RAMA UNIVERSITY

www.ramauniversity.ac.in

FACULTY OF EGINEERING

DATA MINING & WAREHOUSEING
LECTURE-05

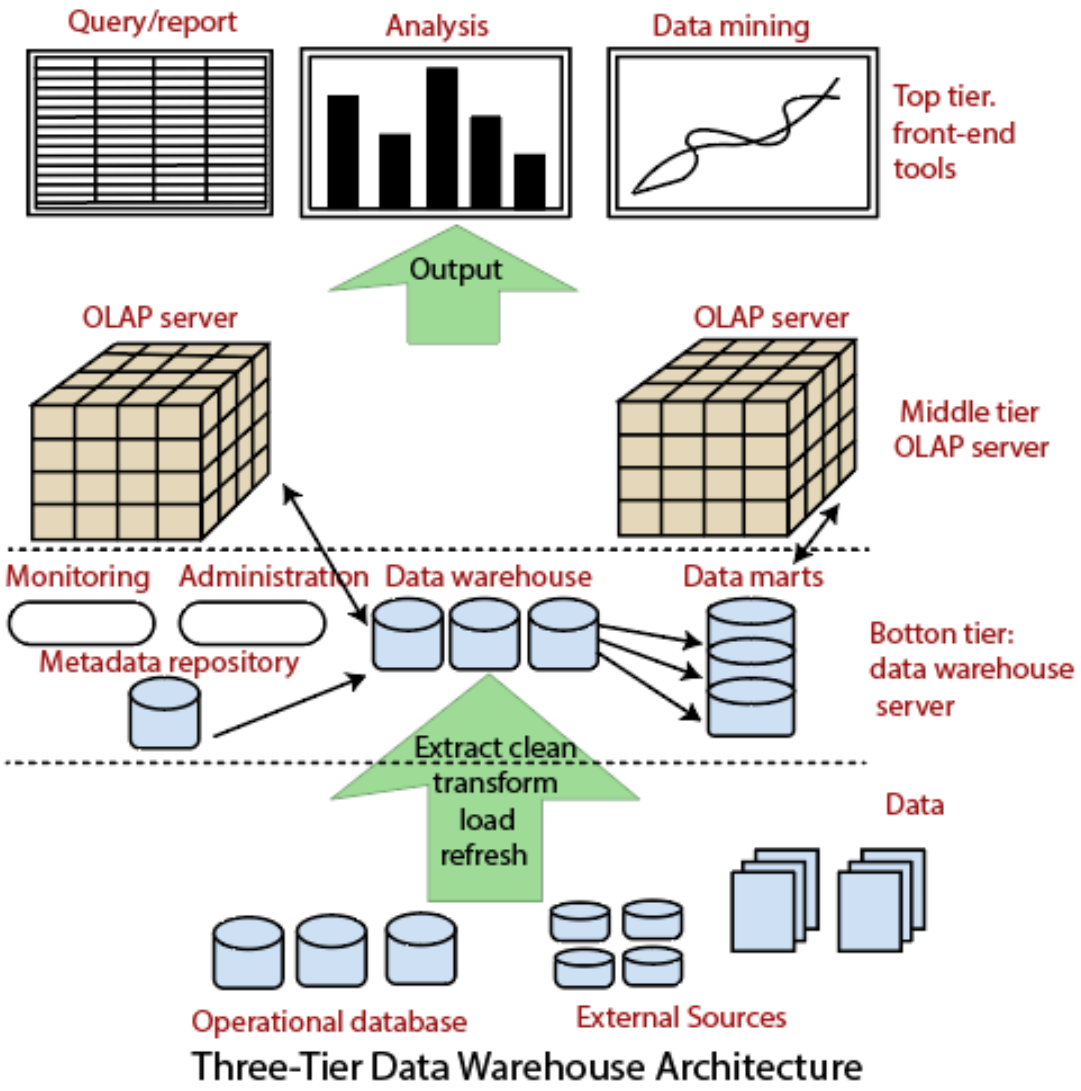
MR. DHIRENDRA
ASSISTANT PROFESSOR
RAMA UNIVERSITY

OUTLINE

- THREE-TIER DATA WAREHOUSE ARCHITECTURE
- BOTTOM TIER (DATA WAREHOUSE SERVER)
- MIDDLE TIER (OLAP SERVER)
- TOP TIER (FRONT END TOOLS).
- PRINCIPLES OF DATA WAREHOUSING
- MCQ
- REFERENCES

RAMA UNIVERSITY
UTTAR PRADESH
KANPUR

THREE-TIER DATA WAREHOUSE ARCHITECTURE



Three-Tier Data Warehouse Architecture

BOTTOM TIER (DATA WAREHOUSE SERVER)

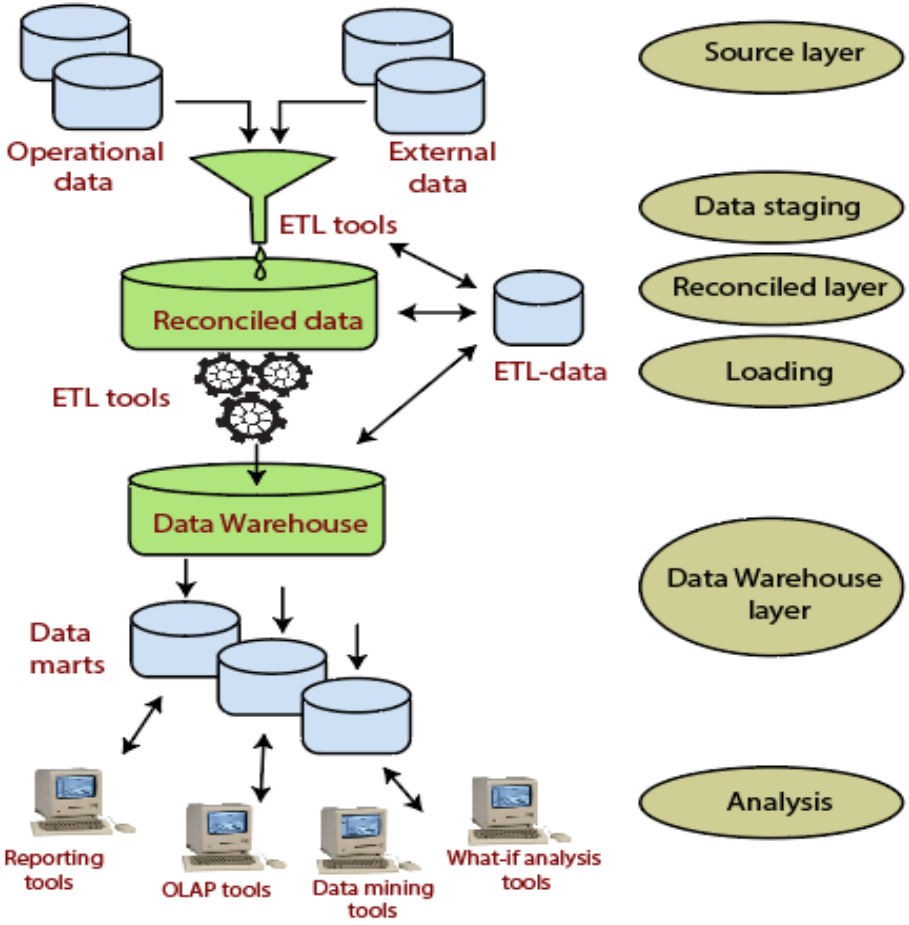
- bottom-tier that consists of the Data Warehouse server
- which is almost always an RDBMS.
- include several specialized data marts and a metadata repository.

- Data from operational databases and external sources (such as user profile data provided by external consultants) are extracted using application program interfaces called a gateway.
- A gateway is provided by the underlying DBMS and allows customer programs to generate SQL code to be executed at a server.

RAMA UNIVERSITY UTTAR PRADESH

ARCHITECTURE OF DATA MINING SYSTEM

- Examples of gateways contain ODBC (Open Database Connection) and OLE-DB (Open-Linking and Embedding for Databases), by Microsoft, and JDBC (Java Database Connection).



Three-Tier Architecture for a data warehouse system

MIDDLE-TIER

A middle-tier which consists of an OLAP server for fast querying of the data warehouse.

The OLAP server is implemented using either

(1) A Relational OLAP (ROLAP) model, i.e., an extended relational DBMS that maps functions on multidimensional data to standard relational operations.

(2) A Multidimensional OLAP (MOLAP) model, i.e., a particular purpose server that directly implements multidimensional information and operations.

RAMA UNIVERSITY UTTAR PRADESH KANPUR

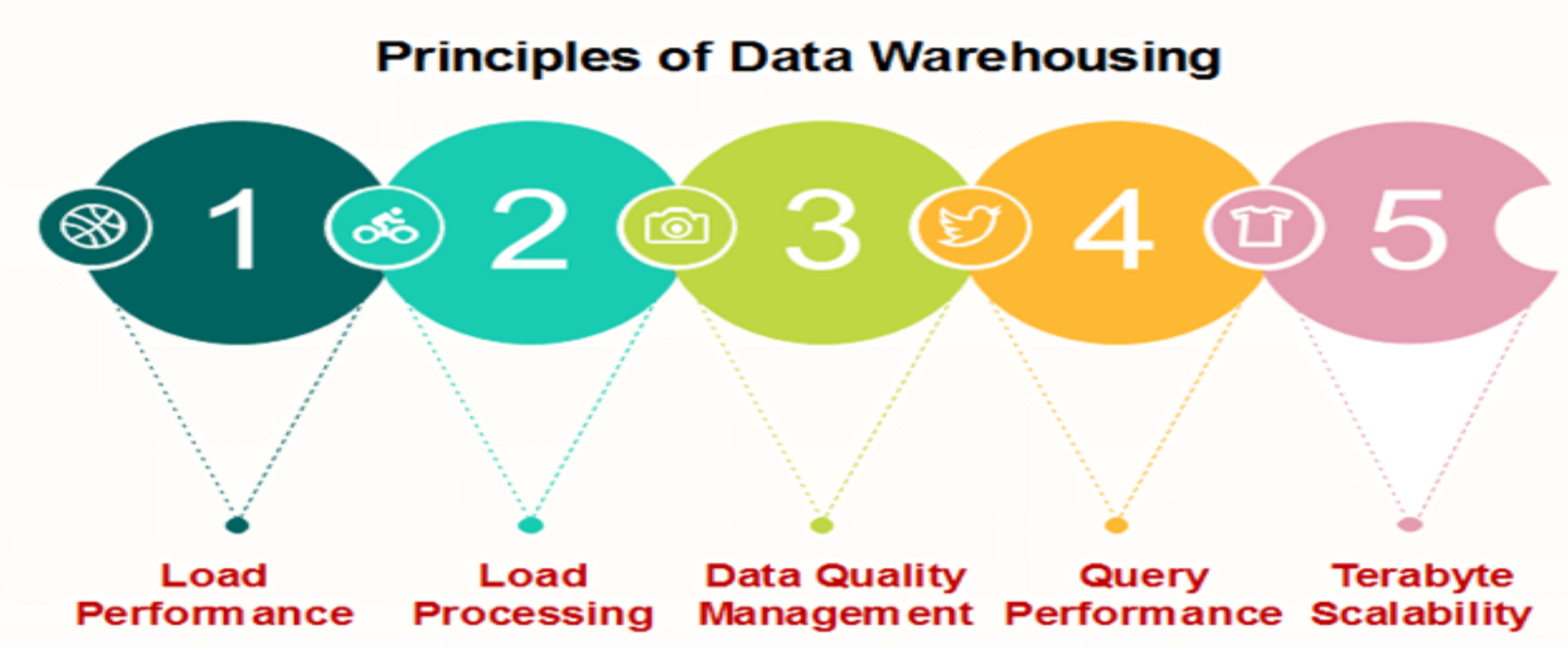
TOP-TIER

A top-tier that contains front-end tools for displaying results provided by OLAP, as well as additional tools for data mining of the OLAP-generated data.

The **metadata repository** stores information that defines DW objects. It includes the following parameters and information for the middle and the top-tier applications:

- A description of the DW structure, including the warehouse schema, dimension, hierarchies, data mart locations, and contents, etc.
- Operational metadata, which usually describes the currency level of the stored data, i.e., active, archived or purged, and warehouse monitoring information, i.e., usage statistics, error reports, audit, etc.
- System performance data, which includes indices, used to improve data access and retrieval performance.
- Information about the mapping from operational databases, which provides source RDBMSs and their contents, cleaning and transformation rules, etc.
- Summarization algorithms, predefined queries, and reports business data, which include business terms and definitions, ownership information, etc.

PRINCIPLES OF DATA WAREHOUSING



Multiple Choice Question

1. _____databases are owned by particular departments or business groups.
A. Informational.
B. Operational.
C. Both informational and operational.
D. Flat

2.. The star schema is composed of _____ fact table.
a) one
b) two
c) three
d) Four

3. The time horizon in operational environment is _____.
a) 30-60 days.
b) 60-90 days.
c) 90-120 days.
d) 120-150 days.

4. The key used in operational environment may not have an element of_____.
a) time
b) cost
c) frequency
d) quality

5. Data can be updated in _____environment.
a) data warehouse.
b) data mining.
c) operational. I
d) nformational.

RAMA UNIVERSITY UTTAR PRADESH KANPUR

REFERENCES

- https://www.tutorialspoint.com/dwh/dwh_overview.htm

- http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf
  DATA MINING BOOK WRITTEN BY Micheline Kamber
- https://www.javatpoint.com/three-tier-data-warehouse-architecture