

Bài 2. Truy vấn căn bản với cơ sở dữ liệu (4 tiết)

1. Mục đích

Sau khi kết thúc bài thực hành này, sinh viên có thể:

- Sử dụng SQL để trả lời các truy vấn cơ bản với cơ sở dữ liệu (CSDL);
- Sử dụng các hàm gộp COUNT(.), MAX(.), MIN(.), AVG(.), và SUM(.) để trả lời các truy vấn cơ sở dữ liệu;
- Biết sử dụng kết hợp khối SELECT ... FROM ... WHERE với các hàm gộp có mệnh đề GROUP BY, HAVING...

2. Nội dung

2.1. Truy vấn đơn giản trên một bảng

- Khối cơ bản của câu lệnh SELECT
SELECT <tên của 1 cột | tên nhiều cột | * | tên hàm gộp | biểu thức>
FROM <tên của 1 bảng | 1 view | nhiều bảng |views>
[WHERE <điều kiện>]
[ORDER BY < tên của 1 cột | tên nhiều cột >]
[GROUP BY <tên cột> [HAVING <điều kiện>]]
[...];
- Hiển thị tất cả cột hoặc một số cột trong bảng;
- Hiển thị cột với tên khác;
- Sắp xếp việc hiển thị dữ liệu trong bảng – ORDER BY;
- Loại bỏ các dòng trùng nhau – DISTINCT;
- Hiển thị một số dòng – WHERE <điều kiện>;
- Tìm kiếm theo nhiều yêu cầu với các phép toán logic NOT, AND, OR;
- Tìm kiếm giá trị trong hoặc ngoài danh sách (list) sử dụng từ khóa IN, NOT IN;
- Tìm kiếm theo khoảng sử dụng từ khóa BETWEEN ... AND, NOT BETWEEN ... AND;
- Tìm kiếm xâu ký tự sử dụng từ khóa LIKE và các ký tự thay thế dấu '%', và dấu gạch dưới '_';
- Tìm kiếm với giá trị NULL, NOT NULL.

2.1. Truy vấn sử dụng các hàm gộp

- Sử dụng hàm gộp COUNT(), MAX(), MIN(), AVG(), và SUM().
- Sử dụng các hàm gộp trên với GROUP BY, GROUP BY HAVING, ...

3. Một số lưu ý

- Luật D' Morgan:
Giả sử p và q là các biểu thức logic, khi đó theo luật D' Morgan:
NOT(p AND q) tương đương với NOT(p) OR NOT(q)
NOT(p OR q) tương đương với NOT(p) AND NOT(q)
NOT(NOT(p)) tương đương với p
- Mức ưu tiên các phép toán logic: NOT, AND, OR. Trong một biểu thức logic nếu có dấu ngoặc (), phép NOT, phép AND và phép OR thì thứ tự thực hiện như sau: thứ nhất là dấu ngoặc (), thứ hai là phép NOT, sau đó đến phép AND và cuối cùng là phép OR.
- Các hàm SUM() và AVG() chỉ áp dụng được với kiểu dữ liệu của cột là kiểu số, các hàm gộp còn lại áp dụng được cho mọi kiểu dữ liệu.
- Trong các hàm gộp, chỉ có hàm COUNT(*) nhận giá trị NULL ở đầu vào
- Các hàm gộp không được đi sau mệnh đề WHERE
- Thứ tự thực hiện các mệnh đề trong khối lệnh SELECT ... FROM ... WHERE... GROUP BY ... HAVING ...

4. Thực hiện

A. Cho cơ sở dữ liệu có tên **HDThucTap** gồm các sơ đồ quan hệ sau:

TblGiangVien(magv int, hotengv char(30), bomon char(50));

TblSinhVien(masv int, hotensv char(30), namsinh int, quequan char(30))

TblDeTai(madt char(10), tendt char(30), kinhphi decimal(10,2))

TblHuongDan(masv int, madt char(10), magv int, ketqua decimal(5,2))

Trong đó:

- GiangVien, SinhVien và DeTai tương ứng là các sơ đồ quan hệ thể hiện các thuộc tính của các kiểu thực thể GIẢNG VIÊN, SINH VIÊN, và ĐỀ TÀI.
- HuongDan là sơ đồ quan hệ thể hiện tập liên kết của 3 kiểu thực thể trên. Cụ thể hơn, sơ đồ này sẽ cho biết, sinh viên nào làm đề tài gì, ai là người hướng dẫn, và điểm thực tập của sinh viên đó là bao nhiêu.
- Magv, hotengv và bomon là mã số, họ tên, và đơn vị chuyên môn của giáo viên tương ứng;
- Masv, hotensv, namsinh, và quequan là mã số, họ tên, năm sinh và quê quán (ghi tên tỉnh hoặc thành phố trực thuộc trung ương) của sinh viên tương ứng;

- Madt, tendt, kinhphi lần lượt là mã số, tên và kinh phí của đề tài;
- Ketqua là điểm thực hiện đề tài của sinh viên do giảng viên hướng dẫn đánh giá.

B. Sử dụng SSMS để tạo lập cấu trúc và các ràng buộc cho cơ sở dữ liệu HDThucTap:

```
CREATE DATABASE HDThucTap;

USE HDThucTap;

CREATE TABLE TBLGiangVien(
Magv int primary key,
Hotengv char(30),
Bomon char(50));

CREATE TABLE TBLSinHVien(
Masv int primary key,
Hotensv char(30),
Namsinh int,
Quequan char(30));

CREATE TABLE TBLDeTai(
Madt char(10)primary key,
Tendt char(30),
Kinhphi decimal(10,2));

CREATE TABLE TBLHuongDan(
Masv int primary key,
Madt char(10)foreign key references TBLDeTai,
Magv int foreign key references TBLGiangVien,
KetQua decimal(5,2));
```

C. Sử dụng lệnh INSERT INTO <tên_bảng>[danh_sách_tên_cột] VALUES <danh_sách_giá_trị > để nhập dữ liệu vào các bảng trên.

Chú ý: Dữ liệu trong các bảng là tiếng Việt, tuy nhiên nên nhập dữ liệu bằng tiếng Việt không dấu để thuận lợi hơn khi truy vấn, ví dụ với ‘Lê Giáp’ thì nên nhập dữ liệu là ‘Le Giap’

- Bảng **GiangVien**:

Mã số giảng viên	Họ và tên	Bộ môn
1001	Lê Giáp	Hệ thống thông tin
1002	Trần Ất	Khoa học máy tính
1003	Bùi Bình	Mạng và Truyền thông
1004	Phan Đình	Kỹ thuật phần mềm
1005	Nguyễn Mậu	Hệ thống thông tin

- Bảng **SinhVien**:

Mã số sinh viên	Họ và tên	Năm sinh	Quê quán (tên tỉnh/tp. trực thuộc trung ương)
1	Lê Văn Xuân	1990	Nghệ An
2	Trần Nguyên Hạ	1990	Thanh Hóa
3	Bùi Xuân Thu	1992	Nghệ An
4	Phan Tuấn Đông		Hà Tĩnh
5	Lê Thanh Xuân	1989	Hà Nội
6	Nguyễn Thu Hạ	1991	Thanh Hóa
7	Trần Xuân Tây		Hà Tĩnh
8	Hoàng Long Nam	1992	Nam Định

- Bảng **DeTai**:

Mã đề tài	Tên đề tài	Kinh phí (triệu đồng)
DT01	Semantic web	100
DT02	Cloud computing	500
DT03	Machine learning	100
DT04	Data mining	300

- Bảng **HuongDan**:

- Viết lệnh SQL nhập đúng 7 dòng dữ liệu vào bảng HuongDan, điều này có nghĩa là, sẽ có 1 sinh viên không tham gia thực tập;
- Xem dữ liệu của 3 bảng kia để nhập dữ liệu vào bảng HuongDan cho hợp lệ

D. Viết lệnh SQL trả lời các truy vấn sau:

- 1) Đưa ra thông tin cá nhân của tất cả sinh viên
- 2) Cho biết mã số, họ tên và quê quán của các sinh viên ở 'Nghệ An'
- 3) Cho biết mã số, họ tên và tuổi của tất cả các sinh viên
- 4) Cho biết họ tên, năm sinh của các sinh viên có mã số là 5, và 7.
- 5) Cho biết họ tên, và quê quán của các sinh viên không ở 'Nghệ An' và 'Hà Nội'

- 6) Đưa ra họ tên, năm sinh và quê quán của các sinh viên ở 'Thanh Hóa' và sinh năm 1990
- 7) Cho biết thông tin cá nhân của các sinh viên sinh năm 1990 và có quê ở 'Nghệ An' hoặc quê ở 'Thanh Hóa'
- 8) Cho biết thông tin cá nhân của các sinh viên có mã số 1, 3, 5, và 7
- 9) Đưa ra thông tin gồm tên và kinh phí của các đề tài có kinh phí từ 200 triệu đồng đến 500 triệu đồng và sắp xếp theo thứ tự giảm dần của kinh phí.
- 10) Cho biết thông tin cá nhân về các sinh viên có họ 'Trần'
- 11) Đưa ra họ tên, năm sinh và quê quán của các sinh viên có tên 'Xuân'
- 12) Đưa ra họ tên, năm sinh và quê quán của các sinh viên mà họ và tên chứa từ 'Xuân'
- 13) Cho biết họ tên, quê quán và mã số của các sinh viên chưa có thông tin năm sinh.
- 14) Cho biết số sinh viên có trong cơ sở dữ liệu
- 15) Cho biết số sinh viên có quê là 'Hà Tĩnh'
- 16) Cho biết số sinh viên của mỗi tỉnh/thành phố có trong cơ sở dữ liệu
- 17) Cho biết tên tỉnh (tên thành phố) có số sinh viên theo học là 1.
- 18) Cho biết mã số của giảng viên hướng dẫn thực tập nhiều hơn một sinh viên
- 19) Cho biết mã số của các đề tài có nhiều hơn 3 sinh viên thực tập
- 20) Cho biết mã số của các sinh viên chưa có điểm thực tập.
- 21) Cho biết mã số, họ tên của các sinh viên có điểm thực tập cao nhất
- 22) Cho biết tên các đề tài không có sinh viên nào thực tập

5. Câu hỏi ngắn

- Cho biết các khóa ngoài trong sơ đồ quan hệ HươngDan?
- Theo sơ đồ của CSDL đã cho, một sinh viên có thể tham gia thực tập nhiều hơn một đề tài không? Tại sao?
- Theo sơ đồ của CSDL Thực tập, một giảng viên có thể hướng dẫn nhiều hơn một sinh viên không? Tại sao?

6. Về nhà

- Vẽ sơ đồ ER cho cơ sở dữ liệu ThucTap.
- Sử dụng kiểu Nchar(), Nvarchar() cho CSDL trên để có thể Nhập và Truy vấn với dữ liệu tiếng Việt (mã hóa Unicode)
- Trả lời các truy vấn sau bằng đại số quan hệ
 - Cho biết tên các đề tài mà giảng viên 'Bùi Bính' hướng dẫn
 - Cho biết tên đề tài không có sinh viên nào thực tập
 - Cho biết tên đề tài có nhiều hơn 2 sinh viên thực tập
 - Đưa ra mã số, họ tên và điểm của các sinh viên