

【107-2 大數據行銷課程個人學期作業】

信用卡公司資料庫數據分析

Building RFM Codes into Database

2019/06/17

生傳三 BXXXXX 李XX

目錄、

一、資料處理過程.....	
1. RFM 基礎數據整理	
2. 五等比例法，最近購買日 (Recency, R)	
3. 五等比例法，購買頻率 (Frequency, F)	
4. 五等比例法，購買金額 (Monetary Value, M)	
5. 五等比例法，RFM 資料統整	
6. Bob stone，最近購買日 (Recency, R)	
7. Bob stone，購買頻率 (Frequency, F)	
8. Bob stone，購買金額 (Monetary Value, M)	
9. Bob stone，RFM 資料統整.....	
二、資料計算結果.....	
三、資料分析與應用.....	
(一) RFM=011、001、100 各客戶群 MLE、WMLE、CAI 分析結果.....	
(二) 資料分析與決策建議.....	
主題一、分析 RFM=011「重要喚回客戶」消費者行為.....	
主題二、分析 RFM=001「重要挽留客戶」消費者行為.....	
主題三、分析 RFM=100「新客戶」消費者行為.....	

一、 資料處理過程

將資料庫運用於 RFM 模型以前，我首先觀察到尚未整理的信用卡資料中，工作表「客戶資料檔」與工作表「信用卡資料檔」之中，相較於客戶 ID 數量，信用卡 ID 數量要來得多，顯示出有單一信用卡公司顧客持有超過一張信用卡的情形出現（例如：客戶 ID 編號 89 號的顧客同時持有與使用信用卡 ID 編號 8905 與 14554 的兩張信用卡）。因 RFM 模型主要作用於協助企業定義顧客價值並發展因應的行銷策略，我決意以人為本，討論單一客戶的顧客價值，故後續討論均將同一客戶所使用的不同信用卡資料合併視為同一類型的消費型態。

資料整理步驟紀錄如下：

- (1.) 將信用卡資料 Excel 檔案中的工作表「信用卡交易紀錄檔」資料全選，插入「樞紐分析表」，並以「客戶 ID（由小到大排序）」、「刷卡日期（最大值）」、「刷卡日期（計數）」、以及「刷卡金額（加總）」為其列標籤，樞紐分析表部分分析結果如下方表 1 所示。

客戶 ID	最大值 - 刷卡日期	計數 - 刷卡日期	加總 - 刷卡金額
89	41225	85	150681
106	41236	75	90192
131	40843	16	69558
179	40954	4	9182
261	40890	13	19815

表 1、樞紐分析表部分分析結果

- (2.) 整理 Recency（最近購買日）資料（五等比例法）：以分析日期 2019/06/10 數值 43,626 減去樞紐分析表中的「刷卡日期（最大值）」，再減去數值 2,300 使數值統一變小，以利驗算與檢查，「刷卡日期（最大值）」數值整理後之部分分析結果如下方表 2 所示。接著，將資料全選，並於「排序與篩選」功能中的自訂排序中設定以「43,626 減 B 列數值再減 2,300」此列數值由小到大排列。並以五等比例法（以購買次數分類）改善五等均分法之謬誤。其中，數值介於 83 與 85 之間者，R 值給分 5 分；數值介於 86 與 87 之間者，R 值給分 4 分；數值介於 90 與 105 之間者，R 值給分 3 分；數值介於 106 與 199 之間者，R 值給分 2 分；數值介於 206 與 598 之間者，R 值給分 1 分，R 值給分（客戶 ID 數值由大到小排列）部分結果如下方表 3 所示。

客戶 ID	最大值 - 刷卡日期	43,626 減 B 列數值再減 2,300
89	41225	101
106	41236	90
131	40843	483
179	40954	372
261	40890	436

表 2、「刷卡日期（最大值）」數值整理後之部分分析結果

客戶 ID	43,626 減 B 列數值再減 2,300	R 值給分
89	101	3
106	90	3
131	483	1
179	372	1
261	436	1

表 3、R 值給分（客戶 ID 數值由大到小排列）部分結果

(3.) 整理 Frequency（購買頻率）資料（五等比例法）：「刷卡日期（計數）」數值相等於累積購買次數，將數值由大到小排列，並以五等比例法（以購買次數分類）改善五等均分法之謬誤。其中，數值介於 955 與 100 之間者，F 值給分 5 分；數值介於 98 與 64 之間者，F 值給分 4 分；數值介於 61 與 26 之間者，F 值給分 3 分；數值介於 24 與 12 之間者，F 值給分 2 分；數值介於 11 與 1 之間者，F 值給分 1 分，F 值給分（客戶 ID 數值由大到小排列）部分結果如下方表 4 所示。

客戶 ID	累積交易次數	F 值給分
89	85	4
106	75	4
131	16	2
179	4	1
261	13	2

表 4、F 值給分（客戶 ID 數值由大到小排列）部分結果

(4.) 整理 Monetary Value（累積購買金額）資料（五等比例法）：將「刷卡金額（加總）」數值由大到小排列。並以五等比例法（以購買次數分類）改善五等均分法之謬誤。其中，數值介於 1,153,973 與 219,911 之間者，M 值給分 5 分；數值介於 218,935 與 97,078 之間者，M 值給分 4 分；數值介於 95,782 與 45,866 之間者，M 值給分 3 分；數值介於 42,222 與 17,533 之間者，M 值給分 2 分；數值介於 17,257 與 583 之間者，M 值給分 1 分，M 值給分（客戶 ID 數值由大到小排列）部分結果如下方表 5 所示。

客戶 ID	累積消費金額	M 值給分
89	150681	4
106	90192	3
131	69558	3
179	9182	1
261	19815	2

表 5、M 值給分（客戶 ID 數值由大到小排列）部分結果

(5.) 將前述 R 值給分、F 值給分與 M 值給分加總得 RFM 模型總分，完成資訊處理，RFM 數值總分部分結果如下方表 6 所示（五等比例法）。

客戶 ID	R	F	M	RFM 加總
89	3	4	4	11
106	3	4	3	10
131	1	2	3	6
179	1	1	1	3
261	1	2	2	5

表 6、RFM 數值總分（客戶 ID 數值由大到小排列）部分結果

(6.) 統整前述以「客戶 ID（由小到大排序）」、「刷卡日期（最大值）」、「刷卡日期（計數）」、「刷卡金額（加總）」、「R 值給分」、「F 值給分」、「M 值給分」與「RFM 數值加總」為列標籤之資料庫資料，建立完整的 RFM 模型資料庫。RFM 模型資料庫部分結果如下方表 7 所示（五等比例法）。

客戶ID	刷卡日期（最大值）	最近一次消費時間間隔(時間-41326)	刷卡日期（計數）	刷卡金額（加總）	R值給分	F值給分	M值給分	RFM數值加總
89	41225	101	85	150681	3	4	4	11
106	41236	90	75	90192	3	4	3	10
131	40843	483	16	69558	1	2	3	6
179	40954	372	4	9182	1	1	1	3
261	40890	436	13	19815	1	2	2	5

表 7、RFM 模型資料庫（客戶 ID 數值由大到小排列）部分結果

(7.) 整理 Recency（最近購買日）資料（Bob Stone）：

以分析日期 2019/06/10 數值 43,626 減去樞紐分析表中的「刷卡日期（最大值）」，再減去數值 2,300 使數值統一變小，以利驗算與檢查，「刷卡日期（最大值）」數值整理後之部分分析結果如前述表 2 所示。接著，將資料全選，並於「排序與篩選」功能中的自訂排序中設定以「43,626 減 B 列數值再減 2,300」此列數值由小到大排列。數值介於 83 與 85 之間者，Bob Stone - R 值給分 24 分；數值介於 86 與 90 之間者，Bob Stone - R 值給分 12 分；數值介於 91 與 106 之間者，Bob Stone - R 值給分 6 分；數值介於 108 與 206 之間者，Bob Stone - R 值給分 3 分；數值大於 206 者，Bob Stone - R 值給分 0 分。Bob Stone - R 值給分（客戶 ID 數值由大到小排列）部分結果如下方表 8 所示。

客戶 ID	43,626 減 B 列數值再減 2,300	Bob Stone - R 值給分
89	101	6
106	90	12
131	483	0
179	372	0
261	436	0

表 8、Bob Stone - R 值給分（客戶 ID 數值由大到小排列）部分結果

(8.)整理 Frequency（購買頻率）資料（Bob Stone）：「刷卡日期（計數）」數值相等於累積購買次數，將數值由大到小排列，延續前述五等比例法 F 值計算結果再乘以 4 為 Bob Stone - F 值。Bob Stone - F 值給分（客戶 ID 數值由大到小排列）部分結果如下方表 9 所示。

客戶 ID	五等比例法 - F 值給分	Bob Stone - F 值給分
89	4	16
106	4	16
131	2	8
179	1	4
261	2	8

表 9、Bob Stone - F 值給分（客戶 ID 數值由大到小排列）部分結果

(9.)整理 Monetary Value（累積購買金額）資料（Bob Stone）：將「刷卡金額（加總）」數值由大到小排列，根據 10% of dollar purchase with a ceiling 9 points 調整為 Bob Stone - M 值。M 值給分（客戶 ID 數值由大到小排列）部分結果如下方表 10 所示。

客戶 ID	累積消費金額	Bob Stone - M 值給分
89	150681	15068
106	90192	9019
131	69558	6955
179	9182	918
261	19815	1981

表 10、M 值給分（客戶 ID 數值由大到小排列）部分結果

(10.)將前述 Bob Stone - R 值給分、Bob Stone - F 值給分與 Bob Stone - M 值給分加總得 Bob Stone - RFM 模型總分，完成資訊處理，Bob Stone - RFM 數值總分部分結果如下方表 11 所示（Bob Stone）。

客戶 ID	Bob Stone - R	Bob Stone - F	Bob Stone - M	Bob Stone - RFM 加總
89	6	16	15068	15090
106	12	16	9019	9047
131	0	8	6955	6963
179	0	4	918	922
261	0	8	1981	1989

表 11、Bob Stone - RFM 數值總分（客戶 ID 數值由大到小排列）部分結果

二、 資料計算結果

將 RFM 模型資料庫以「RFM 數值加總」由大到小排列部分結果如下方表 8 所示（五等比例法）。其中，RFM 數值加總後數值為 3 者，共計有 8 名客戶，佔總客戶數 8%；數值介於 4 與 6 之間者，共計有 20 名客戶，佔總客戶數 20%；數值介於 7 與 9 之間者，共計有 21 名客戶，佔總客戶數 21%；數值介於

10 與 12 之間者，共計有 29 名客戶，佔總客戶數 29%；數值介於 13 與 15 之間者，共計有 22 名客戶，佔總客戶數 22%。

客戶 ID	R 值給分	F 值給分	M 值給分	RFM 數值加總
6118	5	5	5	15
9051	5	5	5	15
13687	5	5	5	15
15195	5	5	5	15
17586	5	5	5	15

表 8、RFM 模型資料庫以「RFM 數值加總」由大到小排列部分結果

其次，於 R 值給分、F 值給分、M 值給分中，令各列大於 R 值給分均值者為 1、小於 R 值給分均值者為 0（R 值給分均值為 3.06）；大於 F 值給分均值者為 1、小於 F 值給分均值者為 0（F 值給分均值為 3.04）；大於 M 值給分均值者為 1、小於 M 值給分均值者為 0（M 值給分均值為 3.03），重新整理資料，並於 EXCEL 中操作使 RFM 值等於 R&F&M，整理結果部分如下方表 9 所示。

客戶 ID	R 值給分	F 值給分	M 值給分	RFM
89	0	1	1	011
106	0	1	0	010
131	0	0	0	000
179	0	0	0	000
261	0	0	0	000

表 9、RFM 模型資料庫以平均數衡量大小部分結果

各 RFM 數值對應的客戶類型：

- （1） RFM=111，重要價值客戶，最近有交易行為、累計交易次數多、累計交易金額高，企業應用心保持的優質客戶，共計有 21 名客戶。
- （2） RFM=011，重要喚回客戶，交易次數和累計交易金額均高，然近期較少交易，企業應喚回之，共計有 11 名客戶。
- （3） RFM=101，重要深耕客戶，交易金額大且最近有交易發生，需重點看待，共計有 4 名客戶。
- （4） RFM=001，重要挽留客戶，交易金額大，但是最近沒有交易而且交易次數較少，需要重點挽留，共計有 5 名客戶。
- （5） RFM=110，潛力客戶，交易次數多，且近期有發生交易，需要挖掘跟進，共計有 6 名客戶。
- （6） RFM=100，新客戶，最近剛開始交易的新用戶，屬於有推廣價值的客戶，共計有 10 名客戶。
- （7） RFM=010，一般維持客戶，交易次數多，但是都是小額消費，且最近沒有發生交易，屬於一般維持客戶，共計有 3 名客戶。
- （8） RFM=000，流失客戶，RFM 各項數值均低，屬於已經流失掉的客戶，共計有 40 名客戶。

三、資訊分析與應用

考量消費者使用信用卡的消費特性，攸關荷包佔有率的消費頻率在信用卡的分析上尤其重要。為具體化分析結果所能應用的行銷決策，我選擇針對信用卡公司應加強爭取的 RFM=011 重要喚回客戶、RFM=001 重要挽留客戶、RFM=100 新客戶此三群信用卡客戶做出以下 MLE、WMLE、CAI 分析與決策建議（為清楚表達此三群客戶之購買頻率差異，以下將分別以客戶 ID89、861、2027、4819、1611、3219 演示我於資料庫作業練習中的操作成果）。

備註：

① MAXIMUM LIKELIHOOD ESTIMATE（MLE、最大相似估計值）：

將單一客戶各段購買時間間隔加總計算平均。對於客戶 A 來說，距離現在最近的一次購買時間間隔能提供的訊息豐富度最為重要（重要性依序為最近、次近…最遠），然而，MLE 將時間間隔加總平均的計算結果，無法顯示較近一次時間間隔的較強作用。

② WEIGHTED MAXIMUM LIKELIHOOD ESTIMATE（WMLE、加權最大相似估計值）：將單一客戶各段購買時間間隔加權後再計算平均。

③ 比較 MLE 方法和 WMLE 方法：

對於消費時間分布平均的客戶來說，MLE 和 WMLE 計算結果無異；對於消費時間間隔越來越長的客戶來說，MLE 方法估算值小於 WMLE 估算值；對於消費時間間隔越來越短的客戶來說，MLE 方法估算值大於 WMLE 估算值。

④ Customer Activity Index（CAI、顧客活躍度分析）

透過 CAI 的運算可進行會員 ARFM 分析，包含會員 ARFM 輪廓基本分析（不活躍/隨機；流失/游離/忠誠；低價值/高價值），與會員 ARFM 之動態分析（CAI 低活躍性/高活躍性；R 最近購買日期遠/近；F 購買頻率低/高；M 購買金額低/高）。以信用卡公司 CAI 分析結果為例：自刷卡間隔天數看來，可分作穩定刷卡客戶、漸趨靜止客戶、漸趨活躍客戶……等等。

（一）RFM=011、001、100 各客戶群 MLE、WMLE、CAI 分析結果
針對分別代表 RFM=011 重要喚回客戶、RFM=001 重要挽留客戶、RFM=100 新客戶此三群信用卡客群之代表客戶之 MLE、WMLE、CAI 分析結果如表 10 所示：

客戶 ID	MLE 數值	WMLE 數值	CAI
89	10.5764706	9.24404925	0.12597977
861	7.5505618	7.97128589	-0.0557209
2027	17.9736842	22.4331984	-0.24811353
4819	17.9230769	18.7769231	-0.04763949

1611	42	21.1666667	0.49603167
3219	4.3846154	4.0276923	0.08140351

表 10、客戶 MLE、WMLE、CAI 部分分析結果

（二）資料分析與決策建議

主題一、分析 RFM=011「重要喚回客戶」消費者行為

針對 RFM=011 之信用卡客群代表，客戶 ID 89 與客戶 ID 861 顧客之 MLE、WMLE、CAI 分析結果，觀察其消費行為發現：客戶 ID 89 號顧客的 MLE 方法估算值稍大於 WMLE 估算值，代表其交易間隔有逐漸縮短、轉好的傾向；CAI 值稍大於 0，表示其為刷卡行為漸趨穩定的、可以稍微樂觀的看待為漸趨活躍型客戶。客戶 ID 861 號顧客的 MLE 方法估算值稍小於 WMLE 估算值，代表其交易間隔有逐漸拉長的傾向，且其 CAI 值小於 0，屬於漸趨靜止的顧客。

客戶 ID	MLE 數值	WMLE 數值	CAI
89	10.5764706	9.24404925	0.12597977
861	7.5505618	7.97128589	-0.0557209

RFM 數值為 011 的此群顧客，儘管保持著優於平均水準的購買頻率和消費金額，卻已有一段時間不再使用該張信用卡進行交易。仔細分析此 11 名客戶的背景資訊與信用卡到期日期，本群消費者之平均年齡達 43.72 歲，為男女比例均衡之壯年人口，職業多半為白領階級。過往最頻繁於國內量販超市和交通上使用信用卡交易，並且以旅遊類型的累積消費金額最高，且其中 65% 為海外刷卡（如下方表 11 所示）。另外，本群顧客不再頻繁以該張信用卡交易的問題點並非出自於信用卡使用權限到期而尚未更新。

針對 RFM 為 011 的信用卡用戶，建議信用卡公司可以因應此群消費者昔日多於日常生活用品和交通上以信用卡交易的消費特性，推出如「六月底前單筆刷卡一千元享一百元現金回饋」、「交通票券刷卡享 89 折折扣」等優惠措施。

刷卡類型	交易次數	平均消費金額	累積消費金額
旅遊	81	15902.58	1288109
量販超市	251	2148.5	539273
交通	253	1771.04	448072
保險	78	4402	343356
3C	87	3616.59	314643
精品	50	5426.55	271327.55
百貨	72	2853.63	205461.36
餐飲	24	2726.04	65425
休閒文化	49	570.2	27939.99
藥妝	23	914.13	21024.99

表 11、RFM=011 客戶群信用卡消費行為

主題二、分析 RFM=001 「重要挽留客戶」消費者行為

針對 RFM=001 之信用卡客群代表，客戶 ID 2027 與客戶 ID 4819 顧客之 MLE、WMLE、CAI 分析結果，觀察其消費行為發現：客戶 ID 2027 與客戶 ID 4819 CAI 值均小於 0，同樣屬於漸趨靜止的顧客。且客戶 ID 2027 號顧客的 MLE 方法估算值遠小於 WMLE 估算值，代表其交易時間間隔越來越長。客戶 ID 4819 號顧客的 MLE 方法估算值則稍小於 WMLE 估算值，代表其交易間隔有逐漸拉長的傾向。

客戶 ID	MLE 數值	WMLE 數值	CAI
2027	17.9736842	22.4331984	-0.24811353
4819	17.9230769	18.7769231	-0.04763949

RFM 數值為 001 的此群顧客，儘管最近沒有交易且累積交易次數低，然而單筆平均交易金額高，值得信用卡公司重點挽留（資料庫中，信用卡公司所有客戶單筆消費金額約為 1,910 元，除卻極端值，此群消費者單筆消費金額可達 6,000 元以上，具有相當的消費潛力）。

仔細分析此 5 名客戶的背景資訊與信用卡消費行為，本群消費者之平均年齡達 48.8 歲，為男女比例均衡之壯年人口，婚姻狀況多為單身。過往最頻繁於交通上使用信用卡交易，並且以旅遊和量販超市類型的累積消費金額最高，且均為國內交易（如下方表 12 所示）。同時，單筆消費超過 30,000 元的交易紀錄以旅遊與休閒文化刷卡類型為主。

針對 RFM 為 001 的信用卡用戶，建議信用卡公司可以因應此群消費者昔日於國內旅遊和休閒活動中願付較高的消費特性，推出如「國內旅展滿三萬送三百」、「線上旅宿平台 OTA 首刷享現金回饋」等優惠措施。

刷卡類型	交易次數	平均消費金額	累積消費金額
旅遊	9	27202	244818
量販超市	28	3902.428	109268
休閒文化	10	9937.7	99377
交通	60	976.98	58619
3C	10	4634.9	46349
精品	4	3677.75	14711
百貨	4	2128.75	8515
公用事業	20	74.75	1495

表 12、RFM=001 客戶群信用卡消費行為

主題三、分析 RFM=100「新客戶」消費者行為

針對 RFM=100 之信用卡客群代表，客戶 ID 1611 與客戶 ID 3219 顧客之 MLE、WMLE、CAI 分析結果，觀察其消費行為發現：客戶 ID 1611 號顧客的 MLE 方法估算值遠大於 WMLE 估算值，代表其交易時間間隔越發密集；CAI 值大於 0，表示其為刷卡行為漸趨穩定的、屬於活躍型客戶。客戶 ID 3219 號顧客的 MLE 方法估算值大於 WMLE 估算值，代表其交易間隔有逐漸縮小的傾向，且其 CAI 值同樣大於 0，屬於漸趨活躍的顧客。

客戶 ID	MLE 數值	WMLE 數值	CAI
1611	42	21.1666667	0.49603167
3219	4.3846154	4.0276923	0.08140351

RFM 數值為 100 的此群顧客，為最近剛開始交易的新用戶，屬於具有推廣價值的客戶。目前 RFM=100 的 10 名客戶中，40% 為年齡低於 35 歲的未婚青年、60% 為平均年齡 58.1 歲的已婚壯年人口，男女比例均衡。目前最頻繁於 3C 居家電信和交通上使用信用卡交易，並且以量販超市和 3C 居家電信刷卡類型的累積消費金額最高（如下方表 13 所示）。

針對 RFM 為 100 的信用卡用戶，建議信用卡公司可以因應此群消費者消費力道集中於量販超市和 3C 居家電信的消費特性，與電信公司和量販超市合作，推出如「指定通路 5% 現金回饋」、「單筆消費滿萬贈送信用卡用戶專屬贈品」等優惠措施。

刷卡類型	交易次數	平均消費金額	累積消費金額
量販超市	23	3518.87	80934
3C 居家電信	67	1191.62	79839
旅遊	5	12682.2	63411
交通	61	926.9	56541
藥妝	9	6195.5	55760
百貨	24	1239.4	29746
休閒文化	20	1074.85	21497
餐飲	6	1411.5	8469
精品	2	1700	3400

表 13、RFM=100 客戶群信用卡消費行為