Virginia Society for Technology in Education
VSTE
www.vste.org

Data Warehousing and Decision Making

by William Flaherty

Can you show standardized test results by certain groups of students for multiple years? Can you provide each teacher with past standardized test results for each student in his/her classes? Can you provide each student with a realistic prediction of what GPA and SAT is needed for admission to certain colleges? If your district cannot allow all administrators to quickly and easily produce the answers to these questions, you might be interested in a new technology now available to schools that can do this and much, much more! This technology is called data warehousing.

An Introduction to Data Warehousing

Hanover County Public Schools (HCPS) has created an Instructional Decision Support System, IDDS. This system combines information from many disparate data systems and incorporates them into a common repository that is available to administrators throughout the school district. The student information system is linked to attendance, discipline, grade point average, transcripts, standardized tests and student surveys.

The data warehousing movement began with a book by Bill Inmon in 1992, *Building The Data Warehouse*. Inmon's definition of a data warehouse is "A subject-oriented, integrated, time variant, non-volatile collection of data in support of management's decision-making process." The last part of this definition should be the most appealing to school administrators today, all of whom are in the midst of high stakes testing and school accountability. The next wave of reform will certainly be data based decision-making. This is a natural evolution that will assist administrators in making decisions for the right reasons. Until now, school leadership has not had the necessary tools to analyze the vast amount of information that is generated from one set of standardized tests. This task grows exponentially when one attempts to examine multiple years of data, multiple tests, specific groups of children, special programs, attendance, discipline, teacher training, etc. The proper data warehouse solution can allow one to easily sift through all this data with a few clicks of a mouse.

A view from the balcony of a data warehouse reveals different data sources (main student database, state standardized tests, Stanford tests, SAT, ACT, district generated data) feeding into a robust database housed on a powerful server that is accessible to administrators throughout the district. The final key component is a user-friendly interface that will allow non-technical people to easily access the information from their computer. Due to recent advances, this solution now has a price point that is affordable for most school districts.

The VSTE Journal is published by the Virginia Society for Technology in Education. Permission is granted to copy and distribute single articles from this publication for non-profit use with copyright notice.

Contents copyright © 2002, VSTE All rights reserved.

Data Warehousing, continued

Data Warehousing in Practice

Hanover County Public Schools' quest began in the summer of 1999 with a simple School Board goal: "Promote instructional programming in the following areas, among others: . . .implications of longitudinal assessment of student achievement." This mandate lead to the formation of a committee chaired by the school superintendent, Stewart Roberson, which examined means of attaining this goal. The composition of this committee consisted of a school board member and members from both the instructional and technology organizations within the school district. David Papenfuse, Assistant Superintendent of Finance and Technology, and Lloyd Jones, Assistant Superintendent of Instruction, headed the team from technology and instruction respectively. This make-up was of vital importance to the success of the project. The data warehouse was operational by the summer of 2000. Virtually all administrators in HCPS's nineteen schools and central office were trained by December 1, 2000. HCPS had the in-house expertise to maintain the warehouse and support the user base.

HCPS took very deliberate steps to accomplish this in a short time frame. These steps included evaluating database software for the warehouse, selecting the database software, training of technical staff for both the database and the data warehousing features of the database, evaluating the end user software, selecting software; training of technical staff in installation, setting up and securing end user software, establishing focus groups with principals, surveying potential users, introducing the Instruction Decision Support System at the annual administrative retreat, and training of administrators. Although all steps were necessary, it was particularly important to involve administrators in the process and to incorporate their suggestions as we developed this information delivery system. HCPS administrators were anxious to have a solution that would aid in the arduous task of examining our state standardized tests.

School districts receive a "composite score" and "individual skills tested scores" for each student. An example of this would be the composite score for the fifth grade math test and its skills of computation and estimation; measurement and geometry; probability and statistics; patterns, functions and algebra; and number and number sense scores. With the new data warehouse, a principal can generate a report in minutes that will produce averages, minimum and maximum scores for the composite and the five categories. The drag of a mouse enhances this report so that it can be examined by gender, socioeconomic background, disability, or any other indicator that is in the database. Stephen Goldsmith, principal of Washington-Henry Elementary School, is examining the Stanford test scores of fourth graders to see if they can be predictors for students at risk of failing Virginia's fifth grade state tests. His plan is to tailor his school's curriculum and special programs to meet the needs of these students. The data warehouse technology allows him and his staff to provide customized curricula according to the skill level of these students. Carol Cash, principal of Lee-Davis High School, provides her Algebra II teachers with a roster of their students with their Algebra I composite scores and all its scores by skill from the state test. This gives the Algebra II teachers a good indication of the skill level of each student in the class. Cash generates these reports in approximately 10 minutes.

Data Warehousing, continued

Required Technology and Training

The two key pieces of software that HCPS selected for this project are Microsoft's SQL Server 7.0 and Business Objects. This month the district will upgrade SQL Server 7.0 to SQL Server 2000. The end user software is called Business Objects (see the end of this article for the URL). Our district believes that it is imperative to the success of the project to have a person dedicated to the building and maintenance of the data warehouse. The training received by the technical staff was invaluable.

The district trainer developed an initial two-day module, complete with manual, for the district administrators. This training was customized to allow each participant to use school data to develop custom reports for his/her educational practices. At every training session at least one participant exclaimed that he generated a report in ten to fifteen minutes that had taken him a week to produce the previous year.

Each participant also received a follow-up day of training. A user group met every other month. The meetings provided the members with new tips and tricks from the district trainer and allowed time for sharing of technical skills learned and insights gained into the instructional program. Carole Urbansok-Eads, Director of Guidance, Testing, and Research, provided a valuable session dealing with the types of statistical analysis that is valid and pitfalls to avoid.

Business Objects is a powerful tool because it allows novices to generate meaningful reports and it allows users with advanced statistical knowledge to produce even more sophisticated information. This product also allows the school district to produce standard reports that are held in a repository and are available to each school. These reports are templates. When templates/reports are used, they prompt the user to select the appropriate school and then generate the report for that school. The administrator can then either use the prepared reports or do additional scenarios with the basic information generated. These reports can be saved at the building level for later use or enhancement.

Although Business Objects may be purchased as a web implementation, HCPS chose to use the client-sever version of the product. A powerful, secure server, located on the district's wide-area network, hosts the SQL database in a Microsoft NT environment. The client software, Business Objects, is loaded on each administrator's desktop. Business Objects, SQL Server, and NT authenticate a required login and password. Information in the data warehouse is read-only. Other than district generated reports, all work is saved at the building level. A report can be saved to a laptop. The computer can then be disconnected from the network. Any reports saved can then be further analyzed in an off-line mode.

Implementation

The district faced several challenges while implementing this solution. Identifying the data elements for the warehouse was key. Each element has associated costs, of both a monetarily and human nature. Great effort has been and is being expended to minimize the effort required to keep information accurate. The initial migrating of data from the different sources into the SQL database was both tedious and time consuming. Staying focused on specifics was difficult when dealing with overwhelming

Data Warehousing, continued

amounts of information. Freeing up all district administrators for two days of training was an intricate task.

The most difficult task in the future will be to maintain current and accurate data. All student information is entered at the building level, twenty schools with multiple individuals inputting information at each site. Data stewards, the persons most responsible for the information, will be charged with insuring the accuracy of particular data sets. Limiting data will always be a challenge. Before any new element is added to the warehouse, a committee will review it. By definition, a data warehouse contains no "live" data. The system is refreshed each night with any time sensitive data. Standardized test scores need only be entered once when they are received. Evaluation of new techniques and user satisfaction will be a continuing effort.

Despite the challenges, there are many advantages of HCPS's Instruction Decision Support System:

- It is easy to use.
- Administrators have a fast and easy way to receive information not available in the past.
- Customized reports can be generated and saved.

The IDSS allows instructional leaders to make decisions based on information. This in turn leads to improved student performance, the goal of educators everywhere.

Article Resource URLs

Business Objects
http://www.businessobjects.com

References

Inmon, W.H. (1992). Building the Data Warehouse. Wiley & Sons, Inc., Hoboken, NJ.

About the Author

William Flaherty was named the Virginia Department of Education's 2002 Region I Educational Technology Leader in late October 2002. He is the Director of Technology Services for Hanover County. He can be reached at bill@hanover.k12.va.us.