

TP n°=0 : Pré traitement des données

Etant donné les données du COVID19 en leur format brute.

Travail demandé :

Partie 1 : Exploration des données

Analyse de forme :

- Analyser la variable cible
- Identifier les lignes et les colonnes
- Identifier les types des variables
- Analyser les valeurs manquantes

Analyse de fond :

- Visualisation de la Target
- Signification des variables
- Relation variable/Target

Partie 2 : Prétraitement des données

- Appliquer une première version pour le prétraitement de données
- Créer un premier modèle et évaluer les résultats obtenus
- Itérer le processus du prétraitement jusqu'à obtenir un model robuste

Partie 3 : démarrer vos projets

- Collaborer en sous-groupes et créer des équipes
- Choisir des données brutes (chaque sous-groupe doit choisir un domaine d'application différent)
- Appliquer les étapes précédentes pour le prétraitement de vos données