**Project Report**

**Air Quality Data of India**

Sri Vani Chaitanya Varma Datla

**Dataset:**

This data is combined (across the years and states) and largely clean version of the Historical Daily Ambient Air Quality Data released by the Ministry of Environment and Forests and Central Pollution Control Board of India under the National Data Sharing and Accessibility Policy (NDSAP).

Data Set Link: <https://www.kaggle.com/shrutibhargava94/india-air-quality-data>

Attributes:

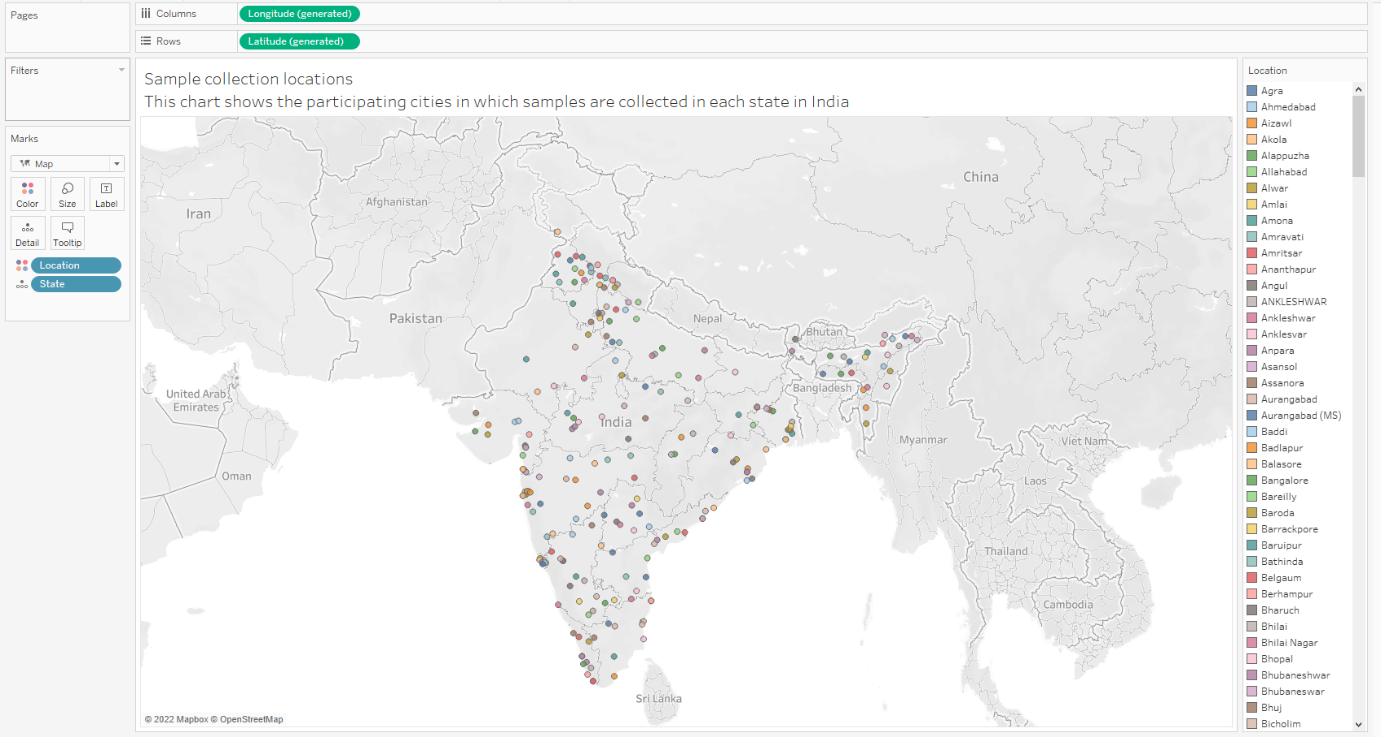
1. Stn\_code (Dimension): Stores Particular Station Code for Sates or cities.
2. Sampling\_Date (Dimension): Stores particular month and air quality report code.
3. State (Dimension): Stores Name of States.
4. Location (Dimension): Stores Name of Cities.
5. Agency (Dimension): Local Authority name which collected report.
6. So2 (Measure): Level of Sulphur dioxide (µg/m3) micrograms per cubic meter air.
7. No2 (Measure): Level of Nitrogen dioxide (µg/m3)..
8. Rspm (Measure): Level of Respirable Suspended Particulate Matter (µg/m3)..
9. Spm (Measure): Level of Suspended Particulate Matter (µg/m3)..
10. location\_monitoring\_station (Dimension): Place from Air quality measured.
11. pm2\_5 (Dimension): Level between 2 and 5.
12. Date (Dimension): Stores dates monthly from Year 1987 to 2015.
13. Type (Dimension): Stores zone in which air quality measures like industrial, residential.

**Visualization tool used**- Tableau

**Visualization – 1:**

**Identify the cities in India from where air quality data was recorded.**

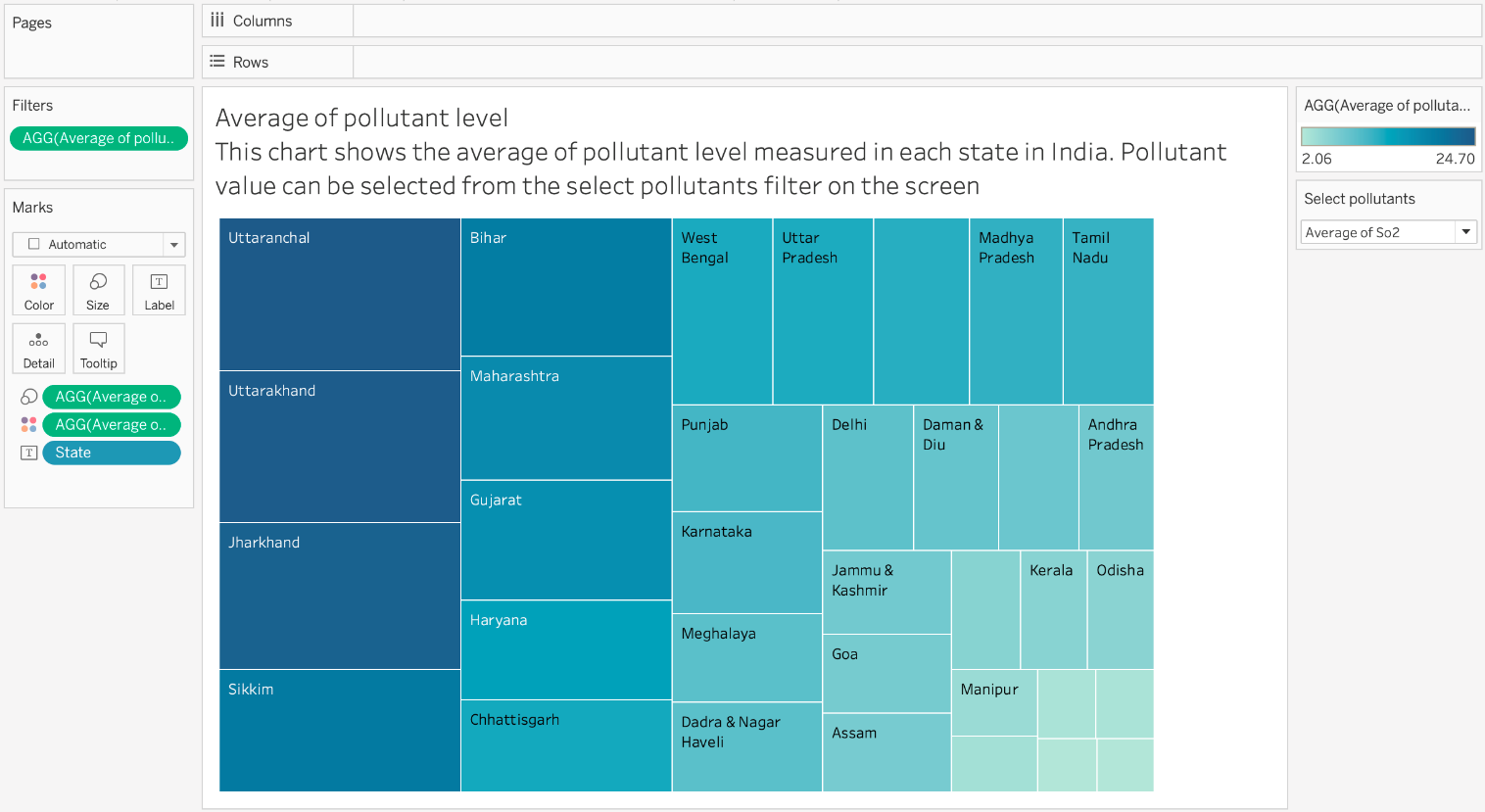
In the visualization chart 1, I have created a **map** showing the participating cities where samples are collected from each state in India.



**Visualization – 2:**

**Identify the average pollutant level at each state in India.**

In the visualization sheet 2, we have created a **parameter** called ‘select pollutants’ and a **calculated field** called ‘Average of Pollutant’ which assigns one of the average values of So2, No2, PM2.5, SPM, RSPM based on the value selected by the user in select pollutants parameter on the screen. Then the chart rearranges itself based on the average of selected pollutant.

******

**Visualization – 3:**

**Identify the average pollutant level at each location type in India.**

In the visualization chart 3, I have created a **horizontal bar chart** to compare the average of pollutant levels at each of the location types in India.

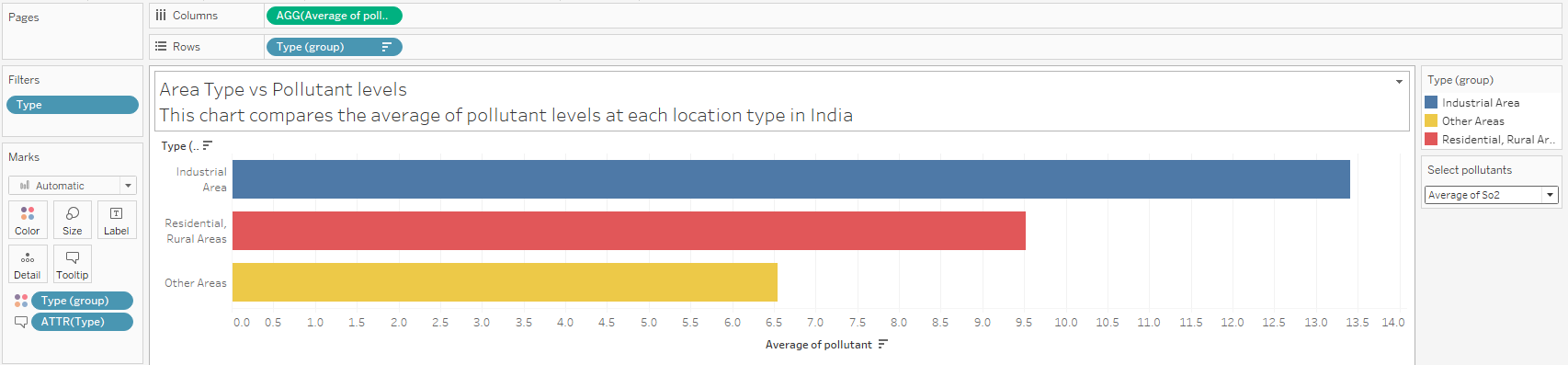
Location types are:

Industrial Areas

Residential and Rural Areas

Other Areas

From the analysis it is evident that average pollution level of all the pollutants is highest at the industrial areas.



**Visualization – 4:**

**Identify which region is most polluted in India.**

In the visualization sheet 4, I have **grouped** all the locations in India into 3 regions –

North India,

South India and

Central India

to find out which region is most polluted with the selected pollutant from the parameter list. I have filtered the data with year starting from 2010 to 2015.



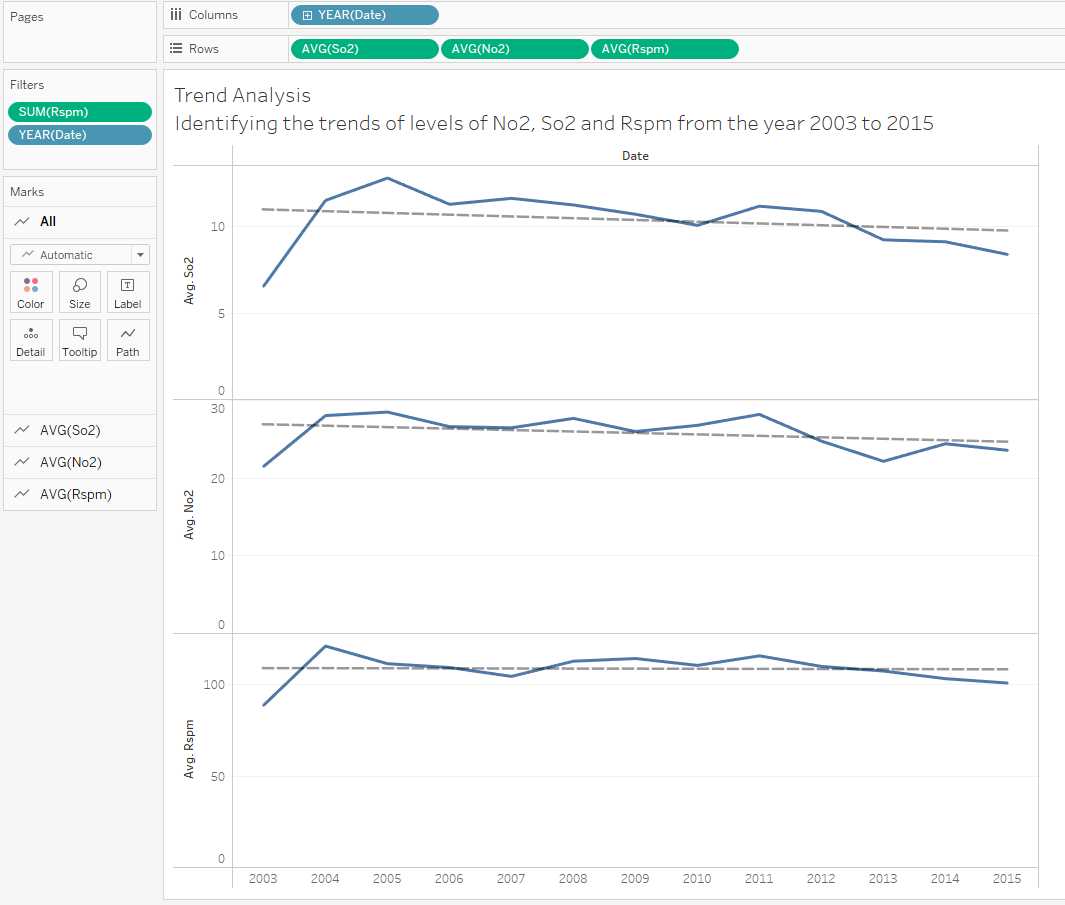
**Visualization – 5:**

**Identify and compare trend of average of pollutant levels from 2003 to 2015.**

Trend lines can be used to illustrate trends in the dataset.

In the visualization sheet 5, I have created **trend lines** for average of So2, No2 and Rspm levels from year 2003 to 2015 in India.

From the analysis it is found that pollution took a rise from the year 2003, maintained the same for some years and showing a downward trend by the end of 2015.



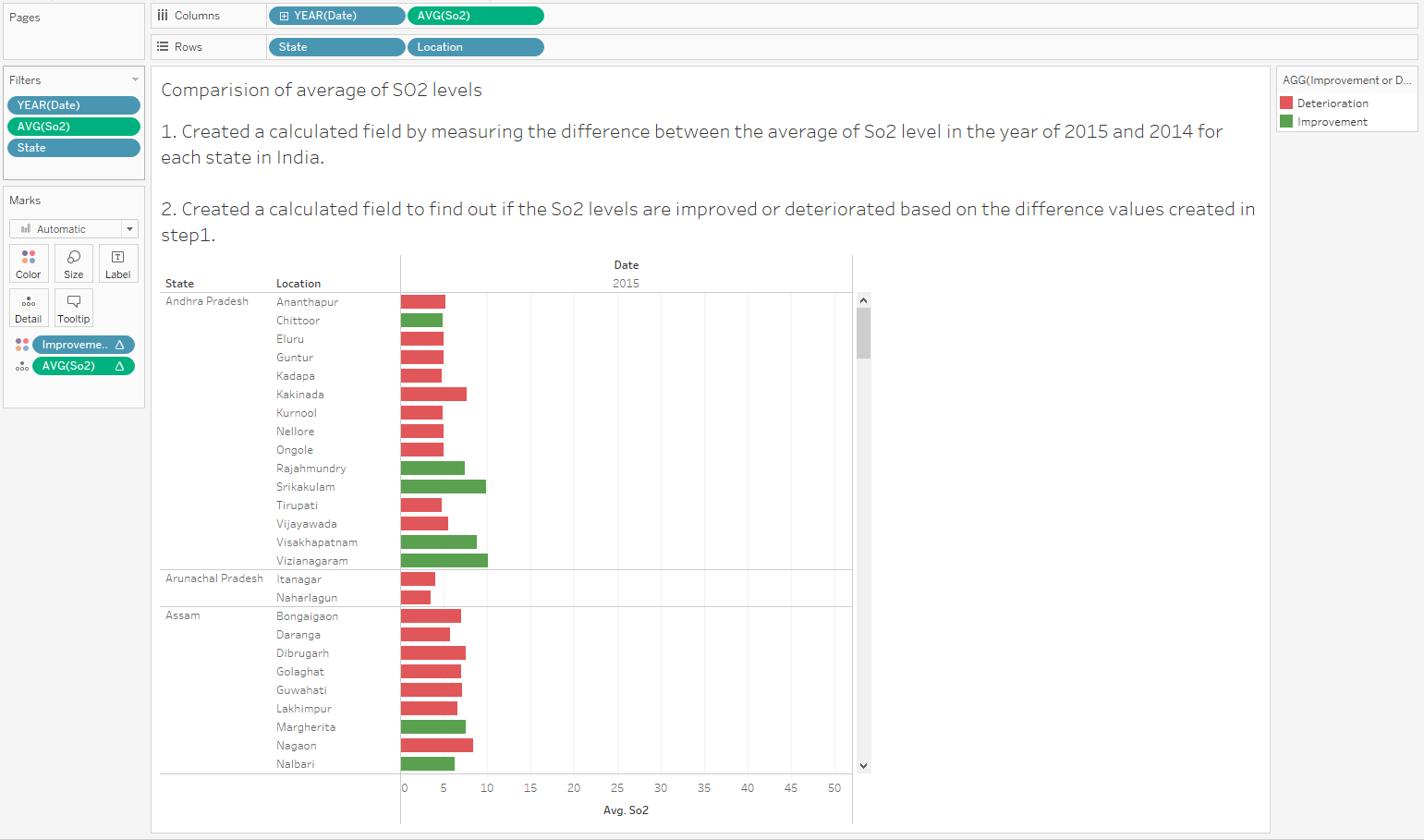
**Visualization – 6:**

**Compare the average of So2 levels between the years 2015 and 2014 for each city in India.**

In the visualization sheet 6, I have created a **calculated field** called ‘Difference between So2 levels of 2015 and 2014’. Then, created another **calculated field** called ‘Improvement or Deterioration’ based on first calculated field to find out if the So2 levels got improved or deteriorated.

If the condition improved bars are coloured as green otherwise red for the year 2015. If we hover over the bars we can see the difference of values.

(Sulfur dioxide (SO2), a foul-smelling toxic gas, is part of a larger group of chemicals called sulfur oxides. These gases, especially SO2, are emitted by the burning of fossil fuels or other materials that contain sulfur)



**Visualization – 7:**

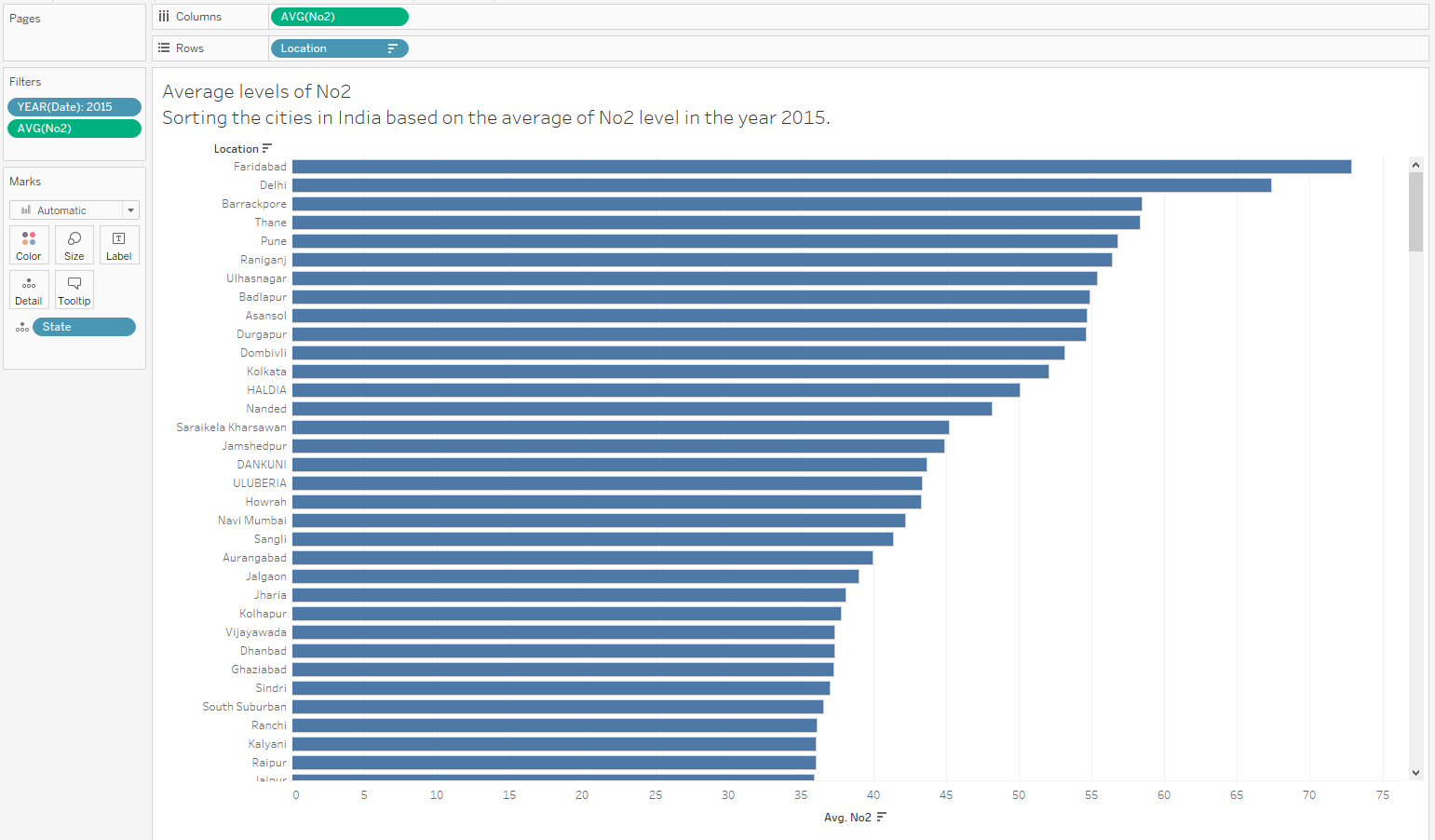
**Identify the city which is most polluted with average of No2 levels in the year 2015 in India.**

In the visualization sheet 7, I have created a chart which shows the average of levels of No2 **sorted** from high to low measured at each city in India in the year of 2015.

From the analysis, it is found that Faridabad from Haryana state in India topped the list with average No2 levels of 72.86 microgram per cubic meter air.

(Nitrogen Dioxide (NO2) is one of a group of highly reactive gases known as oxides of nitrogen or nitrogen oxides (NOx). Other nitrogen oxides include nitrous acid and nitric acid. NO2 is used as the indicator for the larger group of nitrogen oxides.

NO2 primarily gets in the air from the burning of fuel. NO2 forms from emissions from cars, trucks and buses, power plants, and off-road equipment)



**Visualization – 8:**

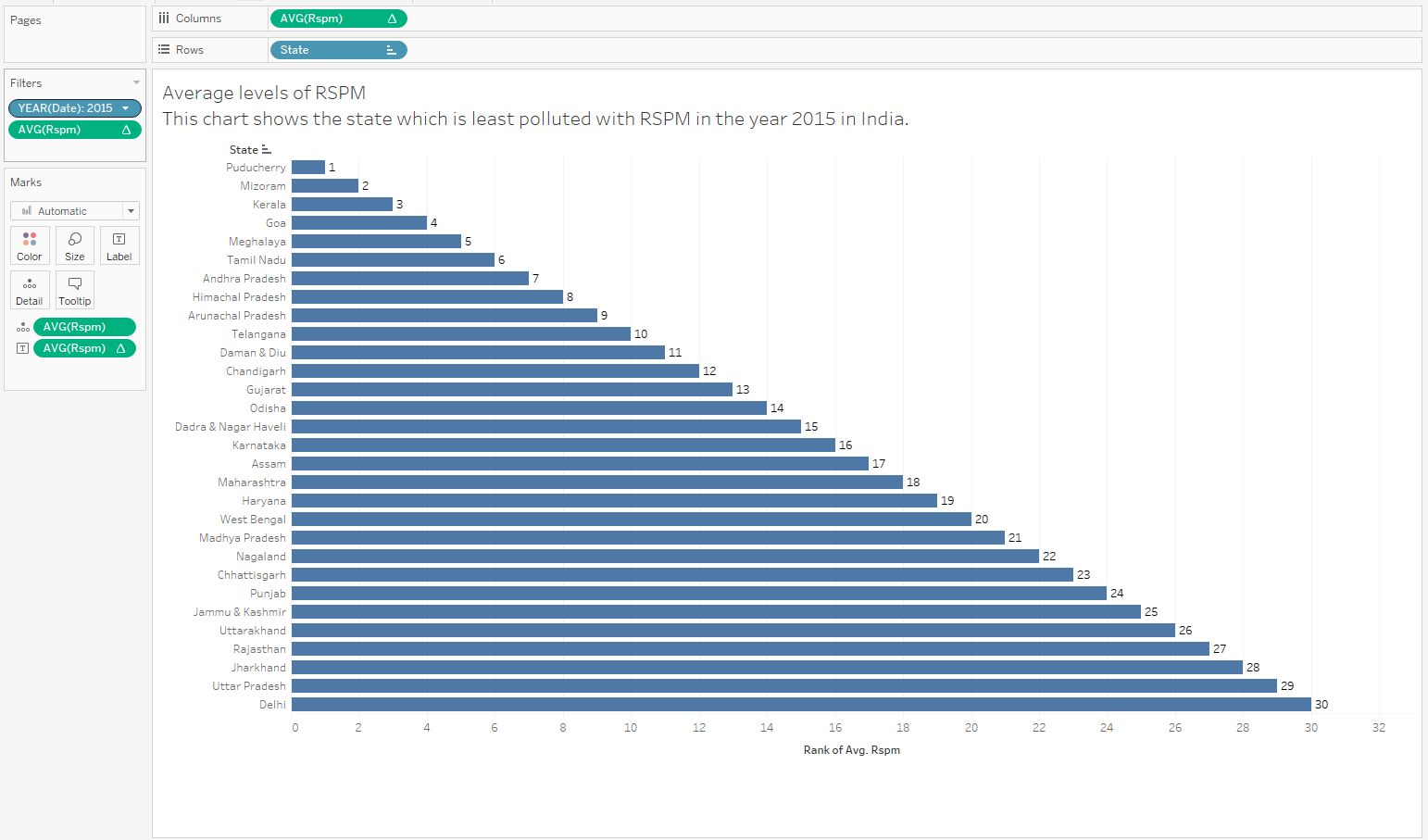
**Identify the state which is least polluted with RSPM in the year 2015 in India.**

In the visualization sheet 8, I have created a chart which implements Rank **table calculation** on the average of levels of Respirable suspended particulate matter (RSPM) which are measured at each state in India in the year of 2015.

From the analysis, it is found that Puducherry ranked first with average RSPM value of 34.8 microgram per cubic meter air which is lowest in India in the year 2015.

(PM stands for particulate matter (also called particle pollution): the term for a mixture of solid particles and liquid droplets found in the air.

RSPM are inhalable particles, with diameters that are generally 10 micrometers and smaller)



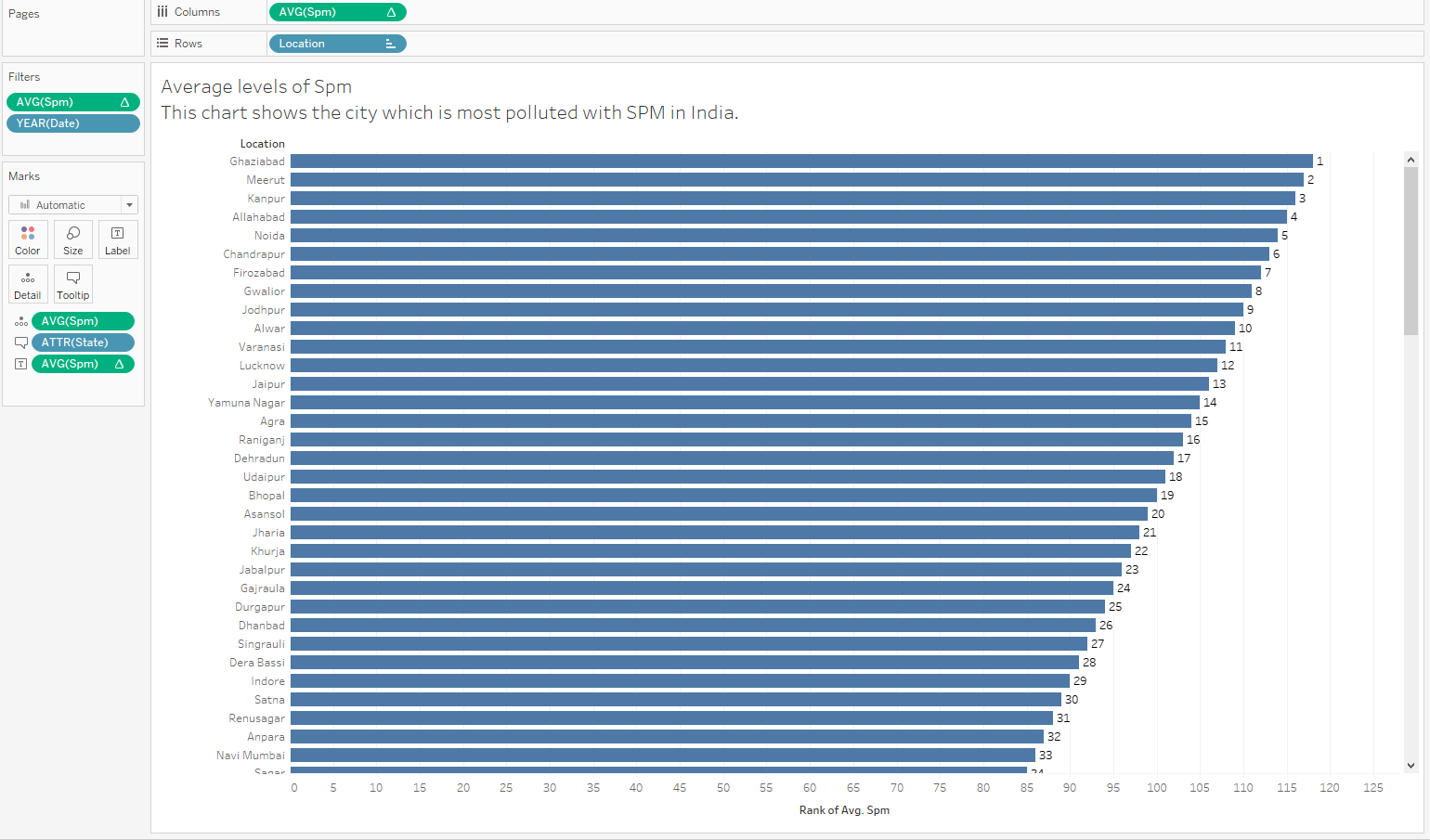
**Visualization – 9:**

**Identify the city which is most polluted with SPM in India.**

In the visualization sheet 9, I have created a chart which implements Rank **table calculation** on the average of levels of suspended particulate matter (SPM) which are measured at each city in India from the year 2010 to 2015.

From the analysis, it is found that Ghaziabad from Uttar pradesh ranked first with average SPM value of 541.8 microgram per cubic meter air which is highest in India.

(Suspended particulate matter (SPM) are finely divided solids or liquids that may be dispersed through the air from combustion processes, industrial activities or natural sources)



**Visualization – 10:**

**Identify the top 5 states polluted with both So2 and No2 from the year 2011 to 2015 in India.**

In the visualization sheet 10, I used **sets** in tableau.

Set 1 – has top 10 most polluted states with average of So2 from the year 2011 to 2015

Set 2 – has top 10 most polluted states with average of No2 from the year 2011 to 2015

**Union of Sets** – top 5 states which are present in both the sets (polluted with both So2 and No2).

The states are:

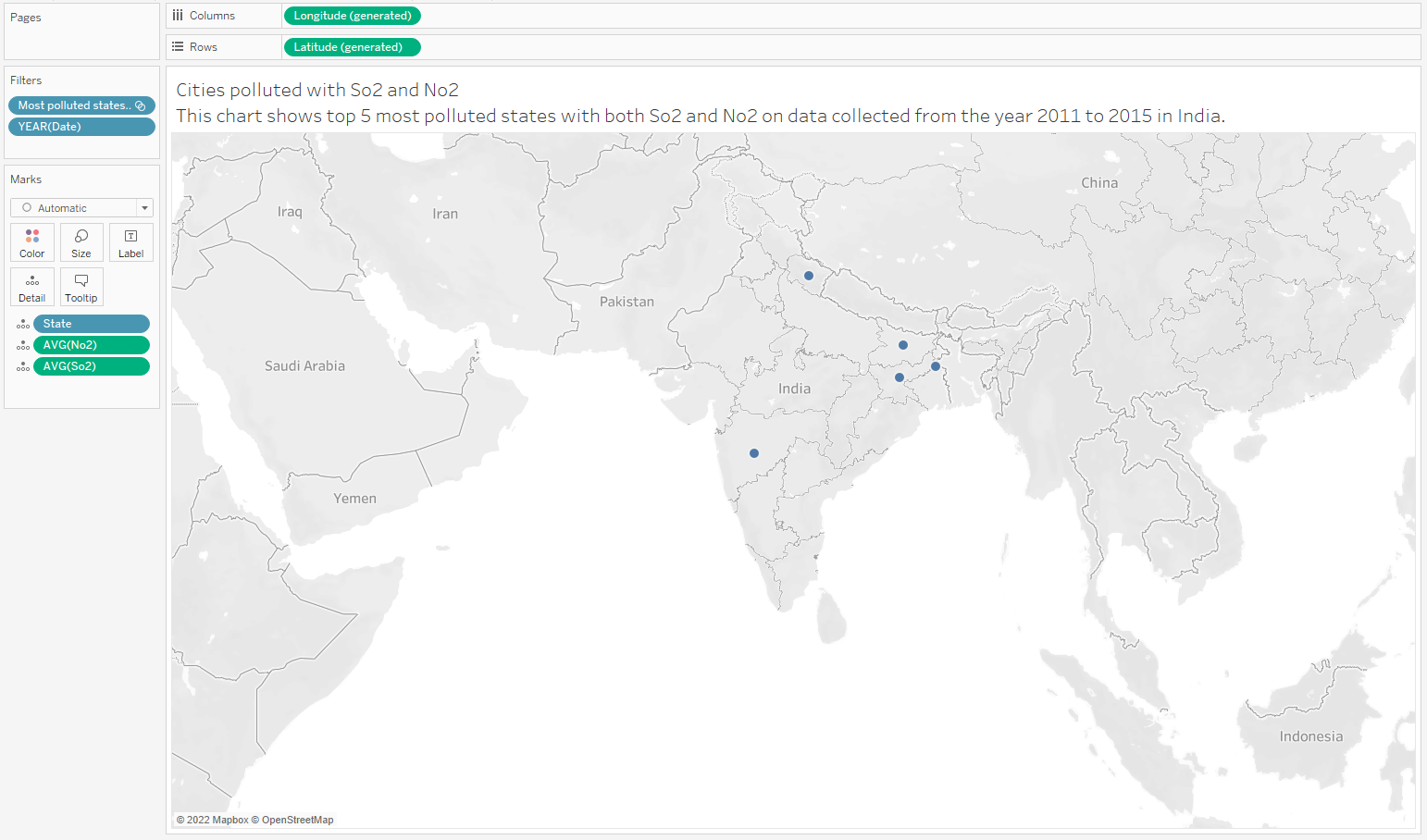
Uttarakhand

Bihar

West Bengal

Jharkhand

Maharastra

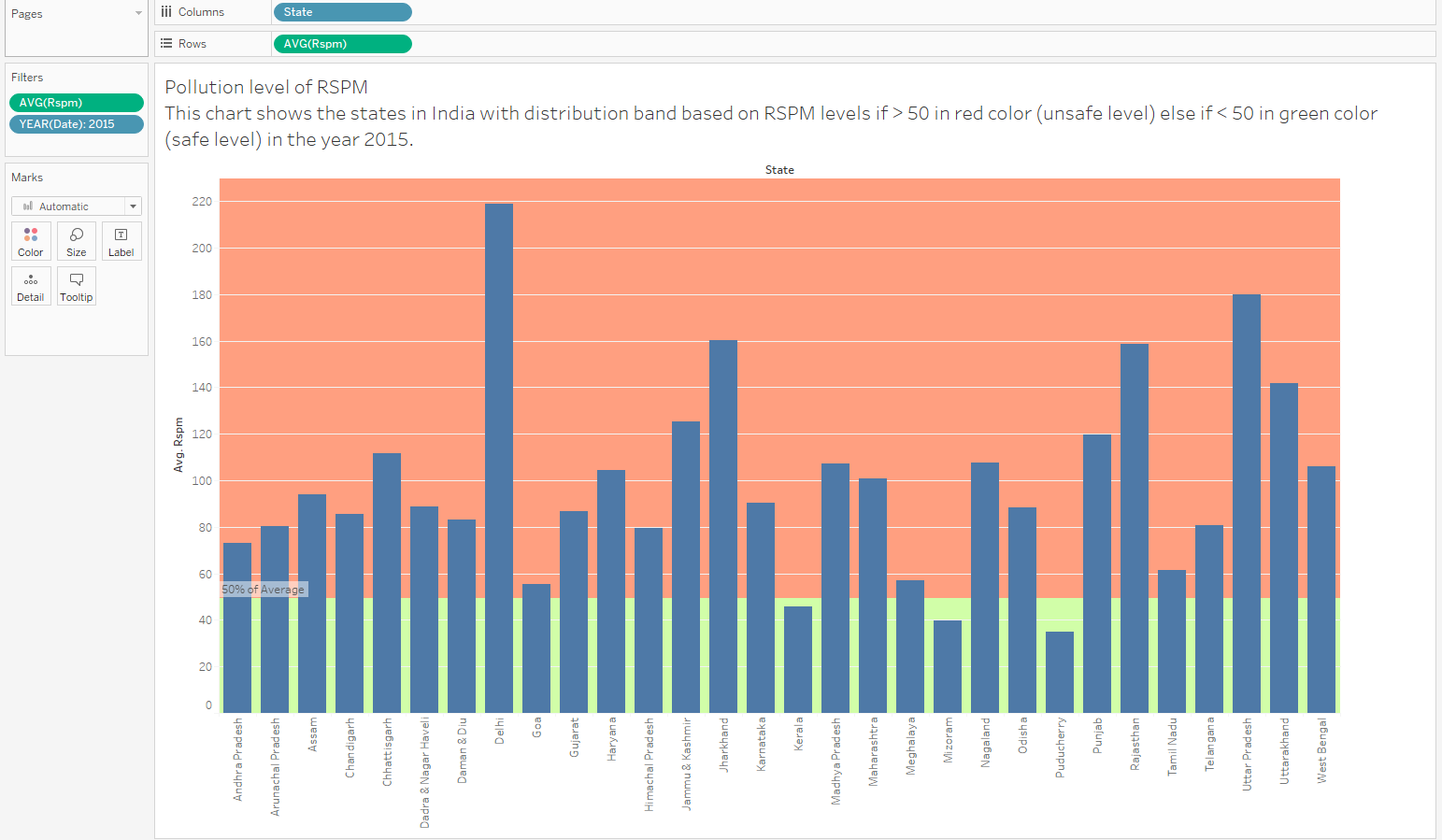


**Visualization – 11:**

**Differentiate the states in India based on safe and unsafe levels of RSPM.**

In the visualization sheet 11, I used **distribution band** to show the states which have RSPM greater than 50 in red colour indicating unsafe levels of RSPM on the other hand states with RSPM less than 50 are covered in green indicating safe levels of RSPM. We did analysis for the year 2015 on the states in India.

From the analysis, we can say capital of India Delhi is most polluted with RSPM.



**Visualization – 12:**

**Predict the future of pollutant levels based on the data available to us.**

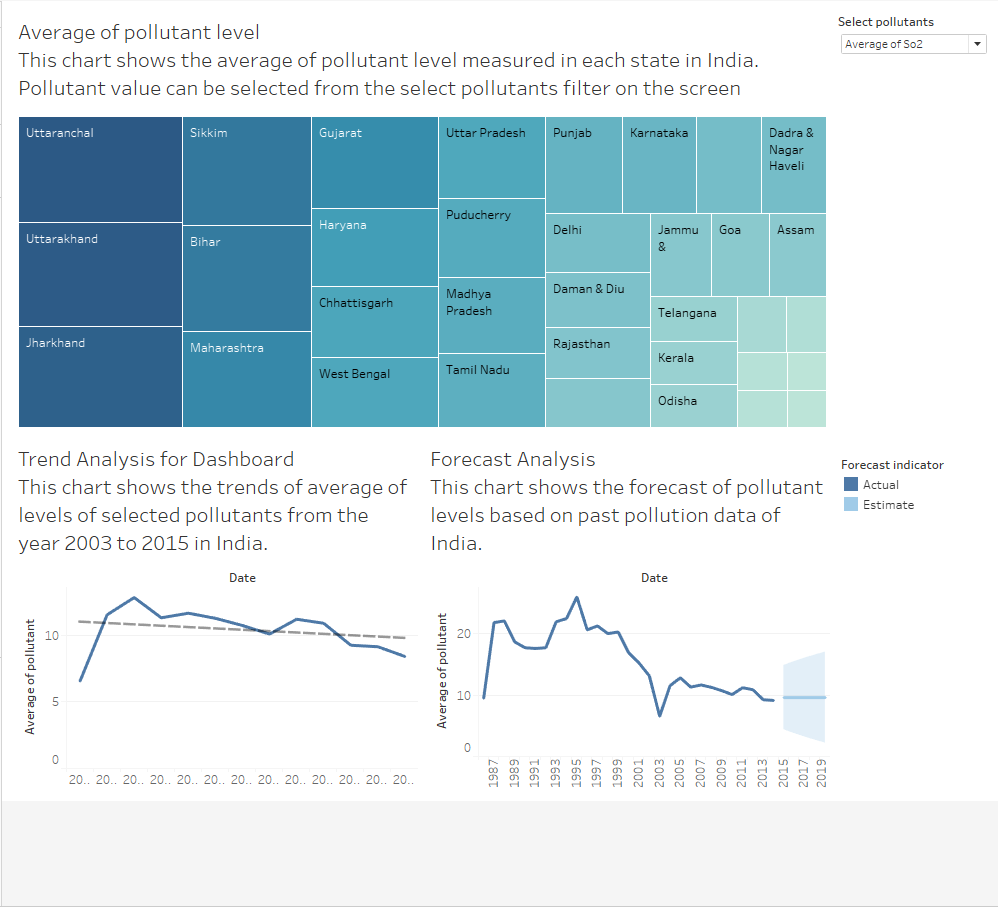
In the visualization sheet 12, we are **forecasting** the pollutant levels (one of average values of So2, No2, RSPM, SPM, PM2.5 selected by the user using select pollutants parameter) based on the pollution data available.



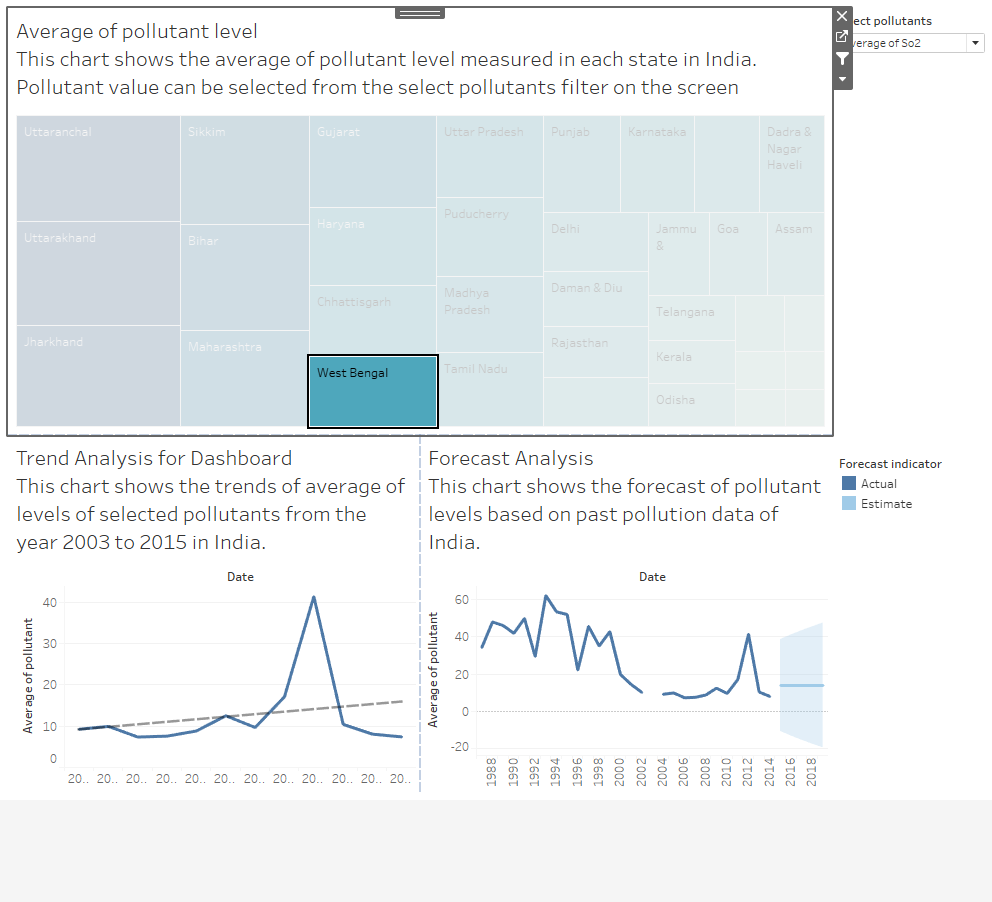
**Dashboard - 1:**

**Current and future analysis of air quality in India.**

I have added ‘Average of pollutant level’ worksheet, ‘forecast of pollutant levels’ worksheet and ‘Trend Analysis for Dashboard’ worksheet into this dashboard. Dashboard will be updated based on the selection made from ‘select pollutants’ parameter.



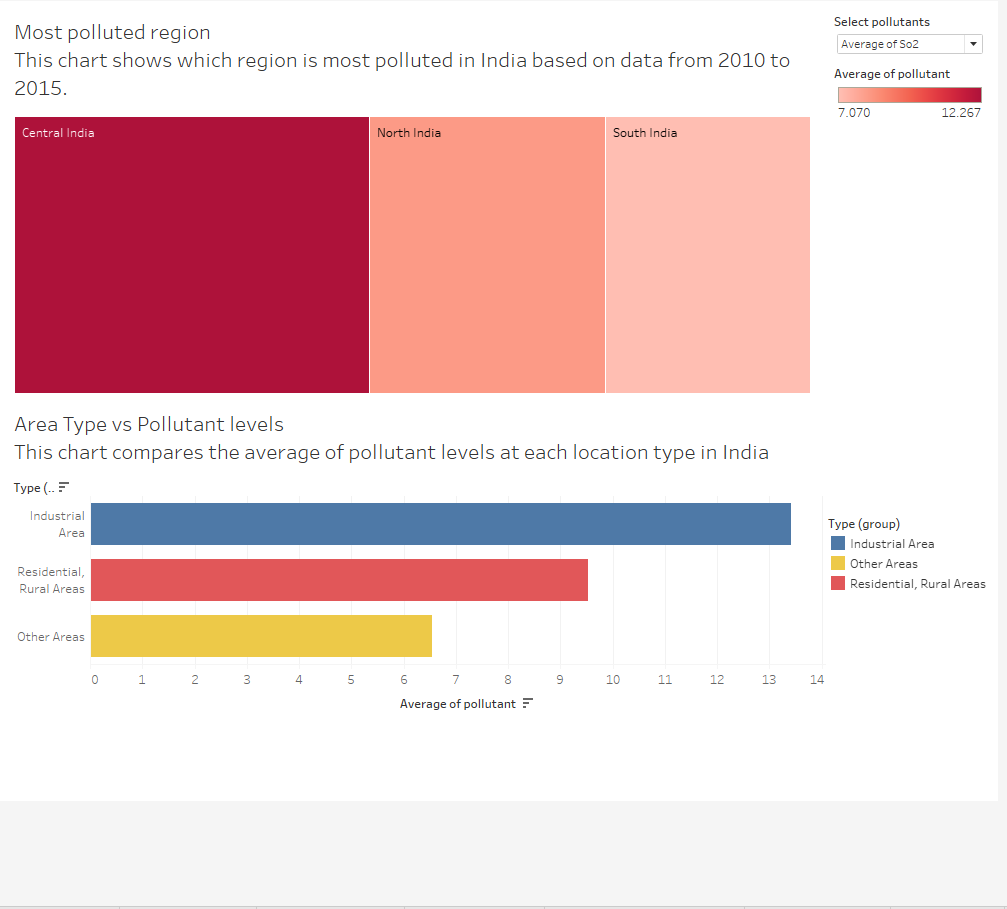
**Dashboard 1 Action** – I have created the dashboard action to see the trend and forecast of the pollutant for each state if we select a state from the top list.



**Dashboard – 2:**

**Pollution analysis based on location.**

Added Area type vs pollutant levels at each location type worksheet and most polluted region worksheet in the dashboard. Based on the value selected on ‘select pollutants’ parameter dashboard will be updated accordingly.



**Dashboard 2 Action** – I have added URL action for the regions to help users identify which states come under that region in India.

If we click on North India, it redirects to - <https://en.wikipedia.org/wiki/North_India>

If we click on South India, it redirects to - <https://en.wikipedia.org/wiki/South_India>

If we click on Central India, it redirects to - <https://en.wikipedia.org/wiki/Central_India>

**Story – 1:**

**Air Quality in India**

I have attached 5 worksheets and 2 dashboards to the story. Added text box to brief about each pollutant in the story. Formatted the story to make it visually appealing.

