IBM CAPSTONE ASSIGNMENT

CO2 Emission Study

IEA data study for Year 2015

Smita Khapre

5/30/2019

Contents

Problem:

Increasing CO2 Emissions across the globe has become a major problem. It has caused Global Warming, This has caused irreversible damage to environment. Most of the CO2 emissions come from fuel combustions in different sectors. The International Energy Agency (IEA), an autonomous agency, was established in November 1974. IEA collects the CO2 fuel combustions data of all the countries in all sectors across the globe. Most of the countries releasing maximum CO2 in atmosphere chose to move their manufacturing to developing countries in Asia and South America to reduce their pollution levels. Globally this does not solve the Global Warming problem. As a matter of fact, it is further enhancing the Global Warming issue. As these countries have no or little regulations on pollutants. The CO2 emissions from transportation (marine and aviation) of the manufactured goods from latter to former countries is also added to the environment.

The Background:

Global Warming and Climate Change are the known issues in current World. We are experiencing more frequent storms, Tsunamis, Blizzards, Tornadoes, extreme temperatures, droughts, floods, Polar ice melting causing sea level rise, Ozone depletion and so on all over the Globe. All of these are marked by human activities causing carbon compounds and CO2 emissions. To view the data from 1971 to 2018, refer to https://webstore.iea.org/co2-emissions-from-fuel-combustion-2018. Same is used in details in the CO2 Data Emisssions Study report. We have used the Data analysis techniques to understand the spread of this issue across the Globe.

Data Analysis Methodology:

1. Choropleth mapping shows the CO2 emission spread in different countries.
2. Data Wrangling, Standardization and Normalization
3. K-Mean Clustering
4. Four Square APIs to check on one of the maximum CO2 emission per capita on the venues available.

Description of the data

I have used the data from IEA website referring to
www.iea.org/publications/freepublications/publication/CO2-emissions-from-fuel-combustion-highlights-2017.html

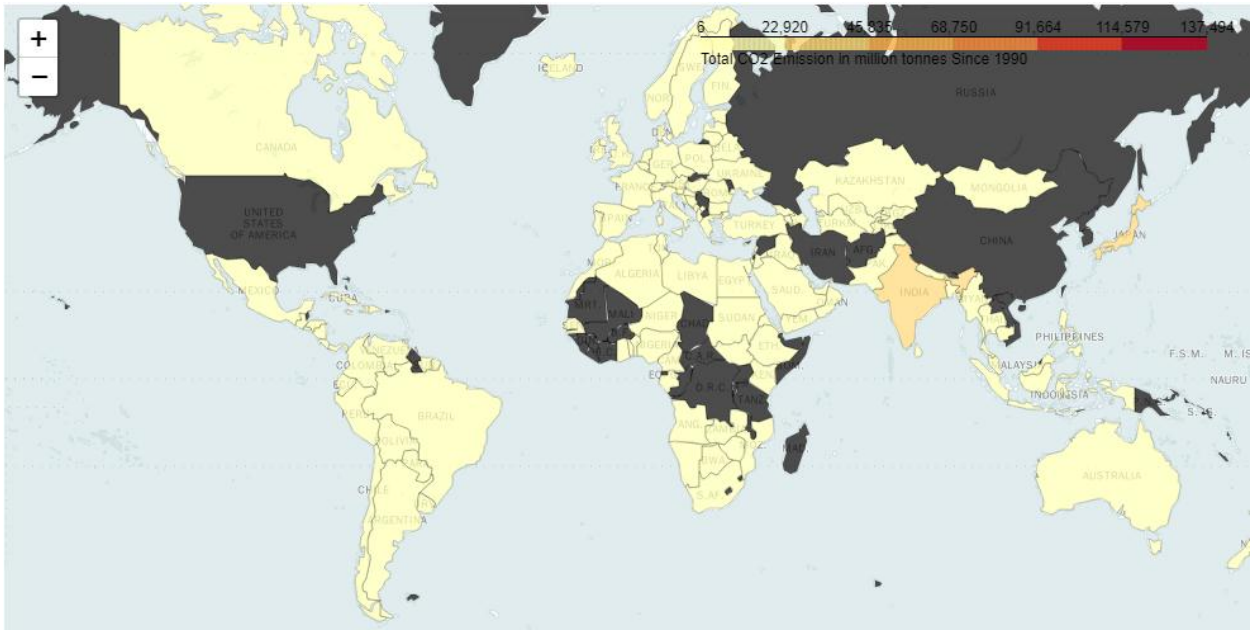
It contains following data for most of the countries across the globe from 1990 to 2015.

1. CO2 emissions total and different sectors – Refer to https://github.com/smitakh/Coursera_capstone/blob/master/CO2FC_90.csv
2. Population – Refer to https://github.com/smitakh/Coursera_capstone/blob/master/POP_90.csv
3. GDP – Refer to https://github.com/smitakh/Coursera_capstone/blob/master/GDP_90.csv
4. Per Capita CO2 emissions Population – Refer to https://github.com/smitakh/Coursera_capstone/blob/master/CO2POP_90.csv
5. Per Capita CO2 emissions sector-wise - – Refer to https://github.com/smitakh/Coursera_capstone/blob/master/SECTOR.csv https://github.com/smitakh/Coursera_capstone/blob/master/SECTPOP.csv

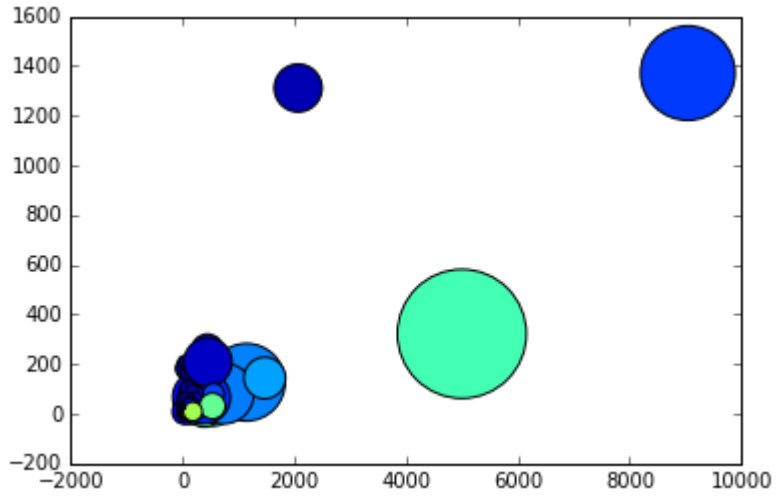
Choropleth Mapping of CO2 Emissions across the world

In Year 2015 CO2 Emission in Million tons is shown below

Total cumulative CO2 Emission in Million tons since 1990 to 2015 is shown below



Scatter Plot



Above clearly shows 3 outliers.

Clustering: Other Factors to be considered

But only looking at country-wise emissions and trying to solve this problem is not a wise option. As we have country-wise population, GDP, per capita emissions also. We must take each of it into consideration. Following table shows the combined data for year 2015 of five countries out of 146.

	CO2 Emission	Population	GDP	CO2POP_KG
Country				
Canada	549.2	35.9	1796.4	15.32
Chile	81.6	18.0	263.1	4.52
Mexico	442.3	121.0	1207.7	3.66
United States	4997.5	321.7	16597.4	15.53
Australia	380.9	24.1	1485.3	15.83

Also looking at the choropleth mapping, we derive that we have a few segments only. Let's run K-means clustering on the data and look at the outliers and other segments as follows:

Clus_km	Total Countries
0	128
1	1
2	1
3	6
4	9
5	1

The Clusters marked in Yellow shows the Outliers. Details are below

	CO2 Emission	Population	GDP	CO2POP_KG	Clus_km
Country					
People's Rep. of China	9040.7	1371.2	8909.8	6.59	1.0
United States	4997.5	321.7	16597.4	15.53	2.0
Japan	1141.6	127.0	5986.1	8.99	4.0

Looking at other 3 segments

	CO2 Emission	Population	GDP	CO2POP_KG	Clus_km
Country					
Russian Federation	1469.0	144.1	1723.9	10.19	3.0
Korea	586.0	50.6	1266.6	11.58	3.0
Canada	549.2	35.9	1796.4	15.32	3.0
Mexico	442.3	121.0	1207.7	3.66	3.0
Indonesia	441.9	257.6	987.5	1.72	3.0
Australia	380.9	24.1	1485.3	15.83	3.0
Turkey	317.2	77.5	1087.6	4.10	3.0
Spain	247.0	46.4	1414.9	5.32	3.0
Netherlands	156.0	16.9	868.3	9.21	3.0

	CO2 Emission	Population	GDP	CO2POP_KG	Clus_km
Country					
France	290.5	66.5	2777.5	4.37	5.0
Germany	729.8	81.7	3696.6	8.93	5.0
Italy	330.7	60.7	2059.5	5.45	5.0
United Kingdom	389.8	65.1	2682.3	5.99	5.0
India	2066.0	1311.1	2296.6	1.58	5.0
Brazil	450.8	207.8	2330.4	2.17	5.0

And 128 more countries in other general cluster ‘0’

Thus there is a need to look at CO2 emission per capita in kilogram data too. Below are the top ten countries producing maximum data. That means while as a country in whole they are not standing out but per population wise it does.

Country	CO2 Emission	Population	GDP	CO2POP_KG	Clus_km
Qatar	79.9	2.2	167.0	35.77	0
Curaçao	4.9	0.2	1.8	30.72	0
Kuwait	85.4	3.9	139.7	21.93	0
Bahrain	30.1	1.4	30.8	21.83	0
United Arab Emirates	180.2	9.2	360.0	19.68	0
Gibraltar	0.6	0.0	1.2	17.39	0
Saudi Arabia	531.5	31.5	672.2	16.85	0
Trinidad and Tobago	22.8	1.4	22.7	16.76	0
Australia	380.9	24.1	1485.3	15.83	4
United States	4997.5	321.7	16597.4	15.53	2

With this we can conclude that along with the outlier, we need to focus on these ten countries too.

Four Square Analysis

When we run Four Square data on Qatar, we just get below

	name	categories	lat	lng
0	Elaaaj E Alam	Afghan Restaurant	25.354826	51.183884
1	Sivas	Pub	25.354452	51.184330

Let's see the sector-wise data for Qatar

Country	Total CO2 emissions from fuel combustion	Electricity and heat production	Other energy industry own use*	Manuf. industries and construction	Transport	of which: road	Other sectors	of which: residential
Qatar	79.9	20.2	30.9	13.6	15	15	0.3	0.3

This shows that other industries and energy production costs 60% of its emission.

Results and Discussion:

We saw that the CO2 emissions concentrations are very high in certain countries. Where as other countries CO2 emission concentrations are very low. If we see the choropleth mapping we see contrasts of colors. This means if we focus on the darker shaded countries and understand the reasons of the emissions and try to reduce them.

Point to be noted is that the higher CO2 emission countries also have higher GDP. Thus based on the Global GDP and Global needs, we need to understand the combustion reduction methods. As transportation and Energy requirements would continue to increase, Utilization and efforts in Green energy is required. Its implementation on these countries and necessary support by the countries who can support is needed to reduce combustion. This would be the way to collectively reduce the CO2 emissions Globally. Only Restrictions on CO2 emissions would not be the fruitful method to do it !!!

Conclusion:

As per our study, China and United States definitely stands out against entire world as it caused nearly 43% of entire world's emission. But they contributed to 33% of worlds GDP.

Nevertheless, we found top ten countries with most CO2 emission which need to focus on the reduction efforts. Best way is to discourage the fossil fuel energy and encourage the green energy !!!

Thank You !!!