

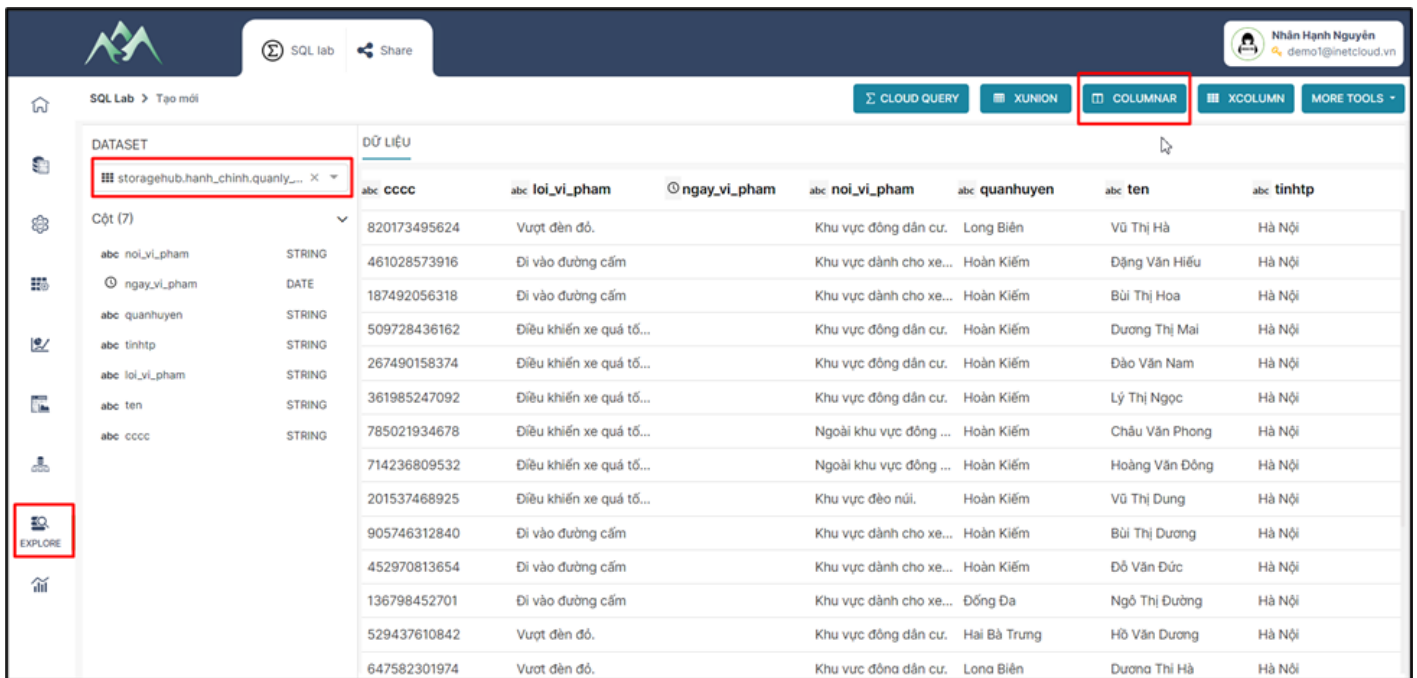
6. COLUMNAR

Mục tiêu COLUMNAR: Chuyển dữ liệu từ định dạng "dòng" sang "cột".

- [Thực hiện trên hệ thống theo các bước hướng dẫn sau:](#)

Thực hiện trên hệ thống theo các bước hướng dẫn sau:

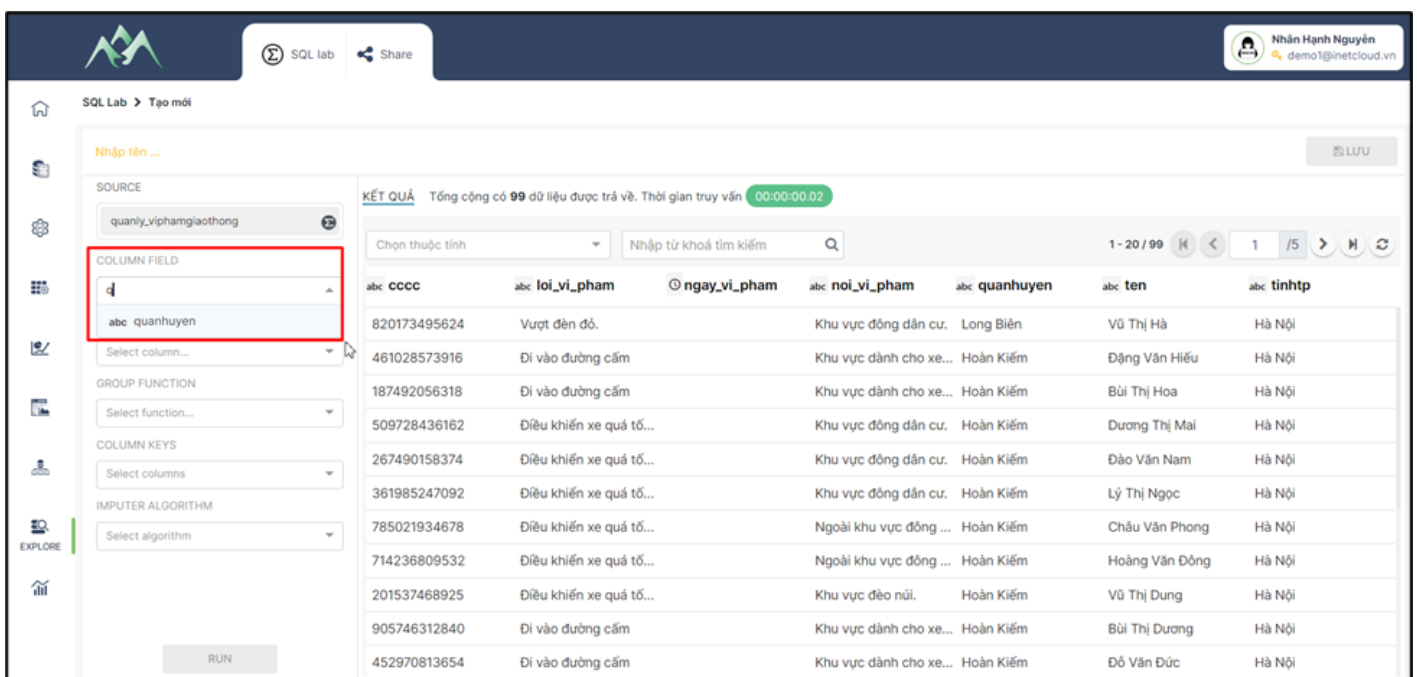
Bước 1: Vào hệ thống chọn module Explore, chọn tập dữ liệu cần tổng hợp và tạo mới một tập dữ liệu tổng hợp từ COLUMNAR.



The screenshot shows the SQL Lab interface. The 'EXPLORE' module is selected in the left sidebar. The 'DATASET' dropdown is set to 'storagehub.hanh_chinh.quanly...'. The 'COLUMNAR' button is highlighted in the top navigation bar. The main area displays a data table with 7 columns and 20 rows of data.

abc cccc	abc loi_vl_pham	ngay_vl_pham	abc noi_vl_pham	abc quanhuyen	abc ten	abc tinhp
820173495624	Vượt đèn đỏ.		Khu vực đông dân cư.	Long Biên	Vũ Thị Hà	Hà Nội
461028573916	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đặng Văn Hiếu	Hà Nội
187492056318	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Hoa	Hà Nội
509728436162	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Dương Thị Mai	Hà Nội
267490158374	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Đào Văn Nam	Hà Nội
361985247092	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Lý Thị Ngọc	Hà Nội
785021934678	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Châu Văn Phong	Hà Nội
714236809532	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Hoàng Văn Đông	Hà Nội
201537468925	Điều khiển xe quá tốc...		Khu vực đèo núi.	Hoàn Kiếm	Vũ Thị Dung	Hà Nội
905746312840	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Dương	Hà Nội
452970813654	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đỗ Văn Đức	Hà Nội
136798452701	Đi vào đường cấm		Khu vực dành cho xe...	Đống Đa	Ngô Thị Đường	Hà Nội
529437610842	Vượt đèn đỏ.		Khu vực đông dân cư.	Hai Bà Trưng	Hồ Văn Dương	Hà Nội
647582301974	Vượt đèn đỏ.		Khu vực đông dân cư.	Lona Biên	Dương Thị Hà	Hà Nội

Bước 2: "COLUMN FIELD" chọn cột dữ liệu cần định dạng "dòng" sang "cột".



The screenshot shows the SQL Lab interface with the 'COLUMN FIELD' dropdown menu open. The 'abc quanhuyen' option is selected and highlighted with a red box. The main area displays the same data table as in the previous screenshot.

Bước 3: "COLUMN VALUE" chọn giá trị cột cần tổng hợp

The screenshot shows the SQL Lab interface. In the 'COLUMN FIELD' section, the 'COLUMN VALUE' dropdown menu is open, and 'loi_vi_pham' is selected. The main table displays data with columns: cccc, loi_vi_pham, ngay_vi_pham, noi_vi_pham, quanhuyen, ten, tinhtp. The table content is as follows:

abc cccc	abc loi_vi_pham	ngay_vi_pham	noi_vi_pham	abc quanhuyen	abc ten	abc tinhtp
820173495624	Vượt đèn đỏ.		Khu vực đông dân cư.	Long Biên	Vũ Thị Hà	Hà Nội
461028573916	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đặng Văn Hiếu	Hà Nội
187492056318	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Hoa	Hà Nội
509728436162	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Dương Thị Mai	Hà Nội
267490158374	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Đào Văn Nam	Hà Nội
361985247092	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Lý Thị Ngọc	Hà Nội
785021934678	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Châu Văn Phong	Hà Nội
714236809532	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Hoàng Văn Đông	Hà Nội
201537468925	Điều khiển xe quá tốc...		Khu vực đồi núi.	Hoàn Kiếm	Vũ Thị Dung	Hà Nội
905746312840	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Dương	Hà Nội
452970813654	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đỗ Văn Đức	Hà Nội

Bước 4: "GROUP FUNCTION" chọn các hàm tính toán, xử lý dữ liệu cho cột giá trị cần tổng hợp.

The screenshot shows the SQL Lab interface. In the 'COLUMN VALUE' section, the 'GROUP FUNCTION' dropdown menu is open, and 'COUNT' is selected. The main table displays data with columns: cccc, loi_vi_pham, ngay_vi_pham, noi_vi_pham, quanhuyen, ten, tinhtp. The table content is as follows:

abc cccc	abc loi_vi_pham	ngay_vi_pham	noi_vi_pham	abc quanhuyen	abc ten	abc tinhtp
820173495624	Vượt đèn đỏ.		Khu vực đông dân cư.	Long Biên	Vũ Thị Hà	Hà Nội
461028573916	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đặng Văn Hiếu	Hà Nội
187492056318	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Hoa	Hà Nội
509728436162	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Dương Thị Mai	Hà Nội
267490158374	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Đào Văn Nam	Hà Nội
361985247092	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Lý Thị Ngọc	Hà Nội
785021934678	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Châu Văn Phong	Hà Nội
714236809532	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Hoàng Văn Đông	Hà Nội
201537468925	Điều khiển xe quá tốc...		Khu vực đồi núi.	Hoàn Kiếm	Vũ Thị Dung	Hà Nội
905746312840	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Dương	Hà Nội
452970813654	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đỗ Văn Đức	Hà Nội

Lưu ý:

Để hiểu và sử dụng từng hàm có thể tham khảo tài liệu chi tiết theo đường dẫn: ...

Bước 5: "COLUMN KEYS" chọn cột trường thuộc tính là khóa xử lý

SQL Lab > Tạo mới

Nhập tên ...

SOURCE: quanly_viphangiaothong

COLUMN FIELD: abc quanhuyen

COLUMN VALUE: abc loi_vl_pham

GROUP FUNCTION: COUNT

COLUMN KEYS: x loi_vl_pham x

IMPUTER ALGORITHM: Select algorithm

RUN

KẾT QUẢ Tổng cộng có 99 dữ liệu được trả về. Thời gian truy vấn 00:00:00.02

abc cccc	abc loi_vl_pham	ngay_vl_pham	abc noi_vl_pham	abc quanhuyen	abc ten	abc tinhtp
820173495624	Vượt đèn đỏ.		Khu vực đông dân cư.	Long Biên	Vũ Thị Hà	Hà Nội
461028573916	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đặng Văn Hiếu	Hà Nội
187492056318	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Hoa	Hà Nội
509728436162	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Dương Thị Mai	Hà Nội
267490158374	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Đào Văn Nam	Hà Nội
361985247092	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Lý Thị Ngọc	Hà Nội
785021934678	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Châu Văn Phong	Hà Nội
714236809532	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Hoàng Văn Đông	Hà Nội
201537468925	Điều khiển xe quá tốc...		Khu vực đồi núi.	Hoàn Kiếm	Vũ Thị Dung	Hà Nội
905746312840	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Dương	Hà Nội
452970813654	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đỗ Văn Đức	Hà Nội

Bước 6: "IMPUTER ALGORITHM" chọn thuật toán điền giá trị thiếu.

SQL Lab > Tạo mới

Nhập tên ...

SOURCE: quanly_viphangiaothong

COLUMN FIELD: abc quanhuyen

COLUMN VALUE: abc loi_vl_pham

GROUP FUNCTION: COUNT

COLUMN KEYS: x loi_vl_pham x

IMPUTER ALGORITHM: interpolate

RUN

KẾT QUẢ Tổng cộng có 99 dữ liệu được trả về. Thời gian truy vấn 00:00:00.02

abc cccc	abc loi_vl_pham	ngay_vl_pham	abc noi_vl_pham	abc quanhuyen	abc ten	abc tinhtp
820173495624	Vượt đèn đỏ.		Khu vực đông dân cư.	Long Biên	Vũ Thị Hà	Hà Nội
461028573916	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đặng Văn Hiếu	Hà Nội
187492056318	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Hoa	Hà Nội
509728436162	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Dương Thị Mai	Hà Nội
267490158374	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Đào Văn Nam	Hà Nội
361985247092	Điều khiển xe quá tốc...		Khu vực đông dân cư.	Hoàn Kiếm	Lý Thị Ngọc	Hà Nội
785021934678	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Châu Văn Phong	Hà Nội
714236809532	Điều khiển xe quá tốc...		Ngoài khu vực đông ...	Hoàn Kiếm	Hoàng Văn Đông	Hà Nội
201537468925	Điều khiển xe quá tốc...		Khu vực đồi núi.	Hoàn Kiếm	Vũ Thị Dung	Hà Nội
905746312840	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Bùi Thị Dương	Hà Nội
452970813654	Đi vào đường cấm		Khu vực dành cho xe...	Hoàn Kiếm	Đỗ Văn Đức	Hà Nội

Trong đó:

- o Thuật toán nội suy (interpolation) được sử dụng để ước lượng giá trị tại các điểm trung gian dựa trên các giá trị đã biết của một tập dữ liệu.
- o Thuật toán Simple được dùng để xử lý các bài toán như tìm kiếm tuyến tính, sắp xếp nổi bọt, hoặc tính giai thừa.
- o Thuật toán KNN là một thuật toán dựa trên khoảng cách để phân loại hoặc dự đoán giá trị của một điểm dữ liệu mới dựa trên các điểm dữ liệu đã biết trong không gian đa chiều.

- o Thuật toán Payesian là một trong những thuật toán phân loại phổ biến nhất trong học máy, đặc biệt là đối với các bài toán phân loại văn bản và lọc thư rác.

Lưu ý: Để hiểu và sử dụng từng hàm có thể tham khảo tài liệu chi tiết theo đường dẫn: ...

Bước 7: Ấn RUN để xem kết quả dữ liệu thêm mới vào tập dữ liệu.

The screenshot shows the SQL Lab interface. On the left, there are configuration options for SOURCE (quanly_viphangiaothong), COLUMN FIELD (abc quanhuyen), COLUMN VALUE (abc loi_vl_pham), GROUP FUNCTION (COUNT), COLUMN KEYS (loi_vl_pham), and IMPUTER ALGORITHM (interpolate). A red box highlights the 'RUN' button at the bottom left. The main area displays a table with 99 rows and 7 columns: abc cccc, abc loi_vl_pham, ngay_vl_pham, abc noi_vl_pham, abc quanhuyen, abc ten, and abc tinhtp. The table contains various data points, including phone numbers and addresses. A 'LƯU' button is visible in the top right corner.

Bước 8: Nhập tên cho tập dữ liệu mới được tổng hợp, chọn Lưu thông tin tập dữ liệu được tổng hợp bằng công cụ COLUMNAR

The screenshot shows the SQL Lab interface with the 'COLUMNAR' tool selected. The 'SOURCE' field is set to 'quanly_viphangiaothong'. The 'COLUMNAR' tool is configured with 'abc quanhuyen' as the COLUMN FIELD and 'abc loi_vl_pham' as the COLUMN VALUE. The 'GROUP FUNCTION' is set to 'COUNT'. The 'COLUMN KEYS' are 'loi_vl_pham'. The 'IMPUTER ALGORITHM' is 'interpolate'. A red box highlights the 'columnar_vpgt' text input field. Another red box highlights the 'LƯU' button in the top right corner. The main area displays the same table as in the previous screenshot. A 'RUN' button is visible at the bottom left.