

大數據行銷期末報告

使用資料庫：信用卡

B07302230 財金三 王博奕

目錄

壹、	前言	5
貳、	資料概況描述	5
一、	客戶資料檔	6
二、	信用卡資料檔	7
參、	顧客靜態價值分析—RFM 法	8
一、	五等均分法	8
二、	Bob Stone 法	8
三、	自定義法	9
四、	小結	10
肆、	顧客動態價值分析—CAI 法、CRI 法	11
一、	CAI	11
二、	CRI	13
三、	小結	16
伍、	顧客集群分析	17
一、	建立市場區隔	17
二、	小結	20
陸、	購物籃分析	21
一、	相關係數矩陣	21
二、	因素分析	22
三、	小結	23
柒、	結語	24

表圖目錄

表目錄：

表格	頁碼	表格	頁碼
表 2-1	5	表 5-1	17
表 3-1	8	表 5-2	17
表 3-2	9	表 5-3	18
表 3-3	10	表 5-4	19
表 3-4	10	表 5-5	19
表 4-1	11	表 5-6	19
表 4-2	11	表 5-7	20
表 4-3	14	表 6-1	21
表 4-4	15	表 6-2	22
表 4-5	15	表 6-3	23

圖目錄：

圖表	頁碼	圖表	頁碼
圖 2-1	6	圖 2-5	7
圖 2-2	6	圖 4-1	12
圖 2-3	6	圖 4-2	13
圖 2-4	7	圖 4-3	15

摘要

本份報告將透過五種不同指標來分析與橫顧客價值，透過不同的分析方式，來對於不同的顧客投放不同的行銷建議，且唯有在考量到不同指標的情況下才能夠給予最為全面的視野，因此最後結語的部分會通盤分析以及給予行銷建議。

- 一、 **RFM 指標**：將使用三種不同的方法，分別為五等均分法、Bob Stone 法、自定義法，建立不同的評分機制，分析各評分法的優劣好壞，設計出最適合信用卡產業的評分機制。
- 二、 **CAI 指標**：先計算各個客戶的 CAI 值，並且結合該客戶的 Frequency 創建出 ARFM，分析四個象限的客戶並且定義第一象限為**活躍忠誠顧客群**、第二象限為**活躍潛力顧客群**、第三象限為**不活躍顧客群**、第四象限為**不活躍重度使用顧客群**。接下來將針對不同族群設計不同方案。
- 三、 **CRI 指標**：計算各個顧客之 CRI 值，投入行銷資源挽回 CRI 值較大、且 RFM 指標較優的顧客。
- 四、 **集群分析**：利用 SPSS 統計軟體中的 K-means 法得到三種族群，分別為**忠實客戶群**、**流失群**、**新興客戶群**，接者再使用 ANOVA 與事後檢定檢驗各項細項之顯著性與關聯性。
- 五、 **購物籃分析**：使用相關係數矩陣和因素分析來找出 Top20 購物籃和五種不同購物籃，針對同個購物籃之商品設計行銷策略以提升營收。

壹、 前言

隨著時代的推演，現在已邁向貨幣極度寬鬆的世代，許多央行利率都到達歷史新低。這使得傳統銀行的主要獲利來源從「利差收入」轉向「手續費收入」。而在手續費收入當中，很重要的一部分是信用卡。往往一家銀行就有許多不同的信用卡等級，且不同信用卡等級所要針對的客群又不一樣。舉例而言，台新的 Richart 卡就打著年輕人的市場、花旗與國內第一大電商平台 Pchome 推出聯名卡、Momo 與自家集團公司富邦推出聯名卡、國泰世華與長榮推出聯名卡等等，上述例子可以看出當今信用卡有「專業分工」之趨勢，不同卡有不同的使用時機，因此如何抓準自身商業優勢就變得十分重要。過往只用一張卡的觀念已過時，根據 2019 年聯徵中心的資料，台灣人均已有 4.3 張信用卡。

隨著金融科技的進步與金融法規的允許，給予了金融機構使用顧客資訊，透過顧客資訊的利用，能夠深入了解顧客背後的潛在價值，也可以針對此來做客製化的行銷與服務。因此本文的目標在於利用此資料集進行精準行銷。確切方法的步驟有三，一為資料概況的描述且進行顧客價值靜態與動態的分析、二為進行顧客分群，劃分出市場區隔、三為用觀察產品間的相關性與設計出購物籃。藉由上述三步驟，期盼設計出適合此家信用卡公司的行銷策略。

貳、 資料概況描述

本資料記錄了 100 位客戶於 2010/12/1 至 2012/11/30 間共 7764 筆交易，其由三個資料檔所組成，詳細內容如表 2.1 所示。

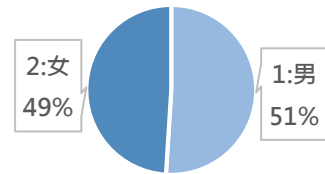
客戶資料檔	客戶 ID、年齡、生日、居住地、教育程度、性別、婚姻狀況、職業、辦第一張信用卡的時間
信用卡資料檔	信用卡 ID、客戶 ID、信用卡開卡日、信用卡到期日、信用額度、卡等
信用卡交易紀錄檔	交易序號、信用卡 ID、客戶 ID、刷卡日期、刷卡產品、產業分類、刷卡類型、刷卡金額、刷卡地點

▲表 2-1、信用卡資料檔內容

一、客戶資料檔：

一、性別：男女比例差不多

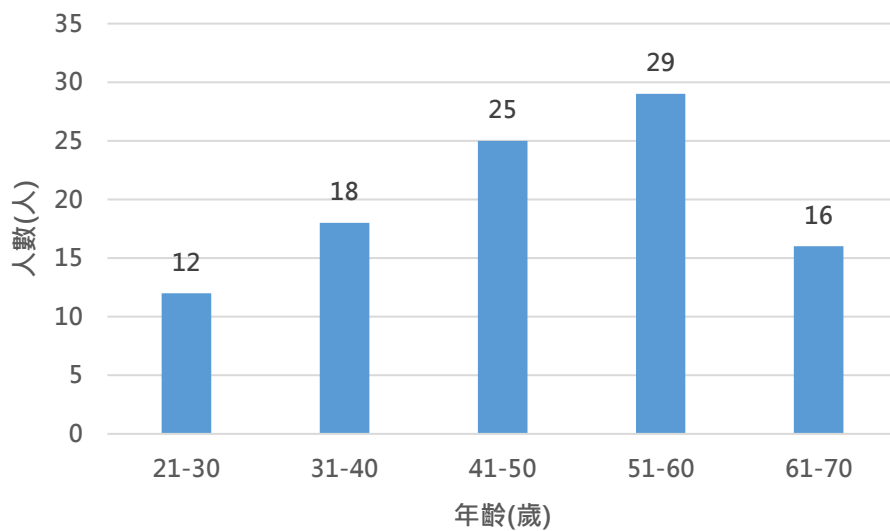
本資料集之 100 位客戶中，男性占 51 位，女性占 49 位，平均年齡為 47.03 歲。



▲圖 2-1、男女客戶比例圓餅圖

二、年齡：以中年為大宗

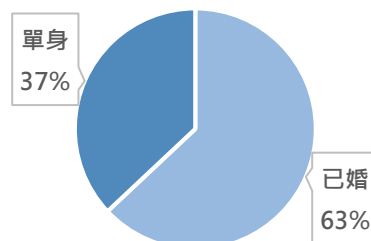
本資料集之年齡以 2012 年為基準來計算，分布如圖 2-2 所示。以 51 至 60 歲為大宗，且在此之前有年齡越高人數越多的趨勢。



▲圖 2-2、客戶年齡分布直條圖

三、婚姻狀況：以已婚為主

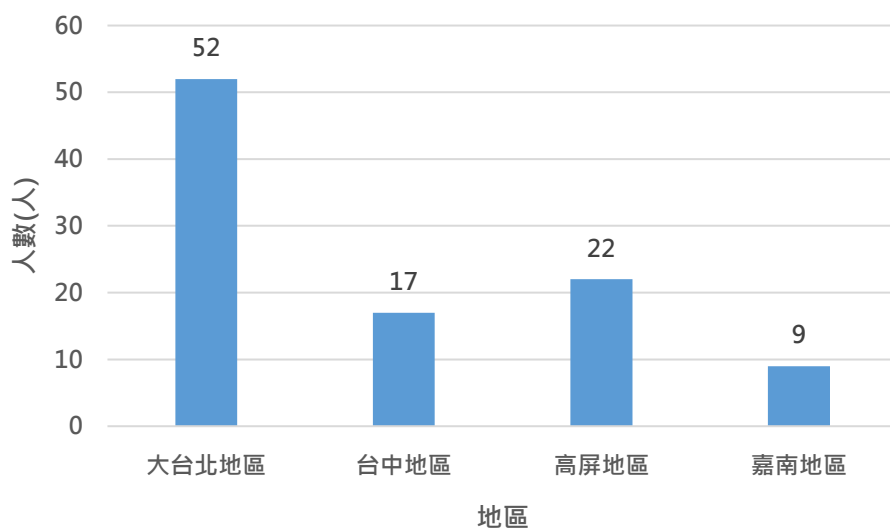
本資料集之 100 位客戶中，單身占 37 位，已婚占 63 位，可能是因為主要客戶為中年人。



▲圖 2-3、婚姻狀況比例圓餅圖

四、居住地：大台北地區客戶最多

本資料僅將客戶粗分為四個區域，詳細分布情形如圖 2-4 所示，由圖 2-4 可視，**大台北地區人數最多**。

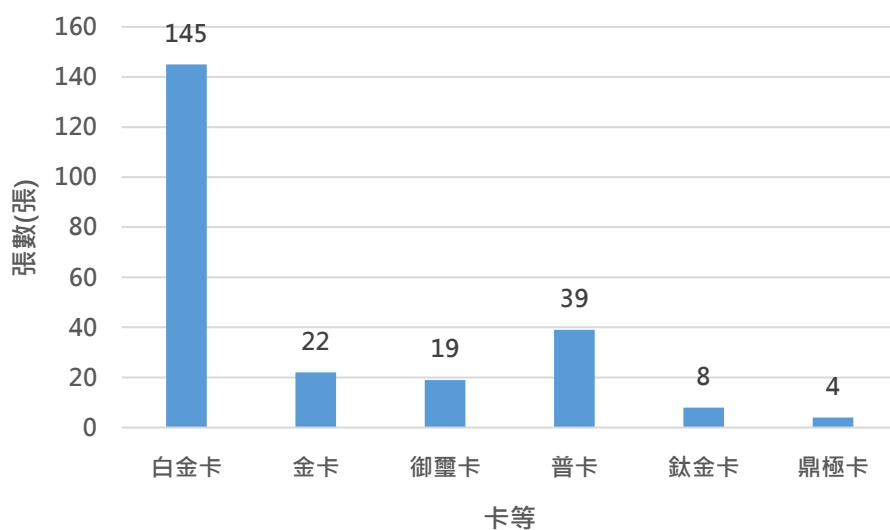


▲圖 2-4、客戶地區分布直條圖

二、 信用卡資料檔：

1. 卡種：以**白金卡**為主

本資料中，客戶所持有的信用卡等級可能重複，詳細如圖 2-5 所示。



▲圖 2-5、客戶所持有信用卡卡等分分布直條圖

參、顧客靜態價值分析—RFM 法

了解完資料概況後，接下來將進一步利用「信用卡交易紀錄檔」來衡量每位顧客對於企業之價值，實際做法為使用 RFM 指標分析。RFM 指標使用三大指標來衡量客戶價值：**上次購買日期(Recency)**、**購買頻率(Frequency)**、**平均購買金額(Monetary Value)**。雖然三大指標定義十分明確，但是在權重方面，卻可以因情況而調整，此處將使用三種不同權重設定，分別為五等分法、Bob Stone 法、以及自定義法來分析顧客價值。

一、五等均分法

五等均分法的給分機制是令排序在前 20% 的客戶獲得 5 分，次 20% 的客戶獲得 4 分，其餘以此類推。表 3-1 將列出排名前 10 名客戶各項指標原始值與得分。

客戶 ID	Recency	Frequency	Monetary Value	R	F	M	RFM Score	RANK
6118	1	250	968004	5	5	5	15	1
9051	1	205	265085	5	5	5	15	1
13687	3	399	1153973	5	5	5	15	1
15195	3	209	711772	5	5	5	15	1
17586	1	237	226459	5	5	5	15	1
4210	4	196	337601	4	5	5	14	6
5287	4	264	914311	4	5	5	14	6
6143	1	64	300735	5	4	5	14	6
6256	4	139	243291	4	5	5	14	6
15353	1	115	154864	5	5	4	14	6

▲表 3-1、五等均分法總分排名前 10 名客戶各項指標原始值與得分

從表 3-1 可以得知，許多客戶有同分的情形發生，且簡單地簡單將三種衡量指標視為一樣重要，並無針對產業而設計不同的評分制度，因此以下將改良。

二、Bob Stone 法

根據五等均分法所產生的 RFM 分數，實質上屬於順序尺度，不適合進行

量化分析。因此 Bob Stone 法針對此缺失改良，將最近購買日期的 R 分數，以最近三個月為一個單位，等比例切割並且依序給予 24 分，三到六個月給予 12 分，接者 6 分 3 分 0 分、購買頻率高低的 F 分數，設定為實際購買次數的 4 倍。M 分數設定為買金額的 10%，再無條件進入為整數，超過 9 分者以 9 分計算。表 3-2 展示了以此法排序之前 10 名。

客戶 ID	Recency	Frequency	Monetary Value	R	F	M	RFM Score	RANK
605	3	955	92259	24	3820	9	3853	1
3044	5	661	218935	24	2644	9	2677	2
13687	3	399	1153973	24	1596	9	1629	3
5287	4	264	914311	24	1056	9	1089	4
6118	1	250	968004	24	1000	9	1033	5
17586	1	237	226459	24	948	9	981	6
15195	3	209	711772	24	836	9	869	7
9051	1	205	265085	24	820	9	853	8
19639	5	201	326970	24	804	9	837	9
4210	4	196	337601	24	784	9	817	10

▲表 3-2、Bob Stone 法總分排名前 10 名客戶各項指標原始值與得分

從表 3-2 可以看出，同分情形減少，且也有針對不同產業展現出其權重，以此為例，其權重大小依序為：購買頻率、最近購買日、最後才為購買金額。

三、 自定義法

自定義法為筆者自己替信用卡公司設計出的評分機制，詳細給分機制與原因如表 3-3、排序前 10 名如表 3-4。

構面	給分機制	權重	原因
Recency	0-30 days 20 pts 31-60 days 18 pts 61-90 days 12 pts 91-120 days 8 pts 121-150 days 4 pts >150 days 0 pts	中	Recency ：考慮到信用卡的特性，一般而言一個月或兩個月基本上都會有一筆定期支出（如電話費）等，因此給那些穩定用戶 20 分/18 分，剩下都 4 分往下扣。

構面	給分機制	權重	原因
Frequency	購買次數*0.5	高	由於信用卡公司最重視刷卡次數，因此此處直接把購買次數乘上 0.5，而這將直接讓 Frequency 影響程度最大，但不乘上 4 的原因是因為怕 Frequency 過度 dominant，忽略其他兩要素的影響。
Monetary Value	>90% 10pts 70%-90% 8pts 40%-70% 6pts 20%-40% 4pts <20% 2pts	低	金額在信用卡公司佔的比重較低，但仍要作出小區別，因此採用權重式五等分法的方式評分。

▲表 3-3、自定義法給分機制與原因

客戶 ID	Recency	Frequency	Monetary Value	R	F	M	RFM Score	RANK
605	3	955	92259	20	477.5	6	503.5	1
3044	5	661	218935	20	330.5	8	358.5	2
13687	3	399	1153973	20	199.5	10	229.5	3
5287	4	264	914311	20	132	10	162	4
6118	1	250	968004	20	125	10	155	5
17586	1	237	226459	20	118.5	8	146.5	6
15195	3	209	711772	20	104.5	10	134.5	7
9051	1	205	265085	20	102.5	8	130.5	8
19639	5	201	326970	20	100.5	8	128.5	9
4210	4	196	337601	20	98	8	126	10

▲表 3-4、自定義法總分排名前 10 名客戶各項指標原始值與得分

對比表 3-3 與表 3-4，前 10 名的名單一致，可以發現經過調整，購買頻率仍然有巨大的影響力，但自定義的方法能夠改善名單中後半部的排名，不會讓單純 F 指標很高、但 R 與 M 指標很低的客戶仍拿到很高的排名。

四、 小結

透過 RFM 分析，可以大致捕捉客戶靜態的貢獻，對於貢獻高的客戶可以將他們設為 VIP，促使他們提高對於信用卡公司的忠誠，然而 RFM 卻無法捕捉顧客的動態行為，因此下部份將介紹兩個指標來增進客製化行銷的成功性。

肆、顧客動態價值分析—CAI 法、CRI 法

在段落三，筆者使用 RFM 來分析顧客靜態價值分析，而在此段，將使用活躍性指標（Customer Activity Index, CAI）與交易穩定度指標（Customer Reliability Index, CRI）指標來捕捉客戶動態活動，增進 RFM 模型的不足。

一、CAI

CAI 的公式為： $CAI = \frac{MLE - WMLE}{MLE} \times 100\%$ 。其精神為：首先計算客戶的兩次購買間隔的時間，稱為 Interpurchase Time，得到平均購買期間的平均 MLE。接者，把較近一次的購買予以較高的權重，較早之前的購買予以較低的權重，計算出 WMLE。比較這 MLE 與 WMLE 大小，即可得到該客戶的活躍性。

值得留意的一點是該概念無法計算只購買一次的客戶（由於無 Interpurchase Time），正好本資料集剛好有一位客戶只購買了一次，故無法計算 CAI 值，人數從 100 降低為 99。表 4-1 整理了 CAI 大小含意，表 4-2 整理了本資料集的三種族群的概況。

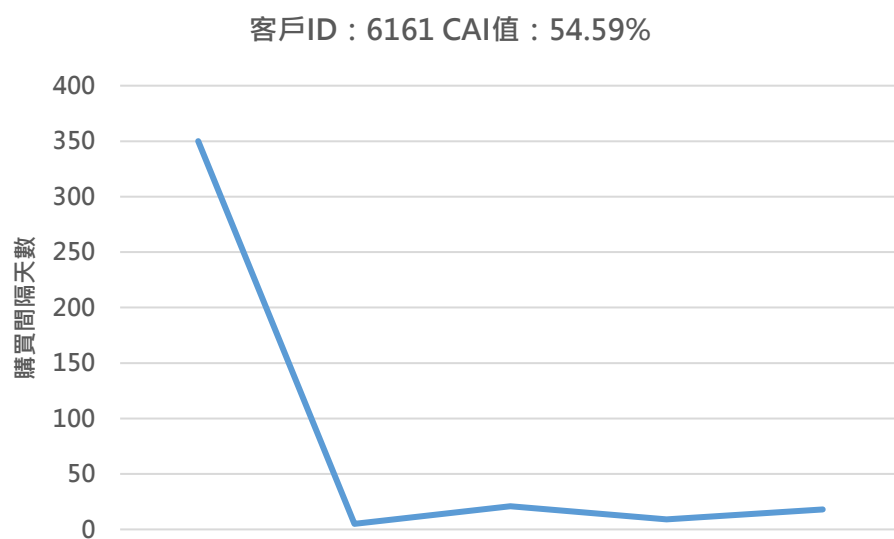
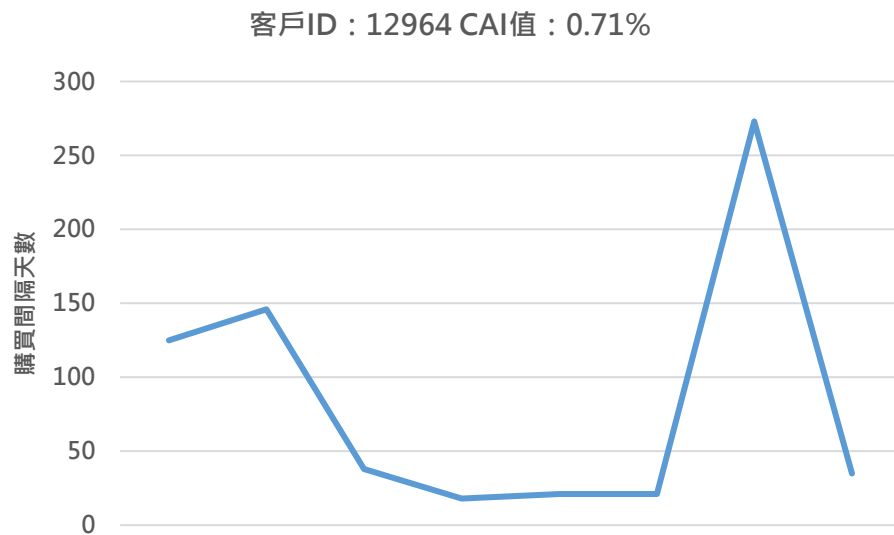
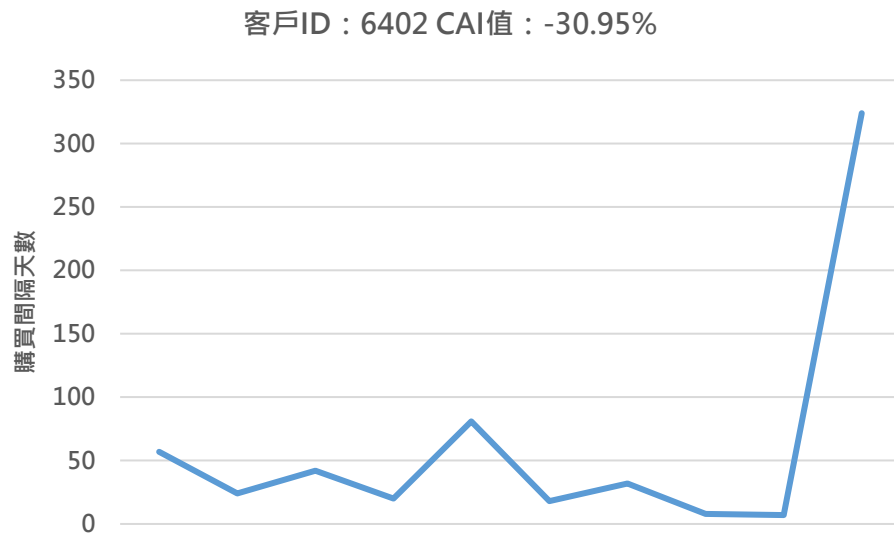
CAI 值	購買期間的算術平均數與加權平均數	代表含意
CAI \approx 0	大致相等	客戶的購買行為非常規律
CAI $>$ 0	大於	客戶的購買行為越來越活躍
CAI $<$ 0	小於	客戶的購買行為越來越低落

▲表 4-1、CAI 值大小不同所代表的不同含意

CAI 群別	客戶人數	人數比例	消費日平均刷卡金額	平均刷卡間隔天數
漸趨活躍群	20	約 20%	2404.845671	39.79029395
穩定刷卡群	59	約 60%	2288.621211	25.35704462
漸趨靜止型	20	約 20%	2543.900679	46.17102092

▲表 4-2、不同 CAI 群別間不同的消費特性與人數

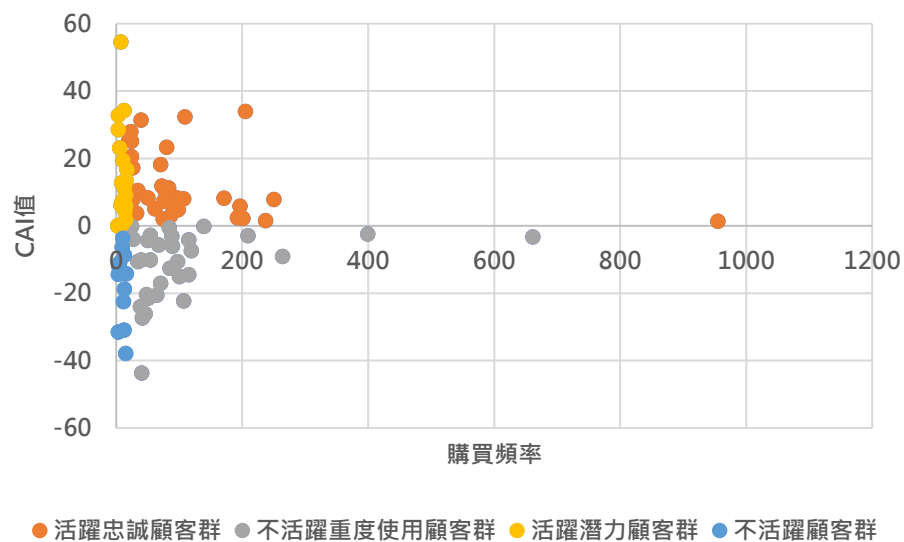
從表 4-2 可以看出其實三種族群的消費日平均刷卡金額並無太大差別，可以得知金額並不是區分客戶活躍性的重要指標。



▲ 圖 4-1、購買間隔天數圖 (由上至下為漸趨靜止型、穩定刷卡群、漸趨活躍群)

從圖 4-1 可以明顯看出三種不同族群的差異，漸趨靜止型在最後的購買件個天數大幅上升，而穩定刷卡型的圖天數較起伏不定，但平均而言維持規律變化，而漸趨活躍群的購買間隔天數大幅下降。

CAI 值可再結合購買金額或購買頻率等 RFM 指標，進行顧客價值結構分析，稱為 ARFM 模型，如圖 4-2。由於本次資料為信用卡公司之資料，因此筆者所納入的指標為購買頻率，並且畫分四個象限，x 軸的上下為 CAI 大於或小於 0，y 軸的左右為購買頻率大於或小於 20。購買頻率門檻設計如此設計是因為取百分比 20%，以符合 80/20 法則。



▲圖 4-2、顧客活躍性分布圖

分析圖 4-2，可得位於第一象限的客戶為活躍忠誠型（橘色部分），最應受到企業禮遇，設定為 VIP，規劃尊榮顧客的企劃。同時最重要的是，要極力拉攏第四象限的客戶（灰色部分），因為他們代表 RFM 分數高，但近期卻漸漸不使用信用卡，公司要極力拉攏那些客戶。可採取的作法為打電話詢問逐一詢問、寄信給予提醒提供獨家刷卡優惠，以免流失公司珍貴的客戶。

二、 CRI

CRI 為測量交易資料的穩定度，其概念與 CAI 相比，最大的不同是加入了「風險」的概念，風險的概念是從標準差而得。有了標準差，就能看出資料的

離散程度，也就是說當資料離散程度小，決策風險小；相反的，資料離散程度大，決策程度大。

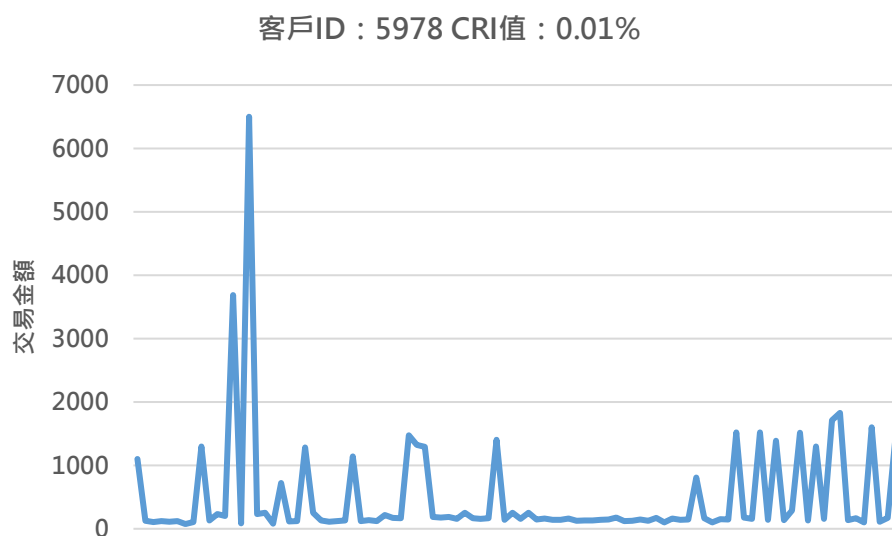
CRI 的公式為： $CRI = \left(1 - \frac{|BE-IE|}{|GE-IE|}\right) \times 100\%$ ，其中 BE 為貝氏估計值，IE 為個人估計值，GE 為群體估計值。有群體估計值即暗示了要將顧客分群，這裡筆者將顧客以性別和教育程度分為了 12 群，但因為有兩群為空，且有一個顧客自成一群，故無法計算出 CRI 值，實際上只有 9 群。

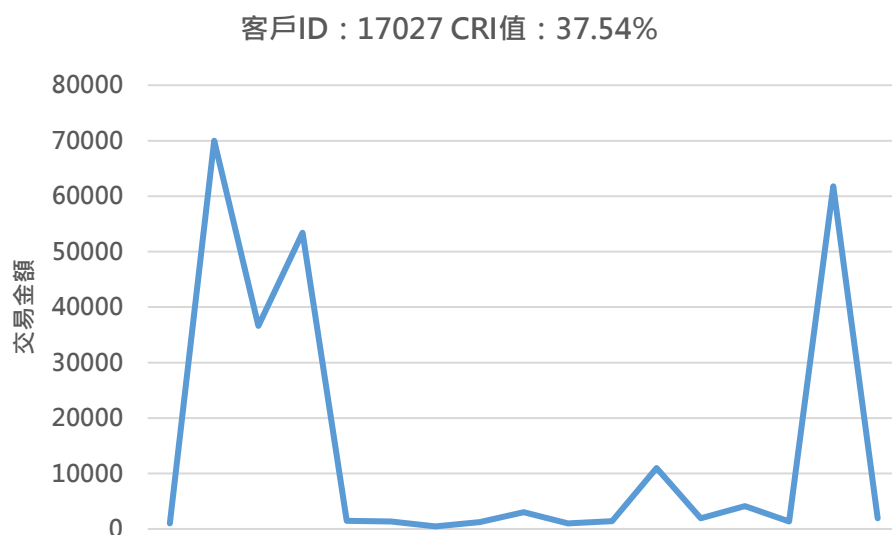
以實際例子來說明，若估計顯示 BE 較接近 GE 而非 IE，代表 GE 穩定度高於 IE。表示個人的交易資料相對於群體而言，尚未形成一個穩定的型態，客戶的交易穩定度較低。從公式可知，若 BE 與 IE 的差距較小，將得到 CRI 的值較大，代表個人交易資料的型態愈穩定。相關整理如表 4-3。

CRI 值	貝氏估計較偏向	代表含意
較大	群體估計	穩定性低
較小	個人估計	穩定性高

▲表 4-3、CRI 值大小之涵義

隨機挑選檢視高 CRI 值（穩定性低）與低 CRI 值（穩定性高）的每項消費，可以發現低 CRI 值的客戶其刷卡金額再 0~7000 以內震盪，消費金額屬於穩定類型，且其刷卡筆數高達 96 筆，而高 CRI 值的客戶其刷卡金額在 0~70000 中震盪，確認 CRI 值確實可以捕捉客戶的刷卡金額穩定性，詳圖如圖 4-3。





▲圖 4-3、不同 CRI 值交易金額變化比較

另外，深究 CRI 值前五高的客戶，發現其教育程度有八成為學士，且刷卡次數皆小於平均數（77.84），可能是因為如此才造成貝氏估計較仰賴群體估計值。並且再將顧客以性別和學歷區分，詳細資料參見表 4-4 與表 4-5。

客戶 ID	性別	教育程度	刷卡次數	CRI 值
17027	男	學士	17	37.53733
13802	女	學士	15	36.4148
13196	男	學士	13	35.9352
4819	女	專科	39	27.29876
5425	男	學士	61	19.75598

▲表 4-4、CRI 值前五高的客戶資料

女&學士以上	男&學士以上
N=21	N=35
CRI = 3.47	CRI = 3.849
女&學士以下(不包含)	男&學士以下(不包含)
N=26	N=16
CRI = 2.001	CRI = 4.598

▲表 4-5、學歷與性別配合 CRI 指標之象限圖

三、 小結

信用卡公司容易面對到靜止戶的問題，**為了避免靜止戶的發生，筆者推薦銀行可以設定一個門檻值，作為靜止戶預警系統的判斷標準。**只要客戶的 CAI 低於該門檻，銀行即可針對這些客戶進行個人行銷，如寄送電子信提醒最新活動，又或給予他們更高的刷卡回饋。相反的，也可以**針對漸趨活躍型的客戶，推行一個「尊榮客戶獨享方案」**，促使他們對於信用卡的黏著度更高。

另外針對客戶穩定度，觀察表 4-5，可以發現**第四象限最不穩定**，而**第三象限最為穩定**。應此可以思考信用卡公司是如何讓**第三象限最為穩定**，分析自身優勢，並且依情況改善套用到第四象限之族群。目標思考在於如何第四象限之購買頻率，提高品牌忠誠度，具體做法之一可以為**發送問卷，詢問該族群之偏好**。

伍、顧客集群分析

本段的目的是要建構顧客的市場區隔，因為在不知道客戶的刷卡習性時就用相同行銷手法會讓效果打折扣。而此處第一階段先將顧客分群，具體做法為使用 SPSS 裡的 K-means，再投入的指標：R 指標、F 指標、M 指標、CRI 指標、CAI 指標。筆者在分群前前移除了兩位客戶，分別為 2027（因為其在算 CRI 時自成一群，故算不出其值）與 6687（由於其只買一筆，故算不出 CAI）。

一、建立市場區隔

此處我設定的群的數目為五群，原因為分三群與四群皆無法有效地分出兩群以上的人，其結果皆集中於一大群。最後結果為表 5-1。

	1	2	3	4	5
R(Recency)	20	20	20	19	12
F(Frequency)	330.5	116.2	477.5	38.2	7.2
M(Monetary Value)	8	9	6	7	4
CRI	0.024695	0.418054	0.001105	2.975688	4.604744
CAI	-3.32415	4.747297	1.368443	-4.6281	6.176954

▲表 5-1、最終集群中心

觀察表 5-1 可知：五群中大致上只能看出三群（另外兩群只有一人），故此分析將基於表 5-1 的結果只對這三群進行特寫檢定，結果如表 5-2、表 5-3。

群	特色	命名
群 2	<ul style="list-style-type: none">● F 與 M 的分數最高，在 RFM 的評分上屬於高分群● CAI 值又是正的，代表該群仍屬於活躍群● CRI 值為最低，代表該群消費行為穩定	忠實客戶群
群 4	<ul style="list-style-type: none">● RFM 皆屬中庸的成績● CAI 值為三群中唯一負數，代表該群可能為正在流失之客戶	流失群
群 5	<ul style="list-style-type: none">● RFM 各個值皆為最低● CAI 值最高，表該群漸趨活躍● CRI 值最大代表該群消費行為不穩定	新興客戶群

▲表 5-2、集群命名原因

全體	組別	1	忠實客戶群	3	流失群	新興客戶群	卡方值 (p 值)
	百分比	1.02%	10.20%	1.02%	41.84%	45.92%	
性別	男	100%	60%	100%	46.30%	53.30%	0.616
	女	0%	40%	0%	53.70%	46.70%	
年齡	21-30	0%	10%	0%	19.50%	6.70%	0.53
	31-40	100%	20%	0%	22%	13.30%	
	41-50	0%	30%	100%	22%	24.40%	
	51-60	0%	30%	0%	26.80%	31.10%	
	>61	0%	10%	0%	9.80%	24.40%	
教育程度	初中及初中以下	0%	0%	0%	2.40%	6.70%	0.931
	專科/高中	0%	40%	100%	36.60%	40%	
	學士	100%	40%	0%	51.20%	40%	
	研究所以上	0%	20%	0%	9.80%	13.30%	
婚姻狀況	已婚	100%	40%	100%	51.20%	26.70%	0.15
	單身	0%	60%	0%	48.80%	73.30%	
居住地	大台北地區	100%	50%	100%	46.30%	55.60%	0.991
	台中地區	0%	10%	0%	22%	15.60%	
	高屏地區	0%	30%	0%	22%	20%	
	嘉南地區	0%	10%	0%	9.80%	8.90%	

▲表 5-3、顧客價值群之條件機率與卡方檢定

第二階段再來進行 ANOVA 分析與事後檢定。在分好群的基礎下，接下來再根據年齡、居住地、教育程度、性別、婚姻狀況、與取完對數後的三個獲利指標（一：平均刷卡金額、二：刷卡總金額、三：RFM 分數）進行 ANOVA 分析。取對數的目的在於使資料常態分佈化。在信心水準 $\alpha = 0.05$ 之下，可能由於樣本數不足的關係，僅有：第四群在婚姻狀況和年齡上與獲利指標 1&2 有顯著差異和第五群在性別上與獲利指標 1&2 有顯著差異，因此我們可以在第四群中，針對婚姻狀況與年齡替顧客分群，給予不同樣態的顧客進行不同的行銷策略、以及在第五群針對兩性採取不同的行銷策略。詳見表 5-4 與表 5-5。

集群四		平方和	自由度	均方	F	顯著性
獲利指標 1* 年齡	群組之間	6.9128	4	1.7282	3.1212	0.0265
	在群組內	19.9331	36	0.5537		
	總計	26.8459	40			

集群四		平方和	自由度	均方	F	顯著性
獲利指標 2* 年齡	群組之間	5.9881	4	1.4970	2.7044	0.0456
	在群組內	19.9281	36	0.5536		
	總計	25.9162	40			
獲利指標 1* 婚姻狀況	群組之間	4.3651	1	4.3651	7.5725	0.0089
	在群組內	22.4808	39	0.5764		
	總計	26.8459	40			
獲利指標 1* 婚姻狀況	群組之間	3.3491	1	3.3491	5.7878	0.0210
	在群組內	22.5671	39	0.5786		
	總計	25.9162	40			

▲表 5-4、集群四與年齡和婚姻狀況

集群五		平方和	自由度	均方	F	顯著性
獲利指標 1* 性別	群組之間	8.612	1	8.6120	5.4950	0.0238
	在群組內	67.391	43	1.5672		
	總計	76.003	44			
獲利指標 2* 性別	群組之間	4.083	1	4.0830	4.7262	0.0353
	在群組內	37.148	43	0.8639		
	總計	41.231	44			

▲表 5-5 群五與性別

最後針對年齡進行事後檢定，發現在第四群當中的獲利指標 1，21-30 歲與其他年齡層有顯著差異，因此可以特別研究如何針對 21-30 歲來進行信用卡的推銷，在獲利指標 2 當中也可以看出 21-30 歲與 41-60 歲有顯著差異，可以推斷在第四群當中 21-30 歲的顧客特別不同，相關資料整理在 5-6 和 5-7。

集群四		平均值差異 (I-J)	標準誤	顯著性	95% 信賴區間	
獲利指標一					下界	上界
21-30	31-40	-.84545*	0.36157	0.025	-1.5788	-0.1121
	41-50	-1.21617*	0.36157	0.002	-1.9495	-0.4829
	51-60	-.91752*	0.34576	0.012	-1.6188	-0.2163
	>61	-.93040*	0.45567	0.049	-1.8545	-0.0063

▲表 5-6、集群四獲利指標 1 與年齡

集群四		平均值差			95% 信賴區間	
獲利指標二		異 (I-J)	標準誤	顯著性	下界	上界
21-30	31-40	-0.699	0.362	0.061	-1.432	0.034
	41-50	-.99565*	0.362	0.009	-1.729	-0.262
	51-60	-1.03850*	0.346	0.005	-1.740	-0.337
	>61	-0.823	0.456	0.079	-1.747	0.101

▲表 5-7、集群四獲利指標 2 與年齡

二、 小結

總結此一小節，透過分群可以讓我們得知實際上是何者變數具有重大的影響力，另外，本研究將顯著水準設為 0.05，因此使得許多結果不顯著，但得到的是較為準確的結果。假設要找更多的差異性可以放寬顯著水準，但必須承擔更多風險。另外，事後檢定可以幫助分析到底是哪一些組之間有顯著差異。舉例來說，只知道年齡有顯著差異並無法得知是誰與誰差異。

陸、 購物籃分析

在前面三段利用 RFM、CAI、CRI、集群分析來進行顧客價值，但除了研究消費者的異質性以外，也可轉換角度，從消費者刷卡產品產業分類來進行分群。本段首先建構相關係數矩陣選出 Top20 購物籃，最後再利用因素分析將眾多品項分群。

一、 相關係數矩陣

為了得到 Top20 購物籃，首先要先有相關係數矩陣。相關係數矩陣的建構是先將顧客刷卡紀錄轉成 1 和 0。若是 1 即代表有刷卡過該產業，0 則相反)。接者觀察產業與產業之間的相關性。結果如表 6-1，表中顏色越深，代表其相關係數越高。表 6-2 整理了 Top20 的購物籃的搭配。

	2	3	5	6	7	8	9	10	11	12	13	14	15	16	X2
2	1.000														
3	0.357	1.000													
5	0.164	0.076	1.000												
6	0.063	-0.063	0.076	1.000											
7	-0.024	0.151	0.162	0.097	1.000										
8	0.003	0.149	-0.013	0.039	0.292	1.000									
9	0.092	0.161	0.069	0.001	0.347	0.316	1.000								
10	0.129	0.272	0.211	0.002	0.321	0.414	0.396	1.000							
11	0.086	0.205	0.066	0.045	0.284	0.259	0.379	0.298	1.000						
12	0.113	0.122	0.115	0.064	0.184	0.231	0.179	0.102	0.149	1.000					
13	0.144	0.118	0.132	0.118	0.102	0.273	0.331	0.286	0.178	0.145	1.000				
14	0.104	0.393	0.077	-0.035	0.230	0.473	0.276	0.268	0.177	0.110	0.260	1.000			
15	0.080	0.089	0.179	-0.018	0.383	0.368	0.319	0.322	0.259	0.118	0.318	0.400	1.000		
16	0.185	0.192	0.251	0.135	0.264	0.255	0.234	0.366	0.305	0.088	0.349	0.177	0.418	1.000	
X2	-0.052	0.076	-0.031	-0.080	0.045	-0.013	0.069	0.211	0.066	0.115	0.132	0.077	0.062	0.126	1.000

▲表 6-1、相關係數矩陣

其中表中的的欄位名稱依序為：02_學費/教育、03_保險、05_捐贈、06_公用事業、07_百貨、08_精品、09_餐飲、10_藥妝、11_3C 居家電信、12_量販超市、13_交通(含加值)、14_旅遊、15_休閒文化、16_無店舖、X2.中信錢加值。

#	購物籃組合		相關係數	#	購物籃組合		相關係數
1	旅遊	精品	0.4725	11	保險	學費/教育	0.3572
2	無店鋪	休閒文化	0.4182	12	無店鋪	交通(含加值)	0.3491
3	藥妝	精品	0.4140	13	餐飲	百貨	0.3465
4	休閒文化	旅遊	0.4001	14	交通(含加值)	餐飲	0.3306
5	藥妝	餐飲	0.3964	15	休閒文化	藥妝	0.3217
6	旅遊	保險	0.3933	16	藥妝	百貨	0.3206
7	休閒文化	百貨	0.3829	17	休閒文化	餐飲	0.3194
8	3C 居家電信	餐飲	0.3786	18	休閒文化	交通(含加值)	0.3184
9	休閒文化	精品	0.3675	19	餐飲	精品	0.3158
10	無店鋪	藥妝	0.3662	20	無店鋪	3C 居家電信	0.3048

▲表 6-2、Top20 購物籃

觀察 Top20 購物籃，第一名為旅遊與精品，可以感覺是有多餘財力的人才會去刷卡的組合，**推測有可能是消費者去旅遊時順便逛精品店，因此在規劃信用卡活動設計時，信用卡公司可以同時與旅行社和精品店合作，藉此機會回饋更多點數，也讓消費者會大幅提升其在旅遊時的消費。**第二名為無店鋪與休閒文化，無店鋪就是現在熟悉的網購，兩者種類也很雷同，可以得知透過相關係數所建構的矩陣雖然很直觀，但其仍有所貢獻，如排列出高低、以及發現人們直覺所想不到的組合。

二、 因素分析

上一段使用相關係數找出刷卡產品產業兩兩之間的關係，而因素分析的目標更深一層。其目的是縮減構面，將原來的多個變數，依其相關結構，縮減為少數幾個具有代表性的共同因素。通常共同性越高的題項，越適合進行因素分析。因素分析的優點就在於可以從大量產品的購買紀錄中，找到一組關聯性最高的產品，一起組成同一因素。還可以調整特徵值的下限，根據不同需求控制購物籃個數。從因素分數也可看出消費者對於該籃子的購買欲望高低。在本次研究將使用 SPSS 進行主成分分析來分出 5 群，分群結果如表 6-3。

	1	2	3	4	5
08_精品	0.772	-0.089	0.073	-0.093	0.049
15_休閒文化	0.641	0.378	-0.038	-0.019	-0.181
14_旅遊	0.639	-0.061	0.398	-0.059	-0.215
09_餐飲	0.62	0.102	0.06	0.154	0.154
07_百貨	0.6	0.175	-0.111	-0.037	0.18
10_藥妝	0.532	0.337	0.158	0.305	-0.049
11_3C 居家電信	0.491	0.132	0.087	0.133	0.228
05_捐贈	-0.017	0.695	0.111	-0.079	0.048
16_無店舖	0.351	0.67	0.071	0.136	0.017
13_交通(含加值)	0.365	0.405	0.088	0.2	0.095
03_保險	0.221	0.005	0.806	0.095	-0.037
02_學費/教育	-0.105	0.316	0.747	-0.09	0.156
X2.中信錢加值	0.019	0.051	-0.025	0.903	0.012
12_量販超市	0.243	-0.129	0.19	0.216	0.713
06_公用事業	-0.016	0.297	-0.144	-0.315	0.611

▲表 6-3、因素分析結果

將因素分析結果與 Top20 購物籃擺在一起對照，可發現結果大致上吻合，但值得注意的是因素分析的結果沒有將無店舖與休閒文化放在同一個購物籃（這兩者由相關係數計算出來的結果之數值為第二高）。另外特別的是 X2 項自成一群，然回顧到相關係數矩陣，該項與其他項相關係數極低，會有自成一群的結果不太意外。

三、 小結

比較因素分析與 Top20 購物籃，其前提皆為要求在同個市場區隔，即顧客消費習性要相同。舉例來說，根據消費者行為理論，消費者對於產品的需求可以分為兩種類型，實用導向型與享樂導向型。在上一段的集群分析中已完成市場區隔，在設計與刺激信用卡客戶刷卡的同時，要先區隔出哪類型的人屬於哪種人。

柒、 結語

回顧前言，此報告的目的在於替企業針對不同的客戶來提供不同的信用卡設計，期盼達到提高市場佔比率之目的。在段落五中，已將這份報告的族群區分為忠實客戶群、新興客戶群、流失群。由於前兩者族群的狀況較為樂觀，因此下文將針對**流失群**來設計信用卡，企圖挽回那群客戶，提高企業營收。

在流失群當中，其人數佔據整體顧客 41.84%，RFM 指標所得分數屬於中庸，且在段落五集群分析中顯示，平均刷卡金額與刷卡總金額在**年齡與婚姻狀況**有顯著，因此下文將針對這兩點來進攻。

整個流失群的平均年齡為 46 歲整，而單身比例為 48.8%，與整體單身比例 37%差了超過 10%。另外流失群當中年齡為在 21-30 歲的族群平均刷卡金額為 1564.32 元，與整體流失群的平均刷卡金額 3063.81 元有近 2 倍之差。此結果又可呼應到段落五做的事後檢定：21-30 歲與其他年齡層有顯著不同。為此，可延伸有兩個問題：**要如何提高單身族群刷卡誘因**以及**要如何吸引年輕人**？

針對問題一，可以回顧段落六中的 Top20 購物籃的第一名：旅遊與精品。首先，**單身族群沒有家庭的束縛，推測較能去國外旅行**，此外，**由於單身族群不用負擔家庭的經濟重擔，較有經濟能力去購買精品**，因此推薦信用卡公司可以與旅行社和精品店合作。面對旅行社，可以找安排國外旅遊居多的旅行社，而針對精品店，可以找常出現在國外經典景點中的精品店，如此雙管齊下，可以藉此捕捉回流失群。

針對問題二的解法為：**設計一個年輕人較容易上手的信用卡**，並且尋找一些年輕人可能會較感興趣的東西合作，如現在流行的網購來攻佔年輕人的市場。此外，與電商平台推行聯名可以帶動不只年輕人的年齡層，更讓拉動整個流失群。

基於以上理由，建議信用卡公司可以採納前面所提到的兩個方案，以捕捉流失的客戶群。