

Introducción:

El agua puede generar paz o provocar conflictos.

Cuando el agua es escasa o está contaminada, o cuando las personas tienen un acceso desigual o nulo, pueden surgir tensiones entre comunidades y países.

Más de 3.000 millones de personas en todo el mundo dependen del agua que cruza las fronteras nacionales. Sin embargo, solo 24 países tienen acuerdos de cooperación para toda el agua que comparten.

A medida que aumentan los impactos del cambio climático y crece la población, existe una necesidad urgente, dentro y entre los países, de unirnos en torno a la protección y conservación de nuestro recurso más preciado.

La salud pública y la prosperidad, los sistemas alimentarios y energéticos, la productividad económica y la integridad ambiental dependen de un ciclo del agua que funcione bien y se gestione de manera equitativa.

La calidad del agua es el conjunto de características químicas, físicas y biológicas del agua en función de las normas de uso. La mayoría de las veces se utilizan en relación con un conjunto de normas cuyo cumplimiento, por regla general, se logra mediante el tratamiento del agua y se puede evaluar. Las normas más comunes que se utilizan para controlar y evaluar la calidad del agua reflejan la salud de los ecosistemas, la seguridad del contacto humano y la salubridad del agua potable. La calidad del agua tiene un impacto significativo en el suministro de agua y, a menudo, determina las opciones de suministro. Los parámetros de calidad del agua están determinados por el uso previsto. El trabajo sobre la calidad del agua tiende a centrarse en el agua que se trata para el consumo humano, el uso industrial o doméstico o la restauración (medio ambiente/ecosistema, normalmente salud humana/vida acuática).

Descripción de la temática de los datos:

El Dataset aborda la temática de la calidad del agua, proporcionando una visión integral y detallada de los parámetros clave que determinan su potabilidad y seguridad para el consumo humano. Este conjunto de datos recopila información sobre diversas mediciones y características del agua, lo que permite analizar y comprender la dinámica de su calidad.

El Dataset contiene una variedad de variables relevantes que capturan diversos aspectos de la calidad del agua, tales como:

* **Valor de pH:** Evalúa el equilibrio ácido-base del agua y es un indicador de su acidez o alcalinidad. Según la OMS, el pH aceptable para el agua potable varía entre 6.5 y 8.5, y los rangos actuales de investigación están dentro de este estándar.
* **Dureza:** Producida principalmente por sales de calcio y magnesio, indica la capacidad del agua para precipitar jabón y está influenciada por el tiempo de contacto del agua con depósitos geológicos.
* **Sólidos disueltos totales (TDS):** Representan la capacidad del agua para disolver minerales y sales inorgánicas, afectando su sabor y color. La OMS recomienda un límite deseable de 500 mg/l y un límite máximo de 1000 mg/l para el agua potable.
* **Cloraminas:** Utilizadas como desinfectantes en sistemas públicos de agua, las cloraminas se forman al agregar amoníaco al cloro. Niveles de hasta 4 mg/l se consideran seguros.
* **Sulfato:** Presente en minerales y rocas, su concentración varía en diferentes fuentes de agua, siendo de 3 a 30 mg/l en la mayoría de los suministros de agua dulce.
* **Conductividad:** Mide la capacidad del agua para conducir corriente eléctrica, relacionada con la concentración de sólidos disueltos. Según la OMS, no debe superar los 400 μS/cm.
* **Carbono orgánico total (COT):** Indica la cantidad total de carbono en compuestos orgánicos presentes en el agua. Según la EPA de EE. UU., el COT en agua potable tratada debe ser menor a 2 mg/l.
* **Trihalometanos (THM):** Sustancias químicas presentes en agua tratada con cloro. Niveles de hasta 80 ppm se consideran seguros.
* **Turbidez:** Mide la cantidad de materia sólida suspendida en el agua. Un valor medio de turbidez recomendado por la OMS es de 5 NTU, siendo 0.98 NTU un valor aceptable.
* **Potabilidad:** Indica si el agua es segura para el consumo humano, donde 1 significa potable y 0 no potable.

Al analizar este Dataset, es posible identificar tendencias y patrones en la calidad del agua a lo largo del tiempo, comprender los factores que afectan su potabilidad, evaluar la eficacia de los tratamientos aplicados y detectar posibles riesgos para la salud pública.

El objetivo principal de este Dataset es permitir a los analistas ambientales, ingenieros de tratamiento de agua y responsables de políticas sanitarias examinar la calidad del agua desde una perspectiva integral. Con la aplicación de técnicas de análisis de datos y herramientas como Power BI, se puede transformar este conjunto de datos en información visualmente atractiva y fácilmente comprensible, lo que ayuda a tomar decisiones estratégicas informadas para garantizar la seguridad y calidad del agua.

Alcance:  
El alcance del proyecto de Power BI centrado en el análisis de la calidad del agua puede abarcar diversas etapas y objetivos clave para obtener una comprensión completa y valiosa de los parámetros de potabilidad y seguridad del agua. A continuación, se describe un alcance típico para un proyecto de este tipo:

1. **Definición de objetivos:** Establecer objetivos claros para el proyecto, como identificar parámetros críticos de calidad del agua, analizar la variabilidad de estos parámetros a lo largo del tiempo y comprender los factores que afectan la potabilidad del agua.
2. **Recopilación de datos:** Obtener y preparar los datos de calidad del agua, que incluirán información sobre pH, dureza, sólidos disueltos totales (TDS), cloraminas, sulfatos, conductividad, carbono orgánico total (COT), trihalometanos (THM), turbidez y potabilidad.
3. **Limpieza y transformación de datos:** Realizar limpieza de datos para manejar valores nulos, datos duplicados o incoherentes. Transformar los datos según sea necesario para crear una estructura adecuada para el análisis, como la creación de tablas de hechos y dimensiones.
4. **Creación de un modelo de datos:** Diseñar un modelo de datos en Power BI que refleje las relaciones entre las tablas y permita un análisis coherente y eficaz.
5. **Desarrollo de visualizaciones:** Crear visualizaciones interactivas, como gráficos de barras, líneas, tortas y mapas, para mostrar la variabilidad de los parámetros de calidad del agua por período, ubicación y otras categorías relevantes. Desarrollar tableros interactivos que permitan a los usuarios explorar los datos y obtener información relevante.
6. **Análisis y descubrimiento:** Identificar tendencias en los parámetros de calidad del agua a lo largo del tiempo y realizar análisis comparativos entre diferentes fuentes de agua o zonas geográficas. Evaluar la influencia de factores ambientales y humanos en la calidad del agua.
7. **Implementación de métricas clave:** Calcular y mostrar métricas clave de calidad del agua, como niveles medios de pH, dureza, TDS, conductividad, entre otros.
8. **Generación de informes y Dashboards:** Diseñar informes ejecutivos y Dashboards que presenten los hallazgos y permitan a los usuarios acceder rápidamente a los datos relevantes.
9. **Capacitación y documentación:** Proporcionar capacitación a los usuarios finales sobre cómo interactuar con el tablero y realizar análisis personalizados. Documentar el proceso de preparación de datos, modelado y visualización para futuras referencias.
10. **Entrega y seguimiento:** Presentar el proyecto a los stakeholders y usuarios finales, recopilando comentarios y realizando ajustes si es necesario.
11. **Mantenimiento continuo:** Actualizar regularmente el tablero con nuevos datos para mantener la relevancia y precisión de las visualizaciones.

Este alcance proporciona una estructura general para abordar un proyecto de Power BI relacionado con la calidad del agua. Sin embargo, el alcance específico puede variar según las necesidades y metas de la organización.

Hipótesis

En el análisis de la calidad del agua, pueden plantearse diversas hipótesis para explorar y validar. Estas hipótesis son suposiciones preliminares que se pueden someter a pruebas utilizando los datos disponibles. A continuación, se presentan algunos ejemplos de hipótesis que podrían ser relevantes en este contexto:

1. **Variación estacional:** La calidad del agua varía según las estaciones del año. Por ejemplo, se podría hipotetizar que los niveles de ciertos contaminantes aumentan durante la temporada de lluvias debido a la escorrentía.
2. **Impacto de la fuente de agua:** El origen del agua (superficial vs. subterránea) tiene un impacto significativo en su calidad. Podría plantearse la hipótesis de que el agua subterránea tiene niveles de dureza más altos debido a su contacto prolongado con minerales geológicos.
3. **Efecto de la urbanización:** La calidad del agua es inferior en áreas urbanas comparadas con áreas rurales, debido a la mayor presencia de contaminantes industriales y domésticos.
4. **Influencia de las prácticas agrícolas:** Las áreas cercanas a zonas agrícolas presentan mayores niveles de nitratos y fosfatos en el agua, debido al uso de fertilizantes y pesticidas.
5. **Eficacia del tratamiento:** Las plantas de tratamiento de agua son eficaces en la reducción de contaminantes como los trihalometanos y las cloraminas. Podría plantearse la hipótesis de que los niveles de estos contaminantes son significativamente menores después del tratamiento.
6. **Variabilidad geográfica:** La calidad del agua varía según la región geográfica. Por ejemplo, se podría investigar si ciertas regiones tienen consistentemente niveles más altos de conductividad debido a la presencia de minerales específicos en el suelo.
7. **Relación entre parámetros:** Existe una correlación entre diferentes parámetros de calidad del agua. Por ejemplo, se podría hipotetizar que altos niveles de TDS están asociados con una mayor conductividad.
8. **Impacto de eventos climáticos extremos:** Los eventos climáticos extremos, como sequías o inundaciones, tienen un impacto significativo en la calidad del agua. Podría plantearse la hipótesis de que las inundaciones aumentan la turbidez del agua debido al arrastre de sedimentos.
9. **Temporalidad y calidad:** La calidad del agua ha estado mejorando o empeorando de manera constante a lo largo de un período de tiempo específico debido a cambios en las políticas de gestión del agua o a factores ambientales.
10. **Influencia de la industrialización:** Las zonas con alta industrialización presentan mayores niveles de contaminantes específicos, como metales pesados o compuestos orgánicos volátiles, en comparación con zonas menos industrializadas.

Estas son solo algunas ideas de hipótesis que podrían ser relevantes para un análisis de calidad del agua. Es importante recordar que las hipótesis deben ser respaldadas por datos y sometidas a pruebas estadísticas para determinar su validez.