# Data Warehouse With OLAP Technology for the Tourism Industry

D

**Preetvanti Singh**
*Dayalbagh Educational Institute, India*

**Vijai Dev**
*Dayalbagh Educational Institute, India*

## INTRODUCTION

The main objectives of this chapter are to present:

1. Overview of Data Warehouse and OLAP Technologies.
2. Literature review of Data warehousing and OLAP technology.
3. Data warehouse architecture and Integration of Data Warehouse with OLAP technology.
4. Implementation for Tourism Industry.
5. Future research direction in this field.

In the modern digital era, we are living in a data-driven world, where an enormous amount of data is collected and stored on a daily basis. It becomes important to have the ability for accessing and analyzing this data in order to use it effectively. The collection of enormous business data is termed a data warehouse that enables organizations in making decisions. A data warehouse is a central repository of integrated data from one or more than one, unlike data sources. It is a data management system designed to enable and support business intelligence activities.

A large amount of data in data warehouses come from different places such as finance, sales, and marketing. A data warehouse periodically pulls data from these applications and processes it to make it ready for access by the decision-makers. Data warehouse technologies are used by decision-makers to build forecasting models, run logical queries, and identify trends in an organization.

Online Analytical Processing (OLAP)is used to analyze and evaluate data in a warehouse. This technology organizes data in the warehouse using multidimensional models. It breaks down data into dimensions; for example, total sales might be broken into dimensions such as geography and time. Breaking the complex data into multiple dimensions enables analysts to apply OLAP technology for organizing information to easily understand and use business data for efficient decision-making. OLAP plays a vital role in meeting organizations' analytical demands by allowing decision-makers to measure facts across the company.

The tourism industry is closely interconnected with the various global industries/sectors, and contributes towards the complete growth of a country by bringing several economic benefits including building brand value and identity of a country. The benefits of tourism on host destinations include boosting the revenue of the economy, creating a large number of jobs, enriching diversity and culture, and developing the infrastructures of a country.

DOI: 10.4018/978-1-7998-9220-5.ch012

## BACKGROUND

To enhance the decision-making capabilities using multi-dimensional data, researchers have developed data warehouses for different real-life problems. Al Faris & Nugroho (2018) developed a data warehouse to enable the company in keeping the delivery service time always on target. Analysis was done using OLAP to build a report and dashboard. The operational database was transformed into a data warehouse through the Extract Transform Load (ETL) process. Cuzzocrea, Moussa & Vercelli (2018) applied Lambda architecture to develop an approach for supporting data warehouse maintenance processes in the context of near-real-time OLAP scenarios. Big summary data was utilized and was assessed via an empirical study that focused on the complexity of such OLAP scenarios. Rahutomo, Putri & Pardamean. (2018) built a data warehouse model for education management support. Data was collected through interviews, questionnaires, observations, and literature review. Models were developed using Visual Basic.Net 2008, SQL Server 2005, and Crystal Reports. Schuetz, Schausberger & Schrefl. (2018) developed a semantic data warehouse to support business intelligence in precision dairy farming based on the sensor data. The authors introduced semantic OLAP patterns to automate periodic analysis. Sutedja, Yudha, Khotimah & Vasthi (2018) designed a data warehouse to integrate various operational databases for providing information about students at a university. The design method used 4 stages: selecting the business process, declaring grain, and identifying the dimensions and the facts. A dashboard was developed for providing the relevant and integrated information of students from different angles. Wang (2018) created a multi-dimensional cube of teaching evaluation and extracted knowledge hidden in the data.

Agapito, Zucco & Cannataro (2020) designed a COVID warehouse to model, integrate, and store the COVID-19 data, and pollution and climate data for Italian Regions. Using ETL this data was integrated and organized as a dimensional fact model while considering time and geographical location dimensions. OLAP analysis was performed to provide a heatmap visualizer and extract selected data for further analysis. Dehdouh, Boussaid & Bentayeb (2020) defined a cube operator, MC-CUBE (MapReduce Columnar CUBE) to build columnar NoSQL cubes by considering the non-relational and distributed aspects while storing data warehouses. Khalil & Belaissaoui (2020) implemented big OLAP cubes based on NoSQL key-value stores. Structures for supporting decision-making were presented by responding to OLAP-based analytical operations. Shelest & Holub (2020) discussed the intellectual analysis of outdoor advertising data and suggested a methodology for correct storage of data which is the basis for clear data analysis. Surarso & Gernowo (2020) integrated data warehouse and OLAP with k-medoids clustering for data prediction and control based on the accreditation self-evaluation report matrix of the study program. Data of 3 years was taken for analyzing new students, study periods, their achievements, and lecturers using OLAP technology and the k-medoids cluster. Microsoft data warehouse services and related software architecture were introduced by Wang, Liu & Wu (2020) to solve the problems of slow analysis queries caused by the huge amount of data. Agricultural Science and technology support data was considered for developing the system. Wu, Zhou, Wang & Jiang (2020) developed an urban flood data warehouse with available structured and unstructured urban flood data. Gradient Boosting Decision Tree, regression model, and deep learning model were applied to forecast the depth of the flooded areas. Flood condition maps were produced based on the different rainfall return periods. Yulianto & Kasahara (2020) developed a data warehouse to manage the tuition-fee-level of higher education institutions in Indonesia for providing information of applicants' tuition fees to the administrators. The data warehouse demonstrated four basic dimensions (faculty, year, entrant type, and tuition fee level) regarding applicants, tuition fee level, and payment status.

Automobile Predictive Maintenance Using Deep Learning
Sanjit Kumar Dash, Satyam Raj, Rahul Agarwaland Jibitesh Mishra (2021). *International Journal of Artificial Intelligence and Machine Learning (pp. 1-12).*
www.irma-international.org/article/automobile-predictive-maintenance-using-deep-learning/279274