

Project Proposal

Introduction:

A data warehouse is a collection of data from one or more sources that can improve the efficiency of operations. We live in a data-driven society where massive amounts of data are collected and stored every day. A data warehouse provides us with a multidimensional view of aggregated and collected information. This project focuses on creating a variety of data warehousing and OLAP approaches, with a particular emphasis on their new standards. The project can utilize Data Warehouse functions like data cleaning, data mining, and data transformation. Data Warehouse also includes OLAP tools for querying and analyzing data from many perspectives.

Motivation:

The team chose this topic because we wanted to learn how data warehousing helps decision-making by facilitating and optimizing access to multiple data sets. In addition, to gain more knowledge about data warehouses and how they are used in a wide range of applications. To develop and run a data warehouse, it's not enough to pick an OLAP server, put up a schema, and run a few sophisticated queries.

Why is it challenging?

- Data accuracy is difficult to maintain, especially in large-scale warehousing operations.
- The data warehouse will eventually need to be updated to meet new requirements. Unfortunately, updating is expensive, resulting in project failure sometimes.
- Architecture is rigid, and structure is complicated.

Who will benefit from a solution to the problem?

- Effective Data Warehousing can benefit the organizations by improved data analytics and effective data retrieval.
- Data Warehousing can help the improving the integrity of data gathered from various platforms
- Scalable data warehouses can handle limitless information with the same instructions hence it can automate data gathering.

Why is it suitable for Advanced Database Topic?

We have chosen Data Warehouse and OLAP for Advanced Database Topics because a data warehouse's OLAP functionality extends beyond simple querying and reporting. It entails combining many enterprise source systems into a vast historical dataset that can be analyzed using advanced queries. An OLAP system is a customized database that can handle multidimensional data cubes.

Initial Ideas:

To develop our project, we will leverage data warehousing and OLAP techniques. We will also do OLAP functions like roll-up, drill-down, slice-and-dice, and pivot. We may also utilize ROLAP or MOLAP models in this project. ROLAP is an upgraded RDBMS with multidimensional data mapping to execute standard relational operations, whereas MOLAP implements procedures on multidimensional data.

Advantages:

- Data warehouses may be used to track and adapt marketing strategies as well as strengthen customer relationships.
- Data sources are added to the data warehouse to ensure that they are constantly updated and available. It enhances data quality and integrity, allowing for better decision-making.
- Furthermore, extracting data based on specified criteria is advantageous.

Related works:

[1]. Chaudhuri, S., & Dayal, U. (1997). An overview of data warehousing and OLAP technology. ACM Sigmod Record, 26(1), 65-74

[2]. Arun Sen and Atish P. Sinha, "A Comparison of Data Warehousing Methodologies", communications of the ACM March 2005/Vol. 48, No. 3.