Phase 1: Project Definition and Design Thinking
Project Definition: The project involves analyzing water quality data to assess the suitability of water for specific purposes, such as drinking,etc.

Water Quality Analysis

Problem Description:

Many communities and ecosystems rely on access to clean and safe water. This project aims to assess and improve water quality in a specific region to ensure safe drinking water, protect the environment, and support various water-dependent activities.

Problem Definition:

Water quality analysis is a critical process used to assess the physical, chemical, biological, and radiological characteristics of water. This analysis is essential for various purposes, including ensuring safe drinking water, protecting aquatic ecosystems, and managing water resources.

Objective:

The project aims to assess the potability of drinking water based on water quality parameters. It employs a data-driven approach, utilizing data preprocessing, exploratory data analysis, data visualization, and predictive modeling to achieve this goal.

Design Thinking Process:

- ★ **Empathize:** Understand the importance of safe drinking water and the factors affecting potability. Identify stakeholders' needs and concerns.
- ★ **Define:** Clearly define the project's objectives, focusing on assessing water quality and determining potability using data analysis.
- ★ **Ideate:** Brainstorm data sources, variables, and analysis methods. Consider potential data sources like water quality measurements, geographical data, and meteorological data.
- ★ **Prototype:** Develop a data analysis plan, including data preprocessing, exploratory data analysis, data visualization, and predictive modeling.
- ★ **Test:** Implement the analysis plan, assess its effectiveness, and make refinements as necessary.

Analysis Objectives:

The primary analysis objectives are to evaluate water quality parameters and determine the potability of drinking water. This includes assessing key parameters like pH levels, turbidity, chlorine concentration, and more. By utilizing data-driven insights, we aim to provide a clear assessment of whether the water is safe for consumption.

Data sources: Data sources for water quality analysis and potability assessment include monitoring station data (e.g., pH, turbidity), meteorological data (e.g., rainfall), geographic information (e.g., location of water sources), historical water quality records, real-time sensor

data and community feedback. These sources collectively provide insights into water quality parameters and help determine the safety of drinking water.

Congos for Water Quality Analysis:

- ★ **Data Integration:** Gather and integrate water quality data from various sources. These sources may include sensors, databases, spreadsheets, or external sources.
- ★ **Data Modeling:** Create a data model within Cognos that reflects the structure of your water quality data. Define relationships between tables or datasets, especially if you're dealing with multiple data sources.
- ★ **Data Visualization:** Use Cognos to create interactive dashboards and reports that visualize water quality metrics. Common visualizations include line charts, bar charts, scatter plots, and geographic maps.
  Include key indicators such as pH levels, temperature, turbidity, dissolved oxygen, chemical concentrations, etc., in your visualizations.
- ★ **Historical Trends:** Utilize time-based data to track historical trends in water quality parameters. This helps in identifying patterns and anomalies over time.Create time series charts to visualize changes in water quality parameters.
- ★ **Real-time Monitoring:** If your data source supports it, implement real-time monitoring of water quality metrics. This can be particularly useful for immediate response to critical changes in water quality.
- ★ **Alerts and Notifications:** Set up alerts and notifications within Cognos to trigger when specific water quality thresholds are exceeded. This can help in proactive decision-making and response.
- ★ **Geospatial Analysis:** If your water quality data is associated with geographical locations (e.g., monitoring stations), use Cognos to create geospatial visualizations to identify spatial trends and hotspots.
- ★ **Parameter Correlation:** Analyze the relationships between different water quality parameters using correlation analysis. Cognos can help you visualize these relationships through scatter plots and correlation matrices.
- ★ **Data Export and Sharing:** Cognos allows you to export reports and dashboards in various formats (PDF, Excel, etc.) for sharing with stakeholders and regulatory bodies.
- ★ **User Access and Security:** Implement user access controls and security measures to ensure that only authorized personnel can view sensitive water quality data and reports.
- ★ **Documentation:** Keep detailed documentation of data sources, data transformations, report generation processes, and any assumptions made during analysis. This is important for audit trails and data transparency.
- ★ **Continuous Monitoring:** Set up automated data refresh and monitoring schedules to ensure that your water quality analysis remains up-to-date.

Design Thinking:

- ★ **Analysis Objectives:** The primary objective of this analysis is to evaluate and continuously monitor the quality of water in a local river. We aim to identify long-term trends, potential anomalies, and seasonal variations in water quality parameters. Additionally, we will develop a predictive model to forecast future water quality conditions, allowing for proactive management and mitigation of any environmental issues.

- ★ **Data Collection:** Data collection for this analysis involves sourcing information from multiple channels. We'll gather historical water quality data from local environmental agencies, real-time data from river sensors, meteorological data from weather stations, and geographic data pinpointing the river's monitoring stations. The collected data will undergo preprocessing, including handling missing values and merging datasets from different sources, ensuring it's suitable for analysis.
- ★ **Data Visualization Strategy:** To gain insights from the collected data, we will employ various visualization techniques. Exploratory Data Analysis (EDA) will encompass time series plots, box plots, and scatter plots to understand data distribution and relationships. Geospatial visualization will involve mapping monitoring stations and creating heatmaps. We will also design interactive Cognos dashboards with dynamic filters for user-friendly exploration of real-time and historical data.
- ★ **Predictive Modeling:** Building a predictive model is a critical aspect of this analysis. We will select relevant features, split the data into training and testing sets, and choose appropriate machine learning algorithms. The model will be trained using historical data, its performance assessed using evaluation metrics, and its predictions visualized against actual data. This predictive component will provide insights into future water quality conditions, aiding in decision-making.

How Insights Help Assess Water Quality and Determine Potability:

- ★ **Detecting Trends and Anomalies:** Exploratory data analysis helps identify trends in water quality parameters and potential anomalies. Sudden changes or deviations from norms may signal water quality issues that require attention.
- ★ **Informing Decision-Making:** Identifying Safe Drinking Water: The predictive model assesses water samples and predicts whether they are potable or non-potable based on historical data. This information directly aids in determining the safety of drinking water.
- ★ **Community Awareness:** Reports and visualizations can be shared with the community to increase awareness of water quality. This transparency fosters trust and encourages responsible water usage.

Conclusion:

This analysis plan integrates data collection, visualization, predictive modeling, and reporting to comprehensively manage local river water quality using Cognos. It empowers decision-makers with timely insights, promoting proactive environmental conservation and ensuring the well-being of both the community and the ecosystem.