RFM Analysis

Manish Gangwar @ ISB

Introduction

How do you find out customers to whom you should offer promotions?

How do you find out the ‘best customers’ , ‘loyal customers’ , ‘big spenders’ etc?

Can we somehow create a rank of customer based on some metrics, to answer above questions?

RFM Analysis

- Key performance Indicators: Elements of past purchase behavior that are predictive of what customer will do in next period. 1960's
  - **Recency** - how long it has been since a customer last placed an order with the company
  - **Frequency-** how often a customer orders from the company in a certain defined period
  - **Monetary value-** the amount that a customer spends on an average transaction

RFM Analysis.

Alternate methods:

- Method 1: Sorting customer data based on RFM, grouping and analyzing results
- Method 2: Computing relative weights for R, F and M and scoring customers
- Method 3: RFM Transition Matrix based predictive analysis

RFM for Targeting

Example:

- Customer base: 400,000 customers
- Sample size: 40,000 customers
- Firm's marketing mailer campaign: $150 discount coupon
- Response rate: 808 customers (2.02%)

- Suppose you have the above data and you want to find a relationship between recency and response rate. What would you do?

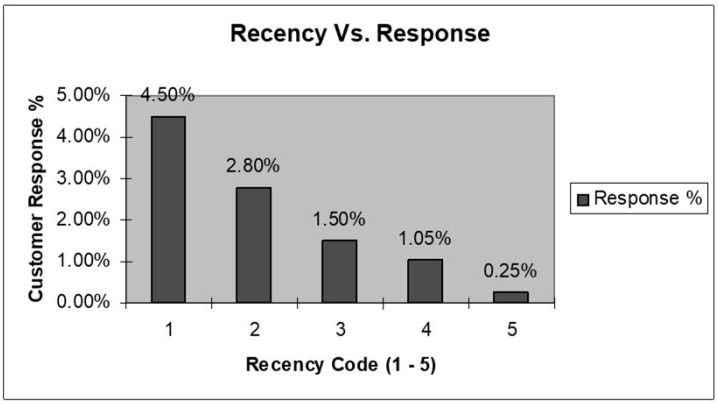
Recency Coding

Recency coding:

- Test group of 40,000 customers is sorted in a descending order based on the criterion of 'most recent purchase date'.
- The earliest purchasers are listed on the top and the oldest are listed at the bottom. The sorted data is divided into five equal groups (20% in each group)
- The top most group is assigned a recency code of 1 and the next group a code of 2 and so on, until the bottom most group is assigned a code of 5
- Analysis of customer response data shows that the mailer campaign got the highest response from customers grouped in recency code 1 followed by code 2 etc

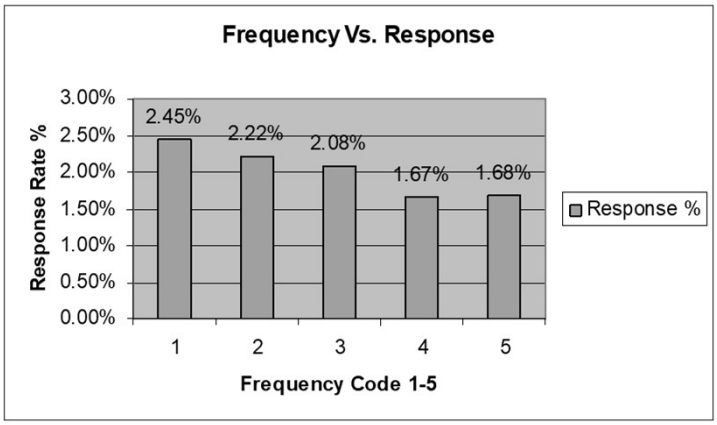
Method-1 Response and Recency

Plot shows the distribution of percentage of those customers who responded fell within the *recency* code grouping of 1 through 5



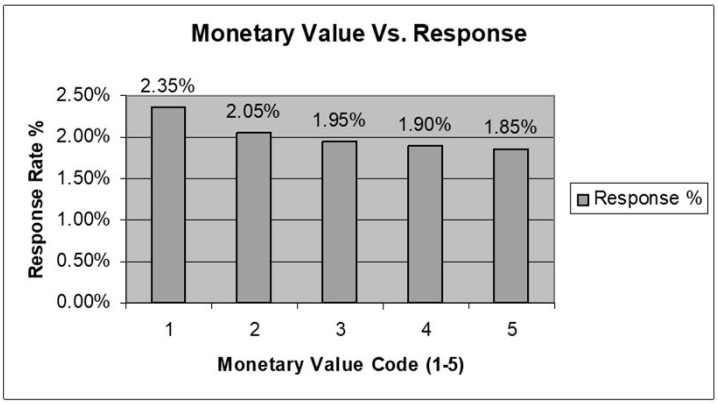
Method-1 Response and Frequency

Plot shows the distribution of what % of those customers who responded fell within the *frequency* code grouping of 1 through 5

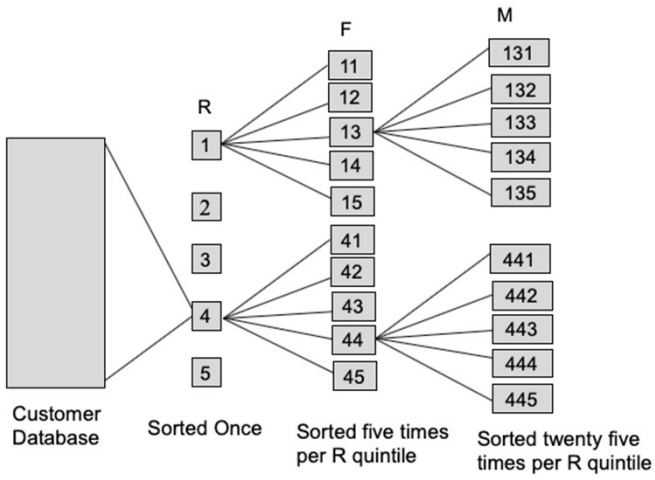


Method-1 Response and Monetary value

The highest response rate (2.35%) for the campaign was from those customers in the test group who fell in the highest *monetary value* quintile (monetary value code =1).



RFM Cell Sorting



RFM Cell Sorting

Example:

- List of 40,000 test group customers is first sorted for Recency and grouped into 5 equal groups of 8000 each
- The 8000 customers in each group is then sorted based on Frequency and divided into five equal groups of 1600 each- at the end of this stage, there will be RF codes starting from 11 through 55 with each group having 1600 customers
- In the last stage, each of the RF groups is further sorted based on monetary value and divided into five equal groups of 320 customers each

  - RFM codes starting from 111 through 555 each having 320 customers
- Considering each RFM code as a cell, there will be **125 cells** (5 recency divisions * 5 frequency divisions * 5 monetary value divisions)

RFM Breakeven Value Analysis

- Breakeven - net profit from a marketing promotion equals the cost associated with conducting the promotion
- Breakeven Value (BE) = unit cost price/ unit net profit
- BE computes the minimum response rates required in order to offset the promotional costs involved and thereby not incur any losses
- Example: In mailing $150 discount coupons,

  - The cost per mailing piece is $1.00

  - The net profit (after all costs) per used coupon is $45,

  Breakeven Value (BE) = $1.00/$45 = 0.0222 or 2.22%

Breakeven Index

- Breakeven Index (BEI) = ((Actual Response Rate – BE)/BE)* 100

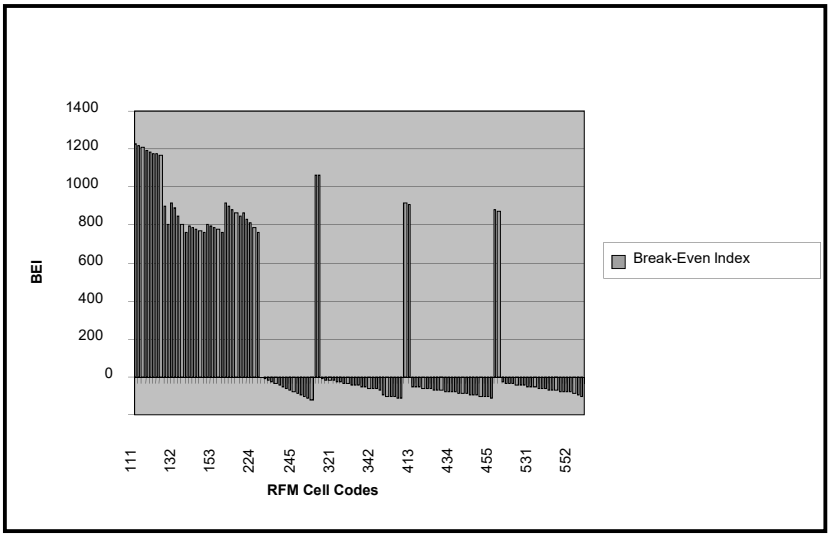
Example: If the actual response rate of a particular RFM cell was 3.5%

BE is 2.22%,

The BEI = ((3.5% - 2.22%)/2.22%)*100 = 57.66

- Positive BEI value some profit was made from the group of customers
- 0 BEI value the transactions just broke even
- Negative BEI value the transactions resulted in a loss

Break Even Index



RFM and BEI

- Customers with higher RFM values tend to have higher BEI values
- Customers with a lower recency value but relatively higher F and M values tend to have positive BEI values
- Customer response rate drops more rapidly for the recency metric
- Customer response rate for the frequency metric drops more rapidly than that for the monetary value metric

RFM and Profitability

	Test	Full customer base	RFM Selection
Average response rate	2.02%	2.02%	15.25%
# of responses	8080	8080	2732.8
Average Net profit/Sale	$45	$45	$45
Net Revenue	$36360	$363600	$122,976
# of Mailers sent	40,000	400,000	17920
Cost per mailer	$1.00	$1.00	41.00
Mailing cost	$40,000	$400,000	$17920
Profits	**($3640)**	**($36400)**	**$105056**

Method 2- RFM Scoring

- Regression techniques to compute the relative weights of the R, F, and M metrics
- Relative weights are used to compute the cumulative points of each customer
- The pre-computed weights for R, F and M, based on a test sample are used to assign RFM scores to each customer
- The higher the computed score, the more profitable the customer is likely to be in the future
- This method is flexible and can be tailored to each business situation

Recency Scoring

- 20 if within past 2 months; 10 if within past 4 months; 5 if within past 6 months; 03 if within past 9 months; 1 if within past 12 months;

  Relative weight = 5

Customer	Purchases (Number)	Recency (Months)	Assigned Points	Weighted Points
	1	2	20	100
JOHN	2	4	10	50
	3	9	3	15
SMITH	1	6	5	25
	1	2	20	100
MAGS	2	4	10	50
	3	6	5	25
	4	9	3	15

Frequency Scoring

- Points for Frequency: 3 points for each purchase within 12 months; Maximum = 15 points;
- Relative weight = 2

Customer	Purchases(#)	Frequency	Assigned Points	Weighted Points
	1	1	3	6
JOHN	2	1	3	6
	3	1	3	6
SMITH	1	2	6	12
	1	1	3	6
MAGS	2	1	3	6
	3	2	6	12
	4	1	3	6

Monetary Value Scoring

- Monetary Value: 10 percent of the $ Volume of Purchase with 12 months; Maximum = 25 points;
- Relative weight = 3

Customer	Purchases (Number)	Monetary	Assigned Points	Weighted Points
	1	$40	4	12
JOHN	2	$120	12	36
	3	$60	6	18
SMITH	1	$400	25	75
	1	$90	9	27
MAGS	2	$70	7	21
	3	$80	8	24
	4	$40	4	12

RFM Cumulative Score

Customer	Purchases (Number)	Total Weighted Points	Cumulative Points
	1	118	118
JOHN	2	92	210
	3	39	249
SMITH	1	112	112
	1	133	133
MAGS	2	77	210
	3	61	271
	4	37	308

- Cumulative scores: 249 for John, 112 for Smith and 308 for Mags; indicate a potential preference for Mags
- John seems to be a good prospect, but mailing to Smith might be a misdirected marketing effort

Example

Customer id, Purchase Date and Purchase Amount

customer_id	purchase_amount	date_of_purchase
760	25	06-11-09
860	50	28-09-12
1200	100	25-10-05
1420	50	09-07-09
1940	70	25-01-13
1960	40	29-10-13
1940	70	25-01-13
1960	40	29-10-13

https://cran.r-project.org/web/packages/rfm/vignettes/rfm-customer-level-data.html

RFM Segmentation

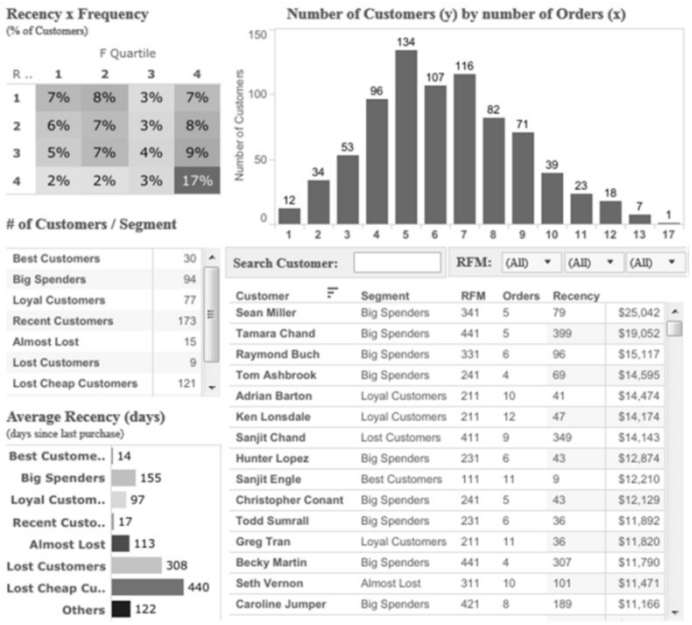
- Create RFM variables
  - Standardize
    - 5 vs 15 and 310 vs 320 ..may be logs
  - Recency, Frequency, Monetary

- RFM Segmentation1.r
  - Cluster analysis
    - K-means
    - Hierarchical clustering
    -
    -
    -

RFM Mangerial Segementation

Segment	RFM	Description	Marketing
Best Customers	111	Customers who bought most recently, most often and spend the most.	No price incentives, New products and loyalty programs
Loyal Customers	X1X	Customers who bought most recently	Use R and M to further segment.
Big Spenders	XX1	Customers who spend the most	Market your most expensive products.
Almost Lost	311	Haven't purchased for some time, but purchased frequently and spend the most.	Agressive price incentives
Lost Customers	411	Haven't purchased for some time, but purchased frequently and spend the most.	Agressive price incentives.
Lost Cheap Customers	444	Last purchase long ago, purchased few and spend little.	Don't spend too much trying to re-acquire.

RFM Dashboard



RFM Advance Managerial Analysis

- RFM_Segmentation2.r
- Classification (contextual knowledge)
  - Recency
    - Less than 6 months - active
    - Between 6 months to year - warm
    - Between 2 and 1 year - cold
    - More than two year - inactive
  - Monetray
    - Less than 60 - low value
    - More than 60 - high value
- Calculate year to year transition probabilities