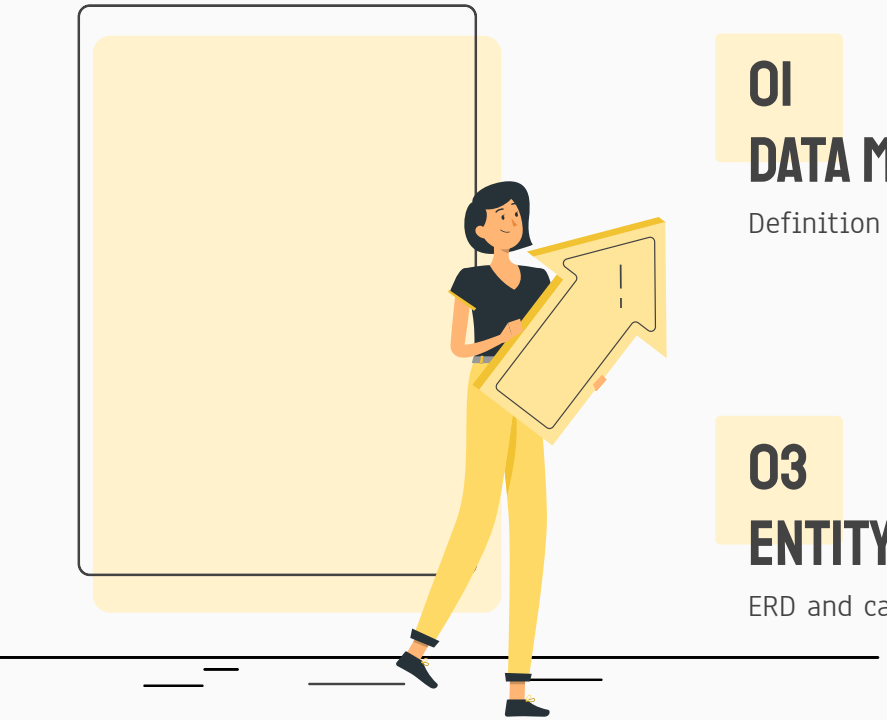
DATA MODELLING

Ari Sulistiyo Prabowo



01
DATA MODELLING

Definition and purpose

02
PHASE DATA MODELLING

Conceptual, Logical and
Physical data modelling

03
ENTITY RELATIONSHIP

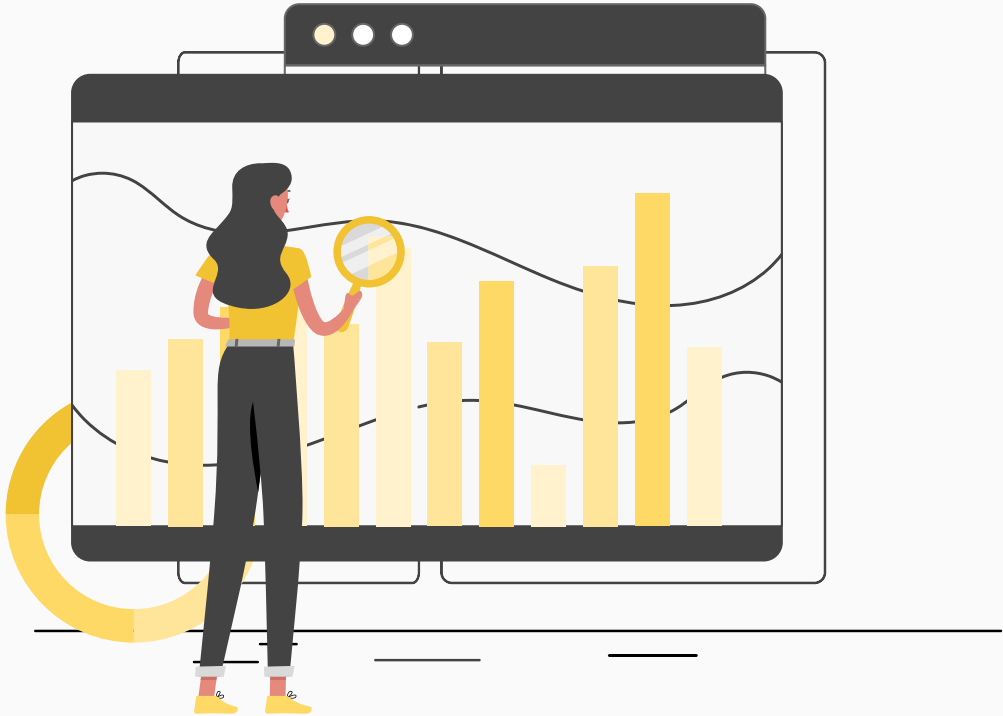
ERD and cardinality

04
TYPES OF SCHEMA

Primary and Foreign key,
Star and Snowflake schema

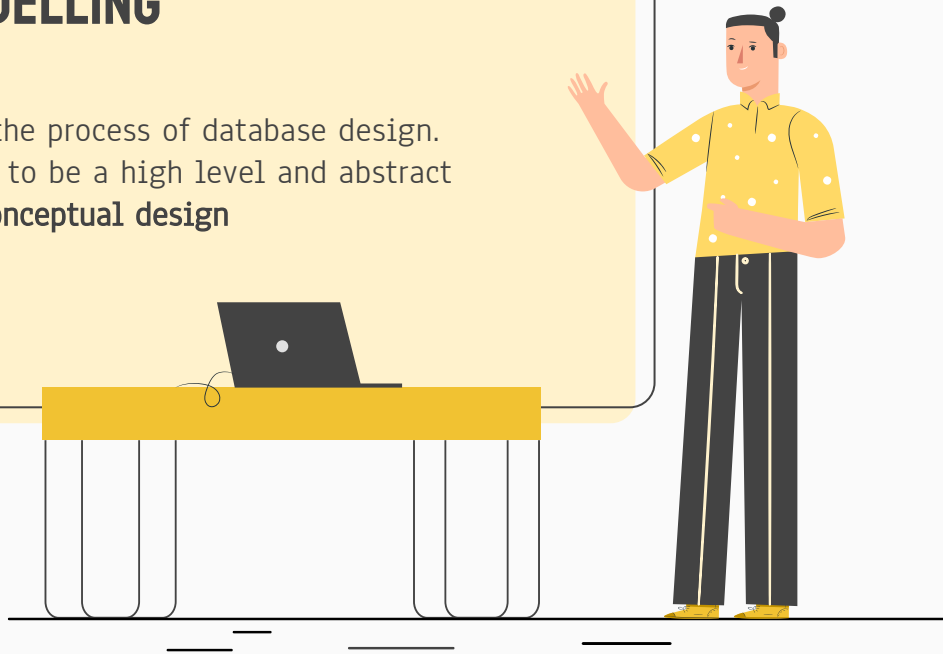
01. DATA MODELLING

Understanding and purpose of data modelling



DATA MODELLING

Data modelling is the **first step** in the process of database design. This first step sometimes considered to be a high level and abstract design, referred as **conceptual design**



THE PURPOSE OF DATA MODELLING

01

- The data contained in database (entities: students, department, courses)

02

- The relationships between data items (students are supervised by lecture)

03

- The constraint of data (student number has eight number, subject has four credits)

BASIC CONSTRUCT OF DATA MODELLING

ENTITY

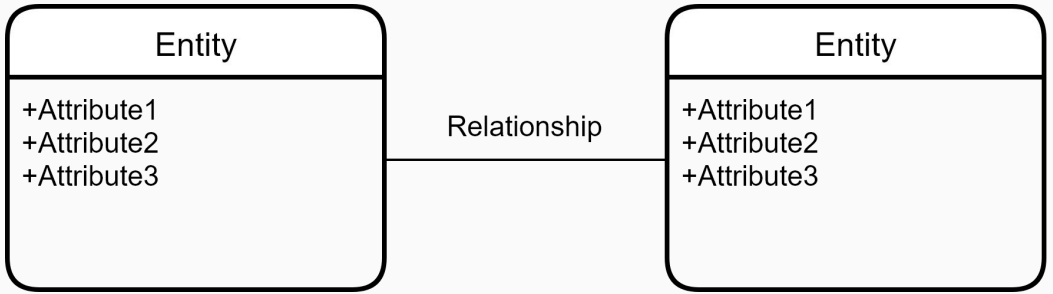
Entity is the principal data object about which information is to be **collected**. It is analogous to **table** in database

ATTRIBUTE

Attribute describes the entity of which they are associated. It is analogous to **column** in database

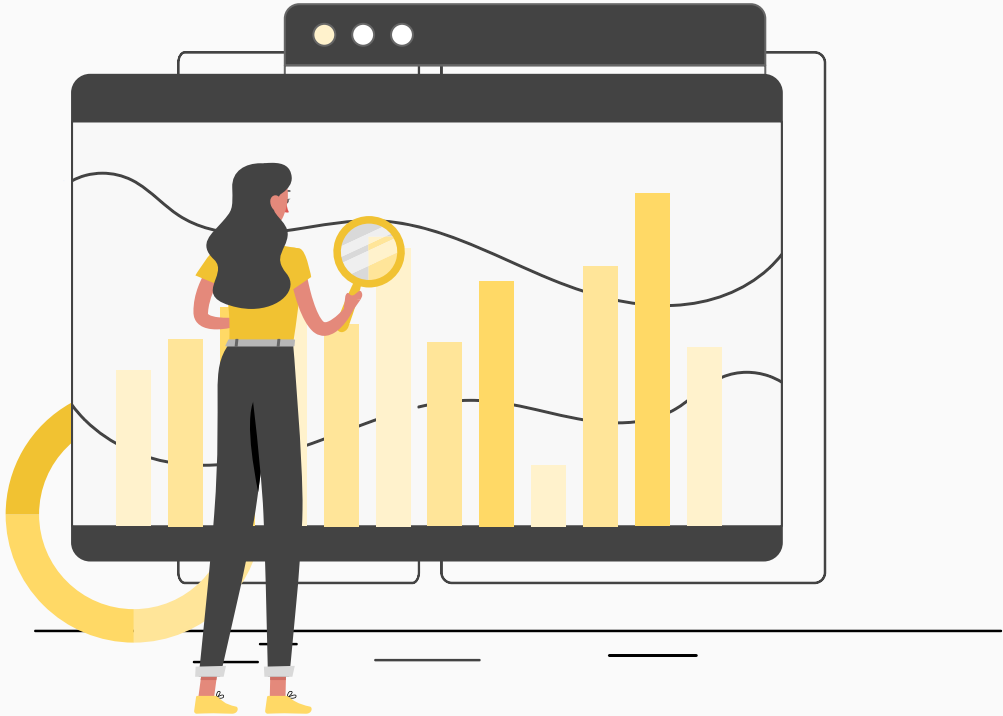
RELATIONSHIP

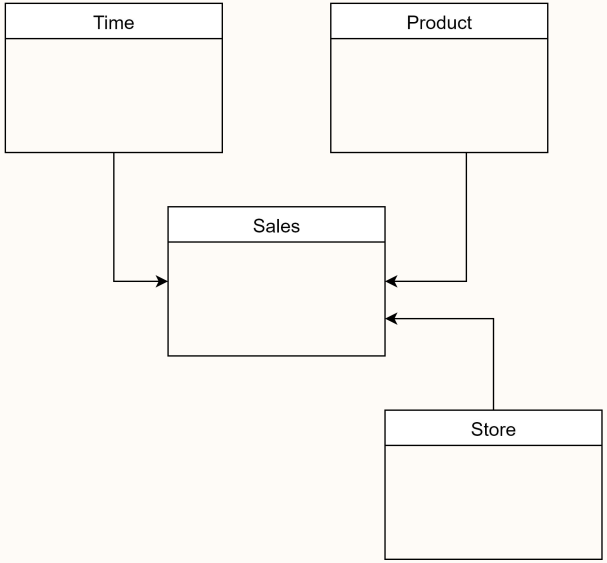
A relationship represents an association between **two or more entities**.
Example: employees are assigned to projects



02. PHASE OF DATA MODELLING

Conceptual, Logical and Physical data modelling

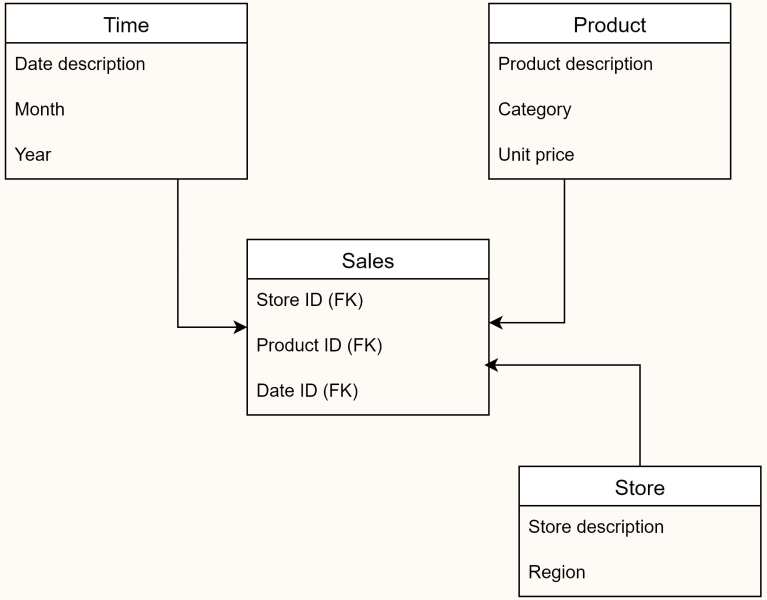




Conceptual data is used by **business stakeholder**. The purpose is to **organize**, **scope** and **define** the business concept

- ❏ Provide flexible data-structuring
- ❏ Easily understood and enhanced
- ❏ Only entities visible
- ❏ Abstract relationship
- ❏ The basis for identification and high-level description of main data objects; they avoid details

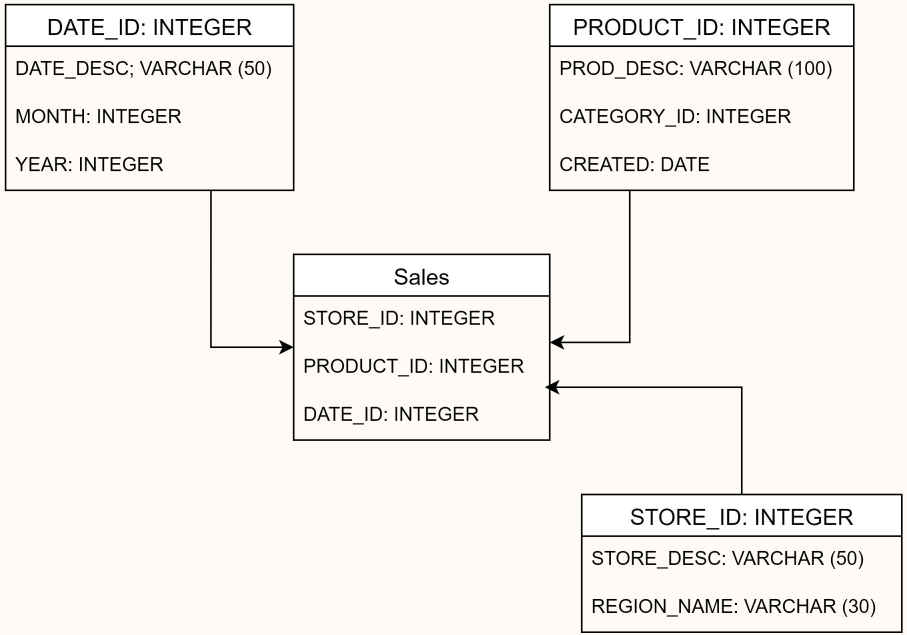
—CONCEPTUAL DATA MODELLING



Logical data is used by **data architects** and **business analyst**. The purpose is to **develop** the technical map of rules and data structure

- Presences of attribute for each entities
- Primary key and foreign key relationship
- User friendly attribute name
- More detail than conceptual model
- Database agnostic
- Bit more effort and required to enhance

—LOGICAL DATA MODELLING



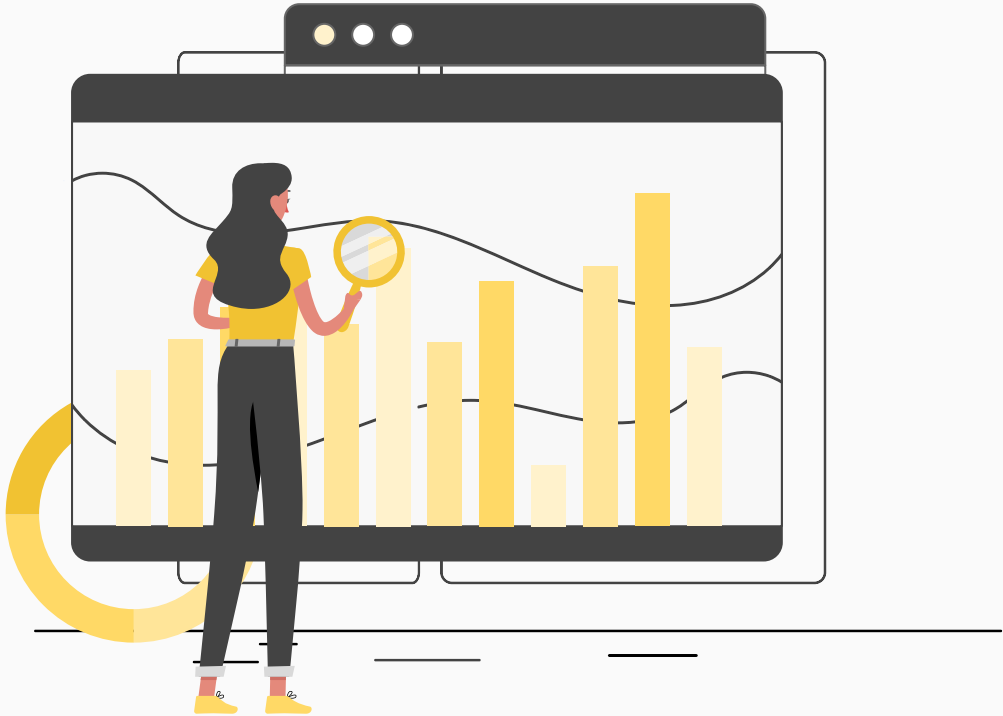
Physical data is used by **developers**. The purpose is actual implementation of database

- Entities referred to as table
- Attributes referred to as column
- Database compatible table names
- Database compatible column names
- Database specific data types
- Difficult for users to understand
- Significantly more effort than logical model

—PHYSICAL DATA MODELLING

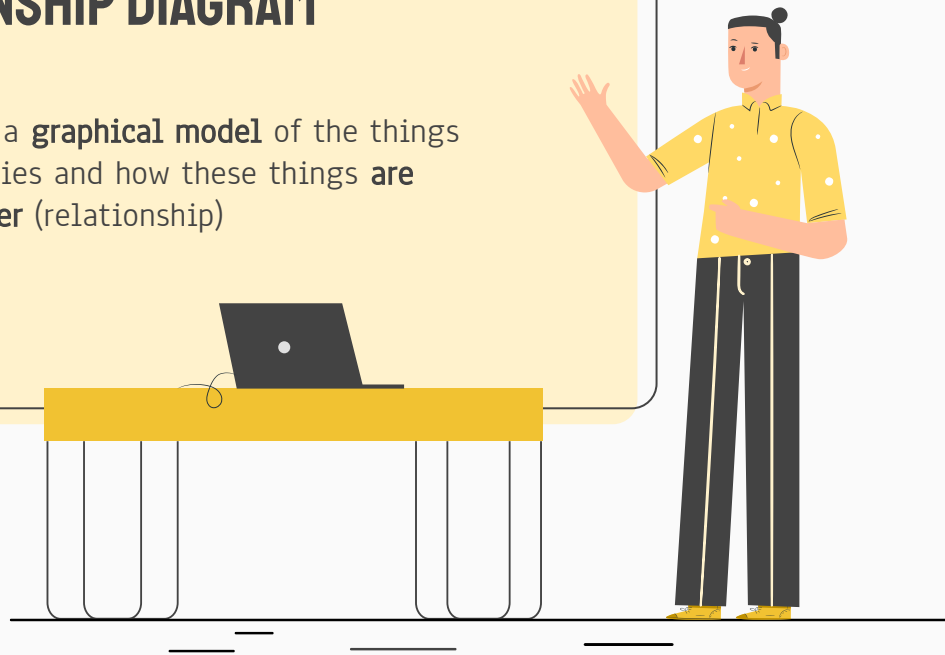
03. ENTITY RELATIONSHIP

Entity Relationship Diagram (ERD) and cardinality

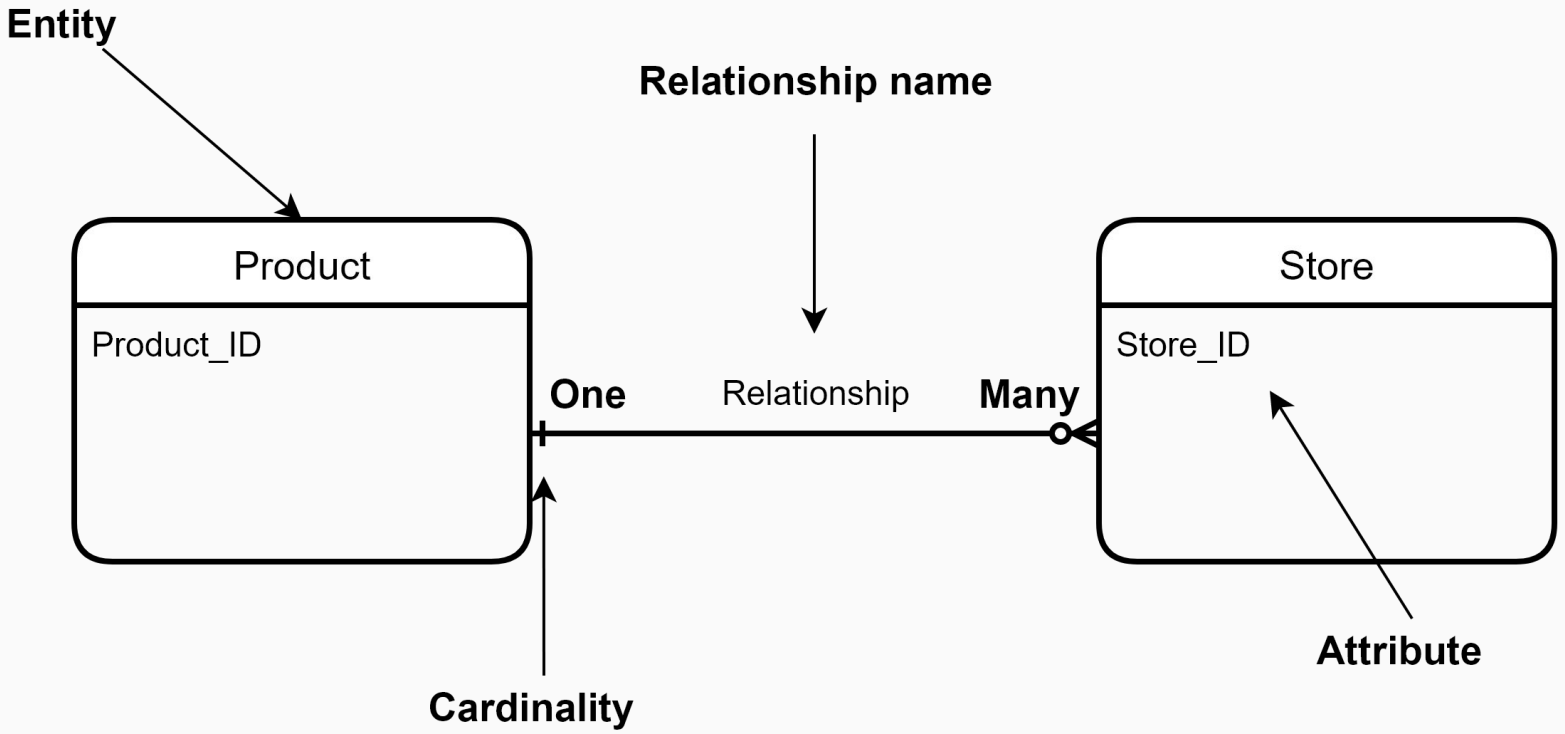


ENTITY RELATIONSHIP DIAGRAM

Entity relationship diagram provides a **graphical model** of the things that organization deals with entities and how these things **are related to one another** (relationship)



ENTITY RELATIONSHIP DIAGRAM



CARDINALITY

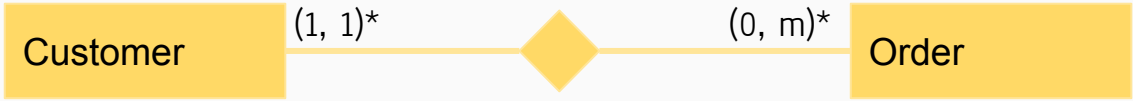
Cardinality is the **numerical mapping between entities**. This describes **“how many”** of one entity are related to **“how many”** of another entity

MAXIMUM CARDINALITY

Maximum value of instance of both sides of a relationship (1 or many)

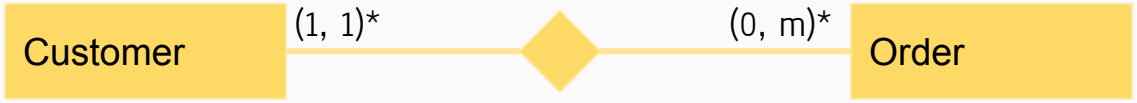
MINIMUM CARDINALITY

Minimum value of instance of both sides of relationship (0 or 1)

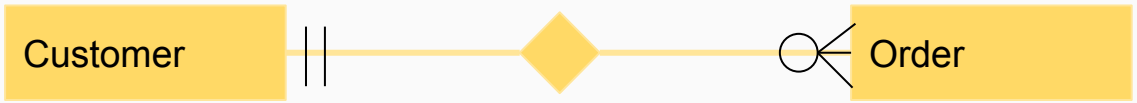


*notes: (min, max)

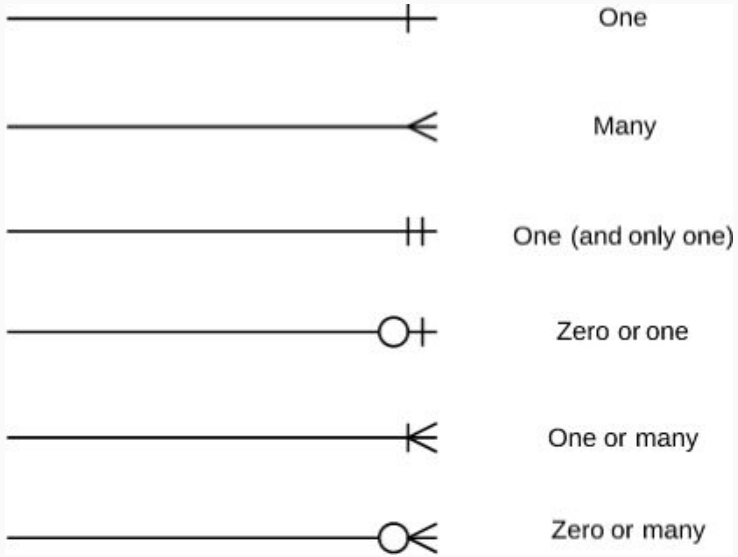
CROW FOOT NOTATION



*notes: (min, max)

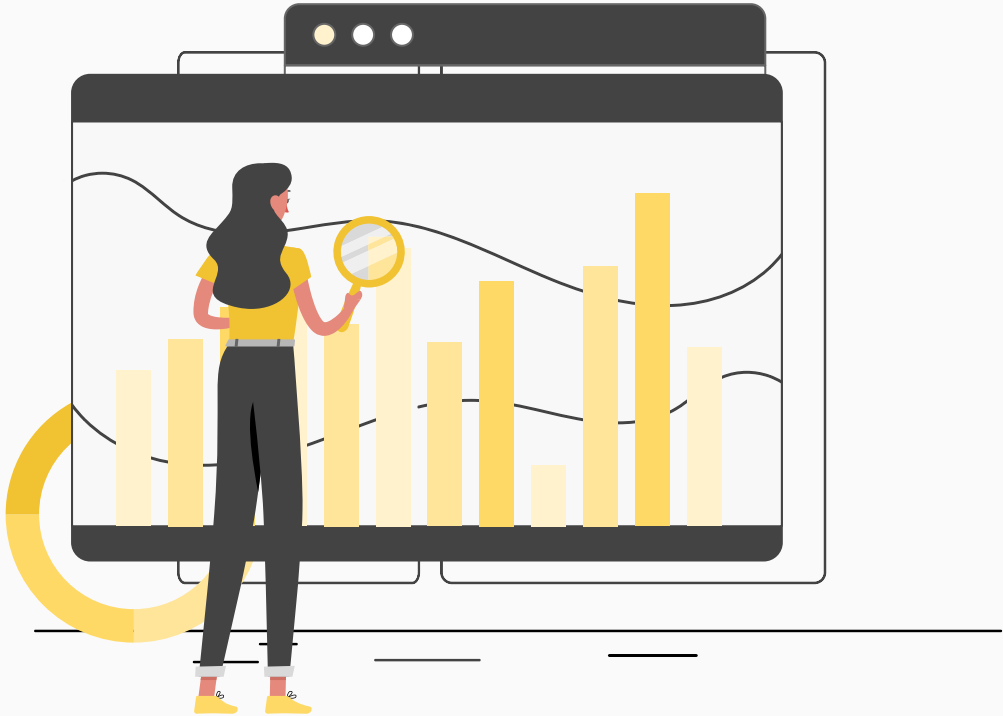


*notes: (min, max)



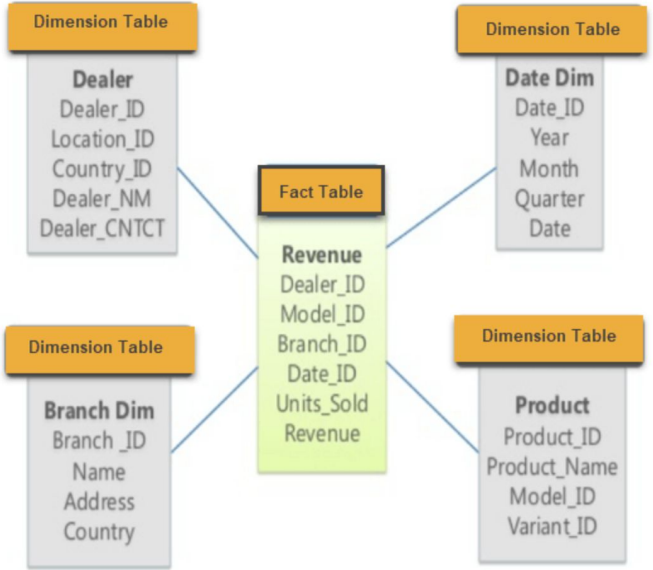
04. TYPES OF SCHEMA

Star schema and Snowflake schema



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.



Example of Star Schema Diagram

Characteristic of star schema:

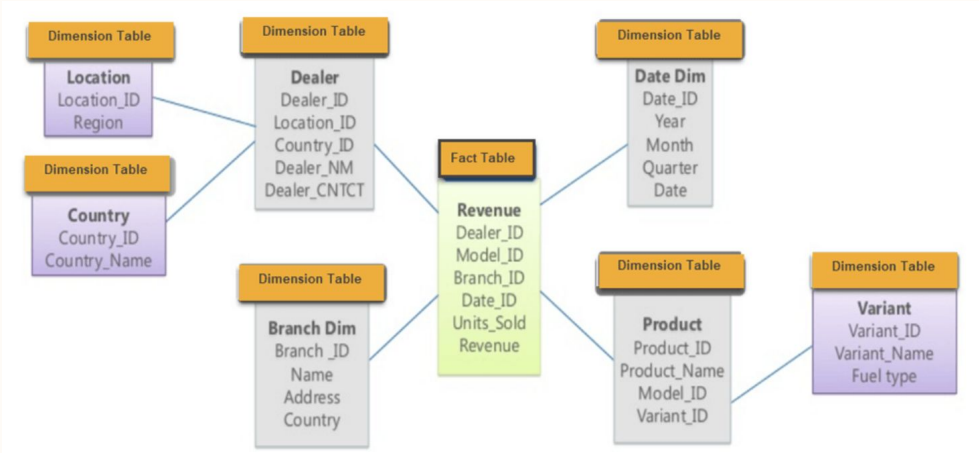
- Every dimension in a star schema is represented with the only one-dimension table.
- The dimension table should contain the set of attributes.
- The dimension table is joined to the fact table using a foreign key
- The dimension table are not joined to each other
- Fact table would contain key and measure
- The schema is widely supported by BI Tools

—SNOWFLAKE SCHEMA

A Snowflake Schema is an extension of a Star Schema, and it adds **additional dimensions**. The dimension tables are **normalized** which splits data into additional tables.

Characteristic of Snowflake schema:

- The main benefit of the snowflake schema it uses smaller disk space.
- Easier to implement a dimension is added to the Schema
- The primary challenge that you will face while using the snowflake Schema is that you need to perform **more maintenance** efforts because of the **more lookup tables**



Example of Snowflake Schema

THANK YOU

LinkedIn: https://www.linkedin.com/in/ariprabowo/

Github: https://github.com/densaiko/

Medium: https://dataimpact.medium.com/