

CSC110 Project Report: Carbon Emission per Capita Ranking and Projection

Kevin Cai, Junsong Guo, Ethan Zhang, Patrick Zhou

Monday, December 14, 2020

Problem Description and Research Question

One factor behind climate change is the emission of carbon dioxide into the atmosphere. Carbon dioxide is one of the most potent greenhouses and can trap heat (thermal infrared energy) on Earth from radiating into space. This process of Earth absorbing energy from the sun and redirecting excess energy back into space is what keeps the balance of the temperature on Earth. However, recent human activities have disturbed this balance. According to NASA, humans have increased atmospheric CO₂ concentration by 47 percent ever since the Industrial Revolution started.¹ As the result, the average annual temperature on Earth has been rising ever since, and over the past 100 years, the temperature has increased by almost 1 degree Celsius.² Although this might look like a big deal, this slight increase has caused catastrophic impacts around the globe, including climate change, risen sea-levels, ocean acidification, and much more.

This is why our group would like to find out, **“Out of 90 countries with available data, in years 1961 - 2014 which countries had the highest average CO₂ emissions per person and will have in the future (2015 - 2034) if the trend continued?”** We think it is important to know this data because we all contribute to the emission of carbon dioxide on a daily basis. This can be from driving cars to simply wasting energy though not unplugging cables when we are not using them, as a portion of electricity is still generated through burning fossil fuel, which releases the most CO₂ into the atmosphere. Hence, calculating the CO₂ emissions per person can offer us an insight into the estimated average carbon footprint for individuals in each country. This estimate in turn reflects the lifestyle of people in different countries and how wasteful they are. This is a much fairer reflection of the CO₂ emissions of each country than simply ranking the total emissions of each country, due to the fact that countries with larger populations are bound to release more CO₂. We are also comparing these data over a span of years to see how the carbon footprint changed over the years. From the trends we see over the past years we also want to calculate the linear regression and estimate the average carbon footprint of people in each country in the future. This estimate could reflect the outcome if we continue to emit CO₂ at the current growth rate.

Dataset Description

1: source: The origination of world bank

(https://data.worldbank.org/indicator/SP.POP.TOTL?name_desc=false)

file format: csv

The first column is the name of the country, the second column is the code of the country, the columns from 4th - 64th are the range from 1960-2019 for each country's yearly population.

2: source: Carbon Dioxide Information Analysis Center

(https://cdiac.ess-dive.lbl.gov/ftp/ndp030/nation.1751_2014.ems)

file format: text

Total fossil-fuel emissions are the sum of emissions from gas fuels, liquid fuels, solid fuels, gas flaring, and cement production. Carbon emissions per capita are measured as the total amount of carbon dioxide emitted by the country as a consequence of all relevant human (production and consumption) activities, divided by the population of the

¹Dunbar, Brian. “The Ups and Downs of Global Warming.” NASA, NASA, www.nasa.gov/topics/earth/features/upsDownsGlobalWarming.html.

²Canada, Environment and Climate Change. “Government of Canada.” Canada.ca, / Gouvernement Du Canada, 28 Mar. 2019, www.canada.ca/en/environment-climate-change/services/climate-change/causes.html.

country. Emission of bunker fuels, which is used especially on ships, and this index doesn't count in the total fossil-fuel emissions

Computational Overview

Description of the major the computations our program performs: Basically, our computation is separated into three parts: read in dataset, data analysis, and visualization. While reading the dataset, since the raw data (carbon_emission.txt) contains redundant information that we will not be using, we performed a filter operation on the redundant data. Also, you may notice in population.csv that the first line contains the header "country name" and the years. Since we only need the data below, we started to read the file from the second line. A transformation of datatypes is also included in our code. Since the raw data is separated by commas, we need to convert them into the dataclasses CarbonEmission and EmissionPerCapita. Then, since we were trying to graph out emission per person, an aggregation of two sets of data is needed here to calculate emission per capita.

Also, machine learning is used in our project. Like the idea in assignment 1, a computational model of linear regression is used here. We generate a stream of increasing arbitrary numbers to predict future growth of emission.

How our program reports the results of your computation in a visual and/or interactive way: When we run the file in the console, a webpage will show up, visualizing two still bar charts showing the data in 1960. You will notice a "play" button on top left. After clicking on it, the graphs will show the dynamic change of data from different countries over several decades over time, all the way to 2014. Also, in the dynamic line chart, you can place the cursor on an arbitrary piece of the line to check out the specific amount of emission in a specific year.

How our program uses new libraries to accomplish its tasks: Libraries we used: dataclasses, re, csv, plotly, plotly.io, plotly.subplots, random and typing. We use the csv module to load in the raw data. Dataclasses is a module that we use to make our work more easily to read and rewrite.

The most important library in this project is plotly since the visualization relies fully on that. The library plotly.objects is to generate the main dynamic bar chart.

The library plotly.subplots is to generate the second dynamic line chart.

Also, we used re to help process the strings in the raw data file (population.csv and carbon_emission.txt). For example, we used `\n(\w+)\n` to process the line with the country name, `\s+Total\s.*\n` to process the line start with at least 1 space and "Total Fossil-Fuel", `Year\s+.*\n\n` to process the line start with "Year Emissions", and `((\d+\s+\d+\s+.*\n)+)` to process the lines with actual data.

We used the random library to build the machine learning data prediction part of our work. Like assignment1, we generated some additional sets of data in the following years for the visualization part.

A Brief Description of Any Changes

In the proposal, we talked about the possibility of importing pandas into our file. But the raw data are already in csv and txt format, so we do not need that module here. Instead, we used a different library re to help process the data.

We used to put the total emissions data just to the right of the amount of emission of the top 10 countries. Later, we found that the total emission data will be much larger than other data shown in the graph, which will make the graph ugly. What is worse, it is no use comparing the total emission to the top 10. So, we separate out the total emission data and drew a new dynamic line chart to show the dynamic change of total emission throughout the years. We believe this idea will make our work more straightforward.

We abandoned the large circular, earth-like image idea in the proposal because we believe it would not be useful to answering the main question we raised from the beginning of the project. The three small icons on the right are also abandoned.

We did not use the idea of transforming the raw data into tabular data. Instead, we used dataclass to help locate and process the data.

Furthermore, we also added inline citations and changed our research question to be more precise.

Discussion of the Results of Our Exploration

The visualization of our project answered the question we raised directly by proving the given and the forecasted data shown on the graph.

We were also limited by the absent data for a lot of countries in the data sets, so we had to filter out countries with incomplete data and only work with the 90 countries that have both 55 years of population and emission data.

The next steps may include more detailed exploration, such as research into the top 2 or 3 countries. By looking for additional datasets, we can analyze the emission structure of a specific country. To be specific, the total emission of a country may include automobile exhaust, Industrial emissions, Agricultural emissions and so on. We can perform a deeper analysis into these data to help identify and optimize the concerned policies. Thus, help in reducing the overall pollution to the planet.

References

Canada, Environment and Climate Change. “Government of Canada.” Canada.ca, / Gouvernement Du Canada, 28 Mar. 2019, www.canada.ca/en/environment-climate-change/services/climate-change/causes.html.

“The Causes of Climate Change.” NASA, NASA, 18 Aug. 2020, climate.nasa.gov/causes/.

“Climate Change: Atmospheric Carbon Dioxide: NOAA Climate.gov.” Climate Change: Atmospheric Carbon Dioxide — NOAA Climate.gov, 14 Aug. 2020, www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide.

Dunbar, Brian. “The Ups and Downs of Global Warming.” NASA, NASA, www.nasa.gov/topics/earth/features/upsDownsGlobalWarming.html.

“How to Reduce Your Carbon Footprint.” The New York Times, The New York Times, www.nytimes.com/guides/year-of-living-better/how-to-reduce-your-carbon-footprint.