Global Historical Emissions Analysis
Kai Murata, Richard Kim, Hana Tafesse

Climate change and extreme weather patterns have disrupted our livelihood. Carbon emission is a leading cause for climate change and in the increase of disastrous climate related events. We collected information from the Global Carbon Project, an organization that seeks to quantify global greenhouse gas emissions. We hope to use the database and analyze these carbon emission patterns to see which countries contribute large amounts of carbon emissions to our planet.

Our data was in the form of a csv with over 1100 rows of data. The data included the country of origin for the emission, the sector at which the emission was formed from (coal, oil, etc.) and the metric ton of each country and sector from 1960 to 2019. From this data, out group decided to answer the following four questions:

1. Which five countries produced the largest volume of carbon emissions from 1960 to 2019? Which five countries produced the least?
2. Is there a country that has a decrease in carbon emissions? If so, what context can be provided to explain the decline?
3. Which country has the largest rate of carbon emissions from 1960 to 2019?
4. Which of the carbon emitting sectors emitted the most CO2?

For these objectives, we created two separate dataframes to use for analysis and plotting. One was separated using each of the countries. Any rows with null values were dropped in order to make sure we have 60 years of data for all countries. The sectors were removed and the years were grouped into decades using nested for loops. We then obtained a 60-year total volume for each country. Using this data, we selected the five highest and lowest emitting countries.

For each of these countries, we created a bar chart **(Figure 1 and Figure 2)** to identify the vast difference in emissions. Smaller less developed countries are in the bottom five emitters (Sao Tome and Principe, Tonga, Dominica, Comoros, Saint Vincent and the Grenadines) produced less than 10 metric tons in 60 years. However, each of the five largest emitters, (United States, China, Russia, Japan, and Germany) produced over 50,000 metric tons of carbon emissions individually. Where the United States and China both emitted over 200,000 metric tons of carbon emissions in 60 years.

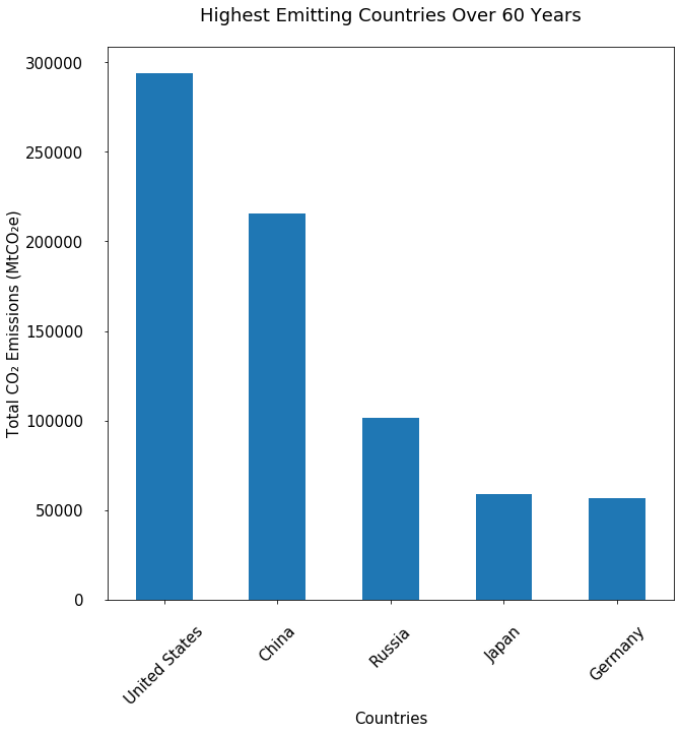

Figure 1. Bar Chart of Top Five Emitting Countries

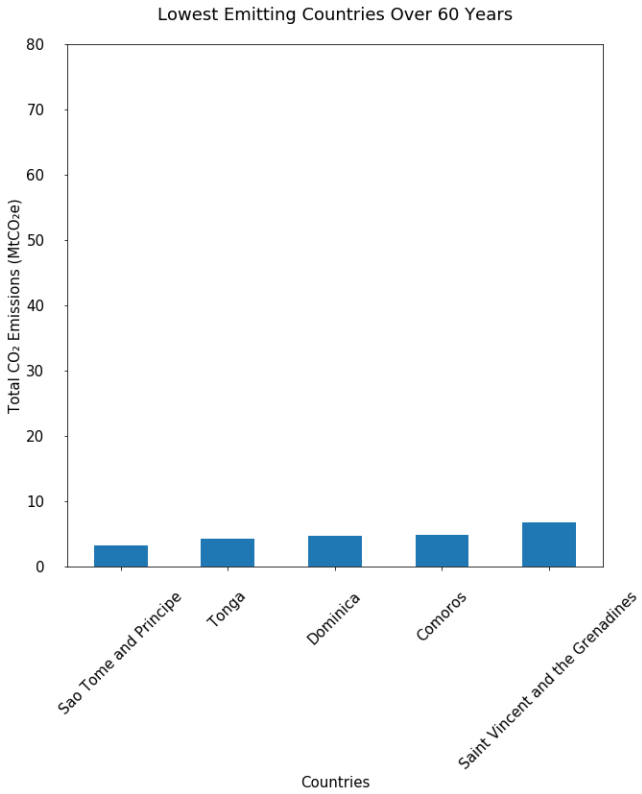

Figure 2. Bar Chart of Bottom Five Emitting Countries

Next, we created a line graph **(Figure 3)** for the top five countries to see trends. We found that there were decreases in carbon emissions amongst the top five countries during the 60 years.

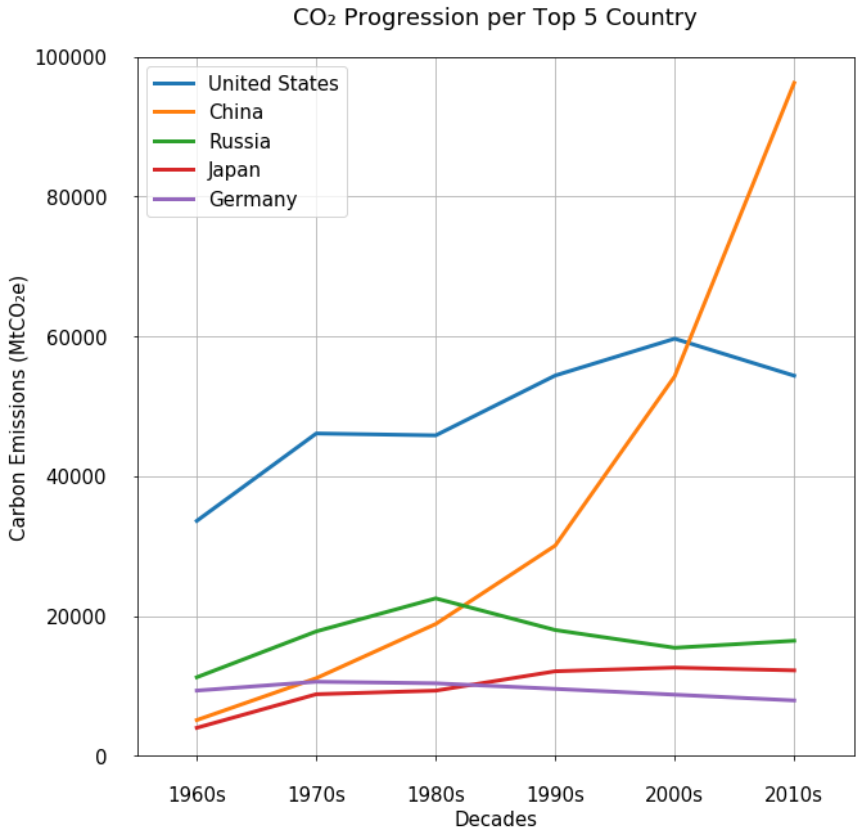

Figure 3. Line Graph for Top Five Emitting Countries

Possible reasoning for the decrease in carbon emission between the countries was identified as below:

- German implementation of nuclear program reduced their reliance on fossil fuels
- Russia experienced the fall of the Soviet Union putting their economic progress to a temporary halt
- USA began to export labor to foreign countries such as China which is a factor in China's massive boost in carbon emissions
- China becoming a major exporter of goods can be considered a reason for other countries decline in carbon emissions

Next, we created a scatter plot and plotted regression lines **(Figure 4)** for the top five emitting countries.

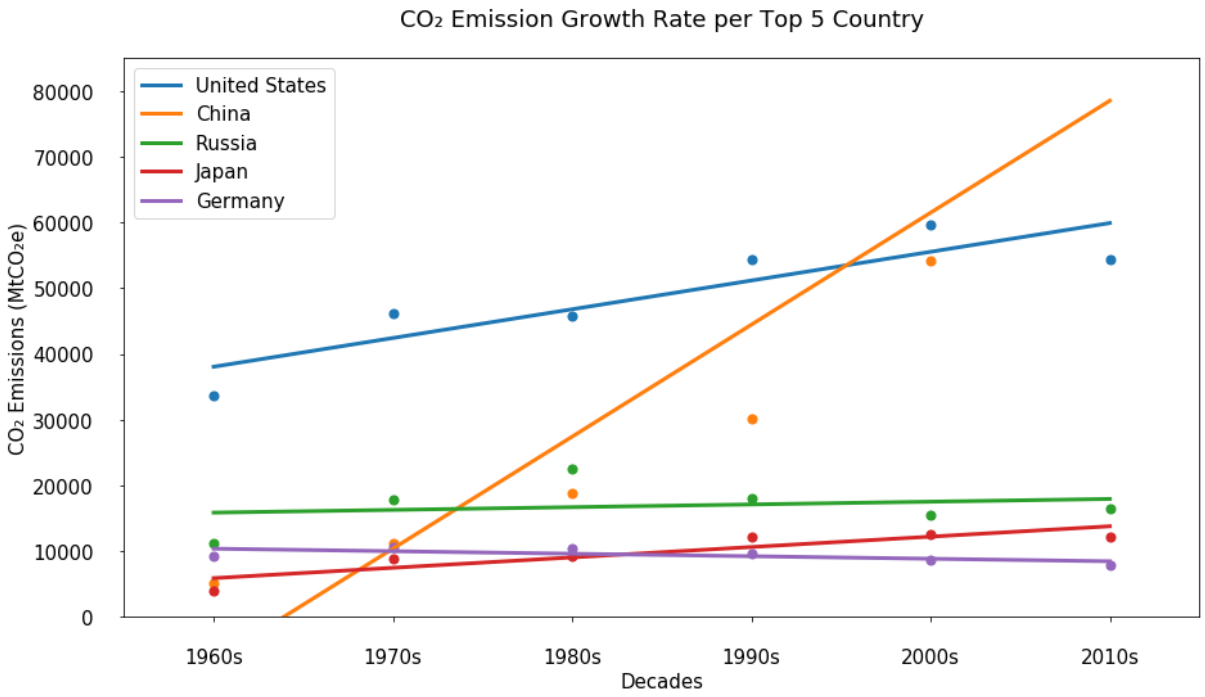

Figure 4. Scatter Plot with Regression Lines for Top Five Emitting Countries

From the regression lines, we were able to identify three facts:

- China began with approximately 5000 metric tons of carbon emissions in the 1960s. However, they emitted approximately 96,000 metric tons in the 2010s
- China's economic boom of exported goods correlates with their jump in carbon emissions from the 2000s to 2010s
- Germany has had a negative slope in carbon emission for the past 60 years most likely due to nuclear power production identified above

Next, our group analyzed the different sectors of emissions. These sectors included, bunkers, cement, coal, gas, gas flaring, and oil. We disregarded country of origin and looked solely at what sector is responsible for the carbon emissions in the past 60 years.

First, we created a pie chart **(Figure 5)** to identify whether there were any sectors taking up the majority of the emissions.

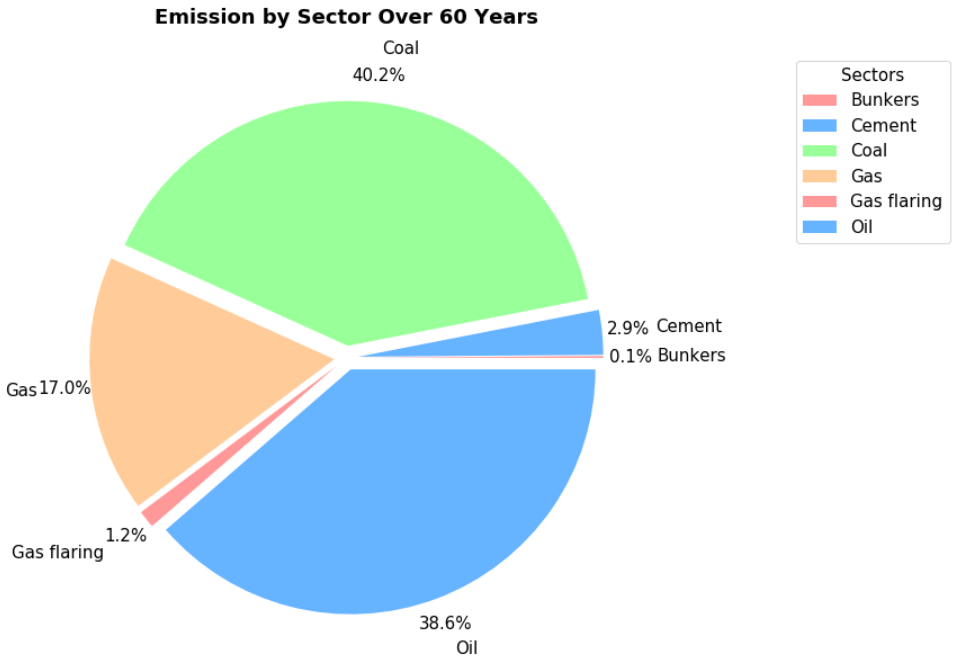

Figure 5. Pie Chart of Different Amount of Emissions by Sector

As expected, coal and oil are the largest sources of carbon emission in the previous 60 years at around 40% each. Gas has emitted nearly half of both coal and oil at 17%. The other sectors are not large contributors due to their nature. Cement, bunkers, and gas flaring are all specific causes of emission. Bunkers from ships, cement from construction, gas flaring from mining for oil. Therefore, they are much less at emitting. Whereas coal, oil, and gas are all means of travel and daily life used by the majority of the population.

Finally, we created a line graph **(Figure 6)** to see if there are any trends in the sectors.

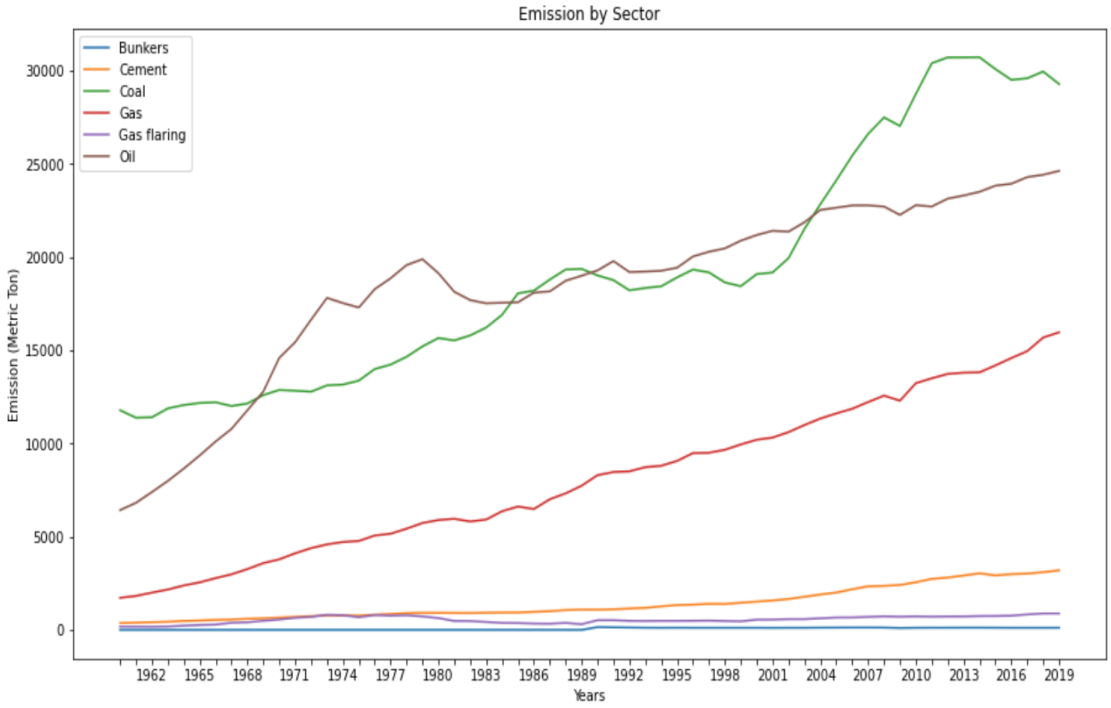

Figure 6. Line Graph of Emissions by Sector over 60 Years

From the line graph, we were able to identify that all sectors have constantly increased their emissions over the previous 60 years. Oil can be seen to have an increase in the 1960s. This was most likely due to normal economical development. However, there is a sudden decrease in the 1980s which is most likely due to conflict in the middle east. Additionally, there was a large increase in coal emissions between the 2000s and 2010s. This was most likely due to an increase in demand for coal from China to run their factories.

During this study, we ran into a few difficulties. However, the largest issue that was present was how transposing the data frame actually leads to a loss in data, which we did not know. When transposing, columns will show up but will not be part of the dataframe. Therefore, calling for a column by name or trying to use row names was not possible.

We figured this problem out when we ran the “columns” call which returned all columns except the index. The index column was lost when we transposed our original dataframe.

In order to solve this issue, we created a new column with duplicate data as the index to create a new column we could call. If we were to repeat this project once again, we would create and duplicate the index column within our dataframe before transposing to prevent the loss of data. Additionally, doing so would prevent the need for hard-coding the column information like we did during the analysis.