

國立交通大學

管理科學系碩士班

碩士論文

顧客價值發現之研究一

整合資料探勘技術與顧客終身價值分析



A Study of Discovering Customer Value:

Integrating Data Mining Technique and Customer Lifetime

Value Analysis

研究生：陳麒文

指導教授：楊 千 博士

林君信 博士

中華民國九十四年六月

顧客價值發現之研究
—結合顧客終身價值分析與資料探勘技術

A Study of Discovering Potential Customer Value—
Integrating Customer Lifetime Value Analysis and Data Mining
Technique

研究生：陳麒文

Student：Chi-Wen Chen

指導教授：楊 千

Advisor：Dr. Chyan Yang

林君信

Dr. Chiunsin Lin



A Thesis

Submitted to Department of Management Science

College of Management

National Chiao Tung University

In partial Fulfillment of the Requirements

For the Degree of

Master

of

Business Administration

June 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年六月

國立交通大學

研究所碩士班

論文口試委員會審定書

本校 管理科學系碩士班 陳麒文 君

所提論文 顧客價值發現之研究—

整合資料探勘技術與顧客終身價值分析

合於碩士資格水準、業經本委員會評審認可。

口試委員

胡千

林君儀

何淑惠

張同忠

指導教授

胡千

林君儀

系主任

李經遠

教授

中華民國九十四年六月三十日

顧客價值發現之研究— 結合顧客終身價值分析與資料探勘技術

學生：陳麒文

指導教授：楊 千 教授
林君信 副教授

國立交通大學管理科學系碩士班

摘要

顧客關係管理已經成為企業在經營上的重要策略，為了能夠掌握具有價值的顧客，使企業的獲利能夠穩定且成長，顧客的潛在價值發現、維持與增強在現今競爭的商業環境中扮演著企業重要的獲利關鍵成功要素；然而，資料庫中的顧客特質、購買行為樣式等行銷觀點知識是未被觸及以及大量被隱藏的，因此如何進一步從資料庫中來發掘出顧客的潛在價值之特徵及潛在購買行為樣式是值得研究的。在探討過去的文獻後發現，顧客終身價值可以去計算出顧客對於企業貢獻程度，資料探勘可以發掘出顧客的潛在購買行為樣式，兩者間具有相輔相成之效。本研究結合顧客終身價值分析以及資料探勘的技術來建立出一套模式來找出每位顧客的潛在價值及特徵，透過顧客終身價值分析找出現有價值，以及每位顧客之預期購買金額以及活動機率，藉以計算出其潛在價值。針對現有價值以及潛在價值為基礎來做集群分析。資料探勘方面則以集群結果為基礎來做關聯分析、OLAP 以及決策樹，瞭解現有價值以及潛在價值最高的群組之特徵及購買行為樣式，並做為新進顧客的類別之預測與瞭解。最後，本研究點出各集群背後的行銷意涵，提供給管理者在應用顧客關係管理時的參考。

關鍵字：顧客價值、顧客終身價值、資料探勘、顧客關係管理

A Study of Discovering Customer Value: Integrating Customer Lifetime Value analysis and Data Mining Techniques

Student: Chi-Wen Chen

Advisors: Dr. Chyan Yang
Dr. Chiunsin Lin

Abstract

Customer relationship management (CRM) has become an important strategy for businesses. In an environment of intense competition, the finding, maintenance and strengthening of customer value is key to a business in making a profit. Past studies have found that customer lifetime value (CLV) can be used to calculate each customer's contribution to the company, and that data mining can be used to extract customers' potential behavioral patterns and characteristics. Though both are complementary, rarely are there studies applying the two methods at the same time. This research has developed a conceptual framework and process that integrates CLV analysis and data mining technique. It uses CLV analysis to calculate customer current value (CCV) and estimate customer potential value (CPV). Next, based on CCV and CPV, clustering techniques were applied for customer groupings. Finally, a data mining method will be applied to discovering the characteristics and the potential purchasing behavioral pattern of the most valuable customer group. By establishing a customer value pyramid combing the marketing implications of each cluster and suggestions, this research served as a reference for managers engaging in CRM.

Key Words: Customer value, Customer Lifetime value, Data mining, Customer relationship management

誌 謝

求學過程中一路走來，經過許多波折，涉獵許多領域。依稀還記得國中時代的調皮與叛逆，只知玩樂，不知求學重要性，無知的選擇到專科就讀，1年的電子工程及4年的機械工程生涯，不斷的在雜亂中摸索思緒，最後插大至農產運銷系就讀，並專攻於行銷領域，終於在大學中得到不少學業成績上的肯定及成就，同時也獲得許多老師的提拔及讚賞，包括：周世玉老師、吳明敏老師、李宗儒老師等，感謝他們無私貢獻，提昇我研究分析與事物判斷的能力及興趣。多樣的領域(工、農、商)也成就我對於不同事物的思維及觀察之興趣。

研究所生涯中，獲得許多知識與寶貴經驗，包括競賽的活動(Tic100、AT&A、夢公園行銷競賽等)、研究的能力與態度、簡報的技巧、做報告的方式、團體討論的觀念、資訊搜尋的技術、生活上的學習以及人生大道理，這都要感謝曾一起學習及打拼的同學們，還有研究所的老師之教學與關懷，使我倍感溫馨。每一段回憶都讓我印象深刻，因為與你們互動與指教，讓整個碩士生涯多采多姿，也在我的旅程中增添色彩。

這篇碩士論文的完成，是從過去的10多年的求學生涯中的學習、成長以及經驗所累積而成的，雖然論文的撰寫一開始所困於黑暗中，無法找到出口，但指導教授楊千老師以及林君信老師給予一盞明燈，引領我至光明大道，點出文章的方向與創意，口試委員張國忠老師、何淑熏老師以及好友彥寧也適時的提供指點及協助，最後提昇論文的完整性及價值貢獻，在此一併感謝。

最後，還要感謝家人、為詩、好友們在生活中的支持，以及陪伴我寫論文的MVP(Marketing Value Power)研究室、伺服器以及小A(筆記型電腦)，因為各方的幫助，讓我感到萬分幸福及感動！

「業精於勤，荒於嬉；行成於思，毀於隨」-韓愈 進學解，是我在做任何事情時所秉持的態度，現在是如此，未來亦是如此...

陳麒文 謹誌
國立交通大學管理科學系碩士班
於風城
2005年6月

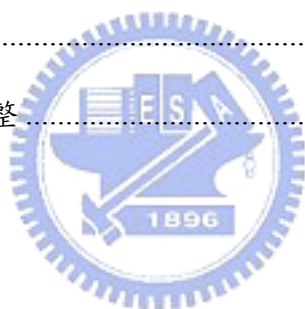
目錄

摘要	I
ABRSTRACT	II
目錄	IV
表目錄	VII
圖目錄	IX
1 第一章 緒論	1
1.1 研究背景	1
1.2 研究動機	2
1.3 研究目的	4
1.4 研究流程與內容	6
2 第二章 文獻探討	8
2.1 顧客關係管理	8
2.1.1 顧客關係管理的重要性	8
2.1.2 顧客關係管理定義	10
2.1.3 顧客關係管理範圍、建構與評估流程	11
2.1.4 顧客關係管理的好處	13
2.1.5 顧客關係管理的成功關鍵	13
2.2 顧客潛在價值	14
2.2.1 顧客潛在價值重要性	14
2.2.2 顧客潛在價值預測	15
2.2.3 顧客潛在價值策略	15
2.3 顧客價值發現－顧客終身價值	17
2.3.1 顧客終身價值的定義	17
2.3.2 顧客終身價值模型簡介	22
2.3.3 顧客終身價值模型的考量因素	25
2.3.4 顧客終身價值模型考量因素的難處	27
2.3.5 顧客終身價值的重要性與好處	28
2.4 顧客價值發現－資料探勘	30
2.4.1 資料探勘定義	30

2.4.2	資料探勘流程.....	30
2.4.3	資料探勘分析方法.....	32
2.4.4	資料探勘的應用.....	35
2.4.5	資料探勘重要性與顧客價值發現之關係.....	36
2.5	顧客關係管理、顧客價值發現、顧客終身價值與資料探勘之關係.....	37
2.5.1	顧客關係管理與顧客價值發現的關係.....	37
2.5.2	顧客終身價值與顧客價值發現之關係.....	37
2.5.3	資料探勘與顧客價值發現之關係.....	38
2.5.4	資料探勘、顧客終身價值、顧客價值發現與顧客關係管理之關係.....	39
3	第三章 研究方法.....	42
3.1	本研究架構.....	42
3.2	本研究流程.....	43
3.3	研究方法.....	45
3.3.1	顧客終身價值分析之個別顧客現有價值.....	45
3.3.2	顧客終身價值分析之個別顧客潛在價值.....	45
3.3.3	資料探勘分析.....	52
3.4	本研究分析環境介紹.....	56
3.5	整體資料庫介紹.....	57
3.6	建立資料探勘前準備.....	58
3.6.1	設立資料探勘目標.....	58
3.6.2	目標資料選擇.....	58
3.6.3	資料探勘資料表建立.....	59
3.6.4	資料庫前處理.....	60
3.7	整理後資料庫描述.....	62
4	第四章 分析－顧客終身價值分析與資料探勘技術.....	65
4.1	顧客終身價值分析.....	65
4.1.1	顧客終身價值之顧客現有價值.....	65
4.1.2	顧客終身價值之顧客潛在價值.....	69
4.1.3	顧客現有價值與潛在價值因素之相關性分析.....	74
4.2	顧客潛在價值分群.....	75
4.2.1	顧客現有價值分群.....	75
4.2.2	顧客潛在價值集群分析.....	76

4.3 顧客現有價值與顧客潛力價值交叉分析.....	76
4.3.1 顧客現有價值與顧客潛在價值個數交叉分析.....	76
4.3.2 顧客現有價值與顧客潛在價值獲利能力交叉分析.....	78
4.3.3 顧客現有價值與顧客潛在價值交叉累積分析-20/80 法則驗證.....	79
4.4 預測結果驗證.....	81
4.4.1 混淆矩陣建立.....	81
4.4.2 預測結果分析.....	81
4.5 資料探勘.....	82
4.5.1 OLAP 分析.....	82
4.5.2 顧客潛在價值購買行為.....	86
4.5.3 顧客潛在價值之人口統計因素發現-決策樹.....	89
5 第五章 結論.....	92
5.1 研究結論與發現.....	92
5.2 顧客關係管理策略應用.....	95
5.3 本研究在學術上的貢獻.....	96
5.4 研究限制與未來研究方向.....	97
5.4.1 研究限制.....	97
5.4.2 未來研究方向.....	97
參考文獻.....	99
附件 1 各集群之 OLAP 分析表.....	106

表 4-11 現有價值集群個數	76
表 4-12 最後集群中心點	76
表 4-13 集群個數	76
表 4-14 顧客現有價值與顧客潛力交叉分析	77
表 4-15 顧客現有價值與顧客潛在價值獲利能力交叉分析	78
表 4-16 各集群顧客價值與顧客數累積	80
表 4-17 訓練集與驗證集比對分析	81
表 4-18 OLAP 分析結果表	84
表 4-19 整體交易資料庫之潛在購買行為	87
表 4-20 黃金顧客群潛在購買行為	88
表 4-21 黃金顧客之決策樹	90
表 5-1 黃金顧客價值發現彙整	94



圖目錄

圖 1-1 本研究流程	7
圖 2-1 顧客關係管理的範圍	11
圖 2-2 顧客關係管理評估模型	12
圖 2-3 顧客現有價值及潛力價值矩陣	16
圖 2-4 顧客金字塔	17
圖 2-5 顧客終身價值與顧客權益之關係	19
圖 2-6 管理顧客終身價值流程與策略執行後的效果	29
圖 2-7 資料探勘流程	31
圖 2-8 資料探勘的分類	33
圖 2-9 從流程導向來分類資料探勘方法	33
圖 3-1 本研究之概念性架構	43
圖 3-2 本研究流程	44
圖 3-3 本研究資料庫結構	60
圖 4-1 顧客價值矩陣	78
圖 4-2 顧客價值與潛力獲利能力交叉圖	79
圖 4-3 20/80 法則驗證圖	80
圖 4-4 OLAP 多構面交叉圖	83
圖 4-5 OLAP 立方體呈現	83
圖 4-6 黃金顧客群關聯語法	88
圖 4-7 顧客價值發現之決策樹	89
圖 5-1 顧客價值族群金字塔	95

第一章 緒論

1.1 研究背景

近年來，顧客關係管理(Customer Relationship Management, CRM)已經成為企業在經營上的重要策略之一，企業認知到顧客關係管理的重要性，因此紛紛將此觀念加以導入至營運模式中；然而顧客關係管理要成功，必須要能夠掌握真正具有價值的顧客，如此企業的獲利才能穩定成長，避免不必要的資源浪費，因此顧客的獲得，價值的掌握、維持與增強在現今競爭的商業環境中扮演著重要的關鍵成功要素。由於現今的顧客之需求難以掌握、顧客異質性高、市場上競爭激烈、整合溝通行銷活動多元化、同儕意見影響力大、行銷資訊複雜且雜亂等情況下，企業要能夠建立與顧客的關係，往往會相當困難，即使企業已朝向顧客為導向，並且把顧客視為資產來與企業流程整合來思考，但是顧客關係管理並非容易的事，因此，如何以顧客為導向來運作企業、與顧客之間建立良好溝通與互動的機制、創造顧客價值以及找出具有價值的顧客，進而達到顧客的維持與信任等相關議題，皆已成為企業在經營顧客關係管理上最主要的考量要素。

顧客關係管理包含許多的程序，包括：設定顧客關係管理目標、思維執行策略、制定衡量顧客的尺度、顧客的價值評估、分析顧客關係管理執行有效性等，而這些運作程序都是必須要加以去衡量及評估的，尤其是顧客的價值評估，因為瞭解「誰才是獲利的顧客」是相當重要的，確定誰才是真正有獲利價值的顧客才能夠精確發展出具有效果的顧客關係管理，而且 20/80 法則也驗證了前 20%的價值顧客可以貢獻 80%的企業獲利，所以掌握精確的顧客價值在顧客關係管理中確實有其必要性，這樣的觀念也因此突顯出顧客終身價值(Customer Lifetime Value, CLV)在顧客關係管理中的關鍵。

近幾年由於資訊科技發達，加上會員制度導入的觀念成熟，使得許多企業皆有個別專屬的資訊系統來運作日常的作業以及輔助決策，而這些資訊系統有部份專門記載顧客的資料，儲放在資料庫中，這些資料對於企業來說或許只是個日常運作所紀錄的歷史「資料」，但是若能夠仔細的去瞭解每筆資料間的關係，反而能夠找出有意義的「資訊」，這

樣的資訊也帶給顧客關係管理上的協助，因為裡頭儲存各種與顧客相關的資訊，包括：購買的商品、時間、地點資料，顧客人口統計資料等，尤其是當企業不斷的成長時，顧客的資料量會相當龐大且雜亂，要能夠適切發展行銷計畫來吸引每個顧客，就必須透過客制化，甚至一對一行銷的方式來滿足與維持顧客，此時，資料庫的分析評估與應用便可以解決這方面的問題，因此在近年來，資料探勘(Data Mining)、資料庫行銷(Database Marketing)、甚至是一對一行銷(One to One Marketing)等相關的議題皆在行銷領域中發燒。

綜合以上所述，在現今的新行銷紀元中所重視的就是如何能夠在紛亂且競爭激烈的商業市場中，找出顧客的價值，精確的顧客區隔，提昇行銷的精確度及效益，並且從企業的資料庫中找出具有潛在價值的顧客，進而提供完善服務、促銷活動、產品、通路配銷來滿足消費者的需求與心理，建立起有效的顧客關係管理，使企業能夠長期獲利。

1.2 研究動機



顧客關係管理所涉及到的層面相當廣泛，包括：顧客資料蒐集、顧客資料庫導入、顧客資料整理、顧客價值分析、顧客關係管理執行策略等。由於企業管理資訊系統(Management Information System, MIS)導入的觀念增加以及科技的技術進步，企業越來越有能力去蒐集以及累積大量的顧客資料於資料庫中，因此資料庫裡的顧客相關知識對於行銷的運用功能上是相當有幫助的，然而，許多在資料庫中有用的顧客特質、購買行為樣式等行銷觀點知識是大量的、隱藏的、以及未被觸及的，要瞭解資料中的潛在關係，用手動的方式來分析出有價值決策制定資訊是幾乎不可能的，因此資料探勘在1994年開始受到注目，是最近發展的方法以及技術(Trybula, 1997)，透過資料探勘的技術將可以從大量的資料去加以整理，並從中發掘出有用的潛在資訊，進一步來瞭解消費者的潛在行為樣式，Davis(1999)根據專業以及商業文獻指出，許多企業使用資料探勘技術來當作協助他們運用計謀來打敗競爭者的工具，協助定義新顧客的方法和更改低成本的策略基礎，因此在近幾年，資料探勘已被廣為注意以及在商業界普及應用(Chye & Grey, 2002)。Goebel & Gruenwald(1999)指出，資料庫的知識發掘是發展迅速的領域，這領域有很強的研究需求、實務上的急迫性以及社會和經濟的需要。

另一方面，為了能夠找出精確且有潛力的顧客，當企業專注在顧客資料庫的導入與蒐集時，去分析顧客終身價值(Customer Lifetime Value)亦是非常重要的，因為顧客終身價值是一種系統化的方式去瞭解和評估顧客與廠商的關係(Jain & Singh, 2002)，可以找出哪些顧客對於企業的貢獻度高與潛在的價值有多大；再者，由於顧客終身價值是去分析顧客停留在企業的期間，因此顧客終身價值模型是去評估顧客整個長期的價值(Hwang et al., 2004)，考量時間的變數；由於它的效益高，所以顧客終身價值分析越來越受到重視。近年來，顧客終身價值的研究也越來越多的，其主要原因有三：1、顧客終身價值為顧客關係管理流程的必要過程，因此企業越來越有興趣瞭解此領域；2、行銷科學機構把此議題列為主要研究的重點；3、顧客終身價值領域的實證研究是特別缺乏的(Reinartz & Kumar, 2000)。Blattberg et al.(2001)也指出，評估行銷價值相關的議題也持續受到注意。然而另人感到意外的是，雖然一直以來，顧客終身價值在一般管理以及行銷管理上都是很明顯的重要觀念，但是專注在以價值為基礎的觀念設計才在近年來出現(e.g. Blattberg et al., 2001; Doyle, 2000)。在實際的研究上，行銷管理在這領域也只有部份學者在探討，原因在於：1、缺少在探討以及計算決定個別顧客的方法，例如交叉銷售(cross-selling)、向上銷售(up-selling)以及口碑傳播(worth-of-mouth)；2、沒有存在一個完整的模型來連結各種相關部份的方法在一起，因此這方面的研究依舊欠缺。

在顧客終身價值分析的領域中，有個議題是相當重要的，但卻少被提及，亦即為顧客潛在價值。顧客潛在價值的衡量在近幾年的研究中陸續出現，顧客潛在價值的考量是具有未來導向的，預測性的，然而價值潛力的因素之衡量並不容易，目前的研究中，建議考量的因素包括：顧客忠誠(customer loyalty)(Hwang et al., 2004; Knox, 1998)、交叉銷售(cross-selling)及向上銷售(up-selling) (Bauer et al., 2003)、推薦行為(recommendation behavior)(Bauer et al., 2003)、人口統計變數與產品使用行為(Jain & Singh, 2002)以及未來的購買行為(Ryals, 2002)等。有些研究雖有考量到許多因素，但仍被建議應考量其它各種因素，包括：向上銷售、參考價值 (見 Berger & Nasr, 1998; Dwyer, 1997; Keane & Wang, 1995; Reinartz & Kumar, 2000; Wang & Spiegel, 1994)、顧客維持率(Mulhern, 1999)，然而，這些因素雖然重要，但都難以去衡量及分析，因此本研究的動機亦即想要試著將資料探勘以及顧客終身價值分析整合在一起來發掘出顧客潛在，藉以解決部份的潛在價值因素及有效的衡量顧客價值。

綜合以上所述，茲將本研究所欲探討的問題及動機包括資料轉換成資訊面、顧客潛在價值分析與評估面以及顧客關係管理策略面，茲彙整如下：

1、資料轉換成資訊面：

- (1)企業如何從龐大的資料庫中發現顧客潛在價值的資訊？應該要何種方法來找出這類的資訊？
- (2)在資料庫中，如何衡量哪一位顧客是具有獲利的？對企業是否有貢獻價值？如何得知顧客的潛在價值有多大？包括：顧客活動力、購買金額以及交叉銷售(cross-selling)的潛力。

2、顧客價值發現之分析與評估面：

- (1)資料探勘以及顧客終身價值在資料庫分析中扮演的角色為何？
- (2)資料探勘技術與顧客終身價值分析對於顧客價值的發現以及對於顧客關係管理有何效益上的提昇？
- (3)顧客價值應如何來呈現？



3、顧客關係管理策略面

- (1)如何將顧客的價值劃分成各種不同特質的群體？每個群體的價值結構如何？有何種特性？
- (2)價值高的前 20%顧客如何來管理，有何策略上的意義，價值為後 80%的顧客怎樣的策略意義於其中，應如何來管理？
- (2)如何根據顧客價值分析的結果發展出有效的顧客關係管理行銷方案？

1.3 研究目的

資料探勘是從大量的顧客資料庫找出潛在的顧客行為樣式，顧客終身價值是從資料庫中去分析出哪些顧客對於企業的貢獻程度高以及顧客潛在價值如何，兩者之間有相輔相成的功效，若將兩者的觀念結合來應用，應用於顧客價值發現應會有不錯的成效，如

此對於顧客關係管理的應用以及策略執行會更具效力，因此值得探討研究。然而這顧客終身價值以及資料探勘對於顧客關係管理的研究通常都是被各別探討的，顯少有研究將此兩概念整合，故本研究的主要目的在於發展有效的顧客關係管理運作模式，專注在發展顧客終身價值的運用，將資料探勘的技術以及流程，導入顧客終身價值的概念，透過顧客的資料庫，分析出具有價值的顧客，透過實証研究，以顧客終身價值為基礎，發展出有效的顧客關係管理分析模式及應用策略。

Hwang et al.(2004)發展出考量過去貢獻獲利以及期望潛在價值的模型來衡量顧客終身價值，利用決策樹來將人口統計變數以及顧客忠誠度整合至期望潛在價值當中，並且根據模型所計算出的價值為基礎來加以區隔分群，提出這些區隔的行銷策略，由於考量到顧客過去貢獻的獲利以及期望潛在價值，因此本研究亦即參考此模型的建構精神來發展本研究的顧客終身價值模型，找出現有價值；在潛在價值方面，本研究參考 Schmittlein et al.(1987)所提的顧客活動率(probability of active customer)以及 Schmittlein & Peterson(1994)所提的預期再購買金額來計算顧客潛力，並且將資料探勘流程及分析技術導入到整個顧客評估的模型上，建構出完整的顧客價值發現之分析模式。此外，將各顧客的現有價值、未來潛在價值以及人口統計變數等相關構面來建構出一個三維的立方體(cube)，發掘出最具有潛在價值的顧客群組特徵，最後，本研究再將找到的顧客價值群組，透過決策樹的分析來瞭解不同顧客群組的樹狀結構，以做為未來新進顧客的分類以及預測基準。

綜合以上所述，本研究主要的目的有以下 5 點：

- 1、探討資料探勘以及顧客終身價值在顧客關係管理上的好處以及幫助。
- 2、將資料探勘的流程以及顧客終身價值的分析技術整合，以發展出完善的資料庫行銷分析模式，找出具有潛在價值的顧客。
- 3、建構顧客價值發現流程模式，並透過實證來檢驗效果，以做為學界及業界參考。
- 4、提出以價值為基礎的顧客區隔模式，考量顧客終身價值以及各種潛在價值因素，並且針對各區隔提出各種減少行銷資源配置以及企業價值提高的行銷策略。
- 5、提供具有價值意義的各種維度，包括：顧客現有價值、潛在價值等來探討顧客價值，以利管理者制定行銷決策。

1.4 研究流程與內容

本研究的流程劃分為 5 個部份，第一部份為本研究的研究背景、動機以及目的，主要在瞭解本研究的原因、目前在此研究領域所欠缺之處、本研究所涵蓋的範圍、本研究問題及運作流程。

第二部份為文獻探討，針對顧客終身價值分析以及資料探勘技術的過去以及最新研究情形做深入探討，並從中瞭解學者在此領域的看法及意見，藉此做為本研究主要核心模型建構以及研究方法之基礎。

第三部份為研究方法，本研究根據之前學者的研究來做為本研究方法的主軸，所採用的研究方法包括：顧客終身價值、顧客潛在價值、決策樹、關聯法則等相關方法來加以探討，資料庫介紹、研究環境等亦會在此章節做說明。

第四部份為實証研究部份，利用本研究所採用的方法，根據實際的資料加以驗證計算。

第五部份為結論、建議以及後續研究方向；將本研究的結果予以呈現，並且提出本研究在管理上的意義及貢獻，最後則提出未來後續可以研究的方向。

整個研究程序如下圖 1-1 所示：

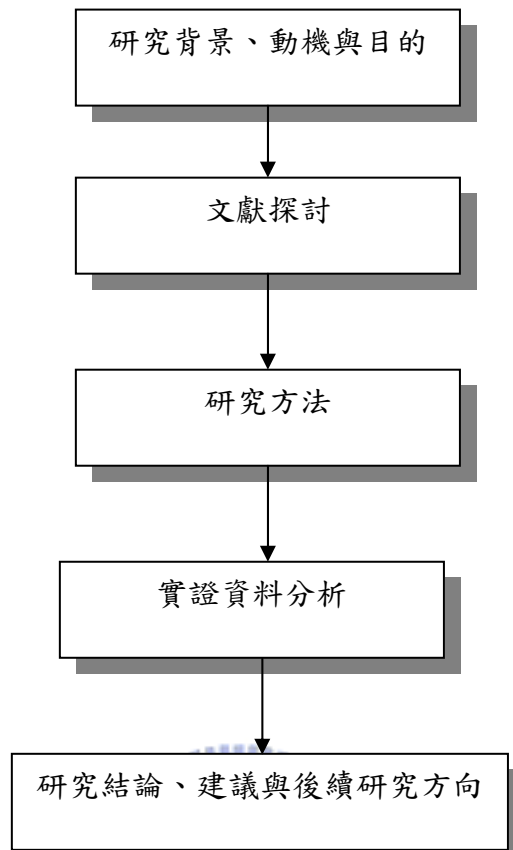


圖 1-1 本研究流程



第二章 文獻探討

要做好顧客關係管理，瞭解顧客的貢獻價值以及顧客的潛在行為是重要的關鍵，而這此兩關鍵值可以透過資料探勘技術以及顧客終身價值分析來獲得，因此本研究將資料探勘技術以及顧客終身價值分析的觀念整合在一起，建立出具有意義的顧客價值發現之評估模式，以提昇顧客關係管理的效率及效果，所以在文獻上，本研究針對顧客關係管理、顧客潛在價值、資料探勘以及顧客終身價值相關文獻做深入探討。

2.1 顧客關係管理

到了 90 年代，行銷研究領域已開始專注在與顧客關係上的建立(Ryals, 2002)，由於顧客關係管理的議題在近幾年持續被受到重視，因此企業對顧客的重心已從原本只是企業流程的個體轉變成企業成功的關鍵要素(Jain & Singh, 2002)。本段落將根據顧客關係管理的重要性、顧客關係管理定義、顧客關係管理範圍、建構及流程、顧客關係管理好處以及顧客關係管理成功因素加以探討。

2.1.1 顧客關係管理的重要性

在競爭環境激烈、消費者需求多樣化以及重視個人化的商業環境裡，顧客關係管理已變的非常重要，企業的獲利來源已不再是從產品為中心的角度來思維，顧客已成為企業運作的核心所在。Jain & Singh(2002)指出，以顧客為中心方法的重要性已逐漸增加，這可由現今盛行的許多顧客關係管理議題來加以驗證，例如：一對一行銷以及資料庫行銷。Kim et al.(2003)認為，在新的世紀，顧客關係管理已成為領導企業的其中一個策略，也因此顧客關係管理已逐漸在公司高層級的議程中被討論。Payne(2002)點出，公司已逐漸將顧客關係管理視為公司策略的主要元素，理由有二：1、新科技使得公司可以精確鎖定所選定的市場區隔、微小的區隔或者是個別顧客；2、新的行銷思維認知到傳統行銷的限制以及瞭解到顧客為焦點的重要性。

過去的企業是以產品中心為導向來運作，產品被視為資產，廠商所專注追求的是銷售更多的產品，如此可以獲得高毛利，也就是說，企業認為增加利潤來自於每個產品銷售(Jain & Singh, 2002)。Ryals(2002)指出，過去行銷策略傾向於產品方向來思維，強調以產品、價格、促銷以及通路等 4Ps 來做為策略發展的基石，在這樣的環境當中，零售商會相信獲利來自於產品，而非來自於顧客，如此使得營運焦點變成以銷售更多產品以及增加市場佔有率為目標，新的顧客與長期忠誠的顧客被視為是相同的價值，但這樣的觀念在現今的商業環境當中已逐漸的被修正，並開始轉向以顧客為觀點來思考。

顧客關係管理的重視也就代表著以顧客為中心的導向思維被受注目，Kim et al.(2003)指出，顧客關係管理的本質是從產品為導向轉換為以顧客為導向的哲學。Jain & Singh(2002)從績效評估的角度來看，行銷活動以及績效評估已經逐漸的重視在顧客的關係上，而非產品上。Jain & Singh(2002)亦指出，隨著逐漸瞭解顧客忠誠的重要性，企業已將以顧客為中心的方法導入到策略規劃中，如此顧客生命週期比過去的產品生命週期在行銷策略中更成為中心的角色，而顧客為中心的思維是鎖定在企業與每個階段的顧客生命週期，增加顧客生命週期以及交易的價值，並且視顧客為資產以及專注在顧客的獲得與維持，這樣的維持可以成為持續性競爭優勢的基礎。因為一般來說都相信，長期終身的顧客對於企業來說是更有利潤的，因此越來越多的企業專注在與企業維持長期關係的顧客上以及之後所創造的獲利以及成長。Gupta & Lehmann(2003)比較線上企業在以顧客為中心以及以產品為中心在策略規劃上的差異，結果如下表所示，由下表 2-1 得知，以顧客為中心與產品中心在各階段的所專注的方向是有差異的，以產品為中心運作的企業為傳統行銷的流程概念，而以顧客為中心的策略規劃則較強調以顧客價值為基礎的發展程序。

表 2-1 線上企業在以顧客為中心以及以產品為中心在策略規劃上的比較

以產品為導向的線上企業在策略規劃的程序		
階段 1: 行銷研究→	階段 2: 規劃→	階段 3: 執行→
<ul style="list-style-type: none"> ● 消費者研究 ● 評估競爭者 ● 評估供應商提供的資源與能力 	<ul style="list-style-type: none"> ● 建立目標 ● 選擇行銷組合 ● 細部運作規劃 ● 任命執行的責任 	<ul style="list-style-type: none"> ● 行銷組合執行 ● 檢查消費者需求 ● 供應產品 ● 後續調查
以顧客為中心的線上企業在策略規劃的程序		
階段 1: 引進 CRM→	階段 2: 規劃→	階段 3: 執行→
<ul style="list-style-type: none"> ● 蒐集歷史資料 ● 顧客終身價值計算 ● 顧客區隔 ● 顧客基本資料 	<ul style="list-style-type: none"> ● 創造跨功能團隊 ● 同意價值創造的真實目標 ● 設計以 CLV 為基礎的目標策略 ● 細部的導入規劃 ● 任命執行每日顧客管理的責任 	<ul style="list-style-type: none"> ● 導入顧客回應的程序 ● 衡量結果 ● 蒐集資料 ● 根據獲得結果調整運作規劃

資料來源：Gupta & Lehmann, (2003). "Customers as assets". *Journal of Interactive Marketing*, 17(1), 9-24.

2.1.2 顧客關係管理定義



Conlon(1999)認為，顧客關係管理代表的是企業體整合銷售、行銷、顧客服務、線上活動以及通路伙伴的一種系統及策略。John(1999)認為，顧客關係管理是持續性關係行銷，其強調的重點是去尋找對企業最有價值的顧客，以微型區隔(micro-segmentation)的概念，界定出不同價值的顧客群。企業以不同的產品、不同的通路，滿足不同區隔顧客的個別需求，並在關鍵時刻，持續地與不同層次顧客溝通，強化顧客的價值貢獻。Kim et al.(2003)指出，顧客關係管理可以定義為企業透過企業流程和科技的結合來瞭解公司的顧客，進而加強與顧客間的互動管理。Kumar et al.(2004)指出，顧客關係管理可以被定義為持續診斷顧客在所有有效的接觸點的流程，透過差異化以及量身定做的服務，建立顧客期望服務的行銷活動，如此讓顧客的貢獻達到獲利最大化。Jenkins(1999)定義顧客關係管理為預測顧客行為的流程以及分析顧客影響企業獲利的關鍵因素。Kalakota & Robinson(1999)則認為顧客關係管理是以顧客生命週期、企業與顧客間的關係進展為主軸，針對與顧客之間的互動關係，建構理想的顧客關係管理架構，透過顧客獲取(acquisition)、顧客增進(enhancement)、顧客維持(retention)三個主要的思維加以實現，而這三大核心觀念則需要透過各種不同管理功能，如運作整合性銷售、行銷與服務的策

略，在發展產品與服務上力求改進，以致力於顧客滿意度與顧客忠誠度的提升。

2.1.3 顧客關係管理範圍、建構與評估流程

Hwang et al.(2004)指出，顧客關係管理所涵蓋的範圍從一開始的吸引顧客到顧客維持、培養、背叛、忠誠，最後成為有潛力的顧客等皆為顧客關係管理的範圍，彼此間存在某種關係，如下圖 2-1 所示：

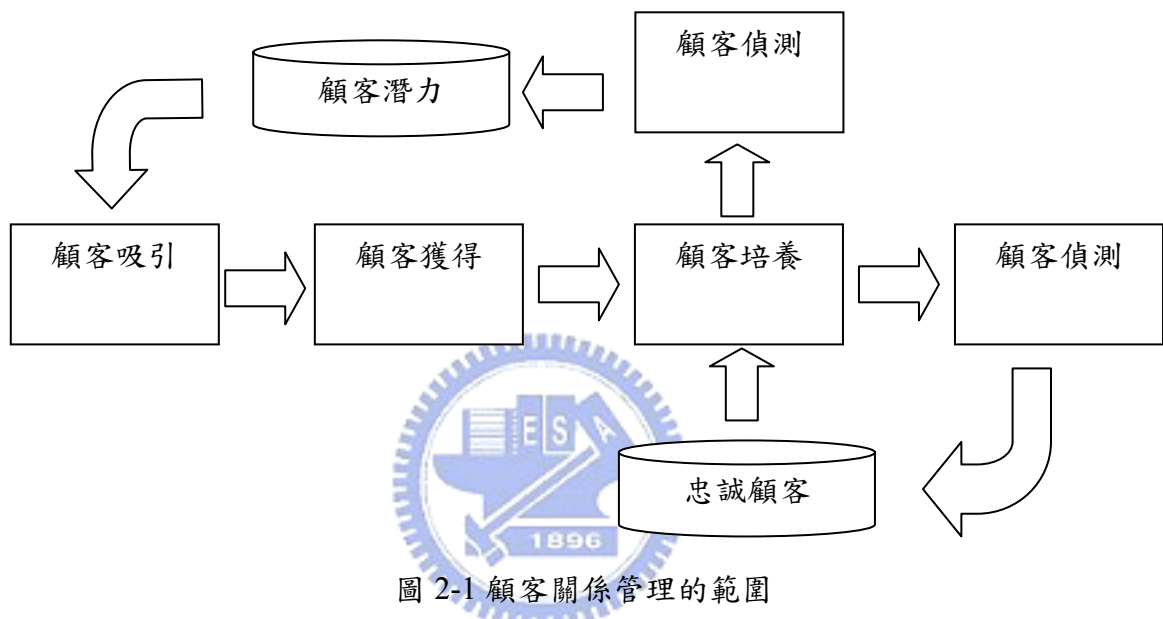


圖 2-1 顧客關係管理的範圍

資料來源: Hwang, H., Jung, T., & Suh, E., "An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry", *Expert Systems with Applications*, 26, 181–188, 2004.

Kim et al.(2003)指出，顧客關係管理的建構流程及範圍如下圖 2-2 所示，首先要先決定顧客關係管理的目標，制定任務及目標，接下來制定顧客關係管理策略，定義出主要的策略因素(principal strategic factors)，然後再分析因果關係，制定觀點及衡量尺度，最後根據結果來瞭解顧客關係管理的執行有效性。其中在制定觀點及衡量上涵蓋四個重要概念，分別為顧客知識、顧客互動、顧客價值以及顧客滿意，其涵意如下：

- 1、顧客知識：顧客區隔的狀態以及顧客資料管理。顧客知識的焦點在於科技的學習，瞭解顧客的需要以及影響顧客互動的因素。
- 2、顧客互動：經營上的優越、顧客服務的通路管理以及顧客流程。
- 3、顧客價值：從顧客上所獲得的利益，例如顧客終身價值以及顧客忠誠。

4、顧客滿意：產品或服務的滿意程度。

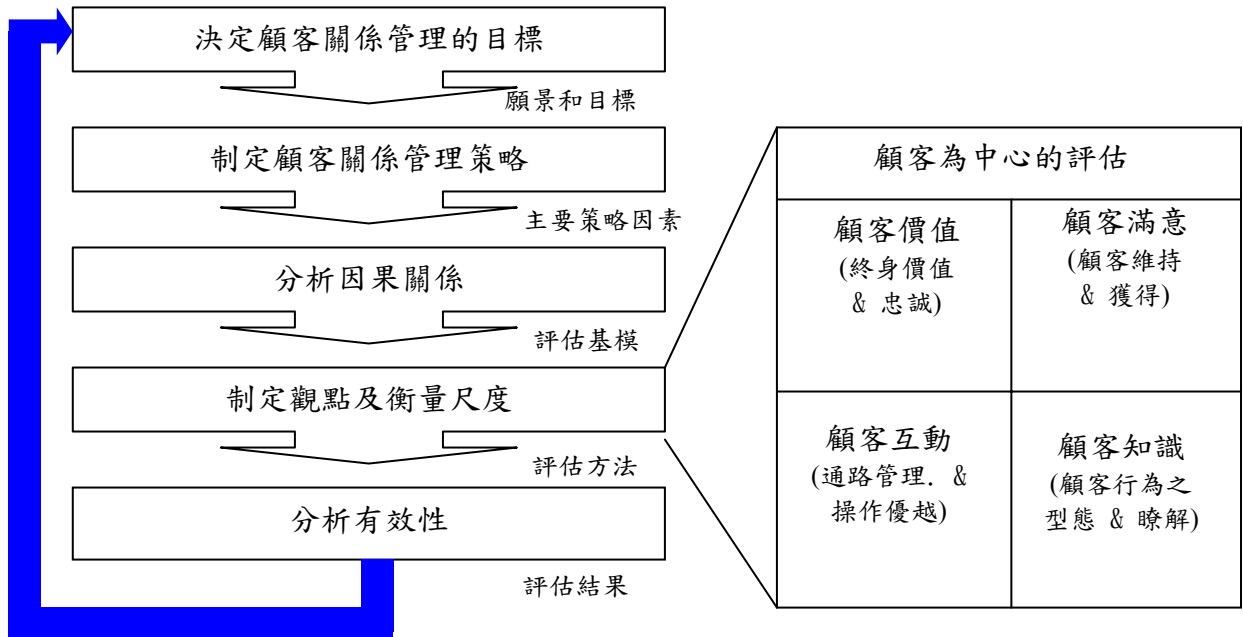


圖 2-2 顧客關係管理評估模型

資料來源：Kim, J., Suh, E. & Hwang, H., "A Model for Evaluating The Effectiveness of CRM Using The Balanced Scorecard," *Journal of Interactive Marketing*, Volume 17, Number2, Spring, pp.5-19. 2003.

Conway & Fitzpatrick(1999)以顧客關係管理流程的角度指出，有效的執行顧客關係管理的策略必須要導入以顧客為中心的組織文化，以這個架構來思維可以知道顧客關係管理範圍包括：

- 1、確認在目標顧客區隔中的每一個顧客之獨特特質。
- 2、模型化每個顧客的現在以及潛在的價值。
- 3、針對高價值的顧客來創造出能夠支援顧客經驗的活動策略、營運計畫或是經營準則。
- 4、重新設計組織、程序、科技以及獎賞系統到關係的策略上。

Gurau et al.(2003)指出，導入顧客為中心的系統必須要包涵幾個重要的步驟：

- 1、蒐集顧客的資訊。
- 2、計算顧客的終身價值。
- 3、根據顧客的價值區隔，並且建立區隔之重要優先排序。

由以上針對顧客關係管理之範圍、建構與評估流程得知，顧客終身價值在運作顧客關係管理時是相當重要的(Kim et al.2003; Conway & Fitzpatrick, 1999; Gurau et al.,

2003)，因為確定哪位顧客是具有潛在價值之後，管理起來會具有效益，對於企業未來獲利及成長也會較有幫助。

2.1.4 顧客關係管理的好處

顧客關係管理的好處其影響力不僅是單一顧客所創造的價值，而是整個企業的價值，有些顧客關係管理的利益則會提昇管理程序的效率，並且從顧客端來學習以及由顧客來測試創新，進而去影響改善經營流程(Ryals, 2002)。Kalwani & Narayandas(1995)指出，長期顧客關係的成功，部份來自於獲利增加，而大部份是來自於關係建立之後所創造的利益。Kim et al.(2003)點出導入顧客關係管理後的潛在利益(Jutla et al., 2001; Stone et al., 1996)，分別為：

- 1、增加顧客的維持與忠誠
- 2、提高顧客獲利率
- 3、創造顧客價值
- 4、產品和服務的客制化
- 5、較低的運作程序、高品質產品和服務



2.1.5 顧客關係管理的成功關鍵

顧客關係管理的成功關鍵部份在於精確評估顧客的價值以及鎖定顧客區隔(Hwang et al., 2004)。Christopher et al.(1991)指出，顧客關係管理不只是在於強調獲得顧客，而且是保留住顧客，因此以顧客維持以及增加獲利為主體關係的研究持續成長當中(e.g. Blattberg & Deighton, 1996; Reichheld, 1996, 1993; Reichheld & Sasser, 1990)。顧客的價值對於企業的成功有正面幫助，使得企業開始朝向與顧客維持長期的關係(Jain & Singh, 2002)，也由於顧客關係管理的重要性，這樣的觀念也促使企業去認知把顧客視為牠們資產的一部份(Kutner & Cripps, 1997)。Wyner(1996)指出，顧客被視為資產，成為經濟的單位，使得傳統的行銷觀念重新被設定。由於顧客是企業重要的無形資產，企業必須要評價並且管理它(Mulhern, 1999)，為了能夠有效的管理資產，因此企業必須要去瞭解哪

一個顧客是有價值的，哪些是價值較低的，如此才可以適當的應用行銷策略(Ryals, 2002)，發揮價值創造的最大效益，由此可知，瞭解顧客的價值以及得知有獲利的顧客是維持顧客的要點。因此，許多顧客關係管理研究人員持續原發展有系統的顧客獲利模型，因為要做好顧客關係管理，誰是有獲利的顧客是很重要的成功關鍵(Hwang et al., 2004)。Gurau et al.(2003)指出，最終，組織在顧客關係管理成功包括對於未來創造出適當的策略性願景，制定適當應用的選擇，有創造力的使用適當的分析技術來挖掘資料以及把顧客為中心程序導入至有彈性的策略規劃程序當中。一旦顧客區隔被定義出來以及根據獲利加以優先排序，如此便可以將以顧客為中心的方法導入之策略規劃程序當中。

由以上顧客關係管理的相關文獻的彙整可以得知，從顧客價值的角度來看顧客關係管理之成功關鍵在於：

- 1、建立適當的策略性願景(Gurau et al., 2003)
- 2、精確的分析顧客終身價值(Hwang et al., 2004; Kim et al., 2003)
- 3、找出顧客區隔，並且加以鎖定(Hwang et al., 2004; Ryals2002)
- 4、將顧客視為企業的資產(Hunt, 1997; Kutner; Cripps, 1997; Wyner, 1996)
- 5、不只強調獲得顧客，更要維持顧客(Hawkes, 2000)
- 6、獲得顧客知識，瞭解顧客潛力(Kim et al., 2003)
- 7、讓顧客感到滿意(Kim et al., 2003)
- 8、與顧客互動良好(Kim et al., 2003)

2.2 顧客潛在價值

顧客潛在價值發現在顧客關係管理的運作上是個重要關鍵，本段落針對顧客潛在價值的重要性、預測以及策略做文獻探討。

2.2.1 顧客潛在價值重要性

要有效的做好顧客關係管理，找到顧客潛在價值是有實質上的必要(Verhoef &

Donkers, 2001)，因為它代表了未來的公司獲利可能性，因此不只是顧客的目前價值資訊是重要的，顧客潛在價值亦是關鍵(Bult & Wansbeek, 1995; Hesket et al., 1997; Hoekstra & Huizingh, 1999)。Verhoef & Donkers(2001)亦指出，經營者不只對於顧客目前價值有興趣，對跨銷售機會的資訊亦高度重視，而且，若將顧客的潛在價值以及目前價值的資訊合併將可以提供顧客關係管理的經理人有機會延伸顧客的區隔。此外，顧客潛在價值是與個別顧客購買產品的數量或是服務有相關(Hoekstra & Huizingh, 1999)，所以它主要是分析個別顧客的購買資料。Blattberg & Deighton(1996)指出，從管理的觀點來看，顧客的潛在獲利是非常有趣的，因為顧客的關係行銷努力所須的特定最適預算是可以從中去分析出來的。Verhoef & Donkers(2001)則指出，顧客潛在價值亦即為藉由顧客行為的理想化來找出獲利以及價值傳送。

2.2.2 顧客潛在價值預測

瞭解顧客潛在價值須要先知道顧客的需求，而個別需求的差異往往會因為一些人口統計變數、態度以及社會階層不同等因素而影響(Engel et al., 1995)。Verhoef & Donkers(2001)指出，在分析顧客潛在價值時，去預測不同產品或服務的購買資料也是重要的預測因素。Spring et al. (2000) 從實務的角度來看，企業的資料庫往往只有顧客資料以及社會人口統計資料，因此從現有的資料中發掘顧客潛在價值對實務來說會有很大的貢獻。目前顧客潛在價值的預測分析方法包括迴歸分析、類神經網路、決策樹、顧客終身價值(Hwang et al., 2004)、Probit Model(Verhoef & Donkers, 2001)、SMC模型(Schmittlein et al., 1987)等。

2.2.3 顧客潛在價值策略

顧客潛在價值的策略方面，Verhoef & Donkers(2001)提出顧客現有及潛力價值矩陣，如下圖2-3所示：

		顧客現有價值	
		低	高
顧客潛在價值	高		
	低		

圖 2-3 顧客現有價值及潛力價值矩陣

資料來源：Verhoef, P.C., & Donkers, B. “Predicting customer potential value an application in the insurance industry,” *Decision Support Systems*, 32 2001 189–199., 2001.

其中，區隔：此區隔並不吸引企業的投資，因為它是低潛在價值以及低現有價值，其策略為減少成本以及增加價格，並且不要試著想去增加此區隔的購買等級。區隔：有高的潛在價值，但是公司並沒有成功的從中獲利，此區隔有許多的向上銷售活動的機會，因此公司應專注此區隔來拓展。區隔：有低的潛在價值，但有高的現有價值，此顧客的向上銷售機會較低，公司的策略應該是要儘可能的維持此區隔的顧客。區隔：此區隔最具有價值，這類顧客為高忠誠而且有大的潛在價值，公司應用各種努力來與此顧客維持好關係，否則損失會很大。

Jay(1998)則提出顧客金字塔的概念，將顧客分為五大類，分別為積極型顧客、非積極型顧客、潛在顧客¹、懷疑型顧客以及其它類，特性分述如下：

1、積極型顧客：在顧客過原特定期間內，曾向你買過產品或服務的個人或企業，而這類的顧客還可細分為四小類，分述如下：

- (1) 主顧客：在積極型顧客中，購買額最多的前百分之一者。
- (2) 大顧客：在積極型顧客中，除了主顧客外，購買額最多的前百分之四者。
- (3) 普通顧客：在積極型顧客中，除了主顧客及大顧客外，購買額段多的前百分之十五者。
- (4) 小顧客：除了上述三者外的積極型顧客。

2、非積極型顧客：過去曾向你買過產品的個人或企業，但在過去特定期間已一陣子沒再購買過。

3、潛在顧客：未曾購買過產品或服務，但卻跟你多少有些關係的個人或企業。

¹此處的潛在顧客類型並非本研究的潛在價值顧客，潛在型顧客是指未購買，但未來很有可能成為顧客的消費者，而本研究的潛在價值顧客是指已購買過，而且未來繼續購買的潛力大，貢獻價值高，因此稱為潛在價值顧客。

4、懷疑型顧客：企業可以提供產品或服務給他們，但卻還未跟這些人建立關係。

5、其他類：不需要購買或使用你產品或服務的個人或企業就屬於其他類。

顧客金字塔繪出來結果如下圖2-4所示，最上面的四層即為積極型顧客群，接下來為非積極顧客群、潛在顧客群、懷疑型顧客群以及其它顧客群。

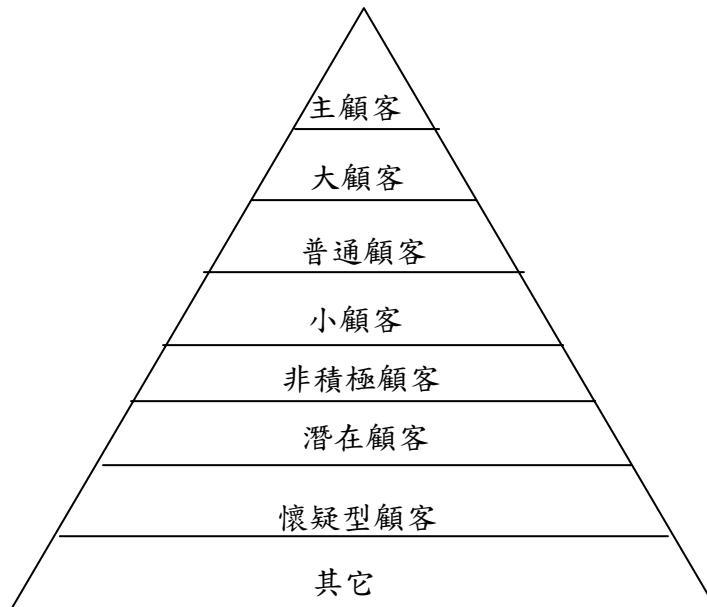


圖 2-4 顧客金字塔

資料來源：Jay C., Wil W., Guido T., & Conny Z., 顧客金字塔的行銷秘密，陳琍玲 譯，美商麥格羅 希爾國際股份有限公司，台北，民國 87 年。

2.3 顧客價值發現—顧客終身價值

顧客價值發現可以透過顧客終身價值分析的方法來找出來，本段將探討顧客終身價值之相關文獻，包括顧客終身價值之定義、顧客終身價值之模型、顧客終身價值之好處與重要性、以及顧客終身價值所須考量的因素與難處。

2.3.1 顧客終身價值的定義

不同的學者對於顧客終身價值會有不同的定義，主要的差異在於研究的方向以及所持的立場不同，因此會有不同的定義出現，但其實學者所提的定義當中，大部份的觀念是相同的-亦即為現值的概念。為了瞭解顧客終身價值一詞的使用，Mulhern(1999)曾加

以討論。Mulhern(1999)根據文獻指出，此研究領域的相關名詞有包括：終身價值(Lifetime value)(Keane & Wang, 1995)，顧客終身價值(customer lifetime value)(Berger & Nasr, 1998)，顧客評價(customer valuation)(Wyner, 1996)，顧客終身評價(customer lifetime valuation)(Dwyer, 1989)，顧客關係價值(customer relationship value)(Wayland & Cole, 1997)以及顧客權益(customer equity)(Blattberg & Deighton, 1996)；在 Mulhern(1999)的研究中使用與 Storbacka(1998)一樣的命名，採用顧客獲利(customer profitability)為名詞，因為獲利分析在許多非直接行銷上的應用並不總是涉及顧客在一段延伸時間的連續購買，而且價值這個術語在行銷文獻上有另外的意義，例如知覺價值，因此常會造成混淆與誤解。

然而，本研究在根據文獻對於顧客權益以及顧客獲利的研究之後發現，認為顧客權益、顧客獲利與顧客終身價值之間在應用上有部份的思維上之差異，此三者之間是呈現相互因果關係。在顧客獲利方面，顧客獲利是個別顧客對於組織的淨貢獻值(Mulhern, 1999)。而Pfeifer et al.(2004)則指出，顧客獲利亦即為在特定期間之內，收入與顧客關係相關成本之間的差異。顧客終身價值方面，Gupta & Lehmann(2003)指出，顧客終身價值為從顧客上所產生的所有未來獲利之現值，因此顧客終身價值有折現之涵意。顧客權益方面，Bayon et al.(2002)指出，顧客權益是包含個別顧客以及投資人，他並根據Blattberg & Deighton(1996)的定義來指出顧客權益亦即為在顧客忠誠於企業的期間之內，所有現在及未來顧客的折現之總和，例如：個別顧客終身價值之總和。由各學者對於顧客獲利、顧客終身價值以及顧客權益的術語之定義可以得知，顧客獲利與會計上的獲利一詞意思相近，因此是把獲利應用在個別顧客的概念上，即個別顧客對於企業的銷貨收入減去銷貨成本，而顧客終身價值中的價值與財務上的價值一詞概念是相似的，亦即把資產的評價由淨現值來表示(Pfeifer et al., 2004)；而顧客權益則是把個別顧客之終身價值以予加總，並與企業價值以及股東權益之觀念銜接起來。Bayon et al. (2002)則點出顧客權益、顧客終身價值以及顧客獲利之間的關係，如下圖2-5所示：

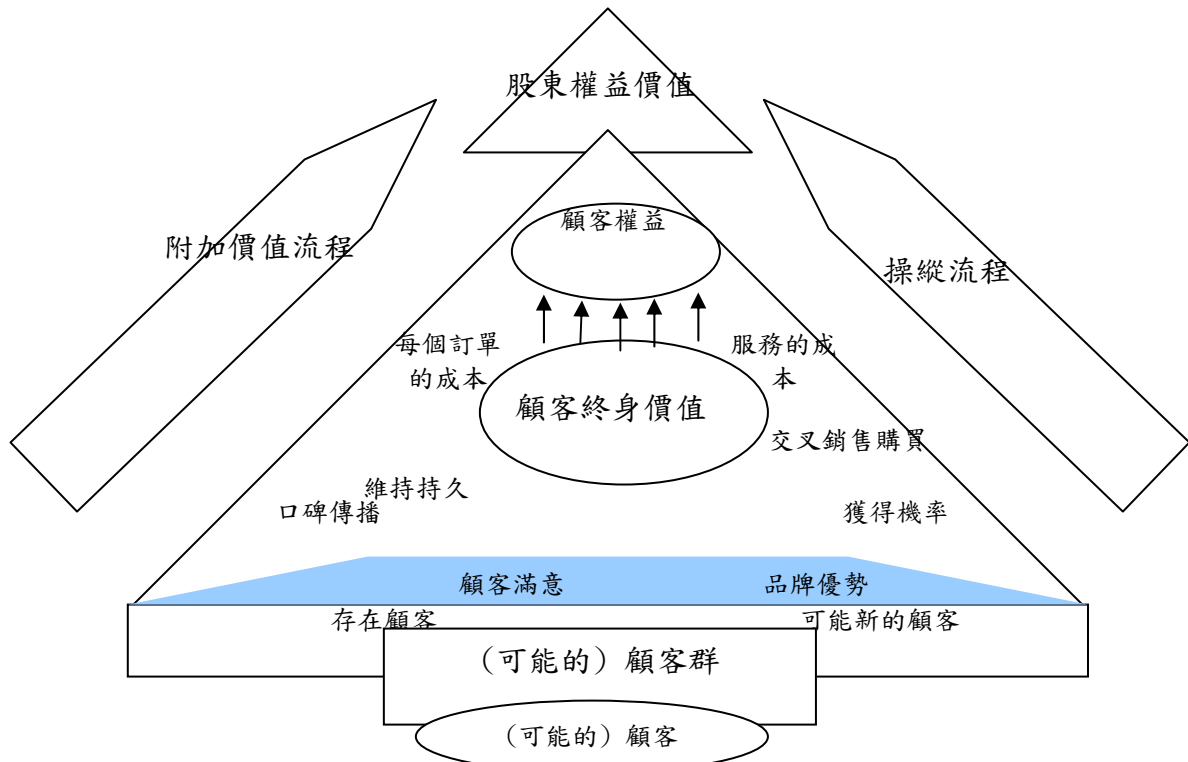


圖 2-5 顧客終身價值與顧客權益之關係

資料來源: Bayon et al., "Customer Equity Marketing: Touching the Intangible," *European Management Journal*, 20 (3), pp. 213-222, 2002.

由上圖 2-5 可知，顧客權益最基石的部份為可能的顧客以及可能的顧客集群，接下來為已存在的顧客以及可能的新顧客，這些顧客可能因為顧客滿意或者是品牌認同而去購買，如此產生歷史交易量，形成折現後顧客終身價值。顧客終身價值因為口碑傳播、獲得新顧客可能、向上(up-selling)或交叉銷售(cross-selling)購買、顧客維持、低訂單成本、低服務成本而創造出顧客的價值，其價值能夠被視為資產，並且對於股東產生價值時，亦即為顧客權益。

經過文獻探討得知，顧客終身價值中的價值概念與一般在行銷中所提的顧客價值、知覺價值等概念是相似的，因為顧客終身價值的概念包括顧客的滿意、獲得與維持，當顧客覺得知覺價值高時，其顧客滿意亦會提高，購買可能性因此上昇，故顧客終身價值也會提高，所以彼此間是有關聯的，而近年來對於顧客終身價值的概念已不再僅侷限在過去的貢獻利潤，還包括潛在價值，再者，近年來許多研究者都以採用顧客終身價值一詞來做為行銷領域研究的名詞，因此，引用「顧客終身價值」一詞來做為此本研究的名詞。

顧客終身價值較為簡單明瞭的定義為從顧客上所產生的所有未來獲利之現值(Gupta

& Lehmann, 2003)。Roberts & Berger(1989)指出，顧客終身價值為所有未來貢獻到經常帳及利潤的淨現值；Courtheoux(1995)稍微修飾 Roberts & Berger(1989)的定義，把從顧客的預期概念導入到定義中，指出顧客終身價值為預期顧客未來貢獻到經常帳及利潤的淨現值。部份的學者則強調顧客與企業間的關係，Kumar et al.(2004)指出，顧客終身價值為顧客與企業間的整個終身之累積現金流量之總額—使用 WACC 折現，而 Bitran & Mondschein(1996)指出，顧客終身價值為顧客在企業當中所產生的總淨獲利之折現值；Jackson(1994)則導入交易以及契約的概念，指出顧客終身價值為所有與企業顧客交易以及契約所獲得的獲利貢獻之淨現值。Knie-Andersen(2001)則認為，顧客終身價值是未來貨幣總額的折現，與未來總額的定義不同；Ryals(2002)點出，顧客終身價值是顧客未來購買的現值，顧客終身價值的計算是基於個別顧客的期望購買，利用折現率調整成現值。到了近期，由於顧客關係管理的逐漸受到重視，因此 Pfeifer et al.(2004)則導入顧客關係的概念，指出顧客終身價值是與顧客關係上的顧客現金流量之折現值。

然而，顧客終身價值較為嚴謹的定義則考量了許多衡量顧客終身價值時的因素到定義當中，Hwang et al.(2004)指出，顧客終身價值為在扣除吸引、銷售、服務顧客的成本之後，企業從顧客的整個交易之終身時間中獲得的收入總額所計算出的貨幣時間價值；Berger & Nasr(1998)指出，企業從顧客交易的整個生命週期之淨獲利或損失，因此顧客終身價值是企業從顧客的整個交易生命週期淨收入去減掉吸引、銷售以及服務顧客的成本所考量的貨幣價值。本研究將學者所定義的顧客終身價值根據年代予以排列出來，彙整如下表 2-2 所示：

表 2-2 顧客終身價值定義

學者名稱	顧客終身價值定義
Pfeifer et al.(2004)	顧客終價值是與顧客關係上的顧客現金流量之折現值。
Hwang et al. (2004)	顧客終身價值為在扣除吸引、銷售、服務顧客的成本之後，企業從顧客的整個交易之終身時間中獲得的收入總額，所計算出的貨幣時間價值。
Kumar et al.(2004)	顧客終身價值為顧客與企業間的整個終身之累積現金流量總額—使用 WACC 折現。
Gupta & Lehmann (2003)	從顧客上所產生的所有未來獲利之現值。
Ryals(2002)	顧客終身價值是顧客未來購買的現值，顧客終身價值的計算是基於個別顧客的期望購買，利用折現率調整成現值。
Knie-Andersen(2001)	顧客終身價值是未來貨幣總額的折現，與未來總額的定義不同。
Berger & Nasr(1998)	企業從顧客交易的整個生命週期之淨獲利或損失，因此顧客終身價值是企業從顧客的整個交易生命週期淨收入去減掉吸引、銷售以及服務顧客的成本所考量的貨幣價值。
Bitran & Mondschein (1996)	顧客終身價值為顧客在企業當中所產生的總淨獲利之折現值。
Courtheoux(1995)	顧客終身價值為預期顧客未來貢獻到經常帳及利潤的淨現值。
Jackson(1994)	則導入交易以及契約的概念，指出顧客終身價值為所有與企業顧客交易以及契約所獲得的獲利貢獻之淨現值(Net Present Value, NPV)。

Roberts & Berger (1989)	顧客終身價值為所有未來貢獻到經常帳及利潤的淨現值。
----------------------------	---------------------------

資料來源：本研究整理

2.3.2 顧客終身價值模型簡介

顧客終身價值的模型相當多樣，在不同的情境、分析需要以及產業差異會有不同的模型型式，本研究針對目前較為常用的終身價值模型加以探討，並且將各種模型予以分類，以做為本研究模型之基礎。顧客終身價值模型基本的型式為顧客終身價值 $=\Sigma(\text{total revenue} - \text{total cost})/(1+d)^t$ ，其中 t =期數， d =每期之折現率，後來的學者則根據此為基礎，在不同的基礎假設之下，考量各種因素於模型中，而產生更具精確及複雜的模型，本段落將根據文獻，將這些顧客終身價值模型加以區分為：顧客終身價值基本模型、顧客終身價值基本模型之延伸，包括整體顧客終身價值以及個別顧客終身價值模型，最後則介紹個別顧客終身價值中的細部元素之分析。

A. 顧客終身價值基本模型

顧客終身價值的基本模型如下所示(Jain & Singh, 2002 ; Hwang et al., 2004) ,

$$CLV = \sum_{t=1}^n \frac{R_t - C_t}{(1+d)^t} \quad (1)$$

其中

t =顧客現金流量的期數

R_t =第 t 期的顧客收入

C_t =在第 t 期所產生的收入 R_t 所產生的總成本

n =顧客所待的總期數

d =折現率

此模型分析的時點是站在顧客過去的歷史資料做折現，找出他進公司之後的貢獻價值。

B. 整體顧客終身價值模型

為了考量顧客之購買行為的差異於模型當中，許多學者都根據Jackson(1985)所提的將工業購買者區分為二類，第一類為不輕易轉變型(lost-for-good)顧客，此顧客若不是完全跟某賣方購買，不然就是完全向另一賣方購買；第二類為多方嘗試型(always-a-share)顧客，此顧客會輕易的轉移賣方，其移轉成本低，因此忠誠度亦不高。許多模型都參考此兩購買者行為來建構整體的顧客終身價值模型，如下所述：

(1)Dwyer(1989)之顧客維持模型(customer retention model)及顧客遷移模型(customer migration model)的顧客終身價值模型

Dwyer(1989)以不輕易轉變型及多方嘗試型的關係型態，發展出兩種顧客獲利性模型的形式：顧客維持模型及顧客遷移模型。前者依據不輕易轉變型態(lost-for-good)，後者依據多方嘗試型態(always-a-share)，來對Jackson(1985)的理論作延伸。

(2)Berger & Nasr(1998)之顧客維持模型及顧客遷移的顧客終身價值模型(延伸Dwyer(1989))

Berger & Nasr(1998)參考Dwyer(1989)的研究，亦針對Jackson(1985)提出的這二類顧客之特色加以整理，對該二類型的顧客之顧客終身價值，在各種的假設情況之下，提出了五種類型的模型，其中四種為顧客獲得模型，一種為顧客遷移模型，這些模型反映了銷售和成本的波動(Blattbeg & Deighton, 1996; Jain & Singh, 2002)，Mulhern(1999)指出，由於考量各種不同的假設，因此這些模型具有參考價值。

(3)Gupta & Lehmann(2003)考慮顧客維持率與成長率模型

以Dwyer(1989)為基礎，將顧客分為活動(active)與非活動(inactive)的兩群，根據基本顧客終身價值模型發展出考慮顧客維持率與成長率的模型。

C.個別顧客終身價值

(1)Mulhern(1999)考慮個別顧客變動行銷成本之顧客終身價值模型

Mulhern(1999)以個人層級為基礎來計算顧客獲利性，利用財務折現的概念，將每一顧客的銷貨收入，減掉產品成本與行銷成本所得出的累積折現值。此模型適用在一系列顧客獲利的間斷時間，所有購買的每段期間都必須被計算，行銷成本的變數在此模型中亦被定義出來，此模型之特色在於易於使用。

(2)Hwang et al.(2004)考慮忠誠顧客以及個別顧客潛力價值模型

Hwang et al. (2004) 提出考慮忠誠顧客以及個別顧客潛力價值之個別顧客終身價值模型，Hwang et al. (2004) 定義潛在價值為當有額外的服務時，從特定顧客所獲得的利潤的潛在價值。目前的顧客終身價值模型都專注根據過去的獲利資料來計算財務上的貢獻，以及轉換成現值，這裡所提的模型，不只是只有獲利貢獻，還包括未來財務的貢獻、獲利潛力以及期望的服務期數。根據顧客的社會人口統計資訊以及交易資料來計算顧客終身價值，找出目前價值、潛在價值以及顧客忠誠。此模型特色在於考量顧客潛在價值，將顧客忠誠的觀念導入顧客終身價值模型中。

(3) Bauer et al. (2003) 的完整個別顧客終身價值模型

Bauer et al. (2003) 探討顧客終身價值的各種成份，將所涉及到的成份予以分為三類，分別為收入、成本以及維持率，針對此三類成份加以深入探討，最後將所有顧客終價值相關的成份全部納入到所開發的顧客終身價值模型中，而形成完整的模型

(comprehensive model)，其涵蓋的分析元素包括：獲得成本、維持率、自動收入

(Autonomous revenue)、向上收入、交叉銷售收入、推薦活動之毛貢獻、銷售成本、關係上的終端成本(Termination costs)、資訊價值(Information value)、合作價值(Cooperation value)、創新價值(Innovation value)等，模型的特色在於考量各種顧客的潛在價值，發展出完整的顧客終身價值模型。

D. 個別顧客終身價值中的細部元素分析模型

(1) Schmittlien et al. (1987) 的預期下期再交易機率之SMC模型：

SMC模型是由Schmittlien et al. (1987) 以統計機率模型來求算出個別顧客的未來活動性機率以及預期購買次數，此模型用來發掘顧客潛在的活動性，適合的應用在百貨公司貴賓卡、一般包裝消費品、郵寄目錄等領域中，詳細的模型運算詳見3.3.2。

(2) Schmittlien & Peterson (1994) 預期再交易金額(延伸SMC模型)：

預期再交易金額是延伸SMC模型所發展出的機率分析模型，它根據每位顧客在資料庫中的平均銷售以及交易量變異情形，根據信賴係數來計算出下期可能的購買金額，因此在計算顧客未來購買潛力上是相當有幫助。

本研究將採用SMC模型來找出未來再交易機率以及利用延伸SMC模型來分析出未來再交易金額來做為顧客潛在價值之衡量。

2.3.3 顧客終身價值模型的考量因素

Mulhern(1999)指出，顧客終身價值的其它研究方向，則必須要去發展三個重要的知識基礎：

- 1、什麼因素使得顧客獲利
- 2、什麼因素決定顧客的獲利程度不等
- 3、怎樣的顧客終身價值衡量是應該被應用在行銷規劃中。

精確分析顧客終身價值是有重要的行銷策略涵意，因此行銷研究人員必須要去考慮許多因素來持續做顧客終身價值分析，適當的顧客終身價值分析項目應該要考量策略情境，而且不同的應用有不同的項目(Mulhern, 1999)，並非僅只有會計面的考量。

Ryals(2002)則針對會計帳面考量來計算顧客終身價值可能的問題加以探討，指出三個衡量顧客終身價值的問題，1、零售商可能會以衡量產品而非衡量顧客之獲利而產生誤導；2、顧客終身價值獲利僅採用會計觀念是錯誤的，這僅能表現出顧客是明顯的獲利，因此在事實上也可能會破壞股東價值；3、顧客獲利可能會有一些隱藏的價值而錯誤的去評估顧客關係利益。而 Stahl & Matzler(2003)指出，精確的評估關係維持的收入以及成本是有許多挑戰，原因為：1、標準會計不予許配置特定顧客關係的成本；2、只有與顧客交易所得的貨幣之利益是被考慮進會計帳戶中；3、收入和成本會隨著時間變化很大；4、現金流量發生在不同的時點以及不同的風險等級。因此，除了考量會計帳面的資訊之外，仍需考量其它的潛在因素。Bauer et al.(2003)討論許多顧客終身價值的文獻後指出，許多顧客終身價值模型沒有提供特定顧客細節的行銷相關資訊，例如期望的交叉銷售收入或是推薦行為，此外，許多模型也沒有考量顧客維持率的結構(Mulhern, 1999)，些方法的前題假設違反顧客移轉以及轉換買方傾向的經濟真實情形。其它的顧客終身價值模型則分別整合顧客維持率以及顧客維持成本，然而，這些模型仍缺少進一步而且更複雜的元素，例如交叉銷售的收入和偏好(see Berger & Nasr 1998; Dwyer, 1997; Keane & Wang, 1995; Reinartz & Kumar, 2000; Wang & Spiegel, 1994)。Stahl & Matzler(2003)指出，僅以交易量為基礎來衡量顧客終身價值會發生誤導，一般來說，高交易量的顧客通常會有較高的議價力，因此享受在較低的售前及售後服務；另一方面，低交易量的顧客通常會付較高的價格，而且會比高交易量的顧客吸收更多銷售及資源。因此，在衡量顧

客終身價值時應該多加考量潛在的價值，而非僅是帳面上的價值，Grembergen & Amelinckx(2002)指出，顧客終身價值評估主要的問題在於專注在現金流量以及會計準則上的評價，顧客終身價值的許多期望利益是無形、間接或策略性的，Bayon et al.(2002)也指出，顧客關係能有效率以及效果的管理主要在於顧客的價值與潛在價值相呼應，因此得知，考量潛在價值是有其必要性的。Hwang et al. (2004)指出，現今顧客終身價值模型沒有考量顧客的潛在價值，無法有效的從過去顧客終身價值獲利貢獻來得知公司的獲利。

本研究在綜合顧客終身價值的文獻之後發現，在文獻中所提及到模型應考量的因素除了會計帳面的資訊之外，還須包括其它因素，大致可分為：顧客維持、顧客獲得、向上銷售(up-selling)與交叉銷售(cross-selling)、顧客離開、推薦行為、顧客滿意與其它因素等，這些因素都是一般會計所不會列入的潛在重要的因素，因此應是須被考量，彙整如下：

(1)顧客維持與獲得

Thoma(2001)指出，獲得成本的因素在決定顧客獲利時是很重要的，顧客終身價值模型若有包含顧客的獲得以及維持則較有意義，許多研究應必須去建立這樣的模型。Hansotia & Wang(1997)指出，管理主要的責任是去維持以及最大化企業的價值，有兩個基礎的策略或關鍵決策必須加以實現：1、獲得正確的新顧客；2、維持和發展正確的已存在顧客。Hansotia & Wang(1997)亦指出，企業的未來現金流量不只決定於不同種類顧客的經濟支出，還包括吸引率以及獲得顧客之間的平衡。Berger & Nasr(1998)指出，在計算顧客終身價值時，考量獲得成本是相當重要的。Thoma(2001)指出，獲得成本是決定顧客淨獲利的重要決定因素，模型考慮顧客獲得與維持將會使模型更有意義。

(2)向上銷售(up-selling)與交叉銷售(cross-selling)

計算顧客終身價值時，考慮交叉銷售以及向上銷售是很重要的(Kim & Kim, 1999；Kim, 2000)，向上銷售意指銷售那些消費者已經購買的商品，在顧客下次光顧時，再次銷售該產品給顧客以提昇銷售量；交叉銷售意指銷售那些顧客未購買過的商品，提供現有客戶新產品或服務的行銷程序，如此來提昇銷售量。

(3)顧客忠誠(顧客游離)

Knox(1998)指出，許多顧客終身價值都是根據過去獲利貢獻來導出未來現金流量，

然而，在一些產業中，考量未來現金流量來計算顧客終身價值是不適當的，因為競爭激烈，結果使得顧客游離，因此游離顧客會影響顧客終身價值。Hwang et al. (2004)亦指出，顧客價值模型應該要考慮到顧客的忠誠與否。

(4)顧客滿意

與顧客終身價值相關的其它因素包括顧客滿意(Mulhern, 1999)。Anderson et al. (1994)發現在衡量個別顧客層級時，顧客滿意和廠商獲利間呈正向關係，此關係隱涵著個別顧客的滿意與獲利間存在著關係，因此，考量到顧客終身價值模型當中是有其必要性的。

(5) 推薦行為(recommendation behavior)

推薦行為和口碑溝通的效果也可視為有潛力的顧客，他們會推薦他們喜歡的產品或零售商的經驗給週遭的人(Murray, 1991)。

(6)其它因素

顧客獲利也有可能被許多其它行為因素所決定，包括價格敏感度、品牌忠誠、購買時機以及知覺因素，例如知覺價值、品牌態度以及對廣告的態度。Mulhern(1999)以及 Bayon et al.(2002)指出，顧客終身價值的方法並非完整的，特別是在分配支出的流動、個別顧客的付出、口碑傳播、領導使用者、偏好、意見價值上等的潛力考量。

因此，綜合以上的探討得知，顧客潛在價值考量因素包括如下：顧客獲得、顧客維持、顧客忠誠、顧客滿意、推薦行為、知覺價值、品牌態度、對廣告的態度、領導使用者、忠誠者、偏好、意見價值、過去交易資料。本研究則將透過資料探勘技術，將顧客獲得、維持、忠誠、交叉銷售等潛在價值因素考量進來。

2.3.4 顧客終身價值模型考量因素的難處

顧客終身價值在衡量時，最困難之處在於很難去獲得個別顧客購買行為的精確資訊(Mulhern, 1999)。由於顧客終身價值計算的主要問題在於企業缺少顧客導向的資料庫創造以及資訊系統，企業很少有如此細節的顧客資訊(Hoekstra & Huizingh, 1999)。

此外，獲得顧客是相當難去衡量的，因為新顧客的資訊是缺少的，如此使得行銷研究人員很難去選擇目標顧客，這將會引起非效率的行銷努力(Hwang et al., 2004)，

Wheation(2000)指出，解決新顧客的獲得資訊問題可以用過去的資料來支持，也就是在資料庫中有較長交易記錄的資料來加以維護，利用可能有較長的生命週期之顧客來與新獲得的顧客來做比較，此外，要達到準確衡量是相當困難的，特定的顧客是否在某一個期間將要去交易是相當具有不確定性的，每個個別顧客有不同的活動以及非活動期間的模式(Kumar et al., 2004)。

2.3.5 顧客終身價值的重要性與好處

近幾年，顧客終身價值已逐漸受到重視(Blattberg et al., 2001；Reinartz et al., 2000；Jain & Singh, 2002)，Jain & Singh (2002)指出，顧客終身價值模型提供系統化的方式來去瞭解所專注的顧客，評估顧客的終身價值以及分析企業在終身價值上的差異影響，並因此間接地創造出企業的價值。Reinartz & Kumar(2000)解釋近幾年顧客終身價值被感到興趣的原因，包括三個理由：

- 1、越來越多的企業瞭解到顧客終身價值在顧客管理程序的觀念。
- 2、行銷科學機構(marketing science institute)把顧客終身價值評估為最重要的研究議題。
- 3、在顧客終身價值的領域當中，實證研究是缺乏的。

確實，由於顧客終身價值分析對於顧客關係管理的效果是正面的，因此顧客終身價值與顧客關係管理逐漸的連接一起，Storbacka(1998)指出，顧客終身價值是顧客關係管理的核心觀點，它提供一些評估顧客獲利分配的衡量方式。而顧客終身價值的影響已成為開發和維持長期獲利顧客關係上的關鍵議題，它在顧客的獲得以及維持上扮演顯著的角色(Ryals, 2002)。Mulhern(1999)指出，當行銷活動透過直接或是互動溝通的形式而須要更加精準的確認出顧客時，顧客獲利便成為行銷策略中的核心角色，因為並非所有顧客都對於企業有財務上的吸引，因此根據顧客終身價值的顧客獲利以及資源配置便是重要的關鍵(Stahl & Matzler, 2003)。Mulhern(1999)指出，個別顧客終身價值的知識可以協助組織市場區隔、行銷資源分配以及行銷組合元素的應用。Jain & Singh(2002)指出，由於網際網路的發展，個別顧客的資訊可以被有效的取得，這樣新的資訊來源也使得行銷研究者重新對於顧客終身價值有興趣。再者，許多網路公司並沒有高價值的實體資本，這樣的公司必須將無形資源正確的評估到會計帳戶當中，因為他們顧客的價值基礎是公司所擁有的最重要無形資產，瞭解顧客終身價值的公司會讓公司更精確瞭解其潛力。

Pfeifer & Robert(2000)指出，顧客終身價值在互動行銷中是個重要且有用的觀念，Jain & Singh(2002)指出，顧客終身價值模型可以協助企業制定策略以及戰略決策，定義出誰是長期的顧客以及他們的特徵，以及行銷資源配置和行銷活動。由於每個顧客對於企業的价值都有差異，企業將可以去評估每個顧客的終身價值以及使用資訊去針對每個顧客來規劃不同的行銷活動(Kumar et al., 2004)，再者，顧客終身價值可以為企業量化顧客的關係以及制定決策結構的架構(Jain & Singh, 2002)，因此，企業應瞭解個別顧客的獲利來去選擇哪一個顧客去溝通以及特定的溝通內容(Mulhern, 1999)，如此可以使企業去瞭解誰是獲利的顧客、配置行銷資源於顧客區隔市場中(Jain & Singh, 2002)以及使企業獲得顧客終身價值的知識，讓企業去發展特定顧客的行銷方案，以加強行銷方案的有效性及其效果。Kumar et al.(2004)指出，計算出顧客終身價值之後可以去根據價值大小來應用行銷策略，如此使得顧客價值最大化，藉以維持顧客生命週期，並且讓企業獲利增加，如下圖 2-6 所示：

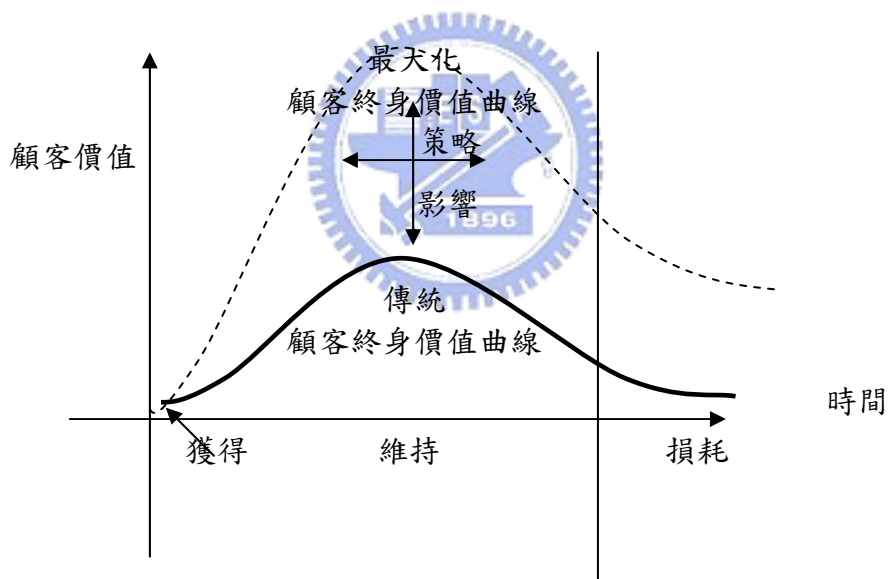


圖 2-6 管理顧客終身價值流程與策略執行後的效果

資料來源：Kumar, V., Ramani, G. & Bohling, T. “customer lifetime value approaches and best practice applications,” *Journal of Interactive Marketing*, Vol.18, No 3, summer. , 2004.

綜合以上各學者對於顧客終身價值的用處及好處來看，結果可歸納出下列幾點：

- 1、使企業獲得顧客終身價值的知識(Jain & Singh, 2002)
- 2、協助企業制定有效率及效果的行銷策略及戰術(Jain & Singh, 2002)

3、瞭解個別顧客之獲利，掌握個別顧客之特徵(Kumar et al.2004；Jain & Singh, 2002)

4、有效配置行銷資源(Mulhern, 1999；Jain & Singh, 2002)

5、提昇企業的獲利(Kumar et al., 2004)

2.4 顧客價值發現－資料探勘

資料探勘在近幾年開始被廣為探討，在行銷的領域裡頭，資料探勘可以找出資料庫裡的顧客資訊，並且進一步分析出顧客的潛在行為，對於顧客關係管理有實質幫助，本段落針對資料探勘文獻加以探討，包括資料探勘定義、資料探勘流程、資料探勘方法、資料探勘應用領域。

2.4.1 資料探勘定義

資料探勘是為了發現隱涵資訊且有潛在價值資訊的一種搜尋以及分析資料的過程，它包括選擇、探索、模型化大量的資料來發現先前未知的樣式以及最後從大量的資料庫中理解資訊(Berry & Linoff, 1997；Fayyad & Piatetsky-Shapiro, 1996; Frawley et al., 1992)。Rygielski et al.(2002)指出，資料探勘的定義為使用統計演算法以及精密的資料搜尋能力來發現資料中的樣式和關聯。Brachman & Anand(1996)指出，資料探勘是從觀察的資料中來粹取樣式或模型。Chung & Gray(1999)認為，資料探勘是從資料中去定義出有效、新奇、潛在有用的關係以及樣式。SAS Institute (1998)定義資料探勘為選擇、探索以及模型化大量資料來發現之前不知道的資料樣式之程序。Han & Kamber(2001)對於資料探勘所做的定義為從龐大的資料中，粹取(extracting)或探勘(mining)出知識。

由以上的定義可以得知，資料探勘可以從大量的資料庫中找出潛在的樣式，進而去瞭解其中的現象，在顧客關係管理中則可以為企業發掘出顧客的潛在行為，以利顧客關係管理策略的制定。

2.4.2 資料探勘流程

Hui & Jha(2000)指出，資料探勘的流程包括建立資料探勘目標、資料的選擇、資料前處理、資料轉換、資料倉儲、資料探勘以及評估資料探勘後的結果，如下圖 2-7 所示。

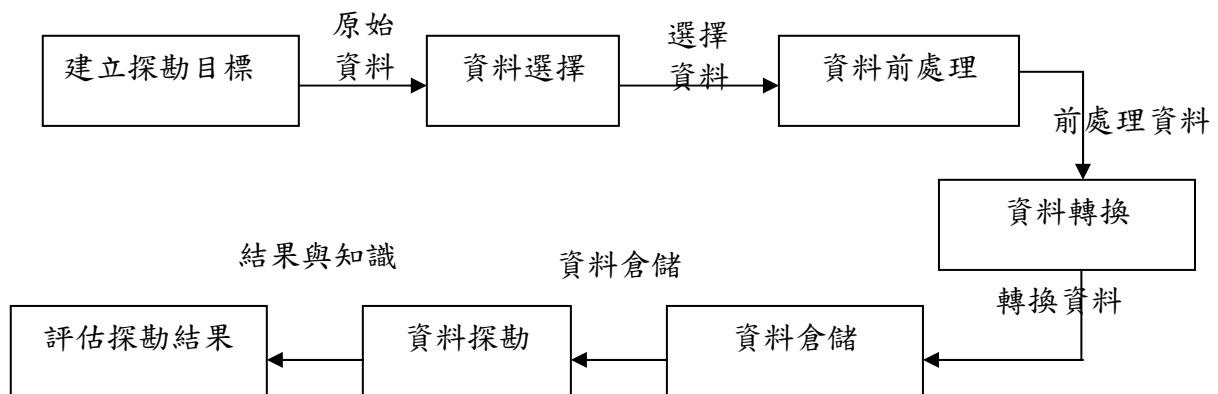


圖 2-7 資料探勘流程

資料來源： Hui & Jha, “Data Mining for Customer Service Support”, Information Management, 38, 1-13., 2000.

Chablo(1999)指出，資料探勘的兩個關鍵流程為：顧客資料轉換以及顧客知識發現。Two Crows Corporation(1999)指出，資料探勘的第一步驟是去描述資料以及彙整統計屬性(例如：平均數以及標準差)，用圖表來視覺化的檢閱，尋找潛在有意義的變數連結；第二步驟必須根據已知結果的樣式去建立預測模型；最後一個步驟是去驗證模型。Two Crows Corporation(1999)亦指出，知識發掘的基本資料探勘步驟為：

- 1、定義企業問題
- 2、建立資料探勘資料庫
- 3、探索資料
- 4、準備建立模型的資料
- 5、建立模型
- 6、評估模型
- 7、展開模型及結果。

其中，在準備建立模型的資料部份有四個主要步驟：1、選擇變數；2、選擇列；3、建立新的變數；4、轉換變數。第 5 步驟的建立模型之任務包括：1、資料蒐集；2、資料偵測；3、選擇；4、資料品質評估以及清理；5、資料合併與整合；6、資料索引(Metadata)

建立；7、裝載資料探勘資料庫；8、維持資料探勘資料庫。

SAS 機構(1998)指出，資料探勘方法由五個步驟，分別為：樣本(Sample)、探索(Explore)、修改(Modify)、模型(Model)以及評估(Asses)。Fong et al.(2002)指出，資料探勘的建立步驟如下：1、建立資料探勘目標：考量成本效益的抵消以及期望精確和結果的價值；2、選擇資料：決定哪個資料是有用的，哪個屬性是有考量價值以及樣本大小應該多大；3、前處理：過濾掉噪音、異常值或不相關資料以及處理遺漏值；4、轉換資料：儘可能減少資料屬性的數量或從已存在的資料屬性中粹取出新的屬性，合併資料表以及將資料調整到適當的的層級；5、儲存資料：整合和儲存資料到單一位置的未歸檔結構中；6、探勘資料：根據資料探勘目標來執行適當的資料探勘功能以及演算法，一般來說，分析人員會先建立資料立方體來提供多構面的資料面，然後執行 OLAP，使用多構面資料立方體結構來發掘知識；7、評估探勘結果：執行各種運作，例如：根據結果來知識過濾、分析粹取知識的有用性、以及呈現結果給使用者，並且促使先前步驟的改變和調整。

由以上的文獻可知，資料探勘流程是一套有系統的資料處理過程，藉以從資料庫中建立起模型及樣式，然後加以應用分析預測，本研究將應用資料探勘的流程，並導入顧客終身價值的分析，以提昇資料探勘後的效果。

2.4.3 資料探勘分析方法

Shaw et al.(2001)指出，各式各樣的資料探勘任務可以廣泛的被區分為 5 類，包括：附屬分析(Dependency Analysis)、分類(Class Identification)、觀念描述(Concept Description)、異常偵測(Deviation Detection)以及資料視覺化呈現(Data Visualization)，如下圖 2-8 所示：

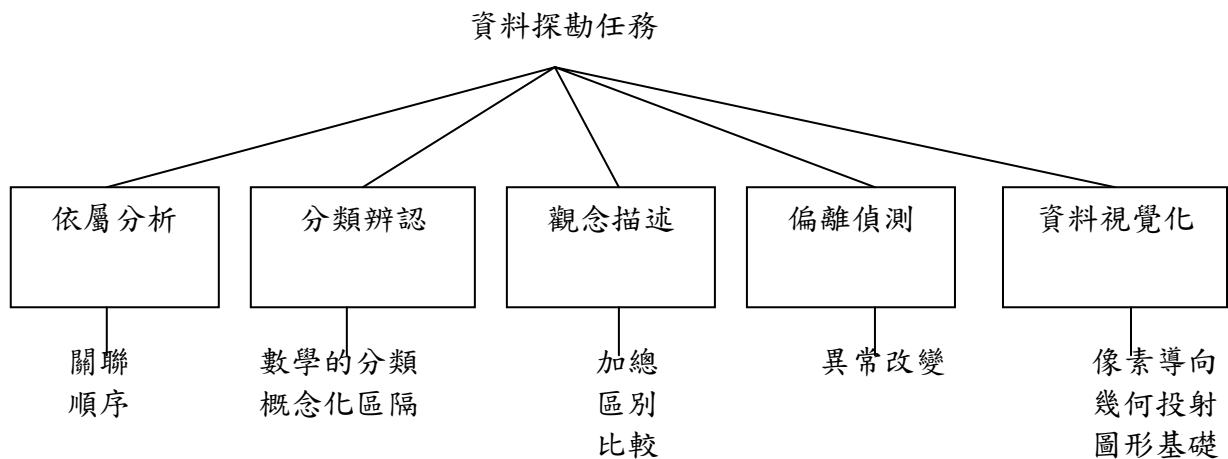


圖 2-8 資料探勘的分類

資料來源: Shaw et. al. "Knowledge management & data mining for marketing". Decision Support Systems, Vol.31 No. 1, pp. 127-137. 2001.

Rygielski et al.(2002)以流程的觀點指出，資料探勘分析包含三類，如下圖 2-9 所示：

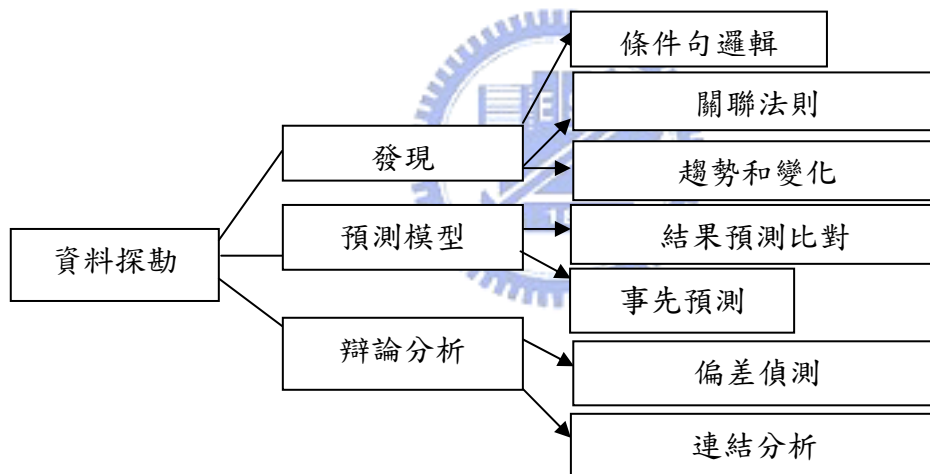


圖 2-9 從流程導向來分類資料探勘方法

資料來源：Rygielski, C., Wang, J.-C., Yen, D.C. "Data mining techniques for customer relationship management." Technology in Society, 24(4), 483-502. 2002.

其中，各分析方法如下：

- 發現(Discovery)：在資料庫中，沒有先決定或是假設樣式是什麼，透過尋找的過程來發現隱藏的樣式
- 預測模型(Predictive Modeling)：從資料庫中找到的模型，並使用此模型來預測未來的流程。
- 辯論分析(Forensic Analysis)：應用粹取樣式到異常或奇特資料元素的流程

根據Han & Kamber (2001) 的分類，資料探勘的功能與所挖掘出來的模式種類共有以下六種，分述如下：

(1)概念/類別描述：特性與區別 (Concept/Class Description: Characterization discrimination) 此一模式是依據分析的對象加以分類，並建立類別 (class)。在描述類別間的特性，有下列兩種方法：

- 資料特性：是指描述類別(class) 內資料一般特性的摘要或目標資料類組的特徵。
- 資料區別：是比較各類組間的特性，以找出各類組間的差異變數為何。

(2)關聯分析 (Association Analysis)

關聯分析是去瞭解X與Y在資料集中的相關性，藉以找出未知的潛在關聯，協助企業在行銷商品上的決策制定，如商品擺設、產品搭銷、交叉銷售與促銷活動的安排等。常用的關聯分析應用為購物籃分析(market basket analysis) 或交易資料分析，從顧客購買的購物籃中之各種項目關聯來分析，瞭解顧客的潛在購買習慣，以幫助企業找出哪些是顧客最常相互購買的商品組合，如買奶油的顧客是否也會同時買麵包?這樣的資訊可以讓商家把奶油與麵包擺在鄰近架上以增加這些商品的銷售量。關聯分析是以規則 (rule) 的方式呈現，亦即“ $X \rightarrow Y$ ”，X、Y皆為資料的集合，此稱為關聯法則 (association rule)，意涵“在資料集中含有滿足條件的X 商品之交易也會同時包含滿足條件的Y商品”，如此的關聯法則存在於資料庫中可能會有上千條(根據資料庫大小及產品結構)，然而並不非每一條法則都是信賴的，因此，有兩種衡量指標必須用來共同衡量關聯法則的可有用性，亦即為信賴度(Confidence)與支持度(Support)，如下表2-3所示。Confidence 一般也可稱為“Accuracy”，指X 如果發生，則Y 也發生的機率，越高的Confidence 值則表示這法則越可靠 (dependable)。

表 2-3 關聯法則之 support 與 confidence 關係

	Confidence low	Confidence high
Support high	準確性低但常發生	準確性高且常發生
Support low	準確性低且不常發生	準確性高但不常發生

因此得知，Support 值是去衡量在資料集中，X、Y產品一起發生的頻率，Confidence 是去衡量X被購買時，Y也同時被購買的機率，簡單寫法為 $X \Rightarrow Y[1\%,50\%]$ ，指X與Y有1%的support及50%的confidence，此兩數值都是衡量法則價值非常重要的指標。

(3)分類與預測 (Classification & Prediction)

分類是指尋找一個能夠描述或區分資料模型(或功能)的流程，進而使用這模型來預測類別未知的資料。在預測的過程中，將以建立好的預測模型(稱為Training data)，透過與其它未知的資料來做預測比較，進而去調整、學習，分析方式有許多不同的方法，如IF-THEN規則，決策樹(Decission Tree)或類神經網路(neural network)。

(4)集群分析 (Cluster Analysis)

集群分析與預測、分類不同點在於不需Training data，資料不須要先知道分類標籤(class label)。在集群分析中相似的資料會群聚在一起(maximizing the intraclass similarity)，但群與群之間極不相似(minimizing the interclass similarity)，而每一群可視為同一個類別等級。

(5)異常值分析 (Outlier Analysis)

資料庫中可能包含與一般資料行為或模型不同的資料，這種資料稱為異常資料(outlier)，而Oullier Analysis亦即利用異常值探勘(Oultier mining)來找出這些異常的資料。大部分的資料挖掘方法都把這類的資料視為噪音或例外而丟棄，但在一些應用中，如保險詐欺的偵測(fraud detection)，這種很少發生的事件會比常發生的事件的案例，使用異常值分析是很有用的。

(6)進化分析 (Evolution Analysis)

資料的進化分析是針對會隨時間改變的模型來做描述與建立趨勢資料，此方法雖可以利用集群、區別、預測或關聯分析來分析時間相關的資料(time-related data)，此外，進化分析還包含時間序列分析(time-series data analysis)、Sequence pattern matching 與 similarity-based analysis。

2.4.4 資料探勘的應用

資料挖掘的應用範圍相當廣泛，不只限於行銷上，其被廣泛應用在生物科技、財務預測、銀行業與醫學等領域。Reinschmidt et al.(1999)指出，目前資料挖掘技術主要應用的類別大致分為市場管理、風險管理及流程管理，如下表2-4所示：

表 2-4 資料探勘應用的領域

市場管理	風險管理	流程管理
目標行銷	預測	存貨最適化
關係管理	顧客維持	品牌控制
通路管理	保險	需求預測
市場籃分析	完成分析	企業評估卡
交叉銷售	健保欺騙	
市場區隔		
網路使用分析		

資料來源：Reinschmidt J., Bhattacharya R., Harris P., & Karanasos A., “Mining Relational and Nonrelational Data with IBM Intelligent Miner for Data”, Appendix B, 189~193. 1998.

本研究主要應用在市場管理中的關係管理領域。

2.4.5 資料探勘重要性與顧客價值發現之關係

Rygielski et al.(2002)指出，應用先進的演算法，資料探勘可以在大量的資料中發現知識，並指出資料中可能的關係，因此資料探勘能夠協助企業指出問題所在。Shaw et al.(2001)亦指出，資料探勘工具可以協助去找出潛在的知識以及更瞭解顧客，系統化的知識管理動力可以傳輸知識到有效的行銷策略中。Chye & Grey(2002)表示，資料探勘用來分析大量的資料以及(或)偵測在資料中的潛在樣式，使得原始資料轉換成有價值的價值資訊。由以上的文獻探討得知，資料探勘可以挖掘出有用的潛在資訊，包括可以協助企業更瞭解顧客的預測模型以及與產品間的關聯性。

綜合以上資料探勘文獻得知，資料探勘對於顧客價值發現具有以下的好處：

- 1、找出資料庫裡的潛在的知識(Shaw et. al. 2001)
- 2、瞭解顧客行為樣式(Shaw et. al. 2001)
- 3、將原始資料轉換成有價值的資訊(Chye & Grey)
- 4、預測顧客行為(Jenkins,1999)
- 5、分析顧客區隔(Han & Kamber, 2001)

2.5 顧客關係管理、顧客價值發現、顧客終身價值與資料探勘之關係

本段落將文獻探討的結果彙整，並且探討顧客關係管理、顧客價值發現、顧客終身價值與資料探勘之關聯性。

2.5.1 顧客關係管理與顧客價值發現的關係

將過顧客關係管理的文獻整理得知，要做好顧客關係管理必須要達到要下 8 點：

- 1、建立適當的策略性願景(Gurau et al., 2003)
- 2、精確的分析顧客終身價值(Hwang et al., 2004; Kim et al., 2003)
- 3、找出顧客區隔，並且加以鎖定(Hwang et al., 2004; Ryals2002)
- 4、將顧客視為企業的資產(Hunt, 1997; Kutner; Cripps, 1997; Wyner, 1996)
- 5、不只強調獲得顧客，更要維持顧客(Hawkes, 2000)
- 6、獲得顧客知識，瞭解顧客潛力(Kim et al., 2003)
- 7、讓顧客感到滿意(Kim et al., 2003)
- 8、與顧客互動良好(Kim et al., 2003)

要達到以上 8 點必須要先瞭解顧客的價值，透過精確的分析顧客終身價值可以瞭解顧客潛在價值及現有價值，此價值是企業的重要資產，也是顧客區隔的重要構面，瞭解有潛力顧客才可以將資源配置完善，以達到維持顧客之效。

2.5.2 顧客終身價值與顧客價值發現之關係

要掌握顧客潛在價值，去瞭解顧客終身價值是非常重要的，因為顧客終身價值模型是一種系統化的方式來去瞭解和評估顧客與企業之間的關係(Jain & Singh, 2002)。顧客終身價值已成為顧客關係管理的中心原則(Morgan & Hunt, 1994)，Mulhern(1999)指出，顧客終身價值的知識可以促使許多行銷方面的決策，包括產品和服務的發展、訂價、整合溝通行銷，最終，顧客終身價值提供行銷資源配置以及市場區隔的量尺，如此可以成

功的發展出有效率及效果的顧客關係管理策略。

根據文獻探討後的結果得知，顧客終身價值的特點有：

- 1、使企業獲得顧客終身價值的知識(Jain & Singh, 2002)
- 2、協助企業制定有效率及效果的行銷策略及戰術(Jain & Singh, 2002)
- 3、瞭解個別顧客之獲利，掌握個別顧客之特徵(Kumar et al.2004；Jain & Singh, 2002)
- 4、有效配置行銷資源(Mulhern, 1999；Jain & Singh, 2002)
- 5、提昇企業的獲利(Kumar et al.2004)

以上特點對於顧客價值的掌握是有正向相關的幫助，因此本研究將採用顧客終身價值分析來找出顧客價值。

2.5.3 資料探勘與顧客價值發現之關係

商場競爭激烈以及顧客可以選擇的項目增加，行銷決策者開始產生新的壓力以及對於維持顧客長期關係產生需求(Shaw et. al. 2001)，由於科技的進步，Berry & Linoff (1997)指出，顧客關係管理在最近科技發展而盛行，特別是資料儲存能力、資料倉儲應用以及資料探勘技術，Rygielski et al.(2002)表示，科技例如資料倉儲、資料探勘以及活動管理軟體使得企業在顧客關係管理獲得競爭優勢有了新的領域，特別是由於資料探勘-從大量的資料中粹取出隱藏的預測資訊-組織可以去確認出有價值的顧客、預測未來行為以及使得企業去制定活動以及知識導向的決策。因此 Rygielski et al.(2002)指出，資料倉儲、資料探勘以及其它顧客關係管理的工具和科技提供新的機會給企業去執行關係行銷的觀念，由此可以得知，資料探勘與顧客關係管理之間密不可分的關係，Rygielski et al.(2002)也指出，資料探勘在整個顧客關係管理過程中扮演關鍵角色。

根據文獻調查後的結果得知，資料探勘應用在顧客關係管理領域的特點有：

- 1、找出資料庫裡的潛在的知識(Shaw et. al. 2001)
- 2、瞭解顧客行為(Shaw et. al. 2001)
- 3、將原始資料轉換成有價值的資訊(Chye & Grey)
- 4、預測顧客行為(Jenkins, 1999)

5、分析顧客區隔(Han & Kamber, 2001)

以上特點對於顧客價值發現之掌握有正相幫助。

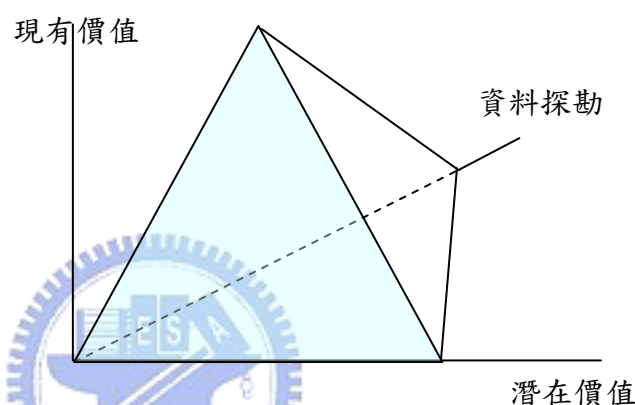
2.5.4 資料探勘、顧客終身價值、顧客價值發現與顧客關係管理之關係

資料探勘以及顧客終身價值在顧客關係管理中扮演重要角色，因為資料探勘可以發現顧客潛在行為樣式，顧客終身價值可以分析出顧客的現有價值及潛在價值，倘若將資料探勘流程與顧客終身價值的分析整合在一起，應將可以相互增進彼此分析顧客價值的效益，因此本研究將此兩種方法的好處相互整合來分析及發現顧客價值，創造出具有吸引力且成功的顧客關係管理策略。彼此間的關係歸納如下表 2-5 所示：



表 2-5 顧客關係管理、顧客價值發現、資料探勘與顧客終身價值之關係

	資料探勘	顧客終身價值分析
定義	為了發現隱涵資訊且有潛在價值資訊的一種搜尋以及分析資料的過程，它包括選擇、探索、模型化大量的資料來發現先前未知的樣式以及最後從大量的資料庫中理解資訊	在扣除吸引、銷售、服務顧客的成本之後，企業從顧客的整個交易之終身時間中獲得的收入總額，所計算出的貨幣時間價值。
目的	從大量資料中找出資訊，並找出行為樣式及模型	計算顧客對於企業的貢獻度
特色	發掘隱藏在大量資料中的模式	以量化的方式衡量顧客的價值，並算出現有及潛在價值
方法	關聯法則、分類、集群等	基本 CLV 模型*、個別顧客之 CLV 模型*、SMC 模型等
分析屬性	人口統計變數、顧客資料、購買商品等	顧客資料、所購買商品的利潤、時間等
與 CRM 的關係	有系統的整理、分析顧客資料庫，藉以瞭解大量資料背後的顧客知識，有效管理顧客	以獲利為基礎，找出對企業有貢獻的顧客以及有潛力的顧客
與顧客潛在價值之關係	找出潛在行為樣式	找出未來可能的交易機率以及可能的交易金額
在顧客潛在價值發現上的缺失	整個資料庫分析，無法精確的以價值為基礎分析出真正的潛在特性。	有些顧客潛在價值無法衡量，例如：交叉銷售、人口統計變數特徵等。

<p>顧客終身價值分析與資料探勘技術整合</p>	<p>改善無法精確分析的問題，透過顧客終身價值分析，找出不同價值族群，並瞭解其潛在行為特性。</p>	<p>改善顧客潛在價值無法衡量的問題，透過關聯法則找出交叉銷售潛力；決策樹來瞭解顧客的潛力特徵。</p>
<p>顧客價值發現整合圖</p>	<p>本研究之顧客價值發現概念如下圖，現有價值以及潛在值找出顧客價值金字塔，資料探勘則是根據各價值集群建立出具有意義的潛在價值發現結構，因此從 2D 平面延伸為 3D 立體圖，清楚瞭解各要素間的關係結構。</p> 	

*詳見第 20-23 頁

資料來源：本研究整理

第三章 研究方法

本章節介紹本研究的架構、方法、模型、顧客資料介紹、分析環境以及資料前處理。

3.1 本研究架構

茲將文獻探討後的結果得知，顧客關係管理要成功必須要達到以下幾點：1、建立適當的策略性願景(Gurau et al., 2003)；2、精確的分析顧客終身價值(Hwang et al., 2004; Kim et al., 2003)；3、找出顧客區隔，並且加以鎖定(Hwang et al., 2004; Ryals, 2002)；4、將顧客視為企業的資產(Hunt, 1997; Kutner; Cripps, 1997; Wyner, 1996)；5、不只強調獲得顧客，更要維持顧客(Hawkes, 2000)；6、獲得顧客知識，瞭解顧客潛力(Kim et al., 2003)；7、讓顧客感到滿意(Kim et al., 2003)；8、與顧客互動良好(Kim et al., 2003)。為了能夠有效的使顧客關係管理運作成功，本研究將資料探勘技術以及顧客終身價值導入，以期能夠有效的發掘潛在顧客。

資料探勘應用在顧客關係管理領域的特點有：1、找出資料庫裡的潛在知識(Shaw et. al., 2001)；2、瞭解顧客行為(Shaw et. al., 2001)；3、將原始資料轉換成有價值的資訊(Chye & Grey, 2002)；4、預測顧客行為(Jenkins, 1999)；5、分析顧客區隔(Han & Kamber, 2001)。

另一方面，顧客終身價值應用在顧客關係管理上擁有以下特點：1、使企業獲得顧客終身價值的知識(Jain & Singh, 2002)；2、協助企業制定有效率及效果的行銷策略及戰術(Jain and Singh, 2002)；3、瞭解個別顧客之獲利，掌握個別顧客之特徵(Kumar et al., 2004; Jain & Singh, 2002)；4、有效配置行銷資源(Mulhern & Francis, 1999; Jain & Singh, 2002)；5、提昇企業的獲利(Kumar et al., 2004)。

因此得知，資料探勘技術能夠瞭解顧客潛在行為樣式，發掘潛在顧客，而顧客終身價值可以以價值為基礎衡量顧客對企業的貢獻，瞭解每位顧客對於企業的重要性，兩者間是相輔相成，本研究亦即將資料探勘技術以及顧客終身價值合

併，有系統的分析顧客價值，並且根據分析結果提出有效的執行顧客關係管理策略，概念性架構如下圖 3-1 所示：

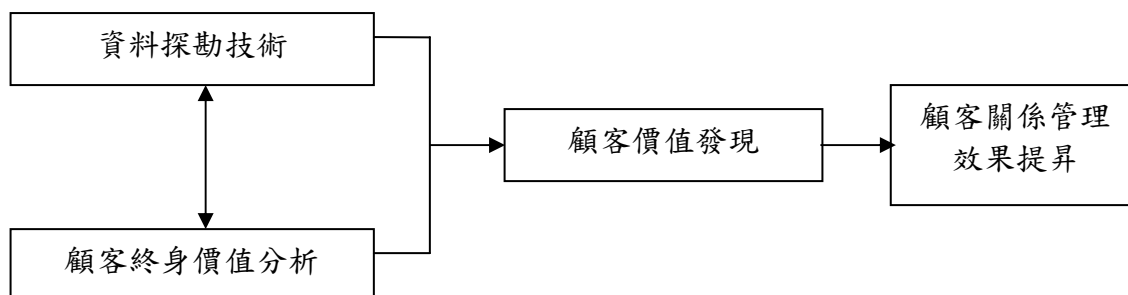


圖 3-1 本研究之概念性架構

3.2 本研究流程

本研究流程係遵循資料探勘步驟，參考圖 2-6 的流程來做為本研究的流程，首先，本研究先將資料庫做前處理，將資料分為訓練集與驗證集，接著利用顧客終身價值分析之顧客現在價值以及利用 SMC 模型來找出訓練集中的每位顧客之預期購買金額以及機率，並計算出其潛在價值，根據潛在價值來做集群分析，得到各個潛在價值顧客群，再利用關聯法則來探討各集群的向上銷售潛力，並且將各顧客的現有價值、未來潛在價值以及其它的人口統計變數等構面建立出一個三維的立方體(cube)，來瞭解潛在價值顧客群組的特性，並根據研究結果來與驗證集做分析預測，探討預測的準精度。此外，本研究以價值集群結果來建立出決策樹，瞭解不同顧客價值群組間的特徵以做為新進顧客的預測與瞭解，最後將這些資料視覺化呈現，並根據結果來研擬行銷策略。整個研究步驟如下圖 3-2 所示：

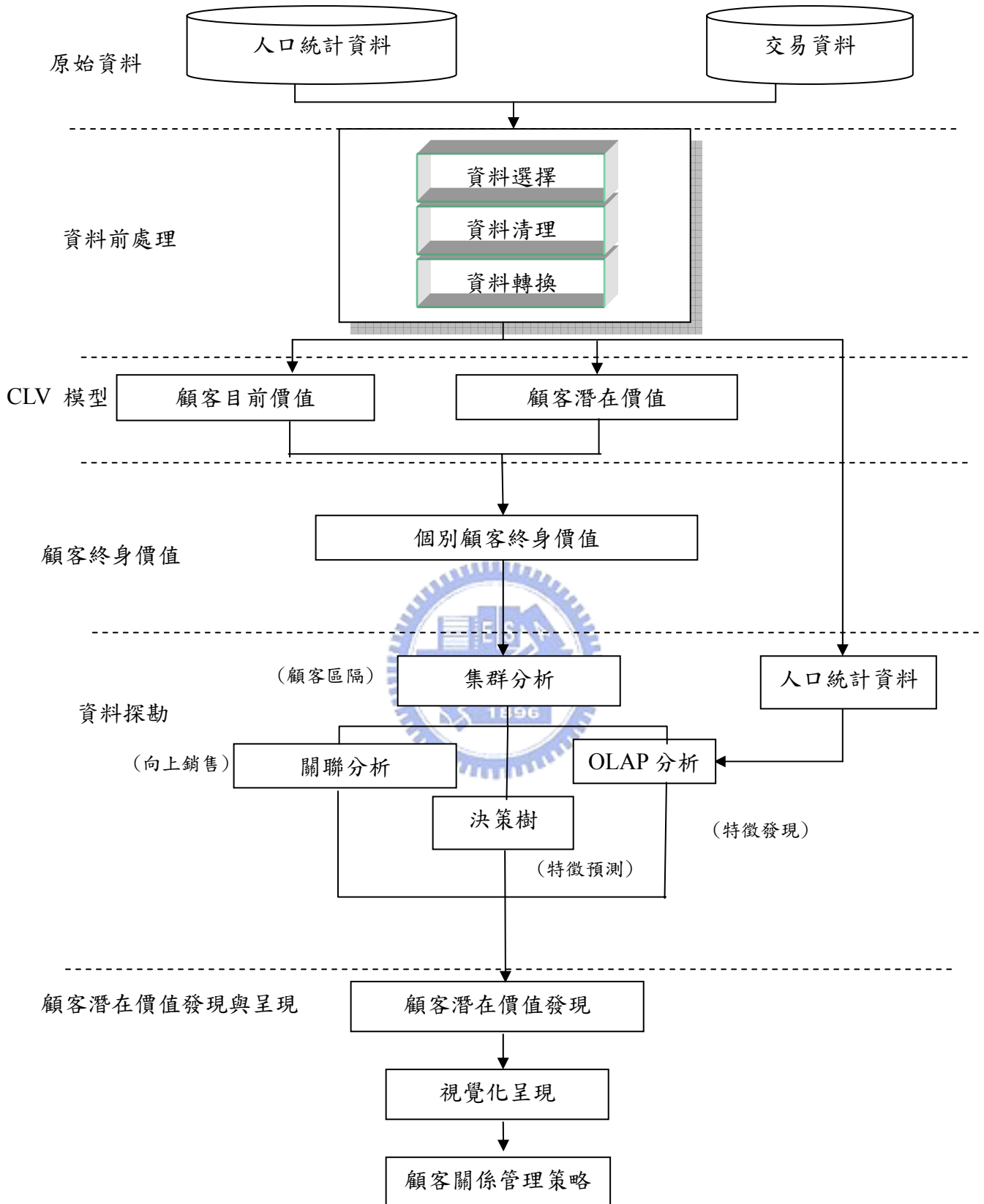


圖 3-2 顧客價值發現分析模式與流程

3.3 研究方法

本研究所使用的方法包括顧客終身價值的分析方法以及資料探勘的技術，其中顧客價值分析包括一般顧客終身價值分析、預期顧客活動率、預期再交易金額，資料探勘技術包括 OLAP 分析、關聯法則、集群分析以及決策樹。

3.3.1 顧客終身價值分析之個別顧客現有價值

本研究在顧客現在價值分析方面，參考 Hwang et al.(2004)所提出的一般顧客終價值分析模型，此模型是以現在的時點來看顧客過去所累積的價值，因此是折現過去的貢獻來瞭解現在時點之價值為多少，公式如下所示：

$$CLV_i = \sum_{t_i=0}^{N_i} \pi_p(t_i)(1+d)^{N_i-t_i} \quad (2)$$

其中

t_i ：顧客 i 的服務期間數

N_i ：顧客 i 的總服務期間數

d ：折現率 interest rate

$E(i)$ ：顧客 i 的期望服務期間數

$\pi_p(t_i)$ ：顧客 i 在 t_i 期間的過去獲利之貢獻程度

$\pi_p(t_i)(1+d)^{N_i-t_i}$ ：表總量代表過去獲利貢獻的淨現值，其中 $\pi_p(t_i)$ 為顧客 i 在 t_i 期間的過去獲利貢獻，而 $(1+d)^{N_i-t_i}$ 為轉換過去獲利到現值的利息率因素(interest rate factor)。

3.3.2 顧客終身價值分析之個別顧客潛在價值

本研究採用 Schmittlien et al.(1987)所提的 SMC 模型以及 Schmittlien & Peterson(1994)所提的 SMC 延伸模型，此兩模型可以計算顧客的預期活動率及購

買金額，適合做為預測顧客潛力的。此兩模型的假設、評估過程及模型如下：

(1)預期再交易機率-SMC 模型

SMC模型是由Schmittlien et al.(1987)以機率模型來求出個別顧客的活動性機率以及預期購買次，當顧客仍具有活動力則稱為活著(Alive)，當顧客離去則稱為死去(Dead)，顧客從第0期被觀察到第T期，在這個期間所購買為X，而最後一次購買時間為t， $0 < t \leq T$ ，因此顧客的資訊包括三個元素：

$$\text{Information}=(X, t, T)$$

而發展的目標在現有的資訊下去計算活著的條件機率：

$$P(\text{Alive} | \text{Information})$$

在計算個別顧客的活動性機率之前，提出五個前題假設，Morrison & Schmittlein(1988)、Uncles & Ehrenberg(1990)與Schmittlien & Peterson(1994)曾對以下的五個假設進行驗證，結果顯示假設是成立的。

假設1.卜瓦松購買：當顧客活著時，每位顧客購買率(λ)符合卜瓦松分配。

當個別顧客具有活動性時，交易率是 λ ，交易次數為X，觀察時期長度為(0, T]，其中0為初次交易的時間，則其未來購買機率呈卜松分配(Poisson distribution)。設顧客不在活動為 τ ，如果顧客仍活動在T(所以 $\tau > T$)，則在(0,T]期間內，購買呈現卜松機率分配，如下：

$$P[X = x | \lambda, \tau > T] = e^{-\lambda T} \frac{(\lambda T)^x}{x!}, x = 0, 1, 2, \dots \quad (3)$$

期望值以及變異數為：

$$E[X | \lambda, \tau > T] = \lambda T; \quad \text{Var}[X | \lambda, \tau > T] = \lambda T$$

假設2.指數生命週期：個別顧客持續具有活動性的時期呈指數分配(exponential distribution)，並且設u為其死亡率(death rate)。

當該顧客不再具活動性的時間(τ)，則其機率密度函數如下：

$$f(\tau | \mu) = \mu e^{-\mu\tau}, \tau > 0 \quad (4)$$

$$E[\tau | u] = 1/u; \quad \text{Var}[\tau | u] = 1/u^2$$

假設3. 個別顧客購買率呈伽瑪分配：不同的顧客之間其交易率(λ)呈現伽瑪分配(gamma distribution)。

每位顧客交易率的不同呈現伽瑪分配(gamma distribution；Schmittlein, Cooper & Morrison, 1993)，其機率密度函數如下：

$$f(\lambda | \gamma, \alpha) = \frac{\alpha^r}{\Gamma(r)} \lambda^{r-1} e^{-\alpha\lambda}; \lambda, \gamma, \alpha > 0 \quad (5)$$

$$E[\lambda | \gamma, \alpha] = \gamma / \alpha; \quad E[\lambda^2 | \gamma, \alpha] = \gamma / \alpha^2$$

參數 γ 是顧客間交易率互異性的不同程度指標。若 γ 值愈大，表示交易率不同程度愈大，反之亦同。上式表示，在顧客全體平均交易率之期望值為 γ/α 的情況下，個別顧客交易率為 λ 之機率。

假設4. 死亡率呈現伽瑪分配：不同的顧客之間其死亡率(u)呈現伽瑪分配(gamma distribution)

顧客失聯率的不同亦呈伽瑪分配(Schmittlein, Cooper & Morrison, 1993)。其機率密度函數如下：

$$g(\mu | s, \beta) = \frac{\beta^s}{\Gamma(s)} \mu^{s-1} e^{-\beta\mu}; \mu > 0, s, \beta > 0 \quad (6)$$

$$E[u | s, \beta] = s / \beta; \quad E[u^2 | s, \beta] = s / \beta^2$$

參數 s 是顧客間失聯率互異性的不同程度指標。若 s 值愈大，表示交易率不同程度愈大，反之亦同。上式表示，在顧客全體平均失聯率之期望值為 s/β 的情況下，個別顧客失聯率為 μ 之機率。

假設5. λ 和 u 的比率相互獨立

(3)至(6)的購買活動是遵循NBD模型(Ehrenberg, 1972)，將購買事件和期間模型合併起來則被稱做Pareto/NBD。Schmittlien & Peterson(1994)將Schmittlien et al.(1987)所發展出的模型命名為SMC模型，並且認為SMC模型能夠掌握交易時的各種情形假設。

預期購買次數

Schmittlien et al.(1987)根據Ehrenberg(1972)所發展出的NBD模型推演出預期

購買次數，如果顧客在期間(T)之前為不活動狀況時，則那段期間稱為 θ ，則在 θ 期間所購買的X期望值為：

$$E[X | \gamma, \alpha, s, \beta, T, \theta] = \frac{\gamma}{\alpha} \theta \quad (7)$$

(7)式的期望可能 θ 值產出為：

$$E[X | \gamma, \alpha, s, \beta, T, \theta] = \frac{\gamma}{\alpha} E[\theta | \gamma, \alpha, s, \beta, T] \quad (8)$$

則 θ 的分配只有根據期間 s 和 β 參數，則 θ 與T的期望值的關係為：

$$E[\theta | \gamma, \alpha, s, \beta, T] = \frac{\beta}{(s-1)} \left[1 - \left(\frac{\beta}{\beta+T} \right)^{s-1} \right] \quad (9)$$

將(9)代入(8)則可以得到期望購買次數，如下：

$$E[X | \gamma, \alpha, s, \beta, T] = \frac{\gamma\beta}{\alpha(s-1)} \left[1 - \left(\frac{\beta}{\beta+T} \right)^{s-1} \right] \quad (10)$$

其變異數為

$$\text{Var}[X | \gamma, \alpha, s, \beta, T] = \frac{\gamma}{\alpha} E[\theta | s, \beta, T] - \frac{\gamma^2}{\alpha^2} (E[\theta | s, \beta, T])^2 + \frac{\gamma(\gamma+1)}{\alpha^2} E[\theta^2 | s, \beta, T]. \quad (11)$$

其中

$$E[\theta^2 | s, \beta, T] = \frac{2\beta}{s-1} \left[\frac{\beta}{s-2} - \frac{\beta}{s-2} \left(\frac{\beta}{\beta+T} \right)^{s-2} - T \left(\frac{\beta}{\beta+T} \right)^{s-1} \right] \quad (12)$$

α 、 β 、 γ 、及 s 參數估計

Schmittlien et al.(1987)發展出一種估計 α 、 β 、 γ 、及 s 參數的三步驟方法。首先假設隨機抽取M個顧客，其中第 i 位顧客在 $(0, T_i]$ 的時期中交易 $X_i=x_i$ 次，而最後一次的交易時間為 t_i 。如果我們得知每一位抽樣顧客的購買資訊 (X_i, t_i, T_i) ，就可以計算在 $(0, T]$ 中的平均交易次數及交易的變異數。為了求算在 $(0, T]$ 中的平均交易次數及交易的變異數，必須選擇 $T_i > T$ 的顧客，並令 $N(T)$ 為 $T_i > T$ 的顧客總數，且 $T(i)$ 為 $\{T_i | i=1, \dots, M\}$ 中第 i 大者。同樣的，令 $(X(i), t(i))$ 為觀察時期為 $T(i)$ 的顧客其交易資訊。完成了上述的假設後，在 $(0, T]$ 中估計的平均交易次數為：

$$\hat{E}[X | T] = \frac{1}{N(T)} \sum_{i=1}^{i_T} X_{(i)} \quad (13)$$

i 是 $T_i > T$ 中最大者。而 $(0, T]$ 中估計之變異數為

$$\hat{Var}[X | T] = \frac{1}{N(T)-1} \sum_{i=1}^{i_T} \left(X_{(i)} - E[\hat{X} | T] \right)^2 \quad (14)$$

因此將(10)與(11)式中 T 值不同，配合所計算出來的 $E[X|T]$ 及 $Var [X|T]$ 值，即可以估計 α 、 β 、 γ 、及 s 之值。在此須瞭解的是，(10)式 T 值的變動所得到的只是 γ/α ， s ， β 的函數。若要再求其餘參數 (α 、 γ) 的估計值，則必須(11)式來求算出來。

Schmittlien et al.(1987)將上述的過程歸納為以下之三步驟：

步驟1：選擇兩個 T 值， T_a 及 T_b ，使 $T_a, T_b \in T(1)$ ， $T_a \neq T_b$ 。利用(11)式，我們可以得到(15)式，如下：

$$\frac{E[X | \gamma, \alpha, s, \beta, T_a]}{E[X | \gamma, \alpha, s, \beta, T_b]} = \frac{1 - (\beta / (\beta + T_a))^{s-1}}{1 - (\beta / (\beta + T_b))^{s-1}} \quad (15)$$

將(15)式等號左方以 $E[X|T_a]/E[X|T_b]$ 取代之，再選取另一對不同的 T 值， T_c, T_d ，則可以經由解以下二式得到 s, β 的估計值 $\hat{s}, \hat{\beta}$ ：

$$\frac{\hat{E}[X | T_a]}{\hat{E}[X | T_b]} = \frac{1 - (\beta / (\beta + T_a))^{s-1}}{1 - (\beta / (\beta + T_b))^{s-1}}$$

$$\frac{\hat{E}[X | T_c]}{\hat{E}[X | T_d]} = \frac{1 - (\beta / (\beta + T_c))^{s-1}}{1 - (\beta / (\beta + T_d))^{s-1}} \quad (16)$$

步驟2：選取 T_e ，使 $T_e \neq T_a, T_b, T_c, T_d$ ，並且 $T_e \in T(1)$ 。利用(16)式為基礎，以下式可求得 γ/α 之估計值 $\hat{r} / \hat{\alpha}$ ：

$$\hat{r} / \hat{\alpha} = \frac{\binom{\hat{s} + 1}{\hat{\beta}}}{\hat{\beta}} \frac{\hat{E}[X | T_e]}{1 - \left(\hat{\beta} / (\hat{\beta} + T_e) \right)^{\hat{s}-1}} \quad (17)$$

步驟3：以 $\hat{r} / \hat{\alpha}, \hat{s}, \hat{\beta}$ 代換(11)式中之 $\gamma/\alpha, s, \beta$ ，等號左方並代換為已求算出來的 $Var [X|T_f]$ 值 ($T_f \in T(1)$)。解方程式便可求出 α 之估計值 $\hat{\alpha}$ 以及 γ 之估計值 $\hat{\gamma}$ 。

將各參數之估計值出來後，亦即可據此估計值來求算個別顧客的活動性機率。

顧客活動性之計算

估計出 \hat{s} ， $\hat{\beta}$ ， $\hat{\alpha}$ 以及 $\hat{\gamma}$ 之後，比須針對 $\hat{\alpha}$ 及 $\hat{\beta}$ 的値之情形來計算出顧客活動性機率，共有三種情況。假設 t 為最近一次的交易時間：

情況1：若 $\alpha > \beta$

$$P[\tau > T | \gamma, s, \alpha > \beta, X = x, t, T] =$$

$$\left\{ 1 + \frac{s}{\gamma + x + s} \left[\left(\frac{\alpha + T}{\alpha + t} \right)^{r+x} \left(\frac{\beta + T}{\alpha + t} \right)^s F(a_1, b_1; c_1; z_1(t)) - \left(\frac{\beta + T}{\alpha + T} \right)^s F(a_1, b_1; c_1; z_1(T)) \right] \right\}^{-1}$$

其中

$$a_1 = \gamma + x + s; b_1 = s + 1; \quad c_1 = \gamma + x + s + 1; \quad z_1(y) = \frac{\alpha - \beta}{\alpha + y} \quad (18)$$

情況2：若 $\alpha < \beta$

$$P[\tau > T | \gamma, s, \alpha < \beta, X = x, t, T]$$

$$\left\{ 1 + \frac{s}{\gamma + x + s} \left[\left(\frac{\alpha + T}{\beta + t} \right)^{r+x} \left(\frac{\beta + T}{\alpha + t} \right)^s F(a_2, b_2; c_2; z_2(t)) - \left(\frac{\alpha + T}{\beta + T} \right)^{r+x} F(a_2, b_2; c_2; z_2(T)) \right] \right\}^{-1}$$

$$\text{其中 } a_2 = \gamma + x + s; b_2 = \gamma + x; \quad c_2 = \gamma + x + s + 1; \quad z_2(y) = \frac{\beta - \alpha}{\beta + y} \quad (19)$$

情況3：若 $\alpha = \beta$

$$P[\tau > T | \gamma, s, \alpha = \beta, X = x, t, T] = \left\{ 1 + \frac{s}{\gamma + x + s} \left[\left(\frac{\alpha + T}{\alpha + t} \right)^{r+x+s} - 1 \right] \right\}^{-1} \quad (20)$$

在(18) 及(19) 式中， $F(a, b; c; z)$ 是高斯超幾何函數(Gauss hypergeometric function)。根據Abramowitz & Stegun(1972)的介紹，高斯超幾何函數為：

$$F(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)} \frac{z^n}{n!} \quad (21)$$

計算出個別顧客的活動性機率以後，只要將這些機率相加，便可以得出具有活動性顧客人數的估計值(Schmittlien et al.,1987)，進而判斷顧客下期購買的可能性機

率。

(2) Schmittlien & Peterson(1994)預期再交易金額

Schmittlien & Peterson(1994)提出SMC延伸模型，此模型主要是去評估下期可能購買的交易量，考量每位顧客過去的平均銷售以及每位顧客的變異量，並計算出信賴係數，其背後的運算邏輯在於：若顧客購買變異大，則顧客的購買變化會很大，因此其下期購買的交易量影響會大，反之亦然。

在計算延伸SMC模型須符合三個前題假設：

- a. 當某特定客戶有X次的再交易紀錄，令 Z_i 為第i次再交易的交易金額($i=1, \dots, X$)。假設 Z_i ($i=1, \dots, X$)為相互獨立的常態隨機變數，其平均值為 θ ，變異數為 σ_w^2 。 σ_w^2 乃為該客戶自身不同的再交易金額的變異數。
- b. 每位客戶的平均交易金額(θ)不同，呈常態分配，其平均值為 $E(\theta)$ ，變異數為 σ_a^2 。 σ_a^2 為客戶間平均交易金額的變異數。
- c. 假設每位客戶的平均交易金額(θ)的分配，乃是與交易率(λ)、失聯率(μ)的分配相互獨立。Schmittlien & Peterson(1994)指出，若某客戶過去只有一次的再交易紀錄，該筆再交易的交易金額，取決於信賴係數(ρ_1)。信賴係數即為客戶間平均交易金額的變異數(σ_a^2)佔總變異($\sigma_a^2 + \sigma_w^2$)的比例，可以決定所要預測數量的大小。

$$\rho_1 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_w^2} \quad (22)$$

若某客戶僅有一次再交易紀錄 ($X=1$)時，則該客戶的預期再交易金額為：

$$E[\theta | Z_i] = \rho_1 Z_i + (1 - \rho_1) E[\theta] \quad (23)$$

若某客戶過去有X筆再交易紀錄($X > 1$)，則先必須求算該X筆交易金額的平均值(\bar{Z})：

$$\bar{Z} = \frac{1}{X} \sum_{i=1}^X Z_i \quad (24)$$

其信賴係數 ρ_x 為：

$$\rho_x = \frac{\sigma_a^2}{\sigma_a^2 + (\sigma_w^2 / X)} \quad (25)$$

因此預期再交易金額為：

$$E[\theta | Z_1, \dots, Z_x] = \left(\frac{X\sigma_A^2}{X\sigma_A^2 + \sigma_w^2} \right) \bar{Z} + \left(\frac{\sigma_w^2}{X\sigma_A^2 + \sigma_w^2} \right) E[\theta] \quad (26)$$

(3)顧客潛在價值計算

本研究將 Schmittlien et al.(1987)所提的顧客活動率(probability of active customer)以及 Schmittlien & Peterson(1994)所提的再購買金額，彼此相乘而形成預測個別顧客未來的可能潛在價值，其模型之各符號的涵意以及模型如下所示。

$$\text{Potential Value}_i = \text{Prob}(\text{Active})_i \times E(\text{Volume})_i \quad (27)$$

其中

$\text{Prob}(\text{Active})_i$ ：顧客*i*之未來再購買機率(活動力)

$E(\text{Volume})_i$ ：顧客*i*之未來再交易金額

Potnetial_i ：顧客*i*之潛力價值



3.3.3 資料探勘分析

本研究所採用的資料探勘分析中，採用關聯分析找出潛在購買行為樣式，OLAP 分析瞭解顧客的特徵，決策樹則是去瞭解顧客的特徵規則，以做為未來預測新來顧客的依據，各方法描述如下：

1、關聯分析

本研究採用關聯分析來找出顧客潛在行為，目的是要瞭解在交易資料庫中項目之間的關係，找出同一個交易中隱含其它項目之存在的項目。表示方式為： $X \rightarrow Y$ ，亦即 $X_1 \wedge \dots \wedge X_m \rightarrow Y_1 \wedge \dots \wedge Y_n$ ，其中 X_i (for $i \in \{1, \dots, m\}$) 以及 Y_j (for $j \in \{1, \dots, n\}$)。透過關聯規則分析商品之間關聯性，以便了解消費者的購物習慣，找出商品項目之間的關聯規則。

關聯法則分析最早是使用 Apriori 演算法 (Han & Kamber, 2001; Agrawal &

Srikant,1994)。Apriori 利用簡單且循序漸進的方式，來找出資料庫項目間的關聯規則。Apriori 演算法缺點為需多次掃描整個資料庫及產生候選項目集，不僅浪費時間也浪費空間；為增加Apriori演算法效率，研究者提出很多改進之演算法。目前大都以FP_growth(Frequent-pattern growth)演算法來減少資料庫掃描次數及不產生候選項目組以增快速度及提高效率(沈清正等，2002)。FP_growth 演算法可以避免多次審視整個資料庫及無須產生候選項目集，本研究所用的DBMiner2.0軟體亦即使用FP_growth 演算法。

FP_growth 的優點：

- (1)建構FP_tree (Frequent-pattern tree) ，用FP_tree 結構將資料庫壓縮。
- (2)完整的找出「常出現的項目集」。
- (3)避免多次審視整個資料庫，可以節省時間。
- (4)可用於時間序列資料的挖掘。
- (5)不產生候選項目集。

FP_growth演算法的輸入為原始交易資料庫及最小支持度門檻值，輸出為「常出現樣式」，主要演算法內容包括下列兩項。

- (1)建立FP_tree：先審視資料庫一次，找出單項目集並計算其出現次數，再依次數以降冪排列。
- (2)從FP_tree去尋找出「常出現樣式」。

FP_growth 的演算法程序如下所示：

```

Procedure FP_growth( Tree, a )
{
if Tree contains a single path P
then for each combination (denoted as  $\beta$ ) of the nodes in the path P do
generate pattern  $\beta \cup \alpha$  with support = minimum support of nodes in  $b$ ;
else for each  $\alpha_i$  in the header of Tree do {
generate pattern  $b = \alpha_i \cup \alpha$  with support =  $\alpha_i$ .support;
construct  $\beta$ 's conditional pattern base and then  $\beta$ 's conditional FP_tree  $Tree_\beta$ ;
if  $Tree_\beta \neq \phi$ 
then call FP_growth( $Tree_\beta, \beta$ ) }
}

```

資料來源：Han, J. & Kamber, M., "Data Mining: Concepts and Techniques," Hardcover Ed: Morgan Kaufmann, 2001.

2、集群分析

本研究採用 K-means 集群分析法，K-mean 模式可以將顧客群集分為指定的群數，而且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。K-means 方法的群的個數是固定的，這個凝聚塊方法是一個自底向上的過程。首先將每個元素作為群，然後一點點的擴大。然後把擴大的過程記錄下來。演算法的大致過程：首先建一個 N*N 的 matrix。每個格子上都是兩個元素之間的距離。然後將最相近的元素組合在一起形成一個新的群。由於 K-means 可以分析出集群的重心，找出組內相似，而組間內相異的情形集群個數，因此適合於本研究來應用。

3、決策樹

本研究是採用 Microsoft Decision tree，它是改良自 C4.5 的演算法，Microsoft® Decision Tree 運算法是以分類概念為基礎。運算法會建立一個根據培訓資料集剩餘資料行來預測某一資料行數值的樹狀結構。因此，樹狀結構中的每一個節點就代表資料行的一個特定 Case。這個節點要擺在那裡，是由運算法決定的，深度與其同層級 (Sibling) 不同的節點代表各資料行不同的案例。它能幫助我們很快速地建立決策樹及利用其結果做推論的工作。其中用來評估一個決策樹分隔變數的衡量標準是分散度 (Diversity)，對於資料的分散度指標有多種計算方式。而分散度指標很高表示這個組合中包含平均分配的多個類別，而分散度指標很低則

表示一個單一類別的成員居優勢。C4.5決策樹演算法，是使用亂度函數(Entropy)作為分散度指標，而Microsoft Decision tree則是使用貝氏分數(Bayesian Score)來作為指標。分析結果將會提供許多資訊，包括：決策樹的結構圖、顯示各樹葉節點的規則(Rule)以及預測結果的機率。

4、結果驗證分析

為了驗證預測結果是否與實際情形有效性，因此採用 Kohavi & Provost(1998)所提出的混淆矩陣(confusion matrix)，此矩陣是用來預測與實際結果之準確性。混淆是由 LxL 的構面所組成，其中 L 為不同的標籤值(例如：大、中、小)，以下以 L=2 的矩陣為例，如下表 3-1 所示：

表 3-1 混淆矩陣表

	預測-N	預測-Y
實際-N	a	b
實際-Y	c	d

在計算上，須找出 Accuracy, True positive rate, True negative rate, Precision, False positive rate, False negative rate，計算公式如下：

(1)準確性 Accuracy

準確性在確定整體的預測結果之準確有多高，值越高表越精確，如公式(28)所示：

$$(a+d)/(a+b+c+d) \quad (28)$$

(2)正確率 True positive rate (Recall, Sensitivity)：

正確率目的為瞭解所實際為Y的，有哪些是被預測到的，值越高表示效果越佳，如公式(29)所示：

$$d/(c+d) \quad (29)$$

(3)正確錯誤率 True negative rate (Specificity)：

正確錯誤率是在瞭解實際為錯誤的，預測上亦是錯誤的比率，其正確錯誤率越高表預測越精確，如公式(30)所示：

$$a/(a+b) \quad (30)$$

(4)精確性 Precision

精確性是在瞭解而預測結果為 Y 的，其實際上亦為 Y 的比率，值越高表越精確，如公式(31)所示：

$$d/(b+d) \quad (31)$$

(5)錯誤正確率 False positive rate

錯誤正確率表實際為 N 的比率，其預測結果為 Y，值越低表越好，如公式(32)所示：

$$b/(a+b) \quad (32)$$

(6)錯誤錯誤率 False negative rate

錯誤錯誤率表預測為 N，但實際為 Y 的比例，值越低表錯誤率越低，如公式如(33)所示：

$$c/(c+d) \quad (33)$$

3.4 本研究分析環境介紹



本研究以Microsoft SQL Server 2000來建構資料庫系統，此伺服器提供許多資料庫應用程式的工具，包括資料轉換服務(Data Transformation Services, DTS)、資料庫複寫工作(Replication)、分析服務套件 (Analysis Service)、English Query、Meta Data Services等技術，提供了企業資料倉儲系統建構的功能，以進行分析及決策的工作，其中分析服務套件提供OLAP技術，讓倉儲的龐大資料經過組織化以後，得以透過用戶端工具快速的分析資料，或是利用複雜的資料探勘 (Data Mining) 技術，分析倉儲中的資料，發掘資料的多面貌。所提供的資料探勘技術包括集群分析以及決策樹。

本研究亦使用DBminer2.0資料探勘軟體來做為關聯分析的工具，DBMine是加拿大Simon Fraser大學Jiawei Han 教授的智慧資料庫系統研究實驗室所發展的資料採礦軟體系統，該系統設計的目的是把關聯資料庫和資料採礦整合在一起。DBMiner2.0是由OLAP系統發展的探勘系統，可處理屬性為多層概念資料。

DBMiner2.0系統具有以下特點：

- 1、完成多種知識的發現：群組化規則、歸類規則、關聯規則等。
- 2、綜合多種資料採礦技術：對於屬性的歸納、統計分析、發現多層規則、規則引導發現等方法。
- 3、提供互動式的SQL 語言：為DMQL (data mining query language) 語法，是對於屬性查詢語言。
- 4、能整合處理關聯資料庫。
- 5、快速有效的執行及回應。
- 6、以圖形使用者介面(graphical user interface, GUI)視覺化的呈現資訊。

DBMiner2.0 的關聯法則分析有兩種應用方式：intra-dimensional 關聯法則分析及 inter-dimensional 關聯法則分析。intra-dimensional 關聯法則分析是從同一屬性中，去探討其屬性內部項目之關聯，而 inter-dimensional 關聯法則分析是從不同屬性去找出其關聯；本研究將利用 intra-dimensional 關聯法則分析進行探勘。

其餘輔助計算過程的軟體包括：SPSS, Excel, Mathematica, C⁺⁺。SPSS主要應用於次數交叉表分析；Excel主要應用於顧客現有價值以及潛在價值上的計算；Mathematica主要應用於 α 、 β 、 γ 、及s參數估計；C⁺⁺主要是設計一套程式能夠針對所有個別顧客的參數來計算出Gauss hypergeometric function，以及計算出每個顧客的終身價值之現有價值。

3.5 整體資料庫介紹

本研究資料庫是來自微軟的SQL2000 樣本資料庫，此資料是連鎖超市量販店資料，裡頭的紀錄包括：銷售情形、顧客人口統計資料、產品資訊、店家、倉儲、薪資、存貨等資訊，介紹如下：


- 銷售資料表：為事實資料表(Fact Table)，主要做為各資料表的衡量值以及各資料表聯結，因此資料屬性有各資料表的id，此外還有每位顧客的各產品之銷售值、成本價以及購買數量，共有1997以及1998兩年的資料。
- 顧客資料方面：裡頭詳載顧客人口統計變數、購買紀錄等資料。

- 產品方面：共有 110 種類別，1560 產品項目，資料詳載各產品的屬性，包括：品牌、產品類別、售價、淨重以及產品大小等資訊。
- 此組織方面，該公司公有 12 個部門，公司各層級的員工總計 1155 人。
- 商店方面，商店總數為 24 家，分佈在美國、加拿大以及墨西哥等地，遍及在美洲、南美洲、北美洲的西半部。

3.6 建立資料探勘前準備

建立資料探勘的步驟，本研究參考 Hui & Jha(2000)所提出的流程，在資料探勘前的步驟為：建立資料探勘目標、資料選擇、資料前處理、資料轉換等程序，本研究亦即遵循這幾個步驟來運作。

3.6.1 設立資料探勘目標



本研究的資料探勘目標在於從資料庫中發掘能夠反映出顧客潛在價值的各種因素，包括：交叉銷售、預期購買金額、數量、終身價值以及促銷關係等，並且將顧客潛在價值予以分群，發展的完整的顧客潛在價值概念，以提出顧客關係管理上的策略。因此，資料探勘目標為：

- 1、顧客面：找出具有潛在價值的顧客
- 2、產品面：找出潛在價值顧客之交叉銷售以及潛在購買行為
- 3、策略面：將潛在價值顧客做集群，並且與人口統計變數做交叉分析，以瞭解顧客的特徵

接下來的資料探勘步驟都會以設立的目標來運作。

3.6.2 目標資料選擇

根據資料探勘的目標，本研究選擇與目標相關的資料來做探勘，此步驟分為

兩個部份，一個是資料表的選擇，另一個是資料屬性的選擇。

- (1) 資料表的選擇：本研究的原始資料庫中有許多資料表，包括account, category, currency, customer, days-check, department, employee, expense_fact, inventory_fact_1997, inventory_fact_1998, position, product, product_class, promotion, region, reserve_employee, salary, sales_fact_1997, sales_fact_1998, sales_fact_dec_1998, store, time_by_day, warehouse, warehouse_class，其中篩選出與本研究相關的資料表，分別為customer, product, product_class, promotion, sale_fact_1997, Sale_fact_1998, time_by_day，以這些資料表來做為本研究的探勘目標。
- (2) 資料屬性的選擇：將資料表中的屬性與本研究探勘目標相關的資料篩選出來，如下表3-2所示：

表 3-2 各資料表所選擇的屬性

Table	Attributes
Dimension_P	Customer id, Age, Marital_status, Salary_income, Gender, Total_children, num_children_at_home, Education, Occupation, House owner, num_cars_owned, Date_accnt_opened
Dimension_P	Product_class_id, Product_id
Fact_Table	Product_id, Time_id, Customer_id, Promotion_id, Store_id, Store_sales, Store_cost
Dimension_T	Time_id, The_date, The_month, The_year, quarter

資料來源：本研究整理

3.6.3 資料探勘資料表建立

經過篩選與資料探勘目標相關的資料之後，將這些資料設計成資料庫，本研究將資料庫設計成星狀結構(Star Schema)，如下圖 3-3 所示，其中銷售為事實資料表，而顧客、產品、時間以及促銷為構面資料表，有些資料屬性是須透過資料前處理的值，包括 First_transaction_time, Latest_transaction_time, Total_number_of_transactions, Expected_active_pro, Expected_amount 以及 Group_no，這些屬性皆是為了計算顧客終身價值分析以及找出顧客集群時所須的資訊。

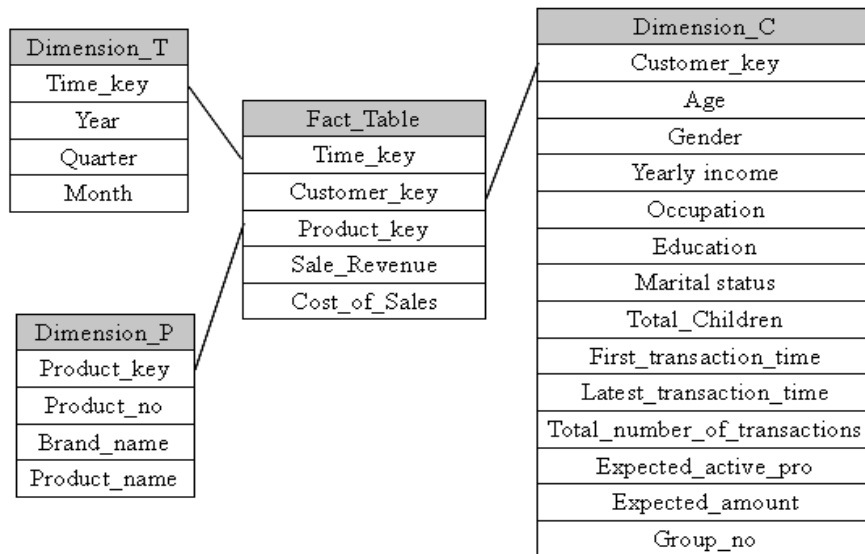


圖 3-3 本研究資料庫結構

3.6.4 資料庫前處理

(1) 資料清理

資料清理的目的主要是為了處理遺漏值，移除錯誤的值以及不一致的值，以避免這些值影響分析結果的精確度；本研究在處理遺漏值、錯誤的值以及不一致的值方面，當屬性的值屬於名目資料則使用 SQL 語言來找出來，使用的語言方式是以 group by 以及計數的方式來找出這些值，以顧客資料表的 Age 以及 gender 屬性為例，SQL 語言如下所示：

SELECT Age, count (gender) as total

FROM Dimension_C

GROUP BY Age

則得到結果為：

country	total
20-29	1717
30-39	1205
40-49	7359

country	total
	1
20-29	1717
30-39	1204
40-49	7359

country	total
20-29	1717
30-39	1204
40-49	7359
111	1

表 3-3 遺漏值偵測

附註：此為舉例，並非資料庫真實內容

由上表 3-3 即可看出 Age 是否有遺漏值，左邊的表無遺漏值，中間的表則是有遺漏值，同理若要查 gender 是否有遺漏值，則 group by gender，即可看是否有遺漏值，以此類推；若要瞭解是否有值錯誤或是不一致，則亦可以找出，如最右邊的表所示，發現 111 是錯誤的值，其它的值以此類推來加以清理。

若為屬性資料為連續的資料，則以遞增排序的方式來找出，若為遺漏值則會排在最前面，由此來找出遺漏值。不一致或是錯誤的值方面，則一樣先以排序的方式來偵測，例如：store_sale 值皆在 \$0.5~\$23.64 之間，若排序發現有低於或高於此區間的值，亦即為錯誤或是不一致的值，若難以判斷時，則以迴歸分析來，輔以散佈圖來判斷值是否有偏離，並且以 R-Square 上昇與否來看刪掉值之後的效果。

本研究將所找出的遺漏值、錯誤值以及不一致的值予以刪除掉。

(2) 資料刪除

由於本研究是資料庫是 2 年，分別為 1997 及 1998 年，其顧客是必須完整涵蓋這兩年的購買資料，若是有顧客在 1997 年以及 1998 年之後才加入的，必須予以刪除，否則無法計算顧客終身價值。使用 SQL 語言來篩選出來，語法如下：

```
DELETE FROM Demension_C
```

```
WHERE date_accnt_opened > 1997/1/1
```

(4) 資料轉換

本研究資料轉換的值包括：birthdate, date_accnt_opened, store_cost, store_sales, time_id 等，轉換過程如下：

- birthdate 轉換成 1997 年的年齡，例如：1977 年的 1 月 1 號，會轉換成 20 歲，以此類推，並產生新的屬性值“age”。

- Date_acct_opened 轉換成進入會員幾年，例如：1994 年 1 月 1 號進入會員，則入會員共 3 年，以此類推，並產生新的屬性值”into_accunt”。
- 由於要計算顧客終身價值，因此 Store_sales 與 store_cost 相減，得到新的屬性”profit”。
- 由於要計算各 SMC 模型時所要用的 T 值，因此 Time_id 將原本 1997 年 1 月轉換成 1，2 月轉換成 2，依此類推至 1998 年 12 月，日期與 T 值的對應情形如下表 3-4 所示：

表 3-4 本研究之時間與 T 值對應表

時間	T 值	時間	T 值
1997 年 1 月	1	1998 年 1 月	13
1997 年 2 月	2	1998 年 2 月	14
1997 年 3 月	3	1998 年 3 月	15
1997 年 4 月	4	1998 年 4 月	16
1997 年 5 月	5	1998 年 5 月	17
1997 年 6 月	6	1998 年 6 月	18
1997 年 7 月	7	1998 年 7 月	19
1997 年 8 月	8	1998 年 8 月	20
1997 年 9 月	9	1998 年 9 月	21
1997 年 10 月	10	1998 年 10 月	22
1997 年 11 月	11	1998 年 11 月	23
1997 年 12 月	12	1998 年 12 月	24

(5)資料分割

本研究以 20/80 比例來將資料分為驗證集以及訓練集，因此訓練集共有 19 個月，驗證集共有 5 個月，屆時將會找出訓練集的潛在顧客模式來與驗證集做預測，以驗證本研究的預測效果。

3.7 整理後資料庫描述

本研究資料庫共有 4 個資料表，分別銷售資料表、顧客資料表、產品資料表、時間資料表，其每個資料表的內容概述如下：

- 銷售資料表：此資料庫裡頭共有 1997 以及 1998 年的資料，1997 年共計有

86837 筆交易資料，1998 年有 182883 筆交易資料，總計 269720 筆資料，資料分割後，訓練集共 189982 筆資料(19 個月)，而驗證集共 79738 筆資料(5 個月)。

顧客交易資料之每個月交易人數以及平均交易次數方面，本研究將 24 月資料予以彙整，設 T 為月，則每個月的交易次數、每月交易人數以及平均每月交易之描述統計結果彙整如下表 3-6 所示：

表 3-5 本研究資料之每個 T 值的平均交易次數

T	交易次數	交易人數	平均交易次數	T	交易次數	交易人數	平均交易次數
1	7034	1677	4.194394753	13	14936	3092	4.830530401
2	6844	1611	4.248292986	14	14392	3007	4.786165614
3	7710	1818	4.240924092	15	14924	3128	4.771099744
4	6590	1562	4.218950064	16	14552	2960	4.916216216
5	6866	1622	4.233045623	17	14556	2983	4.879651358
6	6912	1604	4.309226933	18	14741	3002	4.910393071
7	7752	1830	4.236065574	19	15044	3112	4.834190231
8	7038	1675	4.201791045	20	14517	3016	4.813328912
9	6663	1561	4.268417681	21	15432	3147	4.903717827
10	6479	1533	4.226353555	22	14125	2915	4.845626072
11	8232	1940	4.243298969	23	17339	3653	4.746509718
12	8717	2089	4.172809957	24	18325	3771	4.859453726

資料來源：本研究整理

- 顧客資料表：此資料庫裡共有 10281 位會員資料，其中有 8842 位顧客有購買紀錄，1986 位顧客無購買，無法計算顧客的潛力值，因此予以刪除。資料分割後，訓練集的顧客數為 8295 人，驗證集的顧客總數為 6141 人(僅列出在訓練集中繼續在驗證集中購買的顧客)，其整體會員的人口統計變數資料彙整如下表 3-5 所示：

表 3-6 本研究之顧客人口統計變數資料

gender	frequency	education	frequency	yearly_income	frequency
F	5097	Bachelors Degree	2619	\$10K - \$30K	2222
M	5184	Graduate Degree	539	\$30K - \$50K	3327
Total	10281	High School Degree	3039	\$50K - \$70K	1845
houseowner	frequency	Partial College	990	\$70K - \$90K	1207
N	4094	Partial High School	3094	\$110K - \$130K	493
Y	6187	Total	10281	\$130K - \$150K	506
Total	10281	marital_status	frequency	\$150K +	223
num_cars_owned	frequency	M	5142	\$90K - \$110K	458
0	587	S	5139	Total	10281
1	2266	Total	10281	total_children	frequency
2	3103	member_card	frequency	0	1038
3	2913	Bronze	5703	1	2025
4	1412	Golden	1198	2	2068
Total	10281	Normal	2420	3	2059
occupation	frequency	Silver	960	4	2063
Clerical	205	Total	10281	5	1028
Management	1461	Country	frequency	Total	10281
Manual	2583	Canada	1717		
Professional	3382	Mexico	1205		
Skilled Manual	2650	USA	7359		
Total	10281	Total	10281		

資料來源：本研究整理

- 產品資料表：產品資料表中共有 1560 筆資料，可劃分為食物、飲料以及非消耗品三大類，47 種產品類別，110 種產品子類別以及 111 個產品品牌。
- 時間資料表：時間共 730 天資料，24 個月份，從 1997 年 1 月 1 日至 1998 年 12/31 日。

第四章 分析－顧客終身價值分析與資料探勘技術

分析方面，本研究已將資料前處理完成，因此開始做資料分析，首先先做顧客終身價值分析，找出現有價值以及潛在價值，接下來根據顧客終身價值分析結果分群，找出最具有價值的顧客，針對最具有價值顧客群，透過資料探勘的關聯分析、OLAP 分析以及決策樹來描述顧客群以及分析潛在行為。

4.1 顧客終身價值分析

顧客終身價值分析分為兩部份-現有價值以及潛在價值，現有價值部份，本研究根據(2)式來計算顧客終身價值。潛在價值部份，採用(27)式為來計算，此部份為預期顧客潛在價值，本研究將顧客終身價值分析結果，分此兩部份來說明，此外，本研究亦針對顧客現有價值以及潛在價值的一些因素，例如：最近購買、購買頻率、第一次交易時間等數值來做相關性分析，藉以瞭解哪些因素與顧客現有價值以及潛在價值是有相關性的。

4.1.1 顧客終身價值之顧客現有價值

(1)每位顧客總淨利

為求算顧客現有價值，本研究先找出每位顧客在訓練集中每期對於企業獲利的加總($\sum_{t_i=0}^{N_i} \pi_p(t_i)$)，因此求算出每位顧客在期間內(19 期)所購買的總額以及銷售成本，選擇事實資料表，計算結果彙整如下表 4-1 所示(由於資料繁多，故僅列出前 20 位顧客)：

表 4-1 每位顧客總淨利

單位：美元

顧客編號	個別顧客 總銷售收入(R _t)	個別顧客 總銷售成本(C _t)	淨利 $\sum_{t_i=0}^{N_i} \pi_p(t_i)$
3	\$203.17	\$77.73	\$125.44
5	\$1.08	\$0.52	\$0.56
6	\$112.74	\$43.68	\$69.06
8	\$112.60	\$47.59	\$65.01
9	\$40.17	\$15.10	\$25.07
10	\$162.99	\$67.88	\$95.11
11	\$44.35	\$17.90	\$26.45
14	\$48.94	\$18.52	\$30.42
17	\$50.17	\$20.46	\$29.71
18	\$26.48	\$9.56	\$16.92
19	\$13.37	\$5.94	\$7.43
20	\$59.23	\$24.05	\$35.18
21	\$107.40	\$46.43	\$60.97
23	\$76.96	\$32.33	\$44.63
24	\$91.56	\$35.65	\$55.91
26	\$56.86	\$24.56	\$32.30
27	\$7.52	\$3.38	\$4.14
28	\$62.06	\$23.79	\$38.27
30	\$121.47	\$49.55	\$71.92
31	\$18.24	\$7.10	\$11.14

(2)折現率

在折現率(d)方面，由於本研究是採用美國大型零售店的資料，因此在折現率的選擇上是引用美國在 1997 年 1 月至 1998 年 7 月期間(19 期)的利率值，其各期的平均利率彙整如下表所示：

表 4-2 1997 年~1998 年之美國利率²

期數	利率	期數	利率
24	4.93%	12	5.65%
23	5.10%	11	5.60%
22	5.23%	10	5.58%
21	5.50%	9	5.58%
20	5.55%	8	5.58%
19	5.58%	7	5.59%
18	5.55%	6	5.60%
17	5.55%	5	5.63%
16	5.53%	4	5.63%
15	5.55%	3	5.40%
14	5.53%	2	5.28%
13	5.57%	1	5.35%

註：各期是以月為單位，若單月調整利率二次以上，則計算該月的平均利率。

(3)顧客現有價值計算

在顧客現有價值計算方面，由於要瞭解每位顧客的購買時間，因此找出每位顧客的交易時間，將總服務期數(19)減掉該次的交易時間，成為該次交易折現率(1+d)的次方，再乘上該期的獲利，再將所有期數的獲利值加總來找出顧客現有價值。各期的顧客購買期間結果彙整如下表 4-3 所示(由於資料繁多，故僅列出前 20 位顧客的第一次及第二次交易的時間)：

²資料來源：<http://www.frbdiscountwindow.org/historicalrates.cfm?hdrID=20&dtlID=52>，存取於 2005 年 2 月。

表 4-3 每位顧客被服務期間

單位：月

顧客編號	總服務期間 (N_i)	第一次交易時間(t_1)	第二次交易時間(t_2)
3	19	4	5
5	19	1	0
6	19	7	7
8	19	15	15
9	19	16	17
10	19	1	2
11	19	19	19
14	19	7	8
17	19	8	10
18	19	13	15
19	19	7	11
20	19	1	6
21	19	3	5
23	19	3	8
24	19	17	17
26	19	10	14
27	19	19	19
28	19	1	2
30	19	3	9
31	19	16	16

附註：本研究是以月為一期，而有些顧客會在同一個月交易兩次以上，因此第一次和第二次的交易時間會相同，如顧客 6、8、11、24、27、31。

將每位顧客各期的獲利折現後加總，與顧客總淨利結果做比較，結果彙整如下表 4-4 所示(由於資料繁多，故僅列出前 20 位顧客)：

表 4-4 每位顧客之獲利與現有價值

顧客編號	顧客獲利 $\sum_{t_i=0}^{N_i} \pi_p(t_i)$	總交易次數	第一交易時間	顧客現有價值 $\sum_{t_i=0}^{N_i} \pi_c(t_i)(1+d)^{N_i-t_i}$
3	125.44	30	4	283.37
5	0.56	1	1	1.43
6	69.06	12	7	173.60
8	65.01	19	15	191.65
9	25.07	7	16	65.08
10	95.11	26	1	317.18
11	26.45	7	19	26.45
14	30.42	11	7	19.91
17	29.71	8	8	75.95
18	16.92	5	13	41.12
19	7.43	5	7	32.49
20	35.18	11	1	19.61
21	60.97	20	3	354.97
23	44.63	9	3	178.77
24	55.91	13	17	49.74
26	32.30	7	10	182.24
27	4.14	3	19	21.69
28	38.27	9	1	157.75
30	71.92	19	3	9.86
31	11.14	6	16	118.38

由上表 4-4 得知，若單純以獲利來加總，其結果會與折現後的結果不同，因為不同的期間購買會有不同的價值，因此加總後應該會有所不同，這亦即為顧客現有價值分析的特色。

4.1.2 顧客終身價值之顧客潛在價值

本研究顧客潛在價值計算分為兩部份，分別計算顧客活動性機率以及顧客預期再交易金額，再將兩者相乘即可得到顧客潛在價值，分析結果如下。

(1) 顧客活動性機率

本研究根據 Schmittlien et al.(1987)所發展出的 SMC 模型來計算顧客活動性機率，首先先估計 α 、 β 、 γ 、及 s 參數。估計過程分為三個步驟，如下：

A. 步驟 1：選擇 T_a, T_b, T_c, T_d

本研究訓練集共有 19 期($T=19$)，每個月份代表 1 期，根據表 3-5 的 T 值與每個月份的對照表之編號，本研究選擇 $T_a=17$ ， $T_b=15$ ， $T_c=16$ ， $T_d=14$ ，代入(16)式後求算出 s 與 β 的估計值 \hat{s} 、 $\hat{\beta}$ ，分別為 13.04645531、64.65946206。

B. 步驟 2：選取 T_e

令 $T_e=18$ ，找出平均數為 4.9103(見表 3-5)，變異數為 11.9921，利用(17)式為基礎，可求得 γ/α 之估計值 $\hat{\gamma}/\hat{\alpha}$ 為 1.125086，因此將此結果代入至(11)式得到 $\hat{\gamma}$ 及 $\hat{\alpha}$ ，分別為 0.23044 及 0.20482。

C. 步驟 3：計算活動性機率

本研究估計出 α 值以及 β 的結果，其大小為 $\alpha < \beta$ ($0.20482 < 64.65946206$)，此特性適用於(19)式，因此將估計出的值代入至(19)式中，藉以求出每位顧客的活動性機率，在計算過程中必須要瞭解每位顧客最近一次交易時間以及總購買次數，並且計算(19)式中各個數值，因此利用 SQL 語言找出所須資訊，並加以計算，彙整如下表 4-5 所示(由於資料繁多，故僅列出前 20 位顧客)：



表 4-5 每位顧客活動性機率計算過程

顧客編號	總購買次數(x)	最近交易的時間(t)	$\frac{s}{\gamma + \chi + \beta}$	$\left(\frac{\alpha + T}{\beta + t}\right)^{r+x}$	$\left(\frac{\alpha + T}{\beta + t}\right)^s$	$\left(\frac{\alpha + T}{\beta + T}\right)^{r+x}$
3	30	16	0.308595397	6.06E-19	1	2.083E-19
5	1	1	0.982643533	0.753305	14.64891	0.7124013
6	12	18	0.537402132	7.61E-08	0.726476	6.646E-08
8	19	19	0.417127569	2.23E-12	0.620993	2.233E-12
9	7	16	0.676792352	0.000131	1	0.0001043
10	26	18	0.340844136	1.02E-16	0.726476	7.501E-17
11	7	19	0.676792352	0.000104	0.620993	0.0001043
14	11	18	0.560489496	3.27E-07	0.726476	2.895E-07
17	8	12	0.643414838	4.5E-05	1.941737	2.393E-05
18	5	13	0.755138876	0.00271	1.639672	0.0019784
19	5	7	0.755138876	0.003809	4.681123	0.0019784
20	11	13	0.560489496	6.2E-07	1.639672	2.895E-07
21	20	19	0.404204159	5.13E-13	0.620993	5.125E-13
23	9	17	0.613174766	6.7E-06	0.851502	5.494E-06
24	13	19	0.516141526	1.53E-08	0.620993	1.526E-08
26	7	14	0.676792352	0.000153	1.387596	0.0001043
27	3	19	0.85399913	0.037542	0.620993	0.0375418
28	9	10	0.613174766	1.4E-05	2.74142	5.494E-06
30	19	18	0.417127569	2.78E-12	0.726476	2.233E-12
31	6	16	0.713822293	0.00055	1	0.0004542

高斯超幾何函數(Gauss hypergeometric function)的計算上，利用(21)式來運算，計算結果再與表 4-5 所列之數值代入至(19)式中，得到每位顧客的活動性機率，結果彙整如下表 4-6 所示(由於資料繁多，故僅列出前 20 位顧客)：

表 4-6 每位顧客活動性機率

顧客編號	$F(a_2, b_2; c_2; z_2(t))$	$F(a_2, b_2; c_2; z_2(T))$	顧客活動性機率
3	1.8109	1.652012	0.18347379
5	1.153	1.010496	0.02136552
6	1.224	1.297464	0.96152504
8	1.3663	2.032917	0.9803279913
9	1.1342	1.232249	0.81115477
10	1.5784	1.680811	0.92774641
11	1.1123	1.112324	0.9801818155
14	1.2021	1.253595	0.96285106
17	1.2129	1.131325	0.18488613
18	1.1113	1.07563	0.50805638
19	1.2105	1.07563	0.02687974
20	1.2875	1.832554	0.16495824
21	1.39	1.46388	0.9801119396
23	1.1694	1.150731	0.88326121
24	1.2322	1.623711	0.9803338827
26	1.1542	1.275332	0.53551994
27	1.0411	1.041083	0.9803455147
28	1.2993	1.438832	0.0268351
30	1.3896	1.46388	0.94478605
31	1.1118	1.093747	0.82138743

(2)再交易金額計算

再交易金額計算方面，參考 Schmittlien & Peterson(1994)所提的計算方式，利用(22)、(23)、(24)、(25)、(26)式來求出。首先先計算平均銷售以及變異數，利用 SQL 語法找出其值，分別為平均值 6.54，變異數 11.99212042，此外，為了計算只有 1 次交易的顧客資料，由於此顧客變異數為 0，因此必須計算出所有顧客的平均變異數來做為只有 1 次交易顧客的變異數指標，計算結果為 10.83772799。接下來計算每位顧客的平均交易金額以及變異數，彙整如下表 4-7 所示(由於資料繁多，故僅列出前 20 位顧客)：

表 4-7 預期再交易金額計算

顧客編號	平均銷售 \bar{Z}	變異數 σ^2	信賴係數 ρ_x	預期再交易金額 $E[\theta Z_1, \dots, Z_x]$
3	\$6.77	6.30379092	0.982779699	6.76833588668314
5	\$1.08	10.837727*	0.525282525	3.67295431891114
6	\$9.40	25.17122727	0.884027693	9.06414260492637
8	\$5.93	9.761724561	0.980987263	5.93800804359841
9	\$5.74	2.466547619	0.984424832	5.75111464724340
10	\$6.27	10.25374662	0.973158397	6.27613580996909
11	\$6.34	6.976595238	0.923267874	6.35153751077321
14	\$4.45	5.939869091	0.965829349	4.52061917293355
17	\$6.27	19.18884107	0.833322849	6.31643617243535
18	\$5.30	3.69168	0.942002394	5.36827081723802
19	\$2.67	4.02918	0.93703401	2.91755874705682
20	\$5.38	15.48618727	0.965370975	5.42458655918292
21	\$5.37	13.88265263	0.952079446	5.42616768124628
23	\$8.55	8.378986111	0.927958687	8.40636900147747
24	\$7.04	12.16275641	0.966208394	7.02617040526631
26	\$8.12	11.90255714	0.937937812	8.02479209387423
27	\$2.51	0.406033333	0.98883984	2.55173570837108
28	\$6.90	17.90840278	0.936435305	6.87312988016770
30	\$6.39	10.2675117	0.964110437	6.39854395596930
31	\$3.04	0.95228	0.986938071	3.08574418313605

*表採用整體平均變異數

(3)顧客潛力計算

顧客潛力計算上，本研究將預期再交易金額以及預期相乘加以得到，彙整如下表 4-8 所示(由於資料繁多，故僅列出前 20 位顧客)：

表 4-8 顧客潛在價值

顧客編號	顧客活動性機率 $Prob(Active)_i$	預期再交易金額 $E(Volume)_i$	顧客潛在價值 $Prob(Active)_i \times E(Volume)_i$
3	0.18347379	6.76833588668314	1.241812259
5	0.02136552	3.67295431891114	0.07847459
6	0.96152504	9.06414260492637	8.715400044
8	0.9803279913	5.93800804359841	6.132769516
9	0.81115477	5.75111464724340	4.665044072
10	0.92774641	6.27613580996909	5.822662449
11	0.9801818155	6.35153751077321	6.467018326
14	0.96285106	4.52061917293355	4.352682974
17	0.18488613	6.31643617243535	1.167821412
18	0.50805638	5.36827081723802	2.727384238
19	0.02687974	2.91755874705682	0.078423218
20	0.16495824	5.42458655918292	0.894830239
21	0.9801119396	5.42616768124628	5.486907989
23	0.88326121	8.40636900147747	7.425019635
24	0.9803338827	7.02617040526631	7.26076207
26	0.53551994	8.02479209387423	4.297436144
27	0.9803455147	2.55173570837108	2.639901922
28	0.0268351	6.87312988016770	0.184441106
30	0.94478605	6.39854395596930	6.0452551
31	0.82138743	3.08574418313605	2.534591486

4.1.3 顧客現有價值與潛在價值因素之相關性分析

本研究將顧客現有價值與潛在價值的各個因素做相關性分析結果，彙整如下表 4-9 所示，由下表可以得知，顧客現有價值與獲利(.992)及購買頻率(.978)有相當的關係。而潛在價值與再交易機率(.944)以及最近購買(.794)有關係，而再交易機率方面則與最近購買有相當的關係，因此得知，總購買次數以及獲利影響現有價值，最近購買影響再交易機率及潛力，這些因素彼此間相互影響。

表 4-9 顧客潛力與顧客購買之相關性分析

	現有價值	潛在價值	再交易機率	再購買金額	購買頻率	最近購買	獲利
現有價值	1.000	.254	.194	.250	.978	.353	.992
潛在價值	.254	1.000	.944	.330	.315	.794	.342
再交易機率	.194	.944	1.000	.086	.294	.835	.290
再購買金額	.250	.330	.086	1.000	.137	.153	.228
購買頻率	.978	.315	.294	.137	1.000	.423	.989
最近購買	.353	.794	.835	.153	.423	1.000	.417
獲利	.992	.342	.290	.228	.989	.417	1.000

4.2 顧客潛在價值分群

根據表 4-3 以及表 4-7 所得的顧客現有價值以及潛在價值來做集群分析，利用 K-means 集群分析，將集群數設定為 3，分析結果再做交叉分析，藉以找出各個顧客的區隔。



4.2.1 顧客現有價值分群

K-means 分析結果，彙整如下表 4-10 所示：

表 4-10 最後集群中心點

	集群		
	1	2	3
顧客現有價值	367.59	98.54	784.96

根據上表 4-10 分析結果予以各集群命名，其中集群 3 值最高，因此命名為高現有價值顧客，集群 1 值次高，因此命名為中現有價值顧客，集群 2 值最低，因此命名為低現有價值顧客，其各群之命名及個數彙整如下表 4-11 所示：

表 4-11 現有價值集群個數

高顧客現有價值	303
中顧客現有價值	1340
低顧客現有價值	6652
總數	8295

4.2.2 顧客潛在價值集群分析

顧客潛在價值集群分析結果彙整如下表 4-12 所示：

表 4-12 最後集群中心點

	集群		
	1	2	3
顧客潛在價值	4.212	.612	6.764

根據上表分析的結果，集群 3 的顧客現在價值最高，因此命名為高潛在價值顧客，集群 1 的顧客潛在價值次之，因此命名為中潛在價值顧客，集群 2 的顧客潛在價值最低，因此命名為低顧客潛在價值，其各集群命名及個數彙整如下表 4-13 所示：

表 4-13 集群個數

高顧客潛在價值	3175
中顧客潛在價值	2061
低顧客潛在價值	3059
總數	8295

4.3 顧客現有價值與顧客潛力價值交叉分析

4.3.1 顧客現有價值與顧客潛在價值個數交叉分析

為了能夠找出最具有貢獻價值的顧客，因此將顧客現有價值以及顧客潛力做交叉分析，分析結果彙整如下表 4-14 所示：

表 4-14 顧客現有價值與顧客潛力之顧客數交叉分析

		低潛在價值顧客	中潛在價值顧客	高潛在價值顧客	總計
高現有價值顧客	個數 (百分比)	22 (.3%)	28 (.3%)	253 (3.1%)	303 (3.7%)
中現有價值顧客	個數 百分比	369 (4.4%)	215 (2.6%)	756 (9.1%)	1340 (16.2%)
低現有價值顧客	個數 (百分比)	2668 (32.2%)	1818 (21.9%)	2166 (26.1%)	6652 (80.2%)
總計	個數 (百分比)	3059 (36.9%)	2061 (24.8%)	3175 (38.3%)	8295 (100.0%)

根據上表4-14交叉分析結果得知，具有高現有價值以及高潛在價值共有253位，佔總數的3.1%，此區顧客不但過去對於企業的貢獻是最大，而且預期未來的貢獻也是最高，屬於最值得開發的顧客群，因此命名為「黃金顧客群」。在高現有價值以及中潛在價值顧客方面，共有28位顧客，佔總數的0.3%，此顧客雖具有高現有價值，但潛力方面為普通，因此命名為「一般忠誠，高價值群」。高現有價值，而低潛力的顧客方面，共佔22位，0.3%，此顧客群雖有高現有價值，但未來的價值潛力卻不高，亦即這類顧客過去常光顧，但或許某些因素有可能在未來的購買潛力下降，因此這類顧客應要多加注意，儘可能的找回這類顧客的認同，所以命名為「須挽回顧客群」。在中現有價值以及高潛在價值方面，共有756位顧客，佔總顧客數9.1%，這類顧客過去貢獻的價值普通，但具有潛力，因此相當值得開發，命名為「高忠誠，一般價值群」。在中現有價值以及中潛在價值顧客方面，共215位，佔2.6%，此類顧客命名為「一般忠誠及價值群」。在中現有價值以及低顧客低潛在價值群方面，共有369位，佔4.4%，這類顧客群命名為「低忠誠，一般價值群」。在低現有價值以及高潛在價值方面，這類顧客雖然現有價值低，但未來很有可能成為忠誠顧客，因此命名為「值得開發顧客群」。在低現有價值與中顧客潛在價值群方面，共有1818位，佔總顧客數21.9%，此顧客群命名為「一般忠誠，低價值群」。在低現有價值以及低潛在價值方面，共有2668位，佔顧客總數36.9%，這類顧客不但過去貢獻低，而且未來也不具潛力性，因此這類顧客命名為「游離顧客群」。本研究將命結果彙整如下圖4-1所示：

顧客現有價值	高現有價值顧客	須挽回顧客群(7)	一般忠誠，高價值群(8)	黃金顧客群(9)
	中現有價值顧客	低忠誠，一般價值群(4)	一般忠誠及價值群(5)	高忠誠，一般價值群(6)
	低現有價值顧客	危險游離顧客群(1)	一般忠誠，低價值群(2)	值得開發顧客群(3)

低潛在價值顧客 中潛在價值顧客 高潛在價值顧客

顧客潛在價值

圖 4-1 顧客價值矩陣

4.3.2 顧客現有價值與顧客潛在價值獲利能力交叉分析

顧客現有價值與顧客潛在價值獲利能力交叉分析結果，彙整如下表4-15及圖4-2所示：

表 4-15 顧客現有價值與顧客潛在價值獲利能力交叉分析

	低潛在價值顧客	中潛在價值顧客	高潛在價值顧客	總計
高現有價值顧客 (百分比)	1253475 (1.85%)	1668890 (2.47%)	25958852 (38.40%)	28881217 (42.72%)
中現有價值顧客 (百分比)	5539301 (8.19%)	4401317 (6.51%)	17388459 (25.72%)	27329077 (40.42%)
低現有價值顧客 (百分比)	3475463 (5.14%)	2730436 (4.04%)	5191629 (7.68%)	11397528 (16.86%)
總計 (百分比)	10268238 (15.19%)	8800643 (13.02%)	48538941 (71.79%)	67607822 (100.00%)

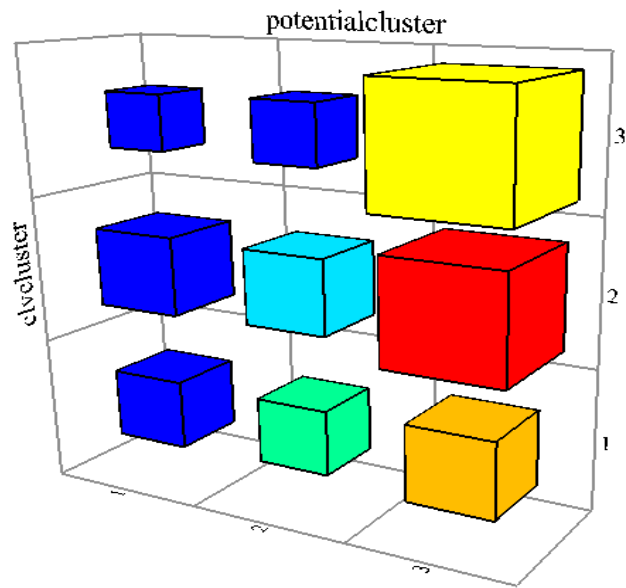


圖 4-2 顧客價值與潛力獲利能力交叉圖

由上表4-15以及上圖4-2可知，黃金顧客群的獲利能力最高，佔總獲利38.4%，而總顧客數僅有253位。其次為高忠誠，一般價值顧客，佔總獲利25.72%，總顧客數達756位，接下來依續為低忠誠，一般價值顧客，佔總獲利8.19%，總顧客數為369位；須開發顧客群，佔總獲利7.68%，總顧客數為2166位，一般忠誠與顧客群，佔總獲利6.51%，總顧客數為215位；游離顧客群，佔總獲利5.14%，總顧客數為2668位(總顧客數最高)；一般忠誠，低價值顧客，佔總獲利4.04%，總顧客數為1818位；一般忠誠，高價值顧客，佔總獲利2.47%，總顧客數為28位(顧客總次數少)；須挽回顧客，佔總獲利1.85%，共22位顧客(總顧客數最少)。

4.3.3 顧客現有價值與顧客潛在價值交叉累積分析-20/80 法則驗證

本研究將各顧客群的價值與顧客數累積起來加以驗證20/80法則(80%的獲利來自前20%的顧客)，將表4-14的總顧客數加總以及表4-15的顧客價值加總起來後，彙整如下表4-16所示：

表 4-16 各集群顧客價值與顧客數累積

	游離顧客群	一般忠誠低價值群	須開發顧客群	低忠誠，一般價值群	一般忠誠及價值群	高忠誠，一般價值群	須挽回價值群	一般忠誠高價值群	黃金顧客群
顧客價值 (累積百分比)	5.14% (5.14%)	4.04% (9.18%)	7.68% (16.86%)	8.19% (25.05%)	6.51% (31.56%)	25.72% (31.56%)	1.85% (59.14%)	2.47% (61.60%)	38.40% (100.00%)
顧客數 (累積百分比)	32.20% (32.20%)	21.90% (54.10%)	26.10% (80.20%)	4.40% (84.60%)	2.60% (87.20%)	9.10% (96.30%)	0.30% (96.60%)	0.30% (96.90%)	3.10% (100.00%)

經由上表彙整的結果，繪出如下圖 4-3 所示，由下圖 4-3 可知，前 3 群顧客 (游離顧客群、一般忠誠，低價值群、須開發顧客群) 為低價值，因此對於企業的貢獻度較低，而此三群的顧客佔所有顧客的總數達 80.2%(6652 位)，而獲利貢獻度卻只有佔總獲利 16.8%(11,397,528 元)，其餘 6 個顧客群，僅佔總顧客數 19.8%(1643 位顧客)，但卻佔總獲利 83.2%(56,210,294 元)，這樣的結果也證實了 20/80 法則的現象。而本研究除了找出前 20% 的顧客，更去深入瞭解此 20% 的顧客中，有哪些是關鍵顧客群(黃金顧客群)，哪些是必須要去挽回的(須挽回顧客群)，此外，另外的後 80% 的顧客群中，有哪一些仍是值得去開發的顧客(值得開發顧客群)。

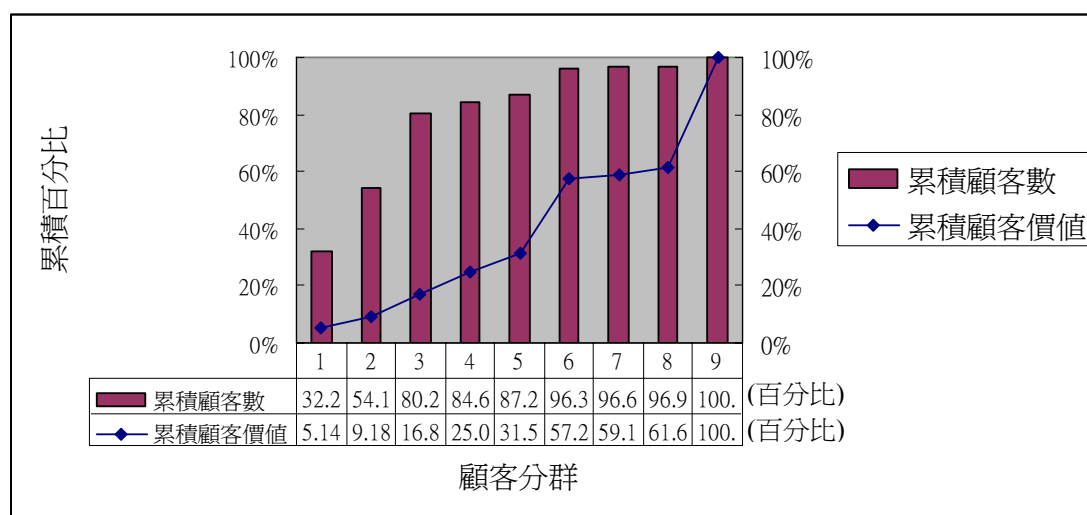


圖 4-3 20/80 法則驗證圖

4.4 預測結果驗證

4.4.1 混淆矩陣建立

本研究將訓練集分析後的結果來預測是否在驗證集上真能夠有效的預測到高價值及潛力顧客，透過混淆矩陣(confusion matrix)來加以驗證是否預測準確。本研究挑選在訓練集中出現以及驗證集中出現的顧客來做比較，結果共有 6141 位在訓練集中出現且在驗證集中出現。針對驗證集來計算每位顧客在驗證集中的顧客終身價值，計算結果再予以分群，根據驗證集的分群以及訓練集的分群結果來找出有多少比例是正確預測的，分析結果，驗證集中共有 1803 位為高價值，2059 位顧客中價值，2279 位顧客為低價值，其與訓練集中的潛力分群做比較，歸納如下表 4-17 所示，其中訓練集中的低潛力亦指集群 1、4、7，訓練集的中潛力為集群 2、5、8，訓練集中的高潛力為 3、6、9 群，加總後來與驗證集做比對。

表 4-17 訓練集與驗證集比對分析

	訓練集-低潛力	訓練集-中潛力	訓練集-高潛力	驗證集總計
驗證集-高價值	106	220	2234	2560
驗證集-中價值	79	1321	355	1755
驗證集-低價值	1467	156	203	1826
訓練集總計	1652	1697	2792	6141

4.4.2 預測結果分析

(1) 準確率(Accuracy)計算如下：

$$= \frac{2234 + 1321 + 1467}{6141} = 81.77\%$$

(2) 真實正確率(True Positive rate)計算結果如下：

$$\text{高價值顧客：} \frac{2234}{106 + 220 + 2234} = 87.26\%$$

$$\text{中價值顧客：} \frac{1321}{79 + 1321 + 355} = 75.27\%$$

$$\text{低價值顧客：} \frac{1467}{1467 + 156 + 203} = 80.34\%$$

(3)精確性(Precision)計算結果如下：

$$\text{高價值顧客} : \frac{2234}{2234 + 355 + 203} = 80.01\%$$

$$\text{中價值顧客} : \frac{1321}{220 + 1321 + 156} = 77.84\%$$

$$\text{低價值顧客} : \frac{1467}{1467 + 106 + 79} = 88.80\%$$

由以上的分析結果得知，本研究所預測的結果與實際情形大致相符，準確率達 80%以上，真實正確率亦皆達 75%以上，精確性也皆達 75%以上，因此可信。

4.5 資料探勘

資料探勘分析上，本研究將利用 OLAP 分析、關聯法則以及決策樹來分析，藉以瞭解顧客的特徵、潛在行為以及預測顧客行為。



4.5.1 OLAP 分析

在OLAP分析方面，本研究針對所有的顧客群來做多維度的描述特徵，包括年紀、所得、職業、促銷活動參與等，此分析主要是利用多構面表來觀察，並輔以圖形來衡量，如圖4-4與4-5所示：

	A	B	C	D	E	F	G	H	I	J	K	
1	Yearly Income	All Yearly Incom										
2	mvpage	所有的 mvpage										
3	Marital Status	All Marital Status										
4	Gender	All Gender										
5	Education Level	All Education Le										
6												
7	Store Sales	Cluster										
8	Media Type		1	2	3	4	5	6	7	8	9	總計 *
9	Bulk Mail		1661.7	2150.9	837.8	280.7	515.9	186.03	57.64	1052	6742.6	
10	Cash Register Handout		1479.3	2115.2	1465	1181.9	2358	345.7	463.98	2171.5	11580.86	
11	Daily Paper		1529.9	2218.9	1734	1474.7	3033	360.25	399.13	3089.6	13839.54	
12	Daily Paper, Radio		1046.5	1678.6	2796	1645.2	4341	101.76	215.99	1940.3	13764.66	
13	Daily Paper, Radio, TV		2861.6	3291.2	2660	1111.4	2569	87.18	92.65	2982.2	15655.82	
14	In-Store Coupon		757.93	601.81	1232	621.09	1340	72.69	198.96	2764.6	7589.06	
15	No Media	102466.33	46683	65447	68404	38485	98158	7776.1	10016	68716	506152.5	
16	Product Attachment		1783	2295.3	2629	1335.3	3457	48.28	189.56	791.1	12529.19	
17	Radio		760.64	1057.9	601.3	246.11	354.2	85.18	45.94	103.97	3255.3	
18	Street Handout		1157.6	2196.6	1058	383.9	1540	21.86	197.32	3033.2	9588.68	
19	Sunday Paper		762.08	1279.7	1223	603.85	2278	105.59	424.98	1099.3	7776.4	
20	Sunday Paper, Radio		738.77	1214.7	2110	1059.6	3101	131.59	280.63	2815	11452.1	
21	Sunday Paper, Radio, TV		415.2	444.04	609.7	195.89	726.1	211.18	34.2	2309.9	4946.11	
22	TV		1571.6	2127.5	605.1	306.81	453.9				5064.9	

圖 4-4 OLAP 多構面交叉圖

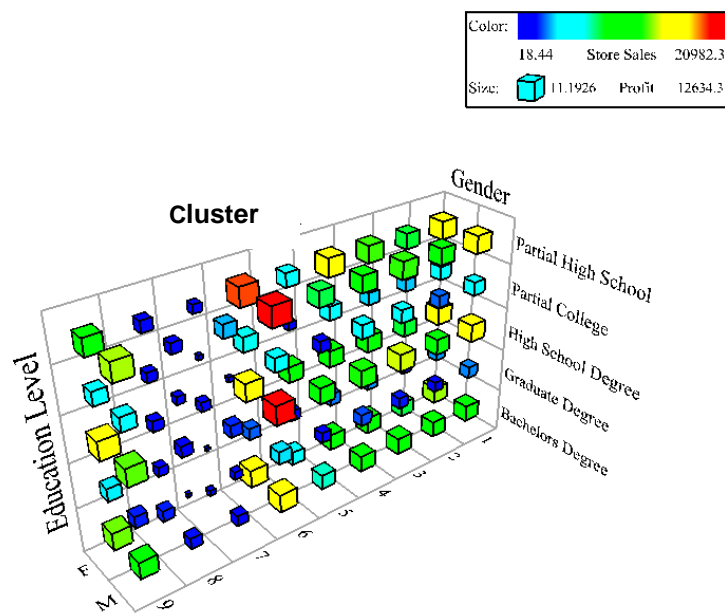


圖 4-5 OLAP 立方體呈現

註：圖中的 cluster 為集群的代號，1 表游離顧客群(最小)，9 表黃金顧客群(最大)

本研究以列出三個構面的資料表及立方體圖來呈現及瞭解各集群的特徵結果。由於 OLAP 分析表所呈現資料繁多且大，因此列於附件 1，在此僅針對立方體圖及明顯差異之處做說明，彙整如下表 4-18 所示。

表 4-18 OLAP 分析結果表

屬性	立方體圖	各集群在各構面明顯差異之處
<p>年紀 性別 集群</p>		<p>游離顧客群與值得開發顧客群男女比例差不多，以40-49人數最多；集群2及6在50-59歲的男性比女性價值高，其餘比例皆差不多；高忠誠、一般價值群，在15-19歲的男性價值比女性高，20-29歲則相反；一般忠誠與價值群在15-19、20-29以及80-89歲的男性價值比女性高；須挽回價值群在所有年齡層中皆以男性價值較高，其中在15-19以及20-29的年齡層是沒有此集群的人。一般忠誠、高價值群的人沒有是15-19歲，20-29及50-59歲的男性價值大於女性，40-49、60-69及80-89歲的女性價值大於男性。黃金顧客群在30-39、60-69及80-89歲的女性價值大於男性，而70-79歲的男性價值大於女性。</p>
<p>所得 是否擁有房子 集群</p>		<p>整體來說，擁有房子的人所貢獻的價值比較高，而所得以30-50K的貢獻最多，150K+貢獻是最少的。集群7在110K-130K、150k+及90K-110K。一般忠誠、高價值群沒有人的所得在150K+，在110K-130K及90K-110K則一定為有房子，在130-150K則一定為沒有房子。黃金顧客群在150K+的等級，一定有房子，30-50K的顧客貢獻最多。</p>

<p>教育程度 婚姻狀況 集群</p>		<p>大致上，已婚與未婚的比例在各集群中是差不多的，學歷比例最多的為Partial High School，而Graduate Degree的比例是最少的。須挽回價值群在研究所等級的都是已婚的，而一般忠誠、高價值群在研究所等級的都是未婚的，而一般忠誠、高價值群的Bachelors Degree是最多的，一般忠誠與價值群的High School Degree是最多的。</p>
<p>擁有小孩人數 居住國家 集群</p>		<p>加拿大的顧客貢獻度都較低，全部集群在游離顧客群、一般忠誠、價值群群、值得開發顧客群。而墨西哥雖然亦集中在游離顧客群、一般忠誠、價值群群、值得開發顧客群，但有部份的顧客都屬於黃金顧客群。整體來說，大部份的顧客都居住在美國，高價值的顧客也大都在美國，其中以擁有4位小孩的顧客群價值較高，其次為3位小孩。0位小孩的貢獻價值最低，其次低的為1位小孩。</p>
<p>職業 會員卡類別 集群</p>		<p>職業以Professional所貢獻價值最高，而大部份的顧客會員卡為Bronze。其中職業為Clerical佔集群須挽回價值群、一般忠誠、高價值群、黃金顧客群的比例低。黃金顧客群大部份的顧客職業為Professional為主。</p>
<p>促銷媒體 集群</p>		<p>游離顧客群都沒有接受過媒體促銷的活動，而TV的促銷在須挽回顧客群以及黃金顧客群皆沒有達到效果，此外，在值得開發顧客群方面，顧客現有價值最高前三個依續為Product Attachment, Daily Paper以及Street Handout；一般忠誠及價值群方面，顧客現有價值最高前三個依</p>

		<p>續為Daily Paper, Radio, Daily Paper, 以及Product Attachment;須挽回價值群方面，顧客現有價值最高前三個依續為Daily Paper, Cash Register Handout, 以及Sunday Paper, Radio, TV；黃金顧客群方面，顧客現有價值最高前三個主要接觸的媒體依續為Street Handout, Daily Paper 以及 Daily Paper, Radio, TV。</p>
--	--	---

4.5.2 顧客潛在價值購買行為

由於本研究針對黃金顧客群來做顧客潛力的深入探討，利用關聯法則來找出各集群的潛在購買行為樣式，此外，本研究亦針對全部的交易資料庫做關聯分析，比較整體交易資料庫與黃金顧客群的交易資料在關聯分析上的差異。在設定支持度(support)以及信賴度(confidence)方面，由於交易量並不一致，因此須設定不同的支持度及信賴度值。

(1) 整體交易資料庫之潛在購買行為

由於交易資料繁多，共 154,515 筆，在設定支持度時，本研究設定為 0.8%，信賴度設為 70%，分析語法如下圖 4-6 所示，而結果如表 4-19 所示，共得到 5 筆關聯規則。

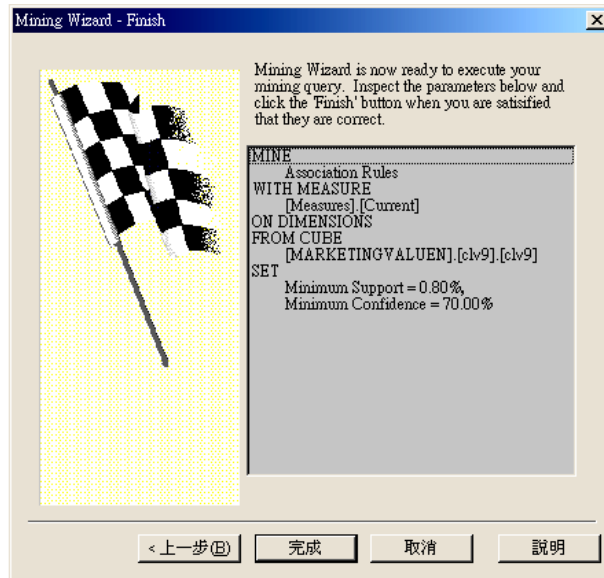


圖 4-6 整體交易資料關聯法則語法

表 4-19 整體交易資料庫之潛在購買行為

Head	Implies	Body	Support	Confidence
prodcut = [505] AND prodcut = [524]	==>	prodcut = [502]	0.849%	83.333%
prodcut = [1075] AND prodcut = [1343]	==>	prodcut = [417]	0.849%	83.333%
prodcut = [1075] AND prodcut = [417]	==>	prodcut = [1343]	0.849%	71.429%
prodcut = [1088] AND prodcut = [1138]	==>	prodcut = [789]	0.849%	83.333%
prodcut = [1164] AND prodcut = [1524]	==>	prodcut = [958]	1.019%	85.714%

(2) 黃金顧客群潛在購買行為

黃金顧客群的交易筆數共有 30,029 筆，support 值設定為 1.5%，confidence 設定為 80%，如下圖所示：

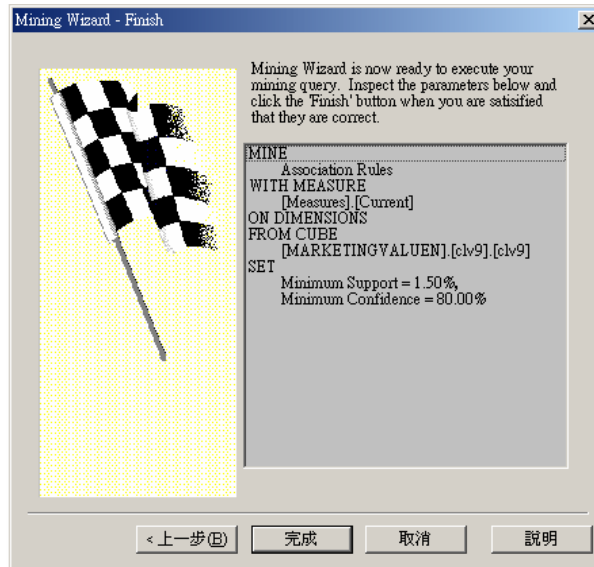


圖 4-6 黃金顧客群關聯語法

分析結果共得到 11 筆關聯資料，如下表 4-20 所示：

表 4-20 黃金顧客群潛在購買行為 單位(百分比)

Head	Implies	Body	Support	Confidence
product = [208] AND product = [695]	==>	product = [811] AND product = [958]	1.581	100
product = [1454] AND product = [215]	==>	product = [21] AND product = [436]	1.581	100
product = [1313] AND product = [177]	==>	product = [1558] AND product = [648]	1.581	100
product = [1208] AND product = [415]	==>	product = [454] AND product = [672]	1.581	100
product = [415] AND product = [454]	==>	product = [1208] AND product = [672]	1.581	100
product = [1153] AND product = [1189]	==>	product = [1196] AND product = [1504]	1.581	100
product = [1454] AND product = [21]	==>	product = [215] AND product = [436]	1.581	80
product = [1313] AND product = [1558]	==>	product = [177] AND product = [648]	1.581	80
product = [1023] AND product = [1251]	==>	product = [519] AND product = [746]	1.581	80
product = [1153] AND product = [1196]	==>	product = [1189] AND product = [1504]	1.581	80
product = [119] AND product = [284]	==>	product = [528] AND product = [611]	1.581	80

由上表 4-19 與 4-20 的結果得知，總體資料庫的關聯法則與黃金顧客群的關聯法則不同，則表示若採用總體資料庫的關聯法則來發展行銷策略，並無法真正

的反應出黃金顧客群的特性，因此本研究結合終身價值分析與資料探勘的技術之好處亦即在此。

4.5.3 顧客潛在價值之人口統計因素發現-決策樹

為瞭解顧客的潛在價值之人口統計變數上的情形，利用決策樹的方法來找出其特徵，本研究針對黃金顧客群的特性來加以分析，藉以找出黃金顧客群的決策樹，分析的構面包括：年紀、性別、婚姻狀態、年所得、教育程度、職業等，分析結果如下圖4-7所示：

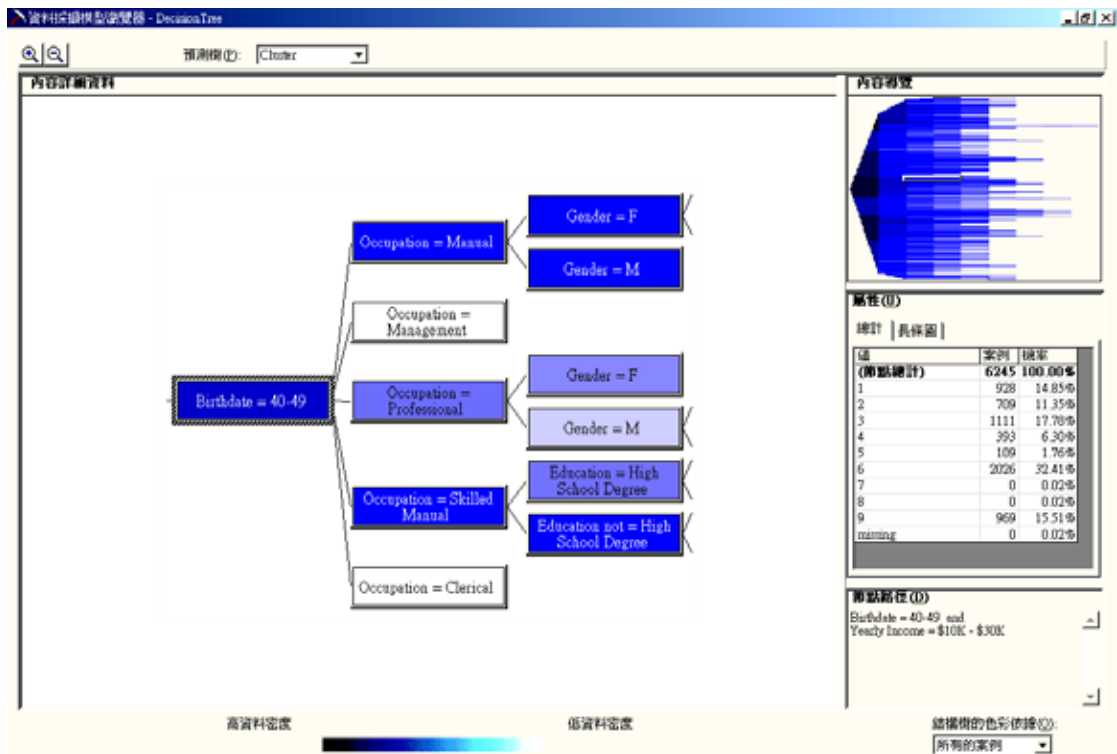


圖 4-7 顧客價值發現之決策樹

註：由於圖形過大，因此僅呈現部份的決策樹

本研究將黃金顧客群發生的可能性較高的規則找出來，選取百分比 70%以上的規則，依百分比由高至低排序，分析結果共得到 28 筆決策規則，如下表 4-21 所示：

表 4-21 顧客潛在價值之決策樹

編號	規則		百分比
	IF	THEN	
1	Age = 70-79 and Yearly Income = \$70,000 - \$90,000 and Education = Partial College and Occupation = Management	黃金顧客群	93.18%
2	Age = 40-49 and Yearly Income = \$130,000 - \$150,000 and Education = High School Degree and Occupation = Management and Gender = F	黃金顧客群	90.48%
3	Age = 30-39 and Yearly Income = \$70,000 - \$90,000 and Education = Partial High School and Occupation = Management and Gender = M	黃金顧客群	89.93%
4	Age = 80-89 and Yearly Income = \$130,000 - \$150,000 and Education = Graduate Degree and Occupation = Management and Gender = F	黃金顧客群	89.29%
5	Age = 50-59 and Yearly Income = \$90,000 - \$110,000 and Education = High School Degree and Occupation not = Management and Gender = F	黃金顧客群	88.06%
6	Age = 60-69 and Yearly Income = \$150,000 + and Education = High School Degree and Occupation = Professional and Gender = F	黃金顧客群	87.31%
7	Age = 80-89 and Yearly Income = \$90,000 - \$110,000 and Education = Graduate Degree and Gender = F	黃金顧客群	86.89%
8	Age = 60-69 and Yearly Income = \$70,000 - \$90,000 and Education = Partial High School and Occupation not = Management and Gender = F	黃金顧客群	85.93%
9	Age = 40-49 and Yearly Income = \$30,000 - \$50,000 and Education = Bachelors Degree and Occupation = Professional and Gender = F	黃金顧客群	85.10%
10	Age = 30-39 and Yearly Income = \$130,000 - \$150,000 and Education = Partial High School and Gender = M	黃金顧客群	84.79%
11	Age = 20-29 and Yearly Income = \$30,000 - \$50,000 and Education = Graduate Degree	黃金顧客群	84.55%
12	Age = 60-69 and Yearly Income = \$10,000 - \$30,000 and Education = High School Degree and Occupation not = S,000illed Manual and Gender = F	黃金顧客群	83.78%
13	Age = 50-59 and Yearly Income = \$130,000 - \$150,000 and Education = Bachelors Degree and Occupation not = Management and Gender = M	黃金顧客群	82.76%
14	Age = 40-49 and Yearly Income = \$30,000 - \$50,000 and Education = Bachelors Degree and Occupation = Management and Gender = F	黃金顧客群	81.77%
15	Age = 40-49 and Yearly Income = \$130,000 - \$150,000 and Education = Bachelors Degree and Gender = M	黃金顧客群	81.29%
16	Age = 20-29 and Yearly Income = \$10,000 - \$30,000 and Occupation = Management and Gender = M	黃金顧客群	80.47%
17	Age = 20-29 and Yearly Income = \$90,000 - \$110,000 and Education = Partial High School and Occupation =	黃金顧客群	79.93%

	Professional		
18	Age = 20-29 or = 30-39 or = 15-19 or = 80-89 and Yearly Income = \$110,000 - \$130,000 and Education = Bachelors Degree and Occupation not = Management and Gender = M	黃金顧客群	78.04%
19	Age = 15-19 and Yearly Income = \$130,000 - \$150,000 and Education = Partial High School	黃金顧客群	77.59%
20	Age = 30-39 and Yearly Income = \$30,000 - \$50,000 and Education = Bachelors Degree and Occupation = Management	黃金顧客群	77.04%
21	Age = 20-29 and Yearly Income = \$130,000 - \$150,000 and Education = Partial High School and Occupation not = Management and Gender = M	黃金顧客群	76.42%
22	Age = 80-89 and Yearly Income = \$130,000 - \$150,000 and Education = High School Degree and Gender = F	黃金顧客群	76.02%
23	Age = 70-79 and Yearly Income = \$90,000 - \$110,000 and Education = Graduate Degree and Gender = F	黃金顧客群	75.90%
24	Age = 20-29 and Yearly Income = \$150,000 + and Education = Bachelors Degree and Occupation = Management	黃金顧客群	73.88%
25	Age = 80-89 and Yearly Income = \$130,000 - \$150,000 and Education = Partial College	黃金顧客群	73.71%
26	Age = 20-29 and Yearly Income = \$70,000 - \$90,000 and Education = Graduate Degree and Gender = F	黃金顧客群	72.01%
27	Age = 60-69 and Yearly Income = \$110,000 - \$130,000 and Education = Partial High School and Gender = F	黃金顧客群	71.52%
28	Age = 50-59 and Yearly Income = \$50,000 - \$70,000 and Occupation = S,000illed Manual and Gender = F	黃金顧客群	70.26%

根據決策樹分析結果得知黃金顧客群的特性，此特性可以做為未來新進顧客的判斷規則，藉以能快速瞭解新進顧客的特性。

第五章 結論

5.1 研究結論與發現

本研究以發掘顧客現有價值以及潛在價值為目標，運用顧客終身價值分析以及資料探勘技術來加以建立出完整的潛在價值及行為分析模式。首先，本研究先確認資料探勘目標，在前處理完成之後，導入顧客終身價值計算，計算出顧客現有價值以及潛在價值，現有價值計算部份是利用一般顧客終身價值計算方式，而潛力顧客計算則是採用 SMC 模型以及再交易金額模型來找出預期購買金額以及預期購買可能性(活動力)，分析結果透過集群分析，以價值為基礎來分群，找出高、中、低的現有以及潛在顧客各 3 群，利用現有價值以及潛在價值兩構面建立出 3×3 的顧客價值策略矩陣圖，找出真正具有貢獻度的顧客群，此 9 個集群分別有黃金顧客群、一般忠誠，高價值群、須挽回價值群、高忠誠，一般價值群、一般忠誠及價值群，低忠誠，一般價值群、值得開發顧客群，一般忠誠，低價值群以及危險游離顧客群，如圖 4-1 所示，此外，本研究根據分群結果，深入探討 20/80 法則之特性，發現在前 20% 的顧客群包括黃金顧客群、一般忠誠，高價值群、須挽回價值群、高忠誠，一般價值群、一般忠誠及價值群，低忠誠，一般價值群，其中有 3.1% 的顧客是屬於高潛在價值以及高現有價值的「黃金顧客群」，這類顧客對企業的貢獻度最大，而且最忠誠，必須加以瞭解此顧客的特徵，本研究也透過資料探勘技術來深入瞭解此顧客群的潛在行為。此外在前 20% 的顧客中裡頭還有一些須注意的顧客群，這些群雖然目前提供的價值是高的，但未來的潛力以及忠誠度可能會不高，因此必須加以注意，分別為「須挽回價值群」以及「低忠誠，一般價值群」，這兩族群須使用能滿足此顧客的行銷策略來找回，避免流失。因此，本研究不但驗證 20/80 法則，而且更進一步去瞭解前 20% 的顧客特性，對於顧客關係管理來說是很重要的，因為若運用相同的行銷策略手法，反而無法很有效的管理前 20% 顧客群，瞭解哪些是最重要，哪些是必須挽回的，如此有助於行銷資源配置及策略應用。

在過去的顧客關係管理的概念中也提及，後 80% 的顧客貢獻度是低的，因此

在配置行銷資源時，應不要太過於投注，以免造成資源浪費，這樣的思維是假設在僅瞭解現有價值的情況，然而，本研究發現，在後 80% 的顧客當中也未必非全然不值得開發的，仍有必須加以留意的族群，亦即為「值得開發顧客群」，雖然目前這群顧客的價值低，但若有效的開發，這群顧客的未來潛力大，有可能會對公司未來的貢獻會很大，因此是值得去重視此類的顧客。

本研究接下來針對黃金顧客群，利用資料探勘的技術，包括關聯法則、OLAP 分析以及決策樹來找出潛在購買行為以及預測模型，瞭解顧客的潛在特性以及隱藏在顧客背後的行為樣式(Patterns)，找出可能購買的產品關聯性，並且將黃金顧客群與整體資料庫的產品關聯性分析相比較發現，產品關聯性是有差異的。而 OLAP 分析後則發現每個族群的在不同構面下的各個屬性是有差異的，決策樹則找出黃金顧客群的預測規則，瞭解黃金顧客群的特徵輪廓。彙整本研究的發現與結論，如下所示：

- 1、結合資料探勘技術與顧客終身價值分析，發掘顧客價值
- 2、找出顧客現有價值以及潛在價值
- 3、建立出策略性的顧客價值矩陣
- 4、發掘前 20% 利潤的顧客集群的策略性應用，哪些是黃金顧客，哪些是須挽回價值顧客
- 5、找出後 80% 利潤的顧客，瞭解其特徵以及有哪些顧客群是值得投資，哪些顧客群是要避免資源過多投入。
- 6、確認出顧客現有值以及潛力與顧客購買之相關性，潛在價值與再交易機率有關，而再交易機率與最近購買有關。現有價值與購買頻率和獲利有關，而獲利與購買頻率也有高度相關。
- 7、挖掘出黃金顧客群的潛在購買行為，瞭解最常購買的產品以及關聯性為何。
- 8、建立出 OLAP 以及 Cube，以找出各集群在不同人口統計變數構面上的差異。
- 9、建立預測黃金顧客群的決策樹，瞭解黃金顧客群的決策規則。
- 10、提出以顧客價值金字塔，並提供顧客關係管理上的策略性應用及意涵。

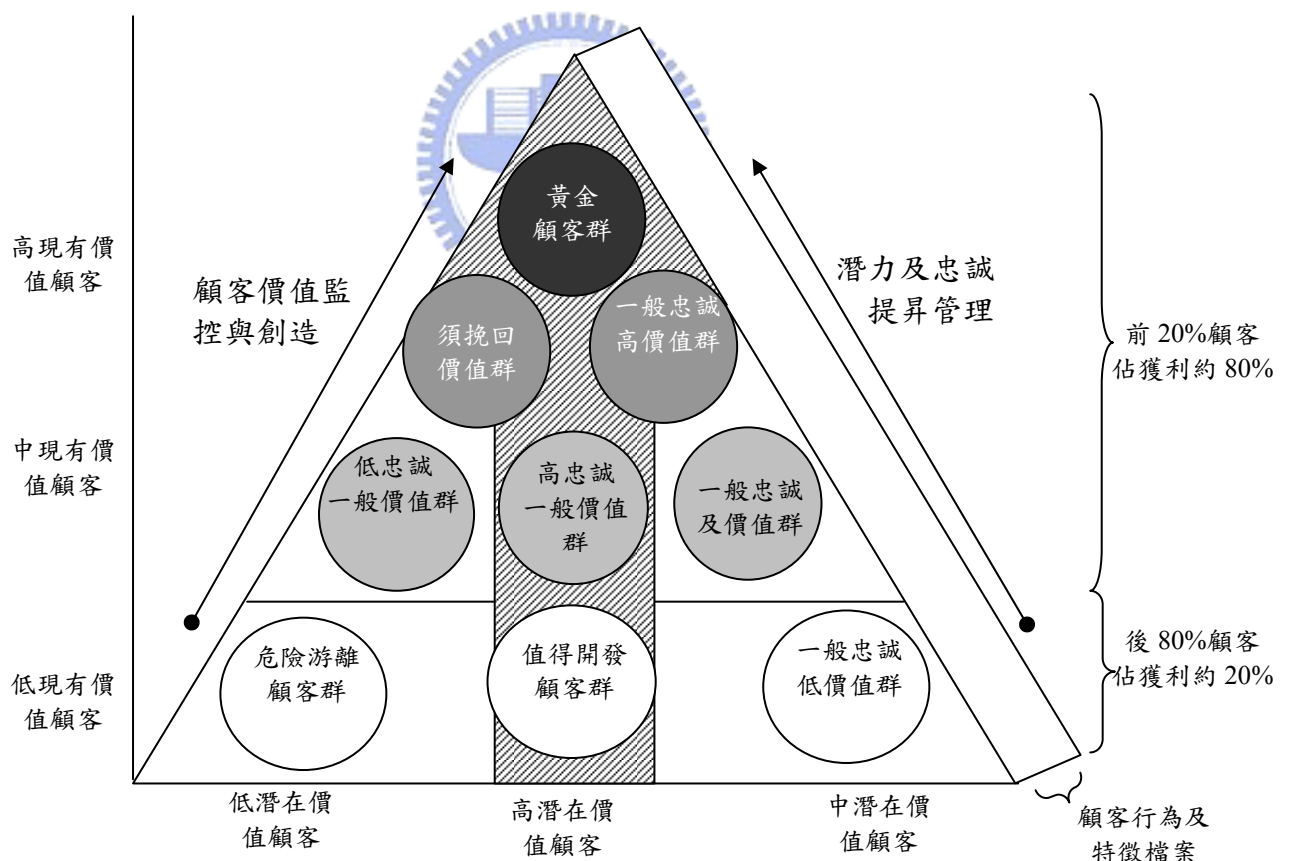
整體來說，顧客價值發現了以下的特徵，彙整如下表 5-1 所示：

表 5-1 黃金顧客價值發現彙整

分析方法		黃金顧客群潛在價值發現		
顧客終身價值分析	現有價值	總購買次數越多以及獲利越多，則現有價值越高。		
	潛在價值	最近有購買及再交易機率則潛在價值越高。		
資料探勘技術	OLAP	<ul style="list-style-type: none"> ● 30-39、60-69 及 80-89 歲的女性價值大於男性，而 70-79 歲的男性價值大於女性。 ● 所得在 150K+ 的等級，一定會有房子，30-50K 的顧客貢獻最多。 ● 已婚與未婚的比例是差不多的，學歷比例最多的為 High School，而 Graduate Degree 學歷最少。 ● 大部份的顧客都居住在美國，擁有 4 位小孩的顧客群價值較高，其次為 3 位小孩。0 位小孩的貢獻價值最低，其次低的為 1 位小孩。 ● 職業大部份為 Professional，以 Bronze 的卡別為主。 ● 前三個主要接觸的媒體依續為 Street Handout, Daily Paper 以及 Daily Paper, Radio, TV。 		
	關聯法則	Head	Implies	Body
		product = [208] AND product = [695]	==>	product = [811] AND product = [958]
product = [1454] AND product = [215]		==>	product = [21] AND product = [436]	
product = [1313] AND product = [177]		==>	product = [1558] AND product = [648]	
product = [1208] AND product = [415]		==>	product = [454] AND product = [672]	
product = [415] AND product = [454]		==>	product = [1208] AND product = [672]	
product = [1153] AND product = [1189]		==>	product = [1196] AND product = [1504]	
決策樹	Age = 70-79 and Yearly Income = \$70,000 - \$90,000 and Education = Partial College and Occupation = Management			
	Age = 40-49 and Yearly Income = \$130,000 - \$150,000 and Education = High School Degree and Occupation = Management and Gender = F			
	Age = 30-39 and Yearly Income = \$70,000 - \$90,000 and Education = Partial High School and Occupation = Management and Gender = M			
	Age = 80-89 and Yearly Income = \$130,000 - \$150,000 and Education = Graduate Degree and Occupation = Management and Gender = F			
	Age = 50-59 and Yearly Income = \$90,000 - \$110,000 and Education = High School Degree and Occupation not = Management and Gender = F			

5.2 顧客關係管理策略應用

根據分析後的集群特性，本研究點出不同集群的特性以及在顧客關係管理上可行的策略應用，繪出如下圖 5-1，本研究所分析出的 9 個集群各有不同的價值程度，以現有價值以及潛在價值為構面，建立出一個金字塔的顧客價值結構，最底層至最高層的顏色由淺入深，代表價值的貢獻度，最底層集群皆為顧客現有價值低的族群，此族群人數雖多(約佔總顧客數 80%)，但對於企業獲利貢獻並不多(約佔總獲利 20%)，因此顏色皆較淺，第二層的價值性次之，最頂層價值性最高。而中間軸(在三角型裡有個鍵頭朝上)稱為顧客關係管理核心軸，所接觸到的顧客群是公司運作的軸心，越往上層越為重要，這些顧客皆是具有潛力，值得去多關心的顧客，而位置越偏離軸心(左右兩邊)的顧客群，則企業並不須要花過多心力在身上，以免浪費資源。



資料來源：本研究整理

位於軸心區域的顧客群有：黃金顧客群、高忠誠一般價值群、一般忠誠及高價值群、須挽回價值群、值得開發顧客群，這些集群須透過資料探勘的模式來更深入的探討其中的潛在特徵。其中必須要針對黃金顧客群建立出最完善的行銷方案，多方位的服務，以顧客關係維持為目標，滿足該族群顧客。須挽回價值群則必須要進一步去深入訪談，盡全力的找出該族群可能離開的原因，並加以改善，以挽回該群顧客的心。值得開發顧客群則是要加以去關心，不可忽略掉此族群對於未來的可能貢獻，建立行銷方案來提供給些族群，如此對於企業的未來獲利有正面幫助。而高忠誠一般價值群以及一般忠誠高價值群則除了維持關係之外，還要思考如何提昇關係的層次。

位於偏離軸心的顧客方面有：低忠誠，一般價值群、危險游離顧客、一般忠誠及價值群以及一般忠誠與低價值群，這些顧客群應儘量避免過多的投資，除了一般忠誠及價值群須儘可能的與目前的關係保持不變之外，其餘的則須減少資源的投資，以免造成浪費，尤其是危險游離顧客，此族群很多都只是來購買過一、兩次，因此在應用行銷方案時，應儘量避免投入到此顧客群中。總體來說，企業應不時的於對顧客群來加強顧客價值監控與創造以及潛力及忠誠提昇管理，以改善顧客關係管理的執行效益及效果。

5.3 本研究在學術上的貢獻

本研究結果，對於學術上有以下的貢獻：

- 1、將顧客終身價值計算概念導入至資料探勘流程中來找出顧客價值
- 2、本研究將顧客終身價值計算後的結果，考慮人口統計變數上的差異，瞭解其中的特性
- 3、提供 20/80 法則中更細部的策略性思考
- 4、建立起顧客價值發現的流程及模式，提供學術界參考
- 5、將顧客終身價值與顧客潛在價值皆計算出來，並且建立顧客價值策略矩陣以及顧客價值金字塔，提供不同層面思考，以提昇顧客關係管理的效益。

5.4 研究限制與未來研究方向

5.4.1 研究限制

本研究主要的限制歸納如下：

1、產業別的差異

本研究主要是針對連鎖超市量販店為主的資料庫，然而，不同的產業會有不同的現象，因此本研究結果會因不同產業而被限制。

2、未考量到其它顧客潛在價值的因素

由於資料庫上無提供其它潛力的因素，例如：顧客滿意度、推薦行為、品牌態度等，因此無法完整的考量所有顧客潛在因素。

3、顧客資料庫僅有兩年的資料

本研究的資料庫僅參考兩年資料，倘若有可長年限的資料考量，或許能更加完臻。



5.4.2 未來研究方向

未來可行的研究方向，本研究歸納如下：

1、可以以本研究所建立的為基礎，建立一套顧客關係管理系統

目前的顧客關係管理系統少有針對顧客獲利為基礎來分群，根據分群後結果來應用顧客關係管理策略，甚至目前沒有一套系統考量顧客潛在價值，因此可以以本研究所建立的為基礎來開發一套顧客關係管理系統。

2、比較以及採用其它的顧客潛在發現的模式

顧客潛在價值分析模式尚有其它種，例如：迴歸分析、類神經網路、probit models等，因此可以使用其它的預測模式來發展出更完臻的結果。

3、找出每位顧客的價格敏感度，建立出第三個構面

由於價格敏感度在顧客潛在價值中亦是重要的變數，是否是黃金顧客群的價格敏感度會很低，而危險游離顧客群的價格敏感度會很高，抑或是黃金客群中，有部份的價格敏感度高，這些問題的釐清對於顧客關係管理上的應用會有正面的幫助。

4、建立自動化更新的分析模式

由於本研究主要是探討模式的可行性，驗證結果是值得開發的，因此未來研究可以針對本研究的模式來提供自動更新、轉換、計算、視覺呈現的即時運作系統，以利管理者得到最新以及最快的資訊。

5、針對其它集群，來找出其潛在顧客行為

本研究僅針對黃金顧客群的潛在行為做深入探討，倘若亦針對其它族群，例如：值得開發顧客群、須挽回顧客群、游離顧客群等的行為做深入分析，應是另一個有趣的研究議題，尤其是值得開發顧客群，若能夠有效的得知此族群的行為，並且加以提供行銷方案，或許可以比較行銷方案執行前以及執行後的有效性，對於此價值發現的效益更能夠去衡量出來。

本研究主要是提出顧客終身價值分析以及資料探勘技術整合的概念，提出流程及架構，利用實證來加以驗證，未來此方面仍須學者來多加研究，以建構出完整的顧客價值發現模式，以提昇顧客關係管理上的效益及效果。

參考文獻

中文部份

1. 王堅斌，「使用分類結合法與決策樹判別忠誠顧客」，國立成功大學資訊管理研究所碩士論文，民國九十二年。
2. 沈清正、陳仕昇、高鴻斌、張元哲、陳家仁、黃琮盛、陳彥良，「資料間隱含關係的挖掘與展望」，資訊管理學報第九卷，專刊，75-100 頁，2002。
3. 林柏甫，「客戶關係管理應用於顧客保留及顧客成長以電子業為例」，國立成功大學高階管理碩士在職專班碩士論文，民國九十二年
4. 陳麗君，「應用資料探勘技術於信用卡黃金級客戶之顧客關係管理」，元智大學工業工程與管理研究所碩士論文，民國九十二年。
5. 鄭傑之，「資料庫行銷：以顧客獲利性為基礎配置行銷資源」，中正大學企業管理研究所碩士論文，民國九十年。
6. Jay C., Wil W., Guido T. & Conny Z. 顧客金塔的行銷秘密，陳琍玲 譯，美商麥格羅 希爾國際股份有限公司，台北，民國 87 年。

英文部份

1. Adrian, P. "A Strategic Framework For Customer Relationship Management", A BT CRM White Paper, pp.2, March, 2002.
2. Agrawal, R. & Srikant, R. "Fast Algorithms for Mining Association Rules", In Proc. of the 20th Int'l Conference on Very Large Databases, pp.487-499 from <http://citeseer.nj.nec.com/agrawal94fast.html> Access on December, 2004.
3. Anderson, E.W., Fornell, C. & Lehmann, D.R. "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden", Journal of Marketing, 58, pp.53-66, 1994.
4. Bauer, H. H., Hammerschmidt, M. & Braehler, M. "The Customer Lifetime Value Concept and Its Contribution To Corporate Valuation", Yearbook of Marketing and Consumer Research, pp.46-67, 2003.

5. Bayon, T., Gutsche J. & Bauer, H. "Customer Equity Marketing: Touching the Intangible", European Management Journal, 20, pp.213-222, 2002.
6. Berger, P. D. & Nasr, N. I. "Customer lifetime value: marketing models and applications", Journal of InteractiveMarketing, 12(1), pp.17–30, 1998.
7. Berry, M.J.A. & Linoff, G. Data Mining Techniques for Marketing, Sales, and Customer Support, Wiley, New York, 1997.
8. Bitran, G. R. & Mondschein, S. "Mailing decisions in the catalog sales industr", Management Science, 42(9), pp.1364–1381, 1996.
9. Blattberg, R. C. & Deighton, J. "Manage marketing by the customer equity test", Harvard Business Review, Jul–Aug, pp.136–144, 1996.
10. Blattberg, R.C., G. Getz, G., & Thmas, J.S. "Managing customer acquisition", Direct Marketing, 64(6), pp.41-54, 2001.
11. Brachman, R. & Anand, T. "The Process of Knowledge Discovery in Databases: A Human Centered Approach, in AKDDM", AAAI/MIT Press, pp.37-58, 1996.
12. Bult, J. R. & Wansbeek, T. "Optimal selection for direct mail", Marketing Science, 14(4), pp.378-381, 1995.
13. Chablo, E. "The importance of marketing data intelligence in delivering successful CRM", <http://www.crm-forum.com/crm—forum—white—papers/mdi/sld01.htm>, Access on Mar. 2005, 1999.
14. Christopher, M., Psyne, A. & Ballantyne, D. "Relationship Marketing", London: Butterworth Heinemann, 1991.
15. Chung, H.M. & Gray, P. "Guest Editors Special Section: Data Mining", Journal of Management Information Systems, Vol.16, pp.11-16, 1999.
16. Chye, K.H. & Grey, C.K.L "Data mining and customer relationship marketing in the banking industry", Singapore Management Review, 24(2), pp.1-27, 2002.
17. Conlon, G. "No Turning Back", Sales and Marketing Management, Vol.151, December, pp.50-55, 1999.
18. Conway, D.K. & Fitzpatrick, J.M. "The customer relationship revolution – A methodology for creating golden customers", CRM-Forum, 1999, http://www.crm-forum.com/crm_vp/crr/sld01.htm. Access on Feb. 2005.

19. Courtheoux, R. "Customer retention: how much to invest. Research and the Customer Lifecycle", New York, NY: DMA, 1995.
20. Davis, B. "Data mining transformed", Informationweek, 751, pp.86-88, 1999.
21. Doyle, P. "Value-Based Marketing", Journal of Strategic Marketing, Vol. 8, pp. pp.299-311, 2000.
22. Dwyer, F. R. "Customer lifetime profitability to support marketing decision making", Journal of Direct Marketing, 3, pp.8-15, 1989.
23. Engel, J.F., Blackwell, R.D. & Miniard, P.W. "Consumer Behavior", The Dryden Press, Forth Worth, 1995.
24. Fayyad, U.M., Piatetsky-Shapiro, G. & Smyth, P. "From data mining to knowledge discovery: an overview, in: U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (Eds).", Advances in Knowledge Discovery and Data Mining, MIT Press, Massachusetts, Chap.1, 1996.
25. Fong, A. C. M., Hui, S. C. & Jha, G. "Data Mining for Decision Support", IT Pro, IEEE, 4 (2), pp.9-17, 2002.
26. Frawley, W.J., Piatetsky-Shapiro, G. & C.J. Matheus. "Knowledge discovery in databases: an overview", AI Magazine, 13 (3), pp.57-70, 1992.
27. Goebel, M. & Gruenwald, L. "A Survey of Data Mining and Knowledge Discovery Software Tools", ACM SIGKDD Explorations, Vol.1, pp.1-20, 1999.
28. Grembergen, V. W. & Amelinckx, I. "Measuring and managing e-business projects through the balanced scorecard", System Sciences, pp.3348-3356, 2002.
29. Gupta, S., & Lehmann, D. R. "Customers as assets", Journal of Interactive Marketing, 17(1), pp.9-24, 2003.
30. Gurau, C., Ranchhod A. & Hackney R. A. "Customer-centric Strategic Planning: integrating CRM in online business systems", Journal Information Technology and Management, Vol 4, No 2/3, pp.199-214, 2003.
31. Han, J. & Kamber, M. "Data Mining: Concepts and Techniques", Hardcover Ed: Morgan Kaufmann, 2001.
32. Hansotia, B. J. & Wang, P. "Analytical Challenges in Customer Acquisition", Journal of Direct Marketing, Vol. 11 (2), pp.7-19, 1997.

33. Heskett, J.L., Sasser, W.E. & Schlesinger, L.A. "The Service Profit Chain", Free Press, New York, 1997.
34. Hoekstra, J.C. & Huizingh, K.R.E. "The lifetime value concept in customer based marketing", Journal of Market Focused Management, 3(3/4), 1999.
35. Hui, S.C. & Jha, G. "Data Mining for Customer Service Support", Information Management, 38, pp.1-13, 2000.
36. Hwang, H., Jung, T. & Suh, E. "An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry", Expert Systems with Applications, 26, pp.181-188, 2004.
37. Jackson, B. "Winning and Keeping Industrial Customers", Lexington, MA: Lexington Books, 1985
38. Jackson, D. R. "Strategic application of customer lifetime value in the direct marketing environment". Journal of Targeting Measurement and Analysis for Marketing, 3(1), pp.9-17, 1994.
39. Jain, D. & Singh, S. S. "Customer Lifetime Value Research in Marketing: A Review and Future Directions", Journal of Interactive Marketing, spring, 16, pp.34-46, 2002.
40. Jenkins, D. "Customer relationship management", Bank Systems and Technology, 36(2), pp.30-34, 1999.
41. John, O. "Successfully development and Implementing Continuous relationship management", e-Business executive report, pp.26-30, 1999.
42. Jutla, D., Craig, J. & Bodorik, P. "Enabling and Measuring Electronic Customer Relationship Management Readiness," Proceedings of the 34th Hawaii International Conference on System Sciences, pp.1-10, 2001.
43. Kalakota, R. & Robinson, M. "e-Business: Roadmap for Success", Addison-Wesley, 1999.
44. Kalwani, M.U. & Narakesari, N. "LongTerm Manufacturer-Supplier Relationships: Do They Pay Off for Supplier Firms?", Journal of Marketing, Vol.59, Iss.1, pp.1-16, 1995.

45. Keane, T. & Wang, P. "Applications for the Lifetime Value Model in Modern Newspaper Publishing", Journal of Direct Marketing, 9(2), pp.59-66, 1995.
46. Kim, J. C., Kim, D. H., Kim, J. J., Ye, J. S. & Lee, H. S. "Segmenting the Korean Housing Market Using Multiple Discriminant Analysis", Construction Management and Economics, 18, pp.45-54, 2000.
47. Kim, J., Suh, E. & Hwang, H. "A Model for Evaluating The Effectiveness of CRM Using The Balanced Scorecard", Journal of Interactive Marketing, Volume 17, Number2, Spring, pp.5-19, 2003.
48. Knie-Andersen, M. "The Relationship between Customer Satisfaction, Customer Loyalty and Customer Profitability", Denmark: University of Aarhus, 2001.
49. Knox, S. "Loyalty-Based Segmentation and the Customer Development Process," European Management Journal, Vol. 16, No. 6, pp.729-737, 1998.
50. Kumar, V., Ramani, G. & Bohling, T. "customer lifetime value approaches and best practice applications", Journal of Interactive Marketing, Vol.18, No 3, summer, 2004.
51. Kutner, S. & Cripps, J. "Management the customer portfolio of healthcare enterprises", The Healthcare Forum Journal, 4(5), pp.52-64, 1997.
52. Morgan, R.M. & Hunt, S.D. "The Commitment Trust Theory of Relationship Marketing", Journal of Marketing, 58, (July), pp.30-37, 1994.
53. Mulhern, F. J. "Customer Profitability Analysis: Measurement, Concentration, and Research directions", Journal of Interactive Marketing, Vol. 13 (1), pp.25-40, 1999.
54. Murray, K.B. "A Test of Services Marketing Theory: Consumer Information Acquisition Activities", Journal of Marketing, Vol. 55, pp.10-25,1991.
55. Payne P. "A Strategic Framework For Customer Relationship Management ", ABT CRM White Paper, pp.2, March, 2002.
56. Pfeifer, P. E., Haskins, M. E. & Conroy, R.M. "Customer lifetime value, customer profitability, and the treatment of acquisition spending", To appear in the Journal of Managerial Issues, 2004.

57. Pfeifer, P.E. & Robert L. C. “Modeling Customer Relationships as Markov Chains”, Journal of Interactive Marketing, Vol. 14, pp43-55, 2000.
58. Reichheld, F.E. & Sasser Jr., W.E. “Zero-Defections Quality comes to services”, Harvard Business Review, September-October, pp.105-111, 1990.
59. Reichheld, F.E. “Learning form Customer Defections”, Harvard Business Review, pp.57-69, 1996.
60. Reinartz, W.J. & Kumar, V. “On the profitability of long-life customers in a noncontractual setting: an empirical investigation and implications for marketing”, Journal of Marketing, 64(October), pp.17-35, 2000.
61. Reinschmidt J., Bhattacharya R., Harris P. & Karanasos A. “Mining Relational and Nonrelational Data with IBM Intelligent Miner for Data”, Appendix B, pp.189-193, 1998.
62. Roberts, M. L. & Berger, P. D. “Direct marketing management”, Prentice-Hall: Englewood Cliffs, NJ,1989.
63. Ryals, L. “Are your customers worth more than money?”, Journal of Retailing and Customer Services, 9, pp.241–251, 2002.
64. Rygielski, C., Wang, J.C. & Yen, D.C. “Data mining techniques for customer relationship management”, Technology in Society, 24(4), pp.483-502, 2002.
65. SAS Institute. “From Data to Business Advantages: Data Mining, The SEMMA Methodology and SAS Software”, SAS Instutitute: Cary, North Carolina, 1998.
66. Schmittlein, D. C., Morrison, D.G., & Colombo, R. “Counting Your Customers: Who Are They and What Will They Do Next?”, Management Science, 33(January), pp.1-24, 1987.
67. Schmittlein, D.C. & Peterson, R. A. “Customer Base Analysis: An Industrial Purchase Process Application”, Marketing Science, Vol.13, pp.41-67, 1994.
68. Spring, P.N., Verhoef, P.C., Hoekstra, J.C. & Leeflang, P.S.H. “The Commercial Use of Segmentation and Predictive Modeling Techniques for Database Marketing”, Working Paper, University of Groningen, 2000.

69. Stahl, H.K., Matzler, K. & Hinterhuber, H. H. "Linking Customer Lifetime Value with Shareholder Value", Industrial Marketing Management, Vol.32, pp.267-279, 2003.
70. Stone, M., Neil W. & Muriel W. "Managing the change from marketing planning to customer relationship management", Long Range Planning, London, Vol.29, Iss.5, pp.675-683, Oct 1996.
71. Storbacka, K. "Customer Profitability: Analysis and Design, working paper. Center for Relationship Marketing and Services Management", Swedish School of Economics and Business Administration., 1998.
72. Thomas, J. S. "A Methodology for Linking Customer Acquisition to Customer Retention", Journal of Marketing Research, 38 (May), pp.262-268, 2001.
73. Trybula, WJ. "Data mining and knowledge discovery", Annual Review of Information Science and Technology, 32, pp.197-229, 1997.
74. Two Crows Corporation. "Introduction to Data Mining and Knowledge Discovery", Third Edition, 1999.
75. Verhoef, P.C. & Donkers, B. "Predicting customer potential value an application in the insurance industry", Decision Support Systems, 32, pp.189–199, 2001.
76. Wang, P. & Spiegel, T. "Database Marketing and Its Measurements of Success: Designing a Managerial Instrument to Calculate the Value of Repeat Customer Base", Journal of Direct Marketing, 8(2), pp.73–81, 1994.
77. Wayland, R. E. & Cole P. M. "Customer Connection: New Strategies for Growth", Harvard Business School press, 1997.
78. Wheation, P. "The Life Cycle View of Customers", U.S. Banker, June, 110, pp.77-78, 2000.
79. Wyner, G. A. "Customer Profitability: Linking Behavior to Economics", Marketing Research, 8(2), pp.36-38, 1996.

附件 1 各集群之 OLAP 分析表

Gender*Age*Cluster 多維度分析

CLV		Cluster									
Age	Gender	1	2	3	4	5	6	7	8	9	總計
15-19	F	4582.61	2854.15	3649.72	2122.10	299.85	3691.20			631.03	17830.66
	M	3972.78	3438.78	3801.65	973.01	755.27	3024.28			1664.21	17629.97
15-19 合計		8555.39	6292.93	7451.36	3095.11	1055.13	6715.47			2295.24	35460.62
20-29	F	16585.23	10322.55	19256.70	6539.98	3461.23	19349.33		853.14	16888.39	93256.54
	M	17721.84	10263.79	18738.52	10137.09	6377.50	17796.22		1510.25	16159.29	98704.49
20-29 合計		34307.07	20586.33	37995.21	16677.06	9838.73	37145.55		2363.39	33047.68	191961.04
30-39	F	18332.61	11607.32	17526.88	10457.05	4355.68	23950.61	712.26	1870.99	18832.93	107646.32
	M	16669.91	12138.39	22291.68	6080.76	7191.92	17227.89	1296.54	1914.22	13441.86	98253.16
30-39 合計		35002.52	23745.71	39818.56	16537.80	11547.59	41178.50	2008.80	3785.21	32274.79	205899.49
40-49	F	18381.95	10689.97	18641.48	8093.05	6401.09	21752.30		1307.27	24750.31	110017.44
	M	18594.35	10740.19	16805.12	7259.40	6240.93	24308.60	1440.04		10103.40	95492.03
40-49 合計		36976.31	21430.16	35446.60	15352.45	12642.03	46060.90	1440.04	1307.27	34853.71	205509.47
50-59	F	18434.02	8393.31	16559.49	10306.48	5488.41	24343.09	626.69	1193.73	12924.60	98269.82
	M	16043.73	13612.28	20306.73	7335.77	7341.15	17570.08	3327.10	1840.74	15923.36	103300.95
50-59 合計		34477.75	22005.59	36866.21	17642.25	12829.57	41913.17	3953.79	3034.48	28847.96	201570.77
60-69	F	17253.28	11723.34	17107.41	12773.58	4544.46	18064.46	657.10	2735.04	16548.05	101406.70
	M	17194.19	11408.93	17386.91	10467.02	6291.87	22178.82	4166.27		7866.64	96960.64
60-69 合計		34447.46	23132.27	34494.32	23240.60	10836.33	40243.28	4823.37	2735.04	24414.69	198367.35
70-79	F	18687.29	11111.72	16353.42	9864.63	3963.80	19962.68	639.34	1305.43	9874.75	91763.06
	M	16296.66	12153.29	19222.22	8396.29	6158.78	22747.71	1256.34	1521.37	17641.28	105393.93
70-79 合計		34983.95	23265.02	35575.64	18260.92	10122.57	42710.39	1895.68	2826.80	27516.03	197157.00
80-89	F	12795.49	8181.34	13281.34	8682.88	2438.55	16014.30		2032.09	12378.42	75804.40
	M	11632.58	8193.73	14569.26	7036.63	6641.85	16104.46	602.37	593.73	8812.53	74187.15
80-89 合計		24428.07	16375.07	27850.59	15719.51	9080.40	32118.76	602.37	2625.81	21190.95	149991.54
總計		243179	156833	255498	126526	77952	288086	14724	18678	204441	1385917

年所得*是否擁有房子*集群

Current		Cluster									
Yearly Income	House owner	1	2	3	4	5	6	7	8	9	總計
\$10K - \$30K	N	22073.42	14189.29	23229.02	11956.69	4772.25	31926.38	2527.81	1445.72	18071.60	130192.18
	Y	30513.95	19757.00	34360.83	16918.68	11444.05	35367.46	1996.98	742.31	19766.82	170868.07
\$10K - \$30K 合計 *		52587.36	33946.30	57589.85	28875.37	16216.30	67293.84	4524.79	2188.02	37838.42	301060.25
\$110K - \$130K	N	2126.60	1063.63	2115.52	1461.42	1144.79	2164.14			607.52	10683.63
	Y	10460.04	5846.91	10409.66	3231.87	3679.83	9479.63	1302.18	2577.26	5074.62	52062.01
\$110K - \$130K 合計 *		12586.64	6910.54	12525.18	4693.30	4824.62	11643.77	1302.18	2577.26	5682.14	62745.63
\$130K - \$150K	N	2364.95	1545.44	1317.49	1158.30	494.37	2316.19	752.56	1254.70	5021.15	16225.16
	Y	8793.34	6654.11	11029.31	5777.38	5446.33	8920.54	626.35		12007.16	59254.52
\$130K - \$150K 合計 *		11158.29	8199.55	12346.80	6935.68	5940.70	11236.73	1378.91	1254.70	17028.31	75479.68
\$150K +	N	1213.24	528.52	1609.58	680.81	267.08	881.47				5180.70
	Y	3364.25	2977.58	4260.30	1257.24	955.68	5100.77	695.88		3301.28	21912.97
\$150K + 合計 *		4577.50	3506.10	5869.88	1938.04	1222.75	5982.24	695.88		3301.28	27093.68
\$30K - \$50K	N	33807.90	20164.29	38568.92	17002.22	12257.29	41982.16	755.68	2789.28	32781.09	200108.84
	Y	43664.67	29271.29	47100.70	22262.91	13887.82	48792.93	1458.54	4531.01	42347.53	253317.41
\$30K - \$50K 合計 *		77472.57	49435.58	85669.63	39265.13	26145.11	90775.10	2214.23	7320.30	75128.62	453426.26
\$50K - \$70K	N	19437.17	13026.70	18641.92	9129.03	5671.68	21678.74	744.16	1198.94	12851.17	102379.52
	Y	26207.42	17475.03	24322.67	11861.08	6479.86	30172.64	1306.83	758.93	14958.25	133542.70
\$50K - \$70K 合計 *		45644.58	30501.73	42964.59	20990.11	12151.54	51851.37	2051.00	1957.87	27809.42	235922.22
\$70K - \$90K	N	13190.89	8641.10	13727.54	6950.52	3168.25	13833.58	626.69	2052.74	9003.44	71194.75
	Y	15143.30	8764.12	15332.49	9569.59	5179.48	22304.19	1291.04	730.66	20215.81	98530.68
\$70K - \$90K 合計 *		28334.19	17405.22	29060.03	16520.11	8347.73	36137.78	1917.73	2783.41	29219.25	169725.43
\$90K - \$110K	N	2157.35	2137.40	1970.51	1472.68	1429.42	1720.06			2226.78	13114.20
	Y	8660.04	4790.65	7502.03	5835.28	1674.16	11445.15	639.34	596.45	6206.83	47349.93
\$90K - \$110K 合計 *		10817.39	6928.05	9472.54	7307.96	3103.58	13165.21	639.34	596.45	8433.61	60464.13
總計 *		243178.52	156833.07	255498.50	126525.70	77952.35	288086.03	14724.05	18678.00	204441.05	1385917.27

教育程度*婚姻狀態*集群

Current		Cluster									
Education	Marital Status	1	2	3	4	5	6	7	8	9	總計
Bachelors Degree	M	30294.90	21607.34	32919.19	16574.21	9797.60	39597.96	1277.29	1304.43	25976.88	179349.8
	S	31806.46	21382.96	33052.14	14149.12	9090.79	35641.07	1384.55	4717.78	26044.04	177268.9
Bachelors Degree 合計 *		62101.36	42990.30	65971.33	30723.33	18888.39	75239.03	2661.84	6022.21	52020.92	356618.7
Graduate Degree	M	6511.90	4824.97	6168.11	2713.28	2438.16	4856.39	752.56		6386.69	34652.06
	S	7853.23	3269.27	6492.09	5177.68	2195.73	7675.22		1277.71	6209.35	40150.29
Graduate Degree 合計 *		14365.14	8094.24	12660.19	7890.96	4633.90	12531.61	752.56	1277.71	12596.04	74802.35
High School Degree	M	33531.78	20226.46	38269.36	18079.53	14546.31	41553.36	1383.50	2702.33	29872.33	200164.9
	S	35819.08	23424.46	38262.33	15089.36	14379.59	39708.59	3386.91	2058.86	34067.81	206196.9
High School Degree 合計 *		69350.87	43650.92	76531.69	33168.89	28925.90	81261.95	4770.41	4761.19	63940.14	406361.9
Partial College	M	11492.76	7619.63	11912.33	9751.03	1838.99	13581.80		652.23	8055.86	64904.64
	S	11611.88	8589.94	11834.47	6218.11	923.77	13092.36		1898.77	9124.20	63293.49
Partial College 合計 *		23104.64	16209.57	23746.80	15969.14	2762.76	26674.16		2551.00	17180.06	128198.1
Partial High School	M	36638.84	23352.62	38105.45	18175.39	10224.74	44444.73	3288.24	1338.75	32002.22	207570.9
	S	37617.69	22535.42	38483.04	20598.00	12516.67	47934.54	3250.99	2727.14	26701.67	212365.1
Partial High School 合計 *		74256.53	45888.03	76588.49	38773.38	22741.41	92379.27	6539.23	4065.90	58703.89	419936.1
總計 *		243178.5	156833.0	255498.5	126525.7	77952.35	288086.0	14724.05	18678.00	204441.0	1385917

擁有小孩數*國家*集群表

Current		Cluster									
Total Children	Country	1	2	3	4	5	6	7	8	9	總計
0	Canada	1133.82	2192.87	3184.97							6511.67
	Mexico	645.29	1701.38	5338.45		1160.32	9198.06			7188.37	25231.87
	USA	21860.02	13509.00	15092.23	10005.59	8624.47	12560.02		1898.03	9034.31	92583.66
0 合計 *		23639.13	17403.25	23615.65	10005.59	9784.79	21758.07		1898.03	16222.68	124327.20
1	Canada	1417.46	4744.08	6397.62							12559.16
	Mexico	563.19	4335.75	11900.04	842.44	1523.74	21478.52			10962.26	51605.94
	USA	46857.85	24442.22	25390.76	21627.55	14591.27	34799.66	1935.38	1841.59	19473.11	190959.38
1 合計 *		48838.50	33522.05	43688.42	22469.99	16115.01	56278.18	1935.38	1841.59	30435.37	255124.48
2	Canada	1984.31	4604.43	7783.42							14372.17
	Mexico	466.41	3777.96	11158.33	686.70	2002.97	22349.61			12491.71	52933.69
	USA	42994.34	24895.45	32295.44	26056.90	16778.27	37634.39	2079.62	1254.06	23036.53	207025.00
2 合計 *		45445.07	33277.85	51237.19	26743.59	18781.24	59984.00	2079.62	1254.06	35528.24	274330.86
3	Canada	2290.71	3846.96	8440.22			250.81				14828.69
	Mexico	911.08	4250.38	11419.74		1673.04	25146.72		637.32	13292.85	57331.13
	USA	48157.93	21413.73	30220.19	26380.40	15705.74	35018.31	4714.95	5958.04	26868.10	214437.40
3 合計 *		51359.72	29511.07	50080.15	26380.40	17378.78	60415.84	4714.95	6595.36	40160.95	286597.22
4	Canada	1894.28	4370.97	8205.44							14470.69
	Mexico	987.34	2754.26	13345.56	263.45	1409.88	25428.70			16383.79	60572.98
	USA	44522.07	21938.96	35037.20	27460.44	10389.28	35175.53	3966.48	4085.19	35647.53	218222.69
4 合計 *		47403.70	29064.19	56588.19	27723.89	11799.16	60604.23	3966.48	4085.19	52031.32	293266.35
5	Canada	842.81	1656.12	5685.79							8184.72
	Mexico	353.21	1874.87	5660.45		802.80	11891.49	712.26		10797.43	32092.51
	USA	25296.40	10523.67	18942.65	13202.24	3290.57	17154.22	1315.36	3003.77	19265.06	111993.93
5 合計 *		26492.41	14054.66	30288.89	13202.24	4093.37	29045.71	2027.62	3003.77	30062.49	152271.16
總計 *		243178.52	156833.07	255498.50	126525.70	77952.35	288086.03	14724.05	18678.00	204441.05	1385917.27

職業*會員卡別*集群

Current		Cluster									
Occupation	Member Card	1	2	3	4	5	6	7	8	9	總計
Clerical	Bronze	3500.57	2221.59	3963.03	1036.94	242.27	2886.81		705.21	2638.37	17194.78
	Golden	307.41	744.02	250.15	244.54		742.91		652.23	613.89	3555.14
	Normal	80.28	71.27	174.57							326.12
	Silver	843.01	431.88	62.64	478.10		807.97				2623.60
Clerical 合計 *		4731.28	3468.75	4450.40	1759.58	242.27	4437.69		1357.43	3252.26	23699.65
Management	Bronze	23129.15	15487.63	21256.72	12081.49	7957.79	29217.21	1441.24	3203.65	22029.68	135804.54
	Golden	5462.60	3734.79	6838.61	3444.72	1242.29	6263.94	674.92	730.66	4321.32	32713.85
	Normal	1795.29	1562.88	1712.36	2262.05	969.10	1144.95			1394.91	10841.53
	Silver	3191.62	2876.44	5359.41	1854.92	1701.28	4529.44			1919.68	21432.80
Management 合計 *		33578.66	23661.74	35167.10	19643.18	11870.46	41155.53	2116.16	3934.31	29665.59	200792.73
Manual	Bronze	28437.51	16059.82	27306.71	11310.00	10417.15	32468.68	1467.95	1267.16	20433.75	149168.73
	Golden	4775.40	2964.62	5767.84	1897.90	2835.01	8347.92		1464.78	7275.39	35328.87
	Normal	27087.42	16239.07	28369.94	14548.50	5153.85	24409.69	2583.36	592.58	14644.70	133629.10
	Silver	3969.72	3539.94	4223.57	715.46	243.66	4868.27			2976.11	20536.72
Manual 合計 *		64270.06	38803.44	65668.06	28471.86	18649.67	70094.56	4051.31	3324.51	45329.95	338663.42
Professional	Bronze	54551.60	35663.28	54413.24	27331.86	18562.67	60871.49	2547.43	5899.53	45173.03	305014.14
	Golden	10830.78	6912.48	11498.46	7475.84	2204.41	15377.85	1249.49	624.08	12911.05	69084.45
	Normal	5037.79	3833.72	4812.86	2129.12	1823.83	6587.90			967.10	25192.31
	Silver	10856.87	5965.62	10068.08	5064.00	2611.64	8306.62	695.88	661.21	8948.83	53178.74
Professional 合計 *		81277.04	52375.10	80792.65	42000.82	25202.55	91143.86	4492.80	7184.81	68000.01	452469.65
Skilled Manual	Bronze	25639.78	17382.84	31787.04	13889.79	9097.64	31783.81	1490.45	1281.49	29584.06	161936.89
	Golden	4729.29	3082.10	6346.34	4613.10	1216.82	8136.90			7922.98	36047.53
	Normal	25452.60	16095.28	25948.48	13325.68	11286.19	37677.67	2573.33		20686.20	153045.44
	Silver	3499.81	1963.81	5338.43	2821.70	386.75	3656.00		1595.45		19261.96
Skilled Manual 合計 *		59321.49	38524.04	69420.29	34650.27	21987.39	81254.39	4063.78	2876.94	58193.24	370291.82
總計		243179	156833	255498	126526	77952	288086	14724	18678	204441	1385917

媒體類別*集群

Store Sales	Cluster									
Media Type	1	2	3	4	5	6	7	8	9	總計 *
Bulk Mail		1661.7	2150.9	837.8	280.7	515.9	186.03	57.64	1052	6742.6
Cash Register Handout		1479.3	2115.2	1465	1181.9	2358	345.7	463.98	2171.5	11580.86
Daily Paper		1529.9	2218.9	1734	1474.7	3033	360.25	399.13	3089.6	13839.54
Daily Paper, Radio		1046.5	1678.6	2796	1645.2	4341	101.76	215.99	1940.3	13764.66
Daily Paper, Radio, TV		2861.6	3291.2	2660	1111.4	2569	87.18	92.65	2982.2	15655.82
In-Store Coupon		757.93	601.81	1232	621.09	1340	72.69	198.96	2764.6	7589.06
No Media	102466.33	46683	65447	68404	38485	98158	7776.1	10016	68716	506152.5
Product Attachment		1783	2295.3	2629	1335.3	3457	48.28	189.56	791.1	12529.19
Radio		760.64	1057.9	601.3	246.11	354.2	85.18	45.94	103.97	3255.3
Street Handout		1157.6	2196.6	1058	383.9	1540	21.86	197.32	3033.2	9588.68
Sunday Paper		762.08	1279.7	1223	603.85	2278	105.59	424.98	1099.3	7776.4
Sunday Paper, Radio		738.77	1214.7	2110	1059.6	3101	131.59	280.63	2815	11452.1
Sunday Paper, Radio, TV		415.2	444.04	609.7	195.89	726.1	211.18	34.2	2309.9	4946.11
TV		1571.6	2127.5	605.1	306.81	453.9				5064.9
總計 *	102466.33	63209	88119	87966	48931	124226.1	9533.4	12617	92869	629937.7

