Data Modeling
ERM & ERD

Outline:
- Entity relationship model (ERM) & ER diagram (ERD)
- Entities in an ERD?
- Attributes in an ERD?
- Relationships in an ERD?
- Relationship constraints in an ERD?
- ERD guidelines

ERM and ERD
- **Entity-Relationship Model (ERM)** is a detailed, logical representation of the data for an organization or for a business area.
  - Expressed in terms of:
    - Entities
    - Attributes
    - Relationships
- **Entity-Relationship Diagram (ERD)** is a graphical representation of a Entity-Relationship Model.

ERD
- The purpose of an ERD is to capture the richest possible understanding of the meaning of data necessary for an information system or organization.

- ERDs are made from Entities, Attributes, and Relationships.

Entity
- An object with physical or logical existence
- Has its own identity that distinguishes it from other entities.
  - Examples:
    - Person: PROFESSOR, STUDENT
    - Place: STORE, UNIVERSITY
    - Object: MACHINE, BUILDING
    - Event: SALE, REGISTRATION
    - Concept: ACCOUNT, COURSE

Entities

Student   Teacher   Projects

Entity
(Entity Type/Instance/Set)

- **Entity Type:** A particular class of an entity

> Student
> **Entity Type**

- **Entity Instance:** is a single occurrence of an entity type.

- **Entity Set:** a collection of entities that share common properties, characteristics or attributes.

> E1
> E2
> E3
> **Entity Set**

Attributes

- Each Entity has a set of Attributes
- Attribute is a property, feature, or characteristic of an entity that is of interest to the organization.
  - Example:
   - STUDENT: studentID, studentName, phoneNumber, programme

---

Attributes

> Name   BirthDate
> Student
> Roll_No.
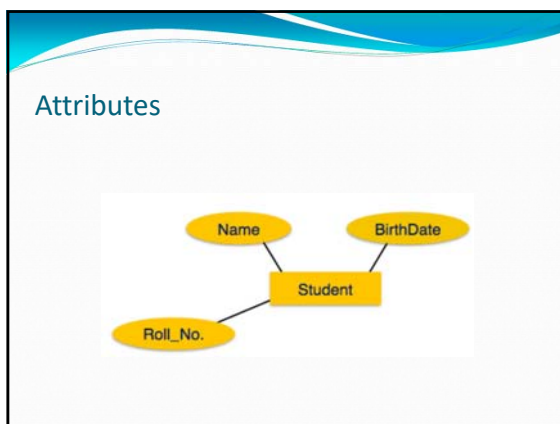
Attributes

- **Candidate Key:** An attribute that uniquely identifies each instance of an entity type.
- **Identifier:** A candidate key that has been selected as the unique, identifying characteristic of an entity type. (Should be underlined in ERD).
- Other types of Attributes: multivalued, composite, and derived.

Criteria for Selecting Attributes
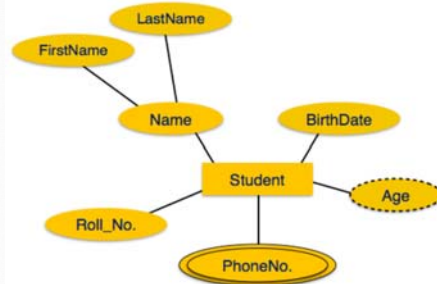
- Choose a candidate key that will not change its value.
- Choose a candidate key that has valid values and not be null.
- Avoid using codes, such as a 2 digit department location

Composite Attributes

> LastName
> FirstName
> Name   BirthDate
> Student
> Roll_No.
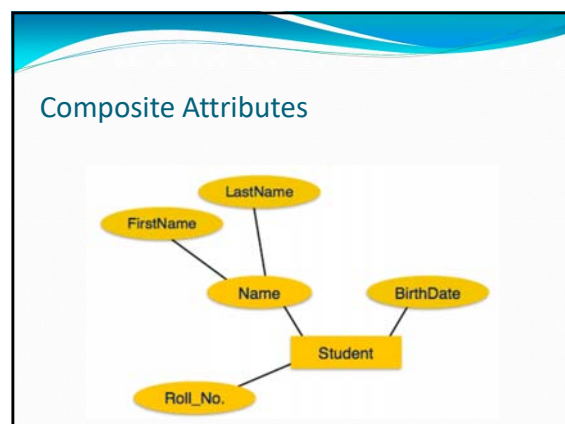
Multi-valued Attributes
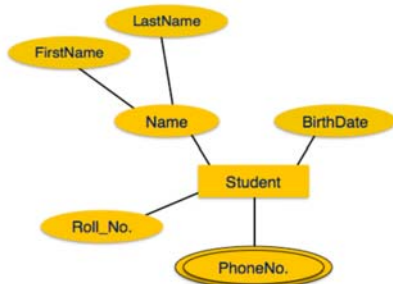
Derived Attributes

EntityNaming Guidelines

- A **singular noun**
- Follow Pascal case (e.g. SalesRepresentative)
- **Descriptive** and **specific** to the organization.
- **Concise**.

Attributes Defining Guidelines

- **An attribute definition should**:
  - State what the attribute is and why it is important.
  - Make clear what is and isn't included in the attribute's value.
  - Define any aliases.
  - Indicate if the attribute is required or not.
  - Indicate any relationships with other attributes.

Attributes Naming Guidelines

- **An attribute name**:
  - Should be a *noun*
  - Use Camel case to distinguish from entity names (e.g., dateOfBirth, phoneNumber)
  - Should be *unique*.
  - Should follow a *standard format*. (e.g, employeeName, not nameOfEmployee)
- Similar attributes of different entity types should use similar but distinguished names.

Relationships

- **Relationships** are associations between one or more entity types.
- Act as the "glue" to hold together different components of an ER model.
- **The degree of a relationship** is the number of entity types that participate in the relationship.
  - There are 3 common relationships:
    1. **Unary** (degree one)
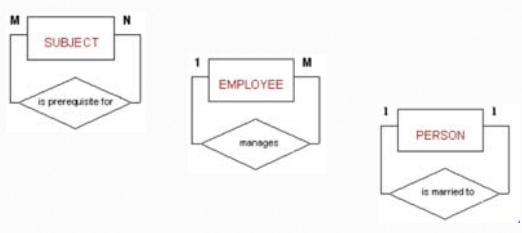   2. **Binary** (degree two)
    3. **Ternary** (degree three)

Relationships
(Naming Guidelines)

- A relationship name should:
  - Be a verb phrase, such as ResidesIn
  - Avoid vague names, such as "Has"
  - Use Pascal case

Unary Relationship

- Relationship between the instances of one entity type.

Binary Relationship

- Relationship between the instances of two entity type.
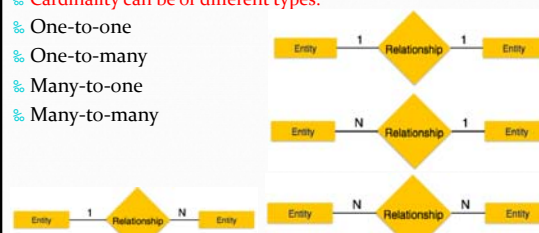
Ternary Relationship

- A simultaneous relationship among instances of three entity types.

Cardinality Ratio Constraint

- The number of times an entity of an entity set participates in a relationship set.
- Cardinality can be of different types:
- One-to-one
- One-to-many
- Many-to-one
- Many-to-many

Participation Constraint

- **Total/Mandatory Participation:** Each entity is involved in the relationship. Total participation is represented by double lines.

- **Partial/Optional participation:** Not all entities are involved in the relationship. Partial participation is represented by single lines.

Participation Constraint

How to Interpret n-to-1 and 1-to-n?

Weak Entity, Identifying Relationship, Identifying Owner

- Entity type for which key attribute can't be defined is called weak entity type.
- Weak entity type has partial key

Min-Max Notation for Relationship Constraints

- ➤ Used to represent both cardinality ratio and participation constraints together
- ➤ (0, Max) means partial participation
- ➤ (1, Max) means total participation

Guidelines for Drawing ERDs

- ❧ Define entities and relationships
- ❧ Add attributes to the entities and relationships (if applicable).
- ❧ Add cardinality/participation constraints to the relationships.
- ❧ Don't forget to use proper naming conventions and symbol representation.
- ❧ Lay out the diagram with minimal line crossing.
- ❧ Group closely related entity types when possible. Try to keep the length of relationship lines as short as possible. Also try to minimize the number of changes of direction in a single line.