



### **Sonoran Lifestyles: Techniques segment customers effectively**

Joan Beckman is increasingly confident that careful analysis of the *Sonoran Lifestyles* customer database will lead to more efficient and effective marketing decisions. As director of marketing for the upscale website specializing in cookware, tableware, linens, and decorative home accessories in the spirit of the Sonora region of Italy, Joan has been pursuing several initiatives in database marketing and customer relationship management.

*Sonoran Lifestyles* is proud of its track record in identifying prospects that become customers. Their historical response rate for direct mail to prospect lists is 2.3% – far above the industry average. However, Joan is convinced that there are opportunities to increase the effectiveness of subsequent email offers. *Sonoran Lifestyles* typically sends emails eight times a year to their full customer list. Yet, as their customer list grows Joan wonders if it may be more effective to further segment their emails – and, through market tests, identify which subset of the full customer list to target.

Marketers have a long history of success with using RFM (for Recency – Frequency – Monetary) analysis to target customers. 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. It’s one of the simplest forms of segmentation. We will learn other segmentation methods in our modeling class.

#### **Decile Analysis**

Although there are several variations of RFM segmentation, they all begin by classifying customers based on the recency of their last purchase, the frequency of their purchases, and the amount spent. The typical first step is to conduct some decile analysis. Decile analysis classifies customers into deciles for each variable – recency, frequency and monetary. For example, for recency, the customer list is sorted in order – from the most recent purchasers to the least recent, then divided into ten equal (or nearly equal) groups. The 10% who comprise the most recent purchasers are grouped into one decile, the 10% who are the next most recent into the second decile, and so on. There is no uniform convention for the coding or labeling of deciles – the top or best decile can be labeled 1 or 10. Regardless of whether 1 or 10

---

*Professor Charlotte Mason prepared this case to provide material for class discussion rather than to illustrate either effective or ineffective handling of a business situation. Names and data may have been disguised to assure confidentiality. The assistance of the Direct Marketing Educational Foundation in supplying the data used for this case is gratefully acknowledged.* Copyright © 2018 by Charlotte Mason and Debra Zahay Blatz.

designates the best decile, the same convention should be used for all decile variables. That is, if the most recent customers are in recency decile 1, then frequency decile 1 should contain the most frequent customers, and monetary decile 1 the largest spenders.

Once customers are coded into deciles, the response rates, purchase amounts or other relevant variables can be compared across the 10 deciles. If recency is indicative of response to a marketing campaign, then the most recent decile will have a higher response rate than the deciles of less recent customers. The same analysis is repeated for frequency and monetary.

## **RFM Classification**

In combination, there would be 1000 ( $10 \times 10 \times 10$ ) potential combinations for the recency, frequency and monetary deciles – ranging from customers who rank in the top decile for all three variables to those in the lowest decile for all three and all possibilities in between. To reduce the number of combinations it is typical to consider quintiles – or five groups – rather than deciles when looking at combinations of RFM. Thus, there are 125 ( $5 \times 5 \times 5$ ) possible RFM groups or segments.

There are at least three variations of how RFM segments are determined:

### **'Independent n-tiles' Approach**

This approach computes quintiles independently for recency, frequency and monetary. That is, the entire customer list is sorted based on recency and divided into recency quintiles. Then, the entire customer list is resorted based on frequency and divided into frequency quintiles. Finally, the entire customer list is sorted one more time based on monetary and divided into monetary quintiles. Then these three quintile variables are combined to form the RFM index. For example, a customer in the 1<sup>st</sup> recency quintile, 2<sup>nd</sup> frequency quintile and 4<sup>th</sup> monetary quintile – is assigned an RFM index of 124.

This approach typically yields unequal numbers of customers across the 125 combinations. For example, some combinations – such as customers who are very recent, very frequent, yet very low on monetary – are less common than others.

### **'Sequential n-tiles' Approach**

This approach computes quintiles first for recency. Then, *within* each of the five quintiles for recency, quintiles for frequency are computed – resulting in a total of 25 recency/frequency combinations. Finally, *within* each of these 25 groups, quintiles for monetary are computed. This method produces approximately equal numbers of customers for each of the 125 RFM cells which is useful for running market tests since it is problematic to compare response rates across groups with widely varying numbers of customers – especially if some groups have very few (say only 5 or 10) customers.

Note, however, that with this approach it is not possible to say in an absolute sense whether a '221' customer or a '411' customer has made more purchases. This is because the frequency quintiles are done *within* each recency group – as an analogy, consider sorting consumers into quintiles based on salary *separately* for males and females. A male in the second salary quintile may have a higher salary (in absolute terms) than a female in the top salary quintile. Similarly a customer in the 2<sup>nd</sup> recency and 2<sup>nd</sup> frequency quintile may have made more purchases than a customer in the 4<sup>th</sup>

recency but 1<sup>st</sup> frequency quintile.

#### 'Intuitive Groupings' Approach

This approach uses intuitive splits rather than quintiles to form the groups. For example, customers may be grouped into recency groups determined by: (1) purchase in last 6 months, (2) last purchase 6-12 months ago, (3) last purchase 12-24 months ago, (4) last purchase 24-36 months ago, and (5) last purchase more than 36 months ago. This approach is 'intuitive' in that it is easy to know what is meant by a customer in the second recency group, but relies on judgment to know where to 'draw the lines.' We will not have exercises on this method, but know this exists.

Following is an outline of what Joan proposes for using RFM analysis for targeting email campaigns to current *Sonoran Lifestyles* customers:

1. Compile the email offering.
2. Select a random sample from the full customer database and send the email.
3. Track the response (% who respond, \$ orders) by RFM classification or 'cell.' (sometimes called a bin).
4. Assess response across recency, frequency and monetary deciles.
5. Assess response as well as revenue by RFM cell.
6. Use results of the test mailing to determine which customers from the remaining customer list (i.e. those not in the test) to mail.

Joan has asked you to conduct an initial assessment of this approach – using data from the results a recent *Sonoran Lifestyles* email. Using data from a random sample of 96,551 customers who received this recent mailing (see Exhibit 1 for a summary and definitions of key variables found in the dataset *Sonoran\_RFM.sav*), Joan has asked you to prepare a report addressing the following questions.

#### **Part I: Preliminary and Decile Analysis**

1. What percent of customers responded (i.e. bought anything) from this email?
2. Of those who bought, what was the average dollars ordered from this email?  
(Hint: Use *Analyze... reports....case summaries*)
3. Create decile variables for recency, frequency and monetary where 1 is highest.  
(Hint: review the Using SPSS for Customer Analysis note if you need help creating the deciles.)
4. Create a bar chart showing the response rate (i.e., the proportion of customers who bought something) to this email by recency deciles. What insights can you derive from this chart?
5. Create a bar chart showing the bar chart of response rate to this email by frequency deciles. What insights can you derive from this chart?
6. Create a bar chart showing the bar chart of response rate to this email by monetary deciles. What insights can you derive from this chart?

7. *Using only those customers who placed an order from this email*, create bar charts showing the average dollars ordered from this email by recency, frequency and monetary deciles.  
(Hint: to select only buyers, go to the Data Editor window and choose *select cases* under the *Data* menu. In the pop-up window, select *If condition is satisfied*. Then specify *buyer = 1*. Click Ok and note that SPSS will 'filter' (but not delete) the unselected cases. Now all analyses will use only customers who bought. When you want to use the entire data set again, you will need to go back to this menu and select 'all cases'. )
8. What do the above bar charts reveal about the likelihood of response and the size of the order across the different recency, frequency, and monetary deciles?

## **Part II: RFM Classification**

Next, Joan is interested in the effectiveness of the 'independent n-tiles' vs 'sequential n-tiles' approach for RFM classification.

Create the RFM values using both the sequential and independent n-tiles approaches using SPSS and the instructions in the lecture slides. Just a reminder, the last problem used only people who placed an order from the email (*buyer =1*). Be sure to set select cases back to ALL before doing Part II of the homework.

1. Create the RFM values using both the sequential and independent n-tiles approaches using SPSS and the instructions in the lecture slides. Label your RFM classification variables as RFM\_IND and RFM\_SEQ respectively
2. Analyze the Bin counts and the Heat maps for each method. What do they tell you about each method? Based on these data, which method provides the best results.
3. Generate a report showing the number of customers, the mean value of dollars spent, minimum, maximum and Sum for each RFM cell for each of the approaches,  
(Hint: use *Analyze ... reports 'N'* will show the number of customers, 'Sum' will show the number of buyers, and 'mean' will give you the mean value of dollars spent per cell.)
4. Which RFM cells had the highest average dollars spent for each method?
5. Generate a report showing the number of customers, the mean value of bought from the last email, minimum, maximum and Sum for each RFM cell for each of the approaches,  
(Hint: use *Analyze ... reports 'N'* will show the number of customers, 'Sum' will show the number of buyers, and 'mean' will give you the mean value of bought for the last email per cell.)
6. Which had the highest percentage response of the total number who purchased from this email?
7. Which cells would you recommend targeting in the future and why?
8. Assume that the break-even response rate is .0192 or 1.92%. Create a graph of the RFM cells versus the mean purchases from the last email campaign for both RFM methods and draw a line showing the break even response on the chart.

Hint: It is sometimes useful to export data from an SPSS output table to Excel for analysis. Here are instructions to do this:

- Click on Table

- Right click and choose 'Copy as Excel Worksheet'

- Open a new worksheet in Excel and 'Paste' to see the data

- You can then sort and do other types of analyses in this worksheet.

**Exhibit 1**

Variable Names and Descriptions  
(*Sonoran\_RFM.sav*  
dataset)

NameDescription*Values of Recency, Frequency and Monetary variables:*

NumOrds	Total Life-to-date # of orders (prior to last email mailing)
TotDol	Total Life-to-date dollars (to last email mailing)
Last	Number of days since last purchase (prior to email mailing)

*RFM values (created by the student):*

RFM_SEQ	RFM cell (using 'sequential' n-tiles approach)
RFM_IND	RFM cell (using 'independent' n-tiles approach)

Note: the RFM cells are coded in SPSS Direct Marketing so that a '5' represents the top or best quintile whereas a '1' represents the bottom or worst quintile.

*Response to Email:*

Buyer	Responded to last email? (1=yes, 0=no)
-------	--

