

輕鬆搞懂數字爆的料

陳怡如
主計局參謀

壹、書籍介紹

一、書名：輕鬆搞懂數字爆的料

二、作者介紹：

(一)戴文波特Thomas H. DAVENPORT知名管理大師及作家，專長領域為數據分析、企業流程創新和知識管理。哈佛大學訪問教授、貝伯森學院（Babson College）資訊科技暨管理學教授、德勤資料分析室（Deloitte Analytics）資深顧問。編、著或合著有16本著作，包括《魔鬼都在數據裡》、《職場必備的分析力》、《大數據@工作力》等。

(二)金鎮浩Jinho KIM華頓商學院博士。韓國國防大學商學及統計學教授、分析研究實驗室總監。在韓國出版六本著作，其中包括暢銷書《100個統計常識》以及《怪誕統計學》。他研發並主持一門教育學程，幫助個人培養分析技巧，也潛心鑽研如何以量化分析解決各種商業與社會問題。

三、出版單位：天下文化

四、出版日期：2015年3月16日



大家什麼是「量化分析」，如何應用於現實生活，並提供建議方法，協助讀者培養分析技巧，加深對分析學的了解，期透過量化統計思維，教導大家善用資料作出更好的決策，並促使生活及工作上能有效地溝通，解決問題。

量化分析的三階段與六步驟分別是建構問題（確認問題、檢視既有發現）、解答問題（建立模型、蒐集資料、分析資料）、溝通與行動（說明結果與採取行動）。

階段一》建構問題

步驟1：確認問題

所謂建構問題，就是界定量化分析所要解答的問題，以及根據分析結果做什麼

貳、內容概要

本書透過思考三個階段、六個步驟，告訴

決策，這個階段很重要，如果定位錯了，再多資料、再複雜的分析，都無法得到正確結果。

步驟2：檢視既有發現

當你已經找出要解答的問題，可能會發現這個問題的某些層面已經有人處理過，這通常能幫助你更恰當的建構問題。

階段二》解答問題

步驟3：建立模型（選擇變項）

模型就是現象或問題經刻意簡化後的呈現。所謂「刻意」是指，建立模型就是為了解決某些問題；「簡化」表示必須去掉不必要的瑣碎細節，只留下重要而有用的獨特點，在這個階段，若你不是量化分析師，可以藉由這些人的合作達到目標，但是有些關鍵的活動，靠的是你專業的認知與見解，有助於解決方案的提出。

模型可比為漫畫。漫畫會選取某些特徵（鼻子、笑容或一綹髮），其他一概省略，模型也一樣，必須專注於某些特徵、有取捨，找出相關面向，其餘則剔除。你必須創造去蕪存菁的模型世界，才能專注處理問題。

步驟4：蒐集資料

首先，將問題化為具備重要變項的模型，接著，變項經過測量就化為資料。若你選的變項經常有人量測或分析，那麼大可沿用別人的測量量度。但是，有些變項，你需要首開先河，決定量測方法。

步驟5：分析資料

資料本身不會透露任何事，所以需要分析，解開其中的意義，例如，透過電話抽樣調查訪問，蒐集選民對總統候選人的

意見，藉由分析資料，我們可以得知：哪個地區、教育程度、收入、性別、年齡和政黨取向的人，會支持哪個總統候選人。

階段三》溝通與行動

步驟6：說明結果與採取行動

你必須對利害關係人溝通分析結果。這個階段非常重要，就算前面的步驟做得再好，差了這一步，一切恐怕白費。

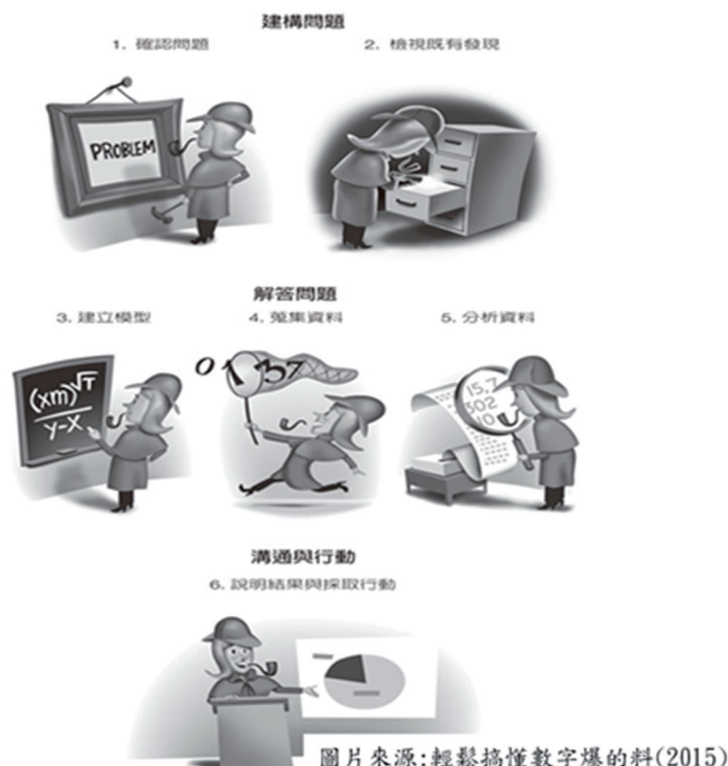
能用資料說故事的分析師，才是優秀的分析師。那些關心自己研究心血是否能付諸實行的分析師，都非常看重這個階段，並在此投注大量心力。

書中列舉南丁格爾的故事來說明這個道理。

1854年10月，英國參與克里米亞戰爭，南丁格爾前往俄國的克里米亞半島，在英國營區裡協助救護和醫療工作。南丁格爾發現，醫院裡的死亡率高達43%，很多軍人沒有戰死，而是受傷被送進醫院後感染而死。於是，南丁格爾開始做統計調查，發現士兵死於院內感染的人數竟然是戰場死亡人數的10倍，她製作「東方部隊死亡原因圖」，創新地透過圓餅圖呈現數據，讓大家清楚看到每月死亡原因的變化情形，數字和圖表一目瞭然，成了強而有力的證據。此舉奏效，南丁格爾不斷回報醫療現場實況給英國政府，鏗而不捨敦促改革，死亡率也逐漸降低。

長久以來，大家不認為這個階段的內涵應納入正式教育學程，傳統上，學術界過於把焦點放在分析方法本身，卻未能透過有效溝通，讓大家了解這些方法，所幸目前情況已改變，漸漸的重視這個步驟的價值。

量化分析：三階段與六步驟



本書中也列舉幾個有趣且實用的案例跟讀者分享，在此擷取其中2則，說明上述三個步驟、六個階段實行的成果。

案例1：FICO評分

所謂的FICO評分，就是以一個介於300到850之間的三位數，簡要說明個人在某個時間點的財務狀況，不管申請信用卡、車貸、房貸或信用貸款，核貸方都會想知道借錢給你所冒的風險，FICO評分會影響到核貸方願意借你多少錢及貸款條件（利率或其他）。FICO評分是將分析結果轉化為行動的絕佳例子，使用情形也愈來愈普遍。

（一）確認問題：

信用分數能讓核貸方快速並客觀的評估客戶的信用風險，在這之前，授信流程相當緩慢、不一致且偏頗，工程師費爾和數學家艾賽克因此心想，若能找出貸款申請人個人背景和財務歷史資料裡的某些因

子，從中算出申貸者的違約風險，應該有助於改善核發貸款的決策。

（二）檢視既有發現：

信用調查報告已有百年以上的歷史，當時有小型零售商聯合販售客戶的財務資料，之後加入電腦設備的運用，發展成大型公司，即便到這時候，泛用型的信用評分系統尚未十分成熟。

（三）建立模型（選擇變項）：

費爾和艾賽克先排除年齡、性別、種族、國籍和婚姻狀態等變項，從信用調查報告裡挑選信用分析變項，並歸納成五大類：

1. 還款歷史

特定種類帳戶的還款資料、不良的公開紀錄、催收項目、延滯的程度及金額、依約付款的帳戶數等。

2 負債額

所有帳戶的負債額、有結存的帳戶數、信用限額的額度、未清償的分期付款比例等。

3. 信用歷史期間

帳戶的開戶期間、最近一次的帳戶活動時間等。

4. 新增信用額度

最近開戶數以及新開戶的類別比例、最近的信用查詢次數、查詢時間、排除過去的帳戶問題後，重建的正面信用紀錄。

5. 動用的信用額度種類

各類型帳戶的數量。

四蒐集資料：

美國人的財務資料不出Equifax、Experian、Trans Union這三大全國信用報告公司之手，財務機構或債權人按月把信用資料送到這幾家公司，資料內容包括消費者的帳戶號碼、信用種類（例如，房貸、信用卡貸款、車貸）、拖欠債款、強制行動、付款歷史。個別消費者在這三家公司所存的資料不盡相同，FICO評分是根據這3家公司所存的個人資料給與分數。

五分析資料：

信用評分會納入信用報告的各種資訊，但每則資訊的重要性並不相同，某些信用歷史會比其他資料重要，因此會得到加權計分，FICO評分是按照費爾和艾賽克開發的專利公式算出的，但每家信用報告公司都有自己獨特的計分演算法。

六說明結果與採取行動：

FICO提供信用報告公司運算軟體，用以計算分數，軟體內含數學公式，公式是藉由抽樣消費者的信用資料而開發出來的。但是，每份信用報告公司都會依據個別考量，小幅修改演算法，因此，各家公

司算出來的分數並不相同。

FICO評分介於300到850之間，得分770以上信用評等最高，725為中等，700為尚可，得分650以下者，通常只能獲得次級抵押貸款，利息會較高。

FICO評分是全世界最廣為使用的信用評分，它提升了風險評估的能力，讓信用市場變得更有效率，信貸機構得以根據消費者的貸款歷史準確預知借款是否會得到償還；此外，愈來愈多跟貸款無關的公司（例如保險公司、手機電話公司、房東、財務金融公司的分支機構）也會善用信用報告，決定是否生意往來，或依據客層訂定費率。有些雇主也會在聘雇之前查核對方的信用評分，協助判斷受雇者會不會是個好員工，簡言之，在分析思考的應用上，FICO評分是成功的範例。

案例2：全視線光學

最常需要分析的一個商業問題就是，到底該在某特定活動上花費多少錢，而行銷開支的最適水準尤其難以定奪。百貨業之父沃納梅克曾有句名言：「花在廣告上的錢有一半是浪費掉的，問題是，我不知道是哪一半。」如今，公司行號可以利用量化分析，判定哪些行銷費用值回票價、哪些不划算，以及哪種行銷組合最有效。這種分析法一般稱為行銷組合分析，這套方法目前廣為從事消費者零售的公司所接受。

（一）確認問題：

全視線光學公司銷售變色鏡片，母公司認為某些特定廣告和行銷活動有待商榷，直覺認為相關費用太高，卻苦無實證資料佐證到底花多少行銷費才合理。全視線的管理階層於是把問題定位為優化行銷開支，以期投注金額能把銷售額提升到相對最大量。

(二)檢視既有發現：

可用的客戶資料散落在各單位。當時只知道可透過這些資料，分析比較不同行銷方式的效果，卻不知道進一步的細節。

(三)建立模型（選擇變項）：

愈來愈多大型機構喜歡用行銷組合優化模型，計算最恰當的行銷開支。這個模型牽涉到行銷回應、行銷成本和產品利潤這幾個變項，它利用線性和非線性規劃法，計算出每週或每月該花多少廣告費、促銷費、該把價錢訂在哪個價位，才能把收入或利潤或兩者最大化。

(四)蒐集資料：

手頭的消費者資料很少，因此無法準確估算到底消費者有沒有看到廣告，或廣告到底有沒有提高銷售額，全視線於是花了幾年時間，成功地把資料整合到資料倉儲系統。

(五)分析資料：

全視線聘用外部顧問來分析資料，排除不重要的解釋因子，確立模型。

(六)說明結果與採取行動：

內部專家從顧問手中取得模型，拿去跟管理高層討論，找出它的意義，然後再融合高層對市場的直覺看法。研究發現，在電視廣告方面投入成本，能讓全視線明顯提高銷售收入。

透過作者列舉的2個故事，我們再次思考三個階段、六個步驟的實行方法，也了解每個步驟的分析層次及重要性，書本內容中尚有許多啟發思考的小故事，字眼淺白卻不乏技術性，甚至列表說明各種變項的差異及測量時機，提供讀者一套完整的思考架構。

參、心得與體驗

在大數據與人工智慧年代，人們期待從傳統上憑經驗和直覺決策，進化到根據數據分析與事實作科學化決策，透過本書的三個階段、六個步驟，不是專業的量化分析家，也能從中培養出分析與推理能力，相關的數字敏感度亦會漸增。

三個階段、六個步驟就像一個指北針，指引我們面對問題的思考方向，按一定的邏輯順序、步驟和方法進行調查、分析及計算，便得以獲取可供決策相關結果，除了作者列舉的例子外，亦有許多與筆者工作業務相關且貼切的實例，如韓榮姿、高志榮（2020）運用多準則決策分析法，找出影響國軍作業維持費編列與分配關鍵因素之關聯性與重要排序，提供各級人員在概算作業時，能快速掌握重點，降低計畫申請退審機率，並提高概算獲賦可能性，本篇文章探討的步驟，如確認問題、檢視現有的資訊、運用分析法找出關鍵因素，到最後說明各因素的重要排序確立模型，便與本書不謀而合。

另楊經武（2017）為避免人員維持費編用差異過大，肇生無效益預算或需動支預備金挹注情形，希望建立較以往精確的推估模型，在確立問題範圍後，依人員維持費組成結構，蒐整影響預測結果的各分項次模型，包含薪餉、各項加給、主副食與口糧等7項相關數據，建立具一定可靠度之模型，後續加以納入不同假設情境，如配合軍公教待遇調整、延長服役年限等檢測，建立的模型已初步具備人員維持費推估的反應能力，可以迅速提供決策參考資訊，上述的2個例子都是本書在工作上最佳運用的例證。

相對地，日常生活中或許不需要如此嚴謹執行數據分析，但是充斥各式各樣的數字，若能藉由簡易量化分析與判斷能力，便能避免資訊誤用或掉入數字陷阱，例如五金店老闆提出「減價50%，再減20%」，他可不是要幫你從定價減70%，而是減掉60%，因為20%的計算基準，已經是減掉50%之後較小的數字，又如使用○牌牙膏蛀牙可以減少23%，會不會只是20個人的少數樣本呢？還有一個知名的例子，百貨公司比爸爸更早知道女兒懷孕的消息，不是黑色幽默，而是塔吉特百貨真實發生的事件。塔吉特怎麼知道顧客懷孕了？說穿了並不神奇，因為該公司設有準媽媽禮物登記處，讓準媽媽們登記自己想要的嬰兒禮物。由於這份清單等於確知已懷孕的顧客名單，塔吉特便依據名單上的消費紀錄，建立起懷孕婦女的購物模型（孕期會購買的商品清單），再用此模型比對其他客人的消費紀錄，找出消費形態類似的顧客，向她們行銷相關商品，達成「未卜先知」的結果。運用這個方法，塔吉特多找出了30%的行銷對象，從降低行銷成本和提高行銷精準度的角度來看，都相當成功。

無論從工作上或生活中的例子得知，量化分析的好處太多了，不但能夠除錯甚至預測，各方面均扮演重要角色。另外，書中所提的數字解讀與呈現，分析結果如何變成故事，也是影響決策的重要一環，甚至用詞不同，解讀也不同，比如「屢戰屢敗」和「屢敗屢戰」每個字長都得一模一樣，但是讀起來的感覺就完全不一樣。後者讓人感覺熱血多了；又比如，根據美國婦女投票行為的調查研究發現，選舉時，每22個老婆裡，只有1個投票和老公投的候選人不一樣。同樣一件事，也可以解讀成，每22個老公裡，只有1個投票和老婆投的候選人不

一樣。前者感覺很沙文主義，後者感覺很女性主義；又比如，有超過四分之三的美國農莊已經可以獲得電力供應…或者解讀成幾乎有四分之一的美國農莊到現在還沒有獲得供電，前者感覺電力公司是負責任的，因此，如何傳達分析結果是不是同等重要呢！

肆、結語

本書是寫給全球不擅長量化分析但具有分析性思維的讀者。作者不預期大家看完後會搖身一變，成為獨當一面的量化分析師或資料科學家，而是，讀完本書後，能理解什麼是資訊、什麼是分析，並渴望自己具備這些能力，對於工作及生活均有所幫助。

筆者身為主財人員，瞭解主財人員主要任務在於協助機關有效分配資源、提供公正客觀的財務資訊，以助機關達成組織目標，我們被賦與專業幕僚的角色，工作中充滿各式各樣的數據，透過本書體會解決問題的客觀條件之一就是運用數據及其分析結果，若主財人員能善用數據，判斷數據的合理性，利用簡單的統計方法，甚或建構模型，讓數據說話，對機關產生的主財問題能夠迅速得到回應及提供具體行動方案，如此，更臻能發揮主財人員的專業立場及服務效能。

參考文獻

1. 楊經武,2017,國軍人員維持費預估模型之建構與運用,主計季刊,第58卷第2期:48-60.
2. 韓榮姿、高志榮,2020,掌握國軍作業維持費編列與分配的關鍵因素,主計季刊,第61卷第2期:16-28.