CHAPTER 4

Secondary data collection and analysis

Objectives

After reading this chapter, you should be able to:

1. define the nature and scope of secondary data and distinguish secondary data from primary data;
2. analyse the advantages and disadvantages of secondary data and their uses in the various steps of the marketing research process;
3. evaluate secondary data using the criteria of specifications, error, currency, objectives, nature and dependability;
4. describe in detail the different sources of secondary data, focusing upon external sources in the form of published materials, and syndicated services;
5. discuss in detail the syndicated sources of secondary data, including household and consumer data obtained via surveys, mail diary panels and electronic scanner services, as well as institutional data related to retailers, wholesalers and industrial or service firms;
6. explain the need to use multiple sources of secondary data and describe single-source data;
7. identify and evaluate the sources of secondary data useful in international marketing research;
8. understand the ethical issues involved in the use of secondary data.

The act of sourcing, evaluating and analysing secondary data can realise great insights for decision-makers. It is also vital to successful problem diagnosis, sample planning and collection of primary data.

Overview

The collection and analysis of secondary data help to define the marketing research problem and develop an approach. In addition, before collecting primary data, the researcher should locate and analyse relevant secondary data. Thus, secondary data can be an essential component of a successful research design. Secondary data can help in sample designs and in the details of primary research methods. In some projects, research may be largely confined to the analysis of secondary data because some routine problems may be addressed based only on secondary data. In addition, given the huge explosion of secondary data sources available, sufficient data may be accessed to solve a particular marketing research problem.

This chapter discusses the distinction between primary data, secondary data and marketing intelligence. The advantages and disadvantages of secondary data are considered, and criteria for evaluating secondary data are presented, along with a classification of secondary data. Internal secondary data are described, and major sources of external secondary data – such as published materials, online and offline databases, and syndicated services – are also discussed The sources of secondary data useful in international marketing research are discussed. Several ethical issues that arise in the use of secondary data are identified.

To begin with, we present an example that illustrates the nature of secondary data, how it may be evaluated, and its relationship to primary data collection.

example

Flying high on secondary data[1]

Money magazine recently published the results of a study conducted to uncover the airline characteristics that consumers consider most important. In order of importance, these characteristics were safety, price, baggage handling, on-time performance, customer service, ease of reservations and ticketing, comfort, frequent flyer schemes and food.

If Air France was considering conducting a marketing research study to identify characteristics of its service that should be improved, this article might be a useful source of secondary data. Before using the data, Air France should evaluate them according to several criteria.

First, the research design used to collect the data should be examined. This *Money* magazine article includes a section that details the research design used in the study. *Money* used a face-to-face survey of 1,017 'frequent flyers'. The results of the survey had a margin of error of 3%. Air France would have to decide whether 'frequent flyers' in the USA could be generalised to the population they wish to understand, whether 1,017 was a sufficient sample size for their purposes and whether a margin of error of 3% was acceptable. In addition, Air France should evaluate what type of response or non-response errors may have occurred in the data collection or analysis process.

The currency of the data and objective of the study would be important to Air France in deciding whether to utilise this article as a source of secondary data. Air France would also need to look at the nature and dependability of the data. For example, they would need to examine how the nine choice criteria were defined. If the criterion price was measured in terms of fare per kilometre, is this a meaningful and acceptable definition to decision-makers at Air France? With regard to dependability, Air France would need to evaluate the reputation of *Money* magazine and of ICR, the research company hired by *Money* to undertake the survey. They would also need to recognise the fact that *Money* used secondary data in its study; how dependable are the sources they used?

The *Money* magazine article might be useful as a starting place for a marketing research project for Air France. It could be helpful in formulating the nature of decision-making problems and associated research objectives. There may be limitations in regard to reliability, dependability or even how generalisable it may be to Air France's target consumers. Many lessons and ideas may be generated from this article that may lead to other secondary data sources and in the design of a well-focused primary data collection. ■

Defining primary data, secondary data and marketing intelligence

Primary data
Data originated by the researcher specifically to address the research problem.

Secondary data
Data collected for some purpose other than the problem at hand.

Marketing intelligence
Qualified observations of events and developments in the marketing environment.

Primary data are originated by a researcher for the specific purpose of addressing the problem at hand. They are individually tailored for the decision-makers of organisations that pay for well-focused and exclusive support. Compared with readily available data from a variety of sources, this tailoring means higher costs and a longer time frame in collecting and analysing the data.

Secondary data are data that have already been collected for purposes other than the problem at hand. At face value this definition seems straightforward, especially when contrasted to the definition of primary data. However, many researchers confuse the term, or quite rightly see some overlap with marketing intelligence.

Marketing intelligence can be defined as 'qualified observations of events and developments in the marketing environment'. The use of the word 'observations' is presented in a wide sense to include a variety of types of data, broadly concerned with 'environmental scanning'.[2] In essence, though, marketing intelligence is based upon data that in many instances have been collected for purposes other than the problem at hand. To clarify this overlap in definitions, Table 4.1 compares secondary data with marketing intelligence through a variety of characteristics.

Table 4.1 A comparison of secondary data and marketing intelligence[3]

Characteristic	*Secondary data*	*Marketing intelligence*
Structure	Specifications and research design tend to be apparent	Can be poorly structured; no universal conventions of reporting
Availability	Tend to have regular updates	Irregular availability
Sources	Generated in-house and from organisations with research prowess	Generated in-house and from unofficial sources
Data type	Tend to be quantitative; many issues need qualitative interpretation	Tends to be qualitative; many issues difficult to quantify
Source credibility	Tend to be from reputable and trustworthy research sources	Questionable credibility; can be generated from a broad spectrum of credibility
Terms of reference	Tend to have clear definitions of what is being measured	Ambiguous definitions; difficult to compare over different studies
Analysis	Mostly conventional quantitative techniques	Opinion based, interpretative
Ethics	In-company data gathering may be covered by data protection acts; externally generated data may be covered by research codes of conduct, e.g. ESOMAR	Some techniques may be seen as industrial espionage – though there is an ethical code produced by the Society of Competitive Intelligence Professionals

Note in the above comparisons the repeated use of the word 'tend'. The boundaries between the two are not absolutely rigid. Consider the example at the start of this chapter, an article published in *Money* magazine. The journalist may have collected, analysed and presented quantitative data to support their qualitative interpretation of the future developments of a market. The data they use and present may come from credible sources and be correctly analysed, but what about their choice of data to support their argument? Other sources of data that may contradict their view may be ignored. The data they present can be seen as a secondary data source and interpreted in its own right by a researcher. The interpretation and argument of the journalist can

be seen as intelligence and have some credibility. In its entirety, such an article has elements of both secondary data and marketing intelligence, and it may be impossible to pull them apart as mutually exclusive components.

As will become apparent in this chapter, there are clear criteria for evaluating the accuracy of secondary data, which tend to be of a quantitative nature. Marketing intelligence is more difficult to evaluate but this does not mean that it has less value to decision-makers or researchers. Certain marketing phenomena cannot be formally measured; researchers may not be able to gain access to conduct research, or the rapid unfolding of events means that it is impracticable to conduct research. The following example illustrates the importance of intelligence to many companies.

example

Behind enemy lines[4]

Robin Kirkby, Director of European Consulting for intelligence specialist Fuld & Company, says there are three principal factors driving investment in intelligence.

'The Internet, globalisation and higher expectations from customers are all putting companies under more pressure to differentiate themselves from the competition. It's frustrating that intelligence gets associated with spying; it's actually a highly ethical activity, focused on underlying competitive dynamics and planning future change.'

According to research by The Futures Group (TFG), 60% of companies have an organised system for collecting competitive intelligence, while 82% of companies with revenues over €10bn make systematic use of it. TFG ranked the leading eight users of competitor intelligence as:

1 Microsoft
2 Motorola
3 IBM
4 Procter & Gamble
5= General Electric
5= Hewlett-Packard
7= Coca-Cola
7= Intel ■

Such widespread use of intelligence in major organisations means it has some role to play in supporting decision-makers, but it has many limitations, which are apparent in Table 4.1. In the development of better-founded information support, credible support can come from the creative collection and evaluation of secondary data. This requires researchers to connect and validate different data sources, ultimately leading to decision-maker support in its own right and support of more focused primary data collection. As this chapter and Chapter 5 unfold, examples of different types of secondary data will emerge and the applications of secondary data will become apparent.

Advantages and uses of secondary data

Secondary data offer several advantages over primary data. Secondary data are easily accessible, relatively inexpensive and quickly obtained. Some secondary data, such as those provided by the National Censuses, are available on topics where it would not be feasible for a firm to collect primary data. Although it is rare for secondary data to provide all the answers to a non-routine research problem, such data can be useful in a variety of ways.[5] Secondary data can help you:

1 Diagnose the research problem
2 Develop an approach to the problem
3 Develop a sampling plan
4 Formulate an appropriate research design (for example, by identifying the key variables to measure or understand)

5 Answer certain research questions and test some hypotheses
6 Interpret primary data with more insight
7 Validate qualitative research findings.

Given these advantages and uses of secondary data, we state the following general rule:

> *Examination of available secondary data is a prerequisite to the collection of primary data. Start with secondary data. Proceed to primary data only when the secondary data sources have been exhausted or yield marginal returns.*

The rich dividends obtained by following this rule are illustrated in the example at the start of this chapter. It shows that the collection and analysis of even one relevant secondary data source can provide valuable insights. The decision-maker and researcher can use the ideas generated in secondary data as a very strong foundation to primary data design and collection. However, the researcher should be cautious in using secondary data, because they have some limitations and disadvantages.

Disadvantages of secondary data

Because secondary data have been collected for purposes other than the problem at hand, their usefulness to the current problem may be limited in several important ways, including relevance and accuracy. The objectives, nature and methods used to collect the secondary data may not be appropriate to the present situation. Also, secondary data may be lacking in accuracy or may not be completely current or dependable. Before using secondary data, it is important to evaluate them according to a series of factors.[6] These factors are discussed in more detail in the following section.

Criteria for evaluating secondary data

The quality of secondary data should be routinely evaluated, using the criteria presented in Table 4.2 and discussion in the following sections.[7]

Specifications and research design

The specifications or the research design used to collect the data should be critically examined to identify possible sources of bias. Such design considerations include size and nature of the sample, response rate and quality, questionnaire design and administration, procedures used for fieldwork, and data analysis and reporting procedures. These checks provide information on the reliability and validity (these concepts will be further developed in Chapter 13) of the data and help determine whether they can be generalised to the problem at hand. The reliability and validity can be further ascertained by an examination of the error, currency, objectives, nature and dependability associated with the secondary data.

Error and accuracy

The researcher must determine whether the data are accurate enough for the purposes of the present study. Secondary data can have a number of sources of error or inaccuracy, including errors in the approach, research design, sampling, data collection, analysis, and reporting stages of the project. Moreover, it is difficult to evaluate the accuracy of secondary data because the researcher did not participate in the research. One approach is to find multiple sources of data if possible, and compare them using standard statistical procedures.

Table 4.2 Criteria for evaluating secondary data

Criteria	*Issues*	*Remarks*
Specifications and research design	■ Data collection method ■ Response rate ■ Population definition ■ Sampling method ■ Sample size ■ Questionnaire design ■ Fieldwork ■ Data analysis	Data should be reliable, valid and generalisable to the problem at hand.
Error and accuracy	Examine errors in ■ Approach ■ Research design ■ Sampling ■ Data collection ■ Data analysis ■ Reporting	Assess accuracy by comparing data from different sources.
Currency	Time lag between collection and publication. Frequency of updates	Census data are periodically updated by syndicated firms.
Objective	Why were the data collected	The objective will determine the relevance of data.
Nature	■ Definition of key variables ■ Units of measurement ■ Categories used ■ Relationships examined	Reconfigure the data to increase their usefulness, if possible.
Dependability	Source: ■ Expertise ■ Credibility ■ Reputation ■ Trustworthiness	Preference should be afforded to an original rather than an acquired source.

example

Number crunch[8]

In December 1997, the Audit Bureau of Circulations (ABC) met UK newspaper publishers and major media buyers from the Institute of Practitioners in Advertising. The meeting aimed to thrash out a formula that could restore ABC's credibility as a trading currency.

Most observers agreed that the ABC's troubles were a direct result of squabbling between media owners. As circulations have continued to slide, the press barons have fought to hold on to their market share through price cuts, promotions and enhanced editorial packages. This has introduced an unprecedented volatility into their sales figures. Not content with trumpeting their own gains, some companies have sought to show up the deficiencies in their rivals' sales figures.

Figures under fire

The argument is best understood through a straightforward example of what is at stake. Let us take the October 1997 ABC figure for *The Times*, which was 814,899. That figure was a monthly circulation average which, prior to the recent dispute, would have been the only official benchmark that agencies used as a negotiating point with press owners (though they turn to data from the National Readership Survey and the Target Group Index to argue their case). At the heart of the dispute has been how that monthly figure is comprised. For example, were all the issues sold at the full price or were some given away cheaply as part of a subscription or promotional offer? Were any sold or given in bulk to an airliner or retailer, and if so how many? What about papers sold to Eire or Spain? Were they included in the total and if so how could that be justified as a piece of credible advertising data?

Another hot issue was the reliability of the monthly figure. Advertisers were dissatisfied with a number that they believe fails to reflect the reality of what they were buying. Director of press buying at The Media Centre, Tim Armes, says 'we'd like to know what each paper sells

daily and we'd like to know week to week fluctuations. The papers all boast about Saturday but keep quiet about Tuesday and Thursday. If one day is dramatically higher than the average, you don't have to be a brain surgeon to realise the others are lower.' ■

As this example indicates, the accuracy of secondary data can vary, particularly if they relate to phenomena that are subject to change. Moreover, data obtained from different sources may not agree. In these cases, the researcher should verify the accuracy of secondary data by conducting pilot studies or by other appropriate methods. Often, by exercising creativity this can be done without much expense or effort.

Currency: when the data were collected

Secondary data may not be current and the time lag between data collection and publication may be long, as is the case with much census data. Moreover, the data may not be updated frequently enough for the purpose of the problem at hand. Decision-makers require current data; therefore, the value of secondary data is diminished as they become dated. For instance, although the Census of Population data are comprehensive, they may not be applicable to major cities in which the population has changed rapidly during the last two years. Likewise, in the GlobalCash Project, the lists of the largest companies in Europe, used to decide who should be surveyed, have to be updated to reflect changes that take place in the two years between each study.

Objective: the purpose for which the data were collected

Data are invariably collected with some objective in mind, and a fundamental question to ask is why the data were collected in the first place. The objective for collecting data will ultimately determine the purpose for which that information is relevant and useful. Data collected with a specific objective in mind may not be appropriate in another situation. In the example at the start of this chapter, the sample surveyed by *Money* magazine was made up of 'frequent flyers'. The objective of the study was 'to uncover the airline characteristics consumers consider most important'. Air France, however, may wish to target 'business class' flyers and 'to uncover perceptions related to trade-offs made in customer service–price–safety'. Even though there may be identical questions used in both studies, the target respondents may be different, the rationale for the study presented to respondents will be different, and ultimately the 'state of mind' respondents may be in when they come to comparable questions will be different. The *Money* survey was conducted for entirely different objectives from those Air France have for their study. The findings from the *Money* survey may not directly support decision-making at Air France, though they may help to define who Air France should talk to and what questions they should put to them.

Nature: the content of the data

The nature, or content, of the data should be examined with special attention to the definition of key variables, the units of measurement, the categories used and the relationships examined. If the key variables have not been defined or are defined in a manner inconsistent with the researcher's definition, then the usefulness of the data is limited. Consider, for example, secondary data on consumer preferences for TV programmes. To use this information, it is important to know how preference for programmes was defined. Was it defined in terms of the programme watched most often, the one considered most needed, most enjoyable, most informative, or the programme of greatest service to the community?

Likewise, secondary data may be measured in units that may not be appropriate for the current problem. For example, income may be measured by individual, family, household or spending unit and could be gross or net after taxes and deductions. Income may be classified into categories that are different from research needs. If the

researcher is interested in high-income consumers with gross annual household incomes of over €120,000, secondary data with income categories of less than €20,000, €20,001–€50,000, €50,001–€75,000 and more than €75,000 will not be of use. Determining the measurement of variables such as income may be a complex task, requiring the wording of the definition of income to be precise. Finally, the relationships examined should be taken into account in evaluating the nature of data. If, for example, actual behaviour is of interest, then data inferring behaviour from self-reported attitudinal information may have limited usefulness. Sometimes it is possible to reconfigure the available data – for example, to convert the units of measurement – so that the resulting data are more useful to the problem at hand.

Dependability: how dependable are the data?

An overall indication of the dependability of data may be obtained by examining the expertise, credibility, reputation and trustworthiness of the source. This information can be obtained by checking with others who have used the information provided by the source. Data published to promote sales, to advance specific interests, or to carry on propaganda should be viewed with suspicion. The same may be said of data published anonymously or in a form that attempts to hide the details of the data collection research design and process. It is also pertinent to examine whether the secondary data came from an original source, one that generated the data, or an acquired source, one that procured the data from an original source. Generally, secondary data should be secured from an original rather than an acquired source. There are at least two reasons for this rule: first, an original source is the one that specifies the details of the data collection research design, and second, an original source is likely to be more accurate and complete than a surrogate source.

Classification of secondary data

Internal data
Data available within the organisation for whom the research is being conducted.

Figure 4.1 presents a classification of secondary data. Secondary data may be classified as either internal or external. **Internal data** are those generated within the organisation for which the research is being conducted. This information may be available in a ready-to-use format, such as information routinely supplied by the management decision

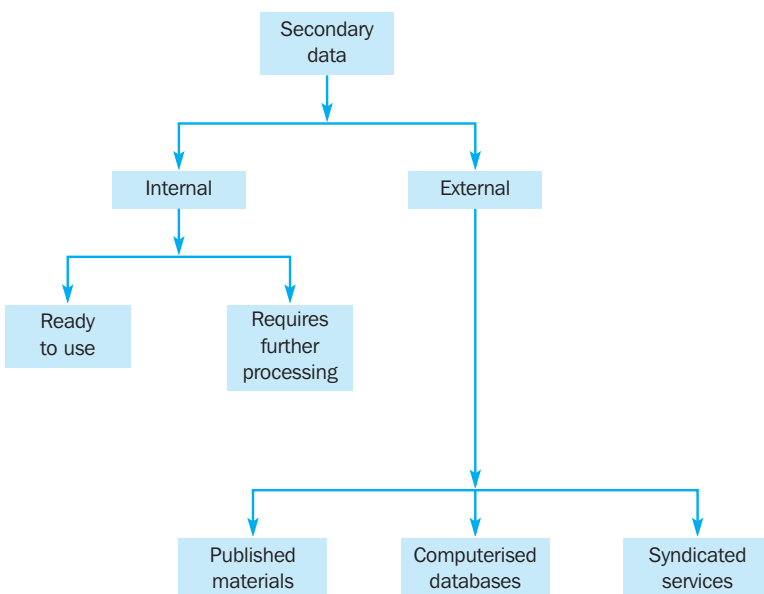

Figure 4.1
A classification of secondary data

support system. On the other hand, these data may exist within the organisation but may require considerable processing before they are useful to the researcher. For example, a variety of information can be found on sales invoices. Yet this information may not be easily accessed; further processing may be required to extract it. Secondary data generated from internal sources will be examined in more detail in Chapter 5. **External data**, on the other hand, are those generated by sources outside the organisation. These data may exist in the form of published material, online databases, or information made available by syndicated services. Externally generated secondary data may be more difficult to access, more expensive, and more difficult to evaluate for its accuracy, in comparison with internal secondary data. These factors mean that, before collecting external secondary data, it is useful to analyse readily available internal secondary data.

External data
Data that originate outside the organisation.

Published external secondary sources

Sources of published external secondary data include local authorities, regional and national governments, the EC, non-profit organisations (e.g. Chambers of Commerce), trade associations and professional organisations, commercial publishers, investment brokerage firms, and professional marketing research firms.[9] In fact, such a quantity of data is available that the researcher can be overwhelmed. Therefore, it is important to classify published sources (see Figure 4.2). Published external sources may be broadly classified as general business data or government data. General business sources comprise guides, directories, indexes and statistical data. Government sources may be broadly categorised as census data and other publications. These data types are discussed further with specific sources used as examples.

General business sources

Businesses publish a lot of information in the form of books, periodicals, journals, newspapers, magazines, reports and trade literature. This information can be located by using guides, directories and indexes. Sources are also available to identify statistical data.

Guides. Guides are an excellent source of standard or recurring information. A guide may help identify other important sources of directories, trade associations and trade publications. Guides are one of the first sources a researcher should consult. The following example illustrates the use of the Electronic Buyer's Guide.

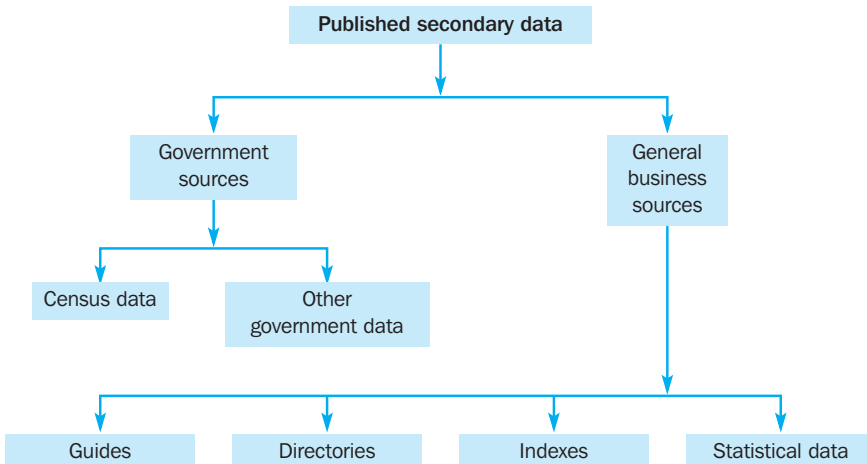

Figure 4.2
A classification of published secondary sources

example

Keeping up to data[10]

The market for business information is growing and the shift towards electronic sources continues. The choice of new media for directory publishers includes offline methods, such as CD-ROM, and online, such as the Internet. In addition, while online systems account for most information sales, the use of CDs is growing faster.

Access to up-to-date information is a big attraction. There is a perception that on-screen information is more up-to-date than a book. However, many CD-ROM directories are not updated any more frequently than their paper equivalents. So, while information on the Internet can be kept up-to-date, CD-ROMs come into their own when storing large amounts of information.

Miller Freeman launched the Electronic Buyer's Guide in 1996 and by October 1997 it was on the Internet (see **www.mfi.com**). There is some information available free of charge; if more information is wanted then there are charges. Alternatively, anyone who buys the guide on CD-ROM is given a password that gives them free access to the Internet directory. ■

Directories. Directories are helpful for identifying individuals or organisations that collect specific data. An example of a directory that you can examine on the Internet is the Central and Eastern European Business Directory. This interactive site provides current information on businesses and organisations in 24 central and eastern European countries (**www.ceebd.co.uk**). Another example is Europages, a reference business directory in Europe that classifies 500,000 companies in 30 European countries. Again, this can be accessed through the Internet and is available in English, French, German, Italian and Spanish versions (**www.europages.com**).

Indexes. It is possible to locate information on a particular topic in several different publications by using an index and abstracts. Indexes and abstracts, therefore, can increase the efficiency of the search process. Several indexes and abstracts are available for both academic and business sources. Examples of newspaper indexes include the *Financial Times Index* (**www.news.ft.com**), *Le Monde Index* (**www.le-monde.fr**), and the Japanese Business News online, *The Nikkei Weekly* (**www.nikkei.co.jp**). These indexes allow researchers to identify sources of particular topics, industries and individuals.

An example of a marketing index is the *Marketing Surveys Index* published by Euromonitor (**www.euromonitor.com**). This is a most comprehensive and up-to-date directory of business research on European and world markets. The index contains details of published research reports, a brief summary of its contents, a keyword index to the markets and products covered, and bibliographic details of the report.

A particularly useful abstract for marketing researchers is the *Market Research Abstracts* published by the Market Research Society in Britain (**www.mrs.org.uk**). Major European and American journals that relate to marketing research are reviewed and an abstract of each article is presented. The abstract is published twice a year. It is divided into sections that cover survey techniques; statistics, models and forecasting; attitude and behaviour research; psychographics, personality and social psychology; advertising and media research; applications of research; industrial market research; market research and general applications; and new product development. Such an abstract allows the researcher to quickly identify and evaluate the worth of journal papers that are relevant to their particular study.

Non-government statistical data. Published statistical data are of great interest to researchers. Graphic and statistical analyses can be performed on these data to draw important insights. Examples of non-governmental statistical data include trade associations such as the Swedish Tourism Trade Association (**www.sverigeturism.se**). The Swedish Information 'Smorgasbord' is a large single source of information in English on Sweden, Swedish provinces, nature, culture, lifestyle, society and industry. Another

example is Euromonitor (**www.euromonitor.com**), which publishes monthly market research journals covering subjects under the headings of Market Research Europe, Market Research GB, Market Research International and Retail Monitor International.

The United Nations provides an example of an organisation with a Statistics Division that provides a wide range of statistical outputs on a global basis (**www.un.org/depts/unsd**). The Statistics Division produces printed publications of statistics and statistical methods in the fields of international merchandise trade, national accounts, demography and population, social indicators, gender, industry, energy, environment, human settlements and disability. The Division also produces general statistical compendiums including the *Statistical Yearbook* and *World Statistics Pocketbook*. Many of the Division's databases in these fields are available as electronic publications in the form of CD-ROM, diskette and magnetic tape and on the Internet.

Government sources

European governments and the EU also produce large amounts of secondary data. Each European country has its own statistical office which produces lists of the publications available (and the costs involved). Examples of national statistical offices include the Centraal Bureau voor de Statistiek Nederlands (**www.cbs.nl**), Danmarks Statistik (**www.dst.dk**), the Federal Statistical Office of Germany (**www.statistik-bund.de**), the French Institut National de la Statistique et des Études Economiques (**www.insee.fr**), and the British Office for National Statistics (**www.statistics.gov.uk**). All of these offices have Internet links that allow you to quickly examine the array of publications that they produce. Their publications may be divided into census data and other publications.

Census data. Most European countries produce either catalogues or newsletters that describe the array of census publications available and the plans for any forthcoming census. In Britain, for example, *Census News* (**www.statistics.gov.uk/census2001/cennews.asp**) is a newsletter that contains the latest information about the 2001 census and previous censuses and is available four to six times a year. Census Marketing in Britain can supply unpublished data from the 1961, 1971, 1981 and 1991 censuses in the form of Small Area Statistics (SAS). SAS are available for standard census areas within England and Wales, such as counties, local government districts, London boroughs, wards, civil parishes and enumeration districts. Maps can also be purchased to complement the data.

Census data can be kept in electronic formats, allowing it to be analysed and presented in a variety of formats at a detailed geographical level. Given the long periods between national censuses and the amount of change that can occur in these periods, other data sources are used to maintain an up-to-date picture of specific regions.

As well as general population censuses, national statistical offices produce an array of industrial censuses. These may include industrial production, housing, construction, agriculture, restaurants and hotels, and financial services, for example **www.statistics.gov.uk/statbase/mainmenu.asp**.

Other government publications. In addition to the census, national statistical offices collect and publish a great deal of statistical data. Examining the Department of Statistics and Research in Cyprus as an example (**www.pio.gov.cy**), major industrial categories such as agriculture, construction, retailing and tourism are classified, with a whole array of available statistics. More generally, demographic, health, household income and expenditure, and labour statistical reports are also available.

Examples of reports from the British Office for National Statistics include *Family Spending* and *Economic Trends* (**www.statistics.gov.uk/statbase/publication.asp**). *Family Spending* provides a snapshot of household spending in the UK, explaining in detail how consumers spend their money. Expenditure patterns are broken down by age, economic

status, income and geography. *Economic Trends* provides monthly macroeconomic statistics of economic trends which include key data such as national accounts, gross domestic product, disposable income, balance of payments, trade in goods, prices, labour market information, industrial output, consumer sales credit and interest rates, as well as extensive commentary to put all this information into context.

In the EC, statistics are collected and published by the Statistical Office of the European Community (SOEC) in a series called Eurostat (**www.europa.eu.int/com/eurostat**). Tables normally contain figures for individual member states of the EU plus totals for all countries. Eurostat divides its publications into themes, which are:

- *Theme 1* – General statistics
- *Theme 2* – Economy and finance
- *Theme 3* – Population and social conditions
- *Theme 4* – Energy and industry
- *Theme 5* – Agriculture, forestry and fisheries
- *Theme 6* – External trade
- *Theme 7* – Distributive trades, services and transport
- *Theme 8* – Environment
- *Theme 9* – Research and development.

It also produces general titles which include *Eurostat Yearbook* (annual), *Basic Statistics* (annual), *Europe in Figures* (annual), *Key Figures* (monthly) and *Eurostatistics* (monthly).

To examine any of the national statistics offices in Europe, visit **www.cso.ie/links/eurolinks.html** and follow a link to the country of your choice. To examine other international statistics organisations visit **www.cso.ie/links/interlinks.html.**

Computerised databases

Most published information is also available in the form of computerised databases. Computerised databases contain information that has been made available in computer readable form for electronic distribution. From the 1980s through to today, the number of databases, as well as the vendors providing these services, has grown enormously. Computerised databases offer a number of advantages over printed data, including:[11]

1. The data are current and up-to-date, as publishers and data compilers are now using computers as the primary production technology.
2. The search process is more comprehensive, quicker and simpler. Online vendors provide ready access to hundreds of databases. Moreover, this information can be accessed instantaneously, and the search process is simplified as the vendors provide uniform search protocols and commands for accessing the database.
3. The cost of accessing these is relatively low, because of the accuracy of searching for the right data, and the speed of location and transfer of data.
4. It is convenient to access these data using a personal computer fitted with an appropriate communication device, such as a modem or a communication network.

While computerised database information can be helpful, it is vast and can be confusing. Thus a classification of computerised databases is helpful.

Classification of computerised databases

Computerised databases may be classified as online, Internet or offline as shown in Figure 4.3. Online databases consist of a central data bank that is accessed with a computer (or dumb terminal) via a telecommunications network. Internet databases can be accessed, searched and analysed on the Internet. It is also possible to download data from the Internet and store it in the computer or an auxiliary storage device.[12]

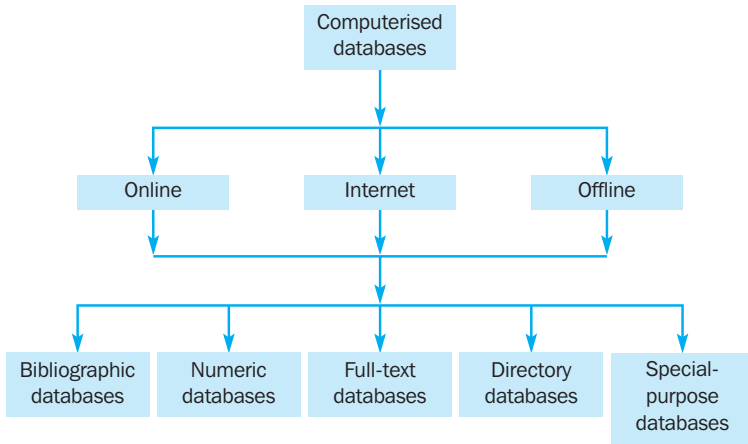

Figure 4.3
A classification of computerised databases

Offline databases make the information available on diskettes and CD-ROM disks. Thus, offline databases can be accessed at the user's location without the use of an external telecommunications network.[13]

Online, **Internet** and **offline databases** may be further classified as bibliographic, numeric, full text, directory or special-purpose databases. **Bibliographic databases** are composed of citations to articles in journals, magazines, newspapers, marketing research studies, technical reports, government documents and the like.[14] They often provide summaries or abstracts of the material cited. The earlier example of *Market Research Abstracts* (**www.imriresearch.com/index.html**) is an example of a bibliographic database. Another example is the *Aslib Index to Theses* (**www.theses.com**); this bibliographic database lists theses at masters and doctoral level and research degrees, including abstracts from Britain and Ireland.

Numeric databases contain numerical and statistical information. For example, some numeric databases provide time series data about the economy and specific industries. The earlier examples of census-based numeric databases using data over a series of censuses provide an example of a numeric database.

Full-text databases contain the complete text of the sources of the database. Examples include *World Advertising Research Center WARC* (**www.warc.com**), *Searchbank: European ASAP* (**www.infotrac.london.galegroup.com**) and *FT Discovery* (**www.lexis-nexis.com**). *WARC* is a supplier of intelligence to the global marketing, advertising, media and research communities. *Searchbank* has over 100 full-text journals on subjects including business, economics, current affairs and new technologies. It includes a spectrum of journals from professional trade publications through to refereed academic journals. *FT Discovery* presents the *Global News* link that gives full access to the Financial Times, World Reporter, Europe Intelligence Wire and Asia Intelligence Wire. It is possible to search on specific countries, sectors and publications.

Directory databases provide information on individuals, organisations and services. *European Interactive Directories* (**www.euroyellowpages.com**) is an example of a directory that has channels based upon EU community activities, country channels and thematic channels such as suppliers, wholesalers and shopping centres. Another example worth examining is the *ESOMAR* directory (**www.esomar.nl**) which provides details of member organisations throughout the world as well as many other publications of value to marketing researchers based in Europe.

Finally, there are **special-purpose databases**. For example, the Non-Governmental Organisation NGO directory (**www.rec.org/REC/Databases/NGODirectory/ NGOfind.html**) helps to track down information about environmental organisations working in central and eastern Europe. It has contact information for over 2,700 organisations from over 15 central and eastern European countries.

Online databases
Databases, stored in computers, that require a telecommunications network to access.

Internet databases
Databases that can be accessed, searched and analysed on the Internet. It is also possible to download data from the Internet and store it on the computer or an auxiliary device.

Offline databases
Databases that are available on diskette or CD-ROM.

Bibliographic databases
Databases composed of citations to articles in journals, magazines, newspapers, marketing research studies, technical reports, government documents, and the like. They often provide summaries or abstracts of the material cited.

Numeric databases
Databases containing numerical and statistical information that may be important sources of secondary data.

Full-text databases
Databases that contain the complete text of secondary source documents comprising the database.

Directory databases
Databases that provide information on individuals, organisations and services.

Special-purpose databases
Databases that contain information of a specific nature, e.g. data on a specialised industry.

In addition, virtually all libraries of major universities maintain special-purpose databases of research activities that reflect the distinct specialisms of that university. Beyond the internally generated, special-purpose databases, university libraries and reference libraries maintain computerised databases with instructions relating to what may be accessed and how it may be accessed. Another library source worth examining for computerised sources is the European Commission's 'Libraries' site (**www.europa.eu.int**). The site, which is multilingual, is distributed by the EUROPA server. EUROPA is the portal site of the European Union. It provides up-to-date coverage of European affairs and essential information on European integration. Users can access Websites of each of the EU institutions.

Syndicated sources of secondary data

Syndicated sources (services)
Information services offered by marketing research organisations that provide information from a common database to different firms that subscribe to their services.

In addition to published data or data available in the form of computerised databases, syndicated sources constitute the other major source of external secondary data. **Syndicated sources**, also referred to as **syndicated services**, are companies that collect and sell common pools of data designed to serve information needs shared by a number of clients. These data are not collected with a focus on a specific marketing problem, but the data and reports supplied to client companies can be personalised to fit specific needs. For example, reports could be organised based on the clients' sales territories or product lines. Using syndicated services is frequently less expensive than commissioning tailored primary data collection. Figure 4.4 presents a classification of syndicated sources. Syndicated sources can be classified based on the unit of measure-

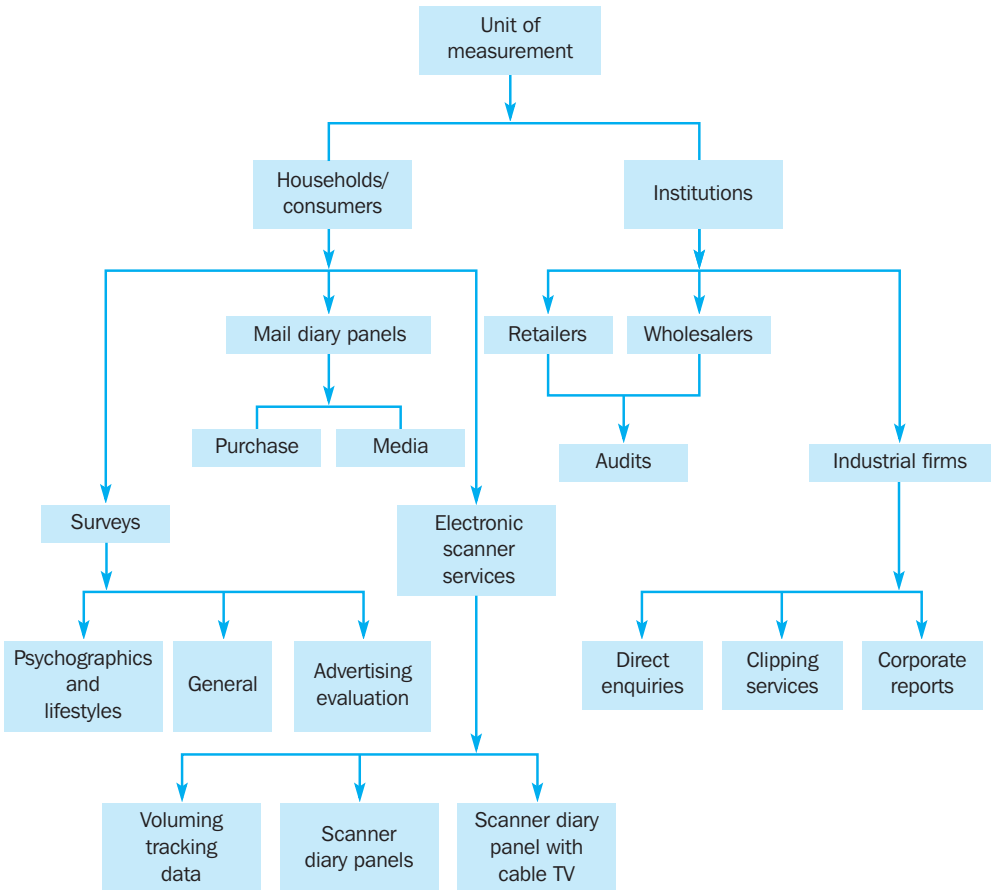

Figure 4.4
A classification of syndicated services

ment (households and consumers or institutions). Household and consumer data may be obtained from surveys, diary panels or electronic scanner services. Information obtained through surveys consists of values and lifestyles, advertising evaluation, or general information related to preferences, purchase, consumption and other aspects of behaviour. Diary panels emphasise information on purchases or media consumption. Electronic scanner services might provide scanner data only, scanner data linked to diary panels, or scanner data linked to diary panels and (cable) TV. When institutions are the unit of measurement, the data may be obtained from retailers, wholesalers or industrial firms. An overview of the various syndicated sources is given in Table 4.3. Each of these sources will be discussed.

Table 4.3 Overview of syndicated services

Type	*Characteristics*	*Advantages*	*Disadvantages*	*Uses*
Surveys	Surveys conducted at regular intervals	Most flexible way of obtaining data; information on underlying motives	Interviewer errors; respondent errors	Market segmentation, advertising theme selection and advertising effectiveness
Mail diary panels	Households provide specific information regularly over an extended period of time; respondents asked to record specific behaviour as it occurs	Recorded purchase behaviour can be linked to the demographic/ psychographic characteristics	Lack of representativeness; response bias; maturation	Forecasting sales, market share and trends; establishing consumer profiles, brand loyalty and switching; evaluating test markets, advertising and distribution
Diary media panels	Electronic devices automatically recording behaviour, supplemented by a diary	Same as mail diary panel	Same as mail diary panel	Establishing advertising rates; selecting media programme or air time; establishing viewer profiles
Scanner volume tracking data	Household purchases are recorded through electronic scanners in supermarkets	Data reflect actual purchases; timely data, less expensive	Data may not be representative; errors in recording purchases; difficult to link purchases to elements of marketing mix other than price	Price tracking, modelling effectiveness of in-store modelling
Scanner diary panels with cable TV	Scanner panels of households that subscribe to cable TV	Data reflect actual purchases; sample control; ability to link panel data to household characteristics	Data may not be representative; quality of data limited	Promotional mix analyses, copy testing, new product testing, positioning
Audit services	Verification of product movement by examining physical records or performing inventory analysis	Relatively precise information at retail and wholesale levels	Coverage may be incomplete; matching of data on competitive activity may be difficult	Measurement of consumer sales and market share; competitive activity; analysing distribution patterns; tracking of new products
Industrial firms	Data banks on industrial establishments created through direct enquiries of companies, clipping services and corporate reports	Important source of information in industrial firms; particularly useful in initial phases of the projects	Data are lacking in terms of content, quantity and quality	Determining market potential by geographic area, defining sales territories, allocating advertising budget

Syndicated data from households

Surveys

Surveys
Interviews with a large number of people using a questionnaire.

Omnibus survey
A distinctive form of survey that serves the needs of a syndicate group. The omnibus survey targets particular types of respondents such as those in specific locations, e.g. Luxembourg residents, or customers of particular types of product, e.g. business air travellers. With that target group of respondents, a core set of questions can be asked with other questions added as syndicate members wish.

Psychographics
Quantified profiles of individuals based upon lifestyle characteristics.

Lifestyles
Distinctive patterns of living described by the activities people engage in, the interests they have, and the opinions they hold of themselves and the world around them.

Various syndicated services regularly conduct **surveys** and **omnibus surveys**. In general, these surveys involve interviews with a large number of respondents using a pre-designed questionnaire. The distinction of the omnibus survey (example at **www.mori.com/charity/omnibus.shtml**) is that it targets particular types of respondents such as those in certain geographic locations, e.g. Luxembourg residents, or consumers of particular types of products, e.g. business air travellers. With that target group of respondents, a core set of questions can be asked with other questions added as syndicate members wish. Other syndicate members can 'jump on the omnibus' and buy the answers to all the questionnaire responses or to specific questions of their choice. Surveys and omnibus surveys may be broadly classified based on their content as psychographics and lifestyles, advertising evaluation, or general surveys.

Psychographics and lifestyles. **Psychographics** refer to the psychological profiles of individuals and to psychologically based measures of lifestyle. **Lifestyles** refer to the distinctive modes of living of a society or some of its segments. Together, these measures are generally referred to as activities, interests and opinions.

Advertising evaluation. The purpose of advertising evaluation surveys is to assess the effectiveness of advertising using print and broadcast media. Television commercials are evaluated using either the recruited audience method or the in-home viewing method. In the former method, respondents are recruited and brought to a central viewing facility, such as a theatre or mobile viewing laboratory. The respondents view the commercials and provide data regarding knowledge, attitudes and preferences related to the product being advertised and the commercial itself. See example at **www.bmrb.co.uk/researchingadvertising/addevelopment.htm.**

In the in-home viewing method, consumers evaluate commercials at home in their normal viewing environment. New commercials can be pre-tested at the network level or in local markets. A survey of viewers is then conducted to assess the effectiveness of the commercials.

General surveys. Surveys are also conducted for a variety of other purposes, including examination of purchase and consumption behaviour. Because a variety of data can be obtained, survey data have numerous uses. They can be used for market segmentation, as with psychographic and lifestyle data, and for establishing consumer profiles. Surveys are also useful for determining product image, measurement and positioning, and conducting price perception analysis. Other notable uses include advertising theme selection and evaluation of advertising effectiveness.

Mail diary panels

Often, survey data can be complemented with data obtained from diary panels (see example at **www.tnsofres.com/about/americas/argentina/argentina.cfm**). Panels were discussed in Chapter 3 in the context of longitudinal research designs. Diary panels are samples of respondents who provide specified information at regular intervals over an extended period of time. These respondents may be organisations, households or individuals, although household diary panels are most common. The distinguishing feature of diary panels is that the respondents record specific behaviours as they occur, in a diary. Typically, the diary is returned to the research organisation every one to four weeks. Panel members are compensated for their

participation with gifts, coupons, information or cash. Based on the content of information recorded, diary panels can be classified as diary purchase panels or diary media panels.

Diary media panels
A data gathering technique composed of samples of respondents whose television viewing behaviour is automatically recorded by electronic devices, supplementing the purchase information recorded in a diary.

Diary purchase panels
A data gathering technique in which respondents record their purchases in a diary.

In diary media panels, electronic devices automatically record viewing behaviour, thus supplementing a diary. **Diary media panels** yield information helpful for establishing advertising rates by radio and TV networks, selecting appropriate programming, and profiling viewer or listener subgroups. Advertisers, media planners and buyers find panel information particularly useful. **Diary purchase panels** provide information useful for forecasting sales, estimating market shares, assessing brand loyalty and brand-switching behaviour, establishing profiles of specific user groups, measuring promotional effectiveness, and conducting controlled store tests.

Compared with sample surveys (see Chapter 10), diary panels offer certain distinct advantages.[15] Panels can provide longitudinal data (data can be obtained from the same respondents repeatedly). People who are willing to serve on panels may provide more and higher-quality data than sample respondents. In diary purchase panels, information is recorded at the time of purchase, eliminating recall errors.[16] Information recorded by electronic devices is accurate because it eliminates human errors.

The disadvantages of diary panels include lack of representativeness, maturation and response biases. They may under-represent certain groups such as minorities and those with low education levels. This problem is further compounded by refusal to respond and attrition of panel members. Over time, maturation sets in, and the panel members must be replaced. Response biases may occur, since simply being on the panel may alter behaviour. Because purchase or media data are entered by hand, recording errors are also possible (see Chapter 3).

Electronic scanner services

The following example illustrates the nature and scope of electronic scanner services as undertaken by A.C. Nielsen, who conduct consumer panel services in 18 countries around the world.

example

A.C. Nielsen – the business economist at work[17]

Most of A.C. Nielsen's revenue comes from selling information on *fast-moving consumer goods* (FMCG). This information is compiled from either scanner data obtained from thousands of supermarkets, pharmacists and department stores, or from a 40,000-household panel who electronically record every aspect of every goods purchase they made using that item's bar code.

They go beyond their standard forms of analysing and presenting existing data to more creative interpretations that attempt to tell the marketer something he or she does not know. Often, creativity is valued more than sophisticated econometric techniques, as illustrated in the following two examples.

Inventory and sales data are available on food store retail sales, so an inventory–sales ratio can be calculated. This ratio had been increasing for several years up to 1992 when it began to shrink. Yet it is generally accepted that the size of the average new store is increasing, implying the need for additional inventories relative to sales. Coincidentally, more sophisticated inventory management techniques have become especially prevalent in food stores in recent years. Apparently, the ratio is being driven more by improved inventory management than by the opening of new, larger stores. A discrete cause and effect cannot be proven, but a linkage between the ratio and better inventory control mechanisms is highly probable.

Consumer spending patterns in Asia are very difficult to analyse, given the paucity of data available and the lack of data comparability across countries or spending components. By calculating consumer spending on food as a share of total consumer spending, and then ordering the results by per capita GDP, a relationship becomes obvious. Poorer countries

spend proportionately more of their resources on food, around 50%. For middle-income countries, this ratio slides from around 40% to 20%, and then holds steady near 20% for the developed economies. The implication is that, as a country obtains a middle-income status, because of this declining ratio, food sales will not grow as fast as other categories of consumer spending. ■

Although information provided by surveys and diary panels is useful, electronic scanner services are becoming increasingly popular. The role of scanned data as a foundation to developing sophisticated consumer databases is developed in Chapter 5. In this chapter we examine scanned data as a distinct source of syndicated data. **Scanner data** reflect some of the latest technological developments in the marketing research industry. They are collected by passing merchandise over a laser scanner that optically reads the bar-coded description (Universal Product Code, or UPC) printed on the merchandise. This code is then linked to the current price held in the computer memory and used to prepare a sales slip. Information printed on the sales slip includes descriptions as well as prices of all items purchased. Checkout scanners, now used in many retail stores, are revolutionising packaged goods marketing research.

Scanner data
Data obtained by passing merchandise over a laser scanner that reads the UPC code from the packages.

Three types of scanner data are available: **volume tracking data**, **scanner diary panels**, and **scanner diary panels with cable TV**. Volume tracking data provide information on purchases by brand, size, price and flavour or formulation, based on sales data collected from the checkout scanner tapes. This information is collected nationally from a sample of supermarkets with electronic scanners. In scanner diary panels, each household member is given an ID card that looks like a credit card. Panel members present the ID card at the checkout counter each time they shop. The checker keys in the ID numbers and each item of that customer's order. The information is stored by day of week and time of day.[18]

Volume tracking data
Scanner data that provide information on purchases by brand, size, price, and flavour or formulation.

Scanner diary panels
Scanner data where panel members are identified by an ID card, allowing information about each panel member's purchases to be stored with respect to the individual shopper.

Scanner diary panels with cable TV
The combination of a scanner diary panel with manipulations of the advertising that is being broadcast by cable television companies.

An even more advanced use of scanning, scanner diary panels with cable TV, combines diary panels with new technologies growing out of the cable TV industry. Households on these panels subscribe to one of the cable or TV systems in their market. By means of a cable TV 'split', the researcher targets different commercials into the homes of the panel members. For example, half the households may see test commercial A during the 6:00 pm newscast while the other half see test commercial B. These panels allow researchers to conduct fairly controlled experiments in a relatively natural environment.[19]

Uses of scanner data. Scanner data are useful for a variety of purposes.[20] National volume tracking data can be used for tracking sales, prices and distribution, for modelling, and for analysing early warning signals. Scanner diary panels with cable TV can be used for testing new products, repositioning products, analysing promotional mix, and making advertising decisions, including budget, copy and media, and pricing. These panels provide marketing researchers with a unique controlled environment for the manipulation of marketing variables.

Advantages and disadvantages of scanner data. Scanner data have an obvious advantage over surveys and diary panels; they reflect purchasing behaviour that is not subject to interviewing, recording, memory or expert biases. The record of purchases obtained by scanners is complete and unbiased by price sensitivity, because the panellist is not required to be particularly conscious of price levels and changes. Another advantage is that in-store variables such as pricing, promotions and displays are part of the dataset. The data are also likely to be current and can be obtained quickly. Finally, scanner panels with cable TV provide a highly controlled testing environment.

A major weakness of scanner data is lack of representativeness. National volume tracking data may not be generalisable to the total population, because only large supermarkets have scanners. In addition, certain types of outlets such as food warehouses, pharmacists and mass merchandisers are excluded. Likewise, scanners have limited geographical dispersion and coverage.

The quality of scanner data may be limited by several factors. Not all products may be scanned. For example, to avoid lifting a heavy item, a sales assistant may use the register to ring it up. If an item does not scan on the first try, the assistant may key in the price and ignore the bar code. Sometimes a consumer purchases many flavours of the same item, but the assistant scans only one package and then rings in the number of purchases. Thus, the transaction is inaccurately recorded. With respect to scanner panels, the available technology permits the monitoring of only one TV set per household. Hence, there is a built-in bias if the household has more than one TV set. In addition, the system provides information on TV sets in use rather than actual viewing behaviour. Although scanner data provide behavioural and sales information, they do not provide information on underlying attitudes and preferences and the reasons for specific choices.

Syndicated data from institutions

Retailer and wholesaler audits

Audit
A data collection process derived from physical records or performing inventory analysis. Data are collected personally by the researcher, or by representatives of the researcher, and are based on counts usually of physical objects rather than people.

As Figure 4.4 shows, syndicated data are available from retailers and wholesalers as well as industrial firms (see the example of the work of German marketing research company GfK at **www.gfkms.com**). The most popular means of obtaining data from retailers and wholesalers is an audit. An **audit** is a formal examination and verification of product movement carried out by examining physical records or analysing inventory. Retailers and wholesalers who participate in the audit receive basic reports and cash payments from the audit service. Audit data focus on the products or services sold through the outlets or the characteristics of the outlets themselves.

example

Retail auditing for retailing information[21]

Introduced in the USA in 1933, A.C. Nielsen pioneered food and drug indices to measure and understand the performance and dynamics of product sales. The Retail Measurement Service of A.C. Nielsen uses store audit data on product movement, market share, distribution, price and other market-sensitive information in over 80 countries across six continents. Using in-store scanning of product codes and store visits by auditors, sample and census information is gathered across the food, household, health and beauty, durables, confectionery and beverage industries. ■

Retail audit data can be useful to consumer product firms. For example, if Colgate-Palmolive is contemplating the introduction of a new toothpaste brand, a retail audit can help determine the size of the total market and distribution of sales by type of outlet and by different regions.

Wholesale audit services, the counterpart of retail audits, monitor warehouse withdrawals. Participating operators, who include supermarket chains, wholesalers and frozen-food warehouses, typically account for over 80% of the volume in the area.

The uses of retail and wholesale audit data include:

1 determining the size of the total market and the distribution of sales by type of outlet, region, or city;
2 assessing brand shares and competitive activity;

Electronic scanning can capture the products used by a consumer far more accurately than manual methods.

3 identifying shelf space allocation and inventory problems;
4 analysing distribution problems;
5 developing sales potentials and forecasts;
6 developing and monitoring promotional allocations based on sales volume.

Audits provide relatively accurate information on the movement of many different products at the wholesale and retail levels. Furthermore, this information can be broken down by a number of important variables, such as brand, type of outlet and size of market. Audits have limited coverage, however; not all markets or operators are included. In addition, audit information may not be timely or current, particularly compared with scanner data. Typically, there is a two-month gap between the completion of the audit cycle and the publication of reports. Another disadvantage is that, unlike scanner data, audit data cannot be linked to consumer characteristics. In fact, there may even be a problem in relating audit data to advertising expenditures and other marketing efforts. Some of these limitations are overcome in computerised audit panels.

Industrial firms

These provide syndicated data about industrial firms, businesses and other institutions. Syndicated data are collected by making direct enquiries to organisations, from clipping services and through the analysis of corporate reports (see example at **www.journalismnet.com/choose/clippings.htm**). The range and sources of syndicated data available to industrial goods firms are more limited than those available to consumer goods firms.

Industrial firm information is useful for sales management decisions, including identifying prospects, defining territories, setting quotas, and measuring market potential by geographic areas. It can also aid in advertising decisions such as targeting prospects, allocating advertising budgets, selecting media, and measuring advertising effectiveness. This kind of information is useful for segmenting the market and for designing custom products and services for important segments.

International marketing research

A wide variety of secondary data are available for international marketing research. As with domestic research, the problem is not lack of data but the plethora of information available. Evaluation of secondary data is even more critical for international than for domestic projects. Different sources report different values for a given statistic, such as the gross domestic product (GDP), because of differences in the way the unit is defined. Measurement units may not be equivalent across countries. In France, for example, many workers are paid a thirteenth monthly salary each year as an automatic bonus, resulting in a measurement construct that is different from those in other countries.[22] The accuracy of secondary data may also vary from country to country. Data from highly industrialised countries in Europe are likely to be more accurate than those from developing nations. Business and income statistics are affected by the taxation structure and the extent of tax evasion. Population censuses may vary in frequency and year in which the data are collected. In Britain, for example, the census is conducted every 10 years, whereas in the People's Republic of China there was a 29-year gap between the censuses of 1953 and 1982. This situation, however, is changing quickly. Several syndicated firms are developing huge sources of international secondary data, as illustrated in the following example.

example

Los Medios y Mercados de Latinoamerica (LMML)[23]

Started in 1994 by Audits & Surveys worldwide, *Los Medios y Mercados de Latinoamerica* (The Markets and Media of Latin America) is the largest multinational survey of media and consumer habits that is conducted in Latin America to provide managers with important information for their marketing strategies. The study, which is repeated every year, aims at tracking the development of media and consumer habits in Latin America.

A recent multinational survey was conducted in 18 Latin American countries, including Argentina, Brazil, Colombia, Mexico and Venezuela. It sampled 6,634 respondents between the ages of 12 and 64. Its probability sample, representing urban as well as rural Latin America, can be projected to 280 million people or 79 million households.

The research design for this survey involved two steps. First, the personal interview technique was used to measure the variety of media, including newspapers, multinational and local magazines, television and radio. Then a 25-page self-administered booklet was passed to the respondent to measure their product consumption and usage in over 100 categories and 800 brands. Demographic data gathered about the respondents included country/region, age, sex, employment status, occupation, education, household size, annual household income, car ownership, household goods owned and services taken.

Companies can easily use these data, as the survey results are provided in a set of 14 printed volumes, and also in computer database formats including an SPSS format. ■

Ethics in marketing research

Possible ethical dilemmas exist when using internal or external secondary data. Some ethical issues that are pertinent include:

- The unnecessary collection of primary data when the problem can be addressed based on secondary data alone
- The use of only secondary data when primary data are needed
- The use of secondary data that are not applicable
- The use of secondary data that have been gathered through morally questionable means
- Compromising the anonymity of customer details held on databases.

As was discussed in Chapter 2, the unnecessary collection of expensive primary data when the research problem can be addressed using only secondary data is unethical. In this case, the researcher is using a more expensive method that is less appropriate. Similarly, the exclusive reliance on secondary data when the research problem requires primary data collection could raise ethical concerns. This is particularly true if the researcher is charging a fixed fee for the project and the research design was not specified in advance. Here again, the researcher's profit goes up, but at the expense of the client.

The researcher is ethically obliged to ensure the relevance and usefulness of secondary data to the problem at hand. The secondary data should be evaluated by the criteria discussed earlier in this chapter. Only data judged to be appropriate should be used.

Internet and computer applications

The World Wide Web as an online source of secondary data

The World Wide Web is a vital source of secondary data and intelligence for the marketing researcher. The speed of the Internet can aid rapid problem diagnosis and data collection at various stages of the research process. Given the global nature of the technology, the Internet is a vital tool for the international marketing researcher. Searching the Web is facilitated by using generalist search engines such as Google, Yahoo! or AltaVista, which require a few keywords to get hundreds of sites related to one subject. One can go directly to the Websites of traditional suppliers of secondary data from government or business sources. Many of those sites also have inside search engines that sort data from the supplier's internal database. Information on the Web is of great value as generally it is current, though care must be taken to note when Web pages have been updated. It should be noted that not all secondary data on the Web is free. The Web may reveal the existence of data on a particular subject or industry, but remember the costs involved in conducting quality research. Hence the Web may be used to give an awareness and a 'taste' of secondary data but it does not necessarily mean 'free' data. See the Companion Website for Professional Perspective 9 by Trevor Fenwick. Trevor's article 'How online retrieval is devaluing research' is a reminder of the quality issues of secondary data gathered from the Internet.

➤➤➤ *See Professional Perspective 9.*

Internal secondary data

Large organisations have intranets, which greatly facilitate the search for access to secondary data. The Coca-Cola Company, for example, has developed powerful intranet applications that enable Coca-Cola managers worldwide to search for past and present research studies and a wide variety of marketing-related information on the basis of keywords. Once located, the information can be accessed online. Even sensitive and restricted information can be accessed by obtaining permission electronically. Visit **www.intranets.com** for a fuller description of intranet technology, evaluations of software and an interactive tour.

External secondary data

Information can be obtained by visiting various business-related sites that provide sales leads and mailing lists, business profiles and credit ratings. Various newspapers, magazines and journals can be accessed on the Web with excellent indexing facilities to locate particular subjects, companies and individuals. Government data for the European Community and for individual countries through to regional and city councils can be accessed via the Web, though the quality and quantity of data available through government sources can vary enormously.

Syndicated sources of information

For syndicated sources of information one can visit the home pages of the various marketing research companies and providers of syndicated information. The A.C. Nielsen home page at **www.acnielsen.com** is a good example. This site provides links to various manufacturers and to various countries such as Britain, Canada and Spain. Other good sources of syndicated data include Mintel, which can be reached at **www.mintel.co.uk**, Taylor Nelson Sofres at **www.tnsofres.com/research/prodfinder.cfm** (it is worth working through the great array of marketing research products they offer), the German marketing research company GfK (English version) at **www.gfk.de/english/presse/broschueren/produkte/BehaviorScan_e.pdf**, and a French version of BehaviorScan at **www.marketingscan.fr/produits/ behaviorscan/produits_behavior.html**.

International secondary data

The Internet has emerged as the most extensive source of secondary information. The utility of the Internet for the marketing researcher is further enhanced due to the easy accessibility and retrieval of information and the ability to cross-validate information from a variety of sources. Most of the Internet links highlighted in this chapter allow for secondary data to be gathered from countries all over the world. Whilst necessarily we have a focus on European matters, it is well worth visiting **www.quirks.com**. The Quirks.com Website is a most thorough source for information on marketing research, including case studies of successful research projects and a comprehensive list of directories.

As a final illustration of the power of the Internet for secondary data collection and analyses, go to the Dun and Bradstreet home page at **www.dbn.com**, and directly access the services and syndicated information for the country of your choice.

Summary

In contrast to primary data, which originate with the researcher for the specific purpose of the problem at hand, secondary data and intelligence are data originally collected for other purposes. Secondary data can be obtained quickly and are relatively inexpensive. They have limitations, and should be carefully evaluated to determine their appropriateness for the problem at hand. The evaluation criteria consist of specifications, error, currency, objectivity, nature and dependability.

A wealth of information exists in the organisation for which the research is being conducted. This information constitutes internal secondary data. External data are generated by sources outside the organisation. These data exist in the form of published (printed) material, online and offline databases, or information made available by syndicated services. Published external sources may be broadly classified as general business data or government data. General business sources comprise guides, directories, indexes and statistical data. Government sources may be broadly categorised as census data and other data. Computerised databases may be online or offline. Both online and offline databases may be further classified as bibliographic, numeric, full-text, directory or specialised databases.

Syndicated sources are companies that collect and sell common pools of data designed to serve a number of clients. Syndicated sources can be classified based on the unit of measurement (households and consumers or institutions). Household and consumer data may be obtained via surveys, diary purchase or media panels, or electronic scanner services. When institutions are the unit of measurement, the data may be obtained from retailers, wholesalers or industrial units. It is desirable to combine information obtained from different secondary sources.

Several specialised sources of secondary data are useful for conducting international marketing research. The evaluation of secondary data becomes even more critical, however, because the usefulness and accuracy of these data can vary widely. Ethical dilemmas that can arise include the unnecessary collection of primary data, the use of only secondary data when primary data are needed, the use of secondary data that are not applicable, and the use of secondary data that have been gathered through morally questionable means.

Questions

1. What are the differences between primary data, secondary data and marketing intelligence?
2. What are the relative advantages and disadvantages of secondary data?
3. At what stages of the marketing research process can secondary data be used?
4. Why is it important to locate and analyse secondary data before progressing to primary data?
5. How may secondary data be used to validate qualitative research findings?
6. What is the difference between internal and external secondary data?
7. How can intranet technology help in the location and dissemination of secondary data?
8. By what criteria may secondary data be evaluated?
9. What criteria would you look for when examining the design and specifications of secondary data? Why is it important to examine these criteria?

10 To what extent should you use a secondary data source if you cannot see any explicit objectives attached to that research?

11 If you had two sources of secondary data for a project, the first being dependable but out of date, the second not dependable but up to date, which would you prefer?

12 Describe, with examples, the main types of government and business secondary data sources.

13 List and describe the main types of syndicated sources of secondary data.

14 Explain what a diary panel is. What are the advantages and disadvantages of traditional and scanner diary panels?

15 What is an audit? Describe the uses, advantages and disadvantages of audits.

Notes

1 Keating, P., 'The best airlines to fly today', *Money* (November 1997), 118–28.
2 Aguilar, F.J., *Scanning the Business Environment* (London: Macmillan, 1967).
3 Adapted from Brownlie, D., 'Environmental scanning', in Baker, M.J., *The Marketing Book*, 3rd edn, (Oxford: Butterworth-Heinemann), 158.
4 Curtis, J., 'Behind enemy lines', *Marketing* (24 May 2001), 28–9.
5 For a recent application of secondary data, see Czinkota, M.R. and Ronkainen, I.A., 'Market research for your export operations: Part I – Using secondary sources of research', *International Trade Forum* 3 (1994), 22–33; Jain, D., Mahajan, V. and Muller, E., 'Innovation diffusion in the presence of supply restrictions', *Marketing Science* (Winter 1991) 83–90.
6 Jacob, H., *Using Published Data: Errors and Remedies* (Beverly Hills, CA: Sage Publications, 1984).
7 Stewart, D.W., *Secondary Research: Information Sources and Methods* (Beverly Hills, CA: Sage Publications, 1984), 23–33.
8 Fry, A., 'Number crunch', *Marketing* (4 December 1997), 29.
9 Fries, J.R., 'Library support for industrial marketing research', *Industrial Marketing Management* 11 (February 1982), 47–51.
10 Miller, R., 'Keeping up to data', *Marketing* (30 January 1997), 35.
11 Post, C., 'Marketing data marts help companies stay ahead of the curve and in front of the competition', *Direct Marketing* 59 (April 1997), 37–40.
12 Notess, G.R., 'Searching the hidden Internet', *Database* 20 (June/July 1997), 37–40.
13 Quint, B., 'Assume the position, take the consequences', *Information Today* 13 (June 1996), 11–13.
14 Notess, G.R., 'The Internet as an on-line service: bibliographic databases on the net', *Database* 19 (August/September 1996), 92–5.
15 Parfitt, J.H. and Collins, B.J.K., 'Use of consumer panels for brand share predictions', *Journal of the Market Research Society* 38(4) (October 1996), 341–67; Ramaswamy, V. and Desarbo, W.S., 'SCULPTURE: a new methodology for deriving and analyzing hierarchical product-market structures from panel data', *Journal of Marketing Research* 27 (November 1990), 418–27.
16 Sudman, S., 'On the accuracy of recording of consumer panels I', *Journal of Marketing Research* (August 1964), 69–83; Sudman, S. and Bradburn, M., 'Response effects in surveys', *Journal of Marketing Research* (May 1964), 14–20; Sudman, S. and Ferber, R., *Consumer Panels* (Chicago, IL: American Marketing Association, 1978); Sudman, S., *On the Accuracy of Recording of Consumer Panels II*, Learning Manual (New York: Neal-Schumen Publishers, 1981).
17 Handler, D., 'The business economist at work: linking economics to market research: A.C. Nielsen', *Business Economics* 31 (October 1996), 51.
18 Andrew, R.L. and Srinivasan, T.C., 'Studying consideration effects in empirical choice models using scanner panel data', *Journal of Marketing Research* 32 (February 1995), 30-41; and Bucklin, R.E., Gupta, S. and Han, S., 'A brand's eye view of response segmentation in consumer brand choice behaviour', *Journal of Marketing Research* 32 (February 1995), 66–74.
19 It is possible to combine store-level scanner data with scanner panel data to do an integrated analysis. See, for example, Russell, G.J. and Kamakura, W.A., 'Understanding brand competition using micro and macro scanner data', *Journal of Marketing Research* 31 (May 1994), 289–303.
20 Examples of scanner data applications include Fader, P.S. and McAlister, L., 'An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data', *Journal of Marketing Research* 27 (August 1990), 322–32; Waarts, E., Carree, M. and Wierenga, B., 'Full-information maximum likelihood estimation of brand positioning maps using supermarket scanning data', *Journal of Marketing Research* 28 (November 1991), 483–90.
21 Based on information obtained from the suppliers.
22 Douglas, S.P. and Craig, C.S., *International Marketing Research* (Upper Saddle River, NJ: Prentice Hall, 1983).
23 Rydholm, J., 'A united effort', *Quirk's Marketing Research Review* (October 1996).