

# The Relational Model

CIS 3730

Designing and Managing Data

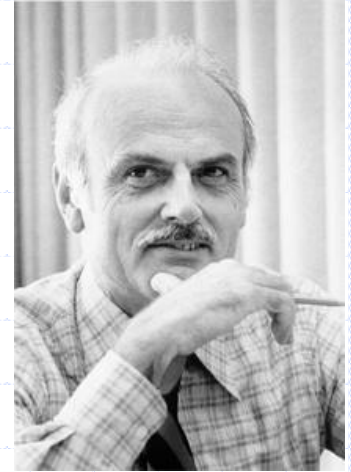
J.G. Zheng  
Fall 2010



# Overview

- ◆ What is the relational model?
- ◆ What are the most important practical elements of the relational model?

# Introduction



- ◆ Edgar F. Codd (IBM), 1970

- ◆ One sentence to summarize relational database model (extremely brief):

Data are organized in *relations* (tables), which are linked by *keys* (constraints)

# Relational Model in MS Access

The image shows three tables in MS Access:

- RETAIL\_ORDER**: Contains columns OrderNumber, StoreNumber, StoreZip, OrderMonth, OrderYear, and OrderTotal. It shows three records with OrderNumber values 1000, 2000, and 3000.
- SKU\_DATA**: Contains columns SKU, SKU\_Description, Department, and Buyer. It shows eight records with SKU values ranging from 100100 to 302000.
- ORDER\_ITEM**: Contains columns OrderNumber, SKU, Quantity, Price, and ExtendedPrice. It shows seven records linking OrderNumber and SKU.

Arrows indicate relationships:

- OrderNumber in RETAIL\_ORDER is linked to OrderNumber in ORDER\_ITEM.
- SKU in ORDER\_ITEM is linked to SKU in SKU\_DATA.
- OrderNumber 1000 in RETAIL\_ORDER is linked to OrderNumber 1000 in ORDER\_ITEM.
- OrderNumber 2000 in RETAIL\_ORDER is linked to OrderNumber 2000 in ORDER\_ITEM.
- OrderNumber 3000 in RETAIL\_ORDER is linked to OrderNumber 3000 in ORDER\_ITEM.
- OrderNumber 1000 in RETAIL\_ORDER is linked to SKU 201000 in SKU\_DATA.

Data are organized in *relations* (tables), and linked by keys.

# Relation

- ◆ A relation is a two-dimensional table that has some specific characteristics:
  1. The table consist of rows and columns
  2. Rows contain data about entity instances
  3. All values in a row describes the same entity instance
  4. Columns contain data about attributes of the entity
  5. All values in a column are of the same kind
  6. Each row is distinct
  7. A cell of the table holds a single value
  8. Each column has a unique name
  9. The order of the rows is unimportant
  10. The order of the columns is unimportant

# Relation Example

EmployeeNumber	FirstName	LastName	Department	Email	Phone
100	Jerry	Johnson	Accounting	JJ@somewhere.com	834-1101
200	Mary	Abernathy	Finance	MA@somewhere.com	834-2101
300	Liz	Smathers	Finance	LS@somewhere.com	834-2102
400	Tom	Caruthers	Accounting	TC@somewhere.com	834-1102
500	Tom	Jackson	Production	TJ@somewhere.com	834-4101
600	Eleanore	Caldera	Legal	EC@somewhere.com	834-3101
700	Richard	Bandalone	Legal	RB@somewhere.com	834-3102

# Non-Relation Examples

EmployeeNumber	FirstName	LastName	Department	Email	Phone
100	Jerry	Johnson	Accounting	JJ@somewhere.com	236-0000
200	Mary	Abernathy	Finance	MA@somewhere.com	444-8898
300	Liz	Smathers	Finance	LS@somewhere.com	777-0098
400	Tom	Caruthers	Accounting	TC@somewhere.com	236-0000, 236-0991, 236-0991
500	Tom	Jackson	Production	TJ@somewhere.com	444-9980
600	Eleanore	Caldera	Legal	EC@somewhere.com	767-0900
700	Richard	Bandalone	Legal	RB@somewhere.com	767-0900, 767-0011

EmployeeNumber	FirstName	LastName	Department	Email	Phone
100	Jerry	Johnson	Accounting	JJ@somewhere.com	236-9987
200	Mary	Abernathy	Finance	MA@somewhere.com	444-8898
300	Liz	Smathers	Finance	LS@somewhere.com	777-0098
400	Tom	Caruthers	Accounting	TC@somewhere.com	236-9987 Fax: 236-9987 Home: 555-7171
500	Tom	Jackson	Production	TJ@somewhere.com	444-9980
600	Eleanore	Caldera	Legal	EC@somewhere.com	767-0900 Fax: 236-9987 Home: 555-7171
700	Richard	Bandalone	Legal	RB@somewhere.com	767-0900

# Terminology Contrast

<i>Database industry</i>	<b>Table</b>	<b>Row</b>	<b>Column</b>
<i>Academic</i>	Relation	Tuple	<b>Attribute</b>
<i>File processing</i>	File	<b>Record</b>	<b>Field</b>

# Key

- ◆ A key is one or more columns of a relation that is used to identify a record
  - Unique keys
    - ◆ Primary key
    - ◆ Candidate key
    - ◆ Alternate key
    - ◆ Surrogate key
  - Foreign key
  - Composite key

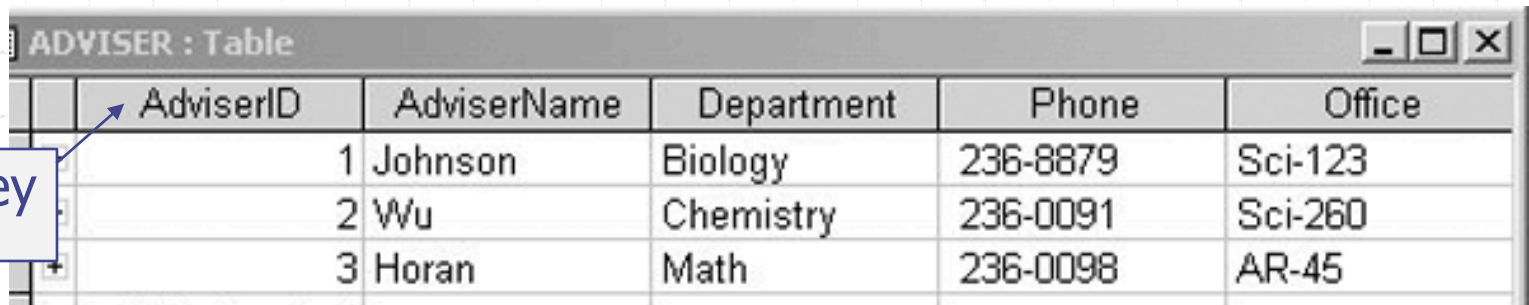
# Candidate Key/Primary Key

## ◆ Candidate key

- The minimum set of column(s) that uniquely identifies a single record (row)
- Each value in this column is unique in this relation

## ◆ Primary key

- Is one of the candidate keys chosen to be the identifying key; others become alternate keys



AdviserID	AdviserName	Department	Phone	Office
1	Johnson	Biology	236-8879	Sci-123
2	Wu	Chemistry	236-0091	Sci-260
3	Horan	Math	236-0098	AR-45

Primary Key

# Primary Key (PK)

- ◆ Primary key is a column/attribute that is used to uniquely identify a record
  - Each value of this key column uniquely identifies a single record (row)
  - There is only ONE primary key for a table
- ◆ Which of the following attribute can uniquely identify a single student?
  - Last Name
  - Full Name
  - Birth date
  - Student Number
  - Social Security Number

# Composite Key

## ◆ Composite key

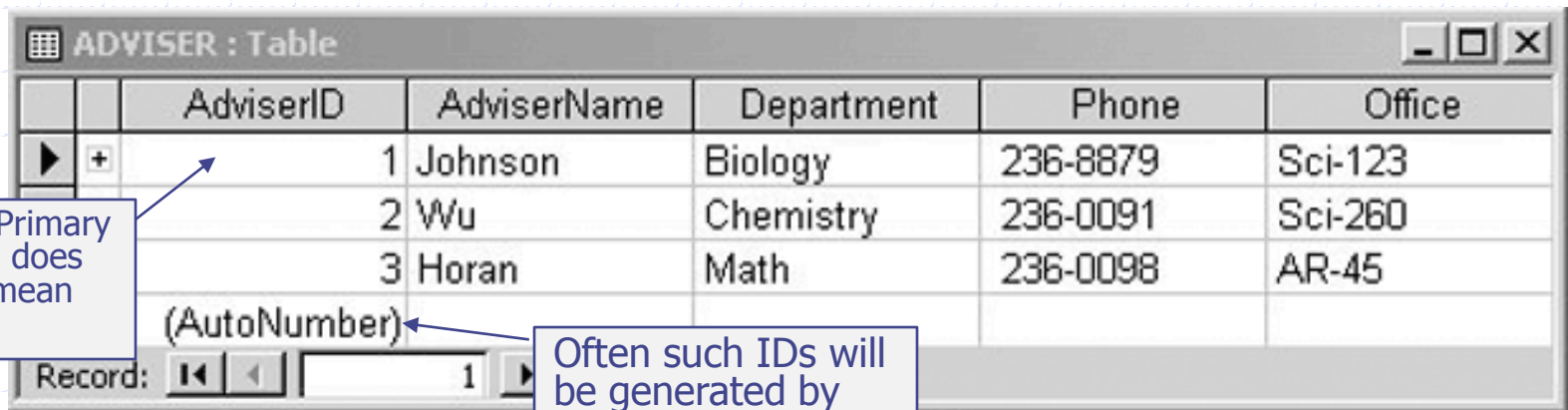
- A composite key contains two or more attributes (columns)
- All keys can be composite keys

## ◆ Example:

- "FirstName" + "LastName"
- "FirstName" + "LastName" + "BirthDate"
- "FirstName" + "LastName" + "BirthDate" + "BirthCity"
- ...

# Artificial Primary Key/Surrogate Key

- ◆ Sometimes it is difficult to find a natural attribute as a primary key, or it is difficult to use a composite key.
- ◆ A column is created arbitrarily and assign each record a unique number/id
  - Product Number, Product Id, Movie Id, Actor Id, etc.



AdviserID	AdviserName	Department	Phone	Office
1	Johnson	Biology	236-8879	Sci-123
2	Wu	Chemistry	236-0091	Sci-260
3	Horan	Math	236-0098	AR-45

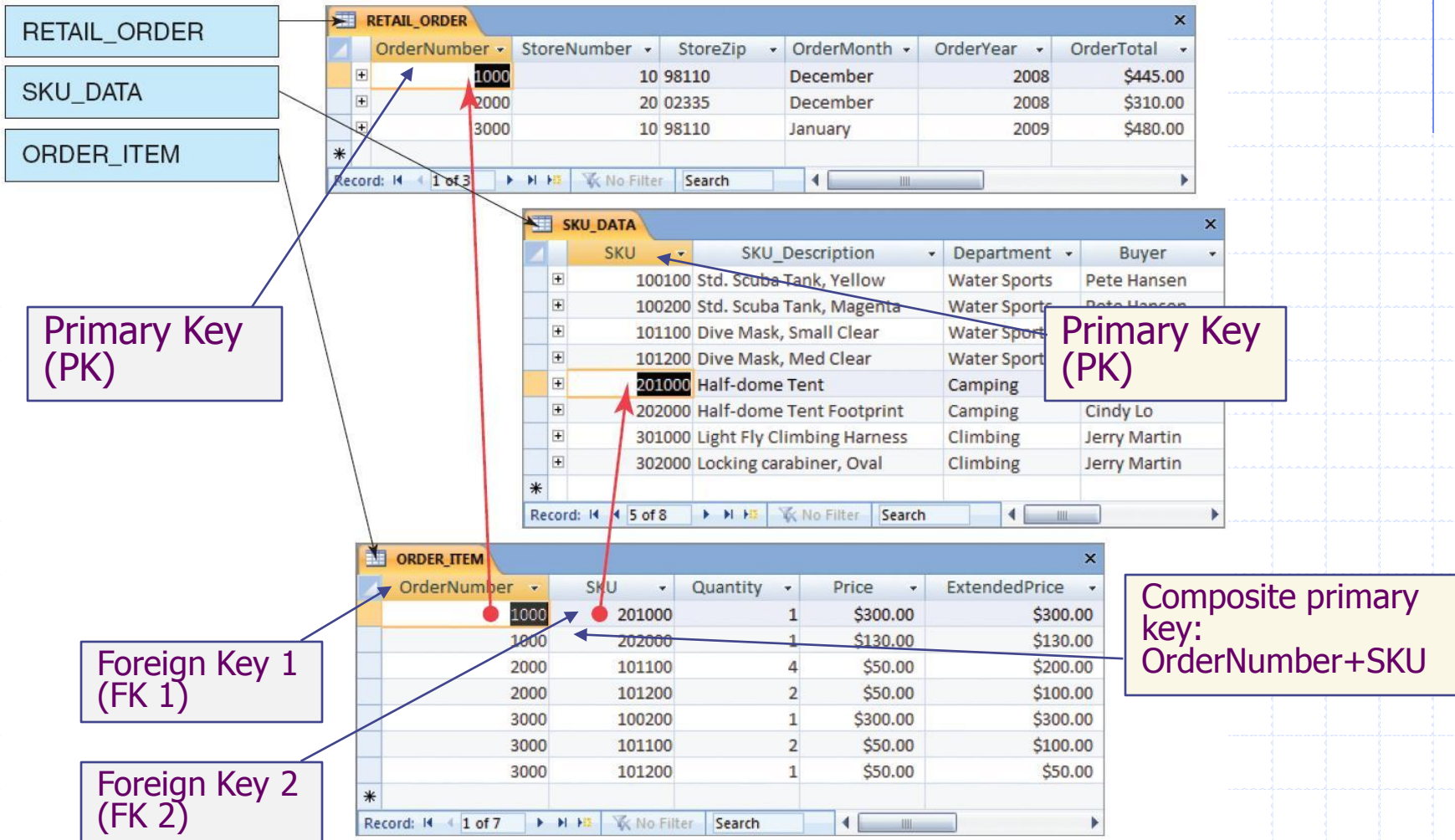
Surrogate Primary Key: the id does not really mean anything.

Often such IDs will be generated by database systems.

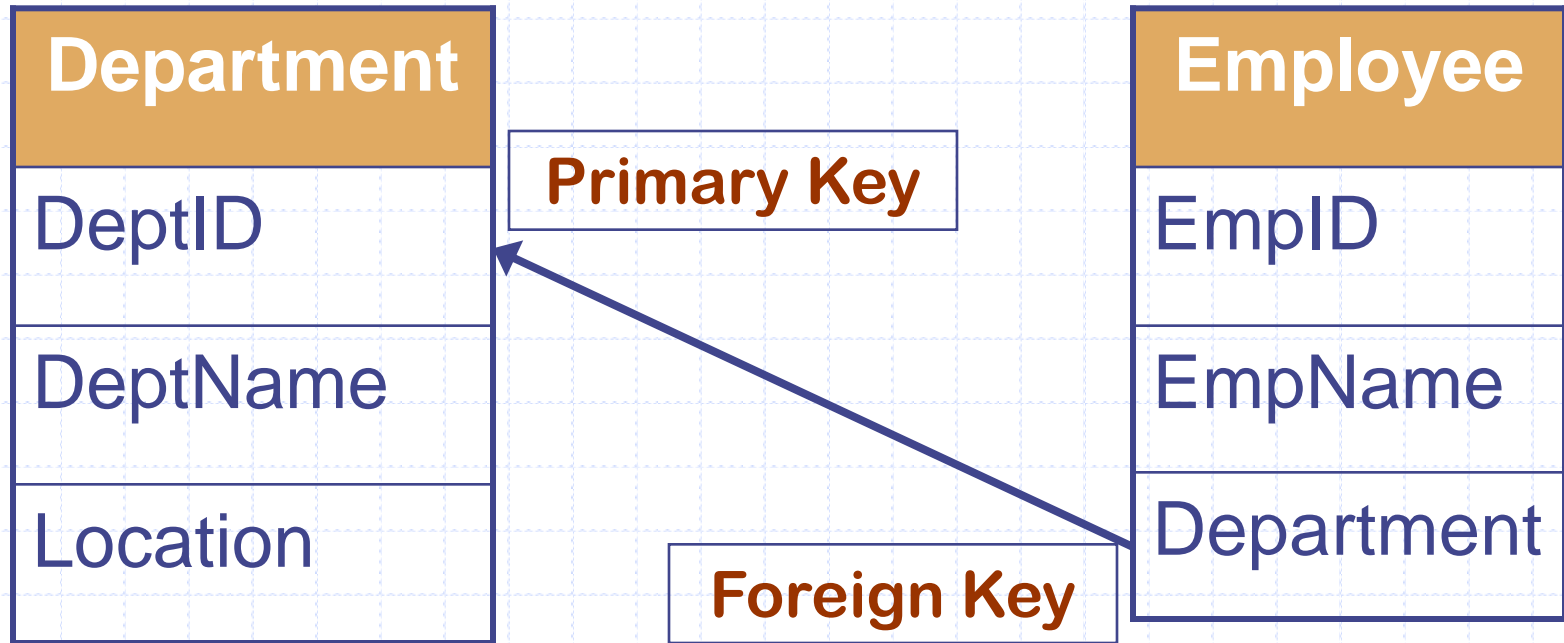
# Relationship and Foreign Key (FK)

- ◆ Relationship is how tables (relations) are linked
  - It is defined by the *foreign key (FK)* constraint
- ◆ A foreign key references a primary key (or any other unique keys) in another table
  - This *pair* of keys are of the *same kind* (may be of different name)

# Relationship and FK Example



# Foreign Key Example

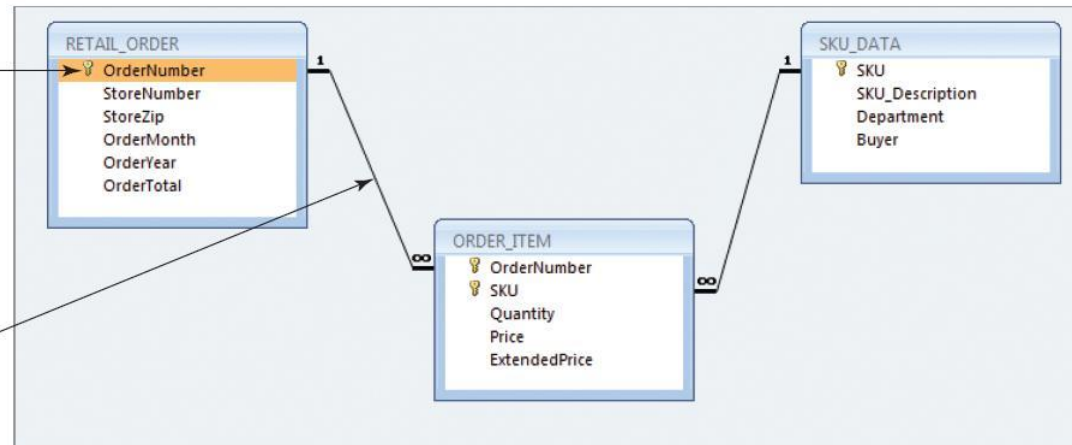


Primary Key and Foreign Key are of the same type (string, number, etc.) and length, but they do not necessarily have the same name.

# Relationship in MS Access

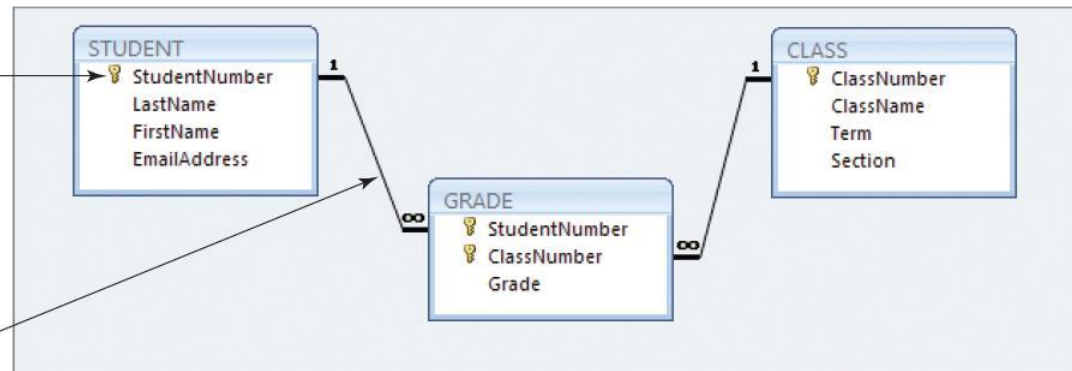
The RETAIL\_ORDER table—the key symbol shows the primary key

The relationship between RETAIL\_ORDER and ORDER\_ITEM—the number 1 and the infinity symbol indicate that one retail order may be linked to many order items by OrderNumber



The STUDENT table—the key symbol shows the primary key

The relationship between STUDENT and GRADE—the number 1 and the infinity symbol indicate that one student may be linked to many classes by StudentNumber

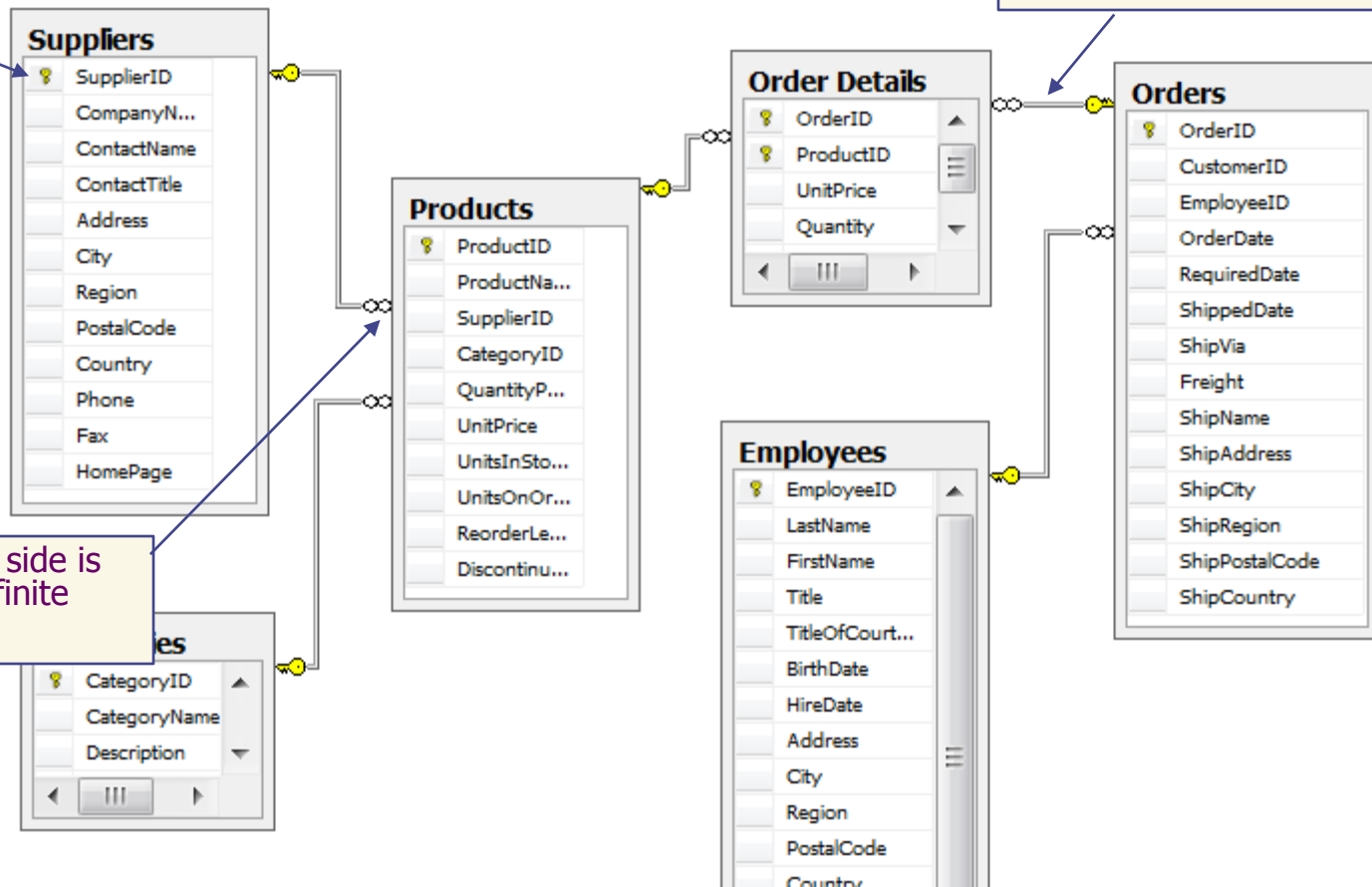


# Relationship in SQL Server Database

HOME-WORK-JACK...2003-mini - ERD

Note: in SQL Server, the line ends do not exactly point to the right column

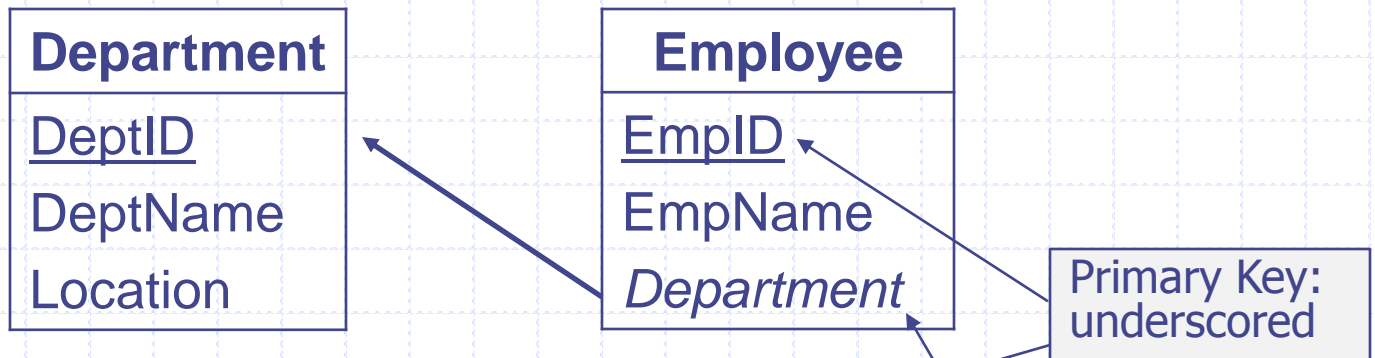
Primary Key (PK) is indicated by a key symbol



Foreign Key (FK) side is indicated by a infinite loop symbol.

# Database Structure Representation

## ◆ Graphic/diagram style



## ◆ Text style

Department (DeptID, DeptName, Location)

Employee (EmpID, EmpName, *Department*)

FK: Department → Department.DeptID

Foreign Key: italicized, and additional definition is just below the table definition

# Schema in Relational Databases

- ◆ Schema is the structure described in a formal language supported by the database management system (DBMS)
  - It's a kind of metadata
- ◆ Relational database schema commonly defines
  - Tables: name, primary key
  - Columns: name, data type, size, value range, etc.
  - Constraints: all kinds of keys
  - Other structures

# Common Data Types

## ◆ Numeric

- Integer: 0, 4, 20, -200
- Decimal: 3.14, 4.0

## ◆ Character/String

- Character: 'A', 't', '4', '\$'
- String/Text: 'cis 3730'

## ◆ Date/Time

- Date: 8/23/2010
- Date/Time: 8/23/2010 12:30:00AM

## ◆ Bit/Boolean

- Yes/No, true/false, 0/1

# “Null” Value

## ◆ “Null” means

- A value is inapplicable: Fax\_Number = Null
- A value is unknown: Stud\_Grade = Null

## ◆ “Null” *does NOT* mean

- Zero, or blank space

## ◆ No primary key (or part of a composite primary key) can be NULL

	EmployeeID	LastName	FirstName	Title	HireDate	City	Region
1	7	King	Robert	Sales Representative	1994-01-02 00:00:00.000	London	NULL
2	8	Callahan	Laura	Inside Sales Coordinator	1994-03-05 00:00:00.000	Seattle	WA
3	9	Dodsworth	Anne	Sales Representative	1994-11-15 00:00:00.000	London	NULL
4	10	Zheng	Jack	Consultant	NULL	Atlanta	GA

“NULL” in SQL Server database tables

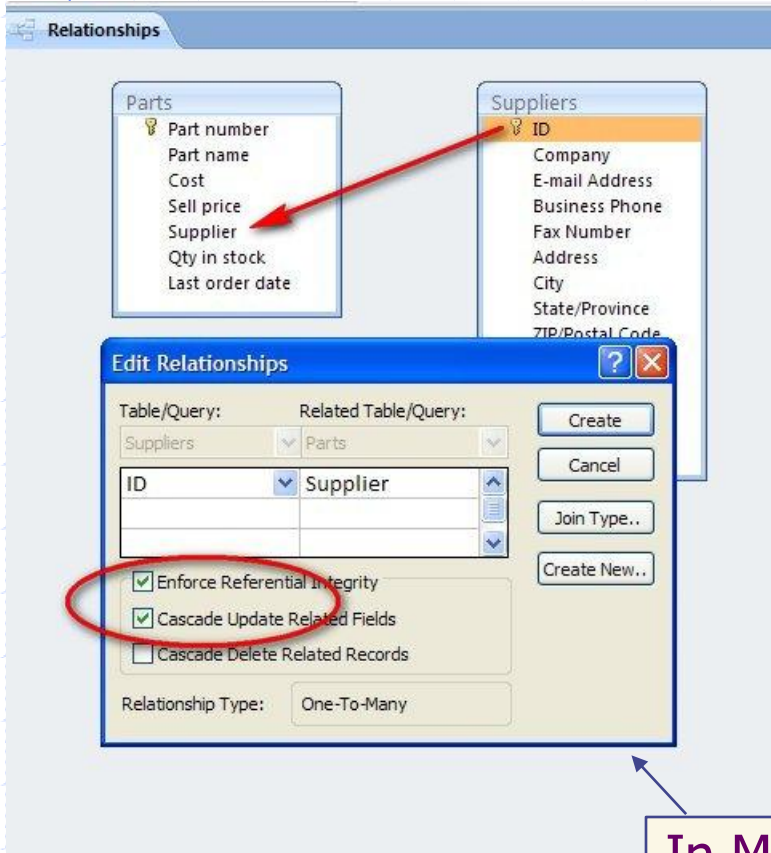
# Data Integrity

- ◆ Data integrity: the correctness or quality of data
- ◆ Relational databases ensure data integrity in three aspects:
  - Entity Integrity: no empty or duplicate rows in a table.
    - ◆ Use primary key constraints
  - Domain Integrity: valid entries for a given column is restricted by the type, the format, or the range of possible values
    - ◆ Use "Check" rules on columns
  - Referential integrity: rows cannot be deleted or modified, if they are referenced by other records in other tables
    - ◆ Use foreign key constraints

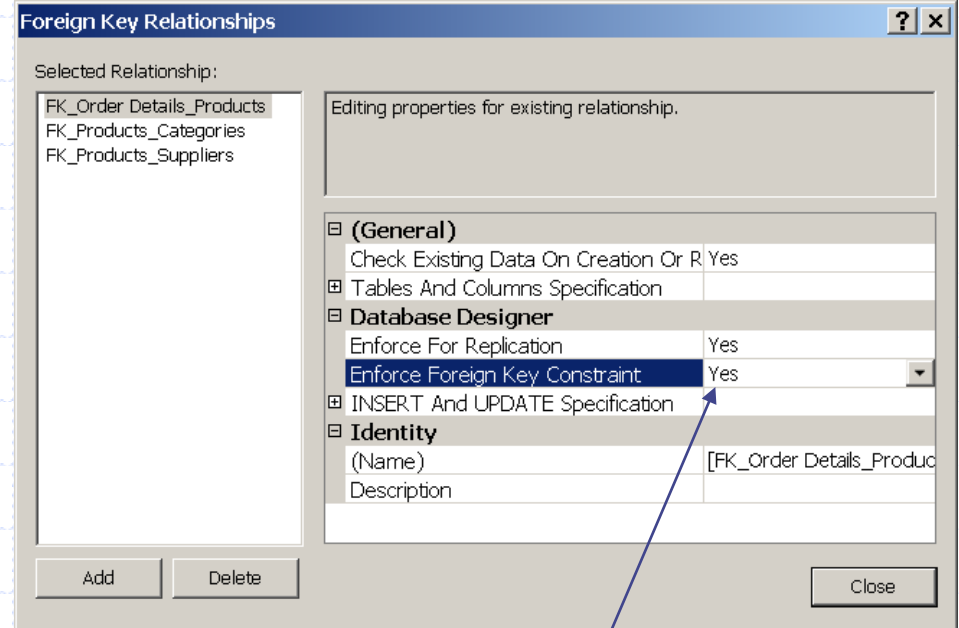
# Referential Integrity

- ◆ Every value of a foreign key must match a value of the primary key or be "NULL"
- ◆ Example (slide #16)
  - In the "Employee" table, "Department" is a foreign key (linked to the "Department" table where "DeptID" is the primary key).
  - Then every value of "Department" in the "Employee" table must exist in the "Department" table first, otherwise the data will be rejected, if the referential integrity is enforced.

# Enforcing Referential Integrity in Databases



In Microsoft Access 2007



In SQL Server

# Summary

## ◆ Key concepts

- Relation (and its 10 features)
- Relational model terminologies: Row, column, record, field, attribute
- Keys
  - ◆ Primary key, candidate key, surrogate key
  - ◆ Composite key
  - ◆ Foreign key
- Database schema and metadata
- Basic data types: number, character, date, bit
- NULL
- Data integrity: entity, domain, referential
  - ◆ How relational database systems ensure these three kinds of data integrity

## ◆ Key skills

- Be able to explain if a table is not a relation
- Identify the primary key, composite primary key, candidate keys, and foreign keys of a given table/relation.
- Understand and can use the graphical and text representation of database structures: tables, columns and keys.