

PROJET 03 : Qualité de l'Air & Impact sur la Santé

OBJECTIF DU PROJET

La pollution de l'air tue des millions de personnes. Comprendre le lien entre les polluants et la santé respiratoire peut guider les politiques publiques et les habitudes personnelles.

FICHIER DE DONNÉES

- qualite_air.csv
- 800 lignes
- Inclut des valeurs manquantes, valeurs aberrantes et déséquilibres pour un apprentissage réaliste

TÂCHES À RÉALISER

■ TÂCHE PRINCIPALE

Prédire Admissions_Respiratoires basé sur la qualité de l'air

■ TÂCHES SECONDAIRES (OPTIONNELLES)

1. Classer les jours comme "Sain" vs "Dangereux"
2. Identifier le "Jour le Plus Propre de la Semaine"
3. Corréler Volume_Trafic avec PM2.5
4. Prévoir les niveaux de pollution pour demain

LIVRABLES ATTENDUS

1. **Exploration des Données (EDA)**
 - Analyse statistique descriptive
 - Visualisations des distributions
 - Analyse des corrélations
 - Identification des valeurs manquantes et aberrantes
2. **Ingénierie des Fonctionnalités**
 - Création de nouvelles variables pertinentes

- Transformation des variables existantes
- Encodage des variables catégorielles
- Normalisation/Standardisation si nécessaire

3. **Modélisation**

- Choix et justification du/des modèle(s)
- Entraînement et validation
- Optimisation des hyperparamètres
- Évaluation avec métriques appropriées

4. **Résultats et Interprétation**

- Performance du modèle principal
- Résultats des tâches secondaires (si réalisées)
- Insights et recommandations
- Visualisations des résultats

CONSEILS

- Commencez par une exploration approfondie des données
- L'ingénierie des fonctionnalités est CRUCIALE pour ce projet
- Documentez votre démarche et vos choix
- Comparez plusieurs approches si possible
- Pensez à l'aspect pratique et utilisable de vos résultats

Bon travail ! ■