Database Management and Warehousing: Key Concepts (Brief Overview)

1. ER Model (Entity-Relationship Model)

Defines data elements and their relationships using:

- Entities: Real-world objects (e.g., Student, Course).

- Attributes: Properties (e.g., Name, Roll Number).

- Relationships: Associations (e.g., Enrolls in).

Example: A Student entity with attributes (ID, Name) and a relationship Enrolls with Course.

2. Relational Model

Data is stored in tables (relations) with:

- Primary Key: Unique identifier.

- Foreign Key: Links to another table.

Example: A Students table with columns (ID, Name) and a Courses table linked by CourseID.

3. Relational Algebra

Query operations include:

- Selection (Selection): Selects rows. Example: Selection(Age>20)(Students)

- Projection (Projection): Selects columns. Example: Projection(Name, Age)(Students)

- Join: Combines tables based on a condition.

4. Tuple Relational Calculus

Specifies what data to retrieve using conditions.

Example: {t | t in Students AND t.Age > 20}

5. SQL (Structured Query Language)

Database Management and Warehousing: Key Concepts (Brief Overview)

Used to manage relational databases:

- DDL: CREATE TABLE, ALTER TABLE.

- DML: SELECT, INSERT, UPDATE, DELETE.

Example: SELECT Name FROM Students WHERE Age > 20;

6. Integrity Constraints

Ensure data accuracy:

- Primary Key Constraint: Ensures uniqueness.

- Referential Integrity: Ensures valid foreign key values.

7. Normal Forms

Eliminate redundancy:

- 1NF: Atomic values.

- 2NF: No partial dependency.

- 3NF: No transitive dependency.

Example: A table split into smaller tables to avoid duplication.

8. File Organization

Methods for data storage:

- Heap Files: Unordered.

- Sorted Files: Ordered by key.

Example: Storing student records in sorted order by ID.

9. Indexing

Speeds up data retrieval:

Database Management and Warehousing: Key Concepts (Brief Overview)

- Primary Index: Based on primary key.

- Secondary Index: On non-key attributes.

Example: Indexing student names for faster search.

10. Data Types

Defines the kind of data:

- Numeric: INT, FLOAT.

- Character: CHAR, VARCHAR.

- Date/Time: DATE, TIMESTAMP.

11. Data Transformation

Prepares data for analysis:

- Normalization: Scales values.

- Discretization: Converts continuous data to intervals.

Example: Normalizing ages to a range of 0-1.

12. Data Warehouse Modelling

Schemas for organizing data:

- Star Schema: Central fact table with dimension tables.

- Snowflake Schema: Normalized dimensions.

Example: Sales fact table linked to Product, Time dimensions.

13. Measures

Types of measures:

- Additive: Can be summed (e.g., Sales).

Database Management and Warehousing: Key Concepts (Brief Overview)

- Non-Additive: Cannot be summed (e.g., Ratios).

Example: Computing total sales by summing sales values.