Water Woes: Unveiling India's Water Quality Crisis

(Because Water Pollution Ain't Water Under the Bridge)

Shreya Singh

October 5, 2024

1 Abstract

This report examines the water quality data for affected areas across India from 2009 to 2012. With increasing industrial activities and development projects, water contamination has become a growing concern. This analysis aims to identify key patterns in water quality degradation, assess the predominant contaminants, and explore potential associations with industrial and developmental efforts. Through exploratory data analysis (EDA), visualizations, and statistical methods, the study aims to shed light on the extent of the problem and offer recommendations for mitigating future degradation.

2 Introduction

Water is a vital resource for both human and ecological sustainability, and the degradation of water quality poses severe risks to public health and the environment. In India, rapid industrialization and development projects have raised concerns about their impact on water sources. The Government of India, through the Ministry of Drinking Water and Sanitation, has released data on affected water quality for the years 2009 through 2012.

This report explores patterns of water contamination at the Panchayat and village levels. The focus is on examining the contaminants reported across different regions, identifying any associations with industrial activities, and offering insights into recurring degradation patterns. With these findings, policymakers and local communities can better address the problem and take preventive measures to protect water quality.

3 Data Description

The dataset used for this analysis comprises records of water quality degradation from various regions of India, reported between 2009 and 2012. The dataset is publicly available on Kaggle.

Numerical Data:

- **Year:** The dataset spans from 2009 to 2012, with each entry representing a report for a specific year.
- **Quality Parameters:** The key parameters measured include salinity, iron, fluoride, arsenic and nitrate

Categorical Data: The key categorical variables consist of:

- **State Name:** 27 states are represented, with Rajasthan having the highest frequency of reported issues (131,417).
- **District Name:** A total of 493 districts are included.
- **Block Name:** 3,577 blocks are represented.
- **Panchayat Name:** The dataset details 37,471 panchayats.
- **Village Name:** 81,703 villages are listed.
- **Habitation Name:** 147,566 habitation names are included.

4 Exploratory Data Analysis with Visualization

4.1 Distribution of Water Quality Parameters

The count plot (Figure 1) illustrates the number of occurrences for each quality parameter in the dataset. From the plot, we observe that the quality parameter labeled 'Iron' has the highest frequency of occurrences, followed by 'Salinity', 'Fluoride', 'Arsenic', and 'Nitrate'. This indicates that iron contamination is a prevalent issue in the water quality data analyzed, warranting further investigation into its sources and impacts on health.

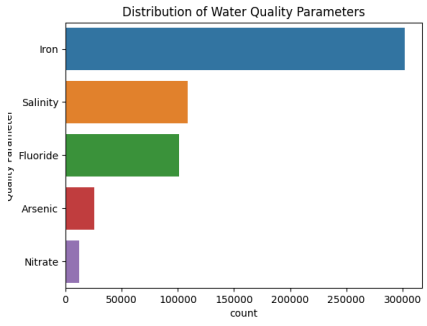

Figure 1: Distribution of Water Quality Parameters

4.2 Number of Affected Areas Over the Years

The plot (Figure 2) illustrates the distribution of affected areas by year. The count plot indicates that the highest number of affected areas occurred in the year 2009. Following this peak, there was a noticeable decline in the number of affected areas, which may reflect improvements in water quality management.

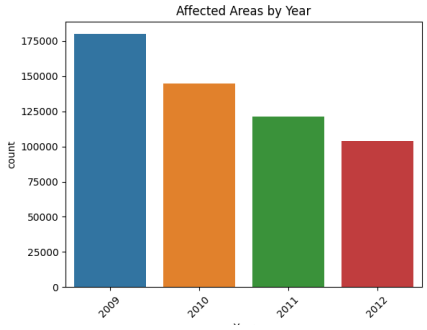

Figure 2: Number of Affected Areas by Year

4.3 State-wise Water Quality Issues

The plot (Figure 3) reveals a clear disparity in the number of water quality issues reported across states, indicating potential areas that may require targeted intervention and management efforts..

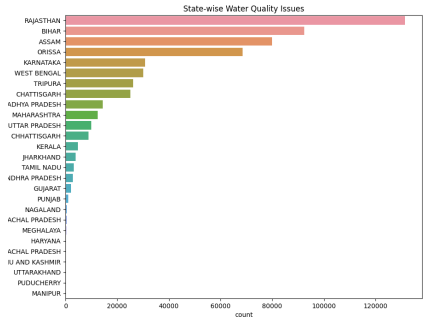

Figure 3: State-wise Water Quality Issues

4.4 Heatmap of Water Quality Parameters by State

The heatmap (Figure 4) displays the count of each quality parameter for every state, providing a comprehensive view of the distribution of water quality is-

sues across the regions. Such insights are crucial for targeted interventions and policy-making aimed at improving water quality in the affected regions.

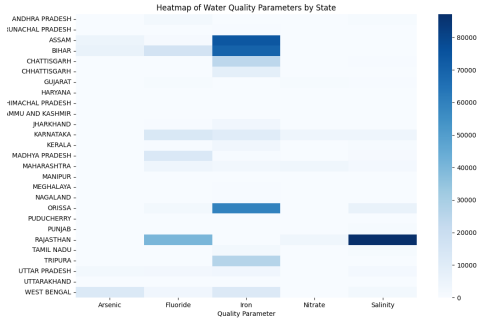

Figure 4: Heatmap of Water Quality Parameters by State

4.5 Water Quality Issues of States Over Years

To analyze the trends of water quality issues across various states over the years, a grouped bar chart was created (Figure 5). This chart depicts the count of water quality issues per state for each year, providing a clear overview of how these issues have evolved over time.

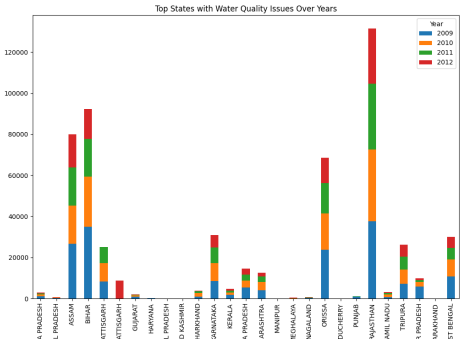

Figure 5: Water Quality Issues of States Over Years

4.6 Top Districts with Water Quality Issues

To identify the areas most affected by water quality issues, we filtered the dataset to display the top 50 districts with the highest number of reported issues. The horizontal stacked bar chart below illustrates the distribution of various water quality parameters across these districts.

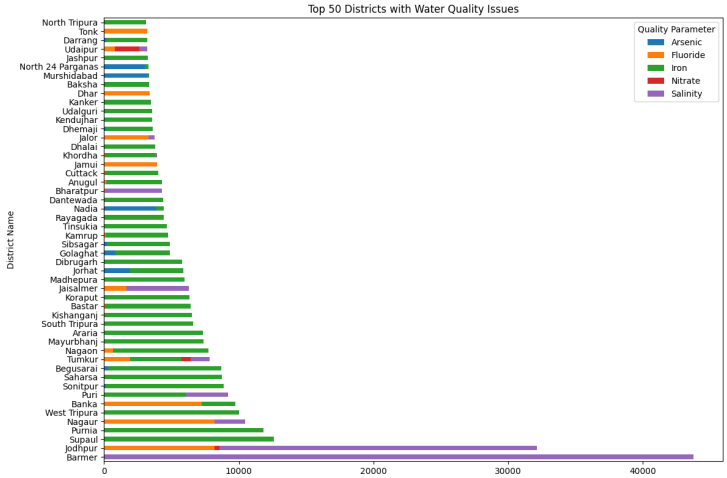

Figure 6: Top 50 Districts with Water Quality Issues

5 Discussion on Results

The analysis reveals several significant findings:

- Trend Analysis: The highest number of affected areas was observed in 2009, with a notable decline in subsequent years. This decline may suggest improvements in water quality management or changes in data reporting practices.
- State Disparities: States like Rajasthan and Bihar show significantly higher counts of water quality issues, indicating a need for targeted intervention in these regions. In contrast, states like Uttarakhand and Puducherry reported minimal issues, possibly due to better resource management or geographical advantages.
- Quality Parameter Variations: The plots illustrate the fluctuating count of various quality parameters across states and years, with parameters such as salinity and iron consistently appearing as critical issues. Understanding these fluctuations is vital for developing effective management strategies.

6 Conclusion

This analysis highlights the significant water quality issues faced by different states in India. The findings of this report underscore the urgency for improved water quality management strategies tailored to the specific conditions of each state. Continuous monitoring and effective policy implementation are essential to improve water quality, alongwith future research to capture evolving trends.