Greenhouse gas emission analysis

--By Satish Gollu

Under the guidance of Dr. Wassnaa Al- Mawee

a. Abstract

Analyzing and finding insights of greenhouse gas emission from all around the world within a stipulated time frame.

Pollutant wise analysis of the entire emissions of the gases from the year 1990 to 2018 and have found that, gases like GHG, GHG Indirect-CO2, direct carbon dioxide emissions and methane-CH4 are the pollutants that have been predominantly emitted into the atmosphere in the last two decades. For the top emitted gases United States of America, the Russian federation, Japan and Germany are being the principal emitters of these pollutants for all the major emitted gases.

By tracking the pollution emission levels across a time period of two decades for the top countries we have found that the emission curve for the United States of America the Russian federation is in the rising manner and in case of Japan and Germany the pollution emission curve is in a falling trend and over all top contributes for the entire emissions.

b. Introduction (Statement of problem, brief description of data):

Over the years, around the world with increasing of technology and ease of human life, there are tons of hazards gases that are emitted to the atmosphere and in order to control or initiate steps to minimize the emissions and need to have an idea which countries are responsible for these high emissions.
We have a data set with information on emissions of major pollutants from all countries and by looking at the data, it has amount of emission which is equivalent in kilotons in just one column and hence in this project our principal focus is to turn the numbers into actionable insights and identify the focus areas like specific gases

for specific countries to reduce the overall emission of the greenhouse gases. Analysis of country wise and gas wise on our planet and make the visualizations and implied insights be available on an easy to access user interface, which would enable to report and suggest to the organizations (like Montreal protocol members) and understand the various greenhouse gas emission levels across the countries of the world.

Brief description of data:

The dataset here consists the values of different gas for different countries from 1990 to 2018. In details, these are information on *anthropogenic emissions* by sources (emissions from fossil fuels, deforestation, land use changes, livestock, fertilization) and removals by sinks of the following GHGs: carbon dioxide (CO2) methane (CH4) nitrous oxide (N2O) hydrofluorocarbons (HFCs) perfluorocarbons (PFCs) unspecified mix of HFCs and PFCs sulphur hexafluoride (SF6) nitrogen trifluoride (NF3) that are not within the scope of Montreal Protocol, an international agreement in 1987, aimed at curbing the emission of ozone depleting substances.

Data Set Characteristics	Multivariate	**Number of Rows**	9688
Associated Tasks:	Exploratory Data analysis	**Number of Columns**	4

c. Methods

Description of the processes of data manipulation:

Initially the data set had just four columns with one of them being the serial number the other column contains the country, year, value and a category variable that contains information about what pollutant it is. In order to perform analysis on the data and visualize it we will need to clean the data.

The analysis using R:

We have used the dplyr package and mutate function with recode as a parameter to rename the category variables according to the gas name using simper terms.

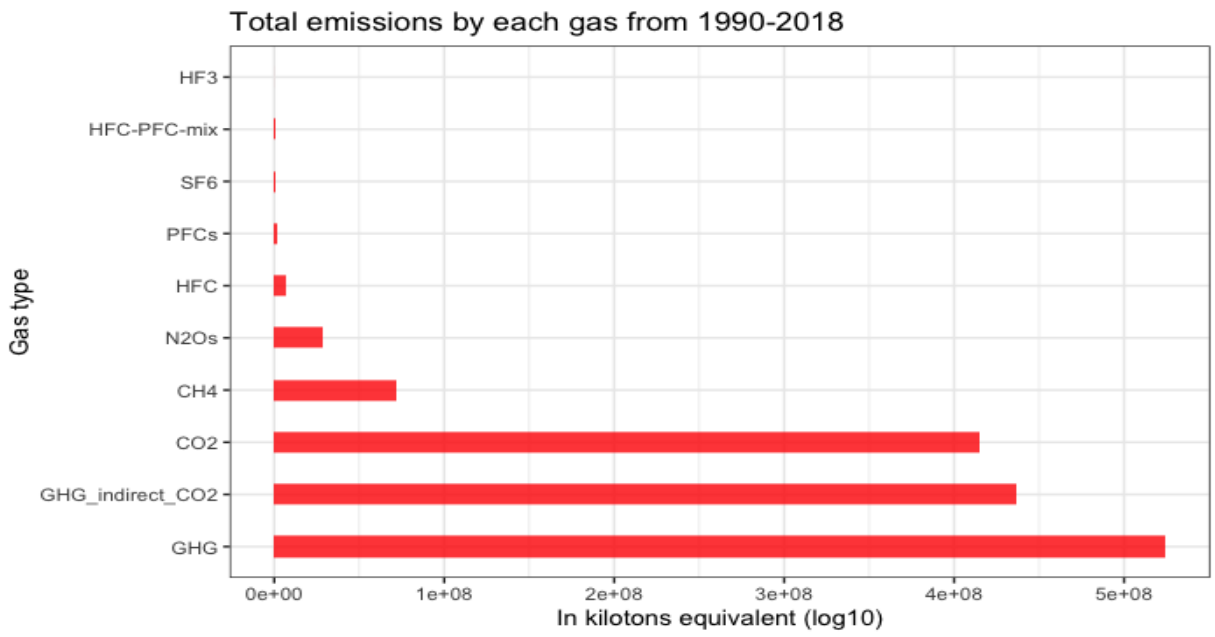
And performed pivoting, grouping, subset selection to regroup the data into designated data frames for further analysis and removed few data points using filter function and while converting data from char to numeric we might get NA

values because some of our input values are not formatted properly, because they contain commas (i.e.,) between the numbers. So, we removed these commas by using the **gsub** function.

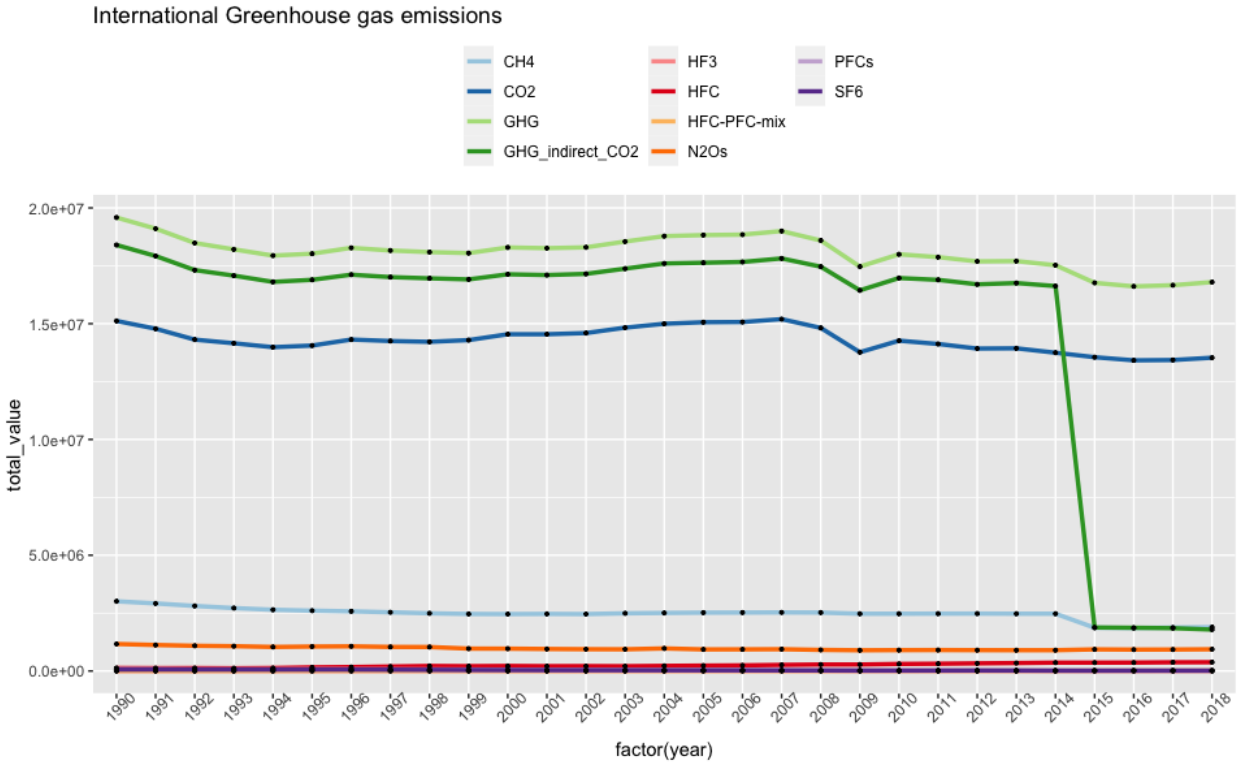
we used ggplot2 library to visualize the filtered data for all of the key insights.

d. Results (Interpretation of plots, web-apps, or computing models):

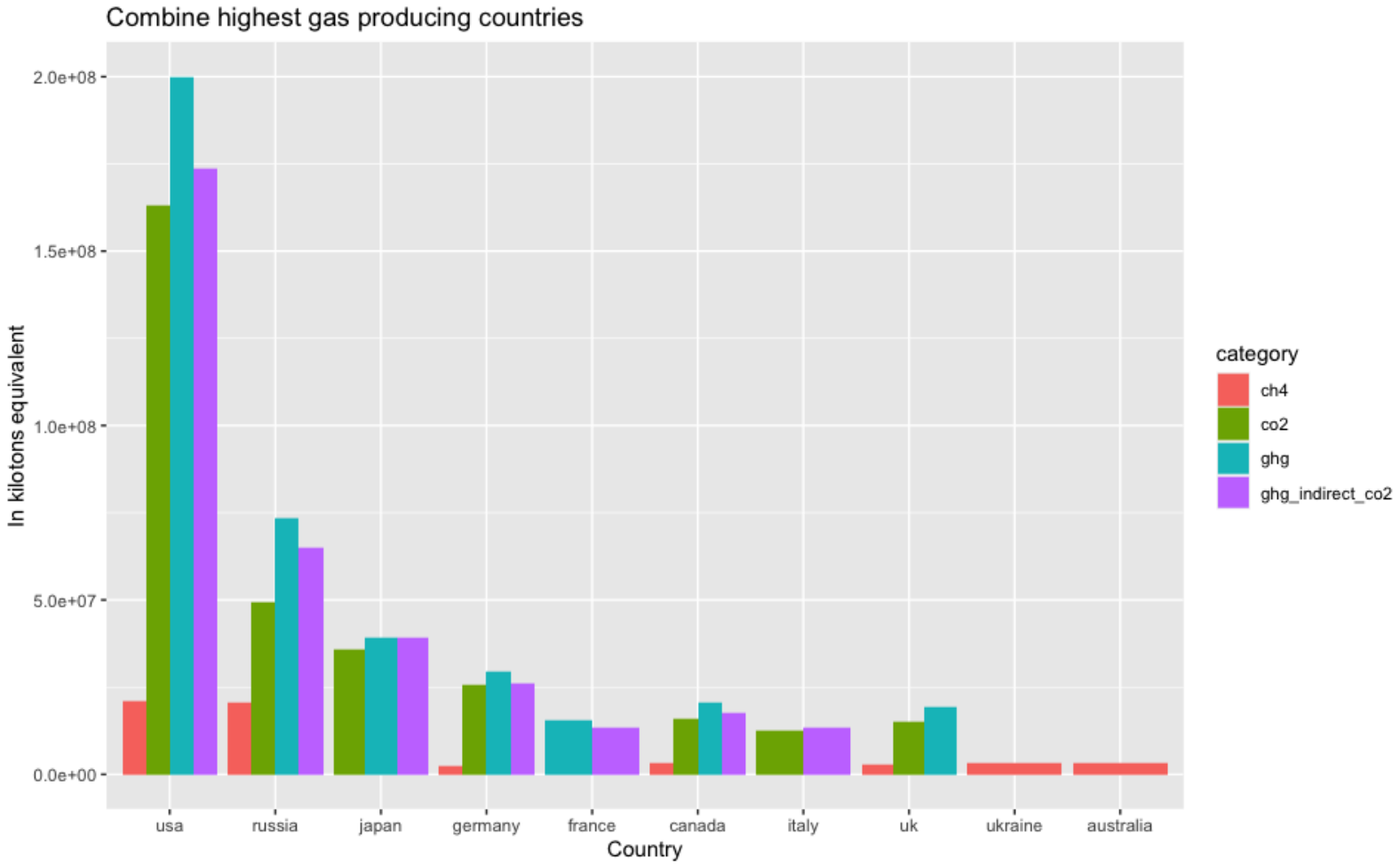
Summary of emissions:



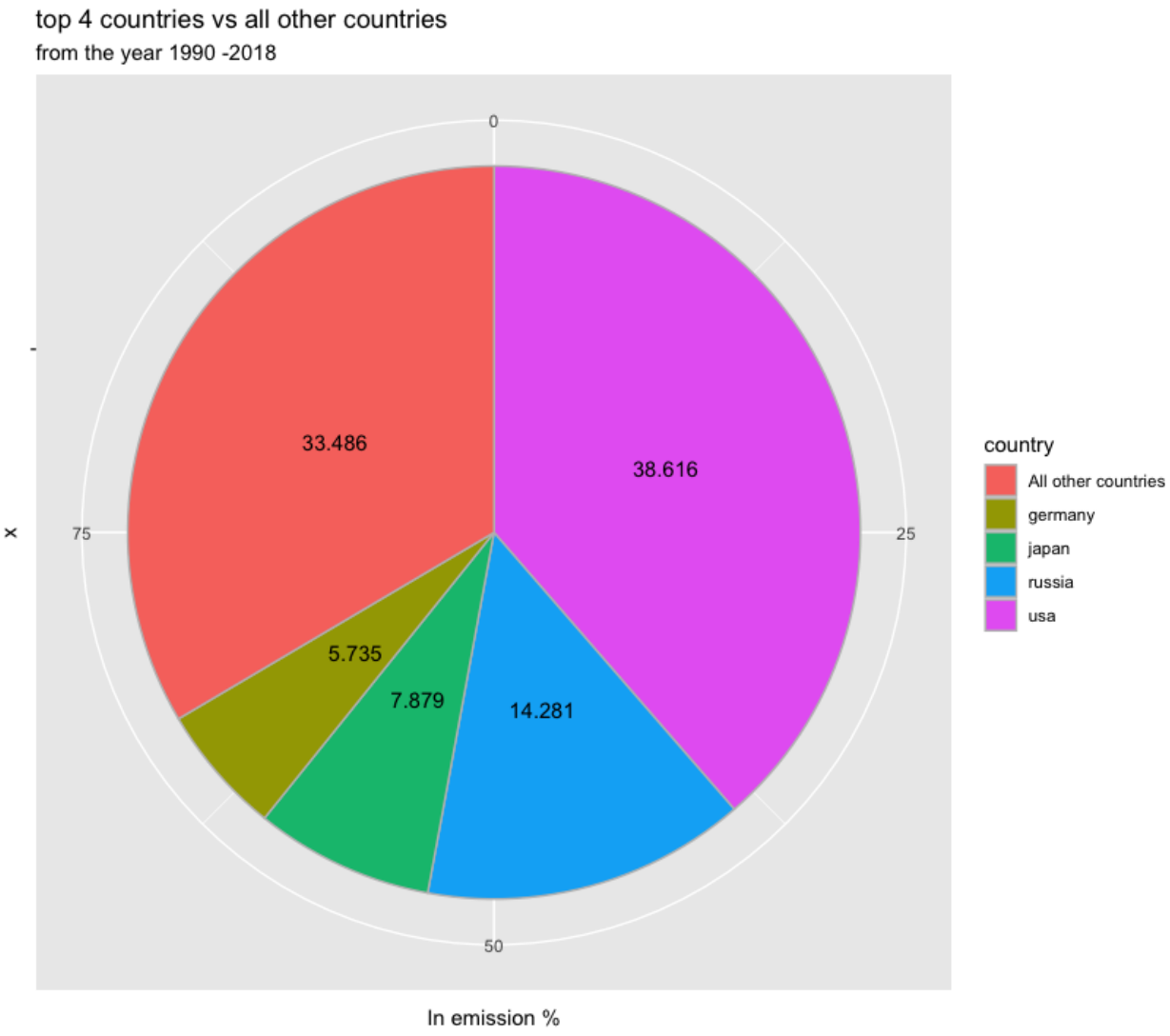
The above plot shows us the gas emitted during the period of 1990-2018 during which the readings have been collected and as per the data GHG, indirect_CO2, CO2 and CH4 are the major pollutants. And we looked at each and every gas year wise emissions from 1990-2018



above figure shows each gas year to year emissions from 1990-2018



from above figure illustrates top countries and it's respective top gas emissions and USA and Russia tops in all gases emission.



key findings and conclusions

From the year 1990 - 2018, around the world the total amount of gas emission is '1486003683' kilotons. from these there are 4 gases which contributes a large amount of emission, they are i) GHG ii) GHG_indirect_CO2 iii) CO2 iv) CH4

"GHG"contributes 35.29% of total emission from the year 1990 - 2018, followed by 'GHG_indirect_CO2' which accounts for 29.41% of total emission and 'CO2' accounts for 27.91% of total emission and 'CH4' contributes to 4.81% of total emissions.

out of total 10 different gases, these 4 gases combined contributing 97.42% of gas emissions where rest of gases contributing 2.58%

GHG -(total contribution 35.29%)

the top GHG producing countries are–'USA','Russia','Japan' and 'Germany'

GHG_indirect_CO2 -(total contribution 29.42%)

the top indirect CO2 producing countries are – 'USA','Russia','Japan',and 'Germany'

CO2 (total contribution 27.91%)

the top CO2 producing countries are – 'USA','Russia','Japan',and 'Germany'

CH4 (total contribution 4.81%)

the top CH4 producing countries are – 'USA','Russia','Ukraine',and 'Canada'

USA, Russia, Japan and Germany these are the common countries that produce top high gases.

out of all countries, these 4 countries together producing 66.52% of total emissions which is huge and out of 66.52% of emissions, USA itself producing 38.62% of total emissions which has major impact to the environment followed by Russia which contributes 14.28% of emission.

in general, the largest source of greenhouse gas emissions from human activities is from burning fossil fuels for electricity, heat, and transportation.

these 4 countries are major manufacturers and exports of industries includes automobiles, consumer electronics, computers, aerospace, semiconductors, and iron and steel.

By initiating controlling of gas emissions with these 4 countries will bring down the overall global impact.

if these 4 countries can adapt to renewable energy to store the power of wind, sun, water, tides and other planetary resources like geothermal heat, which comes from the Earth's core to produce electric power.

Agricultural "biomass" products also can be used to generate electricity and heat.

Renewables generate electricity without producing greenhouse gases—or producing very little when compared to traditional energy sources.

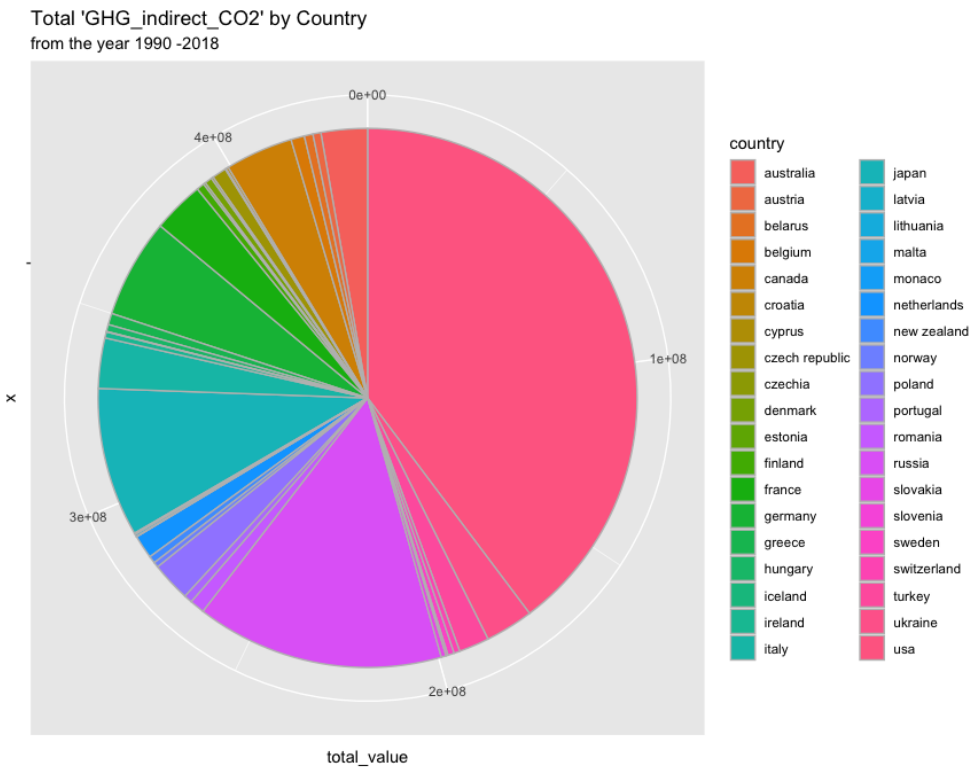
Using recycled materials will also help to minimize the greenhouse gas emissions

This is the best way to work around to minimize the greenhouse gas emissions.

Appendix

	category	total_value	each_emission_percentage
1	GHG	524444808.43	35.292
2	GHG_indirect_CO2	437114065.74	29.415
3	CO2	414828410.77	27.916
4	CH4	71608586.90	4.819
5	N2Os	28275799.26	1.903
6	HFC	7084020.15	0.477
7	PFCs	1301123.03	0.088
8	SF6	1068300.92	0.072
9	HFC-PFC-mix	248135.78	0.017
10	HF3	30431.82	0.002

```
```{r pressure, echo=FALSE}
# 2. Analysis of "GHG_indirect_CO2" gas emission
```{r,fig.width=9}
#selecting a subset of data that only contains "GHG" gas emission
GHG_indirect_CO2 <- dataset[dataset$category=='GHG_indirect_CO2',]
#grouping by each country to know which countries produced highest GHG
GHG_indirect_country <- GHG_indirect_CO2%>% group_by(country) %>%summarise(total_value = sum(value),.groups = 'drop')
GHG_indirect_country <- GHG_indirect_country %>% arrange(desc(total_value))
#visualizing % of contribution of each country
pie <- ggplot(GHG_indirect_country,aes(x="",y=total_value,fill = country))+
  geom_bar(stat = "identity",colour="gray")
pie+coord_polar("y", start=0,direction = 1)+
  ggtitle("Total 'GHG_indirect_CO2' by Country",subtitle = "from the year 1990 -2018")
```
```


```

```{r,warning=FALSE,message=FALSE}
#lets find out the % of emission of each USA gases from 1990 to 2018
usa_gas_wise_total <- usa_emissions%>% group_by(category) %>%summarise(total_value = sum(value),.groups = 'drop')
#find the sum of total country gas emission
usa_sum <- sum(usa_gas_wise_total$total_value)
#create a column name 'emission_percentage and calculate the % of contribution'
usa_percentage <- mutate(usa_gas_wise_total, usa_emission_percentage = (usa_gas_wise_total$total_value/usa_sum)*100)%>%
mutate_at(vars(usa_emission_percentage), funs(round(., 3)))
usa_percentage

```

	category	total_value	usa_emission_percentage
1	CH4	20922010.02	3.646
2	CO2	162992306.72	28.404
3	GHG	200068613.06	34.865
4	GHG_indirect_CO2	173703276.94	30.270
5	HF3	11053.53	0.002
6	HFC	3220566.32	0.561
7	HFC-PFC-mix	181692.62	0.032
8	N2Os	11949725.39	2.082
9	PFCs	320122.91	0.056
10	SF6	471137.03	0.082

	category	total_value	russia_emission_percentage
1	CH4	20462518.47	9.642
2	CO2	49241132.34	23.203
3	GHG	73594670.39	34.679
4	GHG_indirect_CO2	65027126.52	30.642
5	HF3	9.00	0.000
6	HFC	624128.92	0.294
7	N2Os	3009207.35	1.418
8	PFCs	224046.87	0.106
9	SF6	33626.59	0.016

	category	total_value	japan_emission_percentage
1	CH4	1145197.23	0.978
2	CO2	35944194.98	30.701
3	GHG	39022712.33	33.331
4	GHG_indirect_CO2	39031111.23	33.338
5	HF3	16771.92	0.014
6	HFC	704438.43	0.602
7	N2Os	750053.30	0.641
8	PFCs	247704.98	0.212
9	SF6	214350.75	0.183

	category	total_value	germany_emission_percentage
1	CH4	2.270162e+06	2.664
2	CO2	2.553818e+07	29.965
3	GHG	2.959864e+07	34.729
4	GHG_indirect_CO2	2.603061e+07	30.542
5	HF3	5.143645e+02	0.001
6	HFC	2.062569e+05	0.242
7	HFC-PFC-mix	6.291416e+04	0.074
8	N2Os	1.370390e+06	1.608
9	PFCs	3.157764e+04	0.037
10	SF6	1.186500e+05	0.139