

Basic Python - Work with SQLite Database

Hoàng-Nguyên Vũ



1. Mô tả: Thống kê cơ bản trong SQLite

- **SQL** cung cấp nhiều hàm để thực hiện các phép tính thống kê cơ bản trên dữ liệu. Các hàm này được sử dụng để tính toán các giá trị như tổng, trung bình, giá trị tối đa, giá trị tối thiểu, v.v. Dưới đây là một số hàm thống kê cơ bản trong SQL:
 - **SUM()**: Tính tổng của các giá trị trong một cột.
 - **AVG()**: Tính trung bình của các giá trị trong một cột.
 - **MIN()**: Tìm giá trị nhỏ nhất trong một cột.
 - **MAX()**: Tìm giá trị lớn nhất trong một cột.
 - **COUNT()**: Đếm số lượng các giá trị trong một cột.

Ví dụ: Giả sử bạn có một bảng *Product* với các cột *name*, *brand* và *price*. Bạn muốn biết:

- Tổng doanh thu của tất cả các sản phẩm.
- Doanh thu trung bình của các sản phẩm.
- Giá sản phẩm cao nhất.
- Số lượng các sản phẩm khác nhau.

Câu lệnh SQL cho các yêu cầu trên sẽ như sau:

```
1 # Query để lấy tổng giá bán toàn bộ sản phẩm trong bảng Product
2 query = """
3 SELECT SUM(price) AS total_revenue
4 FROM PRODUCT;
5 """
6
7 data_sum = pd.read_sql_query(query, connection)
8 print(data_sum)
9
```

```

10 # Query để lấy thông tin sản phẩm có giá bán cao nhất
11 query = """
12 SELECT NAME, MAX(price) AS PRICE
13 FROM PRODUCT;
14 """
15 data_max = pd.read_sql_query(query, connection)
16 print(data_max)

```

→ Kết quả:

```

data=pd.read_sql_query("SELECT * FROM PRODUCT", connection)
print('=== All data ===')
print(data)

```

=== All data ===

	ID	NAME	BRAND	PRICE
0	1	iPhone 15	Apple	18000000
1	2	Galaxy Z-Fold 5	Samsung	30000000
2	3	Find X	Oppo	20000000
3	4	iPhone 14	Apple	16000000
4	5	Galaxy Z-Flip	Samsung	17000000
5	6	iPhone 15 Pro Max	Apple	48000000

```

query = """
SELECT SUM(price) AS total_revenue
FROM PRODUCT;
"""
data_sum = pd.read_sql_query(query, connection)
print(data_sum)

```

total_revenue
0 149000000

```

query = """
SELECT NAME, MAX(price) AS PRICE
FROM PRODUCT;
"""
data_max = pd.read_sql_query(query, connection)
print(data_max)

```

	NAME	PRICE
0	iPhone 15 Pro Max	48000000

• Hàm thống kê với GROUP BY:

GROUP BY là một mệnh đề trong SQL được sử dụng để nhóm các hàng dựa trên các giá trị chung trong một hoặc nhiều cột và thực hiện các phép tính tổng hợp trên các nhóm đó.

Hàm thống kê được sử dụng để tính toán các giá trị như tổng, trung bình, giá trị tối đa, giá trị tối thiểu, v.v. trên các nhóm dữ liệu.

Ví dụ: Giả sử bạn có một bảng *Product* với các cột *name*, *brand* và *price*. Bạn muốn biết:

- Doanh thu tổng cho mỗi hãng (BRAND) của sản phẩm.

- Giá bán thấp nhất của mỗi danh mục.

Cách thực thi câu lệnh SQL cho 2 yêu cầu trên sẽ như sau:

```

1 # Query để lấy tổng giá bán toàn bộ sản phẩm theo hãng trong bảng
  Product
2 query = """
3 SELECT BRAND, SUM(price) AS total_revenue
4 FROM PRODUCT
5 GROUP BY BRAND;
6 """
7 data_sum_by_brand = pd.read_sql_query(query, connection)
8 print(data_sum_by_brand)
9
10 # Query để lấy thông tin sản phẩm có giá bán thấp nhất
11 query = """
12 SELECT NAME, BRAND, MIN(price) AS PRICE
13 FROM PRODUCT
14 GROUP BY BRAND;
15 """
16 data_min_by_brand = pd.read_sql_query(query, connection)
17 print(data_min_by_brand)

```

→ Kết quả:

```

query = """
SELECT BRAND, SUM(price) AS total_revenue
FROM PRODUCT
GROUP BY BRAND;
"""

data_sum_by_brand = pd.read_sql_query(query, connection)
print(data_sum_by_brand)

```

	BRAND	total_revenue
0	Apple	82000000
1	Oppo	20000000
2	Samsung	47000000

```

query = """
SELECT NAME, BRAND, MIN(price) AS PRICE
FROM PRODUCT
GROUP BY BRAND;
"""

data_min_by_brand = pd.read_sql_query(query, connection)
print(data_min_by_brand)

```

	NAME	BRAND	PRICE
0	iPhone 14	Apple	16000000
1	Find X	Oppo	20000000
2	Galaxy Z-Flip	Samsung	17000000

2. Bài tập:

- Hãy tạo mới bảng có tên *STOCK* có các cột như sau:

Bảng 1: Bảng quản lý cổ phiếu

STOCK		
Cột	Kiểu dữ liệu	Chú thích
ID	INTEGER	Khóa chính
NAME	TEXT	Not null
BUY	INTEGER	Not null
INVESTOR	TEXT	Not null

Hãy thêm mới các dòng có giá trị như sau:

STOCK			
ID	NAME	BUY	INVESTOR
1	ACB	29.45	Nguyen
2	VIC	44.55	Nguyen
3	GMD	74.30	Nguyen
4	ACB	28.45	Vinh
5	VIC	40.55	Vinh
6	GMD	60.30	Vinh

Câu 1: Hãy viết lệnh SQL để truy vấn và in ra kết quả thống kê tổng giá bán (BUY) của bảng STOCK:

Kết quả: Tổng giá bán = 277.69

Câu 2: Hãy viết lệnh SQL để thống kê mã cổ phiếu có giá mua (BUY) lớn nhất theo nhà đầu tư (Investor):

→ **Kết quả:**

	INVESTOR	NAME	MAX_PRICE
0	Nguyen	GMD	74.39
1	Vinh	GMD	60.30