

# Descriptive Statistics

Dataset: 'df\_car'

1. Load dataset all\_cars.csv yang telah Anda buat, simpan dalam variabel bernama 'df\_car'.
  - a. Jika ada kolom bernama 'Unnamed: 0', ubah nama kolom tersebut menjadi 'id'!
  - b. Jika tidak ada, lakukan reset index. Kemudian kolom paling kiri bernama 'index' ubah menjadi 'id'!
2. Tampilkan deskripsi statistik dari dataset df\_car.
3. Kita mengenal bahwa Variabel dibagi menjadi 2, yaitu variabel Kualitatif dan Kuantitatif.
  - Var. Kualitatif dibagi menjadi 2, yaitu: Nominal dan Ordinal
  - Var. Kuantitatif dibagi menjadi 2, yaitu: Continue dan Discrete
  - a. Kolom mana saja yang menurut Anda merupakan var. kualitatif? Mana yang Nominal dan Ordinal?
  - b. Kolom mana saja yang menurut Anda merupakan var. kuantitatif? Mana yang Continue dan Discrete?
4. Apa yg dimaksud dengan Measure of Central Tendency? Tampilkan Measure of Central Tendency dari kolom 'model' dan 'price'! Central Tendency apa yang cocok untuk kolom 'model' dan 'price'?
5. Apa yang dimaksud dengan Measure of Spread? Tampilkan Measure of Spread dari kolom 'price'!
6. Apakah ada Outlier pada kolom 'price'? Kalau ada, berapa banyak data yang merupakan outliers?
7. Variabel/kolom apa saja yang memiliki distribusi skewed ke kanan?
8. Nilai Mean dan Median dari kolom 'mileage' berbeda jauh. Coba buktikan apakah kolom tersebut berdistribusi normal atau tidak!
9. Grafik apa yang cocok untuk menampilkan visualisasi:
  - brand vs price
  - price vs mileage
  - frequency distribution of mileage