

# Indian Water Quality

Anna Maria K V

2022-10-30

## Links

Dashboard

Youtube

## Abstract

Water quality is among our prime concerns as human life is highly dependent on water. In this study we use the Indian Water Quality data set over the years 2009 to 2012 to observe the overall water quality in India over the years. We observe the key parameters affecting water quality across various regions in India. We also analyse the data on various levels including state, district and block. Thus we try to address the popular concern about water quality deteriorating over the years.

## Introduction

Water is existential for life on Earth. It plays an important role in sustaining communities, economies, and societies. The availability of clean drinking water is an essential need for all living beings. In India, most places depend upon ground water for their needs. In fact, we are the largest users of ground water in the world.

Towards the beginning of this century, there was a huge rise in industrial sector. Constructions increased and many manufacturing units were established. Though all these developments led to economic prosperity, there is a growing concern about the underground water quality.

In this project, we try to analyse the water quality affected data from 2009 to 2012 and find out if there is any significant change in the water quality over the years. We will also try to find out the regions most affected and the quality parameters affecting the most. We carry out the study by the following aspects:

1. Observing the quality parameter most affecting the water quality in India.
2. Observing the trend shown by each parameter over the years.
3. Observing the proportion of parameters over the states.
4. Observing the water quality of each state and the parameter that affects maximum.
5. Observing the district wise and block wise water quality of each state.

## Dataset Description

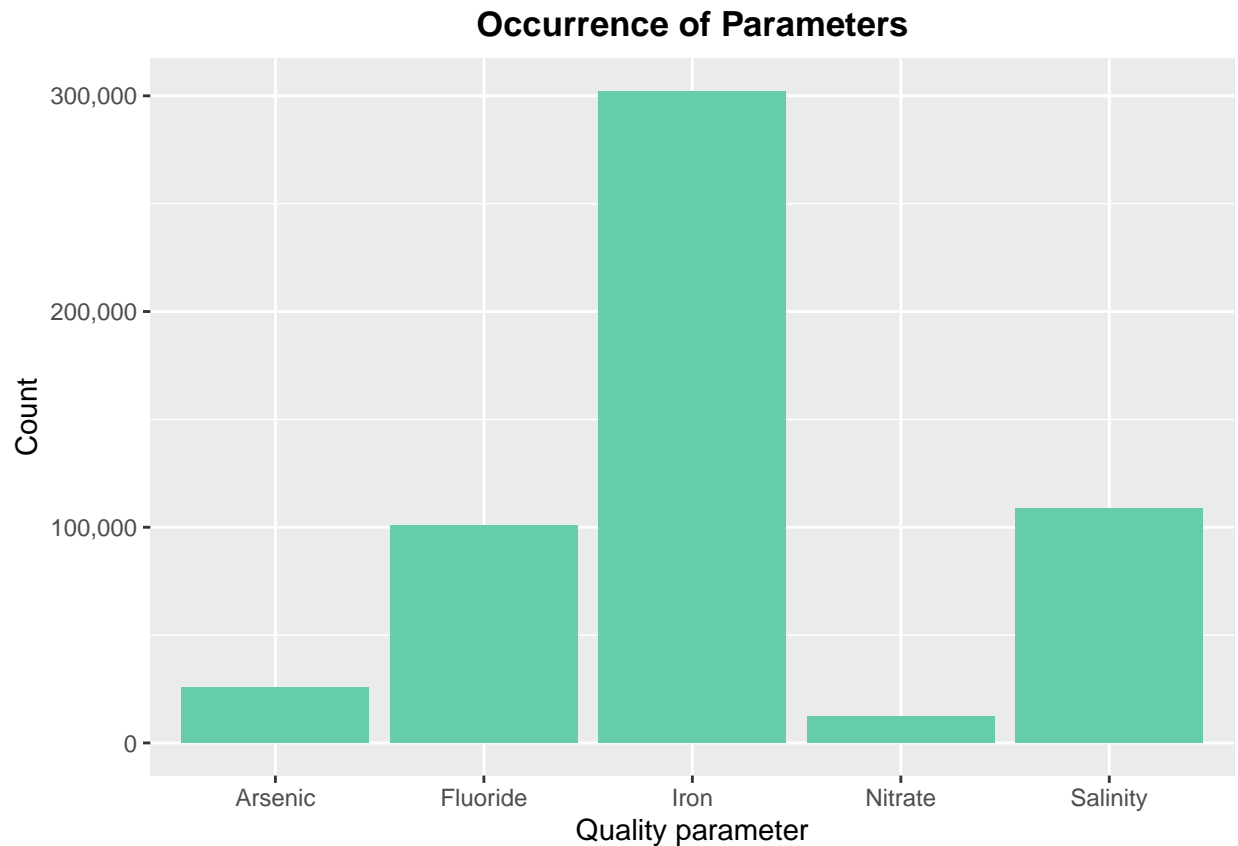
The dataset used in this study is taken from Kaggle. The link to original data is: <https://www.kaggle.com/datasets/venkatramakrishnan/india-water-quality-data>

The data contains areas with affected water quality for the years 2009, 2010, 2011 and 2012. These data sets identify the state, district and specific localities in which water quality degradation has been reported in that particular year. The parameters that were found to be affecting water quality are:

1. **Salinity** : Salinity affects production in crops, pastures, and trees by interfering with nitrogen uptake, reducing growth and stopping plant reproduction.
2. **Fluoride** : Excess amounts of fluoride ions in drinking water can cause health conditions like skeletal fluorosis, arthritis, bone damage, osteoporosis, etc.
3. **Iron** : High amounts of iron in water can cause diabetes, hemochromatosis, stomach problems and nausea.
4. **Arsenic** : Long term exposure to arsenic from drinking water can cause cancer and skin lesions.
5. **Nitrate** : High levels of nitrate in drinking water may increase the risk to colon cancer.

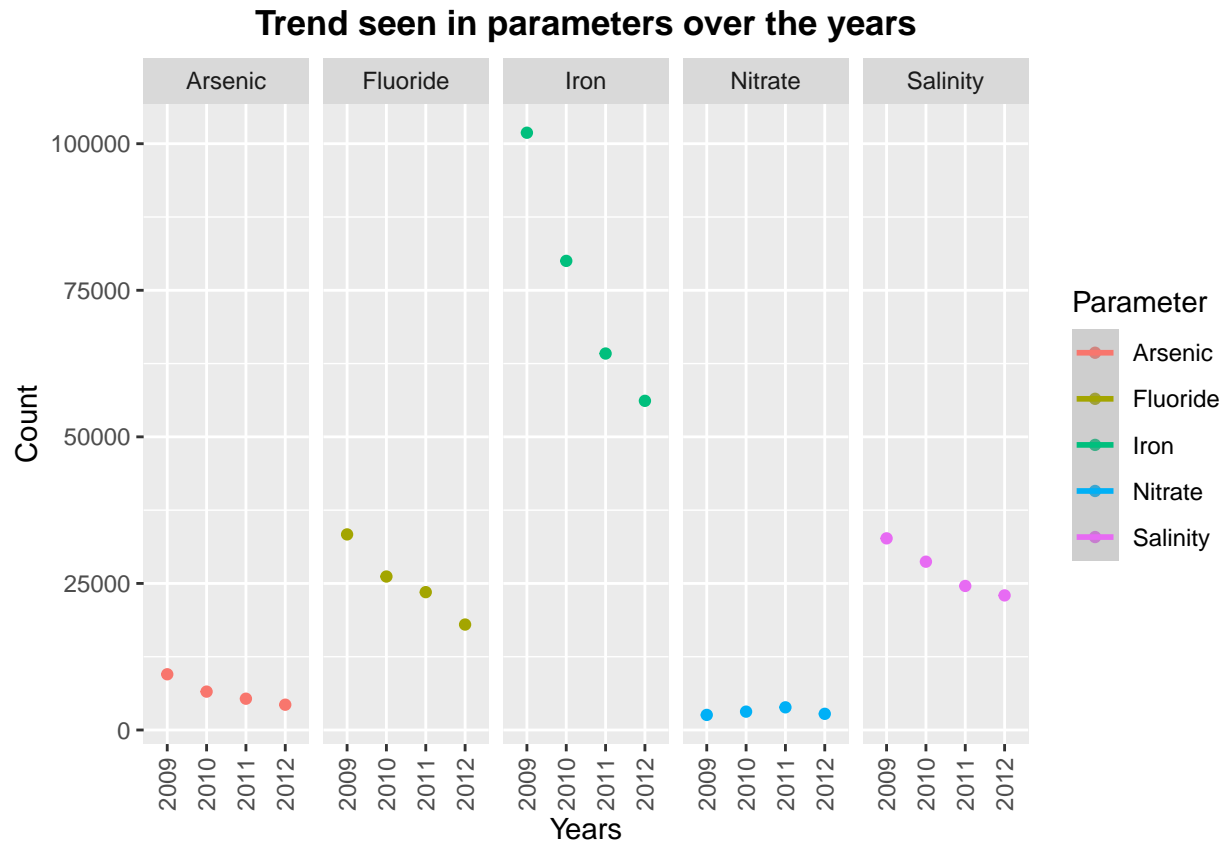
## Analysis

We first clean the data set and to start with analysis, we can find the parameter that affects water quality the most. For this, we plotted the bar graph depicting the count of each parameter over the years.



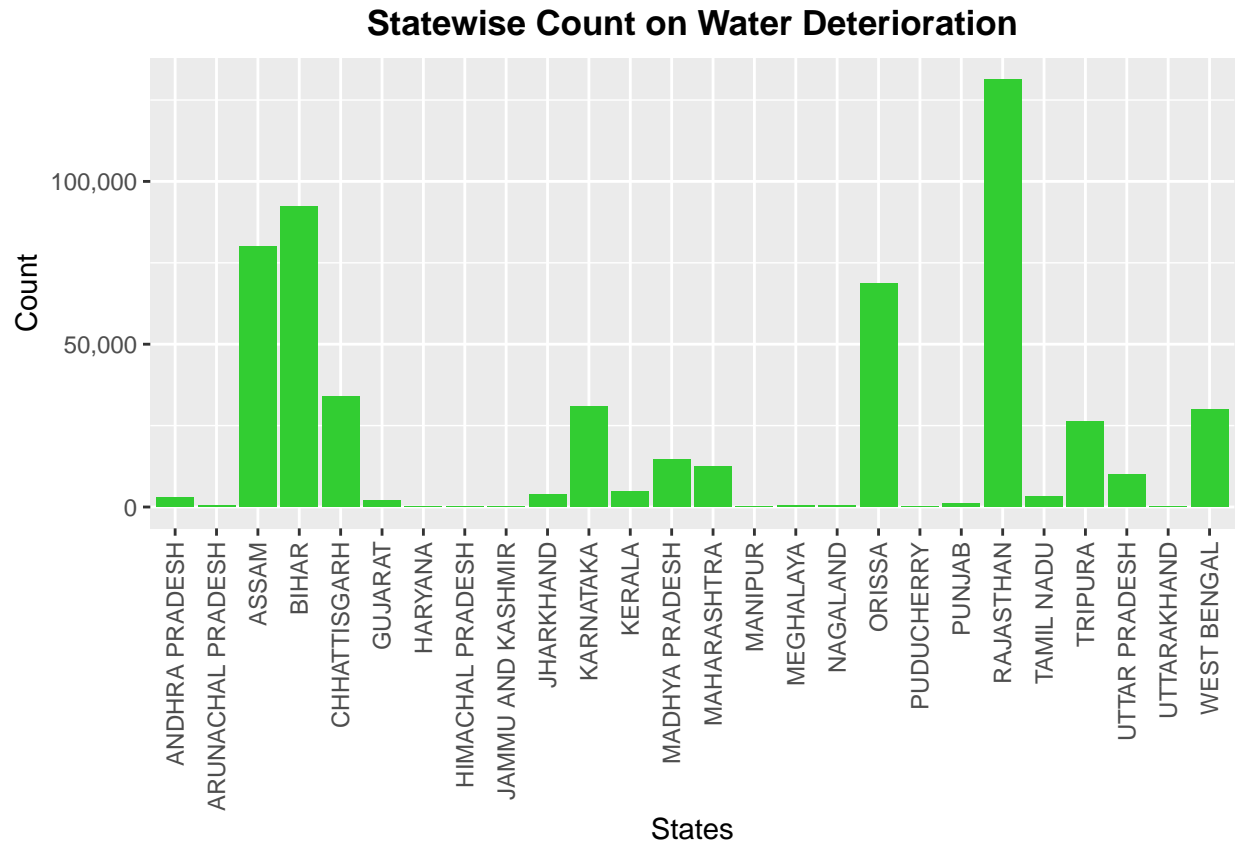
It is inferred from the above graph that Iron is the parameter that occurs the most. Second comes Salinity, followed by Fluoride and Arsenic. The parameter that occurs the least is Nitrite.

We now look at the overall trend in each parameter over the years by plotting the time series data for each variable.



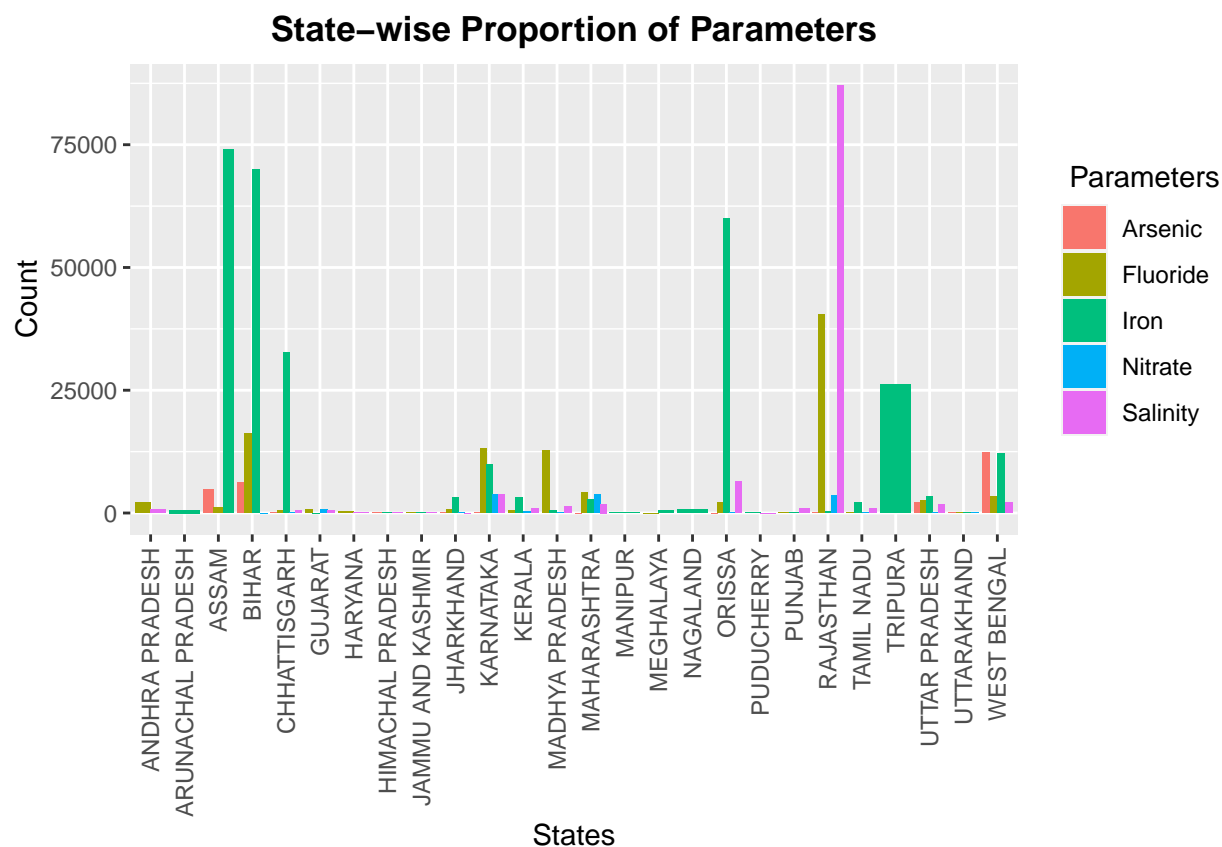
We can observe a general downward trend in the parameters over the years. This is against the popular concern about water quality. But valid conclusions are not possible at this stage. So, we now look into the states.

We can look into the quality affected data we have from each state. For this, we have plotted a bar graph with states against the count on quality affected reports over the years.

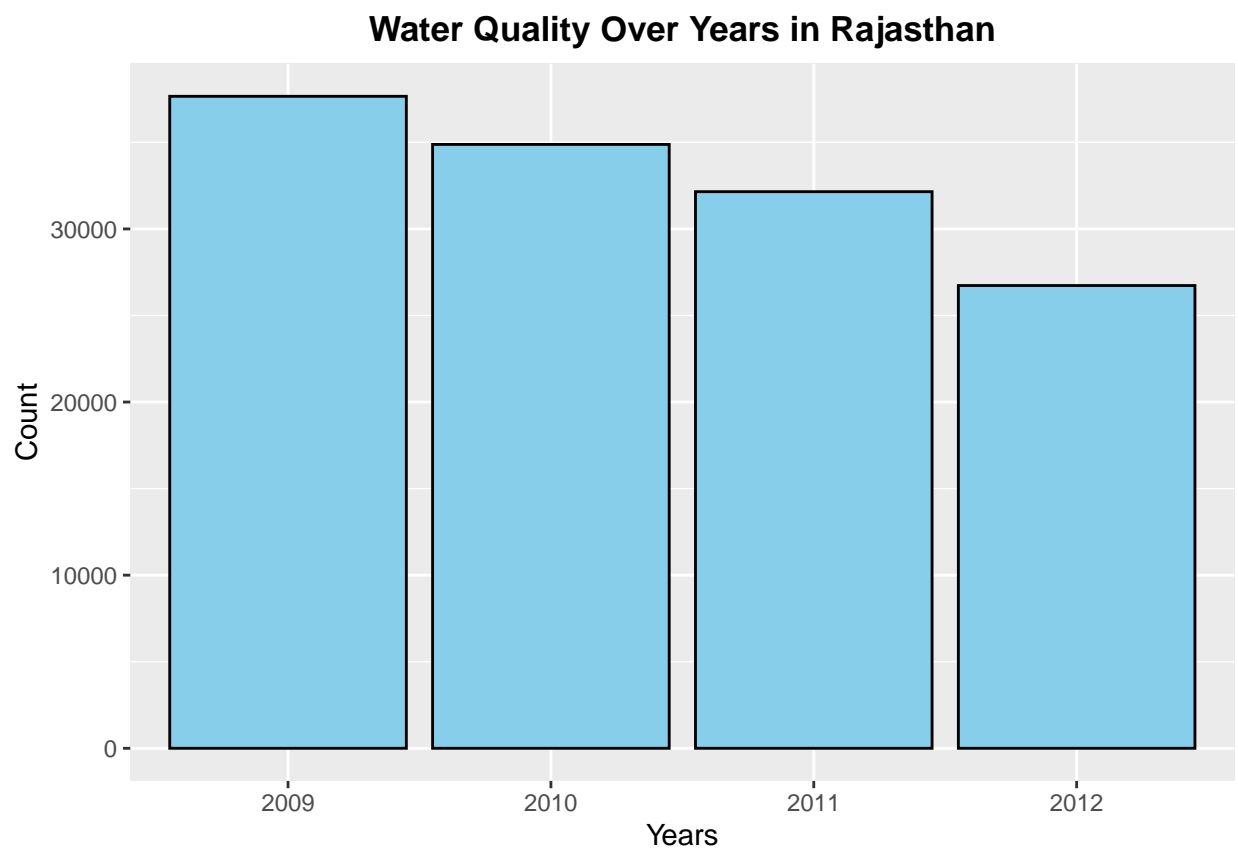


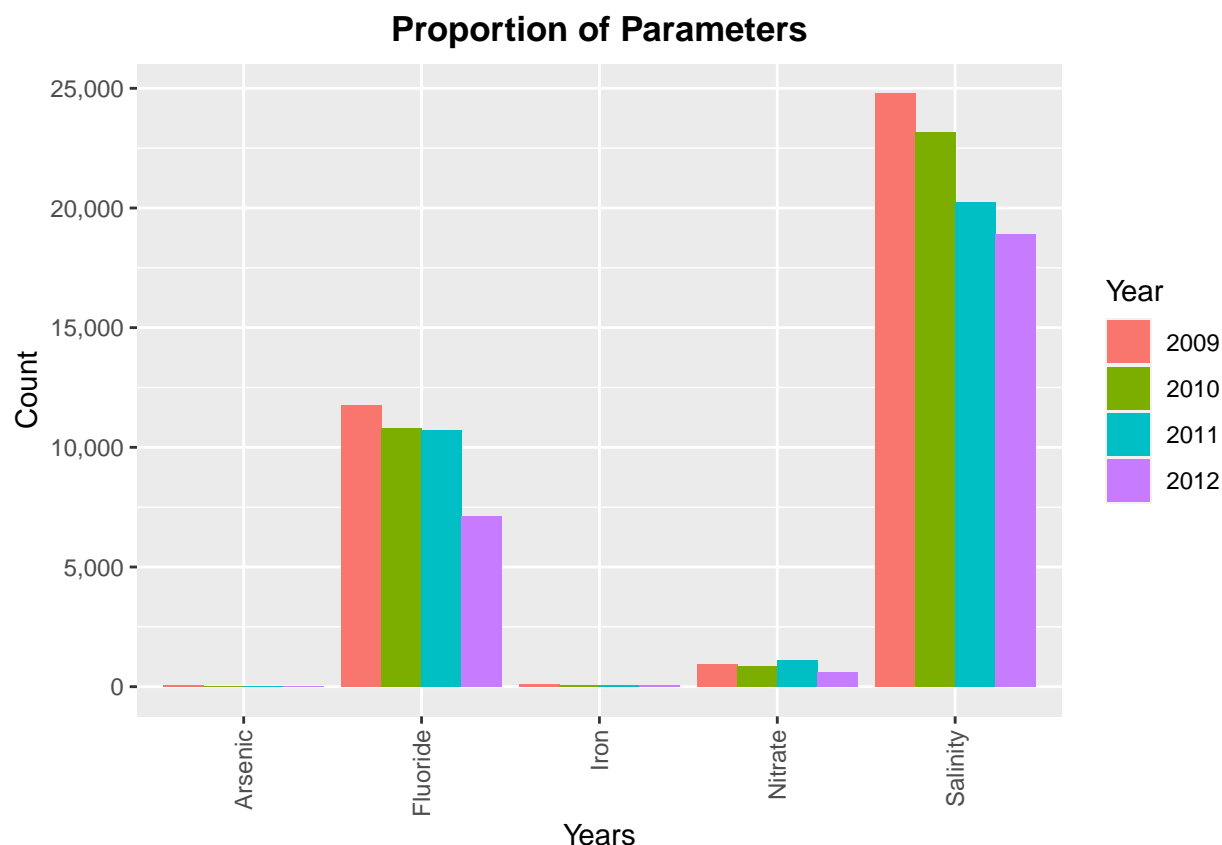
The maximum water quality affected reports are form Rajasthan, followed by Bihar. Assam and Orrisa comes in the third and fourth positions respectively.

We can also plot the state wise water quality proportion for the four years.



We find that the quality parameters affects different states differently. Salinity is the primary concern in Rajasthan but Assam and Bihar is mostly affected by Iron. The proportion of parameters varies from state to state. So, valid conclusions can be made only by analyzing each state separately. Rajasthan being the state with greatest number of quality affected reports over these years, we can now look into this state in detail. We can look at the overall water quality in each year and the proportion of each parameter over the years.





Both graphs display considerable decrease in the parameters affecting water quality, showing that the water quality is actually getting better by years. Each parameter, taken separately, also shows decrease over the years.

The analysis so far indicates that the water quality is not really affected with the developments occurring over time. Against the popular concern, the water quality seems to be increasing at this point. Nevertheless, a proper conclusion can be drawn only by analyzing the data within the sub levels of each state. As the study proceeds, I would like to analyse the quality parameters within each district and block of states. The parameter that occurs the maximum at each levels will be identified and their trend with the time will be analysed. Based on these analyses, the area that is affected at most can be identified and the quality parameter associated can be inferred.

## Conclusion

We analysed the Water Quality data set and arrived at the following conclusions: 1. The parameter that occurs the most is Iron, followed by Salinity, Fluoride and Arsenic respectively 2. An overall decrease is seen in each parameter over the years. 3. The parameter that affects water quality at most varies from state to state. 4. The most affected state is Rajasthan and the parameter the causes maximum deterioration is Salinity. 5. Water quality tends to improve in Rajasthan over the years. 6. Quality parameters shows a downward trend in Rajasthan over the years.

I hope to arrive at more conclusions after analyzing the data over each state, district and block. The overall trend in each states and its sublevels will be obtained as the next step in the study.