

1 Mô tả

Mỗi nhóm thực hiện quá trình phân tích dữ liệu trên một tập dữ liệu liên quan đến ***tai nạn giao thông*** để tìm ra 5-7 insight quan trọng. Ví dụ:

- Đất nước/khu vực nào xảy ra tai nạn giao thông thường xuyên nhất? Giải thích?
- So sánh cách các khu vực đối phó với tai nạn giao thông?
- Loại hình tai nạn giao thông nào gây tổn thất nặng nề nhất?
- Loại phương tiện nào thường xuyên gây tai nạn nhất?
- (*) Xu hướng tăng/giảm của số vụ tai nạn giao thông và dự đoán trong tương lai gần.
- (**) Gợi ý giải pháp giảm tai nạn giao thông và phân tích cụ thể hiệu quả ước tính của giải pháp đó (lợi ích & thiệt hại).

2 Dữ liệu

Bạn có thể sử dụng bất kì tập dữ liệu nào mà bạn tìm thấy, với các điều kiện:

- Liên quan đến tai nạn giao thông.
- Có mô tả dữ liệu.
- Là dữ liệu thực tế.
- Là dữ liệu hợp pháp.
- Có thể truy cập công khai.

Keywords: traffic accident, road accident, car accident, car crash, vehicle crash, etc.

3 Nộp bài

Mỗi nhóm tạo 1 thư mục với tên là **MSSV1-MSSV2-...** bao gồm các file dưới đây:

- Một notebook (**Analysis.ipynb**) chứa các phân tích của nhóm. Hãy nhớ *reset kernel* và *run all* notebook trước khi nộp để đảm bảo notebook chạy được và ra kết quả mong muốn. Bạn có thể sử dụng thêm bất cứ công cụ nào khác đã được học (như Power BI hay Tableau).

- Một thư mục tên **Data** chứa các file dữ liệu liên quan. Nếu dữ liệu quá lớn (hơn 20mb) thì bạn lưu link của data set vào một file text và bỏ vào thư mục này.

Sau đó nén thư mục này thành file **MSSV1-MSSV2-....zip** và nộp trên Moodle. Mỗi nhóm chỉ cần 1 người nộp.

4 Đánh giá

- Mức độ hữu ích của các thông tin rút trích được (40%+ **10%**).
- Trình bày notebook trực quan và thuyết phục (40%).
- Sử dụng hiệu quả các công cụ đã học (20%).

5 Hình thức và Quy định

- Mỗi nhóm 3-5 người.
- Báo cáo đồ án bằng hình thức vấn đáp.
- Mỗi nhóm chuẩn bị 1 slide tóm tắt lại các insight tìm được.
- Các nhóm giống bài nhau hoặc giống nguồn có sẵn (50% nội dung notebook trở lên, không phân biệt ngôn ngữ) nhận **0 điểm môn học**.