Visualising CO2 Pollution Trends

Tyagi Mahir

2022-11-02

I have used data from from Kaggle and The World Bank. Links to the datasets are https://www.kaggle.com/datasets/ruchi798/global-environmental-indicators and https://data.worldbank.org/indicator/NY.GDP.PCAP.CD respectively.

ABSTRACT
In this project, we looked at data pertaining to CO2 emissions of various countries, factors that would predict their levels and possible courses of action. We used tools such as visualization and correlations to draw our conclusions.
We found that prosperity predicts per capita levels of CO2 emission, that the use of renewable energy inversely predicts CO2 emissions and that the Kyoto Protocol has had a definite impact on emissions of greenhouse gases.

INTRODUCTION
One of the biggest problems plaguing the modern world is climate change. It presents many challenges, one of which is that the world's temperature has increased by 1.1°C since 1880. One of the reasons for this is levels of CO2 in our atmosphere. Since 1970, CO2 emissions by humans have increased by about 90%. In this project, we will look at the data to create our conclusions as to why CO2 levels are rising, and what we can do about it.

DATASETS
Here we have used 3 data sets:

-*CO2*: CO2 emissions of 192 United Nations Members, in 1000 Tonnes. Rows are values and columns are years.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
CO2_Emissions2 <- read.csv("C:/Users/heyit/OneDrive/Desktop/Mahir/DS1,2022/visuproject/Air and Climate/
# Cleaning up some data, changing column names to year and
# converting data to numeric.
for (i in 3:32) {
    colnames(CO2_Emissions2)[i] = CO2_Emissions2[1, i]
}
CO2 = CO2_Emissions2[-c(1), ]
```

```
CO2 = CO2[-c(1)]
CO2 = CO2 %>%
    mutate_all(list(~str_replace_all(., ",", "") %>%
        as.numeric %>%
        replace_na(0)))
CO2$Country = CO2_Emissions2$Country[-c(1)]
```

-*GDPPC*: GDP per capita of various countries and others, 1960-2021.

```
GDPPC <- read.csv("C:/Users/heyit/OneDrive/Desktop/Mahir/DS1,2022/visuproject/GDPPC.csv")
```

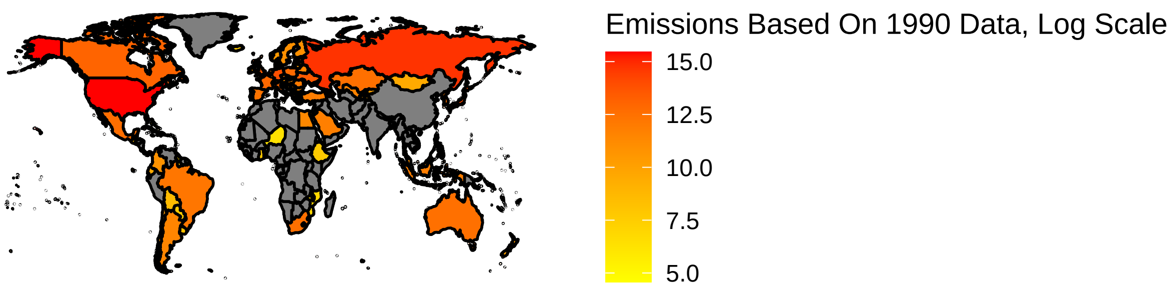
-*renewable*: Proportion of energy produced that is renewable of various countries, in %.

```
renewable <- read.csv("C:/Users/heyit/OneDrive/Desktop/Mahir/DS1,2022/visuproject/Energy and Minerals/Re
```

GRAPH 1 - COUNTRIES WITH HIGHEST EMISSIONS

The first question that piques our curiosity is, what is the distribution of CO2 pollution by country?

```
# Downloading the world map and joining it with our data
mapdata = map_data("world")
colnames(mapdata)[5] = "Country"
mapdata = left_join(mapdata, CO2, by = "Country")
mapdata = mapdata %>%
    filter(!mapdata$Country == "Antarctica")
# Plotting
ggplot(mapdata, aes(x = long, y = lat, group = group)) + geom_polygon(aes(fill = log(mapdata$"1990")),
    color = "Black") + scale_fill_gradient(name = "Emissions Based On 1990 Data, Log Scale ",
    low = "yellow", high = "#ff0000", na.value = "grey50") +
    theme(axis.text.x = element_blank(), axis.text.y = element_blank(),
        axis.ticks = element_blank(), axis.title = element_blank(),
        aspect.ratio = 1/2, rect = element_blank())
```

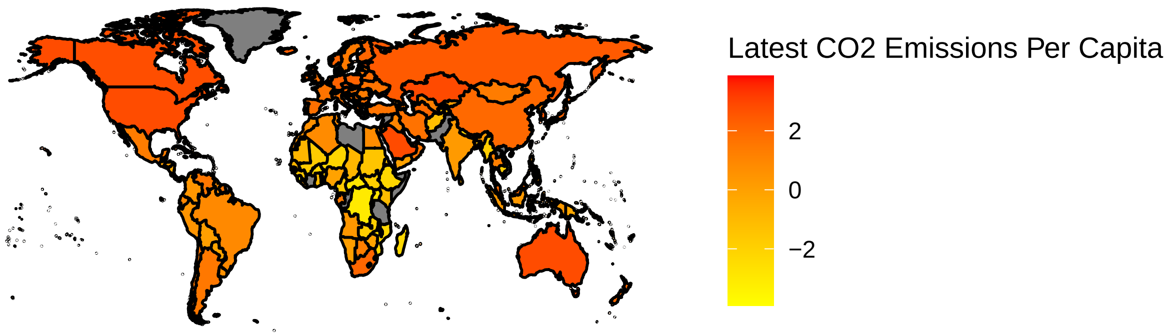


As you can see, many values are missing(shown by gray colour). Also, Countries with a higher population have higher emissions, creating a bias. Let us try something new.

GRAPH 2: PER CAPITA DATA

Using emissions per capita hels us get a clearer picture. Also, there is more data to work with.

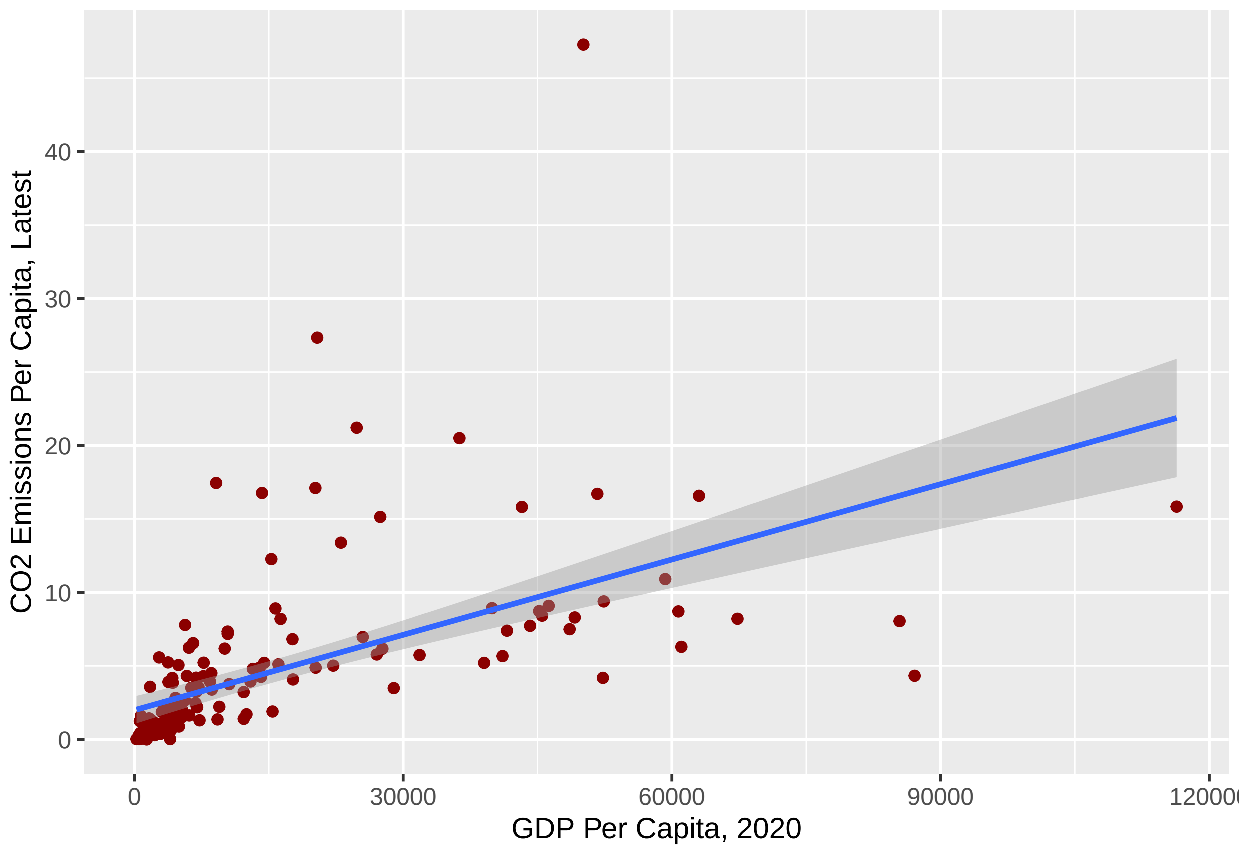
```
# Plotting
co2map = ggplot(mapdata, aes(x = long, y = lat, group = group)) +
    geom_polygon(aes(fill = log(CO2.emissions..per.capita...latest.year)),
        color = "Black") + scale_fill_gradient(name = "Latest CO2 Emissions Per Capita ",
    low = "yellow", high = "#ff0000", na.value = "grey50") +
    theme(axis.text.x = element_blank(), axis.text.y = element_blank(),
        axis.ticks = element_blank(), axis.title = element_blank(),
        aspect.ratio = 1/2, rect = element_blank())
co2map
```



Here, you can see that industrialized nations seem to have more emissions, we will test this hypothesis.

Graph 3 - GDP AND EMISSIONS
We will try to find a relationship between GDP and emissions.

```
# Preparing
colnames(GDPPC)[1] = "Country"
GDPPC = left_join(GDPPC, CO2, by = "Country")
GDPPC = GDPPC %>%
    filter(!is.na(GDPPC$X2020))
GDPPC = GDPPC %>%
    filter(!is.na(GDPPC$CO2.emissions..per.capita...latest.year))
GDPPC = GDPPC[-c(90), ]  #Removing outlier for plotting
# Plotting
ggplot(GDPPC, aes(x = X2020, y = CO2.emissions..per.capita...latest.year)) +
    geom_point(colour = "Dark Red") + labs(y = "CO2 Emissions Per Capita, Latest",
    x = "GDP Per Capita, 2020") + stat_smooth(method = "lm",
    formula = y ~ x, geom = "smooth")
```




```
# Finding correlation
cor(GDPPC$X2020, GDPPC$CO2.emissions..per.capita...latest.year)
```

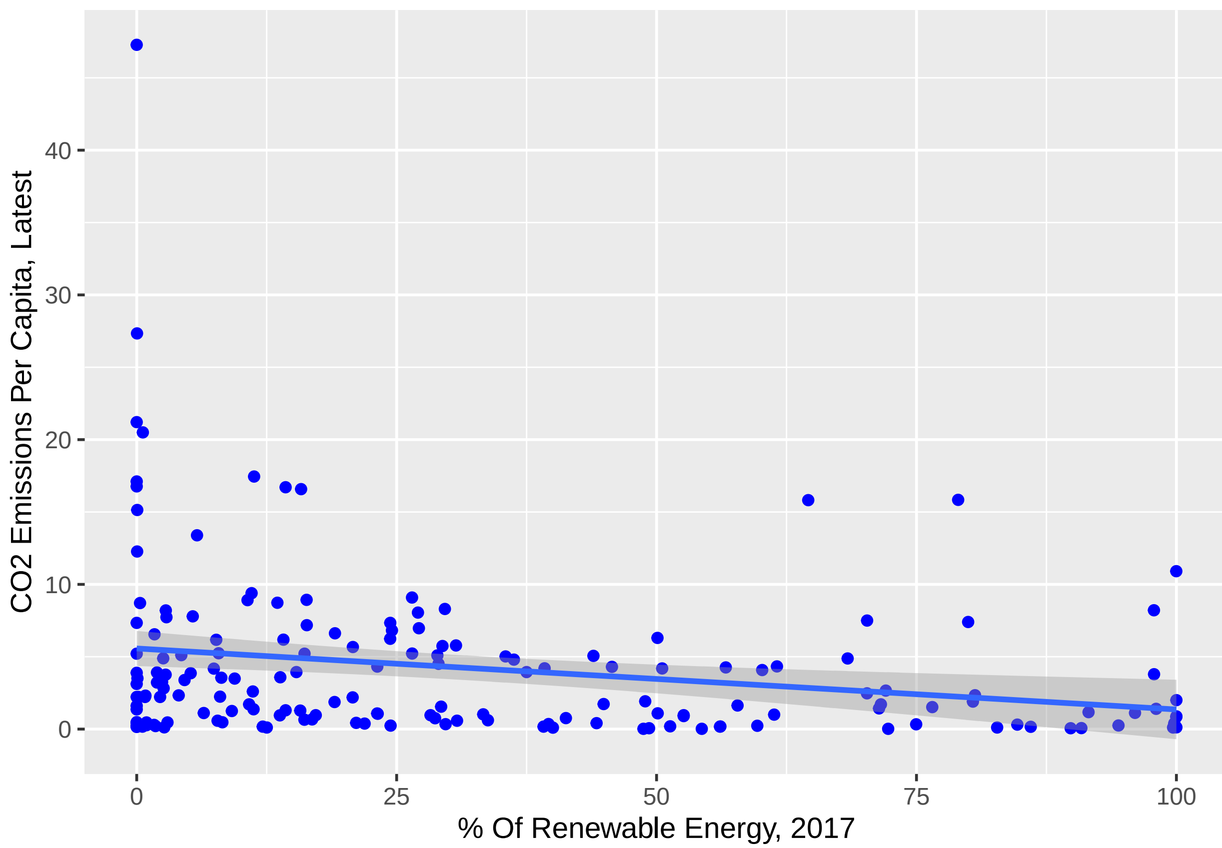
```
## [1] 0.5691422
```

Here we see there is a strong relationship between GDP Per Capita and CO2 emissions.

GRAPH 4- IMPACT OF RENEWABLE ENERGY

We will see if the usage of renewable energy can help us solve the problem indicated by the last graph.

```
colnames(renewable)[2] = "Country"
renew = left_join(CO2, renewable, by = "Country")
renew = renew %>%
    filter(!is.na(X2017))
ggplot(renew, aes(X2017, CO2.emissions..per.capita...latest.year)) +
    geom_point(colour = "blue") + labs(y = "CO2 Emissions Per Capita, Latest",
    x = "% Of Renewable Energy, 2017") + stat_smooth(method = "lm",
    formula = y ~ x, geom = "smooth")
```




```
# Finding correlation
cor(renew$CO2.emissions..per.capita...latest.year, renew$X2017)
```

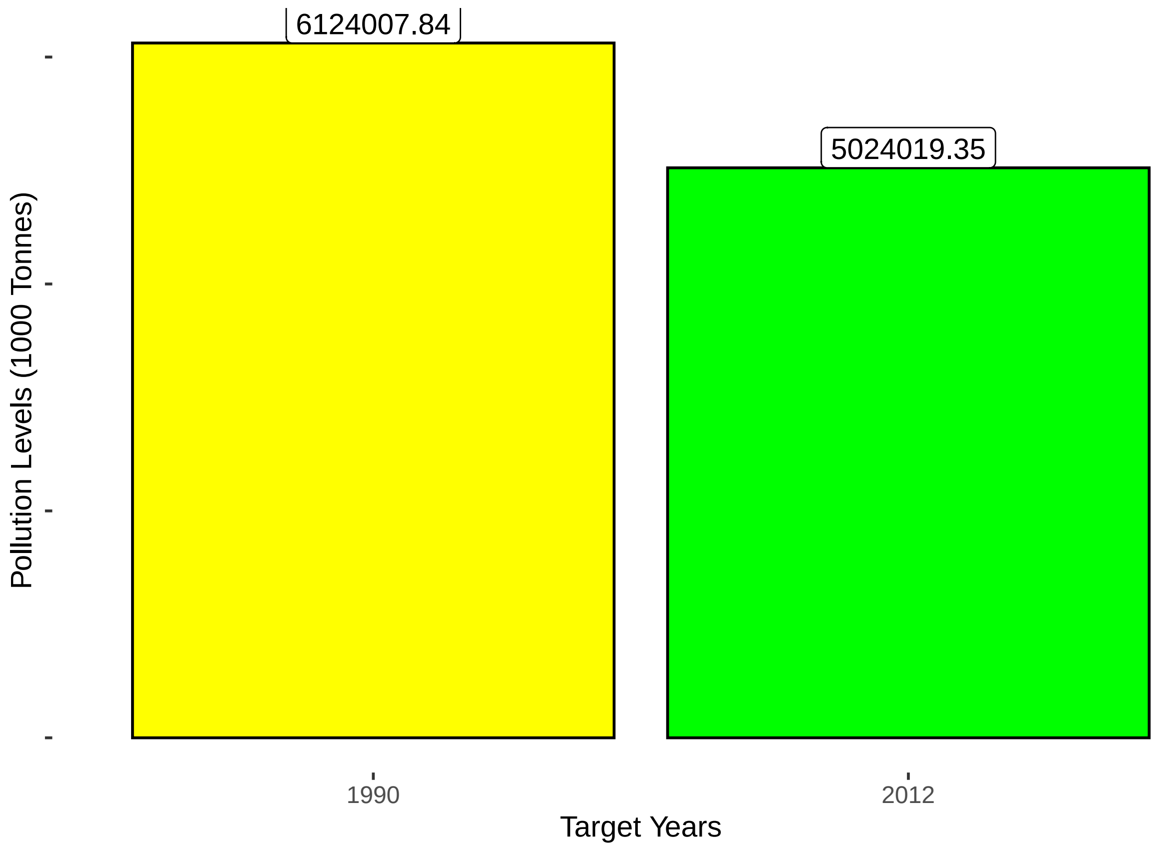
```
## [1] -0.2273103
```

Here lies a weak relationship, indicating that the use of renewable energy can help us in our endeavor.

GRAPH 5- KYOTO PROTOCOL
In its Annex B, the Kyoto Protocol sets binding emission reduction targets for 37 industrialized countries and economies in transition and the European Union. Overall, these targets add up to an average 5 per cent emission reduction compared to 1990 levels over the five year period 2008–2012 (the first commitment period).
We will see if the Kyoto protocol has made an impact in the 20 countries that signed its first commitment period.

```
kyoto = CO2 %>%
    filter(Country %in% c("Canada", "Sweden", "Norway", "Japan",
        "Switzerland", "Hungary", "Poland", "Latvia", "Liechtenstein",
        "Estonia", "Ukraine", "Australia", "Iceland", "Bulgaria",
        "Czechia", "Romania", "Monaco", "New Zealand", "Slovakia",
        "Russia", "Slovenia"))
kyotobar = data.frame(years = c("1990", "2012"), sums = c(sum(kyoto$`1990`),
    sum(kyoto$`2012`)))
ggplot(kyotobar, aes(x = years, y = sums, label = sums)) + geom_col(fill = c("Yellow",
    "Green"), colour = "Black") + labs(x = "Target Years", y = "Pollution Levels (1000 Tonnes)") +
```

```
theme(rect = element_blank(), axis.text.y = element_blank()) +
geom_label(vjust = 0)
```



There seems to be a 17.96% reduction in CO_2 levels in the target period and the chosen countries, indicating the effectiveness of the Kyoto protocol.

CONCLUSIONS

From the following data, we conclude that CO_2 emissions vary between countries, with wealthy countries emitting more CO_2. This creates the problem of conflicting interests, specifically that of bringing disaffected populations out of poverty and of reducing carbon emissions. However, we see that the use of renewable energy and pacts like the Kyoto Protocol are helpful in battling the situation.