

東 海 大 學

工業工程與經營資訊學系

碩士論文

資料採礦於顧客關係管理之應用
-以建材零售業為例

研 究 生：施旻慧

指 導 教 授：洪堯勳 教授

胡坤德 教授

中 華 民 國 一 〇 二 年 六 月

**Applying Data Mining to Customer Relationship
Management
-A study on Construction Materials Retail Business**

By
Min-Hui Shih

Advisors : Prof. Jau-Shin Hon
Prof. Kun-Te Hu

A Thesis
Submitted to the Institute of Industrial Engineering and Enterprise
Information at Tunghai University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in
Industrial Engineering and Enterprise Information

June 2013
Taichung , Taiwan , Republic of China

應用資料採礦於顧客關係管理

-以建材零售業為例

學生：施旻慧

指導教授： 洪堯勳 教授
胡坤德 教授

東海大學工業工程與經營資訊學系

摘 要

由於零售業者面臨動態的競爭環境，再加上消費者對產品需求多樣化的消費模式，因此顧客關係管理逐漸受到企業重視。過去的傳統資料處理方式已跟不上現在的速度，利用資料採礦的方式來做資料的處理及分析，可以有效的發掘出資料隱藏著看不見的重要訊息，將資訊做更大價值的發揮。

本研究以建材零售業為研究對象，針對購買建材的廠商與需要做裝潢的顧客，企圖了解是否能透過結合 80/20 法則與關聯法則提升忠誠顧客比例，以及透過關聯法則分析顧客基本資料與歷史交易資料，了解顧客的購買行為進而為顧客量身訂作客製化產品及服務，進而與顧客建立良好的互動關係。

本研究首先將資料採礦與顧客關係管理之概念結合，以建立出研究之架構。其次，將初始的資料透過資料轉換與定義，轉換為使用者需要的結構化的資料，使分析者與系統皆能判讀。利用二階層群集分析來對於顧客做分群，分群出 80/20 兩群顧客以及流失與未流失兩群顧客。最後透過資料採礦的關聯法則，採用 Apriori 演算法萃取出對於企業有意義的法則，同時根據法則的判讀做出解釋並提供建議及因應的方式，可作為經營決策者制定行銷策略的參考。

關鍵字詞：資料採礦、二階層集群分析、關聯法則、顧客關係管理、建材零售業

Applying Data Mining to Customer Relationship Management

-A study on Construction Materials Retail Business

Student : Min-Hui Shih

Advisors : Prof. Jau-Shin Hon

Prof. Kun-Te Hu

Department of Industrial Engineering and Enterprise Information
Tunghai University

ABSTRACT

As retailers face dynamic competition and continuous consumer's demands for products consumption modes of diversification. Therefore, Customer Relationship Management (CRM) is gaining more and more attention in the enterprise. The traditional ways of data processing is out of date; and have been replaced with Data Mining which is used to. not only extrapolate data but also reveals the meaning behind the raw data collected in such a manner the information can be interpreted and utilized.

This research focuses on the study of building materials retail business, from the perspective of the manufacturers who purchase of the construction materials and the customers who need to do the interior decorating. Attempt to understand the Pareto Principle whether through a combination of the 80/20 rule and association rules to enhance the proportion of loyal customers. As well as through the association rule to analysis the basic information of customer and the transaction data to understand the customer's buying behavior and provide customers tailor-made products and services. And establish a good interaction relationship with customers.

First, we will be combining the concept of data mining and customer relationship management in building the main structure of this study. Transforming the raw data obtained through data switching and data identification into materials which are acceptable by both analysts and the system. Using the Two-Steps Cluster to categorize the customer and then grouping of 80/20 into two groups and the customer churn. Finally, through the Association Rule in data mining, we will calculate the most significant rule for this company by Apriori algorithm. Meanwhile we also provide the explanation of the rule, suggestions and solutions as well as derive an optimal marketing strategy.

Keywords : Data Mining, Two-Step Cluster, Association Rule, Customer Relationship Management, Construction Materials Retail Business

致謝詞

兩年的研究所生涯終於接近了尾聲，這兩年裡，首先要感謝兩位指導教授，洪堯勳主任以及胡坤德老師，除了在學習上鼓勵我多嘗試不同領域外，也鼓勵我多去戶外運動，這樣才能讓頭腦與身體保持健康與清醒，經過兩位老師的細心指導，讓我能更順利的完成論文。

接著要感謝口試委員陳建良老師、周瑛琦老師的指導，對論文詳加閱讀並給予許多建議，使論文的品質得以提昇。在學期間，也承蒙黃欽印老師、鄭辰仰老師在學業上提供寶貴的建議與知識，讓我能更順利完成論文。

在此也感謝個案公司總經理鐘義雄、業務經理洪茉莉以及會計主任陳靖雯所給予的協助與支持，使本研究得以完成。另外，也要感謝在碩士的兩年研究生涯裡，白偲函以及學弟賴彥舟陪著我一起奮鬥、一起研究程式，我永遠不會忘記。謝謝研究室的同學晉韶、泰羽、家佑、學弟宗緯、維誠以及勝彥，有你們熱心的幫助，使我在口試前更能心無旁騖的專心應試。

感謝我的媽媽及妹妹在這段時間給予無條件的支持與鼓勵，沒有他們也不會有今天的我。更感謝系排的學弟妹們，在這段時間的陪伴與鼓舞，這些力量皆為我的後盾與支柱。

最後，僅以這本論文獻給過世的爸爸以及我最愛的家人與所有朋友，謝謝你們。

旻慧 謹誌於

東海大學工業工程與經營資訊研究所

價值鏈創新與管理研究室

中華民國一〇二年七月

目錄

摘要.....	i
ABSTRACT.....	ii
致謝詞.....	iii
目錄.....	iv
表目錄.....	v
圖目錄.....	vi
第一章 緒論.....	1
1.1 研究背景與動機.....	1
1.2 研究問題與目的.....	4
1.3 研究流程.....	5
第二章 文獻探討.....	7
2.1 顧客關係管理.....	7
2.2 資料採礦.....	16
2.3 本章小結.....	26
第三章 研究方法.....	27
3.1 資料收集.....	27
3.2 研究工具.....	27
3.3 研究架構.....	27
3.4 二階段分群法.....	28
3.5 關聯法則.....	31
第四章 實證分析.....	37
4.1 資料採礦.....	37
第五章 結論與建議.....	53
5.1 結論.....	53
5.2 未來研究與建議.....	55
參考文獻.....	56

表目錄

表 1.1 2008~2012 年建材零售業銷售額於零售業所占比重	1
表 1.2 2008~2012 年建材零售商之家數	2
表 1.3 2011 年目前經營上遭遇的困境.....	2
表 1.4 2011 年所採取的經營策略.....	2
表 2.1 顧客關係管理之定義.....	10
表 2.2 資料挖掘實例.....	22
表 2.3 國內資料採礦相關研究整理.....	24
表 4.1 顧客明細.....	38
表 4.2 產品明細(B2B).....	38
表 4.3 交易明細.....	40
表 4.4 群數凝聚過程.....	44
表 4.5 風險顧客符合忠誠顧客要素比例表.....	50

圖目錄

圖 1.1 忠誠顧客的概念.....	3
圖 1.2 研究流程圖(本研究整理).....	5
圖 2.1 市場與銷售之 80/20 法則	8
圖 2.2 顧客關係管理架構主要應用元件.....	9
圖 2.3 顧客關係管理的循環過程.....	11
圖 2.4 顧客關係管理之運作流程.....	13
圖 2.5 Knowledge Discovery in Database Process.....	17
圖 2.6 資料採礦流程.....	21
圖 3.1 研究架構流程.....	28
圖 3.2 高頻項目組之產生過程(第一次掃描).....	35
圖 3.3 高頻項目組之產生過程(第二次以後的掃描).....	36
圖 3.4 最後所得高頻項目組.....	36
圖 4.1 資料採礦流程圖.....	37
圖 4.2 Access 資料表	41
圖 4.3 Excel 資料表	41
圖 4.4 訂購日期轉換.....	43
圖 4.5 交易金額轉換.....	43
圖 4.6 分群結構表.....	45
圖 4.7 Excel 檢定表	45
圖 4.8 Weka 使用介面	46

第一章 緒論

1.1 研究背景與動機

因市場的競爭越來越激烈，企業對於顧客關係的管理越來越重視，因此顧客的開發、保留、流失已成了企業獲利的關鍵因素。同時由於資訊科技的快速進步，使得資料的收集、儲存、分析的工作更得以實現，諸如資料倉儲、資料挖掘技術的應用等。結合顧客關係管理的意涵與資料採礦的技術，進而幫助企業分析流失顧客、保留顧客與開發顧客，以及如何結合資訊技術與專業知識將資源轉變成具決策價值的智慧，成為每個企業最關心的議題。

台灣建材零售業從傳統的中小企業，提供各種三夾板、木心板、原木訂製品以及裝潢用建材零售，到近幾年已有朝裝潢方面發展的趨勢，並且採取提升服務品質增加顧客滿意度。

過去，建材零售業的經營模式是以純零售為主，對於顧客的經營模式一直是處於願者上鈎的被動經營模式，不會主動開發與連繫舊顧客。由於經營模式簡單，對於與顧客之間的交易紀錄均是透過紙本的方式做記錄，此種方式可讓會計人員方便以及快速紀錄，但在交接上難免會有資料遺失的問題。

根據財政部統計月報資料，如表1.1，各個細項批發業銷售額佔總批發業之比率中，其中建材零售業在2008~2012年是呈現逐年下降的趨勢。在表1.2各細項批發商家數之統計數字中可以觀察到，2008~2012年建材零售業的家數也是呈現減少的趨勢，原因可能是經營不當導致失敗或是公司合併等情形，造成家數逐年減少。

表1.1 2008~2012年建材零售業銷售額於零售業所占比重

產業 (單位：%)	2008年	2009年	2010年	2011年	2012年
建材 零售業	6.33	4.74	4.92	4.51	3.74

(資料來源：財政部統計月報)

表1.2 2008~2012年建材零售商之家數

單位： 家數	2008年	2009年	2010年	2011年	2012年
建材 零售業	19370	18333	17721	16328	15301

(資料來源：財政部統計月報)

針對建材零售業銷售額及家數逐年下降的情形，經濟部統計處有觀察且調查在2011年經營上遇到的困境，並且統計出各項原因之比例，如表1.3，面對種種的困境，每個建材零售商採取不同的經營策略來吸引顧客，如表1.4。

表1.3 2011年目前經營上遭遇的困境

困境	資金 取得不易	營業用地價 格或房屋租 金過高	稅負過重	景氣不佳客 源不足	缺乏 專業人才	進貨 成本增加
單位(%)	9.62	1.92	3.85	25	3.85	21.15
困境	銷售商品 價格下跌	消費者偏好 變動大	勞動 成本提高	國際化 能力不夠	市場競 爭激烈	
單位(%)	1.92	1.92	5.77	1.92	23.08	

(資料來源：經濟部統計處)

表1.4 2011年所採取的經營策略

策略	提高 商品價值	推動商品 國際化	開 發 新事業	薄利 多銷
單位(%)	21.15	3.85	1.92	26.92
策略	拓展銷售 通路及規模	加強廣告 促銷活動	提 升 服務品質	創新營 運模式
單位(%)	21.15	3.85	19.23	1.93

(資料來源：經濟部統計處)

由於建材零售產品差異性低，加上面對景氣不佳客源不足、市場競爭激烈等困境，因此為了保持競爭力，許多公司採取提高商品價值、薄利多銷以及拓展銷售通路等不同的經營策略來吸引新顧客。但從銷售額逐年下降的情形來看，這些經營策略對於建材零售商的效果是不佳的。

現在，各類資訊充斥市場促使顧客主動參與性增強，因此不論是品質、價格及服務方式等皆容易被比較，造成顧客有貨比三家的心態，也促使顧客流動率升高。在公司想透過採取不同的經營策略吸引新顧客來提升利潤的同時，極容易忽略現有顧客的重要價值。

Frederick E. Reichheld(1996)指出，企業吸引一位新客戶所花的成本是維繫一位現有客戶的五至十倍。Peppers & Rogers(2001)指出，企業若能將客戶流失率減少5%，利潤就可提昇60%至100%。因此本研究想透過顧客關係管理的方式替公司維持住現有顧客以及建構完整的顧客關係管理。公司現有的顧客可以分為三種層次，分別是使用後非常滿意者，會再次購買者，且不但自己購買還會推薦他人使用者，而第三層者也就是CRM的目標族群，因為這類的人能比其他兩類人為公司帶來更多更多的利潤，如圖1.1。

顧客關係管理是利用資訊科技技術，強化與客戶的關係以鞏固老顧客，吸引新客戶及提高客戶的利潤貢獻度，並且利用與客戶相關的各種資訊，加以分析，來提高對客戶的了解，提供客製化的服務，增加客戶忠誠度和企業營運績效。一般而言，顧客關係管理蒐集之顧客資料包括行銷活動的顧客反應度、運銷和產品供應之相關數據資料、銷售與購買之資料、客戶資料、客戶註冊資料、相關服務之數據、產品市場資料及網路銷售數據。由於各項資料均為顧客與企業間之互動而產生，因此藉由分析顧客之各項資料，得以有效掌握顧客特性，而資料採礦(Data Mining)之技術及應用即可達到此一要求（尹相志，2004）

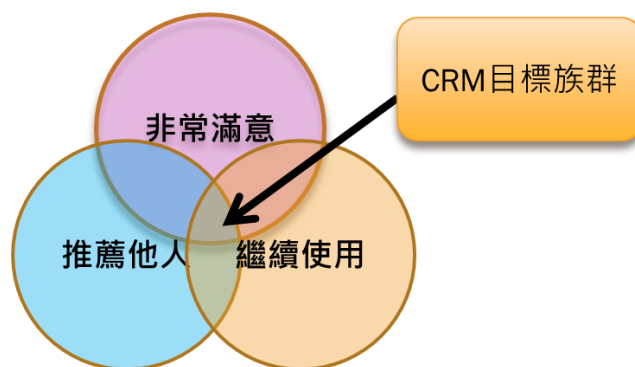


圖 1.1 忠誠顧客的概念(資料來源：本研究整理)

利用資料採礦工具，可從龐雜的資訊中抽取有用的資訊，以客觀的統計分析方法，快速且正確地找出企業需求的經營資訊，得到正確的銷售模式、客戶關係、行銷策略等，以利企業掌握正確經營動態。

隨著資訊科技的發達，企業希望從資料庫中找出顧客的特徵、消費的趨勢、與購買行為的關聯，來提供決策者制定企業的行銷策略。顧客關係管理的觀念即是在這樣的一個環境條件下產生，其主要追求的目標有四：保留既有顧客、獲取更多新顧客、提高顧客對公司的貢獻度、與防止顧客詐欺(陳麒文, 2002)。這些目標均可利用資料採礦的技術來預測未來可能的結果，進而防止顧客流失或建立與顧客之長期關係，以提升企業的營運績效。

1.2 研究問題與目的

由於建材零售業經營模式簡單，對於顧客的交易資料均是使用紙筆作紀錄，雖然此方法可以讓會計人員快速紀錄以及根據顧客不同的要求新增欄位，但在交接時難免會有資料遺失的問題，或是內部與外勤人員資訊錯誤等情形。因此本研究目的之一為將公司與客戶交易的紙本記錄建立顧客交易資料庫，並透過資料轉換為使用者需要的結構化資料，使分析者與系統皆能判讀。

公司的顧客可以分為兩大類，分別為製造商與零售顧客。透過分群法則將製造商的顧客分為80/20兩群，且命名為忠誠顧客與風險顧客，接著將80%的顧客再進行分群動作，分為流失與未流失兩群，進而分析與探討其關聯度。因此顧客面可以分成以下兩個目的：

研究目的一：以80/20法則分析製造商與供應商之行銷關係，比較20%顧客關聯度與80%顧客關聯度是否有雷同之處，期望能從80%的顧客群中再找出潛在忠誠顧客，提升企業營運績效。

研究目的二：透過零售顧客的交易紀錄找出顧客的購買特徵，若有符合特徵的顧客，提供客製化產品及專業性的知識，以達到顧客滿意並且降低顧客流動率，進而提高企業利潤。

1.3 研究流程

本研究分五大流程以圖 1.2 表示：

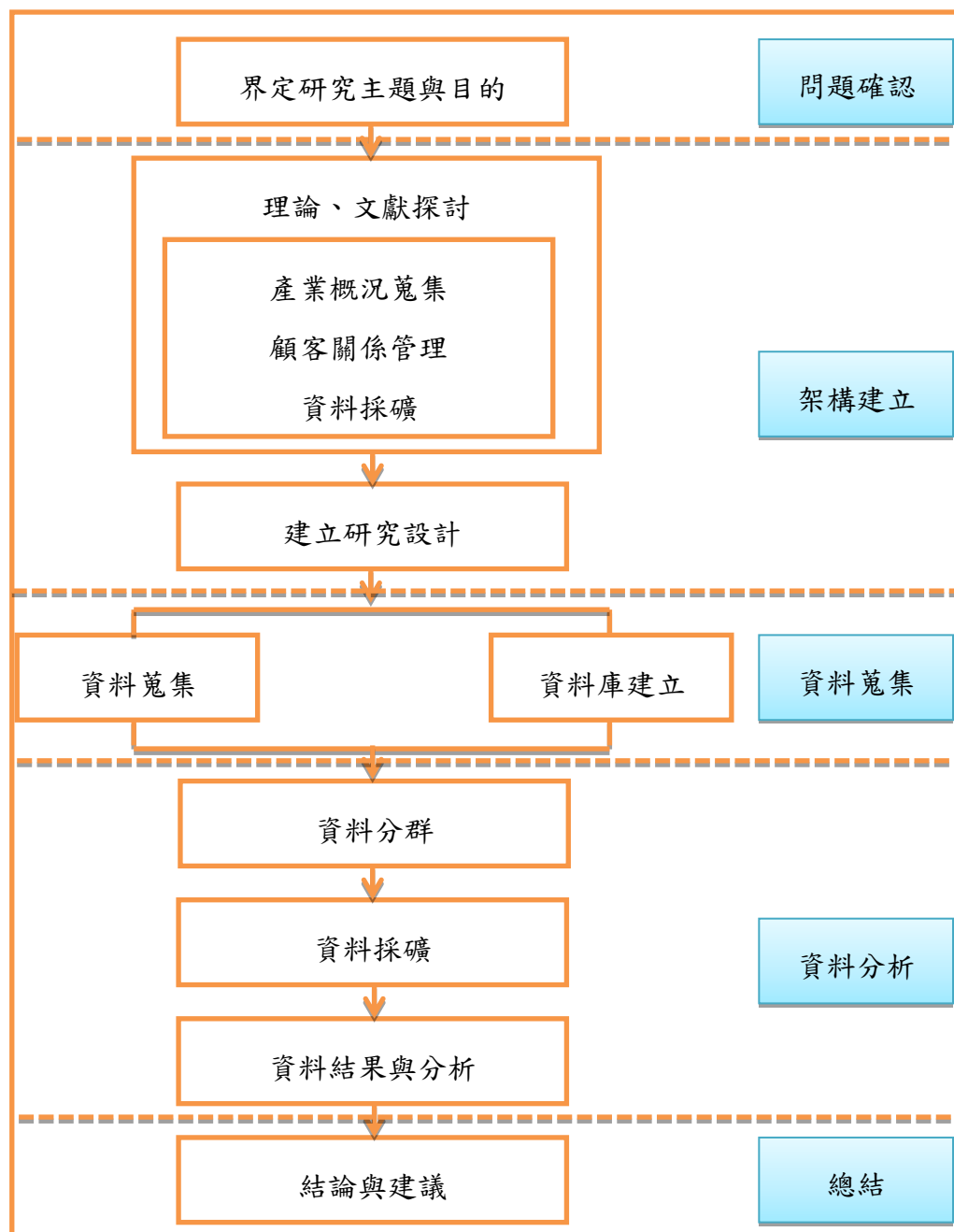


圖 1.2 研究流程圖(資料來源：本研究整理)

1. 問題確認

瞭解及界定研究主題，從觀察到之現象提出研究背景及動機，並瞭解需研究的問題及目的。

2. 架構建立

主要由三大部份之文獻探討以做為歸納架構之建立依據，第一部分產

業概況之蒐集，瞭解建材零售業的營業模式以及歷年交易情形；第二部分顧客關係管理之發展以及應用之探討；第三部分資料採礦之研究，對於其應用以及技巧做一深入之研究，進而決定使用何種探勘方式分析資料。

3. 資料收集

取得研究對象之初級資料，透過資料處理、過濾以及正規化，並建立欲研究資料的資料庫以利後續資料採礦分析。

4. 資料分析

利用資料採礦技術關連規則以及群集分析，企圖在資料庫中找出有價值之資訊。

5. 總結

根據本研究分析發現提出結論，並對於業界及學界提出建議。

第二章 文獻探討

2.1 顧客關係管理

顧客關係管理的英文為Customer Relationship Management(CRM)，最早開始發展顧客關係管理的國家是美國，其於1980年代初期便有所謂的接觸管理，專門收集顧客與公司聯繫的所有資訊；至1990年初其則演變成為包括電話服務中心與支援資料分析的客戶服務功能。

麥肯錫顧問公司董事長 John Ott 則將 CRM 做另一種方式的解釋，他認為所謂的顧客關係管理，應該是「持續性的關係行銷」(Continuous Relationship Marketing, CRM)。其強調的重點：尋找對企業最有價值的顧客，以微型區隔的概念，界定出不同價值得顧客群，企業以不同的產品、不同的通路，滿足不同區隔顧客得個別需求，並在關鍵時刻，持續地與不同層次的顧客溝通，強化顧客的價值貢獻。同時還必須進行反覆測試，進而隨著顧客消費行為的改變調整銷售策略，甚至是更動結構(陳曉開,1999)。

經濟部商業司則指出，顧客關係管理乃技術性策略，將資料驅動決策轉變為商業行動，以回應並期待實際的顧客行為。從技術觀點來看，CRM 代表必要的系統及基礎架構，已擷取、分析與共享所有企業與顧客關係。從策略的角度來看，CRM 代表一個過程，用來評估與分配組織的資源，給那些能帶來最大利益的顧客關係活動。

根據80/20 法則 (80/20 Principle)，企業80%的利潤是來自於20%的顧客(陳麒文, 2002)，如圖2.1所示，因此企業對這20%顧客的關係就顯得格外重要。另一方面，若以成本的角度來思考，企業會從既有的顧客中獲取90%的營收(Curry, J., Wurtz, W., Thys, G., & Zijlstra,1998)。而根據李昇敦的研究調查(李昇墩, 2000)顯示，一般人認為受到不好的服務時，96%的人會默默地離去，其中91%的人日後絕不會再光顧，企業一般只能聽到4%不滿顧客的抱怨；同時調查中也顯示，其中一位不滿的顧客平均會把他的抱怨轉告八到十個人，甚至其中的20%還會轉告20 人之多。由以上種種的省思再次強調出管理大師彼得杜拉克 (Peter Drucker) 所說過的：「每個公司最重要的工作是爭取顧客、留住顧客、和竭盡所能地擴大顧客獲利貢獻度。

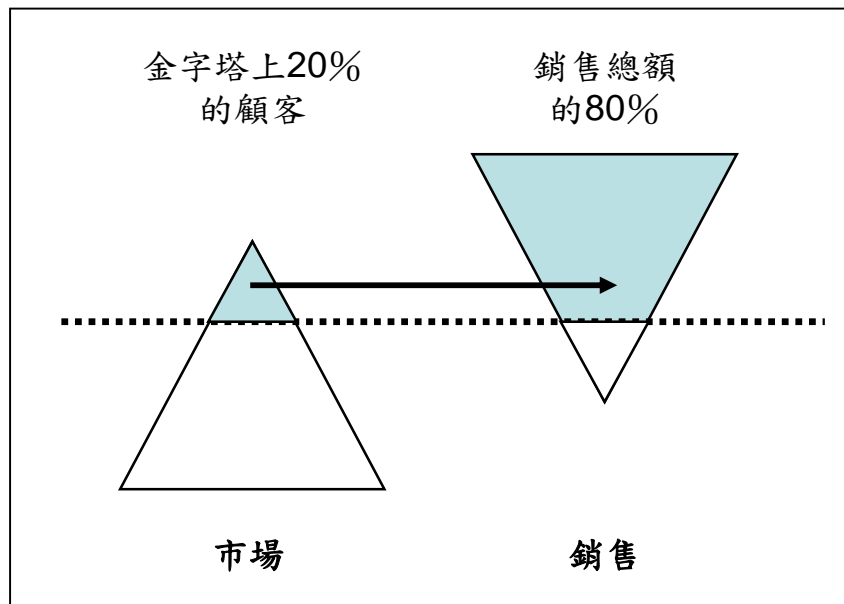


圖 2.1 市場與銷售之 80/20 法則(資料來源：陳麒文，2002)

企業經常把焦點放在「獲取新顧客」上，卻往往忽略了本身原有的顧客群(唐林林，2003)，如此一來，便造成了所謂的「旋轉門效應 (Revolving-door Effect)」(夏載，2001)，亦即費盡心思地將新顧客拉進來時，舊的顧客卻出走了。同樣地，IBM 也表示，必需以六到八倍於留住舊顧客的代價，才能拉進一個新的顧客。這些說明都顯示著，當顧客忠誠度不高或資訊取得不足時，企業的獲利狀況就會面臨威脅，而顧客關係管理便是為了解決企業面臨網路與相關資訊科技的衝擊，及缺乏堅強擁護的顧客群危機。

而顧客關係管理主要目標仍在於即時滿足客戶需求、提高客戶滿意度、與客戶建立長期良好的關係及增加營業利潤。但是，隨著資料倉儲與資料挖掘等知識管理技術的應用，客戶關係漸漸成為 CRM 的核心。如何透過顧客分析找出客戶得消費行為、忠程度、潛在消費群與主關鍵客戶，進而利用促銷管理針對不同市場區規劃行銷活動，以達到建立品牌知名度、改變購買行為或維持客戶忠誠度等目的，是企業對於 CRM 的期許。此外，藉由每次促銷活動的效益與顧客關係的交叉分析，選擇適當的通路與客戶進行互動，更是企業掌握客戶需求變動與提升客戶滿意度的最佳解決方案(1999 年度台灣「顧客關係管理」運用現狀調查報告，2009)。圖 2.2 表示顧客關係管理架構的主要應用元件。

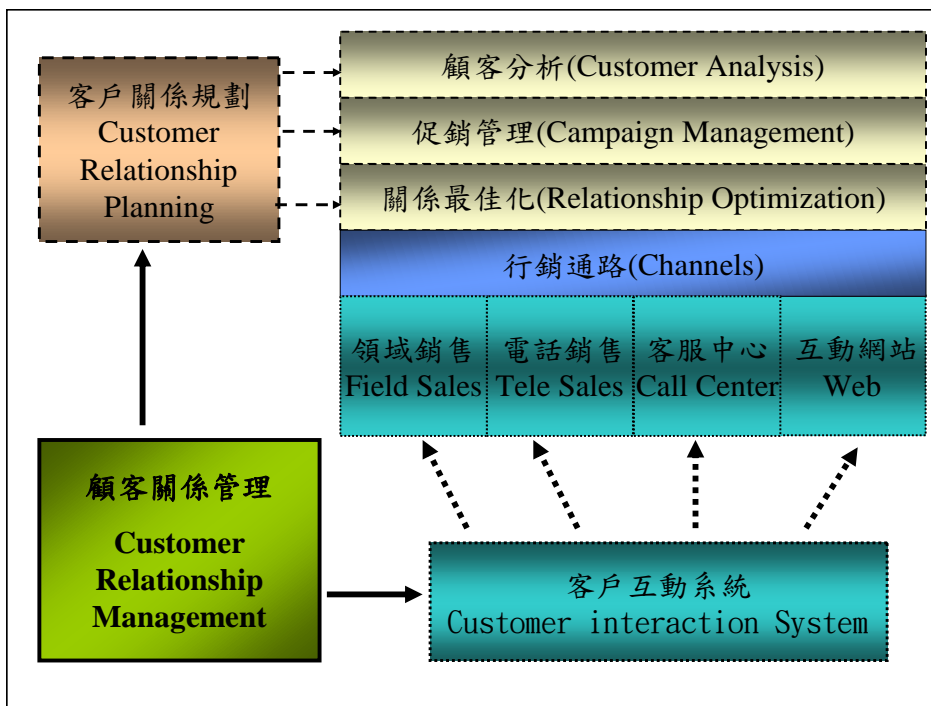


圖 2.2 顧客關係管理架構主要應用元件

(資料來源：1999 年度台灣「顧客關係管理」運用現狀調查報告，2009)

綜合以上論述可以歸結出，今日的顧客關係管理理由由於結合顧客溝通管道及資料分析上的資訊技術的進步，可以分別就技術面與策略面來看。就技術面而言，便是利用資訊庫技術，使企業可以蒐集所有客戶相關資料，加以大量轉換、載入、分析，將這些資料加以預測和分析，以作為行銷策略制定的參考，也就是將資料驅動決策 (Data-driven decisions) 轉變為商業行動，以提高執行成功的機率，進而達到提高利潤及降低成本的目的。就策略面而言，則是透過顧客分析找出客戶的消費行為、忠誠度、潛在消費群與對企業最有貢獻價值的顧客，以妥適分配組織的資源，並以不同的產品、不同的通路，滿足不同區隔顧客的個別需求，以致力於顧客滿意與顧客忠誠的提昇。同時配合策略的改變更動組織結構與流程，獲得員工與顧客的一致支持，才能有效率地因應顧客需求做調整，成為真正以顧客為中心的彈性組織。

2.1.1 顧客關係管理定義

顧客關係管理的定義廣泛，許多學者以及業界均有其解釋，本研究將其整理如表 2.1：

表2.1 顧客關係管理之定義

提出學者、業界	年代	顧客關係管理之定義
David	1999	顧客關係管理就是「關係管理Relationship Management」、「終身價值行銷Life-time Marketing」、「忠誠行銷Loyalty Marketing」。這些策略企圖創造企業與顧客間長期獲利的關係，並發展忠誠關係與創造利潤。
麥肯錫(McKinsey)公司	1999	顧客關係管理應該是持續性的關係行銷(Continuous Relationship Marketing, CRM)，尋找對企業最有價值的顧客。
Alex (Alex, S.,1999)	1999	認為顧客關係管理簡單地定義就是滿足大部份對於企業具有價值(利潤貢獻度)的顧客需要；同時顧客關係管理是許多技術與觀念的集合與發展，牽涉的技術與觀念包含行銷學、一對一行銷、資料倉儲、資料採礦、顧客區隔與顧客忠誠度的維繫等。
NCR安迅資訊公司 (安迅資訊公司,1999)	2000	顧客關係管理為分析、整理代表顧客消費習性的顧客交易資料，歸納出每一位顧客的需求，並針對其做出回應。而其主要的目的為：1.鞏固現有顧客(Customer Retention)；2.贏取新顧客(Customer Acquisition)；3.增進顧客利潤貢獻度(Customer Profitability)。
Jackie Kandell (Jackie Kandell,2000)	2000	CRM的重點在於以客戶需求為核心，視客戶生命週期為重要的企業資產，主要目標在於保有客戶並提高客戶滿意度。在同一時間內提供客戶不同的產品與服務，儘可能滿足客戶所有的需求以謀取最大的收益。主要應用的行銷方式有：一對一行銷、大量客製化等。
Paul Ratcliff (Paul Ratcliff,2000)	2000	認為CRM是以行銷為基礎的新觀念，用來有效的維持並加深企業與客戶之間的關係，資訊科技在CRM中屬於輔助性的角色，讓CRM相關工作更有效率的執行。本質上，CRM透過客服代表進行客戶服務與維持現有客戶的工作，客服代表分析客戶相關資料以從事行銷與交叉銷售的活動，滿足個別客戶的需求。
Newell	2000	顧客關係管理是靠增加顧客與公司之間的連繫來不斷改變顧客行為的整個過程。

表 2.1 顧客關係管理之定義 (續)

提出學者、業界	年代	顧客關係管理之定義
陸承志 (陸承志, 2001)	2001	所謂的顧客關係管理, 係指企業運用整體資源以全面瞭解每位個別顧客, 並運用所有管道與顧客互動, 以提昇顧客的終身價值。
資策會	2002	顧客關係管理應該是持續性的關係行銷 (Continuous Relationship Marketing, CRM), 尋找對企業最有價值的顧客與顧客互動, 並以達成或超越顧客最滿意為目的。並且是需要持續不斷的開創、修正、執行與顧客間的互動。

(資料來源：本研究整理)

2.1.2 顧客關係管理之步驟

顧客關係管理是一種不斷將最新、即時及迅速地將顧客資訊轉變成有益的顧客關係, 而且他是一種循環反覆的過程, 從建立顧客知識開始, 到建立所謂的顧客互動溝通管道, 如此反覆操作以找出最適合自身的最佳管理方法, 如圖 2.3 所示。大概可分為四個階段：

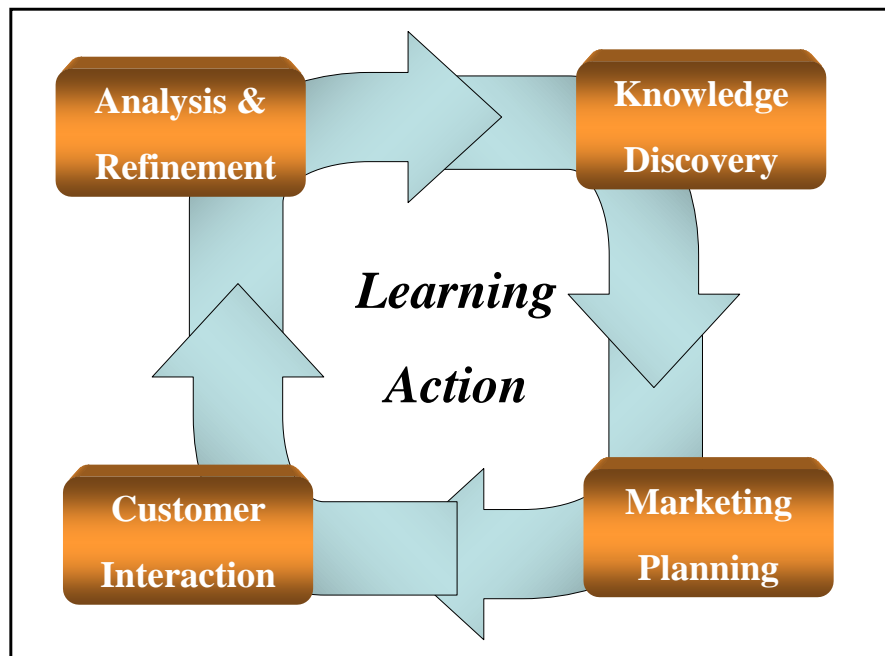


圖 2.3 顧客關係管理的循環過程(資料來源：劉玉萍, 2000)

1. 知識發掘 (Knowledge Discovery)

這是透過顧客確認、顧客區隔以及顧客預測的流程。企業經由各種路徑蒐集各種的顧客互動與交易來源,同時將其轉換成為對管理與規劃有用的資訊與知識,給予行銷人員使用詳細顧客資訊的能力,以做為最佳的歷史資訊與顧客屬性分析,並做為更佳決策。

2. 市場行銷計畫 (Marketing Planning)

這是為特定顧客產品、提供通路、時程以及從屬關係的流程。行銷規劃幫助發展策略性的溝通計劃或活動,掌握與規劃投資機會、服務顧客的計劃以及產品與服務。

3. 顧客互動 (Customer Interaction)

這是運用相關及時資訊,透過辦公室前端應用軟體與顧客之間的互動溝通的流程。顧客互動必須與你的顧客、潛在的延伸企業互動點,以及銷售策略與顧客購買活動的連結相互對應。

4. 分析與修正 (Analysis and Refinement)

這是利用顧客互動資料分析結果,持續修正與管理方法的流程。藉由與顧客對話中持續學習的階段,分析從顧客互動與調整的訊息、溝通方式及時機。

而 Kalakota and Robinson(1999)認為,要妥善的管理顧客生命週期,就必須以三個階段來實行顧客關係管理,這三個階段分別為:獲取新顧客 (Acquisition)、強化既有顧客的獲利性 (Enhancement)、以及維持現有顧客 (Retention),如圖 2.4 所示。

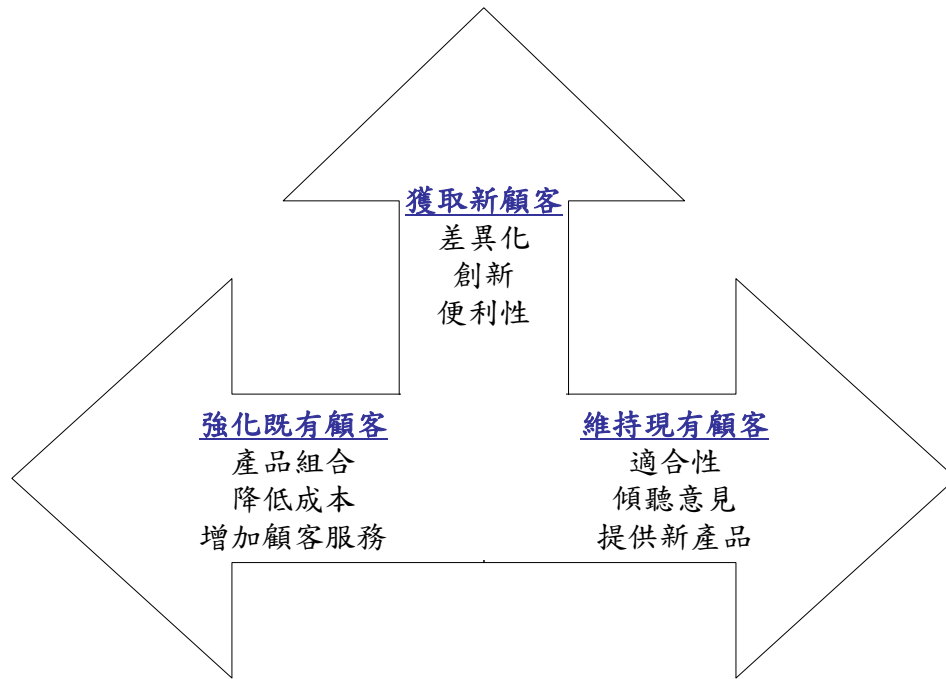


圖 2.4 顧客關係管理之運作流程(資料來源：Kalakota, R., & Robinson,1999)

2.1.3 顧客忠誠度

世界產業的脈動已由生產製造導向轉為以顧客服務為導向，顧客已成為了企業最重要的目標，因擁有顧客即擁有了收益。當產品不斷推陳出新，產品多樣化、微利的時代來臨時，顧客的選擇性因此而大幅增加，對於消費者不啻是一個好現象。但對於企業而言，供應者必須要推出更多樣化的產品來滿足顧客，以維持顧客對於產品的支持，顧客的忠誠度變的難以維持。鑑於上述問題，許多專家學者均有其深入的研究。

Jones and Sasser(1995)認為顧客的忠誠度會影響顧客對產品或服務未來再次購買的意願。其中將顧客忠誠分為長期忠誠和短期忠誠兩種，長期忠誠是指顧客長期購買某特定品牌的產品或服務，不易改變選擇；短期忠誠是指當顧客有更好的店家或產品選擇時，顧客就會立即轉換。造成短期忠誠的原因包括政府法規的限制競爭、限制替代品的專利技術、移轉成本及競爭者強力的促銷活動等。

Oliver(1997)認為顧客忠誠是指顧客雖然受到環境影響或外在行銷手法的誘惑，顧客對所喜好的產品或服務的未來再購買意願與承諾仍不會有所改變。Bowen and Shoemaker(1998)認為顧客忠誠是顧客再次光臨的可能性大小，並且顧客願意成為企業一份子的意向。Singh and Sirdeshmukh(2000)

認為顧客忠誠是消費者願意繼續與服務提供者維持關係的一種傾向。

Fredericks(2001)提出顧客佔有率(share of customer)的觀點。相對於傳統企業追求市場佔有率，所謂顧客佔有率是指針對同一顧客，設法賣給他更多的商品。換句話說也就是對於特定類型的消費，讓顧客將預算花在特定企業商品或服務的比例達到最高。顧客消費愈高，為企業帶來更高的利潤，愈是忠誠的顧客。

顧客忠誠度的維繫可有效的保持住企業獲利的主要顧客群，由於開發新顧客的成本遠高於舊顧客保留所需的成本，而且忠誠度是金錢買不到的，所以企業是否成功的一個重要指標即為顧客的維持率，當顧客忠誠度愈高，顧客的維持率也愈高；相對地，在維持顧客、提高忠誠度的同時，對於顧客的流失也成為了一個重要的議題，本研究將在下節做一探討。

2.1.4 顧客關係管理系統之重要性及效益

Stone, Woodcock & Wilson(1996)關係行銷的利益通常發生在下列一個或一個以上的區域：

- 增加顧客保留及顧客忠誠度—顧客停留越久，買的越多，更常購買，亦即增加生命週期價值。
- 較高的顧客貢獻度，不只是因為每位顧客購買越多，而是因為：
 - ◆ 花費較低成本招收新顧客。
 - ◆ 銷售成本的降低，因為原有的顧客通常更有反應。

增加利潤及競爭力的關鍵可能是擺脫那些永遠不可能具獲益性的顧客並集中於招收與保留具有高終身價值的顧客。

依照 Spengler(2000)調查台灣企業導入 CRM 的狀況而言，僅有 10% 的公司已經完成 CRM 的建置工作，有 28% 的公司正處於建置規畫期，另有 48% 的公司正在評估是否要引進，而沒有引進企劃者企業約占 14%。由此可見 CRM 在國內頗具發展空間。

而其問卷結果中，業者導入 CRM 系統主要有以下四點原因：

1. 推展行銷業務

公司期待導入 CRM 後，能夠發展訂製化的產品給消費者。例如信用卡業者可依不同顧客購買偏好來寄發不同之郵購目錄，保險業者

則可依顧客年齡及婚姻狀況，預測顧客未來可能購買保險之種類。

2. 提升顧客服務的品質

以企業利用電話服務中心(Call Center)功能為例，由於客服人員可以根據電腦資料庫所提供的顧客背景資料，與所使用的服務，直接在電話線上快速回應顧客需求，因此不僅可以減少顧客的抱怨，還能增加顧客忠誠度。

3. 提升經營績效

CRM 強調企業的流程設計應該以顧客為導向，而非以產品為導向，因此台灣企業採用顧客關係管理後，可以減低各項成本，尤其是行銷費用，進而提高公司營運績效。

4. 提升公司形象

銀行、信用卡中心等金融業者率先採用 CRM，除了著眼於創造業績之外，也期待可以藉此將金融業提升至「服務業」的層次，塑造新的企業形象。這種情形造成一個新的趨勢，促使其他產業隨之採用 CRM，來提升公司的形象。

蕭崑龍(2001)認為顧客關係管理可以達成以下目標：

1. 顧客忠誠

為了要長期維繫顧客，企業必須長期投資在其顧客上，一個完善的 CRM 解決方案提供企業一個獨一無二的機會，使其能夠瞭解個別不同的顧客，並進而改進產品與服務，以達成顧客各自不同的特殊需求。這些機會包括了與既有顧客的種種互動，包括推薦顧客採購新產品或附加產品/服務，提供彈性化的服務合約，以及開發新產品以及隨著顧客需求變動而修正的加強措施。

2. 降低銷售成本與增加獲利

由於大部分的企業其銷售與行銷成本佔據了高達 60% 的收益，故須努力降低銷售成本。

3. 積極主動的顧客管理

顧客服務解決方案協助企業得以積極主動管理其顧客並且建議長期的顧客關係。這些 CRM 解決方案也協助企業提供更快捷的服

務、對於顧客問題能提供更精準的答覆，同時提高顧客忠誠度等具競爭力的優點。

4. 改善品質與顧客諮詢

品質保證解決方案是顧客關係管理流程的重要環節，因期能確保對於顧客的產品與加強改善的要求都能夠在一個有效、準確且及時的方式做適當的處理，亦即是公司的產品都能夠做到一致性，符合顧客的需求。

5. 提高顧客滿意度

一個完善的 CRM 解決方案將把企業推向一個與其顧客「一對一」的密切關係。有了 CRM，企業們將可以積極處理顧客的需求。例如：在短時間內聯繫追蹤比平常量大很多的支援需求，接洽顧客以找出一個長期與短期兼具的解決方案。企業亦可採用 CRM 解決方案以辨別出購買方式，並利用這些資訊來針對顧客的要求與需要做行銷處理。

6. 從銷售產品轉型為銷售解決方案

顧客不僅需要產品，也需要完善的解決方案，而關鍵點就在如何從多樣化的產品與服務當中，在有限資源的情況下設計出具體完善的解決方案，使銷售成為一個適合顧客的解決方案，這也是一個極大的挑戰。

2.2 資料採礦

隨著時間的累積，各企業組織的電腦系統中的資料儲存量也隨之增加，然而，這些資料中有用的資訊，往往隱藏著某些特徵（Patterns）以及關係（Relations），使用傳統資料查詢和統計功能，並不容易找出它們。資料採礦（Data Mining）是一種專門的程序，可在大量存放的資料中，找出先前並不知道，但最後可有效理解的資訊，並可利用這些找出的資訊建立一個預測或分類的模型，或識別不同資料庫之相似性，產生的資訊可協助決策者進行更週延的決策。

2.2.1 知識發掘

資料採礦是知識發現的一種過程，但知識發現並不只有資料採礦一項

工具，凡是能夠找出商業知識的方法過程，皆屬於知識發現的範疇。

U. M. Fayyad 在 IEEE Expert Intelligent Systems & Their Applications 中所發表的文章(Fayyad,1996)中曾經描繪出整個資料挖掘的步驟如圖 2.5 所示。在此圖中，我們可以看到以下幾個步驟：

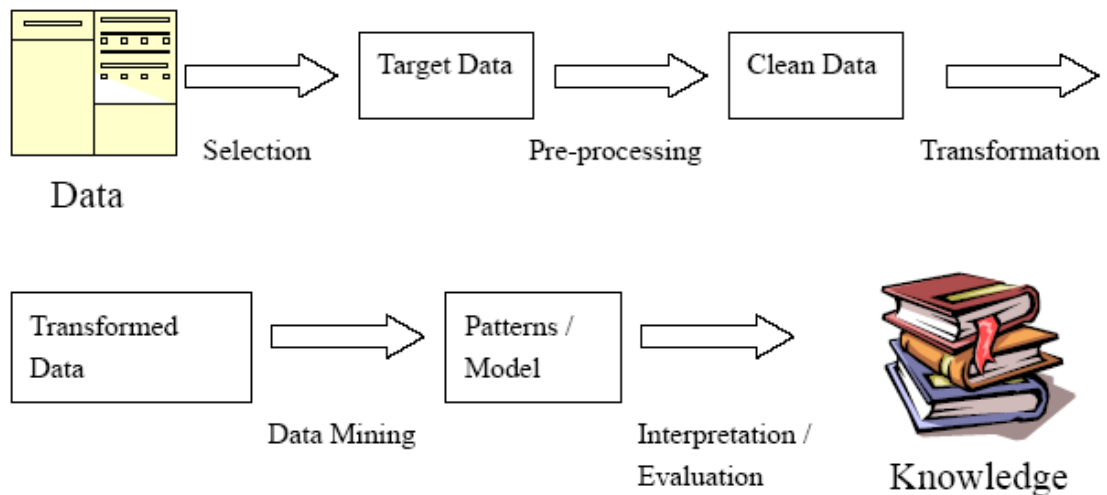


圖 2.5 Knowledge Discovery in Database Process(資料來源：Fayyad,1996)

1. 針對所要開發的應用領域瞭該應用領域、熟悉相關的先導知識，以及最終目標。
2. 建立一個挖掘目標資料集，以選擇資料集或專注在樣本資料子集或變數子集中。
3. 將做前置資料處理，去除錯誤或不一致的資料(Data Cleaning and Preprocessing)。對料做簡化或轉換的工作(Data Reduction and Transformation)。
4. 選擇資料挖掘的工作是：找出分類的規則(Classification Rules)、找出關聯性規則(Association Rules)、找出序列規則(Sequence Rules)、找出同質時序列(Similar Time Series)、找出群集規則(Clustering Rules)。
5. 根據上面的選擇挑選演算法或處理模式。
6. 以上述演算法尋找「樣型」(Pattern)、做迴歸分析(Regression)或找出分類型態等。
7. 評估上述結果，此工作須要專家的參與，以決定是否為一有用的結果。

由此可知，知識發現是一連串的程序，資料探勘是其中的一個步驟而

已。根據 Fayyad 等人對知識發現的定義：「The nontrivial Process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data」，換句話說，就是一個指出資料中有效、嶄新、潛在效益的一個非細瑣流程，而其最終的目標是瞭解資料的樣式。

2.2.2 資料採礦的功能

Michael J.A. Berry & Gordon S. Linoff (2001)認為，Data Mining 的功能包含了六大功能：分類(classification)、推估(estimation)、預測(prediction)、關聯分組或同質分組(association or affinity grouping)、群集化(clustering)、描述及視覺化(description and visualization)。分述如下：

1. 分類(classification)

按照分析對象的屬性分門別類加以定義，建立類組(class)。例如，將信用申請者的風險屬性，區分為高度風險申請者，中度風險申請者及低度風險申請者。使用的技巧有決策樹(decision tree)，記憶基礎推理(memory-based reasoning)等。

2. 推估(estimation)

分類出來的結果是不連續的，而推估所得的結果是連續性的數值。例如按照信用申請者之教育程度、行為別來推估其信用卡消費量。使用的技巧包括統計方法上之相關分析、迴歸分析及類神經網路方法。

3. 預測(prediction)

根據對象屬性之過去觀察值來推估該屬性未來之值。例如由顧客過去之刷卡消費量預測其未來之刷卡消費量。使用的技巧包括迴歸分析、時間數列分析及類神經網路方法。

4. 關聯分組(association or affinity grouping)

從所有物件決定那些相關物件應該放在一起。例如超市中相關之盥洗用品(牙刷、牙膏、牙線)，放在同一間貨架上。在客戶行銷系統上，用來確認交叉銷售的機會以設計出吸引人的產品群組。

5. 群集化(clustering)

將異質母體中區隔為較具同質性之群組(clusters)。同質分組相當於行銷術語中的區隔化(segmentation)，但是，假定事先未對於區隔加以定義，

而資料中自然產生區隔。使用的技巧包括 k-means 法及 agglomeration 法。

6. 描述及視覺化(description and visualization)

有時資料探勘的目的，只是單純的想使用更容易了解的方式，來描述一個複雜的資料庫中，發生了什麼事。對一個狀態來說，一個正確的描述就可以啟發更多對該狀態的解釋。而資料視覺化是一種強而有利的描述是資料庫探勘。要找出有意義的視覺化並不容易，但一張適當的圖表，真比一千條關聯規則更有價值，因為從圖像來萃取有意義的資訊是人類擅長的。

2.2.3 資料採礦的方式

資料探勘不是一種新技術，而是結合多項專業技術的研究 (Pieter,1996)，它包括機器學習、統計方法、資料庫、專家系統及資料視覺化的領域。Curt(1995)認為資料探勘的技術至少包括資料視覺化、機器學習技術、統計技術及資料庫四種技術。Y. Fu(1997)將資料探勘的技術依不同研究領域區分說明：

- 統計方法 (Statistical Approaches)：資料探勘使用許多統計工具，包括 Bayesian 網路、迴歸分析、相關分析及分群分析 (Cluster Analysis)，通常統計模式是經由訓練資料集 (Training Data Sets) 來建立，然後從模式中尋求規則及特徵。
- 機器學習方法 (Machine Learning Approaches)：像統計方法一樣，機器學習方法尋求一個最佳模式來符合測試資料 (Testing Data)。不同之處在於大多數機器學習方法利用資料自動化學習過程，自動歸納出分類規則及建立模組。資料探勘最常使用的機器學習方法包括決策樹歸納。
- (Decision Tree Induction) 及概念式分群 (Conceptual Clustering)。
- 資料庫導向方法 (Database-oriented Approaches)：資料庫導向方法不像前兩個技術在尋找最佳模式，轉而著重在處理現有的資料，是一種屬性導向歸納 (Attribute-oriented Induction)，反覆從大量的資料中找出共同的規則及模式。
- 視覺探索 (Visual Exploration)：將多維度的資料轉換成視覺化物件，

如點、線及區域，讓使用者可以動態檢視及探索有興趣的部分，進而分析出資料的模式。

- 其他方法：如類神經網路（Neural Network）運用於資料的分類及預測，約略集合（Rough Set）運用於分類及分群。這些技術可以互相整合來處理複雜的問題及提供替代方案，如統計方法可以結合視覺探索技術，用以輔助系統處理資料模式及趨勢分析。

另外，常見的資料採礦方法如下：

- 購物籃分析：運用關聯規則（Association Rule），著眼點在於找出可以一起銷售的商品組合，它能顯示商品組合的售出率有多高並且形成規則，當交易是非匿名時，它可以加上一個模組進行跨時性分析。
- 記憶基礎理解（Memory-based Reasoning）：比對新資料和現存資料庫原有資料的相似度，以決定新資料的類別或預測值的工作，它是以距離函數判斷資料間的相似度，再以數個最相似的資料投票來決定歸類結果。
- 群集偵測：針對要分析的資料，利用幾何學、統計、類神經網路等方法，將資料分成多個群內同質、群間異質的群組，使各群組的特徵能有效突顯出來，連結分析搜尋資料間的關聯，並且據此發展出分析模型，這是圖像理論在資料挖掘中的應用。
- 決策樹：主要用在資料分類與預測功能上，它能將訓練資料集的紀錄區分為獨立的子群，每一子群都有自己的規律，彼此是互斥的。利用決策樹技術，容易獲得清楚易懂的分類規則並找出關鍵屬性，所以本研究將加以利用決策樹技術來進行資料挖掘的工作。
- 類神經網路：一種平行分散式的計算模式，以大量簡單的相連人工神經元，模仿生物神經網路的資訊處理系統，使電腦能夠模擬人類的神經系統結構，進行資料的處理。
- 基因演算法：應用選擇、雜交、突變等物競天擇和基因演化的機制，將此機制結合電腦語言，經過世代繁衍，得到最後留下的最佳方案，通常用在找尋預測功能的最佳參數。

本研究應用購物籃分析、集群偵測以及決策樹的方法，透過顧客交易以及維修資料，來做顧客行為特徵的分析。

2.2.4 資料採礦流程

資料採礦流程又可稱為工作循環如圖 2.6 所示，

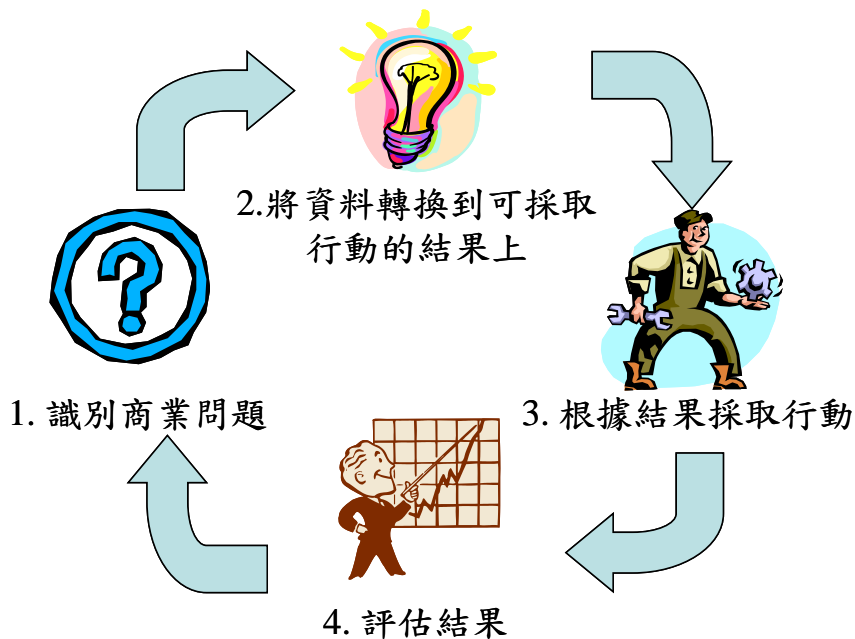


圖 2.6 資料採礦流程(資料來源：IBM,1998)

其包含了四種主要的商業程序：

1. 識別商業問題：瞭解問題的所在，需從什麼方向、角度來分析此問題。
2. 將資料轉換到可採取行動的結果上：使用資料採礦方法，把資料轉換到可採取行動的資訊上。通常此階段最為繁雜，又可分為四個步驟：
 - (1).選取輸入資料：資料探勘的第一步驟，是指定要探勘和分析的資料，資料來源不一定是特定資料庫中的所有資料，探勘的資料可能從一個或多個資料庫中，取得表格、概略表或記錄文字檔。
 - (2).轉換資料：為降低資料量，首先將收集的資料作整理、清除重複或無效的資料記錄，並且確保消費者資料的完整。
 - (3).資料採礦：使用資料採礦方法如分類、趨勢分析、分群、關聯及循序特徵等，從轉換後的資料中發掘存在的多種特徵及資訊。
 - (4).解釋結果：經過資料探勘後，一般以文字及圖形來顯示結果。
3. 根據結果採取行動：依據分析的資訊做出行動。
4. 評估結果：對於行動後所得到回應做一評估，再度定義新的問題，重複此工作循環。

2.2.5 資料採礦技術應用

資料採礦技術的應用相當廣泛，在顧客關係管理方面亦有許多的研究方向。目前在業界應用的範圍可以下表 2.2 表示：

表2.2 資料挖掘實例

	行業別	描述
1	信用卡公司	信用卡公司可使用資料挖掘來從事有關購買授權決定、分析持卡人的購買行為、偵測詐欺行為等，成功的案例有美國聯邦銀行(American Express)及花旗銀行(Citibank)。
2	零售商	可利用資料挖掘來瞭解顧客購買行為及顧客偏好的資訊，例如：透過購物籃分析或是應用其他技術來偵測收銀員詐騙的行為，成功的案例有 Wal-Mart。
3	金融服務機構	使用資料挖掘來分析大量的財務資料，建立交易及風險模式來發展投資策略。
4	銀行	利用資料挖掘以識別顧客的貸款活動、調整金融商品以符合顧客需求，並用來尋找新的顧客及加強顧客服務。
5	電話銷售及直銷	電話銷售及直銷公司因使用資料挖掘已節省許多金錢並且能夠精確的取得目標顧客。直銷公司可依顧客過去的購買資料及地理資料來配置及郵寄他們的產品目錄；對於電話銷售公司而言，不但能減少通話數，而且可以增加成功通話的比率。
6	航空業	對於航空業者而言，瞭解顧客需求已經變得極為重要，主要是藉由資料挖掘獲取顧客相關資訊，來制定因應策略。
7	製造業	資料挖掘已廣泛的被使用在製造工業的控制及排程技術生產程序，可用來偵測潛在的品質問題，以降低產品不良率的發生。
8	電信公司	電信巨人像 AT&T 和 GTE 正應用這些快速偵測不尋常的行為技術來防止竊打。
9	保險公司	提供保險業者從大型資料庫中取得有價值的資訊以進行決策，並利用這些資訊有效的偵測保險詐欺。
10	政府單位	美國聯邦政府利用資料挖掘的技術找出犯罪的關連性及可能性。另外，財政部也利用此方法找出洗錢及詐欺的案件。

(資料來源：Berry, M. J. A.,1997 and 元智大學商業智慧實驗室，2001)

李昇暉在「以資料採礦(Data Mining)深化顧客關係管理」一文中指出，企業可以將資料挖掘技術應用在以下六個方向：

- 獲取新客戶(Customer Acquisition)
- 維繫客戶(Customer Retention)
- 放棄客戶(Customer Abandonment)
- 購物籃分析(Market Basket Analysis)
- 需求預測(Demand Forecasting)與目標行銷(Target Marketing)
- 交叉銷售(Cross-Selling)與主動銷售(Up-Selling)

分述如下：

1. 獲取新客戶(Customer Acquisition)：從第一步開始，可根據顧客屬性來預測其對商品或通路計劃的反應，接著可以比照相對應的實際屬性與反應是否真如預期，並且從中挑選出那些尚未成為我們的顧客，但最有可能會對我們產品感興趣的人來。
2. 維繫客戶(Customer Retention)：當資訊顯示企業的基本顧客已經開始流失到對手陣營時，公司就該採取挽留措施，同時對那些還算穩定的顧客，就該給些誘因使其更願意留下來。
3. 剔除客戶(Customer Abandonment)：當出現顧客是企業投注於其上的費用遠超過他的回饋時，就應該考慮是否停止為這些顧客付出努力與成本。
4. 購物籃分析(Market Basket Analysis)：購物籃指的就是消費者所購買的商品種類及數量，分析消費者購買的產品將會對公司產生多少的收益，即是所謂的購物籃分析，或稱為聯合性分析（Association Analysis）。某些物品經常是同時被購買的，譬如可樂及洋芋片、牛奶和麵包等。從這類的分析我們可以得到對以下問題的解答：如相關產品該如何陳設？該促銷哪些產品？該作什麼行銷策略等。
5. 需求預測(Demand Forecasting)與目標行銷(Target Marketing)：在處理過大量的資料後，當再一次收到一筆新的資料時，電腦系統便會模擬它的結果。換句話說，就是我們能根據某類潛在顧客的特性去預測其需求，從而找出對我們所提供的商品最具有消費傾向的顧客。這方面的分析可加強我們對各種商品其主力顧客的促銷動作，進而提高銷售的成績。

6. 交叉銷售(Cross-Selling)與主動銷售(Up-Selling)交叉銷售是將具有相關性的產品服務加以促銷，像航空公司與租車行、飯店的結盟就是一個例子。主動銷售是根據顧客的消費特性，向潛在顧客介紹適合的產品，如保險公司可以向雙薪並有年幼子女、年收入150萬的保戶提供以下建議：有75%和他們條件相同的保戶除了購買意外險外，也會幫自己的子女購買教育年金。

2.2.6 相關文獻探討

以下為本研究彙總之資料採礦相關文獻，並將研究主題與研究之結果列於表中，其摘要如表 2.3：

表2.3 國內資料採礦相關研究整理

作者	年度	研究主題	研究結論
鍾碧緝	92	資料探勘應用於銀行業信用卡持卡者、帳戶區隔與零售業酒品交叉銷售之探勘研究	利用資料探勘的購物籃分析進行顧客購買行為與偏好研究，以決定最適銷售產品組合與最有可能被推薦於交叉銷售的商品。
陳宜欣	95	以資料探勘技術探討顧客忠誠方案-以某信用卡發卡銀行為例	利用資料探勘的方式探討消費者參與信用卡忠誠方案本身，以及忠誠方案是否能夠為發卡銀行帶來收益。
曾和順	95	應用資料探勘技術提升電話行銷成交率之研究-以國內某郵購公司為例	利用資料探勘技術中之支援向量機及決策樹預測顧客是否接受促銷活動，透過資料探勘技術找出會員特性和投保意願的關聯性，幫助提升電話行銷的成功率。
蔡志豐	98	應用資料探勘技術於資訊能力檢定之學生分群研究-以某國立大學為例	分析及探勘 91 至 93 學年大學部學生資訊能力檢定資料及其學籍資料，進行知識的發現學生特質與資訊能力是否存在潛藏關聯性。

表 2.3 國內資料採礦相關研究整理 (續)

作者	年度	研究主題	研究結論
李明蕊	99	應用資料探勘技術於顧客關係管理之研究個案	透過分群技術將顧客分為 11 群，並透過敘述性統計找出高貢獻度顧客群與潛在顧客群，最終目的是建立良好的顧客關係管理，培養顧客忠誠度增加顧客再購意願，促進企業的可持續發展，發展成為一種新的競爭優勢。
歐宗殷	99	資料探勘為基礎之零售業銷售預測模式-以連鎖超商鮮食商品為例	以資料探勘的方法論建構銷售預測模式，運用灰關聯分析的特性，將影響銷售量的重要因子篩選出來，並以實際的銷售資料進行預測模式的比較。
陳玉珊	99	以資料探勘作中小企業信用保證貸款之信用預測	首先對選出之資料先進行 Data Preprocessing，並就 Data Imbalance 之問題，利用增加樣本數方法(SMOTE)解決。再以是否承保和承保後是否逾期為切入點，使用 ID3、J48、Naïve Bayesian Classifier、SVM、Logistic Regression、Adaboost 和 Random Forest 7 種探勘方法來探勘出承保之准駁及貸放後逾期之規則。
許郁卿	100	整合資料探勘及紅綠配行銷策略於商品銷售之研究	以顧客交易資訊為基礎，分別針對紅配綠行銷策略提出各模式機制的演算法且提出雛型系統。根據系統的銷售數據，可方便產品經理人進行商品的控管，最後制定出適合使用的行銷策略。
蘇俐文	100	應用資料探勘技術於顧客關係管理之研究	利用資料探勘之分群法將顧客進行分群，並分析不同群組屬性，一不同的顧客特性進行不同的策略方法，找出潛在顧客及提升顧客滿意度、忠誠度及企業獲利能力。

表 2.3 國內資料採礦相關研究整理（續）

作者	年度	研究主題	研究結論
周建成	102	應用交通資料蒐集與發佈設備及資料探勘法協助觀光地區交通管理策略之研究	利用資料探勘法之分類決策樹與關連關係演算法，進行策略執行之門檻值設定與策略宣導之效益分析。研究結果發現車輛偵測器對於預測壅塞發生助益良多，且在路況良好時，利用資訊發布設備進行交通策略宣導可發揮較大效益。
Aggarwal Sonal & Bhatnagar Vishal	102	Technological applications of data mining and virtual reality: a literature survey and classification	作者先進行問卷調查收集資料，接著透過分群的文獻進行分群的動作，根據不同的主題提出分類框架，再進一步資料採礦。根據資料採礦之結果，將假設的與實際的進行比較與探討。
Mohammad Siami & Zeynab Hajimohammadi	102	Credit scoring in banks and financial institutions via data mining techniques	利用資料探勘技術挖掘出信用評分，將顧客分為企業、個人，企業再分為中型與小型企業進行公司的信用評分，並且透過類神經網絡以及變數選擇挖掘出個人的信用評分。

（資料來源：本研究整理）

2.3 本章小結

從資料採礦理論集資料採礦類型的探討中，可以知道資料採礦可以用來幫助企業分析顧客的忠誠度或是流失度以及購買特性，從過去的歷史資料中，經過資料採礦方法的演算歸納，推導出強而有力的法則。本研究即是利用此一方法中的購物籃分析，期望能夠建立顧客購買產品的關聯法則，讓企業根據這些法則提供更客製化的產品以及服務，而在進行此方法之前，也先對資料進行分群以求得更正確的結果。

第三章 研究方法

本章將對本研究的研究方法加以闡述，第一部分首先對研究架構加以介紹，第二部分將對於所使用的應用工具做一描述，第三部分將對於本研究應用資料採礦的分群技術以及關連法則的演算法做解說。

3.1 資料收集

本研究取得之初級資料為南部地區某建材零售商之顧客交易記錄，資料選取的期間為民國 97 年 1 月 1 日至 101 年 12 月 31 日，針對公司與顧客的交易記錄及行為做統計，以其公司之廠商與零售顧客為主要研究對象，每一次交易即為一筆交易記錄，共 368,864 筆交易記錄。

- 資料範圍：民國 97 年 1 月 1 日至 101 年 12 月 31 日
- 研究對象：南部地區某建材零售商

3.2 研究工具

本研究將紙本記錄輸入至 Microsoft Excel 作為初步資料的原始檔案來源，並且以 Microsoft Access 作為資料庫作業平台。因為本研究原始輸入資料為 XLS 格式的資料，在此先用 Excel 作為轉換格式的工具，將轉換後的資料輸入至 IBM SPSS Statistics 20 進行分群之步驟。分群後之群體資料導入 WEKA 應用軟體作為關連法則挖掘的平台。

3.3 研究架構

本研究主要的架構，如圖 3.1 表示，首先取得顧客交易資料，包括顧客明細、產品明細、交易明細以及進貨資料，以顧客代號為主要索引鍵，透過二階段分群法分出忠誠顧客以及高風險顧客，透過關聯法則之分析萃取出顧客之購買行為之知識規則。

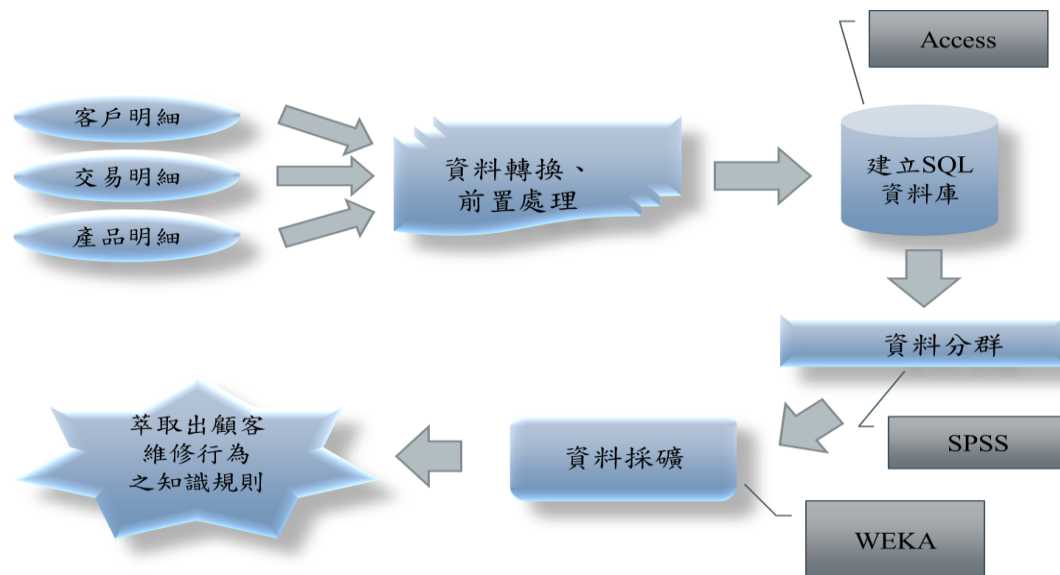


圖 3.1 研究架構流程(資料來源：本研究整理)

就資料採礦技術而言，最廣泛出現在資料挖掘工具裡的技術有決策樹(Decision Tree)、關聯法則(association Rule)、群集(clustering)以及類神經網路(Neural Network)。這些技術都有許多不同的演算法以及實現方式，幾乎每種工具都有些微的差距；然而資料挖掘的範圍太廣，無法在此詳細論述每一種演算法，因此針對本篇研究所運用到的群集演算法以及關聯規則(Association Rule)演算法兩項來做說明。

3.4 二階段分群法

在現今的統計學之中，要將樣本給劃分為幾個集群時，最常使用的一個程序就是集群分析方法(Cluster Analysis)，其中可分為階層式集群分析(Hierarchical Cluster Analysis)和非階層式集群分析(Nonhierarchical Cluster Analysis)，而兩階段方法就是結合了階層式與非階層式的優點，以求得最佳群數。分別說明如下：

1. 階層式集群分析：主要是利用凝聚(agglomerative)與分離(divisive)的方式，將樣本的個體視為一群，然後將相近的個體合成同一群，依次的結合使得群組愈來愈少，最後所有個體結成一群，或是經由相反的程序，將所有個體從同一群，一一分離出為個別的群數。此種方法最大的好處可以使決策者依據實際需要，選擇最適當的群數，作為區隔資料之依據。而群集間距離的判定有下列方法：

(1)最短距離法(nearest – neighbor method)

又稱為單一聯結法(single linkage method)，其採用群集間之最近距離，作為群集間之代表性距離。

$$d_{\min}(G_i, G_j) = \min_{X_i \in G_i, X_j \in G_j} d_{ij} \quad (3.2.4)$$

(2)最長距離法(furthest – neighbor method)

又稱完全聯結法(complete linkage method)，其與最短距離之聚類方式相反，其採用兩群集間最遠點之距離作為兩群集間之代表性距離。

$$d_{\max}(G_i, G_j) = \max_{X_i \in G_i, X_j \in G_j} d_{ij} \quad (3.2.5)$$

(3)群平均法(group – average method)

最短距離法及最長距離法各有其優缺點，觀念上，前者為空間之收縮，後者為空間之擴散。而群平均法又可稱為平均聯結法(average linkage method)，表示兩群間所有元素間距離之平均值。

$$d_{\text{ave}}(G_i, G_j) = \frac{1}{n_i n_j} \sum_{X_i \in G_i} \sum_{X_j \in G_j} d_{ij} \quad (3.2.6)$$

(4)華德法(Ward method)

華德法又稱離差平方和法(squares sum of dispersion method)和最小變異數法 (minimum variance method)，是由Ward (1963) 提出。此法先將每一樣本資料視為一群集，後再依序將群集合併，合併順序的安排是基於方差分析概念，以離差平方和為相似性測度。同群集樣本之間離差平方和最小，群集與群集間的離差平方和最大。假定已將m個變量的n個樣本分成T類， X_{ikt} 表第t類第i個樣本之第k個變量， n_t 表示第t類樣本之個數， \bar{X}_{kt} 表示第t類第k個變量的重心，則第t類中樣本之離差平方和為

$$S_t = \sum_{i=1}^{n_t} \sum_{k=1}^m (X_{ikt} - \bar{X}_{kt})^2 \quad (3.2.7)$$

其中

$$\bar{X}_{kt} = \frac{1}{n_t} \sum_{i=1}^{n_t} (X_{ikt}) \quad (3.2.8)$$

T 個種類之離差平方和為

$$S = \sum_{t=1}^T S_t = \sum_{t=1}^T \sum_{i=1}^{n_t} \sum_{k=1}^m (X_{ikt} - \bar{X}_{kt})^2 \quad (3.2.9)$$

華德法在類和類合併時，是針對某兩類合併後，其 S 為其中增加最小者，予以合併之(高士傑，2001；黃俊英，2003；易任和葉惠中，1990)。本研究使用此演算法代表階層分群方法。

2. 非階層式集群分析：非階層式集群分析中最常使用的方法為K-Means Cluster分群法，與其他的分群法最大的差異在於：若樣本的分群數為已知的情形之下，則K-Means Cluster分群法可以將所有的樣本分為指定的群數，而且K-Means Cluster分群法比較不容易受到不適當之屬性（構面）、異常點與不同的相似度（距離）的影響，因此本研究採用K-Means Cluster分群法與SOM的分群法做比較。

K-means又稱為Forgy's algorithm(Jain,2001)，其主要概念是以反覆疊代的動作，逐次降低目標函數值，直到目標函數不再變化，即代表分群結束。簡單來說，K-means演算法是一種將n筆輸入資料依照其資料特性分群為K類的分群演算法，其中K為一正整數，其演算法的目標在於求取各個輸入資料與其相對應分群中心點距離平方和之最小值。

主要的演算步驟如下：

步驟1：任意由n個資料物件中，隨機選取k個物件視為k個群聚的中心。

步驟2：對於所有n個物件，找其最近似的群心(以距離近者相似度較高)，然後將該物件歸到最近似的集群。

步驟3：計算目標函數(式3.2.10)，若未改變，表示沒有任何物件變換所屬群聚，可以結束分群，若有改變，則進行步驟4。

步驟4：重新計算新的集群中心點(群聚內各物件的平均值)，回到步驟2。平方誤差標準(Square-Error Criterion)之計算如下：

$$E = \sum_{i=1}^k \sum_{p \in c_i} |p - m_i|^2 = 0 \quad (3.2.10)$$

(m_i 代表群聚 i 的中心， p 是群聚 i 內的物件， c_i 則代表群聚 i)

K-means群聚法簡單且有效，其分群結果佳、誤差率低，在屬性增加時仍保持較低的錯誤率(Balakrishnam,1994)，但K-means在使用上的困擾為必須事先設定分群數目 K ，但一般較難從原始資料中判斷適當的群數，且在使用者主觀的認定並不斷測試下，容易造成客觀度不足的問題。因此發展出選取K-means初始群數的方式，分為四種(Pena et.al ,1999)：

(1)Anderberg隨機法：將資料透過隨機方式，分割成 K 個群組。

(2)Forgy方法：參考Anderberg的方法，由資料庫中隨機選取 K 個種子，藉著最近的種子分配到所代表的集群。

(3)Macqueen方法：由資料庫中隨機選取 K 個種子，依據其順序指派剩餘的種子到最接近群心的集群，經過每一個指派，所有群心重新計算完成。

(4)Kaufman方法：最初群數的決定來自連續代表種子的選取，直到 K 個集群都被找到。

3. 兩階段方法：Sgarma(1996)建議如果要求的更精確的分類結果，可以同時使用分層法與非分層法。第一階段以華德法做分群，以決定群組個數，第二階段再以K-Means進行群集，以二階段方法目的是由於第一階段華德法是屬於分層式群集分析，當二個個體一旦被分在一群，則其後永遠就在同一群，而此時在第二階段所使用K-Means就是彌補此種缺點，以達到最佳組內同質、組間異質的群數。

此種分析的手法就是透過兩階段分群方法，修正了分層式集群法中對於集群分析一旦形成，就無法對併入不適當集群觀察值進行重新分群的缺點。同時也克服了非層級式集群法需事先決定集群數目與集群中心點的問題。

3.5 關聯法則(Association Rule)

挖掘關聯式法則(Han, J., and Kamber,2001)為資料探勘中相當重要的一個議題，一般交易資料庫中都儲存著數量龐大的交易資料，而每一筆交易

記錄也都記載著相關的項目(item)，包含使用者相關資料及交易的項目或時間等；而資料探勘的技術便可從這些大量的資料中，挖掘出各個項目之間的關聯法則。

關聯規則是指由於某些事件的發生而引起另外一些事件的發生，發現關聯的目的是為了要在交易資料庫中發現項目之間的關係，找出同一個交易中隱含其它項目之存在的項目。其起源在於分析市場購物籃資料(Market Basket Data)中的大量商品集合(Large itemsets)的關聯程度，因此，也稱為市場購物籃分析(Market Basket Analysis)。關聯法則通常應用在市場購物籃分析上，藉由銷售點系統所紀錄的消費者購買紀錄，分析出購買商品之間的關聯性，根據這些關聯性，可以了解顧客的消費習慣，知道顧客是誰、買了些什麼以及為什麼這樣買。藉由此些資料，企業可以得知哪些商品是會被同時購買，及其被購買的相關先後順序；而這些資訊對於企業在做產品銷售計劃及決策時相當幫助。

關聯法則的主要目的是要找出資料項目間的關聯性，其型式為 $X \rightarrow Y$ ，就是當客戶買了某項產品 X 時，同時會去購買 Y 產品，這兩種產品 X、Y 之間就是一種關聯。其中 X 與 Y 分別表示在資料庫中不同的資料項目組，也就是若 X 出現時，同時也可能出現 Y。例如，顧客買了牛奶和麵包時，也會購買果醬；其表示方式如下：

牛奶&麵包 \rightarrow 果醬 [support=40% ， confidence=70%]

40% 和 70% 則是代表該條規則的支持度(Support)和信賴度(Confidence)。也就是在所有交易資料中，會購買牛奶和麵包的人佔了總交易量的百分之四十；而在這百分之四十的交易中，會再購買果醬的人佔了百分之七十。這種交易行為的百分比，必須要高於某個程度的比例這條關聯性的法則才算成立。因此以下介紹關連法則裡面重要的參數名詞：

- 項目組(Itemset)

相異物品的項目組。當一筆紀錄中有{A, B, C}三個項目，則項目組有{A}、{B}、{C}、{A, B}、{B, C}、{A, C}、{A, B, C}，若物品項目有 n 項，則項目組有 $2^n - 1$ 種。因此，當項目有很多時，項目組是很可觀的。

- 支持度(Support)

即某項目集在資料庫中出現的頻率(frequency)，通常以 $Support(X)$ 表示之， X 為 item set。例如：若資料庫中有 100 筆交易紀錄，其中有 40 筆交易發現有購買牛奶，則此 item set 的支持度為 $40/100=40\%$ 。其支持度越高，代表這個 item set 越值得深入探討

- 信賴度(Confidence)

指兩個項目組之間的信賴程度，通常以 $Support(X \cap Y) / Support(X)$ 表示之， X, Y 為 item set。也就是在 X 出現的機率下， Y 也會出現的條件機率

- 支持度門檻值(Min Support)

若某個項目組的支持度太低，則不具參考性，因此，在進行關聯法則前會先訂定一個支持度的門檻值(threshold)，此支持度的門檻值稱為最低支持度(Min Support)。若項目組超過 Min Support 值則稱為高頻項目組(Large itemset, 或稱為 Frequent itemset)。因此，如何去訂定適當的 Min Support 值是很重要的，因為太低的 Min Support 值會探勘出過多的規則而無所適從，太高的 Min Support 值又容易錯失一些具有價值的規則。

- 信賴度門檻值(Min Confidence)

與 Min Support 意義類似。在進行關聯法則前亦先訂定一個信賴度的門檻值，此信賴度的門檻值稱為最低信賴度(Min Confidence)。若信賴度太低，則不具參考性，且在探勘過程中，未達到 Min Confidence 的關聯法則將會被刪除。

Megiddo and Srikant(1998)指出支持度門檻值及信賴度門檻值之設定是很重要的，當支持度門檻值設太低時，會將重要性較低之項目也包含進來，而設太高又怕因此而失去某些重要規則。信賴度門檻值之設定太低則較不具意義，Han and Kamber(2001)認為門檻值的設定需依據使用者的需求而定。

因此關聯法則是否可被保留使用是透過以上係數的設定，必須要滿足支持度是否超過支持度門檻值(Min Support)及信賴度是否大於最小的信賴度門檻值(Min Confidence)而定(M. S. Chen, J. Han, P. S. Yu,1997)(R. Agrawal, R. Srikant,1994)(R. Agrawal, T. Imielinski, A. Swami,2004)。

3.5.1 關聯法則的演算法

所謂關連法則就是某些項目會引發其他項目出現的規則，例如消費者購買雜誌後，通常有 75% 的人會同時買咖啡，此時「買雜誌→買咖啡」便是一個法則。而法則的挖掘，就是在龐大的資料中，把一些資料項目的相關性找出來的方法。以交易資料庫為例，每天均有大量的交易發生，經年累月累積下來的龐大資訊，無法由人進行分析來找出商品的關連性，然而這些交易記錄事實上隱含了許多有用的資訊(如消費者購買習慣)在裡面，如果能運用適當的方法將它找出來，便可能發現商機，創造利潤，挖掘關連法則的演算法也就是在這種需求驅使下所產生出來的方法。以下就本研究所使用的 Apriori 演算法做一介紹。

3.5.2 Apriori 演算法

Agrawal 等學者於 1994 年提出了 Apriori 演算法(R. Agrawal, R. Srikant,1994)，目前已是探討關聯法則最具代表性的演算法，其後雖又許多針對不同情況所提出的演算法，但大多都依據 Apriori 加以延伸改進。Apriori 演算法的探勘過程主要分成兩個階段：

1. 辨識常出現的項目集合

先找出滿足最小支持度的所有項目組，也就是在交易資料庫中包含項目組的交易數量必須大於或等於所指定的最小支持數量。然後，這些滿足最小支持數量的項目組就稱之為高頻項目組(large itemsets or frequent itemsets)，若一個項目組包含有K個項目，稱之為K項目組(K-itemsets)；若某K-項目組滿足最小支持數量，即稱之為高頻K項目組(large K-itemsets)。

2. 產生關聯法則

根據前階段所找出的高頻項目組及以最小信賴度為條件，計算出所有符合的關聯規則；例如ABC為高頻3項目組，A、B、C \subseteq I，假如關聯規則AB \rightarrow C滿足最小信賴度，則此關聯規則成立。Apriori演算法表示式如下：

$$L_1 = \{\text{Large 1-itemsets}\};$$

For (k=2;Lk-1 \langle >;K++)

{CK=Set of Now Candidates;

```

for all transactions t D
for all K-subsets s of t
if (s ε
CK) s.count++;
Lk={cεCK | c.count>=sinsupp}}
Set of all large itemsets = ∪ K LK;

```

舉個例子來說明 Apriori 演算法推導關聯法則的產生過程。在資料庫 D 中有 4 筆已排序的交易紀錄，如圖 3.所示。另外設 Min Support=50%，也就是超過 $4*50%=2$ 筆的資料可視為 Large itemsets；Min Confidence=80%。在第一次掃描時，如圖 3.2 所示，找出資料庫中所有 1-itemset 並 count 出現次數放在 C1 中，在 C1 中 count 次數超過 Min Support(超過 2 筆紀錄)，則視為 Large itemsets，把結果放在 L1 中。

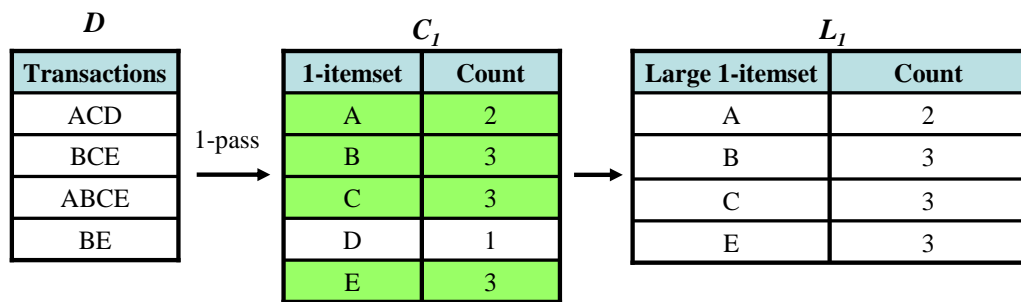


圖 3.2 高頻項目組之產生過程(第一次掃描)

接下來的掃描，如圖 3.3 所示，用聯集(join)的方式，如第二次掃描的過程，即將 $L_1 * L_1$ (*為 join 的運算元)，來產生 C₂ 的 candidate itemsets，並篩選(prune)出超過 Min Support 的 itemsets，放到 L₂；直到 join 出來無法產生新的 Large itemsets。

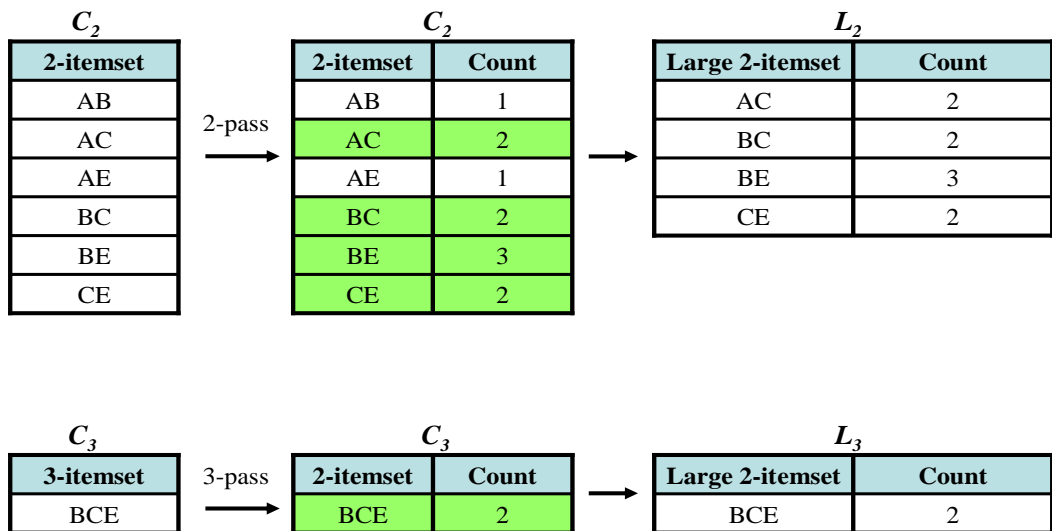


圖 3.3 高頻項目組之產生過程(第二次以後的掃描)

當找完所有的 Large itemsets 後，將 $L_1 \sim L_3$ 中所有的 itemsets 彙整，查若 $(k-1)$ -itemsets 為 k -itemsets 的子集合，則將 $(k-1)$ -itemsets 刪除；例如， $\{A, B\}$ 、 $\{A, D\}$ 、 $\{B, D\}$ 是 $\{A, B, D\}$ 的子集合，所以 $\{A, B\}$ 、 $\{A, D\}$ 、 $\{B, D\}$ 會被刪除。最後加入 Min Support 來考慮，所得之 Large itemsets，如圖 3.4 所示。

Large itemsets	Count
BCE	2
AC	2

圖 3.4 最後所得高頻項目組

第四章 實證分析

4.1 資料採礦

根據文獻探討過的資料採礦流程逐一進行採礦的步驟如圖4.1，先確立採礦的目標，了解其領域，界定資料範圍，然後進行資料的處理、轉換，接著做採礦的動作，將探勘出來的結果做出分析並解釋，萃取出知識。

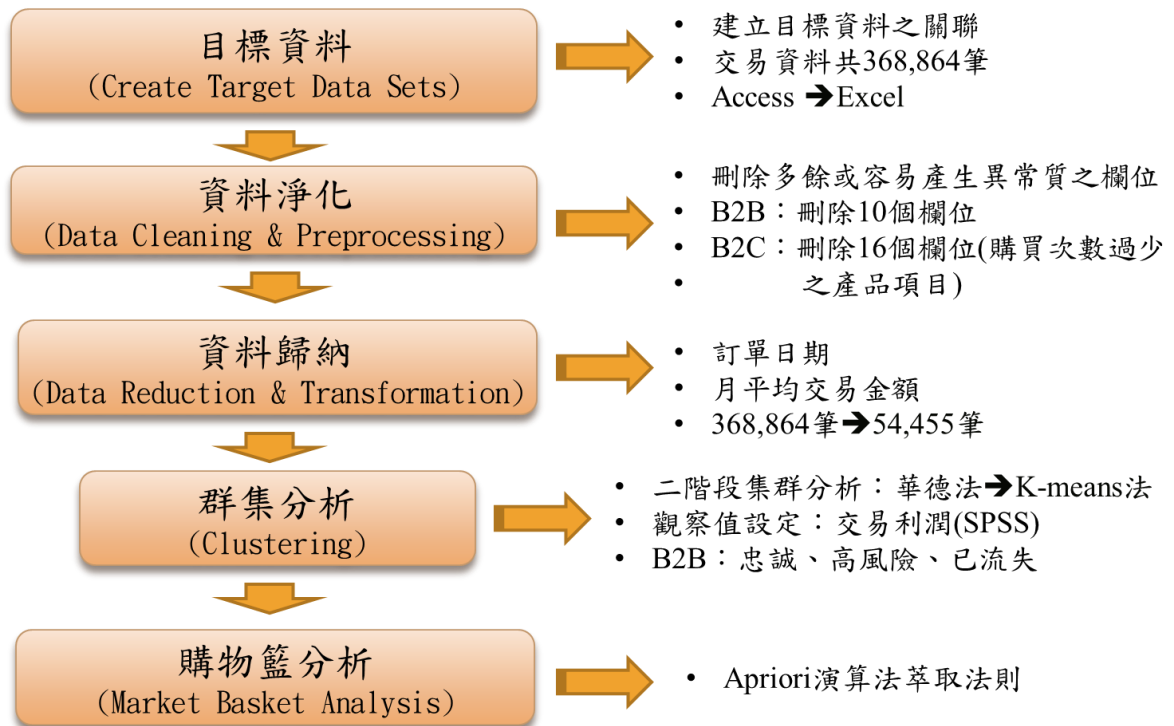


圖 4.1 資料採礦流程圖(資料來源：本研究整理)

4.1.1 目標資料(Create Target Data Sets)

為了能夠正確地挑出適當的資料集合來做採礦的工作，必須先對於取得之資料做了解，假設對於資料來源的應用領域沒有足夠的認知以及充分之了解，往後的工作將會使得最後萃取出來的知識不具可用性。

本研究透過建材零售商取得的顧客交易資料表包括下表4.1所示顧客明細、表4.2產品明細以及表4.3交易明細這三大項的相關資料，樣本共有368,864筆資料，51個欄位。

表4.1 顧客明細

項目	欄位名稱	描述	項目	欄位名稱	描述
1	客戶編號	客戶之唯一鍵值	4	聯絡轄區	公司郵遞區號
2	性別	客戶性別	5	公司地址	公司地址
3	公司型態	公司營業之型態	6	公司年齡	公司年齡

表4.2 產品明細(B2B)

項目	產品名稱	描述	利潤	項目	產品名稱	描述	利潤
1	南亞門片	門片	800	31	木紋門	門片	500
2	PC	門片	1200	32	小百葉	通風片	50
3	H5	門片	1200	33	貼皮胡桃	門片	500
4	2P 米	門片	1200	34	木紋彩雕	門片	500
5	2P 柚	門片	1200	35	南亞門框	門框	200
6	A7	門片	1200	36	烏心石框	門框	200
7	K6	門片	1200	37	越檜玻璃門	門片	200
8	胡 2P	門片	1200	38	南檜玻璃門	門片	600
9	H12	門片	1200	39	越框 A 級	門框	300
10	H3	門片	1200	40	越框 B 級	門框	300
11	保白	門片	500	41	花其松 A 級	門框	300
12	代工白橡	門片	200	42	柚木門框	門框	500
13	木紋彩繪	門片	500	43	柳安紗門	門片	500
14	南亞門組	門框+ 門片	300	44	柳安紗窗	窗戶	500
15	拼花彩檜	門片	500	45	南檜百葉門	門片	500
16	柚木門框	門框	1000	46	胡桃長條 百葉	通風片	400

表4.2 產品明細(B2B)(續)

項目	產品名稱	描述	利潤	項目	產品名稱	描述	利潤
17	香山百葉門	門片	800	47	柚木框 3 號線	門框	1000
18	香山門柱	補強用	0	48	梧桐木門片 胡桃	門片	500
19	香山玻璃門	門片	500	49	南檜原木門 大門片	門片	1200
20	越檜 A 級	門框	350	50	南檜門片	門片	500
21	南檜原木門	門片	500	51	MH 系列	門片	500
22	塑門框	門框	200	52	柚色實木門	門片	500
23	烏心石門框 3 號線	門框	400	53	咖啡色 通風門	門片	1400
24	烏心石門框 雙線	門框	250	54	硫化銅門框	門框	500
25	越檜紗門	門片	500	55	木紋雙色	門片	2400
26	南亞塑鋼 門框	門框	200	56	硫化銅門片 全組	門框+門 片	1000
27	花其松和式 門組	門片	1200	57	硫化銅門片	門片	800
28	烏心石 和式窗	窗戶	300	58	白鐵門	門片	600
29	柳安木 百葉片	通風門 片	200	59	木編瓦門扇	門片	1000
30	越檜和式 門組	門片	3000	60	硫化銅門片 風梨花	門片	1200

表 4.2 產品明細(B2C) (續)

項目	產品名稱	項目	產品名稱	項目	產品名稱
1	鞋櫃	12	強化玻璃	23	白鐵門組
2	TV 櫃	13	塑膠地磚	24	包圓柱
3	展示櫃	14	實木地板	25	櫃台
4	廚房拉門	15	拉門	26	洗手台
5	廚房櫃	16	置物櫃	27	儲藏室隔門
6	衣櫥	17	壁板	28	浴室門
7	床頭組	18	更衣室	29	鋁窗
8	書桌	19	衣架	30	浴室天花板
9	書櫃	20	屏風	31	化妝台
10	浴室拉門	21	木門組		
11	天花板	22	軌道抽屜		

表4.3交易明細

項目	欄位名稱	描述	項目	欄位名稱	描述
1	施工地區	台南市行政劃區	8	流失狀態	流失、未流失
2	施工人員	該次施工人員	9	施工類型	住宅、店面
3	購買產品	同等產品代號	10	建議項目	建議補強項目
4	公司年齡	顧客公司經營年數	11	進貨成本	該筆訂單進貨之成本
5	付款方式	現金、支票	12	銷售金額	該筆訂單之總交易金額
6	訂單日期	下單之日期	13	利潤	銷售金額-進貨成本
7	交易金額	每月平均交易金額	14	抱怨次數	顧客抱怨次數

為了方便資料後續的處理作業，必須做資料轉換的工作。因此本研究利用 Microsoft Office Access 來做關聯式資料的轉換，如圖 4.2 使其呈現為 Excel 的檔案模式以利方便作業，轉換後的資料表如圖 4.3 表現。

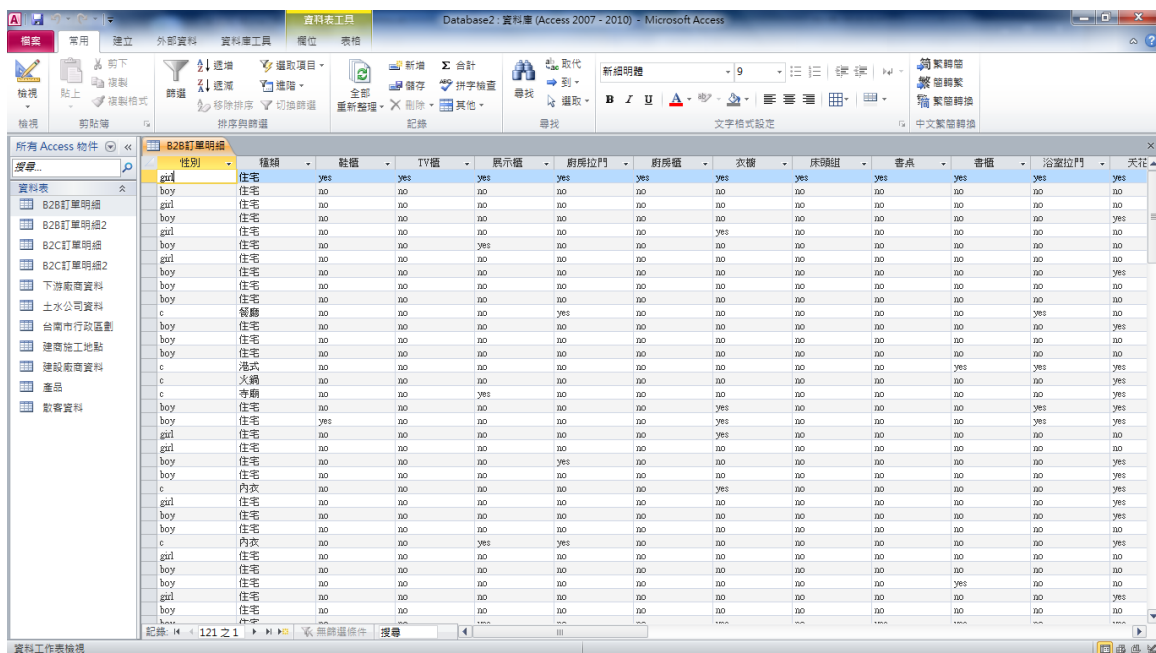


圖 4.2 Access 資料表

The screenshot shows the Microsoft Excel interface with the converted data table. The table has the following columns: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W. The data rows correspond to the Access table, with columns 1-24 representing the features: 性別, 種類, 鞋櫃, TV櫃, 展示櫃, 廚房拉門, 廚房櫃, 衣櫥, 床頭組, 書桌, 書櫃, 浴室拉門, 天花板, 強化玻璃, 塑膠地磚, 實木地板, 拉門, 置物櫃, 壁板, 更衣室, 衣架, 屏風, 木門組, 軌道.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
1	性別	種類	鞋櫃	TV櫃	展示櫃	廚房拉門	廚房櫃	衣櫥	床頭組	書桌	書櫃	浴室拉門	天花板	強化玻璃	塑膠地磚	實木地板	拉門	置物櫃	壁板	更衣室	衣架	屏風	木門組	軌道
2	girl	住宅	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
3	boy	住宅	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no
4	girl	住宅	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no
5	boy	住宅	no	no	no	no	no	no	no	no	no	no	yes	no	yes	no	yes	no	yes	no	no	no	no	no
6	girl	住宅	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	no	yes	no	yes	yes	yes	no	no
7	boy	住宅	no	no	yes	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	yes	yes	yes	no	no
8	girl	住宅	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	yes	no	no	no	no
9	boy	住宅	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	yes	no	no	no	no	yes	no	no
10	boy	住宅	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	yes	no	no	no	no	no
11	boy	住宅	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no
12	c	餐廳	no	no	no	yes	no	no	no	no	no	yes	no	no	no	no	no	no	yes	no	yes	no	no	no
13	boy	住宅	no	no	no	no	no	no	no	no	no	no	yes	no	yes	no	no	no	no	no	no	yes	no	no
14	boy	住宅	no	no	no	no	no	no	no	no	no	no	no	yes	no	yes	no	no	no	no	no	no	no	yes
15	boy	住宅	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	yes	no	no
16	c	港式	no	no	no	no	no	no	no	no	yes	yes	yes	no	no	yes	no	yes	yes	no	no	yes	yes	no
17	c	火鍋	no	no	no	no	no	no	no	no	no	no	yes	no	yes	no	no	no	no	yes	no	yes	no	no
18	c	寺廟	no	no	yes	no	no	no	no	no	no	no	yes	yes	yes	no	no	no	yes	no	no	yes	yes	no
19	boy	住宅	no	no	no	no	no	yes	no	no	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	no	yes	no
20	boy	住宅	yes	no	no	no	no	yes	no	no	yes	yes	yes	no	no	yes	yes	no	yes	no	yes	no	no	yes
21	girl	住宅	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
22	girl	住宅	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	yes	no	no	no	no	no	no
23	boy	住宅	no	no	no	yes	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	no	yes	yes	no
24	boy	住宅	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	yes	no	no	yes	no	no

圖 4.3 Excel 資料表

4.1.2 資料淨化(Data Cleaning & Preprocessing)

一般的原始資料常有無法預期的錯誤或重複的情況發生，此時便要先把資料前置處理，去除錯誤或不一致的資料；本研究中所需剔除資料的欄位有下列幾項：

1. 顧客明細部分

(1) 公司資料：公司資料包括了公司地址及聯絡轄區，一般來說顧客是否下訂單主要依據為施工地點的區域，故公司聯絡轄區不列入採礦範圍。

2. 產品明細部分

(1) 進貨產品明細：6項進貨次數極少之產品項目，對於採礦過程容易產生極端值，故予以刪除。

(2) 銷售產品明細：B2C的零售項目中有16項產品顧客訂購次數較少，再次訂購機率較低，且對於採礦過程容易產生極端值，故予以刪除。

3. 交易明細部分

(1) 利潤：當顧客下訂單，公司會記錄進貨成本與銷售金額，透過計算後新增利潤欄位，且刪除進貨成本與銷售金額此二欄位。

經過初步資料淨化，剩餘之欄位共有 32 欄。

4.1.3 資料歸納(Data Reduction & Transformation)

這個階段是要對資料做簡化或轉換的工作，將資料轉換為適合進行採礦的格式。為簡化資料的複雜性，經過處理的有下列欄位：

- 訂單日期：該筆交易所代表顧客下訂單之日期
- 交易金額：該筆訂單的總銷售金額

因本研究資料是關聯型式，客戶資料共有 54,455 筆，每一筆代表一位客戶，而交易資料則有 368,864 筆，每一筆代表一次的訂購記錄，因此接下來的工作是藉由資料的轉換，以客戶編號為主要索引鍵，將 368,864 筆的資料以增加欄位的方式來縮減為 54,455 筆，如此可大量的減少資料的數目，減輕以後工作時，演算法執行速度的負擔，亦可保留住原始資料的完整性。轉換方式如下：

- 如果顧客在一天內下訂單次數有 2 次以上，必須考慮到一位顧客有可能有很多筆交易紀錄，若以原始型態交易記錄為索引鍵時，資料過於繁瑣與龐大，因此利用客戶編號來當作索引，使其資料型態是以每筆資料代表一位客戶在本研究範圍期間有下過訂單。因此抽出訂購日期此欄位，轉換為該年訂購是否訂購，接著抽出交易金額此欄位，轉換為該年月平均交易金額，如圖 4.4 以及圖 4.5 所示。

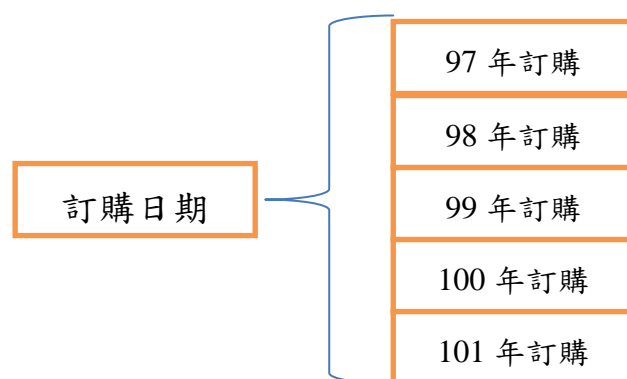


圖 4.4 訂購日期轉換

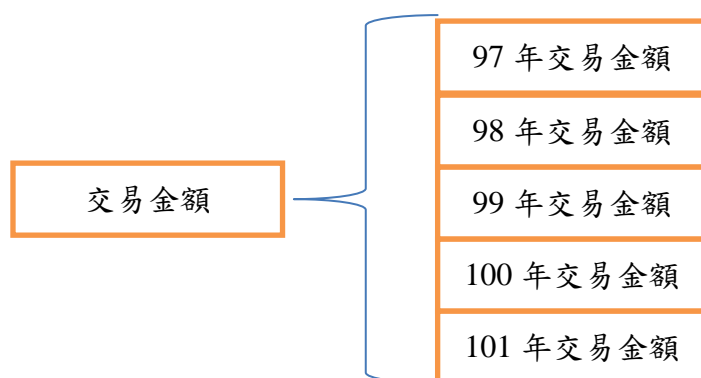


圖 4.5 交易金額轉換

有些欄位裡的變數屬於文字型態，在後續上的處理並不是很方便，因此需做轉換。在施工類型、付款方式以及公司型態此三項欄位裡之變數選項較少，將其文字型式轉換為數字代表。

4.1.4 群集分析(Clustering)

經過上述轉換已經 368,864 筆資料轉換為 54,455 筆，而欄位由原本 51 個欄位轉變為 33 個欄位。由前一章所提及，在處理龐大資料時，可先透過分群的方式，將相似度高的資訊聚集在一起，不相似的資訊分開，進一步從中去找尋規則和關聯。因此透過二階層分群法(華德法→K-means)，

將 B2B 的顧客作分群。

本研究使用 spss20.0 統計分析軟體來對此進行分析，首先進行二階段分群法的第一階段，也就是階層式分群法中的華德法(Ward method)，此方法主要是用來確定資料適合的分群數目，計算區間採用歐基里德直線距離平方(Squared Euclidean distance)來觀察379個顧客由379群凝集至1群的凝集過程，結果如表4.4所示。

表4.4 群數凝聚過程

集群數目	凝聚係數	係數增量	係數增加率
5	1,790,199,468	1,075,014,155	60.04%
4	3,610,213,106	1,820,013,638	50.41%
3	7,957,939,555	4,347,726,449	54.63%
2	13,306,905,106	5,348,965,551	40.19%
1	88,411,584,953	70,104,679,847	79.29%

從資料表中可以看出，全部資料從 379 群凝集至 1 群的過程中，其變異係數(coefficient)在二群要凝集為一群時差異最大，當凝聚係數突然增大，表示組內誤差突增，此時應停止凝聚的程序，該組則為最佳分群數。(陳宜欣，2006)

4.1.5 K-means 法

在初步把資料分為二群之後，利用 K-means 方法對於資料作一些觀察與驗證的動作。以第一階段分群所得到的結果二群做為分群數，剛開始 K-Means Cluster 會隨機指定各個群組中心，並且進行計算初始群集的中心點，隨著計算疊代的變化，各集群的商品向量會向中心點收斂而改變群組的個數，分群結束後得到分群結構表，如圖 4.6，且利用 Excel 統計及繪製圖表檢定是否符合分群需求，如圖 4.7，從圖中可以看出設定利潤為分群設定觀察值分出的忠誠顧客雖僅佔了總顧客人數的 19%，但為公司帶來的利潤已經高達總營業額中的 80%，利潤來源相當集中，加上公司並無法保證忠誠顧客是否會流失，因此必須增加忠誠顧客之比例，減少顧客流失率。因此接著透過關聯度分析觀察忠誠顧客與高風險顧客的關聯度法則是

否有雷同之處，藉以提升忠誠顧客之比例。

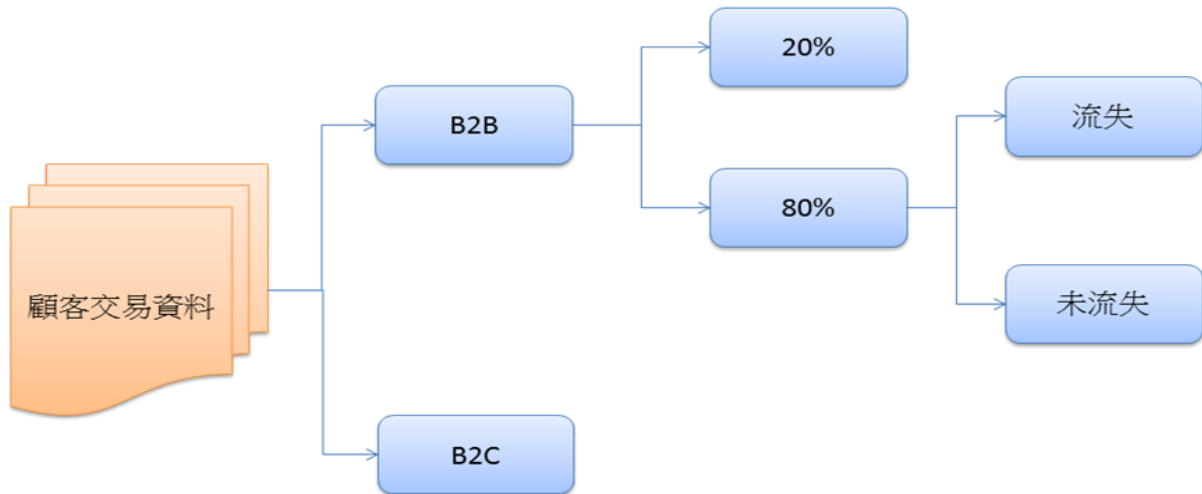


圖 4.6 分群結構表

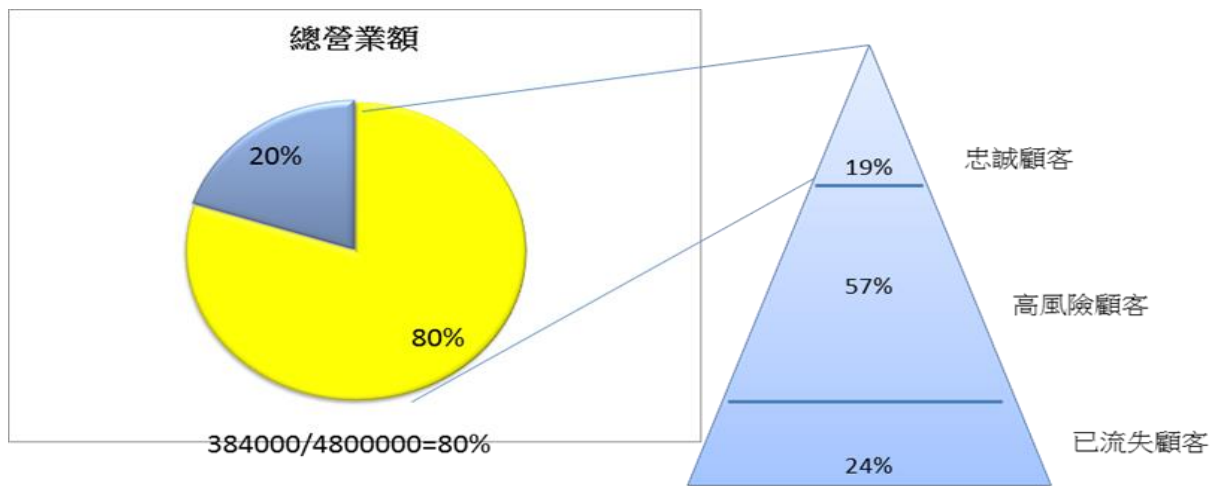


圖 4.7 Excel 檢定表

4.1.6 購物籃法則分析(Market Basket Analysis)

當利用資料挖掘的技術從資料庫中挖掘出許多的知識和法則時，由於產生的知識、法則的數量很多，但其中的有些的知識、法則對使用者而言是具有重覆性的、直覺的或無意義的。故在眾多關聯法則中，須找出具有意義的關聯法則，以提供決策者參考依據。依據 Klemettinen(1994)等學者指出，並不是所有有高信度及高支持度的法則就是有用的，因為有些法則的意義可能重覆或是意料中，有些沒有意義法則是由不相關的屬性所構成，因此關聯法則建立了以後，選出有意義且可解釋的法則。

本研究關聯法則所使用的工具為紐西蘭 Waikato 大學所開發出來的 Waikato Environment for Knowledge Analysis 軟體，使用介面如圖 4.8 所示。

以下將對於挖掘出的關聯法則做一分析。

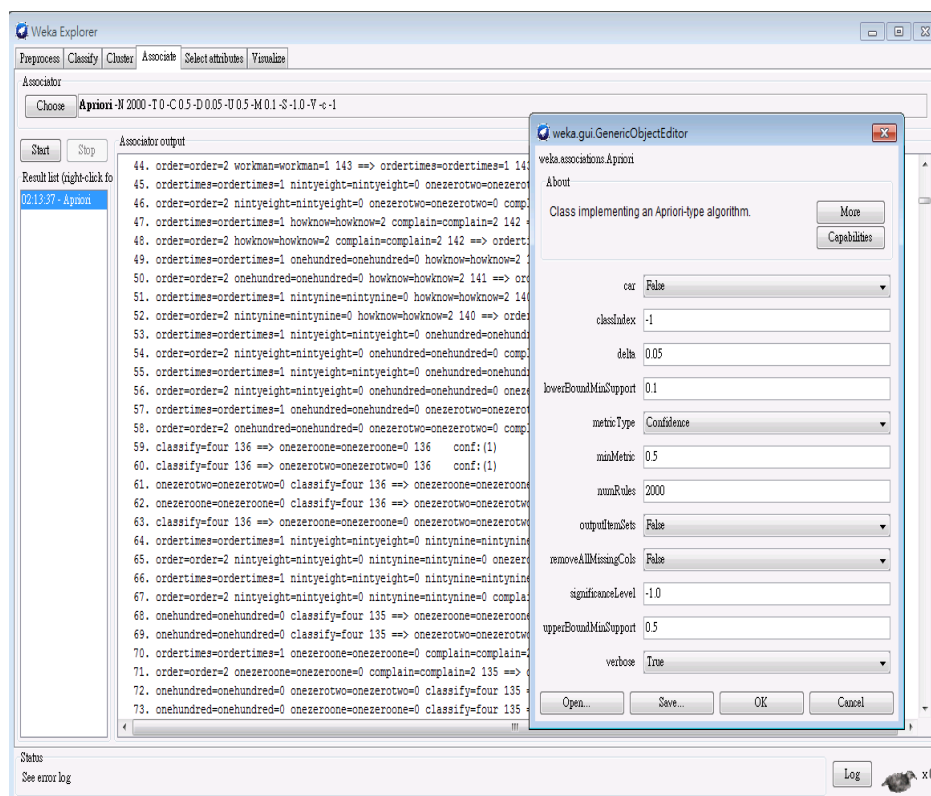


圖 4.8 Weka 使用介面

在做關聯法則探勘時必須先做好參數設定的工作，也就是支持度門檻值(Min Support)與信賴度門檻值(Min Confidence)。此二參數的設定，多憑經驗而得，資料量小時，最小支持度門檻值可比較小，但當最小支持度門檻值設太小時，會使電腦記憶體容量耗用大增而無法執行，所以須依資料之特性來找出關聯法則；反之，最小支持度門檻值應設較高。因為信賴度愈高，表示關係愈強烈。

在軟體的預設值中，支持度門檻值是設定在 1.0，信賴度門檻值在 0.9 的範圍，表示出現的關聯規則是在支持度 1%~100%之間且信賴度要至少 90%的法則才會出現。在這樣的情況下，法則所呈現的大多是屬於無意義的關連，例如：97 年訂購=0 \Rightarrow 97 年交易金額=0 sup:(0.8) conf:(1)；此法則表示 80%的客戶如果 97 年訂購的次數是 0 的狀況，有 100%的可信度該客戶 97 年的交易金額為 0。因此，本研究認為支持度的範圍可做向下修正的調整，經過反覆的測試，支持度在 1%~50%之間，而信賴度至少要在 50%以上此規範下產生的法則在分析上較有實質的意義，才有足夠的採信度來提供給廠商，做為決策依據。以下將有用的關聯法則做一整理及解釋：

1. 忠誠顧客

- (1). 施工區域=5，99 年訂購=1，100 年訂購=1 → 102 年訂購=1
sup:(0.48) conf:(0.94)

在 48%的忠誠顧客群中，若施工區域為中西區、99 年與 100 年有下過訂單，則在 102 年會下訂單的可信度有 94%。

從此法則可以觀察到，若施工地區位於離案例公司較近，則會向案例公司下訂單的機率較高，因此施工地區與下訂單的機率是有關聯度的。

- (2). 購買產品=2，98 年訂購=1，99 年訂購=1 → 102 年訂購=1
sup:(0.37) conf:(0.97)

在 37%的忠誠顧客群中，若購買產品為高級門片類，在 98 年與 99 年皆下有訂單，則在 102 年下訂單的可信度為 97%。

從原始交易資料中發現，部分廠商會固定訂購單價高利潤的產品，而且每年訂購的次數與數量具有周期性，因此累計下來為案例公司帶來的利潤是相當高的。

- (3). 購買產品=1，99 年交易金額=3，公司年齡=10 → 付款方式=2，
100 年訂購=1 sup:(0.38) conf:(0.93)

在 38%的忠誠顧客群中，若購買產品為南亞門片，在 99 年的月平均交易金額為 20000~30000 且廠商公司年齡為 10 年以上，則在 100 年會下訂單且付款方式為支票的可信度為 93%。

從原始資料中發現，有一部分的顧客自身公司的年齡已經超過 10 年且與案例公司下訂單的單筆訂單金額甚高，因此付款方式往往以支票的形式付款，而此種顧客公司經營時間久且有相當的營業規模，因此有時會接到建設公司的大型案子，相對向案例公司下訂單的金額也會提升，因此為案例公司的忠誠顧客。

- (4). 施工區域=6，98 年訂購=1，100 年交易金額=2 → 購買產品=3
sup : (0.39) conf : (0.89)

在 39%的忠誠顧客群中，若施工區域為安南區、98 年有下過訂單且在 100 年交易金額為 10000~20000，則購買產品為高級門片類的可信度為 89%。

從原始資料中發現，部分的顧客雖然施工區域近但不是每年都會下訂單，但是若該年有向公司下訂單則該年的月平均交易金額通常會在 20000 上下，而且訂購的產品項目通常以高單價為主，因此為案例公司帶來之利潤相當高，而此種顧客也被歸類到忠誠顧客群中

(5). 購買產品=35，施工區域=3，101 年訂購=1，101 年交易金額=3 → 付款方式=2 sup : (0.38) conf : (0.8)

在 38% 的忠誠顧客群中，若購買產品為南亞門框、施工區域為南區、101 年有下過訂單以及 101 年平均每月銷售金額為 20000~30000，則付款方式有 80% 為支票付款。

從原始交易資料中發現，購買門框類產品為公司所帶來的利潤比其他產品低，若依照看單筆訂單帶給公司的利潤多寡來區分忠誠顧客是不公平的，由於部分廠商訂購次數每月至少 10 次，每次的數量至少 5 片，每個月為公司帶來的利潤甚至超越其他訂購高利潤產品的廠商，此種顧客對公司有著相當高的忠誠度。

(6). 99 年訂購=1，100 年訂購=1，101 年訂購=1 → 102 年訂購=1，公司年齡=10 sup : (0.45) conf : (0.9)

在 45% 的忠誠顧客群中，若在 99、100 及 101 年皆有以下過訂單，則顧客公司年齡為十年且在 102 年會向公司下訂單的可性度為 90%。

此種類型之顧客由於顧客本身也是經營多年的公司，每年在訂貨的時候會習慣與固定的廠商下訂單，降低風險，因此此種類型之顧客已經屬於案例公司的老顧客及忠誠顧客。

(7). 98 年交易金額=3，99 年訂購=1 → 100 年訂購=1，公司型態=1 sup : (0.37) conf : (0.87)

在 37% 的忠誠顧客群中，若在 98 年月平均交易金額為 20000~30000 且在 99 年有下過訂單，則公司類型為建設公司及會在 100 年下訂單的可性度為 87%。

從原始資料中發現，建設公司每年下單次數雖然比其他類型的公司少，但是訂購量卻比其他公司來的多，使案例公司不僅可賺零

售差額，還可有裝潢的營收。因此此種類型的顧客可帶給公司相當多的利潤，故分類為忠誠顧客。

根據以上七條忠誠顧客的關聯法則，可以歸納成七條可讓風險顧客提升為忠誠顧客之要素，如下表。並且根據歸納出的七條要素與風險顧客關聯度分析做比較，觀察是否有雷同之處藉以提升忠誠顧客之比例。

- 施工地點以南區、中西區以及安南區為主者
- 下單次數少，訂購產品項目利潤高者
- 付款方式大多為公司開支票，此種顧客偶爾會接到大案子
- 每月平均銷售金額為 20000
- 利潤不多，但是訂購次數多者
- 從 97 年至 102 年每年都有持續訂購者
- 顧客公司為建設公司

2. 風險顧客

(1). 公司型態=1，101 年交易金額=1，施工區域=6 → 102 年訂購=1，流失狀態=1 sup : (0.27) conf : (0.86)

在 27% 的未流失顧客群中，若公司型態為建設公司、101 年的月平均銷售金額為 10000、施工區域為安南區，則 102 年會訂購的機率為 86%。

從原始資料中發現，部分建設公司不屬於忠誠顧客群的原因可能是因為此類建設公司剛起步亦或是建案在別地區時向別間建材公司訂購，因此不是每個月都會有訂單。但若施工地點為安南區，符合成為忠誠顧客的兩項要素，因此此類型顧客有機會提升為忠誠顧客。

(2). 97 年訂購=1，98 年訂購=0，99 年訂購=1，99 年交易金額=2 → 100 年訂購=1，流失狀態=1 sup : (0.43) conf : (0.89)

有 43% 的未流失顧客群中，若公司在 97 年有下訂單、98 年未下訂單、在 99 年的月平均交易金額為 20000，則此公司在 100 年會下訂單此行為模式有 67 的可靠性。

從原始資料中發現，此種類型之顧客會根據施工地區不同而選擇對自己有利的建材廠商，對公司來說是屬於流動率高的顧客。

(3). 公司型態=2，98 年訂購=1，施工區域=16 → 100 年訂購=0，流失狀態=2 sup：(0.18) conf：(0.61)

有 18%的已流失顧客群中，若公司型態為土水、在 98 年有下訂單、施工地區為歸仁，則此公司在 100 年會下訂單此行為模式有 61 的可靠性。

從原始資料中發現，此種類型之顧客會根據施工地區不同而選擇對自己有利的建材廠商，若下訂單數量不多而且施工地區又有一段距離的話，公司會為了運送成本問題而提升銷售價格，造成顧客之流失。

接著將忠誠顧客的法則與風險顧客的法則做比較，觀察是否有雷同之處，進而找出潛在的忠誠顧客，如表 4.5。

表4.5 風險顧客符合忠誠顧客要素比例表

忠誠顧客要素	符合要素比例
顧客公司或施工地點以南區、中西區以及安南區為主者	13%(37 間)
下單次數少，訂購產品項目利潤高者	3%(8 間)
付款方式大多為公司開支票，此種顧客偶爾會接到大案子	5%(14 間)
每月平均銷售金額為 20000	2%(6 間)
利潤不多，但是訂購次數多者	5%(14 間)
從 97 年至 102 年每年都有持續訂購者	8%(23 間)
顧客公司為大型建設公司	3%(8 間)

從原始資料中可以發現，雖然某些顧客群符合可能成為忠誠顧客之要素，但是卻無法歸類於忠誠顧客群中，原因可能是在一開始做分群時所設定之觀察值為利潤，因此分群出 19%的忠誠顧客就可提供 80%的利潤。若將觀察值設定為銷售額或訂購次數，可以發現分群後的忠誠顧客比例提升至 34%，但是所提供的利潤占總利潤的 82%，可以發現到雖然增加的 15%忠誠顧客，但是與第一次分群的 19%做比較後，這 34%的忠誠顧客對公司

的貢獻較不明顯，導致一開始分群時會將符合相同要素之顧客分在不同群集中。

因此符合這些要素的風險顧客群中，有可能顧客本身也能歸類於忠誠顧客，只是貢獻度不夠高，在分群時無法與高貢獻顧客分為同一群，對於這些風險顧客，公司可以透過降價或轉移成本等方式來提升這群顧客的貢獻度。

接著設定 B2C 的支持度門檻與信賴度門檻值，由於裝潢類的顧客在下單前都會先請設計師進行設計的事前工作，因此下單的產品項目也比較類似，比較不容易產生無意義的關聯法則，因此在支持度門檻的設定上也可做向上修正的調整，經過反覆的測試，支持度在 1%~75% 之間，而信賴度至少在 90% 以上此規範下產生的法則在分析上較有實質的意義，才有足夠的採信度來提供給廠商，做為決策依據。以下將有用的關聯法則做一整理及解釋：

(1) 分類=1，鞋櫃=1，衣櫥=0，強化玻璃=0，更衣室=1 → 床頭組=1，空間設計=1 sup：(0.45) conf：(0.98)

在 45% 的零售顧客群中，如果施工分類是屬於住宅，有訂製鞋櫃與更衣室，不需要衣櫥與強化玻璃，那他可能也會訂製床頭組且公司會建議顧客與設計師商量空間設計，此行為模式有 98% 的可靠性。

零售客戶若是屬於自宅裝潢，需要強化玻璃的機會就比較小，但是公司可以建議自宅使用別款玻璃。

由於床頭組有收納櫃的設計，因此公司可以建議不訂製收納櫃的客人製作床頭組，既實用又不浪費空間。但是製作櫃子會影響室內空間感的問題，因此公司會建議顧客在詢問設計師的意見。

(2) 分類=2，櫃台=1，收納櫃=1，壁板=1 → 廚房拉門=1，油漆=1 sup：(0.45) conf：(0.98)

在 45% 的零售顧客群中，若施工分類是屬於商店，有訂製櫃台、收納櫃以及使用壁板的話，則有 98% 會訂製廚房拉門且公司會建議壁板顏色可以用油漆與根據室內設計做修改。

在零售顧客群中若分類屬於商店，有大部分是餐廳業者剛開幕需

要裝潢，由於每間餐廳室內擺設不同，所設計的櫃台款式也會不同，因此大都會特別訂製。

若分類屬於商店又是餐廳的話，公司可以建議設置有特色的廚房拉門，不僅可以隔絕廚房與用餐區域，也能為餐廳多添一分設計特色。若是有訂製壁板的店面可以透過油漆的方式且依據店內裝潢更改壁板顏色。

(3) 屋齡=10, TV 櫃=1, 書桌=1, 實木地板=1, 天花板=1 → 床頭組=1, 電燈=1 sup : (0.57) conf : (0.94)

在 57% 的零售顧客群中，若施工的房子屋齡為 10 年，有訂製 TV 櫃、書桌、實木地板及天花板的話，則有 94% 會訂製床頭組且公司會建議可以更換不同款式的電燈。

從原始資料中發現，某些零售顧客若是屋齡已經 10 年或 10 年以上，通常會想進行整修。而通常進行整修的內容為地板、天花板或收納櫃，但在台南市區的舊房子空間並不大，特別是公寓式家庭，因此想訂製收納櫃的家庭公司會建議可以改做床頭組比較不占空間。而有重鋪地板或是天花板裝修的話，公司會建議可以更換燈泡顏色或是燈座款式，避免施工後整體顏色不協調。

第五章 結論與建議

5.1 結論

隨著時代改變，市場已逐漸由生產導向轉變為顧客導向，顧客的維持與流失成了公司在市場上競爭的勝敗關鍵。本研究企圖從顧客交易資料中，將紙本記錄的交易紀錄建構成資料庫，透過分析尋找出顧客消費行為的特徵值，分辨出流失顧客、高風險顧客以及忠誠顧客的特徵，藉以分析如何提升忠誠顧客之比例提出對策，並且針對 B2C 之顧客提供更專業性的知識與服務。

而根據本研究之研究結果，綜合下列結論：

1. 研究目的一

(1). 對於忠誠顧客維持之策略

整體服務策略：由於案例公司的忠誠顧客數量所佔的比例較少，若失去任何一位忠誠顧客則損失嚴重，因此可以嘗試使用整體服務策略，提供顧客全方位的服務組合，包括售前、售中及售後三個服務階段，這種策略幾乎涵蓋了顧客的各種不同需求，同時提升顧客滿意度。

(2). 對於符合可能成為忠誠顧客之要素的風險顧客之策略

功能化策略：由於此類型之顧客所符合的要素並不相同，或是同時符合兩項以上之要素，因此依照符合不同要素之顧客提供不同功能系列之產品或服務滿足不同顧客之需求與消費習慣。

(3). 對於已流失顧客之策略

設計鼓勵消費專案：從分析結果發現部分顧客流失是因為運送成本過高，造成顧客成本壓力的問題，因此可以透過促銷方式使顧客能增加產品數量，藉此轉移案例公司之運送成本。

2. 研究目的二

案例公司可依據實證分析歸納出的法則提供給零售顧客客製化產品、服務與建議，其中建議主要分為三大類：

- 若顧客要求製作置物櫃，由於空間感的設計，公司方面會建議顧客下單前先與設計師討論後再決定。

- 若顧客有訂購壁板且是做為隔板之用途的話，公司方面會建議顧客可透過油漆的方式，將壁板的顏色做更改，使整體色彩更具協調性。
- 若顧客是想將天花板與地板同時做更改與整修的話，公司方面會建議顧客可購買不同款式之燈座或更換不同顏色之燈泡，避免燈光顏色與地板顏色不協調，可能導致房間陰暗等情況發生。
- 對於維持零售顧客之策略：
關係策略：透過公司提供專業性知識、處理顧客意見及抱怨情緒，建立與顧客的依賴關係與長期而穩固的關係，而當依賴關係越深則越信任公司。

5.2 未來研究與建議

本研究對於未來研究或是業界的應用有以下建議：

1. 透過愈大量完整的資料，愈能將資料採礦的功效發揮，以尋找出可靠性越高的法則，獲得更高的應用價值，所以建議企業在建構顧客資料時應注重其資料的完整性。
2. 針對產品區分出更多的定義，協助企業訂定出更適合的行銷策略。
3. 整合智慧型決策系統，提供建議協助決策者擬定適當的商品行銷策略與公司經營方式。
4. 由於僅研究一家個案建材零售商，有代表性不足的問題，因此建議後續研究可繼續採用其他建材零售商，或是比較不同的產品組合銷售方式的接受度。
5. 本研究運用關聯法則來挖掘顧客行為，關聯法則之演算法目前有許多改良式演算法，亦可透過其他的演算法及程式做相關研究。
6. 本研究主要運用顧客的交易行為面的分析變數，進行資料採礦的工作；然而亦有學者認為更必須考量顧客的心理面，才能更清楚掌握顧客對於企業的忠誠度，不過在實務上如何能得到所有顧客的心理面的分析資料，可進行相關的研究。

參考文獻

- 1999 年度台灣「顧客關係管理」運用現狀調查報告”(民 90)，**電子商務導航**，第 2 卷，第 13 期。
- 元智大學商業智慧實驗室(民 90)，<http://140.138.148.80/1507a/DM.htm>
- 尹相志(民 93)。資料採礦-網際網路應用與顧客價值管理，臺北市：維科圖書有限公司。
- 歹年冬-資料採礦度小月 (民 96)。動腦雜誌，304，44-46。
- 安迅資訊公司(民 88)。整合企業經營策略與顧客關係管理。電子化企業經理人報告，20-25。
- 吳旭智、賴淑貞譯(民 90)。Data Mining 資料採礦理論與實務顧客關係管理的技巧與科學(原作者：Michael J.A. Berry，Gordon S. Linoff)。
- 李昇暉(民 89)。以資料採礦深化顧客關係管理。電子化企業經理人報告，7，37-42。
- 唐林林(民 92)。企業建構顧客關係管理之研究—以機械業為例(碩士論文)，中華大學科技管理研究所，新竹市。
- 夏載(民 90)。剖析資料採礦在顧客關係管理中的應用。電子化企業經理人報告，20，71-75。
- 陳佳楨(民 92)。資料探勘應用於就診行為與醫師排班之研究-以埔里基督教醫院為例(碩士論文)，國立暨南國際大學資訊管理研究所，南投縣。
- 陳宜欣(民 95)。以資料探勘技術探討顧客忠誠方案-以某信用卡發卡銀行為例(碩士論文)，國立中正大學行銷管理研究所，嘉義縣。
- 陳曉開 (民 88)，整理 McKinsey & Company, Inc. 董事 John Ott 於台灣第一屆「顧客關係管理研討會」之專題演講：成功地發展及執行持續性的關係行銷，電子化企業：經理人報告，3，26-30。
- 陳麒文(民 91)。健康休閒俱樂部顧客流失分析模式之研究(碩士論文)，私立輔仁大學體育研究所，新北市。
- 經濟部統計處(民 100)，批發、零售及餐飲業經營實況調查報告。
- 網路商業應用資源中心，取自 <http://www.ec.org.tw/>
- 劉世琪(民 92)。應用資料挖掘探討顧客價值-以汽車維修業為例(碩士論文)，私立朝陽科技大學工業工程與管理所，臺中市。
- 劉玉萍(民 89)，整理 NCR 公司助理副總裁 Ronald S. Swift 於遠擎智慧企業部舉辦 2000「一對一行銷」研討會中之專題演講：運用「一對一行銷」執行顧客關係管理已提升企業利潤，電子企業：經理人報告，11，7-10。
- 謝邦昌(民 90)。資料採礦入門及應用—從統計技術看資料採礦。臺北市：資商訊息顧問股份有限公司。

- Alex, S. (1999, October), "Winning CRM strategies". *ABA Banking Journal*, 54-66.
- Bowen J. T. and Shoemaker S. (1998), "Loyalty: A strategic commitment", *Cornell Hotel and Restaurant Administration Quarterly*, 39, 12-25.
- Berry, M. J. A., & Linoff, G. S. (1997) *Data Mining Techniques: for Marketing, Sales, and Customer Support*. New York: Wiley Computer.
- Curry, J., Wurtz, W., Thys, G., & Zijlstra, C. (1998), *Customer Marketing: How to Improve the Profitability of Your Customer Base.*, London: Kogan Page Limited.
- Frederick, E. (1996), Reichheld *Harvard Business Review*.
- Fredericks, J. O., Hurd, R. R. and Salter II, J. M. (2001), "Connecting customer loyalty to financial results ", *Marketing Management*, 26-32.
- Fayyad (1996), "Data Mining and Knowledge Discovery : Making Sense out of Data, " *IEEE Expert*, 11 (5), 20-25 .
- Hall, C. ed.(1995), "The Devil's in The Details:Techniques, Tools, and Application for Database Mining and Knowledge Discovery Part 1", *Intelligent Software*.
- Han, J., and Kamber, M. (2001), *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers.
- IBM (1998), *IBM Intelligent Miner For Data*.
- Jackie Kandell (2000, April), "CRM, ERM, one-to-one Decoding Relationship Management Theory and technology", *Trusts & Estates*, 49-53.
- Jones, T. O. and Sasser, W. E. (1995), "Why Satisfied Customer Defect", *Harvard Business Review*, 88-99.
- Klemettinen, M., Mannila, H., Ronkainen, P., Toivonen, H. and A. I. Verkamo (1994, November), "Finding Interesting Rules for Large Sets of Discovered Association Rules," *Proceedings of the Third International Conference on Information and Knowledge Management, Gaithersburg, Maryland*, 401-407.
- Megiddo, N. and R. Srikan (1998), "Discovering Predictive Association Rules," *IBM Almaden Research Center , KDD-98 Poster Sessions*.
- MIT Technology Review (2001), "The Technology Review Ten".
- M. S. Chen, J. Han, P. S. Yu (1997), "Data Mining: An Overview from Database Perspective", *IEEE Transaction on Knowledge and Data Engineering* , 866-883.
- Newell, F. (2000) *Loyalty.com: Customer Relationship Management in the New Era of Internet Marketing*. New York: McGraw-Hill.
- Oliver, R L, Rust R T, and Varki S. (1997), "Customer Delight: Foundations, Findings and Managerial Insight", *Journal of Retailing*, 73, 311-336.
- Peppers, D., & M. Rogers (1993), *The One to One Future: Building Business Relationships One Customer at A Time*. New York: Currency Doubleday.
- Peppers, D. and Rogers, M. (2001), *One to One B2B: Customer Development Strategies for the Business-to-Business World*, Cahners Business Information, Inc..

- Paul Ratcliff (2000, March), "Managing Deeper Relationship", Mortgage Banking, 94-100.
- Pieter, Z. Dolf (1996), Data Mining, Addison Wesley.
- R. Agrawal, R. Srikant (1994, September): "Fast Algorithms for Mining Association Rules", *Proc. of the 20th Int'l Conference on Very Large Databases*, Santiago, Chile.
- R. Agrawal, T. Imielinski (2004), A. Swami, "Mining Association Rules between Sets of Items in Large Databases", ACM SIGMOD, DC, USA.
- Singh, J. and Sirdeshmukh, D. (2000), "Agency and trust mechanisms in consumer satisfaction and loyalty judgments", *Journal of the Academy of Marketing Science*, 28, 150-167.
- Y, Fu(1997), "Data mining task, technique and applications," IEEE POTENTIALS.