

How To Lie With
Statistics

德瑞爾·赫夫 Darrell Huff — 著
閻翊均 — 譯

統計操控 真相與謊言 的

別再讓數字騙了你！
一眼看穿投資詐局、不實廣告與虛假民調，
打造最強自我保護力與決策判斷力

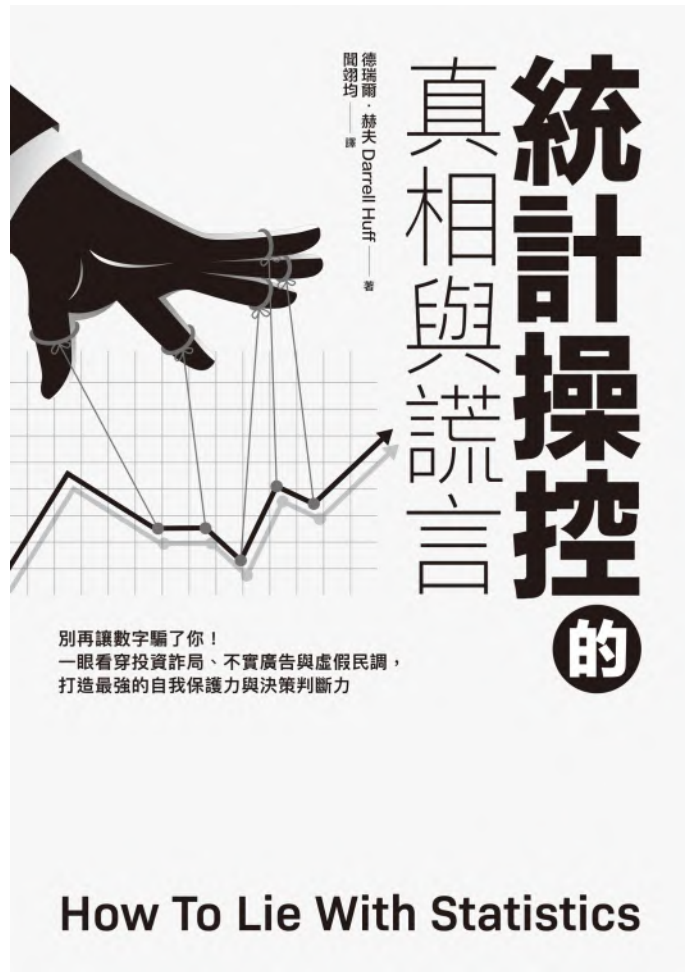
微軟創辦人
比爾·蓋茲「閱讀書單」推薦！
投資大師
肯恩·費雪權威背書！

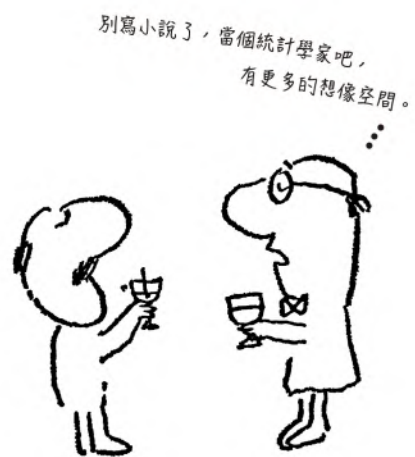
【專業推薦】

FC「FC的碎碎念」版主
劉奕西《高勝算的本事》作者
劉彥廷 教育工作者 / 思辨講師
賴以威 台師大電機系副教授、數感實驗室共同創辦人

暢銷
逾**50萬冊**
全球經典鉅著
★★★★★

TED
超過**260萬**
人次瀏覽
★★★★★





*本書插圖為梅爾·卡爾曼 (Mel Calman) 繪製。

謊言有三種：謊言、該死的謊言和統計數據。

——英國前首相迪斯雷利 (Benjamin Disraeli , 1804 ~ 1881)

總有一天，統計思維會像讀寫能力一樣，成為人民生活的必要條件。

——著名小說家、《世界大戰》 (*The War of the Worlds*) 作者，
H · G · 威爾斯 (H. G. Wells , 1866 ~ 1946)

真正使我們陷入麻煩的，並不是我們不了解的事物；而是我們自以為了解，但事實卻並非如此的事物。

——著名作家，阿特彌斯·伍德 (Artemus Ward , 1834 ~ 1867)

整數永遠都是假的。

——著名作家，塞繆爾·詹森 (Samuel Johnson , 1709 ~ 1784)

我想要編寫一個偉大的主題 (統計學) ，但我強烈的感覺到我的文學能力不足，因為我若想要將統計學寫得容易理解，就不得不犧牲準確性與完整性。

——博學家、心理學家，法蘭西斯·高爾頓爵士
(Sir Francis Galton , 1822 ~ 1911)

前言

隨著英國的英寸、英尺和磅等古老計量單位終結，民意調查機構蓋洛普（**Gallup**）想知道人們對取而代之的「公制」（**metric system**，現行國際化的十進位度量系統）了解多少。他們按照平常的方式詢問調查，發現即使是上過大學的受訪者，也有**33%**的人從未聽說過公制。

隨後，一家週日報進行了自己讀者的調查，並宣稱有**98%**的讀者都了解公制。該報甚至自豪的表示，這顯示它們的讀者比一般人「知識淵博得多」。

兩次調查的結果為何落差那麼大？

蓋洛普的調查精心挑選了不同階層的民眾並與其交談；而週日報則天真且經濟的依賴讀者剪下、填寫並回寄的優惠券廣告回函。

不難猜測，大多數不了解公制的報紙讀者對公制或優惠券沒什麼興趣，因此不會費心剪下回函並參與調查，等於是自行選擇被排除在調查之外。從統計學的角度來看，這種自我選擇產生了帶有偏誤或不具代表性的樣本。多年來，正是這樣的樣本導致了大量具誤導性的結論。

數年前，數十名調查員分別提出關於抗組織胺藥物的數據報告。每個報告的數據都顯示，在使用藥物後，能治療大部分的感冒。這引起一陣騷動，至少在廣告界是如此，同時也推動了醫療產品的熱潮。造成這種現象的原因有兩個，其一是人們找尋治癒感冒的渴望；另一個原因，則是人們在現實中對統計數據有所抗拒，不願真正看見存在已久的科學事實。就像不具醫學背景的幽默作家亨利·G·費爾森（**Henry G. Felsen**）所指出：「適當的治療可以在七天內治癒感冒，

但如果放著不管，感冒則會持續一週。^{*1}」

正因如此，許多你時常讀到或聽聞的統計相關事物，像是平均值、關聯性、趨勢和統計圖表等，其實並沒有表面上那麼單純。在統計的表象背後，可能沒有什麼實質內容，也可能有更多資訊。



在注重事實的文化中，統計學的神祕語言非常吸引人，時常被做為製造轟動、煽風點火、混淆視聽和過度簡化來使用。在報導社會與經濟趨勢、商業狀況、民意調查與人口普查等大量數據時，統計方法與術語已成為必要工具。然而，若報導者沒有充分理解並秉持誠信來使用這些詞彙，加上讀者又不了解詞彙的真實含意時，最終只會得到語意理解錯誤的結果。

在一般談論科學的文章中，被濫用的統計數據，幾乎完全取代穿著白色實驗袍、在昏暗實驗室裡加班工作的英雄形象。因為統計數據宛如化妝工具，讓許多重要事實變得如妝扮後的人，成為截然不同的模樣。事實上，經過良好包裝的統計數據，遠比希特勒的「大謊言」（big lie，扭曲事實做為政治宣傳）更美好；這些數據會造成誤導，卻又不會讓人受到批評或究責。

你可以這麼認為：這是一本用統計學來騙人的初級讀物。整體來說，它看起來可能有點太像詐欺指南了。或許我可以用一名退休竊賊的比喻來為本書辯護。這名竊賊出版的回憶錄，就相當於一門教人如何開鎖與消除腳步聲的專業課程。不過，真正的罪犯早就已經知道這些技巧了；而誠實的一般人則得學習這些技巧做為自衛手段。

編按：這種說法忽略了，感冒經藥物治療和自然病程的時間幾乎沒有差別的事實。

第一章

內建偏誤的樣本

樣本（sample），是統計學中的術語，指從母體隨機抽樣而得的部分集合體。這是因為進行調查時，人力、時間、經費有限，又或是地理、時空受限，難以對母體進行調查的情況下，藉由具有代表性的子群體隨機進行採樣，再將結果統合推論至廣泛母體。在本章，將說明統計操控將如何影響「樣本」：

- 會吹噓自己的「不老實」樣本
- 以偏概全！把樣本當整體
- 「已經跳過」的隨機
- 勸誘樣本答覆的「調查者」



如果你有一桶豆子，裡面有些是紅豆，有些是白豆，你只有一種方法，能準確知道每種顏色的豆子有幾顆：那就是全部數一遍。

不過，你也可以用比較簡單的方法算出大約有多少紅豆：只要抓一把豆子，算出其中有多少紅豆，並假設這把豆子的紅豆數量與整桶的比例相符就行了。如果你採計的樣本數夠多，採樣的方式也適當，在絕大多數狀況下，採樣結果的確可以代表整體。但如果樣本數不夠多，採樣方式也不適當，相較於採樣，理性猜測的結果可能還準確得多；因為這種採樣只能提供虛假的科學準確感，除此之外什麼意義都沒有。



更可悲的是，在我們讀到或自認了解的事情背後，往往是藉由取樣偏誤或數量過少的樣本，又或是兩者兼具所得到的結果。

舉一個極端的例子，可以更清楚看出樣本偏誤是如何產生的。假設你進行了一份問卷調查，問卷中包含這個問題：「請問您喜歡回答問卷嗎？」接著你分析了回收的問卷，得出「大多數人都喜歡回答問卷」的結論，而且你為了讓數據更具說服力，還特別計算到最精確的小數位。當然，真實情況可能是大多數回答「否」的人，都早已將問卷丟進了最近的垃圾桶，從而將自己從樣本中剔除。即使在原始樣本中，進了垃圾桶的問卷占了樣本數的**90%**，但當你宣布了你的驚人發

現時，忽略那些被丟掉的問卷會是一種慣性的傳統作法。

在現實生活中，是否也會產生這樣的偏誤？我敢打包票一定會。

報紙和新聞雜誌在不久前指出，過去十年中，大約有四百萬美國天主教徒改宗（編按：從某宗教或教派轉皈依另一宗教或教派的行為）成新教徒。這個數據是由《基督教先鋒報》（*Christian Herald*）的編輯包霖牧師（Daniel A. Poling）所進行的一項調查。《時代》（*Time*）雜誌為這個數據做了說明：

《先鋒報》的這些數據，是透過對美國新教牧師（總數**181,000**人）的調查中所取得的。回答問卷的**2,219**名神職人員（在**25,000**名受訪者中）表示，在過去十年間，他們總共接待了**51,361**名前羅馬天主教徒進入他們的教會。包霖牧師據此進行推算，推測全國有**4,144,366**名天主教徒皈依新教。聖公會教徒威爾·奧斯勒（Will Oursler）寫道：「即使考慮到誤差，全國改宗的總人數也不會少於兩、三百萬，甚至可能接近五百萬。」

儘管《時代》雜誌未能直接指出錯誤，但它仍值得讚賞，因為它讓我們知道有超過**90%**的受訪者沒有回覆這份問卷。為了徹底推翻這項調查結果，我們只需注意到在這些沒有回覆問卷的**90%**受訪者中，多數人可能沒有遇到任何改宗的天主教徒，因此直接丟棄了問卷。

根據這種假設，並使用與包霖調查時相同的全國任職牧師總數，來做出我們的預估。由於他調查了**181,000**人中的**25,000**人，發現有**51,361**人改宗，因此若詢問所有牧師，應該會得到改宗的人數有**370,000**左右的這個數字。

到這邊，我們用粗略的方法得到一個非常可疑的數字，但它至少和包霖所估算的全國改宗數字（該數字是我們粗估數字的**11**倍，所以更加引人注目）一樣值得信賴。

至於奧斯勒先生「考慮誤差」的自信說法.....好吧，如果他因此

發現了一種補償未知誤差的方法，統計學界將對他的偉大貢獻懷抱無限感激。

吹噓自己的「不老實」樣本

有了這個知識背景，讓我們來研究一下一則幾年前的新聞，該報導指出：「**1924**年畢業的耶魯大學校友的平均年薪是**25,111**美元*1。」

哇，這薪水真是不錯！

但是，等一下，這個驚人的數字到底代表什麼意義呢？是不是就像表面上看起來的那樣，證明若你把小孩送進耶魯大學，當你老了之後就不用工作，甚至當他老了之後也一樣？

若帶著懷疑的眼光觀察此數字，你會發現有兩件事很引人注目。第一，這個數字驚人的精確；第二，這個數字美好到不太像是真的。

在任何廣泛的群體當中，平均收入絕對不太可能精確到個位數字。若要確認自己的去年收入達個位數字的程度，除非所有收入都來自薪資，否則不太可能做到。而收入達到**2.5**萬美元的人，往往不會只賺取薪資；這個階層的人很可能擁有各種投資收入。

此外，這個可愛的平均值無疑是根據耶魯畢業生自己宣稱的收入所計算出來的。即使耶魯大學已在**1924**年於紐哈芬（**New Haven**）建立了榮譽制度，我們也不能確定這套系統在四分之一個世紀之後還能運作良好，足以讓這些耶魯畢業生全都誠實回報收入。有些人在被問及收入時，會出於虛榮或樂觀而誇大數字。而有些人則會低報，尤其是當他們已將稅務申報表上的數字減至最小，後續在提及收入的文件或聲明上，就不太可能提出和申報表不符的數字。

誰知道這些回報收入的畢業生是如何看待這件事呢？此外，雖然機率不大，這些吹噓和低報的收入也有機會可能互相抵銷。不過，比

較有可能的，是吹噓或低報的一方大於另一方，但我們不會知道是哪一種。

為什麼我必須嫁給一個
年薪只有 25,111 美元的普通耶魯畢業生？



剛剛我們已經解釋，從常識來說，這個數字不太可能代表真相。接著，讓我們把焦點放在造成誤差的資料來源。這些資料所計算的畢業生「平均收入」是**25,111**美元，然而這些人的實際平均收入很可能只接近該金額的一半。

這份資料中的耶魯畢業生，其實是來自於某一特定樣本。我們之所以能確定這一點，是因為從邏輯上來說，沒有任何人能找到所有在**1924**年畢業、還活著的耶魯畢業生。而且在**25**年後的今日，一定有很多人的地址已經無從得知。

即使真的能從當時的地址上找到這些耶魯畢業生，他們當中的許多人可能也不會回覆問卷，尤其這還是涉及個資的問卷。以郵件問卷來說，有大約**5%**至**10%**的回覆率就已經算很高了。雖然這份問卷的回覆比例應該更高，但絕對不可能達到**100%**。

於是，我們現在知道這個收入數字的基礎，是以該屆「部分」耶魯畢業生做為樣本，這些耶魯畢業生的地址已知，且回覆了問卷。那

麼，這個樣本具有代表性嗎？也就是說，我們是否可以假設這群人的收入，等同於那些聯絡不上或沒有回覆問卷的群體嗎？

在耶魯大學名冊中，誰是那些「地址無從得知」的小小迷途羔羊？他們是像華爾街上班族、公司董事、製造業主管或公用事業主管的高收入者嗎？當然不是！想要找到有錢人的地址從來就不是件難事。

即使這些人沒有和校友辦公室保持聯繫，多數富有的畢業生的聯絡方式，都可以透過《美國名人錄》（*Who's Who in America*）和其他相關資料找到。所以我們可以合理猜測，那些消失的名字都是從耶魯大學畢業25年後，沒有實現閃亮未來的耶魯人。他們可能是辦事員、技工、遊民、失業的酒鬼、勉強維生的作家與藝術家……他們的收入大概要湊滿半打或更多人，才會達到25,111美元。他們鮮少出現在同學會上，原因甚至有可能只是負擔不起交通費用。

那麼，把問卷丟進垃圾桶的又是誰呢？雖然我們無法確定答案，但至少可以合理猜測，在這些人中，有許多都是收入沒有高到可以吹噓的人。他們可能有點像那位在辦公室第一次收到薪資支票時發現上面夾著紙條的同事，便條紙上建議他把薪水金額視為機密，而非用來建立工作自信。「別擔心，」他會這麼告訴老闆：「我跟你一樣以這個金額為恥。」

我們可以清楚發現到，這些樣本遺漏了最有可能拉低平均值的兩群人。不過，25,111美元這個數字逐漸變得可以解釋了。就算這個數字真的代表什麼，也只能代表1924年那一屆畢業生中的特定群體而已：這個群體的地址找得到人，而且他們也願意站出來告訴眾人他們的收入是多少。即使如此，我們也還需要一個「這些人說的是真話」的先行假設。

但假設絕對不能這麼輕率。根據「市場調查」這類的樣本研究經驗，我們幾乎永遠不應該做出這種假設。過去曾有一項旨在研究雜誌

讀者人口組成的逐戶調查，其中的一個關鍵問題是：「你們家會閱讀哪些雜誌？」其分析過後製成的圖表顯示，有非常多人熱愛刊登內容雖不到艱深嚴肅但較有內涵的《哈潑》（*Harper's*）雜誌，與之相對，只有少數人偏好閱讀登載奇聞軼事的雜誌《真實故事》（*True Story*）。不過，當時出版商的數據卻非常清楚的指出，《真實故事》的發行人比《哈潑》多了數百萬份以上。

設計該份調查報告的人心想，或許我們找錯了提問的群體。但事實並非如此，他們調查的地點已經確實涵蓋了全美各種鄰里。當時唯一合理的結論是，有許多回答此調查的受訪者沒有說實話。這份報告調查出來的只有人們的虛榮心。

到了最後，調查人員發現，如果你想知道特定群體的人在讀什麼雜誌的話，直接詢問他們是無法得到真正的答案。獲得更多且完整資訊的方法是去他們家，告訴他們你想買舊雜誌，問問他們有什麼。接著，你只要計算他們有多少《耶魯評論》（*Yale Reviews*）和《愛情羅曼史》（*Love Romances*）就行了。想當然爾，這個不可靠的調查方法也沒辦法告訴你這些人讀了哪些雜誌，你只能知道他們接觸過哪些雜誌。

我讀了你最近的民調，我愛死了！
尤其是那個美好的結局……



以偏概全！把樣本當整體

同樣的道理，下次你從文章中讀到美國的一般人（你肯定時常聽說過有關這位一般人的各種事蹟，不過多數不怎麼可信）每天刷牙**1.02**次時——這是我剛剛編造出來的數字，但這個數據的可信度和其他人提出的沒有多少差別——請捫心自問一個問題：怎麼有人能調查出這種事？如果有人在無數廣告中讀到「不刷牙就是社會罪人」，你覺得這個人會向陌生人坦白他從不定期刷牙嗎？對於只想知道人們在「口頭上」對刷牙有何想法的人來說，這個統計數據或許有意義，但這個數據不會讓你知道人們用牙刷清潔牙齒的頻率。

人們總是說河流不會高於其源頭。好吧，但如果某處藏著一個抽水站的話，河流的確可以高於源頭。同樣的道理：抽樣調查的結果不會優於該調查所立基的樣本。當這些抽樣數據經過層層統計操控的過濾，並刪減到小數點後一位的平均值時，調查結果會散發出令人信服的光環，但只要你仔細檢視抽樣過程，就會發現這個結果其實無法令

人信服。

為了取得具有價值又可信的結果，以抽樣調查為基礎的調查報告，必須使用具有代表性的樣本，也就是說，該樣本需要全數排除偏誤。這就是為什麼我們的耶魯收入金額毫無可信度。這也是為什麼你在報紙與雜誌上讀到的許多數據本身，其實是毫無意義的。

一位精神科醫師的報告曾指出，幾乎每個人都罹患了精神官能症。這種過度泛化的說法，除了會破壞「精神官能症」一詞的特定涵義之外，還應該檢視這位醫師的樣本。換句話說，精神科醫師觀察的都是哪些人呢？事實證明，他得出這個充滿啟發的結論，其實是透過研究自己的病人而來，但是他的病人與代表所有人的樣本群體相距甚遠。如果一個人是正常的，我們的這位精神科醫師根本不會遇到他。



請用這種方式再次檢視你讀到的各種數據，就能避免再度發生表象與事實不符的事情。

另一件值得記住的事，是樣本的可靠性同樣容易被偏誤破壞，不論這些偏誤是否可以看見。也就是說，就算你找不到明顯的偏誤，只

要你發現偏誤有機率存在，你就應該允許自己對結果懷抱一定程度的懷疑。**事實上，在統計數據中，永遠都有偏誤存在。**如果有人對此感到質疑，**1948年與1952年的美國總統大選**可以證明這一點。

若還需要進一步的證據，我們可以回顧**1936年《文學文摘》（*Literary Digest*）**的著名災難^{*2}。當時的**1,000萬份**電訪與《文學文摘》的訂閱者，全都向即將步向末日的雜誌編輯保證，**阿爾夫·蘭登（Alf Landon，共和黨總統候選人）**將會在選舉人團（**electoral college**，美國選舉制度中代表人民進行投票的選舉人）中拿下**370票**，**富蘭克林·羅斯福（Franklin Roosevelt，民主黨總統候選人）**則會拿下**161票**，而且他們的受訪名單已在**1932年**的大選提供過準確預測了。

一份已經經過如此測試的受訪名單，怎麼會出現如此失準的偏誤呢？這份名單當然會有偏誤，正如各篇大學論文與其他事後分析指出：在**1936年**，能負擔得起電話與訂閱雜誌的人，並不能代表全體選民的「橫斷面」（**cross section**，同一時間或時期的數據資料）。從經濟層面來說，他們是一群特殊、充滿了共和黨支持者偏誤的樣本。因此，樣本群體選擇了蘭登，但全體選民卻做了相反的選擇。

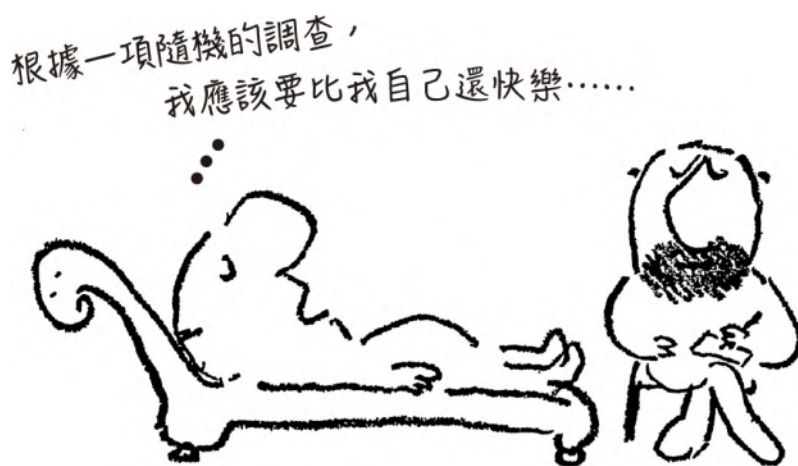
「已經挑選過」的隨機

所謂的基礎樣本，是我們稱作「隨機」（**random**）的一種樣本。隨機的意思，是從「字集」（**universe**，範圍最大的集合）中進行純粹偶然的選擇，在統計學家的用詞中，字集指的是樣本所屬的整個群體。例如，我們從依據索引法排列的檔案夾裡，每隔十個名字就抽取一個做樣本；或者從一個裝滿紙條的帽子裡拿出**50張紙條**；又或是在舊金山的市場街（**Market Street**）上，從每**20個**遇到的人中找一個人進行訪問。

但別忘了，上述最後一個的樣本群體不能代表全世界的人口、美國的人口或舊金山的人口，只能代表當時處於市場街上的人口。曾有一位民意調查的調查人員指出，她在火車站找人採訪是因為「你可以在火車站中找到所有類型的人」。但我們必須向她指出事實並非如此，像是「需要照顧年幼孩子的家庭主婦」，就幾乎不會出現在車站而會在樣本中被忽略。

檢驗抽樣是否隨機的方法是：這整個群體中的每個人或每件事物都有平等的機率成為樣本嗎？

純粹的隨機樣本，是唯一一種可以讓我們充滿信心用統計理論檢驗的樣本。但這種樣本有個問題，那就是這種抽樣方式對於許多調查來說，實在太困難又太昂貴了，光是成本考量就足以讓人放棄這個選項。更加經濟的替代方案叫做「分層隨機抽樣」(**stratified random sampling**)，在民意調查與市場研究等領域，時常會使用這種統計方法。



在取得分層樣本的過程中，你得先根據已知的「樣本盛行率」(**prevalence**，整體或特定群體中某個屬性或特徵的比例)，按照比例將樣本區分成數個小組。但在這裡就會出現問題：你取得的樣本比

例的資訊可能並不正確。

舉例來說，你可能會要求調查人員必須和一定數量的黑人交談、在每個收入階層都和某某百分比的人交談，或與指定數量的農民交談等。與此同時，你則把抽樣群體平等切分成**40歲**以上與**40歲**以下的兩個群體。

這個方式聽起來似乎不錯——但接著會發生什麼事呢？在區分黑人與白人的部分，調查人員通常可以做出正確判斷。但在收入階層的部分，他會犯下比較多錯誤。至於農民的部分——若有一個人既從事農田兼職，同時又在城市裡工作的話，你應該把他分在哪個類別呢？就連以年齡來區分也可能會出現問題，因為調查人員為了符合年齡的分組，而會選擇明顯低於**40歲**或明顯高於**40歲**的人當作調查對象。在這種狀況下，這個樣本群體會因缺乏**30歲**末與**40歲**初的人而產生偏誤。無論怎麼做，都不可能達到完美。

除了這些問題之外，最重要的是，你要如何在這些分層小組中取得隨機樣本呢？最顯而易見的方法，是從一個列出所有人員的名單中「隨機挑選樣本」；但這個方法太貴了。所以你走上街頭——但這麼做的偏誤，就是樣本中沒有待在家裡的人。你或許可以在白天挨家挨戶進行調查——但這麼做就沒辦法調查到在公司工作的上班族。你換成晚上進行挨家挨戶的調查——但這麼做又忽略了那些去電影院和夜店的人。

最後，執行民意調查成了一場持續對抗偏誤的戰鬥，而進行戰鬥的，是一直以來聲譽良好的各種民調機構。閱讀這些報導的讀者必須銘記在心的是，這是一場民調機構永遠無法獲勝的戰鬥。我們在讀到「**67%**的英國人反對」的事情結論時，永遠都應該保持質疑：這**67%**的抽樣群體是哪些英國人？

這個樣本恐怕不如我想像的那麼隨機。



金賽博士 (**Alfred C. Kinsey**) 在其著作《金賽性學報告》的男性篇和女性篇中所提出的數據也是如此。儘管這項研究具有出色的開創性，但卻受到非隨機採樣的詛咒。樣本中有許多特殊的群體，像是受過大學教育者 (女性中有**75%**) 和曾經入監服刑的人。另外一個致命弱點，在於受訪對象的取樣大多屬於願意談論「性相關話題」的人，這些人的生命經歷與那些對性避而不談的受訪者截然不同，因此，在這項研究的樣本中，自然就排除了不願談論性的人們了。

馬斯洛 (**Abraham H. Maslow**) 所做的一項研究，也證實了這一切不僅僅是無妄的猜測。在他的研究對象中，許多人後來自願參加了金賽的研究，馬斯洛發現這些人通常在性方面更加不傳統而且相對成熟。

金賽的研究和最近的性行為研究都一樣，最重要的問題，在於我們解讀數據時 (或者解讀常見的結論時) ，如何避免獲得太多可能不符事實的資訊。基於抽樣的任何研究都存在這種危險，將一本厚重的書籍或龐大的研究報告以簡化的摘要呈現時，這種危險可能會變得更加嚴重。

也許你明年會被選為隨機的樣本……



在金賽的研究中，他的研究數據至少包含了三個階段的抽樣。他的人口樣本群體（第一階段）不是隨機樣本，因此可能對任何人口並不特別具代表性。但另一件同樣重要且應該要記住的是，研究中完成的任何問卷，都只是可能問題的其中一個樣本（第二階段）而已。而受訪者針對每個問題給出的答案，也只是其個人對該問題的態度與經驗所構成的樣本（第三階段）罷了。

動搖樣本答覆的「調查者」

調查人員的組成也有可能會以一種奇妙的方式影響結果。

第二次世界大戰戰爭期間，美國民意研究中心（**National Opinion Research Center**）派出兩組調查人員到美國南部的一座城市，以三個問題訪問了**500**位黑人。其中一組調查人員都是白人，另一組則都是黑人。

他們提出的其中一個問題是：「如果日本人征服了美國，這裡的黑人會受到更好還是更差的對待？」黑人調查人員的報告中，有**9%**的受訪者認為「會更好」；而白人調查人員獲得此類回答的受訪者則是**2%**。同時，在黑人調查人員的受訪者中，有**25%**認為黑人會受到更

差的對待；而白人調查員得到的「會更差」結果更高達**45%**。此外，當他們把問題中的「日本人」替換成「納粹」時，也得到了相似的結果。

這份研究中的第三個問題，則是探討受訪者基於前兩個問題所顯示的感受與態度：「你覺得聚焦於擊敗軸心國比較重要，還是使國內的民主運作得更好比較重要？」根據黑人調查人員所述，有**39%**的人回答「擊敗軸心國」；而根據白人調查人員獲得的結果，這麼回答的人則高達**62%**。

這裡看到的是，未知因素導致的偏誤。對於該調查結果產生最大影響力的因素，可能是我們在解讀民調結果時必定得考慮的一個傾向：**受訪者想要給出讓調查人員滿意的答案**。在戰爭期間，住在美國南方的黑人遇到一名白人提出了一個暗含叛國議題的問題，於是他決定給出好聽的答案，而不是實話，這也沒什麼好奇怪的，不是嗎？另一個可能的影響因素是，這兩組調查人員選擇了不同類型的受訪者。

無論如何，這項調查結果顯然都已經偏頗到毫無價值的程度了。你可以自行判斷，有多少以民意調查當作基礎的其他調查結果也同樣具有偏誤、毫無價值——只不過目前我們還沒有找出檢驗這種偏誤的方法。

如果你懷疑某個民調在整體上來說具有某個特定方向的偏誤，像是《文學文摘》那樣的話，那麼你就應該繼續追查下去。調查人員在選擇代表群體的人時，往往會因為偏誤而偏向那些比人口平均值更富裕、教育程度更高、能獲得更多資訊、更警覺、外表更好看、行為更傳統、習慣更固定的那群人。

只要稍微想一想，你就能知道造成這種結果的原因。讓我們假設你是一名調查人員，現在要在某個街角進行調查。你注意到有兩名男人看起來都符合你必須調查的類別：四十多歲、黑人、城市居民。其中一人的衣著乾淨整潔；另一人則外表骯髒，且看起來脾氣暴躁。為

了完成工作，你肯定會走向看起來比較有意願協助完成調查的那個人。而你在全國各地進行調查的同事們，也會做出類似的決定。

在自由派或左派圈子中（編按：民主黨偏向這派），普遍對民調抱持著強烈的反感情緒，他們認為這些調查通常是受到操縱的。這種觀點背後的實際狀況是，對於思想較不保守的人來說，民調結果往往不符合他們的意見和希望。他們認為，就算之後的選舉結果不是共和黨（編按：共和黨傾向保守主義）勝利，但在選舉前的民調結果總是顯示共和黨會當選。

事實上，正如我們在前面所看到的，調查人員沒有必要操縱民調——也就是說，他們沒有必要為了製造錯誤的印象而刻意扭曲結果。因為，光是樣本方面的持續偏誤就會自動操縱民調結果了。

編按：約合今日 45 萬美元，資料來源：
https://www.bls.gov/data/inflation_calculator.htm，後續相關換算資料來源亦同。
編按：該雜誌因民調與選舉結果誤差太大，於選後幾個月終止發行。

第二章

精心挑選的平均值

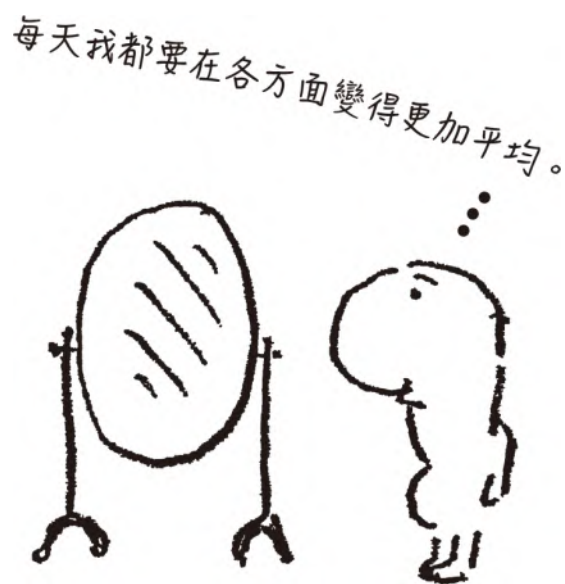
平均值或平均 (average)，可以同時代表統計中的算數平均數 (mean)、中位數 (median) 與眾數 (mode)。算術平均數是將一組數字相加，然後再除以這組數字的個數。中位數則是指一組數字的中間數字，因此，有一半數字的值會大於中位數，而另一半數字的值會小於中位數。眾數則是一組數字中出現次數最多的數字。綜合來說，算術平均數易受資料中的極端值的影響，而中位數、眾數較不受極端值的影響。在本章，將說明統計是如何操控平均這個概念：

- 選「最有利平均」巧立名目
- 故意為之！計算加入極端值



親愛的讀者，我相信你絕不是個勢利鬼，而我當然也不在房地產業工作。但姑且讓我們假設你是個勢利鬼，我也在房地產業工作；而你正在我家住的那一條路上找房子。幫你評估後，我煞費苦心的告訴你，此一區域的年平均收入是**1萬英鎊**^{*1}；或許這一點提高了你住在這裡的意願。總而言之，你買下了這裡的房子，心中牢牢記得我說的那串絕佳數字。我們已經假設你是個有點勢利的人了，所以你很有可能會在告訴朋友自己住哪的時候，毫不猶豫的把這裡的平均年收入告訴他們。

大約一年後，我們再次見面了。我做為某個納稅人委員會的成員，正在分發一張請願單，要求政府維持低稅率、低稅額以及低公車票價。我的請願內容指出，我們負擔不起稅金與公車票價的增加：畢竟這個居住地區的年平均收入只有**2,000英鎊**而已。或許你會支持我和這個委員會推動的請願——你不只是個勢利鬼，而且還很吝嗇——不過，你仍舊在聽到微不足道的**2,000英鎊**時大吃一驚。到底我是在今年請願時說謊？還是去年賣房時說謊？



選「最有利平均」巧立名目

其實在這兩個時間點，你都不能說我在說謊，這就是用統計數據操控的美妙之處。這兩個數據都是我以正當方式得出的真實平均值。它們代表的是相同的數據資料、相同的人和相同的收入。但是很明顯的，其中至少有一個數據具有很高的誤導性，簡直就是公然說謊了。

我的小技巧，是在這兩個時間點使用不同的平均值。要知道，「平均值」這個詞的意義其實非常廣泛。這是一種常用的技巧，有時使用者是無心的，不過通常都是刻意的，這些使用者希望能藉此影響輿論或售出廣告版面。當你聽到某個數字是平均值時，其實你並不知道這是什麼數字，除非你能釐清它是哪種常見的平均值——算數平均數、中位數或眾數。

當我想要獲得較大的平均值時，使用的是**1萬英鎊**，也就是這個住宅區所有家庭收入的「**算數平均數**」。計算這個算數平均數的方法，是把所有收入加起來再除以家庭數量。而當我想要獲得較小的平均值時，我會計算「**中位數**」，它代表的是在這個區域有一半家庭的年收入超過**2,000英鎊**，而另一半家庭則少於這個數字。我也可以使用「**眾數**」，意思是在樣本中最頻繁出現的數字。如果在這個區域裡，年收入**3,000英鎊**的家庭是最多的，那麼這裡的年收入眾數就是**3,000英鎊**。

在這個例子裡，儘管所有的收入金額都是真的，但這種沒有具體說明意涵的「平均值」，其實毫無意義。增加混淆的另一個因素，是在某些特定的情況下，這三種平均值的數字都會十分相近，使得在一般使用上來說，其實不會特別區分這三個數字。

如果你讀到的資訊，是某個群體的男性平均身高只有**5英尺**（約**152公分**），你就能對這群人的身材有相當程度的了解。你不需要詢問這個平均值是算數平均數、中位數還是眾數；因為這三個數值相差

不遠。

因為，當你處理的數據是與人類特徵相關的數據時，數據通常會呈現「常態分布」，此時不同的平均值就會十分相近。如果你畫一條線來表示常態分布的話，將會得到一條鐘形曲線，而該數據的算數平均數、中位數和眾數都會落在同一點。因此，任一平均值都和其他兩種平均值一樣能夠描述一群人的身高。

但在描述一群人的錢包深度時，這些平均值的意義就不一樣了。若你將某個城市中所有家庭的年收入都列出來的話，可能會發現數字落在「沒多少錢」與「**2萬英鎊**」之間，而且你可能會注意到有些數字特別大。大約有**95%**的收入會低於**5,000英鎊**，畫成曲線時，這些人非常靠近曲線的左方。因此，這條曲線看起來並不對稱而像是偏斜的鐘。它的形狀有點像是兒童溜滑梯，左側為樓梯急遽上升到高峰，接著到右側滑梯處慢慢傾斜向下。此數據的算數平均數和中位數將會相距甚遠。你可以清楚注意到，在這種數據中，其中一種「年平均值」（算數平均值）和另一種「年平均值」（中位數）之間的比較會有「效度」（**validity**）上的差別。



在我賣你房產的社區中，這兩個平均值的距離格外遙遠，這是因為這裡的收入數據偏斜特別嚴重。你的多數鄰居都是小規模農夫、在附近村莊接受僱用的工薪階級和領退休金的老年退休者。但是其中有三個居民是週末會來這裡住的百萬富翁，他們大幅提高了收入總額，也因此大幅提高了算數平均數。此區域幾乎所有人的收入都低於這三個百萬富翁拉高的算數平均數。所以，現實中就有這種像是笑話或修

辭法的實例：**幾乎所有人都低於平均值**。

這就是為什麼當你讀到某公司主管或業主宣布，他們公司的平均薪水有多高時，這個數字可能有意義，也可能一點意義都沒有。如果那個數字是中位數，那麼你確實可以從中獲得一些重要資訊：因為至少有一半員工薪水高於此數字，而另一半則低於此數字。但如果該平均值是算數平均數的話（相信我，若公司沒有具體說明其類型，它可能就是算數平均數），你得到的平均值，可能就會包括了一個**2.5萬英鎊**的收入（業主的收入）與大量低薪勞工的收入。「平均年薪**3,800英鎊**」很可能粉飾了**1,400英鎊**的勞工年薪與業主以鉅額工資的形式取得的利潤收入。

我們可以在某些公司的聲明中看到，業主能用多麼簡單的方式把平均值當作讓他們得利的工具，使實際狀況愈糟糕，數字看起來就愈美好。下面讓我們來小試身手吧。

故意為之！計算加入極端值

你是一家小型製造商的三名合夥人之一。時間來到了年終，這是入帳豐碩的一年。你支付了**9.9萬英鎊**的薪水給**90名**員工，他們的工作是製作與運送椅子或公司生產的其他產品。此外，你和你的合夥人每人則獲得了**5,500英鎊**的薪水。接著你發現，你們三名合夥人還可以平分**2.1萬英鎊**的年利潤。你要如何描述這些金額呢？為了便於理解，你可以用平均值來表示這些數字。由於所有員工的工作都差不多，薪水也差不多，所以無論你使用的是算數平均數還是中位數，都不會有太大的區別。你計算出來的結果如下：

員工平均薪資：**1,100英鎊**（ $9.9萬 \div 90 = 1100$ ）；

業主平均薪資與利潤：**12,500萬英鎊**

（ $〔5,500 \times 3 + 2.1萬〕 \div 3$ ）。

這兩個數字看起來遭透了，對吧？那麼，我們可以試試另一種方法。將2.1萬英鎊利潤中的1.5萬英鎊，取出做為薪資獎金分配給三名合夥人。這一次，你在計算平均薪資時，請把你自己和合夥人的薪資與獎金也算進去，而且一定要使用算數平均數：

平均薪資：1,403英鎊

$$([9.9萬 + 1.5萬 + 5,500 \times 3] \div 93) ;$$

業主平均利潤：2,000英鎊

$$([2.1萬 - 1.5萬] \div 3)。$$

哇！這次的數字看起來好多了。雖然你還可以讓數字變得更好看，不過這樣已經夠好了。如果你想要的話，甚至可以進一步證明，在用於薪資與業主利潤的資金比例中，利潤的比率甚至不到6%。總而言之，如今你已經算出可以公開發表、張貼在布告欄或用來談判的數字了。

由於我們簡化了這個例子，所以這種計算方式非常粗糙，但和各個企業以會計名義做的事相比起來，可說是小巫見大巫。有鑑於在複雜的企業中，員工階級包括了從初階打字員到獎金高達數十萬美元的總裁等職位，所以企業可以用這種方式來粉飾各種事。

因此，你在看到平均薪資的數字時，**首先要問的是：這是什麼的平均？有哪些人的薪資包括在內？**美國鋼鐵公司（United States Steel Corporation）曾指出，他們的員工平均週薪在不到十年內成長了107%。這個數字確實沒錯——但如果你仔細留意的話，會發現他們在計算早期薪水的數據時，包含了較多比例的兼職員工，因此這種顯著薪水增加的效果就會打折。如果你在前一年的工作為兼職，後一年的工作是全職，那麼收入當然會翻倍，但這個數字與真實的薪資成長率毫無關聯。

你可能曾在報紙上讀到，某年美國家庭的平均收入是**6,940**美元。除非你知道此調查中的「家庭」所代表的意義，也知道該平均值是哪一種，否則你不該認為這個數字能帶給你有用的資訊（而且你也應該要知道提出這些資訊的人是誰、他是怎麼知道的，以及這個數字有多準確）。

事實上，這些資訊可能來自美國人口普查局（**Bureau of the Census**）。如果你能取得人口普查局的完整報告，你將能輕而易舉的在其中找到你需要的其他資訊：這個平均值是中位數；「家庭」指的是「兩個或兩個以上的人，彼此有關聯且住在一起」。此外，如果你重新閱讀報告中的表格，你會發現此平均收入是基於一個相當大的樣本，也因為有如此大量且足夠的樣本數量，使得這個估算，在**20**次抽樣之中，會有**19**次抽樣落在此平均收入**±71**美元的差額之內。

有了這樣的機率和差額，這個估算值其實可以算是十分不錯。人口普查局的人員同時有足夠的能力與財力，可以在執行抽樣研究時把精確度拉到相當高的程度。我們可以合理推測，他們沒有特別想要證明任何觀點。但你在生活中讀到的大部分數字，通常並不是在這麼幸福的環境下所產出的，而且也不是所有數字都會附上這些資訊，讓你知道它們有多精確或多不精確。我們將在下一章深入討論這個問題。

同時，你可能會想要對《時代》雜誌在〈出版商的一封信〉（**A Letter from the Publisher**）中提到的某些事項提出質疑。這篇文章指出，「新用戶」的「年齡中位數是**34**歲，平均家庭年收入是**7,270**美元」。他們在更早之前對於「舊訂閱戶」的調查中指出，這些訂閱戶的「年齡中位數是**41**歲.....平均收入是**9,535**美元.....」這裡我們很自然會想提問，為什麼在這兩篇文章中，他們都清楚指出年齡的數值是中位數，但卻沒有具體說明平均收入是哪一種平均值。這會不會是因為使用算數平均數會使收入的金額較高，能夠讓廣告商覺得讀者比較富裕呢？

你也可以試著把這種「這是哪種平均值」的遊戲，套用在第一章
1924年耶魯畢業生有多成功的收入數字上。

編按：原書出版的**1954**年，英鎊兌美元的匯率約為**1：2.5**，因此，當時的**1**萬英鎊約合今日**28**萬美元，後文相關英鎊數據可依此進行換算。

第三章

刻意忽略的小數字

多數的統計結果都是由樣本推論母體，因此當然可能會有誤差，其中最常見的誤差就是把對的當成錯的（稱為第一型誤差），而「顯著水準」就是檢測統計中會出現第一型誤差的機率，正常來說合理的顯著水準約為1%~5%，然而這個小數字卻常常會被忽略掉，進而造成統計上的操控性，而其他小數字，還包含了在計算平均中的偏差值、範圍等。在本章，將討論這些「小數字」是如何被操控的：

- 樣本過少，統計不再顯著
- 太簡化！不現實的「平均」
- 看不見的數字，最重要



一位統計學家曾建議，當你得知一項調查的結果時，你應該問：「在取得這個統計結果之前，你調查了多少樣本？」正如先前提到的，具有偏誤的樣本可以被有心人用來產生想要的統計結果。而如果隨機樣本的數量極小時，也可以實現這一點。

樣本過少，統計不再顯著

廣告的大字標題寫著：「根據用戶回報，使用多克斯牌（Doakes）牙膏的用戶減少了23%的蛀牙。」你可能覺得減少23%的疼痛是個不錯的選擇，所以你繼續讀下去。你在廣告內容中發現，此數據來自一個值得信任的「獨立」實驗室，而且這個估值經過了執業會計師的認證。你還能再多要求什麼呢？

自從使用多克斯牙膏，
我們的蛀牙就減少了23%。



然而，如果你不是個特別容易受騙或特別樂觀的人，你就會回憶起過去的經驗，發現各品牌的牙膏效果其實都差不多。那麼，多克斯牙膏公司是如何得出此結果的呢？難道他們能夠用這樣斗大的字眼說謊，又順利避開法律制裁嗎？事實上，他們並沒有說謊，也沒有必要說謊。他們可以用更簡單也更有效率的方式得到這個結果。

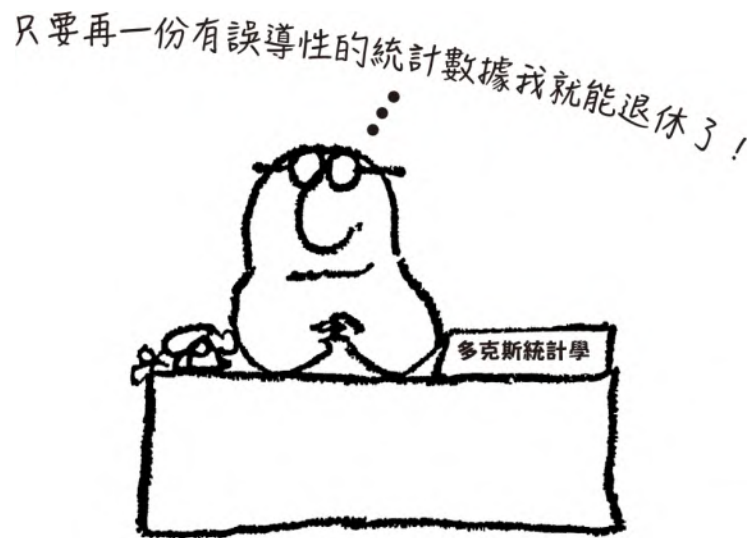
這個廣告中的主要問題，在於樣本數不足——也就是統計學上的不足；不過以多克斯的目標來說，這個樣本數量則是恰到好處。若你注意到廣告上較小的字體，會發現接受該測試的牙膏使用者只有**12**個人。

不得不承認，你其實應該謝謝多克斯給你這個練習機會。有些廣告商會直接省略這些資訊，就連精通統計學的人也只能猜測他們面對的是哪種詭計。而且在牙齒清潔類的商品中，其實**12**人左右的樣本數不算太糟。前幾年有一款名叫康沃爾博士牙粉（**Dr. Cornish's Tooth Powder**）的商品，聲稱他們在「齙齒矯正方面……獲得了顯著的成功」。他們的主要構想是牙粉中含有尿素，而實驗室研究已經證明了尿素對於矯正齙齒是有效的。不過此結論之所以毫無意義，是因為他們只做了初步的實驗研究，而且實驗案例只有六個。



讓我們回過頭來，檢視多克斯如何能這麼輕易的取得毫無做假又值得成為廣告標題的結果，而且還讓所有數字都得到認證。他們的作法是讓一小群人在六個月內不斷計算自己的蛀牙數量，並改使用多克斯牙膏。這項實驗只會有三種結果：蛀牙明顯變多、蛀牙明顯變少或數量相當。只要第一種結果和第三種結果出現，多克斯公司就會直接對數據進行調整（在某些我們不容易注意到的地方），然後再實驗一

次。透過這種「操作偶然」的作法，多克斯遲早會遇到某一個測試組的結果是大幅進步，足以成為醒目標題又值得廣告投放。無論這些受試者使用的是多克斯牙膏、蘇打粉，還是以前用的舊牙膏，結果都會是一樣的。



使用小群體當作樣本的重要性在於：在大群體中，因為偶然出現的任何差異可能都只會是微小的差異，不足以引起大篇幅的關注。宣稱牙膏能「改善**2%**」的蛀牙是沒辦法賣出太多牙膏的。

在純粹偶然的情況下，如果樣本數夠小，就可以產生不帶有任何意義的結果。你可以用很低的成本去測試，只要丟硬幣就行了。硬幣出現正面的比例是多少？當然是**50%**囉。這是每個人都知道的事。

使用多克斯牙膏的醫生中，
其中 85% 都是以免費樣品的方式取得。



好吧，讓我們實驗看看結果如何……我剛剛丟了十次硬幣，其中有八次正面朝上，這證明了丟出硬幣後正面朝上的比例是**80%**。嗯，至少多克斯牙膏的統計數據就是用這種方法證明出來的。現在輪到你測試看看了。你當然有可能會得到**50%**正面朝上的結果，但較高的機率是不會；你的結果很可能和我的一樣，和**50%**相差甚遠。但是，如果你有耐心能持續丟一千次硬幣測試的話，你幾乎（但不是絕對）可以確定你得到的結果會非常接近**50%**的正面朝上——也就是能夠代表真正機率的結果。進行試驗時，唯有在樣本數充足的狀況下，平均數定律^{*1}才會成為有效的描述或預測。

但是，樣本數要多少才算充足呢？這也是一個很棘手的問題。決定樣本數是否充足的其中一個因素，在於你藉由抽樣研究的人口規模多大、人口多樣性多高。此外，有時候樣本中的數字並沒有表面上看起來的那麼簡單。

其中一個值得留意的例子，和前幾年對小兒麻痺症（**polio**，又稱脊髓灰質炎）進行的疫苗研究有關。就醫學實驗來說，這項研究是個

令人印象深刻的大規模實驗：在實驗社群中，有**450**名兒童接種了疫苗，當作對照組的另外**680**名兒童則沒有接種疫苗。不久後，該社群遇上了一場流行病的爆發。根據醫療人員的檢驗，接種過疫苗的兒童中，沒有任何一個感染小兒麻痺症。

不過，控制組的兒童也通通沒有感染小兒麻痺症。實驗人員在規劃此研究時，忽略或不理解小兒麻痺症的發病率很低。按照一般比例，在該實驗的社群規模中，預計只會有兩個病例，因此這個實驗打從一開始就註定毫無意義。若想要取得具有任何意義的結果，參與實驗的兒童數量得增加到**15至25**倍。

許多偉大的（但也轉瞬即逝的）醫學發現，其實也是從類似的方式開始的。正如一位醫師曾說過的：「快點，你得在為時已晚之前先使用這個新藥物。」

然而，不是只有醫學界會出現這種過失。公眾壓力和草率的新聞業常會帶來未經證實的治療方法，在醫療需求極高且統計背景模糊的狀況下尤其容易如此。數年前十分受歡迎的流感疫苗，和晚近出現的抗組織胺藥物就是例子。這些不成功的「治療」之所以會受歡迎，很大的原因在於疾病本身的不確定性，以及邏輯上的缺陷。若你願意給你的身體一點時間，它甚至能自己治癒感冒。

你要如何避免被沒有真正結論的結果所愚弄呢？難道每個人都必須成為統計學家，自己研究原始數據嗎？其實事情沒有那麼困難，你只要運用一種容易理解的顯著水準（**significance**）測試就可以了。這種測試是一種簡單的方法，能讓你了解自己看到的數字，代表的是真實的結果，而非偶然製造出來的表象。顯著水準是一個不存在的小數字——這是因為提出該數據的人認為，身為非專業讀者的你不會理解何謂顯著水準；又或者提出該數據的人想要證明某種特定觀點，而他們認為你理解何謂顯著水準，所以不願寫出這個數字。

若你的資訊來源提供了顯著水準的數字，你將會更加理解你眼前看到的是什麼情況。在描述顯著水準時，最簡單的表達方式是機率。例如，我們先前看到人口普查局的報告就指出，他們提出的數字有 $\frac{19}{20}$ （95%）的機率具有一定程度的準確性。在多數研究中，低於5%的顯著水準就已經足夠好了。有些研究要求的顯著水準則是1%，這個數字代表的是該研究的樣本有99%的機率具有真實的明顯差異或結果。而可能性這麼高的事物有時甚至會被稱作「幾乎完全確定」（practically certain）。

太簡化！不現實的「平均」

此外，還有另一種不存在的小數字，它的缺席也同樣會使研究結果出現大問題。這個數字告訴我們研究中的平均值來自多大的範圍，或者有哪些偏差值。一般來說，無論一個研究是否明確指出平均值種類是算數平均數還是中位數，平均值都是一種過於簡化的數字，而過於簡化比毫無價值更加糟糕。「完全不了解」遠好過於「錯誤認知」，一知半解是一件很危險的事。

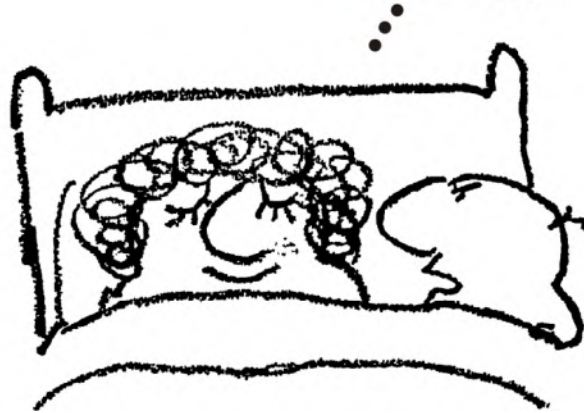
舉例來說，美國近來有太多住宅在規劃時，都按照統計學上的家庭平均人數3.6人來進行設計。把這個數字放進真實世界中，代表的就是一個家庭由三或四人組成，也就是住宅需要兩間臥室。儘管三至四人的家庭是「平均值」，但在實際上卻只占所有家庭的少數。建商指出這是「我們為平均的家庭建造平均的房屋」，但他們直接忽略了大多數家庭的人數其實是比较平均來得多或少的。這使得建商在美國某些區域建造了過多的兩臥住宅，而相對較大與較小的住宅則明顯不足。

這就是一個具有誤導性的不完整統計數據所造成的高昂代價。一個大型的公共衛生團體曾對此做出回應：「當我們不只以算術平均數

來查看這個數字的真正範圍時，我們就會發現，所有家庭中只有**45%**是三至四人的家庭。**35%**是一到兩人的家庭；**20%**是多於四人的家庭。」

在**3.6**這個具有權威性又精確到令人信服的數字面前，常識一敗塗地。這個數字莫名的勝過了每個人都能從觀察中得知的事實：多數家庭中的人數較少，而鮮少家庭的人數很多。

為什麼你總是占用 75.6% 的空間？



另一個以同樣方式忽略了小數字的，是「葛塞爾標準」^{*2}，這些消失的小數字使許多父母陷入痛苦之中。許多父母會在週日報紙的印刷頁面或其他地方讀到：「一名」小孩會在「某某月齡時」學會直立的坐姿，並立刻聯想到自己的孩子。如果他們的孩子在同樣的年齡還沒辦法坐穩，他們必定會得出結論：自己的小孩發展「遲緩」或「不正常」，甚至是其他同樣令人擔憂的狀態。有鑑於一半的小孩必定會在過了該年齡之後才能坐穩，所以許多父母都因此感到沮喪。當然了，有另外**50%**的父母會因為發現孩子發展「超前」而感到的喜悅。從數學上來說，這些喜悅與沮喪的父母數量正好能互相平衡。但是，當這些沮喪的父母努力迫使他們的孩子符合標準且不再落後時，可能反而會造成傷害。

葛塞爾與他提出的方法並沒有問題。有問題的，是這個數字從研究人員手上，轉移到譁眾取寵或資訊不足的作家筆下，再轉移到讀者眼中的過程，沒能領會到中間消失的數字。如果文章能在「標準」或平均值旁加上年齡範圍，就能避免許多誤會了。如此一來，這些父母就會發現孩子的發展落在正常範圍內，不再擔心毫無意義的微小差異。幾乎沒有任何人在任何一方面是剛好正常的，就像在我們扔了一百次硬幣後，不太可能會剛好出現正面和反面各50次一樣。

若我們混淆了「標準」和「期望」，還會使狀況變得更糟。葛塞爾陳述的只是他觀察到的事實；是父母自己在閱讀書籍和文章時得出了結論，認為慢了一天或一個月學會走路的孩子就是有問題的。

許多文章之所以會針對金賽十分著名（但鮮少受到正確解讀）的研究報告提出愚蠢批判，是因為這些文章把「標準」視為好的、正確的、可取的觀點。有些文章指控，金賽給年輕人灌輸思想使其「墮落」，尤其是他把各種常見但未經許可的性行為稱作「標準常態」的論述。其實金賽只是指出，自己注意到這些行為很常見（這才是標準常態所代表的意思），但他並沒有對行為本身表現認可的態度。無論這種行為是否下流，都不是金賽博士該在他的職業領域內決定的。金賽博士遇到的是許多觀察者也同樣會遇到的困擾：在提起敏感的主題時，若你沒有立刻說出自己採取支持或反對的立場，將會是很危險的一件事。

看不見的數字，最重要

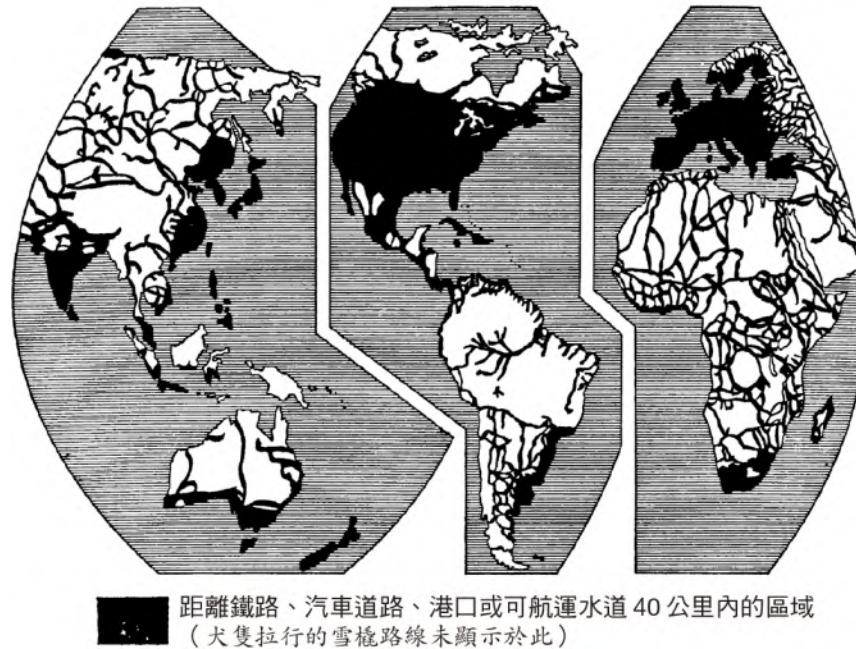
這些不存在的小數字能欺騙我們，是因為它們實在太常缺席，以致我們時常會忽略到它們並不存在。這當然正是不存在的小數字能夠成功的祕訣。部分針對新聞業的批評指出，如今的業界缺乏外出跑新聞這種老派的工作模式，十分令人遺憾。這些批評也嚴厲指責「華盛

頓那些只會坐辦公室的記者」，僅會以缺乏價值判斷的文章重新表述政府發布的新聞稿來維生。有關新聞業缺乏進取心的範例，請見《雙週新聞雜誌》（*Fortnight*）提出的「新工業發展」列表中的資訊：「西屋電氣（Westinghouse）的『新型低溫浸泡法』能使鋼鐵的硬度增加三倍。」

這段描述聽起來像是工業上的一大進展……但當你開始試著理解這句話後，會發現它變得像是一團水銀一樣，令你難以捉摸。這種浸泡法會使所有鋼鐵的硬度都變成浸泡前的三倍嗎？還是這種浸泡法製造出來的鋼鐵，相較於舊方法製造出來的鋼鐵具有三倍硬度？這種浸泡法到底是什麼？很顯然的，這篇報導的記者轉述了一些詞語，但卻沒有探究其意義，該記者覺得這些詞語能帶來學到知識的愉快錯覺，所以預期讀者會以缺乏價值判斷的方式閱讀。這種態度太過熟悉，不經令人想起過去對於課堂教學方法的定義：教學指的是講師把教科書的內容轉移到學生的筆記本上，而不是經過雙方的腦袋轉移知識。

不久前，我於《時代》雜誌上查找有關金賽博士的資料時，發現了另一個在進一步檢視之下潰不成軍的陳述。這段陳述出現在多個電力公司於**1948**年共同刊登的廣告中。「如今，有超過四分之三的美國農場都有電力可用……」這聽起來相當不錯。這些電力公司確實善盡了職責。當然，如果你想故意刁難的話，你也可以把這句話解釋成「如今還有接近四分之一的美國農場沒有電力可用」。然而，這句話真正的詭計在於「可用」（**available**）這個詞，這些公司可以靠著「可用」來提出任何想提出的陳述。很顯然的，「可用」並不代表四分之三的農夫全都真的在使用電力，否則該廣告一定會直接做出這樣的描述。這些農夫只是有電力「可用」而已——根據我的理解，這句話既可以代表電線經過了這些農場，也可以代表電線距離這些農場**16**公里或**160**公里遠。

可購買本書之世界分布圖



且讓我在此引用一篇發表在流行雜誌上的文章標題：〈現在你可以知道你的孩子會長多高了〉。該文章在顯眼的位置放了兩張圖表，一張是男孩的，一張是女孩的，這兩張圖表顯示了孩子在成長過程中每一年的身高，是最終身高的多少百分比。「若想知道你的孩子在長大成人後的身高，」其中一張圖片說明寫道：「請用孩子目前的身高對照此圖表。」

有趣的是，這篇文章本身就會告訴你，該表格的致命缺陷是什麼——前提是你確實閱讀了這篇文章。並不是所有孩子的成長方式都一模一樣。有些小孩剛開始成長較慢，之後才會加速；還有些小孩會在短時間內快速成長，接著趨緩平穩；以及有些小孩的成長速度則相對穩定。你可能已經猜到了，這兩張圖表的資料來源是測量了大量樣本之後取得的平均值。儘管研究人員隨機抽樣了**100**名孩童的總身高，因此該數據無疑足夠準確。但是對於只想知道自己小孩會長多高的父母來說，這兩張圖表幾乎毫無意義。若你想知道自己的孩子會長多高，

可以透過觀察孩子的父母或祖父母的身高來做出更好的猜測。儘管這樣的猜測不像圖表那麼精確而符合科學，但至少準確程度和圖表相當。

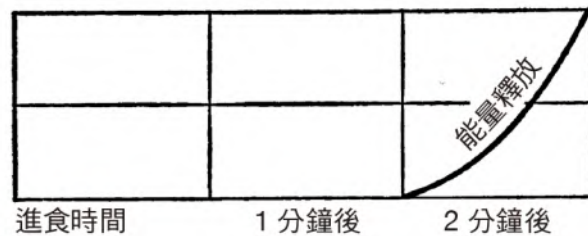
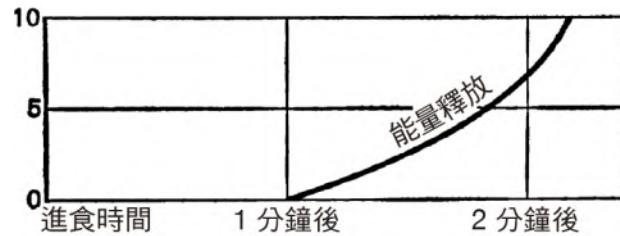
我很高興的注意到，根據我**14**歲在高中軍訓時所記錄的身高，最終我應該只有**173**公分。但我現在的身高是**180**公分。高達**7**公分的人類身高誤差，會被視為非常糟糕的猜測。

我希望我有能量去拿葡萄堅果麥片
並獲得麥片的「能量釋放」。

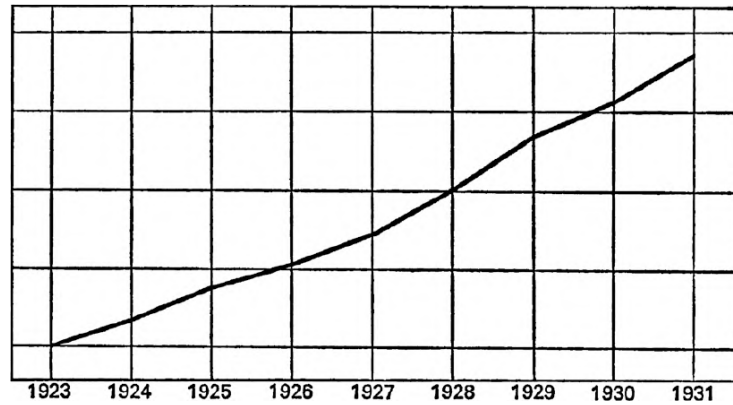


在我面前，有兩盒葡萄堅果麥片的包裝盒。正如盒子上的推薦語所描述的，這兩盒麥片略有不同：其中一盒引用了米老鼠卡通中雙槍彼特（**Two-Gun Pete**）說的話，另一盒則寫道：「如果你想變得像西部牛仔霍比（**Hoppy**）一樣.....那你就要和霍比吃一樣的食物！」不過，兩個盒子都提供了圖表指出（「科學家證明了這是真的！」），這兩種麥片可以「在兩分鐘內就為你補充能量！」。其中一個盒子上，藏身於驚嘆號森林中的圖表標示了縱軸的數字；另一個盒子則直接省略了縱軸的數字。有鑑於兩張圖表都沒有標示數字代表的意義，所以兩者其實沒什麼差別。而兩張圖表各有一條陡峭攀升的曲線（顯示「能量釋放」），只不過其中一條線的起始點是吃了麥片的一分鐘後，另一條的起點是兩分鐘後。此外，其中一條線的攀升速度是另一

條線的兩倍，由此可知，就連繪製圖表的人也不覺得這兩張圖表具有任何意義。

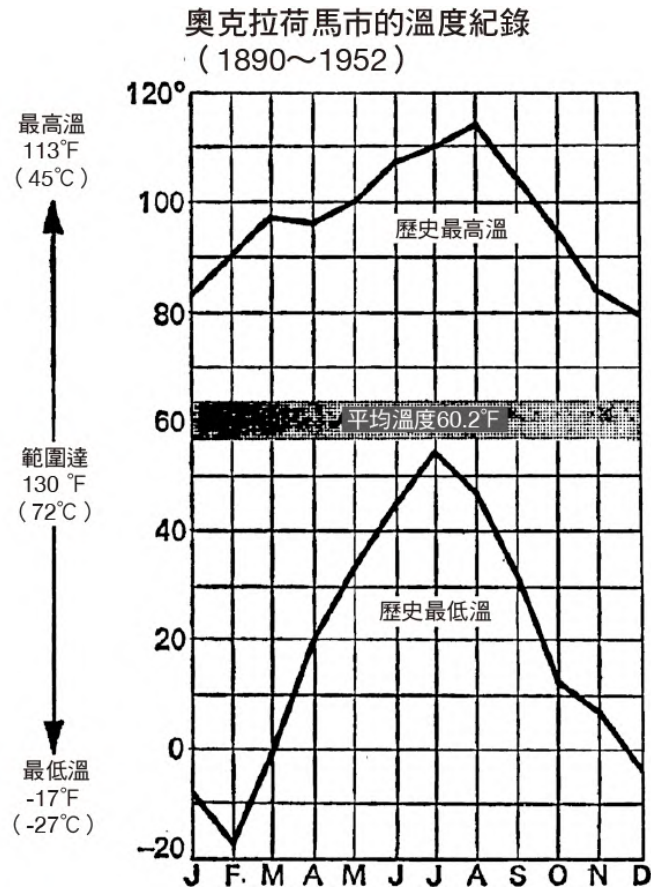


想當然爾，會出現這種愚蠢圖表的商品只會賣給青少年，或一大清早就筋疲力竭的青少年父母。不可能會有人用這種統計數字來侮辱成年生意人的智商.....但，真的不可能嗎？讓我告訴你我在《財富》（*Fortune*）雜誌一個十分特殊的專欄中看到的一張圖表，那張圖表是用來廣告一間廣告公司的（希望這樣的描述不會令你混淆）。圖表中的線條顯示了該公司的業務每年都維持驚人的成長趨勢。不過圖表上沒有數字。這張圖表既可以表示這間公司的業務在一年之內出現了驚人的成長，變成兩倍或者增加了數百萬美元；也可以表示公司在統計方面的成長速度像蝸牛一樣，一年之內進帳只增加了一到兩美元。但不管這張圖表代表的是何者，都維持著同樣的誠實度，使圖表看起來十分不凡。



只要缺少這些重要數字，不管是平均值、圖表或趨勢，我們就不該太過相信它們。若太過相信這些資訊，你的盲目程度也就等同於只靠著氣溫的算數平均數來決定露營地點的人。

你或許會覺得攝氏**16**度的年均溫算是舒適，但用這個數字選到的露營地點有可能是加州的內陸沙漠，或南岸外的聖尼古拉斯島（**San Nicolas Island**）等地區。如果你忽略了溫度的變化範圍的話，可是有可能會被冷死或曬乾。畢竟聖尼古拉斯島的溫度是攝氏**8**至**30**度，沙漠的溫度則是攝氏**-9**至**40**度。



奧克拉荷馬市 (Oklahoma City) 在過去六十年來的平均溫度，也與攝氏16度十分相近：攝氏15.6度 (華氏60.2度)。但是你可以從圖表看出，這個既涼爽又舒適的溫度背後其實隱藏了攝氏72度的變化範圍。

law of average，編按：又稱為大數法則 (law of large number)，在重複試驗中，隨著試驗次數的增加 (樣本增加)，取得的算術平均數就會愈趨近期望值。

Gesell's norms，編按：美國心理學家阿諾德·葛塞爾 [Arnold Gesell] 測量四週到六歲幼兒身體發展，並設計為量表。

第四章

被放大的「差不多」

很多時候，進行統計的目的，是為了要找出兩者的差異。尤其當我們比較兩個數字時，即使只是一點點的不同，就會依此大作文章。但是差異也有大小之分，若此時的差距僅是「統計誤差」(statistical error，指抽樣或檢測等造成測量值與真實值之間的差異)，則根本不是真正的不同或差異。在本章，將說明統計會如何操控這些無論是否有意思的「差異」：

- 說差與差別，不等同
- 毫無意義的「差很大」



喬賽亞·斯坦普 (Josiah Stamp) 爵士描述了藍道夫勳爵 (Lord Randolph) 審查收入報告時的經歷。藍道夫勳爵對海關收入比去年同期成長了**34%**感到滿意，而在勳爵背後觀看的祕書糾正了他，指出這其實只有**0.34%**的成長。

「這有什麼區別？」勳爵問。祕書為他解釋，說明其中一個數字事實上是另一個數字的一百倍，勳爵回應道：「我以前常看到那些該死的小點，但到現在我才知道它們意味著什麼。」

誤差與差別，不等同

在比較數字時，造成困擾的不是點，而是那些該死的小差異。為了看看這是怎麼回事，我們將開始——若你不介意的話，首先我要介紹兩位孩子給你認識。彼得和琳達 (我們姑且用這兩個新潮的名字來稱呼他們吧) 前陣子完成了智力測驗，這也是大部分孩子會在學校課程中所做的事。在現今這個時代，任何類型的心理測驗都會被歸類成巫毒迷信，所以你可能得經歷一番爭論才能取得測驗結果；這是因為這些資訊太過深奧，以至於人們認為只有交給心理學家或教育者保管才是安全的，而這樣的看法很可能是對的。無論如何，你以某種方式得知了彼得的智商是**98**、琳達是**101**，而智商的平均值或「標準值」是以**100**為基準。

啊哈！所以比較聰明的小孩是琳達，且她的聰明程度甚至高於平均水準。而彼得的智商低於平均水準，但我們還是不要深究這個問題了！

所有類似上述的想法，都是純粹的無稽之談。

如果你爸蠢到不知道I.Q是啥，
那I.Q再高又有什麼意義？



首先，容我在此澄清，我們該留意的事，是無論智力測驗到底測量了什麼東西，都不完全等同於我們平時說的「聰明」一詞。智力測驗忽略了領導能力、創意想像力等重要事物。該測驗不會考量到一個人的社會判斷力、音樂才能、藝術能力或其他資質，更不用說勤奮與情緒平衡等性格特質了。其次，最重要的是，學校提供的智力測驗往往是快速又廉價的團體測驗，在很大的程度上必須依賴閱讀能力；因此，無論你聰明與否，只要閱讀能力不佳，也只會得到糟糕的結果。

假設我們已經知道上述所有背景知識，也同意我們只是把智商視為一種衡量方法，用來處理刻板的抽象概念，模糊定義受試者的能力。而彼得和琳達則接受了社會普遍認為最完善的一項智力測試，也就是修訂版史丹佛—比奈智力測驗（**Stanford-Binet Intelligence Scale**，簡稱史比測驗），該智力測驗不需要任何閱讀能力。

接下來必須要理解的是，大致上來說，智力測驗是一種針對智力的抽樣研究。就像其他抽樣方法一樣，智商是一個具有統計誤差的數字——統計誤差會呈現該數字的準確度或可信度。

對受試者提出智力測驗的問題，就像是你為了評估玉米的品質，

而在玉米田裡的各個地方隨機摘下許多支玉米一樣。假設你摘下了**100**支玉米並仔細檢視，你將會對這整個玉米田的狀況有一定程度的了解。你獲得的資訊已經夠精確了，足以讓你把這塊玉米田拿去和其他玉米田比較——前提是這兩個玉米田不是非常相似。如果這兩個玉米田非常相似，你就必須檢視更多玉米，並運用精確的品質衡量方法進行評分。

你可以使用一種測量方法，來表示樣本代表整個玉米田時具有多高的準確度。**這種以數據來表達結果的測量方法，是「可能誤差」(probable error)與「標準誤差」(standard error)。**

假設你的任務是沿著田地柵欄用腳步測量多個田地的大小。你第一件要做的事可能會是走完你認為應該是**100**碼（約**91**公分）的步伐，之後再測量實際距離，藉由多次重複這個步驟來檢查這套測量系統的準確度。你可能會發現，以平均來說，你測量出來的**100**碼與真正的**100**碼相差**3**碼。也就是說，在你所有的測試結果中，共有一半的測試結果距離**100**碼正好不到**3**碼；另外一半的誤差則超過**3**碼。

那麼，你的可能誤差將會是**100**碼中的**3**碼，也可以表達為**3%**。由此可知，你之後用步伐測量出來的柵欄長度若是**100**碼，則應該被記錄為**100碼±3碼**。

現在多數統計學家比較偏好另一種類似的測量法，叫做標準誤差。這種測量方法算出來的數據，代表的不是二分之一樣本的準確度，而是三分之二樣本的準確度，一般認為這種方法在數學上更加便於使用。考慮到我們的目標，我們可以繼續使用可能誤差。目前仍有研究在使用可能誤差，且此計算方法與史比智力測驗有關。

正如我們在先前步伐測量中遇到的狀況，史比智力測驗的可能誤差也是**3%**。這個數字基本上與測驗本身的好壞沒有關係，它只與測試得出結果的一致性有關。所以，若我們用更加完整的方式來表示，彼得的智商應該會是**98±3**，琳達的智商則是**101±3**。

這個數字代表彼得的智商，有一半機率會落在**95至101**之間；另外一半的機率則是高於或低於這個範圍。同樣的，琳達的智商有一半機率會落在**98至104**這個範圍內。你馬上就可以從這些描述中看出，彼得的智商有四分之一的機率高於**101**，而琳達的智商也有四分之一的機率會低於**98**。這麼一來，彼得的智商可能就不會比琳達來得差，甚至還會產生**3**的差距。

歸根究柢，我們會發現看待智商的唯一方法，和看待其他抽樣結果的方法一樣，應該要用範圍來檢視。舉例來說，智商的「標準值」並不是**100**，而是**90至110**的範圍，而且，把「落在這個範圍內」的孩子，和「落在較高或較低範圍」的孩子比較，是有道理的。但是，把差距很小的智商拿來比較是沒有意義的。你必須隨時在心中牢記著正負值的概念，即使是（特別是）在研究沒有特別說明的時候。

毫無意義的「差很大」

所有抽樣研究都隱含這種誤差，過去已經出現過許多因為忽略誤差而導致的極端愚蠢行為。有些雜誌編輯把讀者調查視為聖經，主要原因在於他們根本不了解這些調查。這些編輯一看到調查指出某篇文章的男性讀者占**40%**，而另一篇文章的男性讀者只有**35%**，這些編輯就會開始要求雜誌推出包含更多類似前者文章的內容。

雖然對雜誌來說，**35%**與**40%**的讀者比例差異或許是很重要的，但是調查計算出來的差異有可能並不是真的差異。調查成本往往會使讀者樣本數下降到數百人左右，尤其是在完全淘汰掉那些根本不讀這份雜誌的人後，數字將會更少。依照內容，樣本中的特定群體可能會很少；例如，在讀者多為女性的雜誌中，樣本中的男性人數可能會非常少。等到這些人依照閱讀該文章的程度被區分成「閱讀全部」、「閱讀大部分」、「閱讀一些」和「沒有閱讀」後，**35%**這個結論所

立基的樣本數可能根本不到十人。在這個驚人數字背後，或許隱藏著極大的可能誤差值，大到依賴這個數字的編輯其實就像是在溺水時抓住了一根稻草。

有時候，人們為了數學上真實存在且可以證明的差異而費盡心力，但事實上這種差異有可能小到根本無關緊要。這也違背了古老的諺語：「唯有當差異真的能製造出差異時，差異才算是差異。」最符合這個敘述的例子，是老金香菸（**Old Gold**）依據一個幾乎毫無意義的數字差異所製造出來的大騷動，這場騷動的影響力甚大，且帶來極大的獲利。

這個例子始於《讀者文摘》（*Reader's Digest*）的一位編輯，雖然他有在抽菸，但他不以為然的認為所有香菸都是一樣的。《讀者文摘》進行了一項研究，找了一群實驗室人員分析數個品牌的香菸煙霧。該雜誌在刊出結果時，列出了各品牌香菸煙霧中含有的尼古丁量與其他成分。《讀者文摘》提出了結論並輔以詳細的數據，指出所有品牌的香菸煙霧幾乎完全相同，無論你抽的是哪一個牌子都沒有差別。

你可能會以為，這個研究結果將嚴重打擊煙草製造商和那些在廣告公司負責構思新廣告的人。或許這個結果會戳穿那些用廣告宣稱產品可以舒緩喉嚨，或對T字部位有益的香菸品牌。

但有些人在該研究結果中有了新發現。在這些毒素含量幾乎完全相同的香菸清單中，必定有某個品牌的香菸是排行最後的，而那個品牌就是老金。消息立刻傳了出去，報紙上立刻出現了由最大字體構成的大型廣告。廣告的標題與文案，只說了老金這個品牌在這本傑出的全國性雜誌進行的香菸測試中，其煙霧中不良物質含量最少。該廣告沒有放上任何數字，也沒有任何文字暗示這些香菸的差異小到應該忽略不計。

雖然到了最後，老金香菸被勒令「禁止」（**cease and desist**）刊

登這種誤導性的廣告。但這項禁令沒有帶來任何變化；老金香菸早就已經從這個廣告中得到大量好處了。

第五章

加工吸睛的圖表

人類是視覺化動物，比起文字數字，圖像更能抓住眼球。尤其是直接將統計數據的每個數值轉化為線圖時，會讓接受訊息的人更快的掌握狀況與趨勢。然而，即使原始的統計數據、圖表都沒有任何錯誤，但在統計操控上仍是可以透過「視覺」到「理解」中間的縫隙來插手。在本章，將說明統計是如何加工「圖表」達到操縱的加乘效果：

- 噁一下！視覺效果加倍
- 美化圖表就靠「換單位」



數字使人感到恐懼。

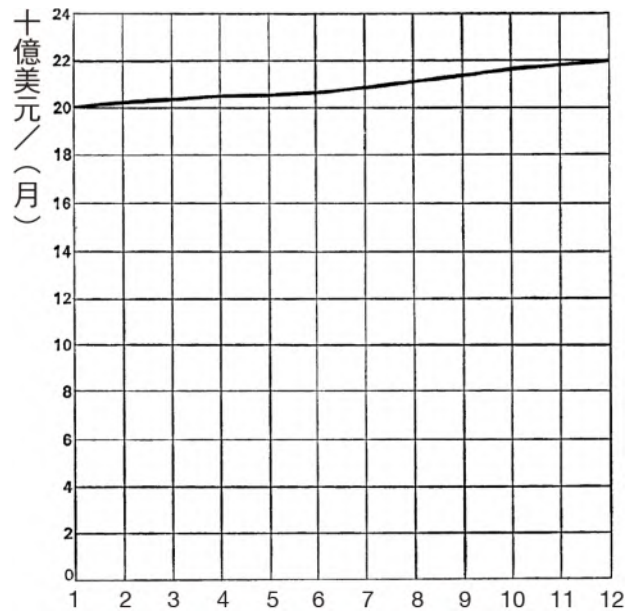
《愛麗絲夢遊仙境》的蛋頭先生（Humpty Dumpty）自信的告訴愛麗絲，他掌握了自己使用的各種字詞（word）。但這種自信對於多數人來說不太能夠擴展到數字上，這可能是因為小時候的數學課留給我們的創傷太嚴重了。

喀嚓一下！視覺效果加倍

無論我們為什麼害怕數字，數字都為那些渴望被閱讀的作家、期許文案能賣出商品的廣告人，以及希望自己的書籍或雜誌能受歡迎的出版商，創造了嚴重的問題。在常見的狀況下，在文章中放進帶有數字的表格是一種禁忌，而文字又無法發揮良好作用時，我們只剩下一個解決方案：畫一張圖表。

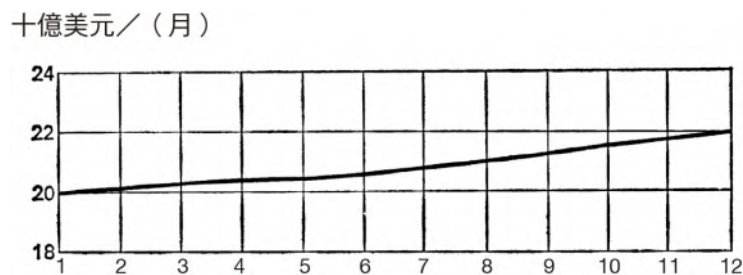
最簡單的一種統計圖表是線條變化圖（折線圖）。這種圖表非常適合用來表現趨勢——趨勢呈現的是所有人都想要展示、理解、聚焦、譴責或預測的事物。下面讓我們用統計圖表來呈現國民所得如何在一年內增加了**10%**。

首先，在紙上畫出方格。沿著方格的底部橫向列出十二個月。接著，在側邊標示十億美元。然後，在表格中標註每個月的點並連線，你的圖表看起來就會像這樣：



這張圖表已經很清楚易懂了。它逐月顯示了一年之內發生的事情。看到這張圖表的人可以很正確理解其中的資訊，這是因為這張圖表是完整且比例正確的，底下還有一條零基線做為對照。你的**10%**看起來就是貨真價實的**10%**——這個上升趨勢雖然十分顯著，但不至於令人感到勢不可擋。

如果你想做的事只有傳達資訊的話，這張圖表已經足夠。但假設你想要證明某個論點，使讀者震驚、推動他們採取行動、賣給他們某些事物的話，請將圖表的底部裁掉會更好。

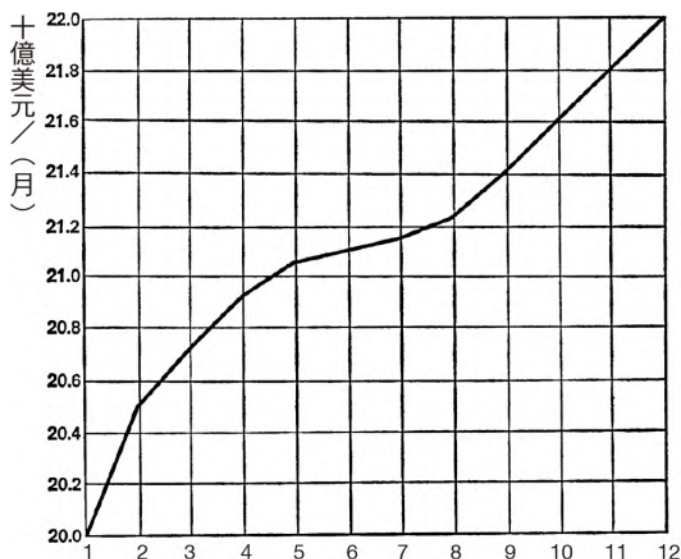


這下子才像樣（如果有人指責你這張圖表具有誤導性，你還可以指出這麼做能節省紙張）！這兩張圖表的數字相同，曲線也相同。它

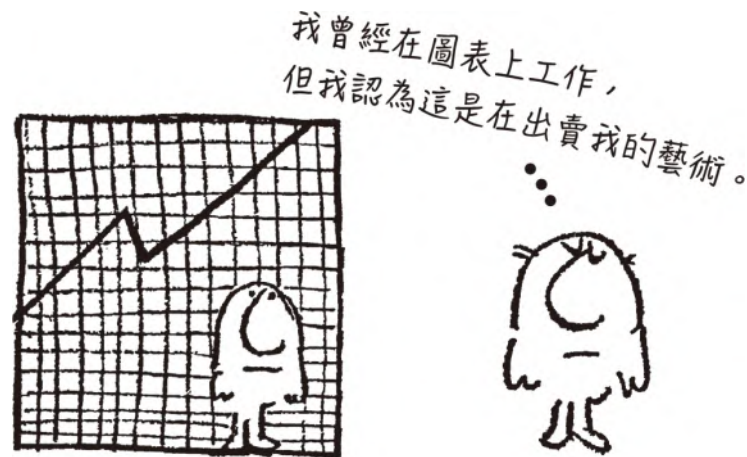
們是同一張圖表。後者沒有任何做假之處——唯有它給人的印象除外。然而倉促閱讀的讀者現在看到的國民所得曲線在十二個月內攀升了圖表一半的高度，這全都是因為圖表的絕大部分都已經不存在了。就像你在英文文法課會遇到一些省略部分字詞的句子一樣，就算少了某些部分，那些句子和這張圖片仍是可以「理解」的。當然，你的眼睛不會「理解」部分圖表已經消失，從視覺上來說，微小的上升變成了急遽的上漲。

美化圖表就靠「換單位」

既然你已經充分練習過如何欺騙了，為什麼不繼續裁切這些圖表呢？這是因為還有另一個進階技巧可使用。這個技巧，會讓**10%**這個適中的上升曲線，變得比**100%**的上升曲線所展現的畫面更有看頭。你只要更改縱座標與橫座標之間的比例就可以了。沒有任何規定指明你不能這麼做，而且你確實可以藉此獲得更漂亮的圖表。只要把側邊標記的所有金額都變成之前的十分之一就可以了。



這張圖確實非常驚人，對吧？所有看到這張圖表的人都可以深切感受到這個國家中生機蓬勃的繁榮。這張圖表等於巧妙的把「國民所得成長**10%**」編輯成「.....攀升了驚人的**10%**」。這種技巧遠比文字編輯還要有效得多，畢竟改變圖片後，你就不需要使用形容詞或動詞來破壞客觀的幻覺，更沒有人可以因此指責你。



而且你在這麼做的同時，你工作的公司很可能是一間優良的公司——或者至少是間可敬的公司。一家新聞雜誌曾用這種方法來呈現「股價創下新高」，這張被截斷的圖表使攀升的股價比實際情況更令人眼花撩亂。

哥倫比亞天然氣公司（Columbia Gas System）的一個廣告中刊登了一張重製的圖表，標明了「來自本公司的新年度報告」。如果你閱讀這些小數字並分析的話，你會發現使用者的生活成本在十年內上升了大約**60%**，但其中的天然氣支出下降了**4%**。雖然這張圖表看起來已經很不錯了，但哥倫比亞天然氣公司顯然覺得這對公司還不夠有利。他們把**90%**的圖表裁切掉了（沒有在圖中標註省略符號，也沒有用任何方式警告你這件事），於是你的眼睛是這麼告訴你的：生活成本增加了三倍以上，而天然氣支出則下降了三分之一！

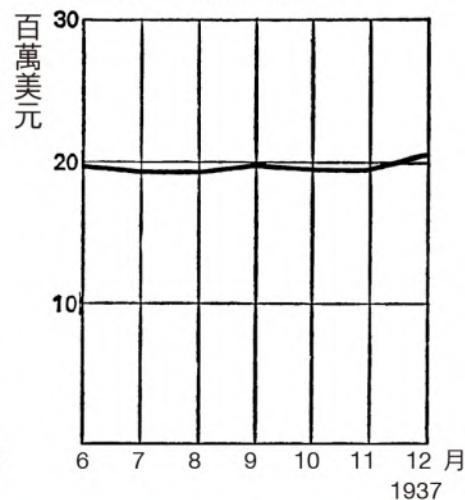
許多鋼鐵公司也使用具有誤導性的類似圖表，希望能拉攏輿論反對加薪。然而，這種編輯圖表技巧早就不是什麼新鮮事了，許久以前就已經有人揭露這種作法的不當之處——而且揭露這點的還不只是在統計學的技術刊物上。

1938年，《鄧氏評論》（*Dun's Review*）的一位社論作者轉載了一則華盛頓特區廣告中的圖表，該圖表的標題完美表達了其論點：「政府工資急遽上漲！」雖然圖表中的曲線像驚嘆號一樣一飛衝天，但曲線背後代表的數字卻並非如此。圖表展現的，是從**1,950**萬美元上升到**2,020**萬美元的增長。從圖表底部飆射出來的斜線一路上升到頂端，使得不到**4%**的成長看起來宛如**400%**。雜誌在一旁附上由一模一樣的數字繪製出來的圖表——這張圖表中的斜線誠實的上漲了**4%**，該圖表的標題如下：「政府工資保持穩定！」。

政府工資急遽上漲！

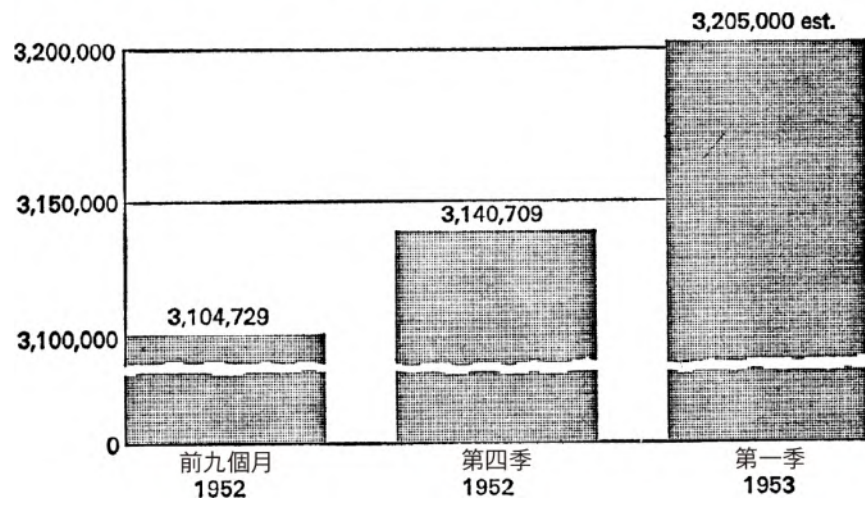


政府工資保持穩定！



《柯利爾百科全書》也在數篇報紙廣告中用相同的方法處理了柱狀圖。需要特別注意的是，圖表的中間已經被裁切掉了：

摘自 1953 年 4 月 24 日的《柯利爾百科全書》報紙廣告



第六章

平面圖的立體想像

為了更吸引人，統計圖通常不會只有乏味的線圖，而是會應用更有創意的立體圖像來呈現統計結果，也能讓人在閱讀時更好聯想到統計的內容。然而這種圖像在印刷上雖然是平面的，在我們生活中卻是立體的，一旦將圖像放大兩倍，效果卻會是立體化的八倍。在本章，將更具體的說明這種「圖像圖表」在統計操控上的使用：

- 「八倍效果」的立體化圖像
- 誤導性圖像，煽動表達意圖



大約十年前，你聽說過很多關於小人物的消息，所謂的「小人物」幾乎可以代表我們所有人。等到小人物這個詞聽起來太過居高臨下後，我們又變成了「一般人」。沒多久後，小人物就被遺忘了，這可能是件好事。但小人物並沒有消失，他仍是表格中的其中一個角色。

「八倍效果」的立體化圖像

你會看到某張圖表上用一名小人代表**100**萬人，用裝了鈔票或硬幣的錢袋來代表**1,000**英鎊或**100**萬美元，用牛隻的輪廓來代表明年的牛肉供應量，這種圖表叫做圖像圖表（**pictorial graph**）。這種圖表是非常好用的工具，但它所具備的吸引力令我擔憂，因為它極有可能成為口若懸河又陰險狡詐的詐欺犯。

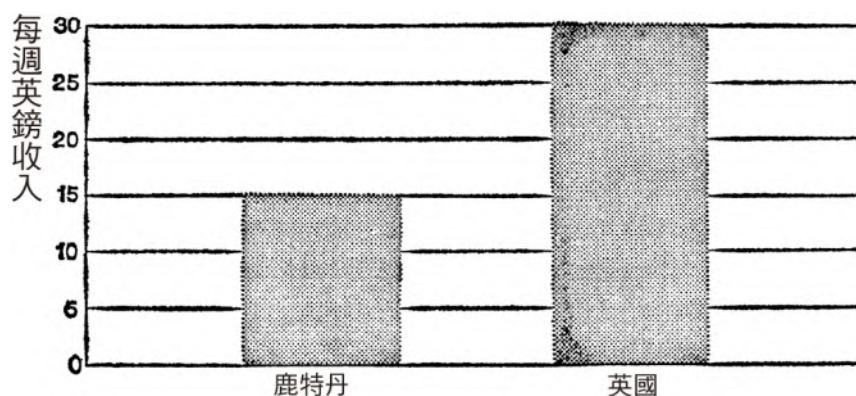
圖像圖表的鼻祖是如今十分常見的「柱狀圖」（**bar chart**）：一種簡單且受歡迎的圖表，可以用來比較兩個或多個事物的數量。柱狀圖也擁有欺騙我們的能力，請在看到柱狀圖時心存質疑，這種圖表可以在代表單一因素時改變其寬度與長度，也可以把圖表內容轉變成難以比較的立體物體的體積。「裁切過的柱狀圖」與我們剛剛討論過的「裁切過的折線圖」同樣聲名狼籍。柱狀圖常出現在地理書籍、公司報表和新聞雜誌上。而其他在柱狀圖之後衍生出來的吸睛圖表也是如此。



舉例來說，假設我想要比較兩個數字，英國和鹿特丹某種工人的平均週薪。這兩者的平均週薪可能是**30英鎊**和**15英鎊**。由於我想要吸引你的注意力，所以我不會只是列出數字，而會繪製一張柱狀圖。

順道一提，如果這個**30英鎊**的數字與你去年夏天新裝的門廊欄杆所支付的鉅額款項不符，請記住，這是因為這個工作所需工人的技術不像你的裝修工作般複雜。而且，我也沒有指明我說的是哪一種平均值，更沒有提出我是如何計算出這個數字的，所以你再怎麼對這個數字吹毛求疵都毫無用處。你發現了嗎？躲到不可信的統計數字背後，是如此簡單的一件事，只要在提供統計數字時別附上其他資訊就好。你可能已經猜到，這個數字其實只是我為了舉例說明而編出來的，不過我敢跟你打賭，如果我提出的數字是**29.35英鎊**的話，你一定猜不到這個數字是我編出來的。

就像這樣，你會看到圖表左側標上了每週英鎊收入。這是一張清楚又誠實的圖表。在這張圖表上，兩倍的收入看起來就是兩倍的大小。



但是，這張圖表不夠吸睛，對吧？我可以輕而易舉的利用比柱狀圖更像是收入的東西來創造吸引力：錢袋。我可以用一個錢袋來代表不幸的鹿特丹人微薄的收入，用兩個錢袋來代表英國人的薪資。又或者用三個錢袋代表鹿特丹人，六個錢袋代表英國人。無論如何，這張圖表仍然十分清楚且誠實，你不會因為只是匆忙一瞥就受到欺騙。這就是製作誠實圖表的方式。

如果我想做的事只有傳遞資訊的話，那麼這樣就已經足夠了。但是，我不只想要傳達資訊，我還想要告訴讀者，英國勞工的處境比鹿特丹勞工好上太多了。當我能愈戲劇化的描繪15英鎊與30英鎊間的不同，我的立論就愈有力。老實說（當然，我可沒打算要老實說），我希望你能用這張圖表推斷出某些事情，並留下深刻的印象，但同時我不想被你發現當中藏有詭計。有一種方法可以做到，而且你每天都會遭受這種方法的愚弄。

我只需要畫一個錢袋來代表鹿特丹人的15英鎊，再畫一個兩倍高的錢袋代表英國人的30英鎊，這樣就成功了。這兩個錢袋確實符合比例，不是嗎？這樣的圖片才能達到我想要讓人留下的印象。英國人的薪資如今使外國人的薪資相形見绌了。



當然了，接下來我要講的就是這張圖片出了什麼問題。由於第二個錢袋的高度是第一個錢袋的兩倍，所以它的寬度也是兩倍。第二個錢袋在圖片中占據的面積不是第一個錢袋的兩倍，而是四倍。雖然數字仍然是二比一，但以多數時候占主導地位的視覺印象來說，兩者的比例是四比一。當然，還可能會出現更糟的狀況。由於錢袋是現實中會出現的立體物品，所以你知道，第二個錢袋的厚度必定是第一個錢袋的兩倍。正如幾何教科書上所寫的，類似固體的體積變化會與任何相同維度的立體相似。所以，此時應該是2乘2乘2，答案是8。如果第一個錢袋裡裝了**15**英鎊，那麼因為第二個錢袋的體積增加了八倍，所以第二個錢袋裡裝的絕對不是**30**英鎊，而是**120**英鎊。

事實上，我巧妙的圖像圖表確實會給人這種印象。雖然我說的是「兩倍」，但在他人眼中留下的，卻是八比一這個壓倒性的深刻印象。

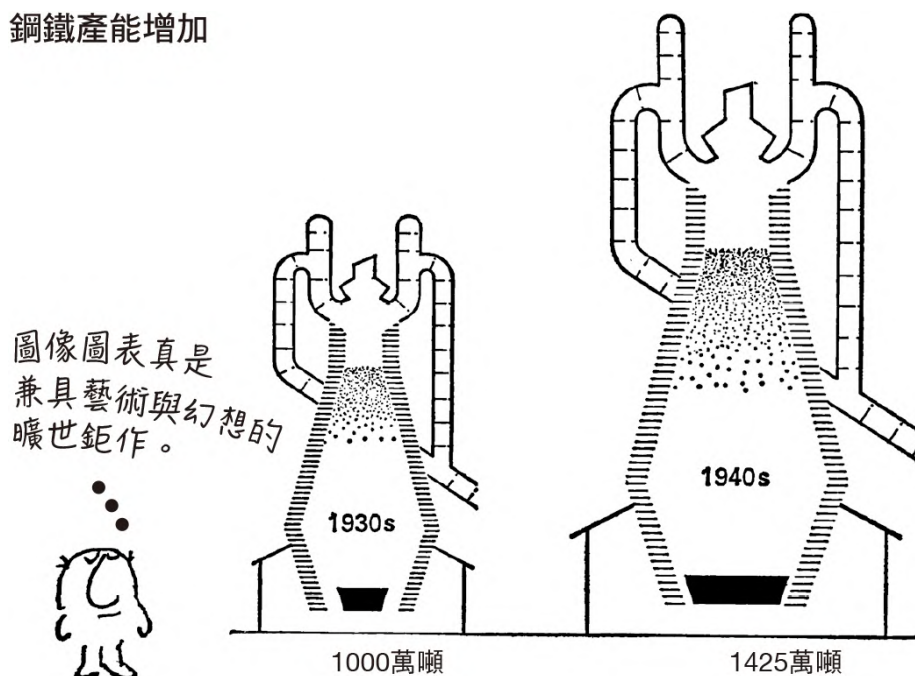
而且，你很難因為這張圖表而責怪我有任何犯罪意圖。我只是在做許多其他人也在做的事罷了。一家知名的新聞雜誌已經做過了——就像我所使用的錢袋一樣。

誤導性圖像，煽動表達意圖

美國鋼鐵協會（American Iron and Steel Institute）也做過了——他們使用的是兩座冶煉鋼鐵的高爐（blast furnace）。他們繪製這張圖片的目的，是展示鋼鐵產業的鋼鐵產量在過去1930至1940年代內急遽增長，藉此表明鋼鐵產業正靠自己做好這項工作，根本不需要政府進行任何干預。使用這種圖像繪製的優點遠比表面上還要更多。第一個高爐代表的是1930年代增加的1000萬噸產能，第二個高爐代表的則是1940年代增加的1425萬噸產能，而第一個高爐大約是第二個的三分之二高。此時，你眼前有兩個熔爐，其中一個幾乎是另一個的三倍大。當原本要表達的是「幾乎達1.5倍」時，其他人看見的卻是「三倍」——這就是平面圖像圖表能成就的事。

在鋼鐵協會製作的圖表中，有趣的還不只是這點而已。不知道為什麼，第二個熔爐變寬了，長寬比例超越了第一個熔爐，而且在第二個熔爐中，用來代表融化鋼鐵的黑色粗線，也變成了十年前的2.5倍長。鋼鐵協會給出的資訊是50%的成長，接著他們畫出了一張成長達到150%的圖，而該圖像創造出的視覺印象則是1500%的成長——除非我的計算尺標和我都失能了，否則事實就是如此。計算數學的技術變成了創造幻想的技術。

鋼鐵產能增加



感謝《鋼鐵之路》(Steelways) 提供圖片。

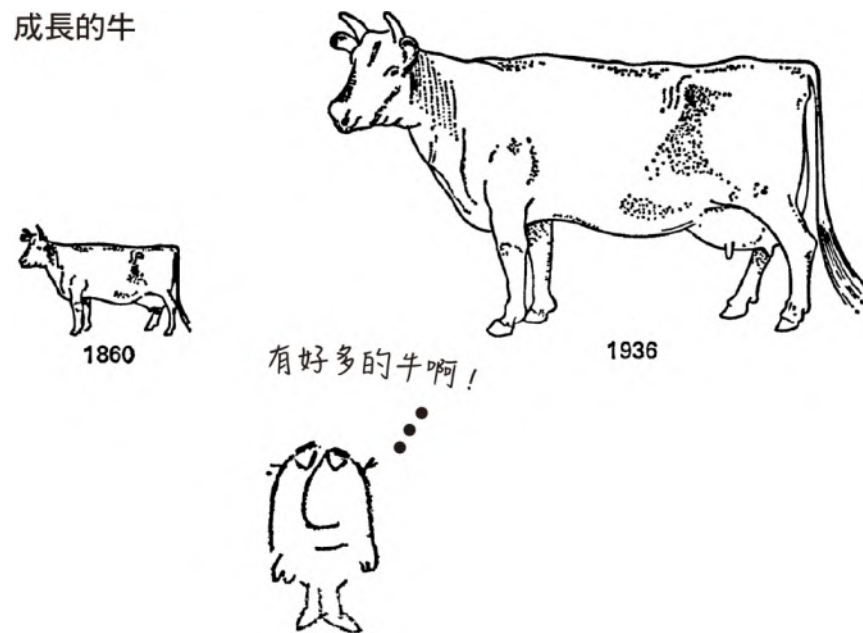
事實上，在同一張四色印刷的圖表頁面上，還很不友善的提供了一張從公平到優質的被截斷的折線圖。因為頁面上的曲線圖只用了該圖表的下半部，誇大了煉鋼產能的成長。這麼做不但能節省用紙，還能使上升率翻倍。

當然了，有些圖表會出現這種狀況，可能只是因為繪製時太過草率，就像是結帳時店員不小心犯錯少找零錢一樣。但如果每次錯誤都對店員有利時，你就會忍不住開始懷疑了。

《新聞週刊》曾用一張圖表展示「老年人如何變得更老」，該圖表上有兩名男性，其中一人代表如今的預期壽命為**68.2**歲，另一人代表**1879**年至**1889**年的預期壽命是**34**歲。這張圖表和前面的例子一樣：其中一人的高度是另一人的兩倍，因此他的體積或重量是另一人的八倍。為了創造出更吸引人的故事，他們利用這張圖表煽動式的傳達事實。我稱呼這種嘩眾取寵的新聞為「黃色新聞」(yellow journalism，指具強烈煽動性、過度渲染的新聞)。此外，該雜誌在

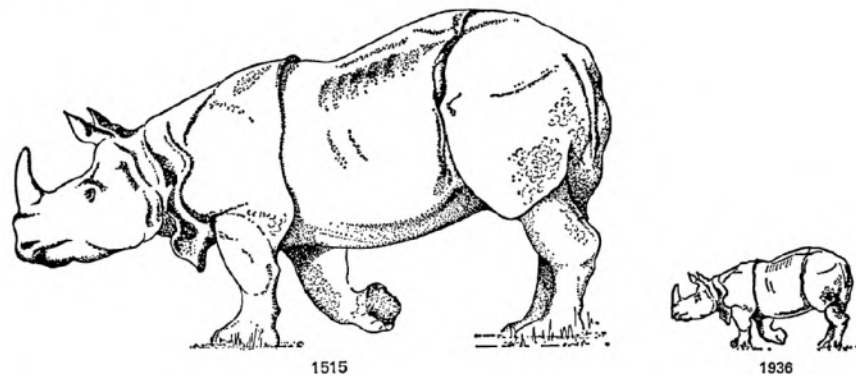
同一篇文章中，還包含了一張被裁切過、誇張的折線圖。

改變圖表中的物件大小還會帶來另一種危險。**1860**年，美國大約有**800**多萬頭乳牛；到了**1936**年，牛隻數量超過了**2500**萬。我們可以畫出兩頭牛來顯示數量上的增加，一頭牛的身高是另一頭的三倍，這麼做將會加深我們剛剛一直在討論的讀者印象。除此之外，這種作法還會使讀者在匆忙瞥見圖表時感到奇怪：他們很可能會以為現在的牛比以前的牛更大。



在你把同樣的欺騙技術應用到犀牛的數量變化上之後，你將會得到這張圖。美國詩人奧格登·納許（**Ogden Nash**）曾在詩中用犀牛（**rhinosteros**）和荒謬（**preposterous**）來押韻。後者也同樣可以用來形容這種欺騙技術。

數量減少的中犀牛



第七章

各自解讀「相關性」

相關性（correlation），又稱為相關、關聯，在統計學中，通常顯示了兩個或數個變數之間線性關係的強度和方向。但是在現實中，比起嚴謹的實驗統計數據，更常使用的是只具有部分相關的統計數據，並以此解釋得到想要的結果。在本章，將說明統計數據與結果之間的相關性，是否真的「有關係」，以及究竟是「什麼關係」：

- 「關係」，是可以創造的
- 財報常用！以相似名目藏錢
- 沒有關聯的前後比較



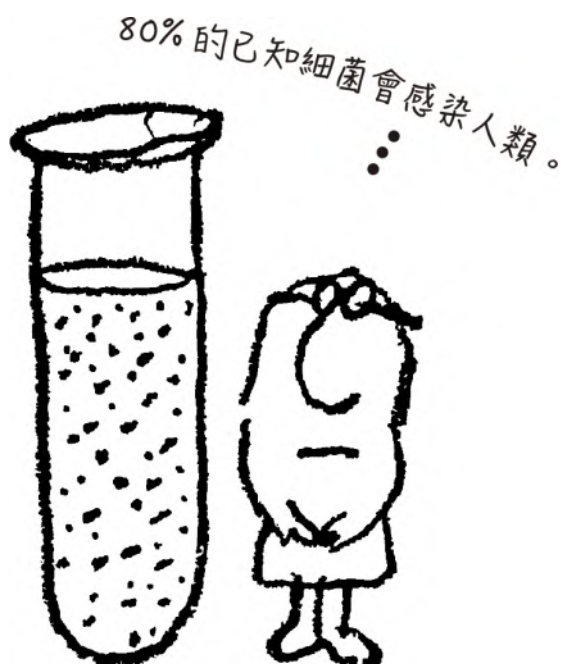
一位印度的法官曾對一位年輕、熱心的英國公務員說：「當你年紀再大一點時，就不會那麼自信的引用印度的統計數據了。因為雖然政府非常熱衷於統計數據——他們蒐集數據，將數據相加、求n次方、取立方根並製作出精彩的圖表。但你絕不能忘記，這其中的每一個數字都來自村莊的守衛（**chowty dar**），他只是寫下了他喜歡的東西。」

「關係」，是可以創造的

如果你沒辦法證明你想證明的，那麼你可以選擇展示其他事物，假裝這兩者是一樣的。由於人類的思想在遇上統計後總會產生茫然的感覺，所以幾乎沒有任何人會注意到其中的差異。只要你使用的是半關聯性數字（**semi-attached figure**）這種工具，就能保證對你有好處，一直以來都是如此。

雖然你無法證明某種藥方能治癒感冒，但你可以（用超大的字體）刊登一則宣稱絕對誠實的實驗報告，告訴讀者只要半盎司（**oz**，約**14**克）的藥方就可以在**11**秒內殺死試管中的**31,108**隻細菌。過程中，請確保你找的實驗室具有良好信譽或聲名遠播。你可以全文轉載那份報告，並找一名模特兒穿上白色醫師袍，拍一張相片，附在文章中。

但請不要提到你在這裡使用的詭計。你不能自行決定（又或者你可以決定？）要不要在文章中指出，試管中效果絕佳的抗菌劑不一定能在人體中發揮作用，更不用說你還得先根據規範，把藥劑稀釋到不至於灼傷喉嚨組織的程度。而且不要告訴讀者你殺死的是哪一種細菌，否則會混淆議題。誰知道感冒是由哪些細菌引起的呢？說不定引起這波感冒的根本就不是細菌？



事實上，試管中的各種細菌與導致感冒的原因沒有任何已知的關聯。無論感冒的成因是什麼，人們在感冒期間進行的推論通常都不會太嚴謹，尤其是在他們鼻塞的時候。或許這個例子太過明顯，人們已經開始有所警覺，但從廣告的頁數來看，這種藥還是賣得很好。

接下來是一個更巧妙的例子。

假設在種族偏見日益嚴重的某段時期，有人僱用你來「證明」相反的事實。這項任務不算困難。你可以進行民意調查，或者採用更好的方法：僱用聲譽優良的機構替你進行調查。從能夠代表社會大眾的人口橫斷面進行調查，詢問他們是否認為黑人找到工作的機會與白人相同。你可以每隔一段時間重複調查一次，如此一來就能在報告中標示趨勢。

普林斯頓大學的公眾意見研究辦公室（**Office of Public Opinion Research**）曾測試過這個問題，最後獲得了有趣的證據，證明了在民意調查中，事情並不總是像表面上看起來那樣。回答此問題的人也回答了其他問題，這些問題的目的是確認受訪者對於黑人是否抱持強烈

偏見。結果證明，偏見最嚴重的人，最有可能在回答「黑人找到工作的機會是否與白人相同？」時回答「是」。

經過計算，在同情黑人的受訪者中，大約有三分之二認為黑人找到工作的機率沒有白人那麼高，而在那些對黑人有偏見的人中，有三分之二認為黑人和白人的機會相等。很顯然的，你幾乎無法靠著這份民調了解任何有關黑人的就業狀況，不過你可能會對人們在種族偏見方面的態度獲得一些有趣的發現。

如上所述，在進行民調的期間，如果偏見愈來愈高的話，獲得的結果中就會有愈來愈多人認為黑人找到工作的機會和白人一樣高。於是你宣布了結果：你的民調顯示，這段時間以來，黑人找到工作的機率變得愈來愈平等。你謹慎的使用了半關聯性的數字，藉此取得了非凡的成就。**偏見愈是嚴重，民調的結果就愈好看。**

又或者我們可以舉另一個例子：「在數量龐大的傑出醫師樣本中，有**27%**的人抽潤喉牌（**Throaties**）香菸——比其他牌子的使用者都還要多。」當然了，從很多方面來說，這個數字本身就有可能是假的，不過無論這個數字是真是假一點都不重要。看到這個毫無關聯的數字時，你唯一能給出的答案是：「那又怎樣？」我們當然要對醫學專業抱持著應有的尊重態度，但是醫師真的比你還要更懂香菸品牌嗎？

他們會因為能夠取得內線消息而選擇危害最小的香菸嗎？當然不會，他們沒辦法獲得任何內線消息，而且你的醫師一定會是第一個這麼告訴你的人。然而，不知道為什麼，你總是覺得「**27%**」聽起來好像代表了什麼有意義的事情。

接著，讓我們稍微減一個百分點，看看某牌果汁機的狀況。大量廣告指出，「經過實驗室測試證明」，某種果汁機能夠「榨出的果汁量多出了**26%**」，而且獲得了「好管家學會（**Good Housekeeping Institute**）的擔保」。

聽起來很不錯。如果你可以選擇效率高出**26%**的果汁機的話，怎麼還會想購買其他品牌呢？我們在此先不討論「實驗室測試」（尤其是「獨立實驗室測試」）證明過的一些怪奇事實，只先討論一件事就好：這個數字代表了什麼意思？比什麼還要多**26%**？你將會在追根究柢後發現，這個數字代表的是這臺果汁機能榨出的果汁比老式的絞刀榨汁機更多。這個數據跟你在購買前想要了解的資訊毫無關聯；這款果汁機甚至有可能是市場上最糟的果汁機。**26%**這個數字不但精確到有些可疑，而且對你來說根本無關緊要。

會用數字愚弄你的，可不只廣告商而已。《本週》（*This Week*）雜誌曾刊登過一篇有關安全駕駛的文章，無疑是為了你好，才會在文章中告訴你若「以時速**70**英里（**160**公里）的速度在高速公路上左搖右擺的飛馳」的話，會發生什麼事。

文章指出，如果時間是早上**7**點，那麼你存活下來的機會將會是晚上**7**點的四倍。他們提出的證據是：「晚上**7**點在高速公路上發生的死亡事故數量，是早上**7**點的四倍。」好的，雖然這個數字大概是真的，但結論並不成立。晚上死亡的人之所以會比早上多，單純是因為晚上在高速公路上的人車比較多，所以死亡人數才會比較多。身為汽車駕駛的你可能會在晚上處於較大的危險，但在「四倍」這個數字之中，沒有任何證據能證明這一點。

按照該篇文章作者使用的荒謬邏輯，你也可以證明在晴天駕駛會比在霧中駕駛更危險。晴天發生的事故數量比較多，是因為晴天的天數比起霧的天數還要多。儘管如此，在起霧的天氣駕駛應該還是危險得多。

你可以用任何交通工具的意外事故統計數字來把自己嚇死.....前提是你沒有注意到這些數字和交通工具本身的關聯性有多麼薄弱。去年死於飛機失事的人比**1910**年還要多，所以現代的飛機比較危險嗎？當然不是！只不過是因為現在搭飛機的人數比以前要多上許多倍而

已。

根據報導指出，某一年因為火車事故而死亡的人數是**4712**人。這個數據聽起來是個遠離火車的好理由，你或許也可以靠著此論點堅持你要繼續開車。但是若你認真了解該數字所代表的意義後，會發現事情和你想的天差地遠。這個數字將近半數受害者是坐在汽車中，和火車在平交道相撞。餘下的受害者中，絕大部分都在搭火車時跨坐在火車車廂之外的鐵桿上^{*1}，只有**132**人確實是火車上的乘客。然而，就算是**132**這個數字，也幾乎無法證明任何比較上的差異，除非你在列出這些數字時，也一併附上乘客的總里程數。

如果你擔心自己在美國東岸與西岸間旅行時有高機率因事故死亡，即使調查去年致死人數最多的交通工具是火車、飛機或汽車，也不會為你帶來太多有用的相關資訊。想知道機率是多少，你要尋找的應該是每**100**萬乘客里程數的死亡人數。這個數字才能以準確性較高的方式告訴你，會在哪裡遇到最大的風險。

財報常用！以相似名目藏錢

除此之外，還有許多其他形式的研究可以在計算某件事物之後，在報告中把該事物說成其他事物。常見的方法是挑兩個聽起來相同、但其實截然不同的東西。假設你是一間公司的人事主管，你的公司和工會起了爭執，於是你「進行了一項調查」，想了解有多少人對工會感到不滿。你不需要考慮詢問工會，除非該工會的成員是一群如大天使領頭的天使，而且他們會完全誠實的回答問題並讓你記錄下來，且提供證據證明多數人都懷抱著不滿的情緒，藉此讓你把蒐集到的資訊做成報告，指出「絕大多數人（**78%**的人）反對工會」。否則你取得此數據的方式，應該是把一大堆毫無差異的抱怨和微小的不滿加總起來，把他們統稱為另一件聽起來很相似的事。雖然你沒有證明任何

事，但聽起來很像有，對吧？

不過，從某種程度上來說，這件事還是很公平的。因為工會也可以輕而易舉的「證明」，幾乎所有工人都反對工廠的運作方式。

如果你想要找出更多半關聯性數字的話，你可以去瀏覽各家企業的財務報表。請特別留意那些因為可能金額太高而隱藏到其他名目之下的利潤。聯合汽車工會（**United Automobile Workers**）的雜誌《彈藥》（*Ammunition*）以下列方式描述這種手段。

該聲明指出，去年某公司的利潤是**3,500**萬美元。每**1**美元的銷售額中，只有**1.5**美分是利潤。你為這間公司感到難過。此時，公司廁所的燈泡燒壞了。為了更換燈泡，公司必須支出**30**美分。**20**美元的銷售額帶給公司的利潤就這樣花掉了。真是讓人不禁想要少用點衛生紙呢。

但是，事實是，公司在報告中列出的利潤其實只有真正利潤的一半或三分之一而已。沒有列出來的那部分利潤隱藏在折舊、特殊折舊以及應急準備金中。

百分比也能為我們帶來同樣的樂趣。通用汽車（**General Motors**）在報告中指出，在最近的九個月內，他們的稅後銷售利潤（扣除稅後）為**12.6%**。但是在同一時期，通用汽車的投資利潤卻來到了**44.8%**，這聽起來糟透了——或者好極了，取決於你想要證明的是什麼觀點。

另一個類似的例子是《哈潑》雜誌的一位讀者在該雜誌信件專欄中為艾普賣場（**A&P**）辯護，指出他們的銷售額淨收益很低，只有**1.1%**。他質問：「有任何美國公民還會公開譴責他們是奸商嗎.....他們若投資了**1,000**美元，在一年後只能獲得**10**美元多的利潤呢。」

我們可能會反射性的認為**1.1%**聽起來小到令人沮喪。而且若將這個百分比，與我們比較熟悉的房貸或銀行貸款等**4%**到**6%**的利率或更高的數字相比，甚至還會認為，如果艾普公司能退出百貨業，將資金

投入銀行並靠利息維生應該能更賺錢吧？

問題在於，年度投資報酬率與總銷售收益完全是兩回事。正如另一位讀者在後來一期的《哈潑》上回覆：「如果我每天早上以**99美分**的價格購買一篇文章，每天下午以**1美元**的價格出售，那麼我的收入只有總銷售額的**1%**，不過，在這整年中我的收入卻占了投資金額的**365%**。」



一般來說，任何數字都具有多種的表達方式。舉例來說，你可以用這些說法來表達同一個事實：**1%**的銷售報酬率、**15%**的投資報酬率、**1,000萬美元**的利潤、**40%**的利潤成長（與**1935至1939年**的平均相比），或者比去年減少**60%**的利潤。取得成功的方法，就是選擇聽起來最適合你當下目標的那個說法，同時你還要相信，讀者鮮少會意識到，這個說法是在利用充滿瑕疵的方式反應出現實狀況。

沒有關聯的前後比較

並不是所有半關聯性數字都源自於蓄意欺騙。許多統計數字（其

中也包括了對所有人來說都很重要的醫療統計數字）之所以會扭曲，是因為數據源頭的報告不合邏輯。你可以在墮胎、非婚生子和梅毒等敏感的議題上，找到矛盾到非常驚人的數字。如果你稍微找一下有關流感和肺炎的最新數據，你可能會得出一個奇怪的結論，那就是這些疾病只出現在美國南部的三個州，這些病例占有通報病例的**80%**。這個比例背後的原因在於，在全美各州不再要求通報相關病例後，只有這三個州仍要求民眾通報。

舉例來說，有些瘧疾的數據幾乎毫無意義。在**1940**年之前，美國南部每年有數十萬的瘧疾病例，而如今的病例屈指可數，這幾年間顯然出現了對民眾健康有益的重要變化。但出現此變化的實際原因是，如今只有確實證明是瘧疾時，人們才會把病例記錄下來，而過去美國南方人常把瘧疾（**malaria**）當作口語用詞來指稱感冒或著涼。

噢不！我們只是有相似的曲線折向。



在美西戰爭期間，海軍的死亡率為**9‰**。同一時期，紐約市民的死亡率是**16‰**。海軍招募人員後來用這些數字來表明，加入海軍遠比在紐約生活更安全。讓我們先假設這些數字是準確的，而且它們也確實可能是準確的。但請停下腳步，試試看你能不能釐清，為什麼這些

數字毫無意義——至少這些招募人員從數字中得出的結論必定是毫無意義的。

這兩組數字是不能拿來做比較的。

海軍的主要成員是身體健康的年輕人；市民的組成則包括嬰兒、老人和病人，無論他們身在何處，他們的死亡率都會比較高。這些數字完全無法證明，符合海軍軍人標準的男性在海軍中的壽命，會比不加入海軍更長。同時這些數字也無法證明相反的論述。

在小兒麻痺症疫苗問世之前，你可能聽說過這種令人沮喪的說法：在醫學史上，去年是小兒麻痺症最嚴重的一年。這個結論是立基於任何人都能取得的證據：那一年的小兒麻痺症病例數字遠比過去還要更多。

但是，專家們在回顧這些數字時，發現了一些足以鼓舞人心的事。第一，那年有較多兒童的年齡都落在最易感染小兒麻痺症的區間，只要這個年齡區間的人數比例這麼高，病例數量就一定會創紀錄。第二，愈來愈多人了解什麼是小兒麻痺症，所以診斷出來的病例變多了，輕度病症的人數也變得更多。最後，由於財務誘因變多了，除了小兒麻痺症的保險變得更多，美國小兒麻痺基金會（**National Foundation for Infantile Paralysis**）提供的幫助也變得更多。出於上述這些原因，人們對於「小兒麻痺症的病例達到高峰」此一說法產生了不小的懷疑，而死亡人數也印證了這種疑慮是正確的。

有趣的是，在衡量某個疾病的發生率時，死亡率或死亡人數往往會比直接的罹病率還要更準確——原因很簡單，我們在遇到死亡時，會以較高的標準進行通報與記錄保存。很顯然的，在這個案例中，半關聯性的數字遠比表面上看起來具有完整關聯性的數字還要更適當。

美國每隔四年就會看到半關聯性數字大量出現一段時間。這並非因為該數字會出現週期性循環，而只是表示選舉時間到了。共和黨在1948年10月發表了一篇競選聲明，該聲明完全建立在看似有關聯，實

則無關的數字上：

杜威 (Thomas Edmund Dewey) 在1942年當選紐約州長時，部分地區的教師最低薪資下跌至每年**900**美元。如今紐約州的學校教師享有全球最高的薪資。杜威州長依據他任命的委員會提出的調查結果給予建議，推動立法機關在**1947**年從該州財政盈餘中撥款**3,200**萬美元，立即增加學校教師的薪資。紐約市的教師最低薪資因此落在**2,500**美元至**5,325**美元之間。

杜威先生可以用各種方法來證明自己是教師的好朋友，但這些數字並沒有做到這一點。這篇聲明是個老套的「之前與之後」技巧，也就是在文中使用許多未經詳細解釋的因素，使得這些因素顯得像是另一回事。

你在這裡得到的「之前」是**900**美元，「之後」則是**2,500**至**5,325**美元，聽起來確實有改善。但是比較小的數字 (**900**) 其實包括了該州所有郊區的最低薪資，而較大的數字 (**2500 ~ 5325**) 則單指紐約市的薪資。或許教師薪資確實在杜威州長的領導下有所改善，但也可能根本沒有。

這份聲明使用的是「之前與之後」的統計方法，而雜誌業與廣告業比較熟悉的手段則是「之前與之後」的相片比較法。你看到了兩張客廳的照片，這兩張照片的目標是讓你看見一層油漆能使環境出現多大的改善。但是在這兩次拍攝之間，客廳裡多了一些新家具，而且「之前」是光線黯淡的小張黑白照片，而「之後」則是大張的全彩照片。

又或者，你會看到兩張圖片展示了某個女孩開始使用某牌潤髮乳之後產生的變化。天啊，她用過潤髮乳之後看起來確實比較好看呢！但你仔細觀察後發現，多數變化其實來自於她新展露的笑容與投射在頭髮上的背光。真正的功臣並不是潤髮乳，而是攝影師 (或許還有修

圖軟體）。

編按：常見於大蕭條時期，是貧窮且須長途跋涉工作的人搭乘火車的模式。

第八章

事後歸因的誤導性

事後歸因 (post hoc) 是一種謬誤，指從結果倒推原因，但事實上，結果和原因彼此之間並無直接關係。這是因為結果和原因可能有相關，卻缺乏明確關係，這中間就出現了可供操縱或造成心理盲點的空間。在本章，將說明這個在日常中尋常可見，把「偶然」當「必然」的假因果關係：

- 同時變化，因果難分
- 有相關不代表有因果



你可以計算荷蘭或丹麥家庭屋頂上的鸛鳥巢，來估計這些家庭出生了多少嬰兒。鳥巢多的家庭可能嬰兒也比較多，藉此進行的估算遠比隨機猜測的結果來得好。

統計術語中，我們會說這兩件事之間存在正相關。這聽起來像是只能證明古代神話真實性的證據，但實際上其實更有價值。它提醒我們一個有用的事實：**兩個因素之間有相關聯，並不能證明其中一個因素導致了另一個因素。**

在鸛鳥和嬰兒的例子中，不難找到可能導致這兩個因素的第三個因素。大房子吸引大家庭，甚至是想要有很多孩子的家庭；大房子有更多的煙囪，鸛鳥可以在上面築巢。但因果關係假設中的缺陷，並不總是那麼容易就被發現，尤其是這種關係似乎很有意義或符合大眾偏見時。

同時變化，因果難分

曾有人費盡心力，想釐清吸菸者的大學課業成績是否低於不吸菸者。事實證明了吸菸者的成績確實較低，許多人對此感到滿意，在那之後一直很重視這個統計數字。表面上看來，取得好成績的途徑就是戒煙；還有些人因此獲得了更進一步的結論，認為抽菸會使思想變遲鈍。

我相信這項研究的調查過程是得當的：樣本數夠大、經過誠實且謹慎的選擇、相關性（**correlation**）具有極高的顯著水準等。

不過，這個研究卻犯了一個古老的謬誤，這個謬誤非常容易出現在統計工具中，被大量的驚人數字隱匿在底下。這個謬誤使我們覺得，如果**B**跟著**A**出現的話，那麼就是**A**導致了**B**的出現。也就是說，人們因為抽菸和較低的成績同時出現，而做出毫無根據的假設：吸菸導致較低的成績。但說不定因果關係其實完全相反，不是嗎？或許較低

的成績會促使學生抽菸而不是喝酒。追根究柢，這兩個結論的可能性一樣高，而且都同樣有證據能支持其論述。但負責宣傳的人當然不會對此感到滿意。

不過，更有可能的是，這兩個事物都沒有導致另一個事物的出現；事實上，這兩個事物都是第三個因素的產物。說不定比較善於社交的人本來就比較沒那麼認真讀書，同時也比較有可能抽較多菸？也或許我們可以從別的地方找到線索：曾有人在外向性與低分之間建立了相關性——這種相關性顯然比成績與智力之間的關係更加密切：或許外向的人比內向的人更常抽菸。重點在於，當某件事擁有非常多種合理解釋時，你沒有權利選擇合你胃口的解釋，並堅持那是對的。但有很多人都會這麼做。

為了避免陷入事後歸因的謬誤，導致最終相信許多並不為真的事物，你必須嚴格檢視所有對關係的陳述。相關性是一個具有極高說服力、似乎可以證明兩件事情具有因果關係的數字，但實際上它有許多類型。

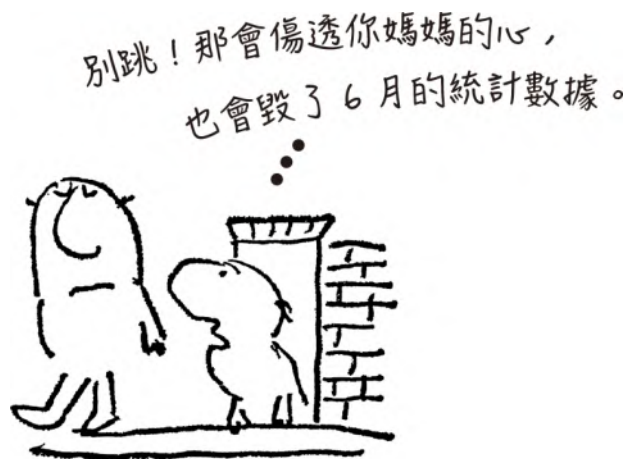
其中一種相關性是由偶然產生的。你可以把某一組數字放在一起，用這種方式證明一些不太可能成真的事，但如果你用同樣的方法再嘗試第二次，你很可能根本無法證明任何事。就像牙膏公司提出的數據看起來似乎可以減少蛀牙一樣，只要把不想要的結果丟掉，刊登想要的結果就行了。或者只要選定的樣本數夠少，就很有可能從特徵或事件之間得到你想要的顯著相關性。

我們常看到的**共變異數 (covariation)**就是一例，在共變異數的關係中，你無法確定哪一個變數是成因，哪一個變數是結果。在某些狀況下，兩個變數的因果關係可能會偶爾彼此對調，又或者兩者同為彼此的因與果。收入與股票所有權之間的相關性就有可能是如此。你賺的錢愈多，買的股票就愈多，而你買的股票愈多，收入就愈多；單純的指出「其中一方導致另一方的發生」是非常不準確的描述。

在各種相關性中，最棘手的或許是一個很常見的狀況：兩個變數對彼此都沒有任何影響，但又確實具有相關性。已經有許多人利用這種相關性完成各種髒活了。吸菸者與成績較差之間的關係就是這種相關性，有太多文章在未經認證的狀況下所引用的醫學統計數據也來自這種相關性，儘管這些相關性確實是真的，但其中的因果關係卻只是一種猜測。舉例來說，曾有人與高采烈的指出一個真實的統計數字，並附上一套毫無道理的虛假相關性描述：麻薩諸塞州長老教會牧師的薪水和古巴哈瓦納（**Havana**）的蘭姆酒價格之間，具有非常密切的相關性。

哪個是成因，哪個是結果呢？換句話說，是這些牧師透過蘭姆酒的交易獲利？還是這些牧師在推動蘭姆酒的銷售？好啦，這種結論太牽強了，光是看一眼就會讓人覺得太過荒謬。但請小心留意其他事後歸因的應用，那些應用與這個例子之間的唯一差異，就是它們比較隱晦。在牧師與蘭姆酒的例子裡，你可以輕而易舉的看出這兩個數字，都是因為另一個因素影響而增長：幾乎全世界的所有事物都出現了史無前例的價格上漲。

讓我們用另一個數字舉例，自殺率通常會在**6**月達到最高峰。是自殺導致**6**月新娘的出現——還是**6**月婚禮推動突然被拋棄的人自殺？另一個更有說服力（不過同樣未經證實）的解釋，是那位花了整個冬季試圖克服憂鬱症的人一直認為到了春天事情就會好轉，但到了**6**月他卻仍覺得一切都糟透了，於是他放棄了。



有相關不代表有因果

另一件需要留意的事，是有些文章會在呈現了數據後，提出超越數據的相關性並由此推斷結論。我們可以輕而易舉的證明，當一個地區常下雨，那裡的玉米就會長得愈高，甚至連作物都會長得愈大。這樣看來下雨似乎是一件好事。但是過度降雨的季節可能會對作物造成嚴重損害，甚至徹底毀掉所有作物。兩者之間的正相關只會維持到某一個程度，之後就會迅速變成負相關。在降雨超過某個量後，雨下得愈多，玉米產量就愈少。

相關性會顯示出某種關係傾向，不過這種傾向往往不會被描述為一對一的理想關係。平均而言，高個子男孩的體重會比矮個子男孩重，因此身高與體重具有正相關（**positive correlation**）。但是，你也可以輕而易舉的找到一位6英尺（180公分）高的瘦男孩，體重比一些5英尺（150公分）高的男孩更輕，因此身高與體重的關聯性小於1。

而**負相關（negative correlation）**是指當一個變數增加時，另一個變數傾向減少。在物理學中，負相關則會變成反比：你距離燈泡愈遠，書上的光就愈少；隨著距離增加，光的強度會降低。這種物理

關係通常會十分友善的表現出完美的相關性，但來自商業界、社會學界和醫學界的數字鮮少這麼漂亮。雖然人們普遍認為教育會增加收入，但教育也很可能變成某人破產的原因。別忘了，**儘管相關性可能是真的，也可能是基於真實的因果關係，但在單一個案的決定性行動中，這種相關性仍舊幾乎毫無價值。**

已經有研究蒐集了厚厚一疊數據，指出大學教育對未來收入的價值，並出版了成堆的小冊子，讓有潛力的學生注意到這些數字——還有以這些數字為基礎做出的結論。我並不打算爭論這種研究的意圖為何。我很贊成教育，尤其是那些包括了初級統計學課程的教育。現在這些數字已經表明了決定性的結論，讓你知道上過大學的人賺的錢比沒上過大學的人更多。社會上當然也有許多例外，但這個傾向既強勢且清楚。

這段論述中，唯一錯誤的，隨著這些數字與事實的出現，得出了一個完全沒有根據的結論。這就是事後歸因謬誤的絕佳案例。該結論與相關數字表明，如果你們（你的兒子或女兒）念了大學，那麼相較於不上大學而把這四年用在別的地方，念大學很可能會讓你們賺進更多錢。這種毫無根據的結論具有一個同樣毫無根據的假設，也就是受過大學教育的人之所以會賺進比較多錢，是因為他們念過大學。雖然我們無法確定，但這些人也有可能是就算沒上大學也會賺比較多錢的那群人——有幾件事相當強烈的表明了事實如此。大學裡面有兩種小孩的數量特別多：聰明的小孩和有錢的小孩。聰明的小孩就算沒有大學知識，也可能表現出優秀的賺錢能力。至於有錢的小孩……有錢人顯然有很多方式可以用錢滾錢。無論有沒有念大學，都很少會出現有錢人的小孩落入低收入階層的狀況。

以下內容摘自一份發行量極大的週日增刊雜誌所刊登的一篇問答式文章。值得令人玩味的是，此文章的作者曾發表過的另一篇文章，名為〈流行觀念：是真是假？〉（Popular Notions: True or

False?) 。

問：念大學對你保持未婚狀態有什麼影響？

答：如果你是女人，你成為未婚女性的機率將直線飆升。但如果你是男人，大學會帶來相反的效果——大幅減少保持單身的機率。

康乃爾大學 (**Cornell University**) 曾對**1500**名大學畢業的中年人進行研究調查。在男性中，**93%**已婚 (相較之下，男性一般群眾的已婚率是**83%**) 。

但在大學畢業的中年女性中，只有**65%**已婚。女性單身的比例，大學畢業的中年女性是一般女性群眾的三倍。

17歲的蘇西·布朗 (**Susie Brown**) 讀到這篇研究時意識到，相較於不念大學，念大學之後的結婚機率會變得較低。這篇文章就是這麼說的，而且文章中附上的統計數字具有信譽良好的資料來源。儘管這些數字與結論相符，但這些數字其實不能支持該結論的正確性；此外我們還要留意，雖然這些統計數字來自康乃爾大學，但結論卻不是，不過，匆促閱讀這篇文章的讀者很可能會覺得結論也來自康乃爾大學。

我們再次看到，真實的相關性被拿來支持未經證實的因果關係。或許這個結論把因果弄反了，就算這些女性沒有念大學，她們也不會結婚。說不定這些念大學的女性如果沒有念大學，結婚率會變得更低。如果這些論述的可能性幾乎等同於作者堅持的結論，那麼這些論述的有效程度想必也和作者的結論相當：全都只是猜測。

請蘇西·布朗特別留意：事實並不一定如此。

一篇醫學文章曾以格外憂慮的態度指出，喝牛奶的人有較高機率罹患癌症。癌症在新英格蘭地區、明尼蘇達州、威斯康辛州和瑞士出現得愈來愈頻繁，這些地方的牛奶生產量與消費量都很大，而牛奶產量極少的錫蘭 (**Ceylon**，斯里蘭卡) 則仍然少有癌症。為了進一步證

明此事，文章指出在美國南方一些牛奶消費量較少的州，出現癌症的機率也比較低。此外，文章還提到，相較於較少喝牛奶的日本女性，常喝牛奶的英國女性罹患某些癌症的比例是日本女性的**18**倍。

只要稍微深入挖掘，就會發現有許多方法都能解釋這些數字。不過，其實只需要一個因素就足以做出解釋了：癌症是一種主要會在中年或中年過後發作的疾病。瑞士和一開始提到的那幾個州，同樣具有相對較長的人口壽命。在進行這項研究時，英國女性的平均壽命比日本女性多出**12**年（編按：現在日本女性平均壽命已超過英國平均）。

海倫·M·沃克教授（**Helen M. Walker**）提出了一個有趣的例證，說明如果每次兩件事都會一起變化時，就假設兩者具有因果關係是多麼的愚蠢。在調查女性的年齡與部分身體特徵之間的關係時，研究人員會先測量女性走路時的腳掌角度。你會發現年齡較大的女性走路時的腳掌角度會比較大。一開始，你可能會覺得這或許代表女性會因為外八而變老，接著你立刻意識到這個結論有多荒謬。那麼，顯然這兩者的因果關係是年齡增加會使得腳掌角度增加，而且多數女人必定會隨著逐漸老化而變得更加外八。

所有類似的結論都很有可能是錯的，而且當然也是沒有根據的。想要得到具有正當性的結論，你必須在一段時間之內研究同一群女性，或盡可能以相等的兩個群體來研究。這麼做才能消除導致此統計數字的真正原因：在較年長的女性成長的時代，人們會教導年輕的女性外八走路，而在如今比較年輕的女性成長的時代，人們則比較不鼓勵外八走路姿勢。

若你注意到有些人（通常可能是因此獲利的一群人），因為某個關聯性而過度激動的話，請先檢視一下那是不是這種因為連續發生的事件或時間趨勢而產生的。在如今這個時代，想要提出兩件事物間的正相關是很容易的一件事：大學學生人數、精神病院收容人數、香菸消費量、心臟病發病率、X光機使用次數、假牙製造量、加州教師薪

資、內華達州賭場的利潤。把這之中的其中一件事稱為另一件事的成因顯然是愚蠢的行為。但如今每天都有人這麼做。



許多人都利用統計方法以及數字與小數點的催眠效果來混淆因果關係，如果你允許這種事情發生，那就跟迷信沒有兩樣。而且，這種混淆所誤導出來的結果往往比迷信更嚴重。這就像是新赫布里群島（**New Hebrides**）的人，堅信體蝨（**lice**）能帶來健康的身體一樣。他們透過數個世紀以來的觀察發現，身體健康的人通常都有體蝨，而病人往往沒有。這個觀察本身既準確又切實，而我們在其他文章中看到的多年非正式觀察往往也準確又切實得令人驚訝。不過，這些原住民從證據中得出的結論倒是令人不敢苟同：體蝨使人保持健康，每個人都應該擁有體蝨。

正如我們剛剛提到的，許多人靠著比新赫布里人更不足的證據——他們只要把數字放進統計學磨坊裡，直到常識再也看不透這些數字就行了——在醫療方面賺進大筆財富，並在雜誌中發表醫學文章，其中也包括專業雜誌。後來終於有判斷力更好的觀察家釐清新赫布里群島到底發生了什麼事：原來這些社群中幾乎每個人都擁有體蝨。或許可以

說，對他們來說有體蝨才是正常狀態。不過，只要有人發燒了（很可能是因為體蝨帶來的疾病），這個人的身體就會因為溫度過高而不適合體蝨居住，於是體蝨就會離開。這正是個成因與結果以無比混亂的方式扭曲、顛倒、交纏的案例。

第九章

統計操控七大手法

統計數據是世上的三大謊言之一。我們必須認知到：讓數字「說話」的是人，而人可以進行操控。在前面幾章我們細數並介紹了統計中的樣本、計算、圖表等基本概念，而在本章一次統合了前述統計內容中的各種操控手法，提醒所有看見統計數據的人們都應提高警覺：

- ◎**手法1**：濫地圖扭曲的真相
- ◎**手法2**：誇大不實的「平均」
- ◎**手法3**：小數點，假精確錯覺
- ◎**手法4**：讓人混淆的基數
- ◎**手法5**：「百分比」與「百分點」
- ◎**手法6**：選擇最有利的數據
- ◎**手法7**：比較基準的設定差異



我們可以把使用統計資料來誤導他人的行為，稱作「統計學式的操控」，並將之簡稱為（雖然不怎麼樣）「統計操控」（**statisticulation**）。

本書的書名與部分內容似乎都在暗示，所有類似的行為都是刻意欺騙的產物。美國統計協會（**American Statistical Association**）的一位分會會長曾為此打電話給我，說這種行為很多時候都不是來自詭計，而是來自無能。雖然他說的話或許蘊含了一些道理^{*1}，但我不確定這兩個原因哪一個比較不冒犯統計學家。相較於此，或許更重要的是記住這件事：有時扭曲與操控統計數據，並不是專業統計學家的工作。就算放在他們桌上的數據符合道德標準，這些數據也可能會被推銷員、公共關係專家、記者或廣告文案寫手扭曲、誇大、過度簡化，並藉由挑選而曲解。

但無論在什麼狀況、有罪的一方是誰，我們都很難認同他們只是粗心的無辜者。雜誌與報紙上的虛假圖表，經常透過誇大其辭來駭人聽聞，而不是輕描淡寫的簡單帶過。就我的經驗，在那些為企業提出統計數據的人之中，也很少有人會呼籲要提供比現實狀況更好的待遇給勞工或客戶，他們提供的待遇往往更糟。此外，你什麼時候看過工會僱用的統計工作者，無能到把勞工的狀況變得更加不利呢？

只要這些錯誤一直都只對發話者有利，我們就很難把錯誤歸咎於粗心或意外。

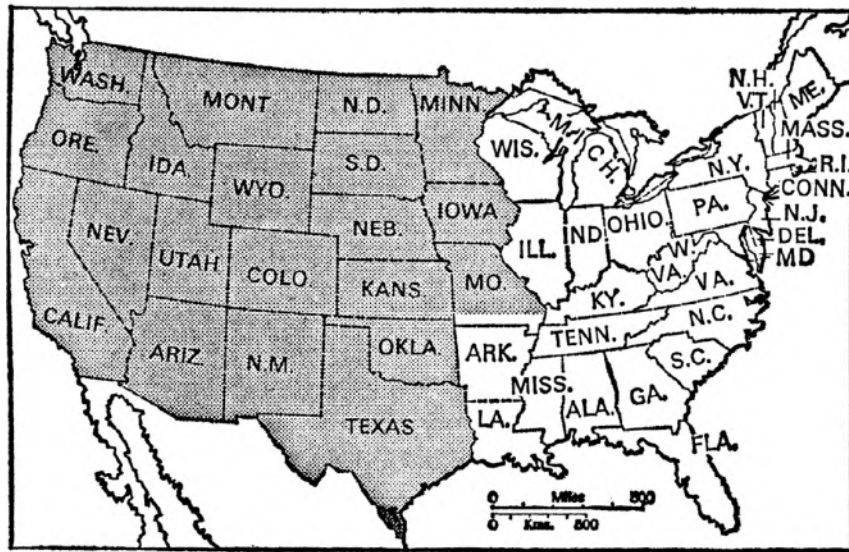
手法1 遭地圖扭曲的真相

在扭曲統計數據的方法中，最棘手的其中一個是使用地圖。地圖會引入一大堆變數，並在其中隱藏事實、扭曲事物之間的關係。在這個領域中，我最喜歡的冠軍是「深色陰影」。該地圖是波士頓國家第一銀行（**First National Bank of Boston**）不久前分發的宣傳品，而納

稅人團體、新聞媒體與《新聞週刊》紛紛廣泛使用了這張地圖。

這張地圖顯示了，美國聯邦政府如今把多少比例的國民所得拿走並花掉。製圖者為密西西比河以西的地區地圖上色（僅阿肯色州、路易斯安那州與部分密蘇里州除外），藉此表示這些地方聯邦政府的支出已經和該州國民所得相等。

深色陰影（西方版）



（東方版）



為了表明我們採用的是沒有作弊的良好測量法，我們特別加上了馬里蘭州（MD）、德拉瓦州（DEL）與羅德島州（R.I.）。

這張地圖的欺騙之處在於，製圖者選擇的都是面積很大、但卻因為人口稀少所以人口所得相對較低的州。如果製圖者秉持著同樣的誠實態度（或者說同樣的不誠實態度）製作另一張圖，他或許得為紐約州或新英格蘭州上色，並因此獲得一片面積小得多也不怎麼令人驚訝的深色區塊。他使用的是一模一樣的數據，但卻會在觀看地圖的讀者腦中留下截然不同的印象。不過，沒人會費心去分發第二張地圖。至少就我所知，沒有任何權力夠大的團體，會希望使公眾支出看起來比第一張地圖更少。

如果製圖者的目標只是傳達資訊的話，其實可以用很簡單的方式做到這一點。他可以選擇一個處於中間的州，這些州的總面積和美國總面積之間的關係，大約等同於這些州的總收入和美國國民所得之間的關係。

這張地圖之所以是一種格外明目張膽的誤導方式，是因為這已經不是新的宣傳伎倆了。這是一種經典手段，或者可以稱作陳腔濫調。波士頓國家第一銀行在許久之前曾出版過這張地圖的各種版本，用這些地圖展現1929年和1937年的聯邦支出，沒多久後，這些地圖便出現在威拉德·柯普·布林頓（Willard Cope Brinton）的典範著作《圖形演示》（*Graphic Presentation*）中，這些地圖是非常可怕的例子。布林頓直截了當的指出，這種方法「扭曲事實」。但是，波士頓國家第一銀行仍舊不斷繪製地圖，而《新聞週刊》與其他應該要更明事理的人（他們很可能確實更明事理）也繼續刊登這些地圖，既沒有附上警告，也沒有為此致歉。

手法2 誇大不實的「平均」

如果你認為現在有通貨膨脹，請思考一下這個狀況。某年人口普查局宣稱該年度「家庭平均收入是3,100美元」。但是，如果你讀過羅

素塞吉基金會（Russell Sage Foundation）刊登在報紙上的「慈善捐贈」報導，你就會知道，該報導指出同一年的家庭平均收入是驚人的**5,004**美元。雖然你可能很高興美國人民的收入這麼好，但你也可能會因為該數字和你在生活中的觀察完全不符而感到震驚。猜想大概是因為自己沒有認識到相符的人吧。

那麼接下來我們要討論的是，羅素塞吉基金會和人口調查局的數字怎麼會相差這麼多？人口調查局提出的是中位數，而且這是理所當然的事，但就算羅素塞吉基金會使用是算數平均數，兩者間的差異也不應該這麼大。事實證明，羅素塞吉基金會之所以能找到這個不得了的繁榮數字，是因為他們製造出了我們只能稱作「假家庭」的家庭單位。他們（被他人要求解釋時）解釋說，計算出這個數字的方法是把美國的總個人所得除以**1.49**億（當時美國人口），獲得每人**1,251**美元的平均值。「因此，」他們補充：「在四人家庭中，平均收入就是**5,004**美元。」

這種奇怪的統計操控在兩個方面做了誇大。**第一，它使用的是算數平均數**，而不是數字較小且資訊量較大的中位數.....我們已在先前的章節討論過此事。**其次，這個統計操控假設家庭的收入與其成員數成正比**。我現在有四個孩子，所以非常希望家庭收入真的和成員數成正比，但事實並非如此。四口之家的富裕程度通常並不是兩口之家的兩倍。

或許羅素塞吉基金會的統計學家並沒有意圖犯下欺騙的罪行。平心而論，他們感興趣的應該是描繪出踴躍捐贈的景象，而不是賺很多錢的樣子，這個有趣的家庭收入數字只是副產品。但是，無論如何，這個數字仍達成了同樣的欺騙效果，也仍然是絕佳例子，可以用來解釋我們為什麼幾乎不該相信這種不合格的平均值陳述。

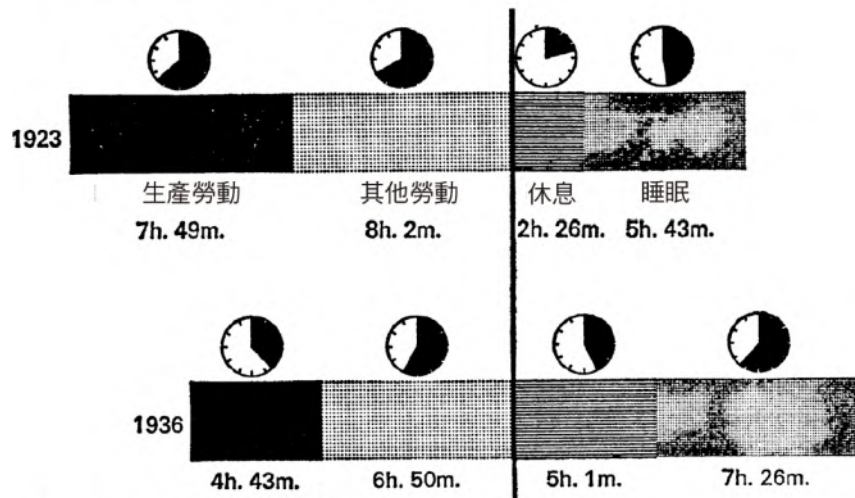
如果沒辦法得到一些讓人印象深刻的統計數據，
你們就會成為統計本月失業率表格的其中一部分。



手法3 小數點，假精確錯覺

如果你想要創造出一種虛假的精準氛圍，藉此為各種糟糕的統計數字提供可信度的話，請考慮使用小數。舉例來說，你調查**100**位市民他們昨晚睡了多少小時，假設得出了**783.1**的總時數。只要是此類的資料，都遠非精確數字。許多人的猜測會與真實睡眠時間相差**15**分鐘或更長的時間，而我們也不能保證這些錯誤會互相平衡。你一定也認識一種人，總是會把尚未睡著的五分鐘記成整整半個晚上翻來覆去的失眠。但是，請繼續計算你的數字，宣布人們一晚的平均睡眠時間是**7.831**小時。這會讓人覺得你聽起來很清楚自己在說什麼。如果你愚蠢到宣布人們每晚睡**7.8**個小時（或「將近八小時」）的話，那聽起來就沒那麼厲害了。**7.8**這個數字聽起來像是真正的事實：一個貧乏的近似值，和隨便的個人猜測所提供的知識量是一樣的。

一名農婦的勞動與休息



圖表改編自蘇聯（圖表統計科學出版研究所〔Scientific Publishing Institute of Pictorial Statistics〕）

卡爾·馬克思（Karl Marx）也以同樣的方式創造出虛假的精準感。在計算工廠的「剩餘價值率」時，他首先大量蒐集了假設、猜測和取整數的數字：「我們假設廢棄物是**6%**.....原料的成本取整數是**342**英鎊。假設每個轉軸**1**英鎊，共有**1**萬個轉軸.....的成本.....我們估計耗損是**10%**.....並假設建築物的租金是**300**英鎊.....」他說：「上述可靠數據是由曼徹斯特紡紗廠提供。」

馬克思根據這些近似值計算出：「因此剩餘價值率是 $\frac{80}{52}$ ，也就是 $153\frac{11}{13}\%$ 。」以一天工作時間十小時來算：「必要勞動是 $3\frac{31}{33}$ 小時，剩餘勞動時間是 $6\frac{2}{33}$ 小時。」

雖然以**33**為分母的數字聽起來有種很準確的絕佳感覺，但這只是虛張聲勢而已。

百分比為混亂的感知提供了肥沃的土壤。正如令人印象深刻的小數一樣，百分比可以為不精確的事物增添精確的光環。美國勞動部（United States Department of Labor）的《每月勞工評論》（Monthly Labor Review）曾指出，在華盛頓特區於特定月份的就業

人口中，提供通勤補助的兼職工作裡面，有**4.9%**的薪資是每週**18**美元。但深究後我們會發現，這個**4.9%**其實是兩份兼職工作，而當時總共只有**41**個職位符合設定的條件。所有以少數案例為基礎計算出來的百分比數字，都可能具有誤導性。直接給出原本的數字其實能提供更多資訊。若你把百分比加上了小數，那麼你的行為就已經從愚蠢逐漸變成欺騙了。

手法4 讓人混淆的基數

你看到一則廣告上面寫著：「現在購買聖誕禮物，省下**100%**的金額。」這聽起來簡直就像聖誕老人開的免費商店，但是，事實上**這只是一種以基數（base）混淆人視聽的方法**。真正減少的金額只有**50%**。你省下的是降價金額的**100%**或新價格的**100%**，雖然這個說法沒有錯，但廣告並沒有寫清楚這件事。

同樣的道理，當某個花農協會的主席在接受報紙採訪時說「鮮花比四個月前還要便宜**100%**」的意思，並不是現在正免費贈送鮮花，但他的遣詞用字聽起來就是如此。

艾達·M·塔貝爾（**Ida M. Tarbell**）在她的著作《標準石油公司史》（*History of the Standard Oil Company*）中則更進一步。她說：「西南地區的降價.....範圍是**14%**至**220%**。」聽起來像是賣家要支付買家一筆可觀的費用，請買家把石油帶走。

《哥倫布電訊報》（*The Columbus Dispatch*）宣稱，某個製成品的銷售利潤是**3800%**，此數字是基於**1.75**美元的成本和**40**美元的售價。在計算利潤百分比時，你可以自行選擇計算方法（而且你有義務指出你使用的是哪一種）。如果按照成本計算，此製成品的利潤是**2185%**；按售價計算則是**95.6%**。《哥倫布電訊報》使用的顯然是自己的一套方法，接著如同我們常常見到的那樣，在報導中提出了一個

過度誇張的數字。

就連《紐約時報》(*New York Times*) 都在刊登美聯社 (*Associated Press*) 一篇來自印第安納波利斯的報導時，輸掉了這場「改變基數之戰」：

不景氣在今天遭受了沉重打擊。水管工、灰泥工、木匠、油漆工和其他隸屬於印第安納波利斯建築工會 (*Indianapolis Building Trades Unions*) 的工人獲得了**5%**的薪資增長，拿回去年冬天減少的**20%**薪資的四分之一。

這樣的計算從表面上看起來很合理。但是減薪時計算百分比的基礎數字是他們原本的薪水；而加薪時計算百分比的基礎數字卻是另一個較小的數字——也就是減薪之後的薪水。

我們可以利用假設來檢視這個統計上的數字錯誤，為了簡單起見，我們假設原始薪資是每小時**1**美元。減去**20%**後，薪資下降到**80**美分。接下來若增加**5%**的薪資，增加的將會是**4**美分，這個金額並不是減薪金額的四分之一，而是五分之一。正如許多我們可以合理推測為無心之過的錯誤一樣，這個錯誤也帶來了誇大的結果，創造出一個更好聽的故事。

上述計算過程能让你了解，當我們想要抵銷**50%**的減薪時，為什麼需要**100%**的加薪。

《泰晤士報》(*The Times*) 也曾在報導中指出，航空郵件在一個會計年度內「因火災而損失了**4,863**磅的郵件，也就是僅僅**0.00063%**。」該報導表示，那年飛機運送的郵件重量是**7,715,741**磅。用這個比例來計算的保險公司很可能會陷入大麻煩。若你認真計算的話，你會發現事實上百分比應該是**0.063%**，遠比報紙所說的高出**100**倍。

改變基數所帶來的錯覺還能為我們解釋疊加折扣的詭計。五金批

發商提供「**50%**加上**20%**的折扣」時，代表的並不是**70%**的折扣。真正的折扣比例是**60%**，這是因為計算**20%**折扣時的基數，是已經折扣了**50%**之後的金額。

常會有人把那些不該加起來、但看起來應該加起來的東西加在一起，許多粗心錯誤和奸巧詭計都來自於此。人們一直在用這種方法來證明過去數個世代的孩子都沒有好好上學。

你可能還記得這件事。首先，在一年的**365**天中，你可以減去三分之一的睡覺時間，也就是**122**天，再減去每天三小時的吃飯時間，也就是**45**天。剩下的**198**天中，還要減去**90**天的暑假和**21**天的聖誕節與復活節假期。剩下的日子甚至連分配給週六和週日都不夠。

你可能會覺得這個伎倆太古老又太容易被拆穿了，不可能會真的用在商業中。但聯合汽車工會在雜誌《彈藥》中堅稱，這種方法仍被拿來用在他們身上：

那種巨大的藍色謊言（**blue lie**，指為保護某個組織或團體聲譽而說的謊）也會在每次罷工時出現。只要出現罷工，商會都會大力宣傳這些罷工每天會造成數百萬美元的損失。

他們在計算這個數字時，會把所有罷工者在同一天全職工作能製造的汽車加起來。他們也用同樣的方式計算出供應商的損失。一切能加進去的事物通通都會被加進去了，其中也包括電車票和店家銷售損失。

此外，也有些人會把百分比當作蘋果一樣隨意相加，用類似的怪異方式來提出反對意見。請看看下面這段來自《紐約時報書評》（*New York Times Book Review*）的論述聽起來多麼有說服力：

書價上漲與作家收入之間之所以會出現差距，似乎是因為製作成本與原料成本出現大幅上升。光是工廠與製造的開支在過去十年中就上漲了**10%**至**12%**，原料上升了**6%**至**9%**，銷售與廣告支出則攀升了

10%。把這些價格上漲加總起來，上升幅度至少有**33%**（對某些公司而言），對部分較小的出版社來說則可能高達將近**40%**。

事實上，如果出版這本書的各種支出項目都上漲了大約**10%**的話，那麼這本書的總支出成本上漲幅度也應該落在**10%**左右。若我們認同這種把所有百分比加在一起的邏輯，最後將可能會製造出各式各樣的幻想。比如說，你今天買了二十樣商品，發現每件商品的價格都比去年上漲了**5%**。把這些百分比「加起來」就是**100%**，於是你的日常支出變成了兩倍——這根本是胡說八道。

這有點像是你可能已經聽過的路邊攤販故事：有位消費者請攤販解釋，為什麼他的兔肉三明治這麼便宜。「這個嘛，」攤販說：「我必須加一些馬肉在裡面。但我用各**50%**的比例混合這兩種肉的：一匹馬搭配一隻兔子。」

曾有一本工會刊物刊登了一張卡通圖，來反對這種毫無根據的統計加法。圖片上的老闆將一般工時的**1.5**美元時薪，加上加班的**2.25**美元時薪，再加上國定假日雙倍時薪的**3**美元時薪，最後得出平均時薪是**2.25**美元。你很難找到任何案例中的平均值比這個數字更沒有意義了。

手法5 「百分比」與「百分點」

另一片適合使人受到愚弄的肥沃土壤，是我們很容易混淆百分比與百分點。如果你的投資年利潤從**3%**上升到**6%**，那麼你可以用「上升三個百分點」讓這個上升聽起來比較溫和。在同樣的效度下，你也可以把這個改變稱作百分之百的增長。若想看到各種文章如何以散漫的方式處理百分比與百分點這兩個令人困惑的數字，你可以特別留意民意調查的相關文章。

百分位數也具有欺騙性。當你聽說強尼在代數課或其他方面的才能比其他同學更好時，用來描述的數字就有可能是百分位數。百分位數代表了他在每一百名學生中的排名。舉例來說，在一個**300**人的班級中，前三名學生的百分位數是**99**，接下來三名學生的百分位數是**98**，以此類推。百分位數的奇怪之處在於，當學生處於百分位數**99**的時候，他的能力可能比百分位數**90**的學生要好得多，而百分位數**40**與**60**的學生的能力則幾乎相等。這是因為許多特質都傾向於在平均值的區間聚集，形成我們在前面章節提到的「正常」鐘形曲線。

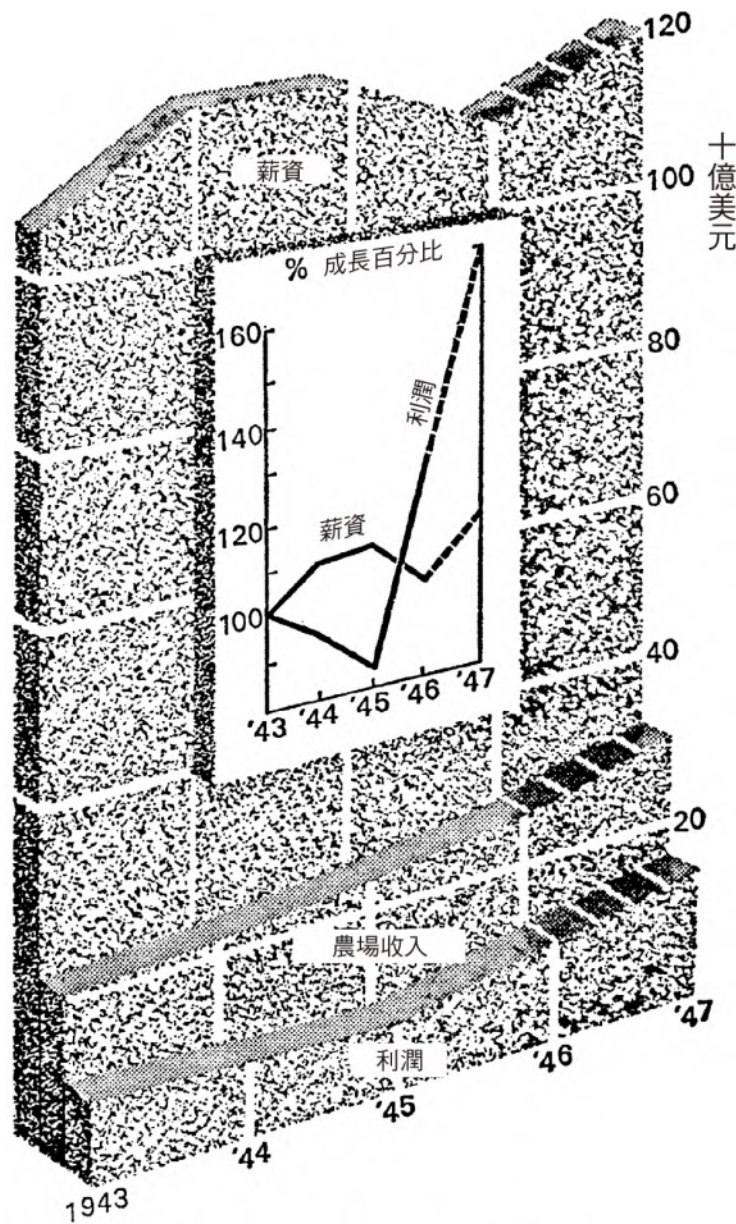
手法6 選擇最有利的數據

統計學家偶爾會互相開戰，就連觀察力最不入微的觀察者也會注意到其中必定有一方是卑鄙小人。在統計人員開始彼此爭執時，老實人就可以休息片刻。

鋼鐵工業委員會（**Steel Industry Board**）指出，鋼鐵公司和工會都沉迷於騙人的伎倆。為了展現**1948**年的鋼鐵業發展得有多好（藉此證明各間鋼鐵公司負擔得起加薪），鋼鐵工會比較了該年與**1939**年的生產率——但**1939**年是生產率特別低的一年。鋼鐵公司在這場欺騙大賽中也不甘示弱，他們堅持應該要拿員工收到的錢來比較，而不是用平均時薪做比較。這麼做的關鍵之處，在於前一年有太多員工都是兼職身分，所以即使薪資率沒有上漲，員工的收入也必定會增加。

《時代》雜誌最著名的特色之一，就是卓越的製圖，該雜誌曾刊登的一張圖表能當作有趣的案例，讓我們知道計算者在使用統計數據時，幾乎有能力以各種妙計達到他們想要的任何目標。在《時代》雜誌面前的是兩個效度相同的方法，一種是對管理者較有利的觀點，另一種則對勞工較有利，而《時代》雜誌決定雙管齊下。這張圖表實際上是兩張圖表，其中一張疊加在另一張上。兩張圖表使用的是相同的

數據。



經《時代》雜誌許可重新繪製，此圖是非說謊圖表的案例。

其中一張圖表用「十億美元」為單位顯示薪資和利潤。儘管數據都在上漲，但去年農場收入的增幅大約是利潤的兩倍，薪資則大約是利潤的六倍。如此看來，巨大的通膨壓力似乎來自薪資。

這張雙圖表的另一部分則以「成長百分比」為單位顯示變化。薪

資的線條相對平穩，利潤的線條則出現急遽上升。我們可以由此推斷，通膨的主要原因應該是利潤。

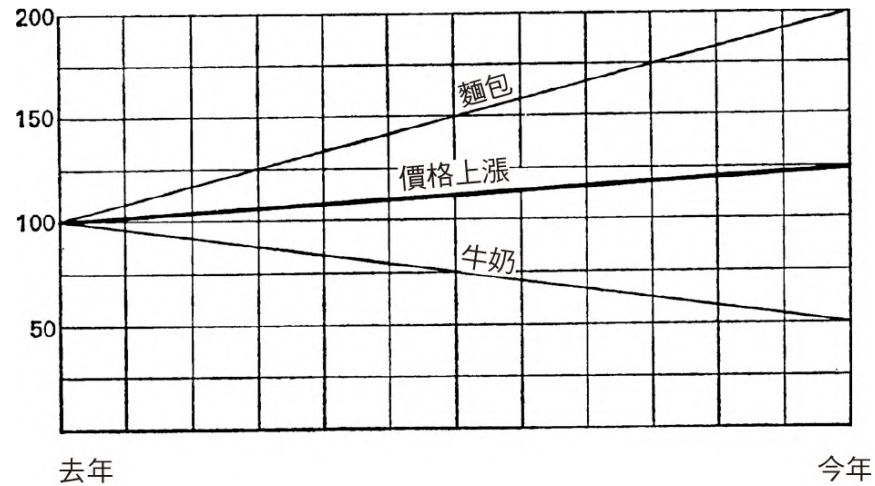
你可以任意選擇其中一個結論。又或者你可以做出更好的決定，輕而易舉的指出這兩個元素都不應該被單獨視為導致通膨的原因。有時候，光是指出「這個主題有爭議」就能帶來重大的影響。

手法7 比較基準的設定差異

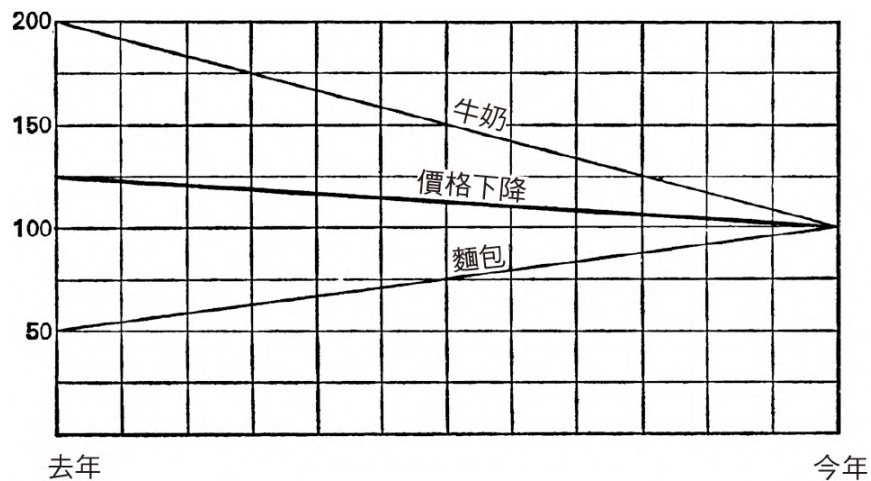
「指數」是一種對數百萬人來說至關重要的數字，這是因為薪資率往往和指數有關。或許值得我們注意的是，我們該怎麼做才能使指數隨著任何人的音樂起舞。

舉一個最簡單的例子，假設去年牛奶的價格是每夸脫（約1公升）10便士（0.1英鎊），麵包是每條10便士。今年牛奶的價格下降了5便士，麵包則上漲到了20便士。接著你要決定的是，你想證明什麼事情？生活成本上升嗎？還是生活成本下降？又或者生活成本沒有變化？

我們先以去年為基數（單位先忽略不計），把當時的價格當作100。由於牛奶的價格下降了一半（50），麵包的價格變成兩倍（200），而50與200的平均值是125，所以價格上漲了25%。



接著我們再試一次，這次以今年為基數。牛奶在一年前的價格是現在的兩倍（200），麵包的售價是現在的一半（50）。計算過後雖然平均值：125，但現在的價格比過去少了25%。



若想證明成本水準沒有任何變化，我們只要切換到幾何平均數就可以了，使用任一年當作基數都可以。幾何平均數和我們平常使用的算數平均數略有不同，不過幾何平均數仍是具有正當性的數字，在某些情況下，甚至是最有效也最能揭露事實的。在計算三個數字的幾何平均值時，你要做的是把它們彼此相乘，並取立方根。遇到四個數字

時，則計算出四次方根；遇到兩個數字時，則計算出平方根，以此類推。

舉例來說，我們以去年為基數，把去年的價格稱作**100**。也就是說，當你把這兩個商品價格視為**100**，相乘並取平方根後，得到的數字就是**100**。今年的牛奶價格是去年的**50%**，麵包是去年的**200%**，所以**50**乘以**200**得到的數字是**10000**。取平方根後你就會得到幾何平均值，也就是**100**。價格既沒有上漲也沒有下跌。

事實上，雖然統計學具有數學基礎，但統計學既是科學，也是藝術。統計學能夠在容許範圍內進行許多操控、甚至扭曲。統計學家往往必須在眾多方法之間選擇（這是一個主觀的過程），並找到他要用以表示事實的那一個方法。依照商業慣例，統計學家不太可能選擇對雇主不利的統計方法，就像是文案撰稿人若能夠把雇主的商品形容成輕巧又經濟實惠的話，它們就不會說商品脆弱又廉價。

即使是從事學術工作的人，也可能會有比較偏好的偏誤（可能是無意識的）、想要證明的觀點、想要達成的目標。

這代表在照單全收前，我們應該要先敏銳的進一步檢視報紙、書籍、雜誌與廣告提供的統計資料、實情描述和數字，謹慎的眯起雙眼會使焦點變得更加清晰。然而，武斷的將這些數字視為欺騙，拒絕所有統計資料也是毫無道理的作法。這就像是擔憂作家有時會用文字隱藏而非揭露事實，就因此拒絕閱讀任何文字一樣。

不過，我們還是得小心，畢竟佛羅里達州的一位候選人，曾因指控競爭對手在「實踐獨身主義」而贏得了大量選票^{*2}。電影《暴君焚城錄》（*Quo Vadis*）的放映商也用巨大的字體引用了《紐約時報》的評論，說這部電影是「史上最驚人的狂妄之作」。專利藥物「瘋狂水晶」（*Crazy Water Crystals*）的製造商更是一直宣傳他們的產品能「快速且暫時的緩解不適」。

據說作家路易斯·布羅姆菲爾德 (**Louis Bromfield**) 會在讀者來函多到難以一一回覆時，用同一封信回覆所有批評信函。這封信既沒有承認任何事，也沒有鼓勵來函者回覆信，但又幾乎同時滿足了所有人。其中的關鍵金句是：「你所說的或許蘊含了一些道理。」這讓我想起了一位牧師，他常會大加稱讚那些由他受洗的嬰兒，所以在會眾中很受母親的歡迎。但是，這些母親們在聊到這件事時，卻沒有任何人記得牧師說了什麼，她們只記得他說的是「好話」。事實上，他每次說的話都一模一樣：「哇喔！（微笑）這真是個小嬰兒呢，不是嗎！」

編按：1950年參議員初選期間，佛州候選人斯馬瑟斯 (**George Smathers**) 用耳語和謠傳打擊對手佩柏 (**Claude Pepper**) 。

第十章

五大招，破解陷阱

世界上充斥著真相被操弄、利用的謊言數據，面對統計操控的無所不在，若還只是天真的相信各種機構組織所發布的數字，就等著被騙吧！本書前述已說明了統計的瑕疵和刻意性，在本章，將提供一套行動綱領，利用五個提問讓破統計的漏見與詭計，學會正確解讀，避免落入陷阱：

- ❶ 說話者「是誰」？
- ❷ 他「如何得知」？
- ❸ 是否有所「缺漏」？
- ❹ 「主題」改變了嗎？
- ❺ 論述「合理」與否？



截至目前為止，我對待你的態度，都好像你是一名充滿熱忱的海盜，需要有人更仔細的指導你如何使用彎刀。在這一章，也就是最後一章中，我將會丟掉這種「教學」說法。我會帶你正視隱含在本書之中的重要目標：說明我們該如何直面假的統計數據。還有，同樣重要的，如何在前幾章討論到的詐欺荒原中，找出哪些數據是可靠與可用的。

在你看到的統計數字中，並不是所有數字都可以透過化學分析或化驗師在實驗室做的實驗來證實其確定性。但是，你可以用五個簡單的問題來測試這些數字，藉由這五個問題的答案，避免將非事實的知識視為真相。

問題1 說話者「是誰」？

你首先要做的事是尋找偏誤——為了證明某個理論、為了聲譽或經費而需要證明某件事的實驗室；以精彩故事為目標的報紙；薪資水準位於風險中的勞工或管理者。

請尋找有意識的偏誤。這種偏誤可能是直接的錯誤論述，或者是確實可以使用但無法令人信服的模糊論述；也可能是選擇有利數據，同時隱藏不利數據。計算者有可能會改變計量單位，或進行比較時雖然都以年為單位，但卻跨越很長的時間選擇了對他們較有利的年份。他們可能還會使用不恰當的衡量方法：在中位數能提供的資訊量比較大時（甚至有可能太大），卻選擇使用「平均值」這個未經認證又能隱藏許多詭計的詞語。

請認真尋找無意識的偏誤，因為這種偏誤往往更加危險。**1928**年，在許多統計學家與經濟學家製作的圖表和提出的預測中，無意識偏誤帶來了許多驚人的結果。這些偏誤愉快的遺漏了經濟結構的裂紋，引用各種證據，以統計方法展示我們如今只不過是處於繁榮發展

的趨勢之中。

你可能必須要進一步檢視，才能找出「誰說的」。這個「誰」可能會隱藏在《高談虛論》（*Lifemanship*）的作者史蒂芬·波特（Stephen Potter）所說的「好聽的名字」（O.K. name）中。

任何帶有醫學專業色彩的事物都是好聽的名字。科學實驗室也擁有好聽的名字。同樣擁有好聽名字的還包括各個學院，尤其是大學，更甚至是那些在技術研究表現突出的大學。

我們在前幾章曾看到一位作者提出「較高等的教育會危害女性的結婚機會」，這位作者就擅用了康乃爾大學的好聽名字。請特別留意，雖然該文章的數據來自康乃爾大學，但結論是作者自己提出來的。好聽的名字會使你產生「康乃爾大學指出……」這樣的錯誤印象。

若發現文章引用了好聽的名字，請確保名字是資訊的後盾，而不是只能寫在資訊旁邊的名稱。

你可能已經讀過芝加哥《商務日報》（*Journal of Commerce*）引以為豪的公告了。該刊物進行了一次調查，主題是哄抬價格和囤積商品，共有**169**間公司回覆，其中有三分之二的公司宣稱他們正在吸收韓戰造成的價格上漲。

「調查顯示（每次看到這些詞，請提高警惕！），」《商務日報》寫道：「這些公司做的事情，和美國商業系統的敵人所提出的指控完全相反。」有鑑於《商務日報》有可能是這項調查的利害關係人，所以我們顯然該在此提問：「誰說的？」這同時也是詢問我們應該要提出的第二個測試問題的好地方。

問題2 他「如何得知」？

事實證明，《商務日報》首先向**1200**間大型公司寄出了調查問

卷。只有**14%**的公司回覆。**86%**的公司不想在公開場合談論他們是否正在囤積商品或哄抬價格。

雖然《商務日報》把這件事粉飾得非常好，但事實仍舊沒有改變，根本沒有任何事情值得吹噓。簡而言之：在接受調查的**1200**間公司中，**9%**表示沒有提高價格，**5%**表示有，**86%**不願意說。這些回覆調查的公司所構成的樣本，或許具有偏誤的可能。

請特別留意把偏誤樣本當作證據的敘述，偏誤樣本包括不當樣本（就像上述樣本一樣）或自行選擇的樣本等。請提出我們在前面章節討論過的問題：樣本數是否足夠大，大到足以得出可靠的結論？

同樣的，這種狀況也和相關性類似：樣本數是否大到足夠具有意義？是否有足夠案例能增加任何顯著水準？身為一名一般讀者，你沒辦法應用顯著性的檢驗，也無法得出有關樣本充分性的確切結論。然而，在你看到的報導中，有很多事情是你無法一眼看出的（就算你花了很長時間看這一眼也一樣），你不會知道其實報導中並沒有足夠的樣本，能說服任何理性的人去相信任何事物。

問題3 是否有所「缺漏」？

報導不會總是告知你他們有多少樣本數。缺少這種數字足以讓我們對整件事產生質疑，特別是當消息來源是利害關係者的時候。同樣的道理，我們也不需要太過認真看待那些沒有提供可信度（可能誤差、標準誤差）的相關性資料。

請小心沒有清楚表明類別的平均值，尤其是在算數平均數與中位數應該會產生巨大差異的事項中。

許多數字都因為缺乏比較性而失去了意義。《展望》（*Look*）雜誌的一篇文章在論及唐氏症時寫道：「一項研究指出，在**2800**個病例中，超過一半病例的母親年齡在**35**歲以上。」能不能從這段描述中得

到任何意義，取決於你是否了解女性通常會在幾歲生孩子，而鮮少有人知道這一類的事情。

下面是一段摘自《紐約客》在**1953年1月31日**刊登的〈來自倫敦的信件〉（**Letter from London**）的污染報告：

英國衛生部（**Ministry of Health**）近期公布的數字顯示，在大倫敦區（**Greater London**）起濃霧的一週期間內，死亡率迅速上升了**2800%**。此數據震驚了社會大眾，一般公眾習慣將英國的糟糕氣候視為一種困擾，而不是死因.....今年冬天的到訪具有出乎意料的致命特性.....

但是，這次寒冬的到訪到底有多麼致命呢？比平常高出這麼多的單週死亡率是不是個例外？這兩個答案確實會使事情有所不同。那麼接下來的幾週呢？是不是死亡率就下降到平均值以下了？若是的話，那是否代表被大霧殺死的人，其實本來就會在短時間內死去？雖然這個數字聽起來令人印象深刻，但是缺乏其他數字會使它失去絕大部分的意義。

有時候，文章提供的是百分比，不會給原始數字，這種作法可能也具有誤導性。

很多年之前，在美國約翰霍普金斯大學（**Johns Hopkins University**）剛開始招收女學生時，有一位並不喜愛男女同校的人提出了令人無比震驚的數字：約翰霍普斯金斯大學的女性中，有**33.3%**的人和教職員工成婚！而原始數字會展現更清晰的全貌：當時有三名女性入學，其中一名嫁給了教職員工。

幾年前，波士頓商會（**Boston Chamber of Commerce**）選出了他們的「美國成功女性」。其中有**16**人中也進入了《美國名人錄》，商會指出，這些人共有「**60**個學位和**18**名孩子」。聽起來商會確實提供了這群女性的一些資訊，但若你進一步探究，就會發現這些女性中包

括了吉爾德斯利夫校長（Virginia Gildersleeve）和莉莉安·吉爾博斯夫人（Lillian M. Gilbreth）。60個學位中有三分之一屬於這兩位女性。想當然爾，18名孩子中有三分之二都是吉爾德斯夫人家的孩子。

任何一間公司都可以宣布共有3003人擁有他們的股票，平均每人持有660股。這項資訊確實是真的。另一件同樣真實的事情是，在公司的200萬股中，有四分之三是由三個人持有的，剩下的四分之一是由3000人持有的。

看到指數時，你可以提出的問題是：「其中少了什麼事物？」缺少的可能是基數，而計算者選擇的基數有可能會扭曲事實的樣貌。

美國曾有一個勞工組織指出，在大蕭條之後，利潤和生產力的指數上升速度都遠遠超過了薪資的指數。該組織用這個論述當作加薪的論據，但接著有人挖掘出了這段描述中缺少的數字，使得該論據失去了效力。我們可以從缺少的數字看出來，利潤的成長百分比幾乎必定會比薪資更高，原因很簡單，因為當時利潤達到了低點，所以利潤的基數較小。

有時候，論述中缺少的是推動改變的因素。這種缺漏會留下一種暗示，讓人覺得改變是人們比較期望的其他因素所造成。例如，我們可以公布某一年的數據，指出四月份的銷售額高於去年四月，藉此表明生意愈來愈好。但這裡缺漏的事實是：去年的復活節在三月，今年的復活節則在四月。

如果有報導指出，過去四分之一個世紀的癌症死亡人數出現大幅增加，那麼這就是一則誤導性的報導，除非報導也指出有多少死亡人數來自以下外部因素：過去罹癌死亡的人被歸入「死亡原因不明」；如今驗屍更加頻繁，使診斷更加準確；今日醫療統計的回報與彙整更加完善；現在人們更長壽，活到容易受癌症影響的年齡。此外，如果你看到的是總死亡人數而非死亡率的話，也不要忽視這個事實：現在的人口比以前更多。

問題4 「主題」改變了嗎？

在分析統計數據時，**請注意原始數字和結論之間的轉換過程**。你會發現人們往往會把某件事講成另一件事。

如前所述，疾病的回報數量增加，並不總是等同於疾病的病例數增加。候選人以毫釐之差勝選的選舉，往往不是民調可以預測得到的。就算雜誌讀者的「橫斷面」所表達的偏好是有關國際事務的文章，也不能當作最終證據，證明雜誌發表了這些文章後，讀者真的會閱讀這些文章。

某一年，加州中央谷回報的腦炎病例數量大幅上升，是過去單年最高病例數的三倍。許多驚慌失措的居民都把孩子送到了其他地方。但是在真正計算時我們會發現，腦炎的死亡人數並沒有大量增加。當時的情況是大量州政府與聯邦政府衛生人員來到此處解決一個長期存在的問題，他們付出努力帶來的結果是記錄了大量輕度病症的案例，而這些案例在其他年份很可能會被忽視，甚至有可能不會被診斷出來。

這不禁讓人想起林肯·史蒂芬斯 (**Lincoln Steffens**) 和雅各布·A·里斯 (**Jacob A. Riis**) 在紐約的報社工作時，是如何掀起了一股犯罪浪潮。當時報紙上的犯罪案件在數量、版面與大型標題這三個面向都達到極高的比例，以至於公眾開始要求政府採取行動。

當時狄奧多·羅斯福 (**Theodore Roosevelt**) 是警察改革委員會 (**Reform Police Board**) 的主席，他感到極為難堪。但他只做了一件事，就輕易結束了這波犯罪浪潮：要求史蒂芬斯和里斯停止報導。事實上，這波浪潮之所以會出現，只是因為以那兩人為首的記者們，彼此競爭誰能挖到最多犯罪案件的新聞。根據警方的官方紀錄統計，案件數量根本沒有增加。

報紙指出：「五歲以上的英國男性每週泡熱水澡的平均次數是冬

天**1.7**次，夏天**2.1**次。」「英國女性每週泡熱水澡的平均次數是冬天**1.5**次，夏天**2.0**次。」該數據的資料來源是英國工程部（**Ministry of Works**）對「**6000**個具有代表性的英國家庭」進行的熱水調查。報導指出此樣本具有代表性，而且數量似乎足以證明《舊金山紀事報》（*San Francisco Chronicle's*）逗趣標題中的結論是合理的：〈英國男性的泡澡次數高於女性〉（**British He's Bathe More Than She's**）。

如果報導能指出這個數字是算數平均數還是中位數的話，該數字將能提供更多資訊。不過，這個數字的最主要問題在於主題的轉變。

工程部真正調查出來的，是人們有多常「說」他們在泡澡，而不是實際上有多常在泡澡。當調查主題是洗澡這一類比較私密的事時，人們嘴上說的與實際做的很可能根本就是完全不同的兩回事，該調查甚至還涉及了英國的泡澡傳統。英國男性的泡澡次數或許比女性更頻繁，也或許更不頻繁；但我們唯一能做出的可靠結論是：他們口頭上是這樣說的。

以下列出更多需要留意的主題轉變。

一次人口普查發現，美國農場比五年前增加了**50**萬座，由此可知美國出現了一波回歸農場的浪潮。但這兩個數字指的其實並不是同一件事。這是因為人口普查局對農場的定義已經改變了；因此被新列入農場的地點至少有**30**萬個，是依據先前的定義時不會被列為農場。

當我們把人們提出的數字當作基數時，就會製造出一些奇怪的現象——就連看似客觀的事實也一樣。舉例來說，許多人口普查的報告都顯示，**35**歲的人口比**34**歲與**36**歲的人更多。之所以會有這種錯誤的數字，是因為回報全家人年齡的，往往是家庭中的其中一個成員，當這位成員不確定其他人的確切年齡時，就傾向於把數字四捨五入至較常見的**5**的倍數。解決這個問題的其中一種方法是：改為詢問出生日期。

中國其中一個大面積區域的「人口」是**2800**萬；五年後這個數字

達到**1.05**億。這種人口數據的成長鮮少會是真的；若要解釋這種巨大的差異，我們必須考慮到這兩次人口普查的目的，以及人們在遇到這兩次普查時，對於「被計算入人口中」比較傾向於出現何種感覺。第一次人口普查是為了收稅和軍事目的；第二次則是為了救濟饑荒。

美國也發生過類似的事。十年一次的人口普查發現，落在**65至70**歲這個年齡段的人比**55至60**歲的更多。我們無法用移民來解釋這種差異。較大的原因可能是人們為了領取社會保險金而大規模謊報年齡。也有另一個可能性是，早期有些人出於虛榮心而低報了年齡。

另一個轉變主題的代表案例是威廉·藍格參議員（**William Langer**）的呼籲：「把囚犯的住處從惡魔島（**Alcatraz**）改成華爾道夫酒店（**Waldorf-Astoria**）的話，費用反而會更便宜。」這位北達科他州人指的是他更早之前的論述：惡魔島收容一名犯人的日均花費是**8**美元，等同於「在舊金山的飯店住一晚的價格」。他把主題從總營運成本（惡魔島）轉變成單純的飯店租金了。

另一種轉變主題的隱晦方式，是各種事後歸因又自以為是的胡說八道。這種胡說八道在**A**事件和**B**事件一起改變的狀況下，把這件事描述成**A**事件因為**B**事件而改變。

《電氣世界》（*Electrical World*）雜誌曾在一篇社論〈電力對美國意味著什麼〉（*What Electricity Means to America*）中提供了一張複合圖表。在圖表中看到，隨著「工廠電力」的上升，「每小時平均薪資」也跟著上升。與此同時，「每週平均工時」則出現下降。這些數字當然都是長期趨勢，且沒有任何證據能表明其中一個趨勢導致了另一個趨勢。

接著，我們還會遇到「第一名」。幾乎任何人都可以聲稱自己在某件事情上是第一名，前提是不要太過講究是哪一件事情。**1952**年底，紐約有兩間報社都堅稱自己在雜貨廣告方面是第一名。某種程度上來說，雙方都對。《世界電報》（*World-Telegram*）解釋說，他們

是第一個全力投放廣告的報社，也就是每一份報紙中都有廣告，這也是他們唯一的一種廣告方式。而《美國日報》（*Journal-American*）則堅稱廣告總行數才是最重要的，而他們在這方面是第一名。正是這種想辦法取得第一名的表達方式，使得收音機中的天氣預報得以把一個普通的日子稱作「**1967**年來最熱的**6**月」。

在你考慮要直接借款或者以分期付款的方式借款時，主題的轉變會使你很難把這兩者的成本拿來做比較。雖然**6%**聽起來就是**6%**，但它很可能根本不是**6%**。

如果你以**6%**的利息從銀行借了**100**英鎊，並花一年的時間每個月等額分期還款，那麼你為這筆借款所支付的費用大約會是**3**英鎊。還有另一種**6%**的貸款，會讓你付上兩倍的錢，這種貸款有時又被稱作花**6**英鎊借**100**英鎊。這就是多數汽車貸款的計算方式，這種計算方法非常微妙。

此處的關鍵是，你借貸**100**英鎊的時間不到一年。在六個月之後，你已經償還了一半的貸款。如果你因為借了**100**英鎊而已支付了**6**英鎊，也就是借貸金額的**6%**，那麼你實際上支付的其實是將近**12%**的利息。

更糟糕的，是許多粗心的冷凍食品計畫買家遇到的事。他們得知的借貸數字從**6%**到**12%**不等。這些數字聽起來像是利息，但事實並非如此。這是一個以美元為單位的數字，最糟糕的是，貸款時間通常是六個月，而不是一年。在半年內因為借**100**英鎊而支付**12**英鎊，計算起來其實相當於**48%**的利息。難怪會有這麼多客戶違約，這麼多食品計畫失敗了。

有些人會利用語意來轉變主題。以下是《商業週刊》（*BusinessWeek*）雜誌的一篇文章：

會計師認為「盈餘」（**surplus**）是一個討厭的字眼。他們會建議公司把這個詞從資產負債表中刪除。美國會計師協會（**American**

Institute of Accountant) 的會計程序委員會 (Committee on Accounting Procedure) 指出：會計用「保留盈餘」 (retained earnings) 或「固定資產之增值」 (appreciation of fixed assets) 來取代此類描述字詞。

以下段落來自報紙的一篇報導，主題是標準石油公司 (Standard Oil) 的破紀錄收入和單日100萬美元的淨利：

公司董事可能正在考慮分割股票，這麼做可能會帶來優勢.....如果每股利潤看起來沒那麼高的話.....



(圖內注釋：本丟·彼拉多^{*1})

問題5 論述「合理」與否？

當一段雜亂論述的基礎是未經證實的假設時，提出「合理嗎？」這個問題往往能降低統計數據的可信度。

你可能曾聽說過魯道夫·佛萊士 (**Rudolf Flesch**) 的可讀性 (**readability**) 公式。這個公式利用單字與句子的長度等簡單且客觀的項目，來測量一篇文章的閱讀難易度。這個想法非常吸引人，正如各種能夠把不可估量的事物簡化成一個數字，並用算數替代判斷的其他方法一樣。雖然許多作家對此並不感興趣，但至少這個想法吸引了那些僱用作家的人，例如報社。該公式的基本假設是，字詞長度等因素決定了可讀性。若要挑毛病的話，我們甚至可以說連這個假設都有待證明。

羅伯特·A·杜福爾 (**Robert A. Dufour**) 把佛萊士公式用在幾篇他覺得方便取得的文章作品上。該公式顯示〈沉睡谷傳奇〉 (**The Legend of Sleepy Hollow**) 的閱讀難度是柏拉圖《理想國》 (*Republic*) 的1.5倍。辛克萊·路易斯 (**Sinclair Lewis**) 的小說《海棠春怨》 (*Cass Timberlane*) 的閱讀難度則高於雅克·馬里坦 (**Jacques Maritain**) 的文章〈藝術的精神價值〉 (**The Spiritual Value of Art**) 。

許多統計數字從表面上看來就是錯的。這些數字之所以會被放上檯面，是因為數字的魔力使人們的常識暫停運作。倫納德·恩格爾 (**Leonard Engel**) 在《哈潑》的一篇文章中列出了一些醫學方面的例子：

一位著名的泌尿科醫師，他計算出美國的前列腺癌病例有**800萬**件——這個數量足以讓落在「易罹患前列腺癌年齡層」的所有男性都罹患**1.1次**前列腺癌！

著名的神經學家估算出，每**12**個美國人中就有一個罹患偏頭痛；由於共有三分之一慢性頭痛源自於偏頭痛，這也就代表了全美國共有四分之一的人口必定罹患了慢性頭痛這個嚴重疾病。

許多文章在提到多發性硬化症時，常提到罹患該病的人數是**25萬**；幸好根據死亡的數據顯示，美國罹患這種麻痺性疾病的病患不超

過3萬至4萬人。

《社會安全法案》（**Social Security Act**）的修正案聽證會，一直因為其他團體用各種方式表達的同一種論點而不堪其擾，不過只要你仔細檢視，就會發現該論點其實不合理。該論點是這樣的：由於美國當時的預期壽命只有大約**63**歲，因此制定退休年齡為**65**歲的社會安全計畫是騙局和詐欺，畢竟幾乎每個人都會在**65**歲之前死亡。

只要環顧四周，觀察你認識的人，就能反駁這個論點。然而這個論點最基本謬誤在於，這個數字指的是出生時的預期，也就是說，我們可以預期大約有一半的嬰兒在出生後的壽命會超過這個數字。

順帶一提，這個數字來自**1939**至**1941**年，只在這段期間是正確的，並且在過時很長時間後仍在使用。最新的預估已經把壽命更正到**69.7**歲，但這個數字仍會製造出一個同樣愚蠢的全新論點：現在幾乎每個人都會活到**65**歲（編按：目前全球人類的平均壽命已達**73**歲）。

前幾年，有一間大型電器公司擬訂了戰後的銷售計畫，該計畫以出生率下降為基礎假設，因為長期以來人們都把出生率下降視為理所當然的事。該計畫強調的是小容量電器，例如，適合公寓大小的冰箱。接著其中一位策劃者的常識突然恢復了：他脫離圖表的時間足夠長，讓他注意到在他、他的同事、他的朋友、他的鄰居與他的老同學之間，多數人都已經生了三、四個孩子或計畫要生三、四個孩子，只有極少數人例外。公司因此展開了更開明的調查和製圖——很快就將盈利重點轉向了大家庭模式。希望以後的規劃者能夠更快的對當前的轉變做出反應。

另一件與常識互相矛盾的，是令人印象深刻的精確數字。根據紐約市數份報紙報導的一項研究表明，與家人住在一起的職業女性每週需要賺進**40.13**美元的薪水才足夠。任何在閱讀報紙時保持邏輯運作的人都會注意到，維持生活所需的金額不可能計算到美分那麼詳細。但

這是一種極大的誘惑；「**40.13美元**」聽起來比「大約**40美元**」還要更詳盡。

你有權利以同樣懷疑的態度看待美國石油業委員會（**American Petroleum Industries Committee**）提出的報告，該報告指出汽車的平均年稅金是**51.13美元**。

外推法（**extrapolation**）是一種很有效的方法，尤其是在我們稱作「趨勢預測」的預言占卜中。但是在觀察外推法產生的數字與圖表時，我們必須時時記住一件事：過去的趨勢或許是真的，但未來的趨勢只不過是一種有根據的推測罷了。

此推測中隱含的預設狀況是「所有條件都相同」和「當前趨勢仍會持續下去」。但事實上，條件不可能會保持相同，否則這樣的生活就太過無趣了。

若想了解不受控制的外推法之中，隱含著怎麼樣的胡說八道，我們可以觀察電視數量的趨勢。

自統計的五年間，擁有電視的美國家庭增加了**10000%**。依照此增加速度預估接下來的五年，你會發現電視的數量很快就會超過數百億臺，也就是每個家庭都擁有數量驚人的**40臺**電視。如果你想顯得更愚蠢的話，你可以把計算電視趨勢的基準年分再往前推，如此一來，你將會「證明」每個家庭在不久後擁有的不再是**40臺**電視，而是高達**4萬臺**。

哈里·杜魯門（**Harry Truman**，民主黨）在**1948**年美國總統選舉大勝湯姆·杜威（**Tom Dewey**，共和黨），而在選舉之前，各種民意調查都顯示杜威將輕鬆戰勝杜魯門。選舉過後，政府研究人員莫里斯·漢森（**Morris Hansen**）稱蓋洛普民意調查（**Gallup**）的選舉預測是「人類史上最廣為人知的統計錯誤」。

不過相較於我們如今最廣泛使用在未來人口上的統計方法（這個方法能惹得全國人民哄堂大笑），蓋洛普的預測方法其實還可說是準

確性的典範。

一直到1938年，由專家組成的總統委員會仍在懷疑，美國的人口是不是真的會達到1.4億；接著在短短12年後，人口數量就增加了1200萬。有一些近期出版、仍有大學在使用的教科書預測，人口的最高峰不會超過1.5億，而且要到1980年左右才會達到。這些可怕的低估之所以會出現，是因為人們都假設趨勢會持續下去，不會變化。

一個世紀前，有一個類似的假設在相反方面表現得同樣糟糕，該預設認為1790年至1860年的人口增長率會延續下去。在亞伯拉罕·林肯（Abraham Lincoln）對國會的第二次演講中，他預測美國人口將會在1930年達到251,689,914人。

不久之後，馬克·吐溫在1874年於《密西西比河上的生活》（*Life on the Mississippi*）中總結了外推法的荒謬之處：

在過去176年的時間中，密西西比河下游（Lower Mississippi）的長度縮短了242英里（389公里）。也就是平均每年只縮短 $1\frac{1}{3}$ （2公里）英里這短短的距離。因此，只要是頭腦冷靜的人，無論他是不是瞎子或笨蛋，都應該能預測到在100萬年前的隔年十一月，也就是古老的鮎粒志留紀（Old Oölitic Silurian Period），密西西比河下游的長度應該長達130萬英里（209萬公里），宛如一根釣魚竿一樣戳進墨西哥灣中。以此類推，任何人都知道，在742年後，密西西比河下游只會剩下 $1\frac{3}{4}$ 英里（2.8公里）長，開羅市和紐奧良市的街道將會連接在一起，由一位市長與市議委員會的治理下安樂的蹣跚前行。科學總是令人著迷。人們可以只要投資如此微不足道的事實，就能獲得如此完整的猜測做為回報。

Pontius Pilatus，判耶穌釘十字架的羅馬總督。

致謝

本書中充滿了各種笨拙又運用了花招的小範例，這些例子是在他人的幫助下蒐集而來。在我向美國統計協會（**American Statistical Association**）提出申請後，許多專業統計學者相信我，他們對於誤用統計學的強烈反對之情和所有人一樣真誠，因此把他們的收藏品寄給了我。我猜這些人應該很開心能在此保持匿名。我也在許多書中找到了一些有價值的樣本，主要書籍是：馬丁·A·布倫博（**Martin A. Brumbaugh**）和萊斯特·S·凱洛格（**Lester S. Kellogg**）的《商業統計學》（*Business Statistics*）、哈德利·坎特里爾（**Hadley Cantril**）的《計量輿論》（*Gauging Public Opinion*）、威拉德·柯普·布林頓（**Willard Cope Brinton**）的《圖形演示》（*Graphic Presentation*）、佛德瑞克·E·克羅克斯頓（**Frederick E. Croxton**）和達德利·J·考登（**Dudley J. Cowden**）的《實用商業統計》（*Practical Business Statistics*）、喬治·辛普森（**George Simpson**）和佛利茲·卡夫卡（**Fritz Kafka**）的《基礎統計學》（*Basic Statistics*）以及海倫·M·沃克（**Helen M. Walker**）的《基本統計方法》（*Elementary Statistical Methods*）。

統計操控的真相與謊言

別再讓數字騙了你！一眼看穿投資詐局、不實廣告與虛假民調，
打造最強自我保護力與決策判斷力

HOW TO LIE WITH STATISTICS

作者：德瑞爾·赫夫 (Darrell Huff)

譯者：聞翊均

責任編輯：楊伊琳、賴玟秀

總編輯：鄭明禮

行銷企劃：徐緯程、林羿君

版權專員：劉子瑜

業務部：葉兆軒、林姿穎、胡瑜芳

管理部：蘇心怡、莊惠淳、陳姿仔

封面設計：職日設計Day & Days Design

內頁排版：王信中

法律顧問：証揚國際法律事務所 朱柏聰律師

出版製作：樂金文化

發行：方言文化出版事業有限公司

發行人：鄭明禮

劃撥帳號：50041064

通訊地址：10046台北市中正區武昌街一段1-2號9樓

電話 / 傳真：(02) 2370-2798 / (02) 2370-2766

出版年月：2024 年 2 月

EISBN：9786267321669 (EPUB)

How to Lie with Statistics

by Darrell Huff

Copyright © The Estate of Darrell Huff, 1954

Illustrations copyright © Mel Calman, 1973

This edition arranged with Peters, Fraser & Dunlop Ltd.,
through BIG APPLE AGENCY, INC., LABUAN, MALAYSIA.

Traditional Chinese edition copyright © 2024, by Babel Publishing Company
All rights reserved.



版權所有，未經同意不得重製、轉載、翻印

Printed in Taiwan