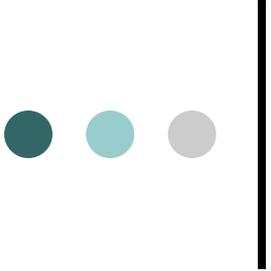


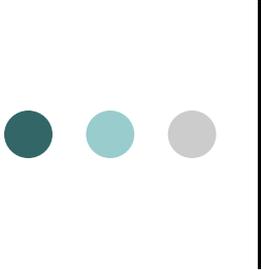
Databases and Data Warehouses





Content

- Concept Definitions of Databases, Data Warehouses
- Database models
- History
- Databases
- Data Warehouses
- OLTP vs. Data Warehouse



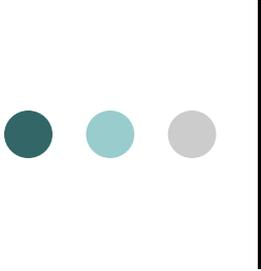
Concept Definition

Database

- a structured collection of records or data

Data Warehouse

- a logical collection of information, gathered from many different operational databases, that supports business analysis activities and decision-making tasks



Database models

- is the structure or format of a database, described in a formal language supported by the database management system

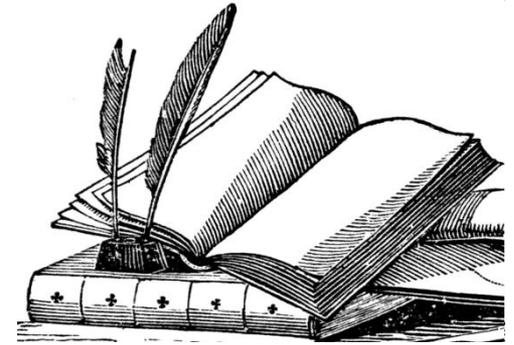
Database models

relational
flat
hierarchical
network
dimensional
object database

Data Warehouse

- relational database model

History

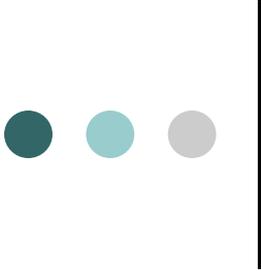


Database

- 1960 - the first database management system
- 1970 - the first relational model
- 1980 - distributed database systems and database machines
- 1990 - object-oriented databases
- 2000 - XML database

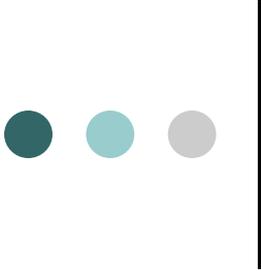
Data Warehouse

It became a distinct type of computer database during the late 1980s and early 1990s



Database

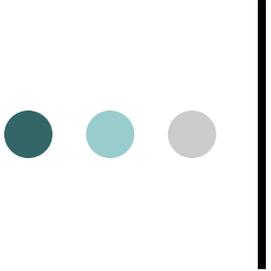
- collection of related data
- database management system (DBMS) is a collection of programs that enables users to create and maintain a database
- used in many applications
- used in all e-commerce sites to store product inventory and customer information



Data Warehouses

“A data warehouse is simply a single, complete, and consistent store of data obtained from a variety of sources and made available to end users in a way they can understand and use it in a business context.”

-- Barry Devlin, *IBM Consultant*



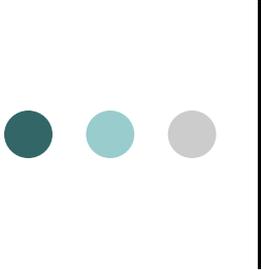
Data Warehouses

- a record of an enterprise's past transactional and operational information
- designed to favor efficient data analysis and reporting
- data warehousing is not meant for current "live" data

Data Warehouses

- large amounts of data – sometimes subdivided into smaller logical units (dependent data marts)
- data storing in a data warehouses are tematically consistent and concern concrete problem or institutions

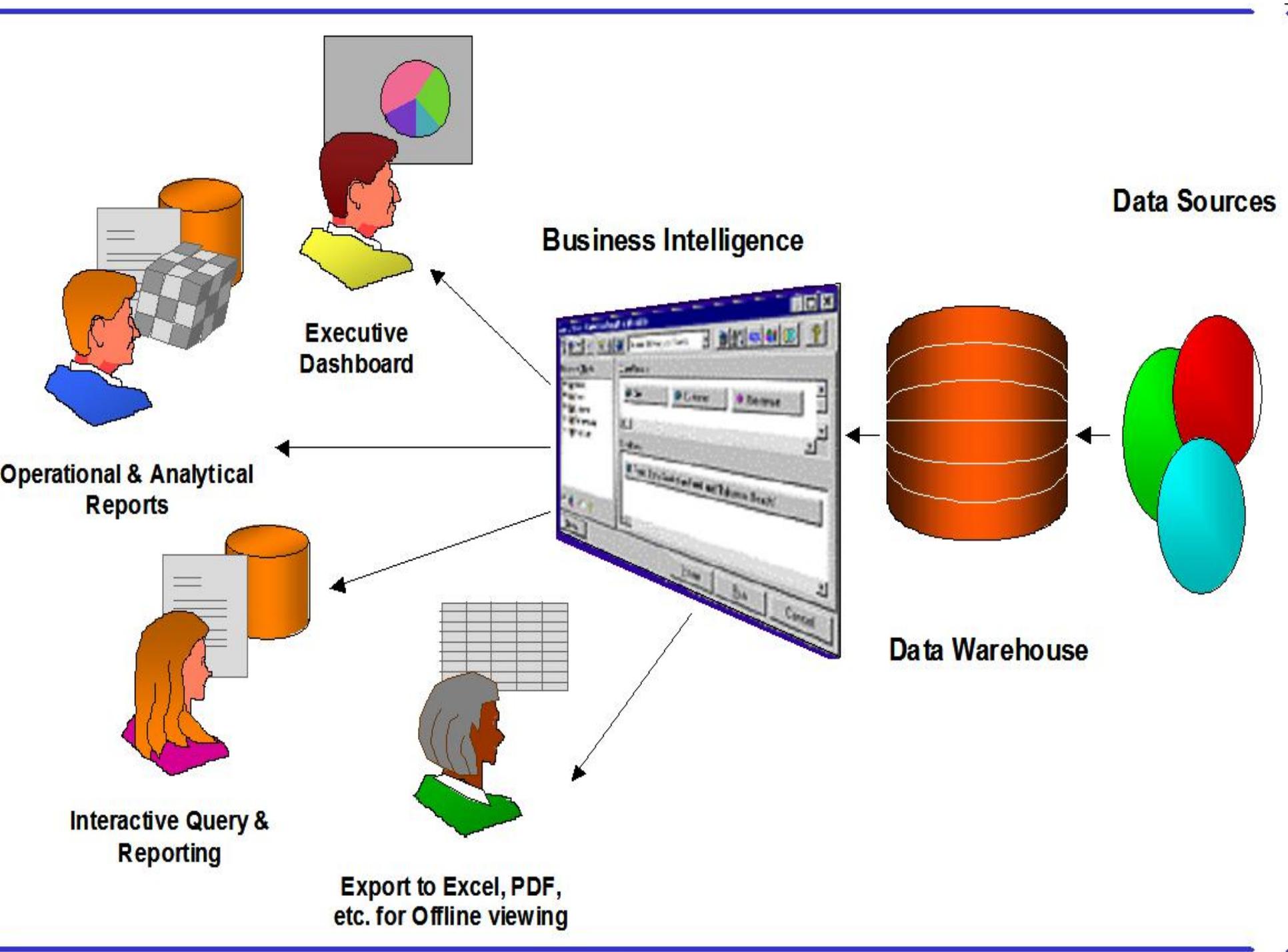




Data Warehouses

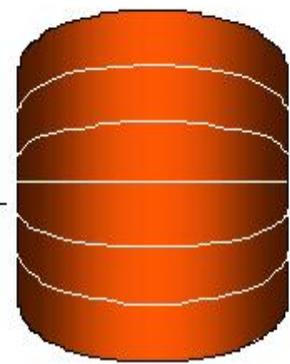
Components of a data warehouse:

- Sources -> Data Source Interaction
- Data Transformation
- Data Warehouse (Data Storage)
- Reporting (Data Presentation)
- Metadata



Data Sources

Business Intelligence



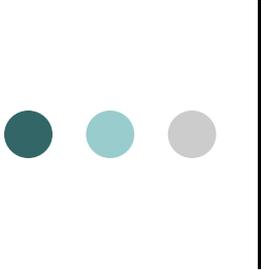
Data Warehouse

Executive Dashboard

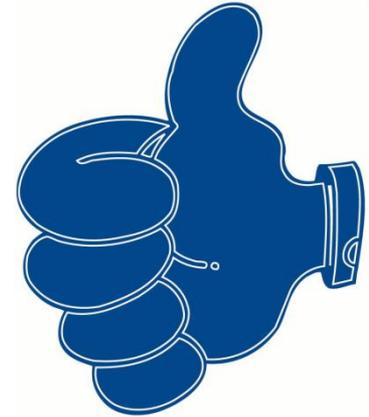
Operational & Analytical Reports

Interactive Query & Reporting

Export to Excel, PDF, etc. for Offline viewing

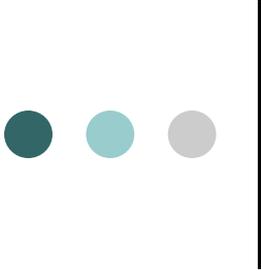


Data Warehouses ADVANTAGES

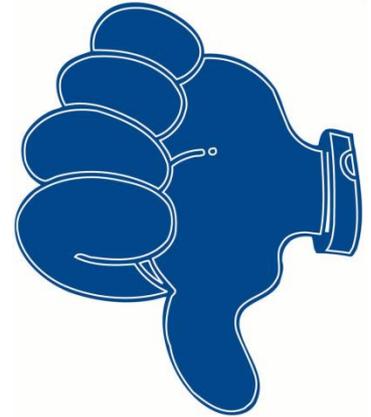


complete control over the four main areas of data management systems:

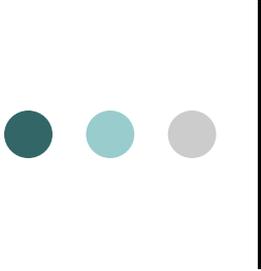
- Clean data
- Query processing: multiple options
- Indexes: multiple types
- Security: data and access



Data Warehouses DISADVANTAGES

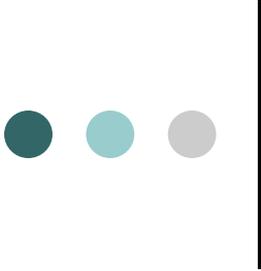


- Adding new data sources takes time and associated high cost
- Data owners lose control over their data, raising ownership, security and privacy issues
- Long initial implementation time and associated high cost
- Difficult to accommodate changes in data types and ranges, data source schema, indexes and queries



OLTP vs. OLAP

- OLTP: On Line Transaction Processing
 - Describes processing at operational sites
- OLAP: On Line Analytical Processing
 - Describes processing at warehouse

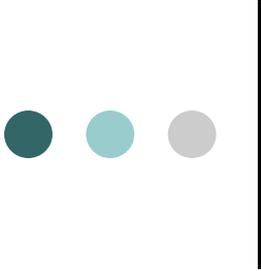


OLTP Database

vs.

Data Warehouse

- **relational databases** - groups data using common attributes found in the data set
- objectives are different

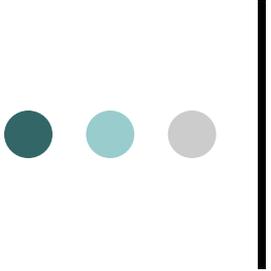


OLTP database

Designed for real
time business
operations

Data Warehouse

Designed for analysis of
business measures by
categories and
attributes

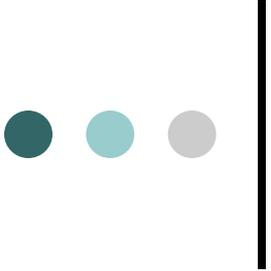


OLTP database

- Mostly updates
- Many small transactions
- Mb - Gb of data

Data Warehouse

- Mostly reads
- Queries are long and complex
- Gb - Tb of data



OLTP database

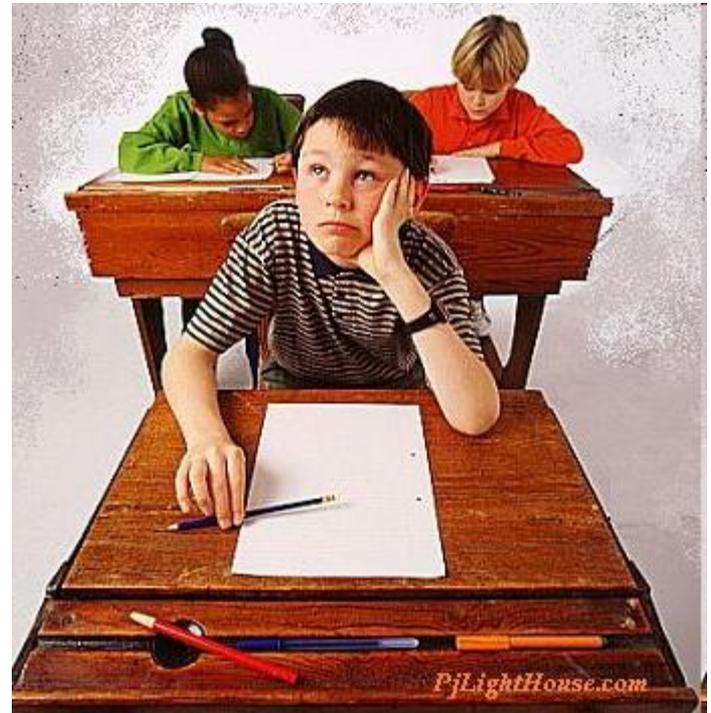
- Current snapshot
- Raw data
- Thousands of users (e.g., clerical users)

Data Warehouse

- History
- Summarized, reconciled data
- Hundreds of users (e.g., decision-makers, analysts)

SUMMARY

four questions for you 😊



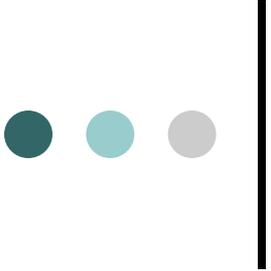


1

Designed for analysis of
business measures by
categories and
attributes

2

Designed for real
time business
operations

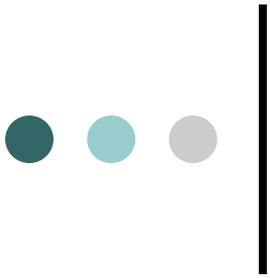


Data Warehouse

Designed for analysis of
business measures by
categories and
attributes

OLTP database

Designed for real
time business
operations

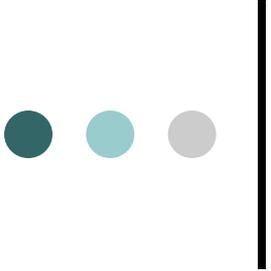


1

Optimized for a common set of transactions, usually adding or retrieving a single row at a time per table.

2

Optimized for bulk loads and large, complex, unpredictable queries that access many rows per table.

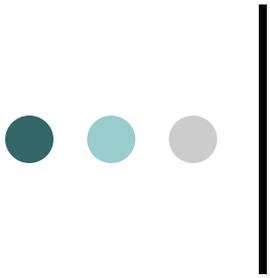


OLTP database

Optimized for a common set of transactions, usually adding or retrieving a single row at a time per table.

Data Warehouse

Optimized for bulk loads and large, complex, unpredictable queries that access many rows per table.

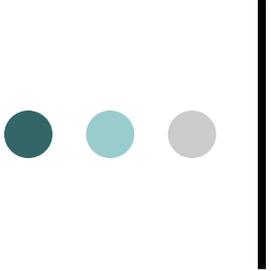


1

Optimized for validation of incoming data during transactions; uses validation data tables.

2

Loaded with consistent, valid data; requires no real time validation.

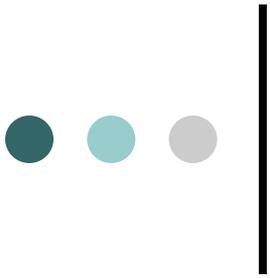


OLTP database

Optimized for validation of incoming data during transactions; uses validation data tables.

Data Warehouse

Loaded with consistent, valid data; requires no real time validation.

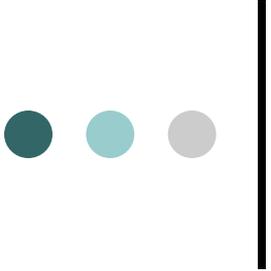


1

Supports few concurrent users relative to OLTP.

2

Supports thousands of concurrent users.

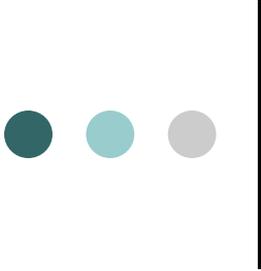


Data Warehouse

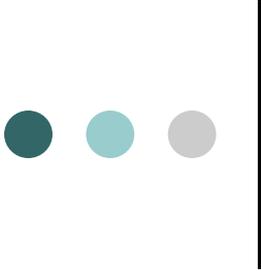
OLTP database

Supports few concurrent users relative to OLTP.

Supports thousands of concurrent users.



Bandingkan databases
dan data warehouses
dengan OLTP and
OLAP. Dalam
Membandingkan gunakan
contoh kasus (bebas)



Sources

- www.wikipedia.com
- www.exforsys.com/tutorials
- www.toolbox.com (blog)
- www.personal.uncc.edu
- www.inf.unibz.it/~franconi/teaching/2001/CS636/CS636-dw-intro.ppt



for your attention!