

# 運用資料探勘技術探討行銷推廣模型之研究 -以家庭清潔用品為例

吳泰熙

國立臺北大學企業管理學系教授

taiwu@mail.ntpu.edu.tw

鍾永富

國立臺北大學企業管理學系研究生

seven.chung7@gmail.com

## 摘要

本研究利用個案公司，透過網路銷售的交易與會員資料，以最近購買日期、購買頻率、購買金額以及顧客關係長度等四個變數，透過分析層級程序法，利用團體決策方法決定上述四個購買行為變數之權重，最後依據加權後的顧客購買行為變數來衡量顧客終身價值。同時利用此四個變數作為顧客分群的依據，以集群分析成功的將顧客分為「新進客戶」、「淺嘗即止客戶」、「潛力客戶」、「預警客戶」、「黏著客戶」及「忠誠客戶」等六群顧客群。

接著透過關聯法則Apriori演算法，分別找出各集群顧客購物籃的關聯規則，最終透過規則的適用篩選，以及領域知識的探討，訂出各集群客製化交叉銷售的行銷策略，供個案公司管理階層擬定行銷決策之建議。

**關鍵詞：**資料探勘，分析層級程序法，顧客終身價值。

**Keywords:** Data Mining、AHP、CLV

## 1. 緒論

面對競爭的環境，企業如何在有限的資源條件下，利用顧客歷史交易資料，透過資料探勘（Data Mining）技術找出潛藏的規則和顧客互動與溝通、保持良好關係、建立穩定的交易模式，協助其保有競爭力，是本文欲探討的議題方向。本章將先針對此議題的研究背景提出說明，再將研究的動機加以描述，進而提出研究的目的與流程。

### 1.1 研究背景

根據模範市場研究顧問公司（world panel, 2012）的調查，台灣家庭清潔用品的整體使用量在近三年呈現持平趨勢（見圖 1-1-台灣家庭清潔用品整體使用量）。在整體市場需求量未提升的情況下，市場競爭相對更加激烈，現有廠商為了讓新進者不易進入此產業及維持其競爭優勢，紛紛採取成本領導事業策略（Cost Leadership Business Strategy），藉由創造以成本為基準的進入障礙，來降低新進入者的威脅。於是，低單價定位策略的家庭清潔用品充斥於整個家庭清潔用品市場。

個案公司在三年前進入家庭清潔用品市場，以天然、無毒為訴求，滿足部份消費者對於安心與洗淨力兼具的需求，驅動了家庭清潔用品市場中，天然商品版圖的成長。面對激烈的競爭，除了新產品的發展外，更應注重顧客關係管理（Customer Relationship Management）。原因有二：一是 Kotler（1994）的消費者行為模式指出消費者行為是由認知進而情感而發生行為，因此和消費者深入的溝通可以提升顧客對於天然、無毒清潔用品的涉入程度。二是因天然商品在成本偏高的現況下，相較一般石化低價的商品，使用的消費者仍佔少數。因此，如何有效的和消費者及顧客溝通、互動，進行促銷客製化，提供價值的交換滿足需求，將是個案公司最重要的工作之一。

由於電腦資訊科技的發展，軟、硬體技術、設備的提升，企業持續累積大量的顧客交易資料，但若無法有效運用，則與資料垃圾並無兩樣。因此，如何利用這些資料成為有用的資訊已成為一重大課題。藉由資料探勘技術的運用，商業智慧（Business Intelligence）的觀念從 90 年代開始發展（尹相志，2010），經由資料探勘技術，從大量的資料中，透過統計相關的演算法，找出潛藏的規則與有用的資訊，加上領域知識的投入，轉化形成商業的智慧，

提供決策者制定決策的參考 (Fayyad, 1996)。由於資料探勘技術提供可預測與資料描述特性，應用的產業領域相關廣泛，包含生物資訊、電信業、網路、行銷、醫學、圖書館等，均可使用資料探勘來挖掘出大量有用的知識 (丁一賢、陳牧言，2003)。究竟，透過資料探勘相關技術的運用，如何協助企業在顧客關係管理的運作與規劃是本研究欲探討與深入分析的主要原因。

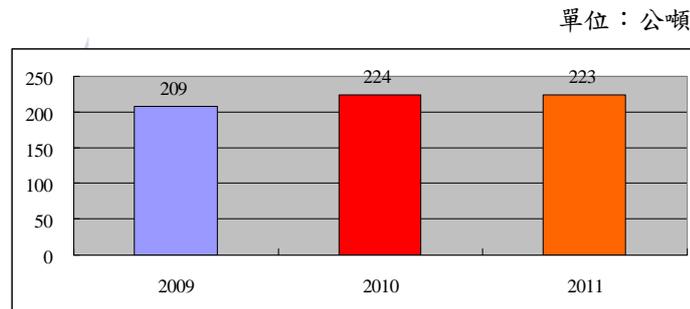


圖 1-1-台灣家庭清潔用品整體使用量

資料來源：引自 World panel，模範市場研究顧問公司、本研究整理

## 1.2 研究動機

現今企業面臨四種主要驅力的挑戰：市場驅力、技術驅力、顧客驅力、及行銷功能的驅力 (張力元等，2007)。面對日趨白熱化的競爭，企業所投入的資源有限。因此如何作有效益的運用，是一個刻不容緩的議題。將資源有效分配運用，除了提升獲利之主要目的外，更應該針對不同顧客需求提供最適化之行銷活動組合，達到有效滿足顧客需求之目的，並藉此建立與顧客的忠誠關係，達到資源分配之最佳效益。因此，本研究希冀利用資料探勘的技術，將個案公司提供的顧客交易資料記錄，以統計演算法將顧客分群，並以顧客關係管理之顧客導向的前提，針對不同集群顧客提出關聯分析之購物籃規則，並提出客製化的行銷策略提供行銷人員擬定決策之建議。

## 1.3 研究目的

根據上述之研究背景與動機，將利用個案公司網路銷售的顧客交易資料，用最近購買日期 (Recency, R)、購買頻率 (Frequency, F)、購買金額 (Monetary, M) 以及顧客關係長度 (Length, L) 等四個變數，經由分析層級程序法 (Analytic Hierarchy Process, AHP)，以專家團體決策方法決定 R、F、M、L 等四個顧客購買行為的權重，再以資料探勘之集群分析 (Cluster Analysis) 將顧客分群，並運用關聯法則 (Association Rules) 找出潛藏有用的規則，加上領域知識 (Domain Know-how) 的探討，提出行銷策略供決策之建議。因此，本研究的研究目的如下：

### 1.以 R、F、M、L 顧客購買行為變數，將顧客分群：

以歷史交易行為記錄，分辨出不同區隔的顧客群，並根據人口統計變數及交易行為找出各集群的特徵，作為客製化活動設計及溝通互動策略擬定的指導方針。

### 2.交叉銷售 (Cross Selling) 購物籃策略建議：

利用關聯法則找出顧客購物潛藏有用的規則，彙整分析後提出購物籃之設計。

### 3.結合領域知識提出相關策略：

針對不同集群客戶提出客製化行銷策略與銷售活動，提升與顧客溝通的效率，進而達到資源有效的分配。

## 1.4 研究流程

為使研究順利進行，首先依據研究背景欲探討的問題所衍生的動機、目的設定研究的主題，再經由個案公司相關人員探討實務操作面臨的問題，做為個案研究素材，進而研讀與研究相關之文獻，以建立研究架構，最後經由實驗驗證與實證分析的結果提出研究的結論與建議等。

## 2. 參考文獻

### 2.1 顧客關係管理

顧客關係管理的核心概念，在於為顧客創造全新的服務與價值，以贏得顧客長期忠誠度 (Stone, Woodcock, and Wilson, 1996; Athaide, Meyers, and Wilemon, 1996; Nan-Hong Lin, Hsin-Ting Ho, 2008)。在資訊科技不發達的時代，有顧客關係管理觀念的雜貨店，會透過認識顧客，長期建立穩定的顧客關係。透過關心顧客的每一筆交易，而非僅在發生交易時，才展現關注，並會適時的提供意見，提醒相關商品的購買，這就是透過與顧客建立良好的關係，達到與顧客互惠互利的結果。隨著資訊相關科技的發展，顧客關係管理藉由資訊系統來規範企業與顧客來往的互動行為與資訊，針對所有顧客進行分層區隔與差異化服務 (Nan-Hong Lin, Hsin-Ting Ho, 2008)。Swift (2001) 提出顧客關係管理流程有四項要素：與顧客互動、與顧客連結、對顧客瞭解、與顧客建立關係。顧客關係管理是一種長期的策略，透過與客戶的互動，瞭解客戶的特性與需求，進而滿足顧客，為顧客提供優質化的服務。利用資訊科技發達的優勢，以科學的方法，透過資料探勘演算法，將顧客分群，從不同群體中找出共同特性與購買行為，在資源有限的情况下，適當的分配與使用資源，和顧客建立長期的關係，提供差異與客製化的服務，滿足不同群體顧客的需求，以達到顧客關係管理的目的。

### 2.2 顧客價值分析模式

顧客關係管理和顧客價值是一體兩面的課題，Berger and Nasr (1998) 認為，顧客關係管理的核心是經由創造顧客滿意度以及更高的顧客價值，和顧客建立以及維持長期關係，並藉此建立忠誠度。許多學者發展出許多顧客價值分析模式，其中以RFM顧客價值分析模型符合本個案研究的資料與特性，若以此模型為基礎，加以衍生，足為衡量顧客價值作為後續研究分析之用，以下將介紹RFM顧客價值分析模型：

#### 2.2.1.RFM顧客價值分析模型：

進行顧客價值分析的工具，RFM是廣為運用的模型。Kaymak (2001) 指出目前企業廣為運用RFM評估消費者價值；根據Hughes and Arthur (1994) 的定義，RFM是指顧客購買行為的三個指標：最近購買日期 (R)、購買頻率 (F)、購買金額 (M)。並認為企業利用此三個指標可以描述出顧客的購買行為。

#### 2.2.2.R、F、M權重

RFM模型分析法主要以最近購買日期 (R)、購買頻率 (F)、購買金額 (M) 等三個顧客購買行為來衡量顧客的價值，然對於這三個變數各自的重要性，許多學者有不同的見解，以Stone (1995) 所提視產業特性差異給予RFM不同權重的觀點較符合本個案研究的方向。

#### 2.2.3.RFML模型及權重：

張心馨、蔡憲富 (2003) 將LRFM變數資料匯入探勘資料庫內，以提供分群作業所需，LRFM資料模型中模型乃修正自RFM模型，起因於Reinartz and Kumar (2000) 透過RFM無法區隔出長期與短期顧客關係的群體。其中「L」指的是顧客關係長度，代表顧客最近一次與第一次購買的時間長度。顧客與企業之顧客關係長度，是顧客價值建立的原因之一。因此參考張心馨、蔡憲富 (2003) 提出之顧客區隔出長期與短期顧客關係的群體考量，將顧客關係長度，運用Stone (1995) RFM模型，提出以RFML做為顧客價值衡量與顧客分群的變數。茲將R、F、M、L四個變數分別說明如下表2-1：

而在權重方面，Stone (1995) 認為應視產業特性的差異，給予RFM不同的權重。因此，吾人認為對於個案公司顧客價值以RFM為基礎，加上顧客關係長度 (L) 來的衡量，其權重應以產業的專業認知來評估，所以本文將運用分析層級程序法，透過專家集體決策來決定RFML權重。

表2-1RFML變數說明

變數	變數說明
最近購買日 (R)	顧客接觸時間，指顧客最後一次購買時間到分析日之間的日數。
購買頻率 (F)	顧客購買次數，資料期間顧客購買次數。
購買金額 (M)	購買金額，資料期間顧客訂購總金額。
關係長度 (L)	顧客關係長度，顧客最近一次與第一次購買的時間長度。

### 2.3 分析層級程序法

利用最近購買日 (R)、購買頻率 (F)、購買金額 (M) 以及顧客關係長度 (L) 等四個變數來衡量顧客價值及作為顧客分群的變數。這四個變數是顧客的購買行為，然而如何決策這四個變數之間的輕重？分析層級程序法 (Analytic Hierarchy Process, AHP) 提供一套運算分析系統，協助人們在面臨多個選擇時做出最佳的決策。經由AHP方法來確認RFML之間的權重，希冀藉此力求客觀，以提升研究的價值。

#### 2.3.1.分析層級程序法：

AHP為1971年匹茲堡大學教授Saaty所發展出來的方法，其屬於一種多目標的決策方法，將複雜且非結構化的問題系統化，將一系統分解為多個層級，由高層次往低層逐步分解，並經過量化的判斷，使得決策者處理多屬性決策問題時，在層次架構中釐清問題，簡化並改進以往依靠直覺的決策程序，求得各方案間的優先權重值，提供決策者選擇適當方案的充份資訊，凡優先權重值愈大的方案表示被採納的優先順序愈高，可降低決策錯誤的風險性 (鄧振源、曾國雄，1989)。

#### 2.3.2.AHP的評估尺度：

分析層級程序法評估尺度的劃分包括五項，即同等重要、稍重要、頗重要、極重要及絕對重要，並賦予名目尺度1、3、5、7、9的衡量值；另有四項介於五個基本尺度之間，並賦予2、4、6、8的衡量值。各尺度所代表的意義如表2-2評比尺度意義及說明所述。

表2-2 AHP評比尺度

估尺度	定義	說明
1	同等重要 (Equal Importance)	兩比較方案的貢獻程度 具同等重要性
3	稍重要 (Weak Importance)	經驗與判斷稍微傾向喜 好某一方案
5	頗重要 (Essential Importance)	經驗與判斷強烈傾向喜 好某一方案
7	很重要 (Very Strong Importance)	實際顯示非常強烈傾向 喜好某一方案
9	絕對重要 (Absolute Importance)	有足夠證據肯定絕對喜 好某一方案
2,4,6,8	相鄰尺度之中間值 (Intermediate values)	需要折衷值時。

資料來源:鄧振源、曾國雄 (1989)。

### 2.3.3.AHP步驟與運算

利用AHP法進行決策問題時，主要步驟如下

- A. 問題界定：界定問題的範圍，此階段需收集資訊及確認問題和方案。複雜問題時可利用層級結構分解，但基於人類無法對七種以上的事物同時進行比較。因此，每一層的要素最好不超過七個。
- B. 層級結構設立：利用階層分析方法，構建整個問題的層級結構。
- C. 問卷設計與調查：對每一成對比較設計問卷，在1—9尺度下，讓決策成員填寫。根據問卷調查所得到的結果，建立成對比較矩陣，再計算求各成對比較矩陣的特徵值與特徵向量，並檢定矩陣的一致性。
- D. 矩陣向量與特徵值計算：假設某層級內要素有A1、A2、A3.....An，建立其成對比較矩陣。
- E. 一致性檢定：根據AHP理論假設，成對比較應滿足偏好關係與強度關係的遞移性，然而，實際評估時，卻很難完全滿足該項假設。因此，Saaty認為必須就成對評比進行「一致性檢定」。

### 2.3.4.AHP決策權重架構

利用AHP進行RFML權重決策，將設立圖2-1之層級結構。以問卷的方式訪談個案公司四位管理階層，以最近購買日（R）、購買頻率（F）、購買金額（M）、關係長度（L）等四個構面，以AHP方法為工具，透過群體決策方式來決定RFML的權重。



圖2-1決策權重架構，修正自張心馨、蔡獻富（2004）

### 2.3.5.本研究問卷設計與調查

透過問卷訪談，內容為『就專業經理人對於顧客對於xx品牌在「最近購買」、「購買頻率」、「購買金額」、「關係長度」四大構面的相對重要性為何?請在以下適合的方格內打勾』。訪談領域專家，分別為行銷主管、銷售業務主管、財務主管共四位專家。

## 2.4 資料探勘

藉由資料探勘的技術，可以增進對顧客需求和行為的瞭解(Witten and Frank, 2005)，企業在經營運作的過程中，透過資訊科技工具的協助，運用資料庫管理系統，累積大量的經營管理相關資料，雖然有大量的資料紀錄，卻無法有效善用資料進行決策，並找出對企業有利的經營法則。為了提升競爭優勢，企業運用資料探勘技術，從過去儲存的歷史資料，經由演算法，從現存資料中剖析出新事實(Grupe and Owrang, 1995)，萃取出有用資訊(Cabena et al. 1997)，結合商業的領域知識，從儲存著大量資料的倉儲中進行挖掘，以發現資料間有意義的新關係、型態和趨勢(Berson et al., 2000)，提供了管理者經營決策方針。因此資料探勘儼然已經成為熱門的資料分析技術之一。以下就資料探勘的定義、功能提出介紹與說明。

### 2.4.1.資料探勘的定義：

Peacock (1998) 認為最廣義的資料探勘等同於知識發現 (Knowledge Discovery in Databases, KDD)。Fayyad (1996) 提出資料探勘是透過演算法，將資料作一分析與應用，以找出其特徵與模式 (Model) 的過程。茲綜合各學者的看法，吾人認為資料探勘的定義是以資料倉儲的能力為基礎，將轉換、清理、淨化過的大量資料，透過演算

法與統計分析，從中挖掘出未知、隱藏且有用的規則，再結合領域知識，提供決策管理者參考的一種過程。

#### 2.4.2. 資料探勘技術的方法：

資料探勘依不同研究目的而對應不同的分析方法，藉由演算規則，從大量資料中挖掘出有用的規則，成為知識的基礎，茲將這些方法及對應演算法說明如下：

- A. 分類 (Classification)：事先預測類別變數的過程稱為分類，從已知的類別變數，依其屬性，建立分類模式。例如，在促銷活動尚未開始前，預測客戶是否會因為產品搭售而加購。分類的資料探勘方法經常運用在預測顧客是否流失、客戶是否購買等。常用到的演算法包括類神經網路、決策樹等。
- B. 推估 (Forecasting)：根據已知的屬性，預測連續變數的過程稱為推估。可運用到如貢獻度預測、金融商品價格變化預測等。常用的演算法包括決策樹、模糊理論、類神經網路與迴歸分析等。
- C. 集群化 (Clustering)：綜合各項屬性根據相似狀況，將相似的事物聚集在一起稱為集群化，目的是找出各集群之間的差異，和群內的相似，也就是群內差異小，群外差異大。例如企業利用客戶的性別、年齡、收入等變數將顧客分群等。常用的演算法如K-means、自我組織地圖 (Self-Organization Map, SOM)。
- D. 關聯規則 (Association Rule)：找出那些事件總是相伴發生，稱為關聯規則或是購物籃分析，例如零售商想瞭解消費者總是會同時買哪些商品?就可以運用關聯規則來獲知答案。因此，商業團體常運用在如網頁結構分析、交叉銷售等。常用的演算法如Apriori、模糊集合 (Rough set) 等。
- E. 時序規則 (Sequential Pattern)：在關聯規則中可找出哪些事件會相伴發生，而時序規則可找出哪些事件發生的順序或週期，例如消費者買了A商品後30~36天會再次購買。時序規則也經常運用在網頁瀏覽順序分析。常用的演算法包括迴歸分析、類神經網路和時間數列分析等。

### 3. 研究方法

#### 3.1 研究架構

本研究欲利用個案公司會員網路購買交易資料，運用AHP專家訪談決策最近購買日 (R)、購買頻率 (F)、購買金額 (M)、關係長度 (L) 等變數的權重，透過資料探勘技術將顧客分群，探討各群的差異。再結合關聯分析之規則，依各集群提出行銷策略建議，供行銷決策制定之參考。本節依據前章文獻探討的AHP方法及資料探勘技術，進一步提出研究架構如圖3-1

- A. 資料表：將個案公司提供之客戶資料表、交易資料表、產品資料表等做適當的前置處理，包括變數代號的轉換，衍生變數的產生、資料適用性的檢查等，以確保資料的可使用性。
- B. 計算顧客價值：由AHP方法，運用專家決策調查，將RFML賦與權重，利用加權後的RFML參數，計算顧客價值。
- C. 顧客分群：以加權後的RFML四個變數，利用SQL Server 2008資料探勘技術的集群分析，以系統自行計算決定群數將顧客分群。
- D. 關聯分析：運用資料探勘技術的關聯分析，分別針對各集群客戶找出產品之間潛藏、有用的關聯規則。
- E. 行銷策略建議：運用關聯分析找出的規則，結合變異數分析 (Analysis Of Variance, ANOVA)，分析各集群在人口統計變數與產品購買行為的差異，提出各集群產品交叉銷售及行銷策略建議，以供行銷人員決策之參考。

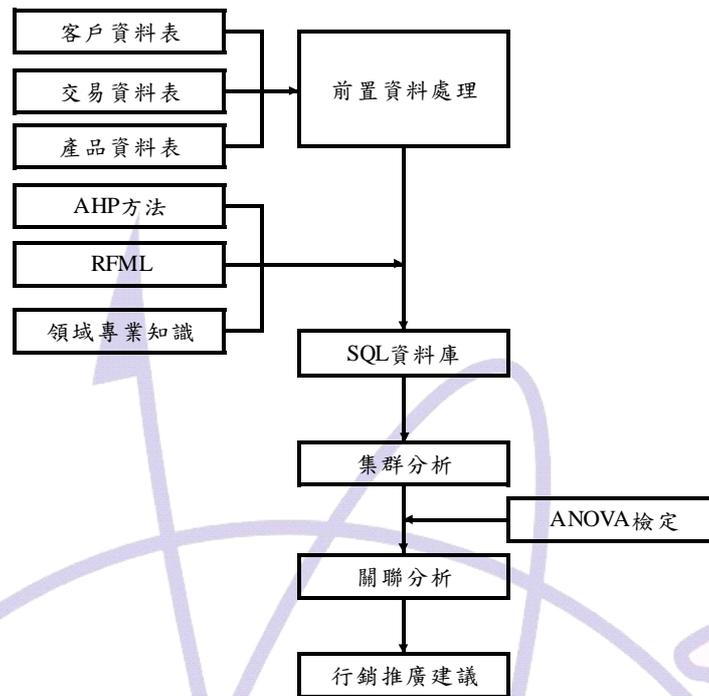


圖3-1 研究架構圖

### 3.2 資料來源

研究所採用的資料，是國內某天然家庭清潔用品會員網路購買交易資料，資料期間從2008年7月到2011年8月共38個月，其中包含商品資料表、會員資料表與交易資料表。交易資料表記錄交易資料，筆數共計31,988筆。因應研究範圍的需要，前述三個資料表需要進行前置處理，主要為名目資料轉換成數值及新的衍生變數產生，新的變數共為11個。透過以上三個原始資料表之欄位項目，重新整理與計算出新的欄位資料，並由交易資料表中的資訊加入在會員資料表中，包括縣市/居住區域/年齡/年齡層...等欄位。經整理後的會員筆數自12,054調整為有交易記錄的5,328筆，後續研究即根據此5,328筆進行分析，見表3-1。

表3-1 資料表說明

資料表	欄位資料說明	筆數
交易資料表	會員編號、訂單編號、訂單日期、商品名稱、訂購數量、訂單金額、折扣金額	31,988
會員資料表	會員編號、性別、出生年月日、縣市、居住區域、資訊來源分類、資訊來源、年齡、年齡層、購買金額、購買頻率、最近購買、關係長度	5,328
商品資料表	商品分類代號、商品分類、商品小分類代號、商品小分類、商品代號、商品名稱	66

研究所運用到的變數共計14個。尹相志（2010）提出Z分數技術除了可以用來定義極端值，還可以將數值標準化，在建立集群分析時，儘量於資料預處理階段利用Z分數進行標準化。因此，變數將採用Z分數標準化將最近購買日、購買頻率、購買金額、關係長度此四變數平滑化，其公式為：

$$Z\text{分數} = (\text{案例值} - \text{母體平均數}) / \text{母體標準差} \dots\dots\dots (1)$$

同時將這些變數的所屬資料表、變數說明、變數名稱、變數值、衡量尺度與變數特性，整理如表3-2變數屬性一覽表。後續研究將利用這些會員購買行為變數與基本資料，作相關的研究與分析，期從中瞭解並掌握有用的資訊。

表3-2 變數屬性一覽表

資料表	變數說明	變數名稱	變數值	衡量尺度	變數屬性
會員資料表	會員編號	No	代號值	名目尺度	質化變數
	性別	Sex	1 為男生、2 為女生	名目尺度	質化變數
	年齡	Age	年齡值	比率尺度	量化變數
	年齡層	Age_level	1 為 30 歲以下、2 為 31 歲~35 歲、3 為 36 歲~39 歲、4 為 40 歲~45 歲、5 為 46 歲以上	順序尺度	質化變數
	購買頻率	ZF	標準化加權值	比率尺度	量化變數
	購買金額	ZM	標準化加權值	比率尺度	量化變數
	最近購買	ZR	標準化加權值	比率尺度	量化變數
	關係長度	ZL	標準化加權值	比率尺度	量化變數
	居住區域	Area	1 為北部、2 為中部、3 為南部、4 為東部、5 為外島	名目尺度	質化變數
	折扣金額	Discount	金額值	比率尺度	量化變數
商品資料表	資訊來源分類	Infor	1 為電視媒體、2 為平面媒體、3 為廣播媒體、4 為網路媒體、5 為人員推廣、6 為實體通路、7 為其它	名目尺度	質化變數
	商品分類代號	Cato1	1 為衣物清潔、2 為個人清潔、3 為家用清潔、4 為廚房清潔、5 為組合類	名目尺度	質化變數
	商品代號	Prdct_no	各代號值	名目尺度	質化變數
	商品名稱	Prdct_name	各商品名	名目尺度	質化變數

### 3.3 AHP方法與RFML顧客價值模型

決定RFML之權重，採用AHP法群體決策的方式，其進行的方式與流程說明如下：

- A. **決定決策層級架構：**RFML權重決策的層級如圖2-1，其是經由最近購買日、購買頻率、購買金額、關係長度四個因素來決定，至於每一個變數的權重決定則透過專家訪談，四個變數兩兩比較重要性。
- B. **專家訪談：**透過表3-3問卷訪談領域專家，分別為行銷主管、銷售業務主管、財務主管共四位專家。
- C. **轉換矩陣表：**經向個案公司管理階層訪談，決定「最近購買」、「購買頻率」、「購買金額」、「關係長度」等四個變數兩兩相比後，建構成對比較矩陣表，以便後續權重的計算。
- D. **一致性檢定：**權重評估的結果要能通過一致性檢定，方能顯示填卷者的判斷前後一致，否則視為無效的問卷。Saaty建議以一致性指標與一致性比例來檢定成對比較矩陣的一致性。檢驗結果個案公司四位管理階層與整合後之CR均小於0.1，符合一致性檢定的標準

- E. **最後權重**：四位專家共同決策，最後整合的權重值分別為最近購買wtdr=0.119、購買頻率wtdf=0.468、購買金額wtdm=0.267、關係長度wtdl=0.146
- F. **RFML顧客價值模型**：四位管理階層決策出來的權重，每位顧客的RFML將乘上相對應之權重，作為分群輸入變數值，以此衡量顧客價值。

$$CLV = 0.119 * ZR + 0.468 * ZF + 0.267 * ZM + 0.146 * ZL \dots\dots\dots (2)$$

### 3.4 資料探勘方法與工具

#### 3.4.1. 演算法

##### A. E-M演算法：

E-M演算法是期望值最大化演算法 (Expectation Maximization) 的簡稱，是分群的方法之一。E-M演算法的開發，主要可改善K-means演算法分群的缺點 (尹相志，2010)。K-means演算法以距離作為相似性分群的依據，因為離案例最近的集群只能有一個，因此集群之間是相斥而不是重疊。若一個案例處在兩個相似的集群間時，只要一點偏差就可能跨越邊界，變成另一個集群的成員，這不穩定的狀況可能造成判斷成員行為模式的誤差。E-M是使用高斯分配 (Gaussian Distribution)，也就是常態分配來描述該案例隸屬於某集群的機率密度，利用此機率函數來取代集群的距離函數。

##### B. Apriori演算法

尹相志 (2010) 指出，Apriori是最普遍的關聯規則演算法，使用了物件組的前置統計量 (Prior Statistics)，包括：

- i. 最大規則物件數：規則中物件組所包含的最大物件數量。
- ii. 最小支援：規則中物件或是物件組必須符合最低案例數。
- iii. 最小信心水準：計算規則所必須符合的最低信心水準門檻。

Apriori演算法是根據資料明細作為基礎，遞迴計算後選的物件組及規則組合，流程如下：

- i. 演算法掃描過全部的紀錄：產生資料中所有包含單一物件數的所有組合 (Item Set)，計算各單一物件的案例數 (支援)。根據排列組合公式，如果有n種商品，應該可以產生n個物件組合。
- ii. 設定參數最小支援：最小支援是指一個規則或物件組出現在總資料中的頻率。如果物件組本身無法達到支援門檻，這個物件組的規則就一定也不能達到門檻。因此，在原來產生的物件組中，會把未達資源門檻的物件組刪除，只保留能達到門檻的物件組合，我們稱為候選大型物件組 (Candidate Large Item Set)。假設最小支援是2，系統會將支援不到2的物件組合予以刪除。
- iii. 再重新以排列組合原理計算候選的物件組：針對刪除後的物件組再重新以排列組合原理計算候選的物件組，如果物件本身都無法達到最小支援的門檻時，包含這些物件的物件組也不可能達到門檻。於是系統會根據上一步驟所過濾的物件產生兩個物件的物件組候選組合，並統計其支援，最後再根據最小支援進行過濾。
- iv. 產生出包含最大規則物件數+1個物件：一直反覆以上步驟，直到產生出包含最大規則物件數+1個物件的物件組為止。舉例來說，設定規則物件數最多為2時，則須要一直迴圈產生物件數為3的物件組為止。
- v. 中止物件組搜尋：物件組產生會根據最大規則物件數+1的門檻中止，或是如果在此步驟中所產生的符合最小支援物件組數為零時，系統也會中止物件組搜尋。
- vi. 將物件組排列組合：計算所有的可能規則，同時計算規則的支援、信心水準。
- vii. 過濾不符合最小支援以及信心水準的規則：此時即完成Apriori演算法的計算過程。

#### 3.4.2. 資料探勘工具

本研究透過SQL Server 2008資料探勘工具的集群分析將顧客分群，再利用關聯分析找出購物籃規則。

##### A. 集群分析

SQL Server 2008資料探勘集群分析是利用E-M演算法，設定探勘結構處理後，完成集群分析的模型處理，可透

過集群圖表、集群設定檔、集群特性檔等檢視器檢視結果。

## B. 關聯分析：

SQL Server 2008資料探勘關聯分析是利用Apriori演算法找出關聯規則，在眾多的規則中可利用以下指標找出有用的規則：

- i. **信心水準 (Confidence)**：信心水準顧名思義就是這條規格的準確度到底有多少，從條件機率的公式看來，等於在A事件發生的前提下，發生B事件的機率。因此公式可以表示如下：

$$\text{Confidence}(A \rightarrow B) = \frac{P(A \cap B)}{P(A)} \dots\dots\dots (3)$$

舉例來說，在某期間內購買A商品的顧客有70人，購買B商品的顧客有90人，而同時買A和B商品的顧客有50，則購買A也會去買B的信心水準為 $50/70=0.714$ 。信心水準愈高，自然這條規則就愈有參考價值（目前SQL Server 2008稱信心水準為「機率」）。此參數即是關聯規則中的最小信心水準，信心水準低於此數值之規則將會被過濾，例如最小機率为0.4，則機率低於0.4的規則會被過濾。

- ii. **支援 (Support)**：指的是符合這個規則的交易次數。支援的公式如下：

$$\text{Support}(A \rightarrow B) = P(A \cap B) \dots\dots\dots (4)$$

舉例來說，每個月的交易中有500次購買A同時又購買B，則支援就是500。信心水準固然重要，但發生的次數不多，則高信心水準就不具意義。

- iii. **最低重要性 (Importance)**：機率高不一定等於有意義的規則。我們必須要比較在有B以及沒有B的條件下，發生A條件的機率比例（依照貝氏機率），由於這個比例可能相當懸殊，因此透過開對數（以10為底數）的方式來取得重要性指標，其公式如下：

$$\text{Importance}(A \Rightarrow B) = \log\left(\frac{P(B|A)}{P(B|\text{not } A)}\right) \dots\dots\dots (5)$$

根據公式，當「在B中發生A條件的機率」高於「在非B的狀況下發A條件的機率」時，則該對數之後會大於零，且此指標愈大，則代表在A條件會造成B出現機率的顯著提升。如果此值為負值，則代表在A條件下越會抑制B的發生。當規則為正值時，SQL Server 2008重要性圖示會是藍色，如果是負值時，重要性規則會是紅色，規則重要性的排序是根據重要性指標絕對值進行排序。例如最低重要性為0.69，則顯示器只會出現重要性高於0.69的規則，其餘的將會被過濾。

- iv. **相依性網路**：圖3-2關聯規則相依性網路，與決策樹的相依性網路相似，只不過是改成呈現產品之間的關聯性，使用者可以透過圖形顏色了解產品間之相依性。圖中顯示購買玻璃清潔劑，會同時購買浴廁、烤爐清潔劑、衣領精、洗手乳與萬用清潔劑等。因此針對玻璃清潔劑作促銷，就可帶動很多品項的銷售。所以玻璃清潔劑可作為犧牲打商品促銷，藉以提高其它品項的購買機率，進而提高整體銷售。

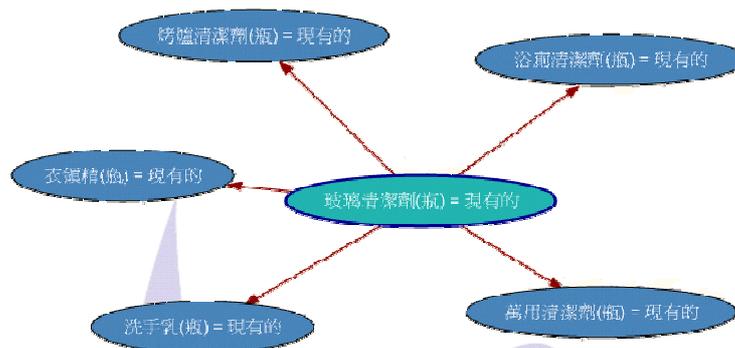


圖3-2關聯規則相依性網路

#### 4.實證結果與分析

本節首先說明利用最近購買日 (R)、購買頻率 (F)、購買金額 (M)、關係長度 (L) 等變數標準化後的數值，再根據群體AHP分析層級程序法，透過專家訪談決定上述變數加權之權重。以此四項加權後的變數，將所有購買顧客以集群分析加以分群，以瞭解各集群顧客其購買行為是否有差異，再根據分群後的每一集群特性予以命名，接著再針對各集群進行購物籃分析，進而提出商品交叉銷售之建議。

##### 4.1 顧客集群命名

經由SQLServer2008應用程式執行，群數由系統自行運算後的結果為十群，應為明顯區隔的分群結果。然在考量行銷資源的有限性，建議將此十群合併縮減。經以前章討論的顧客價值CLV為依變數，透過ANOVA檢定，以Scheffe法進行多重同質分組比較後，可將群數合併為六群，其顧客價值CLV大小依序為集群六 > 集群五 > 集群四 > 集群三 > 集群一 > 集群二。計算出RFML四個變數在六個集群的中心點，再由此中心點繪製成圖4-1群組特性圖。由群組特性圖可明顯的看出每一群的屬性，並將群組特性圖的解釋說明彙整為表4-1。以下將針對各群的屬性說明後，並分別提出命名的建議：

- A. **集群一**：檢視此群顧客，發現其購買頻率、購買金額、關係長度低，最近購買日高，表示此群顧客最近才和個案公司進行交易行為。因此，集群一可命名為「新進客戶群」。
- B. **集群二**：檢視此群顧客，發現其購買頻率、購買金額、關係長度低、最近購買日等均為低，表示此群顧客曾經向個案公司購買過，但購買次數與金額不高，關係長度也不高，且購買發生的時間距離現在已有一段時間。由此可知，此群客戶當時可能被某些原因的吸引，因而產生購買行為，但之後卻沒有產生回購行為。因此，可將集群二命名為「淺嘗即止客戶群」。
- C. **集群三**：檢視此群顧客，發現購買頻率中低、購買金額中低、關係長度低中低，最近購買高中高，表示此群顧客和個案公司有持續性商品交易，且在近期有重覆的購買行為，彼此的關係正在穩定的建立。因此將集群三的客户命名為「潛力客戶群」。
- D. **集群四**：此群顧客的購買頻率、購買金額、關係長度與最近購買均中等，相較於潛力客戶群，集群四客戶在購買頻率、購買金額與關係長度均較高，表示此群顧客已經和個案公司保有一段價值交換的行為；然而在最近購買變數中卻低於潛力客戶群，推論集群四的顧客在過去曾經和個案公司有頻繁的往來，但這些關係距今已有一段時日。由此可知，集群四的客戶對於個案公司的服務已經有所疏離。因此，集群四可命名為「預警客戶群」。
- E. **集群五**：此群顧客購買頻率中等、購買金額中等、關係長度中高，在六群中均排第二，最近購買日得分也高，表示此群顧客和個案公司的關係已有一段時日，且關係穩定成長。因此，集群五可命名為「黏著客戶群」。
- F. **集群六**：此群顧客購買頻率高、購買金額高、關係長度高，最近購買高，表示此群顧客長期以來和個案公司保

持良好關係，也是最有價值的客戶。因此，可命名為「忠誠客戶群」。

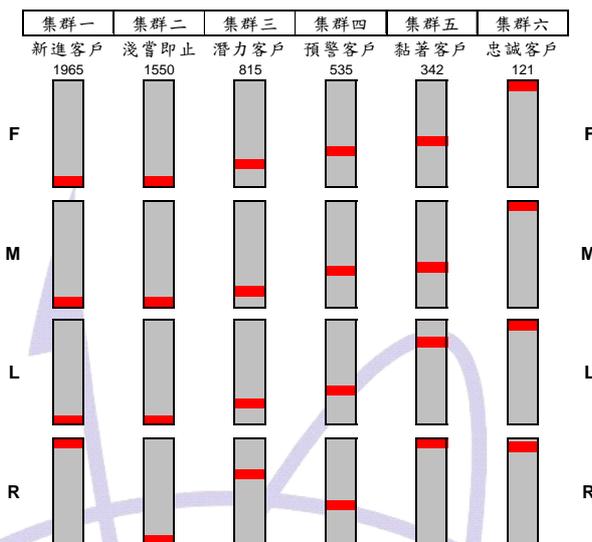


圖4-1群組特性圖

表4-1群組特性說明

群組	群名	說明	會員數量
1	新進客戶	購買頻率低、購買金額低、關係長度低、最近有新交易	1,965
2	淺嘗即止	購買頻率低、購買金額低、關係長度低、長期無交易	1,550
3	潛力客戶	購買頻率中低、購買金額中低、關係長度中低、近期有交易	815
4	預警客戶	購買頻率中、購買金額中、關係長度中、最近無交易	535
5	黏著客戶	購買頻率中、購買金額中、關係長度中高、最近有新交易	342
6	忠誠客戶	購買頻率高、購買金額高、關係長度高、最近有交易	121

#### 4.2 各集群敘述統計描繪：

經由SPSS敘述統計分析，針對各集群的樣本描繪，彙整於表4-2，並說明如下：

- 4.2.1. 新進客戶群：**人數共有1,965名，平均購買金額\$1,154元，為六群中最低；平均每人每次購買金額為1,154元，表示每人平均購買的頻率只有一次，證實其為新進顧客。新進客戶會以嘗試的心理購買試用，在購買的數量不會一次買太多，待有足夠的使用經驗及信心後才會增加購買數量。在性別方面，其中女性佔88.1%遠高於男性的11.9%。一般來說，女性是家庭花費的決策者及主要的購買者，也反應到個案公司的產品。居住區域以北部購買者的62%為最高，年齡層在39歲以下佔67%，推論為是產品的天然無毒訴求較吸引北部、39歲以下消費者。在訊息來源方面，42%的消費者透過電視廣告得知訊息。相較於其它客戶群，新進客戶在30歲以下佔了18%，明顯高於其它集群，表示此產品的使用者有年輕化的傾向。而個案公司曾經在2011年針對此品牌進行電視廣告，推測新進客戶受到電視廣告吸引而產生購買行為。透過電視廣告的訊息，可有效吸引年輕族群。因此，透過年輕族群的培養，並發展長期的關係長度策略。將建議個案公司可持續這方面的廣告投資。
- 4.2.2. 淺嘗即止客戶群：**人數共有1,550名，平均購買金額\$1,228元，平均每人每次購買金額為1,228元，與新進客戶群每人購買頻率一次雷同。但與新進客戶群不同的是：這群客戶的單一購買行為距目前時間較久，不像新進客戶為近期購買。女性佔89%為居多，居住區域中仍以北部購買者63.5%為最高；年齡層31~45歲佔77%；訊息來源部份，消費者透過網路媒體廣告的佔比為33.5%，為六群中最高。而在平面媒體也佔33.5%，僅次於忠誠客戶群。相較於電視媒體，平面媒體對於產品的宣傳可達到更深入的效果。網路媒介則著重於使用者的推薦。因此，此客戶群可多利用這兩個媒介達到深度溝通的效果。

表4-2 敘述統計彙整表

集群名		母體	新進客戶	淺嘗即止	潛力客戶	預警客戶	黏著客戶	忠誠客戶
人數		5,328	1,965	1,550	815	535	342	121
購買金額	人均金額(元)	2,490	1,154	1,228	2,448	5,275	6,083	18,193
	人次均金額(元)	1,402	1,154	1,228	1,224	1,453	1,654	1,862
	購買總額(仟元)	13,267	2,267	1,903	1,995	2,821	2,080	2,201
	百分比	100.0%	17.1%	14.3%	15.0%	21.3%	15.7%	16.6%
性別	男生	10.5%	11.9%	11.0%	8.5%	9.3%	7.6%	9.9%
	女生	89.5%	88.1%	89.0%	91.5%	90.7%	92.4%	90.1%
居住區域	北部	63.6%	62.0%	63.5%	62.7%	65.6%	70.5%	66.9%
	中部	15.4%	16.6%	15.4%	16.3%	12.5%	10.8%	14.9%
	南部	17.5%	17.9%	17.7%	17.3%	18.3%	15.2%	13.2%
	東部	2.8%	2.7%	3.0%	2.8%	2.4%	3.2%	2.5%
	外島	0.8%	0.8%	0.5%	0.9%	1.1%	0.3%	2.5%
年齡層	30歲以下	11.0%	18.3%	8.1%	7.4%	3.9%	5.8%	1.7%
	31歲~35歲	26.9%	29.7%	27.0%	26.3%	24.5%	20.2%	13.2%
	36歲~39歲	22.5%	18.6%	24.1%	23.7%	26.5%	26.6%	29.8%
	40歲~45歲	24.1%	19.9%	25.9%	24.5%	28.2%	28.9%	31.4%
	46歲以上	15.5%	13.5%	14.8%	18.2%	16.8%	18.4%	24.0%
資訊取得	電視媒體	22.1%	42.4%	5.9%	17.5%	11.6%	12.0%	5.8%
	平面媒體	22.7%	9.6%	33.5%	25.6%	29.9%	25.7%	36.4%
	廣播媒體	0.1%	0.2%	0.1%	0.0%	0.0%	0.3%	0.0%
	網路媒體	26.2%	19.7%	33.5%	27.1%	26.9%	26.0%	28.1%
	人員推廣	12.0%	10.9%	11.5%	12.1%	13.5%	17.0%	13.2%
	實體通路	8.5%	8.2%	7.8%	9.0%	9.7%	11.1%	7.4%
	其它	8.4%	8.9%	7.7%	8.6%	8.4%	7.9%	9.1%
購買品類	衣物清潔	42.0%	42.8%	31.8%	40.0%	41.3%	46.9%	48.2%
	個人清潔	18.2%	17.9%	20.1%	17.7%	19.3%	16.8%	16.9%
	家用清潔	16.1%	16.1%	18.8%	16.5%	16.5%	13.6%	15.6%
	廚房清潔	14.2%	16.5%	15.1%	15.2%	13.4%	12.8%	12.6%
	組合類	9.5%	6.7%	14.2%	10.5%	9.6%	10.0%	6.6%

**4.2.3.潛力客戶群：**人數共有815名，平均購買金額\$2,448元，每人每次購買金額為1,224元，表示每人平均購買頻率為2次。在性別方面，女性佔91.5%；居住區域以北部購買者62.7%為最高；年齡層在31~45歲佔74.5%。綜整以上訊息可知女性、居住北部，年齡在31~45歲為潛力客戶的主要購買者。得知訊息方面，除了平面與網路媒體共佔52.8%外，電視媒體也佔了17.5%，推論是個案公司在此品牌上市之初，仰賴電視購物銷售，成功吸引消費者對於此品牌的涉入，進而成為會員消費。

**4.2.4.預警客戶群：**人數共有535名，平均購買金額\$5,275元，每人每次購買金額為1,453元，每人平均購買頻率是3.6次，僅次於黏著與忠誠客戶。在性別方面，女性佔90.7%；居住區域以北部購買者65.6%為最高；年齡層在31~45歲佔79.3%。在平均購買金額與頻率均位居第三，僅次於黏著客戶與忠誠客戶，但在過去多次的購買行為，卻在近期停滯，個案公司應探討此客戶群近期不續購的原因，進而提出改善方案。

**4.2.5.黏著客戶群：**人數共有342名，平均購買金額\$6,083元，每人每次平均購買金額為1,654元，平均每人購買頻率為3.7次。在性別方面，女性佔92.4%；居住區域以北部購買者70.5%為最高，均為六群客戶群佔比最高；年齡

層在31~45歲佔75.5%。在訊息得知方面，透過人員推廣/實體通路的方式為六群中最高，表示這群客戶比較會因為面對面的推廣方式而引發首次的購買行為。黏著客戶群不論在購買頻率、購買金額、關係長度均僅次於忠誠客戶。因此，如何持續保有良好關係，提升顧客價值成為忠誠客戶為首要課題。

**4.2.6.忠誠客戶群：**人數共有121名，平均購買金額\$18,193元，平均每人每次購買金額為1,862元，代表每人的購買頻率為9.8次，為六群中最高。性別方面，女性佔90.1%；居住區域以北部購買者66.9%為最高；年齡層在36歲以上佔85.1%，其中46歲以上佔24%，高於其它各群。訊息得知部份，平面媒體佔36.4%，高於其它各群客戶，推論在於平面媒體的深入報導，教育了此群客戶，強化產品涉入程度。後續可透過平面媒體/DM郵寄方式溝通與教育，建立更深的認同感，同時誘使其對外推廣，培養成為潛在推廣人員。

**4.2.7.小結：**經過集群樣本的描繪，提出以下幾點建議：

- A. 顧客價值群組愈高的人數愈少。因此提升黏著客戶、忠誠客戶群將可有效創造個案公司的利益。
- B. 居住在中南部的消費族群人口比率相對少；換言之，這是開發新客戶的機會點。
- C. 新進客戶群年齡相對低，這應和個案公司產品電視廣告吸引了年輕族群有關。因此，要發展如何培養年輕消費群成為忠誠客戶的計劃，因為相較於年長者，其和個案公司的關係長度可以維持更長久。
- D. 衣物清潔類產品銷售佔比高，建議可利用衣物清潔類，搭售其它品類產品，除了擴展消費者產品線的使用，進而可帶動整體銷售。

### 4.3 關聯規則

資料探勘關聯規則，主要針對個別顧客集群交易資料中的商品品項找出彼此之間的關聯性，作為未來商品交叉銷售之參考依據。除了在規則相關參數篩選設定有兩個重要指標，分別為最小機率與最低重要性的條件設立過濾規則外，本節另外針對個案公司促銷活動所引發非潛藏規則的部分也一併提出說明如下：

#### 4.3.1.規格的適用篩選

- A. **最小機率：**為關聯規則中最小信心水準，關聯分析設定最小機率為0.4，低於此數值的規則將會被剔除。
- B. **最低重要性：**此指標大於0代表規則愈顯著，小於0則有抑制作用，最低重要性設定為0.76，低於此數值的規則將會被剔除。
- C. **支援 (Support)：**指的是符合這個規則的交易次數，若交易的次數不高，就算機率與重要性高，這規則對於購物籃設定並無幫助。因此，本文將支援設為75，若符合這個規則的交易次數不到75，將過濾不計。
- D. **非潛藏有用規則：**
  - i. 個人清潔用品：因個案公司對於這個品類已不再生產，促銷活動以出清庫存量為優先考量。因此所產生之相關規則不予參考。
  - ii. 各樣式組合類：此為季節性短期組合包裝促銷，一次性的搭售正常品促進販賣，售完為止。

個人清潔/組合類的促銷活動所產生的規則是可預見且預期會發生的結果，而非潛藏、有用的規則。這些規則資訊對於後續策略決策的擬定並無幫助，因此忽視不計。

依前述之最小機率及最低重要性之門檻設定及關聯規則忽略不計部份後，總計規則依六個集群數量共129條，每集群之對應數量如下表4-3所示。依上列之條件篩選後，整理每集群機率高前5名規則於表4-4說明如後。

表4-3 關聯規則數量表

群名	規則數量	群名	規則數量
新進客戶	17	預警客戶	33
淺嘗即止	25	黏著客戶	12
潛力客戶	22	忠誠客戶	20
總計	129		

表4-4 各集群關聯規則

客戶群	規則代號	支援	機率	重要性	規則
新進客戶	新-01	176	0.54	0.93	浴廁清潔劑(瓶)->碗盤洗潔精(瓶)
新進客戶	新-02	167	0.53	0.92	洗衣精(瓶)->碗盤洗潔精(瓶)
新進客戶	新-03	138	0.51	1.03	漂白粉(瓶)->洗衣粉(大盒)
新進客戶	新-04	137	0.51	1.10	漂白粉(瓶)->去污粉(罐)
新進客戶	新-05	142	0.50	0.87	烤爐清潔劑(瓶)->碗盤洗潔精(瓶)
淺嘗即止	淺-01	98	0.59	0.93	去污粉(罐)，碗盤洗潔精(瓶)->洗衣粉(大盒)
淺嘗即止	淺-02	151	0.59	0.98	漂白粉(瓶)->洗衣粉(大盒)
淺嘗即止	淺-03	123	0.57	0.92	萬用清潔劑(瓶)，碗盤洗潔精(瓶)->洗衣粉(大盒)
淺嘗即止	淺-04	210	0.55	0.99	去污粉(罐)->洗衣粉(大盒)
淺嘗即止	淺-05	107	0.54	0.89	去污粉(罐)，萬用清潔劑(瓶)->洗衣粉(大盒)
潛力客戶	潛-01	76	0.61	0.93	漂白粉(瓶)，去污粉(罐)->洗衣粉(大盒)
潛力客戶	潛-02	103	0.60	1.65	冷洗精(瓶)->洗衣精(瓶)
潛力客戶	潛-03	89	0.59	0.93	去污粉(罐)，碗盤洗潔精(瓶)->洗衣粉(大盒)
潛力客戶	潛-04	82	0.59	0.93	去污粉(罐)，碗盤洗潔精(瓶)->洗衣粉(大盒)
潛力客戶	潛-05	147	0.59	0.98	漂白粉(瓶)->洗衣粉(大盒)
預警客戶	預-01	102	0.77	0.92	漂白粉(瓶)，去污粉(罐)->洗衣粉(大盒)
預警客戶	預-02	92	0.74	0.90	漂白粉(瓶)，衣領精(瓶)->洗衣粉(大盒)
預警客戶	預-03	91	0.73	0.89	漂白粉(瓶)，碗盤洗潔精(瓶)->洗衣粉(大盒)
預警客戶	預-04	88	0.73	0.89	衣領精(瓶)，碗盤洗潔精(瓶)->洗衣粉(大盒)
預警客戶	預-05	80	0.72	0.88	浴廁清潔劑(瓶)，去污粉(罐)->洗衣粉(大盒)
黏著客戶	黏-01	171	0.67	1.05	去污粉(罐)->洗衣粉(大盒)
黏著客戶	黏-02	145	0.67	1.02	漂白粉(瓶)->洗衣粉(大盒)
黏著客戶	黏-03	132	0.60	0.96	衣領精(瓶)->洗衣粉(大盒)
黏著客戶	黏-04	85	0.57	1.48	烤爐清潔劑(瓶)->浴廁清潔劑(瓶)
黏著客戶	黏-05	148	0.57	0.95	萬用清潔劑(瓶)->洗衣粉(大盒)
忠誠客戶	忠-01	98	0.82	1.02	去污粉(罐)，碗盤洗潔精(瓶)->洗衣粉(大盒)
忠誠客戶	忠-02	210	0.78	1.12	去污粉(罐)->洗衣粉(大盒)
忠誠客戶	忠-03	79	0.76	0.96	浴廁清潔劑(瓶)，萬用清潔劑(瓶)->洗衣粉(大盒)
忠誠客戶	忠-04	123	0.76	0.98	萬用清潔劑(瓶)，碗盤洗潔精(瓶)->洗衣粉(大盒)
忠誠客戶	忠-05	83	0.74	0.96	浴廁清潔劑(瓶)，碗盤洗潔精(瓶)->洗衣粉(大盒)

#### 4.3.2.規則說明

##### A. 購物籃說明

以規則代號新-01為例，關聯內容為『浴廁清潔劑(瓶)->碗盤洗潔精(瓶)』，重要性高達0.93，發生的機率為0.54，支援為176次。表示購買浴廁清潔劑(瓶)後會正向增強(重要性0.93)同時購買碗盤洗潔精(瓶)的可能，機率

有54%，且支援這規則的次數有176次。再以規則代號忠-01為例，關聯內容為『去污粉(罐)，碗盤洗潔精(瓶)->洗衣粉(大盒)』重要性高達1.02，發生的機率為0.82，支援為98次。表示購買去污粉(罐)、碗盤洗潔精(瓶)後會正向增強(重要性1.02)同時購買洗衣粉(大盒)的可能，機率高達有82%，且支援這規則的次數有98次。個案公司可透過這些關聯規則依集群設計購物籃。

## B. 交叉銷售策略訂定之參考：

經過以上的說明刪除並非有用、潛藏的規則後，由每集群中機率高前5條規則看出：新進客戶群，主要以碗盤洗潔精出現在規則內的次數較高，共有3次；其它淺嘗即止、潛力客戶、預警客戶、黏著客戶與忠誠客戶等均以洗衣粉出現次數最高，表示家庭清潔用品銷售以衣物清潔為主。這些規則可供個案公司對於後續交叉銷售策略訂定之參考，在購物籃的商品規劃上可以衣物類產品搭配其它品類共同銷售推廣，以帶動整體產品的銷售。

## 5. 結論與建議

本研究運用修正RFM顧客價值分析模型，加入顧客關係長度(L)以R、F、M、L計算顧客的價值，並使用其為分群變數，成功的把顧客分為「新進客戶」、「淺嘗即止客戶」、「潛力客戶」、「預警客戶」、「黏著客戶」、「忠誠客戶」等六群。接著依據各群的購物籃及人口統計變數，分析出各集群之間的異同，再利用資料探勘關聯分析、找出購買產品的關聯規則，最終透過規則的適用篩選，訂出各集群的交叉銷售的策略。本章將彙整前章實証分析得出的結果，提出以下的結論與建議。

### 5.1 結論

本文旨在探討與顧客溝通、互動、與顧客建立關係，並提供滿足顧客需求的服務，進而達到顧客滿足、企業獲利的目的。而資料探勘的過程主要是找出不同特性的顧客群，讓管理者可以依據這些不同的特性，提出不同的溝通方式與內容來滿足顧客，做到資源分配與使用最佳化的境界，因此吾人提出以下建議供個案公司行銷決策者制定行銷策略之建議。

#### 5.1.1. 分群行銷策略訂定：

依據各群不同的購買行為特性及問題提出行銷策略如下：

##### A. 新進客戶群：

- i. 背景說明：新進客戶群的購買頻率、購買金額、關係長度低，最近購買高，表示此群顧客最近才和個案公司進行交易行為，因此對於此群客戶，首要任務即是吸引其再購買。
- ii. 行銷策略：增加近期再購意願。
- iii. 行銷目的：吸引新進客戶近期再次購買，避免成為淺嘗即止客戶，進而提升為潛力客戶。
- iv. 執行作法：經由郵件通知此集群客戶在期限內再購促銷的折扣或贈品贈送，以增加再購意願。同時在產品組合中，以電視廣告商品為藍心，設計購物籃，因新進客戶群其訊息來源中，有42%的顧客是透過電視廣告得知訊息。因此透過此方式應可以提高其涉入程度，應可提升再次回購機率。

##### B. 淺嘗即止客戶群：

- i. 背景說明：淺嘗即止客戶群的購買發生的時間距今已有一段時間，當時可能因為某些原因的知覺吸引，而產生購買行為，但之後卻沒有再次回購。
- ii. 行銷策略：增加回購意圖。
- iii. 行銷目的：淺嘗即止客戶為買了一次即不再回購，因此增加其回購意圖，吸引其再度消費。
- iv. 執行作法：此客戶群透過網路媒體得知訊息的佔比為33.5%，平面媒體佔33.5%，此兩媒體溝通管道的特性為深入溝通，表示深入溝通可以增加其涉入。因此，透過產品電郵DM寄送，提高對產品功能深入的了解，並搭配產品折價優惠刺激回購意願。

### C. 潛力客戶群：

- i. 背景說明：潛力客戶群顧客已經和個案公司進行商品交易，並在近期發生重覆的購買行為，表示客戶對過去購買的產品有某種程度上的滿意度。因此應該增加其購買產品的多樣化，藉此提高購買單價，並擴大其使用產品的多樣性。
- ii. 行銷策略：提高產品使用的廣度。
- iii. 行銷目的：透過其產品使用的廣度，進而提高購買單價。
- iv. 執行作法：透過產品組合的設計（原使用品項+新品項的搭配），提高單次購買交易單價，並增加不同商品的使用機會，擴大對商品的使用經驗，並累積其滿意度。

### D. 預警客戶群：

- i. 背景說明：預警客戶群顧客和個案公司保有較長期的價值交換的行為，然而在最近購買變數中卻低於潛力客戶群，表示此群顧客在過去曾和個案公司有頻繁的往來，但這些關係距今已有一段時日，由此可知，此群顧客對於個案公司的服務已經疏離。因此，瞭解這群客戶內心潛在想法是首要任務。
- ii. 行銷策略：釐清顧客問題，擬定改善策略。
- iii. 行銷目的：設計問卷調查，瞭解預警客戶群為何在多次購買後不再續購，進而提出改善策略。
- iv. 執行作法：以問卷方式了解預警客戶未持續購買的原因（例如：使用後的認知與期望的落差），並以贈品及購物折扣卷等作為提升問卷回覆率之誘因，並增加其再次購買的機率。

### E. 黏著客戶群：

- i. 背景說明：此群顧客群和個案公司有較長期的關係，在購買金額與頻率僅次於忠誠客戶群。因此關係的維護與提升成為忠誠客戶群為首要任務。
- ii. 行銷策略：升級為忠誠客戶。
- iii. 行銷目的：提高購買頻率與金額、提供誘因以吸引成為忠誠客戶。
- iv. 執行作法：提供購買累積金額誘因活動、傳遞成為忠誠客戶之誘因與好處，並提醒達成忠誠會員之門檻差距，與成為忠誠客戶群的福利。

### F. 忠誠客戶群：

- i. 背景說明：忠誠客戶群長期以來和個案公司保持良好關係，也是公司最有價值的客戶。因此公司應塑造其為備受禮遇的顧客，培養其尊榮之地位，並培養忠誠會員作為公司的潛在開發者。
- ii. 行銷策略：提供尊榮回饋，持續培養其忠誠度。
- iii. 行銷目的：讓忠誠客戶群的成員備受尊榮，提升忠誠客戶成員的涉入程度，並藉其吸引其他客戶群成為忠誠客戶群的一員。
- iv. 執行作法：消費紅利積點激勵、忠誠會員推薦的種子成員、新會員推薦折扣、會員獨享優惠活動、優先參與活動、新產品試用與年終回饋。

#### 5.1.2. 分群客製化行銷購物籃推廣

本研究接續根據資料探勘關聯分析，參考領域知識，提出各群之購物籃銷售促銷。透過關聯分析找出關聯規則，分別針對六大集群進行購物籃設計，見表4-4。個案公司可分別於各集群提出交叉銷售購物籃，提升整體銷售量。

#### 5.1.3. 分群行銷網頁規劃

將以客製化的方式執行新的行銷策略與銷售推廣活動，因此在網頁的設計規則如下：

- A. 分群網頁：會員經由其帳號密碼登入購買網頁後即進入該顧客群特有的行銷推廣活動網頁，載明其客製化的行

K.U.A.S.

銷活動方式。

- B. 推薦商品提醒：除了合購購物籃的設計外，單品點選的購物順序設計應參照各群的購物籃規則。設計點選商品後，頁面推薦其它商品，推薦規則來自於各群的關聯規則。

#### 5.1.4. 媒體投資：

天然、無毒清潔用品的使用方法與消費者購買意願須作深入溝通，經由平面媒體與網路報導而涉入的顧客可以保持較長的關係，這由忠誠客戶群的訊息來源有64%是來自這兩個媒體可看出其端倪。因此，藉由平面媒體與網路廣告深入報導的特性，加強此投資，可強化涉入程度，成為忠誠客戶。

電視廣告可有效吸引年輕族群成為新進客戶。因此持續投入此資源，相較於年長客戶，年輕族群的使用年數也較高。因此，吸引年輕新進客戶購買的同時也須一併思考發展「長期關係」的策略。

#### 5.1.5. 機會點的掌握：

經由各集群內外間的差異分析可觀察出幾個機會點，以下針對這些機會點掌握供個案公司參考。

- A. 中南區消費族群：相較於北區，中區及南區的顧客佔比相對小，表示有很大的機會點，可設計中、南區新客戶的促銷方式提升此地區的來客數。例如，只要是身分證字號第一碼為「T」，即可享有買就送的贈品，諸如此類。
- B. 品類購買：衣物清潔類在各集群的購買比重，除了淺嘗即止客戶群為31.8%，其它五個集群都在40%以上，這佔比和整體市場一致。個案公司可加強在非衣物清潔的品類，多以衣物清潔類品類結合帶動其它品類方式促銷。可將購物籃規則衍生，如規則「忠-03，浴廁清潔劑（瓶）、萬用清潔劑（瓶）、洗衣粉（大盒）」可加其它品項成「浴廁清潔劑（瓶）、萬用清潔劑（瓶）、洗衣粉（大盒）、烤爐清潔劑（瓶）、去污粉（罐）」，並在價格上給予誘因。
- C. 年輕族群：35歲以下顧客和個案公司關係長度較短，由忠誠客戶中35歲以下顧客佔14.9%比率、新進客戶佔48%比率即可看出，如何培養年輕族群成為忠誠客戶，值得深思。但長期而言，仍應持續培養年輕族群，以累積潛在使用者。

## 5.2 未來研究建議

本研究對後續的相關研究，提出下列方向作為日後研究的建議

- A. 客戶終身價值：在衡量顧客價值，僅就RFML四個變數之模型衡量，若能考量相關成本及時間價值，對於客戶價值衡量之研究應會有正面的貢獻。
- B. 資料探勘應結合商業智慧，自動化：資料探勘技術結合商業智慧，以自動化的模式結合領域知識，這方面的研究對於實施顧客關係管理的企業有莫大助益。
- C. 研究對象的擴展：本文研究的對象僅限於網路購物會員資料，研究對象若能擴展到實體通路購物者的研究，對於個案公司產品購買者的行為輪廓可以更有效的掌握。
- D. 研究產品的擴展：供應商受限於通路商的條件限制，必須付出非常大的代價才能將產品上架販售給消費者，雖然許多供應商已經透過顧客關係管理，跳過通路商直接和消費者交換價值。然而礙於產品線的不足，往往無法滿足消費者的需求。因此，建議個案公司可針對會員需求進行研究，找出適合之相關居家用品，並結合其它業者，運用現有會員購買機制，提供多元的產品滿足會員。
- E. 時序分析：可增加時序分析，找出時序的規則，訂出購買品項先後的次序，作為垂直銷售策略之建議。

## 6. 參考文獻

1. 丁一賢、陳牧言 (2003)。資料探勘。滄海書局。
2. 張力元 (2002)。顧客服務管理。華泰書局。
3. 尹相志 (2010)。SQLServer2008 Data Mining。悅知文化。
4. 張心馨、蔡憲富 (2004)。以 Data Mining 技術結合 SOM 和 K-Mean 的消費者分群方法於顧客關係管理和績效獲利性評估。資訊管理學報，11 (4)，77-104。
5. 鄧振源、曾國雄，層級分析法 (AHP) 的內涵特性與應用 (上) (1989)。中國統計學報，27 (6)，5~22。
6. World panel. (2012). [http://www.tns-global.com.tw/market\\_sectors/Default.aspx](http://www.tns-global.com.tw/market_sectors/Default.aspx)。2012 年台灣家庭清潔用品調查。
7. Kotler, P. (1994). Marketing management-analysis, planning, implementation, and control, (8th Ed). Englewood Cliffs,; Pren-tice-Hall.
8. Fayyad, U. M., Shapi, G. P., Smyth, P. & Uthursamy, R. (1996). Advances in Knowledge Discovery and Data Mining. Menlo Park, CA: AAAI Press/The MIT Press.
9. Stone, M., Woodcock, N., and Wilson, M. (1996). Managing the Change from Marketing Planning to Customer Relationship Management, Long Range Planning, 29(5), 675-683.
10. Athaide, G. A., Meyers, P. W., and Wilemon, D. L. (1996). Seller-Buyer Interactions During the Commercialization of Technological Process Innovations, The Journal of Product Innovation Management, 13(5), 406-421.
11. Nan-Hong Lin, Hsin-Ting Ho (2009). The Study on Customer Value and Customer Relationship Management Performance: A Customer-Based Perspective from the Banking Industry in Taiwan, Journal of Customer Satisfaction, 5(2), pp1-36.
12. Swift, R. S. (2001). Accelerating Customer Relationships: Using CRM and Relationship Technologies, 25-26, Englewood Cliffs, NJ: Prentice-Hall.
13. Berger, P.D., and Nasr, N.I. (1998). Customer Lifetime Value : Marketing Models and Applications. Journal of interactive Marketing, 12, 17-29
14. Kaymak, U. (2001). Fuzzy Target Selection Using RFM Variables, IFSA World Congress and 20th NAFIPS International Conference, 25-28.
15. Hughes, Arthur M. (1994). Strategic Database Marketing. Chicago: Probus Publishing.
16. Reinartz, W.J. and Kumar, V.(2000). On the Profitability of Long-Life Customer in a No contractual Setting: An Empirical Investigation and Implications for Marketing. Journal of Marketing, 64, 17-35.
17. Stone, B. (1995). Successful Direct Marketing Methods. Lincolnwood. IL: NTC Business Books, 37-57.
18. Witten, I. H. and Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques, 2nd ed., Elsevier, San Francisco.
19. Grupe, F. H. and Owrang. (1995). M. H. "Data Base Mining Discovering New Knowledge and Cooperative Advantage," Information System Management, 12(4), 26-31.
20. Cabena, P., Hadjinian, P. O., Stadler, R., Verhees, J. & Zanasi, A.(1997). Discovering Data Mining From Concept to Implementation. New Jersey: Prentice-Hall, Inc., 41-59.
21. Berson, A., Smith, S and Therling, K. (2000). Building Data Mining Application for CRM, McGraw-Hill, New York.
22. Peacock, P. R.(1998). Data mining in marketing: Part1. Marketing Management, 6(4), 8-18.