

Rapport - Analyse des Données de Santé Publique via l'API de l'OMS

Girondin Audric, Duckes Jonathan

1. Objectif du Projet

Développer une application web interactive avec **Streamlit** pour analyser et visualiser des données de santé publique issues de l'API du *Global Health Observatory (GHO)* de l'OMS.

L'application permet :

- Récupération dynamique d'indicateurs de santé.
 - Visualisations interactives : graphiques, cartes géographiques.
 - Machine learning :
 - **Régression linéaire** : prédictions des tendances.
 - **Clustering (KMeans)** : regroupement de pays par performances.
-

2. Présentation des Données

Source et Structure

Données issues de l'API **GHO de l'OMS** sur des indicateurs variés (espérance de vie, mortalité infantile, taux de vaccination). Les colonnes principales :

- **Pays, Continent, Année, Valeur numérique.**

Technologies Utilisées

Python, Streamlit, Plotly, GeoPandas, Scikit-learn, Pandas.

3. Variables et Indicateurs Utilisés

- Variables : Tous les indicateurs de santé ayant le mot "Health".
 - Machine learning :
 - **Régression linéaire** : Coefficient, Intercept, Score R2
 - **Clustering** : Nombre d'éléments par cluster
-

4. Difficultés Rencontrées

1. Données Manquantes

Problème : Certaines années/pays présentent des valeurs absentes. *Solution* : Suppression ou imputation des données.

2. Temps de Réponse

Problème : Temps long pour récupérer des données via l'API. *Solution* : Limiter le nombre d'indicateurs chargés.

3. Interface Utilisateur

Problème : Conception d'une interface intuitive. *Solution* : Ajustements constants pour simplifier l'expérience utilisateur.

5. Idées pour Aller Plus Loin

- **Extension des Indicateurs** : Plus d'indicateurs de santé publique.
 - **Automatisation** : Mise à jour automatique des données via l'API.
 - **Hébergement Cloud** : Déploiement pour faciliter l'accès.
 - **Optimisation** : Algorithmes avancés de machine learning.
 - **Personnalisation** : Options d'analyse et visualisation sur mesure.
-

Conclusion

Ce projet combine des outils modernes pour rendre les données de santé publique accessibles et compréhensibles. Les améliorations futures renforceront son utilité pour les chercheurs, décideurs et le grand public.