Experiment No-1

Title:

CASE Study to build Data warehouse construction for star schema and snow flake schema

Objective:

Construct star schema and Snowflake schema.

Theory:

Dimensional Modelling: It is a data structure technique optimized for data storage in a Datawarehouse. The purpose of dimensional modeling is to optimize the database for faster retrieval of data. The concept of Dimensional Modelling was developed by Ralph Kimball and consists of "fact" and "dimension" tables. A dimensional model in data warehouse is designed to read, summarize, analyze numeric information like values, balances, counts, weights, etc. in a Datawarehouse. For instance, in the relational mode, normalization and ER models reduce redundancy in data, on the contrary, dimensional model in data warehouse arranges data in such a way that it is easier to retrieve information and generate reports.

Elements of Dimensional Data Model

Fact

Facts are the measurements/metrics or facts from your business process. For a sales business process, a measurement would be quarterly sales number

Dimension

Dimension provides the context surrounding a business process event. In simple terms, they give who, what, where of a fact. In the Sales business process, for the fact quarterly sales number, dimensions would be

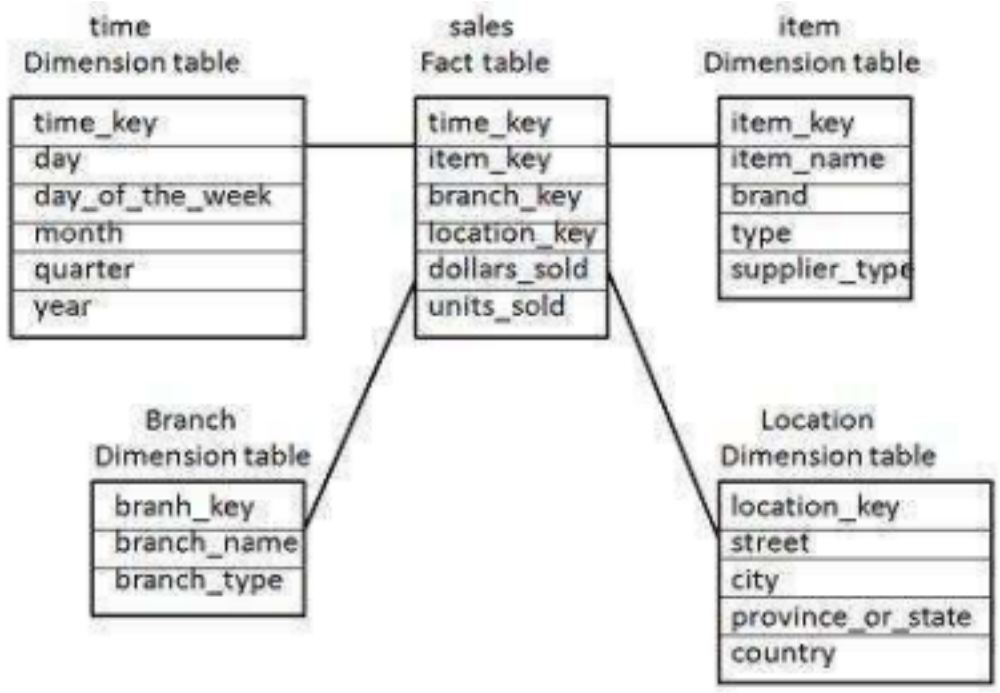
- Who-- Customer Names
- Where – Location
- What – Product Name

In other words, a dimension is a window to view information in the facts.

1. Star Schema:

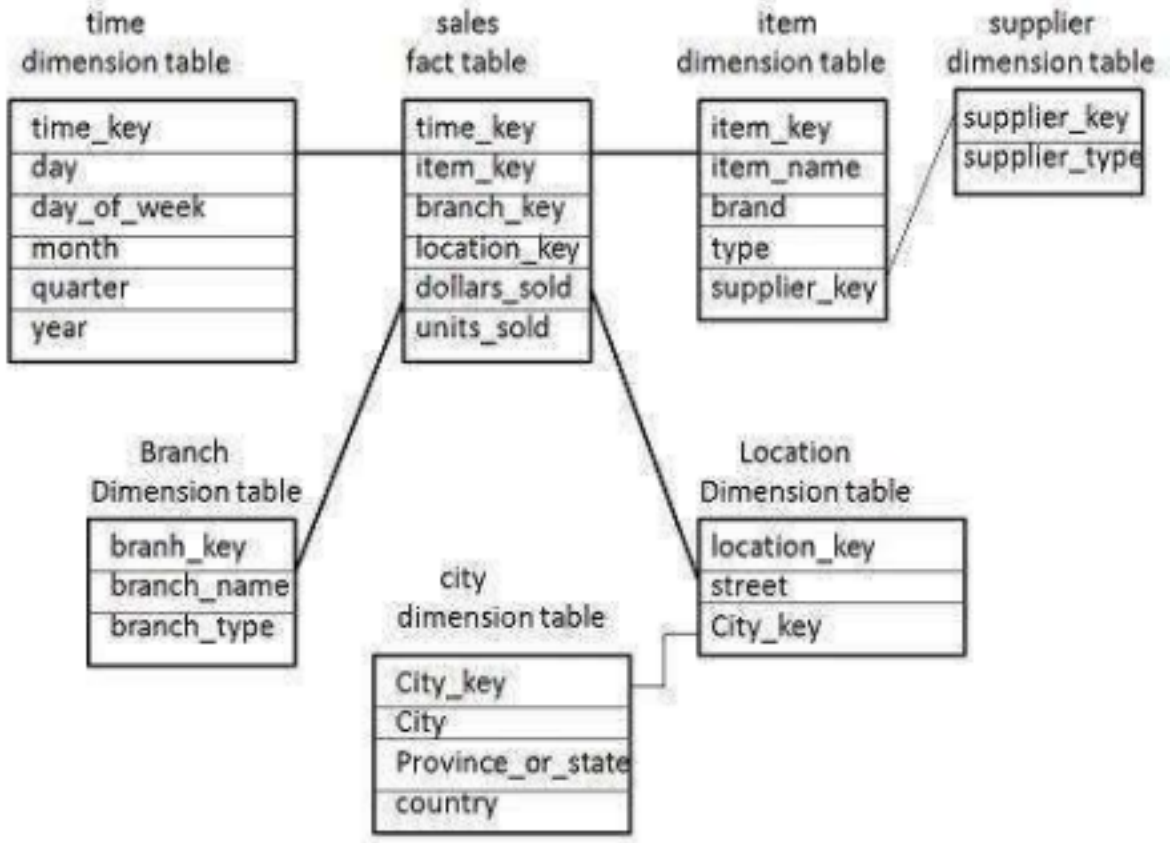
Star Schema in data warehouse, in which the center of the star can have one fact table and a number of associated dimension tables. It is known as star schema as its structure resembles a star. The Star Schema data model is the simplest type of Data Warehouse schema. It is also known as Star Join Schema and is optimized for querying large data sets.

Star Schema of Sales:



2.Snowflake Schema:

- Some dimension tables in the Snowflake schema are normalized.
- The normalization splits up the data into additional tables.
- Unlike Star schema, the dimensions table in a snowflake schema are normalized. For example, the item dimension table in star schema is normalized and split into two dimension tables, namely item and supplier table.



Now the item dimension table contains the attributes item key, item name, type, brand, and supplier-key.

- The supplier key is linked to the supplier dimension table. The supplier dimension table contains the attributes supplier key and supplier type.