

What is BI?

BI is a technology-driven process for analyzing data and presenting actionable information to help corporate executives, business managers and other end users make more informed business decisions. BI encompasses a wide variety of tools, applications and methodologies that enable organizations to collect data from internal systems and external sources, prepare it for analysis, develop and run queries against the data, and create reports, dashboards and data visualizations to make the analytical results available to corporate decision makers as well as operational workers.

Why BI?

To accelerate and improve decision making in today's competitive world.

What is Data Warehouse?

It is the process of constructing an integrated data from multiple heterogeneous system which can support analytical processing.

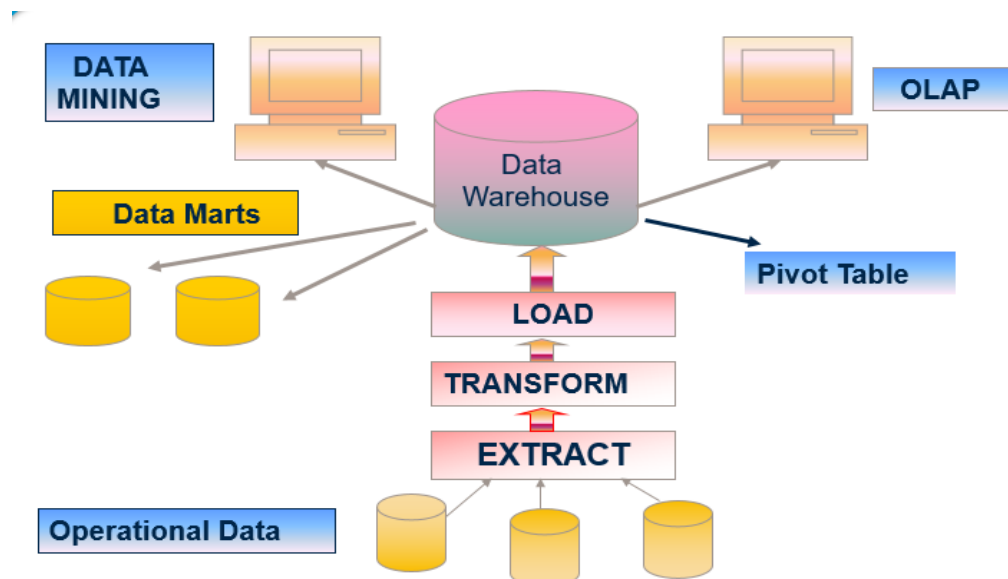


Diagram: Components of BI

What are the characteristics of Data Warehouse?

- Subject Oriented –Data grouped into particular business area instead of business as whole.
- Integrated-Data from heterogeneous sources is integrated
- Time Variant- Data warehouse keep historical data.
- Non Volatile-Data should not change once it is in warehouse.

What is Data Mart?

Data mart is subset of data warehouse which is limited for specific business unit or group of users.

- Dependent Data Mart-Data comes from enterprise data warehouse.
- Independent Data Mart-Data comes from multiple transaction system.

What is Dimension Modeling?

A dimension modeling is database structure used for data warehousing tools. It comprised of fact table and dimension table.

A fact is numeric values that can be count or sum. A dimension is information.

Type of Dimension Modeling

- Star Schema-A central fact table is surrounded by deormalized dimension tables. It may have one or more fact tables.
- Snow Flake Schema-Unlike Star schema, the dimensions table in a snowflake schema are normalized
- Galaxy Schema –Many fact table with common dimension tables

What is ETL?

A data from heterogeneous sources need to be extracted, transform and loaded to target database.

Extract is the process of reading from source database.

Transform is the process of converting extracted data from its previous form to the required form.

Loading is the process of wring transformed data to target database.

To perform ETL we use metadata such as dimension, attribute, fact, measure etc.

What is OLAP?

Online analytical processing is multidimensional analysis data model used for fast access of insight data which will simplify ad hoc queries.

Types of OLPA

- ROLAP- In ROLAP data is stored in relational database.

- MOLAP- In MOLAP data is stored in a multidimensional cube.
- HOLAP- In HOLAP combines the advantages of MOLAP and ROLAP.

What is Data Mining?

Data mining is the process of analyzing large data to discover patterns and trends that go beyond simple analysis. Data mining is also known as Knowledge Discovery in Data (KDD).