International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 5 438 - 442

Enhancing Decision Support for Secondary Education with OLAP

Vaios Papaioannou
Visiting Professor
Hellenic Open University
Patras, Greece
vpapaion@gmail.com

George-Peter K. Economou
Lab Professor,
line 2: CEID, University of Patras
Patras, Greece
econom@upatras.gr

***Abstract*—** Decision-making is one of the most critical processes taking place in a modern school. It is a necessary competence for school administrators and managerial staff especially in Education Directorates who often have to make decisions regarding the implementation of education strategies and policies. It is also important for teaching staff and school curriculum designers in order to plan teaching methods and monitor student performance. Nowadays many school functions are supported by dedicated information systems. Business Intelligence (BI) is a widely used set of techniques and tools for the transformation of raw data into meaningful and useful information for business analysis purposes. They include Online Analytical Processing (OLAP) in order to provide historical, current and predictive views of business operations. Schools in secondary education can be viewed as small organizations where effective decision making is required at many areas and levels. The aim of this project is the research of feasibility of applying OLAP Decision Support Systems in Education and Education Management, any possible benefits as well as possible enhancements. The outcome is the design and implementation of an enhanced OLAP system applied in a specific educational setting based on our case study.

Keywords-*OLAP; Business Intelligence; Secondary Education; Decision Support*

*****

I. INTRODUCTION

Nowadays most of the school functions are supported by dedicated school information systems. Numerous studies deal with decision making and how the data produced from such systems can affect decision making in educational structures raging from single school entities to state educational policies. [1], [2], [3]

The typical data generated in such an environment, enables education decision makers to have access to a vast amount of information concerning students; teachers, administrators. I also involves school finances and operations. Finally there is lots of data which refer to communities served by educational institutions. These data, however, have limited use—and could possibly be detrimental—if decision makers do not understand the benefits and limitations of data, the types of data relevant for the decisions they are confronted with, and how data can be appropriately used for decision making [4].

Decision Support System (DSS) is a computer-based information system that supports business or organizational decision-making activities. DSSs serve the management, operations, and planning levels of an organization (usually mid and higher management) and help people make decisions about problems that may be rapidly changing and not easily specified in advance - i.e. Unstructured and Semi-Structured decision problems. Decision support systems can be either fully computerized, human-powered or a combination of both.

Online Analytical Processing (OLAP) can be regarded as today's modern Decision Support Systems. OLAP is part of Business Intelligence Systems. Business intelligence (BI) is the set of techniques and tools for the transformation of raw data into meaningful and useful information for business analysis purposes. BI technologies are capable of handling large amounts of unstructured data to help identify, develop and otherwise create new strategic business opportunities. The goal of BI is to allow for the easy interpretation of these large volumes of data. Identifying new opportunities and implementing an effective strategy based on insights can provide businesses with a competitive market advantage and long-term stability [9][10] [11].

Schools in secondary education can be viewed as small organizations where effective decision making is required at many areas and levels:

- decisions regarding incoming students e.g. number of students, location of residence, originating schools, number of girls and boys, performance indexes
- current student data, number of girl students, number of boys students, current student overall-subject performance, attendance rates, overall marks, subject marks, absences
- number of graduating students, destination schools, number of failed students

The aim of this project is the research of feasibility of applying OLAP Decision Support Systems in Education and Education Management, any possible benefits as well as possible enhancements. The main approach is the utilization of information processing in a secondary education environment in order to support decision making using specialized software tools. The outcome will be the design and implementation of an enhanced OLAP system applied in a specific educational setting.

II. APPROACH BASICS

Our approach was applied in a real life scenario and concluded in three main steps.

The first step was to analyze the school data. Real data from school activity were gathered in a database. Analysis was necessary in this step in order to determine the relevant data and analyze their structure so that they would be used in the Data Warehouse.

The second step was to design an OLAP multidimensional structure to be used for data analysis. The applicable measures were identified together with dimensions, possible hierarchies, etc and the cubes were processed. Imports of data and DW ROLAP schema were also implemented.

The final step was to utilize these cubes into OLAP reporting of the analyzed data.

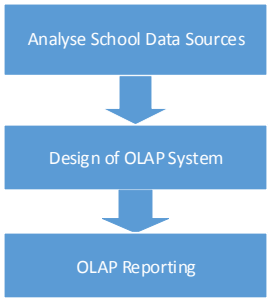

Figure 1. Approach Overview

III. OLAP AND DATA WAREHOUSES

A. *DATA WAREHOUSE*

Data Warehouses (DW) characterize a set of decision support technologies, which together with Analytical Data Processing provide the end user the possibility of obtaining valuable information from data, usually collected from the databases of a company that promotes the strategy decision making.

Using practical terms a data warehouse is a copy of data gathered from transactions of an organization, structured in such a way to accommodate complex queries and analysis of data.

According to Singh data warehouse is a decision support environment, which uses as sources of stored data from different sources. These data are organized to decide the format of information in order to assist him in making decisions. [5][6].

B. *Online Analytical Processing (OLAP)*

In today's business world the volume of data that are stored in warehouses and databases, is growing rapidly. Therefore there is a need for finding a way to manage this data efficiently and without problems in short response times. The client server architecture although solved various problems related to data management, lacked in response to questions issues. With the creation and use of the useful information for business analytical processing OLAP stored in a way that is readily accessible.

Codd et al. were the first to set the requirements to be met by a system to qualify as OLAP system. According to Codd et al. [7] an OLAP system must support multi-dimensional representation of information. OLAP tools are designed to support the decision support process, taking into account data from standard or special procedures of enterprises and organizations. This is accomplished by providing users with different OLAP functions, such as the popular roll up, drill down, slice, and pivot. Typically, each OLAP function is comprehensive and one can easily understand its operation, but the use of sophisticated OLAP tools require complex combinations of different OLAP functions, which are not easily manageable directly to users.

In terms of OLAP implementation advanced OLAP tools exist today raging from the sophisticated Oracle Business Intelligence tools and Microsoft Analysis Services to open source and freeware tools such as Mondrian and ICE cube community edition which were used for the purposes of this project.

IV. DESIGN AND IMPLEMENTATION OF THE OLAP DSS FOR SCHOOL DATA

One of the main deliverables of this project was the design and implementation of an OLAP system for supporting decision making within the school environment. The data sources and data are actual school data sources containing data from the past twelve years. Our purpose was to rely on real existing data sources instead of introducing artificial experimental data. This choice may hinder the overall DSS application but the main benefit was the investigation of the feasibility of the approach to a real situation examining real data. The main components used in this approach were the source data contained in a relational database (MS Access), the OLAP analysis software package and the reporting tools, mainly MS Excel. The source data had to be analyzed depending on their usefulness for analysis and extracted for OLAP processing. The OLAP system was responsible for transforming the input data into OLAP cubes. In fact three different OLAP systems were used, a freeware package, an open source package and a commercial professional level OLAP system. Finally the query results were presented using MS excel in order to demonstrate in our case study actual results and possible uses of the overall system in school decision making.

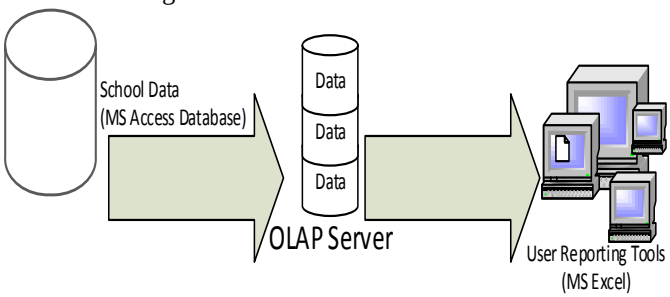

Figure 2. OLAP Implementation Overview

The development of an OLAP model involves three phases, design and development of the database, the OLAP modeling, and finally the implementation of OLAP model via an OLAP application.

A. *Data Source Analysis*

The first step towards the building of an OLAP system is to define the source data and analyze its applicability. In our case the source data is the school database implemented in Microsoft Access. The time span of the data refers to the school years from 2001-2002 to 2011 - 2012. They include information about 802 students who attended in the years mentioned above and also 96 teachers. The database includes the teachers personal information, their specialty, their qualifications etc.

For students the database holds a full list of their personal data, their class of attendance and a section in which each student belongs to each school year for their scores separately for each course for each quarter and final scores, as well as data about whether promoted to the next grade. Also there are data on students that indicate the attendance-absences gathered separately in each school year. Figure 3 depicts a small part of the aforementioned database schema.

 ISSN: 2321-8169
438 - 442

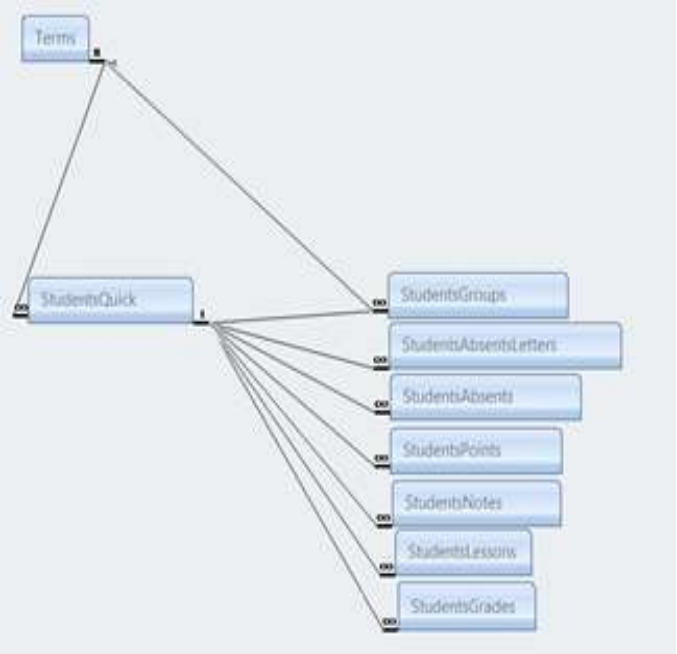

Figure 3. School Database Schema (Data Source)

The architecture of the database and the way it was designed does not follow either of the two main formats of data warehouses, shape or star shape snowflake. We could characterize it more as a relational database rather than multidimensional. However the next stage of the design of the OLAP model we manage to give it a multidimensional character adding a few more relationships between the tables.

B. OLAP Design

After defining the source data for analysis we must import them to an appropriate OLAP package. The OLAP package will create a new database suitable for multidimensional analysis and query rather than two dimensional, relational query. It will prepare the data in order to be readily available for all possible analysis scenarios and feed the reporting software module with respective reports. In order to complete this step three OLAP software packages were used: Microsoft analysis services, Mondrian open source software, ICE cube community freeware edition.

The first step is to define the source of data. The second step is to design a database query which will hold all of the actual data to be processed as shown in figure 4.

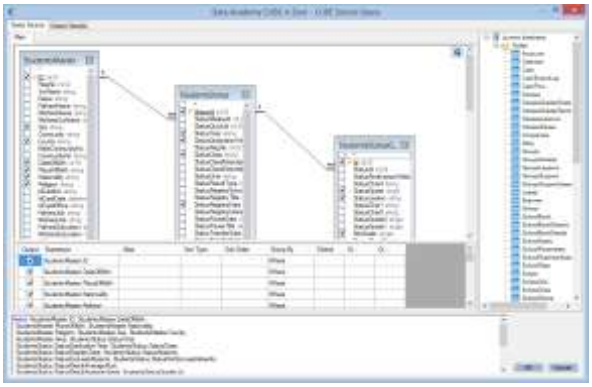

Figure 4. CUBE-it Zero Data Selection

Third step is to design the actual cube structure as shown in figure 5.

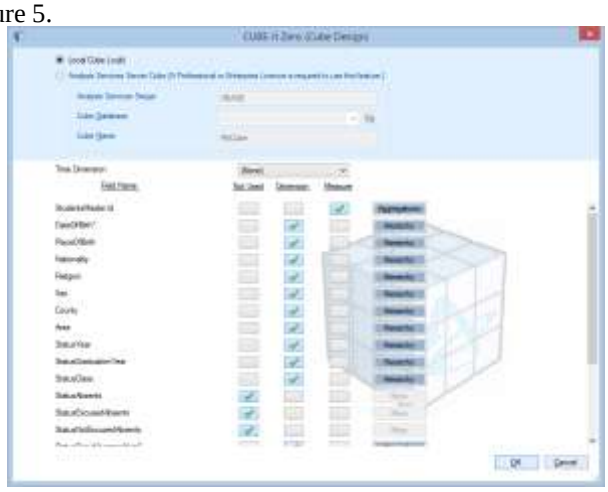

Figure 5. Dimensions and Measures definition in CUBE-it Zero

The results can be examined after OLAP report design as described in the next section.

C. OLAP Reports Design

The use of data in order to plan school related processes and in fact to inform instruction is the critical reason we collect the data. The collection of student performance data without using it to inform instruction would be a waste of valuable teacher time although this is exactly what occurs in many schools. The need for classroom monitoring data satisfies our requirement to understand the students levels against to the content standard indicators teachers are responsible for delivering so that we can do planning based on informed decisions. It also helps planning all of the possible school processes such as term planning, staff allocation, budgeting, allocation of school rooms, school activities, etc.

Although reporting capabilities may be endless we demonstrate just a few examples that demonstrate the usefulness of the approach. For example in figure 6 we can analyze the origins and residency areas of the student population so result can lead to decisions regarding the school start and end times, the decision for routing the transportation options e.g. school bus leasing, the planning of school events, the planning of communication options, the introduction of special care of remote teaching methodologies for isolated students.

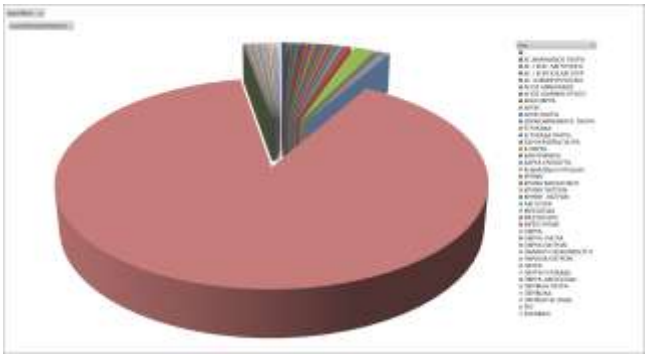

Figure 6. Analysis of Students Residency Areas

Another example is the examination of the allocation of Grades among subjects. This allows us to have a better view of teachers marking habits as well as to examine the performance

of stunts in these subjects. This query report in figure 7 displays the number of students who obtained a specific grade in a specific lesson for the last 13 years.

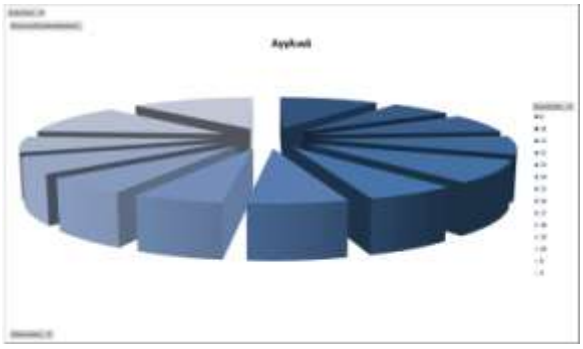

Figure 7. Students Perfomance in a specific subject (English) per Year

Finally A very important area for analysis is the students absences. Absent students can be a very serious problem in some schools and in some situation it can lead to failed class or even abandoning school all together. The following chart shows the allocation of number of students per absences per class. A majority allocation towards high numbers of absences would be alarming for this specific school. Fortunately in our case most students are in the 0 to 15 area as shown in figure 8.

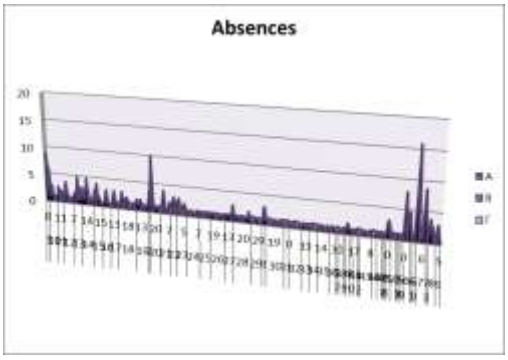

Figure 8. Analysis of student number of absences per class

Finally we can for example examine the number of absences per class per school year as shown in figure 9.

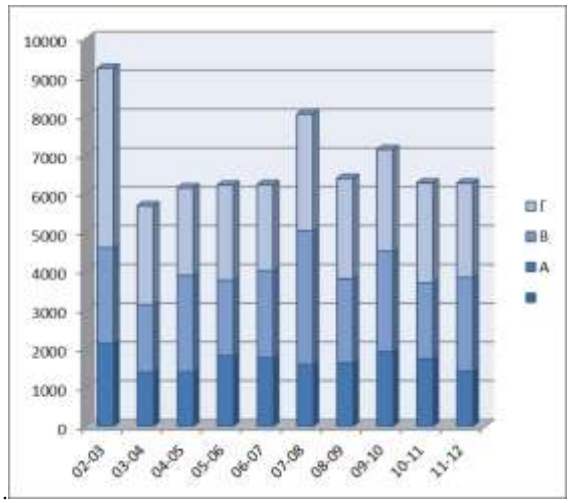

Figure 9. Number of absences per class per school year

School attendance is a very important measure which can indicate problems in teaching, problems with students, health problems and family problems. It allows us for proper school planning and allocation of teams.

V. CONCLUSIONS

Decision-making is one of the most critical processes taking place in a modern school. It is a necessary competence for school administrators and managerial staff especially in Education Directorates who often have to make decisions regarding the implementation of education strategies and policies. It is also important for teaching staff and school curriculum designers in order to plan teaching methods and monitor student performance. Their role is critical because:

- Wrong decisions can be the cause of financial problems because of the large complexity of a school organization
- It is hard to gather all appropriate information in order to make decision
- There are lots of alternatives because of advances in technology

The approach reported in this paper attempts to apply modern Decision Support Systems to assist decision making at the secondary education level. Decision support functionality is realized by offering reporting tools for solving particular school related tasks as well as by allowing users to navigate through the data, query it, generate interactive visualizations and explore those for retrieving interesting details.

The architectural data (Data Warehousing) and On-line data analysis (OLAP), is part of decision support applications and they are a modern approach to the problem of decision support. Although these technologies belong to the Business Intelligence Technologies whose primary target is to support business decisions, we attempted to apply them in an educational environment by viewing it as a special type of organization. Thus the second step was to apply these technologies to a case study involving real school data, kept in databases. The information kept in this database was initially analyzed in order to identify information concerning students, lessons, grades, absences and teaching staff. Once the important information was identified, an OLAP system was used in order to build an OLAP data mart in order to analyze this critical information using analytical queries. The result of these queries were live numerical and graphical reports. These "multidimensional" reports could be used so that we build an information background, capable enough to support us in many facet of school related decision making processes. Decisions, among others, may include the following categories:

- Instructional: Learning resources; student awards & recognition; enrichment activities.
- Resources: Budget development and evaluation; grant writing; employee recognition; recruiting, selecting, recognizing and using volunteers and community resources.
- Governance and management: Discipline plan; attendance policy; community relations and communications; special events; scheduling; school-based fees; school safety issues.

The application of such a model provides a direct analysis of data supports decision-making processes and all of them through actions free of technical issues. However data

selection, cleaning and integration might consume a large proportion of the development effort especially when combining data from various sources and various schools. Better data organization should combine more tables and define dimension hierarchies. In fact School Information Systems are often overwhelmed with data, so that decision-makers cannot effectively synthesize and use them towards a creative decision path. Nevertheless, although the initial data views were very limited, we had the opportunity to apply a system in a real life situation and observe the results. Analysis was performed in real time and end users were required minimal knowledge of computer usage.

From the above we proved that using OLAP systems through their ability to present combined results with the multidimensional character serve the administrative needs of educational units or organizations. Especially the fact that they provide instant information models for changes in source data, enables managers to monitor modern aggregated data at any time to take administrative decisions on the root of the problems.

More specifically, the directors of educational units or organizations by using centralized and up to date data would be able to draw up more accurate financial budgets. Even with regard to the educational process they will be able to identify weaknesses and make structural changes where needed. Educators will be able to group students in relation to a characteristic and discovering weaknesses such as too many absences or low scores in a course or in the general average and to make motions such as creating supportive teaching departments. Finally the use of aggregated data that would have resulted from processing in OLAP systems will contribute to better evaluation of the teaching staff, managers will be able to discern whether the poor performance of some students due to teaching staff or students learning difficulties.

REFERENCES

[1] U.S. Department of Education, Office of Planning, Evaluation and Policy Development, "Implementing Data-Informed Decision Making in Schools: Teacher Access, Supports and Use", Washington, D.C., 2009.

[2] Julie A. Marsh, Caitlin C. Farrell, "How leaders can support teachers with data-driven decision making: A framework for understanding capacity building", Educational Management Administration & Leadership March 1, 2015 43: 269-289.

[3] Jeffrey C. Wayman, "Involving Teachers in Data-Driven Decision Making: Using Computer Data Systems to Support Teacher Inquiry and Reflection.", Journal Of Education For Students Placed At Risk, 10(3), 295–308, Lawrence Erlbaum Associates, Inc, 2005.

[4] Brian Gill, Brandon Coffee-Borden, and Kristin Hallgren, "A Conceptual Framework for Data-Driven Decision Making," Princeton, NJ: Mathematica Policy Research, 2014.

[5] Harry S. Singh, "Data warehousing: concepts, technologies, implementations, and management", Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1998 ISBN:0-13-591793-X.

[6] Inmon, W. H., "What is a Data Warehouse?", Sunnyvale,CA: PRISM Solutions, Inc. Tech Topic 1, no. 1, p. 1.,1992

[7] Codd, E. F.; Codd, S. B.; Salley, C. T., "Providing OLAP to User-Analysts: An IT Mandate", Technical report, E. F. Codd & Associates, 1993.

[8] K. Jayapriya 1., E.Girija,III-M.C.A., R.Uma.,M.C.A.,M.Phil., A Technical Review on On-Line Analytical Processing (OLAP), International Journal of Computer Science and Communication Networks (IJCSCN), Vol 5(5),315-324, ISSN:2249-5789, 2015

[9] S. Poonkuzhali and C.Saravanakumar,. "Indexing OLAP data, OLAP Queries",. Charulatha Publications in Revised. Edition.

[10] Jiawei Han and Micheline Kamber, "Types of OLAP servers, Architecture of OLAP, From OLAP to OLAM", Second Edition.

[11] Kale Sarika Prakash, P.M. Joe Prathap, "Bitmap indexing a suitable approach for data warehouse design",. International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 3 Issue: 2 680 - 683