

# **DATA ANALYTICS WITH COGNOS**

## **WATER QUALITY ANALYSIS**

### **Problem Statement:**

❖ To develop a comprehensive water quality data analysis system to assess potability, identify deviations from standards, and understand parameter relationships. The system should be easy to use and provide stakeholders with actionable insights to improve water quality management and public health.

### **Project Definition:**

❖ The project involves analyzing water quality data to assess the suitability of water for specific purposes, such as drinking. The objective is to identify potential issues or deviations from regulatory standards and determine water potability based on various parameters. This project includes defining analysis objectives, collecting water quality data, designing relevant visualizations, and building a predictive model.

### **Design Thinking:**

#### **❖ Analysis Objectives:**

- Assess potability: Determine whether the water meets regulatory standards for drinking water.
- Identify deviations from standards: Identify water samples that exceed regulatory standards for any of the key parameters.
- Understand parameter relationships: Identify how different parameters correlate with each other and how they affect water potability.

#### **❖ Data Collection:**

- Collect water quality data for the key parameters (pH, hardness, solids, etc.) from multiple sources (government agencies, water utilities, and NGOs).
- Gather the data in a consistent format to make it easier to analyze and visualize.
- Clean and prepare the data by removing outliers, imputing missing values, and transforming it into a suitable format for machine learning.

#### **❖ Visualization Strategy:**

- Use histograms, scatter plots, and heatmaps to visualize the distribution of each parameter, correlations between parameters, and relationships between parameters and water potability.
- Use interactive visualizations to allow users to explore the data in more detail.

## ❖ Predictive Modeling:

- Choose a machine learning algorithm (random forests, support vector machines, or gradient boosting machines) to predict water potability.
- Select the most important features (pH, hardness, and solids) for predicting water potability.
- Train a predictive model on historical data.
- Evaluate the predictive model on new data to see how well it predicts water potability.

## Conclusion:

- The overall goal of this project is to develop a comprehensive water quality data analysis system that can be used to improve water quality management and public health. By using design thinking, we can develop a system that is tailored to the needs of our stakeholders and that provides them with the insights they need to make informed decisions.
- The system will include a visualization strategy to communicate the analysis results to stakeholders in a clear and informative way. It will also include a machine learning model to predict water potability based on the key parameters.
- The system will be evaluated on its usability, accuracy, and usefulness for improving water quality management and public health. We will use the feedback from our stakeholders to refine the system and make it even better.
- We believe that this project has the potential to make a significant contribution to improving water quality for everyone. By providing stakeholders with the insights they need to make informed decisions, we can help to reduce water pollution and protect public health.