

Insights from Climate Change Data

Surabhi S Nath

2016271

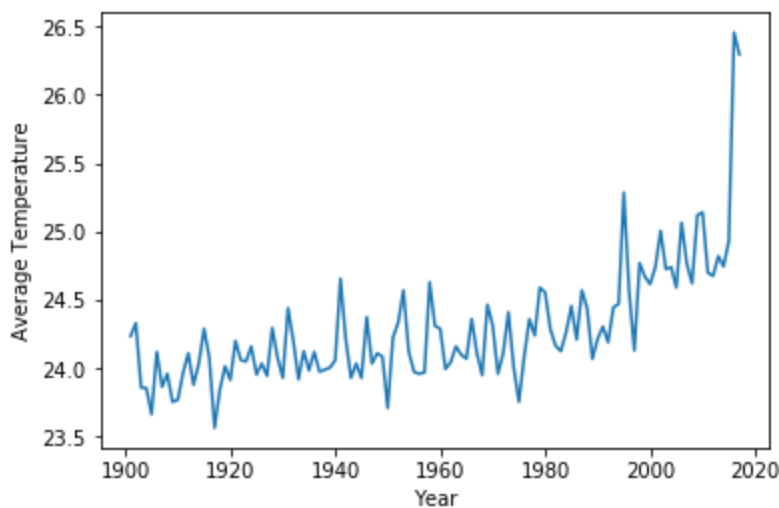
<https://github.com/surabhisnath/Data-Science-Work>

India's Story

We have heard a lot about India's changing climate and weather patterns. Data analysis can validate these claims and make for proper judgement for severity of the situation. From this, we can then plan and take adequate measures in an educated manner.

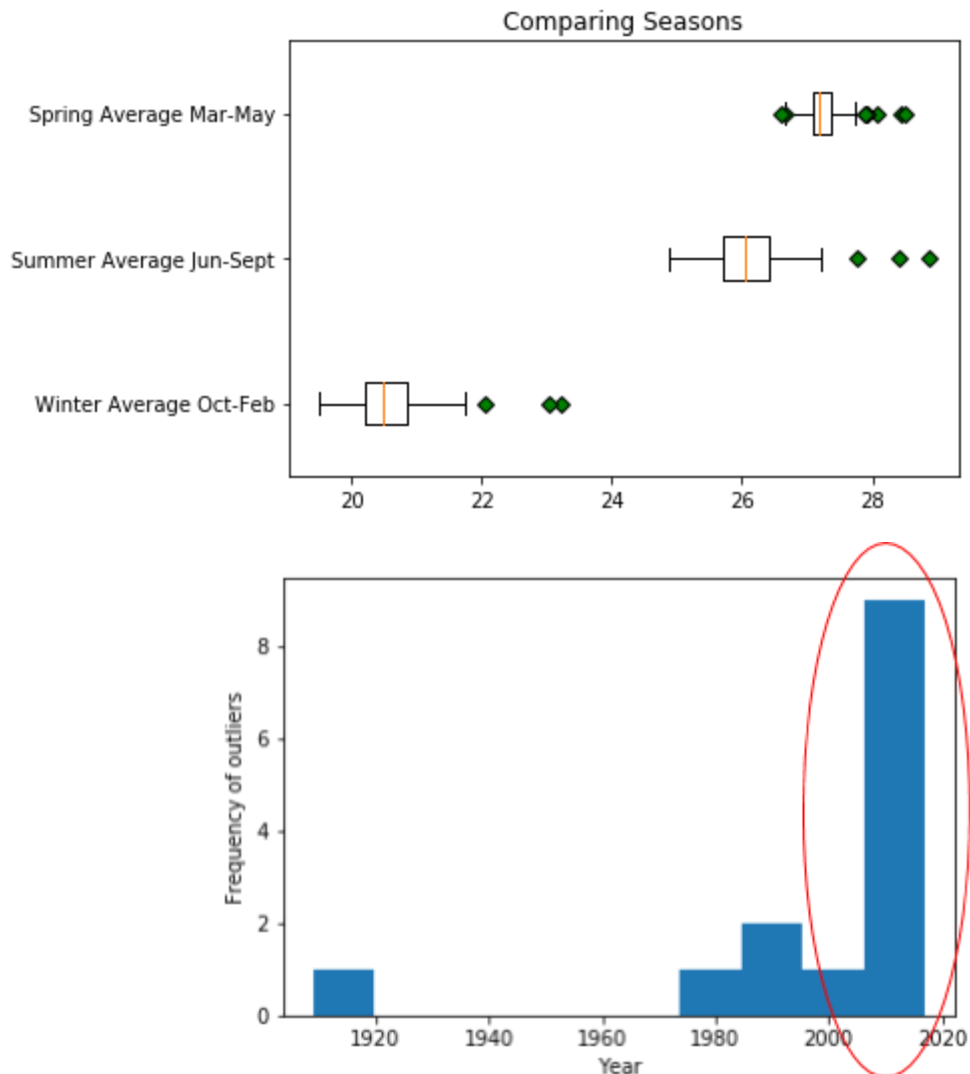
Data sources for this analysis report include datasets published by Government of India and Kaggle.

On plotting the average annual temperature for the whole country over 120 years, the plot indicates that truly the temperatures display an increasing trend. This justifies why Global Warming today is a major concern.

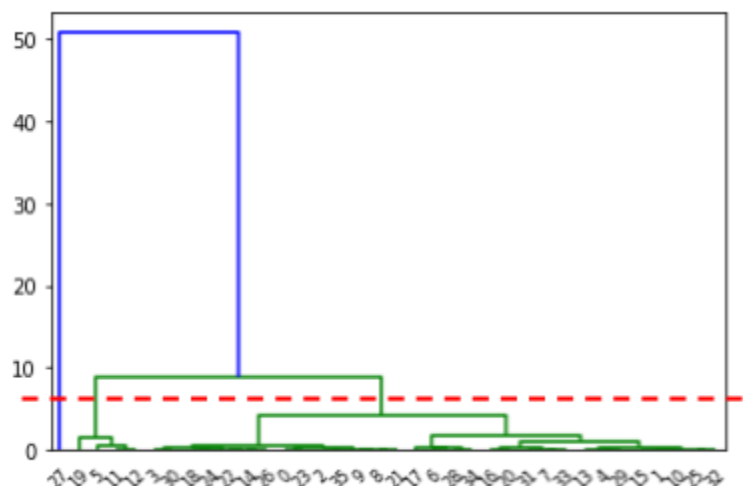


India mainly experiences 3 seasons through the year which are - Winter, Summer and Spring. When we study the variation in temperatures across seasons, we are clearly able to distinguish that temperatures in winter are much lower than in summer and spring. An interesting analysis here will be to build a boxplot. A boxplot can successfully detect outliers. The outlier temperatures are mostly higher than the remaining

distribution. On studying the outliers we see that, they are mainly the temperature data of the most recent years (>2010). This is also an indicator of rising temperatures.



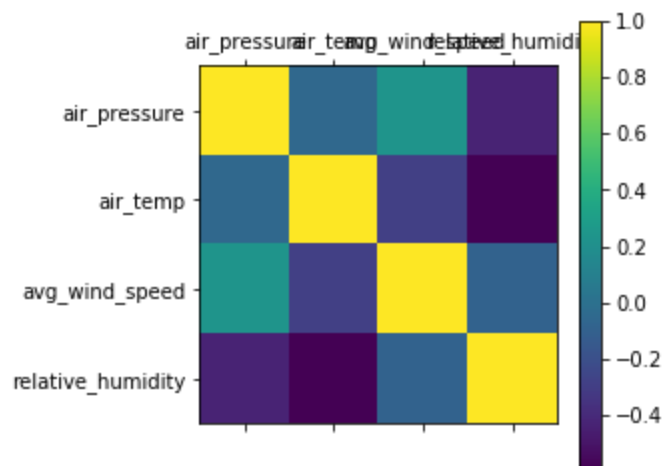
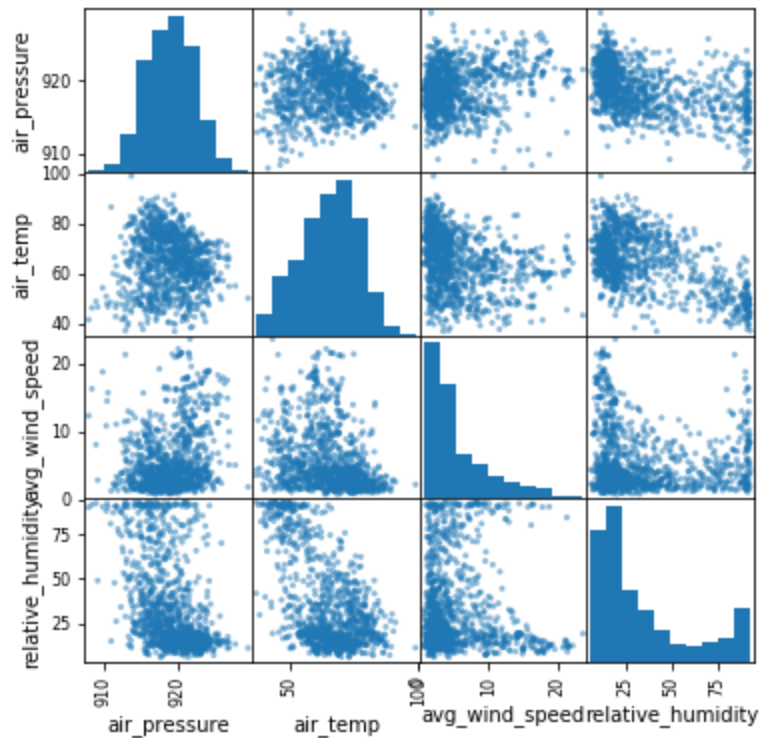
Another major reason for climate change is the major deforestation for construction and development. On studying the % of forest land cleared in each Indian state over 38 years from 1980 to 2018, we see that, on an average, **0.315%** of forested land has been removed. We can group this data into k clusters to identify the more serious states. For choosing the optimal k, we can make a dendrogram and check the number of vertical lines above some threshold. From this plot, and threshold 10, optimal k comes out to be 4.



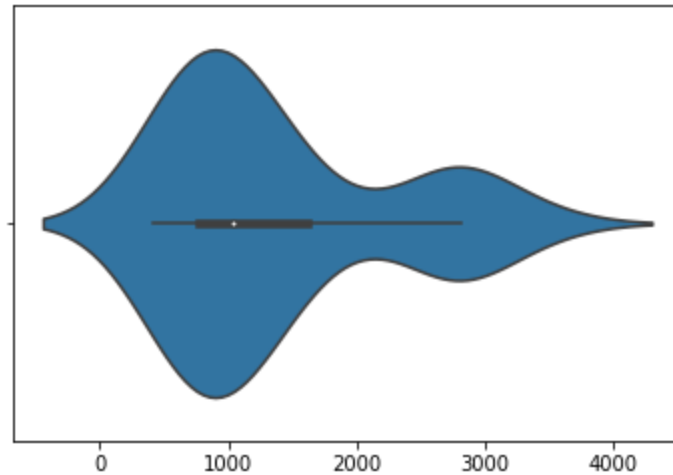
Following this, performing Kmean with $k = 4$ and Agglomerative clustering with $k = 4$ gave the exact same division of states:

Cluster 1 Least deforestation	Cluster 2	Cluster 3	Cluster 4 Max deforestation
'Andaman and Nicobar Islands' 'Arunachal Pradesh' 'Assam' 'Daman and Diu' 'Delhi' 'Jammu and Kashmir' 'Lakshadweep' 'Manipur' 'Meghalaya' 'Mizoram' 'Nagaland' 'Pondicherry' 'Tamil Nadu' 'West Bengal'	'Andhra Pradesh' 'Bihar' 'Chhattisgarh' 'Dadar and Nagar Haveli' 'Goa' 'Himachal Pradesh' 'Jharkhand' 'Karnataka' 'Kerala' 'Maharashtra' 'Orissa' 'Rajasthan' 'Sikkim' 'Telangana' 'Tripura' 'Uttar Pradesh' 'Uttarakhand'	'Chandigarh' 'Gujarat' 'Haryana' 'Madhya Pradesh'	'Punjab'

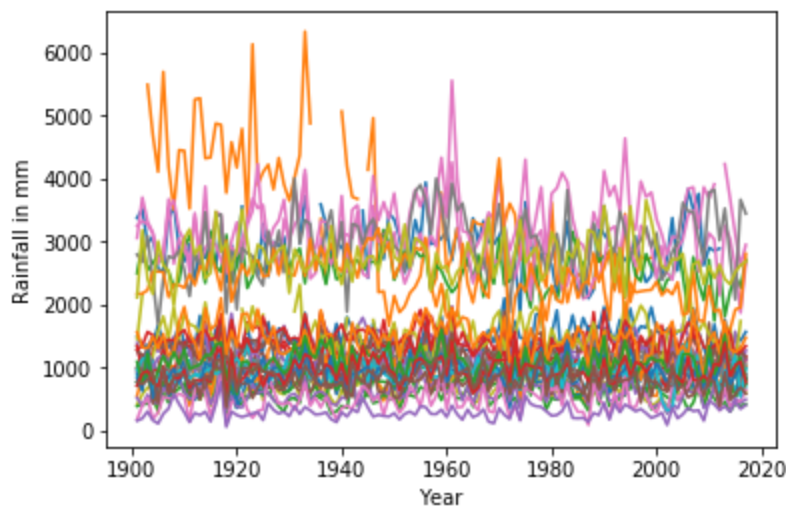
We also want to observe the patterns and correlations between Humidity, Air pressure, Air temperature and Wind speed since these are correlated factors of climate for the same time instant. Through the correlation plot and heatmap, we can see that there is some correlation between Air temperature and Humidity and also some between Air pressure and Humidity.



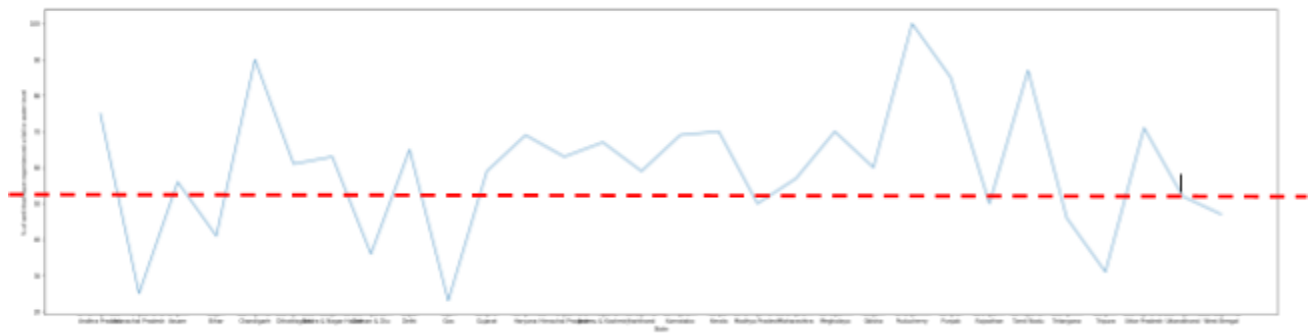
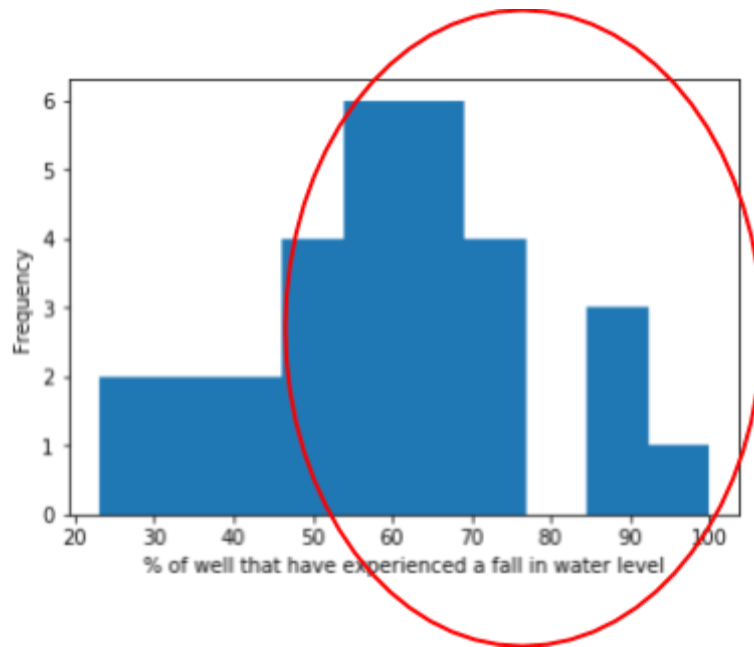
Now, we look at the rainfall in India. We can visualize the distribution using a violin plot. The plot shows the density distribution of rainfall in various states for the year 2017. The multimodes show that there is a high frequency of states with annual rainfall around 1000 and around 3000 mm.



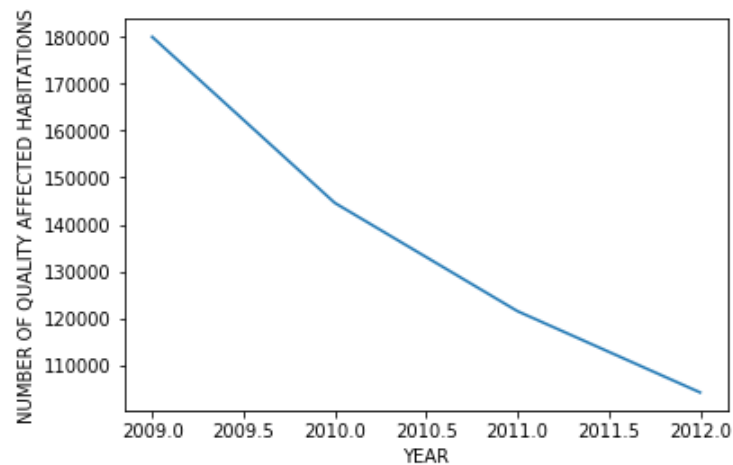
When we plot the rainfall of the various states over 120 years, we don't see too much fluctuation. This means the rainfall has not declined very significantly over time.



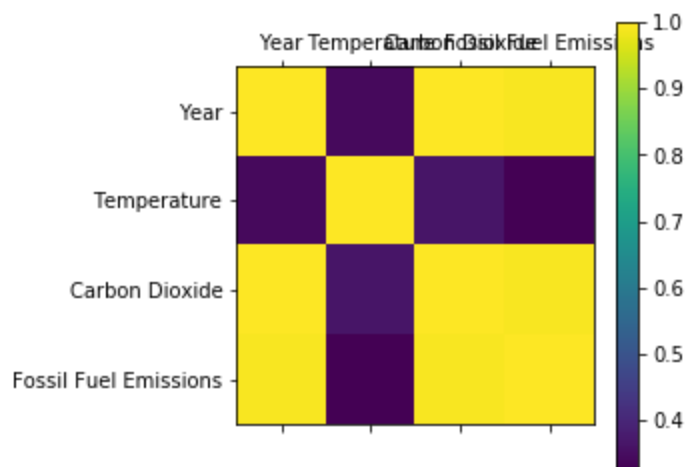
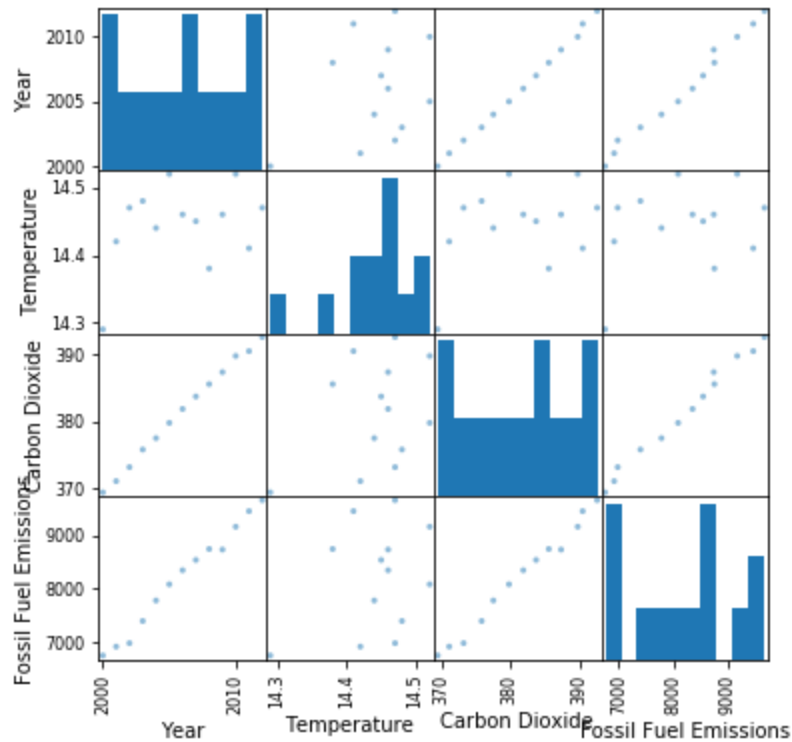
However, groundwater levels have depleted over the years. This could be due to an increase in built-up land, not allowing rainwater percolation or due to rising temperatures. Data on the % of wells with a decline in water level for every state reveals appalling conditions of groundwater depletion. A large fraction of states have over 60% of wells with a decline in water level.



On the other hand, with respect to quality of water, it is pleasing to see that across the last 4 years, the number of habitations with water poor quality have reduced by a large margin.

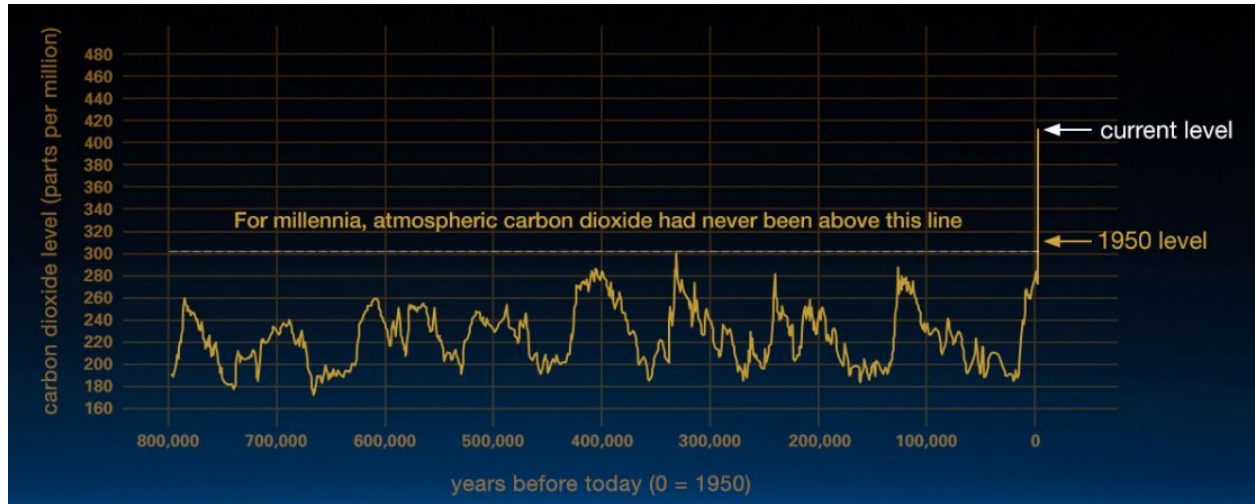


By analyzing pollution levels and air quality, we find that there have been an increase in carbon dioxide levels and emissions fuel across years. This can be a direct influencer for rising temperatures. We see a near perfect correlation between years and CO₂; years and fuel emissions; and CO₂ and fuel emissions.

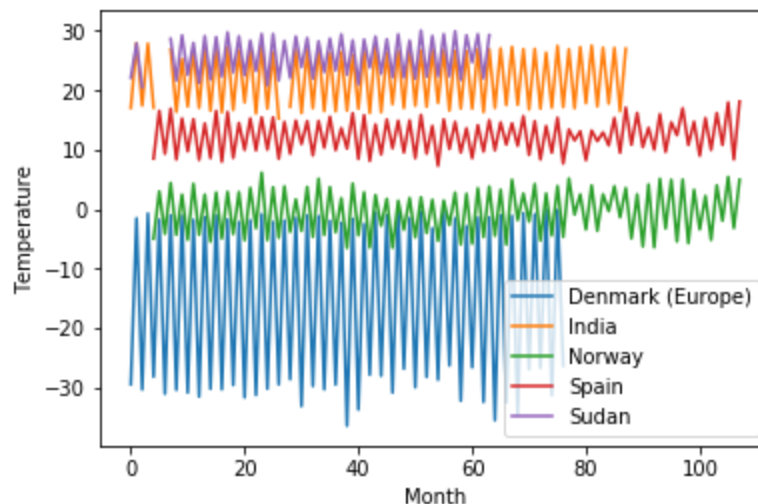


Global Scenario

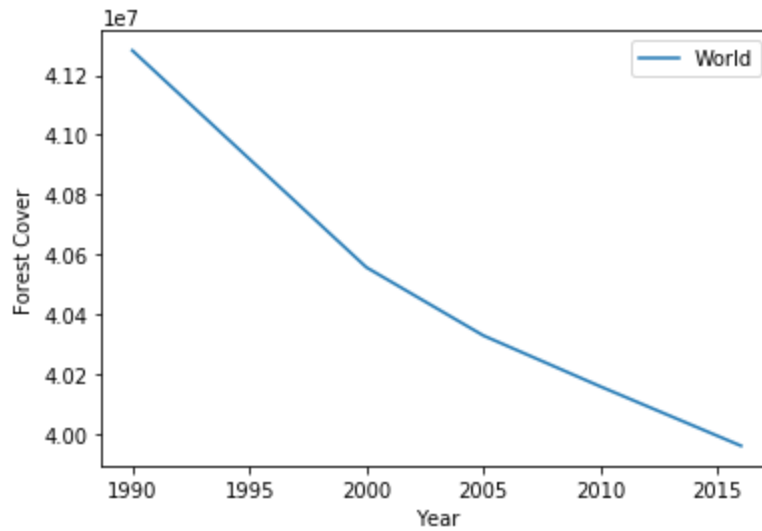
NASA has published that the carbon dioxide content in the air today has surpassed shocking thresholds. Polar ice is melting, ocean temperatures and sea levels are rising, snow cover is depleting. All these are indicators of global warming and climate change.



On plotting temperatures across months in a few of the hot/cold countries, we see that recently, the temperatures are higher than before.



The forest cover in the whole world over 25 years has reduced significantly.



All these indicators imply the seriousness of the issue. We should be conscientious citizens and be careful in handling our resources. Each of our small contributions can make the planet to a safer place to live in.