SQL with Python

Sugarkhuu Radnaa

Py4Econ in Ulaanbaatar

py4econ@gmail.com

Week 4: Learning objectives

Get to know:

1. Basic concepts about SQL and structured database
2. PostgreSQL and PgAdmin
3. Using PostgreSQL in terminal
4. Working with SQL using Python

SQL

SQL is the most popular database language. There are many different DBMS (database management system)s including structured and NoSQL.

Popular structured/SQL databases:

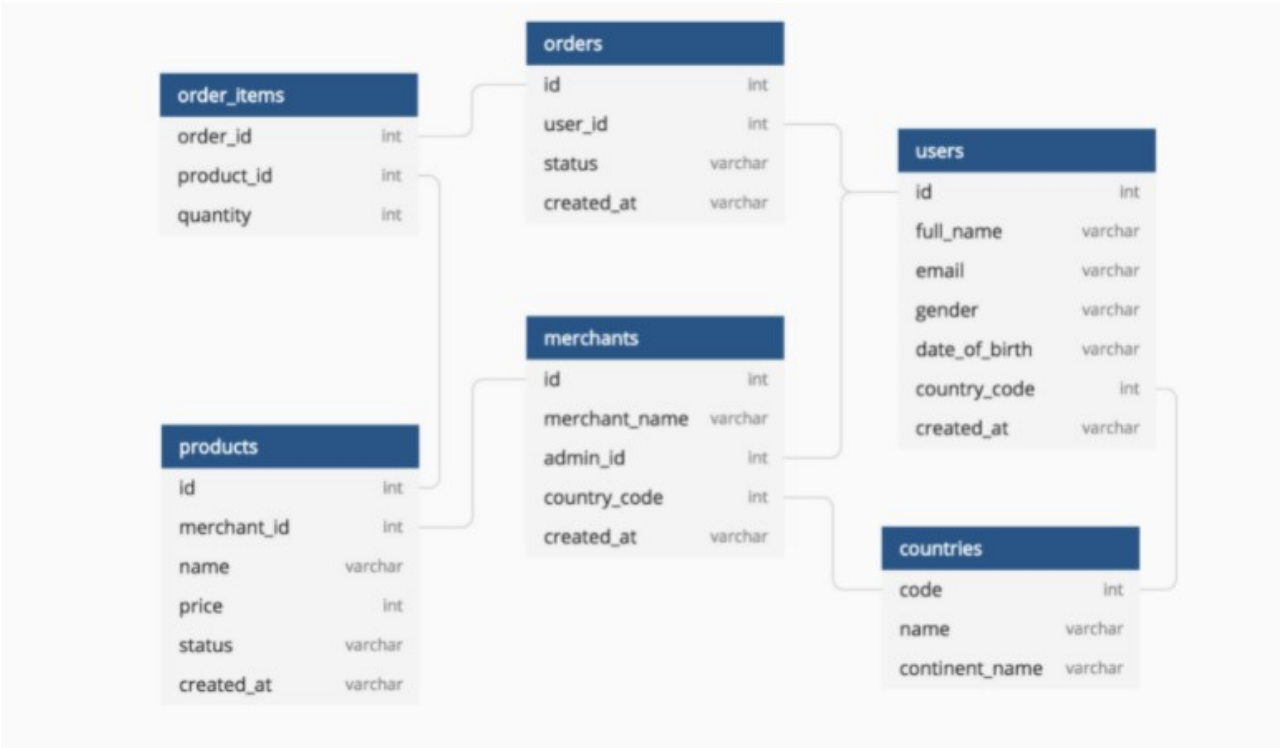
1. Oracle SQL
2. Microsoft SQL
3. MySQL
4. PostgreSQL
5. SQLite

Popular NoSQL databases:

1. MongoDB
2. ElasticSearch
3. Casandra

Sample structured database design

Data architecture is a crucial in any organization. When due implemented, it can become a game changing factor. Efficient ETL, Datawarehouse and data centers are necessary for an efficient DBMS.



SQL: W3SCHOOL examples

1. SELECT CustomerName, City FROM Customers;
2. SELECT * FROM Customers WHERE Country='Mexico';
3. SELECT * FROM Customers WHERE Country='Germany' AND (City='Berlin' OR City='Frankfurt');
4. UPDATE Customers SET ContactName = 'Alfred Schmidt', City= 'Frankfurt' WHERE CustomerID = 1;
5. SELECT * FROM Customers WHERE ContactName LIKE 'a%';
6. SELECT Customers.CustomerName, Orders.OrderID FROM Customers LEFT JOIN Orders ON Customers.CustomerID = Orders.CustomerID ORDER BY Customers.CustomerName;
7. SELECT COUNT(CustomerID), Country FROM Customers GROUP BY Country ORDER BY COUNT(CustomerID) DESC;

Some useful tips

1. Dump database to a file / Restore database from a file
2. SQLite is perhaps quickest and lightest DBMS
3. Run routine tasks in function/procedure

psql basic terminal commands

1. psql -U username - connect to a psql server with username
2. \l - list of databases
3. \c dbname - choose database
4. \dt - list of tables in the chosen database
5. *Command*; - Command to run (Note: a statement must be terminated with semicolon ;)
6. \q - quit psql
7. $pg_dump - Uusername - ddatabase_name - ttable_name > table_dump.sql$

Homework

1. Task 1

   - Push your result into your homework repository
   - Deadline: 1 week

Note: Commit your results step by step.

Task 1

1. Populate a PostgreSQL table with the data in “data.xlsx” (CREATE TABLE)
2. Select 'firstName' and 'lastName' of the first three rows ('LIMIT')
3. Select 'firstName' and 'age' of the last three rows ('ORDER BY, LIMIT')

Tip: Watch the following tutorial for PostgreSQL

- PostgreSQL

Thank you!