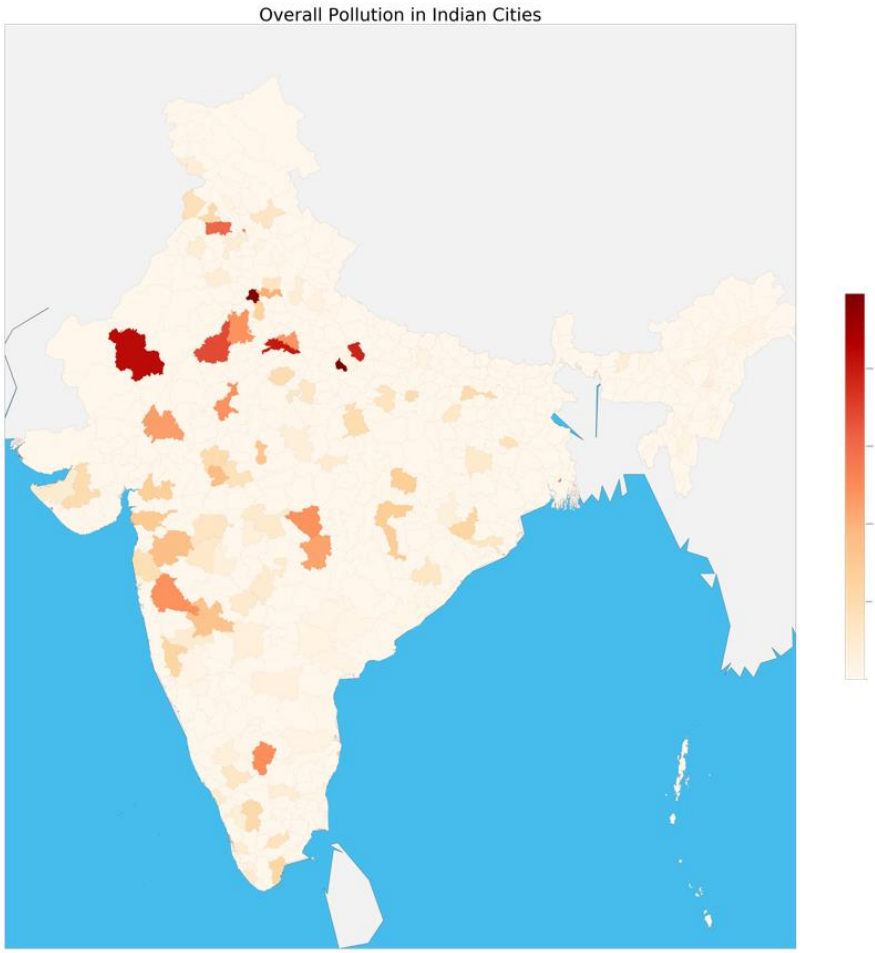
Stating that air pollution is a ‘growing’ concern would be an understatement. 91%of the world’s population lives in places where air quality exceeds WHO guideline limits. 4.2 million deaths occur every year as a result of exposure to ambient (outdoor) air pollution. A fourth of these deaths are from India. The 2018 Greenpeace and AirVisual analysis shows that 22 out of the 30 most polluted cities in the world are in India.

The government launched **National Clean Air Programme** (NCAP), a time-bound national level strategy to tackle increasing air pollution. The Programme targets a 20-30% decrease in pollution (PM10 and PM2.5 concentrations) by 2024 in 102 non-attainment cities.

I downloaded the data set from 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).

The dataset contains the daily readings of pollutants (SO_2, NO_2, RSPM, SPM and $PM_{2.5}$) from different air monitoring stations in cities (304 unique locations, 86 non attainment cities). The source of the pollutant is also stated whether the reading was recorded from a Residential, Industrial or Sensitive (ecologically sensitive zones) part of town.

The pollution levels of all the cities given in the dataset show that a handful of cities (mostly in the Northern and Central Parts of India) are responsible for half the total pollution that has occurred in the country. Non-Attainment Cities contribute to around 65% of total pollution.



To understand the gravity of the situation, I want to see how worse off the non-attainment cities are to the other cities given in the dataset. Below is the average level of the pollutant on a day in that year.

Clearly, the non-attainment cities have higher rates of pollution every year as compared to other cities. In fact Non-Attainment Cities have at an average 35% Particulate Matter than other cities. Particulate Matter have can severely impact human health and are the largest sources of manmade air pollution. It also aggravates environmental damage depleting water bodies and soil from essential minerals and nutrients.

The RSPM and SPM levels (average of 103 and 215 SI units per day) for all cities are greater than WHOs air quality standards (20 SI Units) and even India's laxer standards (40 - 80 SI Units).

What are the sources of Particulate Matter Pollution? According to Wikipedia,

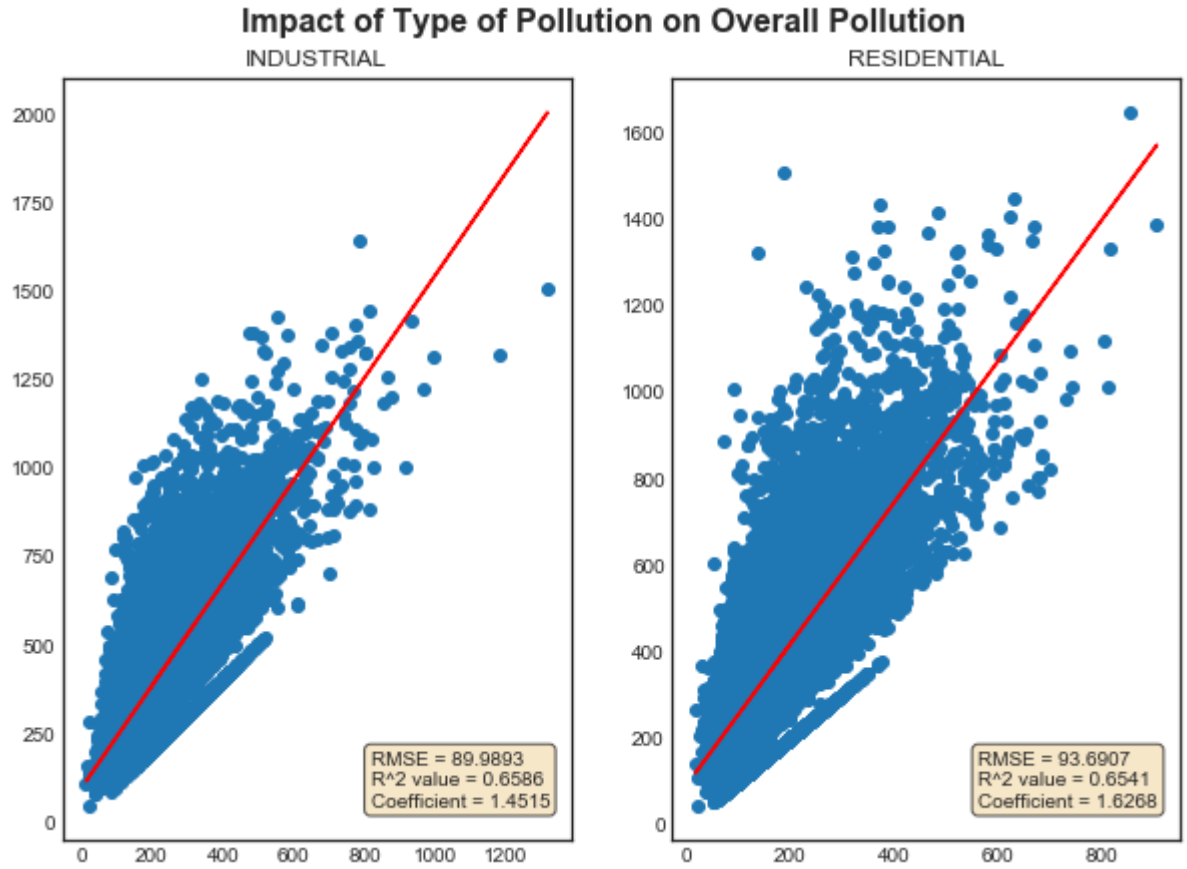
"Some particulates occur naturally, originating from volcanoes, dust storms, forest and grassland fires, living vegetation and sea spray. Human activities, such as the burning of fossil fuels in vehicles, stubble burning, power plants, road dust, wet cooling towers in cooling systems and various industrial processes, also generate significant amounts of particulates."

So there are multiple factors at play that could increase pollution, let's study those!

Which parts of the city emit the most pollutants? If we take the top 30 most polluted cities in the country of all time and plot the total yearly pollution levels from Residential, Industrial or Sensitive Areas, we get some interesting results.

- Nearly all cities have higher pollution levels in industrial areas.
- All cities show a drop in pollution levels after 2013. This could be due to new vehicle emission standards, increasing use of LPG as domestic fuel instead of coal or fuelwood, use of CNG instead of diesel etc
- After 2013, there is no consistent drop in emissions suggesting that the laws have not been stringent enough and governments have not followed through their long term plans of improving air quality in their respective cities.

Intuitively one would think that industrial pollution would be much higher than emissions in residential areas but residential areas are almost at par and even higher in some years so I ran a simple regression model on the data to further analyse the impact of the source on total pollution levels in these cities. I am not considering data from Sensitive areas as it sparse (96% nulls). The results are as follows:



The overall quality of air is more sensitive to changes in pollution levels from residential areas.