Water Pollution and Disease Analysis Report

Dataset Description

The dataset contains environmental and public health indicators across various countries and years. Key variables include:

- Water pollution indicators: Contaminant Level (ppm), pH Level, Turbidity (NTU), Nitrate and Lead concentrations.
- Health metrics: Diarrheal, Cholera, and Typhoid cases per 100,000 people.
- Demographics and infrastructure: Access to clean water, sanitation coverage, GDP per capita, urbanization rate, healthcare access index.

There are 25 columns in total, capturing a comprehensive view of the relationship between water quality and public health outcomes.

Key Findings

1. Higher contaminant levels often correspond with increased disease incidence, especially diarrheal diseases.
2. Countries with lower access to clean water and sanitation show higher values in Cholera and Typhoid cases.
3. Economic indicators such as GDP and healthcare access are inversely related to pollution-related disease rates.

Insights from Visualizations

1. Correlation Heatmap:

Water Pollution and Disease Analysis Report

A heatmap showed strong positive correlations between Contaminant Level, Bacteria Count, and disease incidence, especially diarrhea.

2. Bar Plot of Disease Cases by Country:

Countries with poor water treatment and sanitation systems reported the highest cases of waterborne diseases.

3. Trend Line: Pollution vs. Disease over Years:

In some regions, improved sanitation coverage and water treatment methods over time correlated with a decline in disease rates.

Conclusion

This analysis reinforces the critical link between water quality and public health. Efforts to reduce pollutants and improve infrastructure could significantly lower the disease burden in affected regions.