Chapter 8
Sources of Data

Abstract "Data" is the first word in database marketing with good reason – the quality, impact, and ultimately, ROI of database marketing programs depend on the availability of good data. We discuss the various types of customer data available, e.g., customer demographics, transactions, and marketing actions, and the sources that provide these data such as internal records, commercially processed numbers and segmentation schemes, externally available customer lists, and primary survey data.

8.1 Introduction

The customer information file (CIF) is the building block for database marketing. A customer-focused company always makes its decisions based on the analysis of customer data, and a detailed customer record is the prerequisite for useful analysis.

Customer information may include everything about customers, ranging from their names, addresses, phone numbers, their demographic and lifestyle information, transaction histories, and everything that can be derived from customer contacts. It is true that more detailed information leads to better decision-making. However, there is a trade-off between costs and benefits of collecting data. Therefore, it is a good idea to prioritize each data element once you list out the necessary data elements to be included in your customer information file.

The task of constructing the customer information file ideally begins with the objectives its users set. Hence, its size and specific data elements depend on the decision-making problems managers attempt to solve with the customer information file. Unfortunately, some companies have made mistakes, building their customer databases simply because their competitors have built one or because of some vague vision that it would be good to have all "the data" available. As a result, some companies have spent millions of dollars assembling data and have not seen a clear pay-off. The poor performance of

these investments has given "CRM" a bad name in many companies. In these companies, CRM has become associated with huge information technology investment in assembling data, but no clear plan on how to use it productively.

Accordingly, we recommend the following process for constructing a customer database:

For example, in Step 1, a company's managers may stipulate their desire to focus on campaigns that reduce customer churn, and campaigns that increase cross-selling. In Step 2, they list customer product ownership, customer attrition, and previous campaign histories as the types of data they need to support these activities. In Step 3, they list the exact data they want and how they will obtain it. For example, they list the specific products for which customer ownership is needed and how they will obtain these data.

The rest of this chapter focuses on Steps 2 and 3, the types of customer data that are available, and the sources of these data. A fourth step not covered here but discussed to some extent in Chapters 10 and 25, is the financial evaluation of the data themselves. Many companies may find it is too expensive to compile a specific data field, even though the above process has determined it would be valuable. Sometimes, the company can purchase data from an outside source on a "trial basis," see how well it works in the predictive models that support its marketing activities, and if so, purchase the data for the long term.

8.2 Types of Data for Describing Customers

There are no standard ways of classifying the types of data elements included in customer information file. Moreover, the types of data elements are different across industries and companies. For our convenience, we classify them into (1) customer identification data, (2) demographic data, (3) psychographic or lifestyle data, (4) transaction data, (5) marketing action data, and (6) other types of data.[1] We focus on the common data elements that most companies have in their customer information file.

8.2.1 Customer Identification Data

Customer identification data is the most basic customer information, covering various classification data, the customer's contact addresses, and other useful customer-identification data. More specifically, they include customer's

---

[1] We did not include competitive data in our typology because most companies do not have them. However, in Section 8.3.3 we discuss the use of surveys to acquire competitive information for a small sample of customers and infer from that the competitive behavior of the rest of the customers.

name (first, middle, last, prefix and suffix), a unique id, home and business addresses, home, business and mobile telephone numbers, email addresses, date of birth, and so on. For business customer data, they may also include names of contacts, departments, fax numbers, etc.

Customer identification information is rarely used in building (statistical) response models partially because it is nominally scaled.[2] However, they are critical in maintaining the customer relationship because they provide the means to contact the customer. Catalogers can't send their catalogs to customers without customers' correct addresses. Telemarketers can't initiate selling activities without customer's phone numbers.

There are two customer identification fields worthwhile to mention in greater detail. The first is the customer ID that is uniquely assigned to each customer upon her first contact with the company. Later this customer id is used as a key field to link to other databases. Once the ID is assigned for a customer, the same ID will be used for her repeat visits/contacts. That is, in order to keep track of all customer interactions, we should have a system to identify who a customer is and then pull her customer ID or assign a new customer ID. It is relatively easy to keep track of purchases made by using store (credit) cards that have the ID number on the card. Similarly, online sellers will easily identify customers if they log in with their User IDs. But some customers may pay cash or by a new credit card. Many retailers have difficulty identifying customers without store cards. These retailers attempt to match repeat customers by name, address, phone number or their combination.

Second, identification fields such as address and phone number need to be regularly updated. More than 40 million Americans change their addresses annually. In a year, 17% of consumers move and 22% of businesses move. It is important to keep customers' addresses accurate and up-to-date. A fast and cheap way of updating address is to employ a NCOA (National Change of Address) supplier licensed by the United States Postal Service. To use this service, all customer names with their addresses are sent to the NCOA service provider. Then the data is typically standardized to confirm to the USPS requirements including ZIP+4 code. Next is to match the data against the NCOA file containing records of old and new addresses for people who moved during the last couple of years. The new addresses are provided for those matched customers. Even though it will not guarantee 100% coverage, it is far less expensive and faster than other address updating methods (e.g., correcting by customer survey).[3] Many companies regularly update their customer information file through this NCOA service once or twice a year.

---

[2] Date of birth and address can be converted to "age" and "location" and these factors can be important in predictive models. However, when they have been put in a usable form, we consider them demographic data.

[3] The standard fee is $3/M with an NCOA hit rate of 3% (Robinson 2002). However, some vendors charge a fee on a "hit per name" basis, generally less than 5 cents per name.

8.2.2 Demographic Data

Demographic data spans the kinds of data fields that US Census Bureau collects every 10 years. They include the age of head of household, family income, family size, occupation of head of household, marital status, presence of children, length of residence, education level, own or rent, type of dwelling, car ownership and its types, gender, race, and so on. For business customers, they may also include data ranging from the race of the CEO, to the number of employees, sales volume, years in business, and so on.

Demographic information is especially useful for targeting prospects. Marketing researchers have found that the best predictors for customers' future purchase behaviors are their historical purchase/transaction information. However, transactional data is mainly available for current customers. In order to target prospects without any transactions, we need to utilize their demographic (and/or psychographic) characteristics that are observable to marketers. For example, from current customers we could use a predictive model to identify what types of demographic characteristics a high-value customer has. Then we target prospects whose demographic profiles are similar to those of current high-value customers. And once prospects become customers, transaction data is collected to fine-tune targeting.

Most companies often do not have enough demographic information in their customer information files. It is especially true for companies that have built customer information files based on legacy databases scattered across various departments. As discussed later, those companies can enhance customer information files through overlaying demographic data provided by external data providers. Demographic data are available both at individual or geographical aggregate level. Individual level data are more accurate, but more expensive, and sometimes are not available. The aggregate data are the average demographic values of customers living in the same geographic boundary such as census tracts, ZIP codes, Zip + 4, and postal carrier routes. For example, customer income is usually not available on a customer basis, since IRS filings are private. In that case, income for individual customers may be assumed equal to the average for their census tract.

8.2.3 Psychographic or Lifestyle Data

Lifestyle is a way of life or style of living that reflects the attitudes and values of a consumer while psychographics are psychological characteristics of consumers such as attitudes, values, lifestyles, and opinions. Generally, lifestyles and psychographics are used interchangeably. Specific questions to measure lifestyles consist of three groups: *activities* on hobbies, vacation, entertainment, club membership, sports, etc.; *interests* in family, job, fashion, food,

media, etc.; *opinions* on politics, business, economics, educations, products, culture, etc. (Plummer 1974).

Product usage or ownership can be classified as psychographic data since we can infer customers' attitude and behavior from their product usage. For example, Best Buy would like to know its customer's attitude toward technology, innovativeness, and market mavenism. Best Buy would also like to know its customer's use of various electronic products, as well as other products that may be complementary to electronic products. For example, hiking would be relevant information to know for Best Buy because they might be able to sell hikers electronic gizmos such as GPS locaters. Database marketers often purchase consumer "response lists" to collect these psychographic information. Consumer response lists are lists of individuals who have some identifiable product interest (e.g., martial art equipments or digital cameras) and have a proven willingness to buying by mail (Roberts and Berger 1999). There are several subcategories of consumer response lists. Buyer lists (those who have bought a product or service) and subscription lists (those subscribing to a publication) are most meaningful for database marketers. For example, the *Nordstrom Quality Women Apparel* (buyer) list includes about 500,000 purchasers of apparel from Nordstrom during the past 12 months.

Lifestyle research has grown out of the limitation of demographic variables to explain heterogeneous purchase behavior across consumers. Still few companies are collecting individual lifestyle data for their own database marketing use, partially because of large collection costs, even though they often purchase some lifestyle information from consumer response lists. However, database marketers should note that traditional marketing researchers have successfully used lifestyle data for targeting (or segmenting) prospects for a long time. It may be valuable to conduct a lifestyle survey for a sample of current customers and identify what types of lifestyle characteristics a high-value customer has. Then we might link those lifestyle characteristics to demographic or transaction data that we have available for all customers. Through the chain demographic/transaction data $\Rightarrow$ lifestyle $\Rightarrow$ customer value, we can target the customers we want.

The best-known lifestyle segmentation system is VALS, formerly known as the Values and Lifestyles Program, developed in 1978 by the Stanford Research Institute (SRI) and now owned and operated by SRI Consulting Business Intelligence (SRIC-BI). VALS was one of the first major consumer segmentation programs based on consumer lifestyle characteristics. It correlated people's values about social issues such as abortion rights and military spending with their product and media preferences. In 1989, VALS was revised. Psychological characteristics such as excitement-seeking were found to be more powerful predictors of consumer behavior and more stable over time than social values were. Over the years, many consumer product companies have used VALS for new product development, positioning and effective advertising.

VALS classifies American adults into eight distinctive groups: Innovators, Thinkers, Achievers, Experiencers, Believers, Strivers, Makers, and Survivors. The segments differ in terms of attitudes, decision making patterns, and purchases of products, services, and media. As shown in Fig. 8.1, the VALS segments are defined along two fundamental dimensions: primary motivation (horizontal dimension) and resources available, which is associated with innovative behavior (vertical dimension). VALS points out that people are driven by three powerful primary motivations: Ideals, Achievement, and Self-Expression. Resources refer to education, income, self-confidence, health, eagerness to buy things, and energy level. For example, Achievers have goal-oriented lifestyles and a deep commitment to career and family. They value consensus, predictability, and stability over risk, intimacy, and self-discovery. For more detailed descriptions on VALS, see http://www.sric-bi.com/VALS/types.shtml. GeoVALS estimates the percentage of the eight VALS types by DMA (designated metropolitan area) and zip code. Japan-VALS segments Japanese consumers.
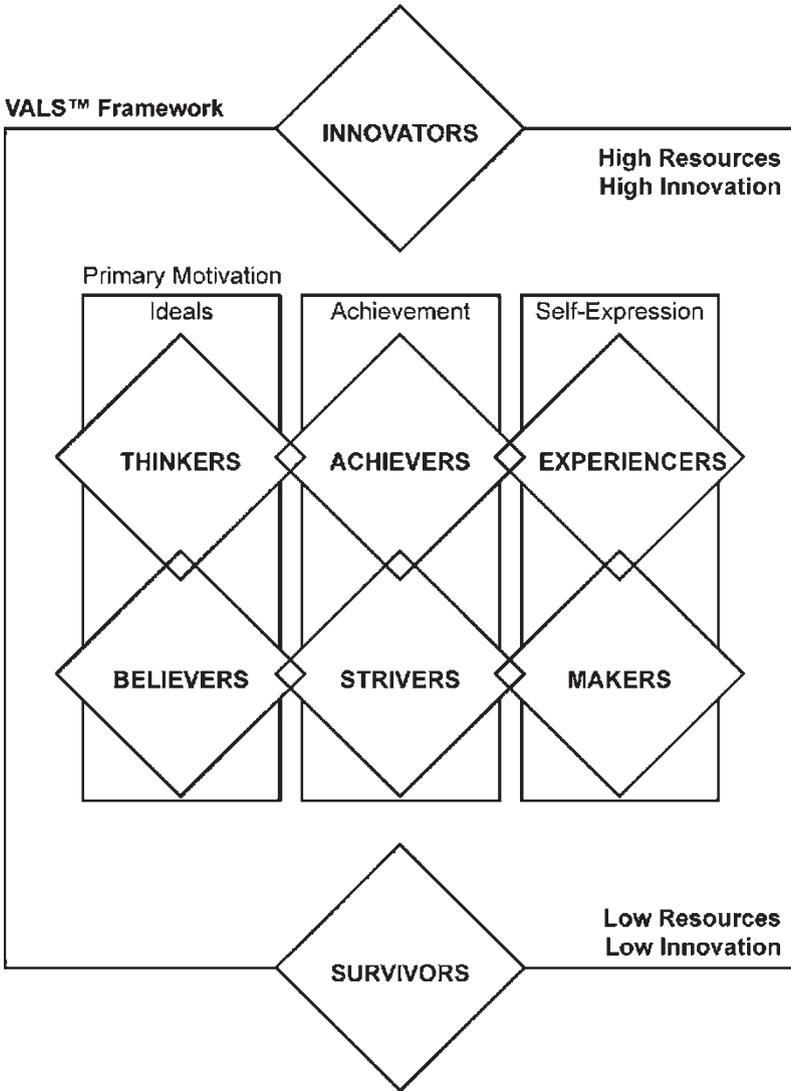
8.2.4 Transaction Data

Transaction data are the most powerful data for predicting future customer purchase behavior. In fact, scanner data researchers for the last two decades have developed various models predicting consumer purchase behavior of packaged goods based on transaction histories. In addition, transaction data cannot typically be purchased outside.[4] Transaction-related data can be collected across various channels through which a company interacts with customers. In addition, valuable transaction information can internally be found in various departments such as customer service, fulfillment, billing, and accounting.

Generally, transaction data include purchase date, purchased items with their product categories, sizes, and prices, purchase amount, method of payment, discount, sales tax, return code, allowances, salesperson ID, and so on. Interpreting customer transactions more broadly, transactions are the outcomes of a process: attention, intention, desire, and action. Hence, we treat any transaction-related information before and/or after the purchase (e.g., product/service inquiry, Web clickstream data, customer complaints, customer satisfaction scores) as transaction data.

Because of data storage and maintenance costs, some companies only save some part of transaction data or its summary. For example, a typical telecommunication company has tens of million customers and each customer makes several calls a day. The size of calling data easily becomes terabytes in weeks.

---

[4] Transaction data defined here do not include product usage or ownership information that can be purchased through consumer response lists. That is, we limit our attention to the transaction-related information specific to the firm's product or service.

VALS™ Framework

INNOVATORS

**High Resources
High Innovation**

Primary Motivation

Ideals   Achievement   Self-Expression

THINKERS   ACHIEVERS   EXPERIENCERS

BELIEVERS   STRIVERS   MAKERS

**Low Resources
Low Innovation**

SURVIVORS

Fig. 8.1 Eight segments of VALS (From *The VALS Segments*, copyrighted graphic used with permission: SRI Consulting Business Intelligence, http://www.sric-bi.com/VALS/types.shtml).

Hence, instead of saving all calling data, the transaction data are recorded in summarized quantities (total number of calls made each day, the most frequently called numbers, etc.). As the data storage and access costs drop, we expect that more data will be stored in raw form.

8.2.5 Marketing Action Data

Marketing action data and transaction data may be the most important types of data for efficient marketing strategies and tactics. Marketing action information covers all marketing efforts directing to customers and prospects whereas transaction information is their responses to the marketing effects. For example, if we send out a catalog and the customer responds, the mailing of a catalog is marketing action data and the response or no response is transaction data.

Remember that our ideal objective in database marketing is to find the optimal set of marketing activities over time that will maximize the customer's long-run profit contribution. The control variables in our hands are marketing activities. Should we send the catalog to a customer? Should the salesperson visit the customer to cross-sell other products? We can make all these marketing decisions efficiently if we have records on historical marketing activities and the corresponding customer responses or performances. Hence, if possible, the two should be linked. Surprisingly, however, not many companies record marketing action information partially because it is time-consuming and cumbersome without an automatic recording system. And it is sometimes difficult to link the marketing activities and the customer response when the firm employs multiple channels. For example, customer may receive a catalog and the catalog encourages the customer to buy on the Internet. So is this purchase matched to the catalog or the Internet? This is important for companies that want to evaluate their marketing efforts. (See Chapter 25 for more discussion.)

Similar to the case of transaction data, marketing efforts are the outcome of an ongoing interaction process with the customer. Accordingly, the gathering of marketing action data is an ongoing process. Consider the case of the life insurance industry. Marketing action data would include the number of communications sent to customers (or prospects), the kinds of promotional offers that were delivered, when the salesperson contacted the customer, when thank you letters were mailed, and any follow-up conducted by the customer-care department or by the sales representative.

Marketing data are also very diverse depending on the type of business. The major marketing activity of a direct marketer is to select a group of customers out of its house list and send mailing packages to them. The direct marketer should record customer lists of who received each mailing, the content of mailings such as types of products and amount of discounts, date of mailing, costs of mailing, etc. Marketing action data can be very complex when the sources of selling products and services are diverse. For example, an insurance company may sell its products through the Internet, in-house or outside telemarketing agency, own sales force or brokers, or alliances with other financial institutions. Marketing strategies differ across these various channels, so the types of marketing action data differ as well. In addition, the difficulty in collecting marketing action data is expected to differ across

various channels. Data collection is relatively easy with the Internet. It may be extremely difficult with brokers because sometimes their sales representatives do not record their selling activities or, even if they do, they are not willing to share this information with the insurance company.

8.2.6 Other Types of Data

Financial data are especially important for financial institutions and service providers who require monthly payments. Some external data providers supply customer-level financial data. For example, Fair Isaac Inc. has developed a credit score called a FICO score (Myers Internet 2005). A FICO score attempts to condense a customer's credit history (e.g., late payments, amount of time credit has been established, length of time at present residence, employment history) into a single number representing the financial risk of the customer. That is, a FICO score is the likelihood that credit users will pay their bills. There are actually three FICO scores computed by data provided by each of the three credit bureaus – Experian, Trans Union, and Equifax.

Some companies also include in their customer information files descriptors they calculate or infer about each customer. These are often derived from statistical models. Examples are lifetime value, values of RFM, credit score, up-selling score, cross-selling score, selling productivity, and so on. The actual score or a decile-ranking may be recorded.

8.3 Sources of Customer Information

Once customer information needs have been defined, we list the specific data fields required to achieve the information objective and determine how to collect them. The goal is to collect the most accurate or valuable information as cheaply as possible. Understandably, there is a tradeoff between the value of information and its acquisition costs.

Marketing researchers have classified marketing research data into primary and secondary data (Malhotra 1993). A researcher collects primary data for the specific purposes of addressing the problem at hand. Collecting primary data is expensive and time-consuming. On the other hand, secondary data are data that have already been collected for purposes other than the problem at hand. Secondary data are easier and faster to collect. Secondary data are classified into two types: internal and external. Internal data are those generated within the organization for which the research is being conducted, while external data are generated outside of the organization.

We use the same topology in classifying sources of customer information: internal (secondary) data, external (secondary) data, and primary data.

As explained, internal data are the cheapest to collect, followed by external data and primary data. In contrast, primary data are the most time-consuming to collect, followed by external and internal data. Hence, when we collect customer data, first we should look internally. If the data are there and its accuracy or quality is acceptable, we stop. But if not, we look for external data sources. And the final medium we rely upon is primary data.

8.3.1 Internal (Secondary) Data

Internal sources should be a starting point for the customer information file. Most companies have vast amounts of internal data, more than they expected to have. Data may have been collected and stored for other business purposes than marketing. For example, order processing or fulfillment system may have compiled order-taking channels, inventory availability, and delivery information. Billing and accounting computer systems may store information on pricing, sales volume, discounts, and net price paid. Contact management systems for the sales force often have valuable information on customer profiles, contact dates, types of promotions offered, and even the content of conversation. Providing information on customer complaints and satisfactions, the customer service department is another valuable internal data source. The marketing and/or sales department stores customer transaction histories and records for various marketing activities. An important but easily ignored source of data that marketing department may have is prospect data. They may come from participants of sweepstakes, names of the gift receivers, or names having called the 800 number.

Unfortunately, managers often do not know exactly where the internal data are stored. Or some of them are not in usable format. This is one of the reasons why the issue of data warehousing is important. A data warehouse is the process of assembling data from various sources, transforming and organizing it into a consistent form for decision-making and providing users with access to this information (Man 1996). That is, a data warehouse is a repository of data gathered from various departments for business decision making. A properly constructed data warehouse provides the company with a more precise understanding of its customer behaviors. For example, Victoria's Secret Stores, the US lingerie chain, used to spend too much time collecting information without thinking about using it (Goldberg and Vijayan 1996). Its data warehousing efforts allowed them to learn that its system of allocating merchandise to its 678 shops, based on a mathematical store average, was wrong. An average store sells equal numbers of black and ivory lingerie while stores in Miami sells ivory by a margin of 10 to 1.

In addition, a data warehouse helps the company to become more customer-centric since data from various functional areas are integrated around customers. With integrated customer databases, many banks such as

BankAmerica Corp. and Bank One Corp. can provide branch and customer service employees with comprehensive customer relationship information (Barthel 1995). As a result, they can improve customer services by satisfying inquiries about various accounts through a single point of contact. In addition, data integration allows banks to identify cross-selling possibilities more readily.

Typically thousands of users with different needs want to use a data warehouse. Hence it is not frequently efficient to meet the needs of all users with a single warehouse. The solution can be a "data mart" that is a kind of a departmental data warehouse. A data mart is a specialized system that stores the data needed for a specific department or purpose (e.g., marketing data mart). The (enterprise) data warehouse is the central repository for all company data with extensive data elements available while a data mart is considered as a subset of the data warehouse specializing in a narrow subject or functional area. For example, the Magazine division of AOL–Time Warner may have its own marketing data mart to satisfy its specific needs. The data mart is created and updated by tapping the enterprise data warehouse.

Compared to a data warehouse, the size of a marketing data mart is relatively small, hence the cost of building a marketing data mart is lower and only several months are required to build it. Recently, however, the size of many marketing databases is increasing rapidly. Several companies have reported that their marketing data marts are approaching the terabyte level (Stedman 1997). For example, Fleet Financial Group is building a 1 TB data warehouse and a 500 GB accompanying marketing database. MCI has 2.5 TB of sales and marketing information and Charles Schwab & Co. has built a customer data mart close to 1 TB. Experiencing performance degradation due to the size of their marketing data mart, they have negotiated with end users to include only critical information in the marketing data marts. Often, they separate the marketing database into even smaller pieces to increase query performance. For example, MCI has built 16 separate marketing data marts such that each is smaller than 100 GB.

8.3.2 External (Secondary) Data

Internal data do not always satisfy the data requirements specified by managers. Therefore, database marketers often purchase external data for two purposes. First, they may rent (targeted) customer lists for prospecting. Second, they may want to augment their customer information files by attaching various demographic and lifestyle characteristics. This is called data enhancement. To enhance your data, you would send your customer list (with identification information such as names and addresses) to a list enhancer and request that a couple of demographic characteristics (e.g., occupation) be attached to those customers. The list enhancer runs your list through its

proprietary database by matching your names/addresses to theirs (Roberts and Berger 1999). Demographic information would be appended to those customers when your customers are on the list enhancer's customer list. The list enhancer returns back the enhanced data disk to you.

For marketing researchers, there are many external data sources including federal and state government, trade associations, commercial data providers, and marketing research firms (Malhotra 1993). In this section, we will limit our discussion on external data sources that are more relevant to database marketers.

8.3.2.1 US Census Data

US Census data are an easy and quick external source for enhancing your customer information file. The Census Bureau conducts a full census every 10 years and now the Census 2000 data is available (www.census.gov). In order to reduce the burden on citizens by asking too many questions, the Census Bureau introduced the long form along with the short form. Five out of six households get a short form containing only the basic demographic questions while the remaining household is asked to fill out the long form containing more than 100 questions about its lifestyles and backgrounds.

The US Census contains very useful information for database marketers. For example, variables in the census include household income, per capita income, education level, types of occupations, home values, rent or own, age of the head of the household, and so on. Individual-level census data are not available due to privacy concerns. However, geographical aggregate (average) information is made available to the public. The census data are available at zip code or block group levels. A block may have as few as 10–20 families in it, and their demographic characteristics in the same block are expected to be similar. Some data elements such as population and income are available at the block group level whereas others (e.g., place of birth by citizenship status, real estate taxes) are only available at the census track level.

Commercial data vendors use census data along with other data to make these data easier and more convenient for database marketers. Considering that the census is conducted every 10 years, they also provide the projected data for each year.

8.3.2.2 Mapping Data

Geocoding is the process of assigning latitude and longitude coordinates to a location. Geomarketing vendors provide geocoding services that can enhance your customer information files by assigning coordinates to each address, ZIP code, or ZIP + 4. Before assigning coordinates, they also verify whether the address is consistent with USPS standard.

Once the location coordinates are attached to the address records, mapping software can display the location on the map. Distances can be calculated between each customer and various store locations. The data can also be useful for presentations and overall understanding of the market from a spatial location perspective. And if it is integrated with marketing information (e.g., customer's demographic characteristics), it can be a valuable tool for both managers and analysts. For example, mapping can help retailers and banks to find a location for a new branch.

8.3.2.3 Individual Level Data

Several external data vendors provide list enhancement services at the individual customer level. For example, Donnelly Marketing, a subsidiary of InfoUSA, has various demographic and lifestyle information for over 225 million individuals and 100 million unique households. For each individual or household, it offers over 250 variables that can be appended to a client's customer information file. They include gender, age, estimated household income, home ownership, estimated home value, dwelling type, credit card information, mail order buyers, auto ownership, ethnicity, marital status, newlyweds, and so on (see http://www.infoUSA.com).

InfoUSA also provides list enhancement service for businesses. Extracted from various sources such as yellow pages, annual reports, leading business magazines and newspapers, and 17 million phone calls to verify information, its business database has over 14 million US businesses. Clients can enhance their business lists with more than 70 variables including type of business, sales volume, number of employees, the race of CEO, the establishment date, credit rating, phone number and fax number, and so on.

There are at least four major vendors competing in the list enhancement business: Acxiom, Experian, Equifax (previously R. L. Polk & Co.), and InfoUSA. Each vendor is slightly different in compiling its lists initially. For example, R. L. Polk & Co. complied its lists from the state motor vehicle registration information along with various public records. Hence, it has some automobile specialized information such as the make, model, and year of the automobile owned. Having sold its Consumer Information Solution group to Equifax, Polk now focuses on the automobile industry (http://www.polk.com). On the other hand, Donnelley Marketing originally compiled its lists through telephone and city directories. Now sources of its lists are expanded to mail order buyers and subscribers, magazine subscribers, credit information and other public sources. In 1999, Donnelley Marketing became the wholly owned subsidiary of InfoUSA, which was already a well-known provider of business databases.

An innovative way of collecting lifestyle data has been crafted by NDL (National Demographics and Lifestyles) in the 1970s.[5] NDL compiles its

---

[5] Afterwards, NDL has been acquired by R. L. Polk.

data by providing warranty card processing services to manufacturers and retailers. The manufacturer inserts warranty cards in its product packaging along with a lifestyle survey provided by NDL. NDL processes the data, provides the warranty information with some lifestyle data to the manufacture, but keeps adding all the information to its customer lifestyle database.

Even large compilers can not achieve 100% matches with a typical customer database. About 80–90% of a company's customer lists are covered by the large compilers (Roberts and Berger 1999). Moreover, for those lists of names matched, some data elements may not be available at the individual level. Compilers provide various geographic average values (e.g., ZIP, ZIP + 4 or block groups) for those data elements. Understandably, it costs more to enhance at the individual level. In addition, some data elements such as income are often the inferred values that are the outputs of a statistical model. For example, we can think of an income prediction model where the dependent variable is income and the independent variables may be the age of household head, automobile ownership, type and make of the car, home value, types of occupation, etc. The model is estimated on a group of individuals for whom individual-level income is available. The model can then be used to predict income for all households.

8.3.2.4 Pre-modeled or Cluster Data

Having applied customer analysis to customer (individual or geographic level) and business data, some external data providers segment the entire US population into several groups. Taking hundreds of demographic and lifestyle variables for each subject, a clustering algorithm attempts to find a number of clusters such that subjects in the same cluster are as homogeneous as possible and, at the same time, subjects in the different clusters are as heterogeneous as possible in terms of their characteristics. Once the plausible number of clusters is determined, each segment is characterized by the average values of its subjects' demographic, lifestyle, psychological, and media interest. See Chapter 16 for further discussion of "cluster analysis."

These segment codes can be said to be pre-modeled information in that they are the output of external data vendors' clustering algorithm. They can save your time and effort in modeling since it is already modeled. Once you have your customer lists cluster-coded, you can tell to which segment each customer belongs and enumerate his/her various characteristics based on the norms for that segment. Clustering information is sometimes criticized because it is too aggregated and/or demographic and lifestyle variables cannot explain the heterogeneity of purchase behavior among customers. For example, you wouldn't expect that the percentage of cat owners for one cluster would be significantly different than that for other clusters. However, clustering information is useful for targeting prospects at least.

Moreover, it helps you to develop appropriate advertising copy and media selections.

PRIZM is one of the first demographic and lifestyle based segmentation schemes (http://www.claritas.com). Developed by Claritas, PRIZM is based on the principle that "birds of a feather flock together." That is, people of similar demographic and lifestyle characteristics tend to live near each other. Applying a series of cluster and factor analyses to US Census data, it clusters census block groups with similar characteristics into 66 neighborhood types, listed in Table 8.1. With similar profiles in terms of demographic and lifestyle characteristics, each cluster has a unique name such as Blue Blood Estates, Furs & Station Wagons, Shotguns & Pickups, and Young Influentials. Once the customer information file of a client is PRIZM-coded, you can improve your targeting decision for prospects.

More recently, Looking Glass has developed a segmentation scheme called Cohorts. Unlike PRIZM, Cohorts is based on self-reported household level

Table 8.1 Brief descriptions of the 66 PRIZM-NE clusters (Courtesy of Claritas, Inc., accessed August 8, 2007. Available at http://www.claritas.com/MyBestSegments/Default. jsp?ID = 30&SubID = &pageName = Segment%2BLook-up)

Segment	Name	Descriptions
GROUP U1: Urban uptown		
04	Young Digerati	With the boom in new computer and digital technology, this cluster represents the nation's tech-savvy singles and couples living in fashionable neighborhoods on the urban fringe
07	Money and Brains	The residents of Money & Brains seem to have it all: high incomes, advanced degrees and sophisticated tastes to match their credentials. Many of these city dwellers, predominantly white with a high concentration of Asian Americans, are married couples with few children who live in fashionable homes on small, manicured lots.
16	Bohemian Mix	A collection of young, mobile urbanites, Bohemian Mix represents the nation's most liberal lifestyles. Its residents are a progressive mix of young singles and couples, students and professionals, Hispanics, Asians, African-Americans and whites. In their funky rowhouses and apartments, Bohemian Mixers are the early adopters who are quick to check out the latest movie, nightclub, laptop and microbrew.
26	The Cosmopolitans	The continued gentrification of the nation's cities has resulted in the emergence of this segment-concentrated in America's fast-growing metros such as Las Vegas, Miami and Albuquerque. These households feature older homeowners, empty nesters and college graduates who enjoy leisure-intensive lifestyles.
29	American Dreams	American Dreams is a living example of how ethnically diverse the nation has become: more than half the residents are Hispanic, Asian or African-American. In these multilingual neighborhoods – one in ten speaks a language other than English – middle-aged immigrants and their children live in middle-class comfort.

(*continued*)

Table 8.1 (continued)

Segment	Name	Descriptions
GROUP U2: Midtown mix		
31	Urban Achievers	Concentrated in the nation's port cities, Urban Achievers is often the first stop for up-and-coming immigrants from Asia, South America and Europe. These young singles and couples are typically college-educated and ethnically diverse: about a third are foreign-born, and even more speak a language other than English.
40	Close-In Couples	Close-In Couples is a group of predominantly older, African-American couples living in older homes in the urban neighborhoods of mid-sized metros. High school educated and empty nesting, these 55-year-old-plus residents typically live in older city neighborhoods, enjoying secure and comfortable retirements.
54	Multi-Culti Mosaic	Capturing some of the growth of new immigrants to the USA – Hispanics now number 38 million people – this cluster is the urban home for a mixed populace of younger Hispanic, Asian and African-American singles and families. With nearly a quarter of the residents foreign born, Multi-Culti Mosaic is a mecca for first-generation Americans who are striving to improve their lower middle class status.
GROUP U3: Urban cores		
59	Urban Elders	For Urban Elders – a segment located in the downtown neighborhoods of such metros as New York, Chicago, Las Vegas and Miami – life is often an economic struggle. These communities have high concentrations of Hispanics and African-Americans, and tend to be downscale, with singles living in older apartment rentals.
61	City Roots	Found in urban neighborhoods, City Roots is a segment of lower-income retirees, typically living in older homes and duplexes they've owned for years. In these ethnically diverse neighborhoods – more than a third are African-American and Hispanic – residents are often widows and widowers living on fixed incomes and maintaining low-key lifestyles.
65	Big City Blues	With a population that's 50% Latino, Big City Blues has the highest concentration of Hispanic Americans in the nation. But it's also the multi-ethnic address for downscale Asian and African-American households occupying older inner-city apartments. Concentrated in a handful of major metros, these young singles and single-parent families face enormous challenges: low incomes, uncertain jobs and modest educations. More than 40% haven't finished high school.
66	Low-Rise Living	The most economically challenged urban segment, Low-Rise Living is known as a transient world for young, ethnically diverse singles and single parents. Home values are low – about half the national average – and even then less than a quarter of residents can afford to own real estate. Typically, the commercial base of Mom-and-Pop stores is struggling and in need of a renaissance.

Table 8.1 (continued)

Segment	Name	Descriptions
GROUP S1: Elite suburbs		
01	Upper Crust	The nation's most exclusive address, Upper Crust is the wealthiest lifestyle in America – a haven for empty-nesting couples over 55 years old. No segment has a higher concentration of residents earning over $200,000 a year or possessing a postgraduate degree. And none has a more opulent standard of living.
02	Blue Blood Estates	Blue Blood Estates is a family portrait of suburban wealth, a place of million-dollar homes and manicured lawns, high-end cars and exclusive private clubs. As the nation's second-wealthiest lifestyle, it's characterized by married couples with children, college degrees, a significant percentage of Asian Americans and six-figure incomes earned by business executives, managers and professionals.
03	Movers & Shakers	Movers & Shakers is home to America's up-and-coming business class: a wealthy suburban world of dual-income couples who are highly educated, typically between the ages of 35 and 54 and often with children. Given its high percentage of executives and white-collar professionals, there's a decided business bent to this segment: Movers & Shakers rank number-one for owning a small business and having a home office.
06	Winner's Circle	Among the wealthy suburban lifestyles, Winner's Circle is the youngest, a collection of mostly 25- to 34-year-old couples with large families in new-money subdivisions. Surrounding their homes are the signs of upscale living: recreational parks, golf courses and upscale malls. With a median income of nearly $90,000, Winner's Circle residents are big spenders who like to travel, ski, go out to eat, shop at clothing boutiques and take in a show.
GROUP S2: The affluentials		
08	Executive Suites	Executive Suites consists of upper-middle-class singles and couples typically living just beyond the nation's beltways. Filled with significant numbers of Asian Americans and college graduates – both groups are represented at more than twice the national average – this segment is a haven for white-collar professionals drawn to comfortable homes and apartments within a manageable commute to downtown jobs, restaurants and entertainment.
14	New Empty Nests	With their grown-up children recently out of the house, New Empty Nests is composed of upscale older Americans who pursue active – and activist – lifestyles. Nearly three-quarters of residents are over 65 years old, but they show no interest in a rest-home retirement. This is the top-ranked segment for all-inclusive travel packages; the favorite destination is Italy.
15	Pools & Patios	Formed during the postwar Baby Boom, Pools & Patios has evolved from a segment of young suburban families to one for mature, empty-nesting couples. In these stable neighborhoods graced with backyard pools and patios – the highest proportion of homes were built in the 1960s – residents work as white-collar managers and professionals, and are now at the top of their careers.

<div align="center">**Table 8.1** (continued)</div>

Segment	Name	Descriptions
17	Beltway Boomers	The nation's Baby Boomers are now in their forties and fifties, and this segment reflects one group of college-educated, upper-middle-class homeowners. Like many of their peers who married late and are still raising children, these Boomers live in comfortable suburban subdivisions and are still pursuing kid-centered lifestyles.
18	Kids & Cul-de-Sacs	Upscale, suburban, married couples with children – that's the skinny on Kids & Cul-de-Sacs, an enviable lifestyle of large families in recently built subdivisions. With a high rate of Hispanic and Asian Americans, this segment is a refuge for college-educated, white-collar professionals with administrative jobs and upper-middle-class incomes. Their nexus of education, affluence and children translates into large outlays for child-centered products and services.
19	Home Sweet Home	Widely scattered across the nation's suburbs, the residents of Home Sweet Home tend to be upper-middle-class married couples living in mid-sized homes with few children. The adults in the segment, mostly between the ages of 25 and 54, have gone to college and hold professional and white-collar jobs. With their upscale incomes and small families, these folks have fashioned comfortable lifestyles, filling their homes with toys, TV sets and pets.

Group S3: Middleburbs

Segment	Name	Descriptions
21	Gray Power	The steady rise of older, healthier Americans over the past decade has produced one important by-product: middle-class, home-owning suburbanites who are aging in place rather than moving to retirement communities. Gray Power reflects this trend, a segment of older, midscale singles and couples who live in quiet comfort.
22	Young Influentials	Once known as the home of the nation's yuppies, Young Influentials reflects the fading glow of acquisitive yuppiedom. Today, the segment is a common address for young, middle-class singles and couples who are more preoccupied with balancing work and leisure pursuits. Having recently left college dorms, they now live in apartment complexes surrounded by ball fields, health clubs and casual-dining restaurants.
30	Suburban Sprawl	Suburban Sprawl is an unusual American lifestyle: a collection of midscale, middle-aged singles and couples living in the heart of suburbia. Typically members of the Baby Boom generation, they hold decent jobs, own older homes and condos, and pursue cocooning versions of the American Dream. Among their favorite activities are jogging on treadmills, playing trivia games and renting videos.
36	Blue-Chip Blues	Blue-Chip Blues is known as a comfortable lifestyle for young, sprawling families with well-paying blue-collar jobs. Ethnically diverse – with a significant presence of Hispanics and African-Americans – the segment's aging neighborhoods feature compact, modestly priced homes surrounded by commercial centers that cater to child-filled households.

Table 8.1 (continued)

Segment	Name	Descriptions
39	Domestic Duos	Domestic Duos represents a middle-class mix of mainly over 55 singles and married couples living in older suburban homes. With their high-school educations and fixed incomes, segment residents maintain an easy-going lifestyle. Residents like to socialize by going bowling, seeing a play, meeting at the local fraternal order or going out to eat.

GROUP S4: Inner suburbs

Segment	Name	Descriptions
44	New Beginnings	Filled with young, single adults, New Beginnings is a magnet for adults in transition. Many of its residents are twentysomething singles and couples just starting out on their career paths – or starting over after recent divorces or company transfers. Ethnically diverse – with nearly half its residents Hispanic, Asian or African-American – New Beginnings households tend to have the modest living standards typical of transient apartment dwellers.
46	Old Glories	Old Glories are the nation's downscale suburban retirees, Americans aging in place in older apartment complexes. These racially mixed households often contain widows and widowers living on fixed incomes, and they tend to lead home-centered lifestyles. They're among the nation's most ardent television fans, watching game shows, soaps, talk shows and newsmagazines at high rates.
49	American Classics	They may be older, lower-middle class and retired, but the residents of American Classics are still living the American Dream of home ownership. Few segments rank higher in their percentage of home owners, and that fact alone reflects a more comfortable lifestyle for these predominantly white singles and couples with deep ties to their neighborhoods.
52	Suburban Pioneers	Suburban Pioneers represents one of the nation's eclectic lifestyles, a mix of young singles, the recently divorced and single parents who have moved into older, inner-ring suburbs. They live in aging homes and garden-style apartment buildings, where the jobs are blue-collar and the money is tight. But what unites these residents – a diverse mix of whites, Hispanics and African-Americans – is a working-class sensibility and an appreciation for their off-the-beaten-track neighborhoods.

GROUP C1: Second city society

Segment	Name	Descriptions
10	Second City Elite	There's money to be found in the nation's smaller cities, and you're most likely to find it in Second City Elite. The residents of these satellite cities tend to be prosperous executives who decorate their $200,000 homes with multiple computers, large-screen TV sets and an impressive collection of wines. With more than half holding college degrees, Second City Elite residents enjoy cultural activities – from reading books to attending theater to dance productions.
12	Brite Lites, Li'l City	Not all of the America's chic sophisticates live in major metros. Brite Lights, Li'l City is a group of well-off, middle-aged couples who have settled in the nation's satellite cities. Residents of these typical DINK (double income, no kids) households have college educations, well-paying business and professional careers and swank homes filled with the latest technology.

(continued)

<div align="center">

Table 8.1 (continued)

</div>

Segment	Name	Descriptions
13	Upward Bound	More than any other segment, Upward Bound appears to be the home of those legendary Soccer Moms and Dads. In these small satellite cities, upper-class families boast dual incomes, college degrees and new split-levels and colonials. Residents of Upward Bound tend to be kid-obsessed, with heavy purchases of computers, action figures, dolls, board games, bicycles and camping equipment.

GROUP C2: City centers

Segment	Name	Descriptions
24	Up-and-Comers	Up-and-Comers is a stopover for young, midscale singles before they marry, have families and establish more deskbound lifestyles. Found in second-tier cities, these mobile, twentysomethings include a disproportionate number of recent college graduates who are into athletic activities, the latest technology and nightlife entertainment.
27	Middleburg Managers	Middleburg Managers arose when empty-nesters settled in satellite communities that offered a lower cost of living and more relaxed pace. Today, segment residents tend to be middle-class and over 55 years old, with solid managerial jobs and comfortable retirements. In their older homes, they enjoy reading, playing musical instruments, indoor gardening and refinishing furniture.
34	White Picket Fences	Midpoint on the socioeconomic ladder, residents in White Picket Fences look a lot like the stereotypical American household of a generation ago: young, middle-class, married with children. But the current version is characterized by modest homes and ethnic diversity – including a disproportionate number of Hispanics and African-Americans.
35	Boomtown Singles	Affordable housing, abundant entry-level jobs and a thriving singles scene – all have given rise to the Boomtown Singles segment in fast-growing satellite cities. Young, single and working-class, these residents pursue active lifestyles amid sprawling apartment complexes, bars, convenience stores and laundromats.
41	Sunset City Blues	Scattered throughout the older neighborhoods of small cities, Sunset City Blues is a segment of lower-middle-class singles and couples who have retired or are getting closed to it. These empty-nesters tend to own their homes but have modest educations and incomes. They maintain a low-key lifestyle filled with newspapers and television by day, and family-style restaurants at night.

GROUP C3: Micro-city blues

Segment	Name	Descriptions
47	City Startups	In City Startups, young, multi-ethnic singles have settled in neighborhoods filled with cheap apartments and a commercial base of cafes, bars, laundromats and clubs that cater to twentysomethings. One of the youngest segments in America – with ten times as many college students as the national average – these neighborhoods feature low incomes and high concentrations of Hispanics and African-Americans.

Table 8.1 (continued)

Segment	Name	Descriptions
53	Mobility Blues	Young singles and single parents make their way to Mobility Blues, a segment of working-class neighborhoods in America's satellite cities. Racially mixed and under 25 years old, these transient Americans tend to have modest lifestyles due to their lower-income blue-collar jobs. Surveys show they excel in going to movies, playing basketball and shooting pool.
60	Park Bench Seniors	Park Bench Seniors typically are retired singles who live in the racially mixed neighborhoods of the nation's satellite cities. With modest educations and incomes, these residents maintain low-key, sedentary lifestyles. Theirs is one of the top-ranked segments for TV viewing, especially daytime soaps and game shows.
62	Hometown Retired	With three-quarters of all residents over 65 years old, Hometown Retired is one of the oldest lifestyles. These racially mixed seniors tend to live in aging homes – half were built before 1958 – and typically get by on social security and modest pensions. Because most never made it beyond high school and spent their working lives at blue-collar jobs, their retirements are extremely modest.
63	Family Thrifts	The small-city cousin of inner-city districts, Family Thrifts contain young, ethnically diverse parents who have lots of children and work entry-level service jobs. In these apartment-filled neighborhoods, visitors find the streets jam-packed with babies and toddlers, tricycles and basketball hoops, Daewoos and Hyundais.

GROUP T1: Landed gentry		
05	Country Squires	The wealthiest residents in exurban America live in Country Squires, an oasis for affluent Baby Boomers who've fled the city for the charms of small-town living. In their bucolic communities noted for their recently built homes on sprawling properties, the families of executives live in six-figure comfort. Country Squires enjoy country club sports like golf, tennis and swimming as well as skiing, boating and biking.
09	Big Fish, Small Pond	Older, upper class, college-educated professionals, the members of Big Fish, Small Pond are often among the leading citizens of their small-town communities. These upscale, empty-nesting couples enjoy the trappings of success, belonging to country clubs, maintaining large investment portfolios and spending freely on computer technology.
11	God's Country	When city dwellers and suburbanites began moving to the country in the 1970s, God's Country emerged as the most affluent of the nation's exurban lifestyles. Today, wealthier communities exist in the hinterlands, but God's Country remains a haven for upper-income couples in spacious homes. Typically college-educated Baby Boomers, these Americans try to maintain a balanced lifestyle between high-power jobs and laid-back leisure.
20	Fast-Track Families	The migration of upscale city dwellers out to the countryside can be seen in the emergence of this exurban cluster. Fast-Track Families is filled with middle-aged parents who have the

Table 8.1 (continued)

Segment	Name	Descriptions
		disposable income and educated sensibility for a granola-and-grits lifestyle: they fish, boat and shop over the Internet – all at high rates.
25	Country Casuals	There's a laid-back atmosphere in Country Casuals, a collection of middle-aged, upper-middle-class households that have started to empty-nest. Workers here – and most households boast two earners – have well-paying blue- or white collar jobs, or own small businesses. Today these Baby-Boom couples have the disposable income to enjoy traveling, owning timeshares and going out to eat.

GROUP T2: Country comfort

Segment	Name	Descriptions
23	Greenbelt Sports	A segment of middle-class exurban couples, Greenbelt Sports is known for its active lifestyle. Most of these middle-aged residents are married, college-educated and own new homes; about a third have children. And few segments have higher rates for pursuing outdoor activities such as skiing, canoeing, backpacking, boating and mountain biking.
28	Traditional Times	Traditional Times is the kind of lifestyle where small-town couples nearing retirement are beginning to enjoy their first empty-nest years. Typically in their fifties and sixties, these middle-class Americans pursue a kind of granola-and-grits lifestyle. On their coffee tables are magazines with titles ranging from *Country Living* and *Country Home* to *Gourmet* and *Forbes.* But they're big travelers, especially in recreational vehicles and campers.
32	New Home-steaders	Young, middle-class families seeking to escape suburban sprawl find refuge in New Homesteaders, a collection of small rustic townships filled with new ranches and Cape Cods. With decent-paying jobs in white-collar and service industries, these dual-income couples have fashioned comfortable, child-centered lifestyles, their driveways filled with campers and powerboats, their family rooms with PlayStations and Game Boys.
33	Big Sky Families	Scattered in placid towns across the American heartland, Big Sky Families is a segment of young rural families who have turned high school educations and blue-collar jobs into busy, middle-class lifestyles. Residents like to play baseball, basketball and volleyball, besides going fishing, hunting and horseback riding. To entertain their sprawling families, they buy virtually every piece of sporting equipment on the market.
37	Mayberry-ville	Like the old Andy Griffith show set in a quaint picturesque berg, Mayberry-ville harks back to an old-fashioned way of life. In these small towns, middle-class couples and families like to fish and hunt during the day, and stay home and watch TV at night. With lucrative blue-collar jobs and moderately priced housing, residents use their discretionary cash to purchase boats, campers, motorcycles and pickup trucks.

Table 8.1 (continued)

Segment	Name	Descriptions
GROUP T3: Middle America		
38	Simple Pleasures	With more than two-thirds of its residents over 65 years old, Simple Pleasures is mostly a retirement lifestyle: a neighborhood of lower-middle-class singles and couples living in modestly priced homes. Many are high school-educated seniors who held blue-collar jobs before their retirement. And a disproportionate number served in the military; no segment has more members of veterans clubs.
42	Red, White & Blues	The residents of Red, White & Blues typically live in exurban towns rapidly morphing into bedroom suburbs. Their streets feature new fast-food restaurants, and locals have recently celebrated the arrival of chains like Wal-Mart, Radio Shack and Payless Shoes. Middle-aged, high school educated and lower-middle class, these folks tend to have solid, blue-collar jobs in manufacturing, milling and construction.
43	Heartlanders	America was once a land of small middle-class towns, which can still be found today among Heartlanders. This widespread segment consists of middle-aged couples with working-class jobs living in sturdy, unpretentious homes. In these communities of small families and empty-nesting couples, Heartlanders pursue a rustic lifestyle where hunting and fishing remain prime leisure activities along with cooking, sewing, camping and boating.
45	Blue Highways	On maps, blue highways are often two-lane roads that wind through remote stretches of the American landscape. Among lifestyles, Blue Highways is the standout for lower-middle-class couples and families who live in isolated towns and farmsteads. Here, Boomer men like to hunt and fish, the women enjoy sewing and crafts, and everyone looks forward to going out to a country music concert.
50	Kid Country, USA	Widely scattered throughout the nation's heartland, Kid Country, USA is a segment dominated by large families living in small towns. Predominantly white, with an above-average concentration of Hispanics, these young, these working-class households include homeowners, renters and military personnel living in base housing; about 20% of residents own mobile homes.
51	Shotguns & Pickups	The segment known as Shotguns & Pickups came by its moniker honestly: it scores near the top of all lifestyles for owning hunting rifles and pickup trucks. These Americans tend to be young, working-class couples with large families – more than half have two or more kids – living in small homes and manufactured housing. Nearly a third of residents live in mobile homes, more than anywhere else in the nation.
GROUP T4: Rustic living		
48	Young & Rustic	Like the soap opera that inspired its nickname, Young & Rustic is composed of young, restless singles. Unlike the glitzy soap denizens, however, these folks tend to be lower income, high school-educated and living in tiny apartments in the nation's exurban towns. With their service industry jobs and modest incomes, these folks still try to fashion fast-paced lifestyles centered on sports, cars and dating.

Table 8.1 (continued)

Segment	Name	Descriptions
55	Golden Ponds	Golden Ponds is mostly a retirement lifestyle, dominated by downscale singles and couples over 65 years old. Found in small bucolic towns around the country, these high school-educated seniors live in small apartments on less than $25,000 a year; one in five resides in a nursing home. For these elderly residents, daily life is often a succession of sedentary activities such as reading, watching TV, playing bingo and doing craft projects.
56	Crossroads Villagers	With a population of middle-aged, blue-collar couples and families, Crossroads Villagers is a classic rural lifestyle. Residents are high school-educated, with lower-middle incomes and modest housing; one quarter live in mobile homes. And there's an air of self-reliance in these households as Crossroads Villagers help put food on the table through fishing, gardening and hunting.
57	Old Milltowns	With the shrinking of the nation's manufacturing sector, America's once-thriving factory towns have aged, as have their residents. Old Milltowns reflects the decline of these small industrial communities, now filled with retired singles and couples living quietly on fixed incomes. These home-centered residents make up one of the top segments for daytime television.
58	Back Country Folks	Strewn among remote farm communities across the nation, Back Country Folks are a long way away from economic paradise. The residents tend to be poor, over 55 years old and living in older, modest-sized homes and manufactured housing. Typically, life in this segment is a throwback to an earlier era when farming dominated the American landscape.
64	Bedrock America	Bedrock America consists of young, economically challenged families in small, isolated towns located throughout the nation's heartland. With modest educations, sprawling families and blue-collar jobs, many of these residents struggle to make ends meet. One quarter live in mobile homes. One in three haven't finished high school. Rich in scenery, Bedrock America is a haven for fishing, hunting, hiking and camping.

survey data rather than neighborhood aggregate data and, hence, it is expected to be more accurate in targeting households (http://www. cohorts.com). The source data are derived from two leading individual data providers, Experian and Equifax. Cohorts ends up with 30 clusters, each labeled with names such as Alex & Judith (affluent empty-nesters) and Chad & Tammie (young families) as summarized in Table 8.2.[6] So far more than 100 consumer marketers from various industries have employed Cohorts.

---

[6] Actually, there are 31 clusters in Cohorts. The last cluster named "Omegas" is formed from statistical anomalies that did not fit into 30 cohesive clusters.

Table 8.2 Brief description of the "2007 Cohorts Segments" (From *2007 Cohorts Segments*, Courtesy of: Looking Glass Inc. Accessed August 8, 2007. Available at: http://www.cohorts.com/pdf/2007_Briefs.pdf.)

Cohort segment name	Description	Median age	Median income
Married couples			
Alex & Judith	**Affluent Eempty-nesters** Dual-income, older couples who use their high discretionary incomes to enjoy all aspects of the good life.	61	$144,000
Jeffrey & Ellen	**Affluent couples with kids** Urban families who, despite having children at home, have sufficient financial resources to own the latest high-tech products and to lead very active recreational and cultural lifestyles.	43	$142,000
Barry & Kathleen	**Affluent professional couples** Educated, dual-income, childless couples who have connoisseur tastes and are focused on their careers, staying fit and investing.	46	$133,000
Stan & Carole	**Upscale middle-aged couples** Unburdened by children, these credit-worthy, dual-income couples divide their time between the great outdoors and domestic hobbies.	50	$75,000
Brett & Tracey	**Hyperactive newlyweds** Young, dual-income, childless couples whose energies are channeled into active sports, outdoor activities, careers and their home lives.	31	$65,000
Danny & Vickie	**Teen-dominated families** Middle-aged, middle-income families whose teen-dominated households keep busy with outdoor activities, computers and video games.	42	$59,000
Burt & Marilyn	**Mature couples** Comfortable, close-to-retirement homeowners who are active investors and who engage in charitable activities, travel, politics and their grandchildren.	67	$58,000
Todd & Wendy	**Back-to-school families** Families with mid-range incomes, pre-adolescent kids, pets, and lots of video, computer and outdoor activities to keep them occupied.	38	$57,000
Chad & Tammie	**Young families** Up-and-coming young families who curtail their lifestyle expenses through less-costly outdoors activities and working around the house.	31	$53,000
Frank & Shirley	**Older couples raising kids** Conservative grandparents, and older parents raising kids, whose home-oriented lifestyles include pets, home workshop, gardening, and sweepstakes.	60	$50,000
Ronnie & Debbie	**Working-class couples** Moderate-income couples with traditional interests including fishing, hunting, automotive work and crafts.	48	$38,000

Table 8.2 (continued)

Cohort segment name	Description	Median age	Median income
Eric & Rachel	**Young, married starters** Young, childless renters whose lifestyle patterns include outdoor activities like camping, fishing and running, as well as automotive work and video games.	28	$20,000
Elwood & Willamae	**Modest-income grandparents** Retired couples with modest incomes who dote on their grandchildren and engage primarily in domestic pursuits.	72	$20,000

Single Females

Elizabeth	**Savvy career women** Affluent, working women with sophisticated tastes, very active lifestyles and good investing habits.	43	$182,000
Virginia	**Upscale mature women** Older women approaching or enjoying retirement, who travel and have upscale interests, including charitable causes and investments.	60	$72,000
Allison	**Educated working women** Childless, professional women building their careers, developing sophisticated tastes and staying fit.	32	$53,000
Andrea	**Single moms with careers** Successful, professional single mothers who balance their careers with the demands of raising their children.	40	$50,000
Bernice	**Active grandmothers** Home-oriented women who enjoy handicrafts, indoor gardening and their grandchildren.	62	$36,000
Penny	**Working-class women** Childless female office workers who are concerned with their appearance; enjoy music, pets and handicrafts; and add intrigue to their lives with the prospect of winning the big sweepstakes.	43	$18,000
Denise	**Single moms on a budget** Single mothers with modest incomes who indulge their kids with video games, movies, and music, and who try to find time for themselves.	36	$17,000
Megan	**Fit & stylish students** Young, fashion-conscious, career-minded female students who enjoy music, aerobic sports and the latest in high tech.	26	$16,000
Minnie	**Fixed-income grandmothers** Older single women who spend lots of time on their grandchildren, handicrafts and religious reading.	73	$11,000

Single males

Jonathan	**Elite single men** High-powered, career-driven men with sophisticated tastes, extensive investments, and the means to travel the world.	45	$186,000

Table 8.2 (continued)

Cohort segment name	Description	Median age	Median income
Sean	**Affluent guys** Affluent, health- and fitness-minded men with investments and upscale interests.	46	$97,000
Harry	**Well-to-do gentlemen** Mature men who are savvy about their investments, travel and politics.	59	$49,000
Ryan	**Energetic young guys** Young, physically active men with strong career drives and upscale interests, including electronics and technology.	33	$48,000
Randy	**Single dads** Single fathers who enjoy outdoor activities, their home workshops and electronic entertainment with their kids.	38	$46,000
Jerry	**Working-class guys** Blue-collar men who spend their free time in the garage or outdoors.	48	$19,000
Jason	**Male students and grads** Physically active, technologically inclined young men finishing school or embarking on their first job.	26	$17,000
Elmer	**Sedentary men** Aging, sedentary men with fixed incomes and few interests beyond their grandchildren and their gardens.	73	$17,000

Households that defy classification

Omegas	Omegas are people who are impossible to classify distinctly. They may be married or single, homeowners or renters, 18–65 years old, have incomes that range from very low to six figures, and enjoy numerous and diverse interests.

Clients have begun to criticize the nature of the clusters developed by the external data providers. Purchase behavior observed in financial services will be different from purchase behavior in groceries. Different industries will have different demographic and lifestyle drivers. In the 1980s, Pinpoint has developed FiNPiN, a consumer classification system designed for the financial service industry (Winters 1993). Several vendors have followed suit. For example, Claritas introduced its own industry-specific segmentation product called P$YCLE. Designed for financial institutions, P$YCLE segments US households into 42 clusters mainly in terms of their financial behavior. Claritas went one step further and develop a segmentation product named LifeP$YCLE for insurance marketers. Claritas gathers household data from a syndicated survey of 90,000 consumers about their use of financial services.

Finally, the segmentation vendors have actively incorporated databases from other specialty research firms. For example, by utilizing the data from Nielsen Marketing Research, Simmons Marketing Research Bureau, credit data, electoral rolls and additional customer survey, vendors can now report more extensive purchase behavior (e.g., usage pattern for particular products and services) for each segment.

8.3.2.5 List Rentals for Prospecting

Database marketers can increase the efficiency of their customer acquisition efforts significantly by carefully selecting the right mailing lists for prospecting. For example, a mail order company of women's petite sizes may want to target only short women (Hatch 1995). Driver's license data from state motor vehicle bureaus have information on drivers' heights and weights. This mail order company may seek and find a list that contains women and their heights and weights.

The list rental industry is very diverse. There is no standard way of categorizing the different types of lists. Roberts and Berger (1999) classify them into consumer and business lists first. Alternatively, lists can be categorized into house lists, response lists and compiled lists. A house list is the list of customers in the company's own customer information file, while (consumer) response lists are some other company's house list or "subscriber" list (Sect. 8.3.2.3). A compiled list is a list of customers compiled from public records, phone directories, or professional associations. The reason these lists are called compiled is that somebody has actually compiled the lists based on data available typically from various sources. Usually the compiled list has some common identifiable characteristics and its size is large with lower unit price. For example, InfoUSA is renting complied lists of 95 million ethnic individuals, 37 million homeowners, 8.4 million new movers, and so on. Hence, a compiled list is appropriate for wide market coverage. On the other hand, a response list is a list of customers who have either purchased or requested information from specific companies. Its mailing response rates are expected to be high since consumers in the response list have previously shown their interests to respond to the mailing. Moreover, there are a wide variety of response lists available, and customers in each response list often show interests on particular products and services. For example, subscribers of a cat magazine will have strong interests in cats. Cat food sellers may want to rent this subscriber list.

There are hundreds of thousands of list buyers and sellers. The "market" for lists is generally organized as follows:

The list buyer wishes to purchase a list say for prospecting. The buyer employs a list broker to find good lists. On the other side, there are owners of lists. This may be a list compiler as described above, or an individual company. Individual companies often employ list managers who are in

charge of selling their customer list. For example, Best Buy may be interested in attracting customers to its stores with an offer for an Apple IPod. Crutchfield, an electronics cataloger, has a list of its recent customers. Best Buy's list broker and Crutchfield's list manager get together and negotiate terms by which Best Buy may rent Crutchfield's list. Crutchfield of course may decide not to rent to Best Buy if it thinks Best Buy will steal sales from them. But if Crutchfield is willing, the broker and list manager negotiate price, terms, etc. The list owner, in this case Crutchfield, then pays both the broker and the list manager a certain percentage. It is an interesting arrangement. Essentially, the list buyer does not pay *directly* for the services of the list broker or list manager. However, one might conjecture that the list buyer *indirectly* pays because undoubtedly the fact that the list owner pays both the broker and the list manager keeps prices relatively high.

For years, the list industry was relatively low-tech. Recently, however, computerized list search engines have appeared. These allow a broker, or the list buyer, to search directly for lists. One such search engine is Nextmark (http://www.nextmark.com). See Roberts and Berger (1999) for more detail about the list rental industry.

Another way to acquire a prospect list is through list exchange. This process is managed by companies such as Abacus (http://www.abacus-us.com). Abacus maintains what they call a cooperative database. Companies contribute names to the database, and in return can obtain new names from the database. List exchanges became more accepted during the cash crunch of the past few years. According to the 2004 *Catalog Age* Benchmark Report on Lists, 30% of all respondents have negotiated list exchanges (Del Franco 2004). They were more willing to exchange with non-competitors than competitors. Chen et al. (2001) showed that information sharing (or list exchange) can be profitable for two competing firms under reasonable conditions.

8.3.3 Primary Data

If the data elements are not available from internal or external sources, they need to be collected directly from a consumer survey. These are costly and time-consuming, but often worthwhile. We will not discuss various statistical issues on collecting primary data (survey, focus group and in-depth interview, observational data, etc.) since they are well documented in traditional marketing research textbooks. Instead, we provide a couple of real world examples how companies collect their primary data.

Traditionally packaged goods manufacturers are mass marketers. However, Quaker has seen the potential of one-to-one marketing. In order to create its own customer list, in 1990 Quaker mailed cents-off coupons, each of which had a unique household number. Analyzing who redeemed coupons and

when they were redeemed, Quaker could learn customer-level purchase behavior for its product. Moreover, this information could be used to customize advertising and promotion to the unique needs of individual households (Mollie 1991).

Philip Morris gives another excellent example. Because of the increasing restrictions on tobacco advertising, it is necessary for Philip Morris to build its own customer information file and reach smokers directly. Customers fill out detailed questionnaires to get free shirts and sleeping bags (Berry 1994). With its 26 million smokers' names and addresses, Philip Morris sends targeted coupons and asks for grassroots support for their lobbying efforts. Similarly, Seagram has built its own customer information file and tracked consumers' names and addresses, the brands and types of alcohol they drink, their sex, birth date, income, and how many bottles they purchase in an average month (Berry 1994).

A strategic alliance with credit card companies or Internet portals often reduces costs of collecting primary data significantly. For example, GM offered GM credit card with MasterCard in 1992, and, as a result, could build a customer database with tens of millions of customers. More recently, several offline-based companies formed strategic alliances with online companies to target online customers and track online customer behaviors. For example, United Artists Theatre Circuit, one of the largest theater chains in the USA, made a long-term strategic alliance with AOL (Time Warner 1999). The alliance allows United Artists efficiently to reach the largest group of moviegoers in cyberspace.

Another important use of primary data, perhaps in its infancy, is to use surveys to gather competitive information (Kamakura and Wedel 2003; Du et al. 2005). For example, a company could survey its customers and ask them how often they purchase from a competitor, what competitive products they own, and what percentage of purchases are from a competitor ("share-of-wallet"). They would have these competitive data for just the sample, say 1,000 customers. The company would then run a predictive model to predict say share-of-wallet as a function of variables it has on all its customers. This model would then be used to score the rest of the customer file. So each customer would be scored in terms of their predicted share-of-wallet, i.e., what percentage of their business is with the company as opposed to the competition.

Du et al. (2005) noted that few firms were collecting competitive information. They suggested that the survey-based approach should be utilized to augment the company's interactions with its customers by adding the inferred competitive interactions. Their empirical analysis indicated that the volume customers transact within a firm has little correlation with the volume they transact with the firm's competitors. In addition, a small percentage of customers account for a large portion of all the competitive transactions, suggesting considerable potential to increase sales if these customers can be correctly identified and encouraged to switch.

8.4 The Destination Marketing Company

Competition among external data providers is getting stiffer as their products and services become less differentiated. In result, data-selling business has become a low margin business. To differentiate from others and increase value-added, external data vendors are beginning to integrate downstream and provide services such as predictive modeling, customer segmentation, cross-selling modeling, and other marketing consulting services. By providing selling and marketing research services, in addition to the data that drive these efforts, these vendors have taken one step further to increase their value share in the total value chain. We observe that a big marketing broker is emerging. We call it "Destination Marketing Company (DMC)." For example, Acxiom's business covers list-selling data enhancement, analytic and marketing services consulting, data quality assessment, and direct mail fulfillment service.

The essential role of the DMC is to connect sellers and buyers. The DMC is like a big marketing broker. A company can outsource its whole marketing function to the DMC. For example, a number of companies employ outside advertising agencies or marketing research companies to assist its marketing department. Similarly, a company may not need a marketing department by outsourcing all of its marketing function from the DMC. The DMC attempts to find potential customers for its products and services, sends communication messages, and closes the sales. The DMC may be compensated on a commission basis.

The DMC has two major resources: customer information and database marketing knowledge. That is, the DMC should have a huge customer information file that consists of individual and business customers, or at least be particularly astute at obtaining databases from external sources. The DMC is also knowledgeable in all aspects of database marketing techniques such as database management and predictive modeling. If a client company asks the DMC to sell its products, the DMC selects a group of customers from its customer information file to be predicted to have high probability of purchasing the product. In searching for the most efficient way of selling the product, the DMC also selects the best communication and sales channel for each customer. Once the sales are closed, the results are recorded back into the customer information file.

Can we find any empirical evidence of an emerging DMC? As discussed, some external data vendors such as Acxiom and Harte-Hanks are moving in this direction. Companies have been willing to outsource their advertising and marketing research function. And some companies use agencies and brokers to sell their products. However, you can argue that while companies may be willing to outsource some part of their marketing function, they should be responsible for the overall marketing strategy, including the target market and product positioning. Without a solid marketing strategy, products become commodities. For example, the profit margins for OEM manufacturers

are very low. They may not have enough resources and capabilities to have their own one-to-one marketing infrastructure but outsource their marketing to the DMC. Therefore, while the Destination Marketing Company has its advantages, it can be a consequence, as well as a cause, of a poor marketing strategy. However, if the DMC is integrated into a firm's marketing group, and that group has a solid grasp of the big picture, the DMC can be extremely valuable.