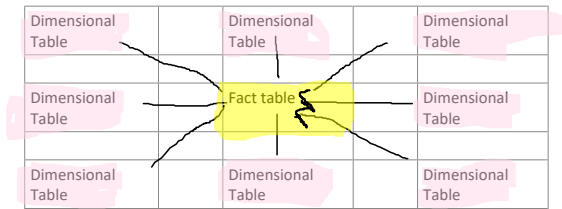# Data warehouse

Tuesday, August 29, 2023   4:08 PM

Used for optimized querying and reporting for analytical purposes
⭐ Loading the data not generating data
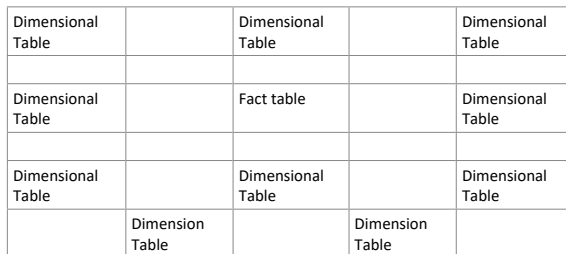
Star Schema
    Characterized by a central fact table connected to dimension tables

Dimensional Table	Dimensional Table	Dimensional Table
Dimensional Table	Fact table	Dimensional Table
Dimensional Table	Dimensional Table	Dimensional Table

What and by What?

Dimension table
    Attribute data (By What)
Fact Table
    Metric data (What)
    Foreign key to dim table

Snowflakes schema

Dimensional Table		Dimensional Table		Dimensional Table
Dimensional Table		Fact table		Dimensional Table
Dimensional Table		Dimensional Table		Dimensional Table
	Dimension Table		Dimension Table	

Dimensional Table

Primary key - unique identification of record in table (generated in warehouse)
    E.g. Sr. No.
Natural key - column used to identify source data uniquely
    E.g. Contact, email
Attribute - help in categorizing and analyzing data
    E.g. Name, Age, ...
...

Audit Column - tracking the data changes and updates
    Eg. Created date, updated date

Fact Table

Primary key
Foreign key of all dim table
    Eg. FK_Cust, FK_Date, FK_Product
Metric value
    Eg. Qty, TotalSales
Audit column
    Create date
    Update date

Galaxy schema

Types of dimension
Rapidly changing dimension

Slowly changing dimension
    Type1 - Changes are overwritten and no history is maintained
        CreatedDate
        UpdatedDate
    Type2 - Preserves history
        StartDate
        EndDate
        Version
        Flag (Valid/Not valid)
    Type3 - Preserves previous value.
        Pre_Attribute cols
        CreatedDate
        UpdatedDate

Unchanging dimension

Denormalization