

推薦序 *Foreword*

馬雲與郭台銘不約而同都說，「新零售」時代來臨了，純電商時代過去了。

新零售其實就是：「電子商務＋大數據＋人工智慧＋虛實整合＋物聯網＋新世代的物流、倉儲」，甚至還有切入到網路金流的機會。最好的案例就是 Amazon 或阿里巴巴。

但是這些新發展背後的重要關鍵，就是大數據。大數據驅動了這些商務公司對消費者的深度瞭解，進一步掌握消費者的型態，再透過新科技去改造物流與倉儲這樣的舊行業，提升服務消費者的效率。最後消費者被無所不在的商務業者掌握，而消費者也享受更多的便利性。說穿了，消費者不知覺的貢獻出自己的行為模式數據，換來更貼心的服務。

因為市場規模的緣故，台灣不會出現像 Amazon 或阿里巴巴這樣規模的業者，而台灣電子商務與零售業者不被取代的競爭力核心，就是大數據。不只是掌握大數據在手上，還要知道深度的運用。只有這樣，才能在國際新零售的潮流下不被淘汰。

我想，這也是作者們辛苦寫出這本書的用心！

前行政院長

張善政

Foreword 推薦序

近年來數位轉型已成為產業界的發展趨勢與重要議題，而數位轉型的核心就是**數據**。隨著大數據和人工智慧技術的發展，如何處理大量倍增的資料儼然是當今企業最重要的一環，而在視覺化分析解決方案中佔有優勢的 Microsoft Power BI 已經成為各個企業與政府機關必備的商業智慧分析工具。

書中以適切的分析方法和範例，說明零售業是如何透過 Power BI 視覺化工具，在營運與管理上進行數據分析，建構管理者需要的儀表板以及客戶 360 度分析，在電子商務與零售業大量數據的環境中，利用簡單上手的工具幫助每個人都能獲得資料分析的能力，從大數據中找出商業價值，進而採取對應的行動，在大數據的時代快速掌握商機。

我與謝邦昌教授認識多年，教授一直致力於大數據和人工智慧教育推廣，也是在台灣第一個與微軟合作，將實體課堂教學搭配微軟線上課程，培養產業 AI 人才。長期以來，老師不但在學校培育英才，更在各個產業應用領域，將一般人可能望之卻步的大數據分析與人工智慧，利用 Excel 與 Power BI 等工具，幫助大家輕鬆駕馭。

這是一本強調實務應用的工具書，希望讀者們能夠透過各位作者們集思廣益的成果，對於建置 Power BI，大數據的操作與應用都能有更深一層的認識與技能提升，並經由謝教授與作者群“全民大數據、產業應用大數據實作”的理念，引領企業成功數位轉型。

台灣微軟首席技術與策略長

丁維揚

8



零售大數據實戰分析

本單元內容要談的是零售大數據－分析架構，屬於相當重要的部份。因為在執行許多數據分析專案時，使用者多半會從資料模型來尋找可用的數據源來解決問題，而資料模型通常就是整個搭建數據架構的主體。

以下就零售產業領域的數據架構做說明，內容涵蓋從數據來源的組成架構、分析指標的建立到分析模型的形成。

8.1 架構數據介紹

► 數據系統架構

通常數據架構會區分不同功能，從系統角度為主的架構，主要有數據蒐集、數據儲存、數據處理與分析、以及重要的數據應用等（如圖 8-1 所示）；其中，數據應用會和視覺化分析最有關係。

- (1) 數據蒐集：該階段要確立蒐集數據的相關文件規範，像是資料格式定義、詞句內容、蒐集的週期，還有數據蒐集遇到異動時，如何進行變更。都是在數據蒐集時就得定義好相關規範，才能確保數據蒐集是處於正確的階段。
- (2) 數據儲存：結構化資料（指有系統性的資料內容），會儲存在具結構化的資料庫中；反之，非結構化資料（指不具系統性的資料內容），常見的有

圖片、聲音等，則要儲存在不具結構化的資料庫中，像是 NoSQL、Hadoop 的環境。

- (3) 數據分析：該階段是數據分析進行的地方，主要步驟依序為數據清理 → 簡單統計分析 → 數據建模。
 - ✓ 數據清理：指的是數據建模前的資料預處理階段，統稱為數據分析做準備的前置。
 - ✓ 簡單統計分析：完成了數據清理之後，必須對於數據有一些初步理解，可利用視覺化來做初步的解讀數據或者建立關聯模型。
 - ✓ 數據建模：不同於關聯模型，數據建模有許多種情境，例如預測模型跟分類模型就可以產生許多不同情境；同時演算法也是數據建模的重點項目。
- (4) 數據應用：基於前面幾個步驟，數據應用多半指將分析結果應用在公司的業務場景。不管是模型、視覺化儀表板等，甚至是說分析報告也是一種數據應用。

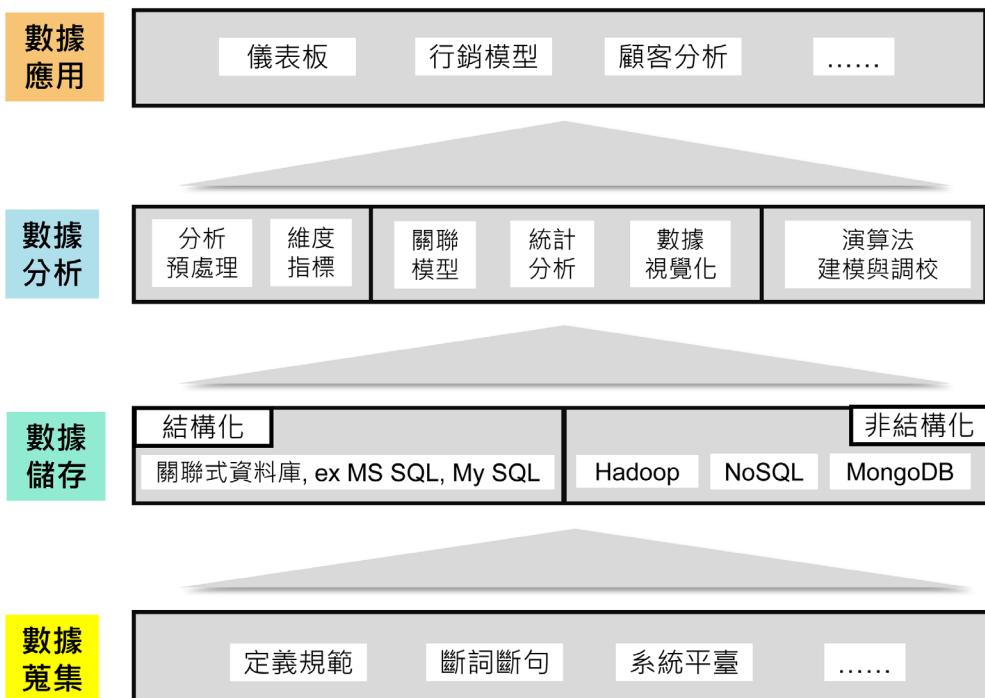


圖 8-1 數據系統架構

► 應用數據架構

現在的零售大數據多半會與電商做互聯網設計的整合，傳統的零售多半是以實體通路為大宗，相對來說也沒有太多的數據，甚至累積的數據多半是「事後數據」，稱為交易（結果）數據。通常這樣的數據都是針對現象去做分析，很難去瞭解到客戶在交易前的心聲，因此傳統零售在「事前數據」的蒐集多半是缺乏的，而形成難以用交易數據去做精準預測模型。

隨著網路跟互聯網等技術的發展之下，傳統零售走向跟電商進行整合，對於蒐集數據這部分，已經可以透過技術來累積更多的「事前數據」，同時運用各種數據分析的手法技術，衍生許多應用場景出來，然而圖 8-2 是實務運用上的一個數據組成架構。

- (1) 營運數據：指的是企業內部數據，以營運為主產生的各種指標數據，像是營收、銷售量、訪客數或轉換率等重要營運指標。
- (2) 行為數據：指的是客戶交易前產生的數據，像是瀏覽時間、瀏覽哪些商品、加入購物車紀錄或分享擴散動作（網路影響力）等動作；這些數據都是會影響是否交易的重要關鍵，也是用來建立數據模型的主要依據。
- (3) 競業數據：指的是競爭對手，像是商品類型、商品名稱、商品價格或者競爭對手的攻擊等數據，透過行為數據和競品數據可以做內、外部的分析，來預防對手的攻擊等。

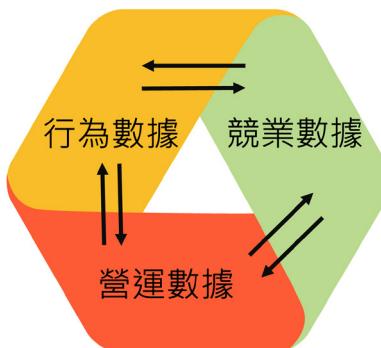


圖 8-2 應用數據組成架構

以上實務運用的架構，一定是來自許多重要元素所組成。圖 8-3 是零售大數據企業與電商整合的標準數據架構，包括虛實整合的成份，總共可以分為 5 大層的數據元素組成。

- 第1層：新零售時代納入虛實整合（O2O）。虛擬世界的數據，從流量開始產生，用戶（或訪客）在瀏覽過程中所發生的各種行為，都會產生電商所要分析的數據，因此流量就是所有虛擬通路的集合數據，沒有蒐集這層數據就不用實現新零售了。
- 第2層：流量數據可以說是交易前的數據，又稱事前數據。在第2層的交易數據，指的是產生轉換行為，用戶下訂單就是產生了交易。而庫存數據代表交易產生就需要進行供給貨物，因此交易數據跟庫存數據是密不可分的關係。
- 第3層：該層數據可以歸納為辨識數據及關聯數據。基於第1、2層數據，可以知道客戶的流量數據和交易數據，串出客戶在虛擬通路的完整行為，以及購買過哪些商品，享受過哪些服務，而產生售後的數據等。
- 第4、5層：財務數據會與收入、支出有關。通常行銷活動、商品採購及軟硬體建設等成本，會歸納在支出；營收、廣告等，會歸納在收入項。透過財務數據和實體通路數據的結合，就能計算出虛實整合（O2O）的實質效益，進而推估出未來需投入的資源比例等重要營運分佈訊息。

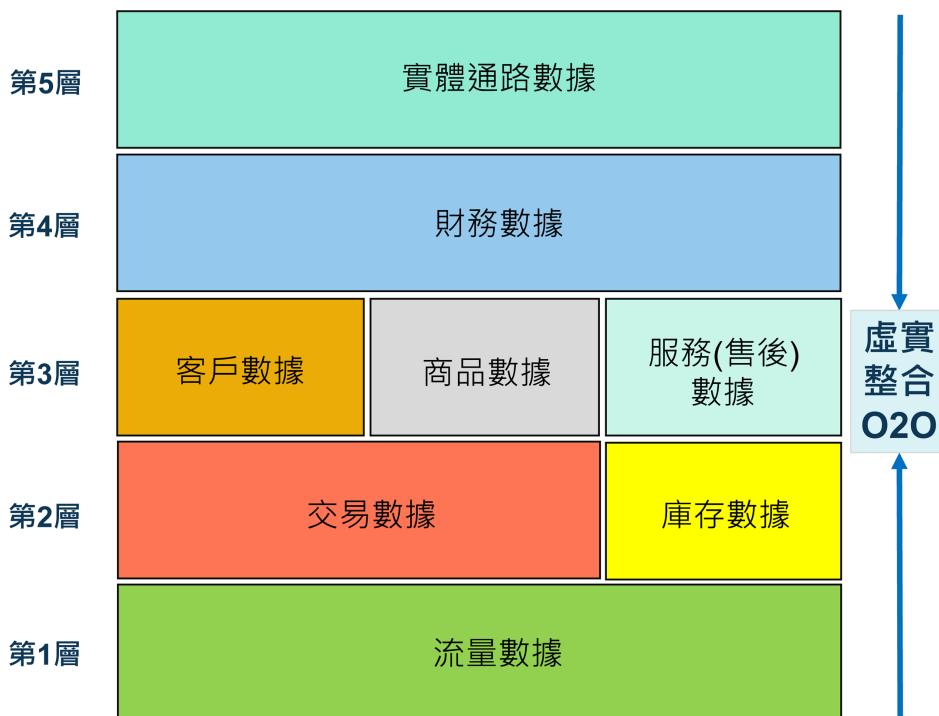


圖 8-3 應用數據組成元素

► 維度 v.s. 指標

我們從零售大數據的數據系統架構、應用數據組成架構及應用數據組成元素等來看，其實是在執行數據分析的維度跟指標變換分析，目的是把過往的策略執行或行銷活動的結果，輔以維度跟指標來理解結果改變的程度，甚至衡量出高或低的比較。

- **維度**：區分 4 大塊，產品、店家、通路及客群，結合起來可以綜合構思分析彼此維度的關聯影響效果。
- **指標**：承上，維度就類似構面的意思，分析彼此的關聯影響效果，必須根據指標來進行衡量比較，而這些在零售大數據裡，不外乎就是使用流量、轉換、銷售到服務等細部指標來比較。

以圖 8-4 來說，從零售大數據的維度及指標來看，初步可以拆分成 16 種分析項目（如圖 8-5）。再分別從這 16 種分析項目進行展開分析，獲取更細部的價值資訊。

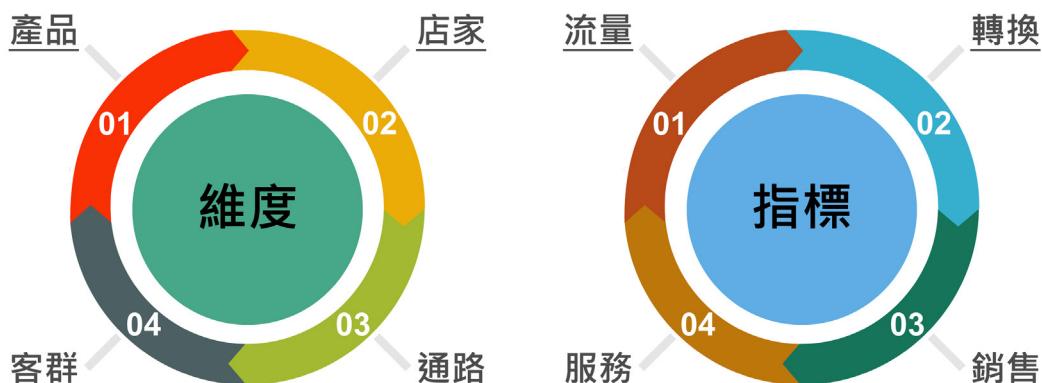


圖 8-4 維度和指標的比較

- 對店家來說：掌握店家的營運管理，助於找出提升銷售機會點。
- 對產品來說：洞察每個產品的表現好壞，助於調整產品的市場策略。
- 對通路來說：洞悉通路的轉換結果，幫助優化通路策略內容。
- 對客群來說：洞悉不同客群的偏好程度，幫助調整內外廣到的行銷策略。



圖 8-5 維度和指標的分析元素

8.2 建立分析指標

在零售大數據領域裡，有很多種分析指標很常使用到。目的就是要從大量且雜亂無章的數據萃取價值訊息，幫助管理階層做出最適判斷及決策；在現實數據分析過程中，分析人員也常遇到使用的數據內容，無法完全滿足分析需求，而產生訊息不足的問題，此時就要思考「增加維度」的方法，通常作法就會是根據需求建立量值計算，把多個維度數據進行組合比較，連結到可以用在解決問題的訊息裡，如此才能減少訊息不足的現象。

接下來我們將說明幾個關於零售業的數據實戰演練案例。

實戰演練 (一) 創建對比指標

通常在執行市場調查分析時，需要瞭解整體業績或某類業績的近兩年的增長狀況。此種數據分析的角度就是「對比」的意思，也就是說假設我們單獨觀察某個數據時並不能看出什麼端倪，此時必須把兩個數據放在一起才能看出兩者之間的差異程度。如同圖 8-6 所示，比較 2016 年及 2017 年的總營收，可以知道市場的增長趨勢。

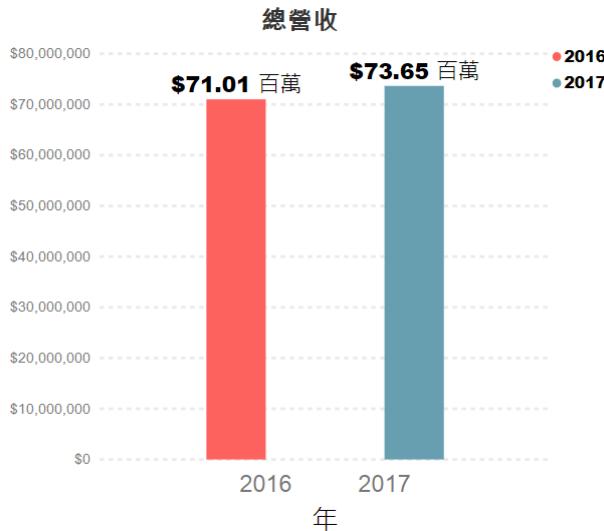


圖 8-6 近 2 年總營收比較

Power BI 的 DAX 語言就能隨時提供計算「對比」比較的函數，透過內建計算資料的高效壓縮功能，快速提供使用者分析「對比」數據指標。此實戰演練將以 Power Pivot 和 Power View 做搭配使用說明。

■ 使用範例：ch8_範例_實戰演練 Part1

資料形式說明：創建「同期比較」指標時，資料表建議存在 "日期參照表"，以利清楚定義內容以及後續創建「同期比較」指標。

■ 演練步驟：切換至資料表區 → 選取 DAX 量值管理表（請讀者先行仿造 CH6.3 建立）→ 滑鼠右鍵 - 新增量值。

Step 1 新增【總訂單金額】量值指標。

DAX 宣告處：總營收 = SUMX('訂單詳細資料表','訂單詳細資料表'[售價]*'訂單詳細資料表'[數量])

Step 2 新增【同期總訂單金額】量值指標。

DAX 宣告處：同期總訂單金額 = CALCULATE([總訂單金額],
DATEADD('時間維度對照表'[日期],-1, YEAR))

Step 3 新增【同期總訂單金額成長率】量值。

DAX 宣告處：同期總訂單金額成長率 = DIVIDE([總訂單金額]-[同期總訂單金額],[同期總訂單金額])

■ Power View 視覺效果 - 設計同期比較

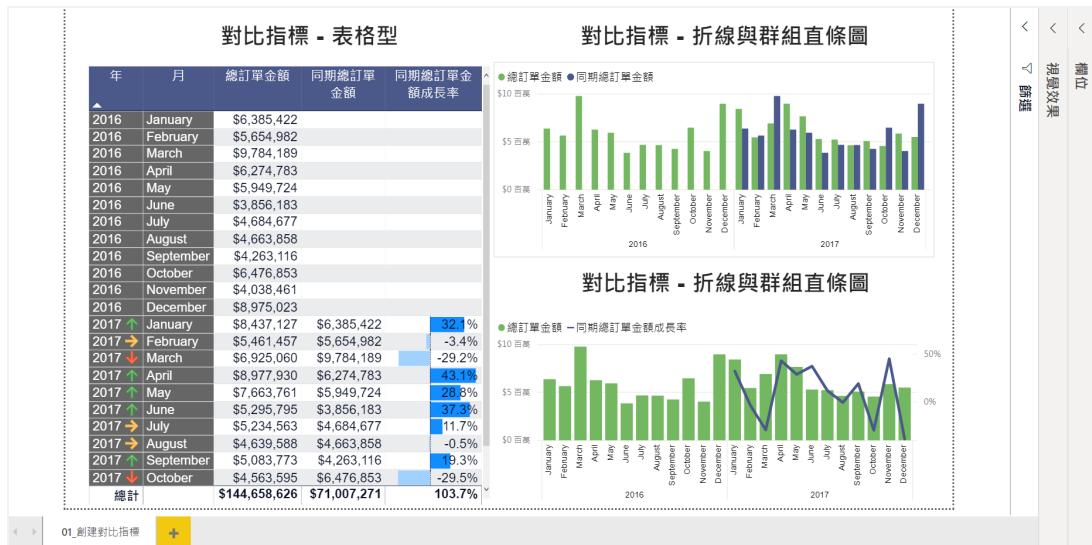


圖 8-7 同期比較指標

實戰演練 (二) 創建品類 (品牌) 熱度指標

品類或品牌熱度通常是用來評估是否能在市場留存的關鍵指標。然而計算品類或品牌熱度的方法有許多種，以下將介紹使用 DAX 函數創建累積佔比形式搭配折線與群組直條圖視覺效果來分析品類（產品）熱度。

■ 使用範例：ch8_範例_實戰演練 Part1

- Step ①** 前置作業。切換至 "訂單細項表"，分別新增 3 個資料行。依序為訂單日期、產品類別及總訂單金額。

DAX 宣告處：訂單日期 = RELATED('日期對照表'[資料日期].[Date])

DAX 宣告處：產品類別 = RELATED('產品資料表'[產品類別])

DAX 宣告處：總訂單金額 = [售價]*[數量]

Step 2 新增資料表。模型 - 計算 - 新增資料表；資料表名稱為品類熱度資料。

DAX 宣告處：品類熱度資料 =
**SUMMARIZE('訂單細項','訂單細項'[產品類別],"總訂單金額指數",
SUM('訂單細項'[訂單金額]))**

產品類別	總訂單金額指數
烘焙食品	\$19,114,485
肉類罐頭	\$10,639,651
飲料	\$22,722,397
油品	\$1,661,818
麥片	\$9,007,711
湯品	\$4,142,364
穀類	\$4,065,277
義大利麵	\$8,840,205
蔬菜水果	\$16,897,666
糖果	\$3,057,010
調味品	\$7,881,755
零食	\$318,770
乳製品	\$2,382,216
乾果	\$16,847,121
果醬	\$6,365,846
醬料	\$10,714,334

圖 8-8 完成產品類別總訂單金額指數

Step 3 新增品類熱度指標。這裡利用建立好的品類熱度資料表，計算分析各個品類的訂單金額指數累積佔比，如此才能知道品類熱度程度。切換至 "品類熱度資料" → 模型 - 計算 - 新增資料行。

DAX 宣告處：品類訂單金額指數累積百分比 = **SUMX(FILTER('品類熱度資料',EARLIER('品類熱度資料'[總訂單金額指數])<='品類熱度資料'[總訂單金額指數]),'品類熱度資料'[總訂單金額指數]/SUM('品類熱度資料'[總訂單金額指數]))**

產品類別	總訂單金額指數	品類訂單金額指數累積百分比
飲料	\$22,722,397	15.71%
烘焙食品	\$19,114,485	28.92%
蔬菜水果	\$16,897,666	40.60%
乾果	\$16,847,121	52.25%
醬料	\$10,714,334	59.65%
肉類罐頭	\$10,639,651	67.01%
麥片	\$9,007,711	73.24%
義大利麵	\$8,840,205	79.35%
調味品	\$7,881,755	84.80%
果醬	\$6,365,846	89.20%
湯品	\$4,142,364	92.06%
穀類	\$4,065,277	94.87%
糖果	\$3,057,010	96.98%
乳製品	\$2,382,216	98.63%
油品	\$1,661,818	99.78%
零食	\$318,770	100.00%

圖 8-9 完成品類訂單金額指數累積百分比

EARLIER 函數

- ✓ 功能：用來進行外部參照計算的實用函數，場景多用於計算排名值。
- ✓ 使用說明：**EARLIER('品類熱度資料'[總訂單金額指數])<='品類熱度資料'[總訂單金額指數]**；表示 **EARLIER('品類熱度資料'[總訂單金額指數])** 為複製一份總訂單金額指數的意思，而 **EARLIER('品類熱度資料'[總訂單金額指數])<='品類熱度資料'[總訂單金額指數]** 表示利用現在真正的總訂單金額指數和複製的所有總訂單金額指數進行比較。

FILTER 函數

- ✓ 功能：表示過濾功能，該函數不能單獨使用，需搭配其他敘述函式。
- ✓ 使用說明：敘述函式 **FILTER('品類熱度資料',EARLIER('品類熱度資料'[總訂單金額指數])>='品類熱度資料'[總訂單金額指數])**，表示將大於等於目前總訂單金額指數的資料行篩選出來。

Step 4 新增品類訂單金額指數排名。切換至 "品類熱度資料"，模型 - 計算 - 新增資料行。

DAX 宣告處：品類營收指數排名 = **COUNTROWS(FILTER('品類熱度資料',EARLIER('品類熱度資料'[總訂單金額指數])<'品類熱度資料'[總訂單金額指數]))+1**

產品類別	總訂單金額指數	品類訂單金額指數累積百分比	品類營收指數排名
飲料	\$22,722,397	15.71%	1
烘焙食品	\$19,114,485	28.92%	2
蔬菜水果	\$16,897,666	40.60%	3
乾果	\$16,847,121	52.25%	4
醬料	\$10,714,334	59.65%	5
肉類罐頭	\$10,639,651	67.01%	6
麥片	\$9,007,711	73.24%	7
義大利麵	\$8,840,205	79.35%	8
調味品	\$7,881,755	84.80%	9
果醬	\$6,365,846	89.20%	10
湯品	\$4,142,364	92.06%	11
穀類	\$4,065,277	94.87%	12
糖果	\$3,057,010	96.98%	13
乳製品	\$2,382,216	98.63%	14
油品	\$1,661,818	99.78%	15
零食	\$318,770	100.00%	16

圖 8-10 完成品類訂單金額指數排名設定

■ Power View 視覺效果 - 設計品類熱度指標

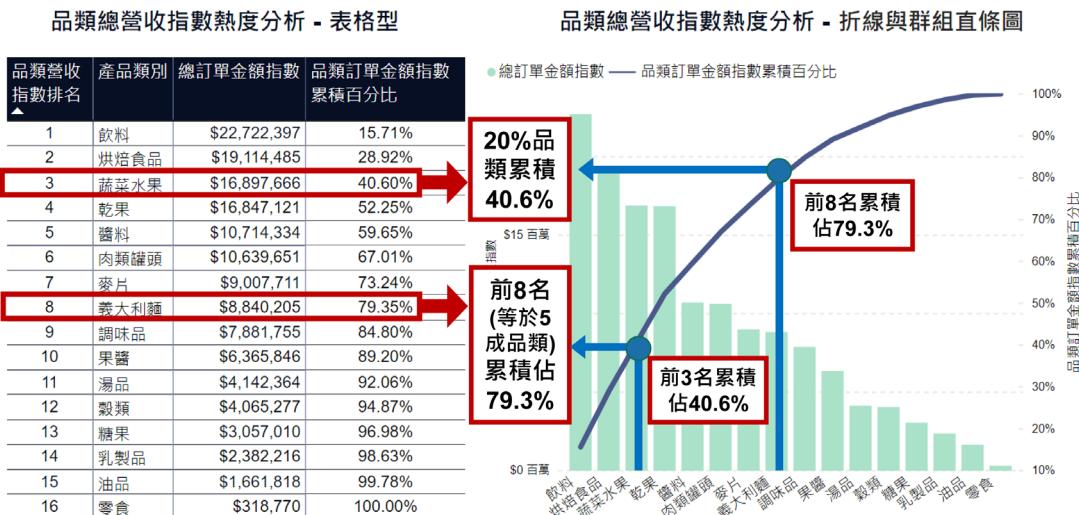


圖 8-11 完成品類（品牌）熱度指標

品類熱度視覺化分析－表格型。進一步將建立完成好的品類熱度資料表，使用 Power View 視覺化分析，如圖 8-11。依據排名排序，可以知道❶集中在前 8 名的品類，佔整個營收指數的 80%左右；或者❷目前 20%的品類約佔營收指數的 40.6%，表示仍有些許成長空間（以 20：80 法則來看）。

品類熱度視覺化分析－折線與群組直條圖。同樣地，使用 Power View 的另一種圖表呈現，如圖 8-11。依據排名排序。用 20：80 法則來看，可以知道集中在前 8 名的品類，佔整個營收指數的近 80%左右；或者目前前 20%（前 3 名， $16 \times 20\% = 3.2$ ）的品類約佔營收指數的 40.6%，表示仍有些許成長空間。

8.3 建立分析模型

實戰演練 (三) 地域分佈四象限應用

通常零售通路還會分析不同區域的會員數，以及瞭解會員的一些貢獻情況。我們接下來將以區域分析模型來比較不同區域之下的會員貢獻，並進一步依據區域顯示出來的會員特徵進行分群定義，分別是「低規模，高貢獻」、「高規模，高貢獻」、「低規模，低貢獻」和「高規模，低貢獻」，如圖 8-12 的四象限說明。因此就能針對不同象限對象給予不同的刺激策略。

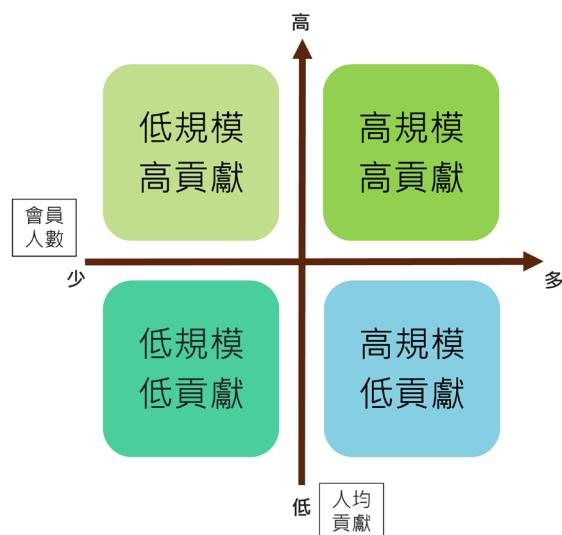


圖 8-12 會員區域貢獻四象限定義

■ 使用範例：ch8_範例_實戰演練 Part2

Step 1 前置作業。若要進行會員區域貢獻四象限來命名會員分群，必須事先新增相關量值函數。切換至資料表區 → 選取 DAX 量值管理表（請讀者先行仿造 CH6.3 建立）→ 滑鼠右鍵 - 新增量值（如下）。

DAX 宣告處：

總消費金額 = **SUM('零售會員交易檔_FIN'[總價格])**

總消費筆數 = **DISTINCTCOUNT('零售會員交易檔_FIN'[訂單編號])**

總消費數量 = **SUM('零售會員交易檔_FIN'[數量])**

總會員人數 = **DISTINCTCOUNT('零售會員輪廓檔_FIN'[會員編號])**

有消費會員人數 = **DISTINCTCOUNT('零售會員交易檔_FIN'[會員編號])**

有消費會員比率 = [有消費會員人數]/[總會員人數]

人均消費金額 = [總消費金額]/[總會員人數]

人均消費訂單數 = [總消費筆數]/[有消費會員人數]

人均消費數量 = [總消費數量]/[有消費會員人數]

平均每筆訂單數量 = [總消費數量]/[總消費筆數]

Step 2 會員區域分析 Power View 表格。



縣市	會員類型	總會員人數	有消費會員人數	人均消費金額	人均消費訂單數	人均消費數量	總消費金額
新北市	一般會員	15,536	14,066	\$5,610.52	1.56	11.1	\$87,165,106.00
台北市	一般會員	11,171	10,231	\$5,898.09	1.60	11.5	\$65,887,589.00
台中市	一般會員	10,947	9,247	\$5,070.64	1.60	10.9	\$55,508,329.00
高雄市	一般會員	10,640	9,284	\$5,032.20	1.53	10.3	\$53,542,651.00
桃園縣	一般會員	6,174	5,672	\$5,643.05	1.58	10.9	\$34,840,198.00
新北市	VIP會員	6,049	6,049	\$10,837.96	3.00	19.3	\$65,558,850.00
臺南市	一般會員	5,311	4,532	\$5,067.57	1.56	10.5	\$26,913,875.00
彰化縣	一般會員	5,031	4,228	\$5,268.63	1.64	11.5	\$306,489.00
高雄市	VIP會員	5,017	5,017	\$10,123.75	2.72	17.0	\$790,860.00
台中市	VIP會員	4,217	4,217	\$10,796.53	3.02	15.0	\$528,971.00
台北市	VIP會員	4,091	4,091	\$10,247.04	2.85	14.0	\$920,657.00
桃園縣	VIP會員	2,903	2,902	\$10,788.60	2.98	13.0	\$319,300.00
臺南市	VIP會員	2,239	2,239	\$10,923.59	3.03	14.0	\$457,912.00
屏東縣	一般會員	1,960	1,652	\$4,543.73	1.49	10.0	\$905,709.00
彰化縣	VIP會員	1,746	1,746	\$11,445.59	3.21	18.0	\$984,005.00
雲林縣	一般會員	1,635	1,266	\$4,890.26	1.50	11.0	\$995,583.00
新竹縣	一般會員	1,566	1,379	\$5,572.61	1.70	12.0	\$726,701.00
基隆市	一般會員	1,529	1,324	\$5,429.61	1.62	10.0	\$301,877.00
新竹市	一般會員	1,463	1,298	\$5,475.49	1.47	9.0	\$601,647.00
南投縣	一般會員	1,413	1,146	\$4,833.58	1.57	10.6	\$6,602,222.00
苗栗縣	一般會員	1,338	1,121	\$4,934.40	1.58	10.3	\$5,624,753.00
嘉義縣	一般會員	1,213	976	\$4,637.06	1.65	13.6	\$784,524,897.00
總計		113,846	103,601	\$6,891.11	2.02		

圖 8-13 會員區域分析 Power View

Step 3 匯入「Quadrant Chart by MAQ Software」(.pbviz) 視覺效果檔案。

該視覺效果可以讓我們快速分析各地區情況。利用本書提供的視覺效果檔案（資料夾名稱「Power bi Visual 線上下載檔案」），使用匯入自訂視覺效果功能，「從檔案匯入」形式選取「Quadrant Chart by MAQ Software」(.pbviz) 視覺效果匯入，如圖 8-13 所示。

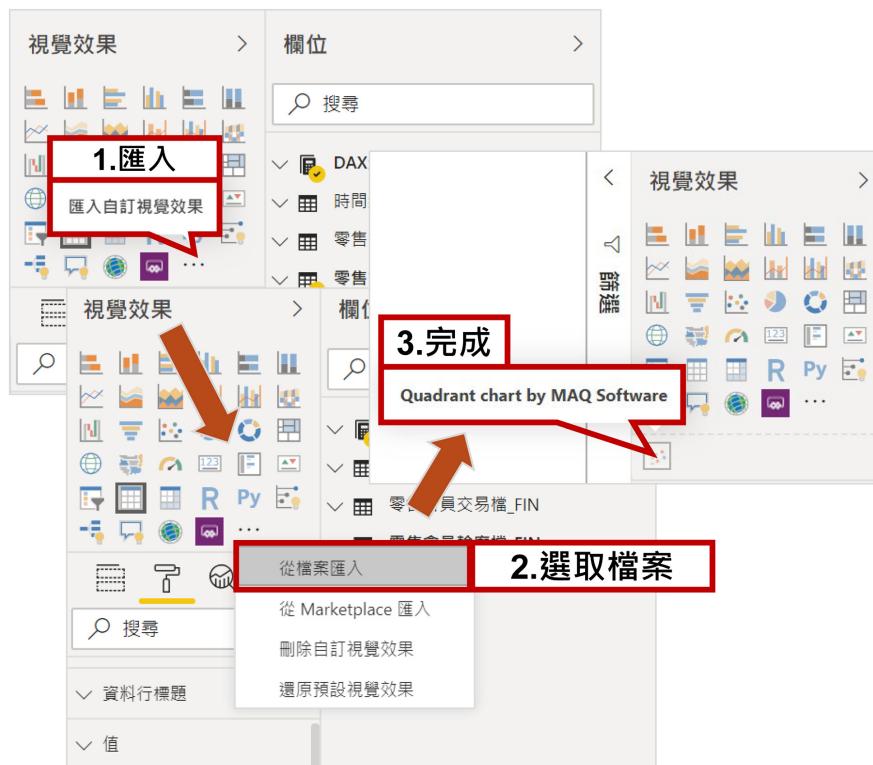


圖 8-14 匯入「Quadrant Chart by MAQ Software」(.pbviz) 視覺效果檔案

Step 4 區域會員四象限視覺化應用分析 – **Quadrant Chart by MAQ Software**。

可以透過客群和入會管道的篩選器搭配，比較各縣市區域在貢獻的差異。而在經營上的任務，可能需要避免這些位於高貢獻區域降級到低貢獻分類中。

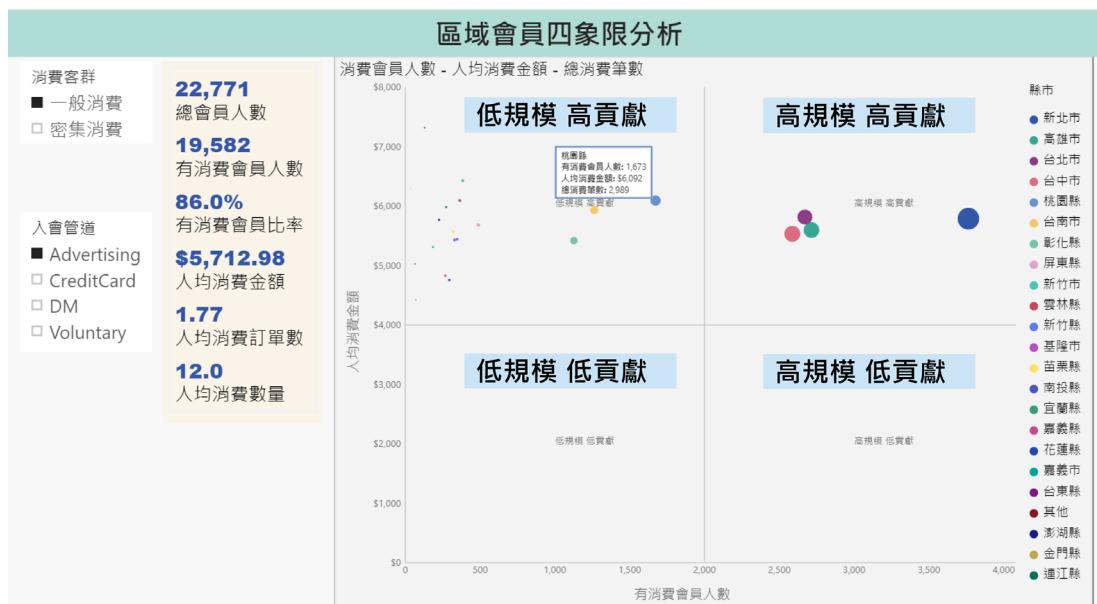


圖 8-15 區域會員四象限視覺化應用分析—Quadrant Chart by MAQ Software

針對區域會員四象限應用分析，筆者整理對應的可能經營策略方向，如表 8-1 說明。

表 8-1 會員四象限經營策略

分群名稱	可能問題	經營方向
低規模，高貢獻	行銷廣告投放資源不足，觸及該市場的客戶較少	投入該市場行銷資源，刺激會員規模成長
高規模，高貢獻	無	屬於重點經營區域，維持市場資源和行銷策略，甚至可以投入更多資源
低規模，低貢獻	已經經營許久，卻未見改善	可能可以採取順其自然的經營方式，倘若在資源充足之下，或許可以研究新的市場拓展方案
高規模，低貢獻	會員的人均貢獻低	優先任務是提高會員人均貢獻，可嘗試優惠套餐的設計跟主題式行銷活動等

實戰演練 (四) RFM 客戶價值分析模型

RFM 客戶價值分析模型是用來衡量客戶價值以及客戶獲利的重要手法。

RFM 客戶價值分析模型的概念，是利用顧客過去的歷史交易紀錄，包括最近一次購買日期指標（Recently）、某時段的購買頻率指標（Frequency）及某時段的購買金額指標（Monetary），透過這 3 個指標來衡量客戶價值。

Recently 跟 Frequency 是用來評估客戶忠誠度指標，Monetary 是評估客戶利益高低之指標。RFM 的用途還可用來做為直效行銷工具，優勢在於可提高❶提高回應率，❷降低行銷成本，❸客戶個人化。

- 使用範例：ch8_範例_實戰演練 Part2

Step 1 計算 R 指標。最近一次購買日期指標（Recently）。

R 表示計算現在和最近一次購買時間的間隔天數，這裡的「現在」，我們設定以「2007-12-31」為基準。切換至零售會員輪廓檔_FIN → 模型 - 計算 - 新增資料行，距今時間間隔 → 模型 - 計算 - 新增量值，量值 R。

DAX 宣告處：距今時間間隔 = DATEDIFF([最近交易日期],"2007-12-31",DAY)

DAX 宣告處：R = CALCULATE(MIN('零售會員輪廓檔_FIN'[距今時間間隔]),'零售會員輪廓檔_FIN'[是否曾消費] = "是")

Step 2 計算 F 指標。購買頻率指標（Frequency）。

F 表示計算頻率次數，這裡定義為客戶的訂單筆數。我們直接使用 DAX 量值管理表的量值總消費筆數 → 並複製新增一量值為 F。

DAX 宣告處：F = DISTINCTCOUNT('零售會員交易檔_FIN'[訂單編號])

Step 3 計算 M 指標。購買金額指標（Monetary）。

M 表示計算累計貢獻金額。我們直接使用 DAX 量值管理表的量值
總消費金額 → 並複製新增一量值為 M。

DAX 宣告處： $M = \text{SUM}(\text{'零售會員交易檔_FIN'}[\text{總價格}])$

Step 4 計算 RFM 資料表。

各別完成 RFM 指標之後，需要建立客戶的 RFM 資料表。切換至資料表區 → 模型 - 計算 - 新增資料表。

DAX 宣告處：客戶 RFM = CALCULATETABLE(SUMMARIZE('零售會員輪廓檔_FIN','零售會員輪廓檔_FIN'[會員編號],'零售會員輪廓檔_FIN'[年齡],'零售會員輪廓檔_FIN'[年齡組距],'零售會員輪廓檔_FIN'[入會管道],'零售會員輪廓檔_FIN'[職業],"R",[R],"F",[F],"M",[M],"數量",[總消費數量]),'零售會員輪廓檔_FIN'[是否曾消費]="是")

另同時新增 RFM 資料表資料欄位，分別為 "年齡"、"年齡組距"、"入會管道"、"職業" 及量值 "總消費數量"。

會員編號	R	F	M	會員價值RFM	數量	年齡	年齡組距	入會管道	職業
DM041796	469	1	2200	4_一般發展客戶	4	33.161643	4.31歲~40歲	DM	服務工作人員
DM059281	419	1	1240	4_一般發展客戶	4	35.084931	4.31歲~40歲	DM	服務工作人員
DM074196	472	1	1700	4_一般發展客戶	4	34.734246	4.31歲~40歲	DM	服務工作人員
DM040988	549	1	2849	4_一般發展客戶	4	35.857534	4.31歲~40歲	DM	服務工作人員
DM073562	563	1	1300	4_一般發展客戶	4	34.358904	4.31歲~40歲	DM	服務工作人員
DM040509	458	1	3100	4_一般發展客戶	4	38	4.31歲~40歲	DM	服務工作人員
DM059202	419	1	1760	4_一般發展客戶	4	38.19452	4.31歲~40歲	DM	服務工作人員
DM042032	487	1	1920	4_一般發展客戶	4	31.104109	4.31歲~40歲	DM	服務工作人員
DM042411	465	1	1150	4_一般發展客戶	4	38.610958	4.31歲~40歲	DM	服務工作人員
DM073900	538	1	4580	4_一般發展客戶	4	33.726027	4.31歲~40歲	DM	服務工作人員
DM039847	524	1	4190	4_一般發展客戶	4	39.720547	4.31歲~40歲	DM	服務工作人員
DM042994	452	1	1260	4_一般發展客戶	4	31.575342	4.31歲~40歲	DM	服務工作人員
DM073134	592	1	1620	8_一般挽留客戶	4	32.843835	4.31歲~40歲	DM	服務工作人員
DM074654	474	1	4060	4_一般發展客戶	4	31.112328	4.31歲~40歲	DM	服務工作人員
DM071466	465	1	1160	4_一般發展客戶	4	34.057534	4.31歲~40歲	DM	服務工作人員
DM060215	392	1	1980	4_一般發展客戶	4	39.205479	4.31歲~40歲	DM	服務工作人員
DM021872	329	1	3490	4_一般發展客戶	4	37.438356	4.31歲~40歲	DM	服務工作人員
DM060268	389	1	3980	4_一般發展客戶	4	31.852054	4.31歲~40歲	DM	服務工作人員
DM060287	388	1	3840	4_一般發展客戶	4	37.049315	4.31歲~40歲	DM	服務工作人員
DM060265	367	1	4060	4_一般發展客戶	4	36.09315	4.31歲~40歲	DM	服務工作人員

圖 8-16 客戶 RFM 資料表

根據客戶 RFM 資料表的定義，設定大於平均值的標籤為「+」，小於平均值的標籤為「-」，因此 RFM 的組合總共會有 8 種，如表 8-2 所示。我們並分別給予這 8 種組合客群不同名稱。

再來我們在下一個步驟將透過 DAX 函數計算每一位客戶的定位，及所屬身分 RFM 標籤命名。

表 8-2 RFM 價值分析模型標籤

R (間隔天數)	F (訂單數量)	M (訂單金額)	客戶標籤身分
+	+	+	重要價值客戶
+	+	-	一般價值客戶
-	+	+	重要保持客戶
-	+	-	一般保持客戶
+	-	+	重要發展客戶
+	-	-	一般發展客戶
-	-	+	重要挽留客戶
-	-	-	一般挽留客戶
+	+	+	重要價值客戶

Step 5 計算 RFM 客戶價值。切換至 RFM 資料表區 → 模型 - 計算 - 新增資料行。

DAX 宣告處：會員價值 $RFM = IF([R] < AVERAGE([R]), IF([F] > AVERAGE([F]), IF([M] > AVERAGE([M]), "1_重要價值客戶", "2_一般價值客戶"), IF([M] > AVERAGE([M]), "3_重要發展客戶", "4_一般發展客戶")), IF([F] > AVERAGE([F]), IF([M] > AVERAGE([M]), "5_重要保持客戶", "6_一般保持客戶"), IF([M] > AVERAGE([M]), "7_重要挽留客戶", "8_一般挽留客戶")))$

會員價值RFM = IF([R]<AVERAGE([R]),IF([F]>AVERAGE([F]),IF([M]>AVERAGE([M]),"1_重要價值客戶","2_一般價值客戶"),IF([M]>AVERAGE([M]),"3_重要發展客戶","4_一般發展客戶")),IF([F]>AVERAGE([F]),IF([M]>AVERAGE([M]),"5_重要保持客戶","6_一般保持客戶"),IF([M]>AVERAGE([M]),"7_重要挽留客戶","8_一般挽留客戶")))

會員編號	R	F	M	會員價值RFM	數量	年齡	年齡組距	入會管道	職業
DM041796	469	1	2200	4_一般發展客戶	4	33.161643	4.31歲~40歲	DM	服務工作人員
DM059281	419	1	1240	4_一般發展客戶	4	35.084911	4.31歲~40歲	DM	服務工作人員
DM074196	472	1	1700	4_一般發展客戶	4	34.734246	4.31歲~40歲	DM	服務工作人員
DM040988	549	1	2849	4_一般發展客戶	4	35.857534	4.31歲~40歲	DM	服務工作人員
DM073562	563	1	1300	4_一般發展客戶	4	34.358901	4.31歲~40歲	DM	服務工作人員
DM040509	458	1	3100	4_一般發展客戶	4	38	4.31歲~40歲	DM	服務工作人員
DM059202	419	1	1760	4_一般發展客戶	4	38.19452	4.31歲~40歲	DM	服務工作人員
DM042032	487	1	1920	4_一般發展客戶	4	31.104109	4.31歲~40歲	DM	服務工作人員
DM042411	465	1	1150	4_一般發展客戶	4	38.610958	4.31歲~40歲	DM	服務工作人員
DM073900	538	1	4580	4_一般發展客戶	4	33.726027	4.31歲~40歲	DM	服務工作人員
DM039847	524	1	4190	4_一般發展客戶	4	39.720547	4.31歲~40歲	DM	服務工作人員
DM042994	452	1	1260	4_一般發展客戶	4	31.575342	4.31歲~40歲	DM	服務工作人員
DM073134	592	1	1620	8_一般挽留客戶	4	32.843835	4.31歲~40歲	DM	服務工作人員
DM074654	474	1	4060	4_一般發展客戶	4	31.112328	4.31歲~40歲	DM	服務工作人員
DM071466	465	1	1160	4_一般發展客戶	4	34.057534	4.31歲~40歲	DM	服務工作人員
DM060215	392	1	1980	4_一般發展客戶	4	39.205479	4.31歲~40歲	DM	服務工作人員
DM021872	329	1	3490	4_一般發展客戶	4	37.438356	4.31歲~40歲	DM	服務工作人員
DM060268	389	1	3980	4_一般發展客戶	4	31.852054	4.31歲~40歲	DM	服務工作人員
DM060287	388	1	3840	4_一般發展客戶	4	37.049315	4.31歲~40歲	DM	服務工作人員
DM060265	367	1	4060	4_一般發展客戶	4	36.09315	4.31歲~40歲	DM	服務工作人員

圖 8-17 客戶 RFM 價值分析模型表

■ Power View 視覺效果 - RFM 價值分析模型視覺化應用分析儀表板。

讀者可參照著圖 8-18 或「ch8_範例_實戰演練 Part2」範例檔案，設計自行所屬企業的 RFM 視覺化應用分析儀表板。



圖 8-18 RFM 價值分析模型視覺化應用分析儀表板

製作 RFM 價值分析模型視覺化應用分析儀表板，會使用到哪些視覺效果及主要欄位呢？請參考表 8-3 及圖 8-19 所示。

表 8-3 RFM 價值分析模型儀表板使用視覺效果

視覺效果名稱	主要資料欄位設計				備註	
	型態：類別或文字	型態：計數、值或量值				
文字方塊	RFM 價值分析模型視覺化應用	-				
散佈圖	會員價值 RFM	R、F、M				
資料表	會員價值 RFM	會員編號計數、平均年齡、R 平均、F 平均、M 平均、數量的平均				
群組橫條圖	年齡組距	會員編號計數				
群組橫條圖	職業	會員編號計數				
群組直條圖	入會管道	會員編號計數				

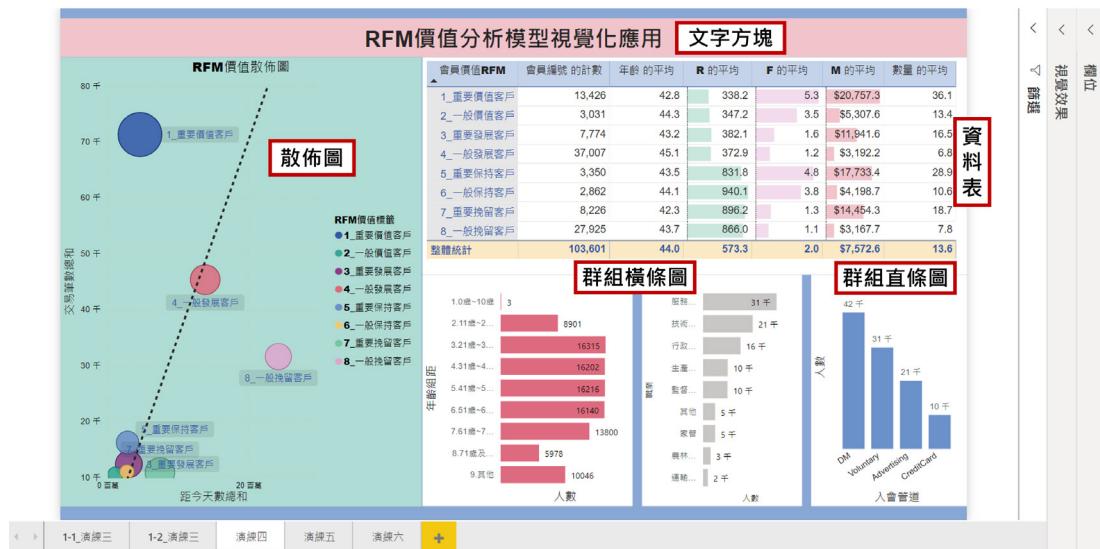


圖 8-19 RFM 價值分析模型儀表板視覺效果設計

可以知道不同客戶身分的 RFM 標籤，代表客戶價值等級不同，然而從標籤身分來看，最直接的就是特徵行為差異。因此應該針對這些特徵行為差異結果，給予不同的客戶經營策略，目的是驅動客戶標籤身分能夠「往上升等」。

針對客戶標籤身分，筆者整理可行的對應經營策略方向，如表 8-4 說明。

表 8-4 RFM 客戶價值分析模型經營策略

客戶標籤身分	RFM	經營策略
重要價值客戶	+++	消費時間最近、次數最多、金額最高。屬於企業最該關心的客群，應施以關心關懷方式，隨時追蹤行為表現，提供高優質服務為主。
一般價值客戶	+ + -	消費次數最多、金額最高。有潛力成為重要價值身分的客群，應施以互動式方式操作，建立黏度習慣，例如遊戲化、猜猜看、MGM 等方式。
重要保持客戶	- + +	消費時間遠、但是以往次數多、金額高。屬於以前的忠誠客戶，應施以主動式出擊方式，保持聯繫溝通的習慣。
一般保持客戶	- + -	消費時間遠、但是以往次數多、金額高。屬於以前的忠誠客戶，應施以主動式出擊方式，保持聯繫溝通的習慣。
重要發展客戶	+ - +	消費時間近、以往次數多、金額一般。屬於潛力客戶類型，首要任務須提升消費頻率，或許可用不定期驚喜來經營該客群。

實戰演練 (五) 分解樹 AI 模型

分解樹（Decomposition Tree）是 Power BI 在 2019/11 發佈的一項新功能，它是屬於一種內建的 AI 模型。分解樹可以依照不同維度進行目標變數的分解，透過內建模型計算來找出影響目標變數的最高或最低因素。

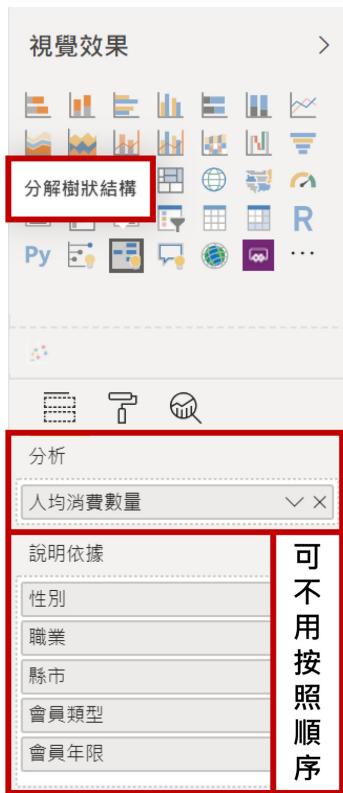
接下來我們說明使用分解樹（Decomposition Tree）AI 模型來進行會影響 "人均消費數量" 指標的變數重要度如何？

■ 使用範例：ch8_範例_實戰演練 Part2

Step 1 啟用分解樹狀結構 視覺化模板。因為是更新版本，目前都是預覽階段，因此需啟用分解樹狀結構 視覺化模板 → 檔案 - 選項及設定 - 選項 → 全域 - 預覽功能 - 勾選 "分解樹狀結構視覺效果" → 確定。



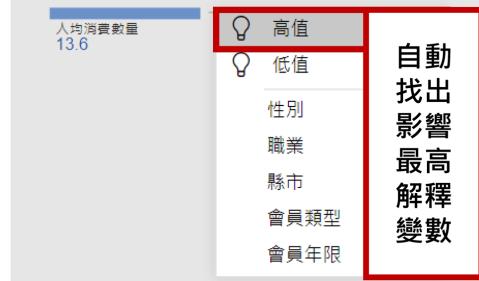
Step 2 設定變數屬性。此時我們可以先設定分析目標變數 "人均消費數量" → 選取解釋變數依據，分別為會員年限、縣市、職業、會員類型、性別。



Step 3 運用分解樹結構 - 啟用 AI 分割。此時我們可以先設定啟用內建的 AI 分割。此為針對置入的解釋變數，透過內建的 AI 演算法進行樹的展開。格式 - 啟用 AI 分割。



Step 4 運用分解樹結構 - 進行影響依據分析。點選分析目標變數的 + 號 → 選擇 "高值" 展開。"高值" 表示 AI 會自動選取的解釋變數中，第一優先影響的解釋變數為何，以此類推。



Step 5 完成分解樹結構 - 進行影響依據分析。依序展開完成後，我們可以試著解釋其中一條規則。

- ✓ 影響 "人均消費數量" 的 變數依序重要性，分別是 會員年限 → 縣市 → 職業 → 會員類型→ 性別。
- ✓ 影響 "人均消費數量" 的 變數依序重要性及其中一條規則，分別是 會員年限（七年期）→ 縣市（新竹市）→ 職業（運輸設備操作工）→ 會員類型（VIP 會員）→ 性別（F）。



圖 8-20 分解樹結構分析

其實從分解樹的功能可以知道，Power BI 後續在 AI 這塊的結合漸漸會提高。分解樹雖然看似簡單，運用也很容易，不過重點仍是 Input 的內容，以及 Output 的解釋。AI 講究的精神之一就是快速運算、易懂的概念。