GREENHOUSE GAS EMISSIONS PREDICTIONS 2023 - 2025 REPORT

Greenhouse Gas (GHG) emissions are a major contributor to climate change: "As greenhouse gas emissions blanket the Earth, they trap the sun's heat. This leads to global warming and climate change. The world is now warming faster than at any point in recorded history".

• United Nations - Causes and Effects of Climate Change



INTRODUCTION

To better understand the amount of carbon sequestration needed to offset increasing global GHG emissions, I have been contracted by CO2BGone, a hypothetical engineering firm, to analyze a database containing GHG emissions data for all countries. In order for the engineering team to be successful, my analysis must include the following points below:



To understand who our market is, determine the top 10 countries influencing climate change.



For these countries, show how their total GHG emissions have changed over time.



Develop a Machine Learning (ML) model for the top 3 countries influencing climate change and predict their 2023 - 2025 emissions.

The following pages detail my findings and recommendations to the team.

*Please note that the accuracy of my findings is limited to the dataset's accuracy. For more details please see the appendix.

TABLE OF CONTENTS

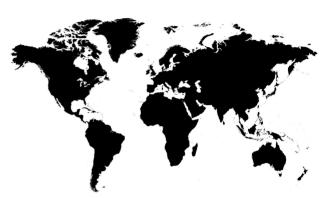


TOP 10 COUNTRIES INFLUENCING CLIMATE CHANGE

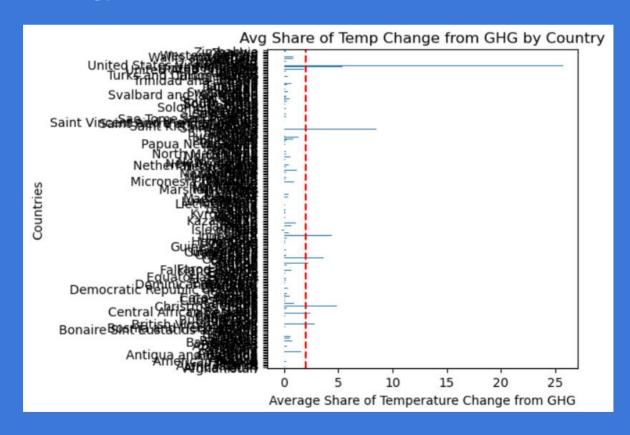
Background

The dataset contains a column titled "Share of Temperature Change from GHG" which shows how much each country's GHG emissions were responsible for the temperature change that year compared to the rest of the world.

For example, in 2019 the United States was responsible for 17.6% of the temperature change due to their GHG emissions compared to the rest of the world.



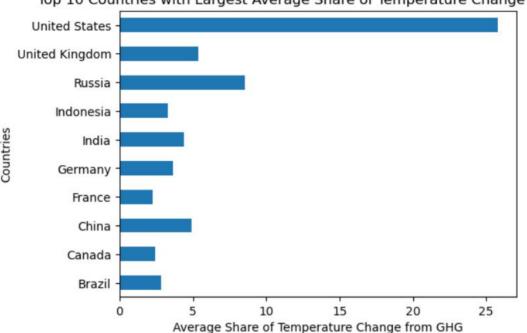
Methodology



To determine which countries had the biggest influence on temperature change I plotted the average share of temperature change from GHG for each country. I visually interpreted that the top 10 countries had a share greater than 2% (shown by the dotted red line). I then used this knowledge to filter the dataset to only show countries greater than 2%.

TOP 10 COUNTRIES INFLUENCING CLIMATE CHANGE

Result

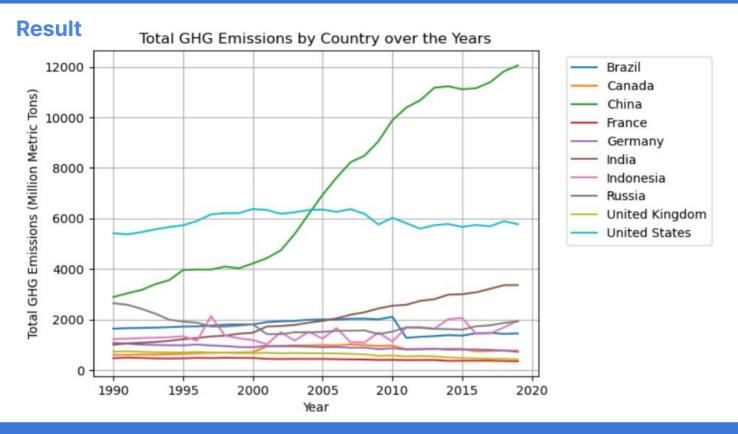


Top 10 Countries with Largest Average Share of Temperature Change from GHG

Summary

The above 10 countries had the largest influence on climate change on average. While this average is useful to help us determine who our key clients should be, we need to also visualize how these countries' emissions have changed over time. This will help us see whose emissions have increased and whose emissions have decreased. For countries with decreasing emissions they may be better to pursue as clients as they are actively working to lower their GHG emissions.

TOTAL GHG EMISSIONS FOR TOP 10 COUNTRIES OVER TIME



Summary

The above plot shows how the total GHG emissions have changed over time for the top 10 countries identified. The time span covers data from 1990 - 2019.

Insights:

- Since the early 2000s China surpassed the US as the leader in GHG emissions
- US emissions have decreased since the early 2000s
- India is the third biggest GHG producer and their emissions have increased since the early 2000s
- The fourth and fifth biggest GHG producers are Indonesia and Russia

Top 3 Countries Influencing Climate Change







2023 - 2025 MACHINE LEARNING (ML) PREDICTIONS

Results (in Million Metric Tons)

	China	United States	India
2023	14,857.8	6825.2	3605.0
2024	16,581.0	7192.1	3630.8
2025	18,943.5	7617.7	3642.2
MSE of ML model	0.004	0.098	0.002

^{*}MSE = Mean Squared Error (lower values = higher accuracy)

Methodology

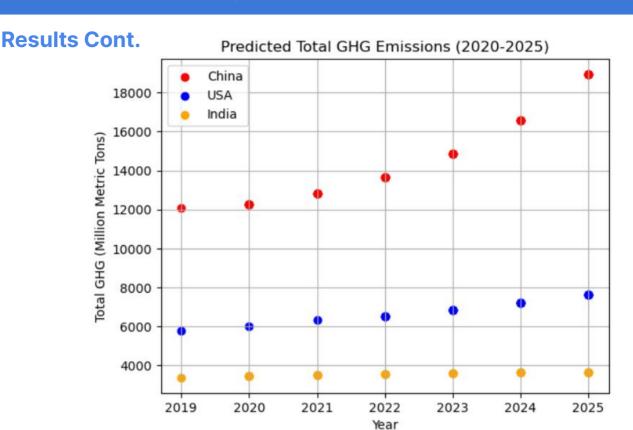
For all countries, I built a polynomial regression machine learning model to predict future data points, which are shown in the table above.

For the US since the dataset only covered data until 2019 I used the model to first predict 2020 - 2021 data and compared it to the actual total GHG emissions for those years according to the EPA. The model was able to predict this with only a 1.2% error. 2022 was also predicted, but there was no public data source that validated actual 2022 data.

I then updated the model to include the 2020 - 2021 data to predict 2023 - 2025 total GHG emissions that are shown above. The model's low percent error comparing its predictions to unseen data coupled with the low MSE value is assuring that the model will give us predictions with higher accuracy.

For China and India I could not find reliable sources to validate 2020 - 2022 total GHG data so I had the model also predict those values.

2023 - 2025 MACHINE LEARNING (ML) PREDICTIONS



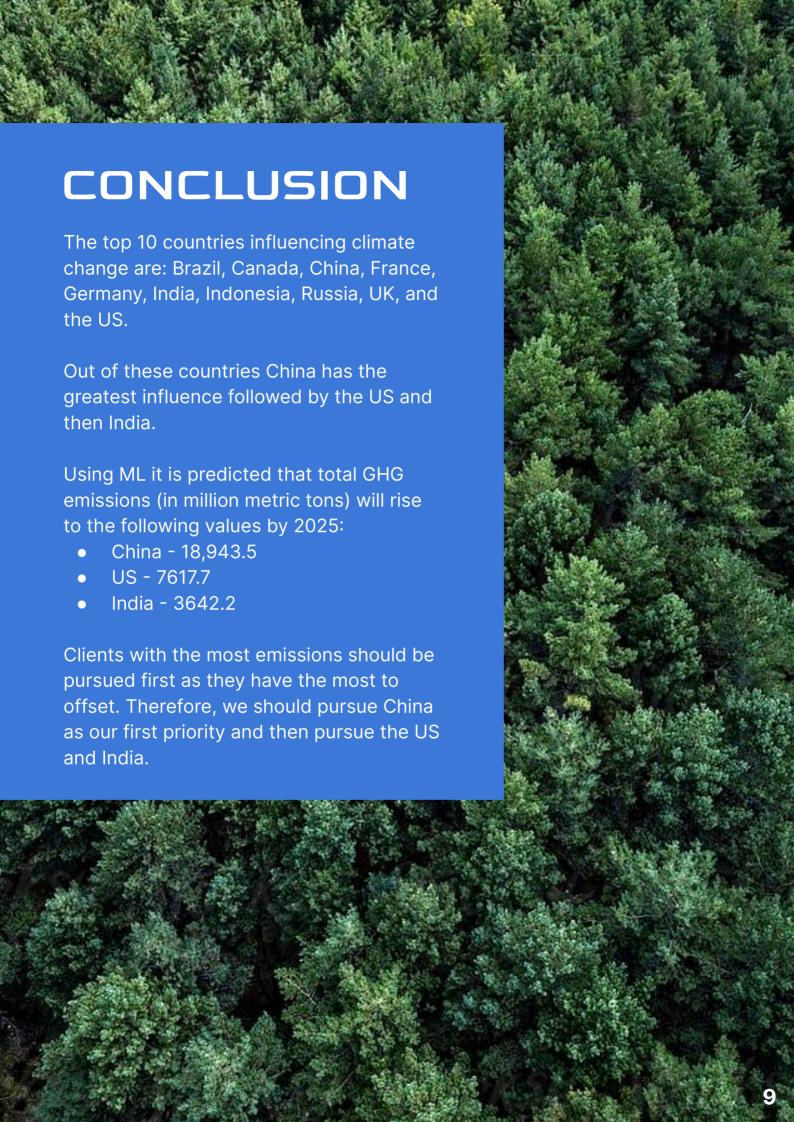
Summary

The above chart shows the predicted data until 2025. Please note that for the US the data from 2019 - 2021 is actual data. For China and India any data after 2019 was predicted using ML.

For all countries we see that their GHG is expected to increase over the next 3 years. China has the fastest growing rate for GHG emissions so they would be a prime potential client to offer carbon sequestration services to if they are looking to offset these emissions. China's emissions are expected to be double those of the US, and over triple those of India.

Please note that China has already begun projects that improve their carbon sequestration, most notably the Great Green Wall, which "aims to plant around 88 million acres of forests" by 2050 according to earth.org. This may decrease their need for hiring outside consultants.

Based on emission trends I would recommend pursuing the clients in the following order: China, US, and then India as the more emissions a country has, the more they would need to offset.



Appendix

DATA SOURCES

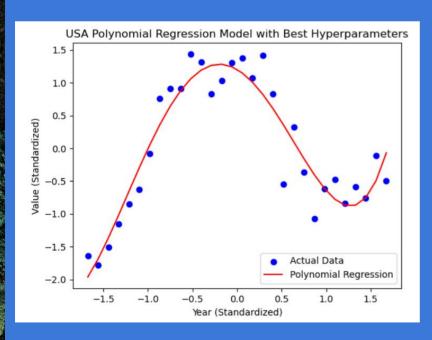
The database is from <u>Our World in Data</u> whose goal is to publish the "research and data to make progress against the world's largest problems" - <u>Our World in Data</u>.

Please note that while they work to ensure their data is of high-quality, Our World in Data cannot guarantee that their data is error-free. Findings from this report are limited to the accuracy of this data.

The United States Environmental Protection Agency (EPA) was referenced to validate total US GHG for <u>2020</u> and <u>2021</u>.



MACHINE LEARNING ALGORITHM



Originally I tried doing a linear regression model, but found its values for MSE were very high. Thus, I proceeded to use a polynomial regression model as shown to the left.

To build my ML model, I performed a ridge regression to prevent overfitting of the model, and also ran a grid search with cross validation

to ensure the parameter alpha (for ridge regression) and polynomial degree were optimized. This optimization was ran separately for each country's data as the parameters could vary since the data changed.

Appendix



If you would like to learn more about how I analyzed the data or how I built my ML algorithms please refer to my Python project code <u>here</u>.

If you have questions or are interested in discussing my findings please feel free to reach out to me on LinkedIn.

CLIMATE CHANGE



Climate change impacts every country, culture, and person across our planet. By understanding the data associated with weather, pollution, and geological features we can work to develop solutions to slow and reverse the effects of climate change.

My hope is that this project inspires you to investigate the ways our planet is changing so that you can be a knowledgeable global citizen.

- Joseph Westover