COMPUTER AIDED SOFTWARE ENGINEERING (CASE)

Dr. Ebadati

Kharazmi University

DATA MODELING

Chapter 6

KEY DEFINITIONS

- ***Data model***
  - A **formal way** of **representing** the **data** that are used and created by a business system
  - **Shows the people**, **places** and **things** about which data is captured and the relationships among them.
- ***Logical data model***
  - **shows** the organization of **data without indicating how** it is **stored**, created, or manipulated.

KEY DEFINITIONS

- ***Physical data model***
  - **shows** how the **data** will actually be **stored** in **databases** or files.
- ***Normalization*** is the **process** analysts use to **validate** data models.
- **Data models** should ***balance*** with process models

THE ENTITY-RELATIONSHIP DIAGRAM (ERD)

WHAT IS AN ERD?

- A picture showing the **information created**, **stored**, and **used** by a business **system**.
- **Entities** generally **represent similar** kinds of **information**
- **Lines** drawn **between entities** show **relationships** among the data
- High level business rules are also shown

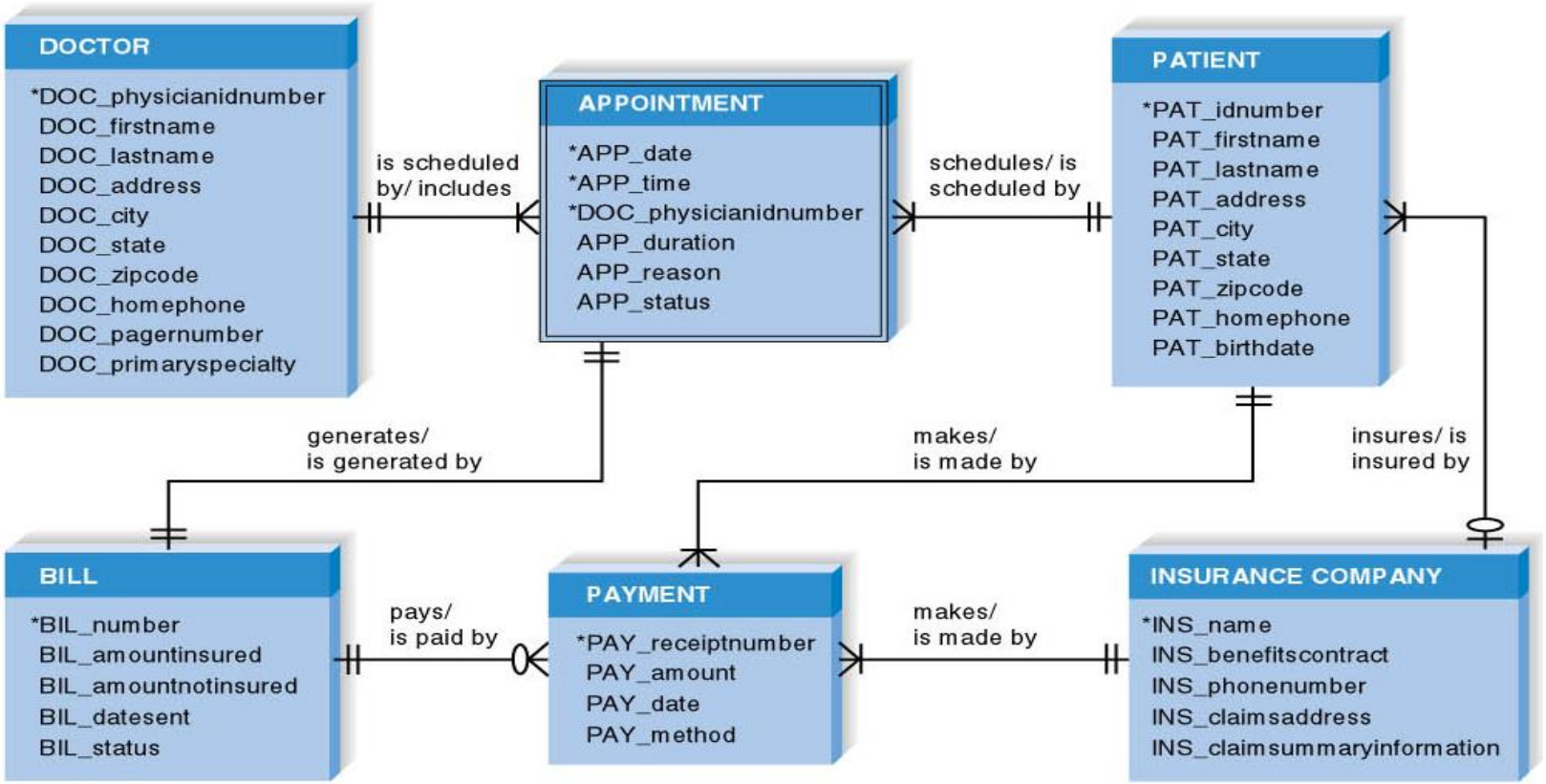
USING THE ERD TO SHOW BUSINESS RULES

- **Business rules** are **constraints** that are followed when the **system** is in **operation**.
- **ERD symbols can show when one instance of an entity must exist for an instance of another to exist**
  - **A doctor must exist before appointments for the doctor can be made**

USING THE ERD TO SHOW BUSINESS RULES

- **ERD symbols** can show when one instance of an **entity can be related to only one or many instances of another entity**
  - **One doctor can have many patients; each patient may have only one primary doctor**
- ERD symbols show **when the existence of an entity instance is optional for a related entity instance**
  - **A patient may or may not have insurance coverage**

AN ERD EXAMPLE



ERD ELEMENTS

	IDEF1X	Chen	Crow's Foot
An ENTITY: ✓ Is a person, place, or thing ✓ Has a singular name spelled in all capital letters ✓ Has an identifier ✓ Should contain more than one instance of data	ENTITY-NAME Identifier	ENTITY-NAME	ENTITY-NAME *Identifier
An ATTRIBUTE: ✓ Is a property of an entity ✓ Should be used by at least one business process ✓ Is broken down to its most useful level of detail	ENTITY-NAME Attribute-name Attribute-name Attribute-name	Attribute-name	ENTITY-NAME Attribute-name Attribute-name Attribute-name
A RELATIONSHIP: ✓ Shows the association between two entities ✓ Has a parent entity and a child entity ✓ Is described with a verb phrase ✓ Has cardinality (1 : 1, 1 : N, or M : N) ✓ Has modality (null, not null) ✓ Is dependent or independent	Relationship-name	Relationship-name	Relationship-name

ENTITY

- A **person**, **place**, **event**, or **thing about which data is collected**
- Must be multiple occurrences to be an entity
  - Example: If a firm has only one warehouse, the warehouse is not an entity. However, if the firm has several warehouses, the warehouse could be an entity if the firm wants to store data about each warehouse instance.

ENTITIES AND INSTANCES

Entity	Example Instances
Patient	John Smith Susan Jones Peter Todd Dale Turner Pat Turner

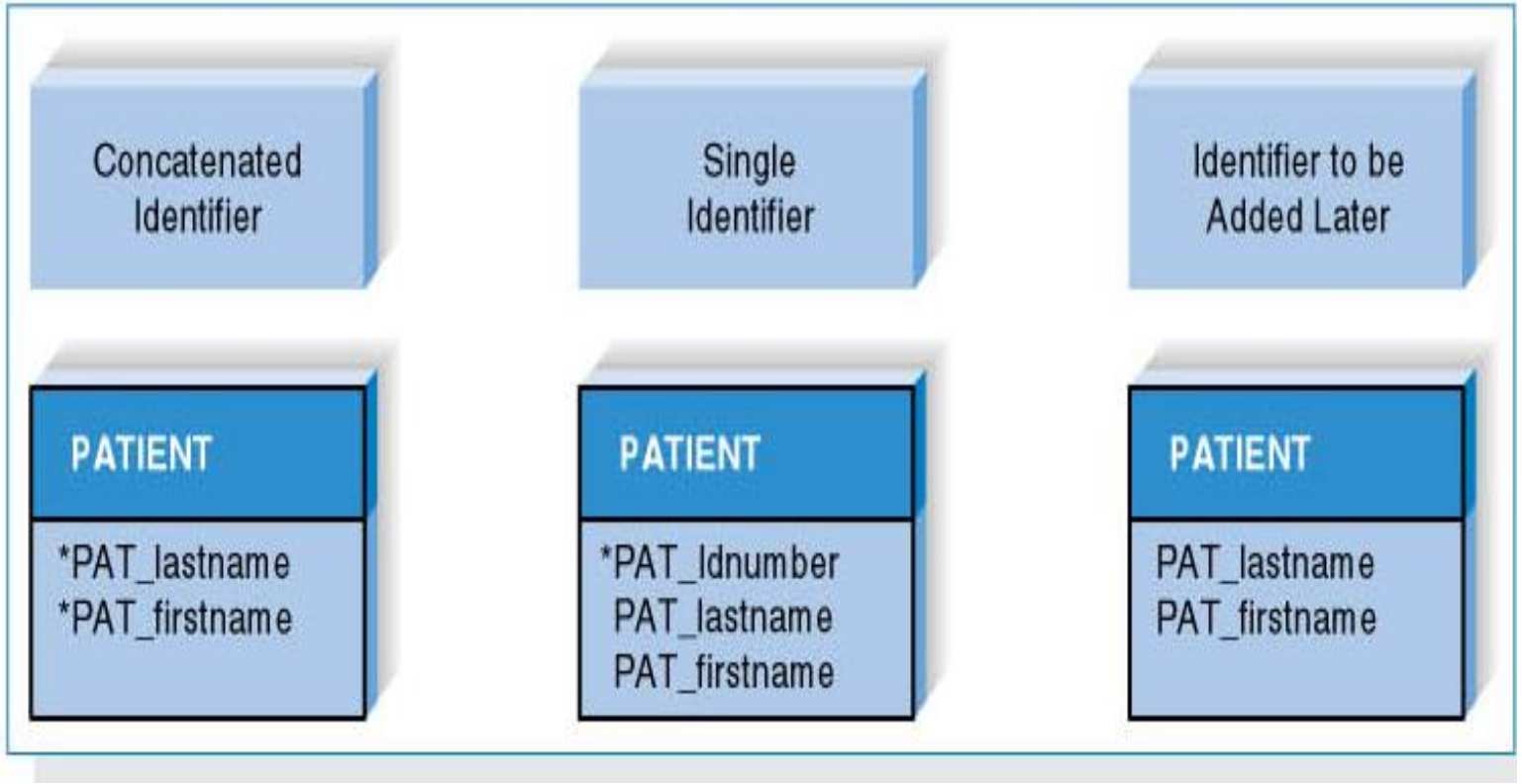
ATTRIBUTES

- **Information captured about** an **entity**
- Only those **used** by the organization should be **included** in the **model**
- **Attribute names are nouns**
- **Sometimes** entity name is added at the **beginning** of the **attribute name** for clarity

IDENTIFIERS

- One or more attributes can serve as the entity **identifier**, uniquely identifying each entity instance
- **Concatenated identifier** consists of several attributes
- An **identifier may** be '**artificial**,' such as creating an **ID number**
- **Identifiers** may not be **developed** until the **Design Phase**

CHOICES FOR IDENTIFIERS



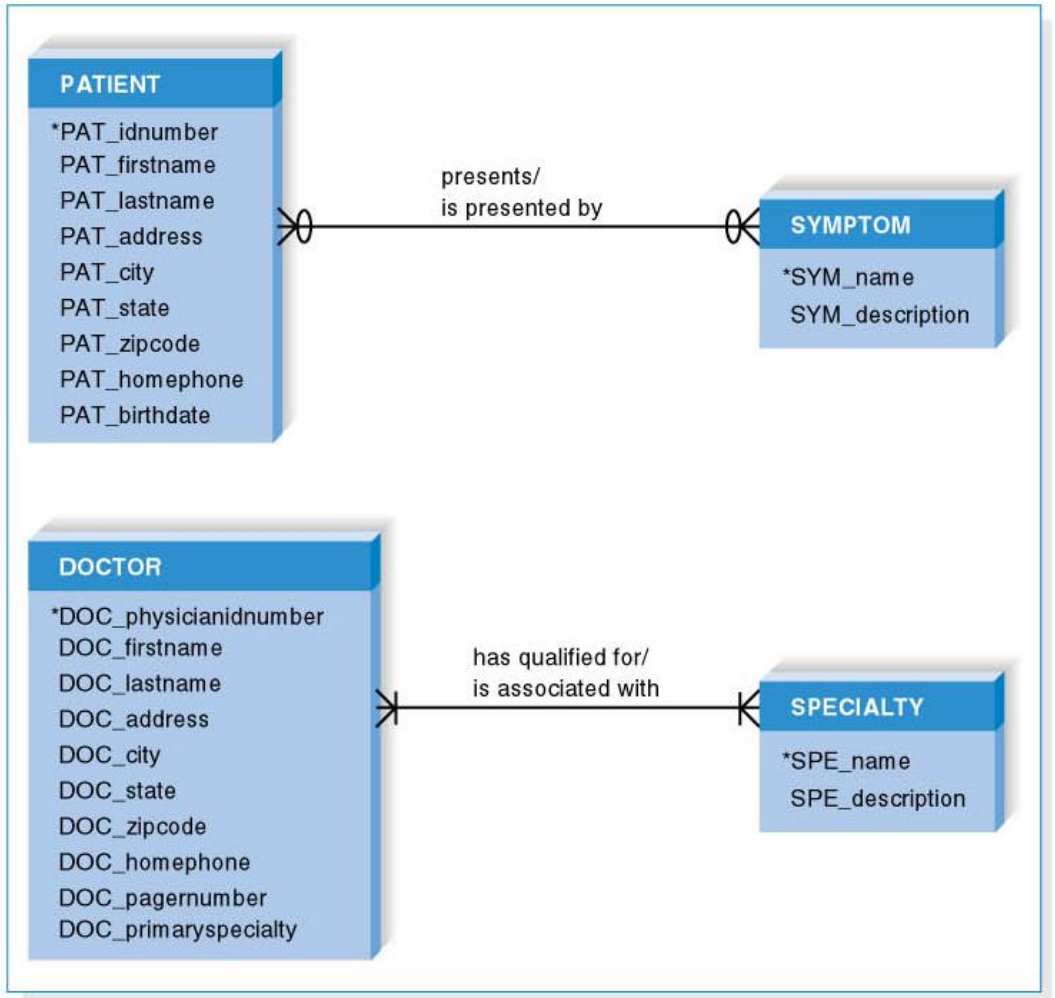
RELATIONSHIPS

- **Associations between entities**
- The **first entity** in the relationship is the ***parent*** entity; the **second entity** in the relationship is the ***child*** **entity**
- **Relationships should have active verb names**
- **Relationships go in both directions**

CARDINALITY

- *Cardinality*
  - **refers** to the **number of times instances in one entity can be related to instances in another entity**
    - **One instance** in an entity refers to **one and only one instance** in the related entity (1:1)
    - **One instance** in an entity refers to **one or more instances** in the related entity (1:N)
    - **One or more** instances in an entity refer to **one or more instances** in the related entity (M:N)

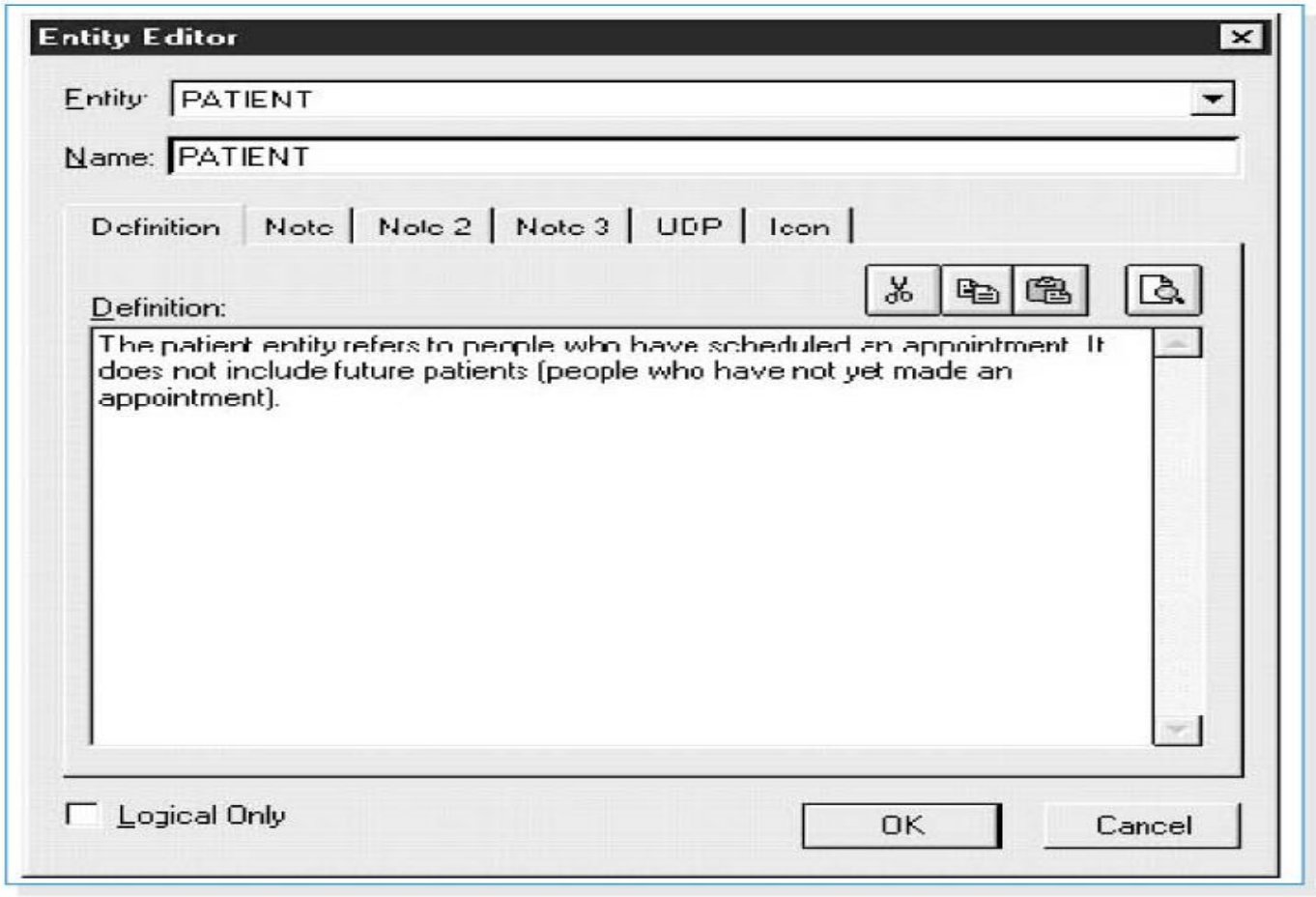
M : N RELATIONSHIPS



THE DATA DICTIONARY AND METADATA

- **Metadata** is **information stored** about **components** of the data model
- Metadata is stored in the data dictionary so it can be shared by developers and users throughout the SDLC
- A complete, shareable data dictionary helps improve the quality of the system under development

DATA DICTIONARY ENTRY FOR THE PATIENT ENTITY (SHOWN USING ERWIN)



BUILDING AN ENTITY-RELATIONSHIP DIAGRAM (ERD)

ERD BASICS

- **Drawing** the **ERD** is an iterative process of **trial and revision**
- ERDs can become quite complex

STEPS IN BUILDING ERDS

- **Identify the entities**
- **Add attributes** and assign identifiers
- **Identify relationships**

IDENTIFY THE ENTITIES

- **Identify major categories of information**
  - If available, **check the process models for data stores**, external entities, and data flows
  - Check the **major inputs and outputs from the use cases**
- **Verify that there is more than one instance of the entity that occurs in the system**

IDENTIFY RELATIONSHIPS

- **Start** with **an entity** and **identify all entities** with which it shares relationships
- Describe the relationship with the appropriate verb phrase
- Determine the cardinality and modality by discussing the business rules with knowledgeable users

ERD BUILDING TIPS

- **Data stores of the DFD should correspond to entities**
- Only include entities with more than one instance of information
- **Don't include entities associated with implementation of the system (they will be added later)**

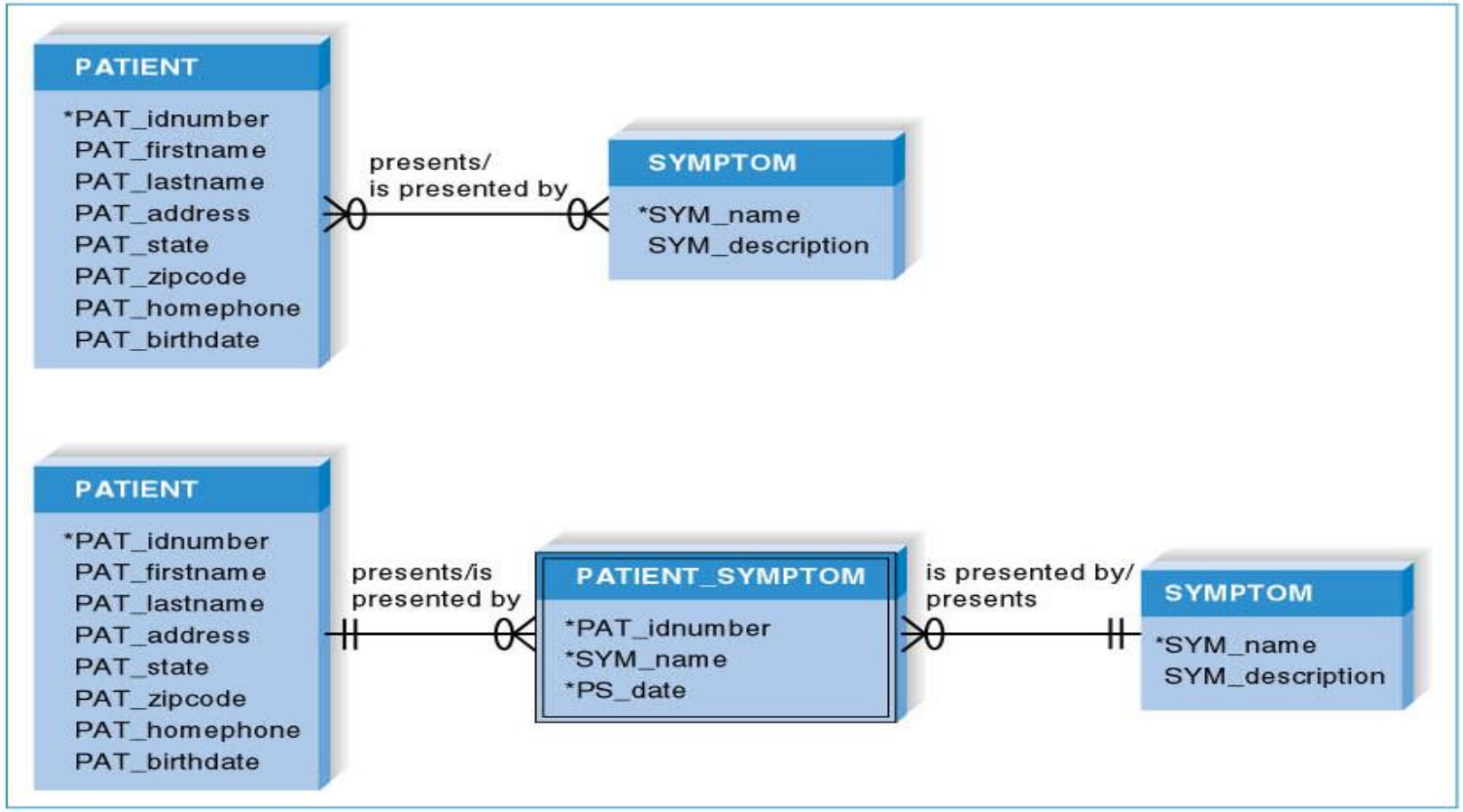
ADVANCED SYNTAX

- ***Independent Entity***
  - **Can exist without the help of another entity**
  - Identifiers created from the entity's own attributes
  - Attributes from other entities are not needed to uniquely identify instances of these entities
- ***Dependent Entity***
  - **Relationships when a child entity does require attributes** from the parent entity to uniquely identify an instance

ADVANCED SYNTAX

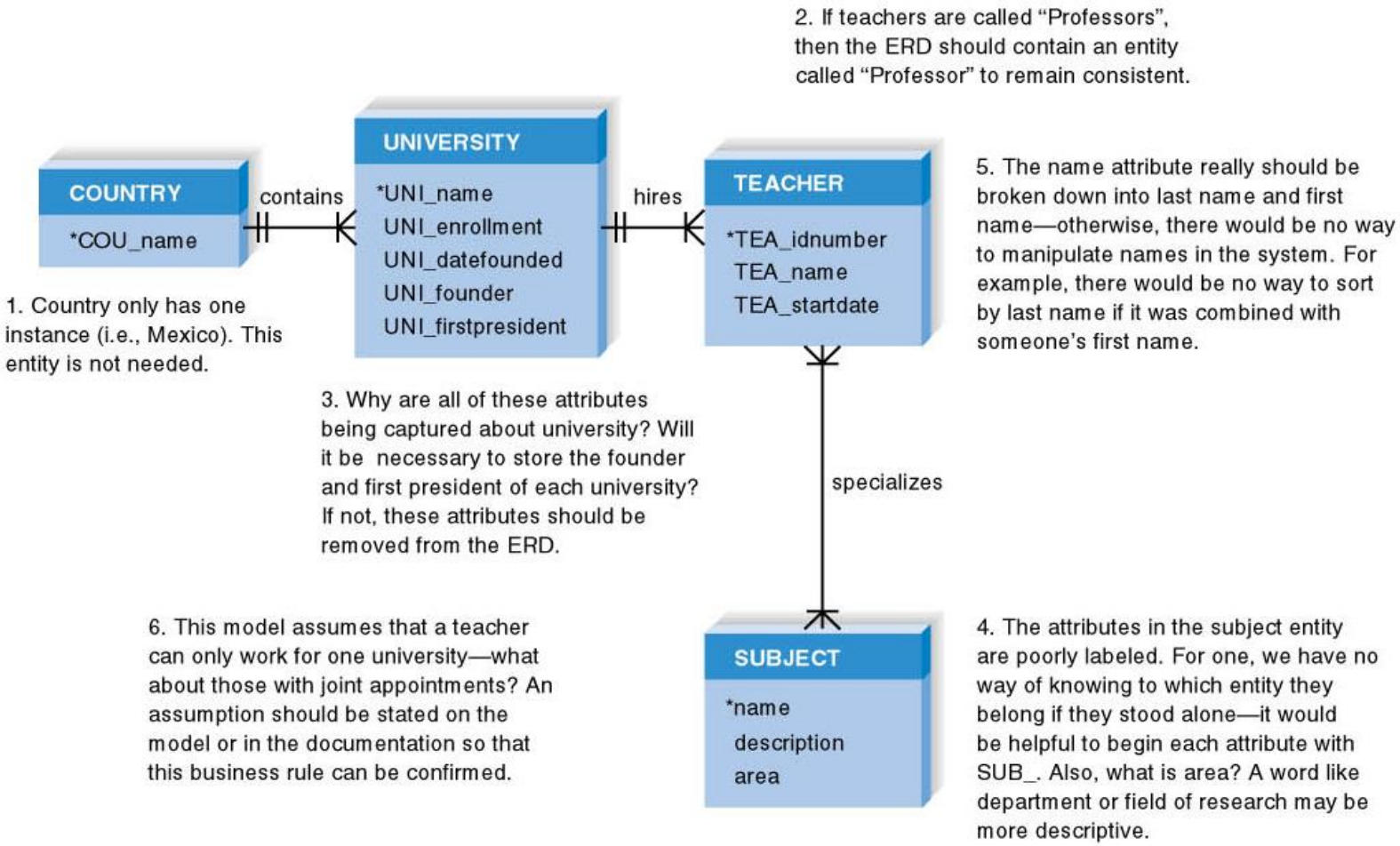
- ***Intersection Entity***
  - Exists in order to capture some information about the relationship that **exists between two other entities**. Typically, intersection entities are added to a data model to store information about two entities sharing **an M : N relationship**.

ADVANCED SYNTAX – RESOLVING AN M : N RELATIONSHIP



VALIDATING AN ERD

DESIGN MODELING GUIDELINES SUMMARY



NORMALIZATION

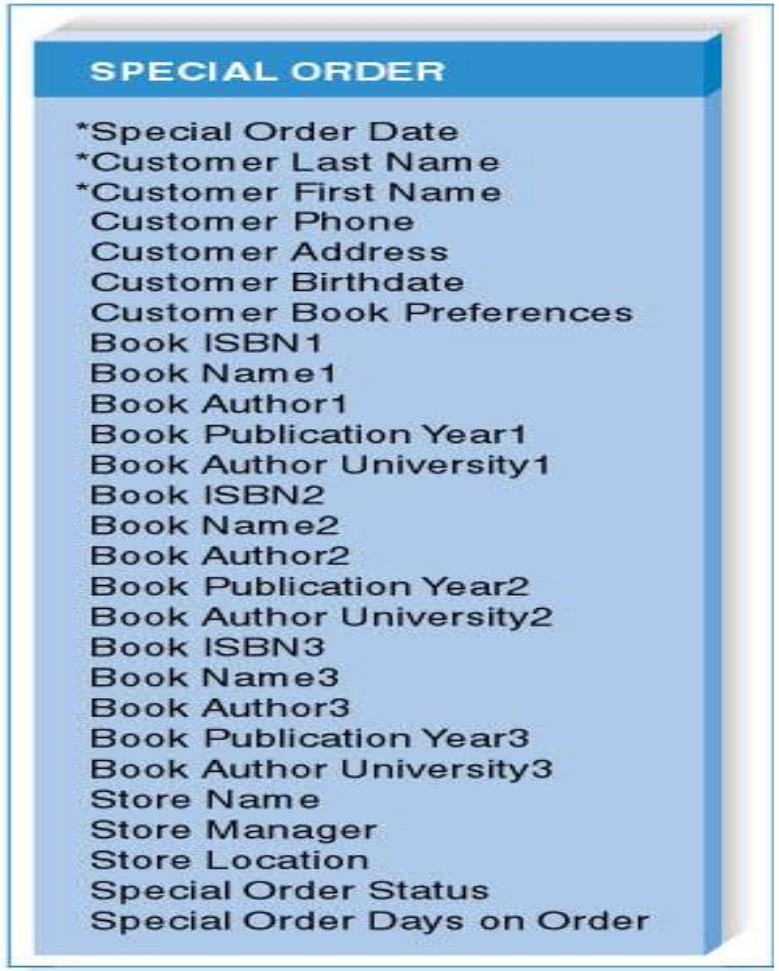
- **Technique used to validate data models**
- **Series of rules** applied to logical data model to improve its organization
- **Three normalization rules are common**

NORMALIZATION STEPS

	0 Normal Form
Do any attributes have multiple values for a single instance of an entity?	Yes: Remove the repeating attributes and repeating groups. Create an entity that describes the attributes. Usually you will need to add a relationship to connect the old and new entities. No: The data model is in 1NF.
	1 Normal Form
Is the identifier comprised of more than one attribute? If so, are any attribute values dependent on just part of the identifier?	Yes: Remove the partial dependency. Move the attributes to an entity in which their values are dependent on the entire identifier. Usually you will need to create a new entitiy and add a relationship to connect the old and new entities. No: The data model is in 2NF.
	2 Normal Form
Do any attribute values depend on an attribute that is not the entity's identifier?	Yes: Remove the transitive dependency or derived attribute. Move the attributes to an entity in which their values are dependent on the identifier. Usually you will need to create a new entity and add a relationship to connect the old and new entities. No: The data model is in 3NF.
	3 Normal Form

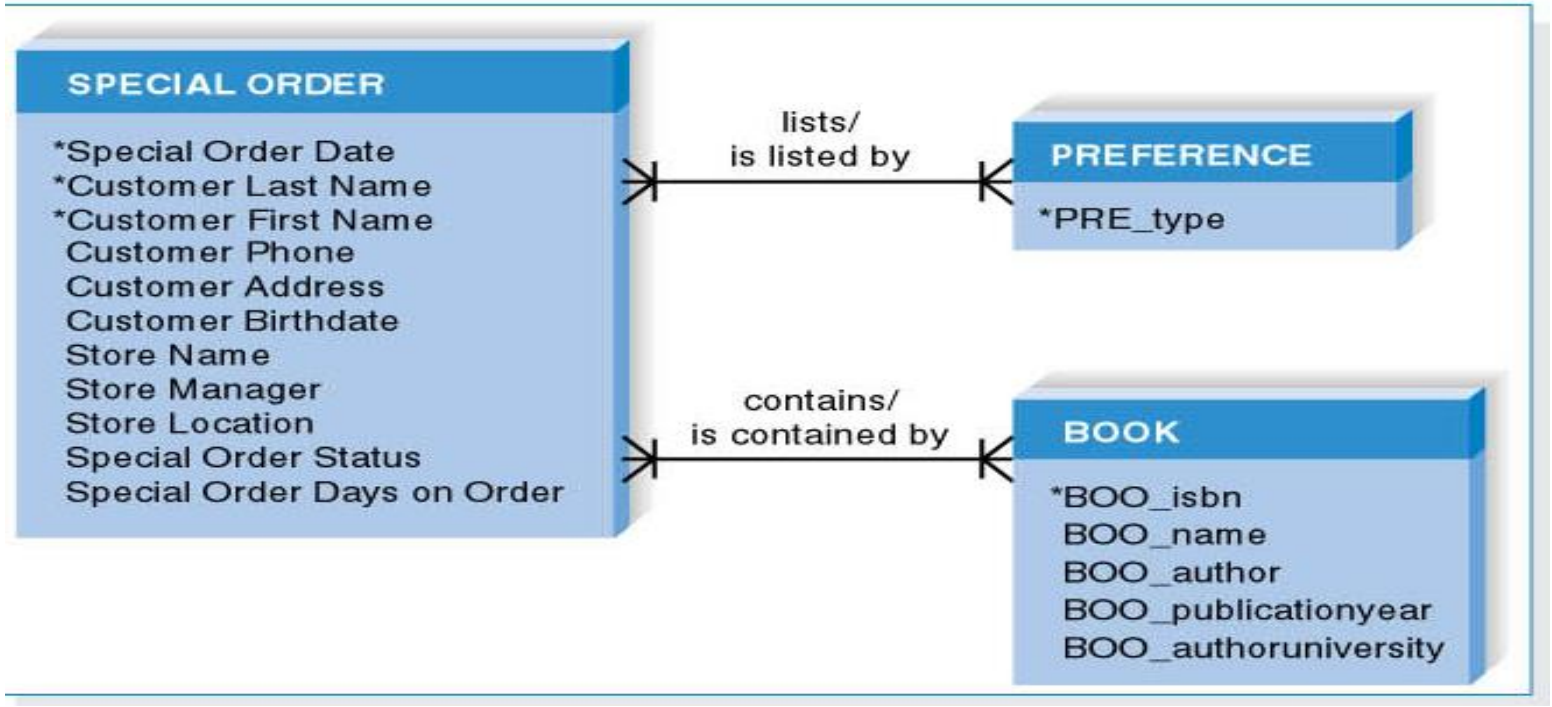
UN-NORMALIZED ENTITY

Begin with an entity from the logical data model



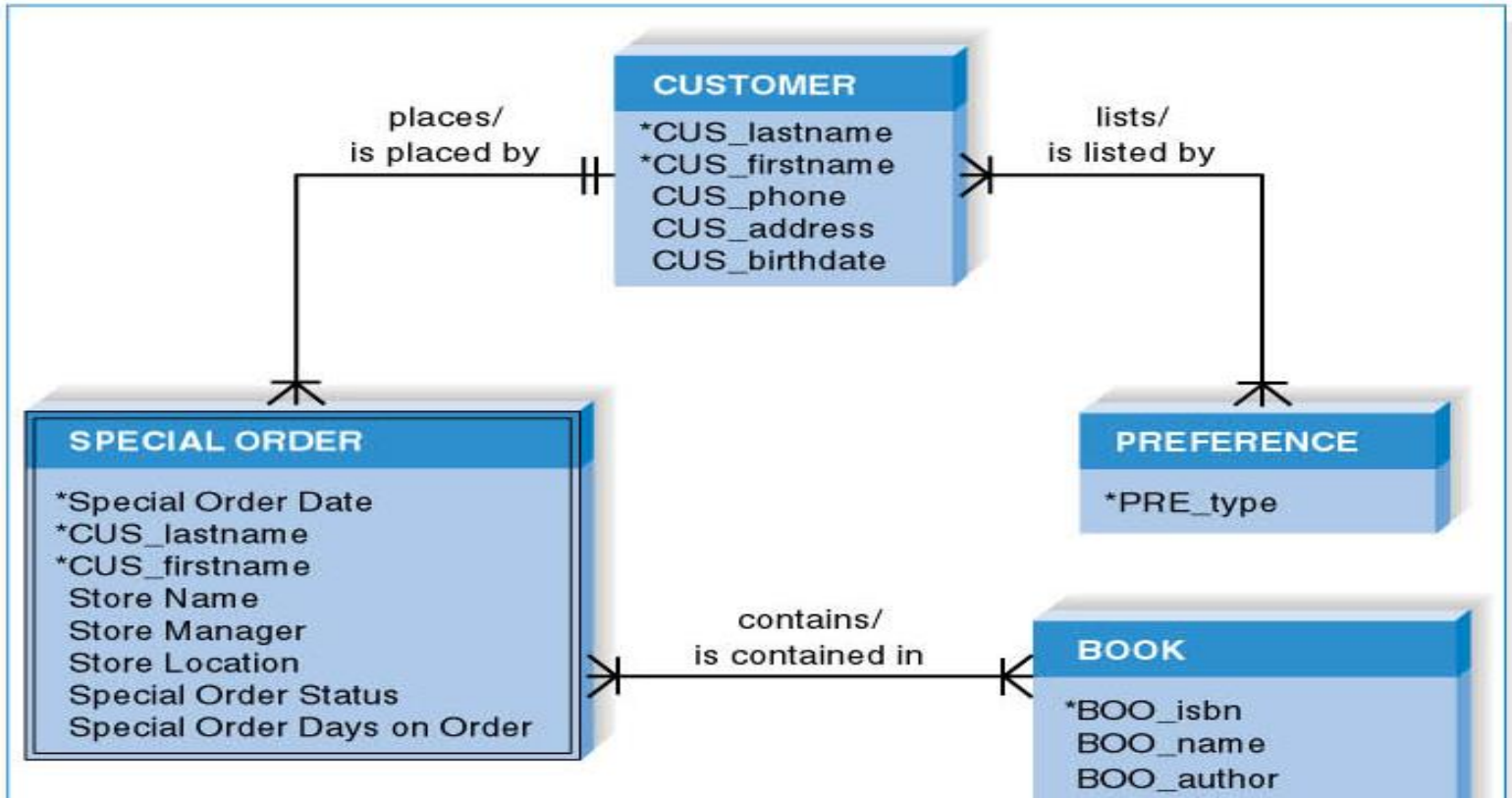
FIRST NORMAL FORM (1NF)

Look for repeating groups of attributes and remove them into separate entities



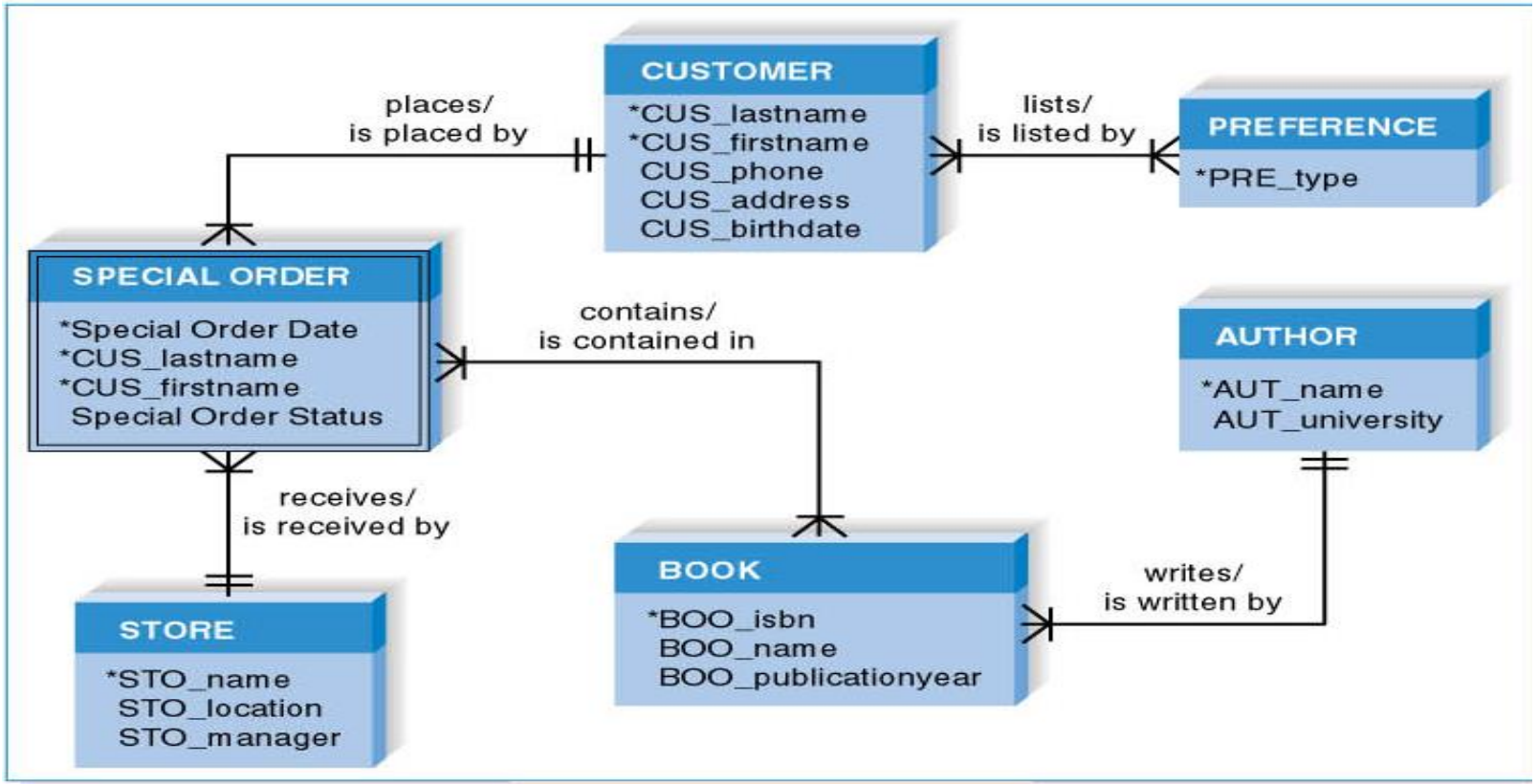
SECOND NORMAL FORM (2NF)

If an entity has a Joint identifier, look for attributes that depend only on part of the identifier. If found, remove to new entity.



THIRD NORMAL FORM (3NF)

Look for attributes that depend only on another non-identifying attribute.
If found, remove to new entity. Also remove any calculated attributes.



BALANCING ERDS WITH DFDS

- **All analysis activities are interrelated**
- **Process models** contain **two data components**
  - **Data flows** and **data stores**
- The **DFD** data components **need to balance the ERD's data stores** (entities) and data elements (attributes)
- Many **CASE tools** provide features to **check** for **imbalance**
- **Check** that all **data stores** and elements **correspond between models**
  - **Data that is not used is unnecessary**
  - **Data** that has been **omitted** results in an **incomplete** system
- Do not follow thoughtlessly -- check that the models make sense!

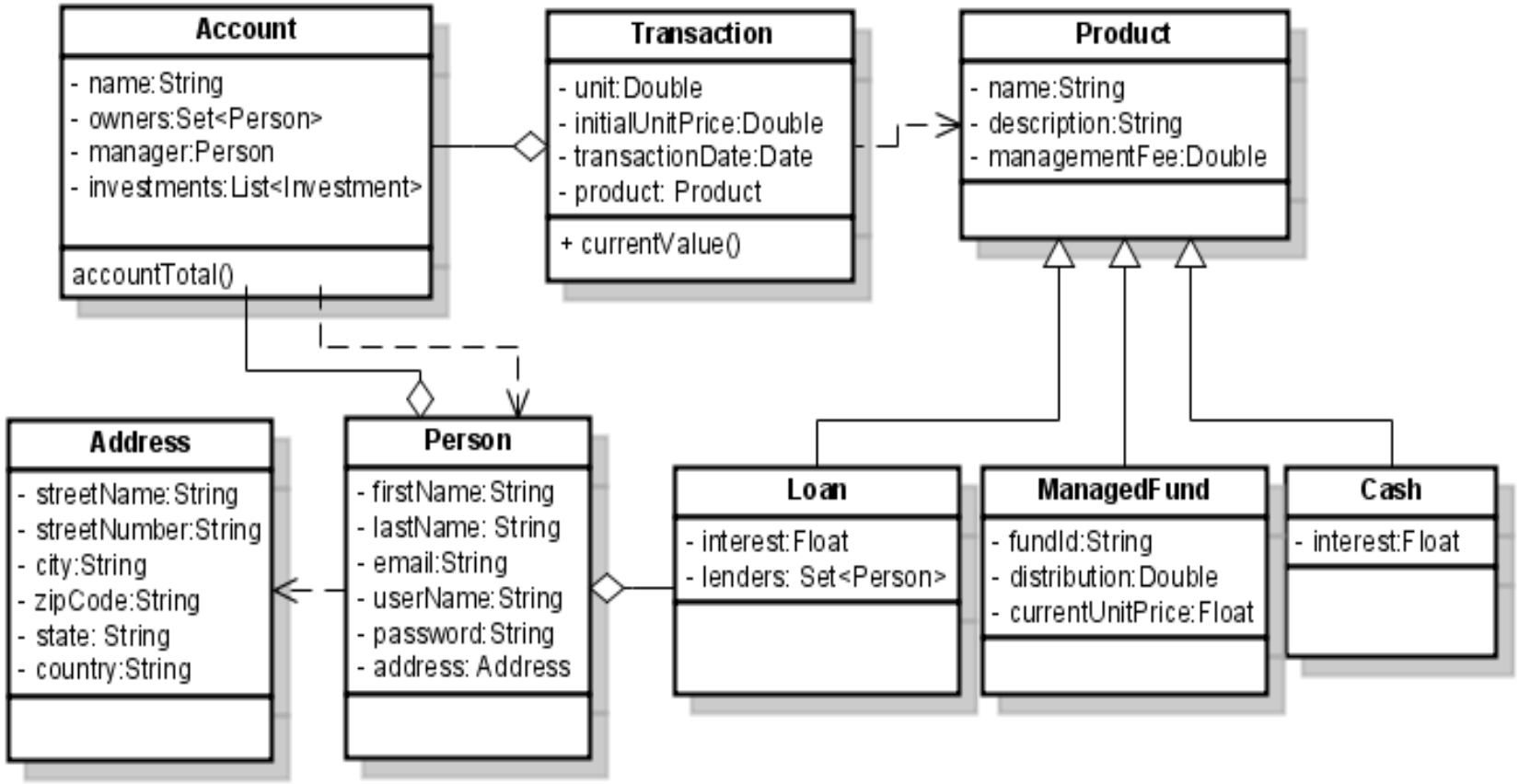
SUMMARY

- The **ERD** is the most common technique for drawing **data models**. The building blocks of the ERD are:
  - **Entities** describe **people**, **places**, or **things**
  - **Attributes** capture **information** about the **entity**
  - **Relationships** **associate data across entities**
- Intersection, dependent, and independent entities must be recognized.
- The **ERD must be balanced with the DFD**.

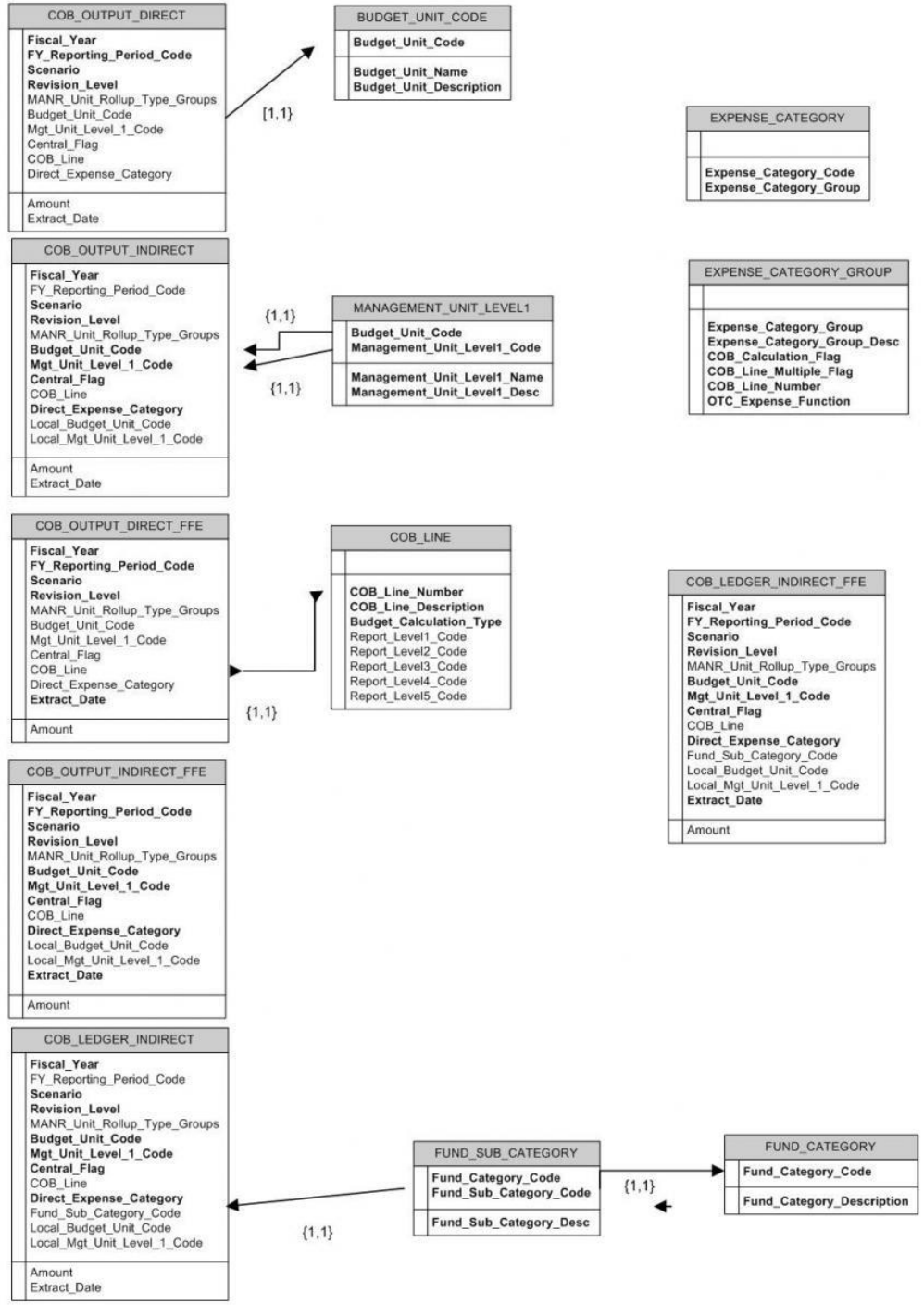
ERD DRAW TOOL

- http://www.aquafold.com/aquadatastudio/er_modeler.html

ERD EXAMPLE

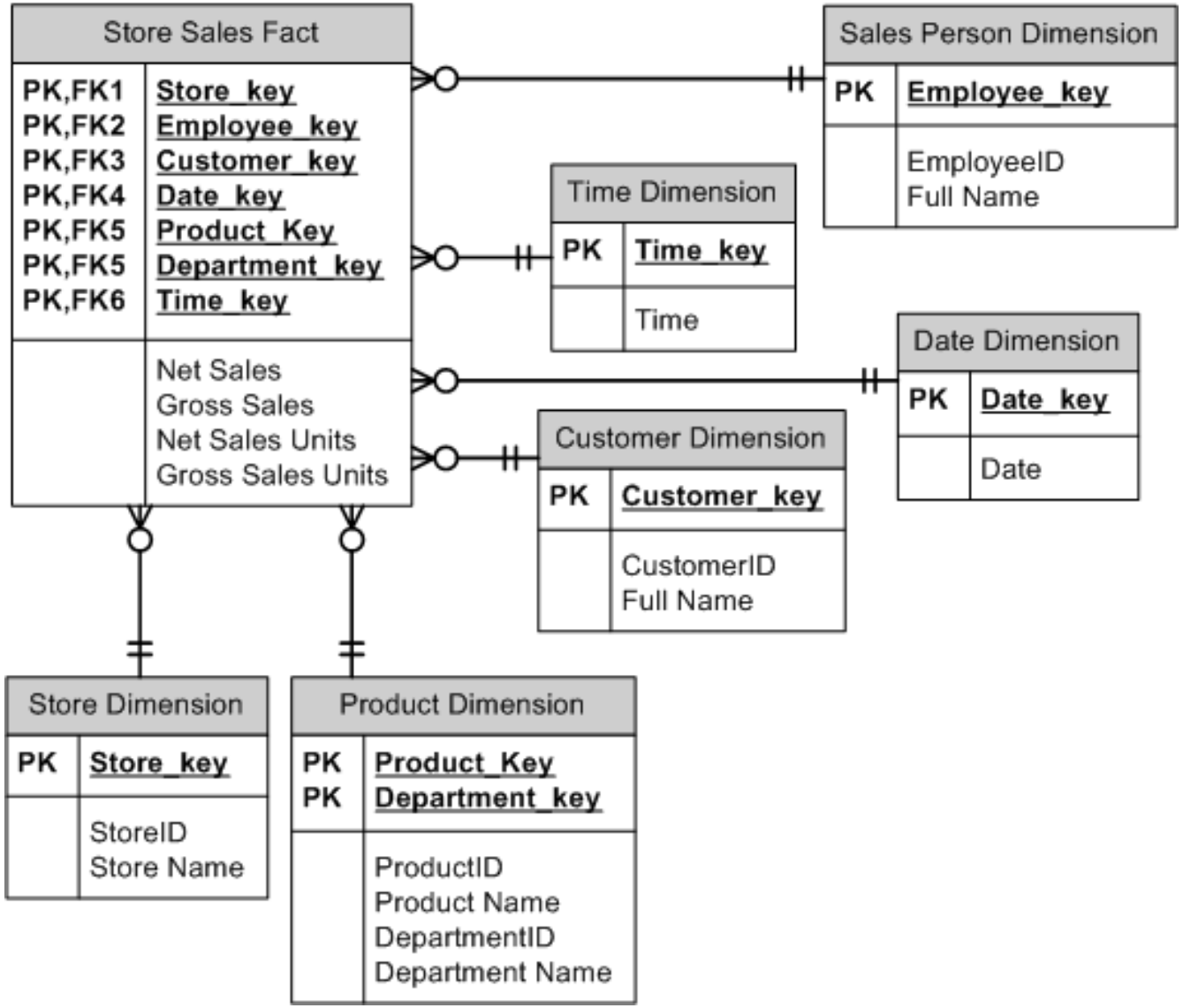

Account
- name:String
- owners:Set<Person>
- manager:Person
- investments:List<Investment>
accountTotal()
Transaction
- unit:Double
- initialUnitPrice:Double
- transactionDate:Date
- product: Product
+ currentValue()
Product
- name:String
- description:String
- managementFee:Double
Address
- streetName:String
- streetNumber:String
- city:String
- zipCode:String
- state: String
- country:String
Person
- firstName:String
- lastName: String
- email:String
- userName:String
- password:String
- address: Address
Loan
- interest:Float
- lenders: Set<Person>
ManagedFund
- fundId:String
- distribution:Double
- currentUnitPrice:Float
Cash
- interest:Float

ERD EXAMPLE

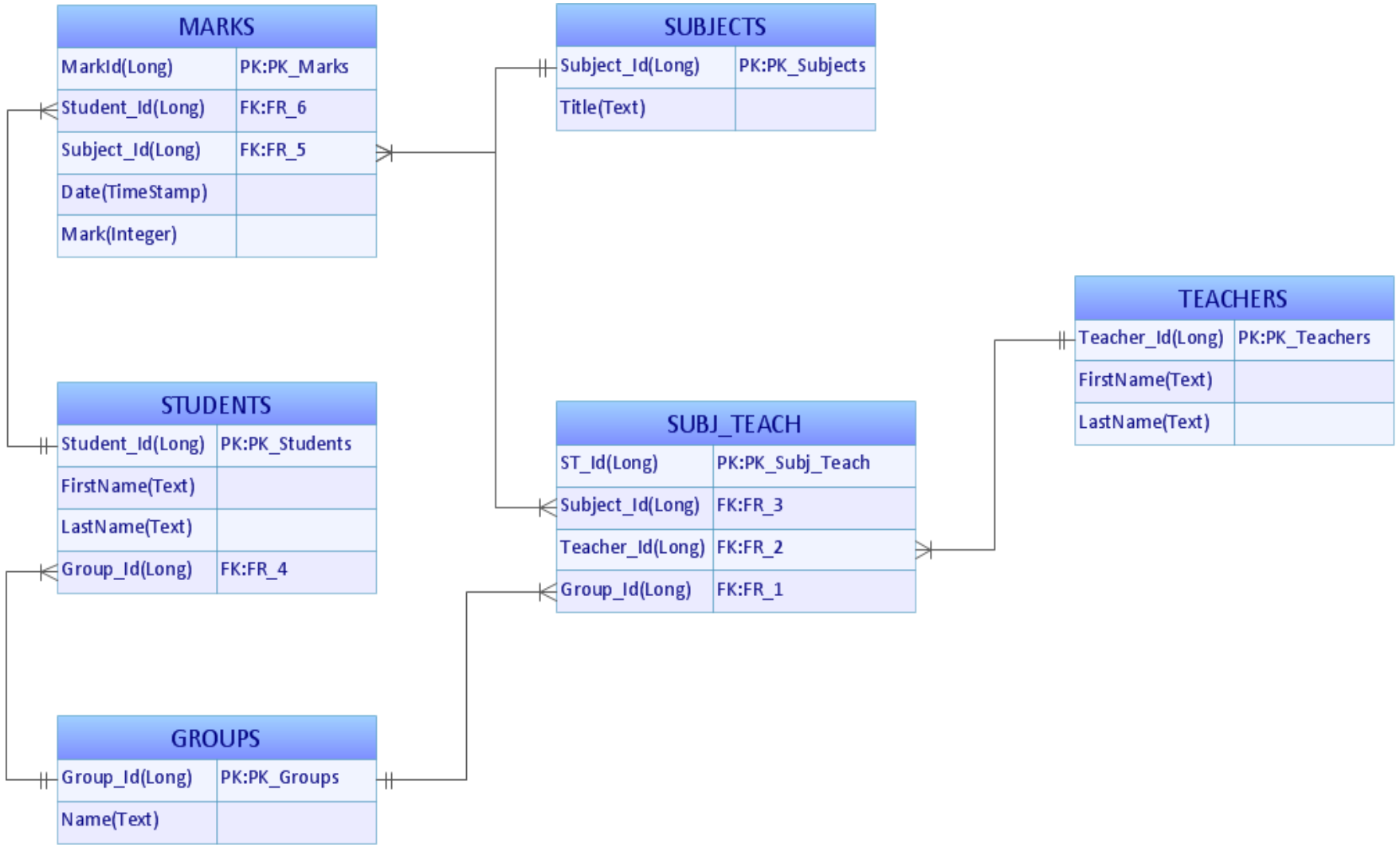

COB_OUTPUT_DIRECT
Fiscal_Year
FY_Reporting_Period_Code
Scenario
Revision_Level
MANR_Unit_Rollup_Type_Groups
Budget_Unit_Code
Mgt_Unit_Level_1_Code
Central_Flag
COB_Line
Direct_Expense_Category
Amount
Extract_Date
{1,1}
BUDGET_UNIT_CODE
Budget_Unit_Code
Budget_Unit_Name
Budget_Unit_Description
EXPENSE_CATEGORY
Expense_Category_Code
Expense_Category_Group
COB_OUTPUT_INDIRECT
Fiscal_Year
FY_Reporting_Period_Code
Scenario
Revision_Level
MANR_Unit_Rollup_Type_Groups
Budget_Unit_Code
Mgt_Unit_Level_1_Code
Central_Flag
COB_Line
Direct_Expense_Category
Local_Budget_Unit_Code
Local_Mgt_Unit_Level_1_Code
Amount
Extract_Date
{1,1}
{1,1}
MANAGEMENT_UNIT_LEVEL1
Budget_Unit_Code
Management_Unit_Level1_Code
Management_Unit_Level1_Name
Management_Unit_Level1_Desc
EXPENSE_CATEGORY_GROUP
Expense_Category_Group
Expense_Category_Group_Desc
COB_Calculation_Flag
COB_Line_Multiple_Flag
COB_Line_Number
OTC_Expense_Function
COB_OUTPUT_DIRECT_FFE
Fiscal_Year
FY_Reporting_Period_Code
Scenario
Revision_Level
MANR_Unit_Rollup_Type_Groups
Budget_Unit_Code
Mgt_Unit_Level_1_Code
Central_Flag
COB_Line
Direct_Expense_Category
Extract_Date
Amount
{1,1}
COB_LINE
COB_Line_Number
COB_Line_Description
Budget_Calculation_Type
Report_Level1_Code
Report_Level2_Code
Report_Level3_Code
Report_Level4_Code
Report_Level5_Code
COB_LEDGER_INDIRECT_FFE
Fiscal_Year
FY_Reporting_Period_Code
Scenario
Revision_Level
MANR_Unit_Rollup_Type_Groups
Budget_Unit_Code
Mgt_Unit_Level_1_Code
Central_Flag
COB_Line
Direct_Expense_Category
Fund_Sub_Category_Code
Local_Budget_Unit_Code
Local_Mgt_Unit_Level_1_Code
Extract_Date
Amount
COB_OUTPUT_INDIRECT_FFE
Fiscal_Year
FY_Reporting_Period_Code
Scenario
Revision_Level
MANR_Unit_Rollup_Type_Groups
Budget_Unit_Code
Mgt_Unit_Level_1_Code
Central_Flag
COB_Line
Direct_Expense_Category
Local_Budget_Unit_Code
Local_Mgt_Unit_Level_1_Code
Extract_Date
Amount
COB_LEDGER_INDIRECT
Fiscal_Year
FY_Reporting_Period_Code
Scenario
Revision_Level
MANR_Unit_Rollup_Type_Groups
Budget_Unit_Code
Mgt_Unit_Level_1_Code
Central_Flag
COB_Line
Direct_Expense_Category
Fund_Sub_Category_Code
Local_Budget_Unit_Code
Local_Mgt_Unit_Level_1_Code
Amount
Extract_Date
{1,1}
FUND_SUB_CATEGORY
Fund_Category_Code
Fund_Sub_Category_Code
Fund_Sub_Category_Desc
{1,1}
FUND_CATEGORY
Fund_Category_Code
Fund_Category_Description

ERD EXAMPLE

Dimensional Model of Store Sales System



ERD EXAMPLE


MARKS
MarkId(Long)
PK:PK_Marks
Student_Id(Long)
FK:FR_6
Subject_Id(Long)
FK:FR_5
Date(TimeStamp)
Mark(Integer)
SUBJECTS
Subject_Id(Long)
PK:PK_Subjects
Title(Text)
TEACHERS
Teacher_Id(Long)
PK:PK_Teachers
FirstName(Text)
LastName(Text)
STUDENTS
Student_Id(Long)
PK:PK_Students
FirstName(Text)
LastName(Text)
Group_Id(Long)
FK:FR_4
SUBJ_TEACH
ST_Id(Long)
PK:PK_Subj_Teach
Subject_Id(Long)
FK:FR_3
Teacher_Id(Long)
FK:FR_2
Group_Id(Long)
FK:FR_1
GROUPS
Group_Id(Long)
PK:PK_Groups
Name(Text)

موفق باشید