Teaching Database Research Methods in Marketing Research Classes: A View from the Trenches

Joseph L. Orsini
California State University, Sacramento, CA, USA

orsinij@csus.edu

Abstract

The opportunities provided by the continuous growth in computer storage capacity have led to the development of methods to investigate large databases. However, while industry is heavily involved in research using databases, primarily for marketing purposes, the university academic component of the marketing research discipline has not yet caught up with industry practice. This study discusses database research methods, along with the impact on the need for the teaching of business and marketing research using databases and their methodologies. Suggestions made, include team teaching and the formation of a university-wide Informing Sciences Department. This organization structure would be analogous to that currently found in most universities: a separate statistics department, while utilizing disciplinary specialists to teach statistical applications within the various disciplines.

Keywords: database, research, teaching, data mining, organization, university, marketing

Background

Overview

The marketing discipline has long considered itself to be the function that provides the interface between the organization and the customer. The "marketing concept", for example, states that an organization "should seek to meet its customers' needs and wants as it strives to achieve its own goals" (Churchill and Peter 1998, p. 12). As such, researching the needs and behavior of the customer, both business and consumer, is an integral part of the discipline. For example, about 40% of the *Marketing News'* featured special sections are related to research methods (*Marketing News* 2001). The *Marketing News* is the publication that goes to all members of the American Marketing Association, the professional organization representing the marketing discipline.

While university faculty are being admonished to "stay in the front of technological change and take a true leadership role in helping students develop business skills necessary for success" (Smart, et al. 1999, p. 206), and to "enhance the relevance of research inquiries" (Day and Montgomery 1999, p. 12), very little is found to accomplish this "true leadership role" relative to teaching Database Research (DR). The purpose of this paper is to discus the evolving utilization of DR methods in marketing and other business disciplines; to indicate barriers to teaching these methods; and to focus on the difficulty of organizing classroom pedagogy in business-related research applications, including the impact on faculty preparedness and composition. It is hoped that raising this issue will help to motivate educators in information technologies to work with their marketing and other business counterparts to arrive at solutions to the problems indicated.

Evolution of Databases

Changes in computer technology have allowed the accumulation of large quantities of data in digital format, and business (and other) organizations have taken advantage of it. As Berry and Linoff (2000) point out, at the midpoint of the 20th century no company had more than the equivalent of 30 or 40 megabytes of data (in paper format) in its ledgers, books and file cabinets. By the end of the 20th century, however, things have dramatically changed; UPS, for example, has 17 million megabytes in the database of package level detail it uses to track shipments. This is about the same amount of storage as contained in all the books of the Library of Congress, the largest repository of information in the world.

The same technology that allows this huge increase in data storage has also led to the development of new methods to analyze this data, with a substantial impact on research efforts. Hair, et al. (2000) contend that the second half of the 20th Century has seen two shifts in the "fundamental character of data analysis" (p. 671). The first was the mid-1970s, when mainframe and (later) personal computers allowed multivariable methods to gain widespread acceptance and use. The second was in the 1990s, with the development of large-scale databases, and the development of new methods to analyze the "information avalanche".

Database Analytical Tools

The following is a brief listing of analytical tools that may be applied after a suitable database has been prepared.

Standard Statistics

A sample of a database may be obtained and inserted into spreadsheet format; therefore, all of the "traditional" statistical techniques are potentially available for use (Hair, et al., 1998). Since there is access to the entire "population", very large samples, to the limits of the capacity of the statistical package employed, may be obtained. As with all samples, however, the existence of very small "segments" may still be too small to yield useful information, even though they may be profitable in the marketplace. Further, the concept of statistical significance becomes less useful with large samples, as findings are increasingly likely to be "significant", even if they are of little managerial importance.

Queries

Queries are investigative questions relating to specific items in the database, and may be in the form of Structured Query Language (SQL) for use in On Line Analytical Processing (OLAP) methods (Berry and Linoff 2000). A simple query may look at the entire database for questions such as "how many times has Brand X been purchased concurrently with Brand Y?" A more complex query may first require the formation of a new and smaller dataset consisting only of the variables of interest, thus allowing faster examination of relationships, and the use of a variety of tools to examine the relationships.

Data Visualization

Methods of displaying data for visual review have been found to be useful, and have been increasingly developed for assisting managerial decisions (Hair, et al., 1998). Visualization methods range from simple graphs and charts to quite complex multidimensional structures, which are combinations of art and mathematics (e.g. Tufte 1983). Shape, color, line and artistic graphics are all used to convey what is hopefully informative material. Query software may use visualization methods to answer the question asked.

Data Mining Tools

Data Mining Tools, as with queries and visualization, are designed to examine the entire database rather than samples of it. Included in these tools are the following methods: Market Basket Analysis, Memory

Based Reasoning, Genetic Algorithms, Cluster Detection, Link Analysis, Decision Trees, and Neural Networks (Berry and Linoff 1997). The tasks performed by these methods may be further classified as Directed methodologies (which are generally referred to in statistical analysis as dependence methods), and Undirected methodologies (interdependence methods). The former include classification, estimation and prediction tasks, and the latter include affinity grouping or association rules, clustering, and description (Berry and Linoff 2000). Each of the data mining methodologies indicated above may be useful in performing both undirected and directed tasks in the search for small but profitable segments which may be found when examining the entire database.

For the purposes of this paper, these four analytical tools will all be included in the term "Database Research" (DR). It should be noted that one of the current definitions of Data Mining, the "extracting meaningful patterns and rules from large quantities of information" (Berry and Linoff 2000) could just have easily been used. However, in order to differentiate between data mining tools and the broader research of databases, the DR term will be used in this discussion.

A review of various professional books on DR applications (e.g. Berry and Linoff 2000, Groth 2000, Westphal and Blaxton 1998) indicates that the fundamental skills needed for successful DR may be found in several base disciplines. Table 1 below indicates the author's conclusions regarding these skill bases, which vary by method of DR. While the appearance of art in the list may be of some surprise, it can be an important component. Tufte (1997) for example, cites how the lack of an appropriate method of data display led to the Challenger space shuttle disaster.

Table 1. Skill Bases for Database Research Application

Skill Bases	Database Research Method			
	Statistical Analysis	Queries	Data Visualization	Data Mining Techniques
Discipline of Application	X	X	X	X
Mathematics/ Statistics	X			X
Information Technology		X	X	X
Software Engineering				X
Art			X	

Database Analysis in Marketing Research

Database research has become a part of the marketing research field. For example, a survey of marketing research consultants concluded that data mining methods were solidly instituted in business practice (Orsini 2001). While not all research firms have this capacity, the trend was clearly on the increase. Another example is the Marketing News' frequent reference to database marketing, including its January 10, 2002 issue featuring the topic (*Marketing News* 2001).

One of the reasons for the growth of DR is the emergence of Direct Marketing, which relies heavily on the existence of databases; it is defined as "an interactive system of marketing which uses … advertising media to effect a measurable response … at any location" (McDonald 1998). The measurability aspect of the sub-discipline allows a more pervasive application of quantitative research designs to be more readily applied than other areas of marketing, and has become an integral part of the sub-discipline (Stone and Jacobs 2001).

Marketing Research Pedagogy

An examination of several current editions of introductory marketing textbooks indicates that they have begun to include mention of database research. For example, Solomon and Stuart (2000) and Czinkota, et al. (2000) discuss it in conjunction with Direct Marketing and Database Marketing, which is the typical context for this type of discussion. The heavy involvement of Direct Marketing in utilization of DR is an evolutionary outgrowth of their historical form of customer contact. Hughes (1996) notes that the Direct Marketing form of retailing began in the 19th century with catalogs, and has continually adopted new approaches as technology has changed. The existence of large databases has allowed the precise segmentation and targeting of those more likely to make purchases of specific types of products, thus Direct Marketing has been among the leaders of database research methods (McDonald 1998).

The current editions of marketing research textbooks, and business research textbooks, have typically begun to contain discussions of databases and DR in their sections on decision support systems, or secondary data, where their previous editions did not. For example, Churchill (2001) vs. Churchill (1996); Aaker, Kumar and Day (2001) vs. their 1998 edition; Cooper and Schindler (1998) vs. Cooper and Schindler (1995); and Zikmund (2000) vs Zikmund (1997) all have DR discussions in their new editions, but not in the previous ones. However, there is typically not enough discussion included to give an instructor confidence in answering student questions on the topic.

Barriers to Teaching Database Research

Developing the expertise necessary to teach basic Database Research concepts in a marketing research classroom context is increasingly necessary. However, there are some important barriers to developing this capacity, including:

- Lack of marketing literature information
- Complexity of dataset preparation
- Complexity of the investigation methods
- Costs of software

Some of these barriers may be addressed with minimal difficulty; others are less tractable. For a more complete discussion of these issues, see Orsini (2001).

The preparation of data warehouses for database analysis is no simple task. In fact, it can be the most time consuming part of any database research process, and is one of the principal causes of these barriers. As noted by Groth (2000):

"The biggest challenge business analysts face in using data mining is how to extract, integrate, cleanse, and prepare data to solve their most pressing business problems. This issue is a formidable one, and can take the bulk of the time in the data mining process."

Insofar as a data warehouse may not even exist in any given research application, one may have to be developed using Information System methods to merge sets of data from a variety of sources, or query tools to create files from existing databases (Groth 2000).

Data quality is also an important issue, so much so that it is advised to "beware the consultant or data-mining product that does not address (or downplays) how to get the data ready for analysis" (Pettit 2000). In addition to redundant data, and incorrect or inconsistent data, there is an issue of data format. Data from operations tends to be either continuous or categorical (ratio or nominal, in typical marketing research terminology), and some of the data mining methodologies do not handle continuous data; they must be made categorical (ordinal or nominal) to be used. All this is assuming that the existing operations database is sufficient to use without adding any additional data from other sources (e.g. consumer credit card information).

Data issues are thus far more complex in database investigation than they are in the survey-based datasets typically used in marketing research. While databases appropriate for analysis are becoming more common in business data repositories, there still may be a substantial amount of data preparation necessary prior to any investigation using either traditional statistics or the newer data mining tools. The net result is that, even when the investigator has a good grasp of the function of the various database research tools, their industry application will require some additional expertise in the area of Information Systems.

Pedagogical Approaches to Teaching Database Research

While there are several barriers to developing the expertise necessary to teach a database research course, as indicated above, this discussion will address only the impact on faculty composition and preparation. The data complexity issue is typically not mentioned in marketing research texts, which generally confine their discussions to data editing, e.g. Aaker, et al. (2001). This is due to the traditional survey orientation of business and marketing research textbooks, where survey costs typically limit the cases in the survey dataset to under 1,000 responses, and questionnaire length limits it to several dozen variables.

Fortunately, the data complexity aspect of DR for the marketing research instructor is potentially not quite so complex as it is in business application. "Groomed" datasets may be developed for classroom purposes, so that the more technical problems involved could have already been rectified, and only research decision issues be unresolved (e.g. where to define category limits in order to make a continuous variable into a discrete variable). While this may be necessary for some classroom purposes, data preparation is still a concern for industry application, thus should be addressed from a pedagogical standpoint. Further, queries are currently a well known aspect of IT, but not well known in the non-IT business disciplines. Moreover, the costly data mining tools are not yet widely available in academic circles in general.

Approaches to Interdisciplinary Pedagogy

The pervasiveness of IT aspects of many university disciplines has raised the issue of how to avoid the problems associated with this fragmentation, including the formation of a new field: Informing Science. Cohen (1999), for example, suggests either a matrix structure, or a virtual organization consisting independent teaching and research centers, as useful alternatives to the traditional functional organization structure of a university hierarchy.

The project management discipline has addressed the departure from the functional organizational structure for some time (e.g. Gray and Larson 2000). While the hierarchical structure of functional organizations provides "…a strong environment in which individuals can share and keep up with the knowledge

and skills of their particular discipline" (Gido and Clements 1999, p. 149), in the case of organizations charged with performing projects, the insularity of the disciplines can be quite detrimental. A project-focused organization structure, conversely, while maintaining smooth interdisciplinary performance of the project, may not provide its members with colleagues within their discipline necessary to provide a "… source of shared functional expertise and knowledge" (ibid, p. 150).

Matrix organizations are the compromise, and have been utilized in the project management discipline for some time (e.g. Gray and Larson 2000). As a young mechanical engineer in a construction organization many decades ago, the author was routinely involved with other engineering disciplines, under a Project Manager, in the design civil works projects. However, while matrix organizations make sense in a project-involved organization, it is not clear how well they would work in the more routine (non-project) teaching functions of a university.

A form of a matrix organization that is well known to academe is the "team teaching" approach to presenting a particular course. In this concept, two or more instructors, typically with somewhat different areas of expertise, combine their efforts to teach one course. In effect, while under the authority of their "home discipline" they join for the "project": a team-taught course. In the author's experience, this approach tends to be an exception, but is a familiar approach to pedagogy.

Virtual organizations are a more recent phenomenon, and would be worth some in depth consideration. Two forms of virtual organizations found in academe are, first, that of "distance education", which itself is a high-tech form of the old correspondence course (e.g. Dacko 2001). A second approach, found in some private universities, is that of a small permanent administration, operating in rented facilities with independently-contracted faculty who are not permanent organization members. Cohen (1999) does not elaborate on his perspectives in this regard, only devoting a paragraph to its discussion. Advances in technology, and acceptance of adjunct professors in basic technology areas (e.g. "computer literacy" courses), points to the possibility of a hybrid approach (e.g. Marold and Fustus 2001).

An Interdisciplinary Example

Interestingly, there is an example of another pervasive discipline that exists as a topic area in many university departments: statistics. Depending on the university, the highest level of statistical expertise may exist as part of the mathematics department, or may be a separate department. However, statistical applications specialists seem to happily exist in a myriad of other departments in the physical sciences, social sciences, and administrative sciences. The author, for example, is one of several of his colleagues in his department who specializes in teaching marketing and business research, which substantially utilizes statistical analysis.

This example would suggest a possible alternative to the matrix or virtual university structures: that of one centralized department of individuals who specialize in the topic area (e.g. Informing Science), with applications specialists in the various departments, as needed. These application area faculty (e.g. marketing research professors) would also be fundamentally grounded in IT, much as they are now fundamentally grounded in statistics, but go to their Informing Science colleagues when presented with situations requiring greater expertise. How these faculty would acquire this fundamental grounding would have to be thoughtfully considered by those academic institutions offering Ph.D.s, but appears feasible.

Conclusions

The last two decades have seen an explosion in the capacity of business organizations to retrieve and store operational data. When combined with customer data, these massive databases constitute a valuable source of information for marketing purposes. Businesses are increasingly using this information in their marketing efforts, particularly in promotion.

Development of methods of researching these databases has paralleled the expansion of storage capacity and access. While it is possible to obtain samples of the databases and use traditional statistical methods, the newly developed database investigation methods have some unique advantages. In particular, they have the capacity to discover very small segments that may not exist in sufficient sizes in a sample to be uncovered, but which may still be profitable.

While it is to industry's benefit to have graduates from marketing and other disciplines who are aware of, and familiar with, the various database research methods, this has not yet been accomplished in any pervasive fashion. While offering some suggestions, the primary purpose of this article is to encourage the many who are addressing database-related pedagogy to continue their efforts. Those of us "in the trenches" need you to continue your help.

As a minimum, those academic institutions which provide the education and training of those who will teach this topic will need to consider changes in their own curriculums. An additional possibility is that it may be necessary to change the very organization structure of the universities. Given the inertia of bureaucratic organizations, however, the evidence would have to be quite strong to prompt widespread fundamental changes from the traditional university organization model.

This would suggest one feasible approach: initially team teach database research classes, while arranging the gradual acquisition of sufficient IT skills by research faculty in their respective disciplines for those application which are infrequent and require specialized expertise. The author is currently pursuing this approach with the Management Information Science Department within his University's College of Business. The development of an Informing Science Department may also be required, much as there currently exists a Statistics Department, to provide resources for the "non-everyday" applications. They may also provide campus-wide database training for Colleges or other units of a university which do not have an IS discipline within them.

References

Aaker, David A., V. Kumar and George S. Day (2001), *Marketing Research*, New York: John Wiley & Sons, Inc.

Aaker, David A., V. Kumar and George S. Day (1998), *Marketing Research*, New York: John Wiley & Sons, Inc.

Berry, Michael J. A., and Gordon Linoff (1997), *Data Mining Techniques*, New York: John Wiley & Sons, Inc.

Berry, Michael J. A., and Gordon S. Linoff (2000), *Mastering Data Mining*, New York: John Wiley & Sons, Inc.

Churchill, Gilbert A. Jr. (2001), *Basic Marketing Research*, Fort Worth: The Dryden Press.

Churchill, Gilbert A. Jr. and J. Paul Peter (1998), *Marketing: Creating Value for Customers*, Boston: Irwin McGraw-Hill.

Churchill, Gilbert A. Jr. (1996), *Basic Marketing Research*, Fort Worth: The Dryden Press.

Cohen, Eli (1999), "Reconceptualizing Information Systems as a Field of the Transdiscipline Informing Science: From Ugly Duckling to Swan", *Journal of Computing and Information Technology* 7 (3), 213-219.

Cooper, Donald R. and Pamela S. Schindler (1998), *Business Research Methods*, New York: Irwin/McGraw-Hill.

Cooper, Donald R. and Pamela S. Schindler (1995), *Business Research Methods*, New York: Irwin/McGraw-Hill.

Czinkota, Michael R., et al., (2000), *Marketing: Best Practices*, Fort Worth: The Dryden Press.

Dacko, Scott G. (2001), "Narrowing Skill Development Gaps in Marketing and MBA Programs: The Role of Distance Technologies for Distance Learning", *Journal of Marketing Education*, 23 (3), p. 228-239.

Day, George S. and David B. Montgomery (1999), "Charting New Directions for Marketing", *Journal of Marketing* 62, Special Issue, p. 3-13.

Gido, Jack and James P. Clements (1999), *Successful Project Management*, Cincinnati: South-Western College Publishing.

Gray, Clifford F. and Erik Larson (2000), *Project Management*, New York: Irwin/McGraw-Hill.

Groth, Robert (2000), *Data Mining*, Upper Saddle River, N.J.: Prentice-Hall.

Teaching Database Research Methods

Hair, Joseph F. Jr., Ralph E. Anderson, Ronald L. Tatham, and William C. Black (1998), *Multivariate Data Analysis*, Upper Saddle River, N.J.: Prentice-Hall.

Hair, Joseph F. Jr., Robert P. Bush and David J. Ortinau (2000), *Marketing Research*, Boston: Irwin McGraw-Hill.

Hughes, Arthur M. (1996), *The Complete Database Marketer*, New York: McGraw-Hill.

Marketing News (2001), November 5, 16.

Marold, Kathryn A. and Janos T. Fustos (2001), "Hybrid Delivery of Computer Literacy Requirements for Undergraduate Business Schools: A New Model for a New Millennium", *Proceedings*, Informing Science 2001 Conference, 359-364.

McDonald, William J. (1998), *Direct Marketing: An Integrated Approach*, Boston: Irwin McGraw-Hill.

Orsini, Joseph L. (2001), "Teaching Data Mining and Other Database Research Methods in Marketing Research Classes: An Interim Assessment", *Proceedings*, Marketing Educators' Association, 121-125.

Pettit, Raymond C. (2000), "Data mining: race for mission-critical info", *Marketing News*, January 3, p. 18

Smart, Denise T., Craig A. Kelley and Jeffery S. Conant (1999), "Marketing Education in the Year 2000: Changes Observed and Challenges Anticipated", Journal of Marketing Education, 21,3, December, 206-216.

Solomon, Michael R. and Elnora W. Stuart (2000), *Marketing*, Upper Saddle River, N.J.: Prentice-Hall.

Stone, Bob, and Ron Jacobs (2001), Successful Direct Marketing Methods, 7th Edition, New York: McGraw-Hill.

Tufte, Edward R. (1997), *Visual Explanations*, Cheshire: Graphics Press.

Tufte, Edward R. (1983), *The Visual Display of Quantitative Information*, Cheshire: Graphics Press.

Westphal, Christopher and Teresa Blaxton (1998), *Data Mining Solutions*, New York: John Wiley and Sons.

Zikmund, William G. (2000), *Business Research Methods*, Fort Worth: The Dryden Press.

Zikmund, William G. (1997), *Business Research Methods*, Fort Worth: The Dryden Press.

Biography

Joseph L. Orsini, Ph.D. is a Professor of Management and Marketing at California State University, Sacramento (CSUS) for eighteen years, Dr. Orsini has taught graduate and undergraduate classes in Business Research and Marketing Research for over twenty years. He is currently active in research consulting in both the private and public sectors, including model building on database samples.

Dr. Orsini began his career as an engineer. He obtained a Bachelor's degree in Mechanical Engineering (Princeton University) and a Master's degree in Engineering Management (University of Alaska) prior to his Ph.D. in Management (UCLA). He also has a Bachelor's degree in Government, and an M.B.A.