database marketing

John A. McCarty

INTRODUCTION

Database marketing involves the analysis of customer transaction data, along with other customer information (e.g., demographics), to segment customers (*see* MARKET SEGMENTATION AND TARGETING), and to develop marketing strategies (*see* MARKETING STRATEGY) to some or all of those segments. A key aspect of this description is that database marketing almost always involves the analysis of transaction data, either alone or with other kinds of customer data. Transaction data include information about the actual purchases made by consumers, such as when a purchase was made, the assortment of items that were purchased, and the total amount that was spent during a particular purchase. Thus, database marketing is generally employed by marketers having access to such information about individual customers, including catalog marketers, retailers, credit card companies, and others that engage in direct contact of one sort or another with customers. Nonprofits (e.g., charities, associations) have also utilized database marketing. A charity, for example, has information about donors, including how much they have donated over the years, the amount of their largest donation, and the number of times they have donated.

Among marketing firms, direct marketers (e.g., L.L. Bean) were the first to employ database marketing techniques. Their earliest efforts were largely informal in that direct marketers tended to note the patterns and preferences of their customers and to utilize knowledge of these in their marketing to repeat customers (Baier, Ruf, and Chakraborty, 2002). As the number of customers for these firms grew to a level that would make such informal methods impractical, computers and analytical procedures were developing that could automate the tracking of customer behavior.

Today, database marketing involves a number of analytical procedures, both basic and sophisticated, that enable direct marketers to track the behavior patterns of their customers. There has also been an increase in the use of database marketing by firms that are not traditional direct marketers, such as credit card providers, automobile rental agencies, brick and mortar retailers, casinos, and hotel chains. Virtually any firm that routinely captures transaction information as part of the way it does business can engage in database marketing.

DATABASE MARKETING AND TRADITIONAL MASS MARKETING

Firms that engage in database marketing differ from most traditional mass marketers that produce packaged goods (e.g., ketchup, paper towels) in two important and related ways. First, database marketers have individual customer-level information. Since they generally deal with their customers in a direct manner, they can identify their customers individually, including customers' names and addresses. Second, database marketers have the information about individual customer's transactions with the firm. That is, every time a specific customer makes a purchase, the marketer collects information about how many different items were purchased, how many of each of these items were purchased, how the customer made the payment (i.e., credit or check), and so on. This individual level information allows database marketers to serve very small segments of their customers and to personalize and customize their marketing efforts to individuals. Furthermore, given that they have the actual transaction data of customers, database marketers can segment and develop strategies based on the actual behavior of customers rather than on characteristics such as demographics.

What database marketers can do with these data contrasts with the abilities of traditional mass marketers. A typical mass marketer, like the ones that produce packaged goods, sells to consumers through intermediaries. Mass marketers do not know the names or characteristics of their individual customers and have no information about the specific purchases, as these transactions occur in supermarkets, drugstores, and/or mass merchandisers. Mass marketers can make inferences about the characteristics of their customers based on survey research and other marketing research studies. Although such studies can provide

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

these marketers with insights about their typical consumer, they cannot tell them about the behaviors of individual consumers. Survey research may inform a mass marketer that the consumers who purchase its brand are generally of a particular age group and gender. Other types of marketing research may tell the mass marketer such things as on what occasions consumers typically use the product, in what quantities they use it, or even how they feel about it (e.g., the pros and cons of its brand relative to other brands). Although this information is extremely valuable to mass marketers in the development of their marketing strategies to the broad segments that they serve, it cannot provide them with ways of marketing to specific customers.

Although it may seem that database marketers have enormous advantages over mass marketers, the ability to market to individual customers comes with certain costs. Developing and maintaining a customer database can be a costly activity. Designing and implementing marketing strategies to small segments of customers and developing customized promotional materials can be very expensive. Therefore, the revenues that firms employing database marketing techniques receive must cover the costs of these efforts. For the vast majority of traditional marketers, the selling of their products through mass marketing outlets such as supermarkets makes sense, given their profit margins. Moreover, it is convenient for most consumers to buy their packaged goods in one place (e.g., local supermarkets); therefore, consumers are generally not interested in purchasing such items as toothpaste, canned beets, and toilet paper in a direct way from the different marketers of these goods. Therefore, for both the firms producing most packaged goods and the consumers of those goods, the way these things are bought and sold makes reasonable sense.

Considering marketing activities at a broad level, it can be said that database marketers and traditional mass marketers engage in similar activities; that is, both types of marketers engage in segmentation and the development of marketing strategies to serve these segments. The crucial difference is the kind of information that is used to engage in these marketing activities. Database marketers have information that allows them to segment on the specific behaviors of their customers and to develop customized strategies to narrow segments or even to individuals. Traditional mass marketers have information from marketing research that allows them to segment on variables that relate to consumption (e.g., demographic information) and develop strategies based on the inferred behavior of large segments of consumers.

Questions that Database Marketing may Address

As noted, database marketing involves the quantitative analysis of customer transaction information in an effort to engage in marketing strategy and tactics with respect to those customers. Three examples aid in understanding the range of questions that database marketing may address.

A common issue for a database marketer is how to segment its database such that it can identify the very best customers, those that it may be able to covert to this top group of customers, and even those toward whom it may want to discontinue marketing efforts. (Hughes, 2005). The very best customers (gold customers, according to Hughes terminology) are those who are doing a lot of business with the company and these are the ones on which the marketer may want to spend more customer service dollars. Those who are just below gold customers are the ones that the company may want to market to heavily, with an eye toward increasing their business to the level of gold customers. At the other extreme, there are customers who do very little business with the marketer. The revenues obtained from these customers may be so little as to not really justify any marketing effort by the firm. Analysis of the revenues obtained from different customers in the data file and their patterns of transactions can help database marketers place their customers in these various segments.

A very common question for catalog marketers is, "who among our total customers should receive an upcoming mailing?" This sort of question is a very serious one for marketers relying on the mail to conduct business in that the postage rate generally makes mailing to the total customer file cost prohibitive. Even if the cost is much less than a dollar per piece, mailing to a file of tens of thousands of customers is an expensive proposition when the typical response rate is less than 5% for a single mailing. This problem is a segmentation and targeting issue in that a marketer would like to mail to the people in the database who are the most likely to respond to the mailing. Identifying those most likely to respond to a particular mailing from transaction variables in the customer data file is a typical database marketing problem. Analysis of the customers who have responded to similar mailings in the past compared with those who did not is one way of evaluating who should receive an upcoming mailing.

A different sort of database marketing effort is in the area of cross-selling. Cross-selling involves efforts toward marketing one product or service to customers currently purchasing other products or services from a marketer. The nature of the information in a customer data file allows marketers to evaluate cross-selling opportunities. Marketers can evaluate the pattern of purchases for the typical customer and determine if there are different products that customers tend to consistently buy together. A supplier of clothing, for example, may note that customers who buy dress shirts from it also tend to buy ties. A cross-selling opportunity would be to identify those customers who have been buying shirts, but not ties, and attempt to market ties to these customers through a promotional effort.

Database Marketing and Data Mining

Database marketing is the application of data mining techniques in the context of marketing strategy and tactics. According to Berry and Linoff (2000), "data mining is the process of exploration and analysis, by automatic or semi-automatic means, of large quantities of data in order to discover meaningful patterns and rules (p. 7)". Data mining, therefore, can be thought of as an approach to understanding and interpreting data; when marketers engage in mining their customer data, it is database marketing. However, data mining has many applications – not only marketing or business applications but also in such diverse areas as health care, engineering, the military, and public policy, to name a few. Data mining is applicable to any situation where there is an enormous amount of quantitative information and there is belief that understanding the patterns of the data will provide useful insights.

Data mining techniques tend to be data driven rather than hypothesis driven (*see* HYPOTHESIS TESTING RELATED TO DIFFERENCES – PARAMETRIC TESTS). In other words, the typical data mining analytical procedure searches for patterns and relationships in the data. In contrast, hypothesis-driven research begins with beliefs about the patterns that may exist in the data. The data are used to test the viability of a hypothesis. The advantage of the data-driven approach is that one need not have a priori assumptions about the relationships; one is searching for any patterns that may exist. Therefore, one may discover relationships that were not anticipated and these serendipitous findings may prove to be meaningful and useful to the database marketer. The very real danger with such data-driven approaches is that one may be capitalizing on chance relationships in the data. Without a hypothesis based on sound theory, it is far more likely that one will find relationships because of the search for any relationship; some of the found relationships may exist because of chance occurrences rather than meaningful and robust relationships that will maintain across time. It is therefore critical that database marketers understand this important aspect of their analysis approach and take steps to evaluate the reliability of their findings.

Database Marketing and Relationship Marketing

Database marketing is closely associated with relationship marketing (*see* CUSTOMER RELATIONSHIP MANAGEMENT). Although definitions of relationship marketing vary from very narrow views to broad conceptualizations (Parvatiyar and Sheth, 2000), a primary focus of relationship marketing is on the long-term interactions between buyers and sellers rather than on the individual transaction event. So, for example, an automobile dealership approaching its business in a relationship manner will think in terms of servicing its customers' transportation needs over years rather than the sale of a car on a specific day. This would entail a consideration of postsale services and attention to the customers such that they will buy future cars from the dealer. Those promoting relationship marketing argue that such long-term focus will be beneficial to both the buyer and seller.

Direct marketers were some of the earliest marketers to think in terms of long-term relationships with their customers. To a great extent, direct marketers' focus on the long term was out of necessity. For a typical mailing, the postal costs are very high relative to the revenues resulting from that specific mailing. This cost-to-revenue relationship dictates that direct marketers think in terms of the long-term profitability of the customers acquired from the mailing in order to justify the mailing. An illustration will help make this point. Imagine that a marketer is going to do a mailing to 50 000 people. The cost of the mailing is \$0.75 per piece; the total mailing cost is therefore \$37 500. The response rate is 2.5% (a typical response rate for a cold mailing), which amounts to 1250 people responding. The average revenue per response (with all costs except for mailing subtracted) is \$30. So, the total revenue before the mailing costs are subtracted is \$37 500. Thus, it appears that the marketer would make no money from the mailing, to the extent that it thinks only in terms of this mailing and the revenues received from this mailing. If the direct marketer thinks about the long term, it will consider how much revenue that the 1250 responders will generate over time. This revenue will not only be in terms of those responders' future purchases but also the revenues generated by people they may refer to the marketer. These 1250 people may generate a lot of profits for the company over the next several years, to the extent that the direct marketer practices sound relationship marketing. This way of thinking relates to the concept of the lifetime value of the customer (*see* CUSTOMER LIFETIME VALUE (CLV)). The lifetime value of a customer is the net present value of the revenues that will be received from a customer over a period of time in the future (usually 3–5 years). In this example, the direct marketer may have only received \$30 in revenues from the average customer for this mailing, but its lifetime value analysis may estimate that it will receive hundreds of dollars from the average responder when considered across time.

Even if a database marketer does not have postal costs, there are very sound reasons for taking the long view with customers. Maintaining a database and engaging in the customized marketing efforts associated with database marketing is expensive. Such expenses will generally only pay for themselves across time. If there is a great deal of churn among a marketer's customer base, the database marketer's analytical procedures may be wasted as the customer database at any given time may contain mostly newly acquired customers.

As a general rule, database marketing works best when the majority of people in the database desire a relationship with the marketer. Alternatively, if a significant number of customers are price-oriented customers, the marketer will likely lose them as soon as another firm offers a lower price. Therefore, engaging in database marketing efforts like loyalty programs when customers are price sensitive may likely not be a worthwhile effort.

The Analytical Procedures of Database Marketing

There are a number of commonly used analytical procedures in database marketing. Although these procedures may differ on the specifics, in general, they are designed to uncover relationships among the various transaction variables in the database. Among other purposes, the analytical procedures can be used to segment customers into groups based on purchase patterns (e.g., frequently purchasing customers, the biggest spenders, and the most loyal customers); identify the customers that are most likely to respond to a specific offer; or identify the best prospects for a cross-selling effort. The analytical procedures of database marketing vary from fairly simple methods that have been around for a long time, such as recency–frequency–monetary (RFM) analysis, to very sophisticated methods that have recently emerged, such as neural network models.

RFM analysis. Years before database marketing was conceptualized in any sort of formal way, direct marketers (e.g., catalog companies) had observed that customers who had recently purchased from them, those who had frequently purchased from them, and those who had spent a fair amount of money with them were the best prospects for new offers (Baier, Ruf, and Chakraborty, 2002). Over time, these informal observations grew into the use of the database marketing analytical procedure known as *RFM analysis*. RFM analysis involves the consideration of how recently customers have purchased from the marketer, measured in number of days, weeks, or months since last purchase (recency); how frequently customers have purchased from the marketer, measured as how many times they have purchased from the marketer (frequency); and the total amount of money they have spent with the marketer (monetary value). Generally, these three variables are used to predict which customers in a data file are the most likely to respond to a particular offer.

There are a variety of approaches used in RFM analysis, differing in terms of how the relative weights of the RFM variables are determined. One common approach to determining this weighting is known as *hard coding* (Drozdenko and Drake, 2002). Hard coding involves the development of a simple linear equation of the RFM variables, where weights for these three variables are assigned as a matter of judgment and past experience. For example, a marketer may reason from past experience with similar offers that recency should be assigned a weight of 2, frequency should be assigned a weight of 1, and monetary value should be weighted 0.5. Therefore, the linear equation provides a weighted score for each person in the data file; those with the higher scores should be the ones who are more likely to respond to a particular offer. The marketer can decide how far down in the data file it should go, for example, mail to 30% of the file, 50% of the file, and so on. Again, past experience would guide the marketer in this decision. Presumably, a greater profit would be realized from mailing to a smaller group of higher probability customers, based on RFM, than mailing to everyone in the data file. The relative weights of the RFM variables could also be determined empirically from the results of a similar mailing in the past. Regression analysis could be used to determine these weights.

An alternative approach to RFM analysis is a method advocated by Hughes (2005). One key feature of Hughes' approach involves the use of a test mailing to a sample of customers in the data file prior to the selection of members of the data file who will receive the mailing. Hughes' approach begins with the sorting of the entire customer data file according to how recently the customers have purchased from the company. The customers are then divided into quintiles according to recency with these quintiles assigned a number from 5 to 1. Thus, 20% of the customers who most recently purchased are assigned a recency score of 5, the next quintile of customers are assigned a recency score of 4, and so on. Next, within each of the recency quintiles, the customers are sorted according to how frequently they have purchased from the marketer, placed in quintiles according to that variable, and given a score of 5 to 1. Finally, each of the frequency quintiles within the recency quintiles are sorted by total amount of money spent with the marketer, placed in monetary value quintiles, and assigned scores of 5 to 1. This procedure results in 125 unique cells with RFM scores of 555 (the most recent, most frequent within recency, and the most money spent within the other two variables) down to 111 (the least recent, the least frequent, and the least money spent).

Hughes suggests that a test mailing be sent to a sample of approximately 10% from each of the RFM cells. An analysis of the revenues and the costs of mailings will provide the response percent needed in order to break even. The response percent in the test mailing for each of the 125 RFM cells is compared to the break-even percent. The marketer then mails to the customers who were not in the test mailing, but only those from the RFM cells that broke even in the test. Typically, the number of cells that are mailed to is far less than the total of 125 cells; however, the profit is higher because of the mailing costs that are saved by not mailing to the unproductive cells.

Although more sophisticated analytical procedures have been developed in recent years, RFM analysis continues to be popular for a number of reasons. Ease of use and simplicity are often cited as reasons for its popularity, as well as the fact that it is a method that decision makers can easily understand (McCarty and Hastak, 2007). The simplicity of RFM is, however, often why it is criticized as a database marketing technique. Critics point out that there are really no sound reasons for limiting the analysis to these three variables when there might be other transaction variables in the customer file that would help identify likely responders.

CHAID analysis. An analytical procedure that has been used by database marketers for segmentation and targeting purposes is chi-square automatic interaction detector (CHAID). Although this analytical procedure was developed for purposes unrelated to segmentation, it has proved to be very useful to database marketers.

CHAID is a classification method that can also be referred to as *tree analysis*. It is often used to find the most likely responders to target for a mailing or other marketing effort. For the purposes of this explanation, let us say that a direct marketer wants to do a major mailing and decides to conduct a test mailing of about 10% of its entire house data file of 1 000 000; the results of the test mailing will be used to identify those in the rest of the data file to receive the mailing. The response results creates the criterion variable to use for the CHAID analysis; this variable is the dichotomous variable of those among the 10% who responded to a test mailing and those who did not respond. In this example, let us say that the overall response rate was 5% of those who were sent the test mailing (i.e., 5000 responders). A number of independent variables are then tested to see which of them are most useful in discriminating between responders and nonresponders. These independent variables can be any variables that the marketer might believe are related to response likelihood, such as the RFM variables, and other variables as well, such as method of payment (i.e., credit card or check), number of different items purchased from the marketer, and so on.

The CHAID procedure begins with a node of all the customers in the test mailing. CHAID searches among the identified independent variables for the one that best discriminates among the customers with respect to the criterion variable (i.e., respond or did not respond to the mailing). For the purposes of this illustration, let us say that the best discriminating independent variable is how frequently customers have purchased from the marketer. The procedure will split the original set of people who have

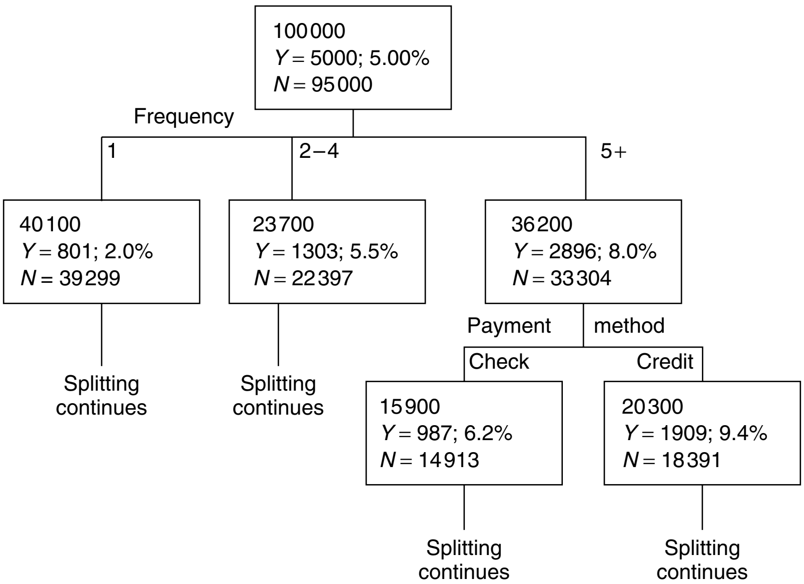

Figure 1 Example of a CHAID classification tree. This is the initial splitting of the CHAID analysis; splitting may continue until no significant differences are found.

sent the test mailing into two or more groups, depending on how this independent variable discriminates with respect to the criterion variable. For example, it may group those who have purchased five or more times together into one node, those who have purchased two to four times in a second node, and those who have purchased only once in a third node, as shown in Figure 1. Although the overall response rate was 5%, the response rate for these three nodes will be different on the criterion variable. Perhaps it is 8% for those who had purchased five or more times, 5.5% for those who had purchased two to four times, and 2% for those who had only purchased one time.

The CHAID procedure continues splitting the groups using this same logic. For each of these three nodes identified by the first independent variable (i.e., frequency), CHAID identifies the variable that best discriminates among the people in each of those groups. So, for example, among those who had purchased five or more times, the best discriminating variable might be whether they purchased with a credit card or check; the procedure will divide this group by the method of payment variable and the two resulting groups will have a different level on the criterion variable. The procedure may divide those who purchased two to four times on a different independent variable, as CHAID considers each of the groups from the original split independently in its search for the next best discriminating variable. This splitting of nodes continues until there are no more splits that would be statistically significant.

The CHAID procedure begins with a node of everyone who received the test mailing and ends with a number of nodes or groups of similar individuals with respect to the various independent variables. In this example, let us say that there are 30 final nodes. The nodes will differ with respect to their response rate. If the independent variables are indeed useful in discriminating between responders and nonresponders, then the response rate of the 30 different nodes may vary from percentages well above the response rate of the total group to some that are well below that of the total. The marketer can then place the rest of the total data file (i.e., the 90% who did not receive the test mailing) into groups or nodes according to their values on the independent variables that defined the 30 nodes. At this point, the marketer can decide which of the nodes should be sent the mailing. This decision may be a function of how many people they can mail to given a certain budget or according to a break-even analysis. In other words, a response percentage that would be necessary to break even can be determined by the revenue and costs of mailing; the marketer can then mail to only those nodes that had that percentage in the test mailing.

CHAID is similar to Hughes' (2005) approach to RFM analysis in that it places customers into cells or nodes; however, CHAID determines these nodes as a function of how the independent variables behave empirically with respect to the criterion variable. Moreover, as noted, CHAID is not restricted the RFM variables as independent variables. CHAID has been shown to be superior to hard coding RFM (Levin and Zavari, 2001; Magidson, 1988); it has been shown to perform better than Hughes' approach to RFM when the overall response rate is low, but similar to RFM when the response rate is relatively high (McCarty and Hastak, 2007).

Multiple linear regression and logistic regression. Multiple linear regression (*see* MULTIPLE REGRESSION) is a general statistical procedure that is used to predict a continuous dependent variable from a set of independent variables. Multiple regression can be useful to database marketers in the instances when they are interested in understanding the relationship among a set of transaction variables in their data file and the criterion variable of interest is not a discrete variable (e.g., response/no response to a mailing).

An example of when multiple regression may be used is when a database marketer would like to know what predicts the dollar amount of business that customers will spend with them. For example, a casino may want to know what variables are related to higher spending among their customers. The reward cards given to customers provide estimates of this information. The criterion variable is amount of money gambled (often estimated for table games; accurately provided for slot machines if the customers insert their card in the machines). There are a variety of independent variables that could be used as predictors including length of stay in the casino hotel, games that a customer plays (e.g., slot machines, table games), number of customer visits in a year, distance from customer's hometown to the casino, and demographic characteristics of a customer's hometown (available from Census Bureau data). The casino database marketer can use these independent variables to predict the criterion variable of money spent; the weights for the independent variables provide insights about the spending habits of its best customers. For example, the marketer may discover that the number of visits a customer makes is not as important as length of stay of visits in predicting spending level, or that bigger spenders play certain table games, and so on.

Another benefit from this sort of regression analysis is that it can provide information about customers who are good prospects for additional marketing. Customers who are predicted to be spending more than they actually are, given their levels on the important independent variables, can be sent promotional materials in an effort to increase their business with the casino.

Logistic regression (*see* LOGIT MODEL) is also utilized by database marketers. Logistic regression models a dichotomous dependent variable with a set of independent variables. Thus, it is useful for database marketing in instances when a marketer would like to predict response/no response to a marketing effort. A particularly useful aspect of logistic regression is that although the criterion variable is dichotomous (i.e., either a customer responded or did not); the resultant predicted values are probabilities ranging between 0 and 1. Thus, if logistic regression is performed on a test mailing or the results of a previous mailing, the analysis provides a weighted function of independent variables that best predict the response variable. The weights of the independent variables provide the information to develop the response probabilities for other customers (e.g., those who were not sent a test mailing). The marketer can decide what level of predicted response probability to use in deciding who should receive the mailing. A break-even analysis may tell the marketer that it will be profitable to mail to anyone in the data file who has a predicted response probability of 30% or higher, for example.

Logistic regression is similar to RFM and CHAID in that it is typically used to determine who should receive a mailing or other marketing effort, based on information from a previous mailing or a test mailing. It differs from the RFM and CHAID in that it does not group people into cells or nodes, rather it provides an individual score (i.e., response probability) for each person.

Neural networks. Neural network analysis is one of the most recent advances in database marketing analytical techniques. These procedures have been widely used in a variety of data mining applications in business and nonbusiness areas (Berry and Linoff, 1997). Neural networks find patterns in data using an approach that mimics the human brain in how it makes connections. The procedure uses an iterative process to "learn" from training data sets; it develops a sense of these patterns and sharpens them during this learning process, analogous to how a human may learn and understand across time. Neural networks have advantages over many other analytical techniques used in database marketing in that they do not follow a specific empirical model (e.g., the linear model) and can be adapted to a wide variety of problems. A main disadvantage of neural networks is that the results of the analyses are not always easily understood (Berry and Linoff, 1997); therefore, they are not always easy to explain to decision makers.

Prospecting for New Customers

Most database marketing activities are in support of marketing efforts directed toward a company's current customers. Companies do need to acquire new customers, however, in that there will always be some attrition of customers, no matter how good are a firm's relationship marketing practices. A key difference between the use of database marketing for current customers and the prospecting for new customers is that a firm does not have transaction information for customers that it is hoping to acquire.

In spite of this limitation, database marketing can help in new customer acquisition and there are a variety of approaches to this. One approach is for a company to develop a demographic profile of their best customers and then acquire lists of people who fit this profile to "cold call." The Census Bureau and commercial vendors (e.g., Nielsen Claritas) provide information about geographic areas that can help database marketers in their prospecting efforts. For example, if a resort hotel finds that its typical customers come from geographic areas with particular demographics and lifestyle characteristics, it can use such information to aggressively market in areas with similar demographic and lifestyle profiles. Therefore, the use of database marketing for prospecting is generally a matter of a firm understanding their own customers, then seeking others similar to their best customers.

Summary

Database marketing involves the application of data mining techniques to the analysis of a company's customer transaction data in an effort to segment customers and to develop marketing strategies to customers. Database marketing is closely aligned with the concept of relationship marketing in that those who engage in database marketing generally need to consider their long-term relationships with customers for the cost of developing, maintaining, and mining a database to be profitable. Furthermore, since the information available in customer databases tends to allow marketers to provide personalized marketing strategies to their customers, such customized efforts tend to foster strong relationships with customers. A number of analytical techniques are utilized in database marketing to help marketers discover meaningful patterns in their customer data; some of these techniques have been available for a long time (e.g., RFM analysis) while others have only recently been made available to the database marketing (e.g., neural network analysis). It is likely that the use of database marketing will grow as more types of organizations learn about its value in developing close relationships with customers.

Bibliography

Baier, M., Ruf, K.M., and Chakraborty, G. (2002) *Contempory Database Marketing: Concepts and Applications*, Racom Communications, Evanston.

Berry, M.J.A. and Linoff, G. (1997) *Data Mining Techniques for Marketing, Sales, and Customer Support*, John Wiley & Sons, Inc., New York.

Berry, M.J.A. and Linoff, G. (2000) *Mastering Data Mining: The Art and Science of Customer Relationship Management*, John Wiley & Sons, Inc., New York.

Drozdenko, R.G. and Drake, P.D. (2002) *Optimal Database Marketing: Strategy, Development, and Data Mining*, Sage Publications, Thousand Oaks.

Hughes, A.M. (2005) *Strategic Database Marketing*, 3rd edn, McGraw-Hill, New York.

Levin, N. and Zahavi, J. (2001) Predictive modeling using segmentation. *Journal of Interactive Marketing*, **15**, 2–22.

Magidson, J. (1988) Improved statistical techniques for response modeling: progression beyond regression. *Journal of Direct Marketing*, **15**, 494–501.

McCarty, J.A. and Hastak, M. (2007) Segmentation approaches in data-mining: a comparison of RFM, CHAID, and logistic regression. *Journal of Business Research*, **60**, 656–662.

Parvatiyar, A. and Sheth, J.N. (2000) The domain and conceptual foundations of relationship marketing, in *Handbook of Relationship Marketing* (eds J.N. Sheth and A. Parvatiyar), Sage Publications, Thousand Oaks, pp. 1–38.


STUDYDADDY
Get Homework Help
From Expert Tutor
Get Help