ASSIGNMENT-11

September 16, 2025

DATA MODELLING:

NORMALIZATION:

Dividing single table into multiple tables with some reference columns.

Example: customer data

CID	CN	CP	CE	PN	PQ	PID	PP
1	A	291	xy@gmail.com	Milk	1 L	P1	52
2	B	.	.	Rice	1 KG	P2	60
3	C	.	.	.	.	P1	.
.	.	.	.	.	.	.	.
100	ABC	892	.	.	.	.	.

- It is the process of splitting the database into multiple tables so that we can reduce redundancy and improve data integrity.
- In powerBI, fact and dimension tables are two types of tables used to build the data model.

Fact: fact tables will have measurable or quantitative data / Numerical data.

Dimension: Dimension table will contain unique categorical data.

Data Model creation: Cardinality is the relationship between two tables in the data model.

It is defined by 4 types:

1. One to one
2. One to many
3. Many to one
4. Many to many

Why data modelling?

- To remove unnecessary columns and rows.
- To perform group by and summarize
- Optimise column and data types

- Use custom columns.

Schema:

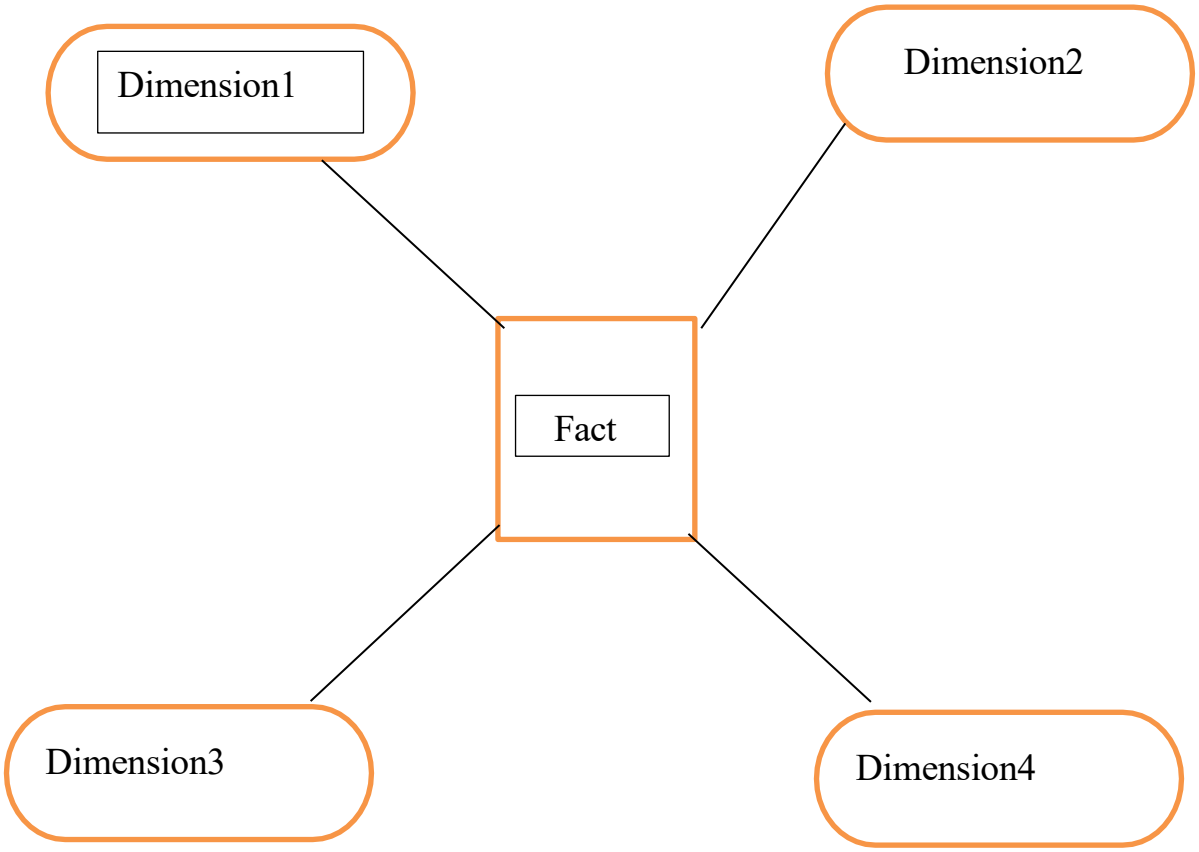
It is the structure of data model that defines how the data is connected and organized.

There are 3 types:

1. Star schema
2. Snow Flake Schema
3. Galaxy schema

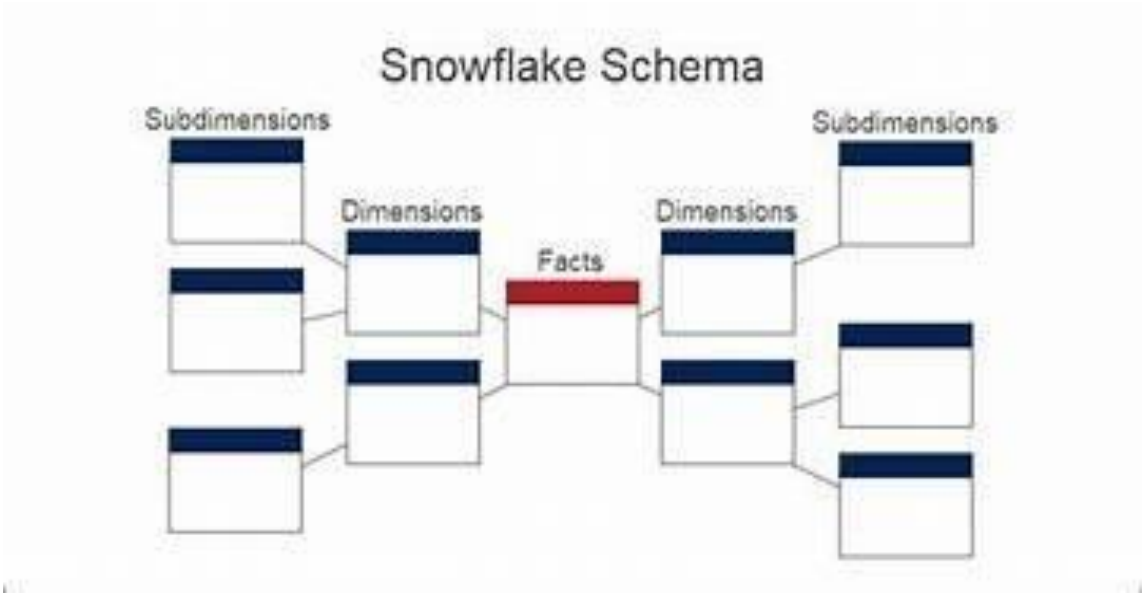
1. Star Schema:

- One fact table to this will connect all dimension table
- Dimension table in a star schema align with fact giving it a star shape.



2. Snow-Flake Schema:

- Dimension table have sub-division table. Snowflake schema use sub-division to represent additional joins in queries.



- It is not compulsory that all the dimensions should have sub-dimensions.

3. Galaxy Schema:

- This schema contains more than one fact table linked to dimension tables.