

* ASSIGNMENT 2 *

Date: 14.11.1

(1) What is Metadata?

- Metadata are data about data. When used in data warehouse, metadata are data that define warehouse objects.
- Metadata are created for data names & definition of given warehouse.
- ⇒ Metadata repository contains following:
 - Description of structure of data warehouse, which include schema, view, dimensions, & derived data defⁿ as well as data mart locⁿ and contents.
 - Operational metadata, which include data lineage, currency of data.
 - Algorithms used for summarization.
 - Data related to system performance, which includes indices & profiles that improve data access & retrieval performance in addⁿ rules for timing & scheduling of refresh, update and replication cycles.
 - Business metadata.

(2) What is Data Mart?

- Data marts contains subset of organization-wide data that is valuable to specific grp of people in an organization.

- A data mart contains only those data that is specific to particular group.
- For eg. marketing data mart may contain only data related to items, customer & sales.
- ⇒ Reasons to create data mart:-

- To speed up queries by reducing volume of data to be scanned.
- To partition data in order to impose access control strategies.
- Easy access to freq needed data.
- Create collective view by group of users.
- Improve end-user response time.
- Lower cost than implementing full data warehouse.
- contains only business essential data and is less cluttered.

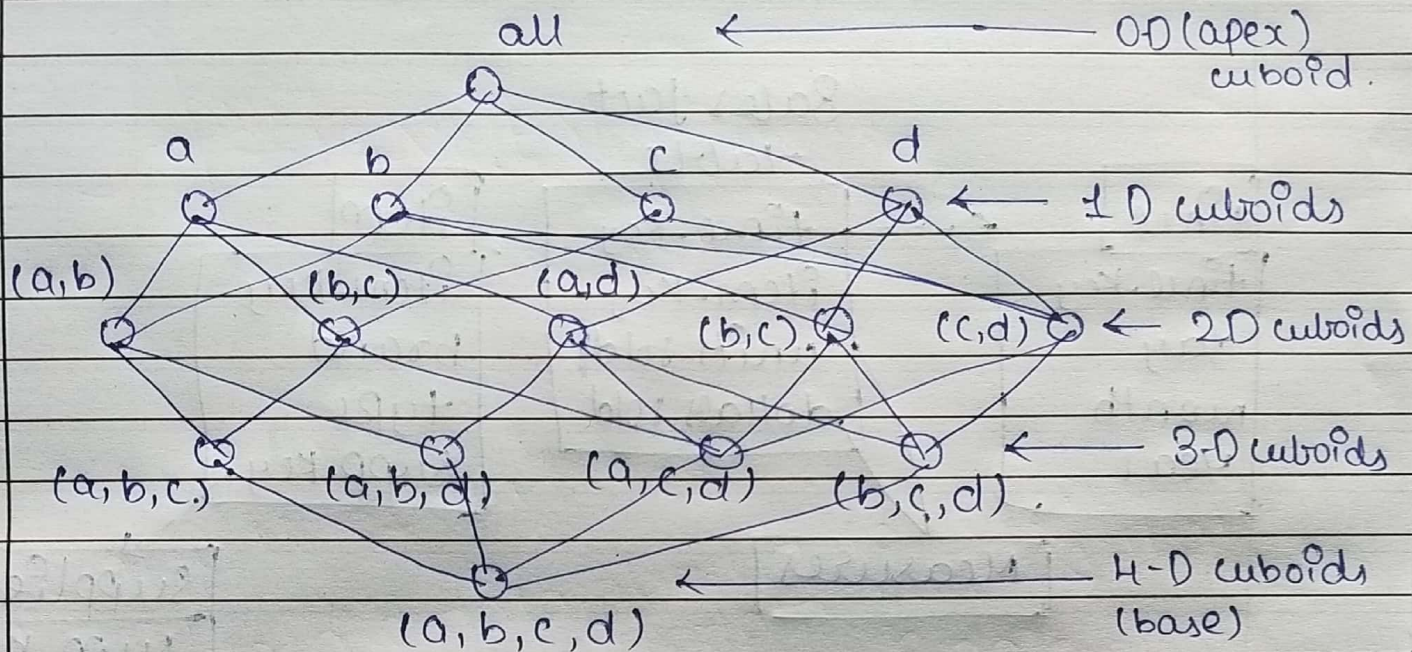
⇒ Types of data marts:

- (i) dependent data mart
- (ii) Independent data mart
- (iii) hybrid data mart

(3) What is cuboids, lattice of dimensions?

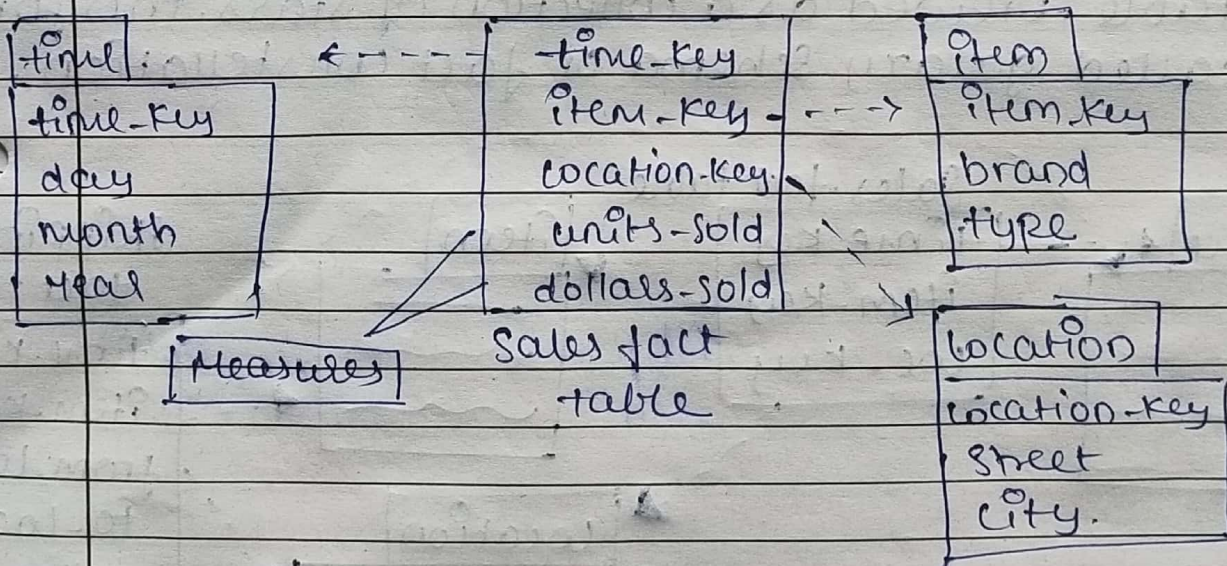
→ There are 3 types of schemas:

- (a) Star Schema
- (b) Snowflake Schema
- (c) Fact constellation



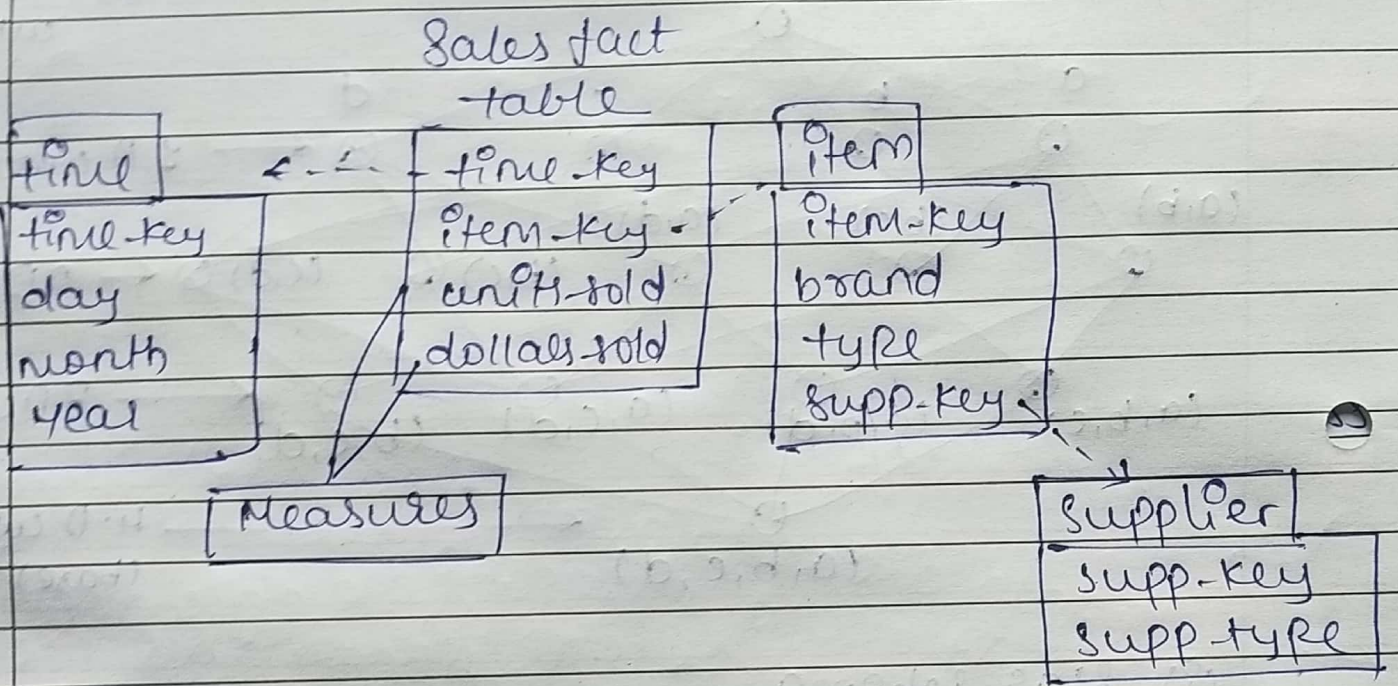
(i) Star schema.

A fact table in middle connected to a set of dimension tables.



(ii) Snowflake schema

A refinement of Star where some dimensions hierarchy is normalized into set of smaller dimension tables.



(iii) Fact constellations

Multiple fact tables share dimension table, viewed as a collection of stars, therefore called galaxy schema or fact constellation.

