Hyperiondev

TASK

Introduction to SQL

Visit our website

Introduction

WELCOME TO THE INTRODUCTION TO SQL TASK!

In this task, you will be introduced to the database language SQL as well as the popular open-source relational database MySQL.

Get in touch

Connect for support

Remember that with our courses, you're not alone! You can contact an expert code reviewer to get support on any aspect of your course.

The best way to get help is to login to **www.hyperiondev.com/portal** to start a chat with a code reviewer. You can also schedule a call or get support via email.

Our expert code reviewers are happy to offer you support that is tailored to your individual career or education needs. Do not hesitate to ask a question or for additional support!

INTRODUCTION TO SQL

According to Encyclopaedia Britannica (2020), SQL (Structured Query Language) is a "language designed for eliciting information from databases". Furthemore, Coronel, Morris, and Rob (2011) expand on this definition by adding that SQL's commands allow a user to create tables of data, from which the data can be manipulated and sorted. We can, therefore, deduce that SQL would be the language that we use to turn data into information.

Like Python, SQL's vocabulary is quite similar to English, which makes it a fairly easy language to learn.

SQL statements fit into two general categories (Oracle, 2017):

1. **SQL as a data definition language (DDL):** includes commands to create, change and drop database objects (such as tables) as well as define access rights to those database objects. They also allow you to add comments to the data dictionary.

2. **SQL as a data manipulation language (DML):** includes commands to manipulate data in existing objects like insert, update, delete and retrieve data within the database tables.

The table below lists the SQL data definition commands (Coronel, et al., 2011)**:**

Command	Description
CREATE SCHEMA AUTHORIZATION	Creates a database schema
CREATE TABLE	Creates a new table in the user's database schema
NOT NULL	Ensures that a column will not have null values
UNIQUE	Ensures that a column will not have duplicate values
PRIMARY KEY	Defines a primary key for a table
FOREIGN KEY	Defines a foreign key for a table
DEFAULT	Defines a default value for a column when no value is given
CHECK	Used to validate data in an attribute

CREATE INDEX	Creates an index for the table
CREATE VIEW	Creates a dynamic subset of rows or columns from one or more tables
ALTER TABLE	Modifies a table (adds, modifies or deletes attributes or constraints)
CREATE TABLE AS	Creates a new table based on a query in the user's database schema
DROP TABLE	Permanently deletes a table
DROP INDEX	Permanently deletes an index
DROP VIEW	Permanently deletes a view

SQL Data Definition Commands (Coronel, et al., 2011, p. 221)

The table below lists the SQL data manipulation commands (Coronel, et al., 2011):

Command	Description
INSERT	Inserts rows into a table
SELECT	Select attributes from rows in one or more tables or views
WHERE	Restricts the selection of rows based on a conditional expression
GROUP BY	Groups the selected rows based on one or more attributes
HAVING	Restricts the selection of grouped rows based on a condition
ORDER BY	Orders the selected rows based on one or more attributes
UPDATE	Modifies an attribute's values in one or more tables rows
DELETE	Deletes one or more rows from a table
COMMIT	Permanently saves data changes
ROLLBACK	Restores data to their original values
Comparison Operators	
=, <, >, <=, >=, <>	Used in conditional expressions
Logical Operators	

AND, OR, NOT	Used in conditional expressions
Special Operators	Used in conditional expressions
BETWEEN	Checks whether an attribute value is within a range
IS NULL	Checks whether an attribute value is null
LIKE	Checks whether an attribute value matches a given string pattern
IN	Checks whether an attribute value matches any value within a value list
EXISTS	Checks whether a subquery returns any rows
DISTINCT	Limits values to unique values
Aggregate Functions	Used with SELECT to return mathematical summaries on columns
COUNT	Returns the number of rows with non-null values for a given column
MIN	Returns the minimum attribute value found in a given column
MAX	Returns the maximum attribute value found in a given column
SUM	Returns the sum of all values for a given
AVG	Returns the average of all values for a given column

SQL Data Manipulation Commands (Coronel, et al., 2011, pp. 221-222)

CREATING TABLES

To create new tables in SQL you use the CREATE TABLE statement. As its arguments, it expects all the columns we want in the table, as well as their data types. The syntax of the CREATE TABLE statement is shown below:

```
CREATE TABLE table_name (
    column1 datatype constraint,
    column2 datatype constraint,
    column3 datatype constraint,
   ....
);
```

In the above, the constraints are optional and are used to specify rules for data in a table. Constraints that are commonly used in SQL are:

- NOT NULL: Ensures that a column cannot have a NULL value
- UNIQUE: Ensures that all values in a column are different
- PRIMARY KEY: Uniquely identifies each row in a table
- FOREIGN KEY: Uniquely identifies a row in another table
- CHECK: Ensures that all values in a column satisfy a specific condition
- DEFAULT: Sets a default value for a column when no value is specified
- INDEX: Used to create and retrieve data from the database very quickly

To create a table called Employee, for example, that contains five columns — EmployeeID, LastName, FirstName, Address, and PhoneNumber — you would do the following:

```
CREATE TABLE Employee (
    EmployeeID int,
    LastName varchar(255),
    FirstName varchar(255),
    Address varchar(255),
    PhoneNumber varchar(255)
);
```

The EmployeeID column is of type *int* and will, therefore, hold an integer value. The LastName, FirstName, Address, and PhoneNumber columns are of type *varchar* and will, therefore, hold characters. The number in brackets indicates the maximum number of characters, which in this case is 255. Note that these are all

within brackets with a semicolon at the end. This is what indicates that the line is concluded.

The CREATE TABLE statement above will create an empty Employee table that will look like this:

EmployeeID	LastName	FirstName	Address	PhoneNumber

When creating tables, it's advisable to add a primary key to one of the columns as this will help keep entries unique and will speed up select queries. Primary keys must contain unique values, and cannot contain null values. A table can only contain one primary key, however, the primary key may consist of single or multiple columns (as you may remember from the previous task).

You can add a primary key to when creating the Employee table as follows:

```
CREATE TABLE Employee (
    EmployeeID int NOT NULL,
    LastName varchar(255) NOT NULL,
    FirstName varchar(255),
    Address varchar(255),
    PhoneNumber varchar(255),
    PRIMARY KEY (EmployeeID)
);
```

To name a primary key constraint, and define a primary key constraint on multiple columns, you use the following SQL syntax:

```
CREATE TABLE Employee (
    EmployeeID int NOT NULL,
    LastName varchar(255) NOT NULL,
    FirstName varchar(255),
    Address varchar(255),
    PhoneNumber varchar(255),
    CONSTRAINT PK_Employee PRIMARY KEY (ID,LastName)
);
```

In the example above there is only one primary key named PK_Employee. However, the value of the primary key is made up of two columns: ID and LastName.

Inserting Rows

The table that we have just created is empty and needs to be populated with rows or records. We can add entries to a table using the INSERT INTO command. There are two ways to write the INSERT INTO command:

1. Do not specify the column names where you intend to insert the data. This can be done if you are adding values for all of the columns of the table. However, you should ensure that the order of the values is in the same order as the columns in the table. The syntax will be as follows:

```
INSERT INTO table_name
VALUES (value1, value2, value3, ...);
```

   So, for example, to add an entry to the Employee table you would do the following:

```
INSERT INTO Employee
VALUES (1234, Smith, John, 25 Oak Rd, 0837856767);
```

2. The other way is to specify both the column names and the values to be inserted. The syntax will be as follows:

```
INSERT INTO table_name (column1, column2, column3, ...)
VALUES (value1, value2, value3, ...);
```

   Therefore, to add an entry to the Employee table using this way, you would do the following:

```
INSERT INTO Employee (EmployeeID, LastName, FirstName, Address,
PhoneNumber)
VALUES (1234, Smith, John, 25 Oak Rd, 0837856767);
```

SELECT

The SELECT statement is used to fetch data from a database. The data returned is stored in a result table, known as the result-set. The syntax of a SELECT statement is as follows:

```
SELECT column1, column2, ...
FROM table_name;
```

column1, column2, ... are the column names of the table you want to select data from. The following example selects the FirstName and LastName columns from the Employee table:

```
SELECT FirstName, LastName
FROM Employee;
```

If you want to select all the columns in the table, however, you use the following syntax:

```
SELECT * FROM table_name;
```

The asterisk (*) means that we want to fetch all of the columns.

You can also use the ORDER BY command to sort the results in ascending or descending order. The ORDER BY command sorts the records in ascending order by default. You need to use the DESC keyword to sort the records in descending order.

The ORDER BY syntax is as follows:

```
SELECT column1, column2, ...
FROM table_name
ORDER BY column1, column2, ... ASC/DESC;
```

The example below selects all Employees in the Employee table and sorts them in descending order by the FirstName column:

```
SELECT * FROM Employee
ORDER BY FirstName DESC;
```

WHERE

Similar to an *if-statement*, the WHERE clause allows us to filter data depending on a specific condition. The syntax of the WHERE clause is as follows:

```
SELECT column1, column2, ...
FROM table_name
WHERE condition;
```

The following SQL statement selects all the employees with the first name John, in the Employee table:

```
SELECT * FROM Employee
WHERE FirstName = 'John';
```

Note that SQL requires single quotes around strings, however, you do not need to enclose numeric fields in quotes.

You can use logical operators (AND, OR) and comparison operators (=,<,>,<=,>=,<>) to make WHERE conditions as specific as you like.

For example, suppose you have the following table which contains the most sold albums of all time:

Artist	Album	Released	Genre	sales_in_millions
Michael Jackson	Thriller	1982	Pop	70
AC/DC	Back in Black	1980	Rock	50
Pink Floyd	The Dark Side of the Moon	1973	Rock	45
Whitney Houston	The Bodyguard	1992	Soul	44

You can select those of them that are classified as rock and have sold under 50 million copies by simply using the AND operator as follows:

```
SELECT *
FROM albums
WHERE genre = 'rock' AND sales_in_millions <= 50
ORDER BY released
```

WHERE statements also support some commands, that allows you a quick way to check commonly used queries. These are:

- IN: compares the column to multiple possible values and returns true if it matches at least one
- BETWEEN: checks if a value is within an inclusive range
- LIKE: searches for a pattern

For example, if we want to select the pop and soul albums from the table above, we can use:

```
SELECT * FROM albums
WHERE genre IN ('pop', 'soul');
```

Or, if we want to get all the albums released between 1975 and 1985, we can use:

```
SELECT * FROM albums
WHERE released BETWEEN 1975 AND 1985;
```

FUNCTIONS

SQL has many functions that do all sorts of helpful stuff. Some of the most regularly used ones are:

- COUNT(): returns the number of rows
- SUM(): returns the total sum of a numeric column
- AVG(): returns the average of a set of values
- MIN() / MAX(): gets the minimum or maximum value from a column

For example, to get the most recent year in the Album table we can use:

```
SELECT MAX(released)
FROM albums;
```

JOINS

In complex databases, there are often several tables connected to each other in some way. A JOIN clause is used to combine rows from two or more tables, based on a column they both share.

Look at the two tables below:

VideoGame

ID	Name	DeveloperID	Genre
1	Super Mario Bros.	2	Platformer
2	World of Warcraft	1	MMORPG
3	The Legend of Zelda	2	Adventure

GameDeveloper

ID	Name	Country
1	Blizzard	USA
2	Nintendo	Japan

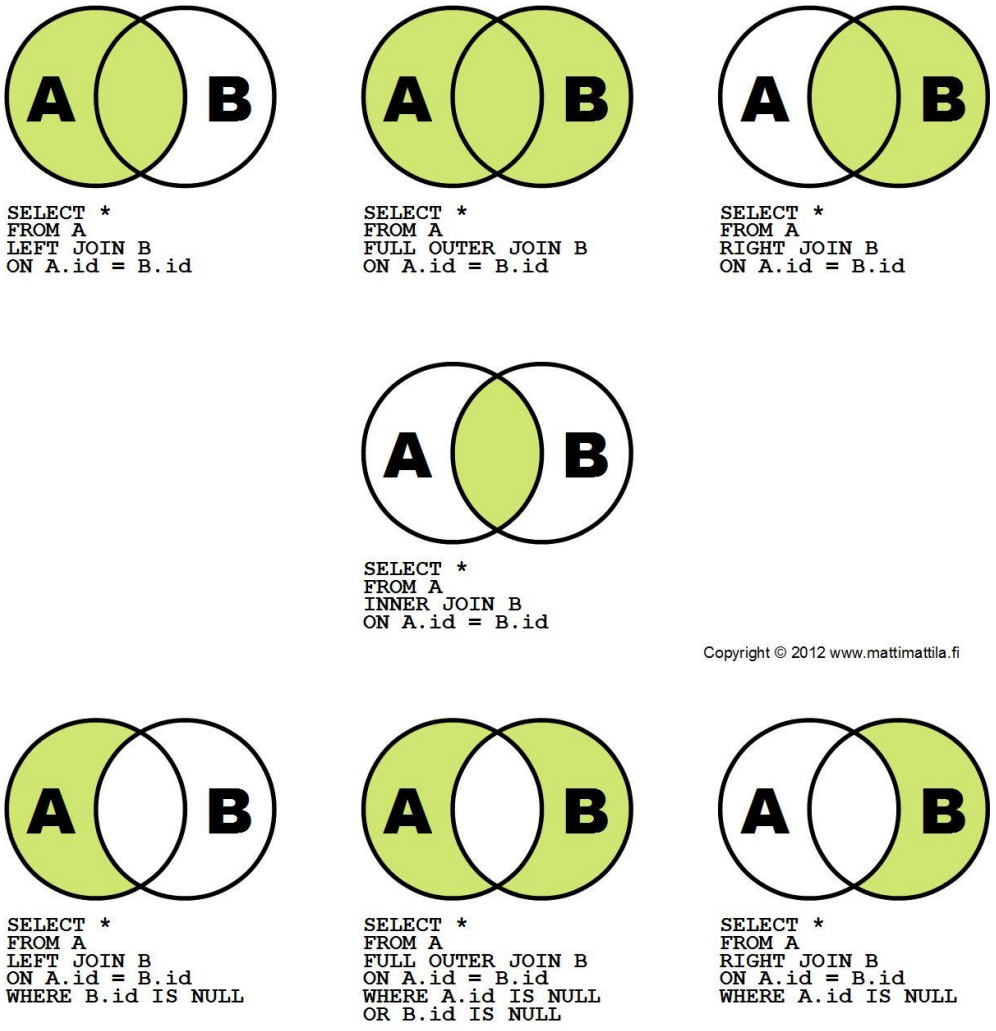
The VideoGame table contains information about various video games, while the Game Developer table contains information about the developers of the games. The VideoGame table has a DeveloperID column that holds the game developer ID and represents the ID of the respective developer from the GameDeveloper table. This logically links the two tables and allows us to use the information stored in both of them at the same time.

If we want to create a query that returns everything we need to know about the games, we can use INNER JOIN to acquire the columns from both tables.

```
SELECT VideoGame.Name, VideoGame.Genre, GameDeveloper.Name,
GameDeveloper.Country
FROM VideoGame
INNER JOIN GameDeveloper
ON VideoGame.DeveloperID = GameDeveloper.ID;
```

The INNER JOIN is the simplest and most common type of JOIN, however, there are many other different types of JOINs in SQL, namely:

- INNER JOIN: Returns records that have matching values in both tables
- LEFT JOIN: Returns all records from the left table, and the matched records from the right table
- RIGHT JOIN: Returns all records from the right table, and the matched records from the left table
- FULL JOIN: Returns all records when there is a match in either left or right table



Graphical representation of common SQL joins ***(Mattila, 2012)***

ALIASES

Notice that in the VideoGame and GameDeveloper tables each has a column called Name. This can become confusing. Aliases are used to give a table or column a temporary name. An alias only exists for the duration of the query and is often used to make column names more readable.

The alias column syntax is:

```
SELECT column_name AS alias_name
FROM table_name;
```

The alias table syntax is:

```
SELECT column_name(s)
FROM table_name AS alias_name;
```

The following SQL statement creates an alias, for the Name column from GameDeveloper table:

```
SELECT Name AS Developer
FROM GameDeveloper;
```

This SQL statement shortens the query drastically by setting aliases to the table names:

```
SELECT games.Name, games.Genre, devs.Name AS Developer, devs.Country
FROM VideoGame AS games
INNER JOIN GameDeveloper AS devs
ON games.DeveloperID = devs.ID;
```

UPDATE

The UPDATE statement is used to modify the existing rows in a table.

To use the UPDATE statement you:

- Choose the table where the row you want to change is located.
- Set the new value(s) for the wanted column(s).
- Select which of the rows you want to update using the WHERE statement. If you omit this, all rows in the table will change.

The syntax for the update statement is:

```
UPDATE table_name
```

```
SET column1 = value1, column2 = value2, ...
WHERE condition;
```

Take a look at the following Customer table:

CustomerID	CustomerName	Address	City
1	Maria Anderson	23 York St	New York
2	Jackson Peters	124 River Rd	Berlin
3	Thomas Hardy	455 Hanover Sq	London
4	Kelly Martins	55 Loop St	Cape Town

The following SQL statement updates the first customer (CustomerID = 1) with a new address and a new city:

```
UPDATE Customer
SET Address = '78 Oak St', City= 'Los Angeles'
WHERE CustomerID = 1;
```

DELETING ROWS

Deleting a row is a simple process. All that you need to do is select the right table and row you want to remove. The DELETE statement is used to delete existing rows in a table.

The DELETE statement syntax is as follows:

```
DELETE FROM table_name
WHERE condition;
```

The following statement deletes the customer Jackson Peters from the Customer table:

```
DELETE FROM Customer
WHERE CustomerName = 'Jackson Peters';
```

You can also delete all rows in a table without deleting the table:

```
DELETE * FROM table_name;
```

DELETING TABLES

The DROP TABLE statement is used to remove every trace of a table in a database. The syntax is as follows:

```
DROP TABLE table_name;
```

For example, if we what to delete the table Customer, we do the following:

```
DROP TABLE Customer;
```

If you want to delete the data inside a table, but not the table itself, you can use the TRUNCATE TABLE statement:

```
TRUNCATE TABLE table_name;
```

MYSQL

MySQL is a popular open-source relational database management system owned by MySQL AB. While it is written in C and C++, it is based on SQL. It also supports many different languages and operating systems. MySQL makes using SQL easy because It has an intuitive interface that allows you to add, remove, modify, search and join tables in a database. You will learn more about MySQL in the next task where we will be installing it and start creating and manipulating databases with it.

Compulsory Task

Answer the following questions:

- Go to the **w3schools website's SQL browser IDE**. This is where you can write and test your SQL code using their databases. Once you are happy with it, paste your code in a text file named **Student.txt** and save in your task folder.
- Write the SQL code to create a table called Student. The table structure is summarised in the table below (Note that STU_NUM is the primary key):

Attribute Name	Data Type
STU_NUM	CHAR(6)
STU_SNAME	VARCHAR(15)
STU_FNAME	VARCHAR(15)
STU_INITIAL	CHAR(1)
STU_STARTDATE	DATE
COURSE_CODE	CHAR(3)
PROJ_NUM	INT(2)

- After you have created the table in question 1, write the SQL code to enter the first two rows of the table as below:

STU_NUM	STU_SNAME	STU_FNAME	STU_INITIAL	STU_STARTDATE	COURSE_CODE	PROJ_NUM
01	Snow	John	E	05-Apr-14	201	6
02	Stark	Arya	C	12-Jul-17	305	11

- Assuming all the data in the Student table has been entered as shown below, write the SQL code that will list all attributes for a COURSE_CODE of 305.

STU_NUM	STU_SNAME	STU_FNAME	STU_INITIAL	STU_STARTDATE	COURSE_CODE	PROJ_NUM
01	Snow	Jon	E	05-Apr-14	201	6
02	Stark	Arya	C	12-Jul-17	305	11
03	Lannister	Jamie	C	05-Sep-12	101	2
04	Lannister	Cercei	J	05-Sep-12	101	2
05	Greyjoy	Theon	I	9-Dec-15	402	14

06	Tyrell	Margaery	Y	12-Jul-17	305	10
07	Baratheon	Tommen	R	13-Jun-19	201	5

- Write the SQL code to change the course code to 304 for the person whose student number is 07.
- Write the SQL code to delete the row of the person named Jamie Lannister, who started on 5 September 2012, whose course code is 101 and project number is 2. Use logical operators to include all of the information given in this problem.
- Write the SQL code that will change the PROJ_NUM to 14 for all those students who started before 1 January 2016 and whose course code is at least 201.
- Write the SQL code that will delete all of the data inside a table, but not the table itself.
- Write the SQL code that will delete the Student table entirely.

Completed the task(s)?

Ask an expert to review your work!

Review work

Rate us

Share your thoughts

HyperionDev strives to provide internationally-excellent course content that helps you achieve your learning outcomes.

Think that the content of this task, or this course as a whole, can be improved or think we've done a good job?

Click here to share your thoughts anonymously.

References:

Coronel, C., Morris, S., & Rob, P. (2011). *Database Systems* (9th ed., pp. 220-297). Boston: Cengage Learning.

Encyclopaedia Britannica. (2020). SQL | computer language. Retrieved 17 February 2020, from https://www.britannica.com/technology/SQL

Mattila, M 2012, Visual representation of common SQL joins, flickr.com, accessed 17 February 2020,< https://www.flickr.com/photos/mattimattila/8190148857>.

Oracle. (2017). Types of SQL Statements. Retrieved 17 February 2020, from https://docs.oracle.com/database/121/SQLRF/statements_1001.htm#SQLRF30042