LAB: Mẫu string, sắp xếp & nhóm

Thời gian làm bài: 30 phút

Các bài toán thực hành cho bài Lab này sẽ giúp bạn có thêm kinh nghiệm thực hành với các mẫu string, sắp xếp và nhóm các tập kết quả.

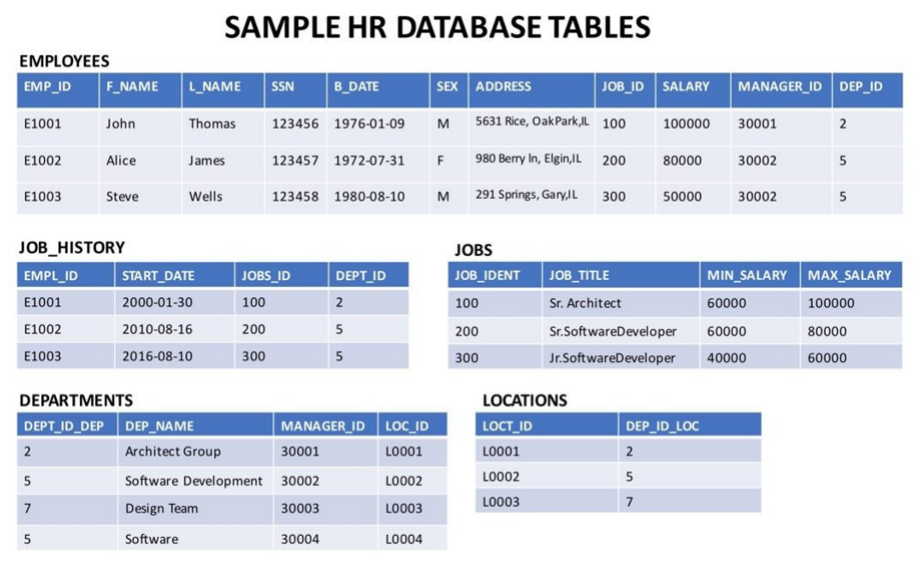
Bạn cũng sẽ học cách chạy các tập lệnh SQL để tạo nhiều bảng cùng một lúc, cũng như cách tải dữ liệu vào bảng từ các file .csv.

Cơ sở dữ liệu HR (Nhân sự)

Chúng ta sẽ thực hành với một cơ sở dữ liệu nhân sự mẫu cho bài Lab này. Lược đồ cơ sở dữ liệu nhân sự này bao gồm 5 bảng được gọi

là EMPLOYEES (nhân viên), JOB\_HISTORY (lịch sử công việc), JOBS (các công việc), DEPARTMENTS (phòng ban) và LOCATIONS (địa

điểm). Mỗi bảng có một số hàng dữ liệu mẫu. Biểu đồ sau đây gồm các bảng cho cơ sở dữ liệu nhân sự.



Để hoàn thành bài Lab này, bạn sẽ sử dụng ứng dụng XAMPP để thực hành với SQL.

Bài Lab này có ba phần:

I. Tạo bảng

II. Nạp dữ liệu vào bảng

III.Soạn và chạy truy vấn

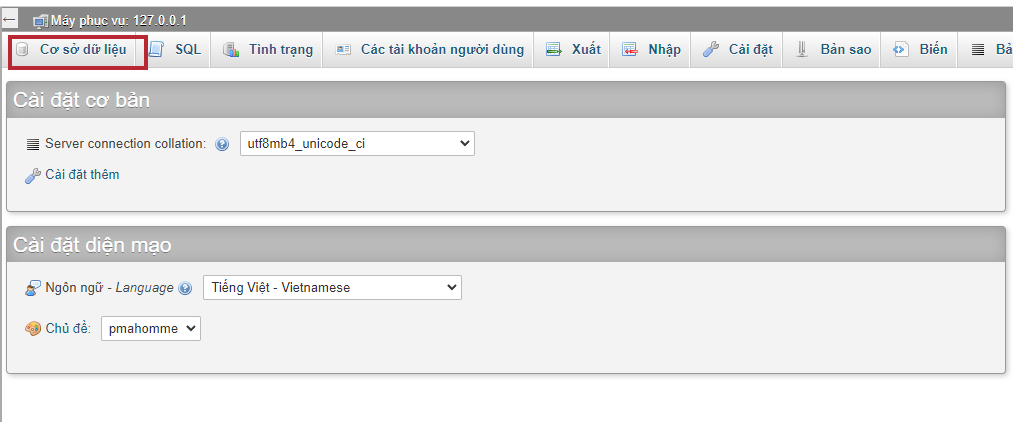
## Phần I. Tạo bảng

Thay vì nhập thủ công các lệnh DLL (Ngôn ngữ định nghĩa dữ liệu) trong editor của SQL để tạo từng bảng, bạn sẽ thực thi một tập lệnh

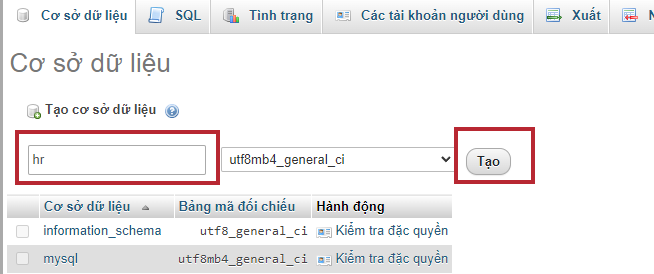
chứa các lệnh tạo bảng cho tất cả các bảng. Hãy xem hướng dẫn từng bước sau để thực hiện việc này:

1. Download file tập lệnh Script\_Create\_Tables.sql ở trang Lab
2. Đăng nhập vào trang quản trị Database của xampp thông qua link <http://localhost/phpmyadmin/>
3. Bạn sẽ thực hiện tạo một Database mới tên là hr

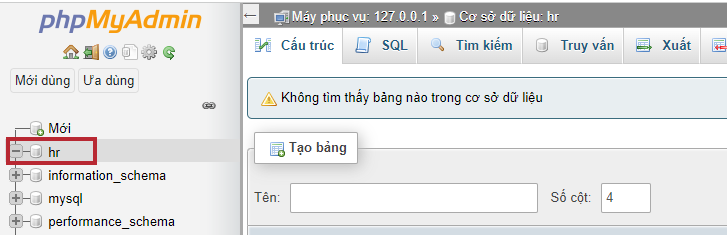
Đầu tiên chúng ta vào tab Cơ sở dữ liệu



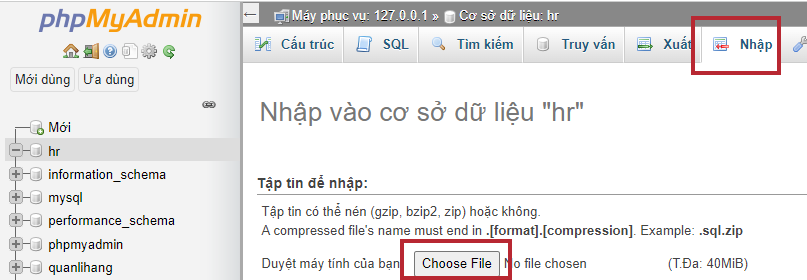
Sau đó thực hiện tạo Database mới với tên là hr,



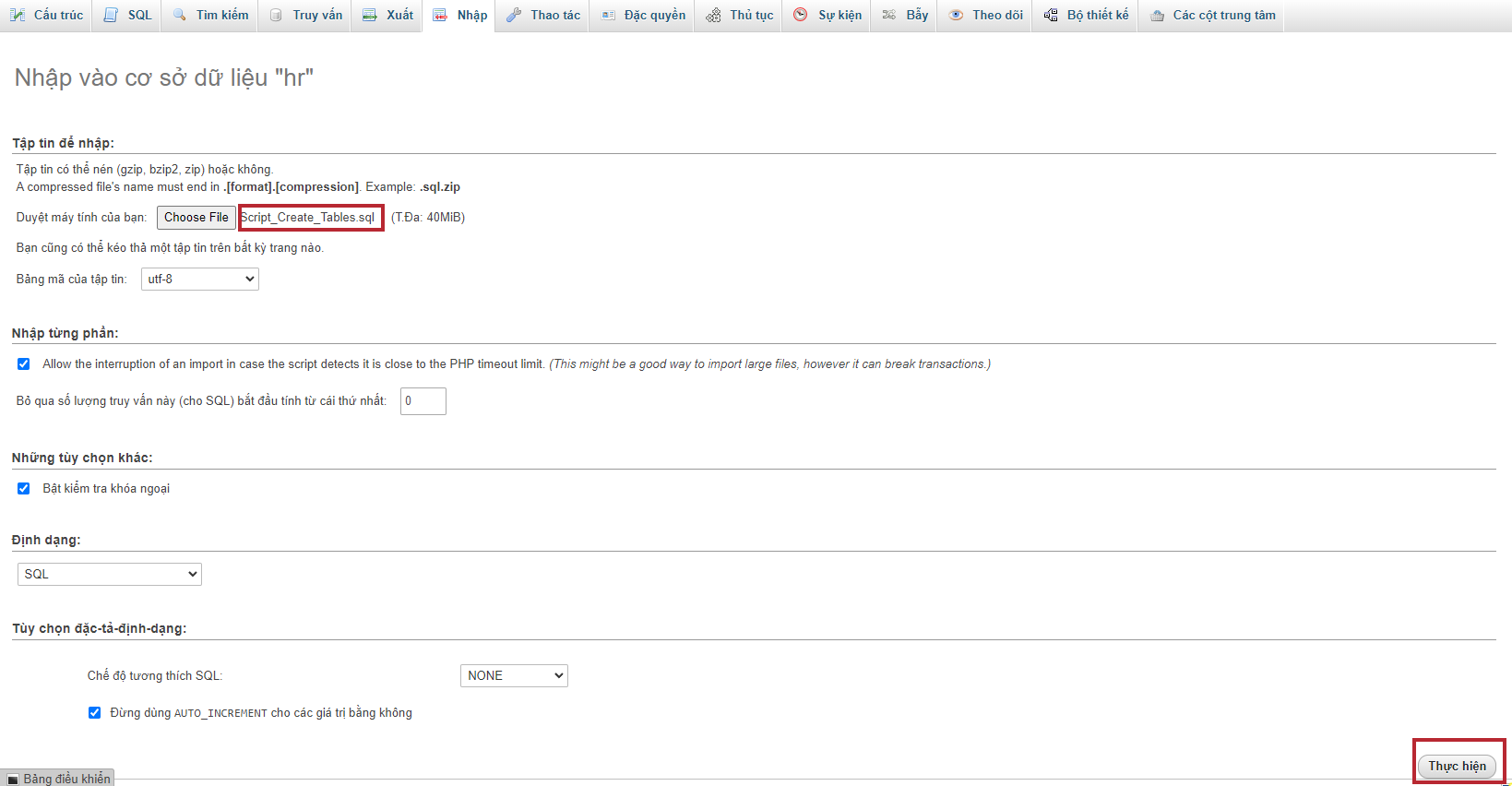
Sau khi ấn “Tạo” các bạn sẽ thấy đã được điều hướng đến Database có tên là hr:



Sau đó chúng ta sẽ chọn vào Tab Nhập (hoặc Inport nếu là tiếng anh) và thực hiện chạy file Script vừa tải về:



Sau khi chọn file xong chúng ta sẽ chọn “Thực hiện”:



Và đây là kết quả khi thực hiện import file script thành công:



Giờ chúng đã tạo thành công 5 bảng, chúng ta có thể chọn vào một bảng bất kỳ sau đó chọn vào tab “Cấu trúc” để xem được kiểu dữ liệu cũng như tên các trường:



## Phần 2. Nạp dữ liệu

Bây giờ, hãy xem cách nạp dữ liệu vào từng bảng. Chúng ta có thể chèn từng hàng vào bảng theo cách thủ công nhưng điều đó sẽ mất nhiều thời

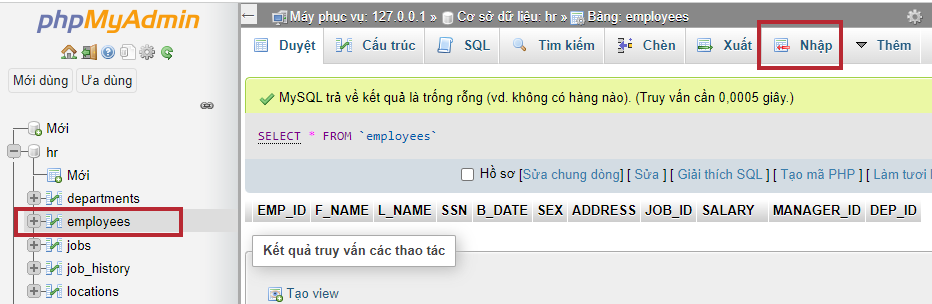
gian. Thay vào đó chúng ta có thể nạp dữ liệu từ các file .CSV.

Hướng dẫn bên dưới giải thích quá trình nạp dữ liệu vào các bảng mà chúng ta đã tạo trước đó, hãy làm theo các bước.

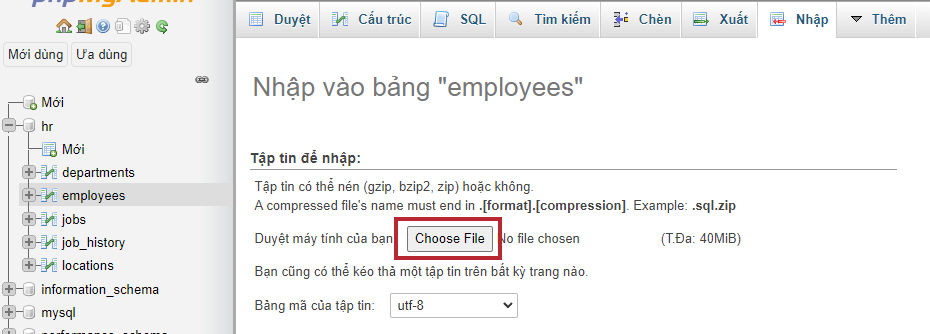
1) Download 5 file nguồn dữ liệu bắt buộc từ trang lab trong khóa học về máy:

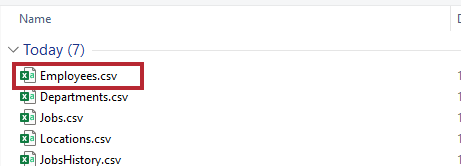
(Employees.csv,Departments.csv,Jobs.csv,JobsHistory.csv,Locations.csv):

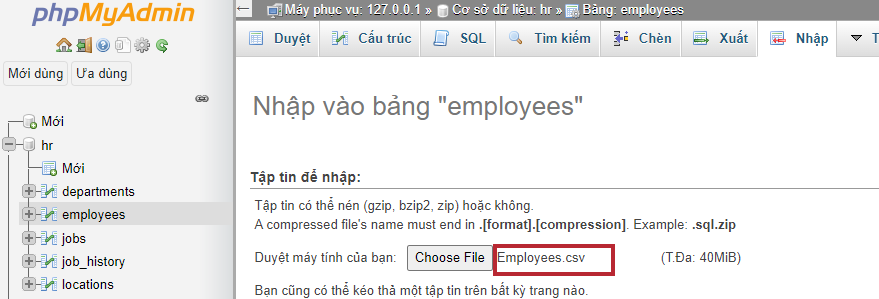
2) Trước tiên, hãy tìm hiểu cách nạp dữ liệu vào bảng Employees mà chúng ta đã tạo trước đó. Từ giao diện chính chúng ta sẽ chọn vào bảng employees, sau đó chọn vào “Nhập”:



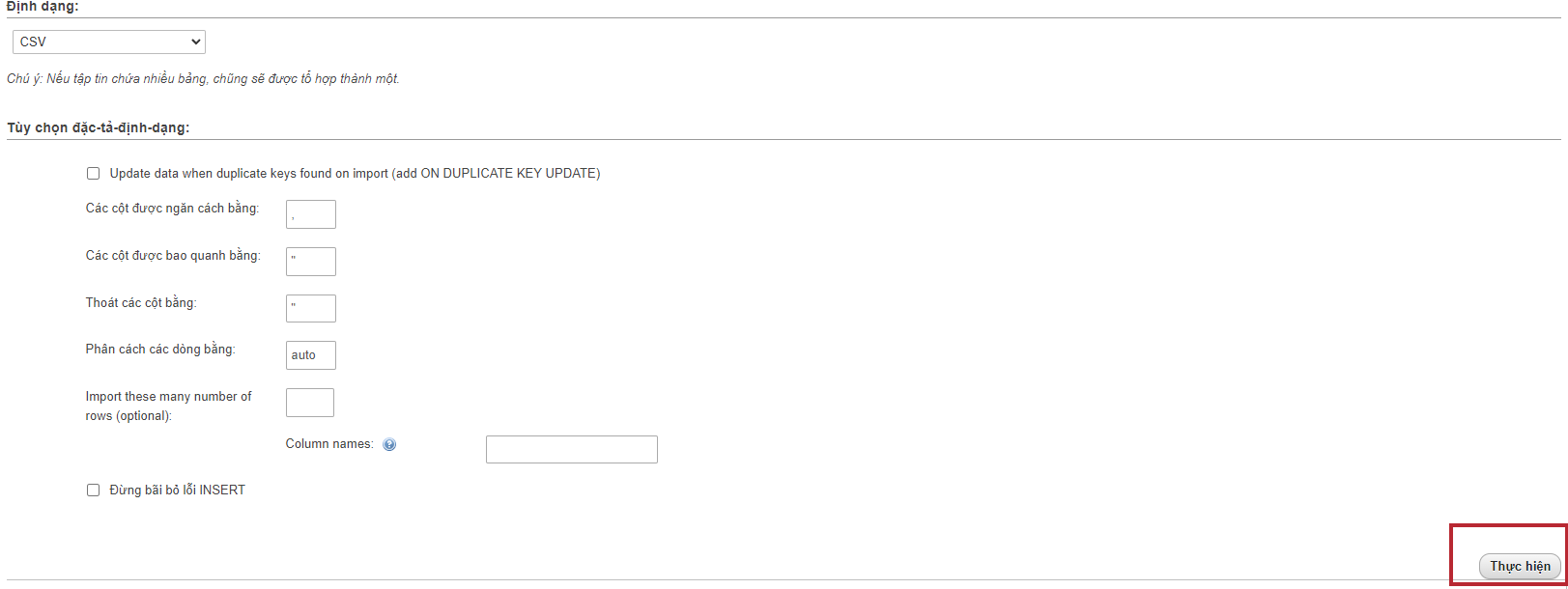
3) Tiếp theo chúng ta sẽ chọn vào “Choose File” sau đó chọn đến file dữ liệu Employees.csv:



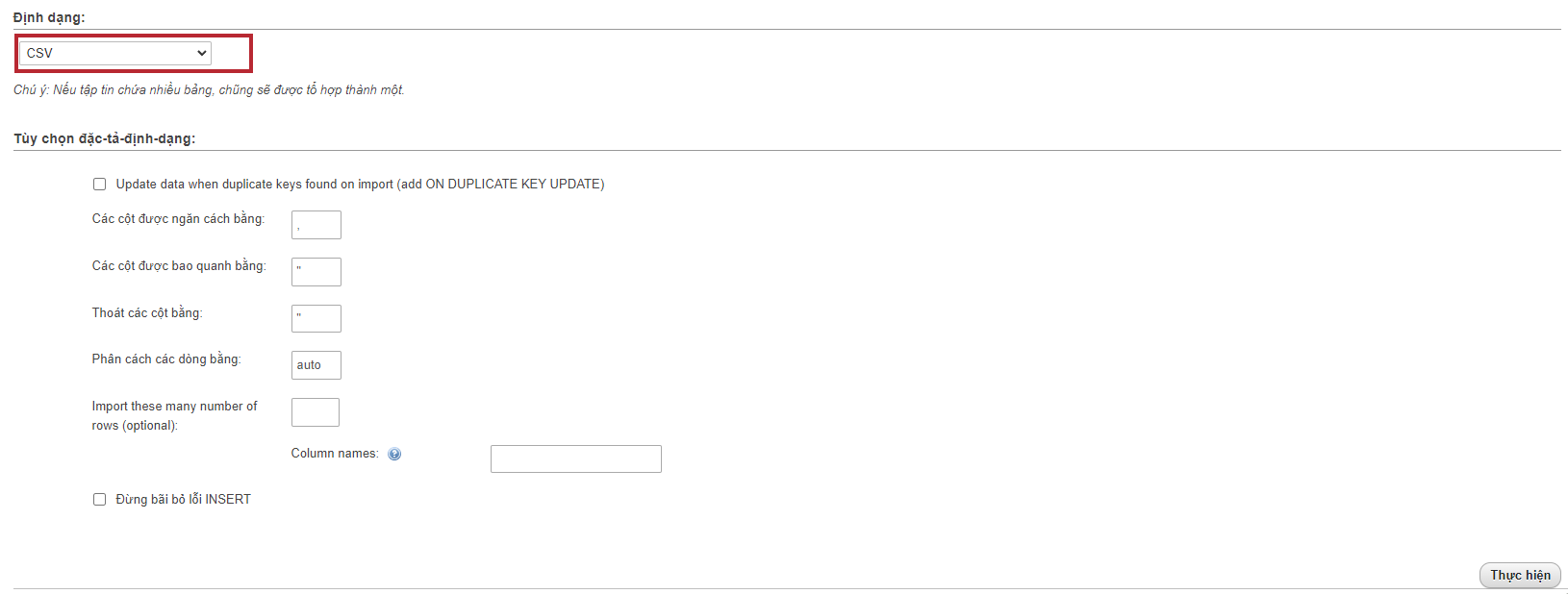




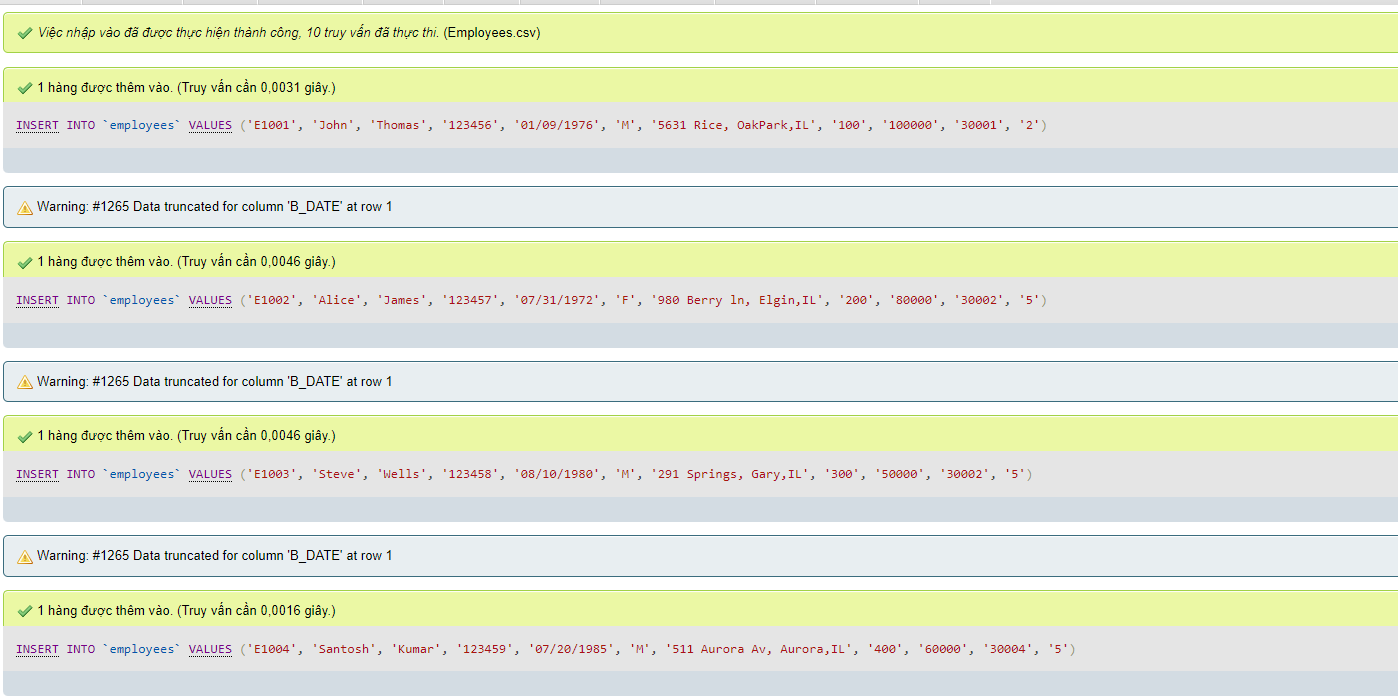
4) Sau khi đã chọn file thành công, chúng ta sẽ kéo xuống cuối trang và thực hiện lệnh nạp bằng cách chọn “Thực hiện”:



Lưu ý ở đây là chúng ta đang thực hiện nạp dữ liệu với được dạng file là .csv

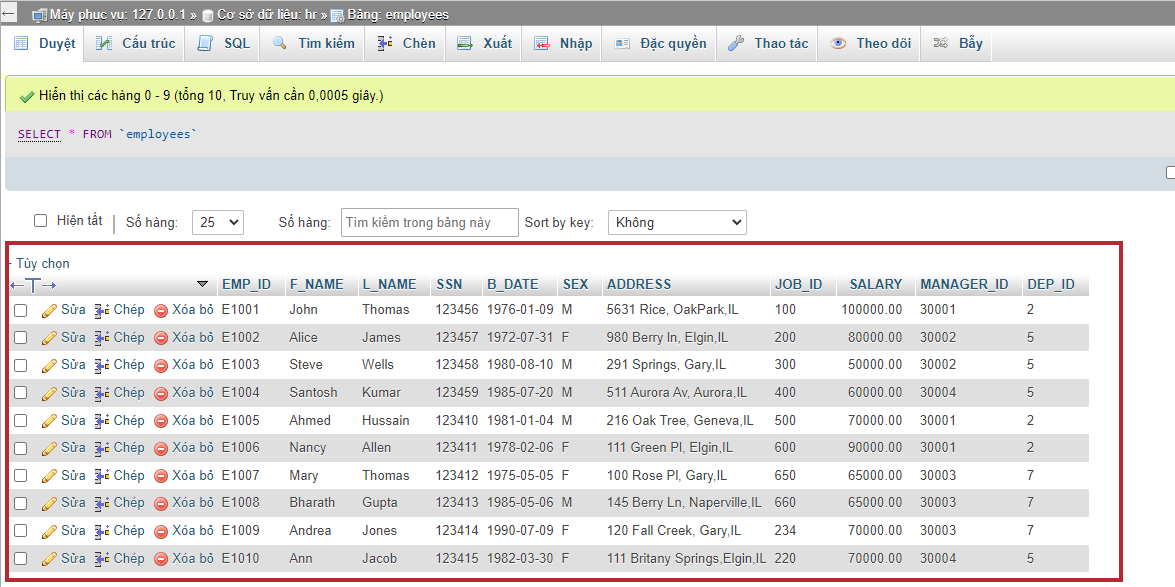


Sau khi thực hiện chúng ta sẽ thấy kết quả nạp thành công như sau:



Nếu việc nạp dữ liệu xảy ra lỗi, thì ở trang này cũng sẽ hiển thị các lỗi. Có thể thấy với bảng employess chúng ta đang thực hiện nạp thành công tất cả các dữ liệu (các hàng).

Chúng ta giờ có thể chọn lại vào bảng employees, và sẽ thấy dữ liệu đã được nạp:



5) Bây giờ, bạn phải nạp 4 bảng còn lại của cơ sở dữ liệu nhân sự - Locations, JobHistory, Jobs, và Departments. Hãy làm theo các

bước ở trên để nạp dữ liệu từ các file nguồn còn lại.

**Câu hỏi 1:** Khi nạp dữ liệu vào bảng JOBS thì có cảnh báo nào không? Có thể làm gì để giải quyết vấn đề này?

Cảnh báo: Data truncated for column 'JOB\_TITLE' at row 1

Gợi ý: Xem dữ liệu được nạp vào bảng này và hãy chú ý đến cột JOB\_TITLE.

**Question 2:** Tất cả các hàng từ file nguồn có nạp thành công vào bảng DEPARTMENT không? Nếu không, theo bạn thì lý do là gì?

Bị trùng lặp giá trị khóa chính

Gợi ý: Nhìn vào cảnh báo. Ngoài ra, hãy lưu ý Primary Key (Khóa chính) cho bảng này và nếu cần thiết hãy thực hiện sửa file dữ liệu để có thể thực hiện nạp thành công.

## Phần III: Soạn và Chạy truy vấn

Bạn đã tạo các bảng cho lược đồ cơ sở dữ liệu hr và cũng đã học cách nạp dữ liệu vào các bảng này. Bây giờ hãy thử và thực hành với một số truy vấn DML (ngôn ngữ thao tác dữ liệu) nâng cao đã được học trong phần này.

Làm theo các bước sau để tạo và chạy các truy vấn được chỉ ra bên dưới:

1) Điều hướng đến cơ sở dữ liệu hr

2) Chọn vào tab SQL và thưc hiện các truy vấn sau

**Truy vấn 1: Truy xuất tất cả nhân viên có địa chỉ ở Elgin, IL**

Gợi ý: Sử dụng toán tử LIKE để tìm các string tương tự

**Truy vấn 2: Truy xuất tất cả nhân viên sinh vào những năm 1970.**

Gợi ý: Sử dụng toán tử LIKE để tìm các string tương tự

**Truy vấn 3: Truy xuất tất cả nhân viên trong phòng ban 5 có mức lương từ 60000 đến 70000.**

Gợi ý: Sử dụng từ khóa BETWEEN cho truy vấn này

**Truy vấn 4A: Truy xuất danh sách nhân viên được sắp xếp theo ID phòng ban.**

Gợi ý: Sử dụng mệnh đề ORDER BY cho truy vấn này

**Truy vấn 4B: Truy xuất danh sách nhân viên được sắp xếp theo thứ tự giảm dần theo ID phòng ban và trong mỗi phòng ban, những nhân viên này được sắp xếp theo họ với thứ tự giảm dần của bảng chữ cái.**

**Truy vấn 5A: Đối với mỗi ID phòng ban, truy xuất số lượng nhân viên trong phòng ban.**

Gợi ý: Sử dụng COUNT (\*) để truy xuất tổng số cột, sau đó dùng GROUP BY

**Truy vấn 5B: Đối với mỗi phòng ban, truy xuất số lượng nhân viên trong phòng ban và mức lương trung bình của nhân viên trong phòng ban.**

Gợi ý: Sử dụng COUNT (\*) để truy xuất tổng số cột và dùng hàm AVG () để tính toán mức lương trung bình, sau đó nhóm kết quả lại

**Truy vấn 5C: Gắn nhãn các cột đã tính trong tập hợp kết quả của Truy vấn 5B là NUM\_EMPLOYEES và AVG\_SALARY.**

Gợi ý: Sử dụng AS “LABEL\_NAME” sau tên cột

**Truy vấn 5D: Trong Truy vấn 5C, hãy sắp xếp tập hợp kết quả Mức lương trung bình.**

Gợi ý: Sử dụng ORDER BY sau GROUP BY

**Truy vấn 5E: Trong Truy vấn 5D, giới hạn kết quả thành ít hơn 4 nhân viên cho các phòng ban.**

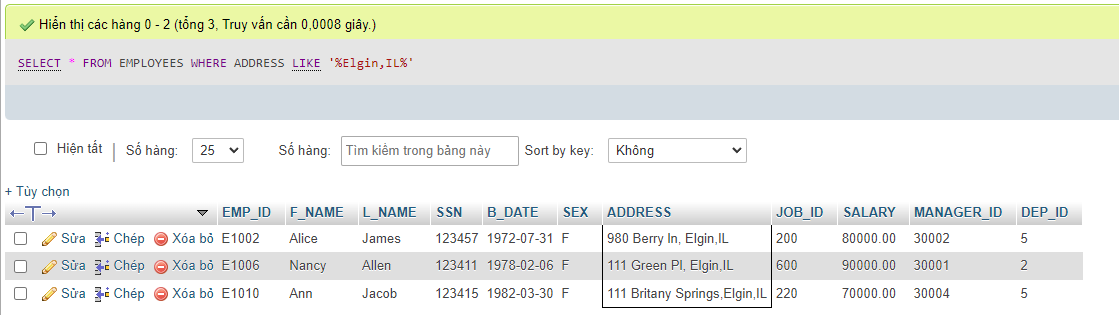
Gợi ý: Sử dụng HAVING sau GROUP BY và sử dụng hàm count () trong mệnh đề HAVING chứ không dùng nhãn cột.

Lưu ý: Mệnh đề WHERE được sử dụng để lọc toàn bộ tập kết quả còn mệnh đề HAVING được sử dụng để lọc kết quả của việc nhóm

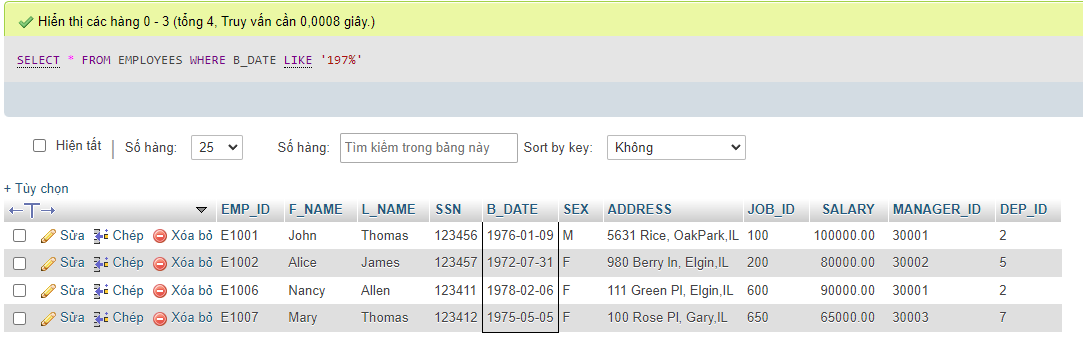
Trong bài Lab này, bạn đã học cách làm việc với các mẫu string, sắp xếp và nhóm các tập kết quả.

Cảm ơn bạn vì đã hoàn thành bài Lab này! Hãy tham khảo lời giải ở dưới đây:

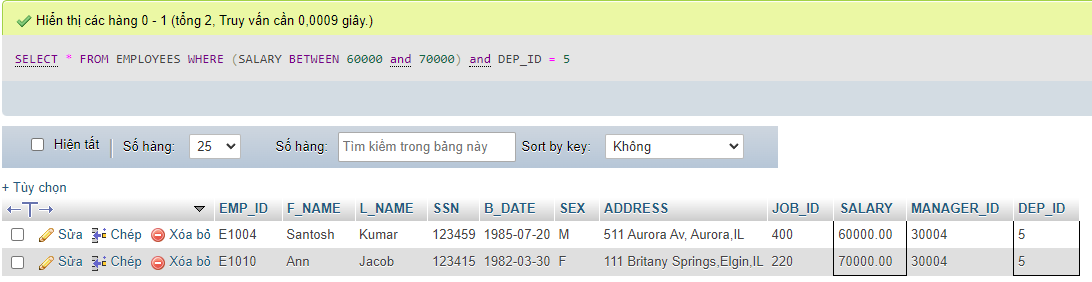
**Truy vấn 1:**



**Truy vấn 2:**



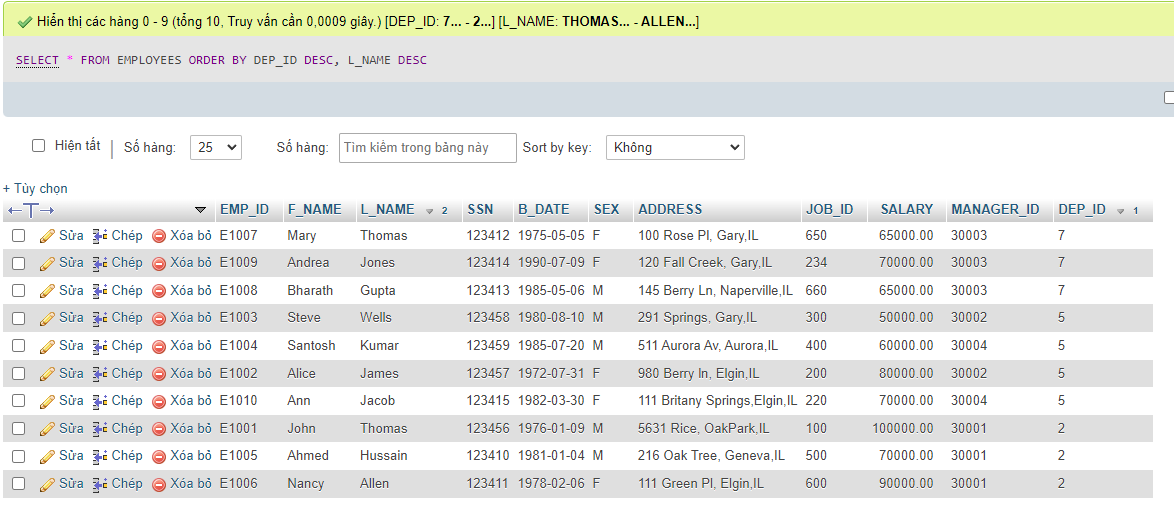
**Truy vấn 3:**



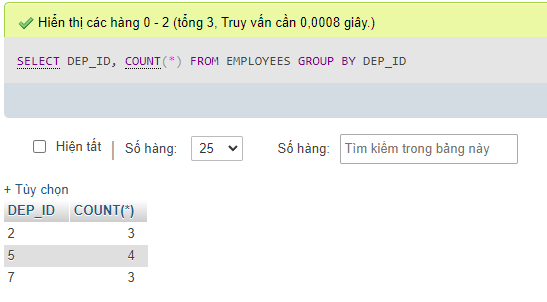
**Truy vấn 4A:**



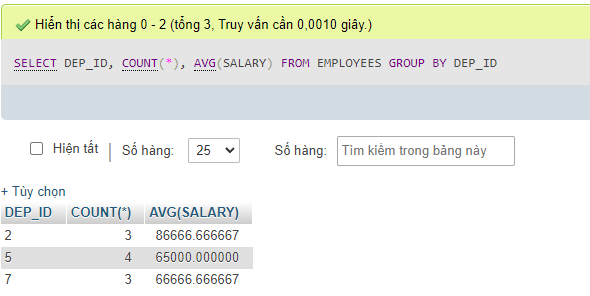
**Truy vấn 4B:**



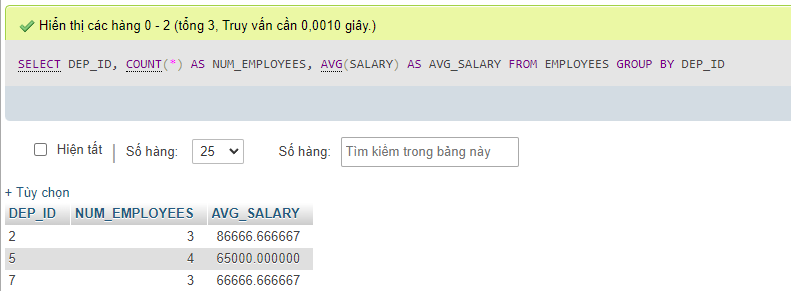
**Truy vấn 5A:**



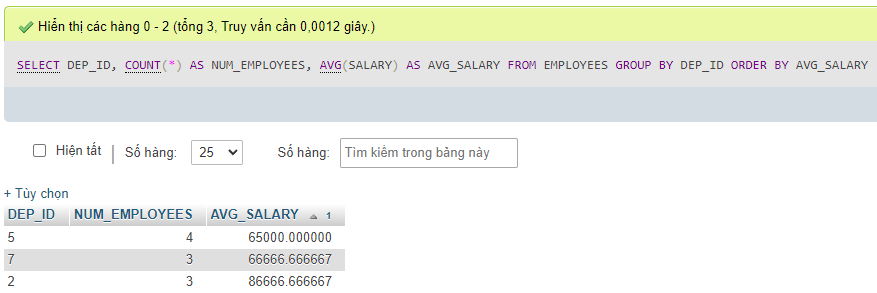
**Truy vấn 5B:**



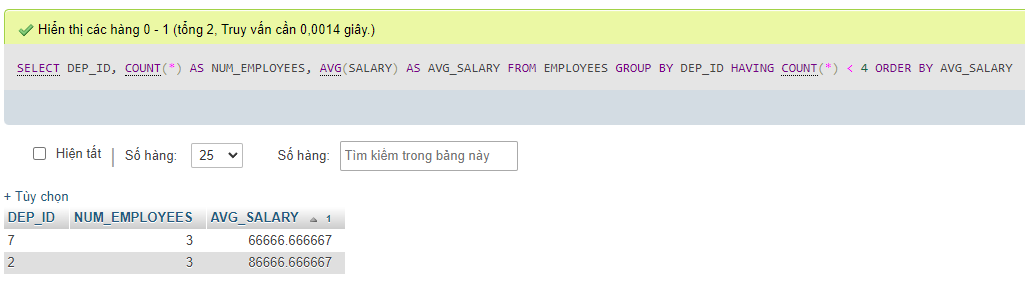
**Truy vấn 5C:**



**Truy vấn 5D:**



**Truy vấn 5E:**



**Tác giả**

**Rav Ahuja**

**© IBM Corporation 2020. All rights reserved.**