

行政院國家科學委員會補助專題研究計畫成果報告

整合資料挖採與關係管理來提昇顧客價值:顧客分析與
互動技術之研究

計畫類別：■個別型計畫

計畫編號：NSC 89-2416-H-002-100-

執行期間：八十九年八月一日至九十年七月三十一日

計畫主持人：陳文華

執行單位：國立台灣大學商學研究所

中華民國九十年十月三十一日

中文摘要

顧客關係管理已被視為行銷的新典範，多數學者認為行銷學的焦點已由傳統的單次「交易」主體移轉為「長期的關係」。而所謂的顧客關係管理乃是一種新的作法，提供顧客優良的服務品質，其目的是為了更有效率地獲取、開發並留住企業最重要的資產——顧客。在和顧客接觸的過程中，針對個別的差異提供和其需求一致的銷售計畫。

在顧客關係管理的議題上，將顧客分群是一個很重要的概念和步驟。本研究以國內某知名銀行之信用卡客戶與某知名化粧品保養品公司客戶為實證研究對象。探討以提供服務為主要的公司及以提供產品為主要的公司，在顧客分群變數選擇上之考量，並比較分群後之結果分析。

研究結果顯示，行為變數與顧客忠誠度間的關連性較人口統計變項與顧客忠誠度間的關連性為高，且顧客與公司往來後的滿意度亦會反應在顧客實際購買行為上。在行為變數方面，RFM (Recency, Frequency, Monetary)變數是直接與忠誠度和滿意度相關的行為變數，故以此三個變數作為顧客分群之區隔變數。

以信用卡公司而言，RFM 之變數定義可以明確地應用在顧客之行為，且顧客之交易資料完整；在化粧品保養品公司方面，由於產品種類繁多，購買型態迥異，且顧客之交易紀錄較不完整，RFM 之變數定義及交易紀錄不若信用卡公司所呈現者來得完整。故以 RFM 變數作為顧客分群之區隔變數時，信用卡顧客的分群效果較化粧品保養品公司顧客的分群效果為佳。

在分群結果方面，信用卡顧客分成 6

群，其中以游離顧客群之人數最多，而較明顯之集群為已流失顧客群、刷卡換贈品顧客群及重要顧客群。化粧品保養品公司顧客群則分成 4 群，其中以游離顧客群之人數最多，而較明顯之集群為已流失顧客群及重要顧客群。

關鍵詞：資料挖採、顧客關係管理、顧客分析、顧客分群

一、背景與目的

從企業競爭的角度而言，今日競爭觀點與昔日迥然不同。過去企業是由製造與成本的立場著眼，透過大量生產的方式為客戶提供所謂價廉又物美的產品或服務。基於這個出發點，企業致力自動化生產，提高產能，甚至建置 ERP(enterprise resources planning)強化供應鏈關係。因此，企業的競爭是製造對製造的競爭，流程對流程的競爭，供應鏈對供應鏈的競爭。

由於自我意識的覺醒，消費者已不再滿足於大量生產的標準化商品或服務，而希望供應商可依據顧客需求提供個人化的商品，而顧客的需求更是時時在變動著，這種客製化的趨勢無論對實務管理者或學術研究者而言均是相當大的衝擊(Pine, 1993)。如 1997 年 Harvard Business Review 的七十五年紀念特刊中，亦認為大量客製化(mass customization)是行銷領域中一個重要的里程碑。在大量客製化的觀點之下，如何瞭解顧客需求以提供其所需之產品及服務，並隨時追蹤消費者滿意度的資訊，變得十分重要。另一方面，長期顧客的重要性也逐漸受到企業的重視，不但開發新顧客的成本超出維繫舊顧客的支出，曾有研究指出此一差距高達六倍至九倍

(Peppers and Rogers,1993)，長期顧客對公司利潤的貢獻亦較短期顧客高 (Reichheld,1994)。因應這樣的趨勢，並結合進展迅速的資訊科技，顧客關係管理的議題受到重視。並被視為行銷的新典範。

在顧客關係管理的議題上，將顧客分群是一個很重要的概念和步驟。將顧客依照正確的區隔變數加以分群，讓公司針對不同的顧客，設計不同的行銷組合，以期有效滿足顧客需求，並使公司易於發掘和比較行銷機會，使產品和行銷訴求更能切合市場需求並根據特定市場區隔反應特性，發展適當的行銷計劃及預算。因此，本研究之目的，在於探討顧客分群之重要區隔變數，並以國內某銀行之信用卡客戶(視為以提供服務為主的公司)及某化粧品保養品公司之客戶(視為以提供產品為主的公司)為實證研究對象。並比較二家公司之分群結果，提供企業界在進行顧客區隔時之參考與運用。

二、文獻探討

市場區隔的概念在行銷界運用已久，Wendell Smith 於 1956 年提出「市場區隔化」觀念，認為市場具有多元性質，行銷人員在發掘市場之歧異相關因素後，可將一個錯綜複雜的市場區隔為幾個同質的小市場，企業可以權衡本身的資源，選擇適當的目標市場，發展行銷組合，以把握這些市場。後續學者對於市場區隔化的定義，與 Smith 的觀念相去不遠，主要在於認定市場歧異性上，有些學者認為以顧客本身需求與特徵之同質性來區分(Dickson and Ginter,1987; Kotler,1992; McCarthy,1993; Schiffman and Kanuk,1994; Engel, Blackwell and Miniard,1995; Sarabia,1996; 洪順慶,1999)，另一群學者則以顧客易受特定行銷活動影響來區分子市場 (Alfred,1981; McDonald,1998; 蔡泰生,1999)。其實這二種概念的基礎是一致的，需求與特徵同質性高的顧客，通常較易被某一種獨特的行銷組合所滿足。因此，所謂市場區隔乃是將

市場中具有相同需求或對公司之行銷策略有相同回應的同質顧客集合成一群，針對不同群之顧客，採取不同的行銷策略，以期達到最大效率。

區隔市場的基礎變數，乃是指可以用來劃分個人、群體或組織機構的某些特徵，也由於這些特徵上的差異，讓不同市場區隔間的消費者，對於企業提供的各種行銷服務會有不同的反應(Dibb et al., 1994)。

市場區隔並沒有唯一的方法，通常會採用數個區隔變數或區隔基礎來畫分市場，期望能真正深入了解市場的結構。一般而言，消費者市場常用的區隔變數有人口統計變數(包括性別、年齡、教育程度、職業、所得、家庭人數、家庭生命週期、宗教、種族與國籍等變數)、地理變數(包括城市規模、人口密度、氣候、國家、國家地區與地形等變數)、心理變數(包括購買者或消費者的社會階級、生活型態與人格特徵等變數)與行為變數(包括產品購買時機、追尋利益、使用者狀況、使用率、忠誠度、購買準備階段與對產品或行銷組合的態度等變數)(Kotler,1992)。

Haley (1968)將市場區隔變數分為描述性變數 (descriptive factors)與因果性變數 (causal factors)二大類。描述性變數包括地理變數、人口統計變數；因果性變數即消費者所追求之利益，導致消費者購買行為之原因，又稱為利益區隔 (benefit segmentation)。Wind (1978)認為消費者行為中大部份的變數都可做為基礎變數，可將其分成一般變數與特定情境變數兩類。一般變數包括人口統計變數、社會經濟特徵、個性、生活型態或其他心理變數等；特定情境變數則包括使用及購買型態、追尋利益、使用頻率、品牌忠誠度、行銷因素敏感度等。Norgan (1994)將市場區隔變數分為產品特性、使用者特性及使用情境三大類。Dibb (1994)等人則將市場區隔變數分為與顧客基本特徵有關以及與產品購買行為有關兩大類。

根據實證結果，顧客購買行為與貢獻

度及滿意度，較顧客基本特徵變數，更能反應顧客間之不同。在行為變數方面，顧客之購買金額、購買次數與購買時間是衡量顧客與公司間關係的重要指標，在實務上亦常利用這三個變數來區隔顧客，並做為直效行銷上，目標顧客選擇之依據，亦即實務界常用之 RFM 分析法。

所謂 RFM 分析法即為以近期 (Recency)、頻率 (Frequency) 及金額 (Monetary) 這三種指標來衡量與顧客的關係。Schjins and Schroder (1996)指出，從行為的觀點來看，RFM 是最常用來測量與顧客關係強度的方法之一。

Kahan (1998)提及 RFM 是應用非常廣泛的行為分析技術，利用 RFM 可以更簡單、更快速分析公司的顧客。其具體作法為：(1)將公司顧客的購買日期按照年代先後順序排列(2)將顧客平均分割成五等分(3)把最近期間購買的百分之二十的顧客編號為 5，依次類推分別為 4、3、2、1，編號 1 為最近購買時間離現在最久。接著，再排列公司顧客購買次數並重複上述的步驟將顧客分類及編號，最後則是將平均購買金額亦納入探討及分類。

透過上述的步驟，可以獲得每一個顧客的 RFM 分數，由最好的顧客(555)到最差的顧客(111)共有 125 群。針對不同群的顧客可以採用不同的行銷策略。在檢視傳統 RFM 分析可以發現，公司在決定分幾群時並沒有一定的衡量標準，不一定是五分法，亦有可能是十分法，採用不同的分法會產生不同的結果。除了分法不同之外，三項指標的權重也可能不同，亦即不一定三項指標所評定的權重都是 5、4、3、2、1。

Kahan (1998)認為 RFM 是非常有用的行為分析工具，不但簡單且具成本效益，提供公司本身顧客的交易資訊。透過整合認知和行為的分析技術，使直效行銷業者可以更有效率地透過電子科技，來獲取資訊，並可增加回應率、降低每一次的訂購成本及獲得更高的利潤。

三、研究結果

以國內某銀行之信用卡客戶(以下簡稱 A 銀行)及某化粧品保養品公司之客戶(以下簡稱 B 公司)為實證研究對象。針對 A 銀行之 534 位信用卡客戶及 B 公司之 350 位客戶，依據行為變數—RFM 將顧客分群，並根據人口統計變項，針對每一個顧客區隔做特性分析，以期對每個顧客區隔的組成有更深入的了解。在分群數目的決定上，若分群數目太多時，會造成各區隔樣本數太少的問題，使模式之效度受到質疑；而分群數目太少時，則會造成顧客分群之效果不佳，無法有效區隔出不同特性的顧客。因此，本研究將顧客分為 4-6 群，並依據分群結果之 Condorcet 值來判斷分群結果的優劣。Condorcet 值愈高，表示分群之結果愈好。一般而言，Condorcet 值高於 0.6 時，即可視為好的分群結果。

3.1 A 銀行顧客分群結果

抽樣樣本中有 50 位顧客在 2000 年 8 月至 2001 年 1 月間皆無消費記錄，所以直接將這群顧客列為同一群，並視為已流失的一群顧客，其他 486 位顧客則依據 RFM 變數進行集群分析。

在進行顧客分群之前，先針對 RFM 變數作一界定。若顧客最近一次消費時間在 2001 年 1 月間，則將 R 定義為 6，時間在 2000 年 12 月間定義為 5，依此類推；F 則是將 6 個月的消費次數加總；M 則為 6 個月的總消費金額。在分群數目的決定上，先將顧客分成 5 群或 6 群，再依據分群之結果，決定最終的集群個數，結果如圖一、圖二與表一所示。在圖一與圖二中，深色圖形部份為 484 位有刷卡顧客在 RFM 三個變數上之分佈情形；淺色圖形部份為各集群顧客在 RFM 三個變數之分佈情形。

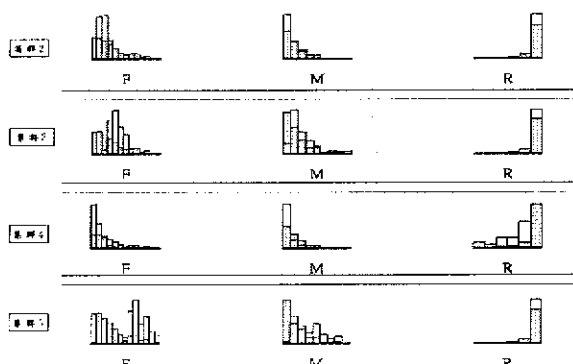
分為 5 群時，整體的 Condorcet 值為 0.7345，分為 6 群時，整體的 Condorcet 值為 0.7586，皆高於 0.6 的標準，表示分成 5 群或 6 群的結果都是可接受的。但在顧客分 5 群的情況下，在集群 5 這個集群

中，總消費金額分佈過於分散，集中度不是非常明顯。顧客分成 6 群時，每一群顧客的特性較為明顯。

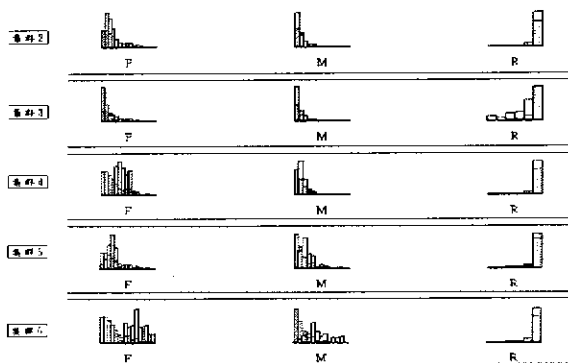
本研究以二種方式對各集群進行人口統計變項分析。第一種為集群與所有人口統計變項變數分析(bivariate analysis)，第

二種方式為顧客分類(classification)，結果如圖三與圖四所示。

由圖三與圖四之結果，發現各集群在人口統計變項上並沒有特別的特徵，因此，行為變數與人口統計變項間並未存在強烈的關連性。



圖一 A 銀行顧客分 5 群之結果



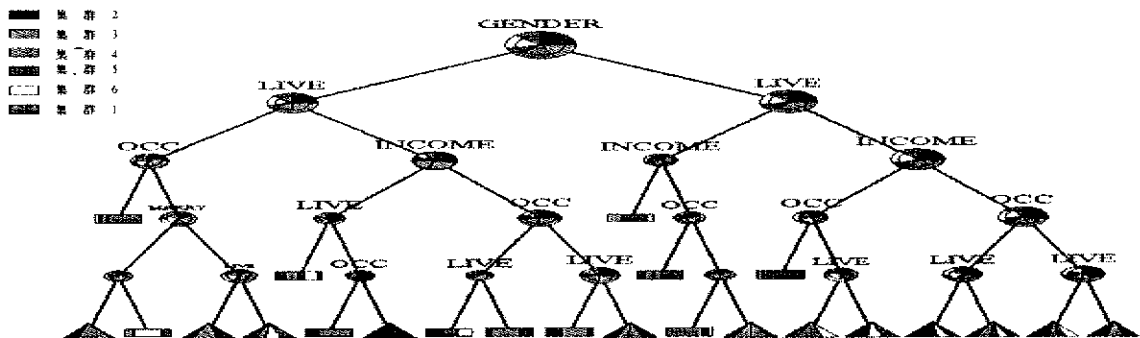
圖二 A 銀行顧客分 6 群的結果

表一 A 銀行顧客分 6 群的結果

		集群 1	集群 2	集群 3	集群 4	集群 5	集群 6
集群大小		50 人 (9%)	134 人 (25%)	100 人 (19%)	96 人 (18%)	85 人 (16%)	69 人 (13%)
Condoirect 值		1	0.8896	0.6696	0.7648	0.6823	0.5531
集中度	R	0 (100%)	6 (100%)	5 (47%)	6 (98%)	6 (87%)	6 (94%)
	F	0 (100%)	0~20 (99%)	0~15 (94%)	20~45 (97%)	5~25 (93%)	35~75 (85%)
	M	0 (100%)	0~40,000 (99%)	0~20,000 (81%)	20,000 ~60,000 (75%)	20,000 ~100,000 (86%)	60,000 ~160,000 (63%)
平均數	F	0	9.545	5.42	31.78	15.8	57.77
	M	0	15,463	11,712	40,605	70,619	140,496
集群定義		此集群之顧客，在 6 個月內皆無刷卡記錄，以信用卡之使用習慣來看，可視為 A 銀行已流失之顧客。	此集群之顧客，在最近一個月內有刷卡記錄，平均而言一個月刷卡次數為 1.5 次，每月金額約 2,600 元，此集群與公司之大多數顧客的行為模式相同，因此視為一般顧客。	此集群之顧客在最近一個月內沒有刷卡記錄，且其頻率與金額皆不高，表示此集群之顧客可能為即將流失之顧客，或是此集群之顧客擁有許多張的信用卡，而 A 銀行之信用卡並非其主要使用之信用卡。	此集群之顧客，平均而言一個月刷卡次數為 5 次，每月金額為 6,500 元。與集群 5 相比，刷卡次數為其二倍，但累積刷卡金額卻較小。	此集群之顧客，平均而言一個月刷卡次數為 2.5 次，每月刷卡金額約為 10,000 元，在累積金額方面，僅次於集群 6，可將此集群視為 A 銀行之次要顧客。	此集群之顧客，刷卡次數與累積金額皆高出其他集群之顧客許多，可視為 A 銀行貢獻度最高，且最為重要之顧客。依其刷卡次數來看，A 銀行信用卡為其主要使用之信用卡。
集群命名		已流失顧客	一般顧客	即將流失顧客	小額消費刷卡顧客	次要顧客	重要顧客



圖三 A 銀行顧客 BIVARIATE 分析結果



圖四 A 銀行顧客分類結果

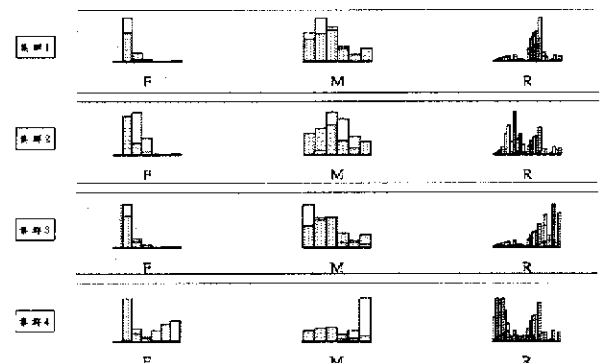
3.2 B 公司顧客分群結果

在 B 公司方面，所有的顧客在 2000 年 2 月至 2001 年 1 月間均有消費記錄，因此，所有的顧客皆依照 RFM 變數進行顧客分群。

在進行顧客分群之前，先針對 RFM 變數作一界定。由於化粧保養品購買習性與產品使用期間的差異性，單純以顧客最近一次消費時間來界定 Recency 並無法有效地區分出顧客在購買時點上的差異。因此，首先針對 R 這個變數進行分類，以最近一次購買日期與倒數第二次購買日期二個維度來區分顧客的 Recency，R 的數字愈高表示顧客最近兩次的購買日期間距愈久；F 則是將 2000 年 2 月份至 2001 年 1 月份的消費次數加總；M 則為 2000 年 2 月份至 2001 年 1 月份的總消費金額。在分群數目方面，將 350 位顧客分成 4 群。分群之結果如圖五與表二所示。在圖五中，深色圖形部份為 350 位顧客在 RFM 三個

變數上之分佈情形；淺色圖形部份為各集群顧客在 RFM 三個變數之分佈情形。

將 B 公司 350 位顧客分成 4 群時，整體的 Condorcet 值為 0.682；就個別集群的 Condorcet 值，集群 1 與集群 3 大於 0.6，但集群 4 的 Condorcet 值則太低，只有 0.28，表示集群 4 內的顧客，在 RFM 這三個變數上的同質性並沒有很高。因此，B 公司的分群結果並不如 A 銀行來的理想。



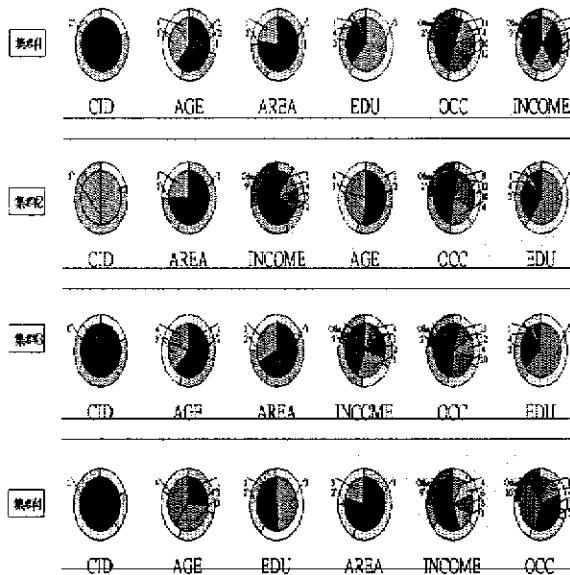
圖五 B 公司顧客分群結果

表二 B 公司顧客分群結果

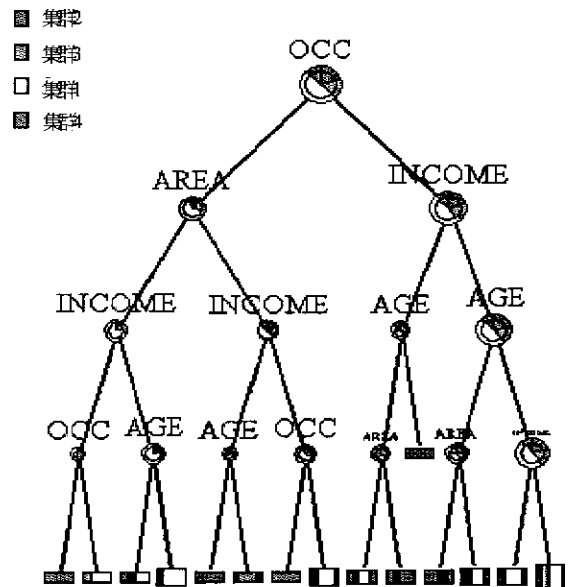
		集群 1	集群 2	集群 3	集群 4
集群大小		178 人 (51%)	68 人 (19%)	63 人 (18%)	41 人 (12%)
Condoirect 值		0.7309	0.5668	0.5971	0.2809
集中度	R	13~18 (99%)	5~11 (85%)	19~24 (86%)	1~6 (78%)
	F	1~3 (100%)	2~4 (97%)	1~3 (98%)	4~12 (76%)
	M	0~9,000 (89%)	6,000~20,000 (84%)	0~9,000 (87%)	20,000 元以上(77%)
平均數	F	1.056	2.338	1.159	5.244
	M	5,070	11,337	4,965	29,584
集群定義		此集群之顧客，R 值偏高，F 與 M 偏低，表示此集群之顧客並非以 B 公司之產品為主要的使用品牌。	此集群之顧客，在 RFM 變數上之情形屬於中間階段，較無明顯的特性，可將此集群視為 B 公司之一般往來顧客。	此集群之顧客，已有相當長的一段時間沒有消費記錄。可將此集群視為 B 公司已流失之顧客。	此集群之顧客在 RFM 三個變數上的表現都是最佳的。且消費金額很高，可將此集群視為 B 公司之忠誠顧客，且為最重要的顧客。
集群命名		游離顧客	一般顧客	已流失顧客	重要顧客

本研究以二種方式對 B 公司各集群進行人口統計變項分析。第一種方式為集群與所有人口統計變項變數分析，第二種方式為顧客分類，結果如圖六與圖七所示。

由圖六與圖七之結果，發現 B 公司各集群在人口統計變項上並沒有特別的特徵，因此，行為變數與人口統計變項間並未存在強烈的關連性。



圖五 B 公司顧客 BIVARIATE 分析結果



圖六 B 公司顧客分類結果

四、研究結論

行為變數與顧客忠誠度間的關連性較人口統計變項與顧客忠誠度間的關連性來得高，且顧客與公司往來後的滿意度亦會反應在顧客實際購買行為上。因此，以行為變數作為顧客分群之區隔變數，更能反應出顧客間的差異。在行為變數方面，RFM 變數是直接與忠誠度和滿意度相關的行為變數，故以此三個變數作為顧客分群之區隔變數。

以人口統計變項對各集群顧客做集群定義與變數分析結果顯示，各集群顧客在人口統計變項上，並沒有明顯的差異，即顧客之行為與其性別、年齡、地區、教育程度、職業及個人月收入等變數間並沒有很強的關連性，顧客並不會因為上述變數的關係而有不同的行為模式。

以 RFM 變數作為顧客分群之區隔變數，信用卡顧客的分群效果較化粧品保養品公司顧客的分群效果好。可能的原因在於，以信用卡而言，RFM 之變數定義可以明確的應用在顧客之行為，且顧客之交易資料完整。而在化粧品保養品公司方面，由於產品種類繁多，購買型態迥異，且顧客之交易紀錄較不完整，因此，RFM 之變數定義及交易紀錄不若信用卡公司來的完整，故以此三個變數來區隔顧客之效果較差。

在分群結果方面，將信用卡顧客分成 6 群，其中以游離顧客之人數最多(佔 25%)，而較明顯之集群為已流失顧客(佔 9%)、刷卡換贈品顧客(佔 8%)及重要顧客(佔 13%)，此三個集群在 F 和 M 二個變數上的歧異性較大。研究另將化粧品保養品公司顧客分成 4 群，其中以游離顧客之人數最多(佔 51%)，而較明顯之集群為已流失顧客(18%)及重要顧客(12%)，此二個集群在 R、F、M 三個變數上的歧異性較大。

五、參考文獻

- 葉怡成 (1992)，類神經網路模式應用與實作，儒林書局。
- 紀志霖 (2000)，"應用人口特質差異提高電子郵件廣告效果之機制-透過類神經網路建構預測及分類模式"，逢甲大學企業管理研究所碩士論文。
- 林建廷 (2000)，"類神經網路在無線通訊市場消費區隔與預測之應用研究"，元智大學管理研究所碩士論文。
- Anderson, J. A. (1995), "An Introduction to Neural Networks," Book News, Inc. Portland, OR.
- Hruschka, H. and Natter M. (1999), "Comparing Performance of Feedforward Neural Nets and K-Means for Cluster-Based Market Segmentation," *European Journal of Operational Research*, Vol.114, pp.346-353.
- Michaelj, A.B. (1997), *Data Mining Techniques For Marketing, Sales and Customer Support*, Wiley Computer Publishing.
- Vellido, A., Lisboa PJG. and Meehan K. (1999), "Segmentation of the Online Shopping Market Using Neural Networks," *Expert Systems with Applications*, Vol.17, pp.303-314.