

**Formation : Python pour la science des données & machine learning**  
**Formateur : Massar Mahamat Ali**

# Pandas: Exploration des données



On se propose d'intégrer des analyses avancées sur un ensemble données en utilisant les opérations de sélection, tri, groupement etc.

Les données source de notre étude se trouvent dans le fichier : “Automobiles.csv”.

La source de données contient des colonnes non significatives à supprimer pour ne garder que les données suivantes : "index","company","body-style","wheel-base", "length", "engine-type", "num-of-cylinders", "horsepower", "average-mileage", "price".

**Question 1 :** Commencer par vérifier le type de données et afficher les 5 premières lignes du jeu de données.

**Question 2 :** Ensuite, nettoyer les données en éliminant les valeurs NaN et 0 et reporter une nouvelle version dans le même fichier.

**Question 3 :** Afficher le détail de toutes les voitures de marque Toyota

**Question 4 :** Afficher le nom de la compagnie, le nombre de chevaux ainsi que le prix de la voiture la plus chère

### Question 5 : Afficher le nombre de voitures par compagnie

**Question 6 :** afficher pour chaque compagnie le prix de la voiture le plus élevé

**Question 7 :** afficher la moyenne de kilométrage par compagnie

**Question 8 :** trier toutes les voitures par prix