Visual Analysis of CO2 Emission and Energy Consumption of Top Emitting Countries

by Christopher Brennan, Pawan Kumar, and Manu Samala

Introduction

The issue of global warming leads to increased sea-level rise, droughts, melting of the Arctic & Antarctic, and environmental migration, among many other serious problems. CO2 is considered one of the major contributing factors leading to Global warming. Though the sources of CO2 released in the atmosphere are numerous, many of them are in balance with the natural carbon cycle. However, the excess consumption of fossil fuels due to increased industrialization and human activities is disturbing the healthy state of our climate through global warming. In this report, we analyze the fossil fuel and renewable energy consumption trend of the top CO2 emitting countries using statistical and visual analysis tools to assess the efforts being made to combat global warming indirectly.

Motivation

Climate change is a large environmental issue spanning decades, resulting in sea-level rise, leading to more uninhabitable land, more intense and frequent hurricanes, and a lack of ice in the Arctic, as some examples. This interactive project aims to present data on CO2 emission and consumption using various visual analytical models to motivate action towards switching energy sources from fossil fuels to renewable energies.

Methodology

In order to determine if emission data follows a Pareto distribution, we utilize descriptive statistics. Under Pareto's 80:20 rule, we assume that 80% of global CO2 emissions come from 20% of the countries. First, we calculate the total global emissions needed by totaling the three CO2 sources. Next, the top 80% of emissions were subsetted, and the 20% highest CO2 emitting countries were identified.

We then investigate which of the three main fossil fuel sources (Oil, Gas, and Coal) contributes the most to CO2 emissions. This was implemented through a Spearman correlation. This method is ideal as the analysis does not assume normality in the data and can determine the strength and direction of the linear correlation between CO2 emission by fossil fuel type and fossil fuel consumption. The total CO2 emissions for all countries were calculated in order to evaluate the impact of each fossil fuel source. It is important to look at the impact of each source because they may not produce CO2 at the same rate per unit, and therefore one or more sources may play a larger role in emission.

We then calculate the percentage each source contributes to the mean and overall CO2 emission. Oil, Gas, and Coal contributed 44.7%, 28.7%, and 26.6%, respectively, whereas their mean emissions were 160, 106, and 209 million tonnes. The impact, or correlation, of Oil being larger than other sources is verified when conducting the Spearman correlation of total CO2 emissions and the consumption of each source.

To best analyse net change of usage in fossil fuels for the top 20% countries (those identified in the Pareto distribution), visual analysis of the time-series chart was used as it not only gives initial and final usage values of fossil-fuel consumption, but also gives trends or sub-trends from 2001 to 2019.

We utilize time-series predictive analysis using ARIMA model for fossil-fuel and renewable energy consumption to show possible changes in trends for top emitting countries. To predict per-capita consumption of fossil fuel and renewable energy for the next 5 years, we use the ARIMA time-series predictive analysis. This analysis method uses lagged observation, and the relationship between observation & residual error to predict the next 5 years.

Visualisation Design

To visualize the CO2 emission of all the countries, we use a bar chart. Through visual analysis, we determined that a majority of countries contribute significantly less CO2 compared to a small subset of countries, and thus we decided to implement the Pareto principle (80:20 rule). Under this constraint, we emphasize the top 20% of countries as they contribute 80% of total emissions. On the bar chart, the top 20% of countries are colored in red while the bottom 80% of countries are colored in turquoise to perceive separation in emissions. With our bar chart, we allow the user to interact with the graph by hovering over a particular bar with their mouse to show the name of the country and the total CO2 emission from 2001-2019. We compare multiple countries at once through this implementation and allow for precision should the user desire that data through hovering, zooming, and scrolling.

Next, we show an interactive map to spatially represent the top 20% of CO2 emitters. The bubbles representing each country are red circles where the relative diameter of each bubble allows for visual comparison of CO2 emissions against the other top 20%. When initializing the interactive map, we start with a view of all countries and can zoom in or out to focus on a particular country or region. Similar to the bar plot, when hovering over the bubble with a mouse, we see the name and total emission of that country.

To show the relationship between emission and consumption of different types of fossil fuels, we implement a correlation matrix and scatterplot. The correlation matrix presents the Spearman correlation values for all interactions of emission and consumption of oil, gas, and coal. These values are color-coded on a spectrum from red to blue, representing -1 to 1, respectively. The color is a visual indicator of which interactions are heavily correlated and the direction of that correlation. We can then investigate this relationship, or other ones, further using the scatter plot. A button on the side of the graph lets the user choose between the matrix and scatterplot. The variables from the dataset were broken down into three categories: consumption data, emission data, and categorical variables. The user can select a variable from each group and display it with emission, consumption, and categorical variables mapped to display on the x-axis, y-axis, and color code, respectively. Clicking the dropdown menu, the user selects a variable to choose the specific interaction they want to view and the desired color to view that data. The categorical variable selection is helpful as the data can be colored/viewed by country or year, thus presenting a large amount of data in a manageable fashion.

For visual analysis of the net change in fossil fuel consumption, data was broken down into two graphs showing fossil fuel consumption or renewable energy consumption. The time-series plot allows us to best visualize change in year-to-year consumption of fossil-fuel/renewable energy in relation to other countries. Time from 2001 to 2019 is on the x-axis, while annual fossil-fuel consumption per capita is on the y-axis. The data is color-coded by country to distinguish each line easily. The user can watch the data plot over time to identify trends or select a particular year of interest.

To show time-series predictive analysis while maintaining visual consistency, we follow a similar style to the previous graph by plotting predictions of fossil fuel consumption and renewable energy usage from 2020 to 2024. The data is displayed as a line graph to show the trends of each country.

Evaluation Plan

This project uses different visual analytical tools to analyze and interpret the data to get the desired information. We use the Pareto plot to determine the top 20% of countries with the highest net CO2 emission. Additionally, the correlation matrix can be used to evaluate the overall accuracy of the emission/consumption relationship as there are perfect correlations with respect to each source's consumption and its emission counterpart. This, along with the presumed positive relationship between emissions and consumption, allows for the evaluation of the accuracy of the analysis.

Predictive analysis is a key analytical tool to help predict the future trend of the fossil fuel and renewable energy consumption of the top 20% net CO2 emitting countries. Here ARIMA (Auto-Regressive Integrated Moving Average) algorithm, which models a linear relationship between the data constituting the series, is used to determine the future trend of the consumption (fossil fuel and renewable energy) by these countries over the next five years.

Discussion

Through our descriptive analysis, we found that the CO2 emission dataset follows a Pareto distribution. This means that a small number of countries are responsible for an overwhelming majority of CO2 emissions and global warming contributions. Therefore, if the policy is instituted in these countries to curb CO2 pollution, it would be more effective in reducing global emissions than the bottom 80% of emitters. From this data, we find that there is a large gap between the the emissions of China and the USA relative to the rest of the world as shown in the bar chart.

We predict that there is a trend of some of the highest CO2 emitting countries reducing the rate at which they will consume energy from fossil-fuel sources (see ARIMA plot). Many countries will have similar consumption levels in the next 5 years as they do currently, or slightly increase compared to the past 5 years. This can also be seen in the fossil fuel line plot as a majority of countries have reduced their consumption in the past year. This can be due to social pressure and policies introduced in order to reduce the consequences of climate change. However, as the population continues to grow, the countries still need to consume energy. This is why many countries are increasing their renewable energy consumption. While this level per capita is low in almost all countries, Canada has a significant reliance on renewable energy compared to others. Except for China, all of the top-emitting countries have had a relatively consistent total CO2 emission over the past decade. However, China has seen a significant increase, more than any other country. This is most likely because they have had large exponential economic growth and energy consumption (Wang et al., 2016).

While it was not surprising that China, USA, United Kingdom, and Canada had high emissions due to their population and GDP, which have positive relationships with CO2 emissions (Pao and Tsai, 2011). It was surprising, however, that Brazil, Iran, and Indonesia had higher emission levels than Australia, which wasn't in the top 20% of emitters. These countries have had rapid economic and/or population growth in recent years (Sasana, Hadi, and Eka Putri, Annisa, 2018), and thus this could explain their presence among top emitters.

Not all countries prioritize the same fossil fuel source. The US had the highest consumption of both oil and gas. However, China has a significant reliance on coal, so much so that they emit around 5 times the amount of CO2 from coal as the US does. When it comes to gas consumption, Russia beats out China and thus is the fourth-largest emitter.

Future Work

1. This study can be expanded by using additional variables like Country GDP, Road Infrastructure, Energy Infrastructure, Energy Source Availability, Literacy rate, etc., to understand its impact on fossil fuel and renewable energy consumption.
2. The predictive analysis can be expanded to include Bi-LSTM, and the results can be compared to the Auto Arima Method used in the analysis here.
3. An analysis of the top 20% CO2 emitting country's budget allocation for renewable energy generation and using the findings in the multivariate forecasting here can help with a more accurate model.

Contributions

Throughout the project each team member focused on certain story points/research questions. Chris wrote up questions 1 and 2, Manu wrote questions 3 and 4, and Pawan wrote questions 4 and 5 for all assignments. For the original project proposal (P1) the introduction and story points were written by Chris, the dataset review was conducted by Pawan, and the visualization inspirations were collected by Manu. Each group member contributed equally throughout all of the projects and with the shiny code and write up for P7.

Bibliography

Ahmad, N., Du, L., 2017. Effects of energy production and CO 2 emissions on economic growth in Iran: ARDL approach. Energy 123, 521–537. https://doi.org/10.1016/j.energy.2017.01.144

de Freitas, L.C., Kaneko, S., 2011. Decomposing the decoupling of CO2 emissions and economic growth in Brazil. Ecological Economics 70, 1459–1469. https://doi.org/10.1016/j.ecolecon.2011.02.011

Pao, H.-T., Tsai, C.-M., 2011. Modeling and forecasting the CO2 emissions, energy consumption, and economic growth in Brazil. Energy 36, 2450–2458. https://doi.org/10.1016/j.energy.2011.01.032

Sasana, H., Eka Putri, A., 2018. The Increase of Energy Consumption and Carbon Dioxide (CO2) Emission in Indonesia. E3S Web Conf. 31, 1–5. https://doi.org/10.1051/e3sconf/20183101008

Wang, S., Li, Q., Fang, C., Zhou, C., 2016. The relationship between economic growth, energy consumption, and CO2 emissions: Empirical evidence from China. Science of The Total Environment 542, 360–371. https://doi.org/10.1016/j.scitotenv.2015.10.027

Christopher Brennan
Texas A&M University

cbrennan@tamu.edu

Pawan Kumar
Texas A&M University

buddha@tamu.edu

Manu Samala
Texas A&M University

msam01@tamu.edu