CHAPTER 6

Customer Value Metrics—Concepts and Practices

Contents

6.1 Overview

In the previous chapter, we examined some traditional marketing metrics and various primary customer-based metrics. This chapter explains some surrogate metrics of customer value used in the industry and proceeds to conceptualize strategic metrics of customer value. Some of the primary customer-based metrics introduced earlier form the inputs to derive customer value—the key metric that drives decision making in the age of data-based marketing.

Size of wallet and *share of wallet* are popular metrics firms frequently apply to evaluate a customer's worth. In a frequently purchased category such as toothpaste, it is possible to observe customers frequently switching between multiple brands. A transition matrix can be used as a customer-based tool which allows us to compute the metric *expected share of wallet* for multiple brands. Metrics have been developed to keep track of marketing actions and costs related to customer acquisition. *Past customer value and RFM value* are more refined methods of computing the relative worth of individual customers. In recent times, *customer lifetime value* has emerged as the key metric needed to manage customers profitably. The aggregate of the customer lifetime value of each customer is referred to as the *customer equity* of the firm.

Customer selection strategies, which are explained in the third section of the chapter, help firms identify the right customers to optimally allocate the available marketing resources. The chapter explains two popular customer selection strategies—namely, decision tree and logistic regression. Typically, both decision tree algorithm and logistic regression identify the predictor variables that best separate the outcome of a 0/1 binary dependent variable. In the last section of the chapter, we discuss techniques companies can use to evaluate alternative selection strategies. The techniques described include lift analysis, cumulative lift analysis, and deciles analysis.

6.2 Popular Customer-Based Value Metrics

Firms have adopted some popular surrogate measures of customer value which they anticipate will be reasonable indicators of the actual customer value. These metrics help firms prioritize their customers in a manner that helps them assign a higher proportion of resources to the customers who they expect will generate greater profits in the future. In the absence of the availability of more sophisticated mathematical techniques, these measures help firms make superior decisions.

We suggest managers attempt to correlate these surrogate measures on a selective basis with more rigorous customer value metrics. Only if these correlations yield satisfactory results (i.e., correlations are substantial) can and should the surrogate measures be used for decision making.

6.2.1 SIZE OF WALLET

Size of wallet is the amount of a buyer's total spending in a category—or, stated differently, the category sales of all firms to that customer. Size of wallet is measured in monetary terms.

$$\text{Size of wallet (\$) of customer in a category} = \sum_{j=1}^{J} S_j \qquad \textbf{(1)}$$

where:

S_j = sales to the focal customer by the firm j

j = firm

$\sum_{j=1}^{J}$ represents the summation of the value of sales made by all firms, J, that sell a category of products to the focal customer.

Example: A consumer might spend an average of $400 every month on groceries, across several supermarkets. Thus, her size of wallet is $400.

Where does the information come from?

Information about the size of wallet can be gathered in many ways. For existing customers, the information can be collected through primary market research (e.g., surveys). A typical question a firm might ask is, "On average, how much do you spend every month on category A?" For prospects, it is quite difficult to obtain the size-of-wallet information on an individual level. Therefore, segment-level information is most appropriate.

Evaluation

Size of wallet is a critical measure of the customer-centric organization. When firms attempt to establish and maintain profitable relationships, the customer's buying potential (i.e., size of wallet) is a critical piece of information. Firms are particularly interested in acquiring and retaining customers with large wallet sizes. The assumption firms make here is that large customers will bring in more revenues and profits.

6.2.2 SHARE OF CATEGORY REQUIREMENT (SCR)

SCR is defined as the proportion of category *volume* accounted for by a brand or focal firm within its base of buyers. SCR is often computed as an aggregate level metric, when individual purchase data are unavailable. However, when such data are available, SCR is computed by dividing the volume of sales (V) of the focal firm in a category by the total category volume *within the firm's buyer base.* SCR is measured as a percentage.

$$\text{SCR (\%) of firm or brand in category} = \frac{\sum_{i=1}^{I} V_{ij}}{\sum_{i=1}^{I} \sum_{j=1}^{J} V_{ij}} \qquad \textbf{(2)}$$

where

j = firm
V = purchase volume
i = those customers who buy brand

$\sum_{i=1}^{I}$ represents the summation of volume purchased by all the i customers from a firm j

$\sum_{i=1}^{I} \sum_{j=1}^{J}$ represents the summation of volume purchased by all i customers from all j firms

The SCR ratio is sometimes calculated simply by using purchase occasions or product units as the unit of analysis. The computation discussed here is for the aggregated case. SCR can also be calculated for individual customers. At the individual level, then, this measures the percentage of a customer's SCR a given brand or store satisfies.

Example: In this example, there are three customers in the category. The category consists of three brands—SAMA, SOMO, and SUMU. Table 6-1 shows the number of purchases during a three-month period.

TABLE 6-1 CALCULATION OF SCR-PURCHASES DURING A THREE-MONTH PERIOD

	Brand SAMA	Brand SOMO	Brand SUMU	TOTAL
Customer 1	2	8	0	10
Customer 2	6	0	3	9
Customer 3	0	4	1	5
TOTAL	8	12	4	***24***

The category volume in the three-month period is 24 units. Brand SAMA has a market share of 33 percent (i.e., 8 purchases out of a total of 24) and a SCR of 42.1 percent (i.e., 8 purchases out of 19, made by its two buyers). This example shows that even though SAMA's market share is already substantial, its SCR is even higher. The high SCR for SAMA indicates that once consumers buy SAMA, they tend to prefer it disproportionately more.

Where does the information come from?

- Numerator: Volumetric sales of the focal firm are readily available from internal records.
- Denominator: The total volumetric purchases of the focal firm's buyer base are typically obtained through market and distribution panels, which are quite common for certain industries (e.g., fast-moving consumer goods (FMCG)). Other industries use mainly primary market research (surveys). Since this information is costly to gather, it is typically collected for a representative sample and then extrapolated to the entire buyer base. Qualitative managerial judgment is another potential low-cost alternative.

Evaluation

SCR is one of the most commonly accepted measures of customer loyalty, at least for FMCG categories. It separates the question, "whether anyone buys the brand" from the question, "how much they buy." An important characteristic of this measure is that it controls for the total volume of the segments/individuals category requirements. In other words, regardless of the total value of purchases per period, in terms of percentage of allocated purchases (loyalty), it puts all customers on the same metric. Overall, it is an important indicator on which firms should strive to deliver. However, this metric does not necessarily indicate whether a high SCR customer will generate substantial revenues or profits—this only comes through knowing something about the size of wallet of this customer.

Case in Point

Suppose a computer manufacturer, say ABC Computers, has the following data about its annual customer purchases on Notebook Computers for the year 2003. Using Table 6-2, it can compute the SCR ratio for each of its customers and identify those customers who have a higher SCR ratio from those with a lower SCR ratio.

From Table 6-2, we can see customer 3 has the highest SCR. Therefore, ABC Computers should identify high SCR customers such as customer 3, and target more of its marketing efforts (mailers, advertisements, etc.) toward such customers and their

TABLE 6-2 SCR RATIOS

	Total Requirement of Notebook Computers per Customer in 2003 A	Total Number of Notebook Computers Purchased from ABC Computers per Customer per Period B	Share of Category Requirement for ABC Computers per Customer per Period B/A
Customer 1	100	20	.20
Customer 2	1,000	200	.20
Customer 3	2,000	500	.25

respective requirements. In addition, customer 3's size of wallet (column A) is the largest, making customer 3 even more attractive.

6.2.3 SHARE OF WALLET (SW)

SW is defined as the proportion of category *value* accounted for by a focal brand or a focal firm *within its base of buyers.* It can be measured at the individual customer level or at an aggregate level (e.g., segment level or entire customer base).

Individual Share of Wallet (ISW)

ISW is defined as the proportion of category *value* accounted for by a focal brand or a focal firm for a buyer from all brands the buyer purchases in that category. It indicates the degree to which a customer satisfies his needs in the category with a focal brand or firm.

It is computed by dividing the value of sales (S) of the focal firm (j) to a buyer in a category by the size of wallet of the same customer in a time period. ISW is measured as a percentage.

$$\text{ISW (\%) of firm to customer (\%)} = \frac{S_j}{\sum_{j=1}^{J} S_j} \tag{3}$$

where

S = sales to the focal customer

j = firm

$\sum_{j=1}^{J}$ represents the summation of the value of sales made by all the J firms selling a category of products to a buyer

Example: If a consumer spends \$400 monthly on groceries, and \$300 of her purchases are with Supermarket BINGO, then BINGO's share of wallet for that consumer is 75 percent in that month.

Where does the information come from?

- Numerator: Typically, sales come from internal records (assuming that information is stored on the individual level).
- Denominator: Typically, sales value across all firms comes from primary market research (surveys), administered to individual customers. Since this information is costly to gather, it is often collected for a representative sample and then extrapolated to the entire buyer base. Sometimes, firms can *infer* the size of wallet for a certain product—for example, in certain business-to-business (B-to-B) contexts. For example, BASF, one of the few manufacturers of car paint, supplies its product to Daimler-Chrysler. Based on its knowledge of how much paint it takes to paint an average sized car, it can infer Daimler-Chrysler's size of wallet for car paint based on its worldwide production output—a figure easily derived from secondary sources.

Evaluation Just like SCR, SW is a measure of customer loyalty and can be an important metric on which firms should strive to deliver. However, SW is unable to provide a clear indication of future revenues and profits expected from a customer.

Aggregate Share of Wallet (ASW) (Brand or Firm Level)

Aggregate share of wallet (ASW) is defined as the proportion of category *value* accounted for by a focal brand or a focal firm *within its base of buyers.* It indicates the degree to which the customers of a focal firm satisfy their needs on average, in a category with a focal firm.

ASW of firm (%)

$$= \frac{\sum_{t=1}^{T} ISW_{ji}}{\text{Number of customers}} \quad \textbf{(4)}$$

$$= \frac{\sum_{i=1}^{I} S_i}{\sum_{j=1}^{J} \sum_{i=1}^{I} S_{ij}} \quad \textbf{(5)}$$

where

S = sales to the focal customer

j = firm

i = those customers who buy brand

Example: Using equation (4), BINGO may calculate its aggregate share of wallet as 60 percent among its 5,000 customers, by summing equation (3) across all customers and dividing by the number of its customers.

Or, using equation (5), the ASW is BINGO's sales in period t ($750,000) divided by the total grocery expenditures of BINGO's customers in the same period ($1,250,000); thus, 750,000/1,250,000 = 60 percent.

Where does the information come from?

- Numerator: This information comes from internal records.
- Denominator: The source is the same as for SCR.

Evaluation SW, like SCR, is an important measure of customer loyalty and an important indicator firms should strive to have.

When to Use SCR and When to Use SW

Information on share of wallet is slightly more difficult to obtain than for SCR. SCR is, in most cases, the preferred measure. This is particularly true for categories where the variance of customer expenditures is relatively small (groceries, for example). If purchases are similar in volume, a customer's lifetime value is primarily driven by his frequency of purchases. Thus, SCR is a fairly appropriate measure of loyalty. However, if the variance of consumer expenditures is relatively high (furniture, cars, or other infrequent purchases), then SW is a better measure of loyalty than SCR. In the former case, the frequency is more easily remembered. In the latter case, the customer more easily remembers the expenditures.

Share of Wallet and Size of Wallet Hold Important Pieces of Information

Even though two buyers might have the same share of wallet, firms might find their attractiveness as customers to be different. Consider Table 6-3.

TABLE 6-3 SHARE OF WALLET AND SIZE OF WALLET

	Share of Wallet	Size of Wallet	Absolute Expenses with Firm
Buyer 1	50%	$400	$200
Buyer 2	50%	$50	$25

Depending on the size of wallet, the absolute attractiveness of buyer 1 is eight times higher, even though the share of wallet is the same as for buyer 2. The example shows it is always important to consider share of wallet and size of wallet simultaneously. Figure 6-1 illustrates this.

The matrix shows the recommended strategies for the various segments differ substantively. The firm makes optimal resource allocation decisions only by segmenting the customers along the two dimensions simultaneously.

Difference of Share of Wallet to Market Share

It is important to recognize the difference between market share (MS) and share of wallet (SW). MS is calculated across buyers and nonbuyers, whereas SW is calculated *only among buyers.*

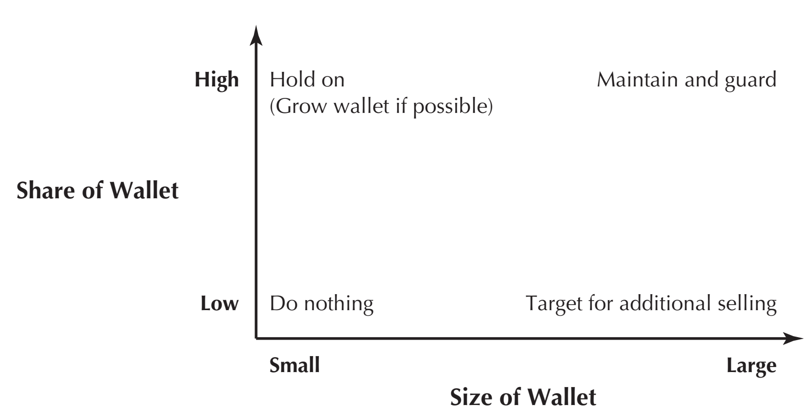
How to Convert SW to MS?

$$\text{MS of firm} = \frac{\sum_{i=1}^{I} (\text{Share of wallet}_i \times \text{Size of wallet})}{\sum_{j=1}^{J} S_j} \qquad (6)$$

where

S = sales to the focal customer
j = firm
i = those customers who buy the brand

FIGURE 6-1 SEGMENTING CUSTOMERS ALONG SHARE OF WALLET AND SIZE OF WALLET



Market share is the share of wallet across all customers in the category divided by the sales across all firms in the category in period *t*. MS is measured on a percent basis and can be computed based on unit volume, \$ volume, or equivalent unit volumes (grams, ounces).

Example: If BINGO has 5,000 customers with an average expense at BINGO of \$150 per month (share of wallet × size of wallet), and the total grocery sales in BINGO's trade area are \$5,000,000 per month, then BINGO's market share is (5,000 × \$150) / \$5,000,000 = 0.15, or 15 percent.

The implication here is that although BINGO has an overall low market share, it has a high share of wallet for those buyers buying at BINGO. This indicates that BINGO is a niche player with a very loyal clientele.

6.2.4 TRANSITION MATRIX

A transition matrix is a convenient way to characterize a customer's likelihood to buy over time or brand's likelihood to be bought. The assumption is that a customer moves over his/her lifetime through various stages of activity. Table 6-4 shows such a transition matrix.

In Table 6-4, the top row indicates the movement for customers who are currently brand A buyers; 70 percent of them will buy brand A next time, 20 percent will buy brand B, and 10 percent will buy brand C. The diagonals (in bold) are customer-retention probabilities computed by the company. However, we see that consumers can switch back and forth from brands. For example, the probability that a consumer of brand A will transition to brand B and then come back to brand A in the next two purchase occasions is 20% × 10% = 2%. If, on average, a customer purchases twice per period, the two purchases could be AA, AB, AC, BA, BB, BC, CA, CB, or CC. We can compute the probability of each of these outcomes if we know the brand the customer bought last. This provides us with the expected share of wallet for the three brands. This process can be continued for as many purchase occasions as desired. This information for the matrix comes from responses to questions as part of routine surveys, for example, "Which hotel did you stay in last?" or "The next time you stay in a hotel, what is the probability that you will stay at each of the hotels that you consider as options?"

6.3 Strategic Customer-Based Value Metrics

6.3.1 RFM

RFM stands for recency, frequency, and monetary value. This technique utilizes these three metrics to evaluate customer behavior and customer value.

TABLE 6-4 TRANSITION MATRIX

		Brand Purchased Next Time		
		A	*B*	*C*
Brand Currently Purchased	*A*	**70%**	20%	10%
	B	10%	**80%**	10%
	C	25%	15%	**60%**

Note: Customer retention probabilities are in bold.

SOURCE: Roland T. Rust, Valerie A. Zeithaml, and Katherine N. Lemon, "Driving Customer Equity" (New York: The Free Press, 2000).

1. *Recency* is a measure of how long it has been since a customer last placed an order with the company.
2. *Frequency* is a measure of how often a customer orders from the company in a certain defined period.
3. *Monetary value* is the amount that a customer spends on an average transaction.

RFM is similar to the switching matrix approach in that it also tracks customer behavior over time in what is called a state-space. That is, customers move over time through a space with certain defined activity states.

Two Methods of Computing and Applying RFM

Two popular methods of computing and applying RFM are in vogue among contemporary marketing professionals. The first method involves sorting customer data from the customer database, using the RFM criteria, and grouping them in equal quintiles and analyzing the resulting data.

The second method involves computing relative weights for R, F, and M using regression techniques, and then using those weights for calculating the combined effects of RFM.

RFM technique helps organizations significantly, not only in clearly targeting valuable customers who have a very high chance of purchasing, but also in avoiding costly communications and campaigns to customers who have a lesser chance of purchasing. RFM technique can be applied only on historical customer data available and not on prospects data.

RFM Method 1 For the following discussions about RFM coding, consider the following example of a firm with a customer base of 400,000 customers. From this customer base, a sample of 40,000 customers is chosen. In other words, every tenth customer from the larger database of 400,000 customers was picked so as to form a test group of 40,000 customers who are representative of the whole customer base.

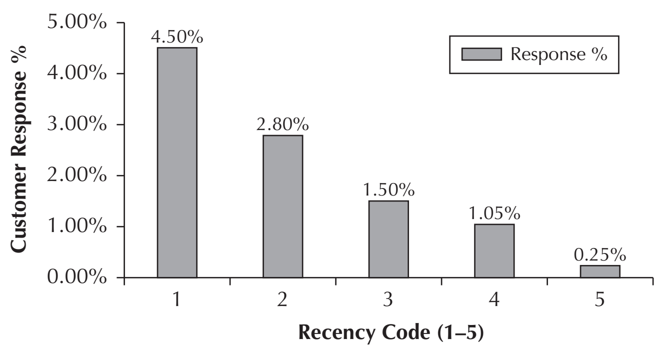
Also, assume this firm is planning to send a marketing mailer campaign of a $150 discount coupon to be mailed to its customers.[1]

Recency Coding Assume this firm sends its $150 mailer campaign to the 40,000 customers in the test group, and assume 808 customers (2.02% of 40,000) responded. In order to determine if there is any correlation between those customers who responded to the mailer campaign and their corresponding historical recency, the following analysis is done.

The test group of 40,000 customers is sorted in 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 are further divided into five equal groups (20 percent in each group). The top-most group is assigned a recency code of 1, and the next group is assigned a code of 2, and so on, until the bottom-most group is assigned a code of 5. An analysis of the customer response data from the mailer campaign and the recency-based grouping data point out that the mailer campaign received the highest response from those customers grouped in recency code 1, followed by those grouped in code 2, and so on. Figure 6-2 depicts the distribution of percentages of customers who responded fell within the recency code grouping of 1 through 5.

Figure 6-2 indicates the highest response rate (4.5%) for the campaign was from those customers in the test group who had the highest recency quintile (recency code = 1). Note the average customer response rate computed for all five groups would be none other than the actual response rate of 2.02 percent achieved by the campaign. That is, (4.50% + 2.80% + 1.50% + 1.05% + 0.25%)/5 = 2.02 percent.

FIGURE 6-2 **RESPONSE AND RECENCY**



At the end of this recency coding exercise we would assign recency values of $r = 1$ through 5 for groups of customers, depending on the quintile that they fall within.

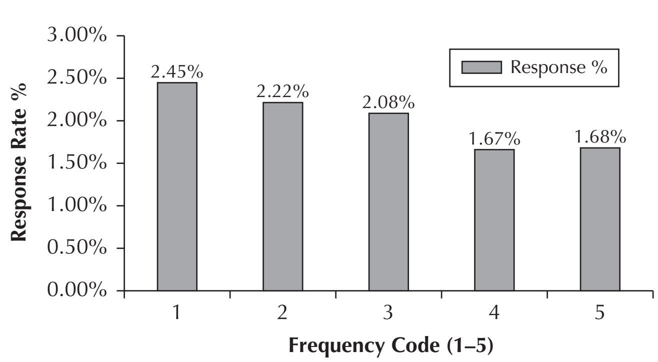
Frequency Coding Frequency coding process is exactly same as the recency coding process just discussed. However, to sort the test group of 40,000 customers based on *frequency* metric, we need to know the average number of purchases made by a customer per month. Of course, the choice of the appropriate time period depends on the usual frequency of purchases. In this case, customers with the highest number of purchases per month are grouped at the top, while those with lesser number of purchases per month were listed below. Here again, the sorted list is grouped into five groups. Those in the top are assigned a code of 1 and those at the bottom a code of 5. The graph below depicts the same.

An analysis of the customer response data from the mailer campaign and the *frequency*-based grouping data point out that the mailer campaign got the highest response from those customers grouped in *frequency* code 1, followed by those grouped in code 2, and so on. Figure 6-3 depicts the distribution of percentage of customers who responded fell within the *frequency* code grouping of 1 through 5.

As it can be seen in Figure 6-3, the highest response rate (2.45%) for the campaign was from those customers in the test group who had the highest *frequency* quintile (frequency code = 1).

At the end of this *frequency* coding exercise we would assign frequency values of $f = 1$ though 5 for groups of customers, depending on the quintile that they fall within.

FIGURE 6-3 **RESPONSE AND FREQUENCY**



Monetary Value Coding The *monetary value* coding process is exactly same as the recency and frequency coding process. However, to sort the test group of 40,000 customers based on *monetary value* metric, we need to know the *average amount purchased per month.* The monetary value is also an important metric driving analysis of customer behavior. As with *recency* and *frequency,* the customer data are sorted, grouped, and coded with a 1 to 5 value.

As can be seen in Figure 6-4, the highest response rate (2.35%) for the campaign was from those customers in the test group who had the highest *monetary value* quintile (monetary value code = 1).

At the end of this *monetary value* coding exercise, we would assign *monetary value* of $m = 1$ through 5 for groups of customers, depending on the quintile that they fall within.

At the end of this whole exercise, you will have individual R, F, and M scores for each customer. Each customer will be assigned to 1 of the 125 groups such as 111, 233, 432 … 555, based on their respective RFM codes.

Limitations This method independently links customer response data with R, F, and M values and then groups customers belonging to specific RFM codes. However, this method may not produce an equal number of customers under each RFM cell. This is because the individual metrics R, F, and M are likely to be somewhat correlated. For example, someone spending more (high M) is also likely, on average, to buy more frequently (high F). However, for practical purposes, it is desirable to have exactly the same number of individuals in each RFM cell. A sorting technique ensuring equal numbers in each RFM cell is described as follows.

RFM Sorting Technique This is a more sophisticated sorting technique which helps to arrive at an RFM code for each customer and ensures the grouping of equal number of customers under each RFM code.

Figure 6-5 depicts a schematic diagram of the logic behind RFM sorting. Consider the list of 40,000 test group customers. The list is first sorted for recency and grouped into five equal groups of 8,000. Therefore group 1 will have 8,000 customers, and so will the other groups through group 5.

Now, take the 8,000 customers in each group and sort them based on frequency and divide them into five equal groups of 1,600 each. At the end of this stage, you will have RF codes starting from 11 through 55, with each group having 1,600 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. Again, we will have RFM codes starting from 111

FIGURE 6-4 RESPONSE AND MONETARY VALUE

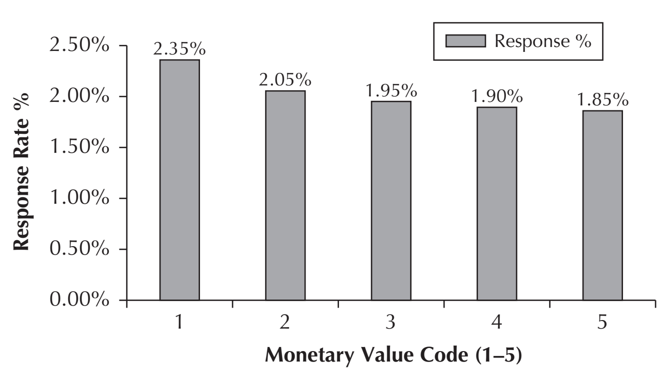
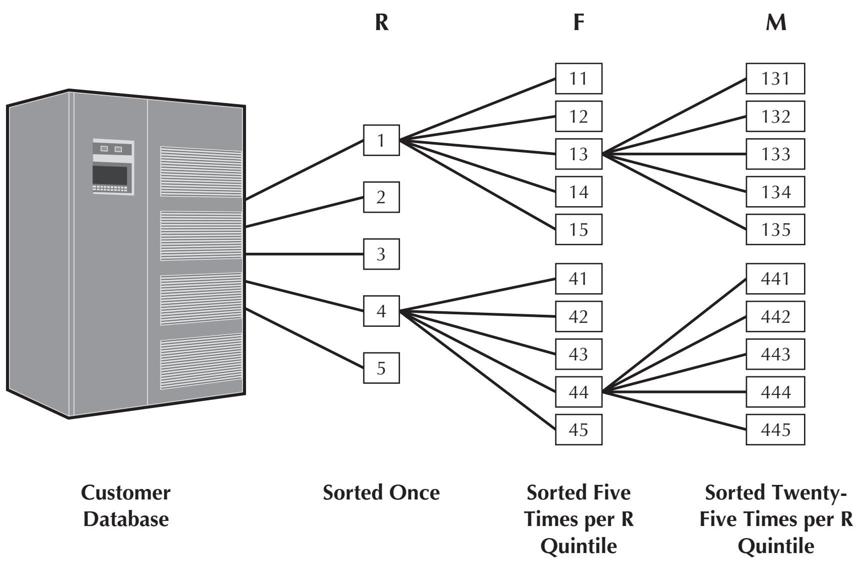


FIGURE 6-5 RFM CELL SORTING



through 555, each having 320 customers. Considering each RFM code as a cell, we will have 125 cells.

5 recency divisions × 5 frequency divisions × 5 monetary value divisions =
125 RFM codes

Breakeven Value Before proceeding to see an application of the RFM codes, it would be useful to review the concept of breakeven value. In marketing literature, *breakeven* refers to the fact that the net profit from a marketing promotion equals the cost associated with conducting the promotion.

$$\text{Breakeven value (BE)} = \frac{\text{Unit cost price}}{\text{Unit net profit}}$$

If this ratio for a particular promotion is 1, then the promotion only broke even and did not generate any net profits. This ratio also computes the minimum response rates required in order to offset the promotional costs involved and thereby not incur any losses.

Consider our example of mailing \$150 discount coupons. Suppose the cost to mail each piece is a dollar, and the net profit (after all costs) is \$45, then the breakeven value or breakeven response rate required can be computed as breakeven value (BE) = \$1/\$45 = 0.0222, or 2.22 percent. This value can be computed for every RFM cell and then can be used to compare it with the actual response rate of each cell.

The breakeven response rate just computed could be used in computing a *breakeven index.* The breakeven index is calculated using the following formula.

$$\text{Breakeven index (BEI)} = \left(\frac{\text{(Actual response rate – BE)}}{\text{BE}}\right) \times 100$$

Therefore, if the actual response rate of a particular RFM cell was 3.5 percent, then BEI = ((3.5% – 2.22%)/2.22%) × 100 = 57.66.

A positive value indicates that some profit was made from the transaction. A BEI value of 0 indicates that the transaction just broke even, and a negative BEI value indicates that the transaction resulted in a loss. See Table 6-5.

TABLE 6-5 COMBINING RFM CODES, BREAKEVEN CODES, BREAKEVEN INDEX

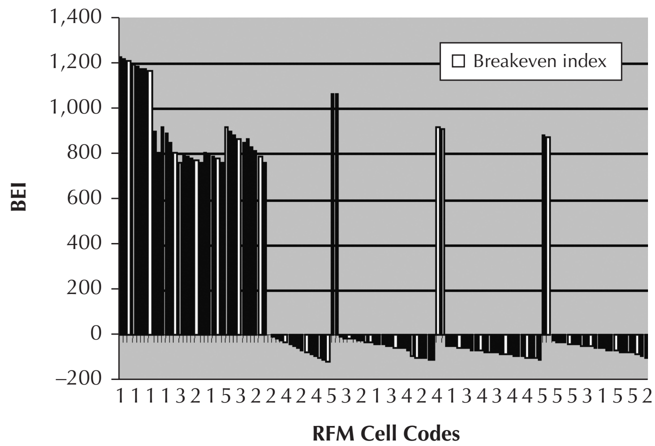
Cell #	RFM Codes	Cost per Mail	Net Profit per Sale	Breakeven %	Actual Response %	Breakeven Index
1	111	$1.00	$45.00	2.22%	17.55%	690
2	112	$1.00	$45.00	2.22%	17.45%	685
3	113	$1.00	$45.00	2.22%	17.35%	681
4	114	$1.00	$45.00	2.22%	17.25%	676
5	115	$1.00	$45.00	2.22%	17.15%	672
6	121	$1.00	$45.00	2.22%	17.05%	667
7	122	$1.00	$45.00	2.22%	16.95%	663
8	123	$1.00	$45.00	2.22%	16.85%	658
9	124	$1.00	$45.00	2.22%	16.75%	654
10	125	$1.00	$45.00	2.22%	16.65%	649
11	131	$1.00	$45.00	2.22%	16.55%	645
12	132	$1.00	$45.00	2.22%	16.45%	640
13	133	$1.00	$45.00	2.22%	16.35%	636
14	134	$1.00	$45.00	2.22%	16.25%	631
15	135	$1.00	$45.00	2.22%	16.15%	627
16	141	$1.00	$45.00	2.22%	16.05%	622
17	142	$1.00	$45.00	2.22%	15.95%	618
18	143	$1.00	$45.00	2.22%	15.85%	613
19	144	$1.00	$45.00	2.22%	15.75%	609
20	145	$1.00	$45.00	2.22%	15.65%	604
21	151	$1.00	$45.00	2.22%	15.55%	600
22	152	$1.00	$45.00	2.22%	15.45%	595
23	153	$1.00	$45.00	2.22%	15.35%	591
24	154	$1.00	$45.00	2.22%	15.25%	586
25	155	$1.00	$45.00	2.22%	15.15%	582
26	211	$1.00	$45.00	2.22%	15.65%	604
27	212	$1.00	$45.00	2.22%	15.55%	600
28	213	$1.00	$45.00	2.22%	15.45%	595
29	214	$1.00	$45.00	2.22%	15.35%	591
30	215	$1.00	$45.00	2.22%	15.25%	586
31	221	$1.00	$45.00	2.22%	15.15%	582
32	222	$1.00	$45.00	2.22%	15.05%	577
33	223	$1.00	$45.00	2.22%	14.95%	573
34	224	$1.00	$45.00	2.22%	14.85%	568
35	225	$1.00	$45.00	2.22%	14.75%	564
36	231	$1.00	$45.00	2.22%	14.65%	559
37	232	$1.00	$45.00	2.22%	14.55%	555
38	233	$1.00	$45.00	2.22%	14.45%	550
39	234	$1.00	$45.00	2.22%	14.35%	546
40	235	$1.00	$45.00	2.22%	14.25%	541

Table 6–5 shows an excerpt from the breakeven index computations for the various RFM cells. The complete table is available in Appendix 2 of this chapter. RFM cells with a corresponding positive BEI value are the groups of customers the marketing campaign should target, and all those RFM cells with corresponding negative BEI values are those customers to be avoided for this promotion.

Rows with positive BEI values have been highlighted. It is interesting to note that of the 125 RFM cells, only customers within 56 cells have a higher chance of offering profitability to the firm, and the rest do not! Therefore, it becomes clear that a firm can save significantly by only focusing on potential profitable customers and leaving the rest.

Figure 6-6 plots the RFM cell codes and their corresponding BEI values. Customers with positive BEI values are to be chosen and the rest are to be left. Note that customers

FIGURE 6-6 RFM CODES VS. BREAKEVEN INDEX



with higher RFM values tend to have higher BEI values. However, at the same time, customers with a lower recency value but relatively higher F and M values also tend to have positive BEI values and hence should be considered for target mailing.

Order of Importance of R, F, and M The industry, in general, uses the RFM technique in the order of recency, frequency, and monetary value. However, the order varies for different industry segments. Although the RFM order is normally acceptable, a more accurate order of coding would depend on the rapidity of customer response rate drops. The metric (R, F, or M) for which the customer response rates drops more rapidly is likely to be the best predictor of future customer response and, hence, could be coded first. Once the metric of highest influence is determined, the same yardstick of measurement can be used to determine the order of the remaining metrics.

Referring to Figures 6-2, 6-3, and 6-4, notice that customer response rate drops more rapidly for the recency metric than the other two metrics. Similarly, the customer response rate for the frequency metric drops more rapidly than the monetary value metric. Therefore, the order of R, F, and M holds good in this case.

RFM and Profitability From Table 6-6, it is clear the firm will benefit significantly more by sending the mailers to select customers within specific RFM cells, than by sending it to their whole customer base.

TABLE 6-6 COMPARISON OF PROFITS

	Test	Full Customer Base	RFM Selection
Average response rate	2.02%	2.02%	15.25%
Number of responses	808	8,080	2,732.8
Average net profit/sale	$45	$45	$45
Net revenue	$36,360	$363,600	$122,976
Number of mailers sent	40,000	400,000	17,920
Cost per mailer	$1.00	$1.00	$1.00
Mailing cost	$40,000.00	$400,000.00	$17,920.00
Profits	($3,640.00)	($36,400.00)	$105,056.00

The loss of $3,650 incurred in conducting the test is offset by the profits generated by by sending mailers to the RFM select customers.

Benefits: RFM techniques provide tangible benefits to the marketing professionals. Organizations need not mindlessly spend on contacting all their customers, but instead can smartly use the RFM selections to identify only those customers with high probabilities of purchase and target their marketing strategies and communications towards that population.

RFM Method 2 In the first RFM method, the RFM values are assigned for each customer by sequential sorting based on the RFM metrics. However, there is an alternative method which uses regression techniques to compute the relative weights of the R, F, and M metrics, and these relative weights are used to compute the cumulative points of each customer. The precomputed weights for R, F, and M, based on a test sample, are used to assign RFM scores to each customer (see Appendix 1). The higher the computed score, the more profitable the customer is likely to be in the future. This method, unlike the earlier one, is flexible and can be tailored to each business situation.

Example[2] Three customers have a purchase history calculated over a twelve-month period (see Tables 6-7 to 6-10). For every customer, numerical points have been assigned to each transaction according to a historically derived R/F/M formula. The relative weight

TABLE 6-7 RECENCY SCORE

Customer	Purchases (Number)	Recency (Months)	Assigned Points	Weighted Points
John	1	2	20	100
	2	4	10	50
	3	9	3	15
Smith	1	6	5	25
Mags	1	2	20	100
	2	4	10	50
	3	6	5	25
	4	9	3	15

Points for Recency: 20 points if within past 2 months; 10 points if within past 4 months; 05 points if within past 6 months; 03 points if within past 9 months; 01 point if within past 12 months; Relative weight = 5

TABLE 6-8 FREQUENCY SCORE

Customer	Purchases (Number)	Frequency	Assigned Points	Weighted Points
John	1	1	3	6
	2	1	3	6
	3	1	3	6
Smith	1	2	6	12
Mags	1	1	3	6
	2	1	3	6
	3	2	6	12
	4	1	3	6

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

TABLE 6-9 MONETARY VALUE SCORE

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

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

TABLE 6-10 RFM CUMULATIVE SCORE

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

based on the importance assigned to each of the three variables, R, F, and M on the basis of an analysis carried out on past customer transactions is as follows:

Recency = 5, Frequency = 2, Monetary = 3

The resulting cumulative scores, 249 for John, 112 for Smith, and 308 for Mags, indicate a potential preference for Mags. In this case, John seems to be a good prospect as well, but mailing to Smith might be a misdirected marketing effort. This example illustrates a simple application of the RFM technique. In practice, however, the number of customers to be analyzed can run into millions. Regression techniques are often employed to arrive at the relative weights for RFM. See Appendix 1 for an overview of regression models.

6.3.2 PAST CUSTOMER VALUE

Past customer value is a model which extrapolates the results of past transactions into the future. The value of a customer is determined based on the total contribution (toward profits) provided by the customer in the past. This modeling technique assumes past performance of the customer indicates the future level of profitability. Since products/services are bought at different points in time during the customer's lifetime, all transactions must be adjusted for the time value of money.

$$\text{Past customer value of a customer} = \sum_{n=1}^{n} GC_{in} \times (1 + r)^{n} \qquad (7)$$

where
i = number representing the customer
r = applicable discount rate (for example 15% per annum or 1.25% per month)
n = number of time periods prior to current period when purchase was made
GC_{in} = Gross contribution of transaction of the ith customer in the nth time period

Example: If we have data on the products purchased by various customers over a period of time, the value of the purchases, and the contribution margin, we can compare the value generated by each customer by computing all transactions in terms of their present value. Table 6-11 illustrates this computation for one customer.

$$\text{Gross Contribution (GC)} = \text{Purchase Amount} \times 0.3 \quad \textbf{(8)}$$

$$\text{Past Customer Value Scoring} = 6(1 + 0.0125) + 9(1 + 0.0125)^2 + 15(1 + 0.0125)^3 + 15(1 + 0.0125)^4 + 240(1 + 0.0125)^5 = 302.01486 \quad \textbf{(9)}$$

This customer is worth \$302.01 in contribution margin, expressed in net present value in May dollars. By comparing this score to a set of customers, we arrive at a prioritization for directing future marketing efforts. The customers with higher values are normally the customers deserving greater marketing resources. This method, while extremely useful, does not incorporate other useful information which could help refine the process of selecting profitable customers. For instance, it does not consider whether a customer is going to be active in the future. It also does not incorporate the expected cost of maintaining the customer in the future. All in all, it is still a backward-looking metric.

6.3.3 LTV METRICS (NPV MODELS)

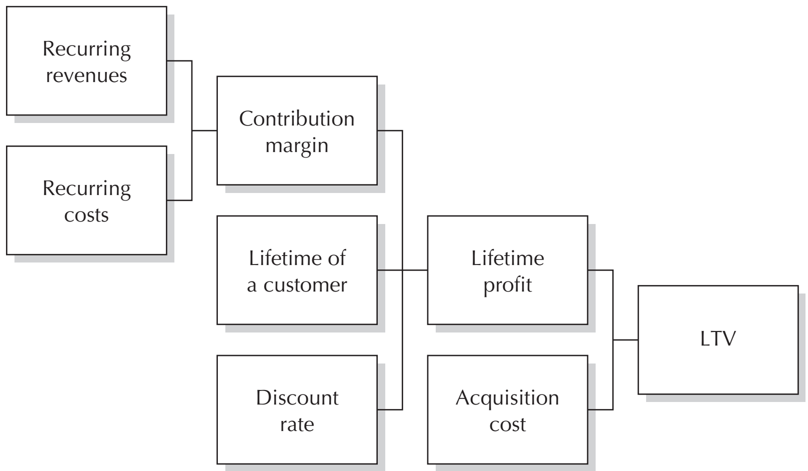
LTV stands for lifetime value and NPV stands for net present value. Evaluating the long-term economic value of a customer to the firm has seen a dramatic rise in interest. This is a direct outcome of the shift from transactional marketing to relational marketing. If a manager wants to evaluate marketing resource allocation plans targeted at improving the long-term value of customers, corresponding control measures must be put in place. Looking at profits on a per-transaction basis is not sufficient. Managers want to have an idea how the value of a client has evolved *over time*. The general term used to describe the long-term economic value of a customer is *lifetime value.* In very simple terms, it is a multiperiod evaluation of a customer's value to the firm. However, the term LTV is not without controversies. Although it is the Holy Grail for some, others call it "an elaborate fiction of presumed precision."[3] In the following section we present some of the most common ways to calculate LTV. Nevertheless, the reader should be aware that there are many specific formulations. It is important to present the principle in such a manner that readers can adapt the calculation to their own requirements. Conceptually, the principle of calculating LTV is represented in Figure 6-7.

As one can easily see, there is no a single way to arrive at each component. For example, are the recurring costs comprised only of direct product costs or also marketing, sales, and service costs? Depending on many factors, such as nature of product, data availability, and statistical capabilities, the inputs for the LTV calculation change. Is this problematic?

TABLE 6-11 SPENDING PATTERN OF A CUSTOMER

	Jan	Feb	March	April	May
\$ Amount	800	50	50	30	20
GC	240	15	15	9	6

FIGURE 6-7 PRINCIPLES OF LTV CALCULATION



It is not. First and foremost, it is important to understand the principle. Based on the general principle, one can then start to adjust the calculation to the reliably available data. Also, one needs to adapt the formulation to the industry and company context. For example, having a defined finite lifetime duration (as for a contractual relationship such as cable subscription) makes for a different formulation as having non-finite relationship (as for a noncontractual relationship such as a supermarket).

After discussing the basic formulations, we will highlight key issues that should be considered when employing the models. In the following discussion, we will present different formulations of the same principle.

Example 1

In the most simple definition, the lifetime value of an individual customer i is the sum of the customer's discounted contribution margins over the respective observation horizon (T).

$$LTV = \sum_{t=1}^{T} CM_t \left(\frac{1}{1+\delta}\right)^t \qquad \textbf{(10)}$$

where

LTV = lifetime value of an individual customer in \$
CM = contribution margin
δ = interest rate
t = time unit
Σ = summation of contribution margins across time periods

The resulting LTV is a measure of a single customer's worth to the firm. The contribution margin CM may vary, of course, across customers and across time units. This formulation is primarily used for pedagogical and conceptual purposes. It is typically based on past customer behavior and may have limited diagnostic value for future decision making.

Where does information come from?

The contribution margin CM and the duration T are derived either from managerial judgment or comes from actual purchase data (see previous section on how to calculate lifetime duration). The interest rate is a function of a firm's cost of capital and can be obtained from the financial accounting function.

Cautionary note: If the time unit is different from a yearly basis, the interest rate δ needs to be adjusted accordingly. For example, if the yearly interest rate is 15 percent, the quarterly interest rate is 3.75 percent.

Example 2

On the next level, one can break down the contribution margin into its constituting elements.

$$LTV = \sum_{t=1}^{T} (S_{it} - DC_{it}) - MC_{it} \left(\frac{1}{1+\delta}\right)^{t} \tag{11}$$

where

LTV = lifetime value of an individual customer in \$

S = Sales to customer i

DC = direct cost of products purchased by customer i

MC = marketing cost of customer i

The cost element in this example is broken down into direct product-related cost and marketing cost. Depending on data availability, it can be enhanced by including service-related cost, delivery cost, or other relevant cost elements.

Where does information come from?

The information on sales, direct cost, and marketing cost comes from internal company records. The key issue with direct cost and marketing cost is that they must be known on a per-customer basis. This knowledge is not necessarily common among many firms. An increasing number of firms are installing activity-based costing (ABC) schemes. ABC methods are used to arrive at appropriate allocations of customer and process-specific costs.

6.3.4 CUSTOMER EQUITY

The sum of the lifetime value of all a firm's customers represents the customer equity of a firm. This metric is an indicator of how much the firm is worth at a particular point in time as a result of the firm's customer management efforts. This metric therefore can be seen as a link to the shareholder value of a firm.

Example 3

Building on the formulation in equation (11), we can aggregate the LTV measure across customers. The resulting quantity is the customer equity (CE).

$$CE = \sum_{i=1}^{I} \sum_{t=1}^{T} CM_{it} \left(\frac{1}{1+\delta}\right)^{t} \tag{12}$$

where

CE = customer equity of customer base in \$ (sum of individual lifetime values)

In this case, the CE measure gives the economic value of an entire cohort or segment of customers—subject to the same restrictions as in example 1.

Example 4

So far, an assumption was that all customers under investigation remain fully active during the period of interest. However, in reality, more and more customers stop their relationship with the firm over time. The next step is therefore to consider customer retention probabilities. This relates to the fact that customers tend to remain in the relationship only with a certain average retention rate *Rr*.

$$LTV = \left(\sum_{t=1}^{T} \left(\prod_{t=1}^{T} Rr \right) CM_{it} \left(\frac{1}{1+\delta} \right) \right)^{t} - AC \qquad \textbf{(13)}$$

where

LTV = lifetime value of an individual customer in \$

Rr = retention rate

Π = Product of retention rates for each time period from 1 to T

AC = acquisition cost

T = total time horizon under consideration

The retention rate in this example is constant (i.e., it does not vary across time or customers). Although this is a common assumption, it is mostly not very realistic, as was previously discussed. Also, the acquisition cost (*AC*) is now subtracted from the customer's value.

Example 5

Equation (13) can be simplified to the following formula under the assumption that $T \rightarrow \infty$ and that the contribution margin CM does not vary over time.

$$LTV_i = \frac{CM}{1 - Rr + \delta} - AC \qquad \textbf{(14)}$$

This formulation is easy for quick calculations and, unless the retention rate is very high, produces results very close to the more precise formulation. One needs only to multiply the contribution margin with the margin multiple $(1/(1 - Rr + \delta))$ and subtract the acquisition cost.

How Long Is Lifetime Duration?

The word *lifetime* must be taken in many circumstances with a grain of salt. Although the term makes little sense with one-off purchases (say, for example, a house), it also seems strange to talk about LTV of a grocery shopper. Clearly, there is an actual lifetime value of a grocery shopper. However, given the long time span, this actual value is not practical. For all practical purposes, the lifetime duration is a longer-term duration used managerially. For example, in a direct marketing general merchandise context, managers do not look beyond a four-year time span. Beyond that, any calculation and prediction may become difficult due to so many uncontrollable factors (the customer moves, new competitors move in, and so on). It is therefore important to make an educated judgment regarding a sensible duration horizon in the context of making decisions.[4]

Comprehensive Example

The following example illustrates some aspects of the previously introduced LTV and customer equity models (see Table 6-12). The observation horizon for this example is five years (column 1). A company targets a list of 10,000 purchased addresses with an acquisition campaign. The company acquires 1,000 customers through the target mailing; thus, the acquisition rate is 10 percent. Once acquired, a customer generates on average \$120 in

TABLE 6-12 CUSTOMER EQUITY CALCULATION EXAMPLE

1 Year from Acquisition	2 Sales per Customer	3 Manufacturer Margin	4 Manufacturer's Gross Margin	5 Marketing and Servicing Costs	6 Actual Retention Rate	7 Survival Rate	8 Expected Number of Active Customers	9 Profit per Customer per Period to Manufacturer	10 Discounted Profit per Customer per Period to Manufacturer
0	120	0.3	36	20	0.4	0.4	400	16	16
1	120	0.3	36	20	0.63	0.25	250	16	14
2	120	0.3	36	20	0.75	0.187	187	16	12
3	120	0.3	36	20	0.82	0.153	153	16	11
4	120	0.3	36	20	0.85	0.131	131	16	9
Total customer equity									

sales (column 2). For simplicity's sake, this level of sales is assumed to be constant over the lifetime of the customers (column 2). The margin of the firm is 30 percent, resulting in a constant gross margin (column 3). Marketing and service cost while alive are constant as well (column 5). The retention rate in the first period is 40 percent and then increases over time, as the loyal customers remain. The resulting number of remaining customers in each period is shown in column 8. The profit per customer (column 9) is computed by subtracting the marketing and service cost from the gross margin. This per-period is discounted to present value with a yearly rate of 15 percent (column 10). The remainder is to multiply yearly discounted profit times the remaining customers in each year. Then sum these values up to the total customer equity of this group or cohort of customers (column 11).

Customer equity share (CES) is an alternative metric to market share that takes customer lifetime value (CLV) into account. CES for each brand *j* can be calculated using the following formula.

$$CES_j = \frac{CE_j}{\sum_k CE_k} \tag{15}$$

where
CE = customer equity
j = focal brand
k = all brands

6.4 Popular Customer Selection Strategies

Customer selection strategies are applied when firms want to target individual customers or groups of customers. The reason for targeting these customers can be manifold, for example, for sending out a promotion or for inviting them to a special event. Finding the right targets for marketing resource allocation is at the heart of any CRM strategy. Smart targeting allows firms to spend resources judiciously and allows customers to receive messages relevant to them. Dumb targeting actions destroy value by over- or underspending from the firm's perspective and by providing undesirable messages (*junk mail*). One step in the successful implementation of CRM is the smart deployment of targeting methodologies to maximize the benefits to firm and customer.

CRM AT WORK 6.1

TESCO

Tesco, the British supermarket chain has a very successful loyalty program in place. In addition, the company has built up distinct analysis and targeting capabilities, which allows Tesco to customize its promotions to a large number of different segments. Consider the following case: Using the loyalty program, Tesco builds a market basket history for each cardholder. It can then analyze basket content, and, based, on the current basket, can propose certain items via its promotions. Tesco observed that a certain group of customers never buy meat from Tesco. The underlying hypothesis could be that these customers buy their meat at the local butcher because they believe it is superior to Tesco's. This would mean Tesco could send them coupons in order to entice them to buy Tesco's meat. However, a completely different interpretation could be that these customers are vegetarians. This, then, would mean if Tesco were to send them coupons, the company would basically show it doesn't really understand its customer. In reality, Tesco plays it safe and doesn't send any coupons for meat. They might forgo some purchases from those who do buy at the local butcher, but Tesco prefers to not harm the relationships with those who are vegetarian.

This is an example where a company combines analytical skills, judicious judgment, knowledge about consumer behavior, and careful targeting of customers. It also shows that targeting always happens in a business context and is not an isolated activity.

6.4.1 DECISION TREES

Using decision trees is a methodology for finding the best predictors of a 0/1 dependent variable. For example, a company wants to know the differential demographic characteristics of loyalty program members versus nonmembers. Decision trees are especially useful when there is a large set of potential predictors for a model. In such a case, it may be difficult to determine which predictors are the most important or what the relationships between the predictors and the target (dependent) variable are. Decision tree algorithms can be used to iteratively search through the data to find out which predictor best separates the two categories of a binary target variable. Typically, this search is performed on two thirds of the available data, with one third of the data reserved for later use for testing the eventual model that develops. Example: Customer data for purchases of hockey equipment from sporting goods catalog.

Step 1: The algorithm is run. It determines that separating the customers based on gender is the optimum first step. See Figure 6-8.

Step 2: The algorithm is run again and determines the optimum approach is to separate the male customers by whether they have bought scuba equipment in the past and to separate the female customers by marital status. See Figure 6-9.

This process can be repeated for each subsegment, until all of the predictors have been applied to the model. When this process is complete, a tree will be developed in which segments are nested within segments. The profitable segments can then be identified for use as target markets.

One problem with the decision tree approach is that it is prone to over fitting, whereby segments are tailored to very small segments, (based upon the dataset that was used to create the tree) and as a result, the model developed will not perform nearly as well on a separate dataset. At this point, the one third of the data left for testing can be analyzed with the decision tree model developed. If the results show a large discrepancy with what was expected, then the model will need to be reevaluated. If the results are within range of what is predicted from the model, it is likely that the model is a good fit.

FIGURE 6-8 STEP 1: CUSTOMER DATA FOR PURCHASES OF HOCKEY EQUIPMENT FROM SPORTING GOODS CATALOG

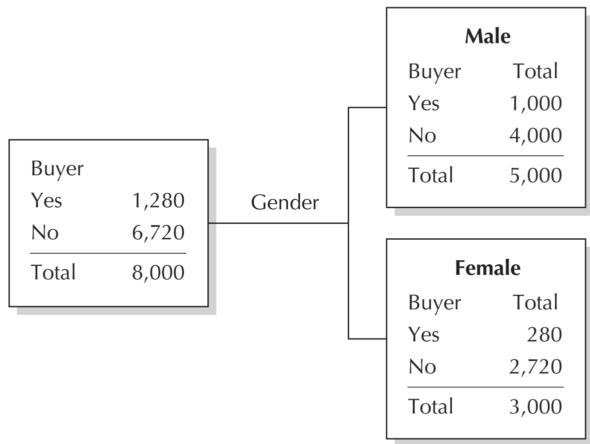
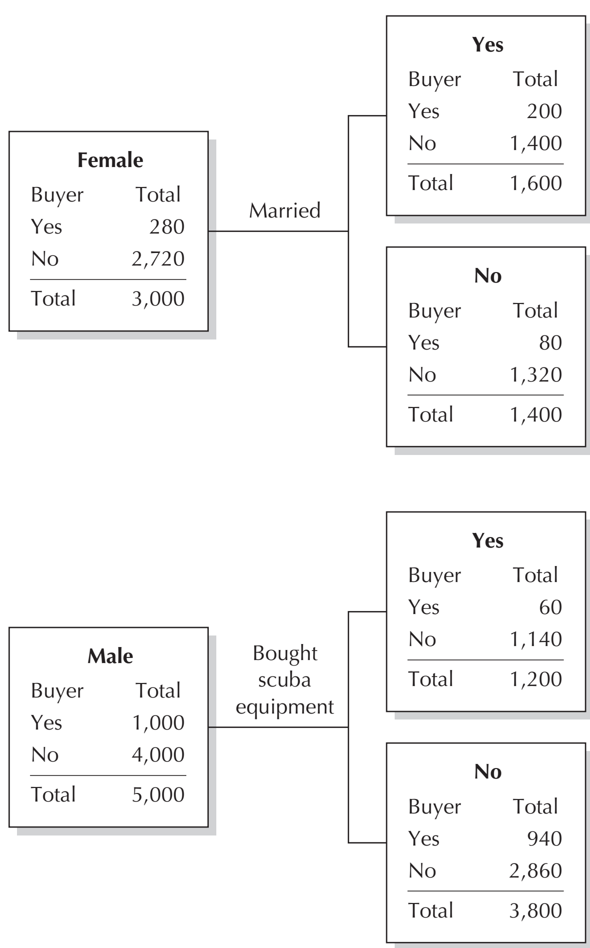


FIGURE 6-9 STEP 2: CUSTOMER DATA FOR PURCHASES OF HOCKEY EQUIPMENT FROM SPORTING GOODS CATALOG



6.4.2 LOGISTIC REGRESSION

Linear regression starts with the specification of the dependent variable and the independent variable. For example the number of people entering a store on a Saturday is the dependent variable of interest and the amount of money the store spent on advertising on Friday is the independent variable. We expect to see a linear relationship between the two variables. A regression analysis creates an estimate of the coefficient that represents the effect of advertising on store traffic. In this situation we note store traffic—the dependent variable—can take on a large range of values. However, in marketing we often come across situations where the dependent variable is binary. For example, in a situation where we are interested in whether a customer bought a product or not, we assign the value 0 to this variable when the customer does not buy, and the value 1 when the customer buys. Linear regression if used when the dependent variable takes on only two values, 0 or 1, and when we interpret the intermediate values (e.g., 0.4, 0.5, etc.) as representing the probability of the desired outcome, is called a linear probability model. However, making statistical inferences using linear probability models poses problems overcome by using logistic regression. *Logistic regression* is the method of choice when the dependent variable is binary and assumes only two discrete values. For example:

1. Whether a customer responded to a marketing campaign or not
2. Whether a person bought a car or not

These observed values for the dependent variable take on only two values and are usually represented using a 0–1 dummy variable. The mean of a 0–1 dummy variable is equal to the proportion of observations with a value of 1, and can be interpreted as a probability. Furthermore, the predicted values in a logistic regression fall between 0 and 1 and are also interpreted as probabilities. For example, home ownership as a function of income can be modeled whereby ownership is delineated by a 1 and nonownership a 0. The predicted value based on the model is interpreted as the probability that the individual is a homeowner. With a positive correlation between increasing income and increasing probability of ownership, we can expect to see results where the predicted probability of ownership is .22 for a person with an income of $35,000, and .95 for a person with a $250,000 income.

Example: Credit Card Offering

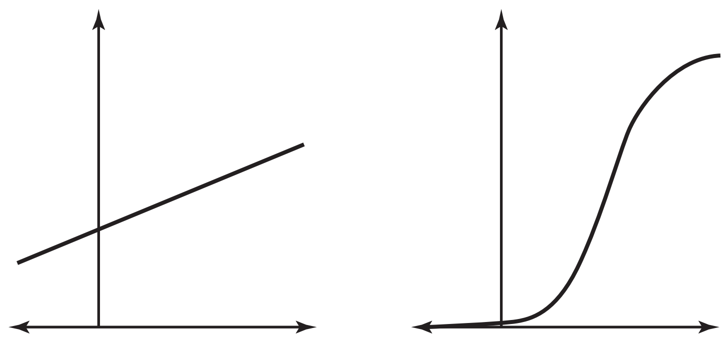
Using logistic regression to identify potential targets for marketing credit cards offers to:

- Dependent variable—whether or not the customer signed up for a gold card offer
- Predictor variables—other bank services the customer used plus financial and demographic customer information

By inputting values for the predictor variables for each new customer, the logistic model will yield a predicted probability. Customers with high predicted probabilities may be chosen to receive the offer because they seem more likely to respond positively.

Unlike in linear regression where the effect of one unit change in the independent variable on the dependent variable is assumed to be a constant represented by the slope of a straight line, for logistic regression the effect of a one-unit increase in the predictor variable varies along an s-shaped curve (see Figure 6-10). This means that at the extremes, a one-unit change has very little effect, but in the middle a one-unit change has a fairly large effect. In the case of the income versus home ownership example, the difference in the likelihood that an individual owns a home may not change much as income increases from $10,000 to $30,000 or from $1,000,000 to $1,020,000, but may increase considerably if income increases from $50,000 to $70,000.

FIGURE 6-10 LINEAR AND LOGISTIC REGRESSION



Mathematically, linear regression takes the form

$$y = \alpha + \beta x + e$$

where y is the dependent variable, x is the predictor variable, α represents a constant, β represents the effect of x on y, and e is the error term. In this sort of regression, y can take on any value between negative infinity and positive infinity. However, as we noted earlier, in many instances we normally observe only binary activity represented by 0 or 1. Hence, if the actually observed dependent variable has to be constrained between 0 and 1 to indicate the probability of an event occurring, a transformation is necessary. This transformation is the basis of logistic regression. The steps in the transformation are given as follows.

Step 1 If p represents the probability of an event occurring, consider the ratio $\frac{p}{1-p}$. Since p is a positive quantity less than 1, the range of this expression is 0 to infinity.

Step 2 Take the logarithm of this ratio $\log\left(\frac{p}{1-p}\right)$.
This transformation allows the range of values for this expression to lie between negative infinity and positive infinity.

Step 3 The value $\log\left(\frac{p}{1-p}\right)$ can now be considered as the dependent variable.

A linear relationship of this value with predictor variables in the form $z = \alpha + \beta x + e$ can be written out. The α and β_s can be estimated.

Step 4 In order to obtain the predicted probability p, the following back transformation is necessary:

$$\text{Since } \log\left(\frac{p}{1-p}\right) = z = \alpha + \beta x + e$$

$$\frac{p}{1-p} = e^z$$

This allows us to calculate the probability p of an event occurring, the variable of interest, as

$$p = \left(\frac{1}{1+e^{-z}}\right)$$

6.5 Techniques to Evaluate Alternative Customer Selection Strategies

These techniques revolve around constructing models to predict or classify customer data, with the goal of discovering a superior business method than what is already being implemented. The most common way to assess a model's performance is to compare its performance to the results expected if no model was used. This gives a baseline measure of how good the model is. The value of a predictive model can then be assessed by using the model to rank or score a sample set of customers and contacting or targeting the customers based on the results of the model.

6.5.1 LIFT CHARTS

Lifts indicate how much better a model performs than the *no model* or average performance. Lifts can be used to track a model's performance over time, or to compare a model's performance on different samples.

To calculate lifts for different deciles of data, we need the following information about a model with regard to the set of customer data.

- Cumulative number of customers: The number of total customers up to and including that decile
- Cumulative percentage of customers: The percent of total customers up to and including that decile
- Cumulative number of buyers: The number of buyers up to and including that decile
- Response rate: The actual response rate for each decile, computed by dividing the number of buyers by the number of customers for each decile

The lift = (Response rate for each decile) ÷ (Overall response rate) × 100. And the cumulative lift = (Cumulative response rate) ÷ (Overall response rate) × 100. The cumulative response rate = Cumulative number of buyers ÷ Cumulative number of customers (see Table 6-13).

TABLE 6-13 LIFT AND CUMULATIVE LIFT

Decile	Number of Customers	Number of Buyers	Response Rate	Lift	Cumulative Lift
1	5,000	1,759	35.18%	3.09	3.09
2	5,000	1,126	22.52%	1.98	5.07
3	5,000	998	19.96%	1.75	6.82
4	5,000	554	11.08%	0.97	7.80
5	5,000	449	8.98%	0.79	8.59
6	5,000	337	6.74%	0.59	9.18
7	5,000	221	4.42%	0.39	9.57
8	5,000	113	2.26%	0.20	9.76
9	5,000	89	1.78%	0.16	9.92
10	5,000	45	0.90%	0.08	10.00
Total	50,000	5,691	11.38%		

Lift Performance Illustration

The decile analysis in Figure 6-11 allows us to distribute customers into ten equal-sized groups. In a model that performs well, customers in the first decile exhibit the highest response rate. The response rate continues to drop as we proceed along the deciles. See Figure 6-12.

For the top decile in this case, the lift is 3.09. This indicates that by targeting only these customers we would expect to yield 3.09 times the number of buyers found by randomly mailing the same number of customers. In contrast, the last decile (decile 10) has only (.08 times) the number of buyers as one would expect in a random sample of the same size. Lift is an index that indicates the model's ability to beat the no model case or average performance. Lifts that exceed 1 indicate better than average performance, and those that are less than 1 indicate a poorer than average performance. Also keep in mind that lift is a relative index to a baseline measure.

The cumulative lifts for the model in Figure 6-13 reveal the proportion of responders we can expect to gain from targeting a specific percent of customers using the model. (If we choose the top 30 percent of the customers from the top three deciles, we will obtain 68

FIGURE 6-11 DECILE ANALYSIS

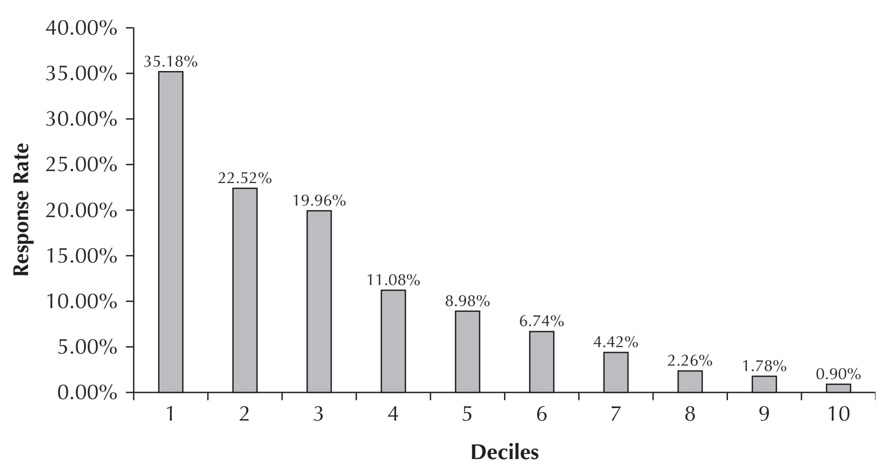


FIGURE 6-12 LIFT ANALYSIS

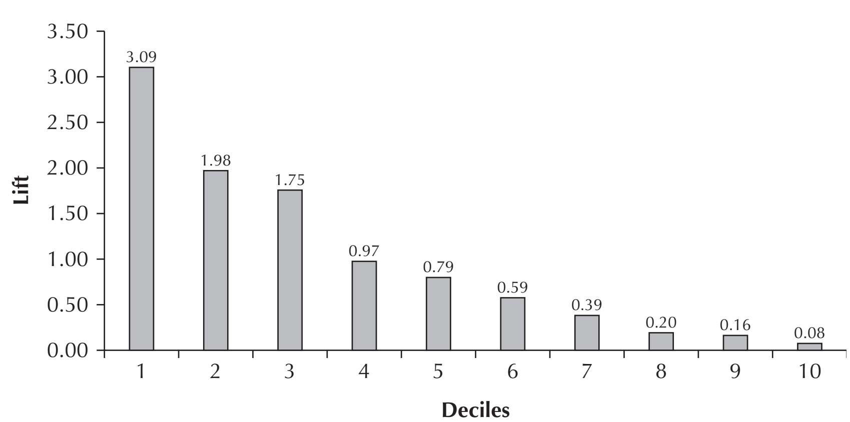


FIGURE 6-13 CUMULATIVE LIFT ANALYSIS

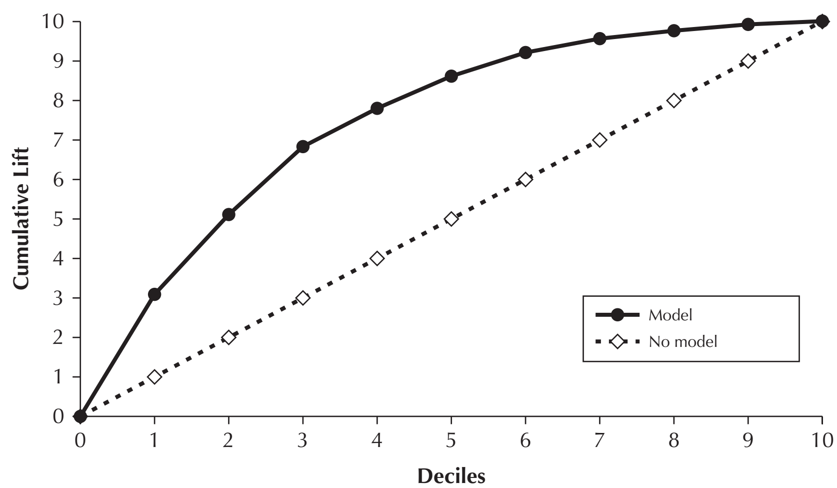
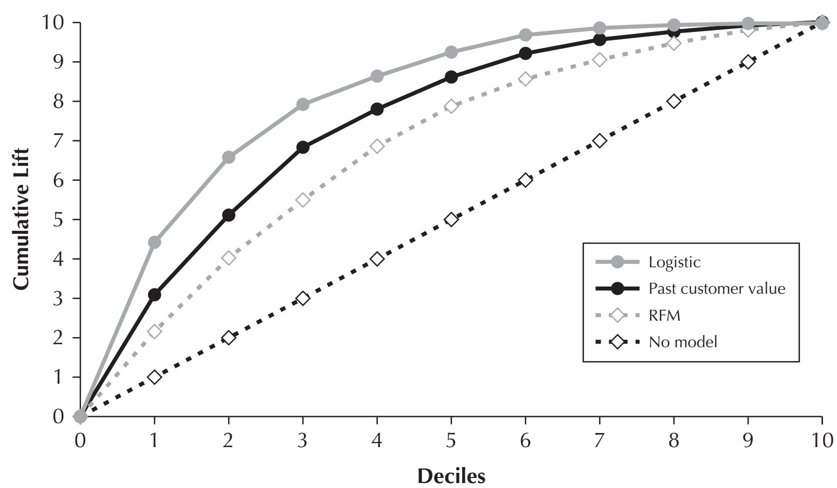


FIGURE 6-14 MODEL COMPARISON



percent of the total responders.) The larger the distance between the model and no model lines, the stronger or more powerful is the model. In a nutshell, lifts can be used to compare two or more alternative models, track a model's performance over time, or to compare a model's performance on different samples.

Past experience shows logistic models tend to provide the best lift performance. This is seen from the topmost curve in the lift chart in Figure 6-14. The model is better able to identify the best customers and group them in the first few deciles. The past customer value approach provides the next best performance, whereas the traditional RFM approach, though by no means redundant, exhibits the poorest performance.

6.6 SUMMARY

Firms use different surrogate measures of customer value to prioritize their customers and to differentially invest in them. Size of wallet is the buyer's total spending in a category and usually firms are interested in acquiring and retaining customers with large wallet sizes. Share of category requirement (SCR) is an aggregate level measure of the proportion of the category volume accounted for by a brand or a focal firm. SCR is one of the most commonly accepted measures of customer loyalty for FMCG categories. Although this is an overall indicator of customer loyalty, it does not necessarily indicate whether a high SCR customer will generate substantial revenues or profits, for which the knowledge about customer's size of wallet is necessary. Share of wallet (SW) is the proportion of category *value* accounted for by a focal brand or firm *within its base of buyers.* At an *individual* level, SW is defined as the proportion of category *value* accounted for by a focal brand or firm for a buyer from all brands the buyer purchases in that category. It indicates the degree to which a customer satisfies his or her needs in the category with a focal brand or firm. Firms can use the information about size of wallet and share of wallet together for optimal allocation of resources. Expected share of wallet can be obtained by using a transition matrix. The transition matrix provides us with the probability a customer will purchase a particular brand if we know which brand the customer purchased previously.

Popular strategic customer-based value metrics are RFM, past customer value, lifetime value, and customer equity. RFM is a composite score of recency, frequency, and monetary value. Two methods used for computing and applying RFM are sorting techniques and calculation of relative weights for R, F, and M. In the sorting technique the RFM values are assigned for each customer by sequential sorting based on the RFM metrics. One of the sorting techniques independently links customer response data with R, F, and M values and then groups customers, belonging to specific RFM codes. However, this method might not produce equal number of customers under each RFM cell. A more sophisticated sorting technique ensures equal numbers in each RFM cell. This is done by first sorting the list of customers by recency and dividing them into equal groups. Each of these groups is then sorted based on frequency and divided into groups of equal size. Each of these subgroups is sorted on monetary value and divided into groups of equal size. This way, each RFM cell will have equal number of customers. Breakeven value and breakeven index can be calculated for each RFM cell. With respect to a promotion, *breakeven* refers to the fact that the net profit from a marketing promotion equals the cost associated with conducting the promotion. Breakeven value (BE) = Unit cost price / Unit net profit. The breakeven value or breakeven response rate can be used in computing a breakeven index, calculated as Breakeven index (BEI) = ((Actual response rate – BE)/BE) $\times$ 100. A positive BEI indicates some profit was made from the transaction. A BEI value of 0 indicates the transaction only broke even, and a negative BEI value indicates the transaction resulted in a loss. An alternate method for RFM metric uses regression technique to arrive at the relative weights of the R, F, and M metrics, and these relative weights are used to compute the RFM score of each customer. The higher the computed RFM score, the more profitable the customer is likely to be in the future. This method has the advantage of being flexible, and it can be tailored to each business situation.

Another important customer-based metric is past customer value, in which the value of a customer is determined based on the total contribution (toward profits) provided by the customer in the past after adjusting for the time value of money. In another metric called lifetime value (LTV), the long-term economic value of a customer is calculated as the sum of the discounted contribution margins over the respective observation horizon. The contribution margin can be computed knowing the values of sales, direct costs, and marketing costs. The sum of the lifetime value of all the customers of a firm represents the customer equity of a firm, another metric which is an indicator of how much the firm is worth at a particular point in time as a result of the firm's customer management efforts. This metric therefore can be seen as a link to the shareholder value of a firm. The formula for computing customer equity incorporates the average retention rate for a firm's customers.

Firms employ different customer selection strategies to target the right customers. Some of the popular customer selection strategies are *decision trees* and *logistic regression.* decision tree algorithm is a recursive partitioning algorithm used for finding the predictors which best separate the two categories of a 0/1 dependent variable. Typically, this search is performed on two thirds of the available data with one third of the data reserved for testing the eventual model that will be developed. Logistic regression is a statistical tool to predict the probability of a binary outcome using predictor variables.

Lift analysis, decile analysis, and cumulative lift analysis are various techniques firms use to evaluate alternative selection strategies. Lifts indicate how much better a model performs than the no model or average performance, and is calculated as (Response rate for each decile) ÷ (Overall response rate) × 100. In decile analysis, customers are grouped into deciles. For a good model, customers in the first decile have the highest response rate, and subsequently the response rate drops. Cumulative lift is calculated as (Cumulative response rate) ÷ (Overall response rate) × 100. It reveals the proportion of responders we can expect to gain from targeting a specific percent of customers using the model.

6.7 EXERCISE QUESTIONS

1. How would you implement the recommended strategies in Figure 6-1? What are some specific marketing actions you would take in the four quadrants?
2. A hotel chain wants to analyze its customer base with RFM. Describe the data fields (variables) in the database necessary to do this.
3. Whatever RFM analysis can do, regression analysis can do as well. Evaluate this statement.
4. How will you use lift charts to determine future marketing action?
5. Describe three business situations where you would consider using logistic regression as the preferred technique for analysis and decision making.
6. What is the link between customer lifetime value and the profitability of an organization?

MINI CASE 6.1

CATALINA IS CHANGING SUPERMARKET SHOPPER MEASUREMENT

Catalina Inc. is a Florida-based company specializing in supermarket shopper tracking and coupon issuing. The company has about 1,600 employees and operates in the United States, as well as in major European countries. The company built its business model on issuing coupons to grocery shoppers online when they check out. The basis for this business model is that traditional print media has long production lead times, and the response to these media is not measurable on the individual customer level. Thus, supermarkets and manufacturers cannot run and track individualized campaigns with traditional media. Catalina's system consists of a printer connected to the cashier's scanner and a database. The information on each shopping basket that checks out via the scanner is then stored in the database. Using the person's credit card number or check number, the database links individual shopping baskets over time. If the person pays cash, the system cannot link the basket. The system then allows both manufacturers and retailers to run individualized campaigns based on the information in the database. For example, Catalina could partner with the retailer to improve its cross selling. A typical issue for any given retailer is that certain customers use the store as their primary shopping location, whereas others use it as their secondary store. To improve the share of wallet with the latter group of customers, Catalina first investigates basket composition of the various buyers. It then finds that certain buyers buy, for example, baby or children products (thus, there is apparently a family behind this shopping basket), yet the number of calories in that basket does not match that of an average family. One explanation for this might be that this shopper uses this outlet as a secondary store. Given this interpretation, the decision then is to allocate to this customer a gift of say $10, for shopping for four weeks in a row spending at least $40 per week in the store. The goal is to selectively target those shoppers of whom the store captures only a low share of wallet, and to entice them to change their behavior.

Questions:

1. Explain why Catalina's approach is superior from a retailer's perspective vis-à-vis the traditional mass media approach.
2. Discuss the role of traditional metrics (such as market share) in this new CRM environment. Should they be discarded?
3. Do you think Catalina's practice (which is entirely legal) is ethically acceptable?

MINI CASE 6.2

DIFFERENTIATING CUSTOMER SERVICE ACCORDING TO CUSTOMER VALUE AT AKZO NOBEL, NV

Akzo Nobel, headquartered in Arnhem, Netherlands, is one of the world's largest chemical manufacturers and also one of the world's largest paint makers. Its chemical unit produces pulp and paper chemicals, functional chemicals (including flame retardants and animal feed additives), surfactants (used in detergents and personal care products), polymers, and catalysts. The polymer division, which serves exclusively the B-to-B market, established a tiered customer-service policy in the early 2000s. The necessity to introduce this policy came out of the recognition that customers are becoming more and more demanding when it comes to asking for additional services around the purchase of polymer products. For example, customers became increasingly aware that Akzo Nobel offers product disposal services, delivers nonstandard pallet sizes, and offers negotiable delivery times. On the one hand, Akzo Nobel did not typically charge for these services, and on the other hand, while extended services were typically offered only to medium- to high-volume clients, the relationship between service offerings and *client profitability* was not clear. When Akzo Nobel investigated more closely who received the service offerings, it found the correlation with client profitability (i.e., contribution margin) was *very* low. This finding led to the development of a formal *tiered service-level framework* within the polymer division.

In the first step, the company developed a thorough list of all possible service activities the company currently offers. Interestingly, the sales force provided a significant number of these services with little further consideration. Since the sales force was compensated on the sales volume it generated, it had all the incentive to lure in clients with little consideration for bottom-line results. Concurrently, to formalize customer service activities, the company implemented a customer scorecard mechanism. The scorecard allowed it to measure and document contribution margins per individual customer. Based on these two pieces of information, the organization then specified which customer type should be eligible for which type of service. In terms of service allocation, it designated certain services would be free for all types of customers, certain services would be subject to negotiation for lower-level customer groups, another set of services would be subject to fees for lower-level customers, and finally, certain services would not be available for the least valuable set of customers. Although the sales force was not very happy at first with these new measures, the new policy had a dramatic impact on the sales force understanding of the drivers of customer profitability.

Questions:

1. How can an organization compute client-level profitability?
2. What types of systems and processes are needed to document client profitability in a systematic and ongoing fashion?
3. Is there a way to move established clients accustomed to receiving a large number of ancillary services for free to paying for these services?

APPENDIX I

Regression Scoring Models

Scoring models is the process of evaluating potential customer behavior on the basis of test results. Typically, a test is conducted in a limited market or in an experimental set up on a small sub-set of customers. This subset of customers is exposed to a marketing campaign and a product offering. The purpose of this test is to assign to each of the remaining customers a *value* which is extrapolated from the results of this test. These values typically reflect the prospective customer's likelihood of purchasing the test marketed product. The process of regression scoring can be represented in the following steps:

1. Draw a random sample from the overall population of prospective customers.
2. Obtain data from the sample that profiles individual consumer characteristics. The R, F, and M scores are variables which profile behavioral characteristics of a customer and are typically used in this procedure, along with other relevant variables.
3. Initiate a marketing campaign directed at the random sample, and record the individuals who become customers.
4. With that information, develop a regression scoring model to obtain a series of weighted variables that either predict which prospects are more likely to become customers or the value of profits that each customer is likely to provide, based on their individual characteristics.
5. By applying these weights to individual characteristics of prospective customers, we can arrive at a value for each customer which indicates how likely it is that the customer will purchase a product, or how much profit the customer will generate, if exposed to the tested marketing campaign.

Often, a model is fitted on a training sample and then validated on a holdout sample.

APPENDIX 2

RFM Values versus BEI Table

Cell #	RFM Codes	Cost per Mail	Net Profit per Sale	Breakeven %	Actual Response %	Breakeven Index
1	111	$1.00	$45.00	2.22%	17.55%	690
2	112	$1.00	$45.00	2.22%	17.45%	685
3	113	$1.00	$45.00	2.22%	17.35%	681
4	114	$1.00	$45.00	2.22%	17.25%	676
5	115	$1.00	$45.00	2.22%	17.15%	672
6	121	$1.00	$45.00	2.22%	17.05%	667
7	122	$1.00	$45.00	2.22%	16.95%	663
8	123	$1.00	$45.00	2.22%	16.85%	658
9	124	$1.00	$45.00	2.22%	16.75%	654
10	125	$1.00	$45.00	2.22%	16.65%	649
11	131	$1.00	$45.00	2.22%	16.55%	645
12	132	$1.00	$45.00	2.22%	16.45%	640
13	133	$1.00	$45.00	2.22%	16.35%	636
14	134	$1.00	$45.00	2.22%	16.25%	631
15	135	$1.00	$45.00	2.22%	16.15%	627
16	141	$1.00	$45.00	2.22%	16.05%	622
17	142	$1.00	$45.00	2.22%	15.95%	618
18	143	$1.00	$45.00	2.22%	15.85%	613
19	144	$1.00	$45.00	2.22%	15.75%	609
20	145	$1.00	$45.00	2.22%	15.65%	604
21	151	$1.00	$45.00	2.22%	15.55%	600
22	152	$1.00	$45.00	2.22%	15.45%	595
23	153	$1.00	$45.00	2.22%	15.35%	591
24	154	$1.00	$45.00	2.22%	15.25%	586
25	155	$1.00	$45.00	2.22%	15.15%	582
26	211	$1.00	$45.00	2.22%	15.65%	604
27	212	$1.00	$45.00	2.22%	15.55%	600
28	213	$1.00	$45.00	2.22%	15.45%	595
29	214	$1.00	$45.00	2.22%	15.35%	591
30	215	$1.00	$45.00	2.22%	15.25%	586
31	221	$1.00	$45.00	2.22%	15.15%	582
32	222	$1.00	$45.00	2.22%	15.05%	577
33	223	$1.00	$45.00	2.22%	14.95%	573
34	224	$1.00	$45.00	2.22%	14.85%	568
35	225	$1.00	$45.00	2.22%	14.75%	564
36	231	$1.00	$45.00	2.22%	14.65%	559
37	232	$1.00	$45.00	2.22%	14.55%	555
38	233	$1.00	$45.00	2.22%	14.45%	550
39	234	$1.00	$45.00	2.22%	14.35%	546
40	235	$1.00	$45.00	2.22%	14.25%	541
41	241	$1.00	$45.00	2.22%	14.15%	537
42	242	$1.00	$45.00	2.22%	14.05%	532
43	243	$1.00	$45.00	2.22%	13.95%	528
44	244	$1.00	$45.00	2.22%	13.85%	523
45	245	$1.00	$45.00	2.22%	13.75%	519
46	251	$1.00	$45.00	2.22%	13.65%	514
47	252	$1.00	$45.00	2.22%	13.55%	510
48	253	$1.00	$45.00	2.22%	13.45%	505
49	254	$1.00	$45.00	2.22%	13.35%	501
50	255	$1.00	$45.00	2.22%	13.25%	496
51	311	$1.00	$45.00	2.22%	12.95%	483
52	312	$1.00	$45.00	2.22%	12.91%	481
53	313	$1.00	$45.00	2.22%	0.98%	–56
54	314	$1.00	$45.00	2.22%	0.94%	–58
55	315	$1.00	$45.00	2.22%	0.90%	–60
56	321	$1.00	$45.00	2.22%	0.86%	–61
57	322	$1.00	$45.00	2.22%	0.82%	–63
58	323	$1.00	$45.00	2.22%	0.78%	–65
59	324	$1.00	$45.00	2.22%	0.74%	–67
60	325	$1.00	$45.00	2.22%	0.70%	–69
61	331	$1.00	$45.00	2.22%	0.66%	–70
62	332	$1.00	$45.00	2.22%	0.62%	–72
63	333	$1.00	$45.00	2.22%	0.58%	–74
64	334	$1.00	$45.00	2.22%	0.54%	–76
65	335	$1.00	$45.00	2.22%	0.50%	–78
66	341	$1.00	$45.00	2.22%	0.46%	–79
67	342	$1.00	$45.00	2.22%	0.42%	–81
68	343	$1.00	$45.00	2.22%	0.38%	–83
69	344	$1.00	$45.00	2.22%	0.34%	–85
70	345	$1.00	$45.00	2.22%	0.00%	–100
71	351	$1.00	$45.00	2.22%	–0.03%	–101
72	352	$1.00	$45.00	2.22%	–0.06%	–103
73	353	$1.00	$45.00	2.22%	–0.09%	–104
74	354	$1.00	$45.00	2.22%	–0.12%	–105
75	355	$1.00	$45.00	2.22%	–0.15%	–107
76	411	$1.00	$45.00	2.22%	11.25%	406
77	412	$1.00	$45.00	2.22%	11.22%	405
78	413	$1.00	$45.00	2.22%	0.55%	–75
79	414	$1.00	$45.00	2.22%	0.52%	–77
80	415	$1.00	$45.00	2.22%	0.49%	–78
81	421	$1.00	$45.00	2.22%	0.46%	–79
82	422	$1.00	$45.00	2.22%	0.43%	–81
83	423	$1.00	$45.00	2.22%	0.40%	–82
84	424	$1.00	$45.00	2.22%	0.37%	–83
85	425	$1.00	$45.00	2.22%	0.34%	–85
86	431	$1.00	$45.00	2.22%	0.31%	–86
87	432	$1.00	$45.00	2.22%	0.28%	–87
88	433	$1.00	$45.00	2.22%	0.25%	–89
89	434	$1.00	$45.00	2.22%	0.22%	–90
90	435	$1.00	$45.00	2.22%	0.19%	–91
91	441	$1.00	$45.00	2.22%	0.16%	–93
92	442	$1.00	$45.00	2.22%	0.13%	–94
93	443	$1.00	$45.00	2.22%	0.10%	–96
94	444	$1.00	$45.00	2.22%	0.07%	–97
95	445	$1.00	$45.00	2.22%	0.04%	–98
96	451	$1.00	$45.00	2.22%	0.01%	–100
97	452	$1.00	$45.00	2.22%	–0.02%	–101
98	453	$1.00	$45.00	2.22%	–0.05%	–102
99	454	$1.00	$45.00	2.22%	–0.08%	–104
100	455	$1.00	$45.00	2.22%	–0.11%	–105
101	511	$1.00	$45.00	2.22%	10.88%	390
102	512	$1.00	$45.00	2.22%	10.85%	388
103	513	$1.00	$45.00	2.22%	0.78%	–65
104	514	$1.00	$45.00	2.22%	0.73%	–67
105	515	$1.00	$45.00	2.22%	0.70%	–69
106	521	$1.00	$45.00	2.22%	0.67%	–70
107	522	$1.00	$45.00	2.22%	0.64%	–71
108	523	$1.00	$45.00	2.22%	0.61%	–73
109	524	$1.00	$45.00	2.22%	0.58%	–74
110	525	$1.00	$45.00	2.22%	0.55%	–75
111	531	$1.00	$45.00	2.22%	0.52%	–77
112	532	$1.00	$45.00	2.22%	0.49%	–78
113	533	$1.00	$45.00	2.22%	0.46%	–79
114	534	$1.00	$45.00	2.22%	0.43%	–81
115	535	$1.00	$45.00	2.22%	0.40%	–82
116	541	$1.00	$45.00	2.22%	0.37%	–83
117	542	$1.00	$45.00	2.22%	0.34%	–85
118	543	$1.00	$45.00	2.22%	0.31%	–86
119	544	$1.00	$45.00	2.22%	0.28%	–87
120	545	$1.00	$45.00	2.22%	0.25%	–89
121	551	$1.00	$45.00	2.22%	0.22%	–90
122	552	$1.00	$45.00	2.22%	0.19%	–91
123	553	$1.00	$45.00	2.22%	0.10%	–96
124	554	$1.00	$45.00	2.22%	0.01%	–100
125	555	$1.00	$45.00	2.22%	–0.08%	–104

ENDNOTES

1. All numerical figures mentioned in the discussions below are hypothetical data created for instructional purposes only. However, due care has been exercised to ensure these data are fairly closer to real life experiences of many firms.

2. David A. Aaker, V. Kumar, and George S. Day, *Marketing Research,* 8th ed., (New York: John Wiley & Sons, 2003).

3. D. R. Jackson, "In Quest of the Grail: Breaking the Barriers to Customer Valuation," *Direct Marketing* 54, no. 11 (1992): 44–48.

4. For details on various forms of LTV models, see: Dipak Jain, and Siddhartha Singh, "Customer Lifetime Value Research in Marketing: A Review and Future Directions," *Journal of Interactive Marketing* 16 (March 1, 2002); and Paul D. Berger and Nada I Nasr, "Customer Lifetime Value: Marketing Models and Applications," *Journal of Interactive Marketing* 12 (Winter 1998): 17–30.