

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

KHOA CÔNG NGHỆ THÔNG TIN

BỘ MÔN HỆ THỐNG THÔNG TIN

MÔN HỌC TRỰC QUAN HÓA DỮ LIỆU

TRỰC QUAN HÓA

TÌNH TRẠNG TAI NẠN GIAO THÔNG

TẠI CÁC TUYẾN ĐƯỜNG LỚN Ở LOS ANGELES

GIAI ĐOẠN 2016-2020

GV hướng dẫn: TS. Lê Thị Nhàn

HCM, 28 tháng 11 năm 2021

Nhóm 1:

- 20C12007 Trần Đình Lâm
- 20C12030 Huỳnh Lâm Phú Sĩ
- 20C12031 Huỳnh Ngọc Đan Thanh

Tổng quan

Phát biểu bài toán

Phân tích dữ liệu

Các bước thực hiện

Các trực quan hóa

Nhận xét & Đánh giá

Phát biểu bài toán

- Lĩnh vực: Quản lý và điều hành giao thông đô thị
- Phạm vi khu vực: Tại các tuyến đường lớn ở thành phố Los Angeles, bang California
- Đối tượng sử dụng công cụ: *Các chuyên viên nghiên cứu về an toàn giao thông*
- Mục tiêu: Cung cấp các trực quan hóa giúp **phân tích nguyên nhân** nào ảnh hưởng nhiều nhất đến tình trạng va chạm xe
- Input: Dữ liệu ghi nhận các vụ va chạm xe ở California, giai đoạn 2016-2020
- Output: Các biểu đồ trực quan giúp đánh giá từng loại yếu tố

Mô tả dữ liệu

Số lượng mẫu: 321,000

Không gian/Thời gian	Con người	Phương tiện
Kinh độ - Vĩ độ	Độ tuổi	Loại phương tiện
Tên tuyến đường	Giới tính	Hãng sản xuất
Năm/Tháng/Ngày/Giờ	Vai trò	Năm sản xuất
Điều kiện thời tiết	Mức độ tỉnh táo (Say rượu/dùng thuốc/buồn ngủ)	
Điều kiện ánh sáng	Mức độ tổn thương	

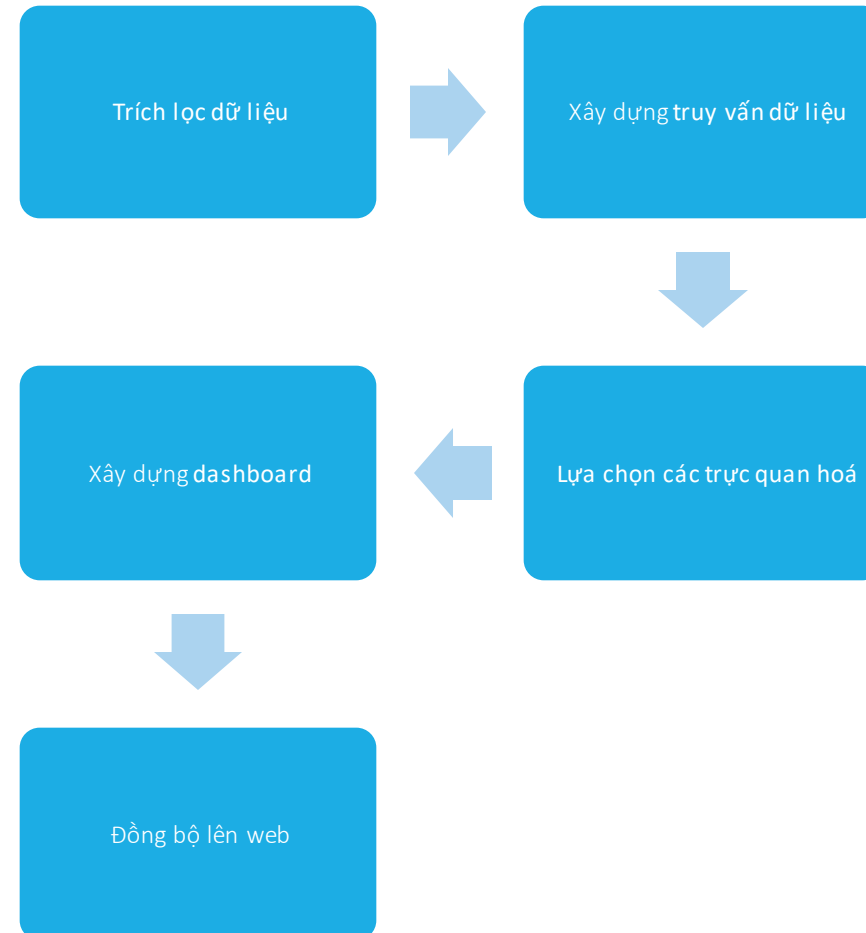
Mục tiêu của các trực quan hóa

Chia nhỏ bài toán thành nhiều câu hỏi:

1. Thời tiết theo mùa có làm tăng khả năng gây tai nạn?
2. Tỷ lệ tử vong hoặc thương tật có cao hay không?
3. Các loại phương tiện gây tai nạn có sự khác nhau hay không?
4. Thời gian nào trong ngày tai xế thiếu sự tỉnh táo nhất?
5. Trong thời gian thiếu tỉnh táo, mức độ tai nạn như thế nào?

Các bước thực hiện:

1. Trích lọc dữ liệu ở Los Angeles, trong phạm vi các tuyến đường lớn, giai đoạn 2016-2020
2. Xây dựng các thành phần giúp truy vấn dữ liệu theo tuyến đường, theo thời gian,....
3. Lựa chọn các trực quan hoá riêng lẻ, mỗi trực quan hóa ứng với một hay vài yếu tố có liên hệ gần với nhau
4. Tổng hợp các trực quan phù hợp nhất thành Dashboard
5. Hiển thị Dashboard lên nền tảng web





Demo các trực quan hóa

Nhận xét & Đánh giá

Nhận xét:

1. Yếu tố phương tiện
2. **Yếu tố con người**
3. Yếu tố ngoại cảnh

Đánh giá:

1. Sử dụng được nhiều thuộc tính, biến hình ảnh
2. Giải thích được hiện tượng
3. Trực quan hóa còn đơn giản
4. Lạm dụng khá nhiều biểu đồ cột



Thank you

Q&A

Tài liệu tham khảo

- Yuan Zhu, Sami Demiroglu, Kaan Ozbay, Kun Xie, Hong Yang, Di Sha, "SAVE-T: Safety Analysis Visualization and Evaluation Tool", *Journal of Advanced Transportation*, vol. 2021, Article ID 5545117, 12 pages, 2021. <https://doi.org/10.1155/2021/5545117>
- Transportation Injury Mapping System (TIMS), *Safe Transportation Research and Education Center, University of California, Berkeley*. 2021
- California Traffic Collision Data from SWITRS. <http://iswitrs.chp.ca.gov/>

Phụ lục

Hãng Toyota là hãng có phương tiện lưu hành lớn nhất, do đó số vụ tai nạn có liên quan cũng chiếm số lượng lớn, tuy nhiên khi so sánh số vụ tai nạn nghiêm trọng (bị thương nặng, có người chết) thì Honda lại vượt trội hơn, sau đó là Toyota.

Harley Davision có số phương tiện lưu hành không lớn nhưng số vụ tai nạn nghiêm trọng cũng ở mức cao.

Các loại hình di chuyển có liên quan đến các vụ tai nạn có sự thay đổi khá rõ trong tuần.

Các loại hình di chuyển bằng phương tiện giao thông (đi xe đạp, đi xe máy và ô tô) thường có số vụ tai nạn trong tuần cao hơn so với cuối tuần.

Tuy nhiên người đi bộ lại có xu hướng dễ gặp tai nạn hơn vào cuối tuần.