A.P. SHAH INSTITUTE OF TECHNOLOGY
PARSHWANATH CHARITABLE TRUST'S
Department of Computer Science and Engineering
Data Science

CSE DATA SCIENCE

Semester : V   Subject :DWM   Academic Year: 2023 - 2024

Module 1

What is a Data Warehouse?

A data warehouse is a relational database system businesses use to store data for querying and analytics and managing historical records. It acts as a central repository for data gathered from transactional databases.

It is a technology that combines structured, unstructured, and semi-structured data from single or multiple sources to deliver a unified view of data to analysts and business users for improved BI. Therefore, it is used for analytical and business reporting purposes, which helps keep past records and analyze data to optimize business operations.
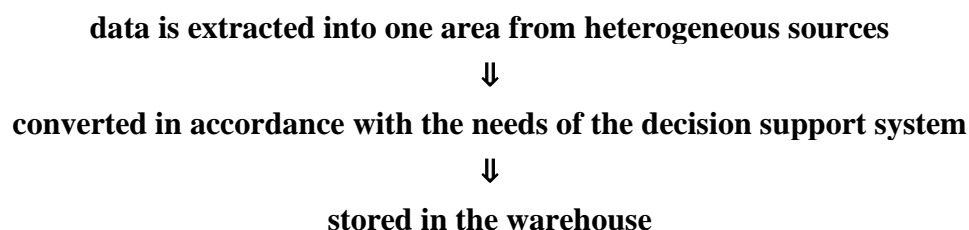
A data warehouse is often confused with a database. However, there is a huge difference between the two.

While a database is merely a conventional technique to store data, a data warehouse is especially intended for data analysis. It keeps everything in a single location from numerous external databanks.

How does it work?

The concept firstly appeared in the 1980s. It was developed to support operational systems' data flow transferred to decision-making systems. These systems required the analysis of large amounts of heterogeneous data accumulated by companies over time.

Data warehouse works on the following principle:

data is extracted into one area from heterogeneous sources
⇓
converted in accordance with the needs of the decision support system
⇓
stored in the warehouse

Thus, the system provides answers for business decisions analyzing all these heterogeneous data. That is why data warehousing primarily aimed to simplify decision-making processes and helps executives to get required information based on the whole data quickly.