

（摘錄 2）第一章 不分析就被淘汰

分析是什麼？

分析（Analytics）是指運用結構化方法，以資料和分析（analysis）解決商業問題，進而產生效益的一門科學。許多公司行號為了削減成本和業務創新等策略性商業問題而搞得焦頭爛額。解決這些問題的行動方案（initiatives）往往會撼動預算和組織結構。資料分析可以是這些決策強而有力的工具。正確的分析方法，是以資料和洞見為依據，在掌握充分資訊的情況下，提出解決方案。

直覺 + 資料 = 強而有力的洞見 => 好決策

不只要產生洞見，更要產生「可採取行動」的洞見——指那種能夠促進新的思考方案、驅動決策，以及推動正面行動的洞見。寶僑公司（Procter & Gamble）、亞馬遜（Amazon）、LinkedIn 和資本一號（Capital One）等業界領導公司，在願景和直覺的支持之下，部署資料領導型（data-led）、假說驅動型（hypothesis-driven）和分析動力型（analytics-powered）策略而稱霸各自的領域。我們將在本書探討這些方法。

歡迎來到商業分析的世界。

2011 年，寶僑的營業額創下超過 826 億美元的新高紀錄。《財星》（Fortune）雜誌的「全球最受讚譽公司」名單上，寶僑從前一年的第 6 名，上升為第 5 名。

寶僑在全球各地有 12.7 萬名員工，於 180 個國家銷售 300 種品牌的產品。它如何能在自 1837 年以來的 177 年中，持續於家居用品產業中保有領先地位？這家年營業額高達 1,400 億美元的複合型消費性產品公司，由於懂得因時因地制宜，運用「假說驅動型分析策略」不斷自我改造，而能繼續維持舉足輕重的地位，並且提供保持領先優勢所需的燃料。

假說驅動型分析是什麼？不妨想像：你要在太平洋中尋找黃金。難道你不想縮小範圍，只找最可能有黃金的地方，而不必從中國到美國，摸遍每一平方英寸的海水？

同樣地，資料有如一片海洋，假說有助於縮小範圍到最可能找到答案的地方。假說是人類根據利害關係人（stakeholders）的集體智慧和經驗，以及他們對於業務與環境的了解所懷有的直覺產生的。我們利用資料來驗證假說，抽絲剝繭找到解決方案。解決方案的強弱取決於資料和假說兩者的好壞。

假說驅動型分析策略是怎麼運作的？

這套策略起始於詢問一組聽起來相當普通的問題：「發生了什麼事？」和「為什麼會發生？」這些簡單的問題協助寶僑之類的公司找到可用於處理某種狀況的行動。行動如果缺乏洞見，就會在不了解哪些路徑能夠產生結果的情況下做決策。

2010 年，寶僑的執行長鮑伯·麥克唐納（Bob McDonald）採取策略，將公司的作業流程從頭到尾徹底數位化，以便利用分析來實現業務優化和提高決策效率的目標，但同時必須縮減成本。2012 年，寶僑遣散了 1,600 位非製造單位的員工，而且資訊科技組織總支出減少 9 億美元以上。儘管精簡人事和縮減開銷，寶僑的總營業收入還是增加。

怎麼辦到的？寶僑投資於分析專業，以發展在難以預測和充滿壓力的環境中，做出明智且即時決策的能力。該分析結果能預測未來一年寶僑的市場占有率和其他績效等統計數字。這項能力的核心，是一連串的分析模型，將在商業組織中有事情發生時被揭露出來，包括績效達到高峰和跌至谷底、找出為何發生的理由，並且點出可以採取哪些行動，以善用或緩和那些事情造成的影響。

寶僑的分析所依據的文化思維，是在深入執行複雜的資料分析之前，先問問題和提出假說。「是什麼」（What）將監控交貨、銷售和市場占有率等領域，密切注意業務的關鍵數值。例外事件發生時，「為什麼」（Why）就會逐層下探動因，從國家、地區、產品線到商店各層級，以了解事件背後的詳情。分析作業會接著找出寶僑可以拉動的「行動槓桿」（action levers），例如訂價、廣告和產品組合，並且估計各項行動可能造成什麼樣的成果。寶僑的決策人員以很高的效率建立起「是什麼」和「為什麼」，所以能夠一針見血解決問題，急劇提升效率和執行業務的步調。

這套策略對於像寶僑那樣的企業巨擘肯定有效。現在，當你面對越來越少的預算和日益膨脹的數據集（datasets），該如何像業界巨擘那樣促進公司的健全和興隆？你該如何以最適當的效率執行商業策略，並做出可能最明智的商業決策？

你可以藉著詢問正確的一組「是什麼」和「怎麼做」（Hows），並使用簡單的結構化流程（即假說驅動型分析，接著是測試洞見）找出可以執行的策略。

這不是深奧的火箭科學，而是常識性的秘訣。就像寶僑，你可以透過如下所述的假說驅動型、資料領導型流程，問一些普通的問題。假說驅動型分析策略可以提供顯著的效益，包括更強而有力的洞見，以及更快取得資訊和採取行動。以假說驅動型方法產生的可行動洞見相當精準，足以明確執行。

(摘錄 3) 第四章 BADIR：以 5 個簡單的步驟執行商業分析

智慧鞋業公司是在時尚鞋類和配件市場區隔逐漸嶄露頭角的一家線上零售商。它的成長速度令人難以置信。但是有一條配件產品線，銷售狀況不只虎頭蛇尾，營業收入更是節節下降。

執行長湯姆去找分析主任傑夫說，「我想看看各部門和各國的 3 年營業收入資料。」一個星期後，傑夫把資訊交給湯姆，並說：「美國的配件營業收入正在下降。」湯姆大發雷霆，並問道：「你就只能發現這個？我老早就知道了，現在沒空。明天早上 7 點再談。」

傑夫的起步並沒有走得很好。他可能知道如何處理資料，卻錯過了極為重要的第一步：「直探真正的商業問題。」這才是湯姆要他做的事。雖然傑夫在運用資料和分析方面很用心，並且花了不計其數的時間與金錢，建立複雜的分析模型，他仍然沒有確認問題出在哪裡和將它處理好。

本章要介紹 BADIR 架構，利用一套 5 個精實、簡化的步驟，帶你從資料走到決策。這些步驟能夠利用簡單卻強而有力的分析，處理 80% 的商業問題。BADIR 是指商業問題 (Business question)、分析計畫 (Analysis plan)、資料收集 (Data collection)、洞見 (Insights) 與建議 (Recommendations)。

任何經理人都能遵循這 5 個步驟，有效提出分析解決方案，而且通常不需要使用複雜的方法。我們的經驗顯示，如果分析沒有在組織中提出成果，通常是因為跳過一或多道步驟，或者並沒有依照適當的順序進行。舉例來說，商業問題可能沒有釐清，或者利害關係人未能提出有意義的假說去探索和測試。看完這一章，你將能依循這套架構，解決需要分析的大部分商業問題。

有效的分析是資料科學和決策科學交會的結果。技術層面需要資料科學方面的相關技術能力，目的是找到洞見。商業層面則需要比較軟的技能，利用決策科學，去影響你的閱聽人和了解業務，目的是產生影響。商業層面往往被急於依據資料產生洞見的組織所忽視。BADIR 架構將這兩個強而有力的層面合在一起，而提高效能和成果。

用於產生影響的分析 = 資料科學 + 決策科學

第一步・商業問題

它能為你做什麼？

- 減少像無頭蒼蠅那樣無所適從。
- 貢獻可行動的建議。
- 獲認可為決策流程的夥伴。

你可能馬上就注意到，雖然 **BADIR** 是資料分析的流程，一開始卻不是從資料著手。這個架構起於了解真正的商業問題。而且，不是任何問題都行。問錯問題，便會對錯誤的商業問題提供無用的解決方案。問題越接近你的分析需要處理的核心商業問題，你會越快取得正確的答案。正確的商業問題會因為做出更明智的決策和產生影響，而直接把你放在通向既定目標的路上。減少像無頭蒼蠅那樣無所適從，加快整個流程，而且會貢獻可行動的建議，交付商業價值。因此，你和你的團隊會被視為業務上有價值的夥伴。這一切都起於需要花時間認清眼前正發生什麼事和問相關的問題。

真正的商業問題，是在了解問題背後的意圖（**intent**）和商業考量（**business considerations**）之後才會發現。

• 從既定目標到真正商業問題的架構

我們先依循傳統的六個問題架構：何事、何人、何地、何時、為何和如何。

一開始就問問題的目的，是確認在它所處的情境中，出了什麼問題。這是偵探福爾摩斯所用的方法。如果不能把問題縮小到真正的問題，那就好像搜尋整個（資料）海洋找黃金，那就是探險家哥倫布所用的方法。你在這裡，應該當起偵探家。為了做分析，一開始就需要問切實相關的問題，協助你了解目前的因素、過去的事件，或未來的策略。

（摘錄 4）案例研究：J.K.羅琳改名出新書

情況：有位第一次出書的作者，名叫羅伯·蓋布瑞斯（**Robert Galbraith**），寫了一本犯罪小說《杜鵑的呼喚》（**The Cuckoo's Calling**）。英國《週日泰晤士報》（**The Sunday Times**）在推特上收到一條匿名消息，說蓋布瑞斯其實就是羅琳，也就是《哈利波特》（**Harry Potter**）小說系列的作者羅琳（**J. K. Rowling**）。

行動：《週日泰晤士報》的美術編輯理查·布魯克斯（**Richard Brooks**）在向出版商查證這件事之前，和他的團隊成員做了一些調查工作，發現兩位原作者的出版商和代理商的確都相同。他們請兩位電腦科學家證實《杜鵑的呼喚》和羅琳的書是否有語言上的相似性。

法律語言學利用統計分析，以採擷文件、部落格、推文和亞馬遜上的書評，尋找作者方面的線索。作家會選用一些字詞以傳達特定的訊息，但這些字詞帶有他們不知道已經在傳達的特有資訊。他們拿了 5 本書的內文給電腦科學家，測試上面的假說，其中 4 本各來自已知的作者，另一本是羅琳寫的《臨時空缺》(The Casual Vacancy)。

第一項分析中，一本書的每一萬字序列都用電腦程式 JGAAP 來跑，以 4 種不同的分析，比較《杜鵑》和其他的書，每一種分析都著重在寫作的不同層面。一種是比較字對 (word pairings) 或者相鄰的字組。舉例來說，它會顯示作者描述為「昂貴」的事物種類：汽車、衣服、食物等。

第二項測試則搜尋「n 個字元」(character n-grams) 或者相鄰字元的序列。舉例來說，搜尋跳 (jump) 的序列，會出現 jump、jumps、jumped 和 jumping。第三項測試則專注於最常用的字，如 a、and、of 和 the 等字。觀察每一本書中 100 個最常用的字，並且比較頻率上的差異，會顯示某個作者使用的某個字出現 5%，另一位作者則出現 7%。

第四項測試只觀察字的長度，以分析一本書有多少百分比是由 3 個字母的字、4 個字母的字等等構成。第四項測試顯示的型態非常具有羅琳的特色。

另一位科學家對 8 本書做了第二項分析，也就是 4 位作者的每一位都加一本書。他以稱作 Signature 的語言軟體程式，分析這 8 本書和《杜鵑》。所有的書都比較六個特色：字的長度、句子的長度、段落的長度、字母的頻率、標點符號的頻率，以及字的用量。《杜鵑》在四項這類測試上，都證明和已知是羅琳寫的書最為相似。

影響：布魯克斯相當肯定推特的匿報者所言不虛之後，便去找羅琳，很快就證實她的確是《杜鵑》的作者。