Visualization of Fossil-Fuel CO2 Emissions Data

Qiulin Luo
College of Engineering
Northeastern University
Boston, MA 02115
luo.qiu@northeastern.edu

Abstract

This work is based on Appalachian Energy Center's global, national and regional fossil-fuel CO2 (Carbon Dioxide) emissions data from 1751 to 2019 for visualization modeling and visualization analysis. The research questions include the percentage of carbon emissions fossil energy breakdown, carbon emission trends, and comparison of carbon emissions by country or region. Dynamic interactive charts were implemented using Javascript and Chart.js, and static analysis charts were implemented using Python.

1 Motivation

Fossil-fuels are non-renewable energy sources that are formed from the remains of plants and animals that lived millions of years ago. They include coal, oil, and natural gas and are used extensively as energy sources for transportation, heating, electricity generation, and industrial processes. Fossil fuels are extracted from the Earth's crust through mining, drilling, and fracking techniques. They release carbon dioxide and other greenhouse gases when burned, contributing to climate change and global warming. Despite concerns over their environmental impact, according to the International Energy Agency, fossil-fuels still account for approximately 80% of global energy consumption.

The increasing levels of carbon emissions have significant consequences for the environment and human health. Climate change, rising sea levels, ocean acidification, and an increase in extreme weather conditions, including heatwaves and hurricanes, are some of the impacts caused by carbon emissions. This emphasizes the need to reduce carbon emissions and shift to cleaner energy sources.

Analyzing and visualizing CO2 emissions data is essential in understanding the current state of carbon emissions and their impacts. It allows policymakers, researchers, and the public to identify the major sources of carbon emissions, track progress, and develop effective mitigation strategies. One of the ways to analyze carbon emissions data is by calculating carbon footprints. A carbon footprint is the total amount of greenhouse gases emitted by a person, organization, event, or product. By analyzing carbon footprints, policymakers and organizations can identify the areas with the highest emissions and work on reducing them.

Data visualization presents complex data in a visual format, making it easier to comprehend and identify patterns. For instance, maps showing the geographical distribution of CO2 emissions can help policymakers and the public understand which regions contribute the most to CO2 emissions. Therefore, my idea is to use the web framework to design interactive and dynamic charts and graphs to visualize and analyze public concerns about CO2 emissions for all those who are interested in this issue.

2 Data set

Table 1: Global, Regional, and National Fossil-Fuel CO2 Emissions: 1751-2019

Nation	Year	Total	Solid	Liquid	Gas	Cement	Flaring	Per Capita	Bunker
AFGH	1949	4	4.0	0.0	0.0	0.0	NaN	NaN	0.0
AFGH	1950	23	6.0	18.0	0.0	0.0	NaN	0.003143	0.0
AFGH	1951	25	7.0	18.0	0.0	0.0	NaN	0.003299	0.0
AFGH	1952	25	9.0	17.0	0.0	0.0	NaN	0.003338	0.0
AFGH	1953	29	10.0	18.0	0.0	0.0	NaN	0.003701	0.0

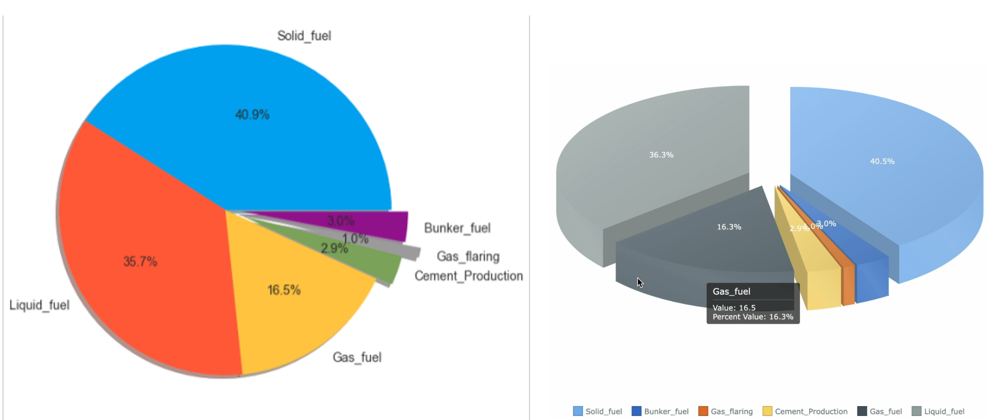
The data set is a time series of Carbon Dioxide (CO2) emissions from fossil-fuel combustion and cement manufacture. The data comes from Appalachian Energy Center, which can be downloaded at `https://energy.appstate.edu/cdiac-appstate/data-products`. Table 1 shows some sample data from this data set, which includes 10 attributes, namely nation, year, emissions from solid fuel consumption, liquid fuel consumption, gas fuel consumption, cement Per capita CO2 emissions and emissions from international trade (bunker fuels). The dimension of this data set is (18547, 10).

However, this data set is not only miscellaneous but also incomplete, so data collation and preparation includes operations such as data cleaning, format conversion, missing value processing, and outlier detection are needed to ensure data quality and consistency. For example, some specific fuel types may correspond to null values, which requires us to use the total emissions minus the existing data to perform the calculation. At the same time, we know that this data set statistics from 1751 to 2019, due to spanning more than 200 years, the world has changed dramatically, there are many countries with great changes in political structure and state structure, more famous time such as the reunification of East Germany and West Germany, the collapse of the Soviet Union, etc., which led to the need to choose a more meaningful year time point as the starting point for consideration. Since the data is about CO2 emissions from fossil fuels, which is closely related to the history of the world industrial revolution, I chose 1950, the year when the second industrial revolution ended, as the analysis point, and performed relevant data cleaning and data pre-processing on the data set for the next data import and visualization analysis.

3 Visualization

3.1 Pie chart

As we can see from the data set, there are many different types of fossil-fuel sources, so the first point I want to focus on is the analysis and comparison of the share of each type of fossil-fuel source in total emissions. This naturally leads to the pie chart used to represent percentages.



(a) static pie chart

(b) interactive pie chart

Figure 1: Emissions for each fuel type from 1950 for all nations

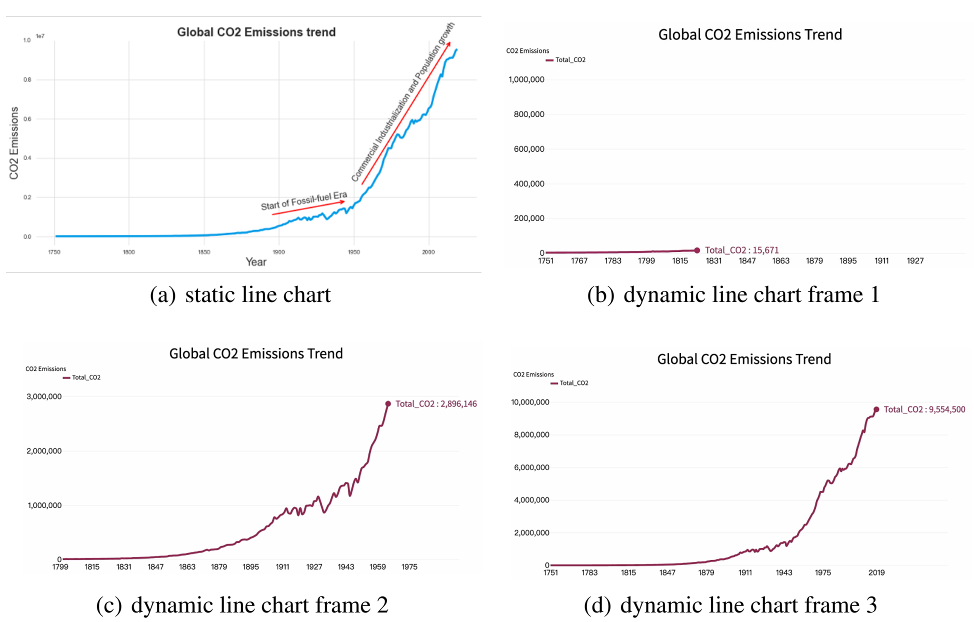
According to Figure 1, the share of solid fossil fuels is the highest, at 40.9 percent, followed by liquid fossil fuels and gas fossil fuels, at 35.7 and 16.7 percent, respectively. This trend is also consistent with the biochemical view that the more carbon-intensive a fuel is, the more it contributes to the industrial structure, and thus the higher the share of emissions.

From a visualization point of view, the static diagram on the left uses different colors to represent different fossil-fuels, and the color differentiation is high enough not to blend together and lead to visual confusion, and secondly, the thickness is just right, not consuming too much data-ink and space, and the smaller scaled blocks are extracted to avoid lack of differentiation. The interactive dynamic diagram on the right not only meets the advantages of the static diagram, but its interactive feature also meets the reader's purpose of obtaining corresponding information. The mouse click on any block will have a drag-out effect for emphasis, and the effect can also be triggered by clicking on the legend below, and repeatedly clicking on the block will have a loading effect. When the mouse is over the block, there will be a hover tab showing the detailed scale values.

My architectural process in this example is to build the front-end using React and call the Chart.js interface. The first step is to create a canvas element, then create a data object for each fossil-fuel block, register multiplexed event listeners for each block, and override both click and hover functions. Hover events use the traditional `div` attribute change method, where the transparent attribute changes from 0 percent to 100 percent whenever the event is responded to, thus the 3d pie block uses the open source pie chart model with manual cutting, registers the animation interface and calls the `update(target, value)` function to trigger the pull-out and put-in effects when the mouse is clicked, and sets the semaphore to mark the state of the specific block.

3.2 Line chart

This time, instead of classifying fossil-fuels, I explore the trend of total emissions from 1751 to 2019, and since it is related to a time series, line charts are used to visualize it. According to the Figure 2, the global CO2 emissions started to increase from 1900 and literally skyrocketed from 1950 and it doesn't look like it is halting. One of the major reasons for this increase in CO2 emissions was the rapid industrialization and economic growth of countries like China and India. These countries began to heavily rely on fossil fuels like coal to fuel their growing economies, which led to a significant increase in CO2 emissions. Additionally, there was a sharp increase in the use of transportation such as cars, trucks, and airplanes around this time. As the world became more interconnected and travel became more accessible, people began to travel more frequently, which in turn led to an increase in emissions from transportation.



(a) static line chart

(b) dynamic line chart frame 1

(c) dynamic line chart frame 2

(d) dynamic line chart frame 3

Figure 2: Global CO2 emission trend

The static figure in Figure 2(a) uses a grid-based graph with red arrows placed at the ends of the curve with the greatest variation, along with text legends to illustrate the reason why fossil-fuel is dominant at a given stage. The Figure 2(b), (c) and (d) show the 3 frames of the dynamic line chart separately. The dynamic curve starts from the origin in 1751 and extends to the right at a fixed rate, while the horizontal and vertical axes change with the dynamic curve. When the upper end of the curve exceeds the vertical coordinate scale, the vertical coordinate is automatically raised one unit upward, while the time scale expands evenly to the right, achieving the effect of a curve trend that changes with time. When the curve trend ends, all horizontal and vertical coordinates will be scaled to the same result as the static chart. This animation effect gives the reader a more intuitive and comprehensive tendency to understand.

To implement the animation, also create an instance of the chart in JavaScript by selecting the canvas element and initializing the chart type, data, and options. Then add data to the chart dynamically by pushing data points to the data sets array and updating the chart using the `update(target, value)` method. Animate the chart by updating the chart data and redrawing the chart at a regular interval using the `setInterval(func, value)` method. To dynamically change the horizontal and vertical coordinates, I add the corresponding coordinate values to the data sets when add the data. Add the timestamps of the horizontal coordinates to the `labels` array and the data of the vertical coordinates to the `data` array.

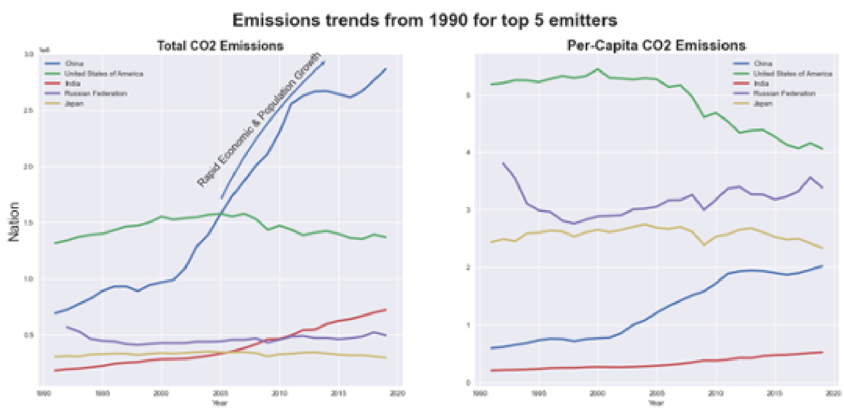


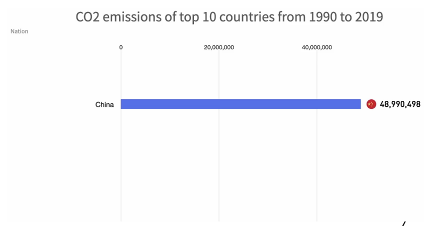
Figure 3: Emissions trend from 1990 for top 5 emitters

Figure 3 shows result of the next problem. The net and average trends since 1990 were compared for the five most prominent emitting countries. The same method of labeling with text narrative is used to illustrate the special slopes. The grid uses a square to make the growth and year interval ratio consistent and more harmonious in appearance. We can see that United States has been the leading emitter till 2005. Whereas for China, the emissions started to skyrocket from 2000. Japan and Russia also seem to have stabilized their emissions. This comparison reflects the different development history and development nodes of these five countries since the 1990s phase, with China's development evident to all.

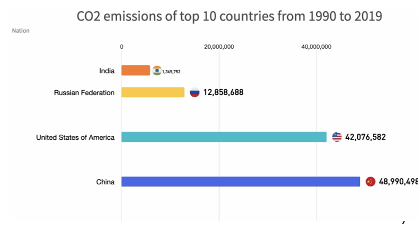
3.3 Bar chart

What are the top 10 and bottom 10 CO2 emitted countries from 1990 to 2019? The problem could probably be expressed as a percentage of a pie chart, but it's not visual enough. I used a dynamic bar chart from large to small to visualize the problem. Firstly, different colors are used to represent different countries. Unlike the static chart that shows all histograms directly, the dynamic chart first pops up the highest emission countries and places them at the top, after which the histograms pop up one by one from top to bottom as the animation plays. Figure 4 shows the 4 frames of this motion picture. The benefit of this dynamic mode is that this real-time visual feedback can help one better understand the trends and distribution characteristics of the emissions, and the number of groupings can be dynamically adjusted as needed to better reflect the distribution of the data, and this flexibility can improve the accuracy and efficiency of data analysis.

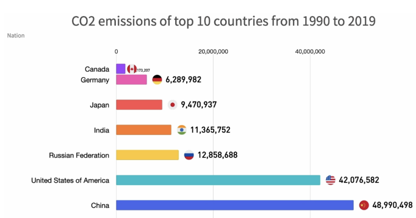
To achieve this animation effect using Javascript, I set up 4 state functions based on the react framework, `initializeBar()` to initialize the histogram, create a data array with default values and an empty canvas object, `updateBar()` to update the histogram and draw a new histogram based on the input data array and the number of groups parameters, `startAnimation()` to set a timer to update the histogram at each time interval and `stopAnimation()` to stop the animation and clear the timer. Bind the text tag as a subclass of bars' `div`.



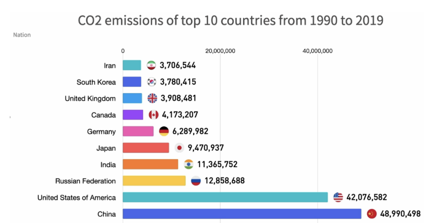
(a) dynamic bar chart frame 1



(b) dynamic bar chart frame 2



(c) dynamic bar chart frame 3



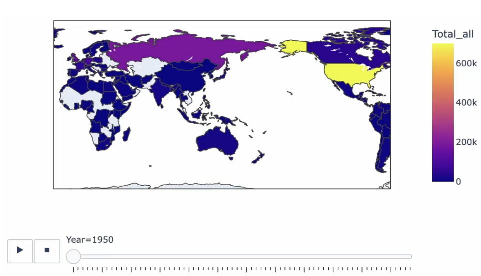
(d) dynamic bar chart frame 4

Figure 4: CO2 emissions of top countries from 1990 to 2019

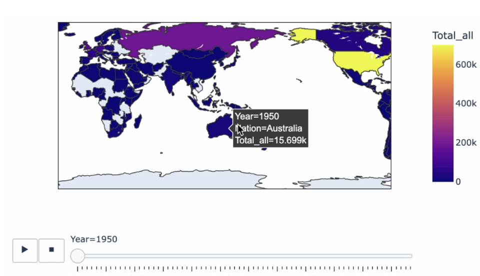
3.4 Map

In order to better represent the global emission trends over time, the visualization is done here in the form of an interactive map. First a variable rectangular range is loaded into the world map using a variable rectangular range and the country blocks on the map are colored, with a temperature scale on the right side to indicate the color size, i.e. the height of the emissions. A draggable timeline is set up at the bottom, and the user can drag the timeline from left to right, or click the play button to make it automatic. Over time, the colors of the blocks within the rectangle are automatically updated according to the total emissions data. One of the maps within the rectangle can also be viewed by dragging and dropping with the mouse, and when the mouse floats over a specific country, a hover tag will be displayed for specific data.

To achieve this effect using Javascript, I set up a div element containing a map, a form controlling color, data, and interaction, and a `div` element containing interactive elements. The timeline module and the temperature scale module I linked to python's `plotly.express` and `plotly.offline` libraries.



(a) original layer



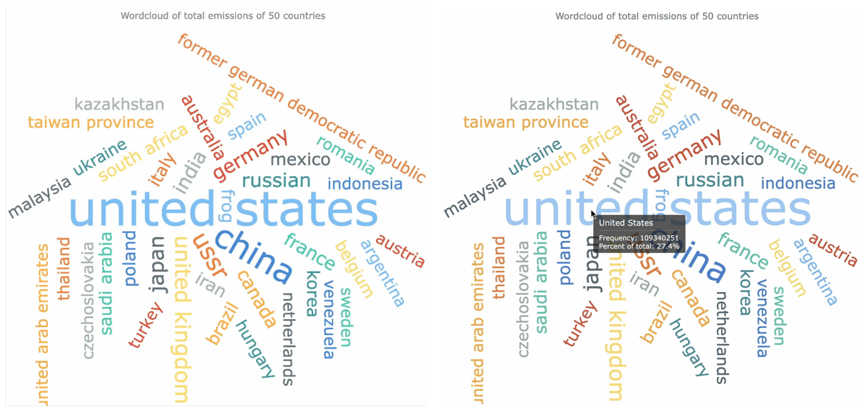
(b) floating effect

Figure 5: Choropleth map for both per-capita and total emissions for all countries in 2019

3.5 Wordcloud

With text-sensitive users in mind, an interactive word cloud is designed here to complement the above visualization. In line with the above map, the larger the font size of the text representing the country, the higher its total emissions.

The implementation of the mouse interaction effect is also basically the same as the above-mentioned interactive graph, i.e., setting the listener functions for the mouse click and hover each to control the `css` attribute value of the target element and the display attribute value of the bound div. The table type data is converted to `JSON` data for the convenience of representation. After loading the word data from the `JSON`, I used the `d3-cloud` layout to generate the word cloud. The layout is configured with various options, such as the size of the layout, the padding between words, and the function to determine the font size of each word. The `draw()` function is called when the layout is complete and creates an `SVG` element to hold the word cloud. It then creates a group element for each word and adds a text element with the word text and font size.



(a) original layer

(b) floating effect

Figure 6: Worldcloud of total emissions of 50 countries

4 Future work

Of course, the visualizations above are based on existing datasets, and forecasting and analysis of future fossil-fuel CO2 emissions is also critical.

To combine reality factors to machine learning methods such as LSTM (Long Short Term Memory) for predicting fossil-fuel emissions, we need to incorporate external factors that may affect the emissions data into the prediction model. These external factors can include economic indicators, energy consumption, population growth, and environmental policies. One approach to incorporating reality factors is to use a hybrid model that combines machine learning methods with expert knowledge and domain-specific information. This can be achieved by including relevant external factors as input variables in the machine learning model.

For example, if we want to predict fossil fuel emissions in a particular country, we may include economic indicators such as Gross Domestic Product (GDP), energy consumption data, and environmental policies as input variables in the prediction model. This can help to improve the accuracy of the prediction by taking into account the impact of external factors on fossil fuel emissions. These parameters can be uploaded and filtered through the visual front-end, after training on the server, the server will return the data and display it in the front-end through the interface, the user can predict and observe the data in multiple directions for different situations, different scenarios, this kind of function requires large web architecture such as front-end and back-end separation, and will be the direction of my efforts.