

Analyse des données publiques sur la qualité de l'eau

Franklin FLAMENT, Mikael TOM, Thomas JOSEPHY et Noé DEPAEPE

Domaine : Informatique

Groupe : G1A

Mots clés : Base de données, Data Science, Machine Learning, Qualité de l'eau, Géolocalisation, Carte, Graphiques, Python, SQL, Objectif de développement durable, IA

Description du projet

La qualité de l'eau est un attribut essentiel à la fois pour l'ensemble des écosystèmes, ainsi que pour la santé humaine. En effet, une eau polluée peut contenir des substances nocives pour la faune et la flore. En outre, la préservation de la qualité de l'eau est un élément majeur du développement durable puisqu'elle conditionne l'accès à une ressource essentielle pour les générations futures. Par conséquent, dans un monde où répondre à nos besoins sans compromettre ceux des générations futures est une priorité, il est nécessaire de surveiller régulièrement les données relatives à l'eau afin d'assurer sa qualité à court et à long terme.

Ce projet vise à réaliser un programme informatique permettant une exploration dynamique des données publiques relatives à la qualité de l'eau. Il s'inscrit principalement dans le domaine de la *Data Science* (science de données), une discipline de l'informatique qui consiste à collecter, filtrer, traiter, analyser et visualiser des données dans le but d'en extraire les plus pertinentes.

Pour ce faire, les données ont été collectées depuis différentes sources publiques puis triées et nettoyées afin de garder seulement les informations utiles à notre projet. Ces données seront ensuite filtrées à l'aide de requêtes qui permettent d'extraire celles nécessaires pour la visualisation des informations liées à un paramètre spécifique. Cependant, le nombre de données étant toujours conséquent, les parcourir à chaque lancement du programme ralentit ses performances. Ainsi, les résultats de ces recherches ont été enregistrés dans des fichiers spéciaux. Ensuite, grâce au langage de programmation *Python* et ses différentes bibliothèques/librairies, les données ont été transformées en graphiques et cartes interactifs sur un site internet.

Par conséquent, le programme informatique contient des cartes interactives et graphiques sur la qualité et l'exploitation de l'eau, incluant les concentrations chimiques des eaux en Belgique, en Europe et dans le monde, un classement des concentrations par pays, une comparaison avec les seuils des *Water Framework Directive* de l'Union Européenne, l'indice *WEI+*, le pourcentage des sites de baignade jugés comme « d'excellente qualité » et des mesures liées à l'Objectif de Développement Durable 6 : eau propre et assainissement.

En outre, du Machine Learning est utilisé pour estimer des futures données concernant les mesures liées à l'ODD 6. Le Machine Learning est un domaine de l'Intelligence Artificielle qui utilise les mathématiques et les statistiques pour permettre à un ordinateur d'apprendre à partir de données et faire des prédictions. Pour ce faire, nous utilisons une Régression Polynomiale de degré 1 à 4. Un principe mathématique permettant d'essayer de suivre les tendances d'un graphe pour estimer les itérations suivantes ; au plus le degré est élevé, au

plus la courbe sera précise, mais le risque est qu'elle prenne trop en compte les irrégularités pas importantes. Il est important de noter qu'il s'agit de prédictions et que d'autres facteurs comme des imprévus peuvent influencer les résultats réels.

Le but du projet est d'avoir une vision locale mais aussi globale de la problématique, tout en parcourant les différentes époques. La visualisation de données concerne donc la Belgique et l'Europe, mais aussi le monde entier. Il est également important d'aborder la qualité de l'eau avec l'aspect le plus large possible, par conséquent, plusieurs indicateurs variés, ainsi que les statistiques concernant le développement durable sont employés.

IMAGES DU PROGRAMME

Et plein d'autres fonctionnalités !

The image displays the AquaData web application interface. The main section features a map titled "Monitoring Sites In Belgium" with a subtitle "Map of the monitoring sites where the data has been collected". The map shows numerous blue dots representing monitoring sites across Belgium. The interface includes a navigation menu with options: "Monitoring Sites Belgium", "Water Composition in Europe", "Water Explicitor And Rating Water Quality in Europe", "Water Composition Worldwide", and "Water Usage Worldwide". To the right, a stack of overlapping screenshots shows various data visualization components: a world map, a line graph titled "Water Population Index (WPI)", a world map with a color scale, a bar chart titled "Analysis of Water Quality and its Compounds in Belgium", and another bar chart titled "The European Countries with the highest Carbon Index".