View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by DigitalCommons@Florida International University

Hospitality Review

Volume 18
Issue 1 *Hospitality Review Volume 18/Issue 1*

Article 2

1-1-2000

Re-engineering Technology: Date Mining for Clubs

Michael L. Kasavana
Michigan State University, shbsirc@msu.edu

Bonnie J. Knutson
Michigan State University, shbsirc@msu.edu

Follow this and additional works at: http://digitalcommons.fiu.edu/hospitalityreview

Recommended Citation
Kasavana, Michael L. and Knutson, Bonnie J. (2000) "Re-engineering Technology: Date Mining for Clubs," *Hospitality Review*: Vol. 18: Iss. 1, Article 2.
Available at: http://digitalcommons.fiu.edu/hospitalityreview/vol18/iss1/2

This work is brought to you for free and open access by FIU Digital Commons. It has been accepted for inclusion in Hospitality Review by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

Re-engineering Technology: Date Mining for Clubs

Abstract
With increasing competition and more demanding members, clubs need a tool to help them belter attract and retain members and predict their behavior. Data mining is such a tool. This article presents an overview of how data warehousing, data marting, and data mining can provide the foundation on which clubs can build strategies to outsmart competitors, build Ioyalty identify new members, and lower costs.

This article is available in Hospitality Review: http://digitalcommons.fiu.edu/hospitalityreview/vol18/iss1/2

Re-engineering technology: Data mining for clubs

by Michael L. Kasavana and Bonnie J. Knutson

With increasing competition and more demanding members, clubs need a tool to help them better attract and retain members and predict their behavior. Data mining is such a tool. This article presents an overview of how data warehousing, data marting, and data mining can provide the foundation on which clubs can build strategies to outsmart competitors, build loyalty, identify new members, and lower costs.

At a time when really knowing members is so crucial to running a successful club, data mining should be at the forefront of club managers' thinking about the future. Clubs have long known that they need to understand the wants and needs of their members in order to keep them happy. In fact, ask club managers what their biggest marketing concern is and, in general, they will say gaining and keeping members.

With increasing competition and more demanding memberships, managers need a tool to help them better attract and retain members, and predict their behavior. Data mining is just such a tool. Having the automated data warehousing and data mining abilities saves the club money and time by having the ability to readily capture and manipulate member information. It also gives clubs the opportunity to be more attentive to their pricing structure and to refine their marketing efforts.

Data mining is simply the process used to get useful information out of the data that the club has stored about its members. "It stands at the intersection of statistics, database technology, and important business concerns."[1] The concept of data mining isn't new for clubs. Historically, clubs have captured and stored information about what members spend (transactions), what they do (activities), and who they are (demographics). Most managers can tell you the profile of their memberships, when the busy meal periods are, and the attendance trends for special events.

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

What is new, however, is a club's ability to quickly extract useful information from large complex membership databases drawn together from a variety of sources. This is fast becoming the next major technology requirement for clubs. To understand how this new technology can make club marketing more effective, managers have to understand the interface among data warehousing, data marting, data mining, and software. Together, they form a powerful basis for re-energizing a club's information system.

Warehouses store data

A data warehouse is simply a place where large amounts of club data are stored. These data can come from within or from outside the club. From a technical perspective, it is a member-oriented collection of raw facts used to support managerial decision-making. A data warehouse can serve as the central focal point of a club information system and provide information unavailable through traditional database search methods. Unlike the legacy of corporate reporting systems so common in club systems today, a data warehouse is created to serve as a facility for integrating data from a variety of internal (club operations) and external (environmental factors) applications. For example, a club warehouse could integrate members' demographic profiles with how often they use the club for various activities, with their spending patterns, with member surveys, and with area census information.

By design, a data warehouse is composed of discrete subject-oriented fields. This means that each member has a unique data field in which all his or her information is stored. By contrast, traditional database systems are organized around club operations, not members. A typical automated club management system may include discrete operational modules like member demographics, accounts receivable, back office accounting, and clubhouse operations. These modules, however, cannot "talk" to each other; they are independent transaction-oriented databases. "Hidden in many of these databases is a wealth of valuable [member] data – data than can be used to enhance marketing programs and add value to [member] service."[2] A club data warehouse structure is designed to hold data organized by member activities, spending patterns, and bundled purchases. Because the member is the focal point of a data warehouse structure, all information can "talk" to all other information in the warehouse.

The primary purpose of a club's data warehouse is to provide a common source for decision support data. A means for making this happen lies in how data are entered. Because a majority of data placed in the warehouse comes from the club's operational applications, there must be a key variable that can link the data together. That link is usually the member

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

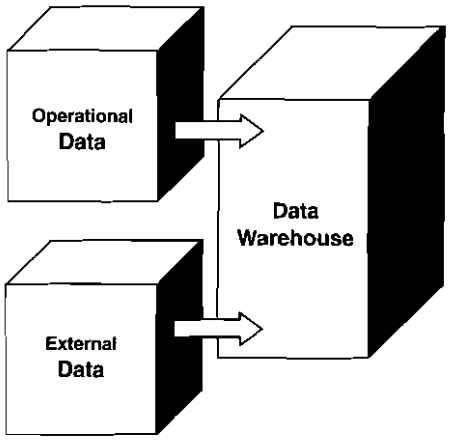
number. The format must also be normalized, meaning that any given data item – Mr. Smith's July account balance or the number of sweaters purchased in the pro shop in September – is entered and stored only once. In addition, the data must be entered using standardized coding procedures and variable identification. This is to ensure uniformity throughout the database management system, thus allowing club management to better understand the content and structure of stored data. It likewise allows easier analysis. Take the example of the member's profile. Coding standards would call for all profiles to be entered the using the same format, i.e., last name, first name, member number, member status, number of children living at home, etc. It is for this reason that data warehousing takes precautions to ensure a high level of conformity and consistency throughout the system. This is often called data cleansing.

Data are centralized

The primary value of a data warehouse is its capacity to organize, store, and process vast amounts of diverse internal and external data within one centralized location. In addition to allowing managers to look at members within the club, it can capture and classify external data so that meaningful comparisons to internal activities can be made. (See Exhibit 1.) For example, suppose a club was preparing a membership drive. It could import external census data from the community and compare it to member profiles. In this way, the club could identify geographic pockets of potential recruits, based on homogenous characteristics. In another case, it could import data about regional or national membership trends and compare them to what is happening within its club.

In a typical data warehouse, there are normally four "storage rooms." They contain historic

Exhibit 1
Data warehousing



FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

data, current data, lightly summarized data, and highly summarized data. Historic data tends to be accessed infrequently and therefore may even be stored on disk or tape. An example might be member accounts for the past decade. Current data reflect the most recent member information and tend to be of greatest value for club marketing. Examples include the number and types of events the members attend, how often they participate in various activities, or which purchases they may have made in the pro shop. It also includes the members' profiles, i.e., membership type and tenure, age, marital status, number of children, and geographic proximity to the club. Lightly summarized data are data gleaned from current data and used to help make a management decision. Weekly aggregation of food and beverage covers broken out by family type might be used to affect menu offerings, pricing, or even board policy on dress codes. Highly summarized data is compact; it is summarized from all other data forms. Here a club might synthesize weekly or monthly food and beverage statements into a single yearly report. Based upon the volume of data that can be stored, ease of access, and the cost of storage media, the operational costs of a data warehouse can be a cost-effective marketing tool for clubs.

Club data enters a warehouse as current data and remains there until it is summarized, archived, or deleted. A data warehouse transforms current data into historic data based upon a chronological timetable. The summarization process uses the detail of current data to compute lightly summarized data that eventually progress to highly summarized data. As a general rule, the more summarized the data, the more likely the club will use it. The obvious explanation of this phenomenon is that highly summarized data should be more easily accessible and content rich.

Data marts are subsets

While the data warehouse is the center of the club's information system, each operational unit contributing data to the warehouse is referred to as a data mart. In essence, the data mart is a subset of the data warehouse and contains the information most relevant to that operational unit, i.e., the pro shop. In general, there are two types of data mart users: "farmers" and "explorers." A farmer is described as someone who knows what data he or she wants and regularly and predictably goes to the same place to find it. In the case of the pro shop, the golf professional may review past history of lessons or rounds played each week. An explorer is an individual who does not know what data is really needed and follows a random, sporadic data hunt. This same golf professional may be looking to increase clothing sales on Tuesdays and Wednesdays. In this case, he or

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

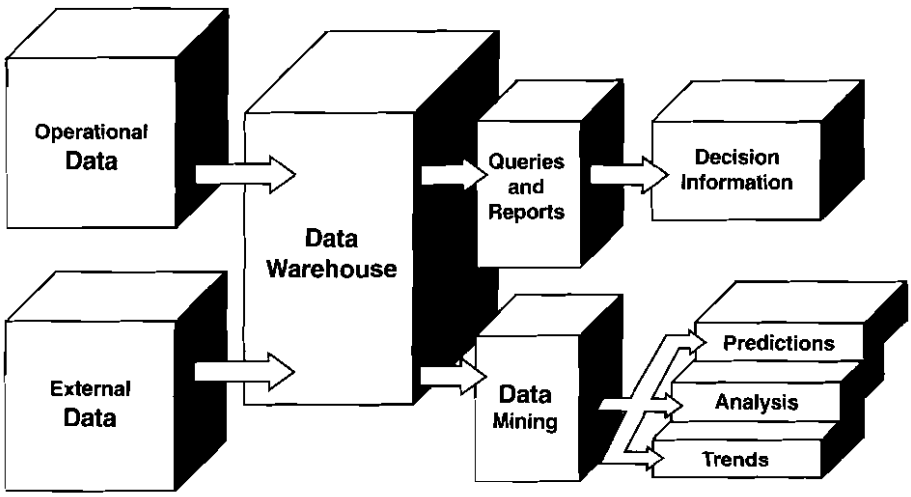
she may explore for information that will identify who is usually playing golf at the club on those days, what their activities are, what they have a tendency to buy, and how he could directly promote to them. By virtue of the fact that a data mart contains fundamental data for the warehouse, it is a natural extension and companion to the data warehouse.

When an operating department – such as membership or accounting -- has its own data mart, it can customize and capture data as it flows into or out of the club's data warehouse. The data in the data mart do not need to service the entire club like a data warehouse, only the department. Therefore, it can be streamlined, summarized, and structured as appropriate. With a data mart, the user can conduct data analyses with little consideration of the impact on resource utilization elsewhere in the data warehouse scheme. It is for these organizational, technical, and economic reasons that a data mart is an attractive outgrowth of a data warehouse. In addition, the department can select data mining software that is tailored to its specific needs. (See Exhibit 2.)

Mining extracts information

Data mining is the process of extracting useful information from a data warehouse or data mart. Clubs can use data mining as the foundation for strategies that help them increase member retention, identify new members, gain a competitive advantage, and lower costs. Simply stated, data mining involves the discovery of new information through an identification and understanding of trends, patterns, and variable correlation.

Exhibit 2
Data warehouses and data mining



FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

The process is intended to turn data into information and information into insight.

Data mining uses mathematical algorithms to search for patterns within large volumes of club data to reach information goals. These include regression and tree models designed to find information buried deep within the club's data warehouse or data mart. While the goals of data mining can be as varied as clubs themselves, they generally fall into four categories:

- Increasing member satisfaction, perceived value, loyalty, and ultimately member retention
- Attracting new members by identifying profiles of most likely candidates
- Cross-selling club services to current members
- Upselling current members to the next level

An example of a question a club might seek to answer in each category follows:

- **Member retention:** How likely is it that "empty nest" members will remain members once their children have grown and moved out from the home?
- **Attracting new members:** What is the ideal profile of a new member, given profiles of the club's "best" members?
- **Cross-selling:** What is the likelihood that a member taking tennis lessons will then buy new equipment at the pro shop?
- **Upselling:** How likely is it that members will upgrade the food service at their business meetings from "coffee break" to lunch?

There are three basic methods to data mining: classification, clustering, and visualization. In general, each of these methods involves a four-state process:

- Identify the right club business problem
- Transform and analyze the data
- Act on the results
- Measure the results[4]

Having a specific question or goal relative to some member group helps define the data mining process. To illustrate the four-stage process, consider the following classification question. Suppose the membership committee wants to know why some members frequent the club more often than others. Ultimately, they want to predict which members can be influenced – and how to best influence them – to increase frequency and thereby promote loyalty and retention. The club can construct a data-mining model based on historical data of two member classifications, active members versus inactive members. A solid model may help man-

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

agement better understand member behavior and therefore enhance member frequency.

Current attempts to find information buried deep within a data warehouse or data mart generally fall into two categories. The first is termed "drill-down" analysis and the second involves the use of member-based "pattern discovery" tools, which often operate independent of a database. Relational database vendors in partnership with online analytical processing front-end tool vendors, such as Oracle, provide drill-down analysis capabilities.

Software provides analysis

Often data mining software vendors feature an Oracle-based information system designed to build and maintain centralized records, i.e., member profiles. Non-database software vendors usually provide member-based pattern discovery tools. These tools are generally workstation based, such as Angoss or IDIS. The tools operate independent of databases and have a self-contained user-interface.

Drill-down data mining allows users to get answers from the general questions posed and then drill down into the underlying data to extract more detail. Recently, many data-mining vendors extended drill-down products to provide additional statistical functions to enhance the analysis capabilities against relational data. Fundamentally, this approach allows users to test hypotheses concerning data, rather than providing a mechanism that automatically uncovers patterns in data.

Client-based pattern discovery tools, on the other hand, automatically detect patterns in data, assuming the data is in the proper format. These tools generally run on a workstation client connected to a database server. Data is extracted from the database, conditioned by the user, and then fed to the pattern discovery tool. After the pattern discovery process, the patterns are presented to the user through a graphical user interface. These tools are generally not scaleable and do not handle enterprise-seized problems. Additionally, they are only loosely integrated into the data warehouse environment.

Data mining software tools have emerged that allow a club to address simplistic levels of data mining, including relational databases, multidimensional analysis tools and statistical analysis packages. Some of these tools, however, allow clubs to perform the more complex levels of data mining, which enable them to find patterns hidden in their large warehouse databases, and make faster, more accurate proactive business decisions. Many software packages employ a drill-down technique that enables analysis of relational databases or standalone automated pattern recognition. The more sophisticated data mining techniques integrate automated pattern discovery and

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

recognition tools with widely accepted interface standards into a comprehensive, scaleable, business prediction system.

An integrated suite of data mining tools is designed to provide managers with automated analysis capabilities to apply a data warehouse in a manner that is natural for the club's natural decision-making process. Data mining doesn't just drill down through data; instead, it continuously and automatically sifts through large amounts of data to find useful information. The key question is: What information can be gained that is not already known? Relational databases, typically built into club systems, are oriented around normalized relational tables with semantics inherent in their structure, whereas pattern discovery tools operate on flat files consisting of a series of independent variables. The discovery tools seek to find relations between data. Data mining is intended to produce sufficient information to guide the user to making more intelligent and insightful queries.

Tools become commonplace

The concepts of data warehousing and data mining were once consigned to researchers with the ability to use the complex technology. With continuing advances in hardware and software, these powerful tools are fast becoming commonplace in the hospitality industry in general, and clubs in particular. Club management system vendors (e.g., Club Data Corporation, Club Systems Group, and Club Technology) recently began supplying the software toolboxes necessary for effective data mining.

Previously equipped with the capability to generate pre-programmed system reports, new application software facilitates dynamic (on-the-fly) database query reporting to assist in market strategy refinement and enhanced product planning. For example, consider new member recruitment and bundled item sales. Using data mining to determine a comprehensive set of current club member characteristics should help define a highly accurate potential member profile. Similarly, mining wine sales may indicate a pattern of purchase corresponding to menu item entrees that can lead to the logical bundling of food and beverage items into an enticing new sales package.[5]

The extension of data mining capabilities to include automated, complex discovery tools for vast quantities of data, combined with the ability to evaluate proposed business strategies, will transform the way clubs do business. But there must be a note of caution here. While data mining has the ability to uncover data patterns, it still requires human skill to interpret the results accurately and develop new club opportunities. In other words, the manager's insights and inherent knowledge of his or her club cannot be replaced by technology. As futurist John Naisbitt quipped: "Hi tech, hi touch."[6]

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

The beauty of data warehousing and mining is that once a club has the information in-house, on most systems, it will stay there until it is purged. Thus, it can give the club an historical perspective that's not always available with hard copy information. "If you mine your data properly, you can find little pockets of [opportunity] when you can augment your revenue with special promotions... That way you don't have to market to everyone. You can increase your hits by marketing to specific [member segments]."[7] But data mining isn't just about getting any member to come to the club more often; it is also about getting the best members to increase their frequency, i.e., those members who generate the most revenue for the club.

Factors influence adoption

There are three major factors influencing the rate of adopting data mining software within the club industry: decreasing cost of computing technology, availability of powerful data mining techniques, and increased knowledge of club management. Recently NCR and Microsoft announced an agreement to link database technologies and jointly provide data warehousing and electronic commerce for hospitality companies, including clubs. This alignment integrates NCR's Teradata relational database management systems with Microsoft's structured query language server for data marting, Internet commerce, guest relationship management, enterprise resource planning, and other applications. Data mining is on the verge of exploding in the club industry. The future indeed looks bright.

Appendix A
Data warehousing for clubs

Sample club management system suppliers

ACES SOFTWARE: www.acessoftware.com

CLUB DATA CORPORATION: www.clubdata.com

CLUB SYSTEM GROUP: www.dmsnet.com

CLUB TECHNOLOGY: www.clubtechnology.com

CSI SOFTWARE: www.csisoftware.com

SMYTH SYSTEMS: INC. www.smythsystems.com

Data mining web sites:

www.datamindcorp.com
www.datamining.com
www.data-mine.com
www.data-warehouse.com

Sample data mining software suppliers

www.apertus.com
www.ibm.com
www.microsoft.com
www.oracle.com
www.neovista.com
www.sargenttech.com
www.sybase.com

Appendix B
Data mining concepts

Relationship marketing: discovering product, promotion, and preferences of member base.

Loyalty analysis: analyzing member information to reveal potential impact of promotions on member frequency.

Promotions forecasting: anticipating the movement of promoted items against traditional product lines.

Seasonal forecasting: predicting movement of seasonal business to optimize

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.

revenue management and related issues.

Revenue management: identification of demand based pricing strategies to balance activities and transactions against revenues.

New product introduction: forecasting sell-through rate of an introductory product or event or activity against database content.

Appendix C
Data mining concepts

Client: a workstation or terminal that is connected to a network file server that is shared by several other clients (i.e., client/server network).

Data mart: a database, or collection of databases, focused on a particular subject, department, or operational function; tends to be a subset of a data warehouse.

Data mining: process of using statistical techniques to discover subtle relationships between data items and the construction of predictive models based on them.

Data warehouse: collection of disparate data designed to support managerial decision making through ad hoc data query.

Database management system (BMS): used to store, process, extract, and manage data in a systematic way.

Decision support system (DSS): application for analyzing large quantities of data and performing a wide variety of calculations and projections.

Drill down: data query routine that involves continuous researching of a data element to its most granular form (i.e. researching a financial statement entry to its component schedule to the individual transactions it incorporates).

Query: a stylized question requesting records meeting specified criteria from a database; rules for constructing queries conform to standardized query language.

Relational database management (RDBMS): used to store, process, and manage data arranged in relational tables. Commonly used for transaction processing and data warehousing.

Scalable: refers to anything whose size can be changed including hardware or software systems and their ability to adapt to increasing demands (e.g. from few network nodes to many or from storing small to large quantities of data).

Structured query language (SQL): standard data structuring and access language used in the formulation of queries for database management systems.

References

[1] Raymond C. Petit, "Data Mining: Race for Mission-Critical Info," *Marketing News* 31 no. 1 (January 3, 2000): 18.

[2] Drury Jenkins, "Customer Relationship Management and the Data Warehouse," *Call Center Solutions* 18 no. 2 (August 1999): 88-92.

[3] Beth Davis, "Data Mining Transformed," *InformationWeek,* (September 6, 1999): 86-88; James J. Vanecko and Andrew W. Russo, "Data Mining and Modeling as a Marketing Activity," *Direct Marketing* 62, no. 5 (September 1999): 52-55.

[4] Gordon Linoff, "Data Mining," *Inform* 13 no. 9 (November/December 1999): 18-24.

[5] Davis.

[6] John Naisbitt, *Megatrends* (New York: Warner Books, 1982).

[7] Frank H. Androka, Jr., "Maze of Information," *Hotel and Motel Management* 214, no. 11 (June 14, 1999): 44, 72.

***Michael L. Kasavana** is the NAMA professor and **Bonnie J. Knutson** professor of marketing in the School of Hospitality Business at the University of Michigan.*

FIU Hospitality Review, Volume 18, Number 1, 2000
Contents © 2000 by FIU Hospitality Review The reproduction of any artwork, editorial or other material is expressly prohib~tedw ithout written permission fmm the publisher.