

Exploratory Research Design:
Secondary Data

The slide features a large, tilted rectangular frame containing abstract geometric shapes: a circle, a rectangle, and several lines. To the right of this frame is a small, square, grayscale illustration of a person with a surprised or excited expression, surrounded by motion lines.



Agenda

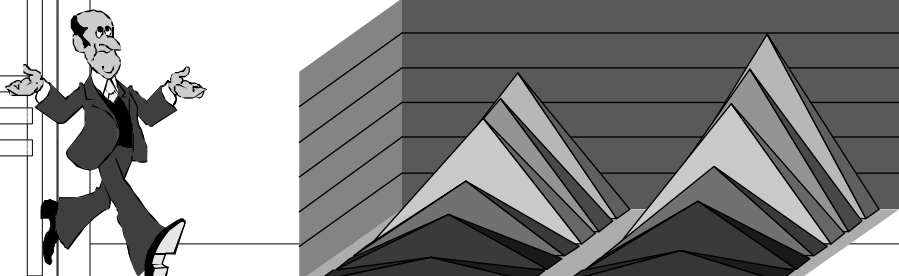
- 1) Overview
- 2) Primary versus Secondary Data
- 3) Advantages & Uses of Secondary Data
- 4) Disadvantages of Secondary Data
- 5) Criteria for Evaluating Secondary Data
- 6) Classification of Secondary Data
- 7) Internal Secondary Data
- 8) Published External Secondary Sources
- 9) Computerized Databases
- 10) Syndicate Sources of Secondary Data
- 11) Syndicated Data from Households
- 12) Electronic Scanner Services
- 13) Syndicated Data from Institutions
- 14) Combining Information from Different Sources: Single-Source Data
- 15) Applications of Secondary Data
- 16) International Marketing Research
- 17) Ethics in Marketing Research

Primary vs. Secondary Data

- **Primary data** are originated by a researcher for the specific purpose of addressing the problem at hand. The collection of primary data involves all six steps of the marketing research process.
- **Secondary data** are data which have already been collected for purposes other than the problem at hand. These data can be located quickly and inexpensively.

A Comparison of Primary & Secondary Data

	Primary Data	Secondary Data
Collection purpose	For the problem at hand	For other problems
Collection process	Very involved	Rapid & easy
Collection cost	High	Relatively low
Collection time	Long	Short



Uses of Secondary Data

- Identify the problem
- Better define the problem
- Develop an approach to the problem
- Formulate an appropriate research design (for example, by identifying the key variables)
- Answer certain research questions and test some hypotheses
- Interpret primary data more insightfully

Criteria for Evaluating Secondary Data

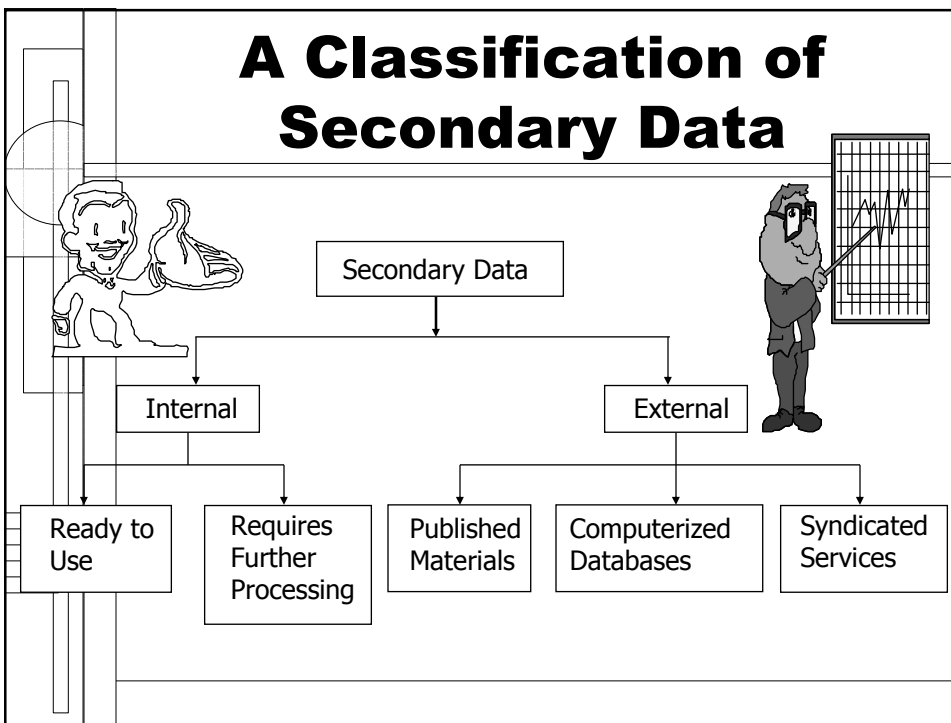
- **Specifications:** Methodology Used to Collect the Data
- **Error:** Accuracy of the Data
- **Currency:** When the Data Were Collected
- **Objective(s):** The Purpose for Which the Data Were Collected
- **Nature:** The Content of the Data
- **Dependability:** Overall, How Dependable Are the Data

Criteria for Evaluating Secondary Data

Table 4.2

Criteria	Issues	Remarks
Specifications & Methodology	Data collection method, response rate, quality & analysis of data, sampling technique & size, questionnaire design, fieldwork.	Data should be reliable, valid, & generalizable to the problem.
Error & Accuracy	Examine errors in approach, research design, sampling, data collection & analysis, & reporting.	Assess accuracy by comparing data from different sources.
Currency	Time lag between collection & publication, frequency of updates.	Census data are updated by syndicated firms.
Objective	Why were the data collected?	The objective determines the relevance of data.
Nature	Definition of key variables, units of measurement, categories used, relationships examined.	Reconfigure the data to increase their usefulness.
Dependability	Expertise, credibility, reputation, & trustworthiness of the source.	Data should be obtained from an original source.

A Classification of Secondary Data



Internal Secondary Data

Department Store Project

Sales were analyzed to obtain:

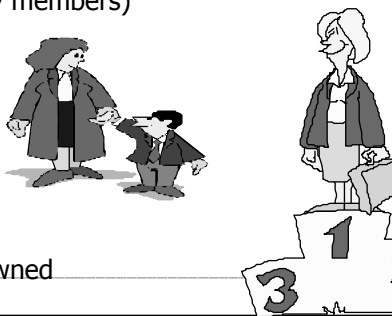
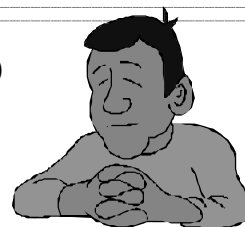
- Sales by product line
- Sales by major department (e.g., men's wear, house wares)
- Sales by specific stores
- Sales by geographical region
- Sales by cash versus credit purchases
- Sales in specific time periods
- Sales by size of purchase
- Sales trends in many of these classifications were also examined.

Type of Individual/Household Level Data Available from Syndicated Firms

I.

Demographic Data

- Identification (name, address, telephone)
- Sex
- Marital status
- Names of family members
- Age (including ages of family members)
- Income
- Occupation
- Number of children present
- Home ownership
- Length of residence
- Number and make of cars owned

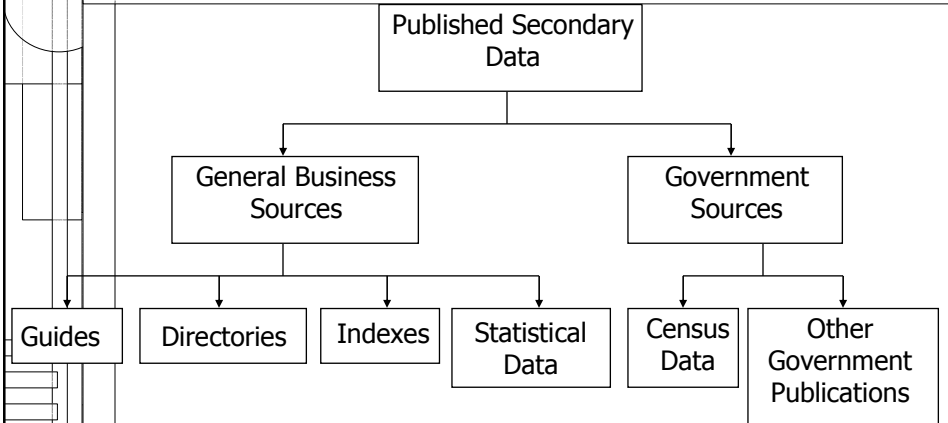


Type of Individual/Household Level Data Available from Syndicated Firms

- II. Psychographic Lifestyle Data
- Interest in golf
 - Interest in snow skiing
 - Interest in book reading
 - Interest in running
 - Interest in bicycling
 - Interest in pets
 - Interest in fishing
 - Interest in electronics
 - Interest in cable television



A Classification of Published Secondary Sources

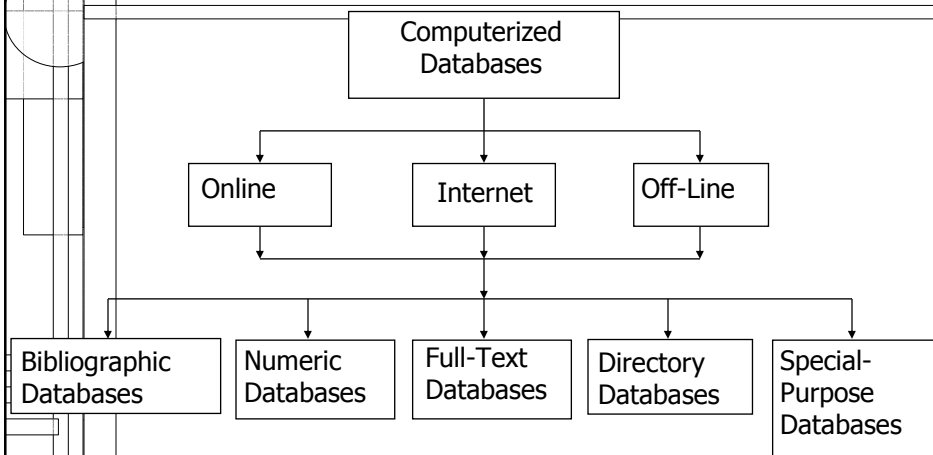


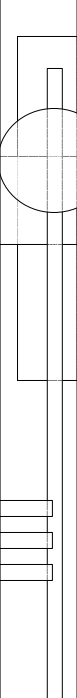
InfoUSA: : Here, There, Everywhere

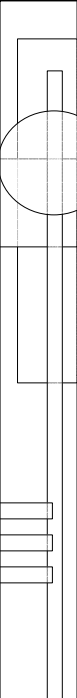
InfoUSA (www.infousa.com) markets subsets of its data in a number of forms, including the professional online services (LEXIS-NEXIS and DIALOG), the general online services (CompuServe and Microsoft Network), the Internet (look-ups), and on CD-ROM. The underlying database on which all these products are based contains information on 113 million residential listings and 14 million business listings, as of 2003. These are verified with over 16 million phone calls annually. The products derived from these databases include sales leads, mailing lists, business directories, mapping products, and also delivery of data on the Internet.



A Classification of Computerized Databases



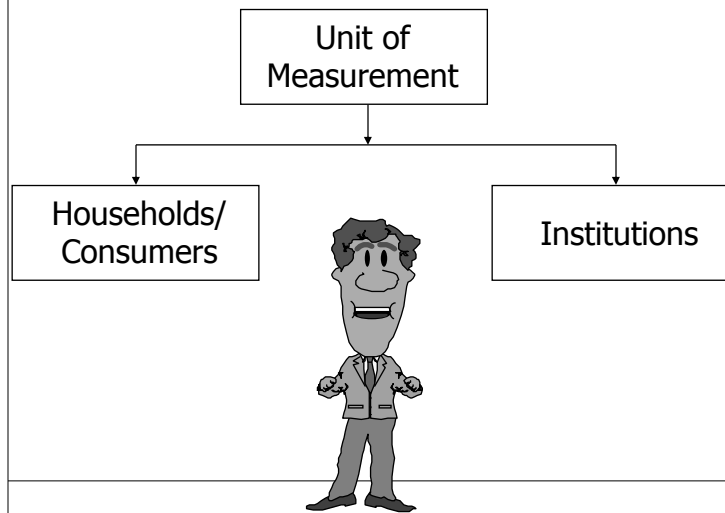
<h2>Published External Secondary Sources</h2>	
	<p>Guides</p> <ul style="list-style-type: none">• An excellent source of standard or recurring information• Helpful in identifying other important sources of directories, trade associations, and trade publications• One of the first sources a researcher should consult <p>Directories</p> <ul style="list-style-type: none">• Helpful for identifying individuals or organizations that collect specific data• Examples: Consultants and Consulting Organizations Directory, Encyclopedia of Associations, FINDEX: The Directory of Market Research Reports, Studies and Surveys, and Research Services Directory <p>Indices</p> <ul style="list-style-type: none">• Helpful in locating information on a particular topic in several different publications

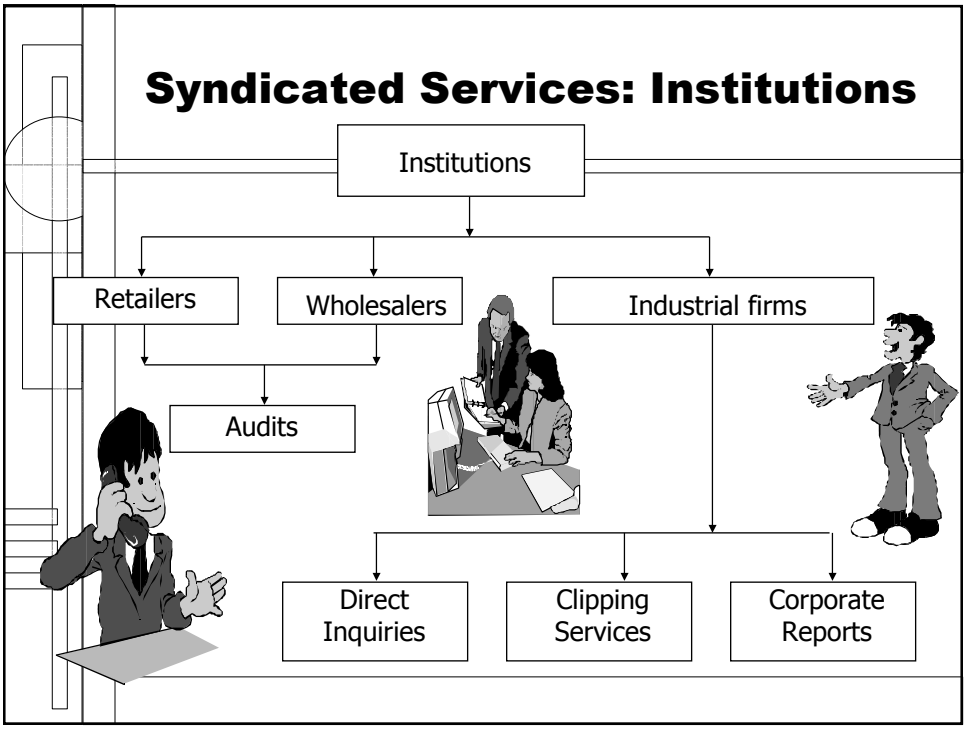
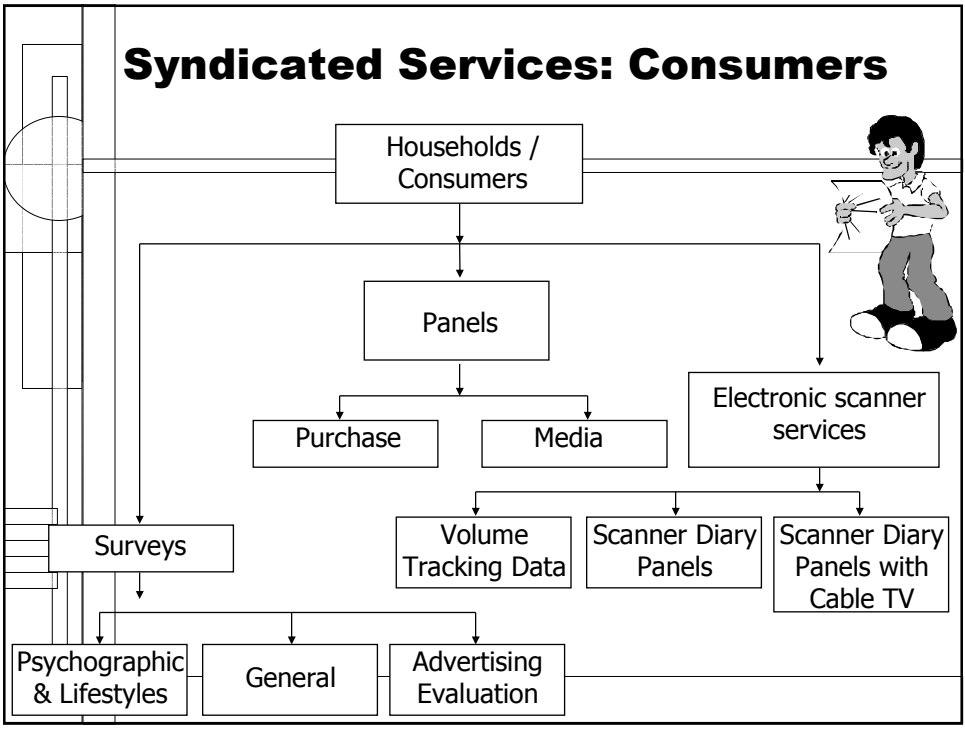
<h2>Classification of Computerized Databases</h2>	
	<ul style="list-style-type: none">• Bibliographic databases are composed of citations to articles.• Numeric databases contain numerical and statistical information.• Full-text databases contain the complete text of the source documents comprising the database.• Directory databases provide information on individuals, organizations, and services.• Special-purpose databases provide specialized information.

Syndicated Services

- Companies that collect and sell common pools of data of known commercial value designed to serve a number of clients.
- Syndicated sources can be classified based on the unit of measurement (households/consumers or institutions).
- Household/consumer data may be obtained from surveys, diary panels, or electronic scanner services.
- Institutional data may be obtained from retailers, wholesalers, or industrial firms.

A Classification of Syndicated Services





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
Purchase Panels	Households provide specific information regularly over an extended period of time; respondent asked to record specific behaviors as they occur	Recorded purchase behavior 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
Media Panels	Electronic devices automatically recording behavior, supplemented by a diary	Same as purchase panel	Same as purchase panel	Establishing advertising rates; selecting media program or air time; establishing viewer profiles

Overview of Syndicated Services

Type	Characteristics	Advantages	Disadvantages
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
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

Overview of Syndicated Services			
Characteristics	Advantages	Disadvantages	Uses
Verification of product movement by examining physical records or performing inventory analysis	Relatively precise information at the 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, analyzing distribution patterns; tracking of new products
Data banks on industrial establishments created through direct inquiries of companies, clipping services, and corporate reports	Important source of information on 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

Single-Source Data	
<p>Single-source data provide integrated information on household variables, including media consumption and purchases, and marketing variables, such as product sales, price, advertising, promotion, and in-store marketing effort.</p> <ul style="list-style-type: none"> • Recruit a test panel of households and meter each home's TV sets. • Survey households periodically on what they read. • Grocery purchases are tracked by UPC scanners. • Track retail data, such as sales, advertising, and promotion. 	

The NYT on the Web: A New Way To Target Customers

To handle alternate forms of interaction and updates, The New York Times created a separate unit, The New York Times Electronic Media Co. *The New York Times* on the Web (www.nytimes.com) has drawn over 10 million registrants as of 2003. The database contains demographic information, such as age, gender, income, and zip code, that ties to an e-mail address for each of the members. This new database marketing system can identify and customize user groups, target Web messages to specific segments of the population, and adjust the message based on audience reaction. It can also increase targeting opportunities through third-party data or additional information supplied by the user.

The NYT on the Web: A New Way To Target Customers

For example, the database enables an automobile firm to emphasize safety to older customers, luxury to affluent ones, and roominess to families. The system is set up so that near real-time data can be received from the Web that indicates how well ads are performing relative to age, gender, and income characteristics. Thus, this system allows a firm to maintain up-to-date information on audiences in order to position its products effectively.

