**PHÂN TÍCH MỨC TIÊU HAO NHIÊN LIỆU**

Dữ liệu ‘mpg.csv’

|  |  |  |
| --- | --- | --- |
| 1 | **mpg** | Miles/(US) gallon số gallon xăng trên 1 dặm |
| 2 | **cylinders** | Number of cylinders số lượng xilanh |
| 3 | **displacement** | Displacement (cu.in.) |
| 4 | **horsepower** | Gross horsepower công suất theo mã lực |
| 5 | **weight** | Weight (1000 lbs) trọng lượng (pound) |
| 6 | **acceleration** | Gia tốc |
| 7 | **model\_year** | Năm sản xuất |
| 8 | **origin** | Xuất xứ |
| 9 | **name** | Tên dòng xe |

1. Đọc dữ liệu từ file ‘mpg.csv’. Hiển thị 10 dòng đầu tiên.
2. Tìm hiểu dữ liệu:
   1. Trong tập dữ liệu, thuộc tính nào định tính, thuộc tính nào định lượng
   2. Hãy cho biết trung bình và trung vị của các thuộc tính: **cylinders, horsepower, weight**
   3. Cho biết miền giá trị và miền phân vị của thuộc tính mpg
3. Hãy so sánh mức tiêu hao nhiên liệu trung bình của các dòng xe của Mỹ, Nhật, và Châu Âu
4. Hãy so sánh công suất trung bình của các dòng xe của Mỹ, Nhật, và Châu Âu
5. Tính tỷ lệ của các dòng xe Mỹ, Nhật, Châu Âu trong bộ dữ liệu.
6. Vẽ đồ thị histogram trọng lượng của các xe.
7. Vẽ đồ thị boxplot so sánh công suất của các dòng xe thập niên 70 và thập niên 80
8. Bằng cách chọn biểu đồ phù hợp bạn hãy khảo sát mức tiêu hao nhiên liệu (mpg) theo các thuộc tính như
   1. Mức tiêu hao nhiên liệu và số lượng xi lanh
   2. Mức tiêu hao nhiên liệu và công suất theo mã lực
   3. Mức tiêu hao nhiên liệu và gia tốc
   4. Mức tiêu hao nhiên liệu theo năm sản xuất

Lưu ý: sinh viên đặt tên theo định dạng sau: **<hoten>\_<msv>.ipynb** (VD: nguyenminhthu\_20127331.ipynb)