**Introduction to data warehouse**

Data warehouse: is a physical repository where relational data are specially organized to provide enterprise-wide, cleansed data in a standardized format.  A data warehouse provides a new design which can help to reduce the response time and helps to enhance the performance of queries for reports and analytics.

**A Data warehouse system has many names…**

**History of Data warehouse**

The Data warehouse benefits users to understand and enhance their organization's performance. The need to warehouse data evolved as computer systems became more complex and needed to handle increasing amounts of Information.  However, Data Warehousing is a not a new thing.

A diagram of data storage

Description automatically generated

Who needs data warehouse?

* Decision makers who rely on mass amount of data
* Users who use customized, complex processes to obtain information from multiple data sources.
* It is also used by the people who want simple technology to access the data
* It also essential for those people who want a systematic approach for making decisions.
* If the user wants fast performance on a huge amount of data which is a necessity for reports, grids or charts, then Data warehouse proves useful.
* Data warehouse is a first step If you want to discover 'hidden patterns' of data-flows and groupings.

**Introduction to Data warehouse**

**Data warehouse examples**

**Airline**: In the Airline system, it is used for operation purpose like crew assignment, analyses of route profitability, frequent flyer program promotions, etc.

**Banking**: It is widely used in the banking sector to manage the resources available on desk effectively. Few banks also used for the market research, performance analysis of the product and operations.

**Healthcare:**Healthcare sector also used Data warehouse to strategize and predict outcomes, generate patient's treatment reports, share data with tie-in insurance companies, medical aid services, etc.

**Retain chain:**In retail chains, Data warehouse is widely used for distribution and marketing. It also helps to track items, customer buying pattern, promotions and also used for determining pricing policy.

<https://youtu.be/U5iF3OdUNNc>

| **Database vs Data warehouse differences** | |
| --- | --- |
| **Database** | **Data warehouse** |
| A database is a collection of related data which represents some elements of the real world. It is designed to be built and populated with data for a specific task. It is also a building block of your data solution. | A data warehouse is an information system which stores historical and commutative data from single or multiple sources. It is designed to analyze, report, integrate transaction data from different sources.  Data Warehouse eases the analysis and reporting process of an organization. It is also a single version of truth for the organization for decision making and forecasting process. |
| It offers the security of data and its access | Data warehouse helps users to access critical data from different sources. |
| It offers a variety of techniques to store and retrieve data. | Data warehouse allows you to store a large amount of historical data to analyse. |
| Acts as an efficient handler to balance the requirement of multiple applications using the same data | Enhances the value of operational business applications and customer relationship management systems |
| A DBMS offers integrity constraints to get a high level of protection to prevent access to prohibited data. | Separates analytics processing from transactional databases. |
| A database allows you to access concurrent data. | Data warehouse provides more accurate reports with reduced TAT (total turnaround time) |

**Dimensionality**

<https://youtu.be/f2kIlKfWzkA>

**Further reading / extension activity**

Sherman, R. (2015). Business intelligence guidebook : from data integration to analytics. Elsevier. Chapter 9