CS 327E Lecture 4

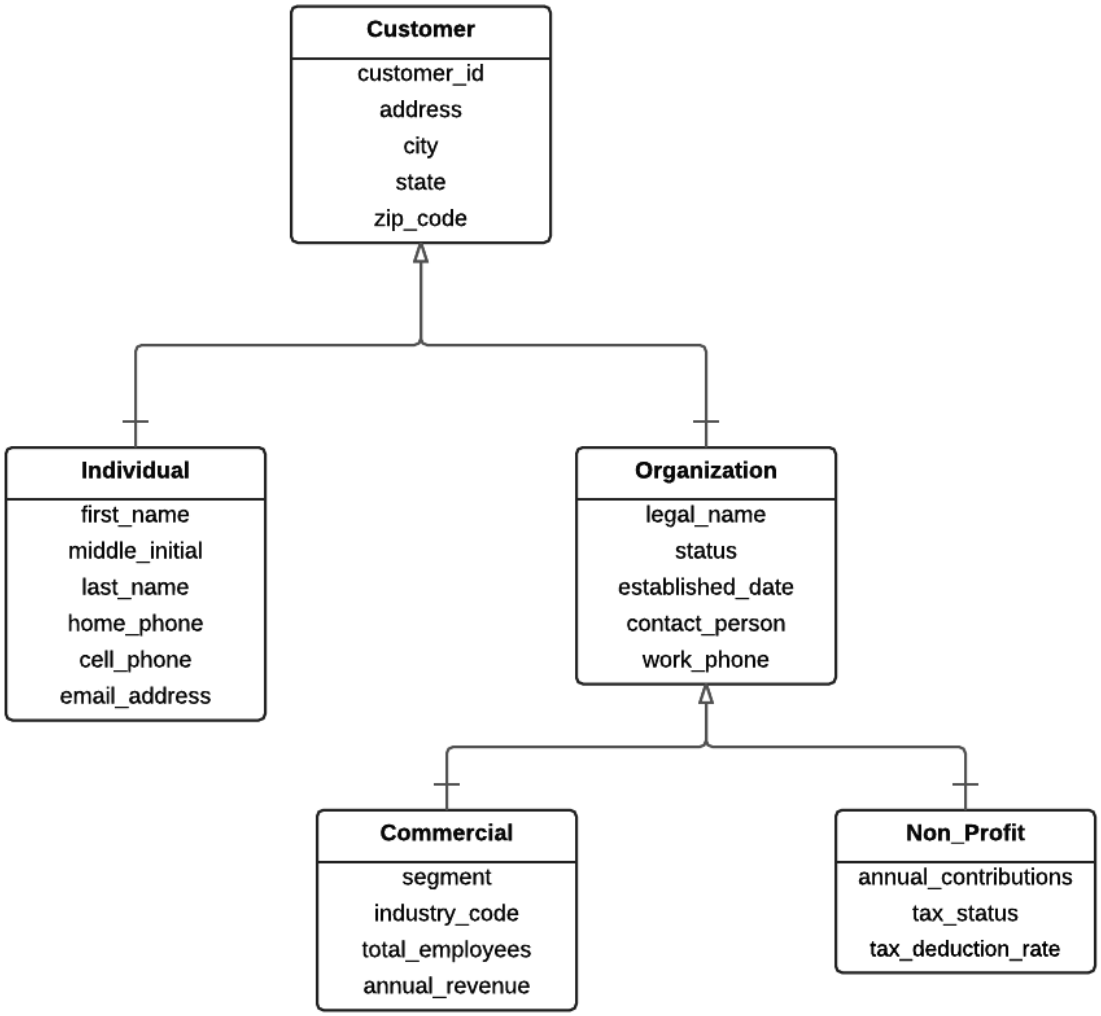
Shirley Cohen

September 12, 2016

Announcements

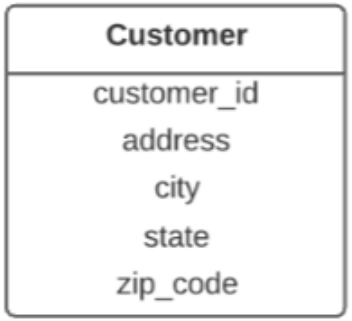
- Lab 1 will be out this Wed and due following Friday
- Setup session for Lab 1 this Wed
- Only one Reading Quiz next week during Lab work

Conceptual Diagram



Concept Question 1

What can go wrong with this conversion?




```
CREATE TABLE Customer
(
  customer_id INTEGER,
  address VARCHAR(50),
  city VARCHAR(30),
  state CHAR(2),
  zip_code CHAR(5)
)
```

A. Table is missing a PRIMARY KEY constraint on `customer_id`

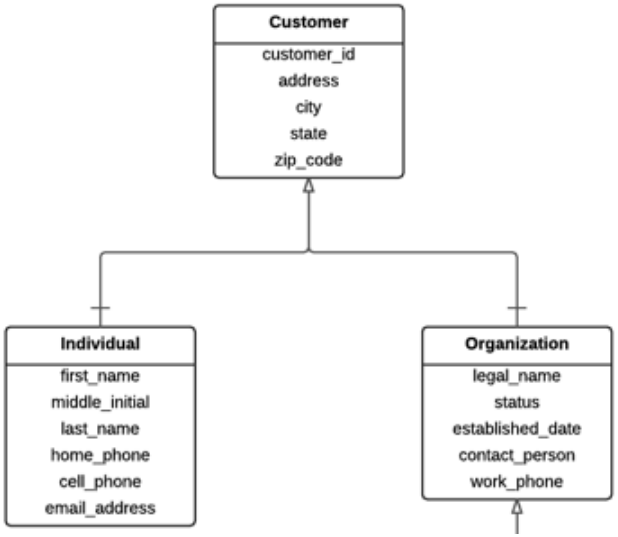
B. Table is missing a UNIQUE constraint `customer_id`

C. Table allows multiple records for the same customer

D. Table is missing a NOT NULL constraint on `customer_id`

Entity Classes to Relations

```
CREATE TABLE Customer
(
  customer_id INTEGER AUTO_INCREMENT PRIMARY KEY,
  address VARCHAR(50) NOT NULL,
  city VARCHAR(30) NOT NULL,
  state CHAR(2) NOT NULL,
  zip_code CHAR(5) NOT NULL
)
CREATE TABLE Individual
(
   customer_id INTEGER PRIMARY KEY,
   first_name VARCHAR(50) NOT NULL,
   last_name VARCHAR(50) NOT NULL,
   home_phone VARCHAR(15),
   cell_phone VARCHAR(15),
   email VARCHAR(50),
   FOREIGN KEY (customer_id) REFERENCES Customer(customer_id)
)
CREATE TABLE Organization
(
   customer_id INTEGER PRIMARY KEY,
   legal_name VARCHAR(50) NOT NULL,
   status CHAR(1),
   established_date DATE,
   contact_person VARCHAR(100),
   work_phone VARCHAR(15)),
   FOREIGN KEY (customer_id) REFERENCES Customer(customer_id)
)
```

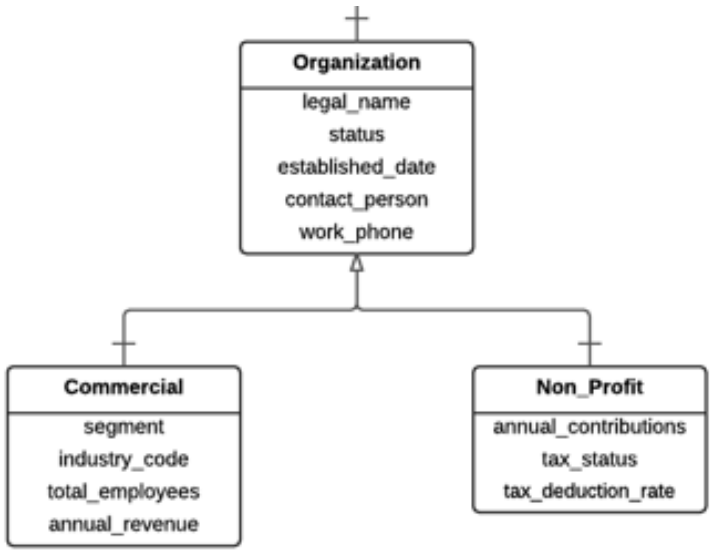


Concept Question 2

```
 CREATE TABLE Organization
(
   customer_id INTEGER PRIMARY KEY,
   legal_name VARCHAR(50) NOT NULL,
   status CHAR(1),
   established_date DATE,
   contact_person VARCHAR(100),
   work_phone VARCHAR(15)),
   FOREIGN KEY (customer_id) REFERENCES Customer(id)
)

CREATE TABLE Commercial
(
   customer_id INTEGER PRIMARY KEY,
   segment CHAR(1),
   industry_code CHAR(5),
   total_employees INT,
   annual_revenue DOUBLE,
   FOREIGN KEY (customer_id) REFERENCES X(customer_id)
)

CREATE TABLE Non_Profit
(
   customer_id INTEGER PRIMARY KEY,
   annual_contributions DOUBLE,
   tax_status CHAR(1),
   tax_deduction_rate DOUBLE,
   FOREIGN KEY (customer_id) REFERENCES X(customer_id)
)
```



What is x?

A. Customer
B. Organization
C. Commercial
D. Non_Profit
E. None

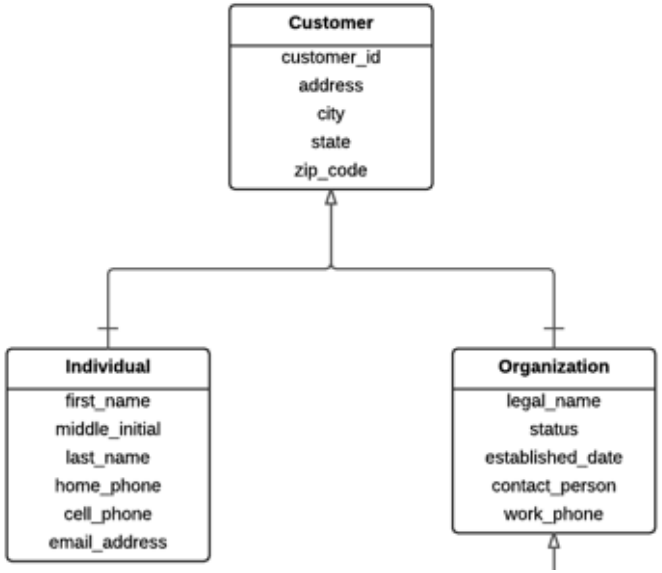
Concept Question 3

Suppose I want to perform certain kinds of look-up queries over the `Customer` schema: 1) look-up all the details of a customer by his/her last name; 2) look-up all the details of a customer by its legal name; 3) look-up all the details for customers who are based in 'Austin'. Which type of look-up query would likely have the longest run time?

```
CREATE TABLE Customer
(
  customer_id INT PRIMARY KEY,
  address VARCHAR(50) NOT NULL,
  city VARCHAR(30) NOT NULL,
  state CHAR(2) NOT NULL,
  zip CHAR(5) NOT NULL
)

CREATE TABLE Individual
(
   customer_id NUMBER(8) PRIMARY KEY,
   first_name VARCHAR(50) NOT NULL,
   last_name VARCHAR(50) NOT NULL,
   ...
   FOREIGN KEY (customer_id) REFERENCES Customer(customer_id)
)

CREATE TABLE Organization
(
   customer_id NUMBER(8) PRIMARY KEY,
   legal_name VARCHAR(50) NOT NULL,
   ...
   FOREIGN KEY (customer_id) REFERENCES Customer(customer_id)
)
```



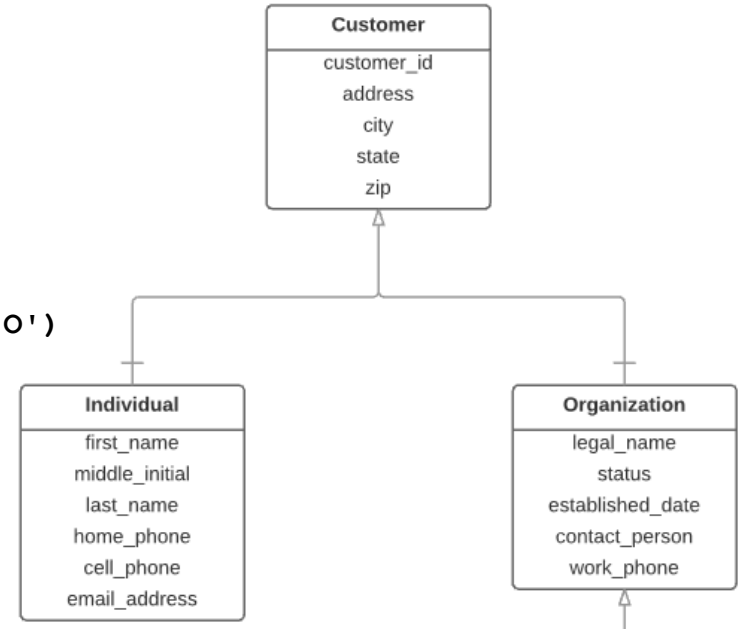
A. Query 2
B. Query 3
C. Query 1
D. Queries 1 and 2
E. All about the same run time

Solution: CO 3

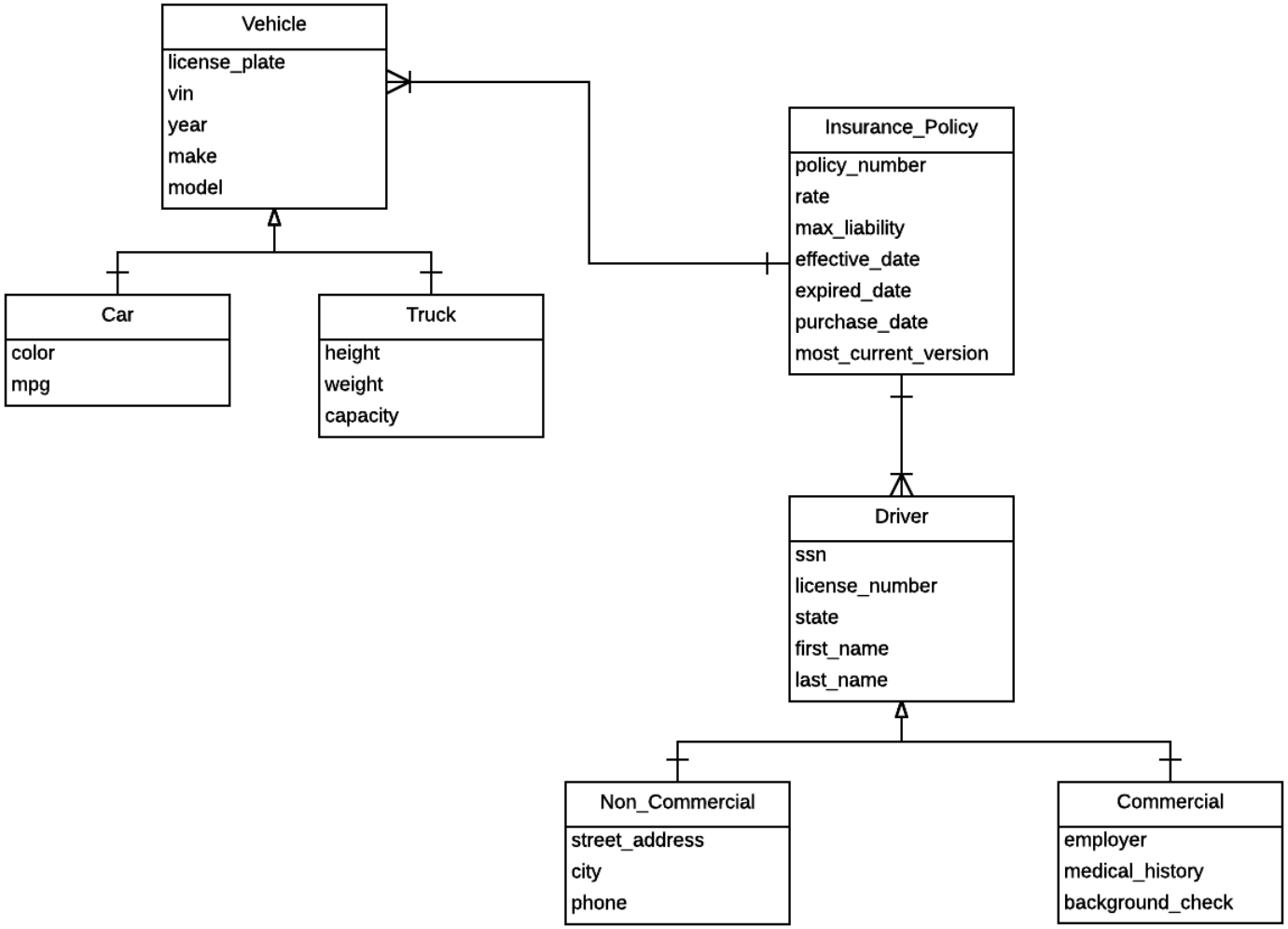
```
CREATE TABLE Customer
(
  customer_id INT PRIMARY KEY,
  address VARCHAR(50) NOT NULL,
  city VARCHAR(30) NOT NULL,
  state CHAR(2) NOT NULL,
  zip CHAR(5) NOT NULL,
  customer_type CHAR(1) CHECK customer_type IN ('I', 'O')
)

CREATE TABLE Individual
(
   customer_id NUMBER(8) PRIMARY KEY,
   first_name VARCHAR(50) NOT NULL,
   last_name VARCHAR(50) NOT NULL,
   home_phone VARCHAR(15),
   ...
   FOREIGN KEY (customer_id) REFERENCES Customer(customer_id)
)

CREATE TABLE Organization
(
   customer_id NUMBER(8) PRIMARY KEY,
   legal_name VARCHAR(50) NOT NULL,
   status CHAR(1),
   ...
   FOREIGN KEY (customer_id) REFERENCES Customer(customer_id)
)
```



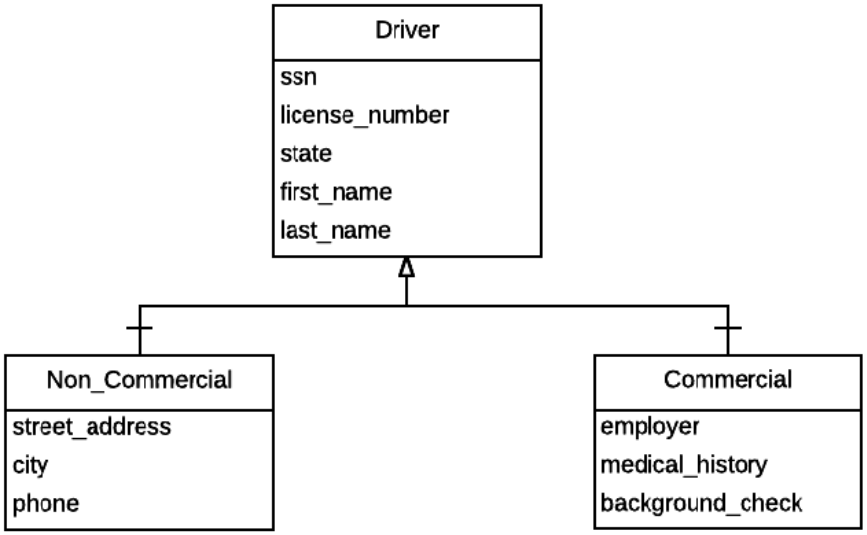
Conceptual Diagram



Concept Question 4

Suppose a `Driver` can be licensed to drive a non-commercial vehicle, a commercial vehicle or both types of vehicles. The conceptual diagram models these Driver types through subtyping and they are translated to the relations below. What can go wrong with this design?

```
CREATE TABLE Driver
(
  ssn INTEGER,
  license_number CHAR(8) NOT NULL,
  state CHAR(2) NOT NULL,
  ...
  driver_type CHAR(1),
  CHECK driver_type IN ('N', 'C'),
  PRIMARY KEY (ssn, driver_type)
)
CREATE TABLE Non_Commercial
(
   ssn INTEGER PRIMARY KEY,
   street_address VARCHAR(50) NOT NULL,
   city VARCHAR(50) NOT NULL,
   ...
   FOREIGN KEY (ssn) REFERENCES Driver(ssn)
)
CREATE TABLE Commercial
(
   ssn INTEGER PRIMARY KEY,
   employer VARCHAR(50),
   ...
   FOREIGN KEY (ssn) REFERENCES Driver(ssn)
)
```



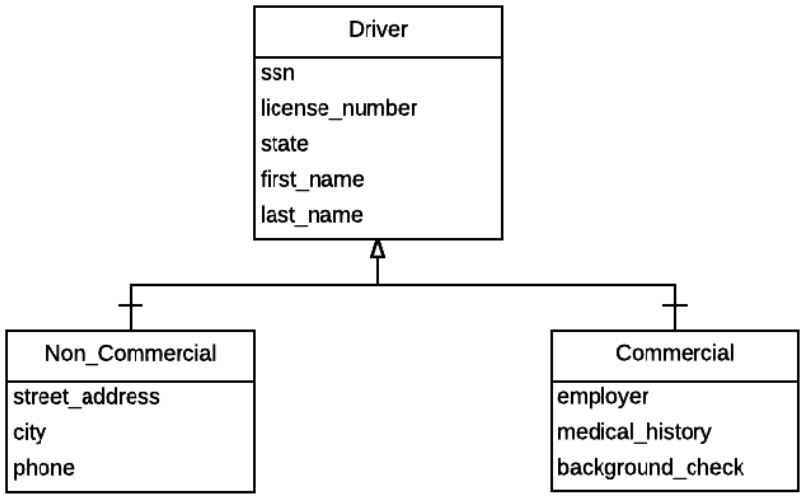
A. The `ssn` pks on the child tables
B. The `ssn` fks on the child tables
C. The composite pk (`ssn`, `driver_type`) on the parent table
D. The check on `driver_type`
E. All of the above

Solution: CQ 4

```
CREATE TABLE Driver
(
  ssn INTEGER,
  license_number CHAR(8) NOT NULL,
  state CHAR(2) NOT NULL,
  ...
  driver_type CHAR(2),
  CHECK driver_type IN ('N', 'C', 'NC'),
  PRIMARY KEY (ssn)
)

CREATE TABLE Non_Commercial
(
   ssn INTEGER PRIMARY KEY,
   street_address VARCHAR(50) NOT NULL,
   city VARCHAR(50) NOT NULL,
   ...
   FOREIGN KEY (ssn) REFERENCES Driver(ssn)
)

CREATE TABLE Commercial
(
   ssn INTEGER PRIMARY KEY,
   employer VARCHAR(50),
   ...
   FOREIGN KEY (ssn) REFERENCES Driver(ssn)
)
```



Quiz Question 1

What is an Update Anomaly?

A. Having to update redundant data across multiple records
B. Not being able to update a record due to a foreign key constraint
C. Being required to delete and insert a record, rather than updating it
D. Not being able to determine the primary key of a table

Quiz Question 2

Normalization is the process of decomposing the relations in a schema with the objective of reducing data redundancies.

A. True
B. False

Quiz Question 3

If the schema is in 2NF, then it must also be in:

A. 1NF

B. 3NF

C. A & B

D. Neither A nor B

Quiz Question 4

The primary key for a *Customer* table should always be the combination of (*customer_id, customer_name*).

A. True
B. False

Quiz Question 5

Which one cannot be used as a primary key for a `UT_Employees` table?

A. SSN

B. UT EID

C. Name

D. None of them above

Unnormalized to 1NF

Rule: A database schema is in 1NF *iff* all attributes have scalar values.

Student_Semester

EID	Semester	GPA	Courses	
alice1	Fall15	3.9	Stats	A
			DB	A
			Alg	A-
bob20	Fall15	3.7	DB	A
			Alg	B+
carol30	Fall15	3.5	Stats	A-
			Alg	B+

unnormalized

Student_Semester'

EID	Semester	Course	Grade	GPA
alice1	Fall15	Stats	A	3.9
alice1	Fall15	DB	A	3.9
alice1	Fall15	Alg	A-	3.9
bob20	Fall15	DB	A	3.7
bob20	Fall15	Alg	B	3.7
carol30	Fall15	Stats	3.5	3.5
carol30	Fall15	Alg	3.5	3.5

1NF

Unnormalized

A pharmaceutical company has a `Drug` table to record the inventory details and price fluctuations of each drug.

`Drug`

drug_nbr	**drug_name**	**drug_qty**	**drug_price**
48	Amoxicillin	500	01/01/13 – 03/31/15 – 0.30 04/01/15 – 01/15/16 – 3.00 01/16/16 – 3.50
50	Lipitor	150	10/01/12 – 03/31/14 – 0.75 04/01/14 – 1.00
72	Singulair	250	01/01/15 – 05/31/15 – 0.20 06/01/15 – 07/31/15 – 0.80 08/01/15 – 2.00

Concept Question 5

The pharma company decides to normalize the `Drug` table. Is the resulting table in 1NF?

drug_nbr	drug_name	drug_qty	drug_price
48	Amoxicillin	500	01/01/13 | 03/31/15 | 0.30 04/01/15 | 01/15/16 | 3.00 01/16/16 | | 3.50
…	…	…	…

UNF

drug_nbr	drug_name	drug_qty	start_date	end_date	drug_price
48	Amoxicillin	500	01/01/13	03/31/15	0.30
48	Amoxicillin	500	04/01/15	01/15/16	3.00
48	Amoxicillin	500	01/16/16		3.50
…	…	…	…	…	…

1NF ?

A. Yes B. No C. Don't know

Functional Dependencies

Definition:

If two records agree on the attributes

$A_1, A_2, \ldots, A_n$

then they must also agree on the attributes

$B_1, B_2, \ldots, B_n$

Formally:

$A_1, A_2, \ldots, A_n \rightarrow B_1, B_2, \ldots, B_n$

FD Example

Which FDs **hold** and **do not hold** on this table?

ID	Name	Phone	City
C0012	Smith	5555	Austin
C3412	Wallace	9876	Houston
C1111	Smith	9876	Dallas
C2323	Johnston	5555	Austin

ID → Name, Phone, City
City → Phone

Not Phone → City
Not Name → Phone

Concept Question 6

Can you find all the FDs that **do not hold** for the *Drug* table?

drug_nbr	drug_name	drug_qty	start_date	end_date	drug_price
48	Amoxicillin	500	01/01/13	03/31/15	0.30
48	Amoxicillin	500	04/01/15	01/15/16	3.00
48	Amoxicillin	500	01/16/16		3.50
50	Lipitor	150	10/01/12	03/31/14	0.75
50	Lipitor	150	04/01/14		1.00
72	Singulair	250	01/01/15	05/31/15	0.20
72	Singulair	250	06/01/15	07/31/15	0.80
72	Singulair	250	08/01/15		0.20

Which ones not FDs?

1. drug_nbr → drug_name
2. drug_nbr → drug_qty
3. drug_nbr → drug_price
4. drug_nbr → start_date
5. drug_nbr → end_date
6. drug_nbr → drug_price

A. 1, 2, 3, 4
B. 1, 2, 3, 4, 5
C. 4, 5, 6
D. 1, 2, 3, 6
E. None are FDs

Homework

- Read chapter 2 from our Learning SQL textbook
- Follow instructions from chapter and install MySQL Server on laptop
- Bring laptop to the next class and sit with your lab partner