

資料視覺化的奇幻之旅

彭其捷

資料視覺化的前置作業：資料準備與資料清洗

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。

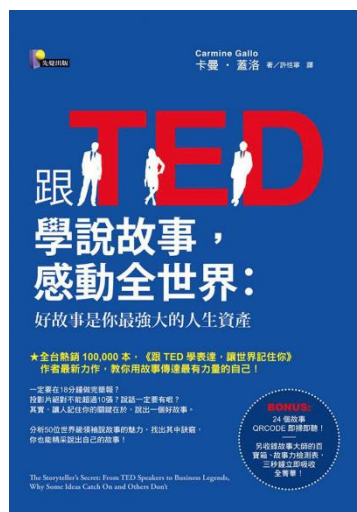
系列文章介紹

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。

雲端時代，有許多好用的線上工具可以視覺化工作流程中給我們許多幫助，本篇內容主要分享幾個好用工具，像是 Google Doc、Google Spreadsheet、Draw.io、Assana 等等，透過這些好用工具的輔助，能夠即時分享視覺化專案進度與成果，協助專案成功。

資料視覺化的奇幻之旅 - 視覺化圖表優化的 10 個絕招



你要說服的是心，不是腦袋 - 跟 TED 學說故事，感動全世界

我們生活當中無時無刻接觸到各種視覺圖表，有的容易理解，有的卻是乍看之下好像看得出什麼，實際上卻又看不太懂，甚至有許多圖表根本看不懂其主張意涵，這到底是什麼原因呢？

將資料進行視覺化涉及到許多能力，除了必備的邏輯思維之外，有時還需要左右腦並用，結合左腦的邏輯天賦與右腦的創意天賦，才能打造優秀的視覺圖表。魔鬼都藏在細節中，看似平凡的視覺圖表，閱讀者常常都在一瞬之間就決定了它的價值，好的圖表能夠輕鬆的引導閱讀，並在過程中加入讓人信任的元素，但有些時候我們卻會因為缺少某些關鍵資訊，又或者是我們可能只是使用了錯誤的顏色暗示，導致閱讀者的混淆，讓信任度大打折扣。

視覺圖表閱覽經驗是需要設計的，好的視覺圖表也需要反覆鋪陳，透過良好的脈絡設計引導流暢閱讀，從標題就讀懂想要表達的訊息，也能夠透過適當的圖像傳達重點，閱讀者可從畫面充分取得所需要的訊息，本篇文章分享的 10 項絕招，看似平凡，卻是讓圖表能夠更加取信於人的好用技巧，強化後的數據圖表，有更佳的說服力，更能夠說服主管、說服客戶，也更能夠影響最終決策。

本篇範例：癌症發生統計資料

本篇案例使用政府部門的『癌症發生統計』資料來進行相關案例解說，可在以下網址下載到本資料集，讀者也可自行嘗試實作看看。

資料素材下載
衛生福利部國民健康署：癌症發生統計資料 https://data.gov.tw/dataset/6399

以下是本資料集的摘要說明：

資料集評分:	★★★★☆ 平均 4 (14 人次投票)
資料集描述:	提供我國癌症發生統計資料，供各界使用
主要欄位說明:	年度、性別、縣市、癌別、WHO2000年人口標準化發生率(每10萬人口)、粗發生率(每10萬人口)
資料資源:	CSV 檢視資料 68-103年縣市別性別癌症別發生率資料.....
提供機關:	衛生福利部國民健康署
提供機關聯絡人:	徐小姐 (0225220888#796)
更新頻率:	每年
授權方式:	政府資料開放授權條款-第1版
計費方式:	免費
上架日期:	2014/04/20
資料集類型:	原始資料
詮釋資料更新時間:	2017/09/04 14:43
關鍵字:	癌症發生 癌症發生率

圖 1、癌症發生統計資料資料集

資料集欄位

以下是本資料集的欄位說明與截圖畫面：

表 2、癌症發生統計資料資料集的欄位、屬性、內容範例說明

	說明	欄位屬性	內容範例
#1	癌症診斷年	時間（年）	1979
#2	性別	文字	不分男女
#3	縣市別	文字	台北市
#4	癌症別	文字	口腔、口咽及下咽
#5	年齡標準化發生率 WHO 2000 世界標準人口 (每 10 萬人口)	數字	3.4
#6	癌症發生數	數字	439
#7	平均年齡	數字	53.37
#8	年齡中位數	數字	55
#9	粗率 (每 10 萬人口)	數字	2.5

癌症診斷年	性別	縣市別	癌症別	年齡標準化	癌症發生數	平均年齡	年齡中位數	粗率 (每10萬)
1979	不分男女	台閩地區	口腔、口咽	3.4	439	53.37	55	2.5
1979	不分男女	台北市	口腔、口咽	3.85	65	52	55	2.96
1979	不分男女	台中市	口腔、口咽	4.16	45	54.18	55	2.88
1979	不分男女	台南市	口腔、口咽	1.58	19	55.79	57	1.24
1979	不分男女	高雄市	口腔、口咽	2.1	32	48.56	50	1.49
1979	不分男女	基隆市	口腔、口咽	3.93	10	56.7	59.5	2.9
1979	不分男女	新竹市	口腔、口咽	4.37	10	48.3	53.5	3.62
1979	不分男女	嘉義市	口腔、口咽	4.54	8	60.75	62	3.16
1979	不分男女	新北市	口腔、口咽	3.26	47	52.85	53	2.2
1979	不分男女	桃園縣	口腔、口咽	2.97	21	55.81	55	2.07
1979	不分男女	新竹縣	口腔、口咽	1.62	4	64.5	64.5	1.11
1979	不分男女	宜蘭縣	口腔、口咽	2.93	10	53	54	2.27
1979	不分男女	苗栗縣	口腔、口咽	0.82	3	55.67	67	0.55
1979	不分男女	彰化縣	口腔、口咽	6.5	55	52.25	54	4.77
1979	不分男女	南投縣	口腔、口咽	2.45	10	52.7	52	1.92
1979	不分男女	雲林縣	口腔、口咽	2.43	15	49.47	50	1.88
1979	不分男女	嘉義縣	口腔、口咽	2.62	12	52.17	53	2.08
1979	不分男女	屏東縣	口腔、口咽	4.08	25	60.2	59	2.83
1979	不分男女	澎湖縣	口腔、口咽	2.31	2	56.5	56.5	1.84
1979	不分男女	花蓮縣	口腔、口咽	2.92	9	47.89	53	2.55

圖 2、欄位資料截圖畫面

資料過濾：台北市＋結直腸癌症

本篇案例我們不會挑選全部資料進行圖表的繪製，減少讀者資訊消化的難度，所以有針對本資料集的其中三個欄位進行過濾，分別是：

- 性別：不分男女
- 縣市別：台北市
- 癌症別：結直腸

性別	縣市別	癌症別
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸
不分男女	台北市	結直腸

圖 3、本篇文章主要描繪的是這個資料結果

單鍵生成圖表

我們可以先透過常用的 Excel 軟體，將此資料集生成一張圖表看看，例如最常用的直條圖：

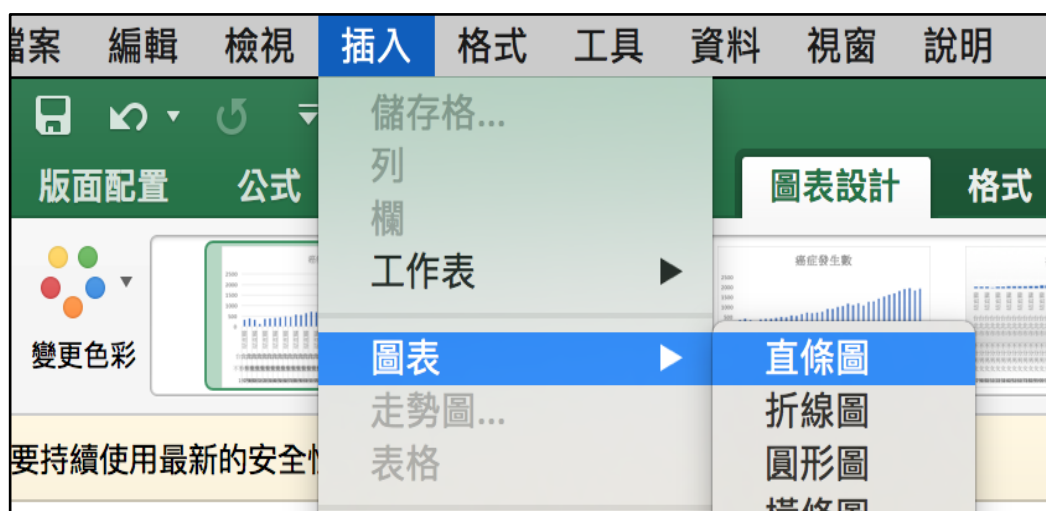


圖 4、超方便的 excel 工具，產出直條圖

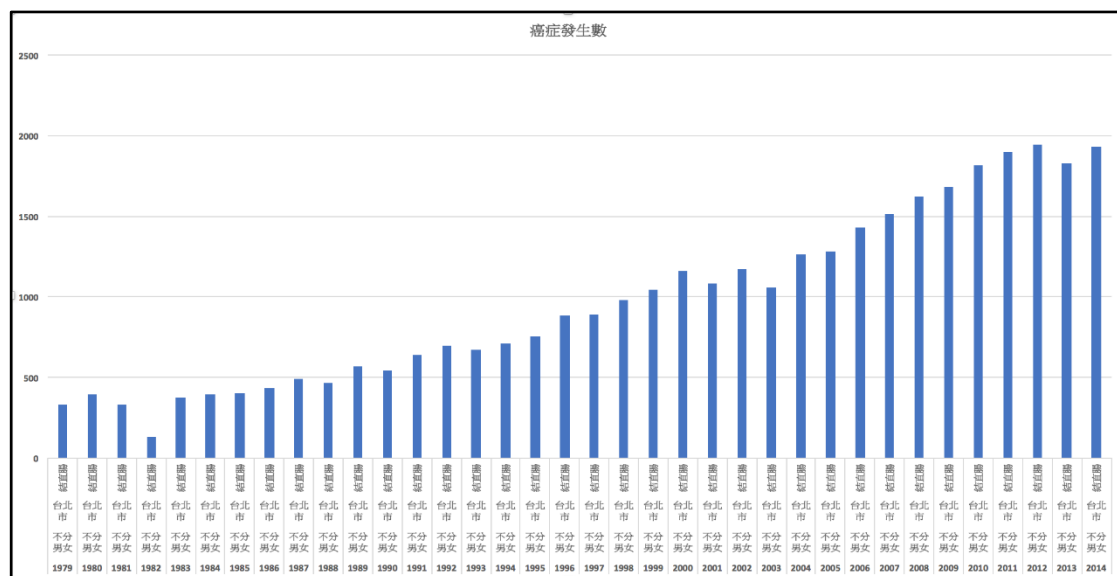


圖 5、一個標準的長條圖

許多人看了視覺化後的結果，會直接將此直條圖應用於簡報或是報告當中，當然透過此圖能夠看出大略的年度數據成長趨勢，但我們有沒有更好的呈現方式呢？或是這張圖還有哪些優化的可能性呢？

單鍵生成的缺點

許多人製作圖表的標準流程，大多是透過軟體的一鍵生成圖表功能開始（例如透過 Excel 等等工具），軟體能夠快速、無痛的產出數據圖表，因此許多上班族都非常依賴 Excel 來生成各種圖表。雖然單鍵生成的功能如此方便，但以分析的目的來說，有時候不一定能呈現出最佳效果，我們能夠在毫秒之間產出大量的長條圖、圓餅圖、線圖，但這些快速生產的結果常常會忽略了許多圖表必須存在的設計細節，也是本文主要分享的內容。

相對於使用軟體，我們有時候用紙筆更能描繪出精彩圖表，紙筆不會特別去限制你的創造力與視覺表達力，更沒有技術上的限制，有的時候我們因為對於軟體的不熟悉，或是基於『這樣就好』的素食概念，因而滿足於各類能夠快速生成的圖表，而非以最佳結果來思考。

以下是透過單鍵生成圖表的可能缺點：

- 無法生成最佳的呈現方式或比例
- 無法進行圖表價值說明，需要人類再行補充
- 有時會有視覺調整困難的技術問題

因此，建議讀者可以自己先思考自己覺得最佳的呈現結果，也可參考此篇文章分享的十項技巧，先畫出理想的圖像，之後再透過軟體實作出來，便可維持原先設計的脈絡主張。

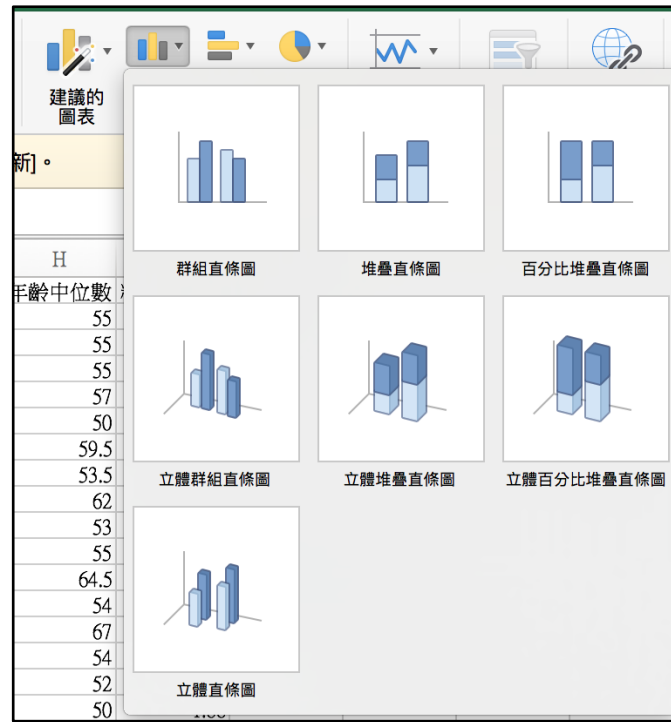


圖 6、透過軟體工具我們能夠輕鬆產出各類視覺圖表，但這真的是最好的作法嗎？

技巧一：【形式】選定想要揭露的資訊，選擇正確呈現方式

第一項技巧，是判斷我們是否使用了合適的圖表呈現方式，以本案例來說，X 軸是時間序列類型的資料時，通常使用線圖是較好的呈現方式，所以同樣的資料我們可以改成線圖呈現看看，更能夠觀察到其趨勢的變化。

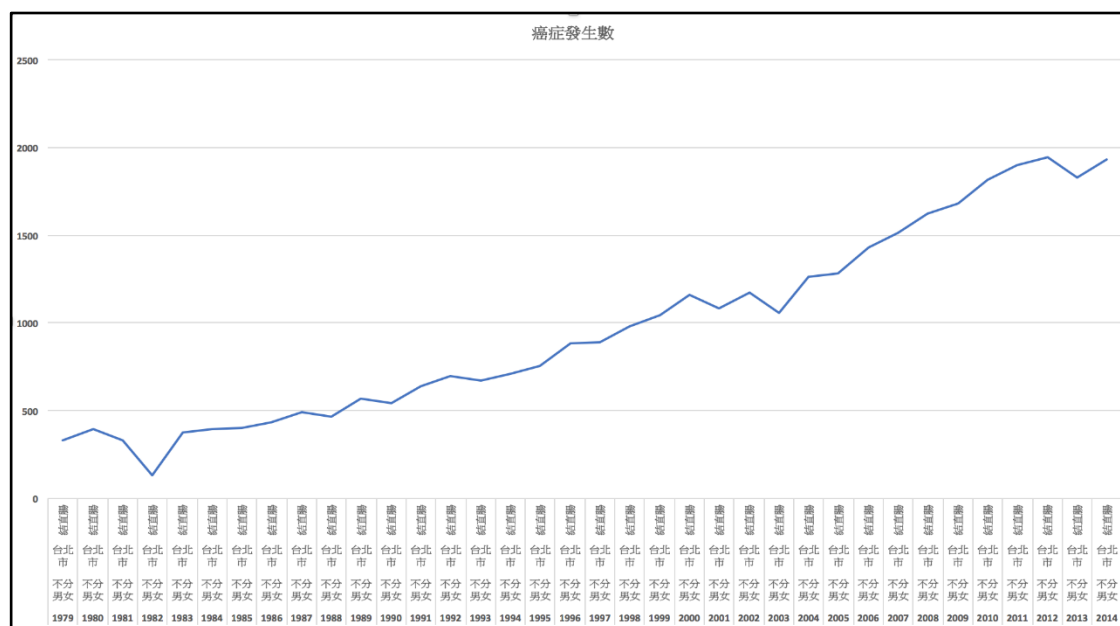


圖 7、以時間序列類型的資料來說，線圖是比較合適的圖像呈現

當然還有許多其他許多的圖表挑選準則，像是視覺化類型與差異等等，有在前幾個月的內容當中有分享過，讀者如果有興趣了解更多，可以前往以下文章進行瀏覽：

本專欄 5 月號文章：資料視覺化的類型與案例

<https://findit.org.tw/newsPage.aspx?pageId=374>

技巧二：【簡單】移除掉不重要、重複的資訊

第二項技巧，是找到畫面當中較不重要的資訊，或是重複性的資訊，如果這些資訊並不能夠帶給閱讀者更多洞見，可以考慮移除或是改變呈現的形式，讓圖表中每個元素均為不可或缺的重要資訊。

透過 Excel 自動產出的圖表，常常會出現許多累贅資訊，以本案例來說，下方自動出現了每個欄位的標籤，但其實描述的都是同樣的資訊，我們就可以直接將其移除掉，讓畫面更清爽，如果移除的是重要資訊，可以在其他的區域再行補充即可。

癌症登記	癌症登記	癌症登記	癌症登記	癌症登記	癌症登記	癌症登記
台北市	台北市	台北市	台北市	台北市	台北市	台北市
不分男女	不分男女	不分男女	不分男女	不分男女	不分男女	不分男女

圖 8、Excel 單鍵生成的圖表，自動在下方加入了重複性的欄位資訊

以下是去除掉重複資訊的圖表，是不是乾淨了很多呢？我們可以記得『LESS IS MORE』這句設計界的經典名言，當視覺被簡化之後，閱讀者可以更專注閱讀重要的資訊。我們可以重複閱讀圖表幾次，來確保沒有多餘的資訊，不過這裡還是要注意，移除掉不代表就不需要呈現，通常狀況是我們需要找另外的說明方式，兼顧畫面的整潔以及閱讀者的資訊需要。

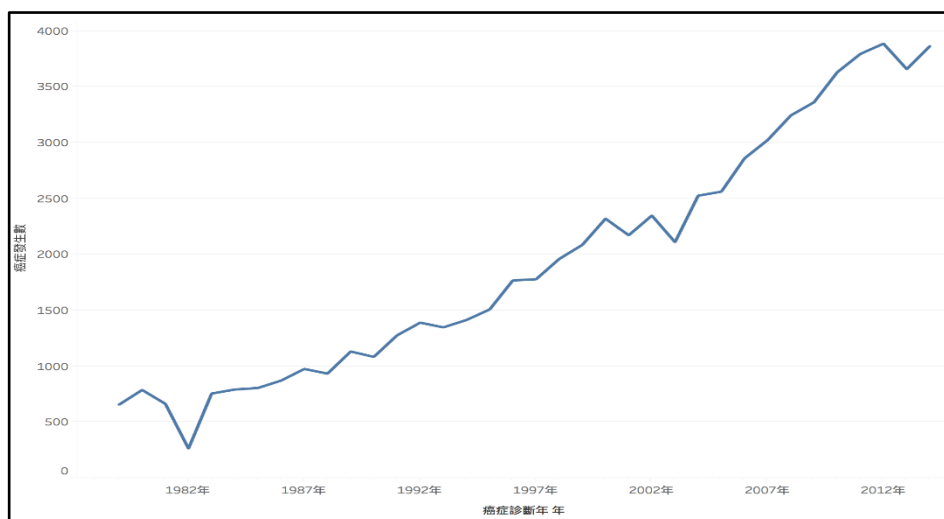


圖 9、移除掉畫面中的重複性資訊，視覺上會清爽很多

技巧三：【顏色】配置合適的、富含暗喻的色彩

另外可以帶來讀者閱讀愉悅感的，是顏色的配置，Excel 單鍵生成的顏色大多使用的是預設的顏色組合，通常不是最佳的結果，甚至有時候會使用了過多的顏色，導致閱讀者必須分神思考每種顏色所代表的意義，相較之下，人類自行配置的方式，較能根據語意或是呈現的隱喻進行顏色最佳化。

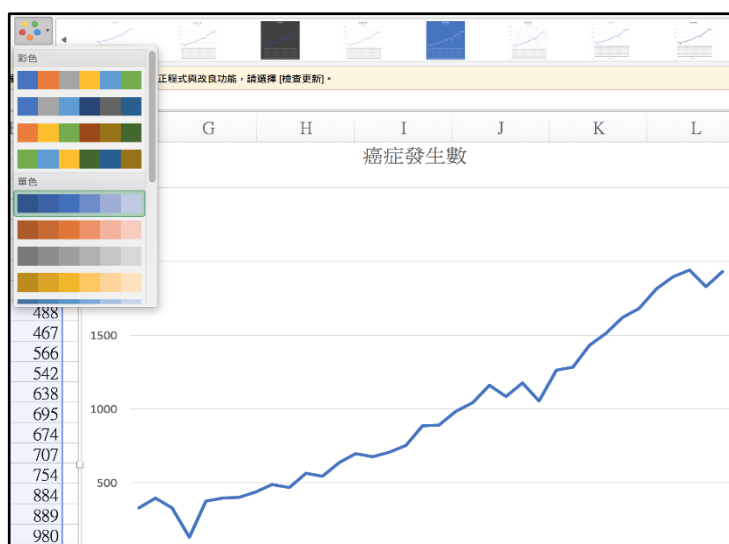
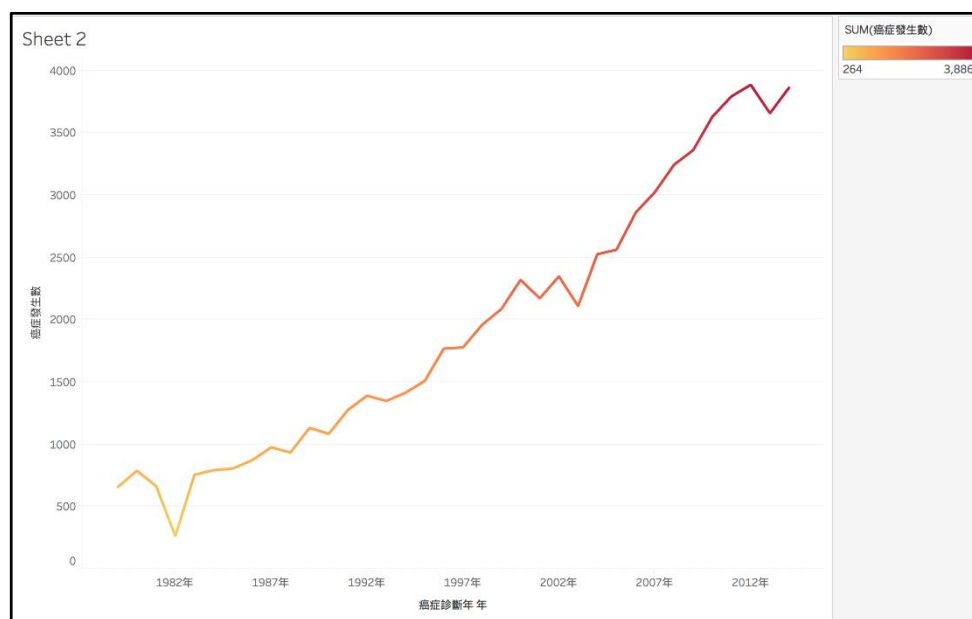


圖 10、透過軟體我們能夠輕鬆配置圖表使用的顏色，但通常非最佳的配置結果

好的顏色設計是有意義的，例如本案例我們使用較深的顏色代表較多的數據量，較淺的顏色代表較少的數據量，這裡有一些顏色配置技巧可以參考：

- 強調資料對比：使用強調、顏色較深的色彩，強調重點
- 強調資訊補充：使用較淡的色彩，灰色就很適合進行資訊補充，不會搶走主角風采
- 強調資訊群組：使用相同色彩或相近的顏色，讓讀者知道是同一組資訊
- 強調數字變化：使用漸層色彩，強化變化的方向與對比

圖 11、本案例在色彩上加入漸層，有暗示使用者數字改變方向的隱喻效



果

技巧四：【重點】用提示資訊 (Tooltip) 引導閱讀

對於設計者來說，他自己通常可以知道這張圖可以從哪些面向進行解讀，也能夠在心中套入相關分析的邏輯，但一般圖表設計很容易踩到的陷阱是，對於其他第一次接觸到這張圖的人來說，不容易抓到視覺圖表的重點，甚至容易解讀錯誤的方向，也造成閱讀上的壓力，如果設計者能夠適當的在圖中加入一些提示資訊 (Tooltip)，能有效的引導讀者進行閱讀。

然而，要注意的是提示資訊不能過多，我們可以考慮在圖中標記差不多 1-5 個重點，更多的話有可能反而讓讀者下意識的覺得此提示非重要資訊，造成反效果。



圖 12、加入提示資訊之後，閱讀者會下意識的關注那些被特別標記的資料點

技巧五：【標題】提供大標題與子標題

提供大標題與子標題能夠有效的提供讀者第一印象，閱讀者可透過標題與子標題概略了解這張圖想要傳達的主題，有種要進入別人家門的第一印象感覺，後續再細細品味圖表想要呈現的資訊，閱讀者可透過標題區啟動閱讀模式，之後再動用到創意右腦跟邏輯左腦。

一個好的標題區設計包括幾個重點：

- 清楚的標題視覺：是否跟內容區有明顯的分界？是否有清楚的文字顯示？對比是否足夠？是否加入了多餘的視覺裝飾（斜體、底線等）？
- 合適的比例配置：是否有設計大標題與子標題的階層？文字大小是否適中？
- 明確的文案設計：標題文字是否容易解讀？是否包括了此圖的重要意涵？是否用了易懂的用語？標題是否都有刪除贅字？

以下圖來說，清楚在標題當中點出了資料區間（1982～2012 年），可讓讀者有一個大概的概念『這是一個 30 年的統計資料』，並且把主詞也列在標題區：『臺北市、直腸癌』，可以有效的引導讀者釐清這並非全部的統計數據加總，而是一個分類之後的數字。

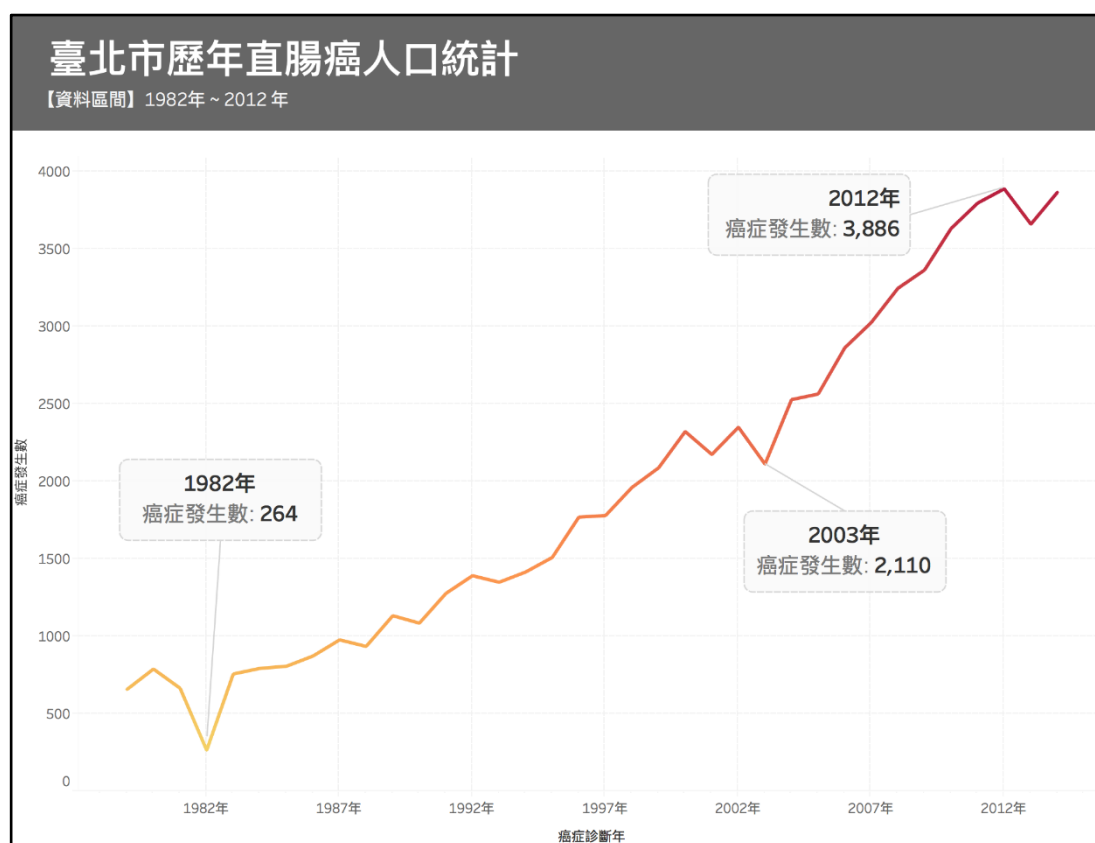
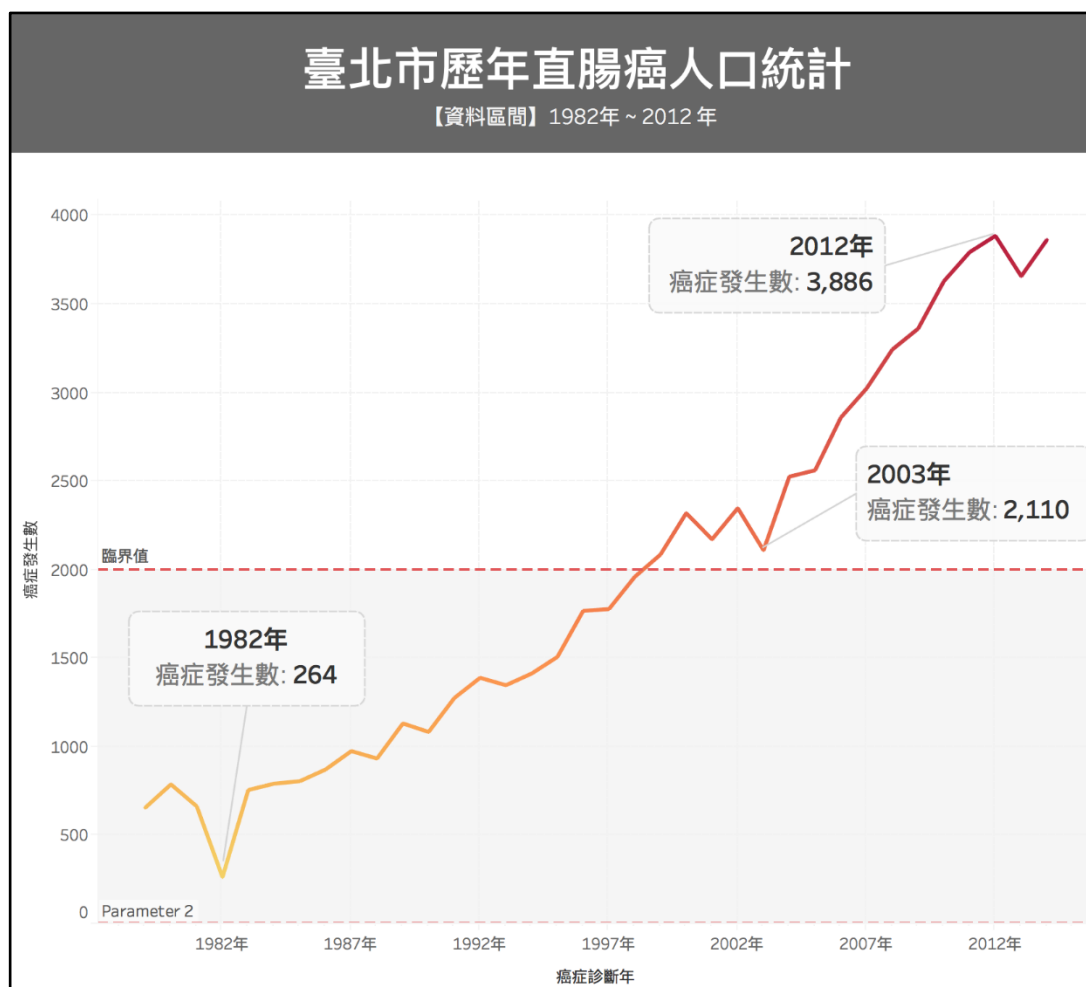


圖 13、加入合適的標題文字，能夠引導讀者掌握圖表的範圍，減少閱讀的壓力

技巧六：【輔助】透過參考線揭露特殊數字或趨勢

如果此批資料當中有一些特殊的歷史事件，或是重要的時間點、資料門檻等等，很適合在圖中加入參考線作為視覺參考，揭露這些數字或趨勢，好處跟提示資訊（Tooltip）的使用有些類似，都能有效的引導使用者進行閱讀，幫助使用者消化資訊。

以此案例來說，我們嘗試在圖中加入了一個『臨界值 2000』，讀者在解析這張圖的時候，就能夠快速的在腦中進行歸類，快速將此圖表切分為：



參考線 2000 以下的區域 (1998 年以前)

參考線 2000 以上的區域 (1998 年以後)

圖 14、透過參考線能夠清楚的輔助讀者進行資料的分類

技巧七：【信賴】標記資料來源

如果希望此圖表能夠更取信閱讀者，有一個非常簡單且高效的重要環節，就是『標記資料來源』，尤其當你的資料是可靠來源的時候，加入資料來源是不可忽略的重要細節。

以本案例來說，由於資料來源是政府單位（衛生福利部），可以將其放置在上方清楚的位置，對於閱讀者來說，知道這個統計並非來自於營利單位，會更信賴其結果，甚至會分享給他們的朋友、同事等等。

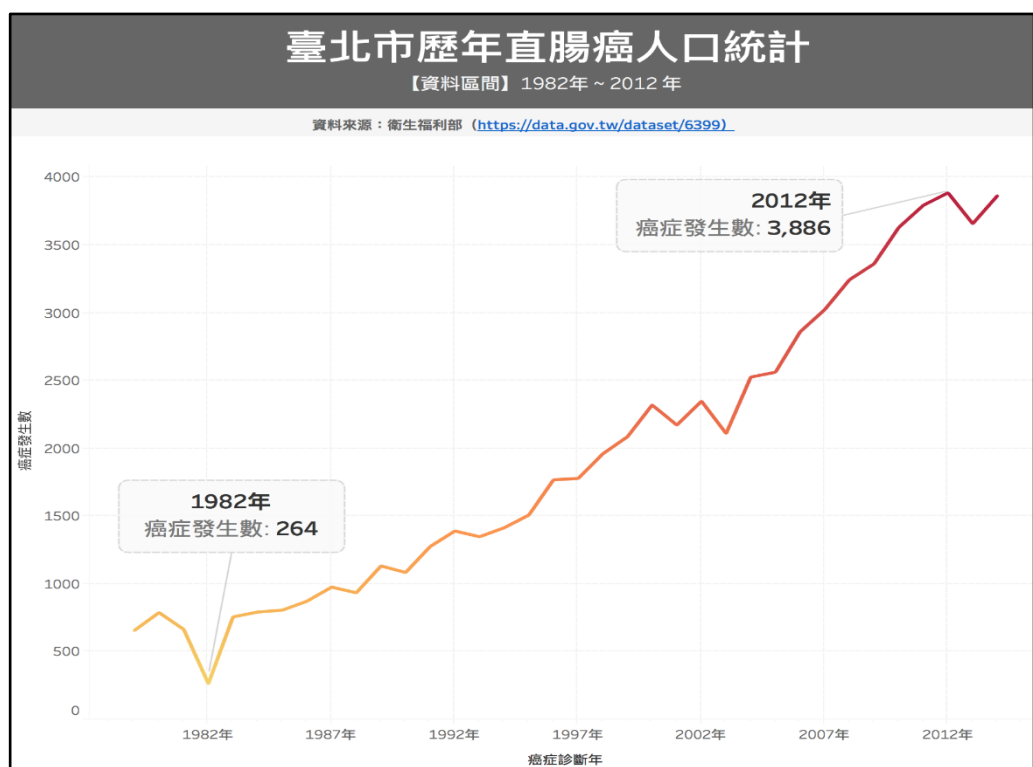


圖 15、僅僅是增添資料來源，就能夠強化此圖表的信賴指數

技巧八：【結構】調整適切的配置比例

一個標準的視覺圖表的結構上通常會採用配置的黃金比例，許多圖表都有參照此規則進行製作，已經變成視覺配置的淺規則，可參考以下的配置比例整理：

表 1、圖表結構配置比例

項目	比例	說明
標題區	10%	通常位於圖表最上方，展示此圖意義的重要區域，文字需要精練且視覺清晰，不宜過大以免搶走視覺區域的風采。
副標題區	5%	用來輔助說明主標題的資訊，通常可以加入一些引導性的文案，幫助使用者理解此圖。
輔助資訊	5%	來源資訊、圖表特殊情境說明等等，通常會用灰色文字來表達，位置可放在視覺內容上方或是下方。

視覺內容	80%	最大比例的區塊，視覺圖表的主角。
------	-----	------------------

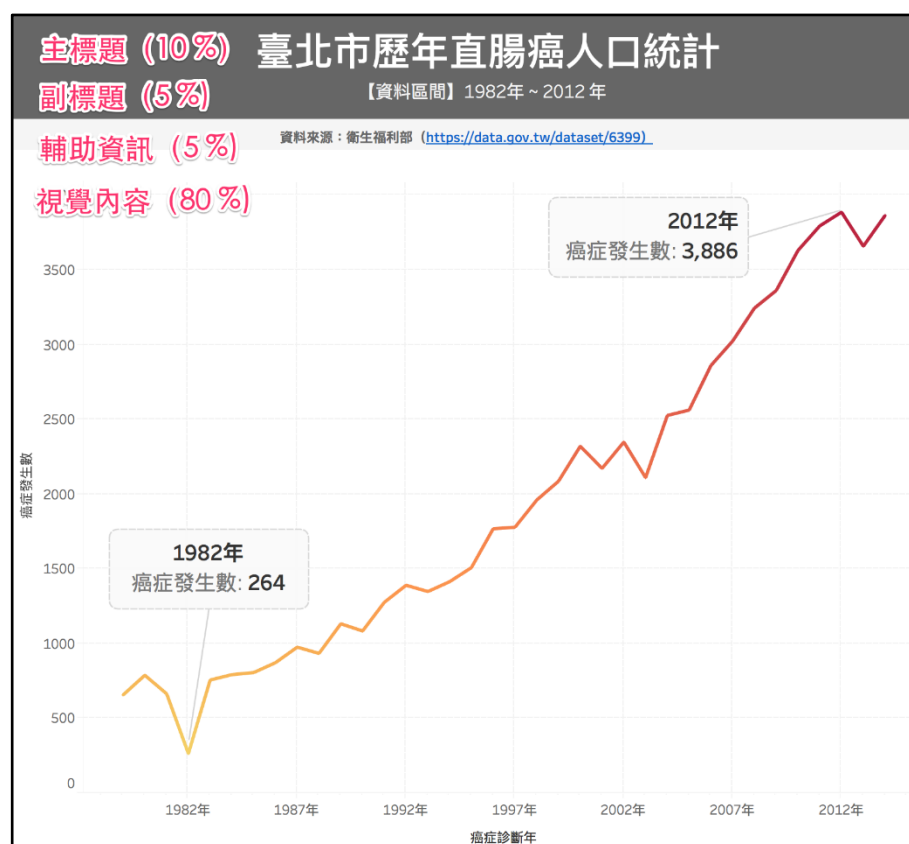


圖 16、一個標準的圖表配置比例（輔助內容可放在圖表上方或是最下方）

不過這裡要特別補充，以上比例只是一個大致的參考，只要記得在多數狀況時能夠參考此黃金結構配置，大概就能做出很不錯的成果，但並不需要每次做圖都拿尺去量測是否有完整的按照此比例設計，只需要記得大概的區域與比例樣態，並且『在必要的時候』可打破此規範，端看真實情境中想要呈現的目標。

技巧九：【趣味】添加互動或生成儀表板

除了一般靜態圖表之外，如果技術與呈現情境可支援的話，可適當的開放一些互動選項給閱讀者，將部份主控權留給閱讀者自由探索，能夠增加圖表的印象與趣味性。

以本案例來說，由於完整資料集包括各類癌症的發生比例，我們可以考慮開放其他病症選項讓閱讀者手動切換，觀察不同癌症的數據差異，甚

至可進一步將其包裝為視覺儀表板。

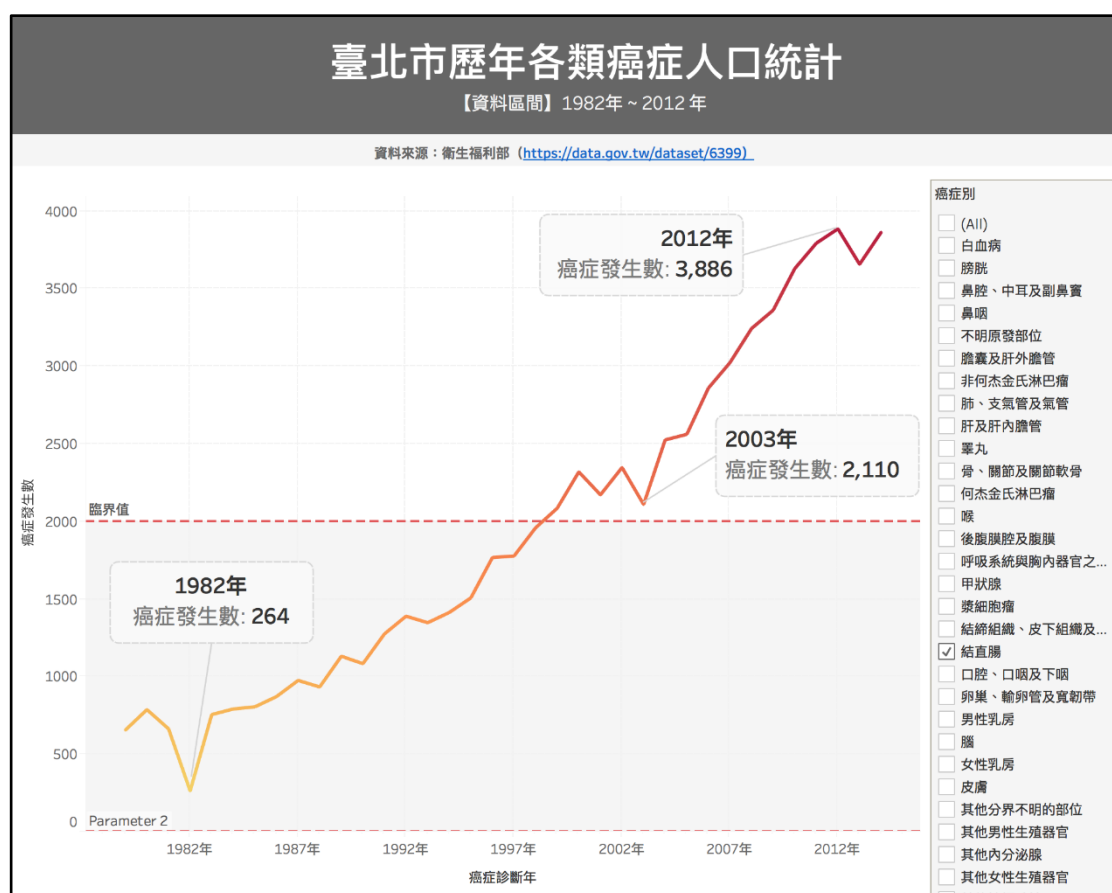


圖 17、開放部分互動選項給閱讀者，有趣又富含探索性

這裡要補充說明，這些互動性或是儀表板選項並非必要，我們可以視情況加入，如果是放在網站上的圖表頁面，還蠻適合加入這樣的選項，但如果最終成果是要直接說服人或是製作為簡報的話，就不建議用這樣的觀念來製作，反而容易混淆重點。

技巧十：【貼心】直接補充想要傳達的想法

最後一個技巧，設計者可直接在圖中補充想要傳達的想法，好處在於更快凝聚與閱讀者的共識，也知道該如何解析這張視覺圖表中的重要訊息，閱讀者可更快速啟動認知大腦與信任度評量，如果相關資訊是可信賴的，便可以快速理解並信任這張圖。

以本案例來說，如果我們想要強調其數字的成長變化，除了給予視覺

上的提升之外，也可以直接做一些數字提示，給予閱讀者一個更清楚的數字概念。例如我們有多種文案可以選擇：

#1：臺北市的罹癌比例不斷提升（讀者：哪種類型的癌症？）

#2：臺北市的直腸癌罹患比例不斷提升（讀者：提升多少？）

#3：臺北市的直腸癌罹患比例，30 年來增長了 14 倍

想法文案的設計是種哲學，也不一定有標準答案，我們可以反覆琢磨，找出最能夠打動人心的文字，找到最能夠取得閱讀者的信賴的語句，這也涉及到設計者與閱讀者之間知識交流的默契。以此案例來說，因為我們希望能夠強調數量的改變，所以乾脆直接算出一個數字比例，並呈現在畫面上，讀者就可以與設計者同步，也能減少資訊消化的壓力成本，也更能夠被這個圖表說服。



圖 18、直接在圖中放上解析想法，閱讀者可更快掌握重點資訊

想法傳達文字不需要太多，重點在於能夠讓閱讀者感興趣，了解並信任你的解讀結果，以下提供幾個重點參考，我們可以透過自問自答的流程來找到好的描述方式：

- 我想要指出這張圖的重點是什麼呢？
- 我希望說服閱讀者的資訊為何呢？
- 閱讀者期待看到的訊息是什麼呢？

結論

視覺圖表有許多呈現方式，透過軟體單鍵生成視覺圖表速度很快，但多數狀況並無法直接產出最佳的呈現結果。

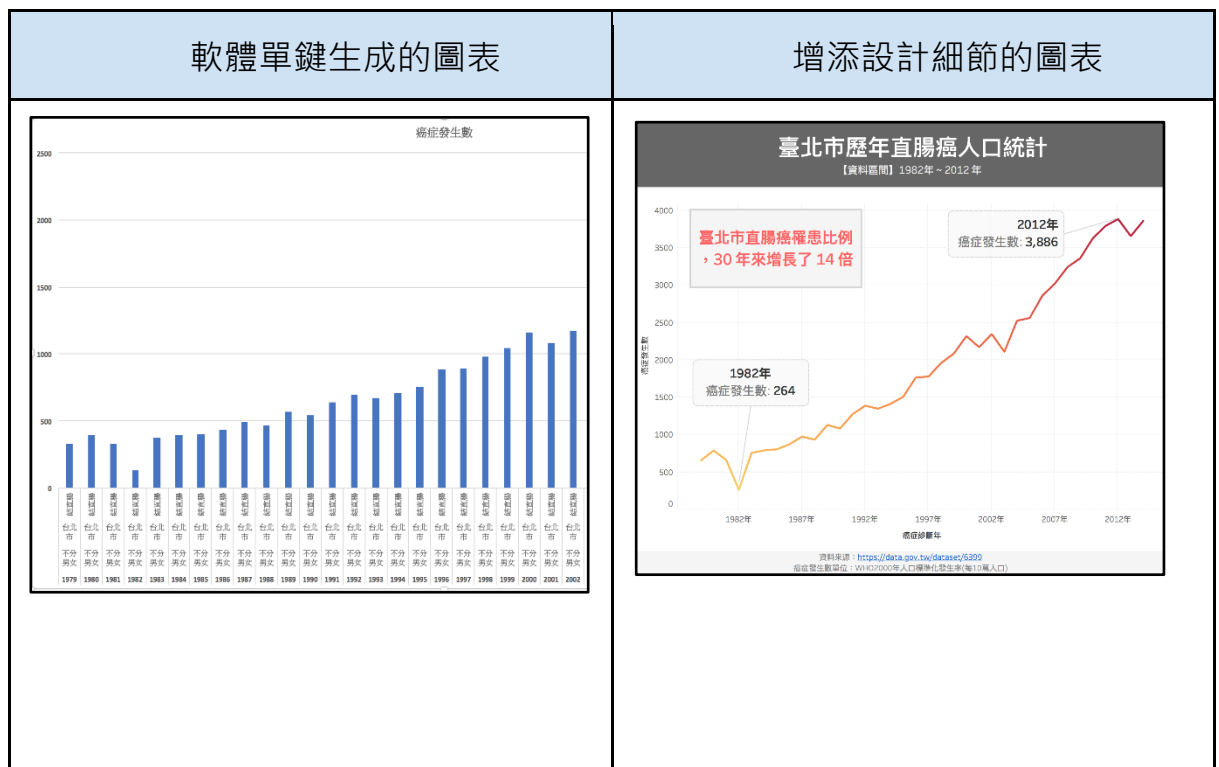


圖 19、軟體生成圖表 v.s. 增添設計細節的圖表

製作視覺圖表的重點包括了許多層面，像是理解溝通的脈絡，究竟我們的對象想要知道的是什麼？我們是否已經最大化滿足他們的期待？本篇文章分享了 10 個視覺化圖表優化的絕招，其實相關技巧都不難，許多內容也都是讀者們已經了解的訊息，但有時候進行圖表編輯的思維之後，有時後會忘記這些簡單但是重要無比的概念，結果做了一個不夠好的呈現結果。一個好的視覺圖表是富含故事性與啟發性的，除了分析師思維之外，我們還需要注入設計師思維的視覺技法，以及傳道者思維的說故事技巧，利用許多視覺隱喻與優秀的文案，一步步引起聽眾的注意，並在最終贏得

閱讀者對於圖表的信任。