助讀者「消化」整理資訊，使圖像豐含大量、多面向、多觀點的資訊，因此，盡可能呈現多重的比較關係，方能充分發揮資料視覺化之效益。

2. 原則二：呈現因果關係

在相關理論或文獻的基礎上，可盡量嘗試將可能相關或具因果關係的變項

納入圖中，讓視覺化成果得以用來解釋關聯，更系統性地了解事物的架構，深化視覺呈現成果。因此，領域先備知識的掌握對於作圖亦相當重要。若能根據相關文獻與理論提出假設，或基於自己的相關知識，則更能掌握哪些欄位變項放入圖中，可能有助釐清因果關係。

3. 原則三：呈現多變量的資料

由於現實世界往往是多變量的（兩個變項以上），作圖時，一般希望盡可能納入豐富的相關資訊，以呈現事物較全面的樣貌。再者，資料視覺化旨在協助讀者「消化」大量而多樣龐雜的資訊，因此，在有限的空間中，設法以簡潔清楚的方式呈現較多變量的資料，是理想的視覺呈現之重要原則之一。

4. 原則四：整合證據

Edward Tufte 也提醒在進行資料視覺化時，切勿過度仰賴工具進行資料分析，整合文字、數字、圖表等各種呈現方式，才能發揮資料視覺化的最大效益。有些時候加上簡單的註解說明，仍然很重要。在各種視覺化工具如此便利的環境下，我們的確更應審慎看待軟體所提供之功能，並加以彈性運用，千萬別讓軟體限縮了我們的視覺呈現方式。

談到整合證據的作法，有一著名的視覺化案例可參考：

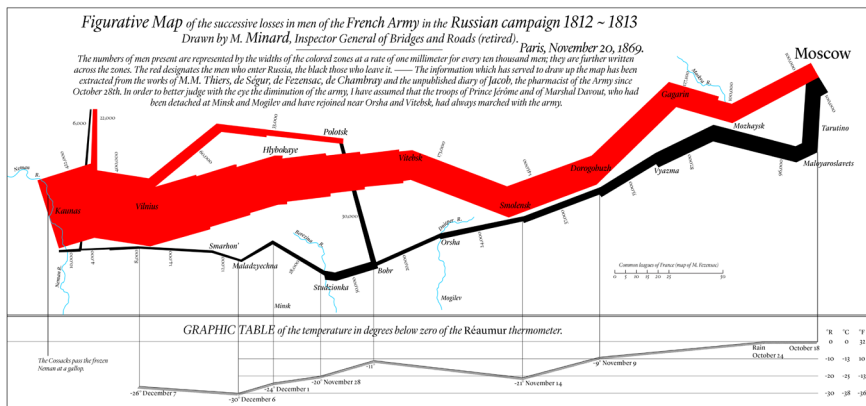


圖 7-3 Charles Minard 's depiction of Napoleon 's ill-fated march on Moscow

資料來源：https://en.wikipedia.org/wiki/File:Minard_Update.png (This file is licensed under the Creative Commons Attribution-Share Alike 4.0 International license)

該圖描繪拿破崙 1812 年征俄失敗的情形，結合了豐富的資訊（即「證據」）與呈現方式，將多方資訊，如：軍隊人數（傷亡情形）、地理位置、行軍距離與方向、氣溫、日期等，全數整合在二維視覺空間中。

5. 原則五：以合適的文字、尺度、資料來源等描述並紀錄各項證據

視覺化的呈現必須如實地說出完整的故事，切勿扭曲事實或刻意操弄。合適的說明與資料來源等描述有助協助理解視覺呈現，例如：圖例說明是相當重要的元素。而合適的尺度更是如實呈現資料的重要關鍵。若以不合適的尺度呈現長條圖，可能讓數值相近的類別看似差異甚大，或讓數值差異甚大的類別看似所差無幾，都可能有操弄視覺成果之嫌。例如：圖 7-4 以相同數據作圖，但在不同尺度下，左圖看似餐廳 A 和 B 的評價落差甚大，而右圖看似餐廳 A 和 B 的評價相近。

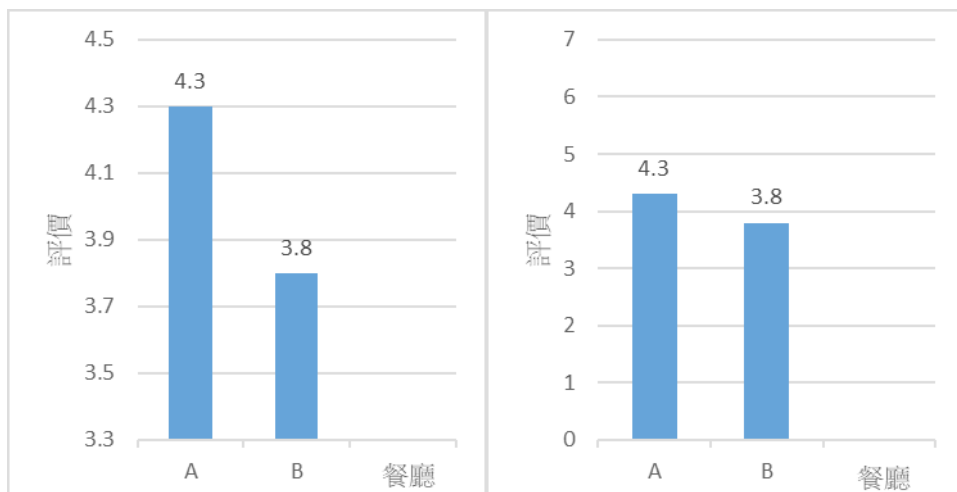


圖 7-4 不合適的尺度將操弄視覺化成果

6. 原則六：內容至上

資料分析與呈現之成敗最終仰賴資訊內容的品質、相關程度及完整性。即使視覺成果做得再華麗、再精美，倘若其內容空洞、不完整、品質不佳或缺乏連貫性，該資料視覺化也稱不上成功之作。一切還是得回歸根本的內容品質控管，才能做出好的視覺成果。

視覺線索 (visual cues) 的運用

除了前述原則，掌握視覺線索 (visual cues) 也是相當重要的。在資料視覺化的範疇中，所謂的「視覺線索」是指人們用以理解圖像呈現的要素。根據 Nathan Yau 在《Data Points》書中的，常見的視覺線索包括：位置、長度、角度、方向、面積、體積、顏色、色調等。一般說來，人們通常只能較精確地感知某些視覺線索（如：位置、長度、角度、方向），而較無法精確感知其他類型的視覺線索（如：面積／體積、顏色／色調）。因此，將資料視覺化時，我們通常要盡量以人們能較精確感知的視覺線索為主要「提示」，避免單獨使用人們較難精確感知的視覺線索。

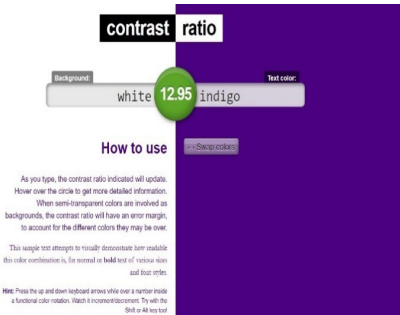
此觀念與「選擇合適的呈現方式」一節之內容可對照理解。許多時候，我們在選擇合適的視覺呈現時，會設法避免讓所選的圖像必須單獨仰賴唯一一種視覺線索，尤其是前述人們難以精確感知的線索。舉例而言，圓餅圖的主要視覺線索仰賴角度，但加上顏色或色調，可提供多重線索；泡泡圖的主要視覺線索是面積，但許多時候也會加上色調，提供多重線索。前述兩例中，圓餅圖較為單純，且其所仰賴的視覺線索是人們較易精確感知的角度，因此，即使以黑白或單色呈現，去除色彩之線索，對於人們理解此圖像的影響不是太大。相形之下，泡泡圖所仰賴的面積與色調都是人們不易精確感知的線索，繪製這類圖像時，提供多重線索益形重要。

以下將進一步探討人們無法精確感知的視覺線索，並簡要說明運用這些線索時特別須注意之事項。

色彩的運用是視覺線索運用中，非常重要的一環。姑且不論人們較難精確感知顏色，根據線上教育百科，男性紅綠色盲人口約佔全人口的 8%、女性約佔 0.5%，且人們對色彩的感知與其所在文化密切相關。因此，我們在規劃視覺呈現時，須慎選顏色。以下選色與配色工具，可供參考：

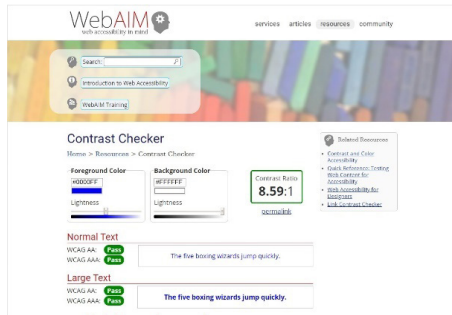
7

數據與地圖資料視覺化概論



Contract ratio

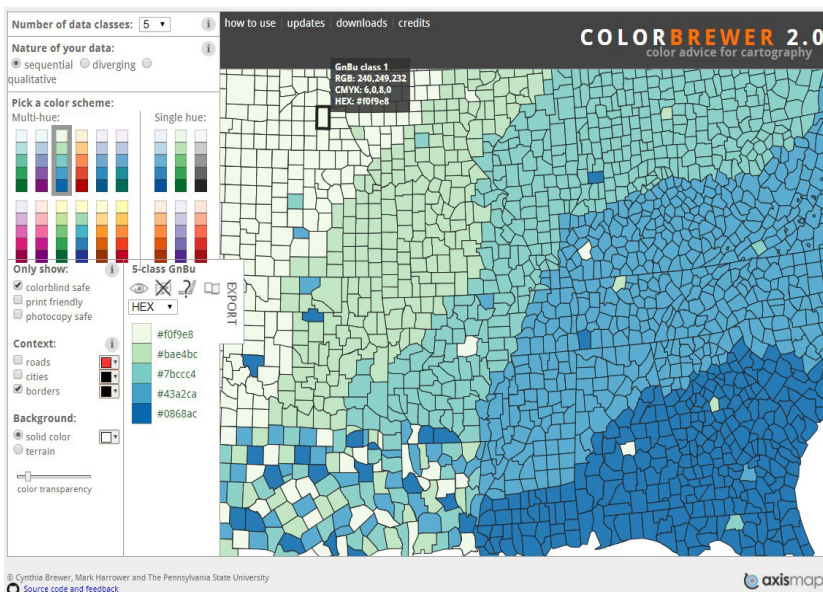
<https://contrast-ratio.com/>



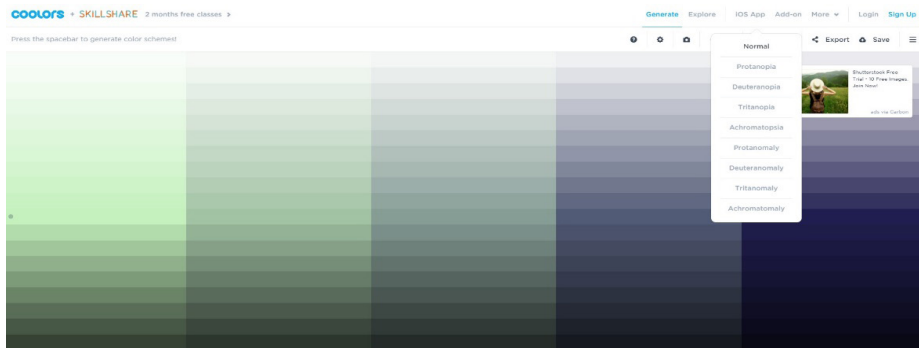
Color Contrast Checker by WebAIM

<https://webaim.org/resources/contrastchecker/>

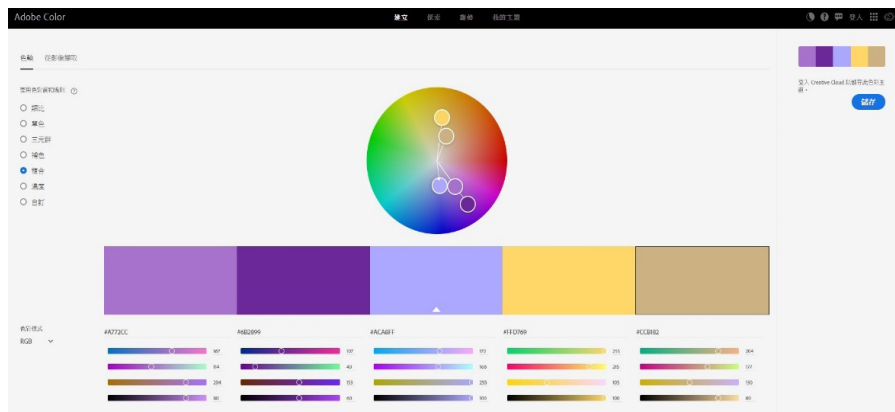
圖 7-5 對比色檢查工具舉隅



Color Brewer <https://colorbrewer2.org/>



Colors <https://colors.co/>



Adobe Color CC

<https://color.adobe.com/>

圖 7-6 選色工具舉隅

透過類似 Color Blind Awareness (<http://www.colourblindawareness.org/>) 這類網站的模擬（如：<http://www.colourblindawareness.org/colour-blindness/colour-blindness-experience-it/>），我們可了解不同類型的色盲人士所見的色彩與一般人有何不同。然而，若難以避開不利色盲人士辨識之顏色，則務必加上其他視覺線索，或同時以色調深淺加強顏色所提供之線索，避免顏色成為單一線索。



圖 7-7 色盲人士所見顏色模擬網站

針對人們難以精確感知的另一項線索——面積與體積，特別提醒大家慎用立體圖。立體圖雖然可能有助呈現複雜的向度，但在資料視覺化講求有效準確溝通的脈絡下，非必要時，建議避免使用立體圖，因為體積也是一項人們較難以精確感知的視覺線索。換言之，除非所欲呈現的維度多到無法透過二維空間呈現，再考慮以二維的平面空間來呈現三維以上的立體圖。當然，若跳脫嚴謹的資料視覺化前提，當你認為讀者無法準確感知立體圖形大小等並不影響你所要傳遞的資訊時，立體圖仍然可使用，一切端看作圖的首要目標為何。

此外，使人分心的背景、顏色與裝飾等皆無益溝通。以圖 7-8、7-9 為例，其用色亦不符合前述原則，其背景與立體長條圖皆對讀者理解圖中資訊無太大助益。尤其此二圖所欲呈現的資訊僅包含兩個向度，以單純的二維平面圖來呈現已綽綽有餘。因此，就資料視覺化原則而言，此二圖有諸多不甚合理之處。其中，圖 7-8 的資料呈現方式可參考本章第四節地圖的作法。

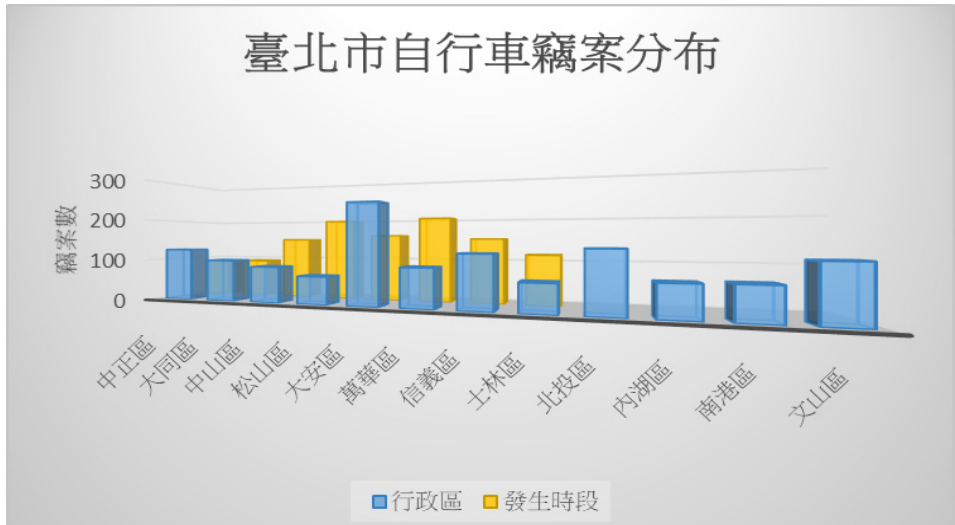


圖 7-8 人們較難精確掌握立體圖

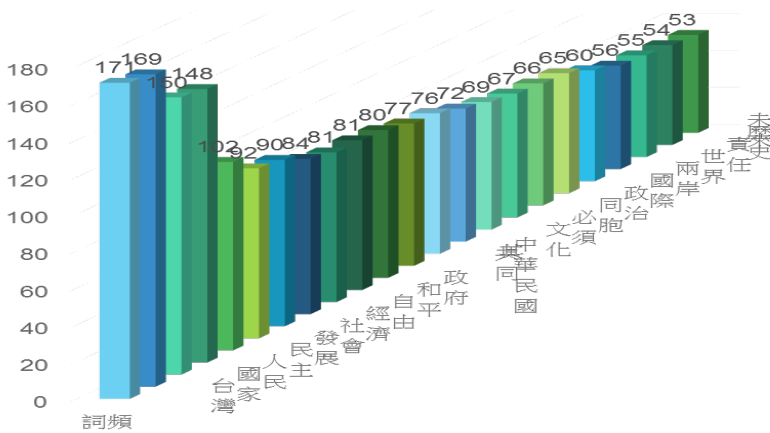


圖 7-9 人們較難精確掌握立體圖

最後，項目順序的重要性也是值得一提的原則。順序的安排可說是一項隱藏的重要視覺線索，順序本身雖非視覺元素，但卻深深影響了我們的視覺理解。例如：只要將簡單的長條圖稍加排序，無論是由高至低或依照年代時序排列，都能讓人更快速理解圖中的資訊，找到最大、最小值或高、低點，有助進行更快速的比較與類聚。總之，清楚掌握自己究竟想要傳遞哪些資訊是最根本的。根據所欲傳達的主軸，對圖表中項目加以排列，將大幅提升該圖表的易讀性。

介紹了這麼多資料視覺化原則與相關叮嚀，有些時候作圖實在難以面面俱到，但有了這層認識與理解，未來在進行資料視覺化時，可更彈性地運用各項原則，並在當下情境中，作出最理想的取捨。例如：倘若為了使用某些代表特定商標的顏色而犧牲了符合色盲人士的需求，則必須同時運用更明顯的色調對比，並搭配其他多重視覺線索來作圖。而平時看新聞或閱讀到包含各種圖表的內容時，可隨時練習用資料視覺化相關原則加以檢視與評論，如此將有助將相關原則內化並自然地運用至自己的視覺呈現中。

選擇適合的呈現方式

當我們要選擇適合的視覺呈現方式時，除了要了解用來作圖的欄位變項之屬性（如：類別變項、連續變項）之外，更要掌握作圖的主要目的。其實，若以簡潔易懂與有效溝通為主要訴求，國民中小學數學課所學的基本概念對於選擇適合的視覺呈現已相當充足。針對變項數量與屬性，選擇長條圖、直方圖、折線圖等呈現，可清楚表達所欲傳遞的資訊內容。

美國天主教大學行銷學系的 **Andrew Abela** 教授自 2009 年以來，製作了幾份協助選擇視覺化呈現方式的決策圖。此決策圖簡明易懂，以下列出視覺化的主要目標及較合適的視覺化呈現方式，供選擇視覺呈現方式之參考：

- 項目比較 → 長條圖
- 歷時比較 → 折線圖、長條圖、雷達圖
- 變數間的關係 → 散布圖、泡泡圖
- 分布狀況 → 直方圖、曲線圖、散布圖
- 靜態比例分布 → 圓餅圖、瀑布圖、100% 堆疊長條圖
- 歷時變化情形 → 100% 堆疊長條圖、堆疊長條圖、堆疊面積圖

上述各種圖形樣式可參考 **Andrew Abela** 在 **The Extreme Presentation**

Method 網站上，根據 Gene Zelazny 的《Saying It With Charts》一書所繪製的「選擇好圖表」（Choosing a good chart）之內容（<https://extremepresentation.typepad.com/files/choosing-a-good-chart-09.pdf>）。而 Juice Analytics 也據此製作了線上圖表選擇工具並提供微軟 Excel 與 PowerPoint 格式的範例案下載（<http://labs.juiceanalytics.com/chartchooser/index.html>；圖 7-10）。

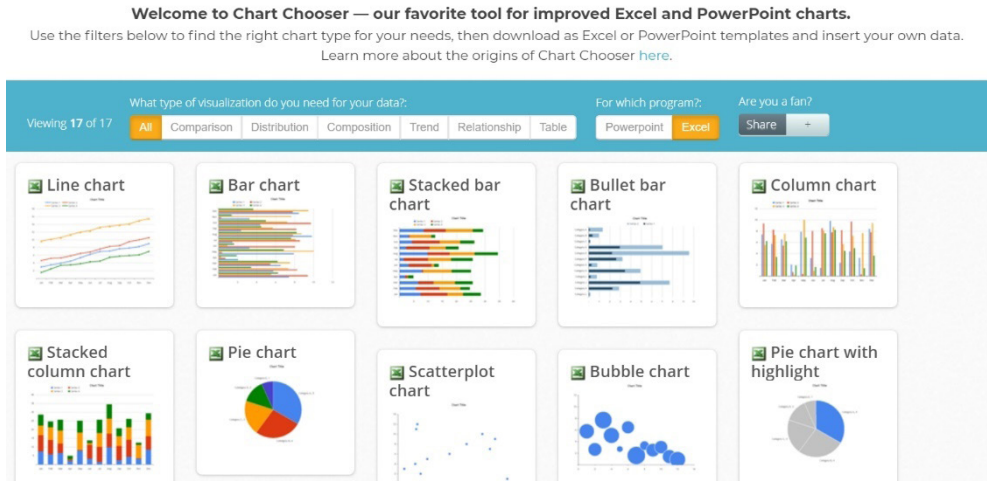


圖 7-10 Chart Chooser

小結

資料視覺化旨在協助我們有效溝通與傳遞資訊，透過各種視覺呈現，協助讀者與受眾減輕認知負荷、消化大量龐雜的原始資料，讓更豐富且有意義的訊息，更快速有效地傳遞。

讀完本節，你應該已掌握資料視覺化的核心精神。當你手邊拿到一份資料，就大膽依循前述資料視覺化原則，適當地運用視覺線索，選擇合適的視覺呈現方式，莫忘初衷，好好地繼續跟隨本章腳步，為資料說出它的故事吧！

7-2 資料視覺化工具選介與準備工作

資料視覺化工具選介

「工欲善其事，必先利其器」。若你透過 Google 搜尋引擎查找資料視覺化工具，檢索結果多得嚇人，各種推薦文論及的工具也是五花八門、不勝枚舉。當我們要進行資料視覺化，究竟該如何選擇工具呢？

研究資料視覺化常用的工具時常與資料型態與資料分析方法息息相關。因此，選擇資料視覺化工具時，首先要了解即將用來視覺化的資料型態（如：文字字串、數值、地理資訊）與主要預計用以呈現的方式為何（如：長條圖、網絡圖、地圖）。一般說來，廣為人知的資料分析與處理工具，如：Excel 試算表、SPSS 統計分析軟體，皆包含製作圖表的功能，可供製作常見的各類型圖表。在需求單純的情境中，這些資料分析與處理的軟體，不失為資料視覺化的簡便工具。

而若你打算製作某些特定形式的圖像時，亦可選擇專門針對此類型的軟體。例如：欲製作社會網絡圖時，往往就可直接選擇專門的軟體來作圖。而較常見的社會網絡圖工具，則有以 Microsoft Excel 插件方式使用的免費開源社會網絡分析軟體 NodeXL (<http://nodexl.codeplex.com/>)、UCInet、Gephi、Pajek 等等。

又如果，你想要選擇可支援多種呈現方式的工具，則可考慮專門視覺化的軟體或平臺，例如：Google 推出容易上手的 Google Data Studio、臺灣資料視覺化團隊自行開發的線上圖表製作平臺 PlotDB，以及本章第三、四節實作部分所欲介紹的資料視覺化軟體 —Tableau 等等。

Tableau 簡介

Tableau 是互動式的資料視覺化軟體，曾獲得諸多獎項，如：Best Overall in Data Visualization、Best of 2005 for Data Analysis、2008 Best Business Intelligence Solution。其主要產品包括：Tableau Public、Tableau Desktop、Tableau Online 和 Tableau Server 等。其中，Tableau Public 為免費版，功能較有限，且須仰賴雲端空間；Tableau Desktop 功能較完整，可離線操作，亦可將成果儲存在雲端。相形之下，若須自動更新資料則可選擇使用 Tableau Online 或 Tableau Server。

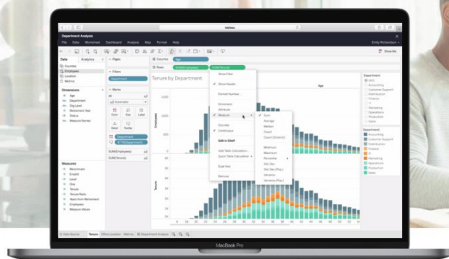
而當我們要選擇適合的 Tableau 版本時，可根據其可支援匯入的檔案格式、視覺化成果的分享方式，以及安全性等面向來衡量。下表綜整 Ryan Sleeper（2018）在《Practical Tableau》一書中的說明，可作為軟體版本選擇時的參考。

	Tableau Public	Tableau Desktop	Tableau Online	Tableau Server
資料匯入支援格式	Excel, text	Excel, text, Access, 統計檔、Tableau 檔	已發布在 Tableau Online 上且已授權之 workbooks	已發布在 Tableau Server 上且已授權之 workbooks
成果分享方式	雲端	離線或雲端	雲端 第三方 off-premise	自家就地部署 on-premise 或雲端
自動更新資料	無法	無法	可	可
安全性	仰賴雲端稍有風險，但可限制他人下載所發布至雲端的成果	視電腦及伺服器本身之安全性而定	視第三方安全性而定	視所在伺服器安全性而定

Changing the way you think about data

THE TABLEAU PLATFORM

SEE IT IN ACTION



Harness the power of your data. Unleash the potential of your people. Choose the analytics platform that disrupted the world of business intelligence. Choose Tableau.

圖 7-11 Tableau 官方網站 (<https://www.tableau.com/>)

而 Tableau 官方所提供的學習資源堪稱豐富，參考相關案例可見 Tableau Gallery (<https://public.tableau.com/en-us/s/gallery>) 或 Viz of the Day (https://public.tableau.com/en-us/s/gallery?qt-overview_gallery=1)。

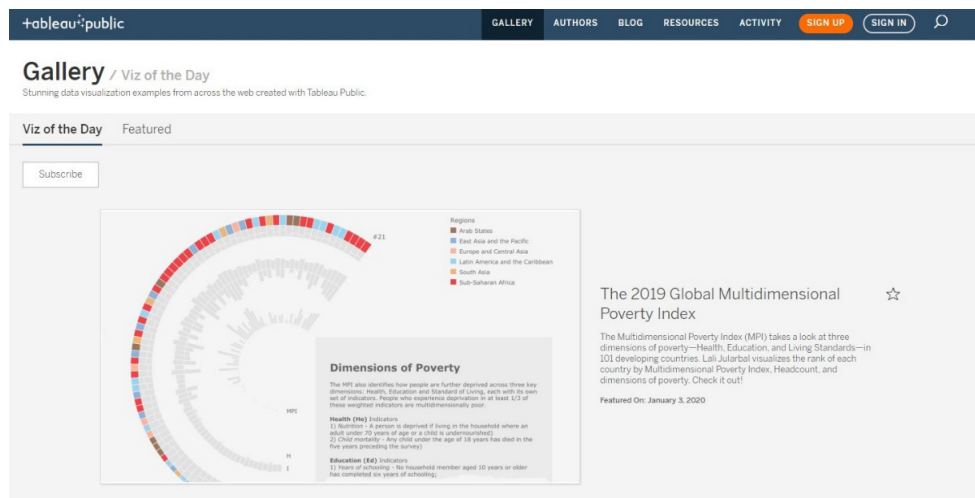


圖 7-12 Tableau Gallery 提供豐富的案例

Tableau 的官方網站也包含了許多學習資源。初學者可根據自己的需求，多加利用 Tableau Public 官方網站上的各項學習資源，透過學習影片（How-to Videos: <https://public.tableau.com/s/resources>）循序漸近地從零開始學習如何操作此軟體，運用官網提供的各種「範例資料」作練習；同時，參考官網整理的「社群資源」，掌握國際間資料視覺化相關社群所舉辦各種視覺化競賽或其他活動、Podcasts，以及社群媒體等。當然，Tableau 的官方網站亦提供軟體相關支援（[Tableau Help](#)），分別針對 Tableau 旗下不同產品作說明。此外，Tableau 還有其知識庫（<https://www.tableau.com/support/knowledgebase>）與論壇社群（<https://community.tableau.com/>），都是學習歷程中遭遇問題的好幫手。無論知識庫或論壇社群皆提供搜尋功能，可直接將目前操作上的問題以關鍵字進行搜尋，查找相關解決辦法，大多數的問題都可在此得到解答。只是對中文使用者而言，以中文查找通常可查得的結果可能不如英文豐富。雖然相信隨著 Tableau 的中文使用者成長，中文資源也將日漸豐富，以目前的情況來看，建議若以中文關鍵字難以獲得解答，可嘗試改以英文關鍵字搜尋；閱讀英文不大費力的話，則可直接以英文關鍵字進行搜尋。

最後，除了運用免費的 Tableau Public 或購買 Tableau Desktop 等版本外，Tableau 亦針對學校師生提供免費授權方式。教師若開設相關課程，於開設該學期可透過 Tableau for Teaching，填妥學校單位以及課程相關資料（<https://www.tableau.com/academic/teaching>），申請教學授權版的 Tableau Desktop；學生在學期間亦可透過 Tableau for Students 檢附在學證明等資料（<https://www.tableau.com/academic/students>），申請為期一年的學生版。本章相關介紹係以 2019 年下半年至 2020 年初筆者掌握的情況來說明。

資料分析與視覺化的基本步驟

一般說來，研究資料視覺化的流程涉及研究資料的整個處理過程，包括：資料蒐集、資料清理與準備、資料分析（包含統計檢定），以及資料呈現（即視覺化）。其中，資料分析與資料呈現有時可來回循環地進行，尤其若你處理的議題或研究愈具探索性質，愈容易在二階段間往返徘徊。因為隨著資料視覺化的探索，可能讓你對資料分析產生更多新的想法；隨著資料分析的深化，可能也讓你對資料視覺化產生更多新的想法。



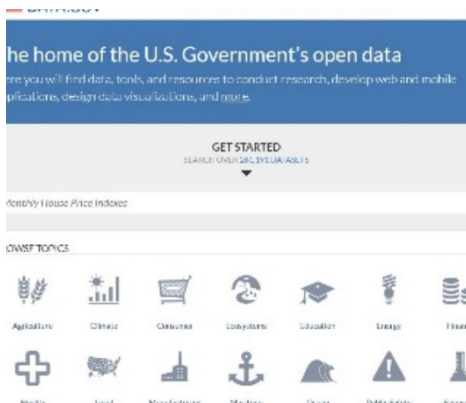
臺北市資料大平臺

<https://data.taipei/>



政府開放資料平臺

<https://data.gov.tw/>



美國政府開放資料平臺

<https://www.data.gov/>

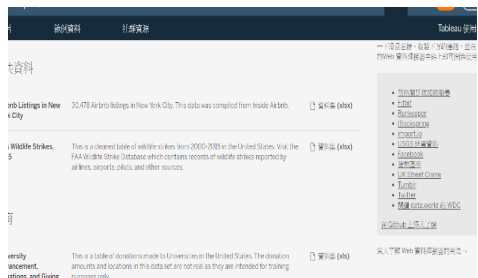


Tableau 範例資料集

<https://public.tableau.com/s/resources>

圖 7-13 政府開放資料與 Tableau 範例資料來源

就資料蒐集而言，研究資料的來源可能是既有的資料，像是政府開放資料或相關軟體的範例資料（如：Tableau 範例資料集）等等（圖 7-13），也可能是自行蒐集（如：問卷調查所得結果）。許多時候，由於單一的資料集未必能滿足當前所欲探討問題的資料需求，因而必須設法蒐集並整併來自不同資料來源的資料，以解答目前所欲探究的問題。例如：當你想製作一張人口密度與自行車竊案的關係圖，以了解人口密度高之處是否較易發生竊案時，可能需要從不同政府部門的統計資料中，取得人口密度與自行車竊案數等資料。然而，不同來源的資料無法任意整併（例如：不同範圍或不同方法的調查無法直接合併比較），如欲嚴謹地蒐集資料，多半須要具備研究方法相關知識與能力，以利判斷如何結合不同來源的資料，以探討相關議題。

完成資料的蒐集後，就進入資料視覺化品質決勝的關鍵階段了。資料清理與準備是資料視覺化非常重要的一環，也是相當費心與費時的階段。資料清理工作做得愈嚴謹，資料的品質也愈理想。而資料分析則是影響資料視覺化深度與廣度的重要環節，倘若能以適合的資料分析方式處理資料，則可強化討論的深度與廣度。然而，資料分析涉及各種學科知識背景與研究方法的知能，且各種不同的資料分析方式皆各有其學問，顯然已超出本章範圍，故對資料分析感興趣的讀者可再鑽研自己感興趣領域的研究方法與資料分析相關書籍。

Information is Beautiful 官方網站指出，繪製資訊圖表的過程當中，大約 80% 的時間都在蒐集與整理資料，只有剩下 20% 的時間花費在規劃與實際將資料轉化為視覺呈現。因此，以下針對資料清理進一步說明。

資料清理

所謂「資料清理」是指將原始資料轉換為可用於資料分析 / 統計分析 / 視覺化的形式，也就是將從資料來源而來而無法直接分析的原始資料（Raw Data）轉化為資料可直接用於分析的乾淨資料（Clean/Tidy Data）之過程。一般說來，可分析的資料結構是以每個變項（variable）為一欄、每個觀察值（observation）為一列的方式呈現的，許多時候可能需要合併、分割等處理（例如：以識別碼合併多個表格，將相同的項目合併、將原本列在「其他」中的項目歸類）。由於資料清理的過程繁複，且如欲改變清理資料的策略，則可能需要退回先前的步驟，故詳實地記錄每個資料清理的步驟是非常重要的工作，在團隊合作中尤然。

若未適當清理資料，則資料分析 / 視覺化的結果無法減輕讀者認知負荷，或無法有效呈現重要發現，甚至可能混淆視聽。圖 7-14 呈現了學生就讀學系分布概況，然而，該圖即為未適當清理資料即將資料視覺化之實例。在資料清理過程中，應先將相同的項目進行合併，才能看出不同學系的分布情形。此圖在製作前並未清理資料，故相同學系的人數散落於不同項目中，讀者須自行將這些相同的項目加總，才能得知特定學系的人數。也就是說，清理資料時，應先合併相同的類目，即合併「資訊管理」、「資管」與「資訊管理學系」、合併「資工」、「資訊工程」、「資工系」與「資訊工程學系」，以及合併「圖書資訊」、「圖書資訊學系」與「圖資」，才能確實呈現出資訊管理學系、資訊工程學系，以及圖書資訊學系的人數分布概況。

學生就讀學系分布概況

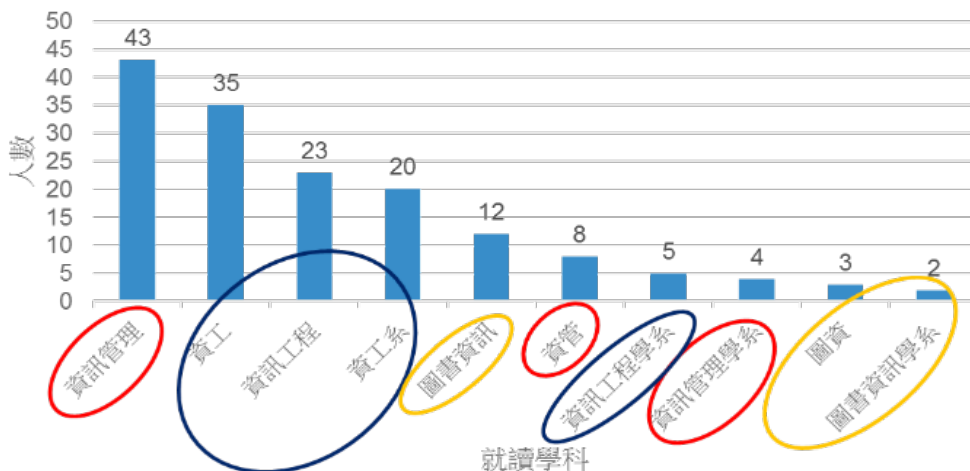


圖 7-14 未適當清理資料則相同的項目散落各項，難以有效溝通

而常見的資料清理工作可能包括：處理遺漏值、檢查拼字與更正錯漏字等細節、檢查並更正資料格式（如：統一日期、電話號碼等格式）、核對範圍（如：確認年齡須為正值、成績須落在一定範圍區間內）、重整變項中之類別（如：將列在「其他」的項目歸入既有類別或另行歸類），以及合併或新增變項等。而 Information is Beautiful 網站也提供了清理後的資料範例，可供參考：<http://www.informationisbeautiful.net/data/>。

小結

資料視覺化工具不勝枚舉，充分掌握資料特性與資料分析與呈現的主要目標，才能幫助自己選擇適合的視覺化工具。而資料視覺化的主要步驟包括：資料蒐集、資料清理與準備、資料分析，以及資料呈現。其中，資料呈現之前的各項準備工作相當費時費力，約佔整個資料視覺化歷程所花費時間的 80%。資料蒐集時，可善用各種政府開放資料，而資料清理的工作務必要嚴謹以對，才能維護資料品質，做出有意義且有效的視覺化成果。

7-3 以政府開放統計資料談數據資料視覺化

本節將帶領 Tableau 新手一步步學習如何運用此軟體進行資料視覺化，先從 Tableau 的初始介面開始談起，並以政府開放統計資料為例，說明如何使用 Tableau 簡易製圖，再進一步介紹 Tableau 的函式語法。

Tableau 初始介面

當你進入 Tableau 軟體，首先，將看到圖 7-15 之初始介面。它與一般人常用的微軟文書處理軟體的初始介面不大相同，然而，其初始介面相當單純，主要分為資料連結、開啟檔案與探索資源三大區塊。

其中，最左上角帶有 Tableau logo 的小花圖示是快速切換資料與視覺化工作表的快速切換鍵，待資料匯入後，如果退回此介面，可直接點擊此快速鍵。而左半部的主要區塊則提供連結與匯入資料的重要功能，可直接點選所欲匯入的資料檔案類型，即可找到資料位置並匯入 Tableau 中。例如：對新手而言，最常見的就是連結 Excel 試算表，故可直接點選 Microsoft Excel 即可匯入資料。

乾淨俐落的中間區塊提供了一項主要功能 — 開啟舊檔，當你開始使用此軟體之後，近期使用的檔案亦將出現在此區塊，直接點擊欲開啟的檔案即可。中間下半部則可開啟 Tableau 提供的範例檔案。至於右側探索區塊則可讓使用者連結到 Tableau 官方網站的各項教學資源，包括：教學影片、每日精選作品等。

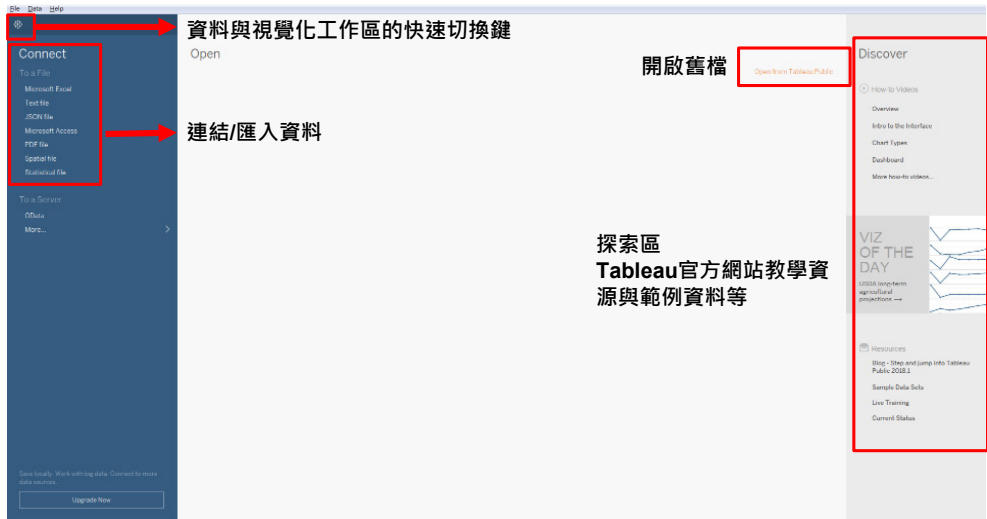


圖 7-15 Tableau 初始介面

Tableau 工作介面

Tableau 工作介面除了與一般軟體相同、置於頂端的工具列之外，大致可分為四區，由左至右，依序包括：資料欄位變項與資料分析區、視覺設定區、工作區，以及圖形選擇區（圖 7-16）。當資料匯入之後，最左側將出現 匯入的欄位變項，我們即可利用這些欄位變項來作圖。而當我們將變項置於中間偏上的欄與行（列）時，Tableau 就會自動幫我們繪製圖形，而此圖形將出現在圖中空白的工作區。透過最右側的圖形選擇及中間偏左的視覺設定區，我們可進一步調整圖形樣式，並根據當前的需求，進一步將 Tableau 自動產生的圖像客製化。

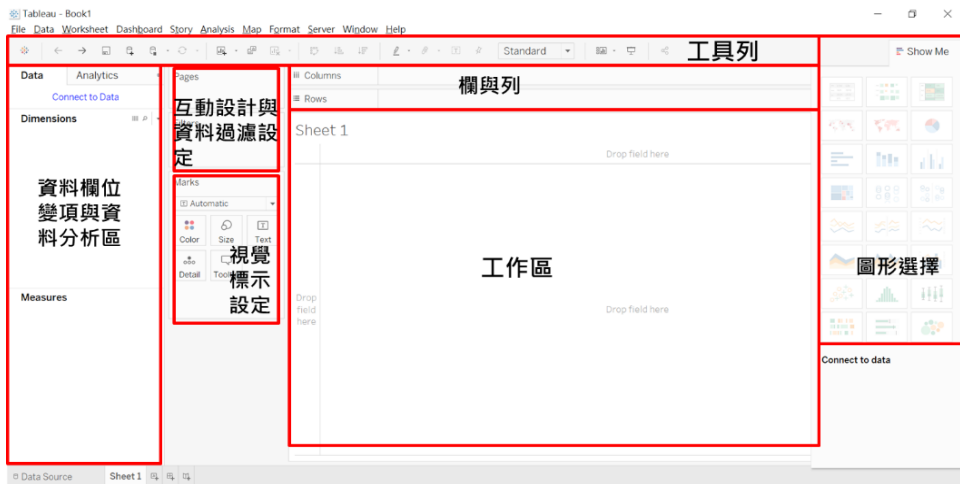


圖 7-16 Tableau 工作介面

Tableau 的基本操作流程

由前述介面說明，我們其實已大致掌握了 Tableau 的基本操作流程了。一般說來，簡易的製圖不外乎進入初始畫面並連結 / 匯入資料，並確保資料型態設定正確（例如：連續變項設為數值、類別變項設為字串、地理資訊設為地理角色）。爾後，再進入工作分頁，準備開始進行視覺化的工作。

在準備資料的過程中，有時也將視需求，運用 Tableau 函式語法新增所需欄位變項等（例如：若資料僅包含學生的單科成績或每學期的成績，可運用公式計算出總平均）。接下來，當所欲用以製圖的欄位變項都準備妥當，就可開始繪圖了！

繪圖時，首先要選擇欲用以繪圖之欄位變項，將之拖曳到欄與行（列）。此時，Tableau 會自動根據所選的欄位變項類型與屬性來判定要幫你繪製什麼圖形。因此，若未確保欄位變項類型與屬性設定正確，Tableau 將難以畫出我們預想的圖。如果對 Tableau 自動繪製的圖形不滿意，則可透過圖形選擇區，直接點選合適的圖形。接著，即可開始進行客製化調整，增加標籤說明、修改顏色等等。最後，再將視覺化成果匯出，就大功告成了。

以下透過實作練習，讓大家可以了解前述流程的確切進行方式並掌握 Tableau 之各項基本功能。

Tableau 實作技巧

接下來，我們將運用美國教育統計開放資料 Integrated Postsecondary Education Data System (IPEDS)，介紹 Tableau 的基本功能。

1. 匯入資料

首先，請進入 Tableau 初始介面，透過連結 Microsoft Excel 的功能，開啟你存放本節欲用來練習的 IPEDS 課堂練習檔案的位置。由於此試算表檔案包含四個分頁，我們僅需要包含實際資料的 Data 分頁，因此，只要以滑鼠左鍵點選欲用來作圖的 Data 分頁不放，並拖曳至右側資料區上半部的空白處，再放開滑鼠，即可匯入該分頁之資料（如圖 7-17）。若資料檔案較大，得稍候片刻，待資料區的下半部出現資料表及其內容，確保所有資料匯入完成。

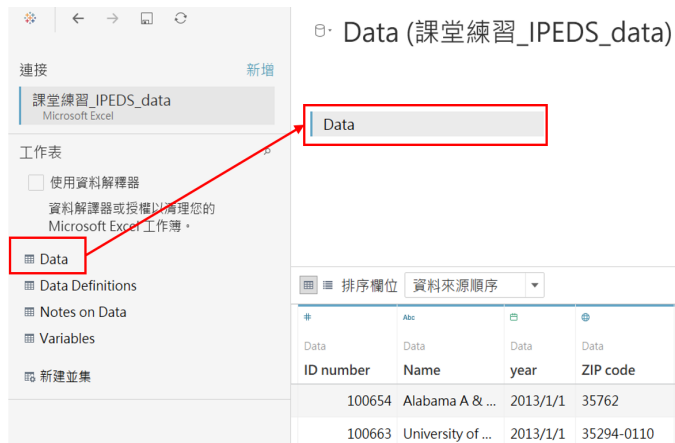


圖 7-17 將檔案匯入 Tableau

此外，雖然本節僅介紹 Tableau 的基本功能，但在此順帶一提，Tableau 也提供同時匯入不同分頁的功能，因此，如果我們想要一次匯入同一試算表的多個分頁，則可將不同的分頁循前述步驟，分別拖曳至右側資料區空白處，並選擇一特定欄位（如：ID）將不同分頁的資料加以連結合併。

匯入資料的過程中，Tableau 會自動判定每個欄位變項的型態與屬性，在每個欄位變項上方，以藍色和綠色分別標示該欄位屬於維度（dimensions）或量值（measures），並以 Abc、# 等符號標示該欄位變項是文字字串或數字等（圖 7-18）。其中，設定成量值的欄位變項可用於一般運算，而設成日期者可以年、月、日等單位處理；設定地理角色者可用以繪製地圖。

此時，我們要確保每個欄位的資料型態設定正確（如：類別變項設為字串、連續變項設為數字），以利 Tableau 自動作圖時，為我們選擇合適的呈現方式。

☰ 排序欄位 資料來源順序 ▼

#	Abc	☰	☰	Abc	☰
Data	Data	Data	Data	Data	Data
ID number	Name	year	ZIP code	Highest de...	County name
100654	Alabama A & ...	2013/1/1	35762	Doctor's degr...	Madison Cou...
100663	University of ...	2013/1/1	35294-0110	Doctor's degr...	Jefferson Cou...
100690	Amridge Univ...	2013/1/1	36117-3553	Doctor's degr...	Montgomery ...
100706	University of ...	2013/1/1	35899	Doctor's degr...	Madison Cou...

圖 7-18 確保資料型態設定正確

2. 簡易製圖：以並列長條圖的製作為例

最簡單的製圖方式，就是直接在連結檔資料後，進入工作表（worksheet），將欲用來繪圖的欄位變項拖曳至工作區的欄與行（列），再選擇圖形呈現方式（如：簡單長條圖、並列的長條圖、泡泡圖）。

一般說來，我們在做製圖的規劃時，可將 Tableau 自動判定成維度的類別變項設想成自變項，將 Tableau 自動判定成量值的連續變項設想成依變項。在這個練習中，我們若想繪製一張並列的長條圖，藉此比較不同學校類型的全職與兼職學生註冊人數，則可直接將欲用以繪圖的欄位變項拖曳至欄與行（列）。此例可將學校分類 [Carnegie Classification] 拖曳至行（列）（Columns）、全職學生註冊人數 [Full-time enrollment] 與兼職學生註冊人數 [Part-time enrollment] 拖曳至欄（Rows），如圖 7-19。

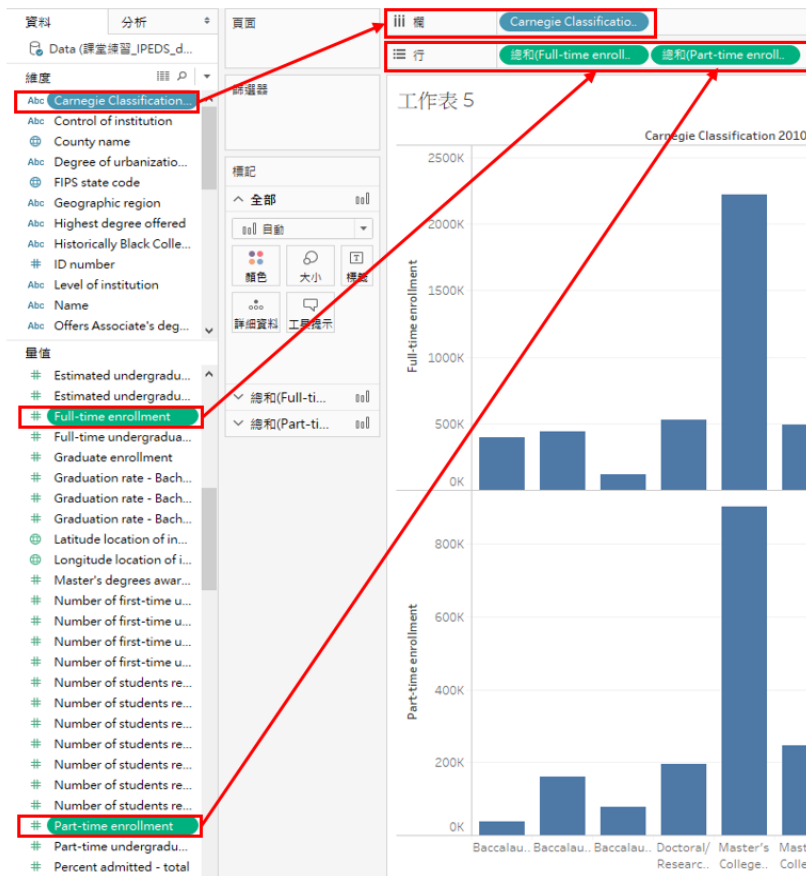



圖 7-19 將欄位變項拖曳至欄與行（列）

接著，透過最右側的圖書選擇功能「顯示」(Show Me)，選擇合適的視覺化呈現方式。此例選擇的是「並列的長條圖」。如此即可初步完成一張並列的長條圖。

3. 簡易客製化

3.1 修改顏色

如果你想進一步作些客製化調整，最常見的可能是修改顏色。在 Tableau 中，只要在標記區點擊「顏色」、「編輯顏色」，即可針對長條圖修改顏色，亦可在此標記區調整長條圖的不透明度與邊界等。

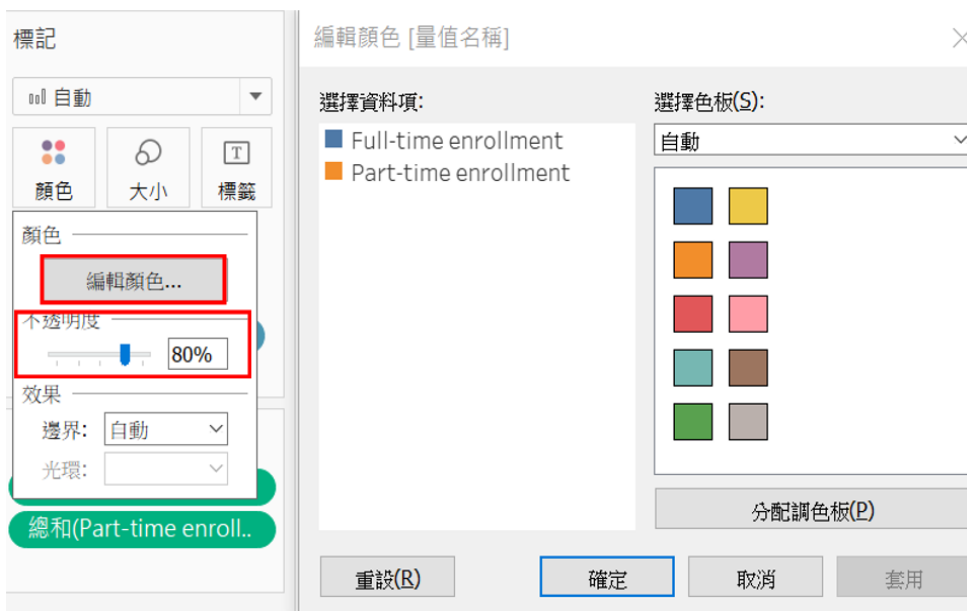


圖 7-20 修改顏色

3.2 新增標籤

若你想要為長條圖增加標籤（labels），以說明每個數值（此例是全職與兼職學生的註冊人數），則可在標記區點擊「標籤」，再勾選「顯示標記標籤」即可（如圖 7-21）。

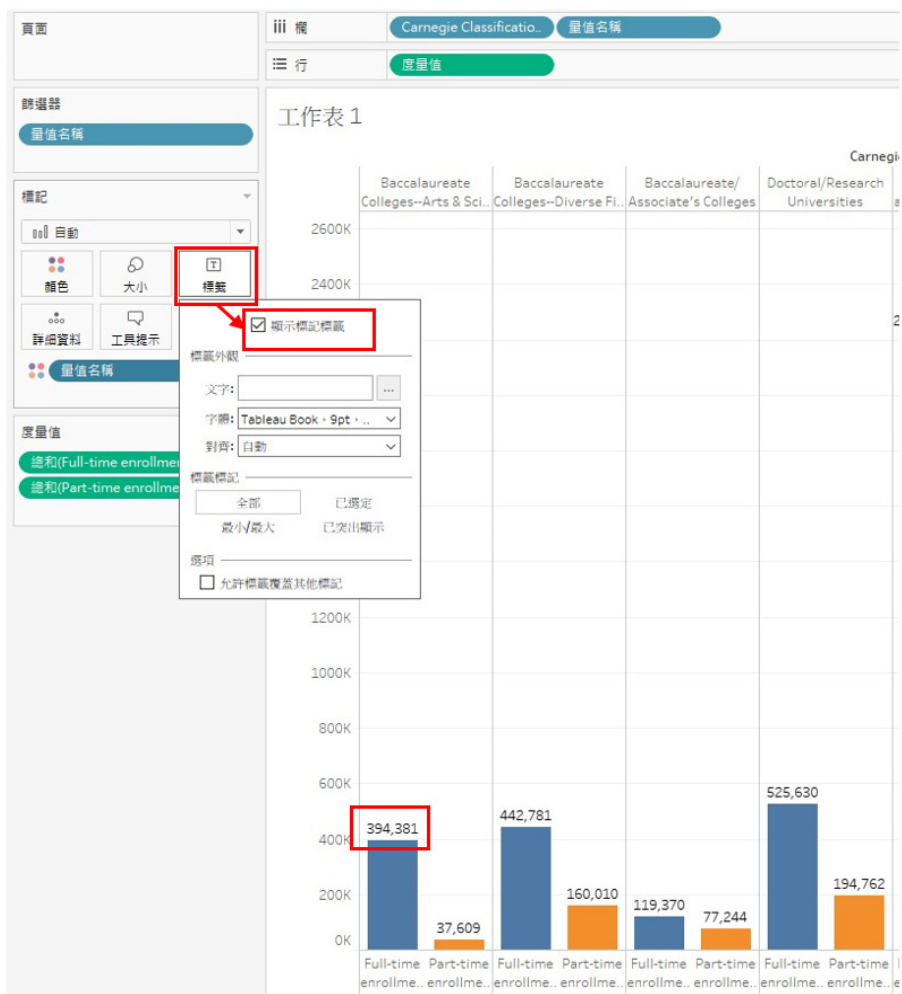


圖 7-21 增加標籤說明

3.3 過濾篩選

此外，Tableau 還提供了過濾篩選部分資料的重要功能。就此例而言，大學分類項目略嫌太多，可能看了眼花撩亂，如僅欲呈現部分大學類型的資料，則可點選大學分類之變項（[Carnegie Classification]）或在工作區右側的空白處按右鍵，選擇「篩選器」，並選擇欲用來篩選之欄位變項（此處為 [Carnegie Classification]）。接下來，再根據你所要選擇的項目作勾選即可（圖 7-22）。

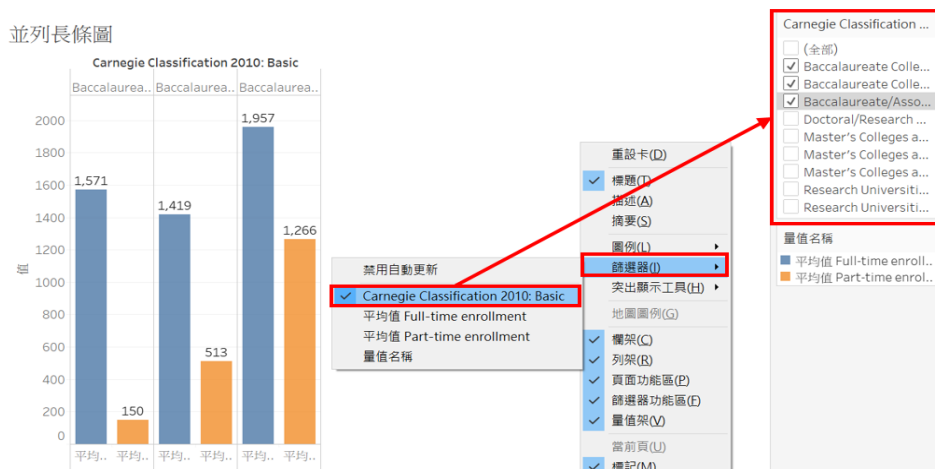


圖 7-22 過濾篩選部分資料

3.4 調整欄位變項的計算方式

最後，Tableau 還提供了一項可直接在視覺呈現中調整變項計算方式之強大功能。就本例而言，我們目前所製作的長條圖是以 Tableau 的預設值——總和來呈現的，然而，若我們認為應以平均數來呈現，方可看出各類型學校平均的招生人數情形，則可直接點選已位於度量值區塊中的欄位變項右緣的小三角箭頭，將其度量中的總和改成平均值（圖 7-23）。而在此並列長條圖中，用來對照比較的兩個欄位變項（即全職學生註冊人數 [Full-time enrollment] 與兼職學生註冊人數 [Part-time enrollment]），二者須修改成相同的計算方式，其對照比較才有意義。所以，此處二者的計算方式須一併由總和改成平均值。

在此順帶提醒，製圖時，須留意 Tableau 預設之數值呈現方式，若其預設值不合理或不合意，則可隨時將其計算方式調整為總和、平均數、中位數、最大值、最小值等。這是一項相當方便而實用的功能。

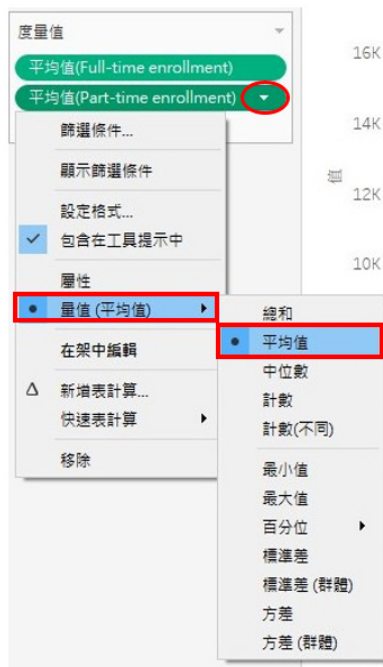


圖 7-23 調整欄位變項的計算方式

4. 管理技巧與匯出視覺化成果

4.1 檢視原始資料

在 Tableau 中，如欲檢視原始資料，可直接退回「資料來源」之頁籤，亦可由工具列的「資料」、「資料 (資料檔案名稱)」、「檢視資料」來對照匯入的資料內容（圖 7-24）。直接退回「資料來源」頁籤固然方便，但有時欲同時檢視視覺呈現與資料，則可運用後者方法來檢視資料。

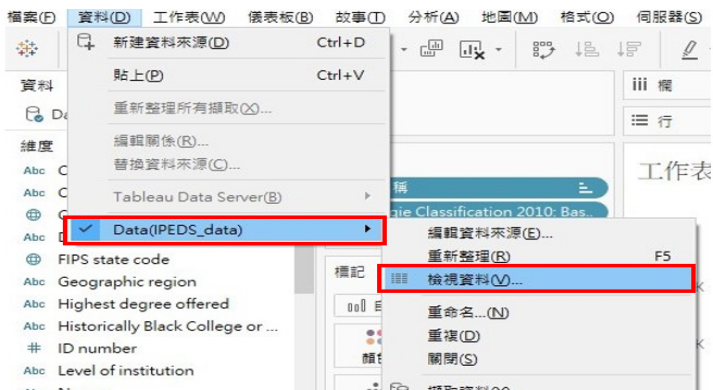


圖 7-24 檢視資料

4.2 管理分頁

Tableau 的工作頁籤分為三種形式：工作表、儀表板與故事。最基本的視覺呈現工作皆是由工作表展開。一般皆在工作表製圖，待所有圖片製作完成後，再視需求，將完成的工作表拼貼整併至儀表板，或將圖片以連續性的故事形式呈現。若要管理分頁，最簡便的方式就是直接於分頁標籤上按右鍵，即可「新建工作表」、「新建儀表板」、「新建故事」、「重複」（複製整個工作表）或「刪除」當前的分頁；如欲修改當前的分頁名稱，則可透過「重新命名」作調整（圖 7-25）。



圖 7-25 管理工作表

倘若你要保留當前的工作表，但想清除工作表中的內容，則可點選工具列的「刪除圖示」，即可透過選擇「清除工作表」來清除內容（圖 7-26）。Tableau 不會再次詢問你是否確定要刪除等，因此，操作類似的動作，須格外謹慎。萬一不慎清除工作表，可盡快點選工具列的復原，或用 **Ctrl+z** 來退回至上一步。



圖 7-26 一鍵清除工作表內容

此外，使用者亦可將 Tableau 的分頁標上不同顏色，加以分類管理。當你在一份檔案中的分頁較多時，不失為一項方便的管理功能。與前述管理分頁的作法相似，只要在分頁標籤上按右鍵，即可將分頁標示成不同顏色（Highlight worksheet），以便管理（圖 7-27）。



圖 7-27 將工作表標上顏色，以便管理

4.3 儲存、分享與匯出工作簿

最後，當你完成視覺化任務後，往往須要存檔或分享。在 **Tableau** 中，儲存整個工作簿就如同一般軟體，透過「檔案」、「儲存」或「另存新檔」即可儲存工作簿。然而，直接儲存下來的工作簿，只能透過 **Tableau** 軟體讀取，如果你希望將工作簿分享給非 **Tableau** 使用者，則須透過「檔案」、「匯出已封裝工作簿（Export Packaged Workbook）」，將檔案存成封裝的工作簿。

而當你希望將特定的工作表匯出，通常可透過以下兩種方式將工作表匯出為 **PDF** 檔或圖檔。不過，若欲匯出 **PDF** 檔，得從「檔案」、「列印為 **PDF**」來執行；若僅欲儲存圖檔，則須從「工作表」、「導出」、「影像」來執行。

4.3.1 匯出 **PDF**

從「檔案」和「列印為 **PDF**」來匯出 **PDF** 檔時，可選擇欲匯出的範圍是整個工作簿（所有的分頁）抑或目前使用中的工作表，亦可指定紙張尺寸（圖 7-28）。若你所作的圖片尺寸大於列表機預設紙張尺寸的大小（如：超出 **A4** 或 **letter size** 的大小），則可考慮將紙張設定為「未指定」。因為當工作表中的圖片超過紙張尺寸時，匯出的 **PDF** 檔會自動將圖片切割成多頁，如此一來，就難以檢視或列印完整的圖片了。因此，許多時候，可考慮直接將紙張尺寸設定為「未指定」，待實際有列印需求時，再透過讀取 **PDF** 的軟體來設定縮放比例即可。

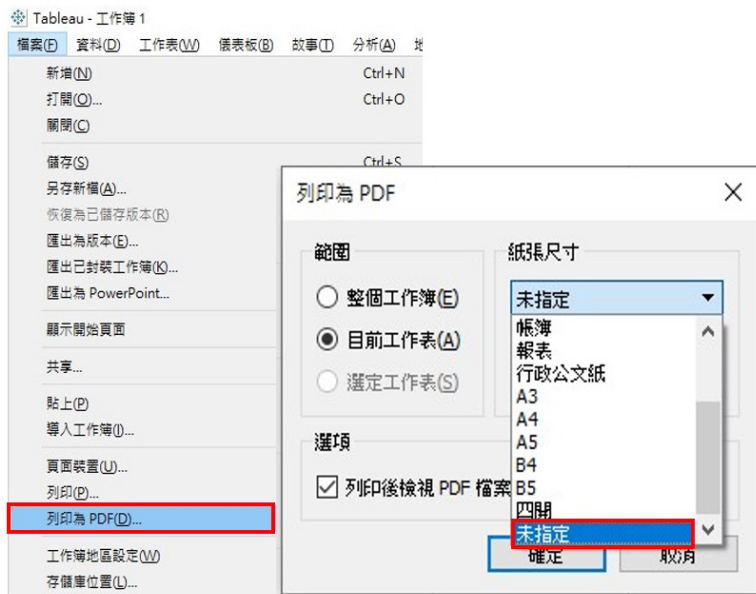


圖 7-28 匯出 PDF 檔

4.3.2 匯出圖檔

在 Tableau 中，匯出圖檔與其他儲存分享功能不同，並非透過「檔案」功能來執行，而須從「工作表」、「導出」、「影像」來執行。這點可能令不熟悉 Tableau 的人感到有些困惑。因為 Tableau 將主要的視覺呈現工作分為工作表、儀表板和故事三大類，所有的工作皆以此為據，所以，當我們只針對其中一個工作表的內容進行任務時，必須直接透過工具列的「工作表」來操作。

圖 7-29 呈現了匯出圖檔的步驟。當你選擇導出影像後，匯出圖像的視窗將會跳出，供你選擇匯出的圖像是否包含標題與描述。最後，在存檔時，可選擇圖片的檔案格式（.jpg、.png、.bmp 或 .emf）。

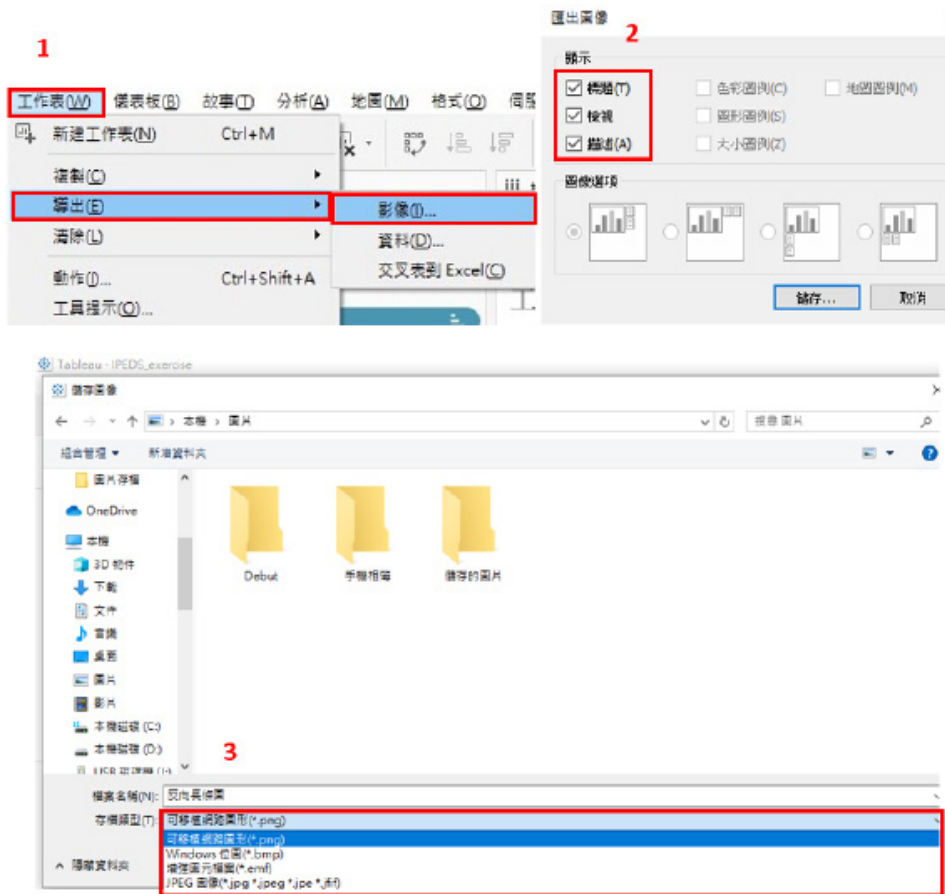


圖 7-29 匯出圖檔

反向長條圖的製作

接下來，我們將在同一個檔案的新分頁中，運用前述 IPEDS 教育統計資料，繪製不同類型大學收受學生數與註冊人數之長條圖，藉此說明反向長條圖的製作方式。

1. 簡易製圖

先前製作並列長條圖時，我們用的學校分類是美國教育統計常用的卡內基高等教育分類 [Carnegie Classification]；現在我們將運用最高授予學位 [Highest

degree offered] 之欄位變項來探討學校類型。當然，你也可以嘗試運用 IPEDS 檔案中的其他有趣變項來製圖。

首先，將收受人數 [Admission total] 和註冊人數 [Enrolled total] 拖曳至欄，並將最高授予學位 [Highest degree offered] 拖曳至行（列）（圖 7-30 步驟 1）。如此，Tableau 已為我們自動產生兩個並排的長條圖。

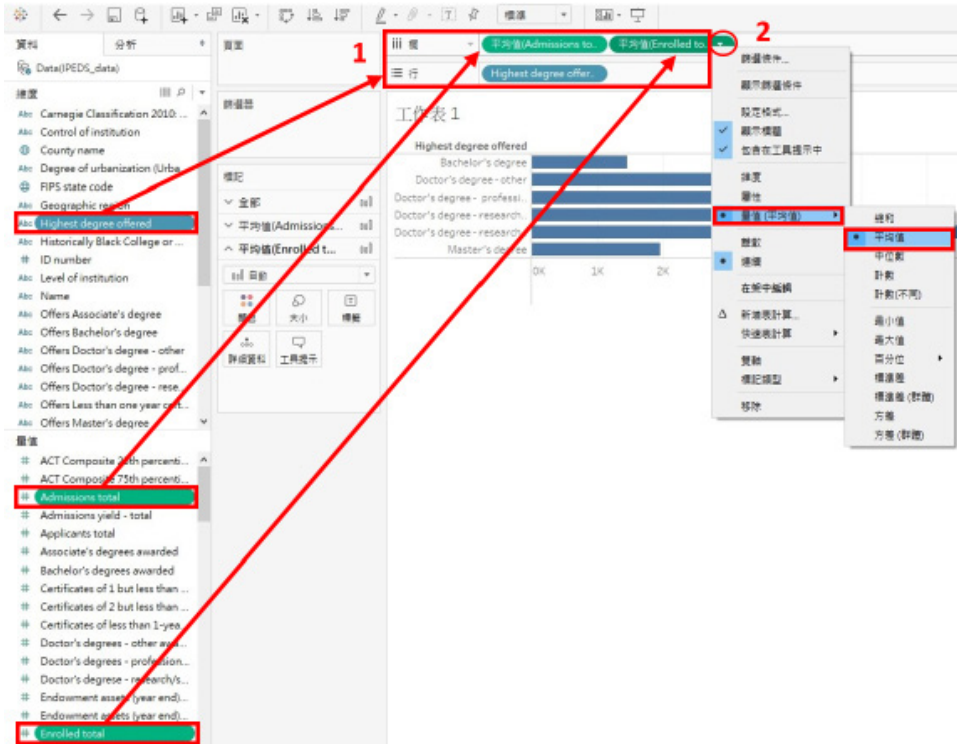


圖 7-30 繪製長條圖

2. 調整欄位變項的計算方式

Tableau 同樣預設以收受人數及註冊人數的總和來繪圖，但我們希望以平均數來呈現不同類型學校的人數情形。此時，我們可運用先前介紹並列長條圖製作的方法，先對此長條圖進行資料呈現上的調整，將收受人數 [Admission total] 和註冊人數 [Enrolled total] 改以平均數呈現（圖 7-30 步驟 2）。

3. 反向軸之調整

接著，我們欲將左側長條圖轉向，讓兩個並排的長條圖分別朝左右反向展開，因而須調整左圖軸的方向。欲調整軸的方向，可直接在左半圖之軸上按右鍵，並選擇「編輯軸（Edit Axis）」（圖 7-31）。此時，編輯軸的視窗將跳出。我們只要在編輯軸視窗中，勾選「倒序」即可。如此一來，左側軸向以倒序（即反向）排列，即可快速完成一反向長條圖。



圖 7-31 設定反向軸

4. 新增參照標準與排序

Tableau 在資料分析方面提供了簡單易用的功能，方便我們為圖像加上參照標準（如：常數線、平均線、四分位差線）。倘若我們欲在剛才繪製的反向長條圖中加上四分位差線作為參照，則可在資料區選擇分析（Analytics）頁籤，再點選欲加到圖中的參照標準，將之拖曳至圖表工作區。

在我們拖曳的過程中，滑鼠移動到工作區時，會出現圖 7-32 之方框，供我們選擇參照樣式，即參照標準將繪製到整個圖表、區段，或儲存格上。此處，我們可選擇「表」，讓四分位差繪製到整個長條圖上。

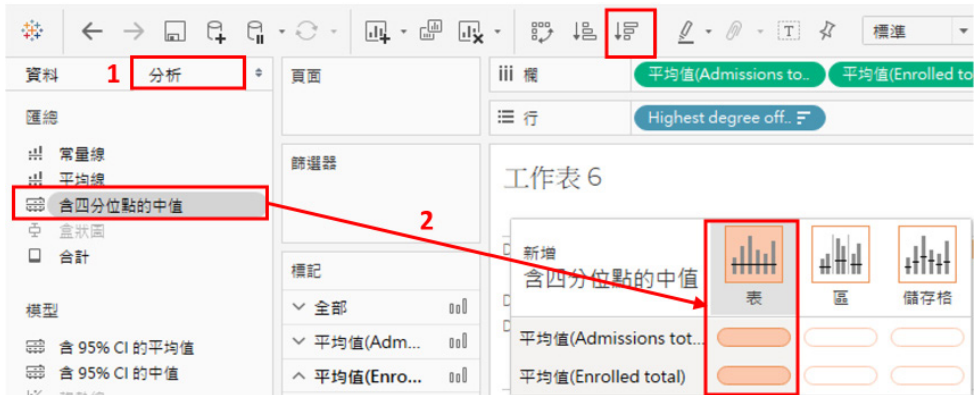


圖 7-32 新增參照標準與排序

此外，我們亦可利用工具列的排序鍵（圖 7-32），進一步針對長條圖進行排序。完成後，即可得到排序整齊並加上參照標準的反向長條圖了（圖 7-33）。

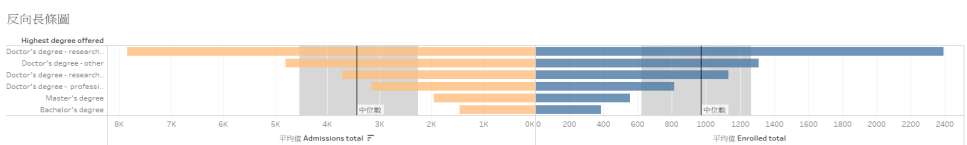


圖 7-33 初步完成反向長條圖

5. 調整尺規刻度

最後，由於此圖左右的尺規刻度不同，可能造成誤導，也難以快速比較左右兩側的數值多寡，因此，我們將進一步將兩軸之尺度調整成相同的，以利比較。作法上，我們可先在數值較大的一軸上透過右鍵點選「編輯軸」，檢視其最大值為何，並複製此數值，以便將另一側的軸也設定為相同的數值。

此例中，我們先在左側軸上，透過右鍵「編輯軸」的功能，檢視其範圍之最大值為何，複製此數值後，再到右側軸上，同樣透過右鍵「編輯軸」的功能，將此數值貼至其範圍的最大值。由於預設的範圍是自動產生的，故我們要將之改為「固定」，並將前述最大值複製到固定結束的空格中（圖 7-34）。



圖 7-34 調整尺度範圍

如此，則可得圖 7-35 之反向長條圖，以便了解不同類型學校的收受人數與註冊人數之落差。

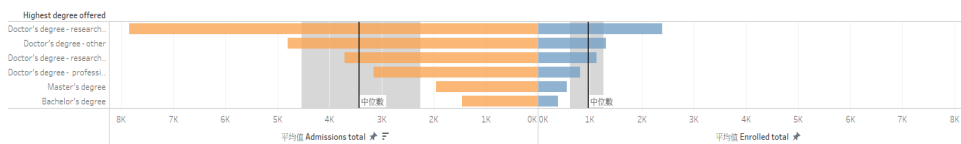


圖 7-35 調整尺度範圍後的反向長條圖

Tableau 函式語法簡介

雖然 Tableau 提供了許多簡單易用的功能，若能運用其函式語法，可大大提升其功能性。以下將簡介 Tableau 函式語法的運作：

在 Tableau 中，函式語法可進行簡單的四則運算（例如：計算總和、平均數），亦可用來取代部分字串、計算日期，或依條件計數等等。無論撰寫何種函式語法，皆是利用新增欄位變項來操作，因此，接下來將先介紹「建立計算欄位」之功能。

我們可以透過點選資料區中，維度或量值右方的的小箭頭，並選取建立計算欄位（Create Calculated Field），亦可透過工具列的「分析」、建立計算欄位來新增變項（圖 7-36）。

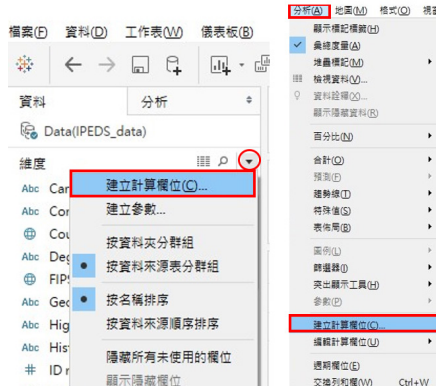


圖 7-36 建立計算欄位的兩種常見方式

接下來，就可為此欄位變項命名，並寫出計算公式或函式語法了。以前述 IPEDS 檔案為例，我們可先試著透過簡單的運算來練習運用此功能。若將註冊總人數 [Enrolled total] 之八成來估算理想的留校率（ideal retention rate），我們可在欄位變項名稱中，輸入「ideal retention rate」，並在計算區寫上算式，以註冊人數乘以 0.8。在 Tableau 中，欄位變項須以方括號 [] 表示，四則運算則可直接以加（+）、減（-）、乘（*）、除（/）表示。故此例中的算式為：[Enrolled total]*0.8（圖 7-37）。按下「確定」，即可新增一計算欄位。

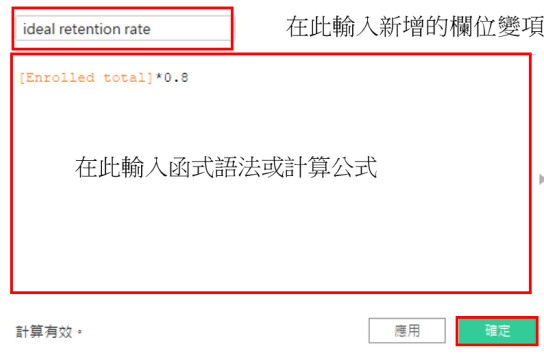


圖 7-37 建立計算欄位

Tableau 函式語法基本概念

如前所述，在 Tableau 中，欄位變項以方括號表示，例如：[年齡]；當表述值包含字串時，通常須加雙引號（""），特定表述值加單引號（' '），視所選函式而定（參見後段函式語法說明）。

7

除了可如前述直接列出算式，針對欄位變項進行四則運算外（例如：[國語]+[數學]+[英文]），亦可利用各種類型的函式語法進行其他運算。簡單的函式語法表述方式多為函數名稱加括號，括號內則須放入變項或條件等，多數情況可與四則運算搭配使用。例如：SUM([人數])、SUM([人數])/COUNTD([ID]) 等等。以下將分別簡介 Tableau 主要的函式類型：

1. 常用數字函式語法簡介

數字函式語法可供回傳絕對值、最大值、最小值等，以下分別說明用以回傳絕對值的 ABS、用來取整數值的 ROUND，以及回傳最大值和最小值的 MAX 和 MIN。

語法說明：ABS(number) 或 ABS([欄位變項])

回傳絕對值，可用來回傳特定數值或特定變項之絕對值
例如：ABS(-2) = 2 或 ABS([歷時變化]) = [變化幅度]

ABS 函數可用於數值或欄位變項。舉例而言，上述函式 ABS([歷時變化]) 可用在了解歷時變化的幅度。也就是說，當我們想要了解某項產品的銷售金額或學生成績的變化幅度，但不在意其變化是增是減時，可利用「歷時變化」之欄位變項，針對每項產品，計算其銷售金額在某段時間的變化（即成長幾元或減少幾元），或針對每位學生，計算其成績的變化（即進步幾分或退步幾分），再運用 ABS 來回傳此歷時變化正、負值的絕對值，以得到其變化幅度。

語法說明：ROUND(number, [decimals])

回傳捨去至指定的小數位數或整數值。
例如：ROUND([特定變項])

ROUND 函數用於取整數或取特定位數的小數，若將其函式省略 [decimals]，則表示取整數。舉例而言，若我們希望將學期成績取至整數位，則可透過 ROUND ([學期成績]) 來捨棄整數之後的小數。

語法說明：MAX (number, number) 與 MIN (number, number)

回傳較大或較小值。

例如：MAX (4, 7) MIN ([變項 A], [變項 B])

而 MAX 與 MIN 函數則是用來回傳較大或較小值。其中，MAX 與 MIN 只能比較兩個相同類型的變項，如：兩者皆為數值。上述實例中，MAX (4, 7) 將回傳 7。而假設我們欲找出班上學生的國語和英文成績中較低的成績時，可利用 MIN ([國語], [英文]) 的函式語法，找出國語和英文兩欄位中較低的成績。

2. 常用字串函式語法簡介

字串函式可供找出是否包含特定用字、取代部分文字等，有助批次處理文字字串資料。以下分別說明可用來找出是否包含特定用字的 CONTAINS，以及用來取代部分文字的 REPLACE。

語法說明：CONTAINS (string, substring)

若字串 string 中包含 substring，則回傳 true

例如：CONTAINS ('Librarians', 'Lib') = true

上述實例中，CONTAINS 語法協助我們找出 Librarians 一字中，是否包含 Lib，得到結果是 true，表示 Librarians 一字包含 Lib。此例看似簡單，但當我們的整份資料包含大量的文字字串時，即可發揮此語法之效用。

語法說明：REPLACE (string, substring, replacement)

在 string 中查找 substring，若查得，則將其取代為 replacement；若未查得，則加以保留

例如：REPLACE ('1st Edition', '1st', '2nd') = '2nd Edition'

上述實例中，REPLACE 函數可幫助我們確認 1st Edition 中是否包含 1st，若包含 1st，則將 1st 取代為 2nd。當我們發現資料中錯誤或需要調整之處，我們可運用 REPLACE 的方式，批次找出所有問題，並加以取代。就本例而言，我們發現原來資料中的 1st Edition 應為 2nd Edition，即可以此函式來處理。

3. 常用日期函式語法簡介

日期函式可供截取部分日期、計算日期／時間區間等。以下介紹兩個常見的日期函式語法。

語法說明：DATENAME(指定日期之部分, #yyyy-mm-dd# 或特定日期變項, [start_of_week]

以字串的形式回傳該日期之部分, 可指定 year、month、week、day 或 hour。Tableau 預設每週由週日起算, 若欲修改可自行設定。

例如：DATENAME('year', #2019-01-15#) = '2019'

例如：DATENAME('month', #2019-01-15#) = 'January'

例如：計算日期變項 [Date] 為該月的第幾週

DATENAME('week', [Date], 'monday') -

DATEPART('week', DATETRUNC('month', [Date]), 'monday') + 1

DATENAME 函數是用來截取部分日期。上述實例中, 第一、二例說明了當我們僅希望截取該日期的年份或月份時, 可以透過此語法, 去除不必要的細節。而第三例則是當我們想要了解該日期是該月第幾週的計算方式。

語法說明：DATEDIFF(date_part, date1, date2, [start_of_week])

計算兩日期或時間之差

例如：以週起始日（此例設定為週一起算），計算兩日期相差幾週

DATEDIFF('week', #2019-01-20#, #2019-01-22#, 'monday') = 1

其中, 2019-01-20 為週日, 2019-01-22 為週二, 以週一起算相隔一週

若以 2019-01-22 週二與 2019-01-25 週五來計算, 則將回傳 0

例如：計算兩日期變項 date1 與 date2 相差幾天

DATEDIFF('day', [date1], [date2])

例如：計算兩時間變項 hour1 與 hour2 相差幾小時

DATEDIFF('hour', [hour1], [hour2])

DATEDIFF 函數是用來計算兩日期或時間之差。假設我們欲知某事件的始末日期相差幾週、相差幾天, 甚或相差幾小時, 則可運用此函式語法來計算。

4. 常用邏輯函式語法簡介

邏輯函式可供設定條件, 針對特定變項新增變項以重新歸類等。以下介紹兩個常見的邏輯函式語法。

語法說明：CASE 針對特定欄位變項設定條件，並回傳相對應之值

```
CASE [欄位變項]
WHEN 'value1' THEN 'return1'
WHEN 'value2' THEN 'return2' ...
END
文字才須加上雙引號
```

例如：欲將 [Type] 中的 A 和 B 歸為第 1 類、C 和 D 歸為第 2 類

```
CASE [Type]
WHEN 'A' THEN 1
WHEN 'B' THEN 1
WHEN 'C' THEN 2
WHEN 'D' THEN 2
END
```

CASE 是資料處理中相當常用的。當我們欲針對特定欄位變項中的內容（值）重新歸類時，CASE 就是最好的幫手。舉例而言，當我們欲將年度變項之 2011、2012、2013、2014、2015 年歸入「2011-2015 年」一類，2016、2017、2018、2019、2020 年歸入「2016-2020 年」一類，則可套用上述實例的語法。

語法說明：IF

針對一系列表述，當表述成立時，則回傳 THEN 之後的值

```
IF <expr> THEN <then>
[ELSEIF <expr2> THEN <then2>...]
[ELSE <else>]
END
```

例如：欲以學校收受學生與實際註冊人數來衡量其表現

```
IF [Enrolled total] = [Admissions total] THEN 'great'
ELSEIF ([Admissions total]-[Enrolled total]) <
[Admissions total]*.6 THEN 'fair'
ELSE 'poor'
END
```

雖然 CASE 和 IF 某種程度都可協助我們「歸類」，但 CASE 是針對某個特定欄位變項中的值來重新歸類（前述實例是將單一欄位變項中的內容進行合併，以重新分類），而 IF 則是針對幾個欄位變項的關係，設定的條件敘述來分級歸類。在這個實例中，我們係以本單元的 IPEDS 教育統計資料來設想，當我們希望透過一所學校收受學生的人數與其實際註冊人數的情況來衡量該校表現時，

可運用上述實例的 IF 語法，將學校表現分級。此函式語法的表述係將學校表現粗分為三級：1. 註冊人數等於受收入數之學校，視為表現優異者；2. 收受人數與註冊人數相差不及收受人數六成者，視為普通；3. 其他則視為不理想。

小結

本節介紹了如何運用 Tableau 進行資料視覺化，以及 Tableau 的基本程式語法。一般說來，將資料匯入 Tableau 後，要記得確認資料格式設定正確（尤其是地理資訊），以利 Tableau 較準確地為我們選擇合適的圖形呈現方式。在工作表中，只要將欲用來繪圖的欄位變項置於欄與行（列），並選擇合適的圖形呈現方式，即可簡易製圖。

除了進一步客製化顏色、標籤等，善用 Tableau 的過濾篩選及相關分析功能，皆有助提升資料探索與呈現的深度。運用過濾篩選功能將有助聚焦特定項目之比較；運用平均數或四分位差等參照功能，則有助理解數值的意涵。而無論是提供比較或幫助理解，皆是資料視覺化的核心價值。

Tableau 的函式語法主要包括：數字、字串、日期與邏輯四大類。其中，數字函式可回傳絕對值、取整數、取較大或較小值等；字串函式可尋找與取代部分文字等；日期函式可截取部分日期或計算日期／時間等；邏輯函式則可協助將欄位變項重新歸類或依設定條件進行分級等。若可進一步運用函式語法，則可更深入地探討相關問題，運用新增的運算欄位，做更多、更有意義的視覺呈現。下節（第四節）將進一步以 CASE 的應用，介紹製作拼貼地圖（small multiple maps）的作法。

7-4 地圖資料視覺化

本節將介紹地圖資料視覺化基本技巧，我們將以臺北市政府開放資料的自行車竊案資料為例，首先，將說明單一地圖的製作方式，接著，再說明拼貼地圖（small multiple maps）的製作方式。

單一地圖的製作

請參考第三節的匯入檔案的作法，將「自行車竊案.xlsx」檔案匯入 Tableau 中。第三節曾強調匯入資料時，須確保每個欄位格式設定正確。由於我們將製作地圖，地理資訊是否設定正確是非常重要的步驟。

以自行車竊案的課堂練習檔案為例，郵遞區號、經度、緯度、縣市和行政區皆須設定為地理角色，其中，由於臺灣的地理角色層級和美國不同，對應上難免有些出入。我們可將縣市設定為「城市」、行政區設為「郡/縣」。不過，無論你將縣市與行政區設定為地理角色中的哪個層級，最重要的是先確保這些地理資訊設定為地理角色（圖 7-38）。如此才能讓 Tableau 判定我們接下來可以繪製地圖。

#	Alt	Alt	Alt	Alt	Alt	Alt	Alt	Alt
編號	案類	發生(現)日期	發生時段	郵遞區號	經度	緯度	縣市	地址
1	自行車竊盜	2013/6/30	13~15	00105	121.557588	25.059991	臺北市	三民路151~180號
2	自行車竊盜	2015/1/1	16~18	00106	121.543445	25.026770	臺北市	住家里四維路124
3	自行車竊盜	2015/1/4	07~09	00106	121.543445	25.026770	臺北市	敦復里仁義路四段
4	自行車竊盜	2015/1/6	13~15	00108	121.497986	25.028590	臺北市	復興路1~30號
5	自行車竊盜	2015/1/8	07~09	00112	121.517799	25.148068	臺北市	...
6	自行車竊盜	2015/1/8	16~18	00105	121.557588	25.059991	臺北市	...
7	自行車竊盜	2015/1/9	04~06	00115	121.609757	25.036009	臺北市	...
8	自行車竊盜	2015/1/9	13~15	00114	121.592383	25.083706	臺北市	...
9	自行車竊盜	2015/1/9	13~15	00112	121.517799	25.148068	臺北市	...
10	自行車竊盜	2015/1/10	04~06	00114	121.592383	25.083706	臺北市	...
11	自行車竊盜	2015/1/10	10~12	00105	121.557588	25.059991	臺北市	...
12	自行車竊盜	2015/1/10	16~18	00108	121.497986	25.028590	臺北市	...
13	自行車竊盜	2015/1/11	07~09	00115	121.609757	25.036009	臺北市	...
14	自行車竊盜	2015/1/11	07~09	00115	121.609757	25.036009	臺北市	...

圖 7-38 確保資料格式設定正確

繪製地圖

地理角色設定完成後，我們就可以進入工作表，開始繪製地圖了。在這個練習中，我們欲繪製一張臺北市自行車竊案地圖，並以顏色來表示竊案數的多寡。

雖然繪製地圖的方式很多。此例中，我們先快速雙擊地理變項[郵遞區號]，

讓 Tableau 先快速為我們產生底圖。接著，將[記錄數]拖曳至標記區的顏色（圖 7-39）。

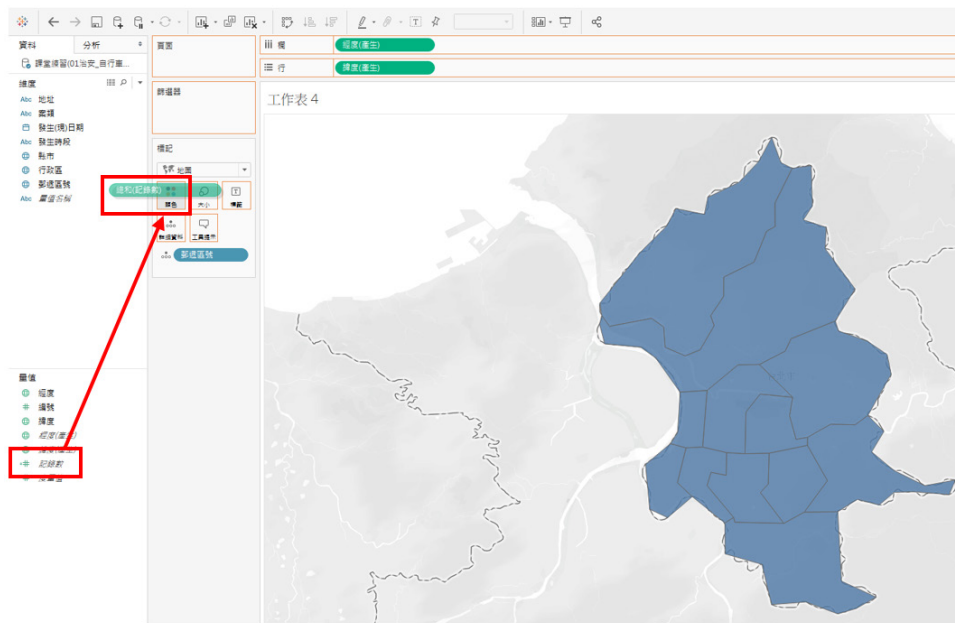


圖 7-39 以郵遞區號快速繪製的底圖

此時，若 Tableau 沒有為我們選擇填色地圖，而是以點狀呈現各行政區位置時，則可參考第三節選擇圖形呈現方式的作法，運用「顯示」（Show Me）功能，點選我們欲繪製的地圖類型。如此，我們便可快速完成一張簡單的地圖了。

然而，我們可能希望進一步客製化地圖的顏色與邊界，甚至加上標籤說明。這些客製化的作法亦如第三節所述，我們可在標記區中，設定顏色及邊界。若要取消邊界，可在標記區的「顏色」中，將「邊界」由「自動」改為「無」（圖 7-40）。

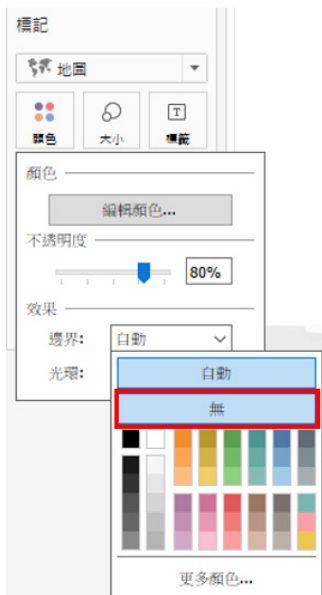


圖 7-40 去除行政區的邊界

若我們要新增地圖標籤，說明行政區的位置分布，則可將「行政區」拖曳至標籤。然而，此時，若 Tableau 無法辨識部分地區，則會出現圖 7-41 的現象。原本完整的臺北市地圖，因為加上 Tableau 無法辨識的行政區標籤而缺了幾塊。此時，則須透過「編輯位置」的方式，協助 Tableau 辨識這些行政區。

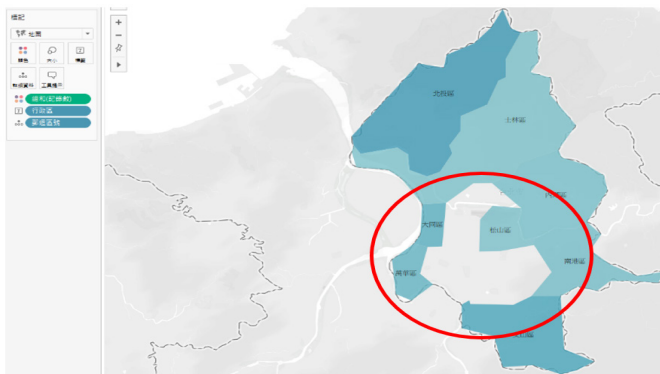


圖 7-41 Tableau 無法辨識部分行政區

我們可透過工具列上的「地圖」，透過「編輯位置」的功能，將「州/省」設定為「Taipei City」，確保 Tableau 可辨識地理行政區，以解決前述問題（圖 7-42）。

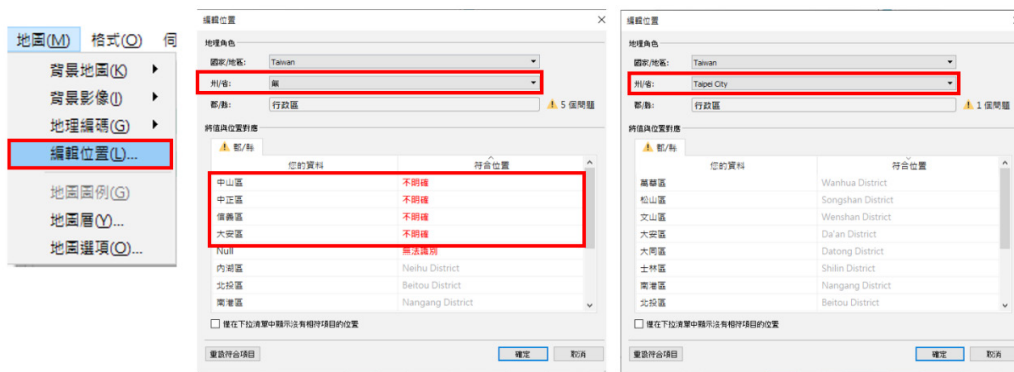


圖 7-42 編輯位置

由此看出，Tableau 將臺北市視為「州/省」。因此，若在設定資料時，直接設定為「州/省」，則在加入標籤時，就不會出現圖 7-41 的現象了。然而，編輯地理位置是運用 Tableau 繪製地圖的一項重要功能，雖然 Tableau 版本不斷更新進步，對於臺灣地圖的辨識也愈來愈理想，但繪製地圖難免遭遇 Tableau 無法辨識地理位置的問題，故前段刻意如此安排，讓大家掌握因應的作法。

如此便完成一簡單的臺北市自行車竊案地圖了（圖 7-43）。

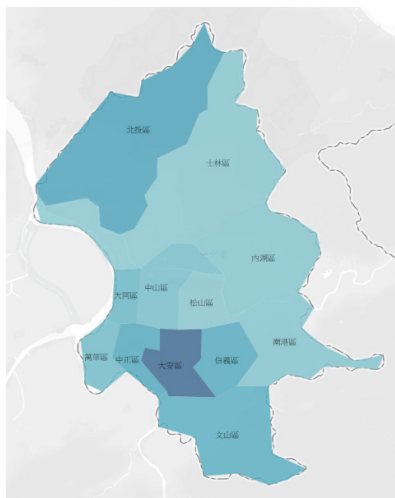


圖 7-43 完成單一地圖

Small Multiple Maps 的製作

第二節談到資料視覺化的基本原則時，我們提到資料視覺化往往希望盡量納入更多變項 / 面向的資料，也提到在非必要時，避免使用立體圖。然而，我們要如何在二維的平面空間中呈現納入許多變項呢？**Small multiples** 將許多小圖拼貼成一大圖的技巧運用往往可以幫助我們達成上述的目標。因此，**Small multiples** 是資料視覺化中，常用的技巧之一。

Small multiples 的中文譯名不一，有人將之譯成「組圖」，有人譯成「分割圖（表）」，也有人譯成「倍製圖」。由於這些名稱似乎不如 **small multiples** 來得貼切，故本節姑且延用原文，並以較直白的「拼貼」來描述其作法。接下來，就讓我們一起來製作 **small multiple maps** 吧。

以 **Tableau** 製作 **small multiple maps** 的兩種常見作法，一是在每個分頁中做一張地圖，再新增一儀表板（**dashboard**），並將前述各圖拼貼至儀表板；一是利用單一分頁製作並排列所有的小地圖。本節將運用較簡單的方式，在單一的工作表分頁中排列小地圖。

排列小地圖的常見方式也有兩種：一是運用既有的欄位變項，以 A 變項為欄、B 變項為列；另一是新增用來指定小地圖排列位置的欄與列變項，直接指定每張小地圖要排在第幾欄、第幾列的位置。由於後者的函式語法相形複雜些，且直接運用既有變項更滿足資料視覺化納入更多變項的原則，因此，本節針對第一種方式作簡介。

7

繪製 small multiple maps 的準備工作

假設我們欲針對年度與最常發生的 5 個時段來繪製 5X5 的小地圖，由於既有的 [發生(現)日期] 已包含年度資訊，故僅須新增 [好發時段] 變項，以指定 5 個好發時段，並用以此二欄位變項來排列地圖（圖 7-44）。

	2014	2015	2016	2017	2018
好發時段 1					
好發時段 2					
好發時段 3					
好發時段 4					
好發時段 5					

圖 7-44 地圖的排列示意圖

此時，我們可利用邏輯函式語法來設定五個好發時段。而為了找出哪五個時段為好發時段，我們亦可充分運用資料視覺化探索資料的方式，透過長條圖找出最常發生的五個時段。圖 7-45 是在新的分頁中，快速製作一張以發生時段為欄、記錄數為行（列）的長條圖，並加以排序。

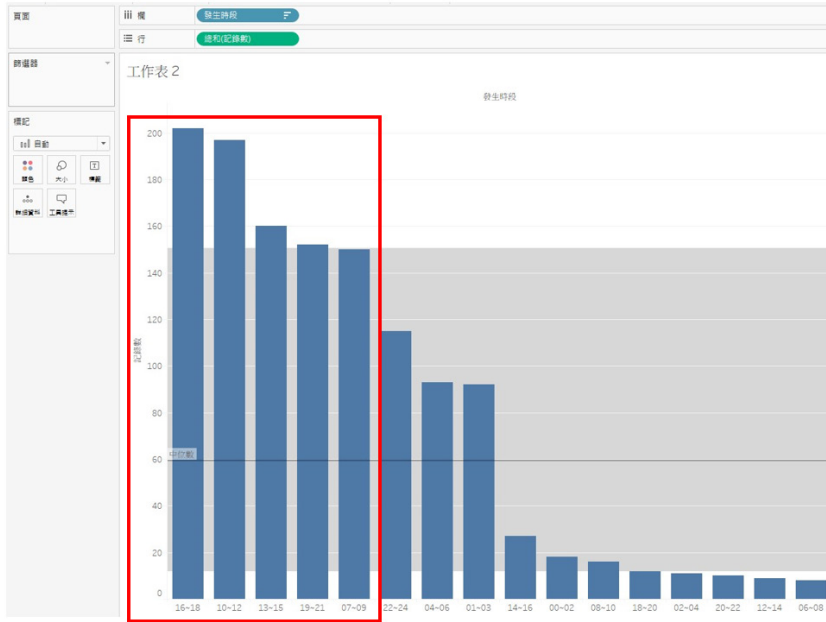


圖 7-45 運用長條圖找出五個好發時段

CASE 函式語法的應用

在維度欄位資料區按右鍵，選擇「建立計算欄位」，並輸入函式語法（圖 7-46）。

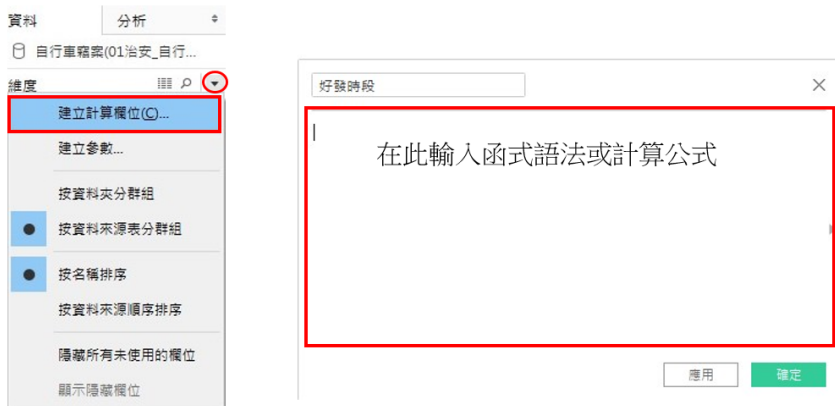


圖 7-46 建立計算欄位

在新欄位名稱輸入「好發時段」，並利用 Case 函數，針對 [發生時段] 設定欲排入前五的好發時段，函式語法內容如圖 7-47，指定當發生時段為「16~18」時，則設定為好發時段的「Top1 16~18」；發生時段為「10~12」，則設定為好發時段「Top 2 10~12」，依此類推，將圖 7-45 的前五名時段設定完成。CASE 函式完成時，其語法必須記得加上 END 作結。



圖 7-47 運用 CASE 函式新增「好發時段」

繪製 Small Multiple Maps

目前我們已備妥要用來排列小地圖的欄位了，接下來，我們直接在已繪製好的地圖工作表中，將 [發生(現)日期] 拖曳至欄，[好發時段] 拖曳至行(列)。若要保留第四節前半段所欲繪製的單一地圖，則可複製工作表（可參考第三節管理分頁的說明，在工作表頁籤上按右鍵，選擇「重複」即可），在新的工作表中進行上述步驟，即可快速繪製出 small multiple maps。

然而，由於竊案並非時刻發生且有些年度的資料可能未必完整，我們以 small multiple maps 打散資料後，部分小地圖的案例更顯稀少，沒有案例的地圖就包含了許多空白。為了讓 small multiple maps 更有效呈現竊案發生數較高的年度和時段，我們可以運用過濾篩選功能，過濾欄與行(列)中的缺值 (Null)，以排除較不明顯的好發時段與資料過於不全之內容。

由於 2013 的資料較不全，我們可點選 [發生 (現) 日期] 欄位變項的小箭頭，選擇「篩選條件 ...」，並取消勾選 2013；同樣地，我們可點選 [好發時段] 變項的小箭頭，選擇「篩選條件 ...」，並取消勾選 Null（圖 7-48）。

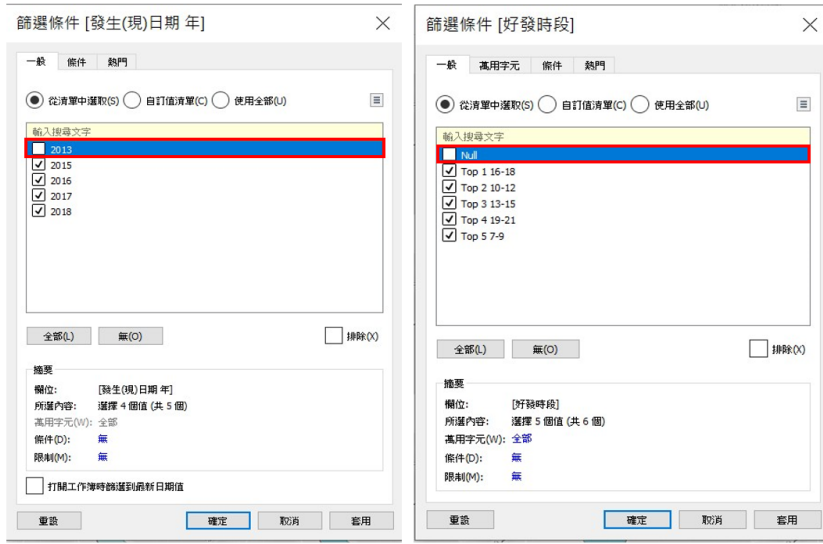


圖 7-48 過濾排除資料不全的部分

如此，我們就初步完成了臺北市自行車竊案的 small multiple maps 了（圖 7-49）。

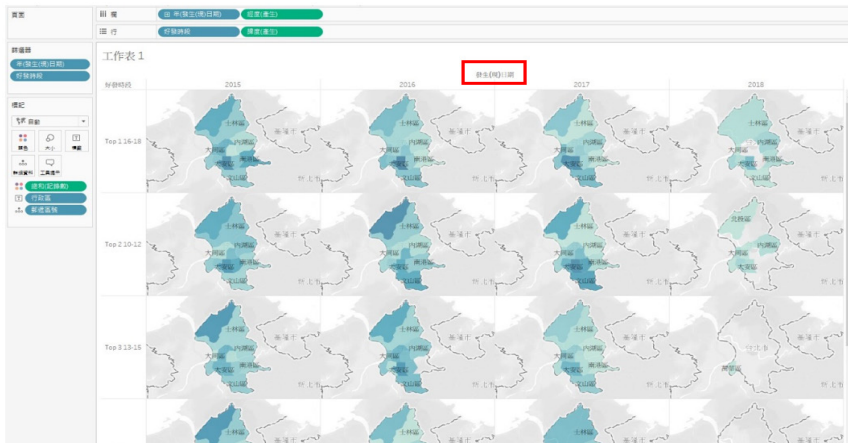


圖 7-49 初步完成 small multiple maps

地圖繪製小技巧

1. 隱藏欄位標籤

如果覺得圖 7-49 的「發生(現)日期」顯得多餘，我們可在欄與行(列)標籤上按右鍵，選擇「隱藏欄位標籤」(圖 7-50)。

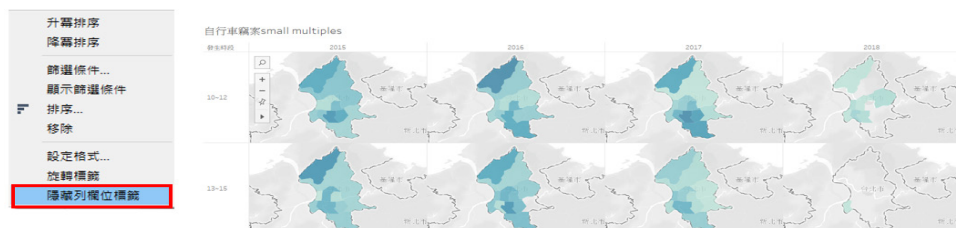


圖 7-50 隱藏 Small Multiple Maps 欄與行(列)標籤

2. 新增註解說明文字

我們亦可針對每張小圖作註解說明。只要在小地圖上按右鍵，選擇「新增註解」、「區域」，即可針對此區塊增加文字說明(圖 7-51)，亦可設定說明文字的字型、顏色等格式。

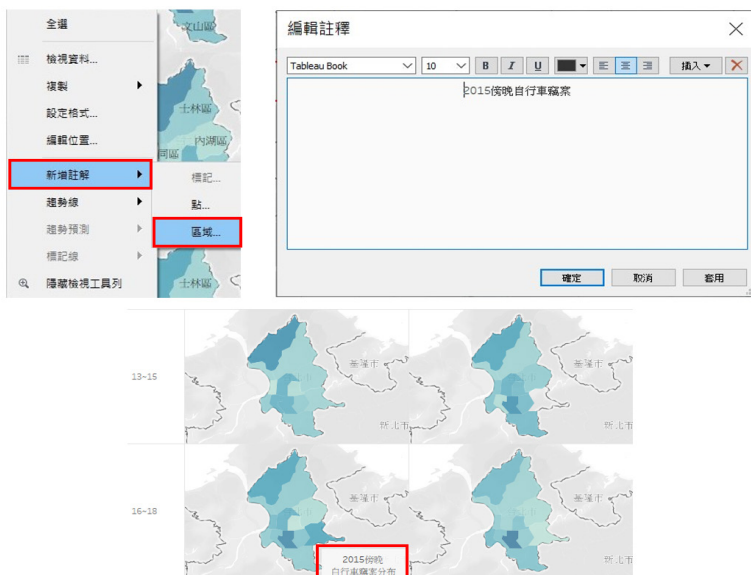




圖 7-51 新增註解說明文字

3. 調整地圖大小與位置

若要調整地圖的大小，可將滑鼠移至地圖上，此時，左上角的工具列即出現（圖 7-52），即運用+、-作放大、縮小等調整。若要調整地圖位置，則可透過+、-下方小箭頭的工具列，點選平移功能（即左數第二個垂直雙向的十字圖示 ）來移動地圖。若要截取部分地圖，則可利用最右側的三個圖示 ，選擇矩形、圓形或其他區塊，再作保留或刪除。

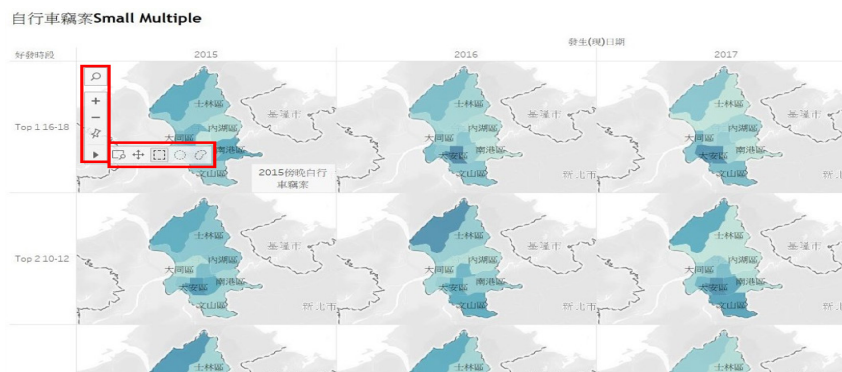


圖 7-52 調整地圖大小比例

4. 新增比例尺

若要新增比例尺，則可透過工具列的「地圖」，開啟「地圖選項」，再勾選「顯示地圖比例尺」即可（圖 7-53）。



圖 7-53 新增比例尺

5. 設定地圖層

最後，若要修改目前的地圖層，亦可透過工具列的「地圖」開啟「地圖層」選項（圖 7-54），即可調整圖層與地圖樣式。地圖層的視窗會開啟在原本資料區的位置，其中，背景樣式深淺有三種選擇，亦可選擇街道圖或衛星圖；地圖層則有許多選項可勾選，主要包括：土地覆蓋情形是否顯示地形或海岸線等，各種邊界是否顯示、各種地理行政區標示名稱等等。完成設定後，點選右上角的叉號，即可關閉此功能。

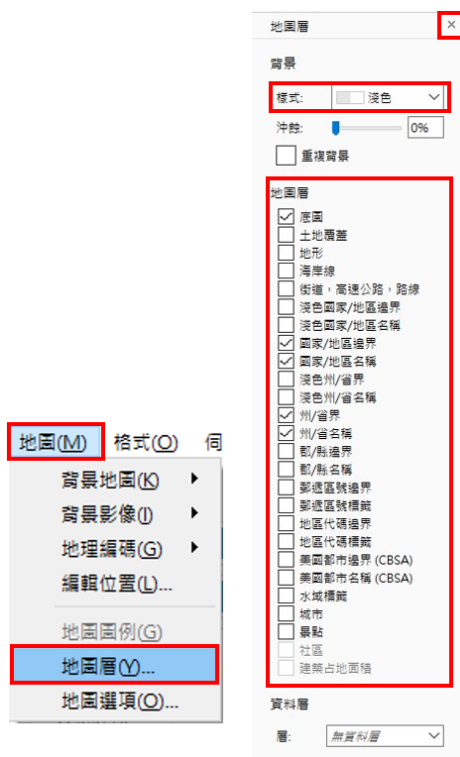


圖 7-54 地圖層的設定

製作互動式地圖

在 Tableau 中可以製作互動圖。本節最後將介紹互動圖的兩種簡單作法：一是運用「顯示篩選條件」的功能，一是運用「頁面」的功能。

第一種方法是針對欲用來互動瀏覽的欄位變項，點選其右側小箭頭，再點選「顯示篩選條件」，右側將出現篩選條件，即可供勾選互動瀏覽。圖 7-55 是針對一張以年度 [發生(現)日期] 和 [發生時段] 製作的臺北市自行車竊案 small multiple maps，並依 [發生時段] 建立互動瀏覽的作法。

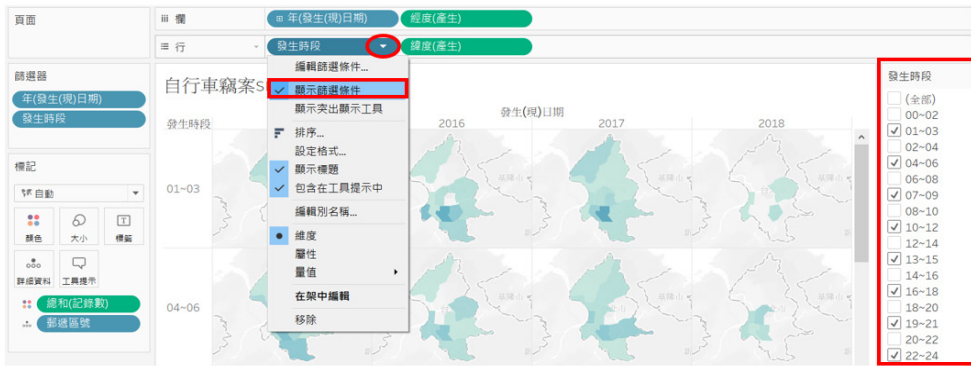


圖 7-55 運用篩選條件製作互動圖

第二種方法則是利用工作區左上方的「頁面」功能來製作互動式地圖。我們若將 [發生(現)日期] 拖曳至「頁面」，即可手動分年瀏覽檢視或自動播放歷年變化。然而，本例已將年度分成小地圖呈現，依年瀏覽意義較不大，故我們可在臺北市自行車竊案的單一地圖中，嘗試這項功能。圖 7-56 是將 [發生(現)日期] 拖曳至「頁面」後的互動圖。

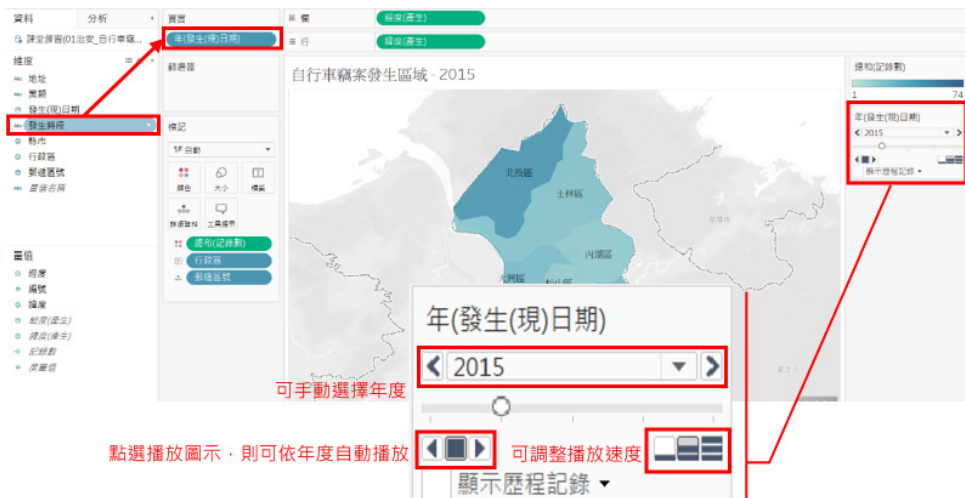


圖 7-56 運用頁面功能製作互動圖

此互動圖中，若手動點選年度（圖 7-56 的 2015）旁的左右箭頭，則可一一檢視不同年度的地圖，或以下拉式選單直接檢視特定年度。若要自動播放，則可點選左下方的左右播放按鍵，讓 Tableau 自動播放。其中，點選向右播放鍵，Tableau 將從 2015 播放至 2018 年；點選向左播放，則將倒序從 2018 年播放至 2015 年。而右下側的三個按鍵代表播放速度，一條線的按鍵是最慢的速度，可自行調整。

最後，在 Tableau 中，快速調整日期呈現單位的小技巧也非常值得一提。就讓我們透過微調先前製作的 small multiple maps，說明調整日期呈現單位的方法。

若我們欲將原本 small multiple maps 中的以「年」為單位呈現的 [發生(現)日期] 改以「季」為單位來呈現，並用「年」來互動瀏覽，可以點選 [發生(現)日期] 右側的小箭頭，並選擇以「季」呈現（圖 7-57）。如此就可快速讓 small multiple maps 改為以第一至四季為欄、以五大好發時段為列來呈現了。接著，我們將 [發生(現)日期] 再次拖曳至「頁面」，即可完成以「季」為 small multiples 底圖並支援分「年」瀏覽的互動圖了。

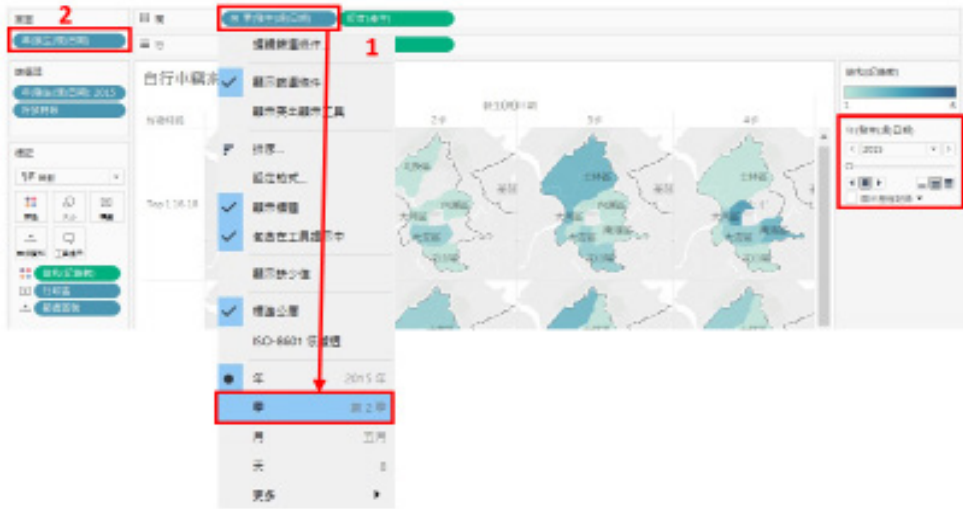


圖 7-57 調整日期呈現單位並製作互動圖

小結

本節介紹了單一地圖與拼貼地圖的製作，也應用了 CASE 函式語法，根據自行車竊案的發生時段，新增一欄位來描述其好發時段。地圖的製作須留意的是欲用來繪製地圖的地理資訊是否設定為地理角色，而倘若 Tableau 無法辨識地理位置，則須透過「編輯位置」的功能來解決此問題。此外，本節也介紹了幾項地圖製作的技巧與互動圖的製作方式等。其中，互動圖的製作與日期的呈現單位亦可應用於各類型的圖形呈現中，皆是 Tableau 相當重要的功能。

雖然本章三、四節著眼於 Tableau 此一資料視覺化軟體，Tableau 也確有其強大功能，然而，沒有任何一項工具是資料視覺化的唯一選擇。藉由介紹此功能豐富的視覺化軟體，希望讀者認識資料視覺化之基本作法，相信學會使用一套工具之後，即可大致掌握資料視覺化的基本操作流程與觀念，未來運用其它工具也能更快速地上手。

參考資料

1. 色盲。教育百科。檢自 <http://163.28.84.216/Entry/WikiContent?title=%E8%89%B2%E7%9B%B2&search=%E8%89%B2%E7%9B%B2>
2. Abela, A. (2009). Choosing a good chart. Retrieved from <http://extremepresentation.typepad.com/files/choosing-a-good-chart-09.pdf>
3. Color Blind Awareness. Retrieved from <http://www.colourblindawareness.org/>
4. Few, S. (2009). Now you see it: Simple visualization techniques for quantitative analysis. Oakland, Calif: Analytics Press.
5. Few, S. (2012). Show me the numbers: designing tables and graphs to enlighten (2nd ed.). Burlingame, Calif: Analytics Press.
6. Kirk, A. (2016). Data visualisation: a handbook for data driven design. Los Angeles: Sage Publications.
7. Sleeper, R. (2018). Practical Tableau: 100 tips, tutorials, and strategies from a Tableau zen master. Sebastopol, CA: O'Reilly Media.
8. Tableau Learning. Retrieved from <https://www.tableau.com/learn>
9. Tufte, E. (2006). Beautiful Evidence. Graphics Press.
10. Weinschenk, S. (2011). 100 things every designer needs to know about people. Berkeley, CA: New Riders.
11. Yau, N. (2013). Data Points: Visualization That Means Something, Wiley.