RFM in the 21st Century
(Recency, Frequency, and Monetary Value)

Jim Porzak, DS4CI.org, El Cerrito, California. Jim@DS4CI.org

Background

RFM was devised in the early 1990's by direct marketers as a way to segment their customer base to target the most valuable (profitable) customers. Arguably RFM is the first use of "data-driven" marketing.

Each customer is scored by:
Recency: # of days since last purchase.
Frequency: # of purchases ever made.
Monetary: total value of purchases.

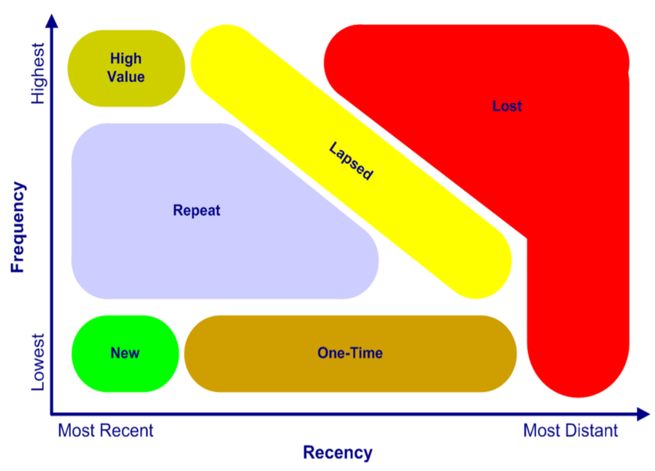
Evolution

Initially, R, F, & M values were binned into quantiles which were combined in one RFM key, e.g. 555 represented the customers in the top quantile for all three metrics. Mail response rates were then calculated for each key which were then ranked.

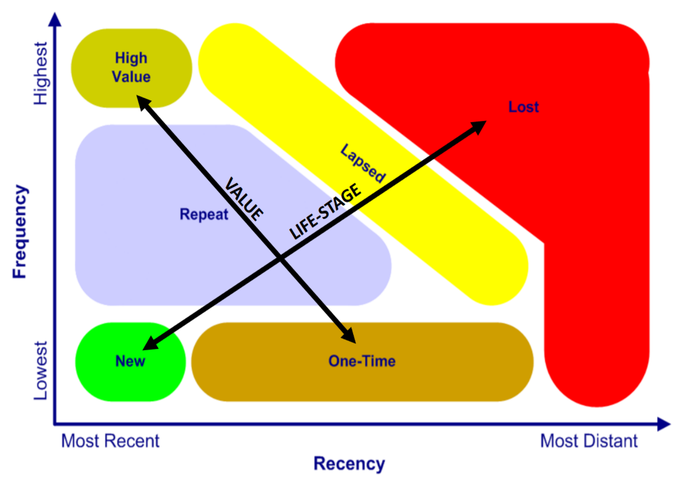
Today, some form of logistic regression on response is based on many predictors giving much better results. *But*, the RFM metrics generally turn out to be important predictors (especially R & F). The intuition of the pioneering marketers was valid.

Actionable RFM Segments

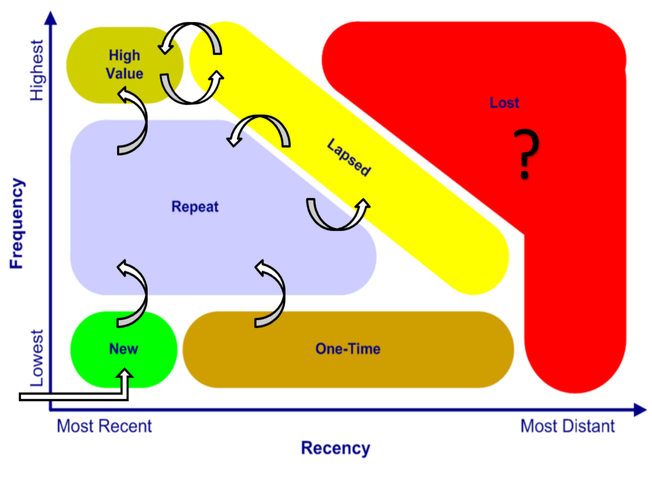
We have proposed the following RFM based segments to many marketing groups. They are intuitive and easy to understand.



Monetary is shown by the colors (gold, silver & bronze) along the Value axis. Life-stage is roughly orthogonal to Value.



Each transition – from customer acquisition through movement between segments – is a KPI and should be actively managed.



rfmr Package - Basics

Supports full RFM analysis

- Input is raw customer transactions
- Calculates RMF metrics for each customer
- Segments based on supplied binning rules
- Calculates segment KPI's

`github.com/ds4ci/rfmr`

Roadmap

- Output customer level transition alerts
- Integrate with DBI supported DBMS platforms
- Make rfmr a service on OpenCPU

rfmr Package – Sample Data

event_orders 20k orders by 4k customers

- *uid* customer ID
- *order_on* date
- *num_seats* quantity
- *value* monetary value

retail_orders 541k order lines in 263k orders by 125k customers

- *CustID*
- *OrderID*
- *OrderDate*
- *OrderChannel* (factor)
- *LineNum* of order
- *SKU* product ID
- *Quantity*
- *Amount*

rfmr Package – Core Function

`rfm(orders_df, cust_id_col, order_date_col, order_value_col, as_of_date = NA)`

Returns tbl with columns:

- *<customer ID column>*
- *last_on* most recent order date
- *first_on* earliest order date
- Frequency number orders (integer)
- Monetary total value (numeric)
- Recency days between last order and data as-of date (integer)
- tenure days between first order and data as-of-date (integer)