



MÃ NGUỒN MỞ TRONG KHOA HỌC DỮ LIỆU

Bài 03. CƠ SỞ DỮ LIỆU NHÚNG SQLite



1. Giới thiệu SQLite

- SQLite là một hệ quản trị cơ sở dữ liệu quan hệ nhúng (embedded relational database management system).
- Đây là một trong những hệ thống cơ sở dữ liệu phổ biến nhất trên thế giới nhờ tính đơn giản và hiệu quả của nó.
- SQLite được thiết kế để sử dụng bên trong các ứng dụng như một thư viện, không yêu cầu máy chủ và không cần cấu hình.

1. Giới thiệu SQLite

- SQLite là một hệ quản trị cơ sở dữ liệu mã nguồn mở (open source). Nó được phát triển và phân phối dưới Giấy phép công cộng không giới hạn (Public Domain License).

2. Một số đặc điểm chính của SQLite

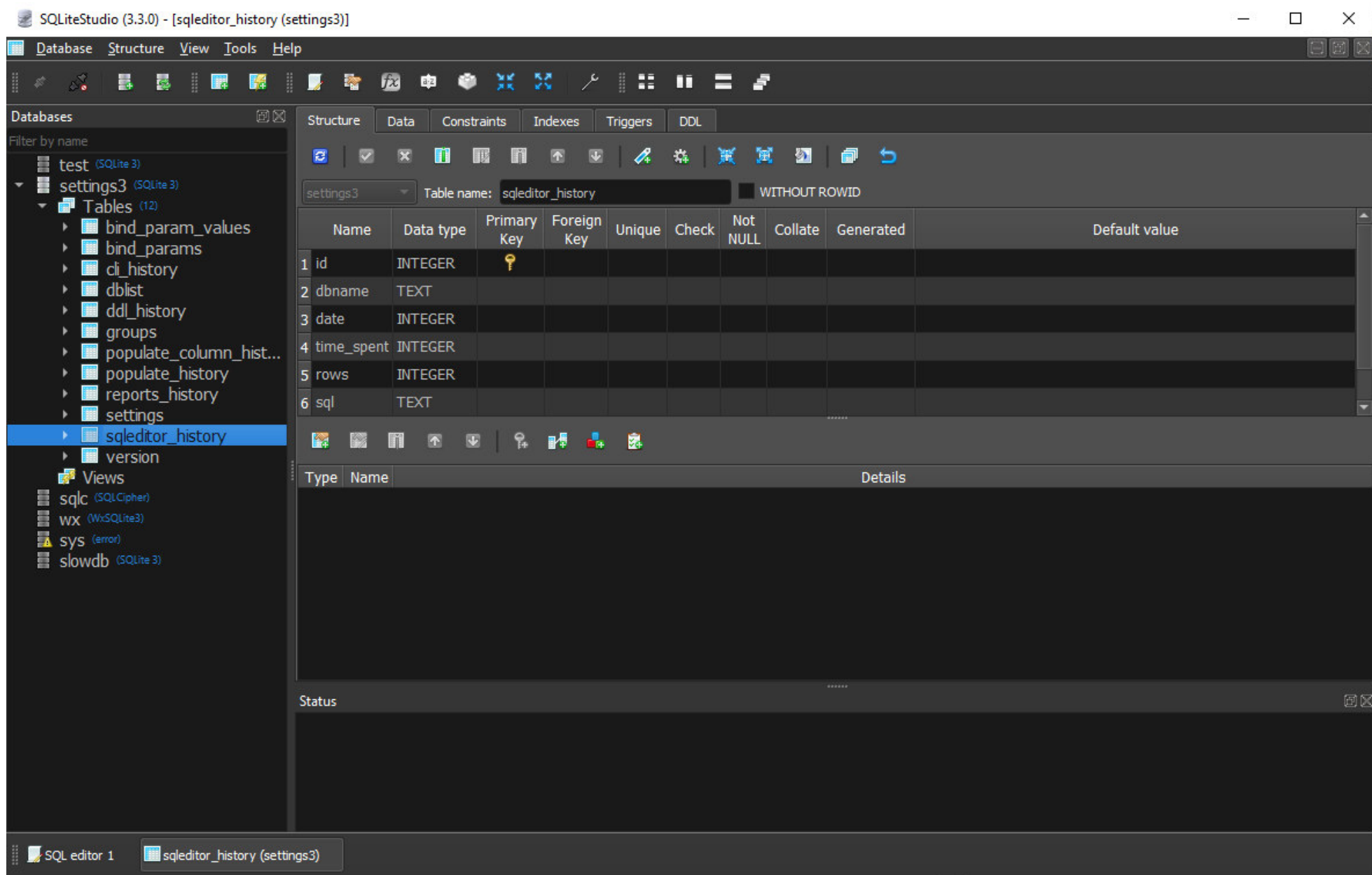
- **Cơ sở dữ liệu nhẹ:** SQLite không yêu cầu máy chủ chạy nền như MySQL hay PostgreSQL. Cơ sở dữ liệu được lưu trữ dưới dạng một file trong hệ thống tệp.
- **Không cần cấu hình:** SQLite rất dễ sử dụng, không cần cài đặt hoặc cấu hình phức tạp.
- **Tính di động cao:** File cơ sở dữ liệu SQLite có thể được sao chép từ hệ thống này sang hệ thống khác mà vẫn hoạt động tốt.
- **Tích hợp rộng rãi:** SQLite được tích hợp vào nhiều ngôn ngữ lập trình và ứng dụng, như Python, Android, iOS, v.v.

3. SQLiteStudio

- SQLiteStudio là một phần mềm mã nguồn mở được thiết kế để quản lý cơ sở dữ liệu SQLite.
- Đây là một công cụ mạnh mẽ và dễ sử dụng cho các lập trình viên, nhà phát triển và người dùng muốn làm việc với SQLite mà không cần phải sử dụng dòng lệnh.

3. SQLiteStudio

- freeware, portable (no installation needed)
- <https://sqlitestudio.pl>



Thực hành 1

1. Tải xuống SQLiteStudio từ <https://sqlitestudio.pl>.
2. Tạo cơ sở dữ liệu có tên là **ex02_mssv_hovaten**.
3. Tạo các bảng và chọn kiểu dữ liệu phù hợp cho các cột của chúng:
 - a. bảng **student** (student_id, name, surname, dateOfBirth, yearEnrolled)
 - b. bảng **course** (course_id, name, creditPoints, yearCommenced)
 - c. bảng **staff** (employee_id, name, surname, jobTitle)
 - d. bảng **program** (program_id, name, creditPoints, yearCommenced)
4. Viết các truy vấn chèn và điền dữ liệu vào các bảng với dữ liệu ngẫu nhiên (tối thiểu 10 dòng cho mỗi bảng).

Thực hành 1

5. Hiển thị tất cả sinh viên có tên bắt đầu bằng chữ „H”.
6. Hiển thị tất cả sinh viên hiện đang học năm thứ 4.
7. Hiển thị tất cả các khóa học từ bảng **course** bắt đầu với những khóa học có số điểm tín chỉ cao nhất.
8. Đổi tên sinh viên có student_id thấp nhất thành **Adam**.
9. Đổi tất cả giá trị trong cột **name** của bảng **course** thành chữ hoa.
- 10.Xóa sinh viên lớn tuổi nhất trong bảng **student**.
- 11.Loại bỏ cột **yearCommenced** khỏi bảng **course**.
- 12.Đổi tên bảng **staff** thành **employee**.
- 13.Nộp bài tập và Github Repo.

4. Sử dụng SQLite trong Python

- Python cung cấp một thư viện tích hợp sẵn tên là **sqlite3**, giúp làm việc với SQLite dễ dàng. Bạn có thể tạo cơ sở dữ liệu, thêm, sửa, xóa và truy vấn dữ liệu thông qua các lệnh SQL.

4. Sử dụng SQLite trong Python

- ***Bước 1: Cài đặt SQLite***
- SQLite thường được cài đặt sẵn với Python. Bạn có thể kiểm tra bằng cách thử nhập mô-đun sqlite3 trong một trình biên dịch Python:

```
import sqlite3
```

4. Sử dụng SQLite trong Python

- ***Bước 2: Kết nối đến cơ sở dữ liệu***
- Sử dụng mô-đun sqlite3 để kết nối với một cơ sở dữ liệu. Nếu cơ sở dữ liệu không tồn tại, SQLite sẽ tự động tạo mới.

```
# Kết nối đến cơ sở dữ liệu  
conn = sqlite3.connect('example.db')
```

4. Sử dụng SQLite trong Python

- ***Bước 3: Tạo con trỏ***
- Tạo một con trỏ để thực thi các câu lệnh SQL.

```
c = conn.cursor()
```

4. Sử dụng SQLite trong Python

- ***Bước 4: Tạo bảng***
- Sử dụng câu lệnh SQL để tạo một bảng trong cơ sở dữ liệu.

```
# Tạo bảng
c.execute('''
    CREATE TABLE IF NOT EXISTS users (
        id INTEGER PRIMARY KEY,
        name TEXT NOT NULL,
        age INTEGER
    )
''')
```

4. Sử dụng SQLite trong Python

- ***Bước 5: Chèn dữ liệu***
- Chèn dữ liệu vào bảng mà bạn đã tạo.

```
# Chèn dữ liệu
c.execute("INSERT INTO users (name, age) VALUES ('Alice', 30)")
c.execute("INSERT INTO users (name, age) VALUES ('Bob', 25)")
conn.commit()
```

4. Sử dụng SQLite trong Python

- ***Bước 6: Truy vấn dữ liệu***
- Truy vấn dữ liệu từ bảng và xử lý kết quả.

```
# Truy vấn dữ liệu
c.execute("SELECT * FROM users")
rows = c.fetchall()

for row in rows:
    print(row)
```


4. Sử dụng SQLite trong Python

- ***Bước 7: Cập nhật dữ liệu***
- Cập nhật dữ liệu trong bảng nếu cần.

```
# Cập nhật dữ liệu  
c.execute("UPDATE users SET age = 31 WHERE name = 'Alice'")  
conn.commit()
```

4. Sử dụng SQLite trong Python

- ***Bước 8: Xóa dữ liệu***
- Xóa một bản ghi từ bảng.

```
# Xóa dữ liệu
c.execute("DELETE FROM users WHERE name = 'Bob'")
conn.commit()
```

4. Sử dụng SQLite trong Python

- ***Bước 9: Đóng kết nối***
- Đóng kết nối sau khi bạn đã hoàn tất.

```
# Đóng kết nối  
conn.close()
```

```
import sqlite3

# Kết nối đến cơ sở dữ liệu
conn = sqlite3.connect('example.db')
c = conn.cursor()

# Tạo bảng
c.execute('''
    CREATE TABLE IF NOT EXISTS users (
        id INTEGER PRIMARY KEY,
        name TEXT NOT NULL,
        age INTEGER
    )
''')

# Chèn dữ liệu
c.execute("INSERT INTO users (name, age) VALUES ('Alice', 30)")
c.execute("INSERT INTO users (name, age) VALUES ('Bob', 25)")
conn.commit()
```

```
# Truy vấn dữ liệu
c.execute("SELECT * FROM users")
rows = c.fetchall()

for row in rows:
    print(row)

# Cập nhật dữ liệu
c.execute("UPDATE users SET age = 31 WHERE name = 'Alice'")
conn.commit()

# Xóa dữ liệu
c.execute("DELETE FROM users WHERE name = 'Bob'")
conn.commit()

# Đóng kết nối
conn.close()
```

Thực hành 2

- Xây dựng chương trình quản lý sổ địa chỉ với các tính năng:
 - Thêm bản ghi
 - Hiển thị bản ghi
 - Xóa bản ghi
 - Chỉnh sửa bản ghi

The screenshot shows a software window titled "Hệ thống quản lý địa chỉ" (Address Management System). It contains a form with the following fields and buttons:

- Form fields:
 - Họ (Surname)
 - Tên (Name)
 - Địa chỉ (Address)
 - Thành phố (City)
 - Tỉnh/Thành (Province/City)
 - Mã bưu chính (Postal Code)
- Buttons:
 - Thêm bản ghi (Add record)
 - Hiển thị bản ghi (Display record)
 - Chọn ID để xóa (Select ID to delete)
 - Xóa bản ghi (Delete record)
 - Chỉnh sửa bản ghi (Edit record)

At the bottom, a table header is visible with columns: ID, Le, Họ, Tung, and Tên. The first row of the table shows the value "1" under the ID column.

Thực hành 3

- Xây dựng chương trình quản lý sinh viên với các tính năng:
 - Thêm bản ghi
 - Hiển thị bản ghi
 - Xóa bản ghi
 - Chỉnh sửa bản ghi

Hệ thống quản lý sinh viên

Mã sinh viên

Họ

Tên

Mã lớp

Năm nhập học

Điểm trung bình

Thêm bản ghi

Hiển thị bản ghi

Chọn ID để xóa

Xóa bản ghi

Chỉnh sửa bản ghi

Mã sinh viên	Họ	Tên	Mã lớp	Năm nhập học	Điểm trung bình
1	Lê	Nhật Tùng	OSDS	2050	8.0

Thực hành 4

- Từ các bài tập trước đây:
 - Thu thập dữ liệu họa sĩ
 - Thu thập dữ liệu ca sĩ
 - Thu thập dữ liệu sản phẩm từ website
- Hãy viết lại các bài tập để lưu trữ dữ liệu đã thu thập được vào cơ sở dữ liệu SQLite

Thực hành 5

- Chọn một mã cổ phiếu của Việt Nam
- Tìm kiếm một trang có chứa thông tin lịch sử giá cổ phiếu đã chọn
- Tiến hành thu thập dữ liệu bằng Selenium về giá cổ phiếu trong 30 ngày gần nhất với các thông tin sau

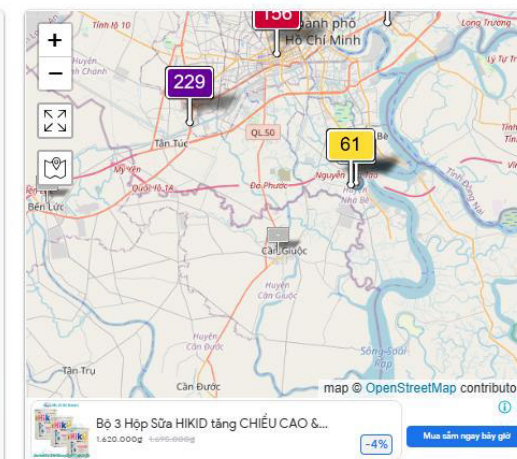
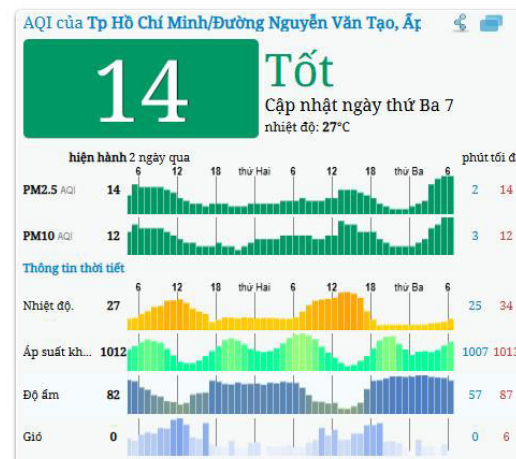
Ngày	Giá mở cửa	Giá cao nhất	Giá thấp nhất	Giá đóng cửa	Thay đổi giá	% thay đổi	Khối lượng
07/10/2024	56,600	57,100	56,500	56,700	+100	0.18%	414,800
04/10/2024	56,800	57,000	56,600	56,600	-200	-0.35%	381,200
03/10/2024	57,200	57,700	56,800	56,800	-100	-0.18%	606,400
02/10/2024	57,800	57,800	56,800	56,900	-900	-1.56%	635,600

Thực hành 5 (tt)

- Lưu trữ toàn bộ dữ liệu đã thu thập được vào SQLite
- Trả lời các câu hỏi liên quan đến dữ liệu:
 - Cho biết ngày và giá cổ phiếu cao nhất vào thời điểm đóng cửa
 - Cho biết giá cổ phiếu thấp nhất vào lúc đóng cửa
 - Trung bình giá cổ phiếu trong 30 ngày vào lúc đóng cửa

Thực hành 6

- Xây dựng một hệ thống ghi nhận chỉ số môi trường của TP. HCM
- Hệ thống chạy liên tục hàng ngày, mỗi ngày thu thập 2 lần, cần thu thập được các chỉ số như sau
- Tiến hành một số phân tích cơ bản sau 30 ngày thu thập



Gợi ý nguồn: <https://aqicn.org/city/vietnam/tp-ho-chi-minh/duong-nguyen-van-tao--ap-3--nha-be-kttv-tram-cam-bien/vn/>

Thực hành 6

- Gợi ý máy tính nhúng chạy hệ điều hành mã nguồn mở



Gợi ý nguồn: <https://aqicn.org/city/vietnam/tp-ho-chi-minh/duong-nguyen-van-tao--ap-3--nha-be-kttv-tram-cam-bien/vn/>