**KHÁC NHAU GIỮA WHERE VÀ HAVING**

**BÀI THAM KHẢO SỐ 1**

WHERE - filter kết quả theo dòng

HAVING - filter kết quả theo GROUP

Ví dụ: bạn có 1 bảng như sau

|  |  |  |
| --- | --- | --- |
| **Student** | **Subject** | **Point** |
| A | Math | 8 |
| A | Physics | 5 |
| B | Math | 4 |
| C | Physics | 9 |

Câu lệnh:

**SELECT Student, sum(Point)**

**FROM Point**

**GROUP BY Student**

*Kết quả trả về:*

|  |  |  |
| --- | --- | --- |
| **Student** | **Point** | **Chú thích** |
| A | 13 | Nó kiếm Student giống nhau ở đây 2 dòng đều có A nên lấy SUM |
| B | 4 |  |
| C | 9 |  |

Câu lệnh:

**SELECT Sudent, sum(Point)**

**FROM Point**

**WHERE Point > 5**

**GROUP BY Student**

*Kết quả trả về:*

|  |  |
| --- | --- |
| **Student** | **Point** |
| A | 8 |
| C | 9 |

*=> Nó sẽ check WHERE từng dòng trước, sau đó mới GROUP BY*

Câu lệnh:

**SELECT Student, sum(Point) as S**

**FROM Point**

**WHERE Point > 5**

**GROUP BY Student**

**HAVING S > 8**

*Kết quả trả về:*

|  |  |
| --- | --- |
| **Student** | **Point** |
| C | 9 |

*=> HAVING thì mới dùng với các điều kiện của cả nhóm (GROUP) như max, sum, count, …được (gọi là các hàm tổng hợp (AGGREGATE function) được define trên 1 tập hợp)*

**Định nghĩa:**

- Cách 1:

+ Where thường dùng để tìm điều kiện cho thuộc tính của Bảng.

+ Having thường dùng để tìm điều kiện trên các Function (SUM,AVG,COUNT…) trên các thuộc tính của Bảng.

- Cách 2:

+ Where : Là câu lệnh điều kiện trả kết quả đối chiếu với từng dòng .

+ Having : Là câu lệnh điều kiện trả kết quả đối chiếu cho nhóm (Sum,AVG,COUNT,…)

***=> Vì vậy nên sau “GROUP BY” thì sẽ chỉ dùng được “HAVING” . Còn “WHERE” thì KHÔNG dùng được sau “GROUP BY”.***

***=> “HAVING” có thể thay thế vị trí dùng cho “WHERE” NHƯNG ngược lại “WHERE” thì KHÔNG thể thay thế vị trí cho “HAVING”.***

***=> “WHERE” phải trước “GROUP BY”.***

**BÀI THAM KHẢO SỐ 2**

**MỆNH ĐỀ WHERE**

*Chức năng chính:* Mệnh đề WHERE ***dùng để kiểm tra các điều kiện cho các thuộc tính của bảng, trả kết quả đối chiếu với từng dòng.***

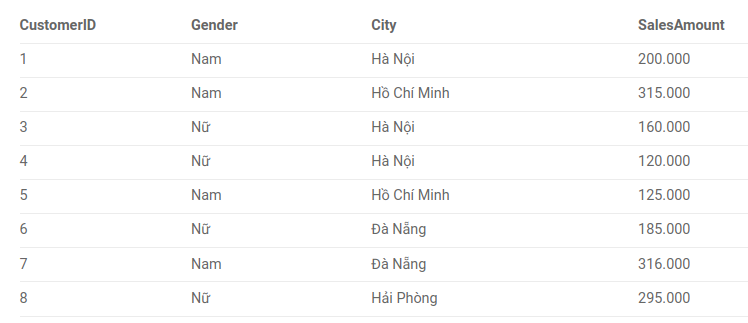
Syntax:

**SELECT** column\_lists

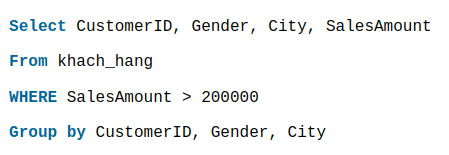
**FROM** table\_name

**WHERE** conditions

**GROUP BY** column\_lists

Ví dụ: Bảng khach\_hang

*Ví dụ khi thực hiện câu truy vấn:*

*Kết quả trả về:*



**MỆNH ĐỀ HAVING**

*Chức năng chính:*

- Mệnh đề HAVING ***được dùng để giới hạn nhóm các hàng trả về trong bảng***, ***trả kết quả đối chiếu cho các nhóm (Sum, Count, Average,…) được tạo bởi Group by.***

- Mệnh đề HAVING ***được thêm vào SQL vì mệnh đề WHERE không thể sử dụng được với các aggregate functions.***

=> ***Có thể sử dụng mệnh đề WHERE và HAVING cùng nhau trong quá trình lựa chọn.*** Trong trường hợp này, ***đầu tiên mệnh đề WHERE lọc các hàng riêng lẻ,*** ***sau đó các hàng được nhóm lại, thực hiện các phép tính tổng hợp*** và ***cuối cùng mệnh đề HAVING lọc các nhóm.***

Syntax:

**SELECT** column\_lists, aggregate\_function (expression)

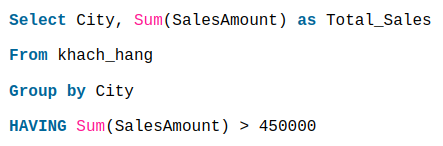
**FROM** table\_name

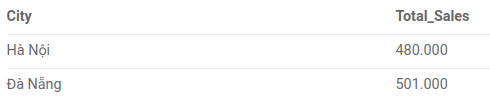
**WHERE** conditions (điều kiện)

**GROUP BY** column\_lists (danh sách các cột gom nhóm)

**HAVING** conditions (điều kiện trên nhóm)

*Ví dụ khi thực hiện câu truy vấn:*

*Kết quả trả về:*



**PHÂN BIỆT MỆNH ĐỀ WHERE VÀ HAVING TRONG SQL**

|  |  |  |
| --- | --- | --- |
| **Nội dung** | **WHERE** | **HAVING** |
| Cách thức Filter | Từng dòng riêng lẻ | Từng nhóm |
| Dùng với aggregate functions | Không | Có |
| Dùng với các câu lệnh | Select, Update, Delete | Chỉ dùng với Select |
| Dùng với Group by | Dùng trước Group by | Dùng sau Group by |

=> **Mệnh đề WHERE** khi được **sử dụng với GROUP BY**, **xuất hiện trước GROUP BY**. Điều này biểu thị rằng **các hàng bộ lọc WHERE trước khi tính toán tổng hợp được thực hiện**. Mặt khác, **HAVING** **xuất hiện sau GROUP BY**, có nghĩa là **nó lọc sau khi tính toán tổng hợp được thực hiện.**

=> Kết luận: **Cả 2 mệnh đề này đều có chức năng lọc dữ liệu**, tuy nhiên vẫn có một số điểm khác biệt ở cách thức lọc, kết hợp với các câu lệnh khác. Nhìn chung **mệnh đề HAVING được trang bị thêm một số tính năng mà mệnh đề WHERE không có**.

**KHÁC NHAU GIỮA DISTINCT VÀ GROUP BY**

**DISTINCT**

Chức năng:

Lệnh này cho phép bạn loại bỏ những bản ghi trùng lặp khi select. Điều này đặc biệt có ý nghĩa khi bảng được select bị trùng lặp dữ liệu trên một vài đầu thuộc tính hoặc khi join các bảng lại với nhau, làm tăng thêm những kết quả không thực sự cần đến.

Syntax:

SELECT DISTINCT Columns

FROM Table\_name

WHERE Where\_conditions

***Chú ý: Cả hai mệnh đề DISTINCT và GROUP BY đều làm giảm số lượng bản ghi trả về trong tậ kết quả bằng cách loại bỏ các bản ghi trùng lặp. Tuy nhiên, nên sử dụng mệnh đề GROUP BY khi muốn áp dụng các hàm tập hợp trên một hoặc nhiều cột.***

***Sử dụng DISTINCTnếu bạn chỉ muốn loại bỏ các bản sao. Sử dụng GROUPY BY nếu bạn muốn áp dụng các toán tử tổng hợp ( MAX, SUM, GROUP\_CONCAT, ... hoặc một mệnh đề HAVING).***

***Nếu bạn muốn nhóm các kết quả của mình, hãy sử dụng GROUP BY, nếu bạn chỉ muốn một danh sách duy nhất của một cột cụ thể, hãy sử dụng DISTINCT. Điều này sẽ giúp cơ sở dữ liệu của bạn có cơ hội tối ưu hóa truy vấn cho nhu cầu của bạn.***

***GROUP BY được sử dụng để nhóm các hàng mà bạn muốn tính toán. DISTINCT sẽ không thực hiện bất kỳ phép tính nào.***

**CÁC LOẠI RÀNG BUỘC (CONSTRAINT) PHỔ BIẾN**

|  |  |
| --- | --- |
| **Loại ràng buộc** | **Ý nghĩa** |
| NOT NULL | Sử dụng để đảm bảo dữ liệu của cột không được nhận giá trị NULL |
| DEFAULT | Gán giá trị mặc định trong trường hợp dữ liệu của cột không được nhập vào hay không được xác định. |
| UNIQUE | Sử dụng để đảm bảo dữ liệu của cột là duy nhất, không trùng lặp giá trị trên cùng 1 cột. |
| PRIMARY KEY (Khóa chính) | Dùng để thiết lập khóa chính trên bảng, xác định giá trị trên tập các cột làm khóa chính phải là duy nhất, không được trùng lặp. Việc khai báo ràng buộc khóa chính yêu cầu các cột phải NOT NULL. |
| FOREIGN KEY (Khóa ngoại) | Dùng để thiết lập khóa ngoại trên bảng, tham chiếu đến bảng khác thông qua giá trị của cột được liên kết. Giá trị của cột được liên kết phải là duy nhất trong bảng kia. |
| CHECK | Bảo đảm tất cả giá trị trong cột thỏa mãn điều kiện nào đó. Đây là hình thức sử dụng phổ biến để kiểm tra tính hợp lệ của dữ liệu (validate data) |

RÀNG BUỘC UNIQUE VÀ RÀNG BUỘC PRIMARY KEY

KHÁI NIỆM PRIMARY KEY

- Một cột có thể được gọi là primary key của bảng nếu nó xác định duy nhất từng bộ (hàng) trong bảng đó. Nó thực thi các ràng buộc toàn vẹn đối với bảng. Chỉ có một primary key trong một bảng. Primary key không chấp nhận các giá trị trùng lặp và NULL. Primary key được chọn cẩn thận khi các thay đổi có thể xảy ra theo cách hiếm khi xảy ra, có nghĩa là khóa chính trong bảng rất hiếm khi thay đổi.

- Primary key có thể được tham chiếu bằng foreign key. Nó tạo ra một chỉ mục nhóm duy nhất trên bảng. Trong một chỉ mục được phân nhóm, các hàng dữ liệu được sắp xếp và lưu trữ trong một bảng hoặc các views trên cơ sở các giá trị chính của nó. Chỉ có thể có một chỉ mục nhóm trong một bảng, lý do đằng sau điều này là một hàng dữ liệu trong bảng chỉ có thể được sắp xếp theo một thứ tự.

KHÁI NIỆM UNIQUE

- Tương tự như primary key, ràng buộc unique key cũng xác định duy nhất một bộ giá trị riêng lẻ trong một mối quan hệ. Nhưng, có những khác biệt nhất định giữa chúng. Một bảng có thể có nhiều hơn một khóa duy nhất. Ràng buộc unique key chỉ có thể chấp nhận một giá trị NULL cho một cột.

- Các ràng buộc duy nhất cũng được tham chiếu bởi foreign key. Nó có thể được sử dụng khi ai đó muốn thực thi các ràng buộc trên một cột và một nhóm các cột không phải là primary key. Không giống như primary key, nó tạo chỉ mục không phân cụm. Các chỉ mục không phân cụm có cấu trúc khác biệt với các hàng dữ liệu. Mỗi mục nhập khóa-giá trị trong đó trỏ đến hàng dữ liệu chứa giá trị khóa do đó nó sử dụng con trỏ.

|  |  |  |
| --- | --- | --- |
| **Tiêu chí** | **Primary Key** | **Unique Key** |
| Công dụng | Nó được sử dụng để làm định danh duy nhất cho mỗi hàng trong bảng. | Nó cũng xác định duy nhất một hàng, nhưng không là khóa chính. |
| NULL | không chấp nhận giá trị NULL | chấp nhận giá trị NULL |
| Số lượng khóa có thể được xác định trong bảng | Chỉ có duy nhất 1 primary key trong 1 bảng | Có thể nhiều hơn 1 |
| Index | Tạo chỉ mục theo nhóm | Tạo chỉ mục không phân cụm |
| Tự động tăng | Hỗ trợ giá trị gia tăng tự động. | Không hỗ trợ giá trị gia tăng tự động. |
| Sửa đổi | Không thể thay đổi hoặc xóa các giá trị được lưu trữ trong các khóa chính. | Có thể thay đổi các giá trị khóa duy nhất. |

- Giống: PRIMARY KEY, IDENTITY, và Unique đều có thể được sử dụng để đảm bảo tính toàn vẹn của dữ liệu.

- Khác:

+ PRIMARY KEY(PK): ràng buộc PK được dùng để định nghĩa khóa chính của bảng. Nó đảm bảo rằng không có giá trị của hàng nào trong cột là giống nhau. Hay là nói 1 cách dễ hiểu ra là nó xác định được tính duy nhất của bản ghi. Ràng buộc PK không cho phép giá trị NULL

+ Unique: Để thay cho việc không sử dụng khóa chính, bạn có thể dùng ràng buộc Unique. Nó cũng xác định được tính duy nhất của bản ghi. Nhưng bạn lưu ý rằng, nó có thể xác định tính duy nhất của bản ghi chứ nó không phải là khóa chính. Ràng buộc Unique cho phép giá trị NULL

+ Identity: Theo lý thuyết thì Identity được sử dụng để tự động sinh ra các giá trị của 1 cột 1 cách tuần tự.

Thực hành thì dễ hơn nhiều. )

IDENTITY(giá trị đầu,giâ trị tăng)

Ví dụ: IDENTITY(1,1) --> Nghĩa là: bắt đầu từ một và sau mỗi bản ghi được chèn vào giá trị nó tăng lên 1

Cũng như Unique, Identity nó cũng đc sử dụng để xác định tính duy nhất của bản ghi và nó không phải là khóa chính.

**SQL SEVER VÀ POSTGRE SQL**

SQL SEVER

Ngôn ngữ truy vấn có cấu trúc hoặc SQL, thường được biết đến nhiều hơn, là ngôn ngữ lập trình được sử dụng để quản lý, truy vấn và truy xuất dữ liệu trong cơ sở dữ liệu quan hệ . Đây là ngôn ngữ tiêu chuẩn được sử dụng bởi các hệ thống quản lý cơ sở dữ liệu quan hệ (RDBMS), bao gồm PostgreSQL, SQL Server, MySQL và Cơ sở dữ liệu Oracle.

SQL thường sử dụng các lệnh được viết ở định dạng câu lệnh cho các truy vấn và các hoạt động cơ sở dữ liệu khác, cho phép người dùng thao tác dữ liệu trong các bảng cơ sở dữ liệu quan hệ. Mặc dù ban đầu được tạo cho cơ sở dữ liệu quan hệ, nhưng SQL đóng vai trò là nền tảng cho nhiều khả năng công nghệ ngày nay, khiến kiến ​​thức SQL trở thành một kỹ năng thiết yếu cho nhiều vai trò công nghệ ngày nay, bao gồm nhà phân tích dữ liệu, kỹ sư cơ sở dữ liệu và thậm chí cả lập trình phụ trợ.

POSTGRE SQL

PostgreSQL là một hệ thống quản lý cơ sở dữ liệu quan hệ đối tượng nguồn mở, được phát hành theo Giấy phép PostgreSQL. Nó hỗ trợ cả quan hệ (SQL) và không quan hệ (JSON) và cung cấp các hàm SQL nâng cao, bao gồm khóa ngoại, truy vấn phụ và trình kích hoạt. PostgreSQL cũng có khả năng mở rộng cao, cho phép bạn xác định các kiểu dữ liệu và tạo các hàm tùy chỉnh.

SỰ TƯƠNG ĐỒNG GIỮA SQL SEVER VÀ POSTGRE SQL

Nói chung, SQL Server và PostgreSQL là một trong những hệ thống quản lý cơ sở dữ liệu quan hệ phổ biến nhất hiện có dựa trên hiệu suất, bảo mật, khả năng mở rộng và khả năng sử dụng của chúng. Cả hai đều bao gồm các tính năng cơ sở dữ liệu quan hệ và cung cấp khả năng tương thích với nhiều ứng dụng doanh nghiệp lớn và nhỏ.

Thông thường, SQL Server là lựa chọn cho các tổ chức lớn hơn phụ thuộc vào các sản phẩm của Microsoft. Tuy nhiên, PostgreSQL đã tạo ra không gian cho chính nó như một hệ thống quản lý cơ sở dữ liệu miễn phí, dễ triển khai, mang lại tính linh hoạt và chức năng tối đa.

SỰ KHÁC BIỆT GIỮA SQL SEVER VÀ POSTGRE SQL

|  |  |
| --- | --- |
| **SQL Sever** | **PostgreSQL** |
| Hệ thống quản lý cơ sở dữ liệu quan hệ | Hệ quản trị cơ sở dữ liệu quan hệ đối tượng |
| Sản phẩm thương mại của Microsoft | Mã nguồn mở (Hoàn toàn miễn phí) |
| Chỉ chạy trên Microsoft hoặc Linux | Chạy trên hầu hết các máy và hệ điều hành |
| Sử dụng Transact-SQL hoặc T-SQL (SQL tiêu chuẩn + chức năng bổ sung) | Sử dụng SQL chuẩn |

Bảng so sánh 2:

|  |  |  |
| --- | --- | --- |
|  | **SQL Sever** | **PostgreSQL** |
| **SELECT** | Select [col1], [col2] | SELECT col1, col2 |
| **Dữ liệu từ các bảng có phân biệt chữ hoa chữ thường không?** | WHERE name = ‘John’ WHERE name = ‘john’  => Không giống nhau | WHERE name = ‘John’ WHERE name = ‘john’  => Không giống nhau |
| **Sử dụng dấu ngoặc kép** | Name = ‘John’ | Name = ‘John’ |
| **Bí danh cho cột và bảng** | SELECT AVG(col1)=avg1 | SELECT AVG(col1) AS avg1 |
| **Làm việc với ngày tháng** | GETDATE()  DATEPART() | CURRENT\_DATE() CURRENT\_TIME() EXTRACT() |