

# 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. On the left sidebar, the 'EXPLORE' icon is highlighted with a red box. The main area displays a dataset named 'storagehub.hanh\_chinh.quanly...' with a table of data. The table has columns: 'cccc', 'loi\_vl\_pham', 'ngay\_vl\_pham', 'noi\_vl\_pham', 'quanhuyen', 'ten', and 'tinhtp'. The data rows show various entries related to vehicle accidents and locations in Hanoi.

cccc	loi_vl_pham	ngay_vl_pham	noi_vl_pham	quanhuyen	ten	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
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.	Long 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 menu lists the columns from the dataset: 'cccc', 'loi\_vl\_pham', 'ngay\_vl\_pham', 'noi\_vl\_pham', 'quanhuyen', 'ten', and 'tinhtp'. The 'quanhuyen' column is highlighted with a red box, indicating it is the selected field for transformation.

cccc	loi_vl_pham	ngay_vl_pham	noi_vl_pham	quanhuyen	ten	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 3: “COLUMN VALUE” chọn giá trị cột cần tổng hợp

SQL Lab > Tạo mới

Share

Nhân Hạnh Nguyễn

demo1@incloud.vn

Nhập tên ...

LƯU

SOURCE

quanly\_viphamgiaothong

COLUMN FIELD

abc quanhuyen

COLUMN VALUE

abc loi\_vl\_pham

COLUMN KEYS

Select columns

IMPUTER ALGORITHM

Select algorithm

RUN

KẾT QUẢ

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

Chọn thuộc tính

Nhập từ khóa tìm kiếm

1 - 20 / 99

1 / 5

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 Đồng	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.

SQL Lab > Tạo mới

Share

Nhân Hạnh Nguyễn

demo1@incloud.vn

Nhập tên ...

LƯU

SOURCE

quanly\_viphamgiaothong

COLUMN FIELD

abc quanhuyen

COLUMN VALUE

abc loi\_vl\_pham

GROUP FUNCTION

co

COUNT

COUNTMISSING

SECONDQUARTILE

COUNTDISTINCT

COUNTEXISTING

RUN

KẾT QUẢ

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

Chọn thuộc tính

Nhập từ khóa tìm kiếm

1 - 20 / 99

1 / 5

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 Đồng	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\_viphamgiaothong

COLUMN FIELD: abc quanhuyen

COLUMN VALUE: abc loi\_vi\_pham

GROUP FUNCTION: COUNT

COLUMN KEYS: loi\_vi\_pham

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_vi_pham	© ngay_vi_pham	abc 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 6: "IMPUTER ALGORITHM" chọn thuật toán điền giá trị thiếu.

SQL Lab > Tạo mới

Nhập tên ...

SOURCE: quanly\_viphamgiaothong

COLUMN FIELD: abc quanhuyen

COLUMN VALUE: abc loi\_vi\_pham

GROUP FUNCTION: COUNT

COLUMN KEYS: loi\_vi\_pham

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_vi_pham	© ngay_vi_pham	abc 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

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 sidebar, the 'EXPLORE' tab is active. The main area is divided into two panels. The left panel contains configuration options: SOURCE (quanly\_viphamgiaothong), COLUMN FIELD (abc.quanhuyen), COLUMN VALUE (abc.loivipham), GROUP FUNCTION (COUNT), COLUMN KEYS (loivipham), and IMPUTER ALGORITHM (Interpolate). The right panel displays the 'KẾT QUẢ' (Results) table with 99 rows of data. The 'RUN' button at the bottom left of the configuration panel is highlighted with a red rectangle.

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 after clicking 'RUN'. The 'Nhập tên...' (Enter name) field at the top left is now populated with 'columnar\_vpgt' and is highlighted with a red rectangle. The 'LƯU' (Save) button at the top right is also highlighted with a red rectangle. The 'KẾT QUẢ' (Results) table remains visible on the right, showing the same 99 rows of data.