

圖解資料學

超實用!

資料蒐集與分析的方法
一次學會

- ✓ 從實體到虛擬，一次學會在資訊海洋中尋到寶的know-how。
- ✓ 從蒐集到分析，一次學會將訊息歸整為有力情報的技術。
- ✓ 從二手到唯一，一次學會在有限素材外創造獨家內容的方法。



作者／繪圖 王友龍

圖解資料學

超實用！

圖解資料學分析與應用
——從數據到決策



圖解資料學分析與應用



圖解資料學分析與應用

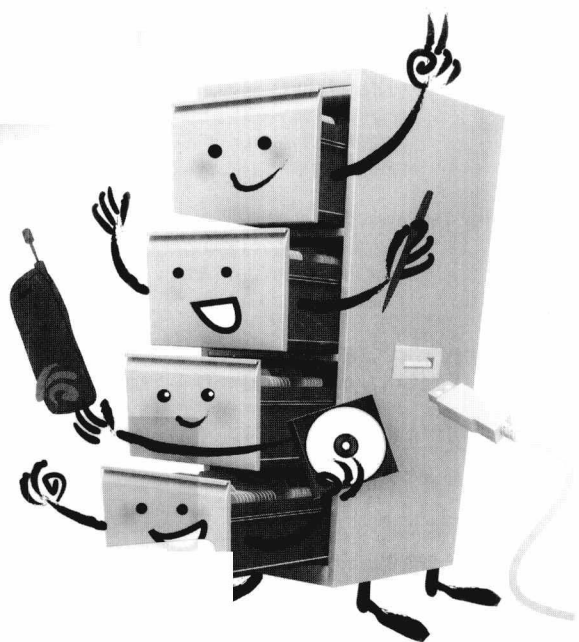


圖解資料學分析與應用



—— 王友國

圖解資料學



作者／繪圖 王友龍

國家圖書館出版品預行編目資料

圖解資料學／王友龍著．——一版．——臺北市：臉譜，
城邦文化出版：家庭傳媒城邦分公司發行，2010.03
面；公分．——（企畫叢書：FP2203）
ISBN 978-986-235-093-5（平裝）

1. 資訊蒐集 2. 資料處理

023.6

99002533

企畫叢書 FP2203

圖解資料學

作 者 王友龍
副總編輯 劉麗真
主 編 陳逸瑛、顧立平

發行人 涂玉雲
出版 臉譜出版
城邦文化事業股份有限公司
台北市中山區民生東路二段141號5樓
電話：886-2-25007696 傳真：886-2-25001952

發行 英屬蓋曼群島家庭傳媒股份有限公司城邦分公司
台北市中山區民生東路二段141號11樓
客服服務專線：886-25007718；25007719
24小時傳真專線：886-25001990；25001991
服務時間：週一至週五上午09:30-12:00；下午13:30-17:00
劃撥帳號：19863813 戶名：書虫股份有限公司
讀者服務信箱：service@readingclub.com.tw

香港發行所 城邦(香港)出版集團有限公司
香港灣仔駱克道193號東超商業中心1樓
電話：(852) 25086231 傳真：(852) 25789337
E-mail：citehk@hknet.com

馬新發行所 城邦(馬新)出版集團【Cite (M) Sdn. Bhd. (458372U)】
11, Jalan 30D/146, Desa Tasik, Sungai Besi,
57000 Kuala Lumpur, Malaysia
電話：(603) 90563833 傳真：(603) 90562833

初版一刷 2010年3月23日

城邦讀書花園

www.cite.com.tw

ISBN 978-986-235-093-5

版權所有·翻印必究 (Printed in Taiwan)

售價：280元

(本書如有缺頁、破損、倒裝，請寄回更換)

目錄

	自序	0 0 6
1	概說	
1	資料的定義與三個要素	0 1 2
2	資料蒐集分析的目的與重要性	0 1 4
3	資料蒐集與分析在提案流程中的位置	0 1 6
4	資料蒐集分析與知識管理的關係	0 1 8
5	資料蒐集與分析是知識經濟的基礎	0 2 0
2	資料的樣貌與分類	
1	從資料到企業情報	0 2 6
2	資料的分類	0 2 8
3	資料分類——從企畫與行銷的角度	0 3 0
4	資料分類——從本質與型態的角度	0 3 4
5	定量與定性情報之間的層級關係	0 3 6
3	資料蒐集的方法	
1	資料蒐集的IPA與5W1H法則	0 4 0
2	資料蒐集的原則	0 4 4
3	建立資料蒐集評估表	0 4 6
4	資料蒐集的流程	0 4 8
4-1	確認研究主題或解決問題的假設	0 5 0
4-2	設定時間範圍	0 5 2

4-3	描述型研究資料的蒐集	0 5 4
4-3.1.1	先蒐集次級（二手）資料——從網路搜尋	0 5 8
	• Google 與 Bing 搜尋結果分析	0 6 8
	• 分析與討論	0 7 4
4-3.1.2	先蒐集次級（二手）資料——從實體資料搜尋	0 7 6
4-3.2	蒐集初級（一手）資料——市場調查	0 8 0
4-4	探索型研究資料的蒐集	0 8 2
4-4.1	進行特定對象的訪談	0 8 4
4-4.2	人員情報型的資料蒐集	0 8 6
5	資料不足時的因應方法	0 8 8

4

資料分析的方法

1	文字型與數據型資料的分析	0 9 4
2	從資料的倉儲化與觀察開始	0 9 6
3	文字型資料的分析流程	0 9 8
	• 描述型案例研究：電子書產業分析（K.J.法的原理運用）	1 0 4
	• 探索型案例研究：店頭訪查資料彙整分析	1 3 4
3-1	分析資料，形成觀點與結論	1 3 8
	• 案例演練：時間序列分析的排列	1 4 2
3-2	呈現分析後的資料，進行提案	1 4 4
4	數據型資料的分析方法	1 4 6
4-1	認識數據型資料的尺度類型	1 4 8
4-2	數據資料分析法	1 5 2
4-3	超級數據分析	1 5 6
5	資料分析的大躍進——高價值情報的產生	1 6 0
	• 分析與討論	1 6 4

5

交叉表與記錄表的運用

1	交叉表與記錄表的由來	1 6 8
2	認識交叉表的形成原理	1 7 2
3	交叉表的運用	1 7 6
4	認識記錄表的形成原理	1 8 4
5	記錄表的運用	1 9 0
	• 分析與討論	1 9 2

6

建立情報競爭力

1	情報與資料的界限	1 9 6
2	建立個人情報系統	1 9 8
3	建立企業的情報競爭力	2 0 0
4	建立家庭的參考資料圖書館	2 0 2

自序

「資料海洋」是我對資料的一種比喻與敬畏，既然是海洋，它就顯得既廣又深，對於沒有經驗的資料蒐集者而言，很容易在這片海洋中沒頂，但懂得資料蒐集的策略與方法，就能駕馭資料，彷彿在海上玩起衝浪板的遊戲，甚至潛入海中，好好觀察海底的世界，讓其他在海上攀著浮木的資料蒐集者只能望洋興嘆。

在現實世界的職場上，我們經常要蒐集與分析資料，才能做出好的決策，決策的擬定很大一部分來自資料的即時與正確，說資料是決策之母，一點都不為過。

每個蒐集資料與分析資料的人員都屬於整個決策系統的螺絲釘與幕後功臣，如果你懂得資料蒐集與分析的方法，有一天，當你坐上決策者大位，你較不會因做錯了決策，而從椅子上摔下來（被請走路），你甚至可對部屬提出資料蒐集政策的指導方針，這代表你從基層一路升上來的紮實功力，可惜，台灣社會重決策而輕忽基本的資料操作工夫（看看書市上一大堆的決策相關書籍即可知），這本書是寫給有心從資料中獲益並想學習蒐集技巧與分析方法的讀者。

現實世界中的資料數量卷帙浩繁，我們一般將資料分為第一手資料（初級資料）與二手資料（次級資料），前者類似別人所沒有的獨家報導，後者是指一般的公開性資料——已經在市場上披露過的資料，第一手資料有可能整垮一個政府，較知名的案例是一九七二～一九七四年美國華盛頓郵報的兩名記者所撰寫的一系列關於水門案件的調查報導，最後導致尼克森總統下台；但如果你認為二手資料不如第一手資料有價值，小看二手資料的威力，那你就錯了，只要懂得整合與分析，創造話題性，二手資料一樣可以嚇嚇叫。一九三九年二戰爆發前夕，英國倫敦的一家出版社出版了一本《歐洲各國軍力評鑑》，對歐洲各國的陸海空三軍戰力做了詳細評比，當然也包括當時野心勃勃的納粹德國，當時的德國元首希特勒看了之後異常震怒，下令參謀本部徹查洩密者，最後查無結果，因為這本書其實是出版社匯集歐洲各國的公開資料，再加以分析與整合而成。

我喜歡拿廚師與食材的例子來比喻資料的蒐集與分析過程，首先，廚師要先準備食材與調味料（資料蒐集），經過烹煮炒炸的料理過程（資料分析）後，再端出一道道的菜色（分析結果）出來，問題來了，如果食材不夠，怎麼辦？廚師是否就做不出好菜？答案是否定的，食材數量少有數量少的作法，一個好的廚師會因應食材的內容與數量而自創不同的菜色，這也是厲害的資料分析者的因應之道，當面臨資料不足時，有兩種方法：另闢蹊徑或就地分析，數年前，我因為要寫一篇關於代糖的市場分析報告，但二手資料很有限，於是我到賣場記錄相關產品的品牌與陳列方法，並訪問一些消費者，也買了一些產品回來加以分析與拍照，最後將產品資料分門別類地整理好，總算交出一篇讓主管放心的報告。

本書的四大特色

四大特色

①雙頁式主題

← 增加讀者的理解與吸收

半邊文字 (800~1,000字) + 半邊圖解 (關聯圖式圖解)

②從實務入手

作者資料蒐集與分析經驗的總結報告

③詳盡的案例演練說明

第三章、資料蒐集的方法

- ◎下載網頁資料的七個步驟
- ◎Google與Bing的搜尋結果比較分析

第四章、資料分析的方法

- ◎運用七個步驟進行文章的裁切與整合
- ◎描述型案例 (六篇電子書產業的報導)
- ◎探索型案例 (機能飲料的店頭訪查資料)

第五章、交叉表與記錄表的運用

從形成的原理、範例說明講到運用的層面

④以「分析與討論」延伸想法與視野

(分布於第三、四、五等三章的內容中)

這本書除了承襲我一貫的雙頁式主題寫作規格——半邊文字（八百～一千字）、半邊圖解（powerpoint關聯圖式圖解），以增加讀者的理解與吸收外，還有以下特色：

一、從實務入手

全書的內容都是我在職場上進行資料蒐集與分析經驗的歸納與整合，算是一本總結報告，並結合一些社會科學的重要專有名詞，以求讀者融會貫通，這是因為我發現這些職場實務上的資料分析方法與社會學的研究方法有頗多相似之處，例如：企畫法中的「K.J.法」就與「紮根理論」是一體的兩面。

二、詳盡的案例演練說明

本書列舉大量的實例說明蒐集與分析資料的步驟與技巧，這些案例主要分布在下列的章節中：

1. 資料蒐集的方法（第三章）—— Google 與 Bing 的搜尋結果分析

以相當的篇幅介紹網路資料的蒐集方法，不僅說明下載網頁資料的七個步驟（從第一步的關鍵字詞輸入技巧講到第七步的以檔案夾儲存下載的網頁資料），並以「電子書」三個關鍵字，比較 Google 與微軟的 Bing 搜尋引擎的搜尋結果，看何者能獲得較佳的資料訊息。

2. 資料分析的方法（第四章）——「描述型」與「探索型」的案例演練

本章說明文字型資料（ex. 一篇文章）的分析方法，先運用七個步驟進行文章的裁切與銜接，而後形成一篇整合後的新文章作為後續的分析基礎（素材），接著再形成各章節的觀點與最後的結論，才算大功告成。

為使讀者容易了解，本章後面以描述型與探索型的案例做說明，描述型案例是以六篇電子書產業的報導為例，講解如何以 K.J. 法的原理進行各篇文章的整合、分析與提案，共有七個演練步驟；探索型案例則以機能飲料的店頭訪查資料為例，說明彙整與分析資料的技巧。

我還增加一篇時間序列分析的排列演練題，以銀行財富管理部門因金融風暴所受到的影響的七段文句為例進行說明；另一方面，也介紹數據資料的尺度類型與分析的方法。

3. 交叉表與記錄表的運用（第五章）——市場調查部分

本章介紹量化與質化調查結果的基本素材——交叉表與記錄表，這是市場調查的輸出產物，從它們形成的原理、範例說明講到運用的層面，這是坊

間一般類似的書籍較少著墨之處，也是讀者不可錯過的內容。

三、以「分析與討論」延伸想法與視野

在第三、四、五等三章的後面，另有「分析與討論」的部分，屬於相關題材的延伸性想法，提供讀者另一種解讀與省思。

目前的台灣社會已將領導、管理、策略、決策與高科技視為長期的顯學，而需要有如工蟻般精神趴趴走的資料蒐集與分析的手工業往往被忽視，有時我們會陷入一種迷思——認為資料到時是從天上掉下來的禮物，船到橋頭自然直，有了這種阿Q心態，自然不會將資料的來源與品質視為一個問題，事實上，資料在組織作業中，一直都是個問題，只有認真做好準備的人才會加以嚴肅看待與解決；如果資料缺乏或錯誤，就不可能轉為有用的情資，自然不可能做出正確的決策。

PART.

1

概説

1

資料的定義與三個要素

「資料」是依附在特定的研究目的下，所蒐集到的所有相關的人類知識集合體，它有三個要素：「形式」、「載體」與「內容的深度」；資料再經過整合分析後，就成為有用的情報。

從廣義的層面而言，因特定的研究目的所蒐集的所有人類知識的集合體都可視為「資料」，特定的分析目的是蒐集資料的前提，由此形成你蒐集資料的動機，例如：你想寫一篇關於台灣飲料產業或人類考古學的報告，那麼你就必須去蒐集相關的資料，而資料包括三個要素：

A · 資料的形式

資料有三種形式（型態）：文字、數字與圖像，有時是片段零散的寥寥數語的文字型態，有時是一篇完整的文章或整理分析過的正式報告，內含大量的文字與統計圖表。

B · 資料呈現的載體

指資料輸出的兩種形式：

(1)實體的紙張形式

可能是手寫的便條紙或資料卡，也可能是電腦打字、再列印輸出的一疊紙張，或是以印刷機印出的書籍（ex. 百科全書）。

(2)數位形式

運用電腦蒐集數位資料並進行儲存（不進行列印），此時，資料是以0與1的電子訊號形式儲存在電腦中。

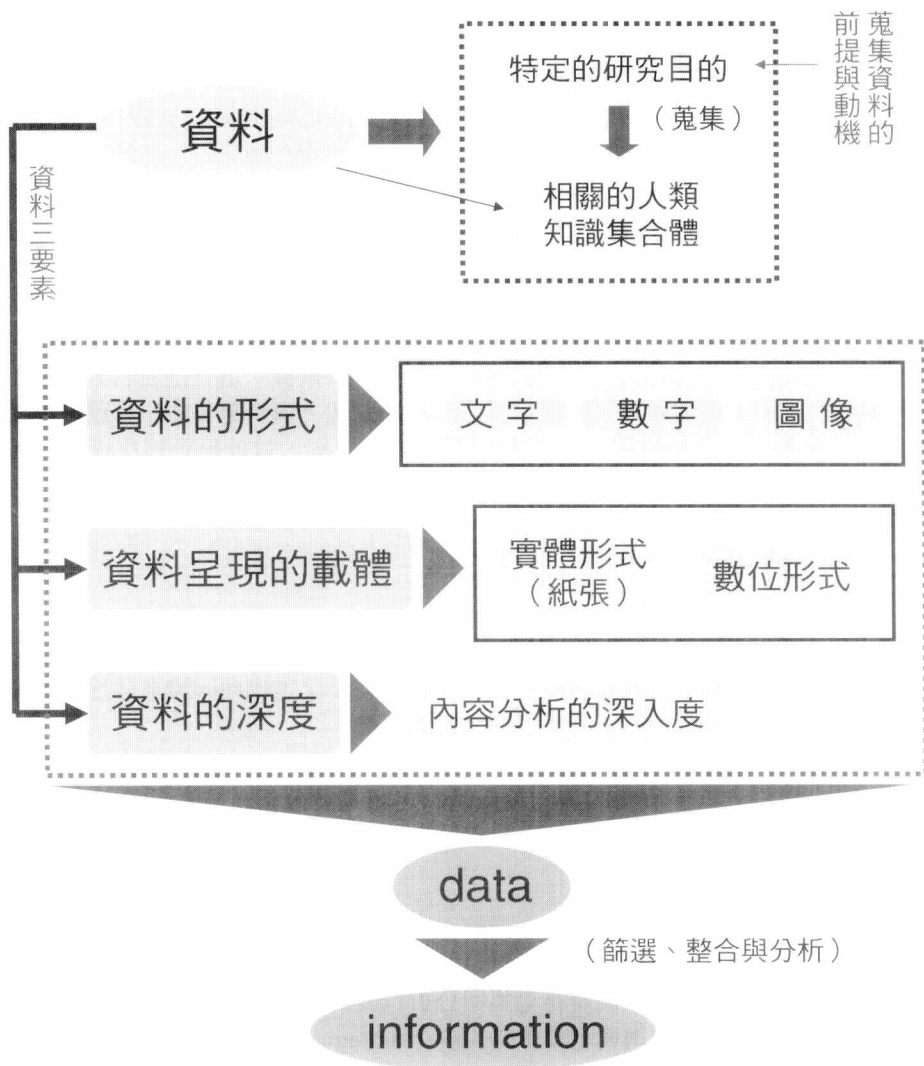
C · 資料的深度

指資料內容的分析深入度，內容多又有深度分析的資料，對研究目標的達成愈有幫助。

這些因特定的研究目的所蒐集的資料，因尚未被整理分析，可視為一般的資料（data），當你進行篩選、整合與分析後，就轉換為有用的情報（information）。

我們所處的世界正被資料的洪流所造成的氾濫所淹沒，依據美國加州柏克萊大學的研究，估計全球每年會產生1~2EB（exabyte）的資訊，相當於10的18次方位元組，而《紐約時報》單日所包含的資訊數量，約等於一個十七世紀英格蘭人一生中所接收到的資訊；如果我們不想被這股資料洪流所淹沒，就應採取一種有效的資料蒐集「策略」，擬定資料蒐集的方向與重點，以提升資料蒐集的效率。

資料的三個要素



2

資料蒐集分析的目的與重要性

你無法憑空杜撰報告書的內容，因此，你必須先蒐集資料，作為資料分析的參考素材，而後製作提案書到最後進行提案簡報，就是資料從誕生、分析、呈現與發表的四個完整歷程。

蒐集資料就像運用你頭上的大型天線截收來自四面八方的訊息，如何篩選正確有用的訊息，考驗你對資料的敏銳度，而資料分析如同啟動你大腦中的超級電腦，對篩選過的訊息進行解碼處理；如果你擅長蒐集與分析資料，將使你擁有中上水準的情報力，搶先別人一步發現新商機，或是掌握了某項新發現的話語權或詮釋權，這些都代表你在職場或學術領域中，有較別人更強的競爭力，也成為你一部分的核心價值。

前面有提過，你因為特定的研究目的去蒐集與分析資料，而資料蒐集是資料分析的前奏曲，提供資料分析所需的參考素材，這是資料蒐集的主要目的，因為你不能憑空杜撰報告內容，一切要有明確的事實作為依據，這就是資料蒐集的工作所存在的價值；當命題與研究目標確定的那一刻起，資料蒐集的工作就已開始啟動，而後經過分析與呈現，進而向相關當事人提案，這就是資料從產生（蒐集）、分析（資料整合與觀點的形成）、呈現（製作提案書）到提案（發表）的完整藍圖（四部曲）。

資料的蒐集與分析與兩個層面有關：

A．與研究有關

研究（research）是人類以系統化的方式探索事物，進而累積知識的心智活動，例如：你了解地球的暖化現象是否存在？又是哪些因素造成地球的暖化？這都是因探索與好奇心而引發了研究動機。

B．與問題的解決有關

先提出假設，依循假設的方向蒐集資料，再進行分析，分析的意義，除了可驗證假設的正確性，也能幫助你釐清問題，並與現狀結合，選擇方向與目標，以增加成功的機率。

一九五〇～一九六〇年代，為企業解決問題與挖掘商機的美國新興行業——企業管理顧問公司大量興起，對資料質量的需求日益殷切，有關資料蒐集與分析的理論與架構也逐漸完備，而資料蒐集與分析也成為撰寫商業類文章／報告的重要第一步，也是提案／企畫與企業策略規畫（strategic planning）的基礎。

個人情報力與資料分析的完整藍圖

