Kenan-Flagler Business School
The University of North Carolina

Recency, Frequency and Monetary (RFM) Analysis

RFM is widely used by direct marketers of all types for selecting which customers to target offers to. The fundamental premise underlying RFM analysis is that customers who have purchased recently, have made more purchases and have made larger purchases are more likely to respond to your offering than other customers who have purchased less recently, less often and in smaller amounts. RFM analysis can also be used to target special offers to 'welcome' new customers, encourage small purchasers to spend more, to reactivate lapsed customers, or encourage other marketing initiatives.

RFM analysis uses information about customers' past behavior that is easily tracked and readily available. *Recency* is how long ago the customer last made a purchase. *Frequency* is how many purchases the customer has made (sometimes within a specified time period, such as average number of purchases per year). *Monetary* is total dollars spent by the customer (again, sometimes within a specified time period).

RFM by Example: The BookBinders Book Club

The BookBinders Book Club sells specialty books and selected other merchandise through direct marketing. New members are acquired by advertising in specialty magazines, newspapers and TV. After joining, members receive regular mailings offering new titles and, occasionally, related merchandise. Right from its start, BookBinders made a strategic decision to build and maintain a detailed database about its club members containing all the relevant information about their customers.

Initially, BookBinders mailed each offer to all its members. However, as BookBinders has grown, the cost of mailing offers to the full customer list has grown as well. In an effort to improve profitability and the return on his marketing dollars, Stan Lawton, BookBinders marketing director, was eager to assess the effectiveness of database marketing techniques. Because of direct marketers' long history of success with RFM and its relative ease of use compared with more sophisticated modeling approaches, Stan decided to test the RFM approach.

Professor Charlotte Mason prepared this note to provide material for class discussion rather than to illustrate either effective or ineffective handling of a business situation.

Stan proposes to conduct live market tests, involving a random sample of customers from the database, for new book titles in order to analyze customers' response and calibrate a response model for the new book offering. The response model's results will then be used to "score" the remaining customers (i.e. those not selected for the test) and to select which customers to mail the offer to.

BookBinders' customer database provides a complete record of purchasing history for each customer. This includes how long they have been a customer, the specific titles ordered and summary totals by category such as cooking or children's books. Of direct relevance for RFM analysis, BookBinders keeps a record of the number of months since last purchase, the total number of purchases made as well as the total dollars spent by each customer. With these three pieces of information for each customer, Stan can easily test the RFM approach.

The Art History of Florence Offer

To test the RFM approach, Stan conducted a test. He had a random sample of 50,000 customers drawn from BookBinders customer database. By selecting a random sample of customers, Stan could be confident that all types of customers would be represented: both recent and not-so-recent purchasers, frequent and infrequent purchasers and customers spanning a range of total dollars spent.

This random sample of customers was mailed an offer to purchase *The Art History of Florence* and their response – either purchase or no purchase – was recorded. Now for each customer in the test, Stan knew his or her values for the recency, frequency and monetary variables at the time the offer was mailed and he knew the response. (Note that the recency, frequency and monetary values are at the time the offer was sent – and, for this analysis, have not been updated for those who did buy *The Art History of Florence*. Had they been updated then all the buyers would fall into the most recent category! What we want to know is whether recency (and frequency and monetary) *values at the time the offer was mailed* are useful for predicting who will respond.)

Stan's objective is to use the results of the test mailing to identify which groups of customers are more likely to respond. Then, for the 'rollout' mailing, he will only target customers who fit the profile of those more likely to respond. By carefully targeting which customers to mail the offer to, Stan hopes to reach the majority of the responders while significantly reducing costs by not mailing to those with a low likelihood of responding. A secondary benefit is that customers with little interest in a title such as *The Art History of Florence* will not get the mailing – which, had they received it, may leave them wondering why they are getting such unappealing offers.

The Results

Stan began by comparing those who bought *The Art History of Florence* to those who didn't in terms of the recency, frequency and monetary variables. Exhibit 1 reports the averages for number of months since last purchase (recency), total number of purchases (frequency) and total dollars spent (monetary) for the two groups of customers: those who did buy *The Art History of Florence* and those who did not. The results are consistent with what proponents of RFM analysis would predict. Those who did respond to the offer were more recent purchasers (8.6 months compared with 12.7), more frequent purchasers (5.2 purchases compared with 3.8) and had spent slightly more in total ($234 versus $206).

This suggests that the RFM variables are indicative of whether or not customers respond. The next step is to assess the response rate by decile for each RFM variable separately.

Exhibit 1 Average Values of R, F, and M Variables for Buyers and Non-Buyers

Mean

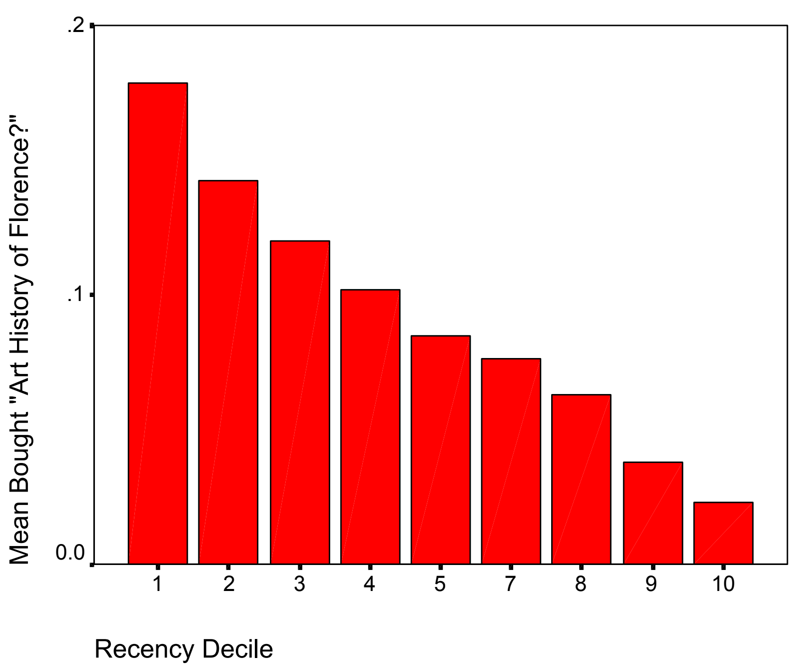
Bought "Art History of Florence?"	Months since last purchase	Total # purchases	Total $ spent
No	12.73	3.758	205.7348
Yes	8.61	5.219	234.3014
Total	12.36	3.890	208.3183

Decile Analysis

To form the deciles, the 50,000 test customers are ranked by the number of months since last purchase (recency) and then divided into ten equal sized groups. The most recent customers are in decile 1, the second most recent in decile 2 – down to the least recent who are in decile 10. (Note that it is arbitrary whether decile 1 or decile 10 reflects the 'best' customers. However, it is important to be consistent. So whichever convention is selected – 1 for best or 10 for best – follow that for all decile analysis.)

The bar chart in Exhibit 2 shows the response rate for each recency decile. The response rate is simply the proportion of customers in each decile who purchased *The Art History of Florence*. Exhibit 2 shows a strong relationship between recency and response rate. The response rate for the most recent decile is nearly .18 or 18%, whereas the response rate for the least recent decile is about .02, or 2%.

Exhibit 2 Response Rate by Recency Decile

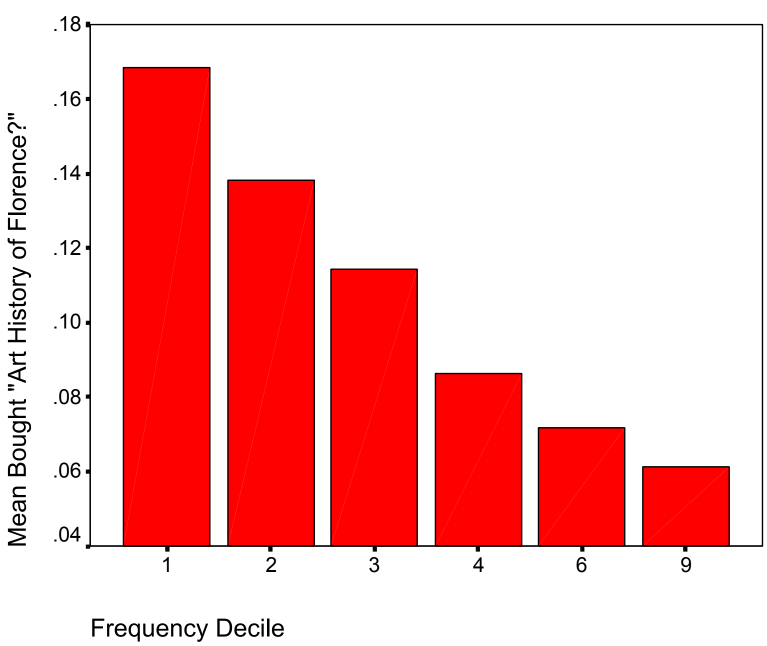


The decile analysis was repeated for total number of purchases to create frequency deciles. The response rate by frequency decile is shown in Exhibit 3. Note that there are only six, rather than 10, groups for frequency. This can happen for variables such as frequency that take on relatively few discrete values. In particular, for BookBinders – and many other retail businesses – there are a large number of one-time purchasers. Some of these are new customers who haven't yet made a second purchase. Others have only been interested enough to make one purchase – and may never become repeat customers.

In this instance, more than one-tenth of the 50,000 customers in the test had made just one purchase. Rather than arbitrarily assigning some of these one-time purchasers to decile 1 and others to decile 2, the SPSS software keeps all of them together in the same decile. Similarly, if there are a large number of two-time purchasers – larger than one-tenth of the total – they too will all be grouped into the same decile. The final result is that it is possible to end up with less than ten deciles. (As a technical note, this varies with the software package. Some software packages will create deciles each containing one-tenth even if it means putting customers with the same value into different deciles. Other software packages including SPSS, which was used here, may create 'deciles' of uneven sizes or even fewer than 10 deciles.)

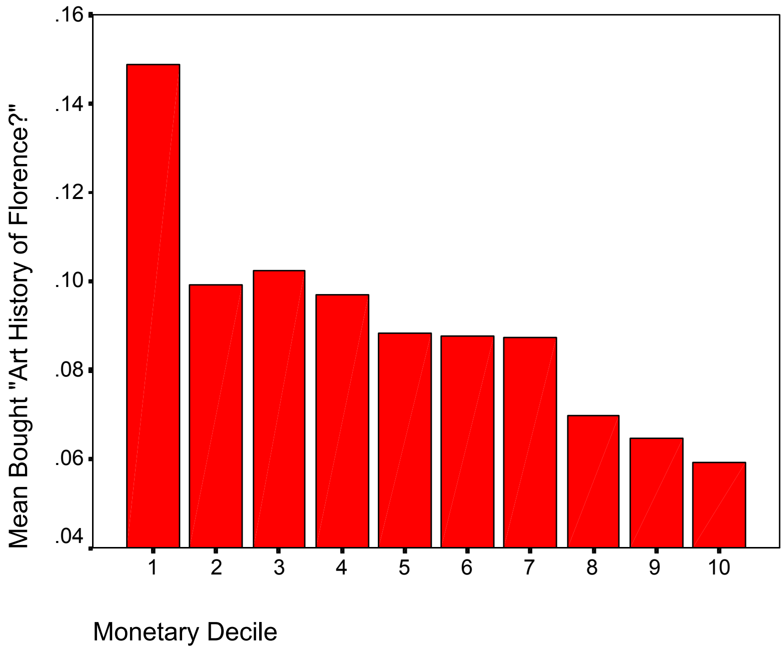
As with recency, the response rates for the frequency deciles show a clear pattern: more frequent customers had a much higher response rate to this offer compared with less frequent customers. Comparing the results for recency and frequency, it appears that recency is somewhat more predictive of response since there is a wider range of response rates across the decile groups.

Exhibit 3 Response Rate by Frequency Decile



Finally the same approach was used for total dollars spent to create monetary deciles. The response rate by monetary decile is shown in Exhibit 4. Once again a similar, though somewhat different pattern, emerges. The customers in the top monetary decile (decile 1) had a much higher response rate than the remaining deciles. However, there are only small differences in the response rates for deciles 2 through 7 and for deciles 8 through 10.

Exhibit 4 Response Rate by Monetary Decile



Putting it all Together with RFM

The decile analysis results show that, individually, recency, frequency and monetary are all predictive of response rate. The next step is to create a combined RFM index. This combination RFM index will divide the customer group into smaller subgroups compared to the decile analysis. In addition, these smaller groupings will reveal larger variations in response rate. Using decile analysis, the largest differences appeared for recency – with response rates ranging from approximately 2% to 18%. Both frequency and monetary showed smaller differences in response rates.

Using RFM analysis, Stan should see even larger differences in response rate that will enable him to more precisely target the likely buyers. Since all possible combinations of recency, frequency and monetary deciles would result in as many as 1000 (=10 x 10 x 10) groups, the combination RFM index is normally done using quintiles or 5 groups.

In practice, there are several variations in the approaches used to assign customers to RFM groups or cells. The approach that Stan decides to use is the one espoused by Arthur Hughes, a widely recognized expert in the field of database marketing. This approach involves sorting customers into 5 recency quintiles first. Then each of the recency quintiles is sorted by frequency and divided into 5 groups – resulting in 25 groups. Then each of these 25 groups is sorted by monetary. A key advantage of this approach is that it yields RFM cells of equal (or more nearly equal) size compared with other approaches.

Thus, the RFM classification is created using the following steps:

1. Rank all 50,000 customers according to recency and divide into 5 quintiles (here we will designate quintile 1 as the most recent and quintile 5 the least recent)
2. For customers in recency quintile 1: rank by frequency and divide into 5 quintiles
3. Repeat step 2 for recency quintiles 2 through 5 – this results in 25 (=5 x 5) groups
4. For each of the 25 recency-frequency groups: rank by monetary and divide into 5 quintiles
5. Now each customer has a quintile value for recency, frequency and monetary which determines his or her RFM group.

Exhibit 5 shows the response rates for the 125 RFM cells. As expected the response rates vary significantly across cells – ranging from a low of less than 1% to a high of approximately 24%. Exhibit 6 shows an excerpt from a report listing the number of customers for each RFM cell, the number of buyers for each RFM cell and the response rate for each RFM cell.

Exhibit 5 Response Rate by RFM Index

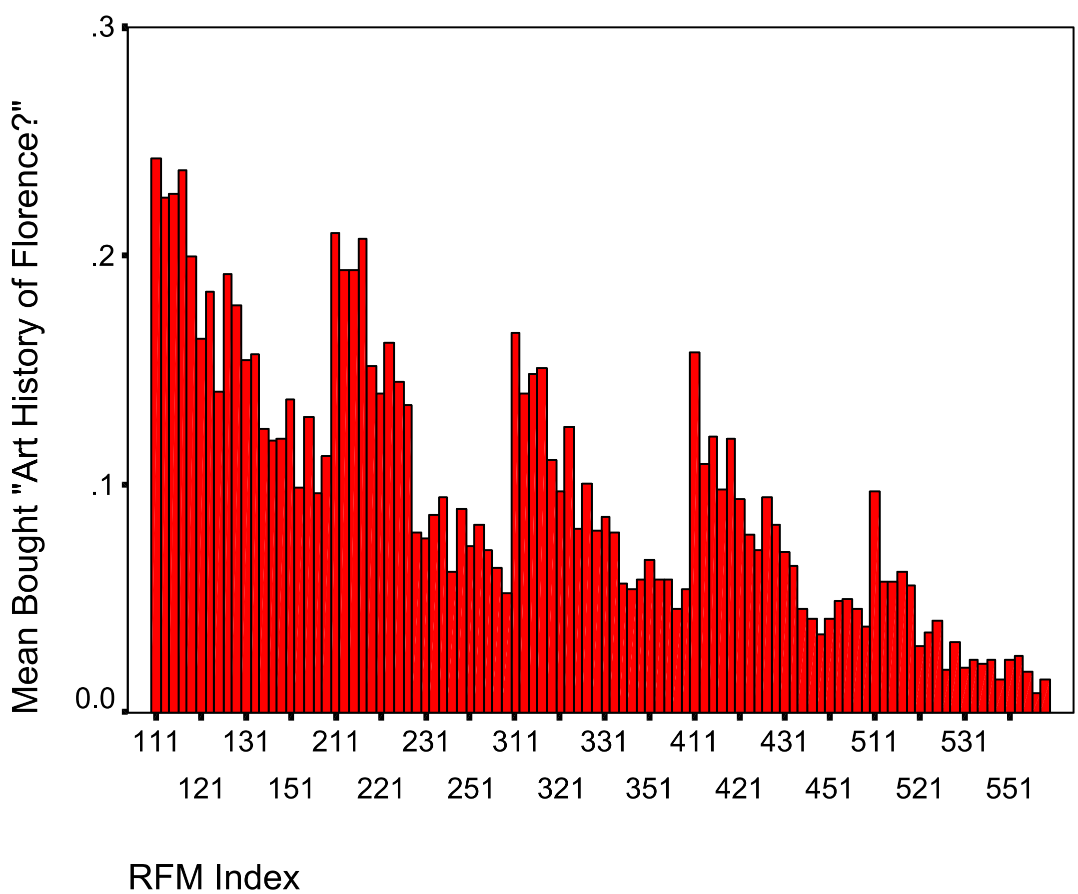


Exhibit 6 Customers, Buyers and Response Rate by RFM Index

RFM Index	# Customers	# Buyers	Response Rate
111	450	109	24.2%
112	452	102	22.6%
113	454	103	22.7%
114	451	107	23.7%
115	455	91	20.0%
121	446	73	16.4%
:	:	:	:
231	606	46	7.6%
232	601	52	8.7%
233	607	57	9.4%
234	603	37	6.1%
235	607	54	8.9%
:	:	:	:
551	599	14	2.3%
552	606	15	2.5%
553	608	11	1.8%
554	596	5	0.8%
555	605	9	1.5%

Putting RFM to Work

The results from Exhibit 5 and 6 clearly show that the response rate to *The Art History of Florence* offer varies across RFM cells. In general, customers who are more recent, more frequent and have spent more were more likely to respond to the offer than customers who were less recent, less frequent and have spent less.

Stan can now use the results of this test to select which of BookBinders remaining customers (who were not part of the test) to target with this mailing. Because the test was done using a random sample, Stan can expect very similar response rates from the full customer list. Using the report that Exhibit 6 is based on, Stan can rank the RFM cells from best to worst in terms of response rate.

To decide what cutoff point to use in determining which RFM cells to mail to, Stan needs to compute the breakeven response rate. Then, for the rollout mailing, he can target customers in RFM cells with a response rate that is greater than or equal to the breakeven response rate. Stan computes the breakeven using the following cost information:

Cost of Mailing the Offer	\$0.50
Selling price (including shipping) paid by customer	\$18.00
Wholesale price paid by BookBinders	\$9.00
Shipping costs	\$3.00

To exactly breakeven, the cost of the mailing is equal to the net profit from the mailing. If X is the number of customers mailed (at a cost of \$0.50 each) and Y is the number who respond (netting \$18 - \$9 - \$3 = \$6 each), then BookBinders will breakeven if X(0.50) = Y (6), or Y/X = 0.50/6. More generally, the following formula is used to compute the breakeven response rate:

Breakeven = (cost to mail the offer) / (net profit from a single sale)
= $0.50/($18 - $9 - $3)
= $0.50/$6
=.083, or 8.3%

Profitability Analysis

Exhibit 7 shows a profitability comparison between two scenarios:

- Mailing to all 50,000 customers in the test
- Mailing only to customers in RFM cells with response rates that are at least 8.3%

Of course, prior to conducting the test, Stan cannot have known which of the 50,000 customers to target. However, it is useful for comparison purposes to evaluate what the profitability *would have been* if Stan had known which RFM cells to target. This allows Stan to compare the mass mailing or 'no targeting' approach to a targeted mailing based on RFM.

Exhibit 7 Profitability Analysis

	Mass Mailing	Targeted RFM Mailing
Number of Customers	50,000	50,000
Number mailed	50,000	22,731[a]
% of customers mailed	100%	45.46%
Number of buyers	4522	3214[b]
Response rate	9.04%	14.14%
Gross Revenues	$81,396	$57,852
COGS	$40,698	$28,926
Shipping	$13,566	$9,642
Mailing	$25,000	$11,365.50
Gross Profit	$2,132	$7918.50
Gross Profit/Sales	2.62%	13.69%
Return on Marketing Expenditures [c]	8.53%	69.67%

[a] A total of 22,731 customers were in RFM cells that had a response rate greater than 0.083.
[b] The RFM cells with a response rate greater than 0.083 contained 3214 buyers.
[c] Return on marketing Expenditures = Gross Profit/Marketing Expenditures

Exhibit 7 shows that a mass mailing approach to the 50,000 customers yields a small gross profit of $2,132 or 2.62% of gross sales. In contrast, targeting only those customers likely to respond yields a gross profit of $7918.5 or 13.69% of gross sales. The RFM approach improves profitability by capturing 71% of the buyers (=3214/4522) while mailing only 46% (=22,731/50,000) of customers.

The mass mailing approach spends $25,000 on marketing (i.e. cost of the mailing) to generate $2,132 – a return on marketing expenditures of 8.53%. In contrast, the RFM targeted approach spends $11,365.50 on marketing to generate $7918.50 – a return on marketing expenditures of nearly 70%.

With these results Stan can anticipate the response for the rollout mailing. If BookBinders has another 450,000 customers who were not part of the test, then approximately 204,570 (45.46% of 450,000) would fall into 'profitable' RFM cells and would be sent the offer. The expected response rate would be 14.14%, or 28,926 buyers.

As in numerous other direct marketing situations, RFM appears to be a useful tool for BookBinders for identifying which customers are more (or less) likely to respond to a specific offer. By conducting a test using a random sample of customers, BookBinders is able to identify and then target only the customers who fall into profitable RFM cells. Profitability is improved by significantly reducing mailing costs – while still capturing the majority of buyers.

RFM in Summary

In the BookBinders example, as in many direct and database marketing situation, RFM was shown to be effective approach for predicting response and improving profitability. RFM analysis offers many advantages: it is simple, relies on data that even the most basic of customer databases can track, it is intuitive and it does not require sophisticated analysis, software, or statistical experts. On the other hand, RFM analysis is not without limitations: it relies on only three variables and may ignore other potentially important predictors.