

Filière : 2IA
Semestre 3 : Ingénierie des données
Tuto1 : Visualisation des données

1. Chargement des bibliothèques python.

Plusieurs bibliothèques Python sont nécessaires pour maintenir et organiser les données et la visualisation :

- a. NumPy : permet d'effectuer des calculs numériques avec Python.
 - b. Pandas : est une librairie python qui permet de manipuler facilement des données à analyser.
 - c. Matplotlib : permet de tracer et visualiser des données sous formes de graphiques.
 - d. Seaborn : permet de créer des graphiques statistiques e. Elle est basée sur Matplotlib, et s'intègre avec les structures Pandas. Cette bibliothèque est aussi performante que Matplotlib, mais apporte une simplicité et des fonctionnalités inédites. Elle permet d'explorer et de comprendre rapidement les données.
2. Téléchargez les données de covid'19 (Voir le dossier partager via Teams).
 3. Affichez les 20 premières lignes des données.
 4. Décrivez les colonnes et leurs types.
 5. Les données sont de type numérique ou catégoriel, séparez ces deux types de variables. Quel est le nombre de chaque type ?
 6. Affichez les données (Confirmed, Deaths, Recovered, Active, New cases, New deaths, New recovered) sous format histogramme pour le Maroc.
 7. Visualisez les données de COVID'19 pour les décès dans les 20 premiers pays sous format diagramme à bandes (horizontal, vertical).
 8. Regroupez (grouped) et empiler (stacked) les barres et mapper les Confirmed, Deaths, Recovered sur les axes x et y pour les 10 premiers pays et pour chaque région WHO (WHO Region).
 9. Visualisez les données en utilisant graphiques à secteurs pour tous les pays d'Afrique (Confirmed, Deaths, Recovered).
 10. Affichez les données pour les décès sous format boîtes à moustaches et tracés de violon pour chaque région WHO.

RENDU DU TP : Document électronique à rendre via le formulaire (voir groupe Teams) à la fin de la séance.