Introduction to SQL

SQL is used to communicate with a relational database and is an industry-wide standard, based on the rules developed in the 1970's by IBM's E.F. Codd. SQL is a declarative language that allows users to interact with the Database Management System (DBMS) to insert, update, and delete data. SQL includes commands to create database structures. SQL is a nonprocedural language; the user specifies what must be done, but not how it is to be done. SQL does not allow you to code logical statements such "IF" statements, loops, and other control structures.

SQL fits into two broad categories – data definition language (DDL) and data manipulation language (DML).

Data definition language – DDL is used to create and define database objects such as tables, indexes, and views. DDL also contains commands to define access rights to those database objects.

Data manipulation language – DML is used to manipulate data in the database and includes commands to insert, update, delete, and retrieve data within the database tables.

SQL is a relatively easy language to learn. It contains a basic command set of less than 100 words. SQL is a nonprocedural language; the user specifies what must be done, but not how it is to be done. SQL does not allow you to code logical statements such "IF" statements, loops, and other control structures.

The SQL language is sanctioned by the American National Standards Institute (ANSI). ANSI prescribes a standard SQL – the most recent version is is SQL-2003. The ANSI SQL standards are also accepted by the International Organization for Standardization (ISO). ANSI determines all aspects of the SQL language. Most RDBMS vendors add their own enhancements to SQL and provide extensions to the ANSI standard that provide additional SQL functionality.

Just as a side note, PL/SQL (PL – Procedural Language) is an Oracle proprietary programming language. PL/SQL is a procedural language that allows programmers to code a logical sequence of steps to make decisions in the application. A procedural language offers developers control structures such as conditional 'IF…THEN' statements and various types of loops to incorporate in applications. SQL can be embedded within PL/SQL blocks to interact with the database. PL/SQL allows developers to codify business rules through the creation of procedures, functions, triggers, etc.

Data Definition Commands

There are several data definition commands offered by SQL. These typically include CREATE, ALTER, and DROP commands. The CREATE commands build the objects in a database such as tables, sequences, indexes, views, etc. The ALTER commands modify database objects in some way such as adding a column to a table. The DROP commands permanently delete objects from the database.

Creating the Database

A DBMS is a collection of programs that manage a database structure. The database is a shared structure that houses a collection of metadata and end-user data, which are stored in database tables. Before you can use a new RDBMS, two tasks must be completed: first, create the database structure and second, create the tables that will hold the end-user data.

The RDBMS creates the physical files that will hold the database. The physical files are stored on the file system supported by the operating system. The database administrator can interact with the DBMS to create the physical files and customize these files.

The Database Schema

Authentication is the process through which the DBMS verifies that only registered users are able to access the database. To be authenticated, you must login to the RDBMS using a user ID and a password created by the database administrator. The user ID is associated with a database schema.

A schema is a group of database objects – such as tables, indexes, views, sequences, etc. – that are related to each other. Usually the schema belongs to one database user or application. A single database can hold multiple schemas belonging to different users and applications. A schema is a logical organization unit within the database. Object names only have to be unique within the schema. Objects can be referenced with the name of the schema followed by the name of the object.

Data Types

Each column defined in a database table will contain a certain data type. All values in that column must conform to the defined data type.

Below are some sample data types:

- Char(n) – stores alphanumeric data, with n representing the number of characters. The character stores n number of characters regardless if it is being used because it pads the data with blanks.

- Varchar2(l) – stores alphanumeric data, with l representing the length. Varchar2 only stores the number of characters needed, unlike the Char data type.
- Number (p,s) – stores numeric data. P stands for precision and s stands for scale. Precision is the total number of digits and scale is the number of digits to the right of the decimal point.
- Date – stores date and time values.

Creating Table Structures

To create a table in a database, use the CREATE TABLE SQL statement. Below is an example of the CREATE TABLE SQL statement:

CREATE TABLE PRODUCT (

P_CODE VARCHAR(10) NOT NULL UNIQUE,

P_DESCRIPT VARCHAR(35) NOT NULL,

V_CODE INTEGER,

PRIMARY KEY (P_CODE),

FOREIGN KEY (V_CODE) REFERENCES VENDOR

ON UPDATE (CASCADE);

The table PRODUCT is created with 3 columns – P_CODE, P_DESCRIPT, and V_CODE. The primary key is P_CODE. The column V_CODE is a foreign key to the VENDOR table.

A table is unique within a database schema.

SQL Constraints

Integrity constraints ensure that the data inserted into the table is to a certain standard. Constraints ensure the integrity and consistency of the data in a table. There are numerous constraints supported by most RDBMSs.

Constraint types:

NOT NULL – ensures a column does not accept null values.

UNIQUE – ensures all values in a column are unique.

DEFAULT – assigns a value to an attribute when a new row is added to a table.

CHECK – validates data when an attribute value is entered. For example, an employee's salary cannot be a negative number.

Common SQL Data Manipulation Commands

Below you will find several common SQL data manipulation commands.

INSERT – Adding table rows

COMMIT – Saving table changes

SELECT – Listing table rows

UPDATE – Updating table rows

ROLLBACK – Restoring table contents

DELETE – Deleting table rows

Adding Table Rows

The INSERT DML statement is used to populate rows into a database table. Below is an example of the INSERT statement:

INSERT INTO bb_ProductOptionCategory

(idOptionCategory, CategoryName)

VALUES (1, 'Size');

Saving Table Changes

It is important to understand when the manipulated data becomes permanent. DML statements are controlled within the context of a transaction. A transaction is DML statement or a group of statements that are logically grouped to do some type work. A COMMIT command will permanently save any changes made to any table in the database.

The ROLLBACK command will undo any DML statements back to the last COMMIT.

Listing Table Rows

The SELECT command is used to list the contents of a table. The syntax is the following:

SELECT columnlist

FROM tablename;

The columnlist represents one or more columns. An asterisk can be used as a wildcard character to list all attributes.

Updating Table Rows

An UPDATE statement is used to modify existing data in a table. The syntax is as follows:

UPDATE tablename

SET columnname = expression [, columnname = expression]

[WHERE conditionlist];

Restoring Table Contents

It is important to understand when the manipulated data becomes permanent. DML statements are controlled within the context of a transaction. A transaction is DML statement or a group of statements that are logically grouped to do some type work. The ROLLBACK command will undo any DML statements back to the last COMMIT.

The ROLLBACK command is only applicable if the COMMIT command has not been used to permanently store the changes in the database. COMMIT and ROLLBACK only work on data manipulation commands that add, modify, and delete table rows.

Deleting Table Rows

The DELETE command is used to remove or delete a table row. The DELETE statement is part of a transaction is not final until the database either issues a COMMIT or ROLLBACK. The syntax of the DELETE statement is below:

DELETE FROM tablename

[WHERE conditionlist];

The WHERE condition is optional. IF the WHERE clause is omitted, all rows in the table are deleted.

Selecting Rows with Conditional Restrictions

You can select partial table contents by placing restrictions on the rows to be included in the output. This is done by adding a WHERE clause to the SQL statement. The syntax is as follows:

SELECT columnlist

FROM tablelist

[WHERE conditionlist];

For example,

SELECT *

FROM PRODUCT

WHERE V_CODE = 21344;

Summary

Finally, SQL fits into two broad categories – data definition language (DDL) and data manipulation language (DML).

Data definition language – DDL is used to create and define database objects such as tables, indexes, and views. DDL also contains commands to define access rights to those database objects.

Data manipulation language – DML is used to manipulate data in the database and includes commands to insert, update, delete, and retrieve data within the database tables.