SQL: Concepts

Todd Bacastow
IST 210: Organization of Data
Design questions

- How many entities are there?
- What are the major entities?
- What are the attributes of each entity?
- Is there a unique identifier for one or more of these entities?
- How are the entities related to each other?
- What queries and/or reports will be produced with the data?
Design Process

- The entity becomes a table
- The entity name becomes the table name
- Each attribute becomes a column
- The instance identifier becomes the primary key
- For the one to many relation, an additional column is added to the table containing the foreign key
SQL

- SQL = Structured Query Language
  - Data Definition Language (DDL)
  - Data Manipulation Language (DML)

- Standards
  - Current - SQL2
  - Emerging - SQL3
  - Most DBMS implementations have a variation

- Syntax = How the language is used
DDL commands

- Create the structure for database objects - the database itself, tables and views
- Create, Alter, Truncate, Drop, and Delete
- Insert and change data are Insert and Delete
DDL Commands

- CREATE TABLE - creates a new database table
- ALTER TABLE - alters (changes) a database table
- DROP TABLE - deletes a database table
- CREATE INDEX - creates an index (search key)
- DROP INDEX - deletes an index
Data Types

- The most familiar data types are
  - Numeric
  - Text

- Other types of data are
  - time and date
  - logical
  - graphic representations of non-Roman character sets
  - binary large object (or BLOB)
“Typical” Syntax

- INT or INTEGER
- REAL or FLOAT
- CHAR(n)
  - Fixed-length character string
- VARCHAR(n)
  - Variable-length strings up to n characters
- DATE
  - In ‘yyyy-mm-dd’ format
- TIME
  - In ‘hh:mm:ss’ format
DEFINING A DB SCHEMA

CREATE TABLE NameOfTable
(Column Name data type null value,
Column Name data type null value,
Column Name data type null value);

**Semicolon after SQL Statements?**
Semicolon is the standard way to separate each SQL statement in database systems that allow more than one SQL statement to be executed in the same call to the server.
Example

CREATE TABLE Customer
(fName varchar (20) Not Null,
 lName varchar (25) Not Null,
 phoneNumber varchar (10),
 address varchar (40) Not Null,
 state varchar (20) Not Null,
 zipCode integer Primary Key);
Another Example

CREATE TABLE tablename
(list of elements)

- Principal elements are attributes and their types, but key declarations and constraints also appear

DROP TABLE tablename

- Delete the created table

CREATE TABLE Sells
(bar CHAR(20),
beer VARCHAR(20),
price REAL);

DROP TABLE Sells;
NULL

In place of a value for an attribute

- Interpretation is not exactly “missing value”
- There could be many reasons why no value is present, e.g., “value inappropriate”
Keys

- A primary key is a column or set of columns that uniquely identifies the rest of the data in any given row.

- A foreign key is a column in a table where that column is a primary key of another table.
  - Means that any data in a foreign key column must have corresponding data in the other table where that column is the primary key.

- In DBMS-speak, this correspondence is known as referential integrity.
Declaring Keys

PRIMARY KEY

- SQL2 does not allow nulls (empty values) in primary key
- Two places to declare:
  - After an attribute’s type, if the attribute is a key by itself
  - As a separate element, when key is > 1 attribute

CREATE TABLE Sells
(bar CHAR(20),
beer VARCHAR(20),
price REAL,
PRIMARY KEY (bar,beer));
Other Properties of Attributes

- **NOT NULL**
  - every row (tuple) must have a real value for this attribute

- **DEFAULT value**
  - A value to use whenever no other value of this attribute is known
Populating a Database

- INSERT statement can be used to add just a single row of data or multiple rows of data into the database.

- The INSERT query is constructed as follows:

  `INSERT INTO nameOfTable (column names) VALUES (data to be added to each of the above column names)`
Example

CREATE TABLE Customers
(name CHAR(30) PRIMARY KEY,
addr CHAR(50) DEFAULT '120 Oxford St',
phone CHAR(16));

INSERT INTO Customers(name) VALUES('Clinton');

<table>
<thead>
<tr>
<th>name</th>
<th>addr</th>
<th>phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton</td>
<td>120 Oxford St</td>
<td>NULL</td>
</tr>
</tbody>
</table>
Notes

- Primary key is by default NOT NULL
- In INSERT, it is OK to list a subset of the attributes and values for only this subset
- But if we had declared “phone CHAR(16) NOT NULL” then the insertion could not be made
Altering a Table

- The Alter Table command allows the database designer to make changes to a table, to add columns as well as add or drop constraints such as primary and foreign keys.

ALTER TABLE ADDRESS ADD (Zip_plus4 VARCHAR2(4))
Changing Columns

ALTER TABLE $R$ ADD column declaration;

- Add an attribute of relation $R$
- Example:

  ALTER TABLE Customers ADD education CHAR(16) DEFAULT 'high school';

ALTER TABLE $R$ DROP column;

- Drop columns of a table
Views

- **VIEW** = an expression that describes a table without creating it = a virtual table
- **TABLE** = a physical table on the hard drive
- View definition form is:
  ```sql
  CREATE VIEW name AS query;
  ```
- Views can be queried as if they existed physically
- Views may be modified, in some cases
Example

- The view VIPCustomers is the set of customers that have at least $10,000
  Customers (name, balance);

  CREATE VIEW VIPCustomers AS
  SELECT * FROM Customers
  WHERE balance >= 10000;

  SELECT * FROM VIPCustomer
  WHERE name LIKE '% Clinton';

Note: A "%" sign can be used to define wildcards (missing letters in the pattern) both before and after the pattern.
Indexing

- Indexing improves the storage and retrieval of database attributes and should be considered for any other fields that are likely to experience heavy use.
- Primary keys form the basis for indexes on the databases and the DBMS will automatically create indexes for fields identified as primary keys.
- Indexes also may be built for fields that are not key values or for fields whose values are not expected to occur uniquely on the database.
  - These are called secondary keys and are explicitly built.