





П

X∄

**= 2 = 0 =** 

**▼** 

1026 May

(A)

c002

a05

p03

800

744

10:56 PM

## **Key Distinctions**

The primary key declares the unique column in a table. It can be declared after the data type of the attribute, or at the end of all the attributes. It cannot have a NULL value. UNIQUE is synonymous and can be used interchangeably with primary key. An example would be cid in the Customers table.

A super key is a column or set of columns that ensures that every row will be unique. An example would be name, city and priceUSD in the Products table.

A candidate key is a superkey with the minimal number of columns. An example would be aid in the Agents table.

A foreign key is a value in a one table that must match the primary key of another table. An example is the foreign key pid in the Orders table that matches the primary key pid in the Products table.

## **Data Types**

All attributes must have a data type. According to our text book, there are six primitive data types which I have listed below. According to the documentation of PostgreSQL, there are many, many more.

- 1. Character strings
  - a. CHAR(n) character strings that are fixed length up to n characters. Short strings are padded to make n characters.
  - b. VARCHAR(n) character strings that imply a string length is used.
- 2. Bit strings
  - a. BIT(n) bits of strings of length n.
  - b. BIT VARYING(n) bits of strings of length up to n.
- 3. BOOLEAN denotes a logic with the possible values being: TRUE, FALSE, or UNKNOWN.
- 4. Integer values
  - a. INT or INTEGER typical integer values.
  - b. SHORTINT the number of bits permitted depends on the implementation.
- 5. Floating-point numbers
  - a. FLOAT or REAL typical floating-point numbers.
  - b. DOUBLE PRECISION used to obtain a higher precision number.
  - c. DECIMAL(n,d) or NUMERIC(n,d) values consist of n decimal digits, with decimal point d positions from the right.
- 6. Dates and Times character strings in a special form.
  - a. DATE for example 'YYYY-MM-DD'.
  - b. TIME for example 'HH:MM:SS'.

## **DVDs That I Own Table**

text	not NULL,		
text,			
INT	not NULL,		
text,			
text,			
INT	not NULL,		
text,			
DATE,			
PRIMARY KEY (title, year);			
	text, INT text, text, INT text, DATE,		

## **Relational Rules**

The <u>"first normal form" rule</u> (1NF) is that all columns must be atomic or indivisible. Each row must be unique, as duplications can cause problems. For every row-by-column position there can be only one value, no structure. If there were a list of values, there would be no easy way to manipulate them and the data would be much harder to retrieve.

This example would make it very difficult o find out how many hammers were sold:

<u>orderId</u>	customerId	<u>items</u>
1	4	5 hammer, 3 wrench
2	23	1 hammer
3	15	2 wrench, 1 key

The <u>"access rows by content only" rule.</u> If you try to point at a row-by-column position by using column1, row A for example, you most likely won't get the data you were looking for. For one thing, you will only get the information that is in column1, row A. In the above example, the result would be 1 if you were querying the snapshot. You must ask for what is there not where it is, because the data may not be in the same position anymore in the database as it is in the snapshot.

The <u>"all rows must be unique" rule</u>. In theory, you can't have two rows with all the same data in it. This ensures that duplication and inconsistency are avoided. This is the rule that is most often violated. If you have for example:

<u>customerID</u>	name	City
1	Bob Smith	Poughkeepsie
2	Bob Smith	Hyde Park

This could be two different Bob Smith's or he could have moved and someone added another record for him. How would you know without a lot more data? How would you know which one was correct or if they both are?