

資料感知

圖 1.2 表現的是根據所學領域劃分葡萄牙的大學生¹。現在，問題來了：這張表有味道嗎？是苦的、甜的還是酸的？

葡萄牙高等教育學生
所學領域比例（%）

| 領域 | 1998 | 2012 |
|------------|------|------|
| 教師訓練與教育科學 | 11.9 | 5.7 |
| 人文學科和藝術 | 8.4 | 9.5 |
| 社會科學、商學和法律 | 38.3 | 31.3 |
| 科學、數學和資訊 | 8.9 | 7.2 |
| 工程、製造業和建築 | 18.9 | 21.9 |
| 農業和獸醫 | 3.0 | 1.9 |
| 健康和福利 | 6.9 | 15.9 |
| 服務業 | 3.7 | 6.4 |
| 不詳 | 0.0 | 0.1 |

高等教育：ISCED97 5-6 級
來源：Eurostat

葡萄牙高等教育學生
所學領域比例（%）

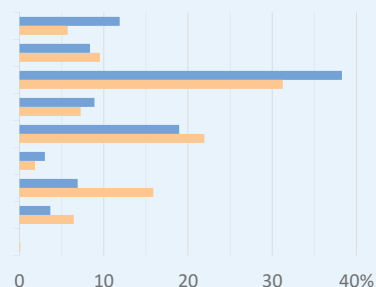


圖1.2 各學習領域的學生比例。

我換個方式來說，讓你更容易理解：是否有可能拿食物來重製這張圖表，讓我們感受到每個數值的不同味道？我可憐的烹飪本領告訴我，可以，我們可以把較低的數值做成苦味，較高的數值做成甜味，然後布置成一張，嗯，披薩。雖然我們可以用其他感官（如聽覺、味覺或觸覺）來製作量性資料的感覺，但這樣好嗎？

不怎麼好。視覺是我們最成熟的感官，在大腦所有的刺激訊號中，佔大部分。換句話說，比起其他感官，大腦用大部分的資源來處理視覺資料，因此利用視覺來閱讀圖表會比用味覺、嗅覺或觸覺更簡單、更精確。

人是視覺動物。我們很清楚眼腦系統的威力，所以我們在想像外星人的型態時，我們通常會想像出有大眼睛與大腦袋的生物，鼻子消失、耳朵不見，味覺與觸覺低下（圖 1.3）。這些從外太空來的生物，遠比地球深海底下視覺退化的生物，還要像我們人類。



圖1.3 我們想像的外星人：大眼睛、大腦袋。

1 本書中使用真實資料，我選擇資料序列的標準，是其變化是否有趣，而非其代表的意義。我很想選一個不為人知、有複雜統計系統的的加勒比海國家作為主要的資料來源，以求在不受成見限制下，得到全新觀點真實資料的好處。然而我做不到，找不到這樣的資料來源，我只能選用美國與歐盟的資料。

那如果我們用座標系統來畫出地形呢？

這張新的照片（圖 1.5），看起來像是前一張照片的簡化版。我們可以看到山與河，但就像 **Magritte** 所說的，這不是煙斗。事實上，這既非風景亦非煙斗；這是 **Excel** 中做出的累計面積圖。我們知道，這種圖常用來表示某些市場的銷售變化。

我們把照片中的藍色線條解讀成河流，所以我們同樣可以在這張圖表中辨識出圖案，以便閱讀、認識與行動。形狀的辨識與位置的意義，在兩張照片、兩幅風景中相似。從這個觀點來看，資料視覺化充其量不過是依據資料表格，建立圖形地貌。

如果說兩張圖片中有什麼差別，這個差別就是我們所製作出的圖形地貌，在刺激訊號與細節上，無法和第一張照片一樣豐富細緻。添加有意義的細節、妥善管理，是資料視覺化最大的挑戰之一。

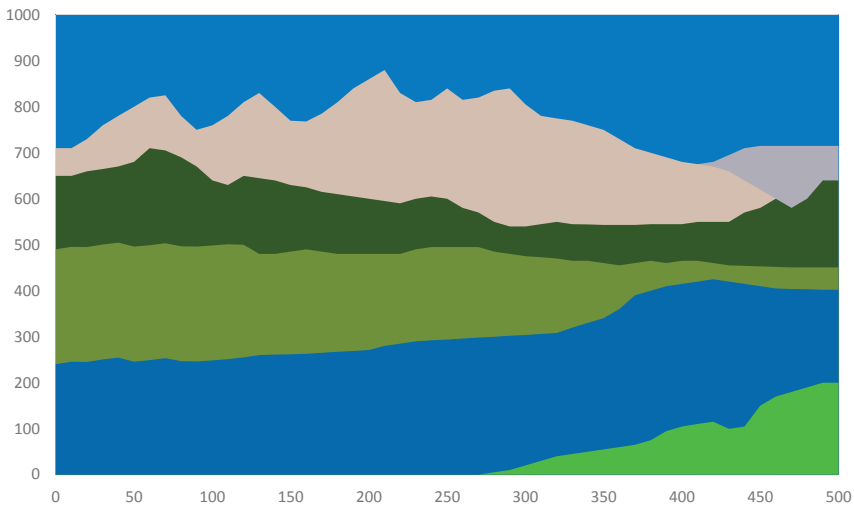


圖1.5 Excel面積圖。

檢查圖表符合特定需求的程度，最簡單的方法之一，就是它是否提供良好的創見，或者我們一眼就可以解讀出來？我們來試試看。

大專學生人數在 1998-2010 年間成長了 9%

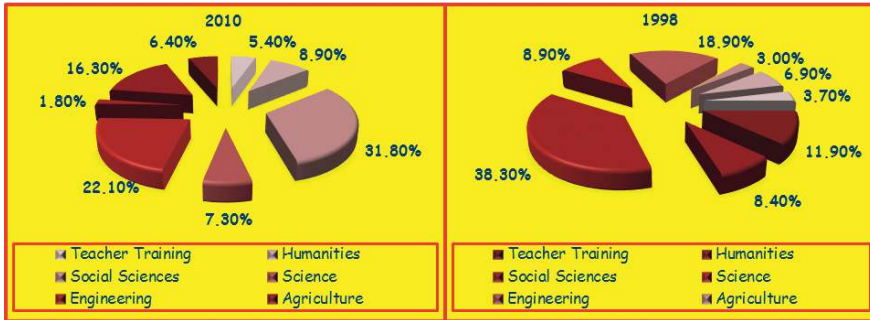


圖1-16 我稱此為圖表怪人 (Graphenstein) - 一張因各種不良設計而生的圖表，這些不良設計，來自各種真實的圖表例子。

圖 1.16 是一個常見（且醜陋）的圖形呈現形式。它應當可以讓我們比較兩年之間的比例。依據你的解讀，回答以下這些任何人都可能提出的問題：

- 哪個領域成長最多？
- 哪個領域衰減最多？
- 人文的表現怎樣？
- 有多少領域成長？又有多少領域衰退？

難嗎？需要多花點時間？信不信由你，這兩張圓形圖包含了以上所有問題的答案。

然而圓形圖有其矛盾之處。支持圓形圖的常見論點之一，是可以用閱讀標記彌補難以準確閱讀圓形圖的問題。從這些圖表可以看出來，這種說法並非祖護圓形圖，而是為視覺化的退步而辯護。難道我們必須先解讀所有標籤，然後才能閱讀圖表？如果我們必須閱讀標記加上圖表，那圖表就變得沒有重點，因為標記應當是補充作用，而非絕對必須的。

不良的圖表選擇與審美選擇，會使得圖 1.16 的呈現喪失作用，但傷害可能不僅於此。資料收集、資料準備、圖表設計，以及觀眾為此花費的時間，雖然難以估計，但卻是真實、無法回收的。如果我們考慮到圖表作者可能會花更長的時間去為圖表「調味」，而非從頭開始做出好圖表，這就更荒唐了。

這個範例說明了，使用相同的資料表格，你仍可創作出不同的圖形呈現、產生截然不同的成效。至於找出我們所接受任務的目標與答案、找出最適當的視覺化方式，還是得靠我們自己。

當你對原形圖表加上變化，這就是在進行設計選擇。變化的標準應當要清楚，且契合任務與觀眾。根據經驗法則，你加上越多的設計維度與變化，圖表的效力就會越差。

但就像許多資料視覺化法則，你一定可以找到例外：當一張擬 3D 圓形圖與一張樹狀圖（圖 1.18）都加上大量變化，這些變化會使圓形圖變成含糊的漸層，而樹狀圖則可以表現出明顯的用意。⁵如果能遵循人類感官的運作去添加變化、為圖表轉化成更好的悟性與意義，就會是良好的影響。負面的變化與設計選擇，通常是因為欠缺這些感官機制的悟性，或者為了所謂「高衝擊性圖表」之類的口號而強加過度的需求。

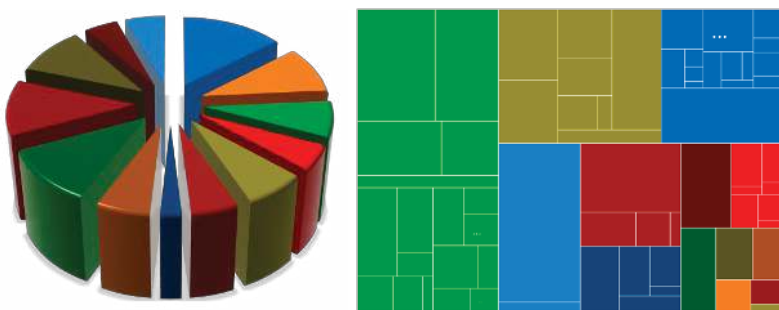


圖 1.18 過多變化可能會造成負面影響（圓形圖），也可能有正面影響（樹狀圖）。

5 樹狀圖與圓形圖類似，用於局部分析，但由於長方形的面積比扇形更容易掌握，能用的資料點更多。不同於與傳統圓形圖，你可以安排資料階層。把長方形與所有資料點比較，或者與其本身的分枝比較。在大多數實作狀況下，你可以將填上的顏色與連續性變數建立關連，以便觀察如產品市價（長方形面積）與成長率（長方形填入的顏色）。

眼睛生理結構

現在我們來談點眼睛的生理結構，進一步瞭解該如何設計視覺呈現。

當物體發出光，或者表面反射光線，就成為可見物。光，可讓我們分辨物體、識別其性質。眼睛是我們捕捉光刺激訊號的器官。圖 2.3 是基礎流程。

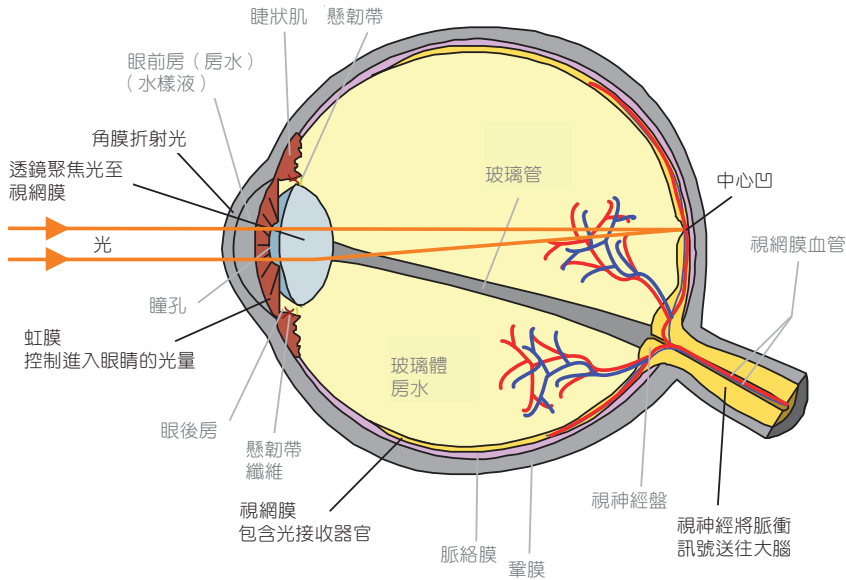


圖2.3 眼睛捕捉光，轉換成電化學脈衝訊號

來源：Wikicommons。修改自 https://commons.wikimedia.org/wiki/File:Schematic_diagram_of_the_human_eye_en.svg

視網膜

視網膜是眼睛的感光區域，覆蓋眼球內部表面超過三分之二。光折射聚焦在視網膜弧形中央的 18° ，也就是黃斑部，在此處從外圍的低解析度視覺逐漸轉換成高解析度的視覺。我們可以把更靠中央的弧形處定義為盾心，也就是眼球內面最中心的位置。

圖 2.4 的圖表中，橫軸表示這些弧面的弧度與其相對比例。圖表中也畫出了視網膜中，中心點兩邊各 50° 範圍的光感細胞分布。

完形法則 (Gestalt Laws)

如果仔細觀察圖 2.16，你會發現三角星座（左上角插圖）準確地描述了三顆星星的位置，與圖中其他部份格格不入。至少對我們來說，圖中其餘部位都顯得有些古怪。試試看找到蒼蠅（Musca）星座吧。⁶如果你不知道這些星星要怎樣才能變成蒼蠅，沒關係，這不是你的問題，但你一定同意，這是因為這幾顆星星太過緊密，構成了一個與其他星星分隔的群組。

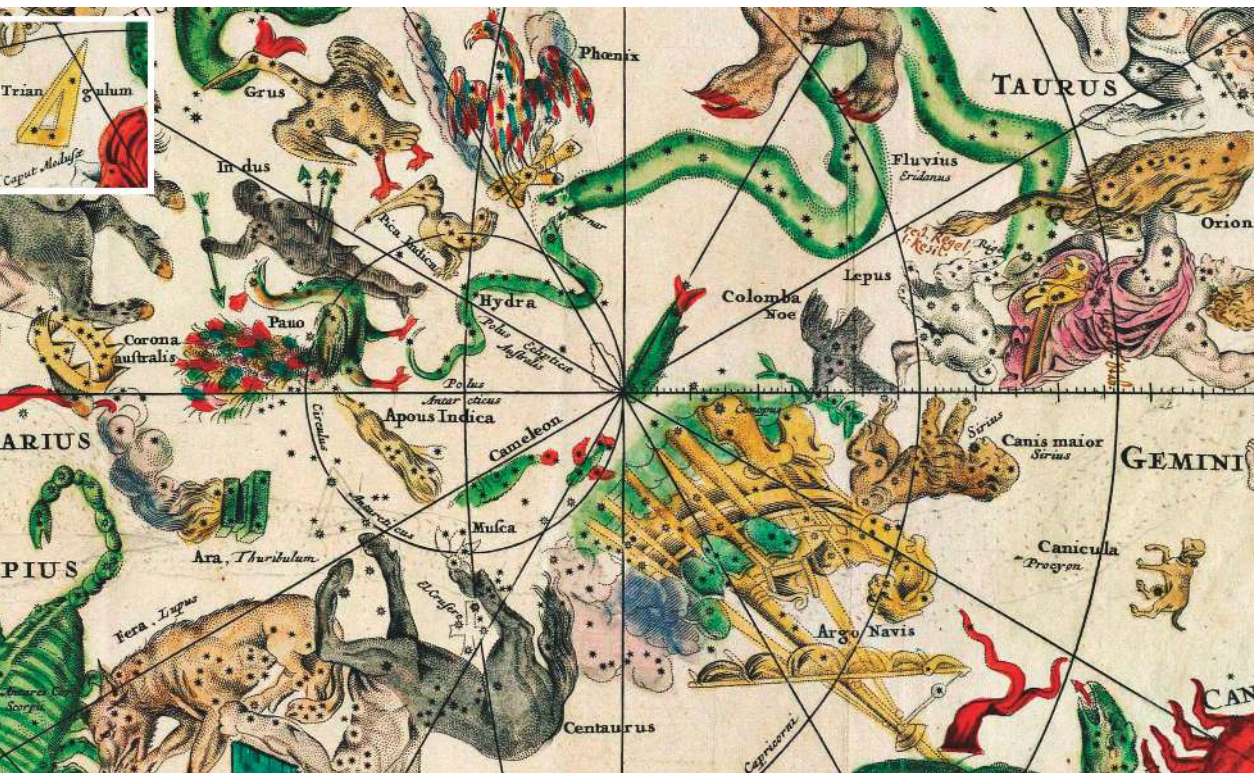


圖2.16 平球論 (Planisphere) 星座圖。17世紀荷蘭繪圖師Frederik de Wit製作的天文圖。來源：Wikicommons。

⁶ 提示：位於中央圓圈底部，靠近人馬座 (Centaurus) 的一個小星座，讀音近於「Mufca」。

呈現時間

在西方文化中，呈現時間的方式在這個社會中是共通的。時間從左（舊）到右（新）流動。在圖 3.4 的左邊，我刪除了所有標記。假設「正常」的時間流動是從左到右，你可能會做出兩個序列都有向下趨勢的結論。事實上你猜錯了。右邊的版本加上標記後說明結果正好相反；兩個序列都呈現向上的趨勢。

在這個範例中，讀者可能會懷疑圖表違背了慣例，因為寬頻網路連線不可能有向下的趨勢。在沒有預期的狀況下，圖表可能沒有這個警告符號，使我們造成與資料所述相反的結論。所以確實有此風險，讀者無法依據感官做出正確的結論（在此範例中，若從左到右閱讀，資料值是下降的），直到閱讀橫軸的感知作業，才會改正。這是資料視覺化的原罪，你應當想盡辦法去避免：在一張圖表中，感知作業負責彌補感官作業，但絕對不是加以修正。唉，如果讀者無法信任你的圖表，那做什麼簡單易讀的圖表還有意義嗎？如果你的讀者抱持了「信任，但要確認」的想法，那你就必須確保「信任」這部分絕對成立。



下載原始圖表

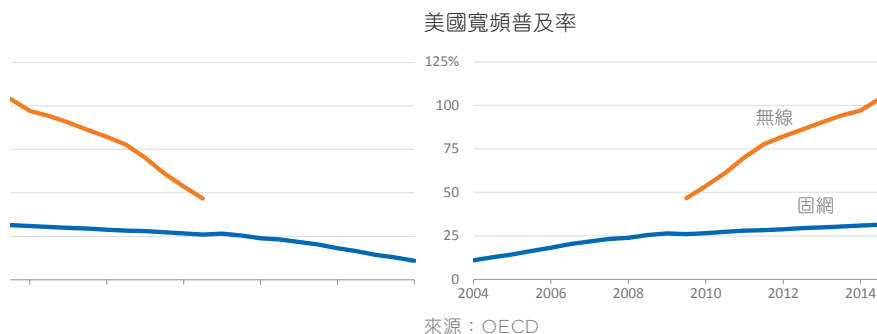
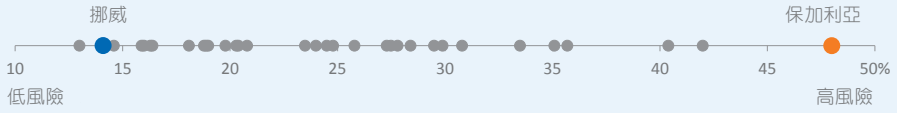


圖3.4 時間往左流動，可能會誤導讀者。

陷入貧窮或社會脫節的人口



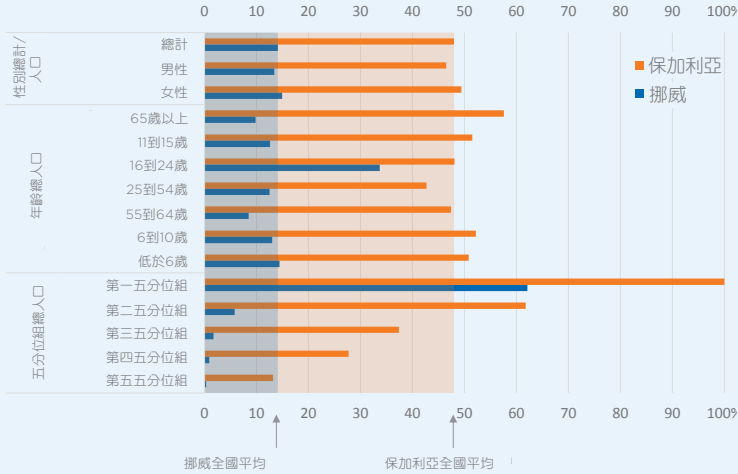
Source: Eurostat

圖3.6 沿橫軸比較各國。

下載原始圖表

2013年貧窮或社會排斥的風險

挪威與保加利亞的比較



來源：Eurostat

圖3.7 具備詳細資料的長條圖。

下載原始圖表

陷入貧窮或社會脫節的人口



圖3.8 竹管圖：兩性與全國平均值的差。

下載原始圖表

陷入貧窮或社會脫節的人口

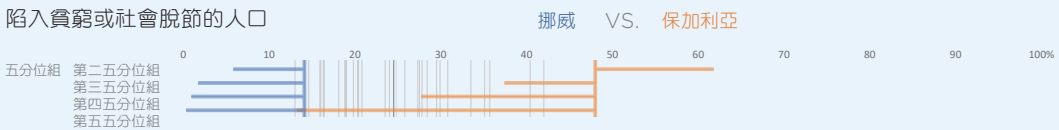


圖3.9 竹管圖：收入多寡與全國平均值的差。

下載原始圖表

最重要的一點，一定要在投影機與螢幕上測試你的視覺化作品，才能用於簡報。（順便一提，儀表板與其他會分享到其他電腦上的視覺化，也是如此。）

智慧型手機與直幅螢幕

請上網

紐約時報的資深美術編輯 Hannah Fairfield，表示該報網站流量有過半來自智慧型手機，這迫使他們得重新思考其報導的結構，與印刷品或大型電腦螢幕不同。

那些設備上可以同時呈現大量資料讓用戶探索，小螢幕代表空間必須謹慎使用，首先就是刪除任何不需要的功能。分層（layering）與定序（sequencing），代表從一個畫面移動到下一個畫面，就像是閱讀一段一段的文字。其間的轉換必須平順，連接也必須自然。

在資料視覺化，直幅顯示很好。並非因為我們要改變一般的長條圖或折線圖長寬比，而是因為直幅顯示的寬度足以顯示這些圖表（圖 7.6），把下半部螢幕留給文字或第二幅圖表。至於橫幅的顯示，我們會放大圖表，塞滿螢幕。

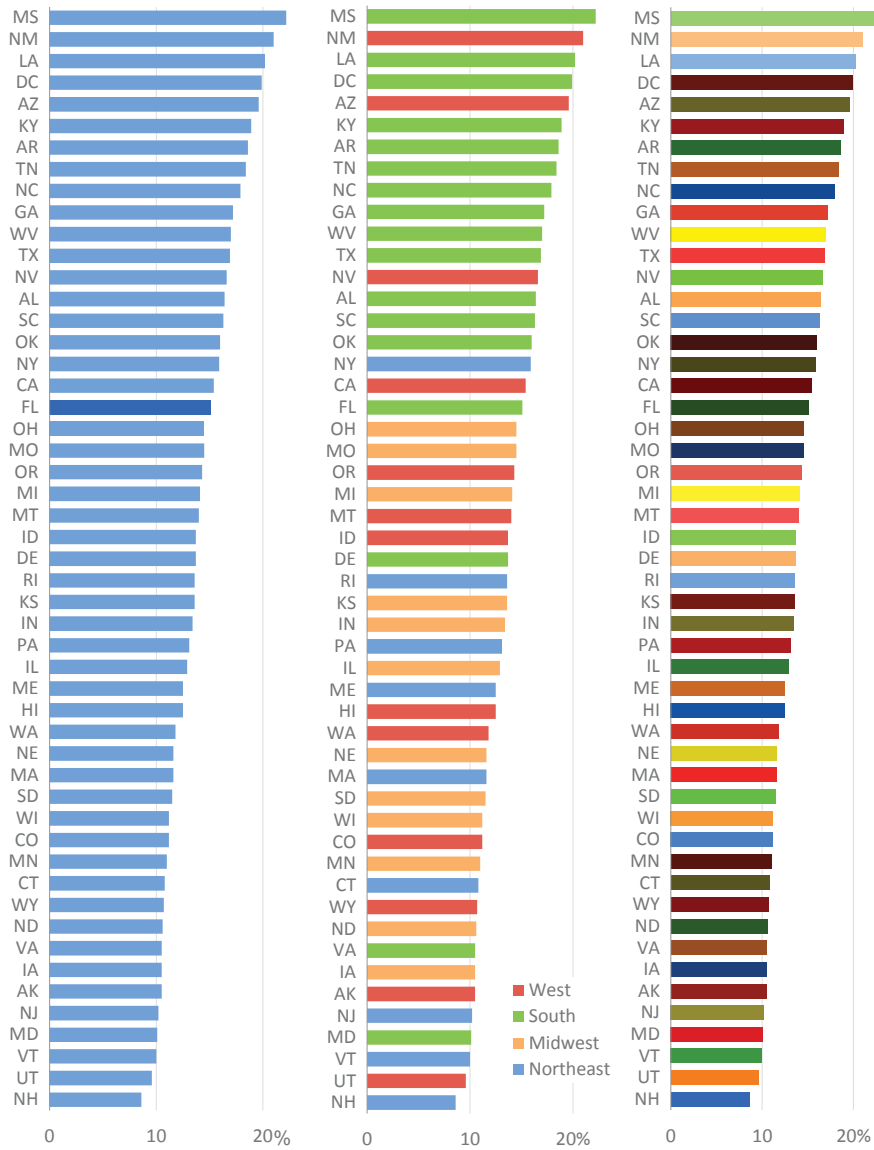


圖 7.6 直幅螢幕可以放超過一張圖表。

PDF 檔

我們在第 4 章探討了用 PDF 分享表格與原始資料這種沒有意義的舉動，但如果你想要分享靜態視覺化，且要與你原本設計的模樣相同，那 PDF 是個很好的選擇。PDF，至少比 PNG 或 JPEG 之類的圖檔格式要好多了。

2012-2013 各州貧窮人口比例



下載原始圖表

來源：美國人口調查局

圖8.4 長條圖中使用顏色的三種方式。

細條圖會自動比序，所以這又是一個不用擔心的問題。細條圖佔據的空間也比長條圖少得多，雖說標記文字有點棘手。我相信細條圖會越來越流行，只要我們能解決每個資料點都需要標示的問題（互動性可以視需求提供標示，讓我們著重於比較有趣的資料點或叢集）。

碼表圖

開車，也許就像管理企業組織一樣，但你若是一切都比照辦理，那就有問題了。要是在總儀表板上用了碼表圖與其他不適合的圖表，那問題可就大了。如果你用 Google 搜尋「總儀表板」(executive dashboard)，就可以看出問題有多大（圖 8.15）。



圖8.15 你不需要用這種儀表板來管理公司吧！

Google 圖片搜尋：Google 與 Google 標誌是 Google, Inc. 的註冊商標。

累積長條圖 (Stacked bar chart)

請上網

圖 9.16 中的累積長條圖，是歐盟的一份報表。這張圖表有許多問題，包括效率等級的反向次序、說明文字放在圖表中央、配色，以及欠缺資料。除這些問題之外，這張圖表示範了累積長條圖中的疊加式長條有多難閱讀與比較，不光是這張圖表，而是所有具備超過兩個分類的組成圖表。

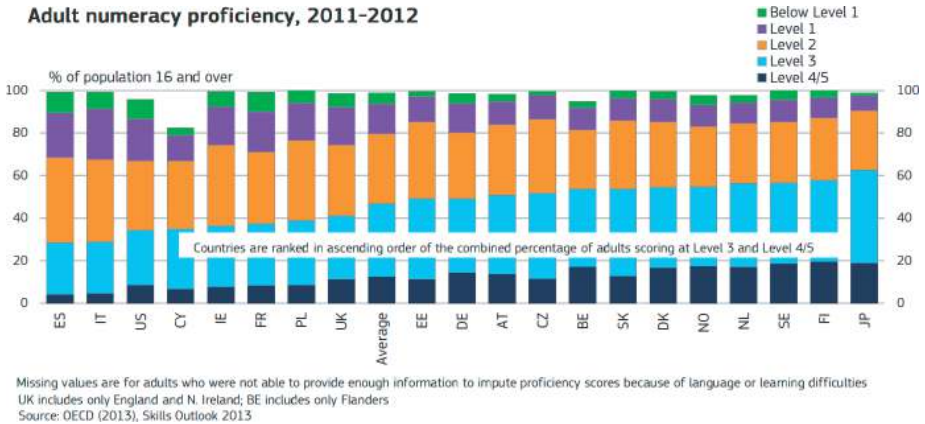


圖9.16 在這張累積長條圖中有許多問題。

在大部分狀況下，我們沒興趣將一個資料點與整體比較，反而比較有興趣把各資料點互相比較，這在本範例中變得更明顯了。這張資料表格的正確圖表，必須適合比較各素養等級，並能夠在各個等級之內進行比較，如圖 9.17。

這張面板圖 (panel chart)⁸，除了不需要對每個等級標上不同顏色，也顯示了只有在各序列對齊時才能看出來的差異。利用顏色，一如先前標示東歐國家的作法，我們就可以看出在等級 2 與等級 3 的值較高，等級 4/5 較低。

8 在第 13 章會探討面板圖與類似圖表類型。