FACULTY OF EGINEERING

DATA MINING & WAREHOUSEING
LECTURE-08

MR. DHIRENDRA
ASSISTANT PROFESSOR
RAMA UNIVERSITY

OUTLINE

❖ **NEED A DATA MART**

❖ **COST-EFFECTIVE DATA MARTING**

❖ **IDENTIFY THE FUNCTIONAL SPLITS**

❖ **IDENTIFY USER ACCESS TOOL REQUIREMENTS**

❖ **DESIGNING DATA MARTS**

❖ **COST OF DATA MARTING**

❖ **MCQ**

❖ **REFERENCES**

Need a Data Mart

- To partition data in order to impose access control strategies.

- To speed up the queries by reducing the volume of data to be scanned.

- To segment data into different hardware platforms.

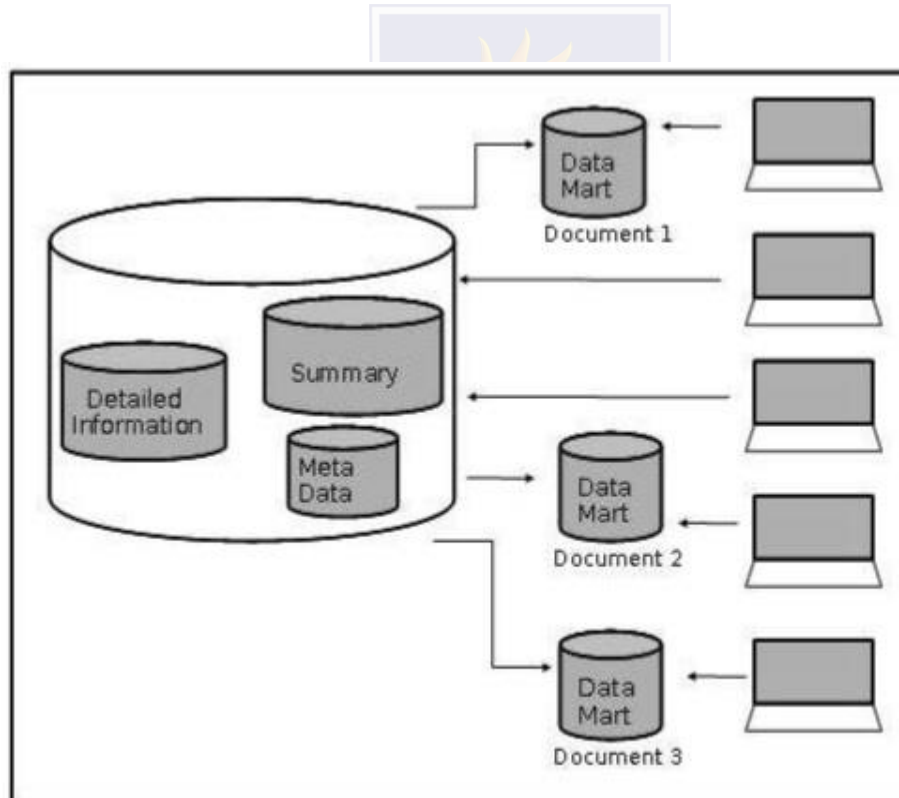- To structure data in a form suitable for a user access tool.

Cost-effective Data Marting
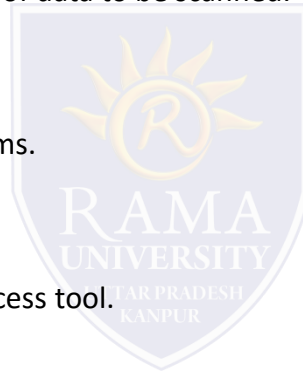
- **Identify the Functional Splits**

- **Identify User Access Tool Requirements**

- **Identify Access Control Issues**

- sales transaction on a daily basis

- sales forecast on a weekly basis

- stock position on a daily basis

- stock movements on a daily basis

Identify User Access Tool Requirements

- We need data marts to support user access tools that require internal data structures. The data in such structures are outside the control of data warehouse but need to be populated and updated on a regular basis.

- There are some tools that populate directly from the source system but some cannot. Therefore additional requirements outside the scope of the tool are needed to be identified for future.

- In order to ensure consistency of data across all access tools, the data should not be directly populated from the data warehouse, rather each tool must have its own data mart.
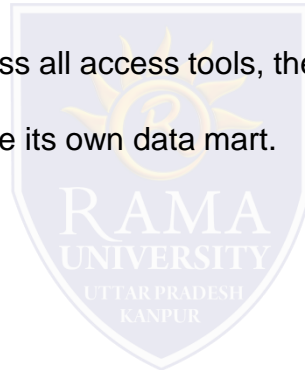
Identify Access Control Issues

- There should to be privacy rules to ensure the data is accessed by authorized users only. For example a data warehouse for retail banking institution ensures that all the accounts belong to the same legal entity. Privacy laws can force you to totally prevent access to information that is not owned by the specific bank.

- Data marts allow us to build a complete wall by physically separating data segments within the data warehouse. To avoid possible privacy problems, the detailed data can be removed from the data warehouse. We can create data mart for each legal entity and load it via data warehouse, with detailed account data.
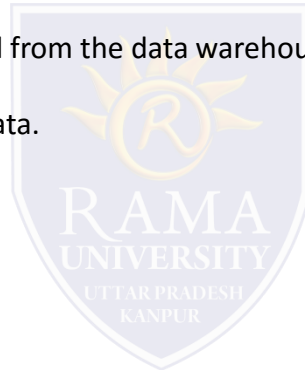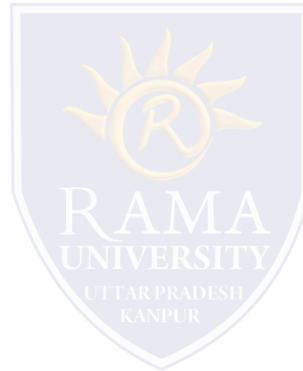
Designing Data Marts

- Data marts should be designed as a smaller version of starflake schema within the data warehouse and should match with the

  database design of the data warehouse. It helps in maintaining control over database instances.
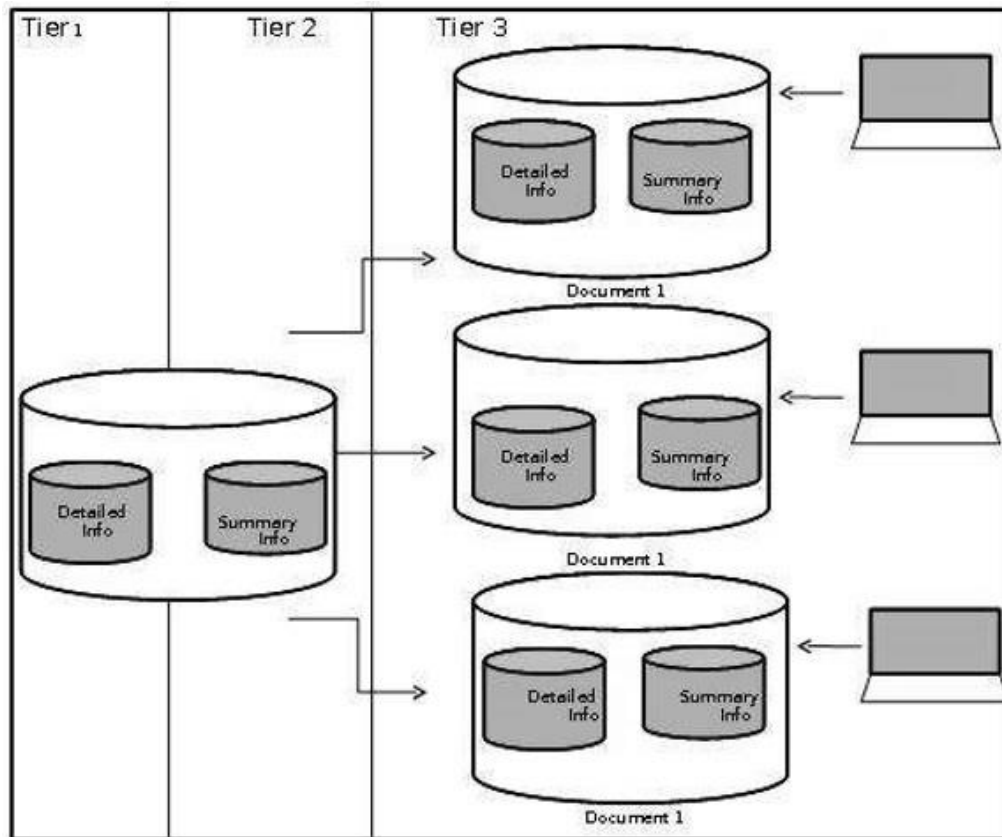
Cost of Data Marting

- **Hardware and Software Cost**

- **Network Access**

- **Time Window Constraints**

Multiple Choice Question

1. The type of relationship in star schema is
_____.
a)   many-to-many.
b)   one-to-one.
c)   one-to-many.
d)   many-to-one.

2.. Fact tables are _____.
a)   completely demoralized.
b)   partially demoralized.
c)   completely normalized.
d)   partially normalized.

3. _____ is the goal of data
mining.
a)   To explain some observed event or
     condition.
b)   To confirm that data exists.
c)   To analyze data for expected
     relationships.
d)   To create a new data warehouse.

4. Business Intelligence and data warehousing is used
for _____.
a)   Forecasting
b)   Data Mining.
c)   Analysis of large volumes of product sales data.

d)   All of the above.

5. The data administration subsystem helps you
perform all of the following, except_____.

a)   backups and recovery.
b)   query optimization.
c)   security management.
d)   create, change, and delete information.
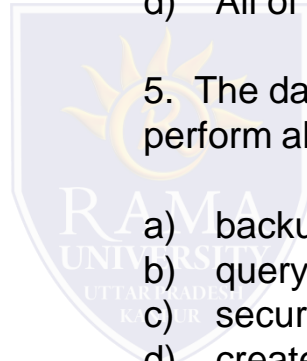
REFERENCES

- 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