# 5

# Internal secondary data and the use of databases

> " *If all you do is stick to research in an organisation like this, then you're dead really.*[1] "

## Objectives

After reading this chapter, you should be able to:

1 describe the nature and purpose of internal sources of secondary data;

2 describe how different technological developments have increased the array of internally generated secondary data;

3 understand how databases are developing into powerful means to understand consumer behaviour through 'electronic observation';

4 understand how databases support traditional forms of marketing research to build up behavioural and attitudinal 'pictures' of target markets;

5 understand how geodemographic information systems can help in integrating data sources and in the graphical display of findings in a non-statistical manner;

6 describe how the link-up of different databases and survey data can be developed through the use of datawarehouses and be analysed through data mining techniques;

7 understand international data capture issues;

8 understand the ethical problems of having individual consumer data held on databases.

STAGE 1
Problem definition

STAGE 2
Research approach developed

STAGE 3
Research design developed

STAGE 4
Fieldwork or data collection

STAGE 5
Data preparation and analysis

STAGE 6
Report preparation and presentation

# Overview

Marketing research as a function does not support marketing decision making in isolation. As discussed in Chapter 1, it is seen by many industry commentators as a component of strategic marketing intelligence. Many information technology and system advances have been made in recent years that have fundamentally changed the way that marketing decisions are supported. For example, significant developments in database technology have meant that scanning systems in retail stores, loyalty card data, store panel data and survey data can be fused together to present very clear and up-to-date 'pictures' of consumers. As well as giving direct support to the marketer, these systems give more focus to marketing research activity and direct support to many stages of research.

This chapter describes how internal secondary data and databases have developed to make major impacts upon how decision-makers are supported. The data collected and analysed through database marketing can be seen as secondary data sources. As with all good secondary data sources, they have a major impact upon the conduct and direction of primary data collection, analyses and interpretation. There are also many ethical issues related to utilising internally generated customer data and the use of databases.

We introduce our discussion with an example of the use of databases. Databases generated within companies or bought in from specialist sources are primarily viewed as a tool to generate direct sales and target promotion activities. However, internally generated customer databases are a secondary data source of value to marketing researchers. Technological developments in the collection, analysis and presentation of data present great opportunities to researchers. There are also potential conflicts with the core premise of anonymity in marketing research. The following example illustrates how a customer database can reveal behavioural characteristics of customers, valuable information to decision-makers in its own right. From an understanding of who Homebase's customers were and how they were behaving, marketing research could be designed to understand the underlying motives for this behaviour. The database analysis acted as an excellent foundation of a research design in deciding who to research and what issues to focus upon.

**Example**  Database mining meets marketing research[2]

The Save and Spend loyalty card of DIY retailer Homebase helped to create a database of more than 4 million active customers and their sales transactions. Homebase could see there were many different types of purchase patterns emerging and asked its marketing research agency for help to understand them. The database told it where people lived and what they were spending, but not things like what sort of house they lived in or their attitudes towards DIY. If Homebase could understand what lay behind patterns of DIY purchasing behaviour for customer types, it could use this information to target its advertising and promotions. Data from store exit surveys were overlaid onto purchasing behaviour from the database. Factors such as the frequency of store visits dominated a base level of segmentation. Statistical modelling then revealed the importance of factors such as distance to store, lifestage and attitudes. This resulted in a clear and actionable segmentation approach. For example, from taking one segment of people looking to decorate their homes in a period style, and overlaying that on the level of disposible income, strong differences were found between men and women. Advertising and promotions could then be tailored to

three core types of cardholder: those who dislike DIY and just want value for money; those doing DIY for recreation; and those who want to distinguish their property and hence want the latest gadgets. Homebase calculated that the project enabled it to generate £10 return for every £1 marketing spend.

## Internal secondary data

Chapter 4 described the nature, purpose and value of secondary data to marketing researchers. A vital source of secondary data comes from within organisations that commission marketing research, namely internal secondary data. These data are generally seen as being 'operational data', i.e. data that represent the daily activities and transactions of a business. Daily transactions may be held in different departments such as sales, accounts or human resources and stored in different manners. The use of operational data has presented opportunities to researchers for as long as businesses have been recording their daily transactions. Even in the days of transactions being recorded manually, it has been a task for marketing researchers to track down different sources of data and analyse them. Locating and analysing internal sources of secondary data should be the starting point in any marketing research project. The main reasons are that, as these data have already been collected, there are no additional data collection costs, there should be no access problems (though individual managers may make access difficult for personal or political reasons) and the quality of the data should be easier to establish (in comparison with externally generated data).

Most organisations have a wealth of in-house information even if they are not marketing or customer focused, so some data may be readily available. For example, imagine a timber merchant that sells wood to builders and cabinetmakers. It creates invoices for all its sales. Its accounts department handles this process and maintains the data that it generates. Yet, there exists much consumer behaviour data in these invoices. They could be analysed by:

- What products customers buy.
- Which customers buy the most products.
- Which customers repeat purchases.
- Which customers appear only when there are special offers.
- Where these customers are located.
- How these customers pay – by cash or credit.
- Which customers are the most profitable.
- Seasonal patterns of purchasing behaviour by product types and customer types.

There may also be data that relate to promotional activities such as spending on advertising, trade fairs, sponsorship deals or personal selling. The researcher could look for details of spending in these areas and seek correlations with any of the analyses of customer behaviour. The task facing the marketing researcher is to search for such data, conduct analyses and present these to decision-makers to interpret. Such a process may focus the thoughts of decision-makers, by realising the potential that lies in these data. With this focus, other types of **operational data** and their value may be realised, managers in other parts of the organisation may release data that they guarded, and connections to sources of intelligence may be generated. Here lies the basis of generating clearly focused primary data collection, and effective marketing research.

**Operational data**
Data generated about an organisation's customers, through day-to-day transactions.

More marketing decision-makers have realised the benefits of analysing customer data. This realisation and technological developments in collecting, analysing and presenting customer data have given birth to a concept known as customer relationship management (CRM). The growth of CRM has emerged from the use and subsequent integration of various direct marketing channels (direct mail, telemarketing) and the rise of e-business, e-communication and the increasing use of the Internet as a conduit for customer care and sales.[3] The main challenge of implementing CRM is to integrate customer data from the post, telephone, personal visits and the Internet into a central database so that it holds all transactions and contacts with each customer allied with the ability to update the data constantly and to access it immediately whenever necessary.[4] An illustration of the use and effect of CRM is presented in the following example.

**Example** | **Getting to know clients lifts profits[5]**

The principles of CRM are simple. Businesses gather accurate information about customers and prospects. Having identified the customers or segments that account for the highest profits, they devise marketing strategies that differentiate between different groups. Greater resources are focused on higher value customers. Every opportunity is used to amass additional information about each client to personalise sales messages and build a closer relationship.

When Mercury Asset Management began to experiment with CRM, it was able to demonstrate that 59% of its profits came from 1% of its customers. Over six months it moved from having a standard type of literature for all its customers and prospects to no fewer than 7,700 types of literature. Digital printing made this personalisation relatively simple. The first stage of this personalisation process was the compilation of comprehensive information about customers.

**Customer database**
A database that details characteristics of customers and prospects that can include names and addresses, geographic, demographic and buying behaviour data.

Many companies see the benefits of compiling comprehensive information about their customers and invest great amounts in developing and maintaining a **customer database**. The customer database for many companies is used to drive all marketing strategies. Customer data can be created by companies from past records, promotional devices such as competitions or direct response advertising. The database is used to stimulate marketing activities, and the response from these activities is fed back to improve and update it. Database marketing is a circular activity where every iteration improves the total value of the database. So, when consumers 'hook up' to an online company, through their PC, their TV or even their mobile phone, they help to develop the customer database.[6] They supply personal details, their choice of products or services and their means of payment. From the knowledge gained from these transactions, new targeted offerings can be formulated, and the nature of the customers' response can be recorded. As the decision-maker learns more about the customers from transaction data, the awareness of gaps in knowledge becomes more focused. Where those gaps cannot be filled with transaction data, the marketing researcher plays a vital role in the generation and interpretation of bespoke primary data. In the development of good research design, the customer database can be seen as a resource to the marketing researcher when conducting internal secondary data searches.

There is a whole array of different means to capture electronically customer transaction behaviour and even potential customers through their search for information to buy services and products. It is beyond the scope of this text to describe the array of CRM technologies, Internet trading and online business. We therefore will just concentrate on a concept introduced in the last chapter, concerning scanned data. From a basis of scanned data we illustrate how other data sources, including primary data from marketing research studies, can be integrated. This serves as a link to examine how decision-makers and researchers make sense of the masses of customer data that may be collected.

## Scanning devices

**Scanning device**
Technology that reads the UPC from merchandise by passing it over a laser scanner.

One of the most fundamental technological breakthroughs that has allowed the monitoring of product sales has been the bar code. With **scanning devices** to read bar codes we now have the ability to count and analyse sales quickly. If a new product is launched, scanning data can monitor sales on a daily basis, breaking down the sales by advertising region and the type of outlet. The scanning device is an electronic means of observation. Consumers do not answer any questions, do not identify themselves; they merely enjoy the benefits of supermarket queues moving far more quickly compared with the days of checkout assistants manually entering the prices for individual goods.

What product bar codes and scanning devices do not do is classify consumers. Classification is fundamental to marketing research, marketing segmentation techniques and ultimately a vast array of marketing decisions. Is the new brand of yogurt more popular with younger age groups compared with older groups? Have more Calvin Klein shirts been sold to male or female buyers? The following example of the Tesco loyalty card illustrates perhaps one of the most sophisticated means of understanding the characteristics of consumers and of linking their characteristics to their purchasing behaviour. When consumers sign up for a loyalty card, they supply Tesco with personal details that enables them to be classified in a number of manners. Using their loyalty card when making a purchase allows Tesco to make the link between what has been bought with who made the purchase. The example shows what impact this means of electronic observation had upon the performance of Tesco.

## Relating customer data to scanning systems

**Example**  **Tesco Clubcard[7]**

On 13 February 1995, the face of retailing in the UK started to change, with the national rollout of Tesco's loyalty scheme, the Clubcard. It began as an attempt to maintain market share by sustaining and developing the value of Tesco's customer base against the then market leader Sainsbury's and fending off threats lower down the grocery food chain, from price cutters such as Asda. The Clubcard enables it to track £4 out of every £5 spent in its stores and helped to transform its entire retail strategy, helping it to leapfrog over Sainsbury's to be the market leader. The Clubcard has 10 million active members and generates more than £100 million in incremental sales every year. The primary reason for its success is the rich seam of data it provides for the chain. This, says a Tesco spokesman, 'helps us to understand what customers want and develop products and services that are relevant'. The transactional data generated by Clubcard gives Tesco almost unparalleled customer insight. This, combined with the intelligence generated from the past use of the data, makes it one of the most successful schemes of its sort. The card is helping Tesco stay well ahead of its rivals and has been instrumental in winning £1 out of every £8 spent nationally and helping it break the £2 billion profit barrier.

The essence of the Tesco example is that the scanned data of products sold in Tesco's stores is linked to known customers: the system links customer identification to product usage. Any promotional offers, competitive activity, new product offerings, discounts, where to locate a new store, to name but a few marketing activities, can be analysed and related to classifications of customer. Many businesses have seen the transformation in the fortunes of Tesco and have tried to replicate the performance through the introduction of their own loyalty card schemes.

The **loyalty card** is the device that supermarkets, pharmacists, department stores, petrol stations and even whole shopping centres and towns have developed in order to link customer characteristics to actual product purchases.

The loyalty card may be offered to customers as they make a purchase in a store. They normally complete an application form which may include their name and address, demographic details, household details, media usage and even some lifestyle characteristics. Once the customers use their loyalty cards, the products they have purchased are scanned and a link can be made through the 'swiped' card to their characteristics that can then be related to the scanned data of their product purchases. In return, the customers earn 'points' for their total spend and may earn additional points for buying particular products. The points gained may be redeemed for cash, additional purchases or even goods and services in other retailers or restaurants.

From the marketing decision-makers' perspective, many benefits accrue from a loyalty card and product scanning system. The following list summarises the benefits to the marketer:

1 *Profiles of customers can be built up.* The types of individual that are being attracted to a store can be monitored. The returns and contributions made by particular types of customer can be measured. Profiles of the 'ideal' customer type can be built up, and plans developed to attract that type of customer.

2 *Products used and not used.* The types of product that are being bought or not bought can be monitored. From the customer profile, other types of product can be added to the range offered. Cross-selling of related products can be undertaken. Linked to the customer profile, actual customer behaviour can be understood more fully.

3 *Communications that have worked and not worked.* Merchandising displays, money-off coupons, three for the price of two, or a clip-out coupon from a local newspaper, for example, can be linked to individuals and products. The effectiveness of particular types of communication for particular types of consumer can be developed. Reassurance that the customer has made the right decision can be given where the size of purchase warrants it.

4 *Distribution methods can be tailored.* Certain customer types may prefer the convenience of a small store that they visit more than once a week for small purchases of 'staple' goods. Other customer types may shop once a month for the total household. Retailers can have different shop formats for different customers, may develop home delivery programmes or even develop Internet shopping systems.

The above four factors interact to allow marketing decision-makers to redefine their market(s) and the offerings they make to those markets. The iteration of target market definition and marketing mix tailored to those markets is at the heart of strategic marketing.

From the marketing researchers' perspective, many benefits also accrue from a loyalty card and product scanning system. The following list summarises the benefits to the marketing researcher:

1 *One big laboratory.* Experimental methods will be described in Chapter 11, but, in essence, the monitoring of customers, markets and interrelated marketing mix activities allows for many causal inferences to be established. For example, what is the effect, and upon whom, of raising the price of Häagen Dazs ice cream by 10%? What is the effect of inserting a cut-out coupon to give a discount on after-sun lotion, placed in *Cosmopolitan* magazine?

2 *Refining the marketing process.* With time series of responses to planned marketing activities, statistical models of consumer response can be built with associated probabilities of a particular outcome. Likewise, models of the consumers over their lifetimes can be built. Again, statistical models can be built with associated probabilities of particular types of product being bought at different stages of a consumer's life.

3 *Developing a clear understanding of 'gaps' in the knowledge of consumers.* The scanner and loyalty card electronically observe behaviour but do not encapsulate attitudinal data. The nature and levels of satisfaction, what is perceived as good-quality service, or what brand image is associated with a particular brand of vodka are examples of attitudinal data. The use of the database helps to identify target populations to measure and the attitudinal data that need to be collected. In all there can be a much greater clarity in the nature of primary marketing research that tackles attitudinal issues.

4 *Linking behavioural and attitudinal data.* If attitudinal data are elicited from consumers, the data gathered can be analysed in their own right. It is possible, however, to link the gathered data back to the behavioural data in the database. The term 'fusing' the data from different sources is used. The key to fusing lies in identifying individual respondents so that one large dataset is built up. The notion of fusing together databases and survey data from different sources is at the heart of building a strong understanding of consumers. The analytical power that emerges from linking behavioural and attitudinal data has been realised by SPSS as illustrated in the following example.

Source: © Science Photo Library

---

**Example** | **The art of SPSS[8]**

The days of pure data processing are over at SPSS. The company has embarked on a strategy to combine the art of working with massive amounts of transactional data, which provide insight into the behaviour of consumers, with the art of working with attitudinal insights, which are typically derived from marketing research. The combination of database analytics and marketing research results in what SPSS calls predictive analytics. The software company decided to take this route when it realised that, for decades, the craft of analysing transactional data was wrongly separated from the marketing research profession. In many companies' headquarters, the departments dealing with data mining and CRM were strictly separated from the other field of providing business intelligence: marketing research. SPSS sees huge opportunities in this convergence both for its clients and for itself. It has started to develop software products that embody this convergence between transactional data analysis and marketing research. SPSS has several products that embody the convergence of database analytics and marketing research, such as Predictive Text Analytics. This application enables companies automatically to process large volumes of answers to open-ended questions, collected through surveys, and combine these with transactional data as well as analyses of a company's own call centre transcripts.

---

The benefits listed above show why many marketers and marketing researchers welcome the power of building an iterative customer database, through scanned product purchases and knowledge of customers who make those purchases. There are drawbacks, however,

that focus on the nature of the 'loyalty card'. Loyalty card schemes may be viewed more as a sales promotion technique in much the same manner as giving trading stamps, a dividend or coupons to be redeemed after a period of saving rather than as a means to capture customer data. Many loyalty schemes fail because they have been founded upon short-term goals and do not capitalise on the consumer insight generated through analysis of the data. Loyalty cards should not be seen as a short-term means to boost sales, because, compared with other sales promotion techniques, they incur huge operating costs. Proficient companies that use loyalty cards do so to identify profitable consumers and their needs and tailor the entire shopping experience to them, rather than merely deliver discounts to bargain-hungry fickle consumers.[9]

There are few questions about the huge costs involved in developing and administering loyalty card schemes. For many retailers, however, the investment allows many marketing and marketing research benefits to be realised. It is only when viewed in the light of offering strategic decision-making power and a complement to an integrated marketing information system that such an investment makes sense. One means of integrating scanned data with knowledge of customers, and of presenting the relationships and analyses in a spatial manner, is through the use of geodemographic information systems.

# Geodemographic data

**Geodemographic information system (GIS)**
At a base level, a GIS matches geographic information with demographic information. This match allows subsequent data analyses to be presented on maps.

**Thematic maps**
Maps that solve marketing problems. They combine geography with demographic information and a company's sales data or other proprietary information.

One of the main elements of database power illustrated in the preceding section is the linking of different data sources from both scanner data and customer databases. The ability to create those links and to display analyses graphically has been achieved with the development of **geodemographic information systems** (GISs). At a base level, a GIS matches geographic information with demographic information, allowing analyses to be presented on **thematic maps**. This base can be built upon with data from customer databases, databases from other sources and surveys. The combined data again can be presented on maps and in conventional statistical tables.

The geographic dimension is vital as a base to the system. Growing up and living in different geographical locations has an effect upon what we buy and the nature of our lifestyle. Look at the huge diversity of consumers around Europe! It is easy to see differences in consumers and their spending habits, between countries and regions within countries, between cities and towns and areas within a town, and even between different sides of a street. These differences emerge from a variety of factors. The following list summarises the main factors, using extreme examples in places. With closer analysis, more subtle differences can be seen which will be illustrated later in this chapter.

1 *Physical geography and climate.* Consumers living in hot Mediterranean climates in villages close to the sea may have many different needs and wants compared with consumers in Scandinavian inner cities.

2 *Economic history, working opportunities.* Consumers who are primarily semi-skilled, working in a declining manufacturing sector, may have many different needs, wants and spending priorities compared with those in a region that attracts recent graduates to work in a burgeoning financial services sector.

3 *Political and legal differences.* Locations with a history of political and legal domination can affect the types of property and subsequently the types of people who live there. The differences may be national, e.g. with policies that encourage state ownership of property, or tax breaks and discounts so that a rented property may be bought by its tenant. The differences may be regional, e.g. a local council may have structural plans to allow the building of new housing estates for families on greenfield sites on the outskirts of cities.

4 *Demographic make-up.* Regions made up of consumers living in predominantly retirement areas, such as seaside towns, will have many different requirements from regions that are heavily populated with single young people.

5 *Infrastructure links.* Infrastructure can include the means of travelling around an area as well as the nature and quality of leisure, sports and shopping facilities. Areas with different levels and quality of infrastructure attract different types of consumer. Families with two cars who can comfortably drive to facilities have different needs and wants when compared with individuals living alone who own a bicycle but not a car.

6 *Property types.* In different locations particular styles of property may dominate: flats rather than houses, multi-storey rather than low-rise, detached rather than terraced, bungalow rather than house. The type, size, quality and costs of property within an area attract different types of consumer.

Thus, differences can be seen between geographic locations that affect the lifestyle of residents, the array of products and services they buy, their ability to buy different types of products and services, and their hopes, fears and aspirations. The founding premise of a GIS is that the type of property a consumer lives in says much about the consumer's lifestyle and consumption patterns. Property type also encapsulates the other five factors that discriminate between consumers living in different geographic regions. For example, consumers living in small one-bedroom flats over shops in a city centre will tend to have very different lifestyles and consumption patterns from those living in large detached rural properties. Consumers in different property types have different propensities or probabilities of buying particular goods and services and of undertaking activities that make up their lifestyle. They also have different propensities to use and be exposed to different types of media.

From a marketing decision-making perspective, geography also plays a vital role. Knowing where one's consumers are located affects the means and costs of distribution. For example, should a retail outlet be built to gain the most returns? Which customers will have to pass our competitors in order to get to us? What features and facilities should the outlet have? The location of consumers also affects the means to communicate with them. Are consumers dispersed over a wide area or tightly clustered together? Do they read the same type of newspaper or magazine? Do they watch the same TV programmes or films at the cinema? The following example of Fred Olsen Cruise Lines illustrates how the firm used geodemographic analyses; it also reveals the links from the firm's customer database to its survey work.

**Example** **Crusing for customers**[10]

Founded in 1848, Fred Olsen Cruise Lines (**www.fredolsencruises.com**) has built a reputation for luxurious voyages in ships that are smaller and more intimate than the norm. With some extremely loyal customers, until recently it required little in the way of marketing to fill the relatively small number of cabins available during the cruising season. This all changed with the launch of *Braemar* in 2001 and increased competition from P&O, Cunard and Carnival. The company has invested in means to create greater customer insight through its customer database, website and marketing research practices. Geographic analysis has helped to pinpoint locations to target for prospective customers. The company has selected catchments around travel agents to target by direct mail. If it runs a 'cruise evening' then it does targeted mailing to publicise it. Depending upon where customers are located, possible attendees of exhibitions are also selected by location and other segmentation criteria and mailed or emailed with free ticket offers. The company's website works directly with the customer database so that segmentation codes and other preference data can be used to target returning customers with cookie-based personalised offers on the home page. In the

future, customers will be able to update their own preferences, including preferred contact channel, online. Though the company already keys in data from post-cruise questionnaires, another goal of the site is to gather extra personal attributes each time a customer visits and also to perform short surveys. Overall, the site helps to give a much deeper understanding of why customers respond to the Fred Olsen offerings.

A map therefore forms the foundation of a GIS – a map that can identify all properties in a country, all roads, shopping centres and major facilities in towns and cities. On top of a base map can be laid a range of statistical measures. They typically originate from a number of sources and have the common feature of being able to relate to a specific postcode or zip code. An example of such a system is one produced by Experian (**www.experian.com.**) It has developed systems specifically tailored for a number of countries throughout the world. The sources and details of data available in each of the above countries differ, as does the legislation that determines what can be stored and analysed on databases. Typically for each country, statistics can be gathered and used to develop individual GISs based upon census data, postal address files, electoral registers, consumer credit data, directories of company directors, mail order purchase records, car registrations and data on access to retail outlets.

From the data collected, the purpose is to classify consumers on a geodemographic basis. Experian defines a **geodemographic classification** as follows:

*Geodemographic classification groups consumers together based on the types of neighbourhood in which they live. If a set of neighbourhoods are similar across a wide range of demographic measures, they will also offer similar potential across most products, brands, services and media.*

**Geodemographic classification**

This groups consumers together based on the types of neighbourhood in which they live. If a set of neighbourhoods are similar across a wide range of demographic measures, they may also offer similar potential across most products, brands, services and media.

With the variables chosen for a particular country, i.e. the types of data that are available to build a GIS, cluster analyses are performed (Chapter 23 details the nature and purpose of cluster analysis). These analyses help to create consumer classifications, based upon the types of property the consumers live in and the propensity of consumers to have certain lifestyles and behave in particular manners. The analyses ensure that each of the descriptions used is reasonably homogeneous in terms of demographic measurements and consumer behaviour. As well as being able to discriminate and describe distinctive groups of consumers, the analyses have to produce 'pictures' of consumers that are meaningful to marketing decision-makers.

Experian has also produced 'Mosaic Global' as a consistent segmentation system that covers over 284 million of the world's households. It is based on the proposition that the world's cities share common patterns of residential segregation. Each country has its ghettos of Metropolitan Strugglers, suburbs of Career and Family, and communities of Sophisticated Singles. In terms of their values and lifestyles, each type of neighbourhood displays strong similarities in whichever country it is found. Using local data from 16 countries and statistical methods, Experian has identified 10 distinct types of residential neighbourhood, each with a distinctive set of values, motivations and consumer preferences, which can be found in each of the countries. Mosaic Global uses the data from the national Mosaic classification systems for the following countries: Australia, China (Beijing, Guangzhou, Shanghai), Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, the Netherlands, New Zealand, Norway, Spain, Sweden, the UK and the USA.

The resulting analyses have produced a classification of 10 main consumer types. Table 5.1 lists these types and the percentages of each type in the populations of Australia and Sweden. The following example describes characteristics of the Mosaic Global group labelled as *Sophisticated Singles.*

**Table 5.1** Mosaic Global classification of the global, Australian and Swedish populations

| Classification descriptor | % of global population* | % in Australian population | % in Swedish population |
| --- | --- | --- | --- |
| Sophisticated Singles | 7.90 | 10.40 | 4.34 |
| Bourgeois Prosperity | 9.20 | 22.40 | 18.61 |
| Career and Family | 8.60 | 3.50 | 7.13 |
| Comfortable Retirement | 2.90 | 2.00 | 2.54 |
| Routine Service Workers | 9.30 | 3.80 | 11.40 |
| Hard Working Blue Collar | 10.90 | 18.90 | 5.59 |
| Metropolitan Strugglers | 18.50 | 5.00 | 26.02 |
| Low Income Elders | 6.20 | 4.20 | 10.25 |
| Post Industrial Survivors | 12.20 | 16.80 | 2.89 |
| Rural Inheritance | 14.60 | 13.30 | 11.22 |

* The countries where Experian is able to access data and build its groups.

**Example**  **Sophisticated Singles**

Sophisticated Singles contains young people, mostly single and well educated, who positively enjoy the variety and stimulation afforded by life in large cities. Typically international in their outlook and with a rich network of personal contacts, they are quick to explore and adopt new social and political attitudes and are important agents of innovation, in terms of both lifestyles and the adoption of consumer products. Most are at the stage of their lives when the development of 'human' capital, i.e. skills, contacts, knowledge, continues to take precedence over the maximization of their incomes or indeed the accumulation of financial assets and much of their income is spent on 'experiences', such as entertainment, eating out, travel, books and magazines, rather than on equipment. They exhibit a variety of household arrangements and typically marry and have children late in their lives. Such people gravitate towards the smarter downtown areas of major cities where they spend short periods of time living in small, rented apartments.

With a GIS, it is possible to pinpoint where the *Sophisticated Singles* are in any of the countries analysed, whether they are clustered in particular regions or cities, or whether they are dispersed. From Table 5.1 differences between the countries in the distribution of groups are clearly seen. If one were to dig deeper within each country there would be clear differences in the proportions of these groups between cities and regions. From these classifications and the data that can be added from other databases, models of consumer behaviour can be developed. Customers can be mapped out to see how far they live from a retail outlet or to see whether they pass a competitor's store to reach a retail outlet. The profile of customers that a company has can be compared with national, regional or city profiles. Data that are captured on customer databases can be mapped out. For example, the ABN AMRO bank can map out which customers have responded to an offer to take out a personal loan at a discounted rate, as well as building up a profile of those who respond. The following example illustrates how Experian's data and systems are merged with customer databases.

**Example**    **Building long-lasting partnerships for Parfümerie Douglas**

Parfümerie Douglas is the largest subsidiary of the Douglas Lifestyle Group in Germany. It is the market leader in retail perfumery in Europe, with more than 430 stores in Germany and a further 300 in the Netherlands, Italy, Austria, France, Switzerland, Spain and Portugal. Retail perfumery is a highly competitive market with little customer brand loyalty. The challenge facing Parfümerie Douglas was to maintain its market leadership and achieve its growth ambitions by encouraging greater customer loyalty and increasing average customer spend. Parfümerie Douglas launched its loyalty card in 1995 and by 2001 had issued 1.4 million cards. Since the card's inception, Experian has partnered with Parfümerie Douglas, merging the data generated from the card with its data and systems to provide a full-service account processing function. The Douglas card has met the company's objectives with regard to customer loyalty, sales per customer, payment security and bad debt management. Douglas had its account processing system already in place when German regulations regarding discounts were liberalised in 2001. This gave it an advantage over its competitors by having the ability to test market special offers and discounts through the card's existing database.

In addition to customer behaviour being added to the GIS, survey data can also be added. The key that would link the survey to the customer database may be either a named customer or a postcode. The following example illustrates a major survey that utilises the power of Experian's Mosaic to enable more detailed analyses of target markets. Table 5.2 illustrates and provides links to a number of established and well-respected surveys that connect to and use geodemographic classifications to generate more insightful analyses of target markets.

**Table 5.2** Example of marketing research surveys that are linked to geodemographic classifications

| Marketing research survey tions | Markets covered | Linked geodemographic classifica- |
|---|---|---|
| FRS – NOP's Financial Research Survey (**www.gfknop.co.uk**) | Financial | Acorn (**www.acorn.caci.co.uk**) |
| MFS – MORI's Financial Survey (**www.mori.com**) | Financial | Mosaic (**www.experian.co.uk**), CAMEO (**www.eurodirect.co.uk**) |
| TNS – Taylor Nelson Sofres' Superpanel | Fast-moving consumer goods | Acorn (**www.acorn.caci.co.uk**) |

**Example**    **The MORI MFS Continuous Tracking Survey**

The MFS (**www.mori.com**) Continuous Tracking Survey is widely syndicated to many of the UK's leading banks, building societies, insurance companies and other organisations operating in the personal finance sector. All major markets are measured and the questionnaire is constantly updated to reflect developments in this rapidly changing sector; 2000 different adults are interviewed each fortnight in-home, face to face, by MORI's fieldforce. Quotas are set to reflect the British population by age, sex, class and region. This large annual sample size of 48,000 enables detailed analyses to be run. The types of questions asked by the MFS Omnibus cover the following topics: credit cards, personal loans, life & pensions, technology tracker, current accounts, general insurance (home & motor, pet, travel & medical), future buyer behaviour, savings & investments, loyalty/store cards. All MFS respondents are postcoded and therefore geodemographic analyses can be run against any group of respondents. This can then be compared with clients' own marketing database for any industry classification.

In this example, an insurance company can map out who bought a new policy from it. It may be able to profile and map out the types of individual who bought different types of financial services beyond insurance. The company may be interested about the levels of satisfaction related to the different financial service companies used by the respondents. The results of the survey can be analysed by the different Mosaic classifications, and levels of customer satisfaction can be mapped out. Additional purchases of insurance related to satisfaction or customer loyalty can be captured. It can be seen from the above that through the use of a GIS, profiles of target markets, measures of the success of marketing decisions and the means to model consumer behaviour can all be achieved. Graphical representations can be made of customers' behaviour, their attitudes and their levels of satisfaction. Using these data, the insurance company additionally has the potential to measure the propensity of potential customers in new locations to buy particular types of insurance policy and other financial services. Using the Mosaic Global classification, the propensity of potential customers could lead to new locations throughout the world.

## Linking different types of data

The previous example illustrated how different types of data can be merged and mapped out to represent customer characteristics. One of the main applications of collecting customer data from different sources and linking the data together would be to perform segmentation analyses. Examining the means by which target markets can be segmented, it is clear to see that the five methods as illustrated in Figure 5.1 can be individually utilised or combined to build clearer 'pictures' or profiles of target consumers.

Figure 5.1 gives examples of where data may be obtained from, to help build up profiles of customers and markets. In the example of 'psychographics' or lifestyle measurements, data may be generated from electronic point of sale (EPOS) systems or surveys. In the case of the EPOS collection, the purchasing of particular types of products can indicate characteristics of a lifestyle. In a more direct manner, questions in a survey can help to build a profile of lifestyle behaviour. In its own right, 'lifestyle' can be a valid means of segmenting a market, perhaps positioning products and services to consumers who aspire to a particular lifestyle. However, being able to combine demographic measurements, broader behavioural characteristics and a knowledge of where these consumers live helps to build a 'picture' of consumers that facilitates strong marketing decision-making support.

Figure 5.1 indicates that as one moves from the demographic through to psychological characteristics the measurement process becomes more difficult. Putting aside the differences in techniques to capture 'demography', 'behaviour' or 'psychology', what is being captured becomes more difficult as one moves towards psychological variables. If one
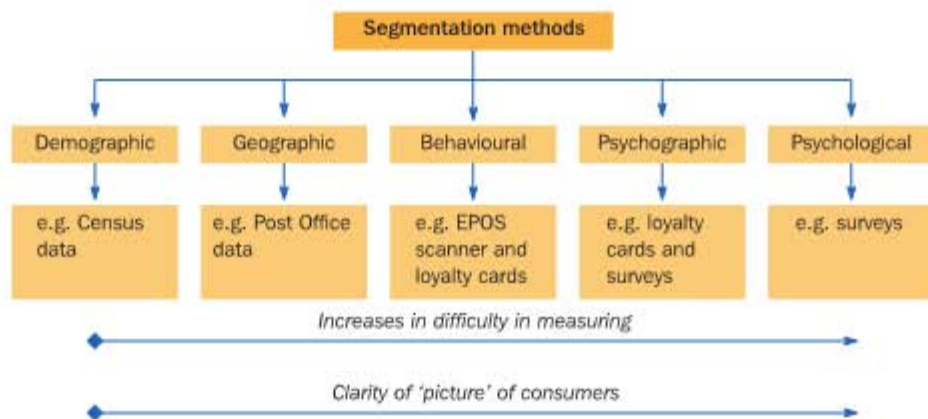


**Figure 5.1**
Methods of segmenting markets

considers psychological variables that are vital to marketing which could be captured, examples such as satisfaction, loyalty, trust and quality are not as easy to capture as questions such as gender, age or where one lives. Chapter 12 will explore the concept of measurement in more depth, but at this stage consider what 'satisfaction' actually means, and then the problems of measuring that concept in a valid and consistent manner.

Conversely, as the measurements become more difficult to conduct, they add more to the 'picture' of consumer and market profiles. To say that a market is primarily female, aged between 25 and 40 and lives in a detached property with a mortgage starts to build a 'picture' of target consumers. To add details of their media behaviour, the array of products and services they buy, characteristics of their lifestyle and their expectations helps to build up a rich and, for decision-makers, very useful 'picture' of target consumers.

Examining the variety of data sources that can be used in the interrelated variables that build market profiles, it is clear to see a role for traditional survey work, scanned data, customer data, externally generated secondary data and the use of loyalty cards. There is a clear interdependence among the different data sources with the increased sophistication of decision support systems that allow 'fusing' of the data to be conducted.[11]

## Stages of development in using databases and survey data to build profiles of consumers and model marketing decisions

The last section discussed how different data could be combined to build strong 'pictures' of consumers. Reflecting upon the role of the marketing researcher in supporting the marketing decision-maker as detailed in Chapter 1, it is clear that the combination of survey data and databases plays a major role in fulfilling the following, helping to:

- describe the nature and scope of customer groups;
- understand the nature of forces that shape the needs of customer groups and the marketer's ability to satisfy those groups;
- test individual and interactive controllable marketing variables;
- monitor and reflect upon past successes and failures in marketing decisions.

The actual implementation of the decision support systems that allow the combination of data sources to be used in supporting decision-makers can take a great deal of time, expense and organisational learning. It is not the intention here to go through the planning, training and organisational issues in making the systems work, but broadly to summarise the stages that an organisation may go through in combining survey and database data. Figure 5.2 summarises the stages of integration. The following descriptions develop the summarised stages in more detail:

1 *Analyse existing consumer database.* These data could include the daily operational transactions or enquiries made to a company. As an internal secondary data source these are the cheapest and most readily available data – providing the organisation culture allows access and analysis to marketing researchers.

2 *Use supplied geodemographic profiles.* There are a growing number of GIS vendors, some of which have been in operation for over 25 years.[12] In this time they have been able to refine the data they collect and the analyses they produce to build consumer profiles. Companies can buy a base system 'off the shelf' from systems vendors, and add a variety of different databases.

3 *Combine existing consumer data with geodemographic profiles.* Using the mapping functions of the GIS, existing customer data can be analysed using the profiles supplied with the system. Maps can be used to illustrate the catchment and types of customer and then to evaluate potential in new locations.
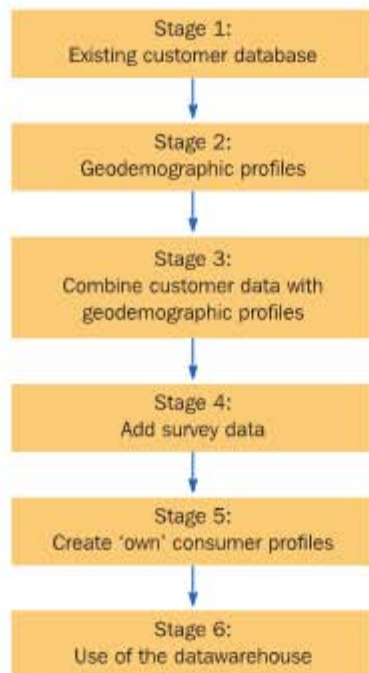
**Figure 5.2**
Stages of development in using databases and survey data to build profiles of consumers and model marketing decisions

4  *Use other surveys (either own or from external sources) that build on geodemographic sources and customer database.* Surveys conducted by a company where either a customer identification or postcode is recorded can be added. Survey data can be analysed using the geographic profiles and analyses represented using maps.

5  *Use combined data sources to create own profiles of customers.* Companies gain experience from using the geodemographic profiles and adding their own data. Over a period of time they may see that the generalised definitions of consumers from the GIS do not accurately represent their existing and target customers. With the benefit and use of their own data, they may take the raw data from the geodemographics vendor and produce their own classifications.

6  *Datawarehouse analytics.* Essentially, this is using many database sources to build one huge database that may be accessed, allowing data to be fused and analysed. The more sophisticated developments of datawarehousing would include the capture and integration of qualitative data that can emerge from intelligence and primary data sources. The analyses that emerge from datawarehouse analytics would be done to suit particular reporting requirements or specific queries from either marketing research or marketing managers. With the growth and significance of this development in decision support, the next section describes the datawarehouse in more detail.

# The datawarehouse

**Datawarehouse**
This may be seen as a 'super-database', but more specifically it may be defined as a process of gathering disparate data from database and survey sources, and converting the data into a consistent format that can aid business decision making.

One of the most prolific users and innovators in **datawarehouses** is the banking industry. The following example illustrates the problems and opportunities for the banking industry of having many different departments with quite distinctive databases.

**Example**

**You say 'warehouse', I say 'database'...[13]**

Three major changes are sweeping through bank marketing: banks are becoming even more customer driven, they are becoming increasingly information rich, and they are now dependent on constantly evolving computer technologies.

Being more customer driven results in a breaking down of the previously hermetically sealed functional areas of banking. Synergistic marketing and sales is now the name of the game. But as these invisible walls come tumbling down, banks are confronting the unintended result of departments with unique informational needs and personal computers: databases of highly valuable information that have no connection with each other. It is as if the tide has gone out, leaving tidal pools teeming with rich data separate from one another along a beach. In this case, the whole of the data really is greater than the sum of its parts. These pools of customer data are not just valuable in and of themselves. The greatest value is in the across-the-board juxtaposition of all the data pools with one another. That's where the confusing conceptual model of a datawarehouse comes in. 'Datawarehousing', then, is simply about the creation of a super-database.

The datawarehouse may be seen as a super-database, but more specifically it may be defined thus:[14]

*A datawarehouse is as much a process of gathering disparate data, converting it into a consistent format that can aid business decision making, as it is a configuration of software and hardware. Datawarehouses empower users by providing them with access to a whole array of information in an organisation, making it available for use in other applications.*

From this definition, the datawarehouse can be described as having the following three qualities:

1 It is a collection of integrated databases designed to support managerial decision making and problem solving.
2 It essentially becomes a giant database that can include survey data held in a database format.
3 It physically separates an organisation's operational data systems from its decision support systems.

At its most fundamental level, the datawarehouse has three components.

1 *Acquisition.* This includes all the programs, applications and various interfaces that extract data from existing databases. It continues with preparing the data and exporting them to the datawarehouse.
2 *Storage.* This is synonymous with any database. It simply involves a storage area to hold a vast amount of data from a variety of sources. The storage area is organised to make it easy to find and use the data. It will be updated from a variety of sources which could be through scanner and loyalty card data on customers, or through the use of intranet data as described in Chapter 4 when examining the compilation of competitor data.
3 *Access.* This encompasses both set reporting of predetermined events and the means of performing individual analyses, querying 'what-if' scenarios. The process of exploring the databases uses **data mining** techniques. As marketing researchers and decision-makers learn about markets and their effects upon those markets, the development of predictive models is facilitated. Predictive models are built and tested using historical customer and transactional data, and then used to predict how customers might respond in any number of future situations. As well as the classic marketing application of who may respond to a particular offer, they help to evaluate how likely different individuals or groups are to end their relationship with a company, default on a loan or

**Data mining**
The process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition as well as statistical and mathematical techniques.

buy an extra product. The use of predictive modelling started out in credit risk: who is most likely to pay up rather than moving out and disappearing? Analysts found that by appropriately weighting different variables such as income they could predict with some accuracy the degree of risk that customers presented. Awareness of the power of modelling outside credit and direct marketing rose dramatically with the use of techniques like collaborative filtering by online vendors such as Amazon. If you buy certain books, they you may well be interested in other books bought by customers with a similar history to yours.[15]

## Data mining

Data mining is a process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories. The process uses pattern recognition as well as statistical and mathematical techniques.[16] Data mining should not be confused with datawarehousing. The datawarehouse could be termed a 'repository' or a place where large amounts of sometimes disparate sources of data are stored; data mining is a process that depends upon access to the data held in that repository.

Examples of what data mining aims to do are as follows:

- Classify customers into specific categories that are meaningful to decision-makers.
- Identify potential target markets that possess the characteristics that decision-makers seek.
- Forecast sales or the use of services.
- Discover which types of products or services are purchased together.
- Discover patterns and trends over time, such as 'after graduation, students take a holiday', and be able to show the probabilities associated with them buying different types of holiday.

Data mining is a way of exploiting the data held by organisations to help discover and develop specific information or knowledge. As well as using proprietary software to perform pattern recognition, statistical and mathematical techniques, data mining can also be seen as a mental process undertaken by decision-makers. The decision-makers who interact with large datasets using data mining are generally not specialists in statistics, data analysis, datawarehousing and other data tools; they are information users, seeking support for their decision-making processes. With the aid of data mining software, the decision-maker is encouraged to think in new ways and ask new questions. The decision-maker discovers more from the data and explores in new areas, integrating other sources of data, going through iterative processes to dig deeper. The exploration process involves the discovery of non-trivial relationships of dependence or associations, non-trivial clusters, factors or trends and an understanding of the managerial significance of these discoveries. A data mining process must be 'user oriented' and that user is typically the decision-maker. The following example illustrate the processes described above by showing how data mining has been used to explore the customer database of one of Europe's leading banks.

**Example** **Credit Suisse**

As one of the world's leading financial services companies, Credit Suisse Group provides banking and insurance solutions for private clients, companies and institutions. Based in Zurich, Switzerland, Credit Suisse employs 80,000 people worldwide. Competition in the financial services industry is intense, and obtaining new customers is expensive. In order to maximise profitability, Credit Suisse focused on three areas. First, identifying profitable cur-

rent customers. Second, managing customer relationships to ensure longevity. Third, retaining profitable customer accounts. In 1997, Credit Suisse started the 'Loyalty Based Management' programme, with the primary goal of retaining profitable customers. It invested in a six-member 'data mining team' that used the Clementine tool (**www.spss.com/clementine**) to analyse a data warehouse of its 2.5 million customers, each with more than 400 attributes. The analysis was used to identify potential leads among Credit Suisse's customers and intelligently market to them based on their individual preferences and histories. With the help of SPSS, Credit Suisse's data mining activities, analysis and modelling have been fully integrated into its business processes and have proven their value in many different applications. As a result of the success of the Loyalty Based Management project, Credit Suisse consultants began to see data mining's benefits and started to use it to sell specific customers targeted services. Credit Suisse can now identify customers, typically the top 1%, who are extremely likely to buy a service, thus increasing the opportunities for cross-selling and retaining customers. Detailed segmentation of its vast customer base allows Credit Suisse to develop targeted solutions for its customers. This segmentation is executed inductively using the cluster algorithm and the dimensions are tailored directly to the customer requirements. Each cluster serves as a starting point for individual marketing campaigns. This hierarchical system is advantageous because the customer database is continually researched and monitored. As a result, changes in the cluster structure are quickly identified and appropriate responses are triggered.

## Databases and marketing research

There has been a phenomenal growth in the use of databases to support marketing decision making. In larger organisations with many divisions or where mergers and acquisitions have taken place, the datawarehouse has facilitated 'fusing' of data from many sources. Such developments are seen as a threat by many in the marketing research industry. However, many marketing research companies and marketing research departments within companies are embracing database techniques, utilising the synergistic benefits of matching database analyses with traditional survey data through data mining. To illustrate this point, consider the following quote from Greg Ward, Development Director for Taylor Nelson (**www.tns-global.com**), the largest marketing research company in the UK:

*The marketing research industry needs to acknowledge that databases are serious products and that both types of information have benefits. If you take the best of both – what we call information based marketing – you get something that is significantly more powerful. The 'them and us' situation does nobody any favours and the idea that the two disciplines bear no resemblance to each other is wrong.*

As marketing decision-makers become more willing and able to interrogate databases and to generate creatively their own decision support, this does not mean the end of 'traditional' marketing research. As illustrated earlier when examining types of data that are used to build consumer and market profiles, psychological data play a vital role that is fulfilled by qualitative and quantitative marketing research. The marketing researcher needs to develop a greater awareness of both how data captured through traditional methods can be integrated with data held on databases *and* how the combined data creatively support decision-makers.

Databases, the development of datawarehouses and the use of data mining techniques are allowing a wider and shared use of data. The graphical formats of presenting data, especially using maps, break down many barriers in decision-makers who resist formal statistical analyses. They encourage managers to tailor output to meet their individual

needs. The creativity that is the hallmark of marketing decision making is supported by the creative collection and connections between data. Where there are gaps in decision-makers' knowledge, they can be more focused and precise in determining what marketing research support they need. Many marketing researchers are rising to meet this challenge.[17]

The next example illustrates how data mining can provide in-depth analyses of survey data at great speed or it can help combine research data with other types of data to provide a much richer source of insight. It illustrates how companies can integrate databases and traditional marketing research, seeing the disciplines as complementary, not competitive.

---

**Example**    **TrustMark CFI**

TrustMark CFI, part of the worldwide CFI Group, is a market research and management consulting business in Zollikon, Switzerland. In a project for a banking customer, TrustMark CFI consultants used Clementine (**www.spss.com**/**clementine**) to analyse customer satisfaction data with operational data, and then build models that would forecast future consumer activity. By combining the two data sources, the models produced by Clementine were more accurate – and comparing the predictions from the Clementine model with reality proved this. When using operational data only, Clementine correctly identified 65.4% of the customers who had increased their asset volume significantly, and 56.9% of those who had greatly reduced it. However, using the two data sources together, the model was correct in 67.9% and 60.7% of cases respectively. According to Leonie van de Vijfeijken, a research analyst at TrustMark CFI:

> the results would have been even better if there hadn't been a large time lag between gathering the operational and survey data. Although the difference in the figures does not sound remarkable, it is significant to the bank's business and helps reduce its marketing costs.

Using these models on current data allows the bank to predict whether an account is likely to be profitable, at risk, or inactive. This information enables it to target marketing resources more effectively and, in the final analysis, improve financial performance. Interestingly, had the bank not used predictive models and simply run standard analyses on the survey research data, a very different picture would have emerged. When past research data was compared with reality, it was shown that only 21% of the respondents who said they were going to increase their asset volume actually did so.

---

There is a blurring of the lines between traditional marketing research and other corporate intelligence activities, including data mining. Whilst some research departments are reinforcing the line that separates the marketing research department from other functions, many are embracing the change. The reason that this integration matters is that it should be the case that the marketing research department is in the best position to drive the understanding of the 'why' of behaviour, rather than just describing the behaviour.[18] Given the different but complementary roles that marketing research and database analytics perform, put together there can be a great synergistic effect. If a company has a customer transactional database, it is very unlikely that it can tell the company why its customers use the products they do and what media they consume. What marketing research can do is overlay those attitudes, identify segments and make the whole database actionable.[19]

# International marketing research

Linking databases and survey data is transforming international marketing research. Within individual companies, customers may be analysed from operational data within a country, showing different patterns of behaviour between different regions or cities and relating that behaviour to their marketing activities. When a company operates across borders, country differences become just another geographical variable.

In deciding to operate or develop in a particular country, companies may buy a geodemographic information system (should one be developed for that country). A GIS may be used as a foundation to add to their operational data. From this point they may go through the stages of database development as laid out in Figure 5.2. However, international marketers face a dilemma in choosing between 'consistency' and 'richness of detail'.

It is possible to create systems GISs to individual countries across the globe based upon the data sources available within that country, its data protection laws and how current and accurate those data sources are. Such systems work extremely well if one's marketing activities are based within that country. The problem arises when, as an international marketer, one wishes to have consumer classifications that are consistent across borders. Mosaic Global is one of many solutions to this problem. However, in resolving the problem of consistency across countries, an amount of the richness of individual country detail may be lost. To illustrate this dilemma, examine the following summaries of Mosaic Global, UK and London. See how, as the system becomes more focused geographically, there is a greater richness of detail but a loss of consistency in the measurements and classification across borders

**Example** | **Mosaic classifications in London, the UK and across the globe**

Mosaic Global is a consistent segmentation system that covers over 400 million of the world's households. Using local data from 25 countries, Experian has identified 10 types of residential neighbourhood that can be found in each of the countries, each with a distinctive set of values, motivations and consumer preferences. Mosaic Global is based on a simple proposition that the world's cities share common patterns of residential segregation. In terms of their values and lifestyles, each type of neighbourhood displays strong similarities in whichever country it is found.

Mosaic UK is the latest version of the market-leading consumer segmentation product. It classifies all 24 million British households into 11 groups, 61 types and 243 segments, and is updated each year. A development team of over 30 staff took over two years to build Mosaic UK. The result is a classification that paints a rich picture of British consumers in terms of their socio-demographics, lifestyles, culture and behaviour, providing the most accurate and comprehensive view of British society at the start of the twenty-first century.

Mosaic London is a geodemographic classification of London. The classification covers 5.9 million households (25% of the UK) in approximately 380,000 postcodes. Mosaic London describes Londoners in depth: each postcode has been allocated to one of 41 types and 12 groups that are specific to London. Groups of people are identified that are important to the London community but may not occur elsewhere in the country.

Performing analyses within countries can be most fruitful, provided a base GIS has been established. Problems start where there is no base GIS. In many countries there are great problems in tracking down and combining data sources that can be relied upon. Further,

even if reliable data can be located, legislation may make the use of certain data types illegal. In many developing countries, the data needed to build a GIS are sparse. With the data that are available, much experimentation is needed to enable valid classifications that reflect consumer types which are useful to marketers and marketing researchers.

## Ethics in marketing research

Marketing researchers are confronted by problems posed by the wording of ESOMAR and individual country marketing research associations' codes of conduct.[20] The codes specify that the compilation of lists, registers or databanks of names and addresses for any non-research purpose shall in no way be associated directly or indirectly with marketing research. The view is that whilst marketing research and marketing database analytics are complementary, because the two activities are subject to different legislative requirements and in the interests of transparency for respondents, the ICC/ESOMAR International Code states:

> When acting in their capacity as researchers, the latter must not undertake any non-research activities, for example database marketing involving data about individuals which will be used for direct marketing and promotional activities. Any such non-research activities must always, in the way they are organized and carried out, be clearly differentiated from marketing research activities.
>
> For more details see www.esomar.org/codesandguidelines.

However, the examples detailed in this chapter show that supporting marketing decision-makers through databases and marketing research can be seen as part of a total information industry. Evidence of the many leading marketing research agencies involved in data collection and analyses through databases illustrates that databases need not be unethical. With due care it is possible to combine marketing research ethics and databases generated through database marketing. There are a growing number of companies that have used marketing research for many years that now combine the traditional role of marketing research manager with a wider role including database management. An essential part of this combined role lies in the management of customer databases, adding survey details to respondents' individual details, at either an individual or an aggregated level.

Given the phenomenal growth of databases in marketing and the support they offer to marketing decision makers, they are here to stay. With well-planned 'traditional' marketing research integrated into database analyses, the strategic power of consumer and market analyses is phenomenal. If marketers abuse their knowledge of consumers, they stand to do great harm to their brands and corporate image. For example, in bank databases there are many opportunities for the cross-selling of products. Rather than welcoming the approach from another division of a bank trying to sell insurance to an investment client, there can be a reaction against the approach, affecting the original business. Consumers are now more aware of how valuable knowledge of their behaviour is and how it is used by marketers. They are willing to trade this knowledge for the kind of rewards that are gained from the use of their loyalty cards. Marketers are aware of the dangers of abusing the knowledge that their customers impart to them. However, there are issues of civil liberties that cannot be ignored. These are touched upon in the following example.

## Loyalty for sale[21]

Provided that shoppers like the benefits and do not object to a system which records every bar of chocolate and bottle of gin purchased, no great harm will be done. However, there is a danger that, despite the safeguards of Data Protection Acts, this mass of information on consumers' habits could leak across the networks into unscrupulous hands. Issues of civil liberty would be raised if, for example, insurance companies could use the data to identify people whose purchases indicated an unhealthy lifestyle; or if the police could draw up a list of suspects by monitoring the purchase of specific items or unusual consumption patterns.

One of the benefits of the use of GISs is that in many cases the individual does not have to be identified; the postcode is a sufficient key to make a link between databases. This maintains the marketing research industry's maxim of respondent confidentiality.

## Internet and computer applications

At the start of the 1990s, this chapter would not have existed in a marketing research text. The idea of conducting internal secondary data searches and analyses would have merited a paragraph or two as part of the process of developing primary data collection. Since then, the massive technological changes that have made the global use of the Internet commonplace, the increased storage space, speed and analysis capabilities of computers, and the increased sophistication of software collectively have made fundamental changes to the environment in which marketing researchers operate. The collection and analysis of customer data held within businesses have developed enormously and marketing researchers cannot ignore these developments.

In Chapter 1 we discussed how, for many years, marketing researchers have recognised the competition they face from an array of management consultants, but more recently competition has emerged from raw data providers such as call centres, direct marketing, database marketing and telebusinesses.[22] Much of the new competition has emerged from organisations that have utilised and developed Internet and computer applications and have been able to offer support to decision-makers, in faster, cheaper and more user-friendly formats, though not necessarily more rigorously or ethically. Leading marketing researchers have developed means of integrating many of the new formats for supporting decision-makers with traditional marketing research methods and have maintained their customary rigour and ethical behaviour. In order to get a feel of how decision-makers may be supported by some of the means discussed in this chapter, we recommend that you explore the following websites. These websites contain case studies of the applications of decision support. When working through these cases, consider what 'gaps' may still exist in the knowledge of decision-makers that may be filled only by the use of traditional marketing research techniques. The following example illustrates such a case. It shows how SPSS is using the integration of different data sources to build consumer behaviour models and predict behaviour. The SPSS concept of predictive analytics was presented earlier in the chapter; this example illustrates its application in combining data mining and marketing research (**www.spss.com**).

**Example**   **Top Dutch insurer Interpolis selects SPSS software for lead generation**

Interpolis (www.interpolis.com) is the largest single-brand consumer property and casualty insurance company in the Netherlands. The company is also active in Luxembourg and Ireland. It has about 6,000 employees, and realised a turnover of €5 billion in 2004. Interpolis has purchased SPSS's predictive analytics software to grow its business further through lead generation. Harnessing SPSS's PredictiveMarketing, Interpolis wants to predict which customers are most likely to buy its insurance products. With SPSS, Interpolis aims to increase sales to its existing customers and broaden its customer base, which can amount to millions of euros in additional revenue. Interpolis generates 85% of its revenues from Rabobank's branch network. The insurer selected SPSS's analytics software better to understand and predict customer needs and feed this information to branch staff, providing a clear indication of opportunities to sell its insurance products. Interpolis's marketing department will also utilise SPSS software to identify which customers are most likely to defect once an insurance contract has ended, the reasons for defection and the best actions to minimise it. 'Interpolis is a great example of a predictive enterprise – leveraging and acting upon customer knowledge across multiple channels: at its marketing campaigns, in branch offices, and in the call center,' added Marcel Holsheimer, Vice President of Marketing for SPSS's Platform and Applications, and, 'Deploying predictive analytics gives it the ability to improve across several business processes and considerably enhance bottom line performance.'

Two major suppliers of software that help with customer relationship management mentioned at the start of this chapter are SAP (www.netsuite.com) and AIT (www.ait.co.uk). They also show how operational data generated in organisations may be integrated. The main GISs to examine individual countries and across boundaries were listed earlier in the chapter, but as a reminder, they were: Acorn (www.acorn.caci.co.uk), Mosaic (www.experian.com) and CAMEO (www.eurodirect.co.uk). Numerous data mining software programs are now available. Kdnuggets (www.Kdnuggets.com) is an online resource for data miners which has identified the most popular data mining software packages. The most popular vendors and their datamining software programs include the following:

SPSS (www.spss.com/clementine),
SAS (www.sas.com/technologies/analytics)
IBM (www-306.ibm.com/software/data/db2bi)
Angoss (www.angoss.com)
Megaputer (www.megaputer.com).

Each of these programs can run on client–server platforms and includes a wide range of analysis tools and techniques, including neural networks and decision trees.[23] A French datawarehouse supplier with a case based on the Renault Company can be found at www.decisionnel.fr. Finally, visit www.crisp-dm.org to see an organisation devoted to the development of data mining methodology.

## Summary

The overall tone of this chapter has been to demonstrate that internally generated secondary data offer great opportunities not only for decision-makers but also for researchers. As with all good secondary data sources, they have a major impact upon the conduct and direction of primary data collection, analyses and interpretation.

Databases, including customer operational data, geodemographic data and survey data, are radically changing how marketing decision making is being supported. There is much debate as to whether the use of databases is compatible with traditional techniques of marketing research. With the junk mail connotations of databases and compromises of respondent anonymity, many marketing researchers may seek to keep the marketing database at arm's length. However, handled with the professional acumen that marketing researchers have displayed for many years, the database presents great opportunities for the marketing researcher. In Europe, many of the leading marketing research agencies and research functions within companies have embraced the marketing database as an essential component of their desire to create 'customer insight'.

For the marketer, databases help to build profiles of consumers, linked to the products, communications and distribution methods those consumers favour. For the marketing researcher, databases can present the opportunity to experiment in 'one big laboratory', build models of consumer behaviour, develop an understanding of the gaps in knowledge of consumers and make links between behavioural and attitudinal data.

Much of the data that offer these benefits have been gained using data that capture customer buying behaviour. The use of the 'loyalty card' is one example. Different types of data, including scanner data, loyalty card data and survey data, may be combined using geodemographic information systems (GISs). Using base geographic and demographic data, existing customers can be analysed and mapped out.

Disparate database sources are pulled together through the use of datawarehouses. The datawarehouse integrates databases and survey data, allowing creative connections between data to be explored. Integrated databases and survey data can be explored using data mining techniques. These processes involve the use of proprietary software and inquisitive, creative decision-makers who search for trends and patterns in customer behaviour and attitudes. The development of datawarehouse and data mining expertise especially helps to cope with the problems of disparate databases and survey data from different countries.

The ethics of using databases provokes much debate in the marketing research industry. As many research practitioners grow more accustomed to using databases, marketing research guidelines and codes of practice are being developed to reflect the good practices that exist in many companies.

## Questions

1 How may 'operational data' held in organisations help to build up an understanding of customer behaviour?

2 What is a customer database? Why may a marketing researcher wish to analyse the data held in a customer database?

3 What kinds of data can be gathered through electronic scanner devices?

4 What other sources beyond electronic scanner devices electronically observe customer behaviour?

5 Describe the benefits to the *marketing decision-maker* of being able to capture data that identify characteristics of consumers and their shopping behaviour in a store.

6 Describe the benefits to the *marketing researcher* of being able to capture data that identify characteristics of consumers and their shopping behaviour in a store.

7 Why may the characteristics of consumers differ, based upon where they live?

8   What is a geodemographic classification of consumers?

9   How can the graphical representation of consumer characteristics using maps help marketing decision making?

10  What benefits may be gained from fusing together customer characteristics held as internal secondary data, with a proprietary geodemographic information system held as external secondary data?

11  How does the compilation of different types of data help to build a strong 'picture' of consumer characteristics?

12  Describe the stages of development in using databases and survey data to build profiles of consumers and model marketing decisions.

13  What is a datawarehouse?

14  What is the difference between a datawarehouse and data mining?

15  Why may there be a difference between a individual country geodemographic classification and a European or global classification of that same country?

# Exercises

1   Visit the websites of Acorn (www.acorn.caci.co.uk), Mosaic (www.experian.com) and CAMEO (www.eurodirect.co.uk). Imagine that you have been commissioned to select a geodemographic system to help a newspaper publisher in a major European city.

    a For such a business, how may a geodemographic system be used for marketing decision making?
    b For such a business, how may a geodemographic system aid marketing research design?
    c Present the case of which of the above systems would best suit such a business.

2   Call in at a supermarket or store that operates a reward or loyalty card scheme that requires you to apply for membership. Pick up an application form and examine the nature of questions you are expected to answer.

    a What marketing research use can be made of the data collected from this application form?
    b Evaluate the design of this form and make recommendations on how the nature of questions could be improved.

3   You are a marketing manager for Carlsberg beer. One of your major customers is a supermarket that uses a loyalty card scheme to electronically observe its customers.

    a What would this supermarket know about beer buying behaviour through its scheme?
    b If it would not share this with you, evaluate any marketing research techniques that you think could generate the same knowledge.

4   Visit the SPSS website www.spss.com and follow a link to its 'predictive analytics' products. Write a report on how marketing research may feed into and/or feed from predictive analytics in either a bank or major retailer.

5   In a small group discuss the following issues: 'What ethical problems exist with the use of marketing databases for marketing researchers?' and 'If, on ethical grounds, marketing researchers refused to utilise the benefits of marketing databases, what inherent weaknesses may exist in their research designs?'

## Video Case Exercise: Subaru

How could Subaru's survey analyses and internal and external secondary data collection help them to profile and target female drivers?

video case

download from
www.pearsoned.co.uk/
malhotra_euro

## Notes

1  McElhatton, N., 'Customer insight in stereo', *Research World* (December 2004), 21.

2  Ibid., p. 20.

3  Evans, M., Nancarrow, C., Tapp, A. and Stone, M., 'Future marketers: future curriculum; future shock?', *Journal of Marketing Management* 18 (5–6), (2002), 579–596.

4  Dawe, A., 'Integration is the thing', *The Times – Special Edition e–CRM* (11 April 2001), 2–3.

5  Sumner–Smith, D., 'Getting to know clients lifts profits', *The Sunday Times* (26 September 1999), 17.

6  Fletcher, K., 'The evolution and use of information technology in marketing', in Baker, M.J. (ed.), *The Marketing Book*, 3rd edn (Oxford: Butterworth–Heinemann, 1994), 352.

7  Mistry, B., 'A question of loyalty', *Marketing* (2 March 2005), 42.

8  Havermans, J., 'Do-it-yourself surveys offer opportunity for MR', *Research World* (June 2004), 20–21.

9  Kilby, N. and Bedwell, R., 'Women stand by their stores', *Marketing Week* (2 June 2005), 32–33.

10  Lawson, J., 'Cruising for customers', *Database Marketing* (December 2003), 29–30.

11  An example of fusing research data with databases is presented in Leventhal, B., 'An approach to fusing market research with database marketing', *Journal of the Market Research Society* 39 (4) (October 1997), 545–561.

12  Sleight, P., 'Newest neighbourhoods', *Database Marketing* (June 2005), 31–34.

13  Man, D., 'You say "warehouse", I say "database" …', *Bank Marketing* 29 (4) (April 1997), 37.

14  Man, D., 'The faqs on datawarehousing', *Bank Marketing* 29 (4) (April 1997), 38.

15  Lawson, J., 'No rocket scientists required', *Database Marketing* (June 2003), 30.

16  Jambu, M., 'Data mining for better knowledge: a case study for the telecommunications industry', ESOMAR, Power of Knowledge Conference, Berlin (September 1998).

17  McElhatton, N., 'Raiding the data bank', *Research* (September 1999), 28–30. See also the arguments in Brand, C. and Jarvis, S., 'Mind games – the new psychology of research', Market Research Society Conference, 2000.

18  Pinnell, J., 'Customer relationships – manage, measure or just understand?' ESOMAR (January 2003), **www.warc.com/print/78475p.asp**.

19  McElhatton, N., 'Customer insight in stereo', *Research World* (December 2004), 21.

20  Hodgson, P., 'Databases: the time for decisions is nigh', *Research* (October 1993), 19.

21  'Loyalty for sale', *Financial Times* (18 September 1998).

22  Smith, D.V.L. 'Bridging the gap between data and decision', *Research World*, (November 2004), 20–21; Savage, M., 'Downstream danger', *Research* (May 2000), 25–27.

23  For illustrations of neural networks and decision trees in marketing, see Garver, M.S., 'Using data mining for customer satisfaction research', *Marketing Research* (Spring 2002), 8–12.