

Relational Database Management for Epidemiologists: Entity-Relationship Diagrams and Planning Table Structure

Wayne Enanoria, PhD, MPH
Center for Infectious Disease Preparedness
University of California at Berkeley

Phases of Database Design

- Define Mission Statement and Objectives
- Analyze the current database
- Create data structures
- **Establish table relationships**
- Define business rules
- Determine and establish views
- Review data integrity

From Last Time...

- Engaged users and management in the design process
- Reviewed existing databases and organizational structure
- Lists
 - ◆ Subjects
 - ◆ Preliminary Fields (Characteristics)
 - Fields with Value Lists
 - Calculated Fields

Outline

- E-R Diagrams
- Developing Tables
 - ◆ Elements of a Table
 - ◆ Developing from a Field List
 - ◆ Relationships

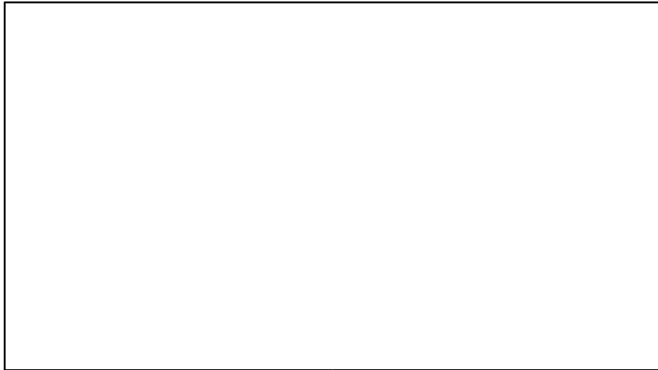
Purpose of ERDs

- An E-R Diagram shows a database's entities and the relationships among the entities in a symbolic, visual way.
 - ◆ Documents a project or application and serves as its foundation
 - ◆ Clarify the system/application's features
 - ◆ Provide a basis for evaluating development options

Specifications

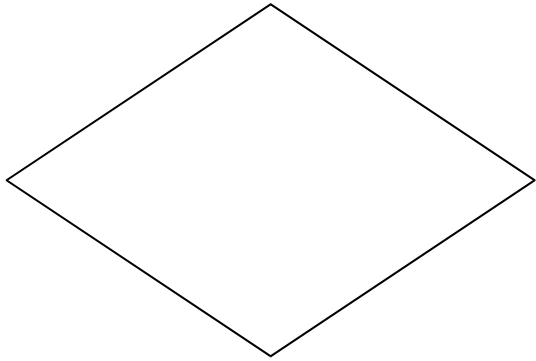
- The E-R model is a detailed, logical representation of the data for an organisation or business area
- It should be understandable to both the user and to the IT technologist
- The model must be as ‘open’ as possible and not tied to any technology or to any particular business methodology
- It must be flexible enough so that it can be used and understood in practically any environment where information is modelled

Entity



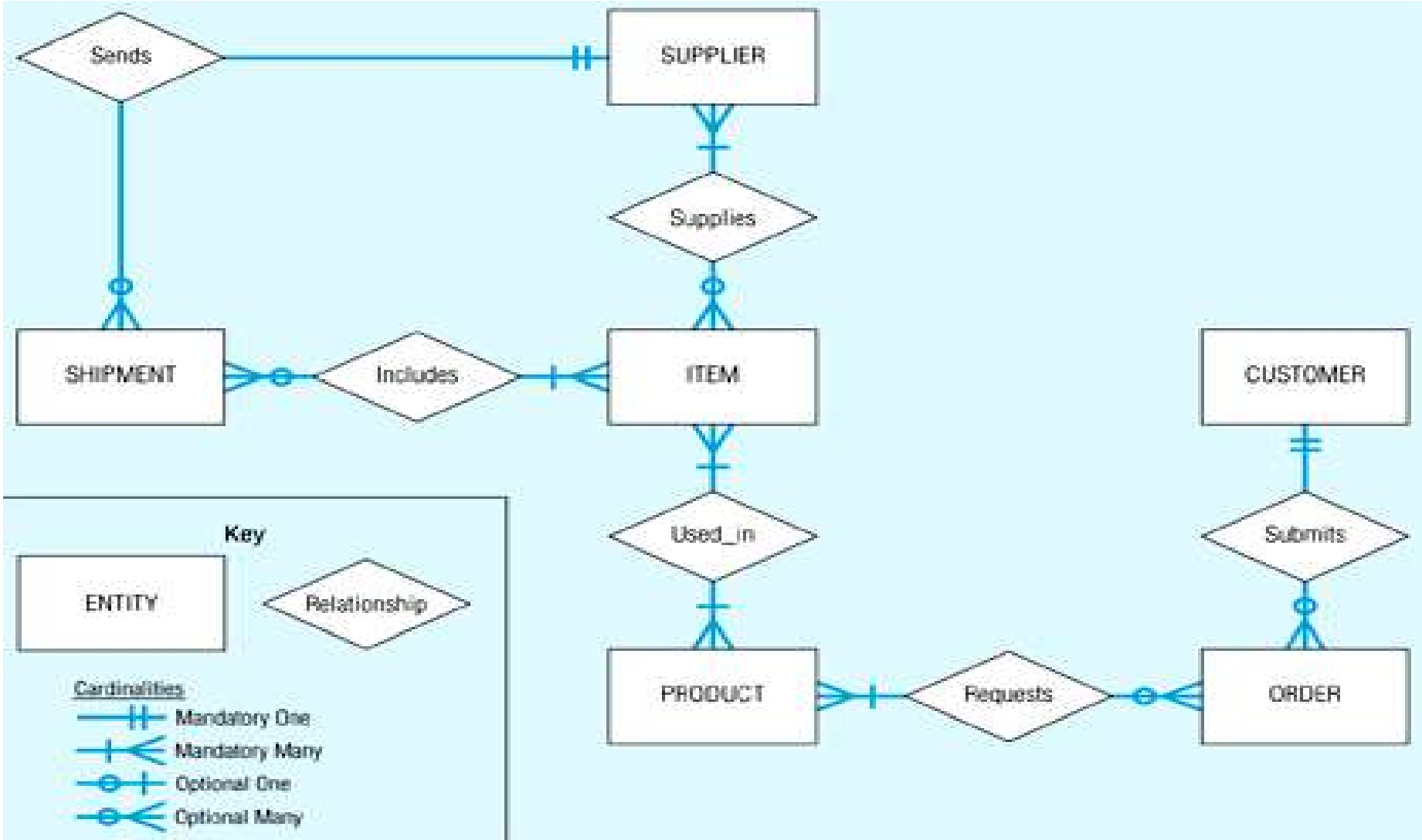
- A box represents an entity; a person, place, object, event, concept.
- Usually written in singular (eg, employee, case, occupation, etc.)

Relationship



- A diamond represents the relationship between entities.

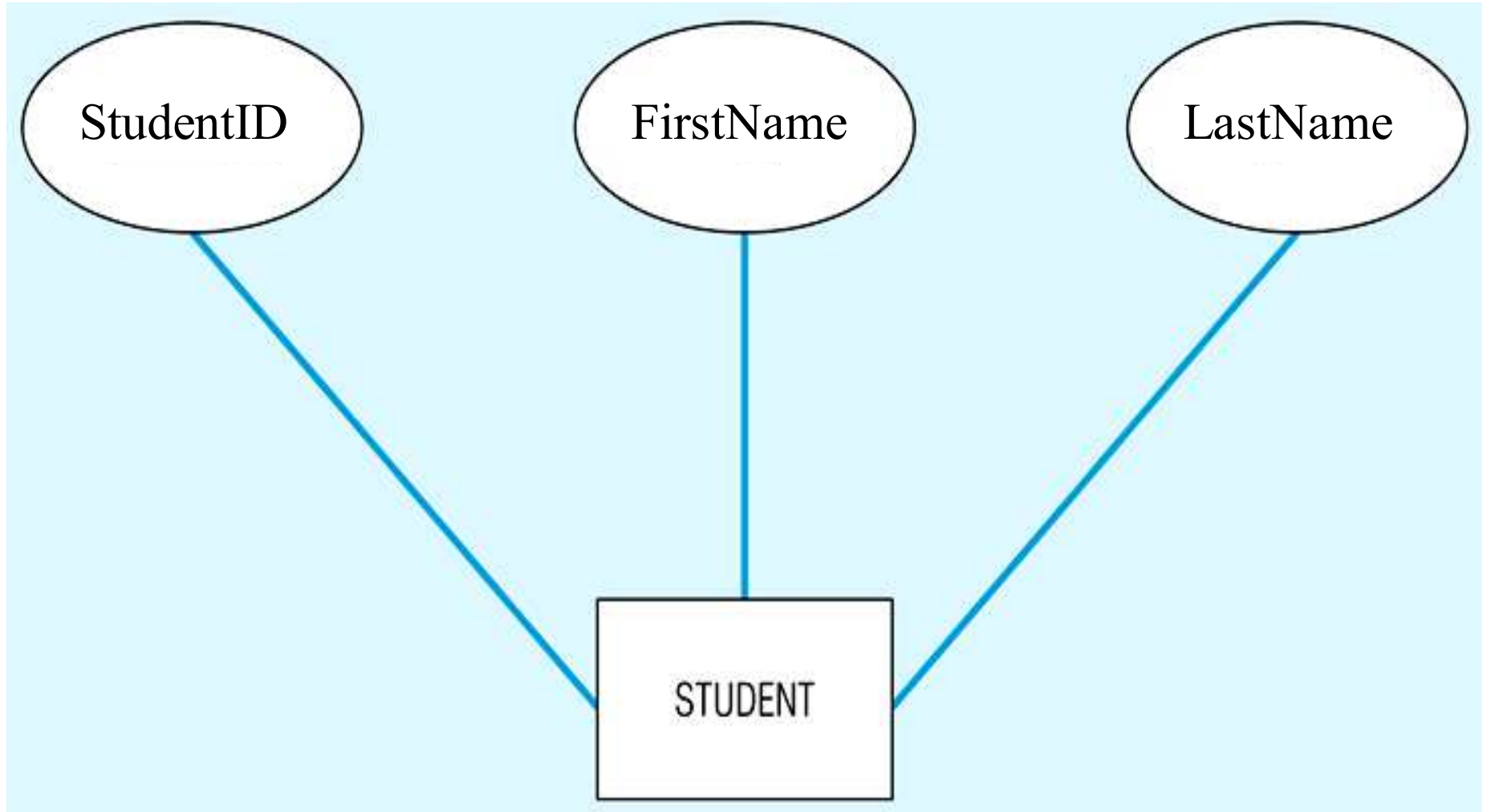
Sample E-R Diagram



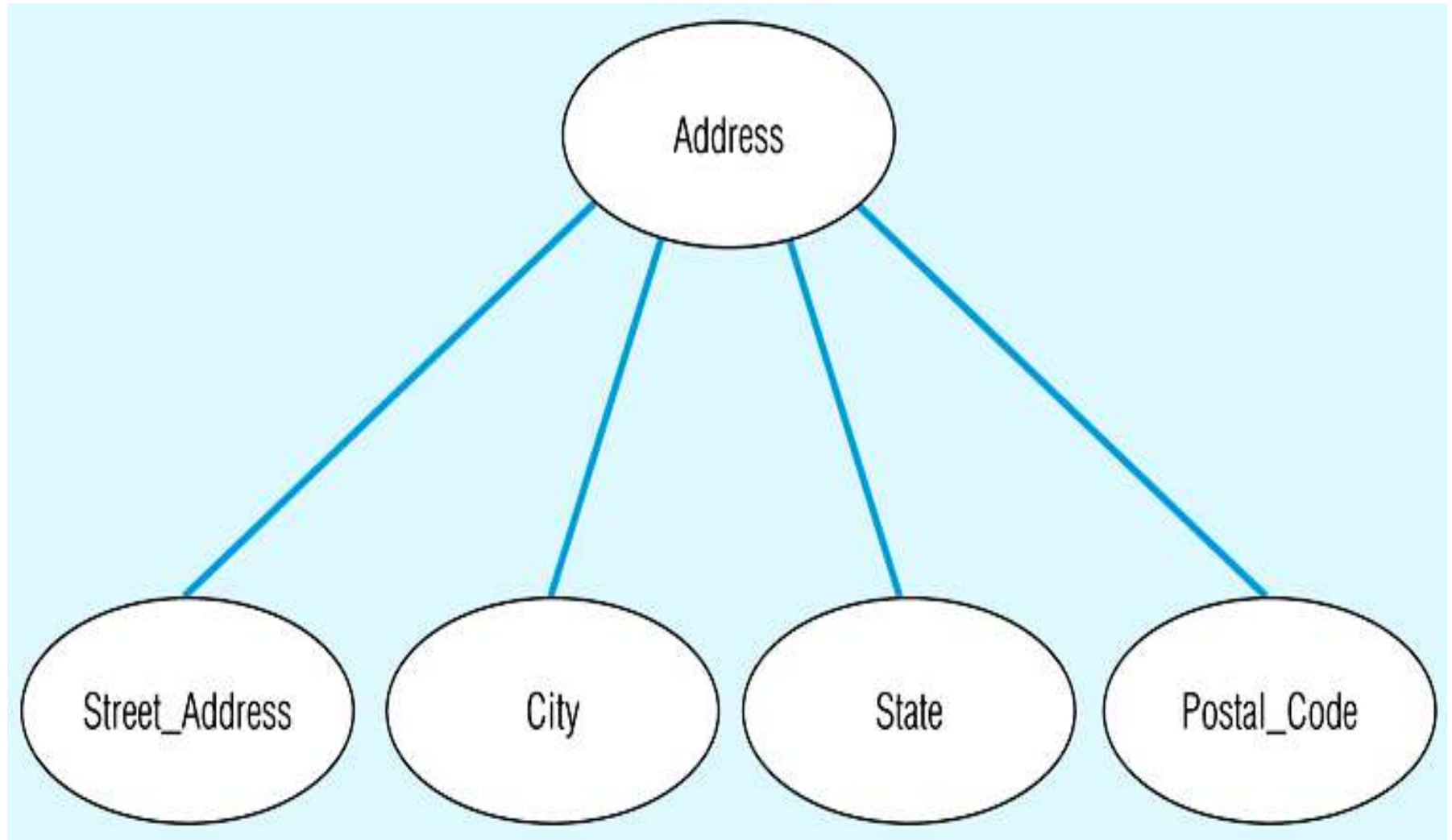
Attribute

- An attribute is a property or characteristic of an entity type, for example the entity EMPLOYEE may have attributes Employee_Name and Employee_Address.
- In ER diagrams place attributes name in an ellipse with a line connecting it to its associated entity
- Attributes may also be associated with relationships
- An attribute is associated with exactly one entity or relationship

Simple Attributes



Composite Attribute



How do we use the E-R diagram?

- Review the diagram with the client or business professional (the person who knows the data).
- Create tables from the entities.
- Create fields (attributes) from the entity's attributes.
- Transform the ERD relationships into keys linking the tables.

Elements of a Table

- Name
- Type
- Description
- Primary Key
- Fields

Table Names

- unique, descriptive name that describes the subject
- concise; use the minimum number of words to convey the subject
- do not use acronyms or abbreviations
- no plurals; a table represents a single subject (object or event)

Types of Tables

- Data Table
- Linking Table
- Validation Table
- Subset Table

Table Descriptions

- Description defines the table and states its importance
- Focused, concise, unambiguous, and clear!
- Example:

“Student Schedule: Those classes that the student is scheduled to attend during this school year. The information provided by this table helps the student implement effective time management and enables the school to figure class loads and student loads.”

--Michael J. Hernandez, Database Design for Mere Mortals.

Primary Key

- The attribute of an entity that uniquely defines it throughout the database structure; it helps establish a table's relationship with other tables.
- A primary key uniquely identifies a given record within a table and represents that record throughout the database.

Elements of a Primary Key

- Unique values
- Comprises a minimum number of fields
- Identify each record in the table
- Not a multipart field
- Cannot contain null values
- Not optional

Each table should have one and only one primary key!

Example

caseid	casefirst	caselast	Dob	Phone
1	joe	smith	2000-01-14	555-5467
2	alan	smith	1968-03-14	555-5467
3	maria	smith	1966-02-15	555-5467
4	henry	smith	1998-08-15	555-5467

Foreign Key

- Foreign keys are attributes that do not belong to an entity, but are used by a database to link back to another (foreign) entity.
- Foreign keys have the same name as the primary key from which it was copied.
- Foreign keys use a replica of the field specifications for the primary key.

Example

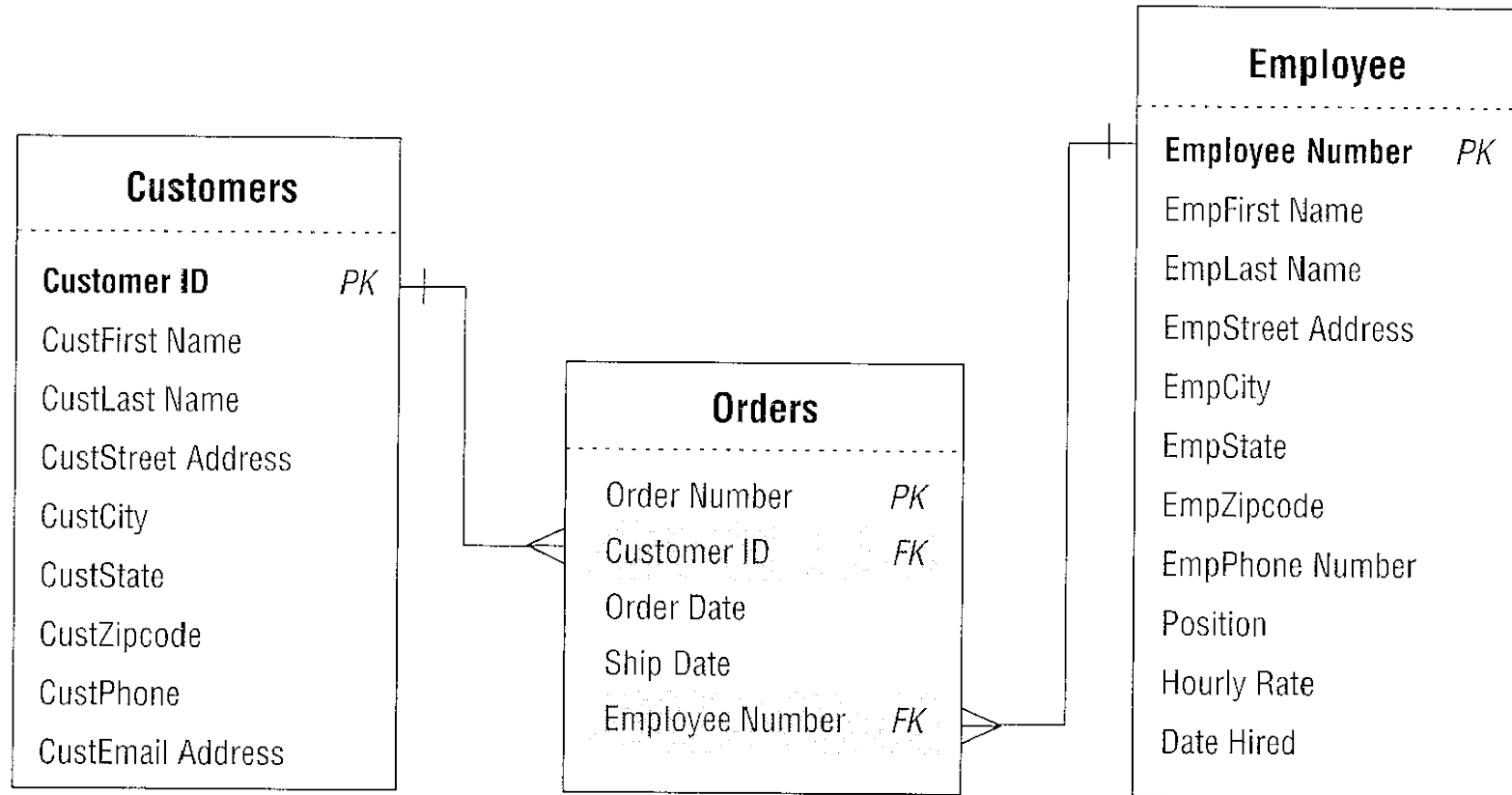


Figure 10.52. Foreign keys that comply with the first element of a foreign key.

SOURCE: Michael J. Hernandez. Database Design for Mere Mortals. Second Edition, Addison-Wesley 2003.

Fields

- Determine which fields best represent characteristics of the table's subject and assign them to that table.

Field List

- Firstname
- Lastname
- Ageat onset
- City
- County
- State
- Race
- Ethnicity
- Date of Symptom
- Fever
- Chest Xray
- Travel Hx
- Date of Arrival
- Date of Departure
- ContactFirst
- ContactLast
- Classification
- Specimen
- Date Collected
- Test Requested
- LabSource
- Result
- Sent to CDC

Field List

CASE

- Firstname
- Lastname
- Ageat onset
- City
- County
- State
- Race
- Ethnicity
- DateofSymptom
- Fever
- ChestXray
- TravelHx
- Date of Arrival
- Date of Departure

- ContactFirst
- ContactLast
- Classification
- Specimen
- DateCollected
- TestRequested
- LabSource
- Result
- Sent to CDC

Field List

- Firstname
- Lastname
- Ageat onset
- City
- County
- State
- Race
- Ethnicity
- DateofSymptom
- Fever
- ChestXray
- TravelHx
- Date of Arrival
- Date of Departure
- ContactFirst
- ContactLast
- Classification
- Specimen
- DateCollected
- TestRequested
- LabSource
- Result
- Sent to CDC

Clinical Information

Field List

- Firstname
- Lastname
- Ageat onset
- City
- County
- State
- Race
- Ethnicity
- Date of Symptom
- Fever
- Chest Xray
- Travel Hx
- Date of Arrival
- Date of Departure
- Contact First
- Contact Last
- Classification
- Specimen
- Date Collected
- Test Requested
- Lab Source
- Result
- Sent to CDC

Travel History

Field List

- Firstname
- Lastname
- Ageatonsset
- City
- County
- State
- Race
- Ethnicity
- DateofSymptom
- Fever
- ChestXray
- TravelHx
- Date of Arrival
- Date of Departure

- ContactFirst
- ContactLast
- Classification
- Specimen
- DateCollected
- TestRequested
- LabSource
- Result
- Sent to CDC

Contact

Field List

- Firstname
- Lastname
- Ageat onset
- City
- County
- State
- Race
- Ethnicity
- Date of Symptom
- Fever
- Chest Xray
- Travel Hx
- Date of Arrival
- Date of Departure

- ContactFirst
- ContactLast
- Classification
- Specimen
- DateCollected
- TestRequested
- LabSource
- Result
- Sent to CDC

Specimen

Field List

Case

Clinical

Travel

Contact

Specimen

• Firstname	• DateofSymptom	• TravelHx	• ContactFirst	• Specimen
• Lastname	• Fever	• Dateof Arrival	• ContactLast	• DateCollected
• Ageatonsset	• ChestXray	• Date of Departure	• Classification	• TestRequested
• City			•	• LabSource
• County				• Result
• State				• Sent to CDC
• Race				
• Ethnicity				

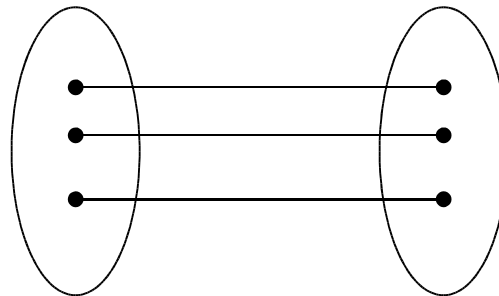
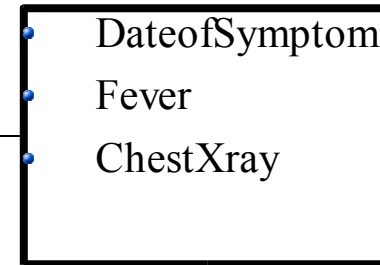
Table Relationships

- One-to-One
- One-to-Many
- Many-to-Many

Example of One-to-One

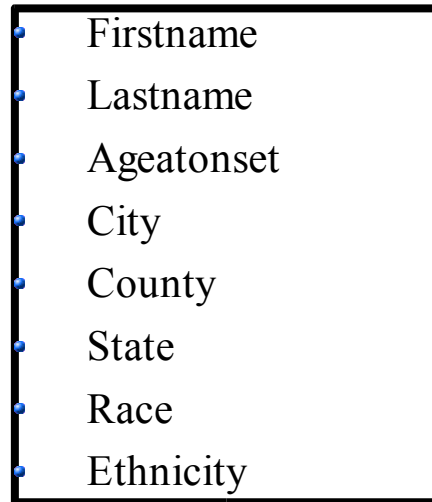
Case

Clinical

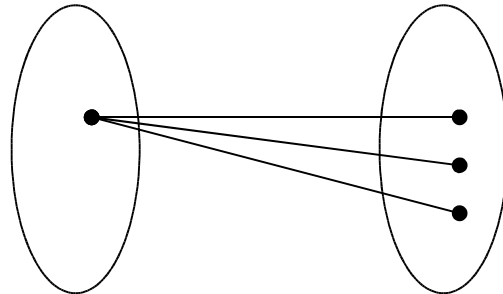
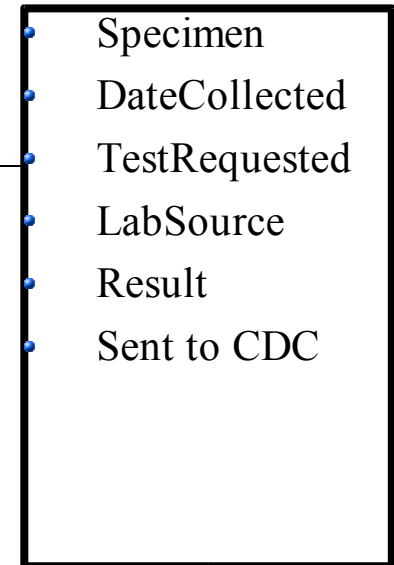


Example of One-to-Many

Case



Specimen

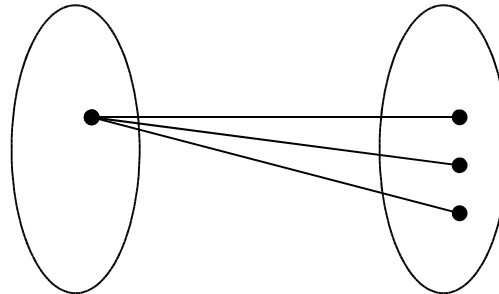
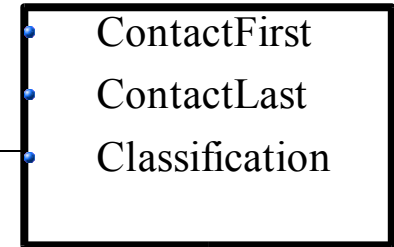


Example of One-to-Many

Case



Contact

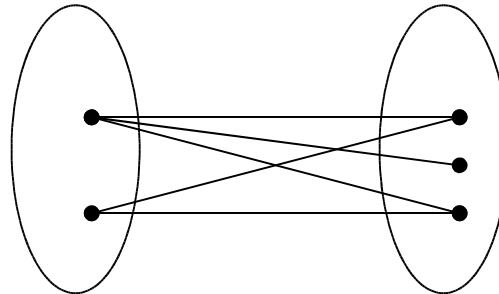
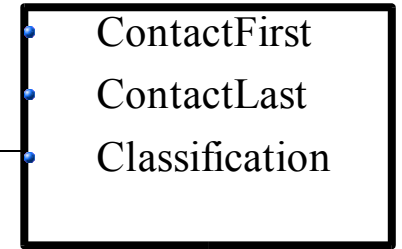


Example of Many-to-Many

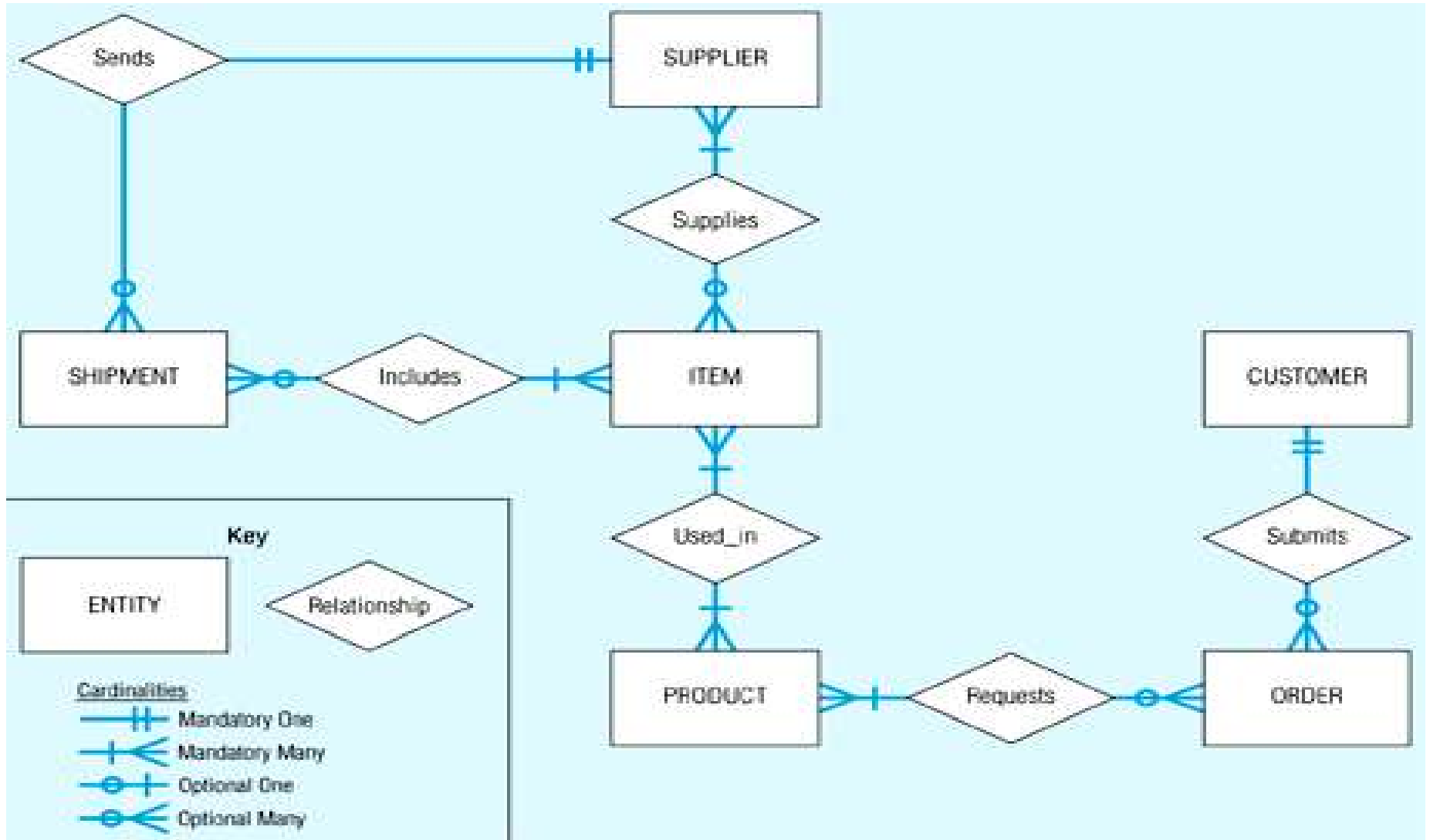
Case



Contact



Sample E-R Diagram



At the End of the Day...

- Reviewed the ERD with key personnel
- Developed a list of fields and tables
- Developed preliminary table relationships

Next Time

- Normalization
- Field Specification