CHAPTER 8

Q 1) Define Secondary Data. What are the advantages and disadvantages of secondary data?

Secondary data, which are gathered and recorded by someone else prior to (and for purposes other than) the current project. Secondary data usually are historical and already assembled. They require no access to respondents or subjects.

Advantages of Secondary Data

The primary advantage of secondary data is their availability. Obtaining secondary data is almost always faster and less expensive than acquiring primary data. This is particularly true when researchers use electronic retrieval to access data stored digitally. In many situations, collecting secondary data is instantaneous.

Consider the money and time saved by researchers who obtained updated population estimates for a town during the interim between the 2000 and 2010 censuses. Instead of doing the fieldwork themselves, researchers could acquire estimates from a firm dealing in demographic information or from sources such as Claritas or PCensus. As in this example, the use of secondary data eliminates many of the activities normally associated with primary data collection, such as sampling and data processing.

Secondary data are essential in instances when data cannot be obtained using primary data collection procedures. For example, a manufacturer of farm implements could not duplicate the information in the *Census of Agriculture* because much of the information there (for example, amount of taxes paid) might not be accessible to a private firm.

Disadvantages of Secondary Data

An inherent disadvantage of secondary data is that they were not designed specifically to meet the researchers' needs. Thus, researchers must ask how pertinent the data are to their particular project. To evaluate secondary data, researchers should ask questions such as these:

- Is the subject matter consistent with our problem definition?
- Do the data apply to the population of interest?
- Do the data apply to the time period of interest?
- Do the secondary data appear in the correct units of measurement?
- Do the data cover the subject of interest in adequate detail?

Even when secondary information is available, it can be inadequate. Consider the following typical situations:

- A researcher interested in forklift trucks finds that the secondary data on the subject are included in a broader, less pertinent category encompassing all industrial trucks and tractors. Furthermore, the data were collected five years earlier.
- An investigator who wishes to study individuals earning more than $100,000 per year finds the top category in a secondary study reported at $75,000 or more per year.
- A brewery that wishes to compare its per-barrel advertising expenditures with those of competitors finds that the units of measurement differ because some report point-of-purchase expenditures with advertising and others do not.
- Data from a previous warranty card study show where consumers prefer to purchase the product but provide no reasons why.

Q 2)Explain Objectives for Secondary-Data Research Designs.

Broad Objective Specific Research Example

1)Fact-finding: Identifying consumption patterns Tracking trends.

2)Model building: Estimating market potential Forecasting sales Selecting trade areas and sites.

3)Database marketing: Enhancing customer databases

Developing prospect lists.

Secondary research designs address many common business research problems. There are three general cat-egories of secondary research objectives: fact-finding, model building, and database marketing. A typical fact-finding study might seek to uncover all available information about consumption patterns for a particular product category or to identify business trends that affect an industry. Model building is more complicated; it involves specifying relationships between two or more variables. The practice of database marketing, which involves maintaining customer databases with customers' names, addresses, phone numbers, past purchases, responses to past promotional offers, and other relevant data such as demographic and financial data, is increasingly being supported by business research efforts.

Q 3) What is Data Mining? Write a note on market-basket analysis.

The term data mining refers to the use of powerful computers to dig through volumes of data to discover patterns about an organization's customers and products. As seen in the Research Snapshot on the next page, this can even apply to Internet content from blogs. It is a broad term that applies to many different forms of analysis. For example, neural networks are a form of artificial intelligence in which a computer is programmed to mimic the way that human brains process information.

Market-basket analysis is a form of data mining that analyzes anonymous point-of-sale transaction databases to identify coinciding purchases or relationships between products purchased and other retail shopping information.

Consider this example about patterns in customer purchases:

Osco Drugs mined its databases provided by checkout scanners and found that when men go toits drug stores to buy diapers in the evening between 6:00 p.m. and 8:00 p.m., they sometimes walk out with a six-pack of beer as well. Knowing this behavioral pattern, supermarket managers may consider laying out their stores so that these items are closer together.

Q 4) Describe the impact of single-source data and globalization on secondary data research.

The marketing of multiple types of related data by single-source suppliers has radically changed the nature of secondary-data research. Businesses can measure promotional efforts and related buyer behavior by detailed customer characteristics. As business has become more global, so has the secondary-data industry. International researchers should watch for pitfalls that can be associated with foreign data and cross-cultural research, such as problems with the availability and reliability of data.

Q 5) Give examples of various external sources of secondary data.
External data are generated or recorded by another entity. The government, newspaper and journal publishers, trade associations, and other organizations create or produce information. Traditionally this information has been distributed in published form, either directly from producer to researcher, or indirectly through intermediaries such as public libraries. Modern computerized data archives, electronic data interchange, and the Internet have changed the distribution of external data, making them almost as accessible as internal data. Push technology is a term referring to an Internet information technology that automatically delivers content to the researcher's or manager's desktop. This service helps in environmental scanning.

6) What are the sources of secondary data? Explain.

Secondary data can be classified as either internal to the organization or external. Modern information technology makes this distinction seem somewhat simplistic. Some accounting documents are indisputably internal records of the organization. Researchers in another organization cannot have access to them. Clearly, a book published by the federal government and located at a public library is external to the company. However, in today's world of electronic data interchange, the data that appear in a book published by the federal government may also be purchased from an online information vendor for instantaneous access and subsequently stored in a company's decision support system.

■ LIBRARIES
Traditionally, libraries' vast storehouses of information have served as a bridge between users and producers of secondary data. The library staff deals directly with the creators of information, such as the federal government, and intermediate distributors of information, such as abstracting and indexing services.

■ THE INTERNET
Today, of course, much secondary data is conveniently available over the Internet. Its creation has
added an international dimension to the acquisition of secondary data. For example, Library Spot,at http://www.libraryspot.com, provides links to online libraries, including law libraries, medical libraries, and music libraries.

■ VENDORS
The information age offers many channels besides libraries through which to access data. Many external producers make secondary data available directly from the organizations that produce the data or through intermediaries, which are often called vendors.

■ PRODUCERS

Classifying external secondary data by the nature of the producer of information yields five basic sources: publishers of books and periodicals, government sources, media sources, trade association sources, and commercial sources.

■ Books and Periodicals

Some researchers consider books and periodicals found in a library to be the quintessential secondary data source. A researcher who finds books on a topic of interest obviously is off to a good start.

■Government Sources

Government agencies produce data prolifically. Most of the data published by the federal government can be counted on for accuracy and quality of investigation. Most students are familiar with the U.S. Census of Population, which provides a wealth of data.

■Media Sources

Information on a broad range of subjects is available from broadcast and print media. CNN Financial News and BusinessWeek are valuable sources for information on the economy and many industries. Media frequently commission research studies about various aspects of Americans' lives,such as financial affairs, and make reports of survey findings available to potential advertisers free of charge.

Q 7) Explain various internal and proprietary sources of secondary data.

Managers often get data from internal proprietary sources such as accounting records. Data mining is the use of powerful computers to dig through volumes of data to discover patterns about an organization's customers and products. It is a broad term that applies to many different forms of analysis.