

資料視覺化的奇幻之旅

彭其捷

Tableau 資料視覺化工具 (上): 特色介紹



作者介紹

彭其捷鑽研網路服務多年,出版過 3 本 UX 使用者經驗專書,曾大量接觸設計、工程,使用者體驗等相關工作,近年因大數據 & 物聯網概念蓬勃發展,觀察到越來越多數據導向的服務興起。然而,艱澀的數據需要良好的設計輔助,才能創造良好的閱讀體驗,其中特別依賴資料視覺化的相關能力。因此,本專欄特別針對各類網路服務的資料呈現提出美學觀點,分享一些國內外資料視覺化的概念、工具與案例。

系列文章介紹

FINDIT 的目標是『發現趨勢,看見未來』,事實上,眾多的趨勢就隱藏在眾多的數據當中,等待著人們去發現、去解讀。透過『資料視覺化』的輔助,能夠把冷冰冰的數據圖像化,協助人們掌握趨勢,更能夠協助新創企業在創新創業過程中找出正確方向,或是幫助投資人找到潛在投資標的,是一項強大的武器。

本系列文章將從資料視覺化的概念開始著手,之後會陸續分享視 覺化的經典案例,相關工具介紹等等,希望讀者們能夠在大數據時代, 透過視覺化的輔助,說出好的故事。

接下來兩個月的主題是 Tableau 資料視覺化工具, Tableau 近幾年被許多商界、研究者人士推崇, 許多智能化功能可引導操作者進行更多的資料洞察。本期的內容介紹 Tableau 的八大特色, 也比較了 Excel 與 Tableau 的使用情境, 有興趣的讀者可以嘗試使用看看。

Tableau 資料視覺化工具(上):特色介紹

Tableau 簡介

在前幾個月的內容有提過,Tableau 軟體近幾年異軍突起,成為許多企業指定使用的資料視覺化工具,Tableau 從北美的商業分析市場起身,2004 年於美國西雅圖成立,十年後在納斯達克上市,成為軟體界的翹楚。Tableau 近期甚至連續五年(2013~2017)被 Gartner商業智能平台報導為領先的分析工具 (Gartner's Magic Quadrant for Analytics)。



資料來源: https://www.tableau.com/about/blog/2017/2/tableau-five-years-leader-gartners-magic-quadrant-analytics-66133

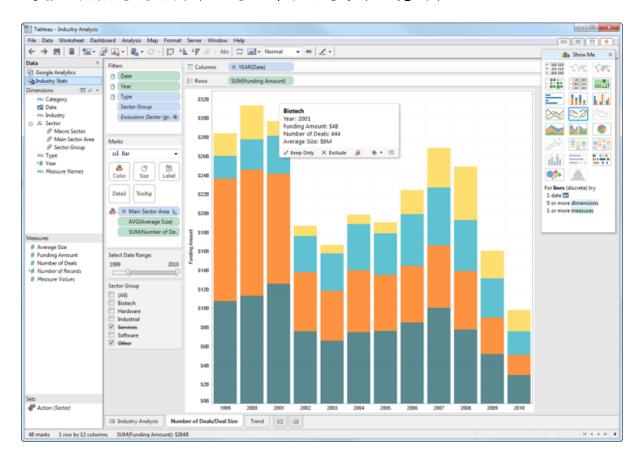
圖 1、Tableau 於 2013~2017 連續五年被 Gartner 商業智能平台 選為分析領先工具

Tableau 提供了極佳的操作經驗與直覺的拖拉手勢操作,即使是新手也能快速學會基本操作,在不撰寫程式的情況下完成洞察任務, 例如以下行為都可透過滑鼠簡單點擊完成,不用撰寫程式,不需要定 義行為,也不用安裝外掛。

- 透過 Tableau 連結資料庫
- 透過 Tableau 產生視覺圖表

- 透過 Tableau 分享設計成果
- 透過 Tableau 過濾數據
- 透過 Tableau 深入資料層級

值得特別一提的是 Tableau 豐富圖表功能,相較於其他軟體能夠 提供更多視覺化輔助,即使是新手也能夠過智慧化功能自動生成報表, 隨著工具熟悉度的提升,更能創作出超專業視覺圖表。



資料來源:https://docs.treasuredata.com/articles/tableau-desktop-odbc

圖:Tableau 軟體的基本操作介面與視覺圖表呈現功能

Tableau 版本介紹

『Tableau』是一系列產品的總稱,旗下主要有四大產品線,分別是:『Tableau Public』、『Tableau Desktop』、『Tableau Server』、『Tableau Online』,除了 Server 版之外,其他版本的編輯功能大致相似,只是滿足不同的使用需求。此外,也有一款免費下載的『Tableau Reader』工具,主要提供 Tableau 桌面軟體的瀏覽體驗。

各版本的介紹如下:

1.Tableau Public

• 網址:<u>https://public.tableau.com/s/</u>

• 特色:可使用 Tableau 多數功能

• 費用:免費

因為如果只是個人的分析需求,Tableau Public 絕大多數已經可以滿足,功能跟付費的 Tableau Desktop 類似,但 Tableau Public 的主要限制是資料成果一定要公開,且不能存檔在自己的電腦,資料也有 100 萬筆資料上限的限制。

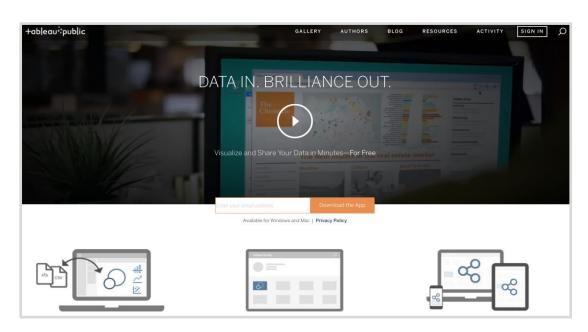


圖 2、Tableau Public 版本可以免費使用,讀者可以下載來試玩看看

透過 Tableau Public 製作的作品,如果進行上傳,將會被上傳到 Tableau Public 的線上視覺圖庫 (Gallery) 當中,供全世界瀏覽。



圖 3、透過 Tableau Public 設計的作品,將會被上傳到雲端,供全 球觀看,我們也可查看別人設計的成果

2. Tableau Desktop

• 網址: https://www.tableau.com/products/desktop

• 特色:擁有商用分析的完整功能

• 費用:付費

Tableau Desktop 是一個完整的桌面版本設計環境,能夠匯入許 多資料來源,也能做出各類分析圖表,可本地儲存,很適合公司進行 商業化導入。

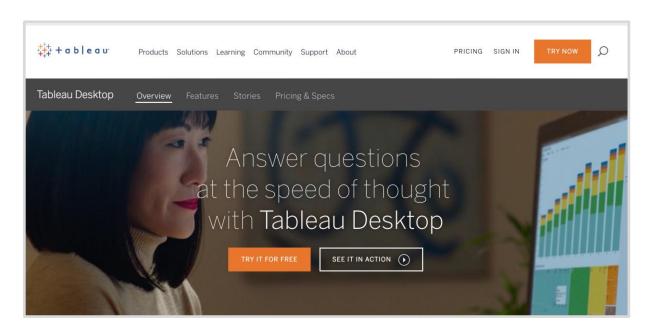


圖 4、Tableau Desktop 是專門設計給企業使用的個人套裝軟體

3. Tableau Server

• 網址: https://www.tableau.com/products/server

• 特色:可供企業私有雲的方式儲存資料與圖表

• 費用:付費

Tableau Server 可部署在企業的內部環境中,或是放置在公開的雲端伺服器如 Amazon、Azure、GCP 主機等等,提供專屬的企業環境使用,可讓 Tableau 桌面版本上傳設計成果到此,完全的私密保存,擁有中央權限管理相關機制,像是使用者、群組管理等等,也可線上管理 Tableau 相關設計成果 (Projects, Workbooks, Views 等等),以及資料來源 (Data Sources)的管理。

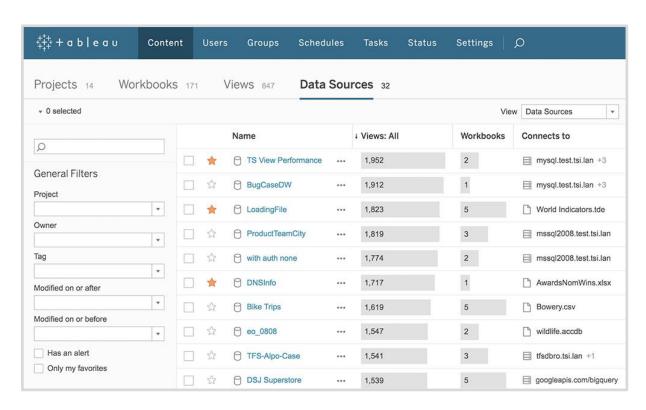


圖 5、Tableau Server 的資料來源管理介面

4. Tableau Online

• 網址: https://online.tableau.com/

• 特色:可直接透過瀏覽器進行視覺圖表的編輯

• 費用:付費

Tableau Online 提供一個月租型的線上分析環境,可於雲端集中管理資料,將檔案線上分享給老闆或同事,也可透過公開分享讓別人查看視覺化報表,不過因為是在雲端上執行,編輯效能較 Desktop 版本來的低。



圖 6、Tableau Online 提供了線上數據分析的環境

Tableau 線上教學

讀者看完以上介紹有意願使用的話,建議可先下載免費的 Tableau Public 版本玩玩看,如果需要較深層的企業分析應用,或是 有內部不能曝光的數字,則可購買 Tableau Desktop 或是 Tableau Server 版本,但費用不低,建議真的有需求再開始相關採購行為。

另外,Tableau 系統軟體對多數人來說都是新的嘗試,但網路上已經擁有豐富的線上教學課程與社群,有興趣的讀者不仿可以先在線上免費學習。

• Tableau 線 上 課 程 網 址 : https://www.tableau.com/learn/training

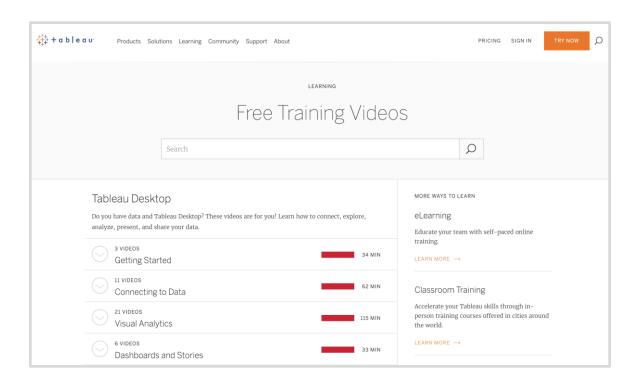


圖 7、Tableau 提供免費線上課程

Excel / Tableau 使用時機

相較於 Excel, Tableau 是一個比較新的軟體 (但其實也已經發展超過 10 年),相信許多讀者都會好奇,我們到底何時適合使用 Excel,又是何時適合使用 Tableau 呢?其實兩者有許多功能是重複的,整體來說 Excel 對於數據的處理功能更為完整,而 Tableau 則是視覺功能使用起來非常順手,但兩者之間並非全然的取代關係,甚至可以協同使用。以下介紹兩者之間的特色與合適的使用時機。

Tableau v.s. Excel

現在全世界最普及的數據分析工具依然是 Excel,跟隨著 Windows 作業系統深入全球每個人的家中,許多人從小到大已經用的非常習慣,然而 Tableau 優異的設計依然擁有其獨特性,值得嘗試看看。下表做了一些 Tableau 與 Excel 的比較:

表 1、Tableau & Excel 比較表

	Tableau	Excel			
基本資料檢視功能(排序、 列表、過濾器)	有完整功能	有完整功能			
軟體普遍性	較低(但逐漸提升)	較高(全球都在使用)			
視覺圖表豐富度	圖表豐富度高,且許多設計 細節可客製化	圖表也很豐富,但設計限制 較多			
視覺圖表互動性	提供大量互動性設計	較少互動性功能,或是需要 擴充設計模組,較為麻煩			
視覺圖表洞察性	提供智能功能,可幫助使用 者判斷使用哪種視覺圖表	需要有經驗的判斷與操作, 才能製作出好的成果			
承載數據量	能夠輕鬆的匯入幾百萬筆 資料量	在龐大資料量時,軟體穩定 性會降低			
學習門檻	可拖拉完成許多任務,且容 易學習	基本功能容易,但進階操作 學習門檻較高			
資料介接來源 (靜態)	可使用許多種類的資料,像是 csv, excel, pdf 等等	支援較少的資料格式(以 Excel 本身格式為主),須 透過外掛擴充			
資料自動更新(動態)	可直接對接許多種類資料 庫,點選完成,流程直覺	也可對接部分資料庫,但是 流程較為複雜			
數學公式使用	也有許多公式可使用,但尚 不及 Excel 的完整性	擁有大量豐富公式可使用			

資料融合	可用手動拖拉的方式,做資 料融合	需要自己下公式 (Ex. VLOOPUP) 融合資料
發佈到網路上	可將圖表發佈為網站格式, 也可嵌入其他網站	可將圖表發佈為網站格式, 也可嵌入其他網站,但流程 較為複雜
價格	如果想使用非 Tableau Public 版本,需要額外採購 支出	

資料來源:彭其捷。

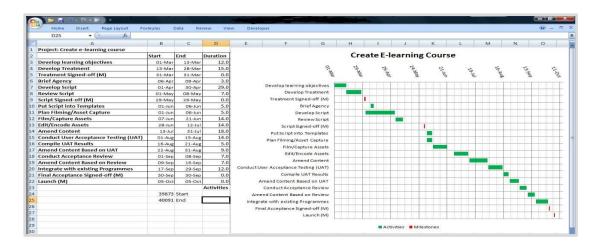
何時使用 Excel?

雖然本篇文章推薦讀者使用 tableau,但某些情境之下依然建議使用 Excel 工具,如下:

- 考量組織特性與既有業務流程
- 與相關工具體系的整合高
- 能夠快速創作小型業務表格
- 適合用於進行大量資料清理(ETL 程序)作業時

1.考量組織特性與既有業務流程

使用 Excel 的第一個考量是組織特性, Excel 在市場上存在已久, 已經成為各組織中的文書編輯首選,多數公司都有採購其授權,所以 如果組織提供 Excel 執行相關文書工作的話,考量的很可能不只工 具因素,有可能 Excel 已經融入整體業務流程當中,此時就不一定適 合更換使用的工具。

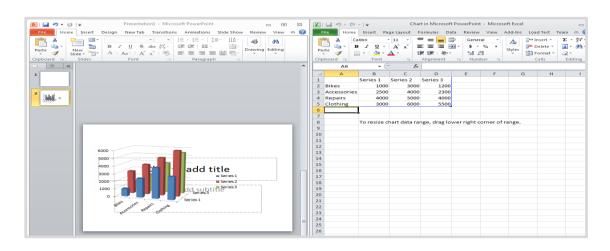


資料來源: https://ianjseath.wordpress.com/2009/06/03/how-to-create-a-half-decent-gantt-chart-in-excel/

圖 8、在組織中 Excel 常常跟業務流程已經相互綁定,此圖為 Excel 甘特圖案例

2.與相關工具體系的整合高

使用 Excel 的第二個原因,是相關支援工具體系的整合,由於微軟體系的軟體相當全面,除了 Excel 之外, Word 與 PowerPoint 都是辦公軟體的主要首選,三者之間有許多功能可以互相介接使用與交互編輯,格式之間彼此也能連動,各自的編輯成果也能直接貼到另外一個工具上面展示,十分方便。



資料來源:https://social.msdn.microsoft.com

圖 9、微軟旗下的文書軟體之間很能夠互通有無,例如 PowerPoint 跟 Excel 就有許多協作關係

3.能夠快速創作小型業務表格

Excel 依然很適合快速創作小型、單次性的業務表格,實際商務世界許多資料不一定是龐大數據,而是一些基本業務數據,只是需要改成表格的形式呈現。Excel 提供了重要的『儲存格』概念,可針對原始資料或是設計成果快速編輯,也可馬上進行輸出為 PDF 等等主流格式,而同樣的流程較不適合在 Tableau 工具上執行,因為Tableau 的強項更在於將現有的大量數據進行一致的數據邏輯整理,而非針對特定資料列表進行編輯。

Rep	Team	January	February	March	April		
George	SG	\$4,100	\$3,200	\$1,000	\$2,900	$\stackrel{\frown}{\boxtimes}$	
Jacob	JJJ	\$3,100	\$2,400	\$3,500	\$2,400	$\stackrel{\frown}{\boxtimes}$	~
Schmidt	SG	\$2,600	\$1,800	\$1,000	\$2,000		<u></u>
Jinglehiemer	JJJ	\$3,700	\$2,200	\$3,000	\$1,200	닸	\searrow
John	JJJ	\$1,200	\$2,000	\$4,000	\$1,100	닸	$\overline{}$
Total		\$14,700	\$11,600	\$12,500	\$9,600		

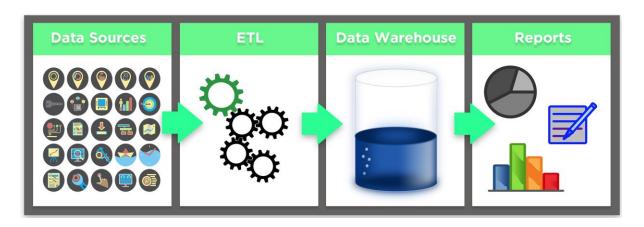
資料來源: https://blogs.msdn.microsoft.com

圖 10、Excel 在編輯小型資料表格的時候非常好用

4. 適合用於進行大量資料清理 (ETL 程序) 作業時

第四個時機是當需要進行許多細緻資料清理程序(ETL)時,ETL 是英文 Extract-Transform-Load 的縮寫,用來描述將資料從來源端經過萃取(extract)、轉置(transform)、載入(load)至目的端的整個過程。資料清理通常需要進行許多複雜的公式設計,或是針對特定的字串(String)格式、時間(Time)格式,地理(Location)格式資料進行轉置處理。舉例來說,我們可能會想要調整原本的時間資料,將『2017/3/1,上午8點39分』這句話轉換成『2017/03/01 08:39』的標準時間格式,此類操作在 Excel 當中有提供大量的公式或是字串處理功能可供使用,且 Excel 線上社群與 Blog 也提供了大量的教學說明,讓我們可以參考前人的經驗進行複雜的資料清理作業。雖然Tableau 也提供許多資料清理的功能與公式使用,但發展比較晚,對

於需要龐大複雜操作的資料清理程序來說,相關支援功能尚不夠完整, 建議可先將資料於 Excel 處理好之後,再丟進 Tableau 來發揮它的 強項:資料視覺化。



資料來源:https://datafloq.com/read/how-to-offload-the-etl-bottleneck-with-hadoop/2333

圖 11、ETL 程序重點在於將不同資料來源整併成乾淨、一致的格 式,供後續視覺化報表或是資料倉儲格式使用

使用 Tableau 的八大原因

Excel 與 Tableau 都是強大的資料視覺化工具,且 Excel 幾乎可以完成所有 Tableau 能夠完成的工作,但何時我們適合使用 Tableau 工具呢?首要考量跟 Excel 相同,需要先了解組織的文化,當組織鼓勵使用較新的工具、或是業務流程容許使用新工具時,甚至是組織正好採購了 Tableau 工具的話,將是使用 Tableau 的合適時機!

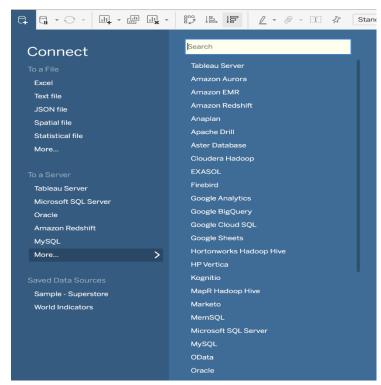
以下內容介紹使用 Tableau 的八大原因:

- 1. 能夠輕鬆執行資料庫連動作業
- 2. 能夠快速成為數據洞察高手
- 3. 能夠利用工具主動提供視覺洞見
- 4. 能夠進行『顆粒級』資料互動檢視作業
- 5. 能夠設計強大的資料視覺成果

- 6. 能夠輕鬆建立組織商業智慧 (BI) 分析平台
- 7. 能夠利用強大的地理視覺引擎
- 8. 能夠輕鬆將設計成果發佈到網路上

原因1:能夠輕鬆執行資料 庫連動作業

一般來說,設定資料連 結在多數軟體上都是複雜 的經驗,例如 Excel 就能跟 連結許多種類的資料來源 但是不同設定起來相當 雜,相較之下 Tableau 提供 設計良好的資料庫連結流 程,能夠輕鬆串連各類資料 來源。



12、Tableau 所提供的資料連結選項, 相當多元且囊括所有主流資料來源

原因 2:能夠快速成為數據洞察高手

讀者也許會好奇,其他軟體也有提供許多圖表洞察功能,Tableau 的特色在哪呢?以上問題主要跟『操作的經驗』有關係,Tableau 花了非常多心力在使用者經驗(UX)的改善,以求設計出邏輯直覺的操作介面,不論是組織新人或是老手,都能透過所見即所得的拖拉介面,快速進行數據洞察工作,自由嘗試、快速迭代,不斷透過各種方便的操作功能分析數據,一直到深刻了解數據為止,在 Tableau 工具的輔助之下,可幫助更多人成為數據洞察高手。

圖

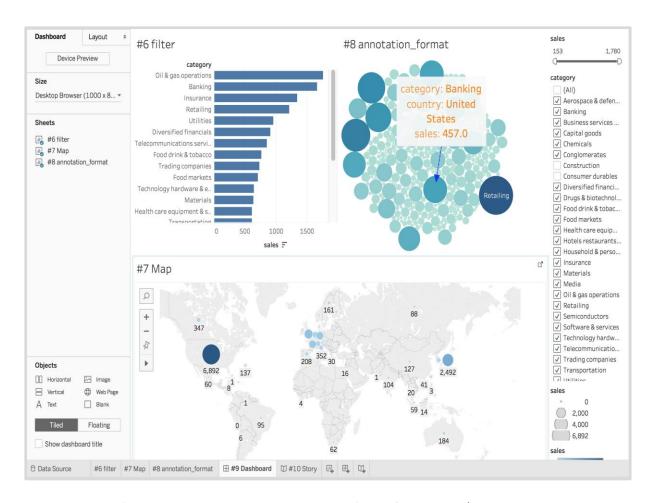


圖 13、操作 Tableau 的過程就像在創作一樣,可在資料之間自由探索,進行各類數據洞察

原因3:能夠利用工具主動提供視覺洞見

使用 Tableau 的第三個原因,是超強大的圖表智慧化提示功能 (SHOW ME 功能)。Tableau 提供許多智慧化行為預測,例如:當我們選擇特定的幾個資料欄位,Tableau 就會自動告知我們有哪些視覺圖表可供使用,如果我們點選兩下,Tableau 甚至可以自動幫我們生成一張美麗的圖表,重點是通常都還蠻漂亮的!但有需要的話,我們還是可以自行調整設計細節。Tableau 就像一個很棒的軍師,不斷給予我們各種設計建議。

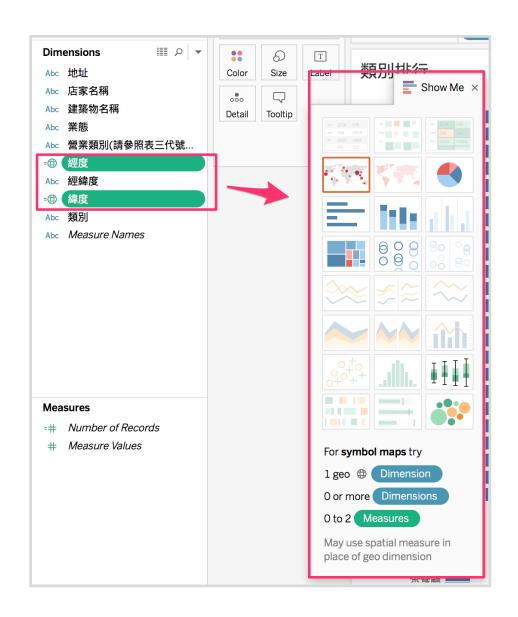


圖 14、當我們選擇部分欄位時,Tableau 即會自動建議合適圖表

原因 4:能夠進行『顆粒級』資料互動檢視作業

在資料量較大時,一般的軟體工具只能看一個大概的趨勢性,但 Tableau 提供了『顆粒級別』的資料節點洞察功能,例如:我們可先 查看整個國家的數據圖表,接下來可持續 zoom-in 查看城市的大小 級別、甚至是街道的大小級別,且此過程是動態互動的,只要透過簡 單的滑過手勢或是點選手勢即可完成。

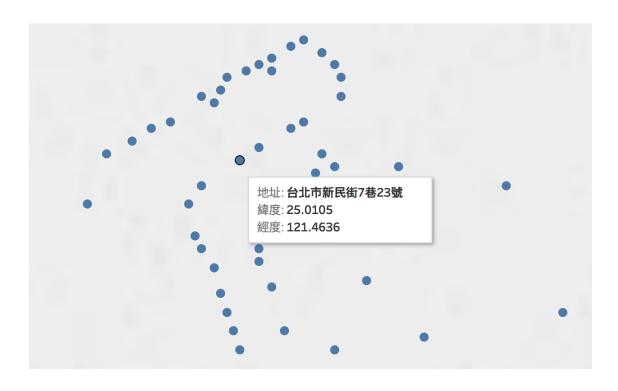


圖 15、Tableau 提供了顆粒等級的資料細部互動檢視,既可看整體 資料趨勢分佈,若有需要的話,也可查看細部資料節點

原因5:能夠設計強大的資料視覺成果

Tableau 提供了強大的視覺編輯引擎,相較於 Excel 能提供更細緻的畫面編輯功能,像是背景設定,內容對齊、字型設定、背景調整、顏色調整等等,也能夠上傳自己的圖片,針對畫布的地方也提供許多參數設定,像是解析度、瀏覽裝置、自訂版型等等,相當方便,能以更符合視覺設計邏輯的方式,創造精彩作品。

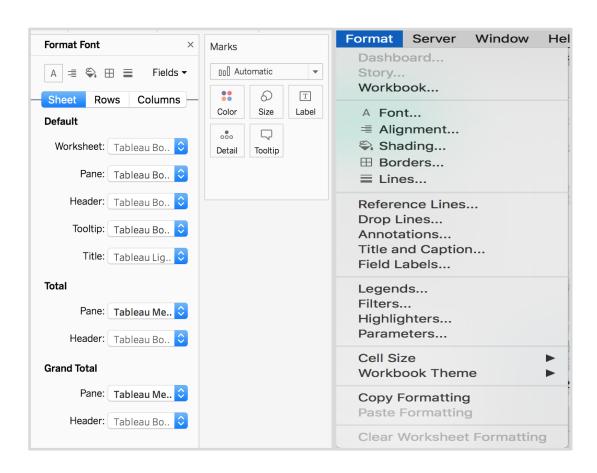


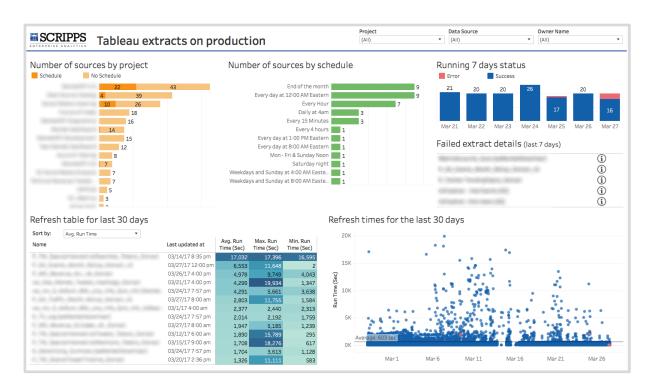
圖 16、Tableau 提供了大量視覺格式編輯工具,提供很大的設計彈性

原因 6:能夠輕鬆建立組織商業智慧 (BI) 分析平台

傳統的商業智慧 (BI) 平台,常常會隨著時間演進以及業務流程的改變而變得異常複雜,導致真的能夠高效使用該 BI 的人數大幅減少,造成組織效率降低。在 BI 領域常見的難題在於多重資料庫的導入,在大型組織當中,各單位各自擁有不同的資料格式甚至是資料庫,Tableau 提供了無痛的資料整合功能,可將不同來源的資料進行融合,或是進行跨資料表的計算等等。

此外,由於 BI 系統資料格式的限制,導致許多人變成資料處理人員,讓各單位的資料都可相容於 BI 系統,但這樣會導致一個現象,就是許多專業分析人員花了 80% 的時間進行資料清理作業,結果真正用在分析數據的時間只剩下 20%,大幅削減了生產力。Tableau 希

望避免這個情況發生,轉而用最友善的方式建構組織商業智慧平台。

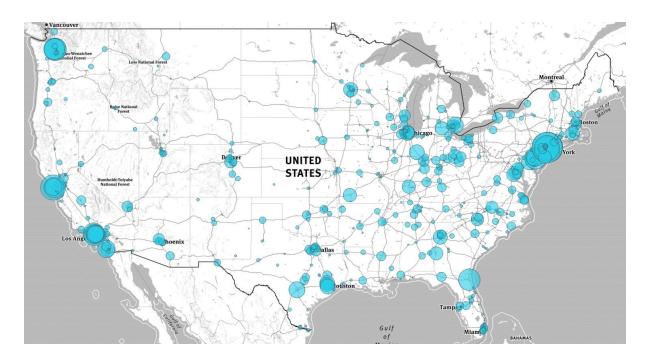


資料來源: http://www.bfongdata.com/2017/03/custom-tableau-server-repository.html

圖 17、透過 Tableau 可方便建立輕量的商業智慧 (BI) 環境,提升組織效率

原因7:能夠利用強大的地理視覺引擎

一般的視覺化工具如果要繪製地理資料,常常需要安裝外掛或是與其他軟體交互使用,但是 Tableau 本身便提供了強大的地理視覺引擎,甚至支援第三方地圖圖層,像是 Google Map、Mapbox、Leaflet等等。Tableau 也內建了地理字串處理引擎,例如『Taiwan』這個詞就能夠被 Tableau 判斷是一個『國家』,進而做出後續的資料推論行為。



資料來源:http://www.evolytics.com/blog/tableau-201-make-symbol-map-mapbox-integration/

圖 18、Tableau 的地理視覺引擎,能夠輕鬆做出華麗的地理視覺化圖表

原因8:能夠輕鬆將設計成果發佈到網路上

一般我們都是透過靜態圖片分享數據分析成果,但常常對方收到後,還是會持續回來詢問更細部資訊,或是相關的連動資訊,像是想要知道某個細部情境之下的數據,這樣不斷重複的一來一回溝通,可能也導致組織浪費了大量的時間。

Tableau 可輕鬆一鍵將設計成果發佈到網路上,或是放置到公開的 Tableau Public 環境,取得分享網址,我們可將該網址提供給客戶、老闆、同事等等,讓對象透過瀏覽器查看你精心設計的數據分析報告,也能自行操作互動洞察或是查看連動資料等,可大幅減少雙方溝通時間。

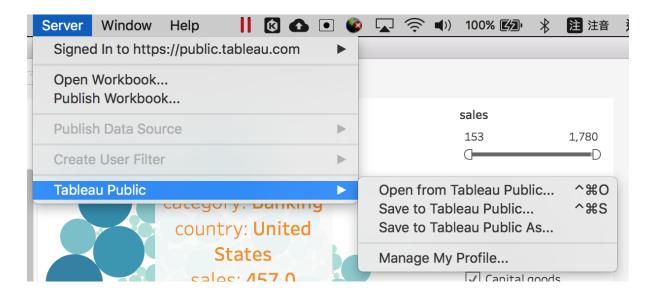


圖 19、Tableau 提供了多種上傳功能,可馬上產生可線上瀏覽網址

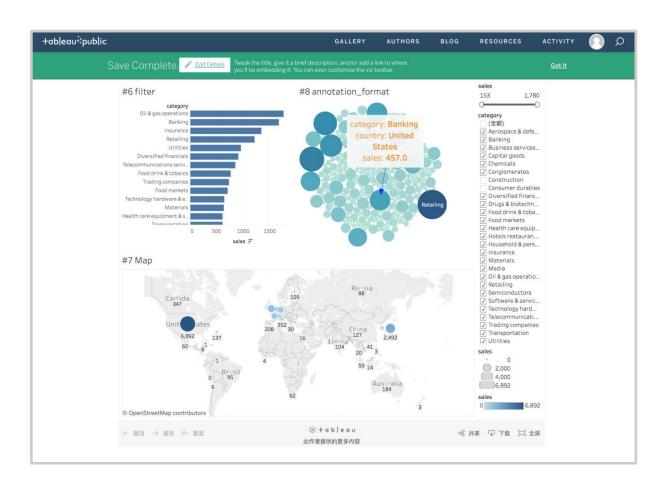
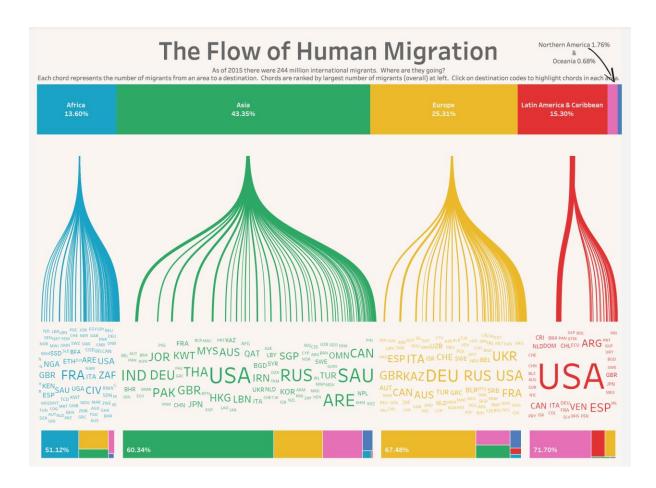


圖 20、單鍵即可將設計成果上傳到網路上,讓對象透過瀏覽器查看

Tableau 精彩作品分享

為了讓讀者了解 Tableau 的能耐,以下介紹幾張於 Tableau Public 公開展示的作品,可以看到透過 Tableau 可做出設計感十足的資料視覺化成果,除了顏色與排版都很專業之外,這些圖表都可以互動,讀者有興趣的話可以回到原本網站上,查看更多細緻資訊。



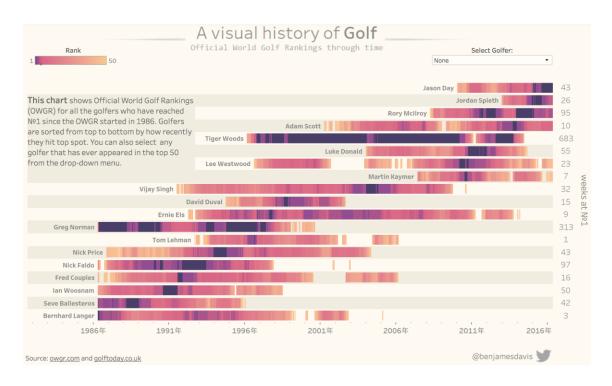
資料來源:https://public.tableau.com/en-us/s/gallery/flow-human-migration?gallery=featured

圖 21、The Flow of Human Migration (人類遷移數據視覺圖)



資料來源:https://public.tableau.com/en-us/s/gallery/urban-forest-nyc?gallery=featured)

圖 22、Urban Forest of NYC (紐約樹資料視覺化)



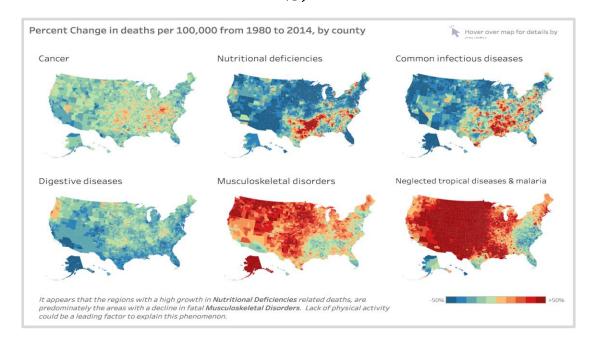
資料來源: https://public.tableau.com/en-us/s/gallery/world-golf-rankings?gallery=featured)

圖 22、World Golf Rankings (世界高爾夫排名視覺化)



資料來源:https://public.tableau.com/en-us/s/gallery/banksy-graffiti-around-world?gallery=featured

圖 23、Banksy Graffiti Around the World (班克斯世界塗鴉地圖視覺化)



資料來源:https://public.tableau.com/en-us/s/gallery/changing-diseases?gallery=featured

圖 24、Changing diseases (1980 到 2014 年之間的死亡比例改變數據視覺化)

結論

大數據時代來臨,全世界的資料量不斷增長,催生了許多資料分析需求,Tableau 入門容易,只要將資料製作成指定的 key-value 格式(例如 .xls.csv 數據)後,丟進軟體就可以自動生成各類視覺圖表,對於新手來說能快速產生成就感,甚至輸出為網頁格式,馬上分享給其他人瀏覽,加強組織生產力。讀者有興趣可以免費下載 Tableau Public 版本把玩看看。

相較於 Excel, Tableau 是一套新興的好用工具,兩者之間可以相互使用,不一定是完全的取代關係, Excel 擁有完整的資料處理機能,而 Tableau 則能夠用更直覺的視覺邏輯創造精彩圖表,相信各位都會找到適合自己的工作流程。

下一期的內容,會實際介紹 Tableau 的實作流程,敬請讀者期待。