**Báo cáo kết quả thực hành phân tích dữ liệu tuần 4**

1. Họ và tên: Nguyễn Hữu Vượt

Nhóm: 3

MSSV : 19020489

1. Nội dung thực hành

* Tham khảo bộ dữ liệu về giá ô tô trên trang kaggle <https://www.kaggle.com/dronax/car-prices-dataset>
* Áp dụng thư viện Pandas để lọc và tách dữ liệu phân tích
  + Tách dữ liệu từ các cột : ‘price, year, color, brand, model’
  + Dữ liệu đc phân tính xoay quanh chủ thể chính là các hang sản xuất ô tô
  + Các thông tin tách được từ bộ data:
    - Các màu đc các hãng sản xuất -> màu ưa thích của người tiêu dùng
    - Mặt bằng chung về giá của các hãng xe và số lượng xe đã sản xuất
    - Độ chiếm lĩnh thị phần ô tô của từng hãng trong giai đoạn được xét
* Khó khăn:
  + Chưa tận dụng hết được các thông tin trong bộ dữ liệu
  + Biểu diễn đồ thị trực quan dữ liệu chưa hợp lí
  + Em không dùng bộ data đã crawl được từ các tuần trước vì một số phần của data bị missing dữ liệu, và chưa tìm hiểu được mình sẽ lấy được thông tin gì từ bộ data đó