

Họ tên: **Nguyễn Hữu Hiệu**
MSSV: **20520506**

CAR DESCRIPTION

1. Code book

| Thông tin | Nội dung |
|--------------------------|--|
| Tên bộ dữ liệu | Car Evalevaluationuation Database |
| Nguồn thu thập | M. Bohanec and V. Rajkovic: Knowledge acquisition and explanation for multi-attribute decision making. |
| Số thuộc tính | 7 |
| Số mẫu dữ liệu | 1728 |
| Thông tin các thuộc tính | 1) buying (int): v-high, high, med, low 2) maint (int): v-high, high, med, low 3) doors (int): 2, 3, 4, 5-more 4) persons (int): 2, 4, more 5) lug_boot (int): small, med, big 6) safety (int): low, med, high 7) class (int): unacc, acc, good, vgood |
| Thông tin tác giả | Marko Bohanec |

2. Raw data

Raw data gồm các file **car.c45-names**, **car.data** và **car.names**

3. Tidy data

Tidy data được lưu thành file **car_evaluation.csv**

4. Instruction list

```
# Import libraries
import os
import time
import pandas as pd

# Insert the directory path in here
path = 'car.data'

col_name = ["buying", "maint", "doors", "persons", "lug_boot", "safety", "class"]

# Add column names into dataset
df = pd.read_csv(path, names=col_name)

print('Task finished!')

df.to_csv('car_evaluation.csv', index=False)
```