

## 1. Phân biệt và trình bày nguyên lý hoạt động của: JOIN, LEFT JOIN, RIGHT JOIN.

Trả lời:

- a. JOIN (INNER JOIN): Trả về tất cả hàng khi có ít nhất một giá trị ở cả 2 bảng

Ví dụ:

Table 1	Table 2	JOIN
A	C	C
B	D	D
C	E	E
D	F	
E	G	

- b. LEFT JOIN: Trả về tất cả dòng bảng bên trái và các dòng đúng với điều kiện từ bảng bên phải

Ví dụ:

Table 1	Table 2	LEFT JOIN
A	C	A
B	D	B
C	E	C
D	F	D
E	G	E

- c. RIGHT JOIN: Trả về tất cả dòng bảng bên phải và các dòng đúng với điều kiện từ bảng bên phải

Ví dụ:

Table 1	Table 2	RIGHT JOIN
---------	---------	------------

A	C	C
B	D	D
C	E	E
D	F	F
E	G	G

**Vậy tại sao chúng ta lại phải thêm Right Join trong khi có thể đổi vị trí bảng khi thực hiện Left Join ?**

Lý do chính là do Left Join thường dễ đọc và hiểu hơn. Bên cạnh đó, nhiều ngôn ngữ truy vấn SQL được thiết kế để hỗ trợ Left Join làm mặc định, trong khi Right Join không phải lúc nào cũng được hỗ trợ tốt.

**2. Thu thập thông tin về yêu cầu tuyển dụng của một công ty mà anh/chị yêu thích và so sánh tương quan của các yêu cầu đó với chương trình học của bản thân. (không bắt buộc)**

**Trả lời:**

**- Thông tin về yêu cầu tuyển dụng của Ecv Technology Company Limited :**

- Có kinh nghiệm phân tích dữ liệu vững vàng và thành thạo SQL / Python / R
- Khả năng hiểu Kiến trúc dữ liệu từ đầu đến cuối bao gồm Khai thác dữ liệu, Chuyển đổi dữ liệu và Mô hình hóa dữ liệu.
- Kinh nghiệm thực tế để thiết kế nền tảng và giải pháp dữ liệu doanh nghiệp kết hợp Dữ liệu lớn, Hệ thống AI và Đám mây (AWS là một điểm cộng).
- Thành thạo Kho dữ liệu đám mây hiện đại (ví dụ: Snowflake, Redshift)

- Có kinh nghiệm về AI/ML là một lợi thế
- Kỹ năng giao tiếp bằng văn bản và nói tốt bằng cả tiếng Việt và tiếng Anh.

- So sánh tương quan:

- Phân tích dữ liệu, kỹ thuật lập trình Python, phân tích và trực quan dữ liệu
- Điện toán đám mây
- Máy học, học máy thống kê
- Quản lý thông tin
- Phân tích dữ liệu lớn

### 3. Tìm hiểu, phân tích điểm giống và khác nhau của **SELECT**, **SELECT \***, **SELECT DISTINCT**

- **SELECT** được sử dụng để truy vấn dữ liệu từ một bảng trong cơ sở dữ liệu. **SELECT** cho phép chọn các cột cụ thể muốn trả về trong kết quả truy vấn.
- **SELECT \*** sẽ trả về tất cả các cột có trong bảng hoặc các bảng đã chỉ định.
- **SELECT DISTINCT** sử dụng để truy vấn dữ liệu, nhưng nó chỉ trả về các giá trị duy nhất từ một cột cụ thể, nó sẽ loại bỏ các giá trị trùng lặp và chỉ trả về mỗi giá trị một lần duy nhất trong kết quả.

**Giống nhau:** Cả 3 đều dùng để truy vấn dữ liệu từ các bảng/view trong cơ sở dữ liệu.

**Khác nhau:**

- **SELECT:** Chỉ lấy ra các cột cụ thể được chỉ định sau mệnh đề **SELECT**.
- **SELECT \*:** Lấy ra tất cả các cột trong bảng.

- **SELECT DISTINCT:** Loại bỏ các bản ghi trùng lặp và chỉ lấy ra các giá trị khác nhau (unique).

#### **4. Ưu nhược điểm của exists (linh hoạt, dễ dùng, dễ đọc – khó hiểu)**

##### **Ưu điểm:**

- **Linh hoạt** - EXISTS có thể được sử dụng với các câu lệnh SELECT, INSERT, UPDATE, DELETE để kiểm tra sự tồn tại của dữ liệu, kiểm tra sự tồn tại dữ liệu dựa trên điều kiện phức tạp, nào đó.
- **Dễ sử dụng, dễ đọc** - Cú pháp đơn giản.
- **Hiệu quả** - EXISTS sẽ dừng việc kiểm tra ngay khi tìm thấy bản ghi đầu tiên thỏa mãn điều kiện. Do đó nó được xem là hiệu quả hơn so với IN trong một số trường hợp.

##### **Nhược điểm:**

- **Khó hiểu** - Với những người mới học SQL, EXISTS có thể khó hiểu hơn so với các toán tử khác.
- **Khó tối ưu hóa** - Việc tối ưu hóa câu lệnh sử dụng EXISTS có thể phức tạp hơn so với IN do phải tối ưu phụ thuộc bên trong EXISTS.

#### **5. Thực hiện 06 bài tập về truy vấn SQL tại slide 33**

- Tìm số lượng nhân viên của từng phòng ban  

```
SELECT PB.TENPHONG, COUNT(NV.MANV) AS "SLNV"  
FROM PHONGBAN AS PB, NHANVIEN AS NV  
WHERE PB.MAPHG = NV.PHG  
GROUP BY PHONGBAN
```

- Tìm số lượng nhân viên nam của từng phòng ban  
SELECT PB.TENPHONG, COUNT(NV.MANV) AS "SLNV Nam"  
FROM PHONGBAN PB JOIN NHANVIEN NV  
ON PB.MAPHG = NV.PHG  
WHERE NV.PHAI = 'NAM'  
GROUP BY PB.
- Tìm phòng ban có từ 2 nhân viên nam trở lên  
SELECT PB.TENPHONG  
FROM PHONGBAN PB JOIN NHANVIEN NV  
ON PB.MAPHG = NV.PHG  
WHERE NV.PHAI = 'NAM'  
GROUP BY PB.MAPHG  
HAVING COUNT(\*) > 2
- Tìm phòng ban có đông nhân viên nhất  
SELECT TOP 1 WITH TIES PB.TENPHONG  
FROM PHONGBAN PB JOIN NHANVIEN NV  
ON PB.MAPHG = NV.PHG  
GROUP BY MAPHG  
ORDER BY COUNT(NV.MANV) DESC
- Tìm đề án có ít nhân viên nữ tham gia nhất  
SELECT TOP 1 DA.TENDA, COUNT(\*) AS "SO NV NU"  
FROM PHANCONG PC JOIN NHANVIEN NV ON  
PC.MA\_NVN = NV.MANV JOIN DEAN DA ON  
DA.MADA = PC.SODA  
WHERE PHAI = 'NU'

GROUP BY DA.TENDA  
ORDER BY [SO NV NU] ASC

- Tìm 3 nhân viên thuộc phòng số 4 có lương thấp nhất  
SELECT TOP 3 NV.HOVN + ' ' NV.TENLOT + ' ' NV.TENNV AS  
"HO TEN"  
FROM NHANVIEN NV JOIN PHONGBAN PB ON  
NV.PHG = PB.MAPHG  
WHERE PB.MAPHG = '4'  
ORDER BY LUONG DESC