Lecture SQL01
Introduction to SQL

Unless otherwise noted, lecture notes are derived from
Visual Quickstart Guide: SQL, Third Edition, by Chris Fehily

UAH
CPE 353

Why Databases?

- Data is everywhere
  - Library catalogs and holdings
  - Financial data
  - Geo-tagged information
  - Surveillance camera footage from highways and buildings
- Computers have facilitated the collection and archiving of large quantities of data
- Data is useless without efficient, reliable access

What is a Database? - 1

- ***Data Model***
  - Conceptual framework for describing the data, operations on the data, and consistency constraints
- ***Relational Database***
  - Data stored in a collection of one or more tables
- ***Database Management System (DBMS)***
  - Program(s) that govern creation, use, and administration of databases

What is a Database? - 2

- Why does one need a DBMS?
  - Efficient access
  - Manage shared access
  - Facilitate concurrency
  - Provide reliability in case of HW/SW failure
  - Ensure privacy of stored data
- ***Database System***
  - Database with a DBMS

What is SQL - 1

- ***SQL***
  - Standardized language that allows one to create, modify, and extract information from a ***relational database***
- SQL is standardized but vendors often add custom extensions which are not compatible with DBMS systems from other vendors
  - We will focus on ***standard commands***

More on SQL - 2

- **SQL** is a ***declarative language***
  - In a procedural language, there are constructs for flow of control
  - In a declarative language, the desired objective is described and the DBMS is responsible for determining how to achieve the objective

More on SQL - 3

- Three types of SQL statements
  - Data manipulation language (DML)
  - Data definition language (DLL)
  - Data control language (DCL)
- ***SQL is NOT case sensitive***
  - By convention, use ALL CAPS for SQL reserved words

Relational Model - 1

- Proposed in 1970 by E. F. Codd (IBM) as a model for large shared databases
  - See

    “A Relational Model of Data for Large Shared Data Banks”,

    Communications of the ACM, vol. 13, no. 6,

    June 1970, pp. 377-387.

Relational Model - 2

- Set theoretic approach
- A ***domain*** is a set of values
- A ***relation*** is a subset of the Cartesian product of two or more domains

Relational Model - 3

- Suppose domain **D_1 = {1, 2, 3}** and domain **D_2 = {red, blue}**
- The Cartesian product

$$D_1 \text{ X } D_2 = \{ \ (1, \text{red}), (1, \text{blue}),$$
$$(2, \text{red}), (2, \text{blue}),$$
$$(3, \text{red}), (3, \text{blue}) \ \}$$

Relational Model - 4

- Relation ==> table
- Tuple ==> table row
- Component/Attribute ==> table column
- A table with no rows is an ***empty table***
- Example: Suppose

  D1 = States

  D2 = Area

  D3 = Population

Relational Model - 5

State	Area	Population
AL	52,419	4,661,900
CA	163,696	36,756,666
GA	59,425	9,685,744
NC	53,819	9,222,414

Relation

Relational Model - 6

State	Area	Population
AL	52,419	4,661,900
CA	163,696	36,756,666
GA	59,425	9,685,744
NC	53,819	9,222,414

Tuple

Relational Model - 7

State	Area	Population
AL	52,419	4,661,900
CA	163,696	36,756,666
GA	59,425	9,685,744
NC	53,819	9,222,414

Attribute

Relational Model - 8

- Additional Terminology
  - # of columns = **Arity** or **Degree**
  - # of rows = **Cardinality**
  - ***Schema*** – description of the relation
    - Relation name
    - Column name
    - Column datatype (atomic data types, no structs)

Relational Model - 9

Arity or **Degree = 3**

Cardinality = 4

State	Area	Population
AL	52,419	4,661,900
CA	163,696	36,756,666
GA	59,425	9,685,744
NC	53,819	9,222,414

Relational Model - 10

- Observations
  - *Order of the rows makes no difference*
    - AL has a given area and population regardless of the row in which it appears within the relation
    - No concept of a row index as with an array
  - *Rows of the table must be different from one another for the table to represent a relation*
    - A set contains no duplicates

Relational Model - 11

- Rows are unnamed and unordered but are uniquely identified by a ***primary key***
- An attribute is a ***key*** for a relation if
  - No two rows have the same values in all key attributes
  - The above is not true for any proper subset of the key
    - Exception: ***Superkey***

Relational Model - 12

- Observations about ***primary key***
  - Every relation has exactly one primary key
  - No two rows may have the exact same primary key
  - Cannot be NULL
  - May consist of one or more columns but only those needed to be unique
  - Does not change with time and is not reused

Relational Model - 12

- Observations about ***primary key***
  - Every relation has exactly one primary key
  - No two rows may have the exact same primary key
  - Cannot be NULL
  - May consist of one or more columns but only those needed to be unique
  - Does not change with time and is not reused

Relational Model - 13

What is the primary key for the following relation?

State	Area	Population
AL	52,419	4,661,900
CA	163,696	36,756,666
GA	59,425	9,685,744
NC	53,819	9,222,414

Relational Model - 14

- ***Artificial keys*** are often introduced
  - A# in Banner system
  - rowid
- May be several different candidate keys
- Only one primary key

Relational Model - 15

- Data may be stored across multiple relations
- ***Foreign Keys***
  - Set of attributes that reference values in another relation
  - Restricted to parent key values
  - Can be NULL
  - To maintain ***Referential Integrity***, if one relation is modified, the other may have to be modified as well. If not, you may have a ***dangling reference***

Relational Model - 16

Primary Key

State	Area	Population
AL	52,419	4,661,900
CA	163,696	36,756,666
GA	59,425	9,685,744
NC	53,819	9,222,414

City	State	Population
Birmingham	AL	228,798
Atlanta	GA	537,958
Charlotte	NC	687,456
Greensboro	NC	250,642

Foreign Key

Relational Model - 17

- ***Integrity Constraints***
  - Conditions that must be true for any database relation
  - Specified when schema defined
  - Checked when relations are modified
  - Relation legal if satisfied
- DBMS only allows legal relations to be stored

Relational Model - 18

- ***Relationships***
  - ***One-to-One***
    - A tuple in relation X may have at most one matching tuple in relation Y

      and
    - A tuple in relation Y may have at most one matching tuple in relation X

Relational Model - 19

- ***Relationships - continued***
  - ***One-to-Many***
    - A tuple in relation X may have zero or more matching rows in relation Y

      and
    - Each tuple in relation Y has only one matching tuple in relation X

Relational Model - 20

- ***Relationships - continued***
  - ***Many-to-Many***
    - Each tuple in relation X may have zero or more matching rows in relation Y

      and
    - Each tuple in relation Y may have zero or more matching rows in relation X

Relational Model - 21

- ***Important Note:***
  - ***SQL does not completely adhere to the Relational Model previously described***