

SOME ASPECTS OF DATA WAREHOUSING IN TOURISM INDUSTRY

PhD. Mirela DANUBIANU

"Stefan cel Mare" University of Suceava, Romania
mdanub@eed.usv.ro

PhD. Candidate Tiberiu SOCACIU

"Stefan cel Mare" University of Suceava, Romania
socaciu@seap.usv.ro

PhD. Candidate Adina BARILA

"Stefan cel Mare" University of Suceava, Romania
adina@eed.usv.ro

Abstract:

In this paper we make a study of the possibility and necessity to deploy a Data warehouse for the tourism industry that aims to support decision makers by giving them different views for the same piece of data. We first point out the three sides of tourism: economic, social and cultural. Then the focus is moved on the information systems that are actually used in domain and we show that the advanced information technology is still lower used than in other economic and industrial sectors. We continue with some concepts regarding data warehouses. Finally we try to demonstrate the need of data warehousing for industry tourism in Romania.

Keywords: data warehouse, data warehousing, information systems, multidimensional model, on-line analytical processing

JEL Classification: C61, C81, C89

INTRODUCTION

Romania has a huge tourist's potential, unfortunately, too little valued and exploited. As a result, one of the strategic developments of the economy aimed the tourism industry.

But strategies are based on different trends obtained from sophisticated analysis of data. Providing the managers in the tourism industry with information about and insight into the existing data is the key function of the data warehouse systems.

A data warehouse (DW) is a collection of technologies aimed to enabling the decision maker to make better and faster decisions. It is designed to support On Line Analytical Processing (OLAP). There are some previous works that are related to the general architecture of a data warehouse and the development of data warehouse prototypes in fields like telecommunication, banking, and insurance. However, to our knowledge, there are no previous works related to data warehousing for the tourism industry. It is the reason for in this paper we study the need and the opportunity to design and to implement a data warehouse in this field.

First we focus on the three dimensions of tourism: economic, social and cultural and we point out the major implication of tourism industry in economy. Then we make a short analyze of the information systems actually used in this area and we show that advanced technology used in the different components of this sector is still low. Section 3 is dedicated to general concepts regarding data warehouses. Here we point also the advantages of using these projects. In Section 4 we demonstrate the need of using a data warehouse in tourism industry.

TOURISM – ECONOMIC, SOCIAL AND CULTURAL PHENOMENON

Branch of the national economy and part of the tertiary sector, tourism has an important role in economic and social life acting as an element that stimulates the global economic system, as a means of diversification of economic structure and as an element of education. Due to the relationship with many other sectors, tourism is an industry of interference and as a result of that is based on results obtained in other branches of activity is an economic branch of consequence.

Tourism demand leads to a continuous adjustment of the offer. This is materialized in the development of tourism structures which entail increased production industries that are directly or indirectly involved in supporting these structures.

In a study conducted by the United Nations World Tourism Organization (UNWTO) effects on the economy can be grouped into following categories:

- global effects on the national economy by stimulating the production and use of labor;
- partial effects on the balance of payments equilibrium, on the level of the exchange rate, on money supply and movement of money, on the distribution of income, regional development, rural environment and population movements and
- external effects on the quality of the environment, on professional training, on habits of consumption, on education and on cultural and social changes.

The economic importance of tourism is underlined by the following aspects [1]:

- tourism is a creator and user of national income;
- it is a means to exploit resources;
- supports diversifying economic structures;
- generate new jobs;
- is a stimulating of investment.

In addition to the economic consequences, tourism has a strong social implication. The specific actions have a direct effect on tourists and on population and environment of the areas visited.

On the international level tourism is an important contributor to the creation of gross world product (7%), and in terms of exports, tourism is ranked on the first places.

OVERVIEW OF INFORMATION SYSTEMS ACTUALLY USED IN TOURISM INDUSTRY

Information technology was initially viewed by the tourism industry as a back-office function that supports the finance and accounting areas. The industry has advanced far beyond this view during the past decade. In some sessions tourism industry leaders pondered the role of technology. Among the conclusions reached were: "Going forward, technology will be the most competitive weapon for any touristic company. If touristic organizations want to compete successfully, they must do so by using technology to drive value to both the customer and to the firm." [4]

Actually, the most used information systems in tourism industry are the front-office systems and the reservation systems.

Front Office Information Systems

Front-office information systems are those data processing systems that provide reports in visual or written form. They are used mainly in the management of tourist accommodation (hotels, motels, hostels or cruise ships) or in the travel agencies activities.

These systems may be used for:

- tourists registration when the personal data about tourists are collected;
- marketing of various tourism products, such as rental cars;
- rooms management, when are collected and processed data regarding the rooms status, (allows instant viewing of room availability for all room types, indicates whether rooms are dirty or clean, allows rooms to be placed out of inventory or out of order to restrict rental)
- tracks of revenues, providing transaction processing and obtain information about any debts and credits in relation to customers

Information Systems Used for Reservations

This kind of systems provides rapid access to information and ensures the accuracy of this information. They bring information services, booking and selling and are used both by individual tourists and travel agents or commissioners.

Such systems can be classified into the following categories:

- information systems that function as data banks accessible through the transmission systems for consultation;
- availability systems that provide information on the status of free or completely occupied a location at a time;
- computerized reservation systems.

Most often this type of systems uses Web technologies. These systems use hardware and software specific to conduct them activities. Although providers of tourist services in Romania currently use such systems for ticketing most, is well to remember that these systems can be used for marketing or management activities.

Despite the fact that tourism is a dynamic industry with important implications on the economy to adopt advanced technology of the different components of this sector is still low. For example, according to a study conducted by the magazine e-Business Watch, the percentage of tourism organizations adopting and using application in different areas (such as customer relationship management CRM, enterprise resource planning ERP or supply chain management SCM) is sensibly lower than in other economic and industrial sectors (Fig. 1).

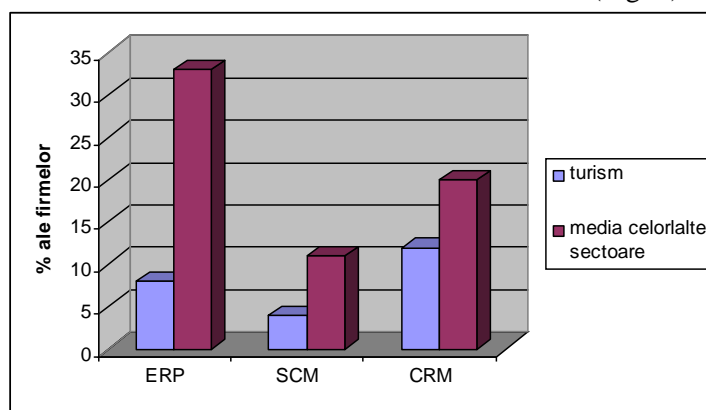


Figure no. 1. Diffusion of advanced software solution in European tourism industry compared to the average diffusion in seven industry sectors (e-Business Watch, 2004)

Information Systems Using Data mining Techniques

In the tourism industry knowing the guests - where they are from, how much they spend, and when and on what they spend it- can help a company to formulate marketing strategies and maximize profits. Due to technological development touristic companies have accumulated large amounts of customer data, which can be organized and integrated in databases that can be used to guide marketing decision [5].

Since identification of important variables and relationships located in these consumer - information systems can be a difficult task, some companies have attempted to raise the power of information by using data mining technologies that exploits the data regarding the consumer. Such data-mining technology allows these companies to predict consumer -behavior trends, which are potentially useful for marketing applications.

WHAT IS A DATA WAREHOUSE

Data warehouse is a relatively new concept, appeared from the need to place a source of coherent and consistent data at managers' disposal to enable obtain information necessary for the substantiation of decisions. Data warehouse has emerged as a result of economic necessity but it was a challenge for Information Technology professionals.

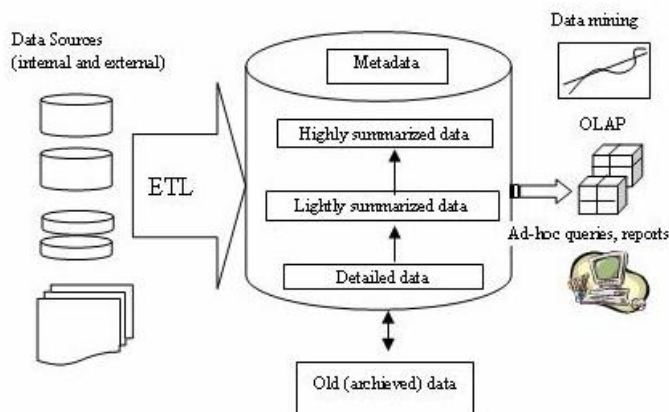
So, what is a data warehouse? William H. Inmon [2] defines a data warehouse as "a collection of integrated, subject-oriented databases designed to supply the information required for decision-making".

A data warehouse contains data extracted from the many operational systems of the enterprise, possibly supplemented by external data. Each of these operational systems records different types of business transactions and enforces the policies of the enterprise regarding these transactions. If each of the operational systems has been custom built or an integrated system was not implemented as a solution, then it is unlikely that these systems are integrated.

For example, Customer A in an operational system and Customer B in another operational system may be one and the same person—but there is no automated way for anyone in the organization to know this. A data warehouse brings together data from the various operational systems to provide an integrated view of the customer and the full scope of his or her relationship with the organization.

Traditional operational systems focus on the data requirements of a department or division. With the advent of business process reengineering, enterprises began espousing process-centered teams and case workers. Modern operational systems, in turn, shifted their focus to the operational requirements of an entire business process and aim to support the execution of the business process from start to finish. A data warehouse goes beyond traditional information views by focusing on enterprise-wide subjects such as customers, sales, and profits. These subjects span both organizational and process boundaries and require information from multiple sources to provide a complete picture.

The data warehouse is an environment, not a product. It is an architectural construct of



information systems that provides current and historical decision support information that is hard to access or present in traditional operational databases. The generic architecture of a data warehouse is presented in Fig. 2.

Figure no. 2. The general architecture of a data warehouse

In practice is used the term *data warehousing technologies* to refer to the gamut of technology components that are required to plan, develop, manage, implement, and use a data warehouse, the term *data warehouse* itself refers to a large, read-only repository of data (Fig.2). At the very heart of every data warehouse lie the large databases that store the integrated data of the

enterprise, obtained from both internal and external data sources. The term *internal data* refers to all data that are extracted from the operational systems of the enterprise. External data are data provided by third-party organizations, including business partners, customers, government bodies, and organizations that choose to make a profit by selling their data.

An important feature of Data Warehouse is that store both detailed data and data on different levels of summarization. *Granularity* refers to the level of detail or summarization of the units of data in the data warehouse. Granularity is the single most critical design issue in the data warehouse environment because it profoundly affects the volume of data that resides in the data warehouse and the type of query that can be answered. The volume of data in a warehouse is traded off against the level of detail of a query. The lower level of granularity corresponds to the more versatile query that can be issued. The higher level of granularity corresponds to the less versatile query that can be issued.

Also stored in the databases are the metadata that describe the contents of the data warehouse.

The Extraction, Transformation, and Loading (ETL) processes include operations for data unification, integration, cleaning, and data transferring from data sources into data warehouse. On the other side different types of software are required to access, retrieve, distribute, and present warehouse data to its end users.

Using DW offers a wide range of benefits [3].

The data warehouse provides access to integrated enterprise data previously locked away in unfriendly, difficult-to-access environments. Business users can now establish, with minimal effort, a secure connection to the warehouse through their desktop PC. Security is enforced either by the warehouse front-end application, by the server database, or both. Because of its integrated nature, a data warehouse spares business users from the need to learn, understand, or access operational data in their native environments and data structures.

The data in the data warehouse are consistent and quality assured before being released to business users and data warehouse becomes the common information resource for decisional purposes throughout the organization. This is the way to provide a single "version of the truth".

Many of the figures and numbers that managers receive have little meaning unless compared to historical figures. Actual historical values are not stored on the operational system nor derived by adding or subtracting transaction values against the latest balance. Historical data are loaded and integrated with other data in the warehouse for quick access.

The ready availability of different data views also improves business analysis by reducing the time and effort required to collect, format, and distill information from data.

At the end of each business process reengineering (BPR) initiative come the projects required to establish the technological and organizational systems to support the newly reengineered business process. Although reengineering projects have traditionally focused on operational processes, data warehousing technologies make it possible to reengineer decisional business processes as well. Data warehouses, with their focus on meeting decisional business requirements, are the ideal systems for supporting reengineered decisional business processes.

NEED FOR A DATA WAREHOUSE IN ROMANIAN TOURISM INDUSTRY

In order to correct any deviations in performance, managers in the tourism industry often need timely analysis reports to measure and monitor the performance rate, increase and decrease in tourist numbers, tourism nights, and percentage of hotel occupations, visits to monument places, and the total revenue from the tourism sector at the national level. They also need timely analysis reports to assist in making long-term decisions. It has been observed that most of the reporting and analysis, time was spent on collecting data from the various systems before the analysis can be made. Managers want and need more information, but analysts can provide only minimal information at a high cost within the desired time frames [6]. In order to provide information for

predicting patterns and trends more convincingly and for analyzing a problem or a situation more efficiently, a data warehouse for this particular purpose is needed.

Evaluation of current tourism industry activity is done by means of statistical indicators such as number of visitor arrivals (total and area of origin - for foreign tourists), spent in hotels and similar establishments, the average length of stays, etc. These indicators allow only an analysis of tourist activity, but without offering the possibility of strategic decisions made in tourism or for those areas with direct or indirect impact on tourism. For this purpose the managers or, on a higher level, the governors may have the possibility of analysis from different perspectives, such as the type visited tourist areas (mountains, seaside resorts spas ...) the type of accommodation establishments, the type of tourism practiced, the profile of tourists who prefer certain destinations, so on. These tests are complete if allow comparisons over time for those indicators.

The multidimensional model, characteristic of data warehouses allows such analysis on different dimensions and hierarchical levels.

For example, data stored in a Data warehouse allow in every moment to quickly answer to questions of following type: *What was the activity (expressed in number of tourists' arrivals) for the 2 stars, 3 stars or 4 stars hotels in the mountain area for the first quarter of 2008 compared with the same period in 2007.*

In this case one can make an analysis from the perspective of follows dimensions: establishment of tourist accommodation, tourist area and time.

In the above example we refer the quarter level for time dimension. But for this dimension can define multiple hierarchies, such as: day-week-month-year, week-month-quarter-semester-year, so on. Starting from the query above, if it is necessary a more detailed analysis (e.g. related to months) there is the possibility of finding immediate response by a drill-down operation.

So decision makers in the tourism industry need more and more analytical information to capture the whole picture of their tourism environment, and it is exactly the role of data warehouse to give them this global view and wide capability for analysis.

CONCLUSION AND FUTURE WORK

Since the tourism industry has become one of the major objectives of development, especially for Bucovina we think that deploying a data warehouse project, even for this area is really useful. Based on this idea we intend to develop a prototype for a data warehouse to allow the on-line analyze of tourist activity, like numbers of arrivals, spent in hotels or similar establishments and the average length of stays from the perspective of establishment of tourist accommodation, tourist area and time. Also we intend to develop some data mining tools in order to describe the profile of tourists interested on Romanian tourism.

REFERENCES:

1. N. Dumitru, *Contribuția turismului în procesul creșterii economice*, *Journal of tourism*, Nr 3.
2. W. H Inmon, *Building the Data Warehouse*; QED Technical Publishing Group, 1992, ISBN: 0-89435-404-3
3. M. Humphries, M. Hawkins, M. Dy, *Data Warehousing Architecture and Implementation*, Pearson Education,
4. Olsen, M., Connolly, D., (1999) *Antecedents of Technological Change in the Hospitality Industry*, *Tourism Analysis*, Vol. 4, p. 29
5. M. Danubianu, V. Hapenciuc *Improving Customer Relationship Management In hotel industry by Data Mining Techniques*, *Proceeding of "Competitiveness and Stability in the Knowledge-Based Economy"*, Vol: CD, 30-31 Mai, 2008, Craiova, Romania, ISSN/ISBN: 978-606-510-162-3, Pagini: 2444-2452

6. J. Dyche. *e-Data Turning Data into Information with Data Ware housing*, Addison-Wesley, Reading, 2000.
7. K. Chau, Y. Cao, M. Anson, J. Zhang. "Application of data warehouse and Decision Support System in construction management. *Automation in Construction*". Elsevier, 2003.