**Data Visualization Project**

**Carbon Dioxide Emissions with Respect To Different Factors**

**Rekkala Venkata Satya Sai Lakshmi Lavanya-U72254994**



**Motivation behind the Project**

According to the data extracted from the UN world bank, air pollution stands third among the risk factors that impact the death rate for almost two decades. Therefore, it is time to give some thought to air quality and it’s importance. CO2 is not considered as a classic air pollutant, as it contributes to the sustainability of living forms by helping in combustion, photosynthesis etc. However, the measures to control the concentration of CO2 still receive a lot of attention because increasing trends in CO2 emissions are now impacting the global climate and they are adversely affecting the human health just like the traditional air pollutants. Also, the Union of concerned scientists in USA published in their website that “Though there are many heat-trapping gases in the atmosphere CO2 puts us at the greatest risk of irreversible changes as it sticks around in the atmosphere for a longer time.” The long-lasting negative impacts of CO2 made me choose “CO2 Emissions W.R.T Different factors” as my topic of discussion.

**Introduction**

I started off by supporting the points mentioned in the Motivation section. Further, I chose relevant tables by taking into consideration the data that has its units as KMTCO2e (thousand metric tons of CO2 equivalent). The datasets suggest the atmospheric impact of different greenhouse gases in terms of the impact created by 1000 metric tons of CO2. Further different factors and the possibility of relation between these factors and the gas emissions has been checked through different visualizations. These visualizations do not imply any causation i.e., they only suggest possible correlation. Finally, different visualizations have been clubbed onto dashboards in order to avoid hustling through different worksheets.

**Methodology**

I picked up four datasets to begin with and then I included more datasets. I used a total of 7 datasets in this project and the sources of these datasets are as follows:

1] <https://ourworldindata.org/causes-of-death>

2] <https://ourworldindata.org/causes-of-death>

3] <https://data.worldbank.org/indicator/EN.ATM.CO2E.KT>

4] <https://data.worldbank.org/indicator/EN.ATM.METH.KT.CE>

5] <https://data.worldbank.org/indicator/EN.ATM.GHGO.ZG>

6] <https://data.worldbank.org/indicator/EN.ATM.METH.KT.CE>

7] <https://data.worldbank.org/indicator/SP.POP.TOTL>

I initially cleaned all the datasets and removed the countries that had null values for all the years. I noticed that the countries with null values are mostly countries that had very less population and Industrialization. Hence, I assumed that they wouldn’t impact my end results much. Further, I combined the data related to different types of greenhouse gases (namely Methane, CO2, Nitrous Oxide and other greenhouse gases (Perfluorocarbons, Hydrofluorocarbons and SF6 etc...)) and placed them in a single worksheet in my excel workbook. I pivoted the columns as my years were populated as Row headers. Also, I made sure that different gas emissions in same countries are summed up while plotting the pie chart in order to avoid any bias. Also, I normalized my greenhouse gas emissions with respect to population and this makes my visualizations more reliable. All the required data is placed in different sheets of a single excel workbook. The data related to population, total greenhouse gas emissions and GDP per capita was cleaned in a similar fashion before finalizing my processed dataset. All the data preparation was done using the Tableau Prep Builder before moving onto the visualizations. All relevant connections as shown below were made on the Data source tab of the Tableau application before proceeding further.

Data\_Table connection

Graphical user interface

Description automatically generated

**Analysis**

**Visualization-1**

Timeline

Description automatically generated

The above visualization backs up the points mentioned in the Motivation behind the project. It can be seen from the column chart that Air pollution is responsible for huge number of deaths every year. The color palette in Tableau has only 20 colors. Therefore, only top 20 risk factors are included in order to avoid any confusion caused by repetitive colors. Also, the animation (play and stop buttons) available on the right side of the worksheet will help us navigate through years. In this case, I have considered data from 1990 to 2019 (i.e., for 29 years) and almost in all the years number of deaths due to air pollution stands at the 3rd position.

**Visualization-2**

Chart, line chart

Description automatically generated

In this Visualization, I plotted the CO2 concentrations in different years from 1991 to 2022 and from this visualization I could conclude that the CO2 concentration in the atmosphere has been rapidly increasing from the year 1991 to 2022.

**Visualization-3**

Map

Description automatically generated

In visualization 3 of my project, I plotted normalized total greenhouse gas emissions in different parts of the world with respect to the GDP per capita. The emissions were normalized with respect to population since it is a known fact that countries with high population will have larger greenhouse gas emissions. As per the map shown above, I figured out that countries with high GDP will have more emissions and this can be due to various factors like Industrialization, automobiles, population etc. that impact the GDP.

**Visualization-4**

**Chart, line chart

Description automatically generated**

In Visualization 4, I tried to prove the conclusions drawn from visualization 3 by plotting the total greenhouse gas emissions against GDP per capita value of the world from 1992 to 2012. From this graph also I could say that total greenhouse gas emissions increased as the GDP per capita increased.

**Visualization-5**

**Chart, pie chart

Description automatically generated**

As per our readings we could say that other greenhouse gases and nitrous oxide cause negative CO2 e emissions. This means that CH4 and other greenhouse gases try to reduce the impact of positive CO2 emissions in the atmosphere by eliminating the main greenhouse gas CO2 from the atmosphere. However, it is not suggested to encourage the emissions of greenhouse gases in order to balance out the CO2 concentration in the atmosphere as these are equally harmful. Therefore, the most appropriate way to eliminate CO2 from the atmosphere is to plant more trees that break down CO2 into oxygen through photosynthesis. From the above pie chart, it can be concluded that positive CO2 are way greater than the negative emissions. Hence, it is important to curb the levels of CO2 in the atmosphere by adapting more eco-friendly.

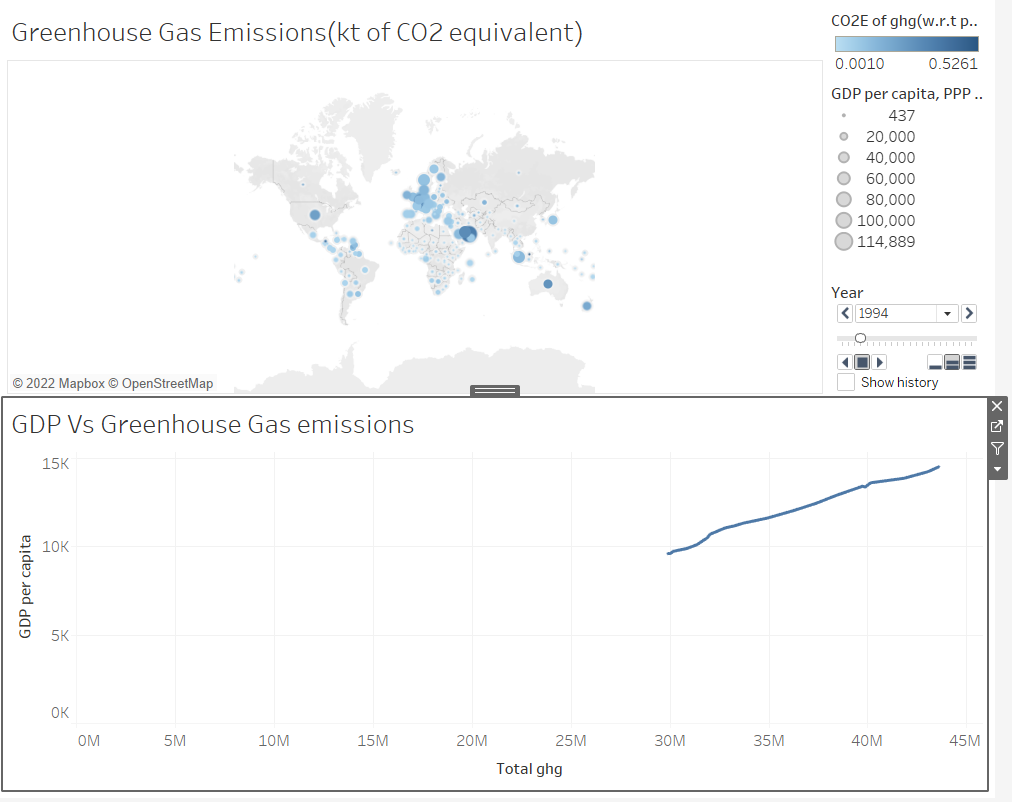
**Visualization-6**

**Chart, line chart

Description automatically generated**

The visualization 6 shows the path followed by different greenhouse gas emissions. This visualization supports the conclusions drawn through visualization 5. The rate at which CO2 concentration increases in the atmosphere is greater than the rate at which it is eliminated through greenhouse gas emissions.

**Dashborad-1**



This dashboard displays the correlation between GDP and greenhouse emissions.

Chart

Description automatically generated

This dashboard displays the measure of different gas emissions from 1991 to 2012.

**Conclusions**

From this project we can conclude that the amount of CO2 being added to the atmosphere is dramatically increasing over the years. Also, we can say that there is a linear relationship between the GDP and the CO2 emissions.

**Future Research questions**

What is rate at which global temperature increases when CO2 is added to the atmosphere?

What is the impact of industrialization on CO2 emissions?

How will deforestation impact the levels of CO2 in the atmosphere?

What is the impact of fossil fuel burning on CO2 levels?