Climate Change and Global Warming: The Role of Carbon Dioxide Emissions

1   Introduction

Climate change is arguably the biggest environmental threat that the Earth is facing in the 21st century. It refers to changes in the Earth's climate system, contributed by both natural processes and human activities. These natural processes include cyclical ocean patterns such as the El Niño, volcanic activities and variations in solar energy from the Sun [1]. On the other hand, human activities, particularly since the Industrial Revolution at the beginning of the 20th century, have been causing irreversible damage to the Earth's climate system through emissions of greenhouse gases and the corresponding impact of global warming. This report provides a data-driven analysis of the primary greenhouse gas, carbon dioxide (CO_2), and its role in global warming and climate change.

1.1   Scope

In Section 2, we will investigate the authenticity of the global warming phenomenon using various key indicators. We will provide a scientific explanation of global warming in Section 3. In Section 4, we will analyse trends in levels of global CO_2 emissions and discuss which countries ought to be most responsible. In Section 5, we will apply a statistical model to understand the relationships between CO_2 emissions per capita and various socio-economic variables. We will discuss limitations of our analyses in Section 6, before proposing potential areas for future research in Section 7. All analyses underlying this report were performed using the R statistical software [2] and are reproducible using the code available at `www.github.com/zeyalt/climate-change`.

1.2   Datasets

The main datasets used in this report include the global temperature and sea level datasets from National Aeronautics and Space Administration (NASA), carbon dioxide emissions datasets from Global Carbon Project (GCP) and Carbon Dioxide Information Analysis Centre (CDIAC), as well as various demographic datasets from the World Bank and United Nations (UN) Population Division. A full list of the datasets used and their descriptions is presented in the Appendix. These datasets are also available at `www.github.com/zeyalt/climate-change`.

2   Is Global Warming Real?

The authenticity of global warming can be validated by analysing changes in the Earth's climate system over time. In this section, we will focus on two key indicators—global temperatures and sea levels.

2.1   Increasing Global Temperatures

The first key indicator is changes in global temperatures, measured in terms of temperature anomalies[1]. A time-series plot showing variations in global land-ocean temperature anomalies[2] from 1880 to 2019 with data from NASA [5, 6] is shown in Figure 1. These temperature anomalies were calculated based on the average temperature of the baseline period between 1951 and 1980.
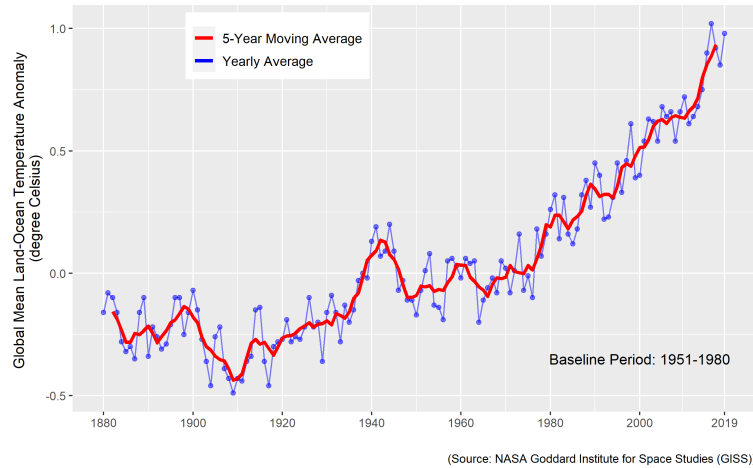

Figure 1: Variation of global mean land-ocean temperature anomaly, 1880-2019

Figure 1 shows that since the 1970s, temperature anomalies have consistently been positive[3], indicating that the Earth has been warmer than the baseline period. Since then, within a span of 50 years, the temperature anomaly has risen by 1 degree Celsius, a clear evidence of global warming. In fact, global temperatures had been projected to continue increasing in the future, regardless of any global intervention measures [7].

2.2   Rising Sea Levels

Rising global sea levels provides another strong evidence of global warming. Figure 2 shows a time-series plot of the global mean sea levels (GMSL) since 1993 with data from NASA [8]. Between 1993 and 2019, the GMSL increased by almost 10 centimetres. This was largely caused by melting glaciers and ice sheets in Greenland and Antarctica as a result of a warmer Earth [9]. In addition, a vast amount of heat trapped by greenhouse gases in the atmosphere is absorbed by the oceans,

---

[1]Temperature anomaly is the difference between an observed absolute temperature and the average temperature over a baseline period. It is typically preferred over absolute temperatures in climate studies because the latter is more susceptible to variations in measurements due to locations and altitudes of weather monitoring stations [3].

[2]Land-ocean temperature anomalies take into consideration both land and sea surface temperatures [4].

[3]A positive temperature anomaly means that the observed temperature was higher than the average temperature of the baseline period; a negative anomaly indicates otherwise [3].

causing thermal expansion of the ocean water and consequently a rise in sea levels [10].
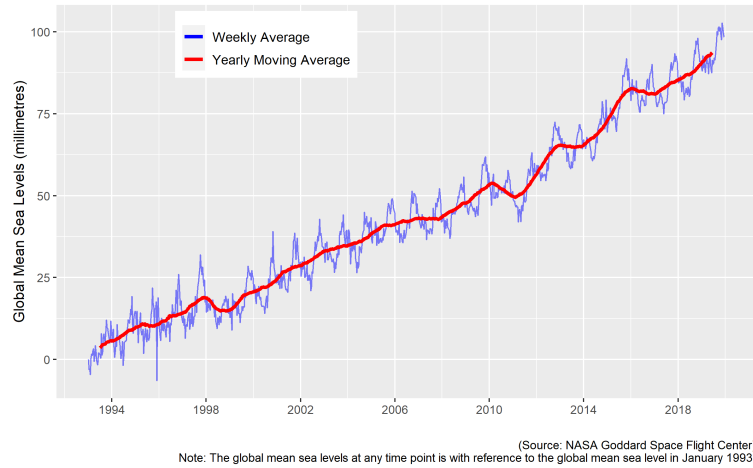

Figure 2: Variation of global mean sea levels, 1993-2019

The threats of higher global temperatures and rising sea levels cannot be understated. One manifestation of these phenomena is the occurrences of catastrophic weather events such as floods, typhoons and tsunamis [11], which lead to losses of lives, disruptions to livelihoods and destruction of natural habitats of animals and plants. Such events are not only increasing in frequency [12], but also in intensity. For example, a study by Kossin in 2018 found that hurricanes and typhoons since the 1980s had been causing greater damages such as unprecedented rainfalls than they did before [13].

3   What Causes Global Warming?

Having demonstrated that global warming is real, we will, in this section, provide a scientific explanation of what causes it. In essence, global warming is the result of an "enhanced greenhouse effect". When solar energy from the Sun hits the Earth's surface, most of it is absorbed but the rest is reflected off. Greenhouse gases in the Earth's atmosphere, such as CO_2, nitrous oxide (N_2O), methane (CH_4) and water vapour (H_2O), trap some of the reflected solar energy and dissipate them within the Earth's atmosphere. This is known as the "greenhouse effect" [14].

The right equilibrium of greenhouse gases in the atmosphere is necessary for survival on Earth. Without these greenhouse gases, the Earth would be significantly colder and many life forms would not survive. However, this equilibrium had been disrupted, as more greenhouse gases were introduced into the atmosphere in the last century. This enhances the greenhouse effect, causing Earth to heat up [14].

There has been an overwhelming consensus in the scientific community that global

warming had been a direct result of human activities [15]. A study by Cook et al. found that this consensus was shared by "90% to 100% of publishing climate scientists" [16]. These human activities are being pursued by countries in the interest of urbanisation and economic growth. Examples include fossil fuel combustion, deforestation, urbanisation and trade. An unintended consequence is that more fuel are being burned to meet the world's increasing energy needs and standards of living, thereby releasing more greenhouse gases into the atmosphere.

4   Key Contributors of Carbon Dioxide Emissions

The primary greenhouse gas is CO_2 and it is emitted mainly through human activities. In this section, we will analyse trends of two major contributing factors of CO_2 emissions—fossil fuel combustion and deforestation.

4.1   Fossil Fuel Combustion

Fossil fuels are derived from fossilised remains of plants and animals that lived on Earth millions of years ago [17]. As a result of how they were formed, fossil fuels contain high amounts of carbon. When they are burned to meet the world's energy demands, tremendous amounts of stored carbon are released into the atmosphere as CO_2. These increase the levels of greenhouse gases in the atmosphere, thereby exacerbating the greenhouse effect.

(Source: GCP and CDIAC.)
Note: Data before 1959 was sourced from CDIAC; data from 1959 onwards was obtained from GCP.
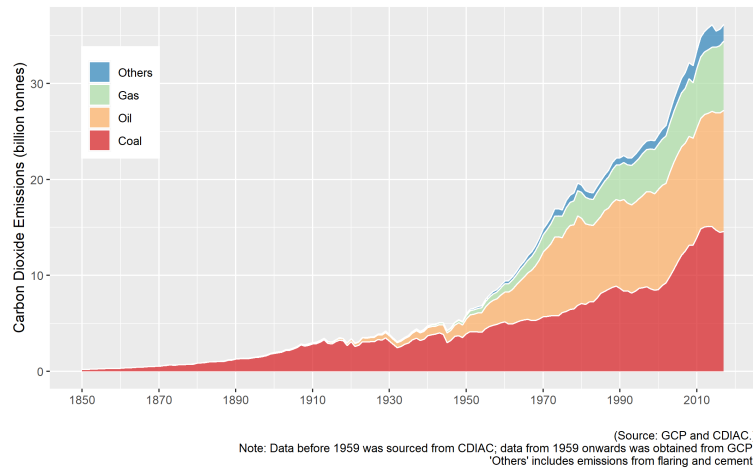'Others' includes emissions from flaring and cement.

Figure 3: Global carbon dioxide emissions by fuel types, 1850 to 2017

There are three main types of fossil fuels—coal, oil and natural gas. Figure 3 shows how global CO_2 emissions by these fuel types have changed since 1850. Until the early 20th century, coal was the predominant source of fuel and the only contributor to global CO_2 emissions. Around the mid-20th century, oil and gas emerged as alternative fuels. Consumption of these alternative fuels released more CO_2, accelerating the increase in global emissions. In 2017, coal and oil were the

dominant contributors of CO_2 emissions, accounting for more than three-quarters of global emissions.

4.2   Deforestation

Table 1 summarises the changes in the global forest cover area since 2000 over 5-years intervals[4]. These numbers indicate that deforestation is happening, evidenced by the decreasing area of global forest cover over the last two decades. It can also be inferred that an average of 19.7 million hectares of forest cover was lost every five years since 2000. This translates to a yearly average of 3.93 million hectares.

Year	Global Forest Area (billion hectares)	Absolute Change (billion hectares)	Percentage Change
2000	4.05	-	-
2005	4.03	-0.024	-0.580%
2010	4.01	-0.018	-0.438%
2015	3.99	-0.017	-0.427%

Table 1: Changes in the total area of global forest covers, 2000-2015

(Source: UN Statistics Division)
Note: Factor of change was calculated from dividing the area of forest cover
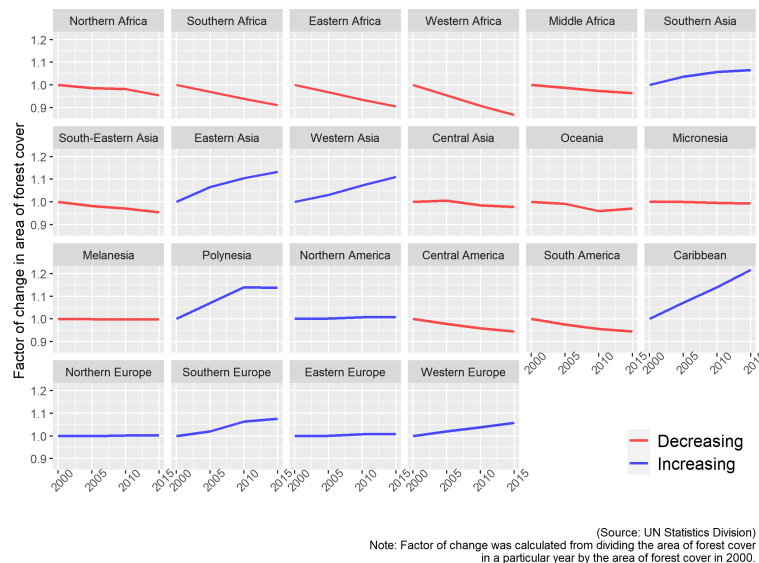in a particular year by the area of forest cover in 2000.

Figure 4: A breakdown of deforestation trends from 2000 to 2015 by geographical regions

A breakdown of the deforestation trends by geographical regions is shown in Figure 4. Evidently, deforestation had been more rampant in some regions than others. For instance, Africa saw decreases in forest covers in every geographical region since 2000, with the steepest decrease in West Africa. Other regions with losses in forest covers include South-Eastern Asia and South America, with the latter being home to Amazon rainforest. On the flip side, regions like Europe, Polynesia, Caribbean and most of Asia had been showing encouraging signs of growing forest cover. In

---

[4]This data was provided by the Food and Agriculture Organisation (FAO) of the UN, who collects and analyses data on forest area at intervals of 5 to 10 years.

Europe, for example, this had been due to increased forest protection, decline in farming and reforestation efforts [18].

Deforestation has been happening for a few reasons. One key driver is the production of timber and wood. Vast amounts of trees are also cut down and burned for agriculture reasons, such as the production of palm oil and soy [19]. In addition, forests are cleared to make way for infrastructures such as roads and buildings.

Deforestation has detrimental effects on the Earth's natural environment. The most immediate consequence is the destruction of eco-systems of many species of insects, plants and animals living in forest environments. It also contributes to more CO_2 in the atmosphere. Given that forests play an important role in regulating CO_2 and oxygen in the atmosphere, deforestation diminishes this ability of forests to regulate CO_2 levels as there are lesser trees to remove CO_2 from the air. Moreover, when cut trees are burned, they release stored carbon as CO_2.

5 An Analysis of CO_2 Emissions Levels

Besides fossil fuel combustion and deforestation, many other human activities such as trade and transportation also emit CO_2. In this section, we will analyse the total global CO_2 emission levels and how they have changed over time. We will also discuss which countries should be most responsible for CO_2 emissions.

5.1 How Have Emission Levels Changed Over Time?

Figure 3 presented in Section 4.1 also depicts the change in global CO_2 emissions over the years. It shows that the global CO_2 emissions had been on an upward trend since 1850, increasing by 180-folds from about 0.2 billion tonnes (bt) in 1850 to more than 36 bt in 2017. The rate of increase of CO_2 emissions was much steeper from the 1950s onwards, after oil and natural gas emerged as alternative fuels. This also coincided with the post-war period, which suggests that more CO_2 could have been emitted as many countries sought to rebuild their economies after the war.

5.2 Who Should Be Most Responsible?

Since climate change is a global problem, every country ought to exercise a shared responsibility to tackle it. In reality, every country is contributing to CO_2 emissions in different magnitudes, some much greater than the others. One common argument is that countries which are the biggest emitters should bear the greatest responsibility. However, there are different interpretations of the phrase "biggest emitters".

One interpretation is based on the share of global CO_2 emissions in a recent year.

For instance, the share of global CO_2 emissions in 2017 is depicted by the treemap visualisation in Figure 5a. It shows that Asia accounted for more than half of global CO_2 emissions. China was the biggest emitter with a global share of 26.4%. The next major emitters were United States (14.2%), India (6.6%) and Russia (4.6%).
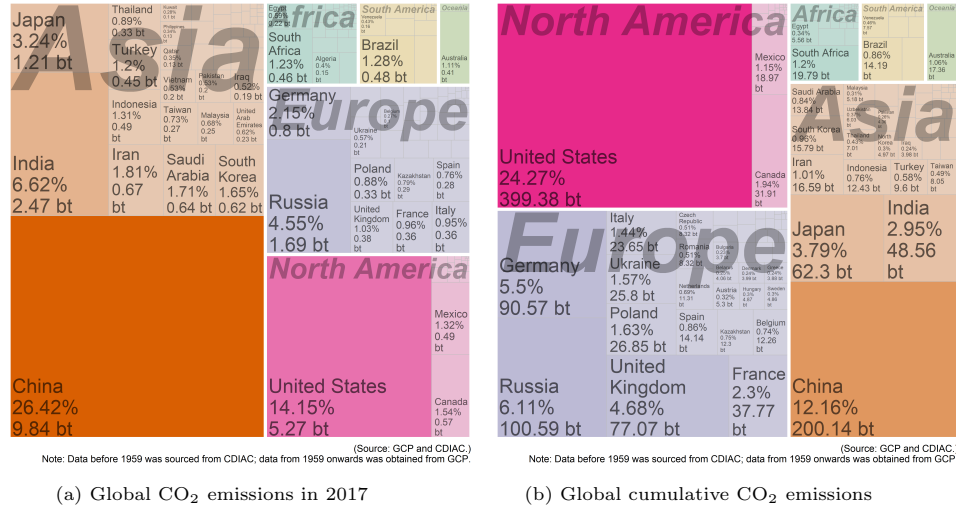

(a) Global CO_2 emissions in 2017

(b) Global cumulative CO_2 emissions

Figure 5: Global share of CO_2 emissions in 2017 and cumulative CO_2 emissions

The share of global CO_2 emissions changes from year to year as countries' levels of economic activities vary. Therefore, an alternative interpretation is based on cumulative CO_2 emissions. A treemap visualisation of the global share of cumulative CO_2 emissions from 1750 to 2017 is shown in Figure 5b. This yields a different picture. Evidently, United States is the leading emitter, accounting for almost 400 bt of emissions or a quarter of all CO_2 that the world has ever emitted. United States is, in fact, the biggest emitter by a significant margin. The next biggest emitter, China, is responsible for more than 12% of cumulative CO_2 emissions.

Several major emitters in 2017, such as Iran, Saudi Arabia and South Korea, were considered small contributors in terms of cumulative emissions. The converse was also true, that is, countries with the highest CO_2 emissions over history were not necessarily the biggest contributors in 2017. An example is the United Kingdom, which accounted for 4.68% of cumulative CO_2 emissions but only 1.03% of the annual CO_2 emissions in 2017. Another observation from Figure 5a and Figure 5b is that Africa, South America and Oceania had consistently been small CO_2 emitters relative to Asia, North America and Europe.

It is arguable that major emitters like China, United States and Russia emit more because of their bigger population sizes. Therefore, another interpretation is based on the average amount of CO_2 emitted per person in the population, also known as "CO_2 emissions per capita". Figure 6 shows a treemap visualisation depicting

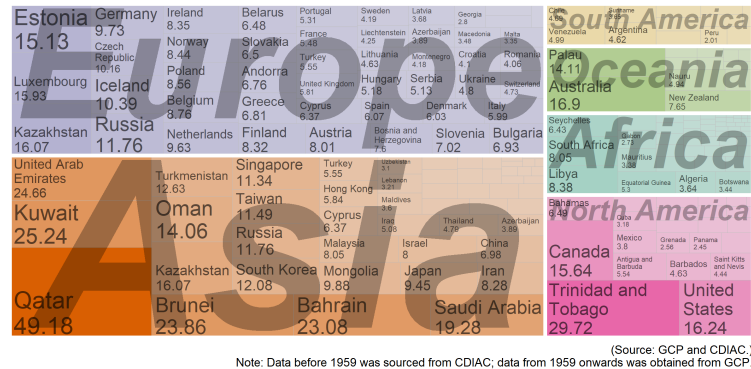the CO_2 emissions per capita of countries based on data in 2017.

Figure 6: Global share of CO_2 emissions per capita in 2017

Figure 6 tells very different stories from Figure 5a and Figure 5b. Major emitters like China, United States, Russia and India, pale in comparison with countries like Qatar, Trinidad and Tobago and Kuwait. In fact, countries with the highest per capita CO_2 emissions in 2017 happened to be oil-producing countries in the Middle East. These countries also tend to have relatively smaller population [20]. Another observation from Figure 6 is that there were huge variations in CO_2 emissions per capita across continents as well as across countries within each continent. The variations across continents can be further visualised using boxplots, as shown in Figure 7.
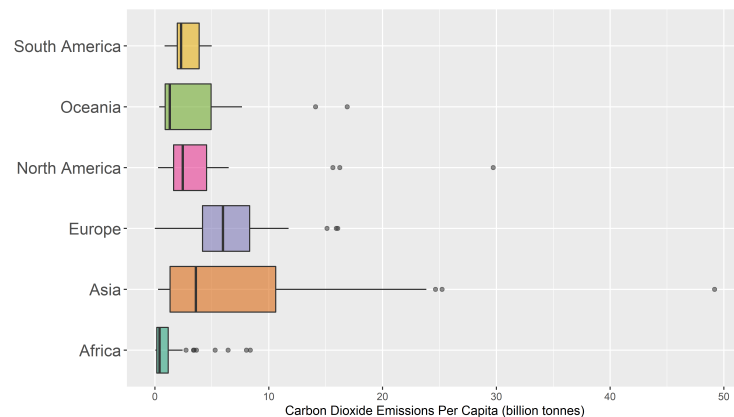

Figure 7: A graphical summary of CO_2 emissions per capita in 2017 by continents

Figure 7 shows several outliers[5]. These outliers correspond to countries with significantly higher CO_2 emissions per capita than other countries within the same continent. Examples include Qatar in Asia (49.2 bt), Trinidad and Tobago in North America (29.7 bt) and Australia in Oceania (16.9). Asia had the largest spread of

---

[5]In this context, outliers were defined as countries whose CO_2 emissions per capita is greater than the upper quartile value within the same continent and 1.5 times of the inter-quartile range of that continent.

CO_2 emissions per capita, with an inter-quartile range (IQR) of 9.28 bt. This suggests that Asia had countries with both very high and very low CO_2 emissions per capita, an observation which could also be inferred from Figure 6.

The continent with the highest median CO_2 emissions per capita was Europe (5.99 bt), followed by Asia (3.60 bt) and North America (2.45 bt). Africa, on the other hand, had the lowest median (0.42 bt) and smallest spread (1.00 bt) of CO_2 emissions per capita amongst all regions, albeit several outliers. This implies that there was a significant number of African countries with very low CO_2 emissions per capita. These observations suggest that there could be some correlation between a country's CO_2 emissions per capita and how developed it is or its standard of living. Correlations between CO_2 emissions per capita and such socio-economic factors will be further discussed in the next section.

6 Modelling CO_2 Emissions Per Capita

The analysis from the previous section suggested that an average person in a developed country emits more CO_2 emissions than one in a less developed country. In this section, we will further explore the relationship between CO_2 emissions per capita and standard of living, typically measured by gross domestic product (GDP) per capita[6]. In addition, we will investigate other socio-economic factors— life expectancy, urban population share and unemployment rate—with data from 2017, and how they correlated with CO_2 emissions per capita in 2017.

The key objective is to find a model that can best predict the response variable, CO_2 emissions per capita in 2017. Such modelling would be useful for international organisations like the UN in understanding key predictors of CO_2 emissions per capita, in order to coordinate global action in mitigating CO_2 emissions. Since CO_2 emissions per capita is a numerical continuous variable, a regression model would be appropriate. A multiple regression analysis will be performed to model CO_2 emissions per capita in 2017 with these four predictor variables. We will also use the cross-validation technique to guide us towards choosing the best model.

6.1 Visualising the Correlations

The relationships between CO_2 emissions per capita and each variable are shown in Figure 8. Each data point represents a country, colour-coded by its continent. Life expectancy and urban population share appear to have similar trends with CO_2 emissions per capita. Their relationships are very different from those of CO_2 emissions per capita with GDP per capita and unemployment rate. A common

---

[6]GDP is the sum of all goods and services produced by a country within a specific time period. It is a measure of a country's economy. GDP per capita represents the average income of a person in a country within a specified time period. The higher a country's GDP per capita, the better its standards of living.

observation across these scatterplots is that they all appear to have a certain degree of non-linearity in the trends.
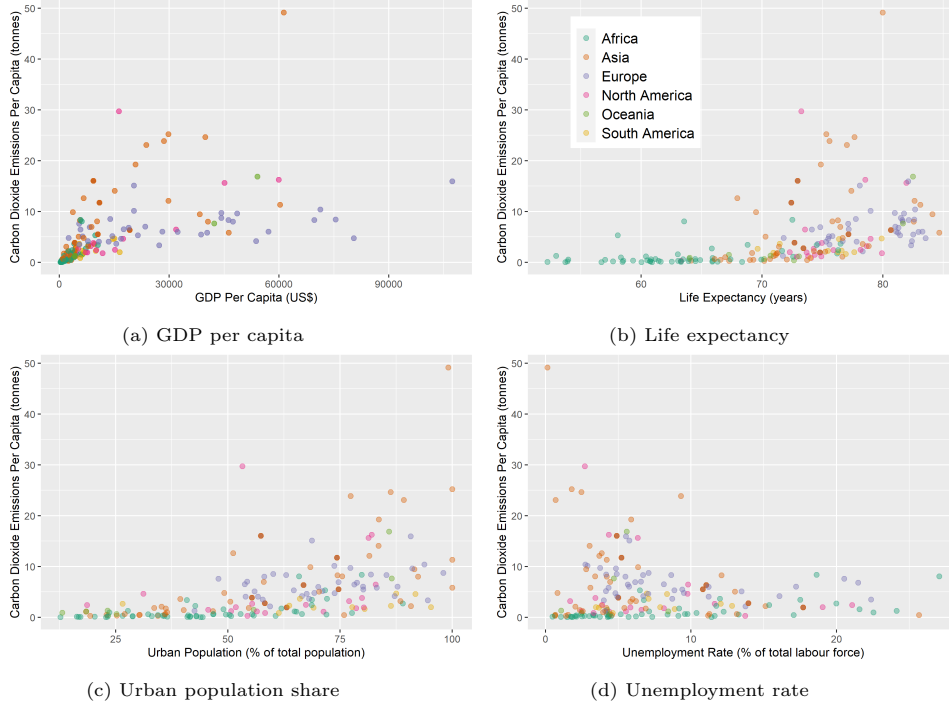

(a) GDP per capita

(b) Life expectancy

(c) Urban population share

(d) Unemployment rate

Figure 8: Correlations of CO_2 emissions per capita with socio-economic variables

In linear regression, it is assumed that linear relationships exist between the predictor variables and the response variable [21]. Therefore, the non-linear patterns in Figure 8 needed to be transformed into linear forms before linear regression could be applied. A logarithmic transformation was assessed to be suitable because it is a monotonically-increasing function and it preserves the order of the data points. Figure 9 shows the scatterplots after logarithmic transformation was applied to the response variable.

It is evident from Figure 9 that logarithmic transformation linearised the relationships of CO_2 emissions per capita with life expectancy, urban population share and, to a weaker extent, unemployment rate, but not GDP per capita. The original graph in Figure 8a shows many data points clustered around the lower left of the graph. A logarithmic transformation of CO_2 emissions per capita spread these data points out along the y-axis. However, the resulting graph shown in Figure 9a was still far from a linear relationship. In fact, it closely resembled a logarithm graph. Applying a logarithmic transformation on GDP per capita produced the linear graph in Figure 10. These linearised forms thereby make it appropriate for further analyses and modelling with linear regression.

It can also be observed from Figure 9 that every scatterplot had best-fit regression
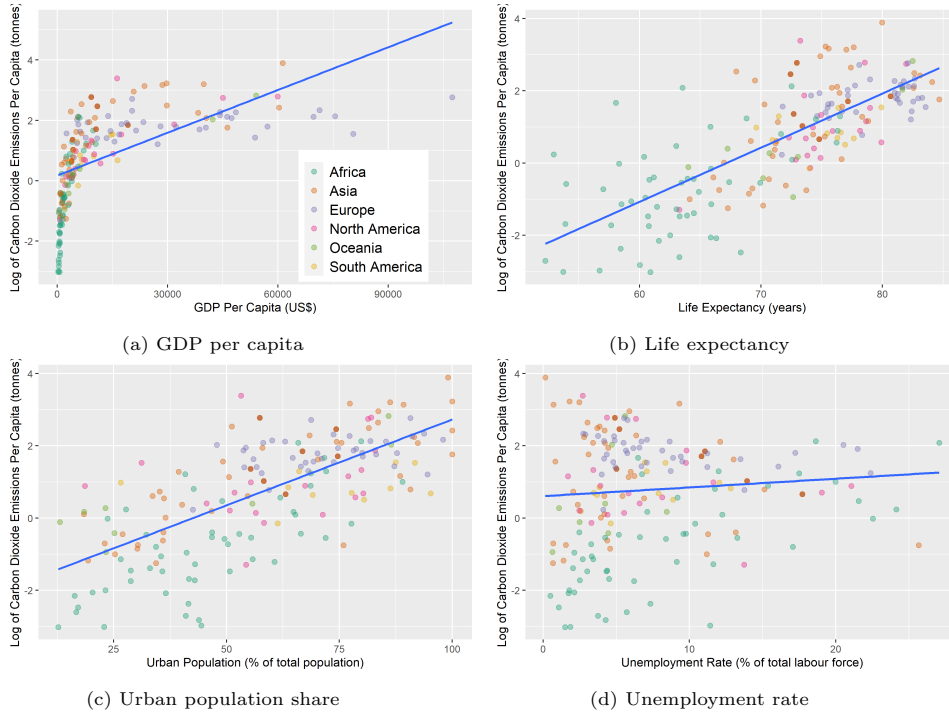
(a) GDP per capita

(b) Life expectancy

(c) Urban population share

(d) Unemployment rate

Figure 9: Correlations of CO_2 emissions per capita with socio-economic variables after logarithmic transformation

lines with positive gradients, suggesting some form of positive relationships between CO_2 emissions per capita and the predictor variables. This implies that countries with higher GDP per capita, life expectancy, urban population share and unemployment rate tended to emit more CO_2 emissions per capita. However, the variable, unemployment rate, appeared to have a very weak positive relationship with the response variable, indicating that unemployment rate alone was not a good predictor of CO_2 emissions per capita in 2017.
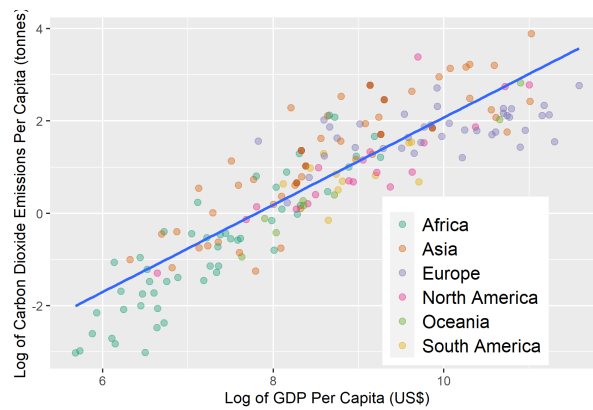

Figure 10: Correlation of logarithm of CO_2 emissions per capita and logarithm of GDP per capita

6.2   Multiple Linear Regression

Using the linearised forms, the regression model can be represented as,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4, \tag{1}$$

where Y is a random variable representing the logarithm of CO_2 emissions per capita, and X_1, X_2, X_3 and X_4 are random variables representing logarithm of GDP per capita, life expectancy, urban population share and unemployment rate respectively. β_0 is the intercept of the regression line, and β_1, β_2, β_3 and β_4 are the coefficients associated with the each predictor variable. Multiple linear regression was performed in the R statistical software and the results are shown in Table 2.

Input variables	Coefficients	Estimates	P-value	Lower CI	Upper CI
(Intercept)	β_0	-7.77	8.13e-29	-8.91	-6.64
Log of GDP per capita	β_1	0.788	1.25e-17	0.625	0.951
Life expectancy	β_2	0.0163	0.214	-0.0095	0.0421
Urban population share	β_3	0.00626	0.0839	-0.000849	0.0134
Unemployment rate	β_4	0.0198	0.0447	0.000478	0.0391

Table 2: Results of multiple linear regression

With the estimates generated from this analysis, the regression model can be represented as,

$$Y = -7.77 + 0.788X_1 + 0.0163X_2 + 0.00626X_3 + 0.0198X_4. \tag{2}$$

All coefficients associated with the predictor variables had positive values, consistent with the trends in Figure 9. The relationship between the logarithm of GDP per capita (X_1) and the response variable (Y) was particularly strong with a coefficient of 0.788, which was evident from Figure 10. Other predictor variables, especially urban population share, had significantly smaller coefficient values. The p-values for the logarithm of GDP per capita and unemployment rate were less than 0.05. This suggests that both variables were statistically significant at 5% significance level, that is, they were significant predictors of the response variable.

The uncertainties around the true values of the coefficients are represented by the 95% confidence intervals in Table 2. The 95% confidence intervals of the coefficients for logarithm of GDP per capita and unemployment rate were entirely in the positive range and did not contain zero, which implies that there is 95% probability that their correlations with the response variable were positive. However, unlike the logarithm of GDP per capita, unemployment rate had a confidence interval which was much closer to zero and is therefore more unstable. On the other hand, the confidence intervals for life expectancy and urban population share contained zero, which means that the hypothesis that these variables did not have an effect on the response variable cannot be rejected.

6.3   Choosing the Best Model

In the previous section, we fitted a multiple linear regression model using all four predictor variables. The results showed that not all variables shared equal importance in terms of statistical significance. With four predictor variables, there are, in fact, 15 different combinations of variables which can be used to model CO_2 emissions per capita. This section further explores these combinations and seeks to find the model that best explains CO_2 emissions per capita in 2017.

For each combination of predictor variables, the original data was randomly split into two subsets—75% as training data and the remaining 25% as test data. A linear regression model was fitted with the training data. Using the fitted models, predictions of the response variable were made given input values in the test data. Finally, these predictions were compared against the actual values in the test data using root-mean-squared error (RMSE)[7]. The best model is the combination which produced the lowest RMSE.

Since the training data was randomly sampled, repeating these steps might generate a different "best" model. Thus, we performed cross-validation with 1,000 iterations to determine the combination that produced the lowest RMSE most frequently. In general, cross-validation also helps to improve the predictive performance of a model by preventing over-fitting of data, a common problem that diminishes a model's ability to generalise to unseen data. The results of cross-validation are summarised in Figure 11. Based on these results, the best model was Model 15. This was the model with all four predictor variables explored in the previous section and is represented by Equation (2).
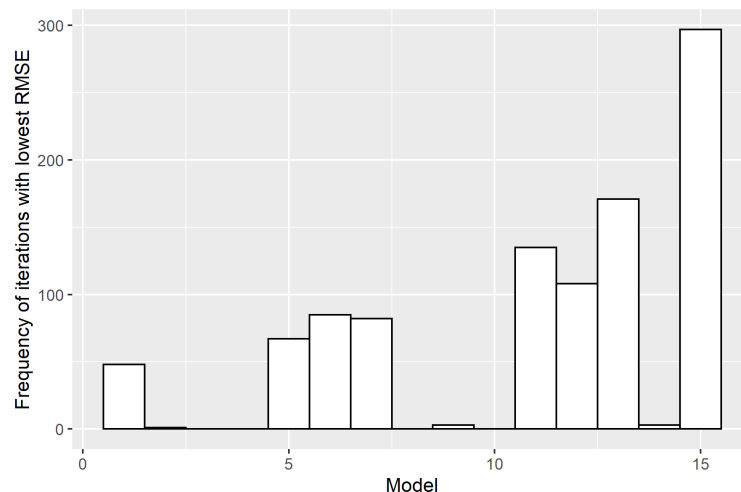

Figure 11: Frequency distribution of models which yield predictions with lowest RMSE

---

[7]RMSE is the standard deviation of the difference between predicted and actual values. It is a measure of how close the fitted regression line is from the data points.

6.4   Correlation and Causality

It is a common pitfall to equate correlation with causality. Correlation represents the size and direction of relationships between variables, whereas causality implies some form of cause-and-effect relationships [**?**]. Despite the need to distinguish the two, correlations provide useful information and can help guide further research into the causal analysis between variables.

In our statistical model from the previous section, it was found that the socio-economic variables had positive correlations with CO_2 emissions per capita in 2017. For instance, our model suggested that countries with higher urban population share emitted more CO_2 emissions per capita. It would be erroneous to conclude that higher urbanisation levels necessarily caused higher CO_2 emissions per capita in 2017, but further research can be conducted to investigate why and how more urbanised countries could have potentially emitted more.

7   Limitations of Analyses

The analyses presented in this report were not without limitations. One key aspect was data quality. Given that most datasets used in this report contain data collated from different countries, the methodologies in data collection and quality assurance practices would hardly be consistent. The datasets on CO_2 emissions came from two different sources: the CDIAC before 1959 and the GCP from 1959 onwards. This meant that there would be differences in how CO_2 emissions were measured and recorded. Such inconsistencies could also stem from the fact that some data used in this report, for instance the global temperature anomalies and CO_2 emissions, dated back to the 1800s when equipment and methods of measurements were not as advanced as today.

Our analyses were also limited by availability of data. An example was the global forest area data collected by the FAO of the UN, which was only available at intervals of 5 years. As much as it would be challenging to obtain precise forest area data from all countries on a yearly basis, such limitation hampers deeper analyses of specific deforestation trends in different regions. Limited availability of data also affected our regression analysis in Section 6. There were about 50 countries without data on at least one of the predictor variables. To ensure that the regression analysis was performed with complete data, these countries had to be omitted.

8   Areas for Future Research

The regression model in Section 6 was a simple model with several underlying assumptions. A potential area for future research would be to investigate the

validity of these assumptions to improve the current model. One assumption was that multi-collinearity between the predictor variables was low. There is scope to ascertain if this assumption was valid and how it affects the model's performance. Other possible areas of future studies include modelling CO_2 emissions per capita by continents, since continents vary in terms of emissions per capita as we saw in Figure 7, modelling CO_2 emissions per capita with more predictor variables, as well as exploring other supervised modelling methods such as random forest and polynomial regression.

9   Conclusions

The evidence from data is irrefutable. The Earth's temperature is rising. Ice sheets are melting and sea levels are increasing. Extreme weather events are unleashing upon the planet more frequently and intensively than before. It is undeniable that the Earth's climate is changing, a result of global warming caused by increasing greenhouse gases in the Earth's atmosphere. In this report, we looked at the primary greenhouse gas, CO_2, and demonstrated with data how human activities have contributed to CO_2 emissions. We developed a regression model to predict CO_2 emissions per capita of each country. We analysed how emission levels had changed over time and learned which countries were the biggest emitters. We also saw that every country contributes to CO_2 emissions in different magnitudes.

Notwithstanding, climate change is a global problem and every nation should share a common responsibility in tackling it. A monumental milestone in the world's fight against climate change was the signing of the Paris Agreement by 197 countries in 2016. The Paris Agreement aims to strengthen global response to climate change by keeping a global temperature rise "well below 2 degrees Celsius above pre-industrial levels" in this century [22]. Amongst other obligations, the Paris Agreement requires countries to cut pollution and limit greenhouse gas emissions. Beyond such global action, societies and individuals should also change their mindsets and be more conscious about how their actions can impact the environment. Only then can the world become a cleaner and a more sustainable place to live in.

References

[1] Holly Shaftel. Overview: Weather, Global Warming and Climate Change. https://climate.nasa.gov/resources/global-warming-vs-climate-change.

[2] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.

[3] National Oceanic and Atmospheric Administration. Anomalies vs. Temperature. https://www.ncdc.noaa.gov/monitoring-references/dyk/anomalies-vs-temperature.

[4] NASA Goddard Institute for Space Studies. GISS Surface Temperature Analysis (GISTEMP). https://data.giss.nasa.gov/gistemp/faq/#q103.

[5] GISTEMP Team. 2020: GISS Surface Temperature Analysis (GISTEMP), version 4. https://data.giss.nasa.gov/gistemp/.

[6] N. Lenssen, G. Schmidt, J. Hansen, M. Menne, A. Persin, R. Ruedy, and D. Zyss. Improvements in the gistemp uncertainty model. *J. Geophys. Res. Atmos.*, 124(12):6307–6326, 2019.

[7] Hannah Ritchie and Max Roser. CO_2 and Greenhouse Gas Emissions. *Our World in Data*, May 2017.

[8] B. Beckley, N.P. Zelensky, S.A. Holmes, F.G. Lemoine, R.D. Ray, G.T. Mitchum, S. Desai, and S.T. Brown. Global Mean Sea Level Trend from Integrated Multi-Mission Ocean Altimeters TOPEX/Poseidon Jason-1 and OSTM/Jason-2 Version 4.2, 2016.

[9] A. Dutton, A. E. Carlson, A. J. Long, G. A. Milne, P. U. Clark, R. DeConto, B. P. Horton, S. Rahmstorf, and M. E. Raymo. Sea-level rise due to polar ice-sheet mass loss during past warm periods. *Science*, 349(6244), 2015.

[10] NASA Sea Level Change. Understanding Sea Level - Thermal Expansion. https://sealevel.nasa.gov/understanding-sea-level/global-sea-level/thermal-expansion.

[11] Christina Nunez. Sea level rise, explained. https://www.nationalgeographic.com/environment/global-warming/sea-level-rise/, February 2019.

[12] European Academies Science Advisory Council. Extreme weather events in Europe. Technical report, March 2018.

[13] James P. Kossin. A global slowdown of tropical-cyclone translation speed. *Nature*, (558):104–107, June 2018.

[14] NASA Global Climate Change. The Causes of Climate Change.
https://climate.nasa.gov/causes.

[15] NASA Global Climate Change. Scientific Consensus: Earth's Climate is
Warming. https://climate.nasa.gov/scientific-consensus.

[16] John Cook, Naomi Oreskes, Peter T. Doran, William R. L. Anderegg, Bart
Verheggen, Ed W. Maibach, J. Stuart Carlton, Stephan Lewandowsky,
Andrew G. Skuce, Sarah A. Green, Dana Nuccitelli, Peter Jacobs, Mark
Richardson, Bärbel Winkler, Rob Painting, and Ken Rice. Consensus on
consensus: A synthesis of consensus estimates on human-caused global
warming. *Environmental Research Letters*, 11(4):048002, April 2016.

[17] June 29 and Melissa Denchak. Fossil Fuels: The Dirty Facts.
https://www.nrdc.org/stories/fossil-fuels-dirty-facts.

[18] Johnny Wood. Europe's forests are booming. Here's why.
https://www.weforum.org/agenda/2019/07/forest-europe-environment/, July
2019.

[19] The World Counts. Why is Deforestation a Problem?
https://www.theworldcounts.com/challenges/planet-earth/forests-and-
deserts/why-is-deforestation-a-problem.

[20] Gapminder Data. Data. https://www.gapminder.org/data/.

[21] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An
Introduction to Statistical Learning with Applications in R*. Springer, 2013.

[22] United Nations Climate Change. The Paris Agreement.
https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-
agreement.

Appendix

Raw Filenames	Description of Dataset	Source
GLB.Ts+dSST.csv	Global mean land-ocean temperature anomaly, 1880-2020	NASA Goddard Institute of Space Studies
GMSL_TPJAOS_4.2_199209_202003.txt	Global mean sea level variations, 1993-2019	NASA Goddard Space Flight Center
CO2-by-source.csv	Annual carbon dioxide emissions by fuel types, 1751-2017	Primary: GCP, CDIAC, Gapminder, UN Population Estimates Secondary: Our World in Data
co-emissions-per-capita.csv	Annual carbon dioxide emissions per person by country, 1751-2017	Primary: GCP, CDIAC, Gapminder and UN Population Estimates Secondary: Our World in Data
annual-co-emissions-by-region.csv	Annual carbon dioxide emissions by country, 1751-2017	Primary: GCP, CDIAC, Gapminder and UN Population Estimates Secondary: Our World in Data
Forest_area_thousands_of_hectares.csv	Global forest area, 2000-2015	UN Statistics Division
country-and-continent-codes-list.csv	List of countries and their continents	Primary: John Snow Labs Secondary: DataHub
API_NY.GDP.MKTP.CD_DS2_en_csv_v2_936013.csv	GDP per capita, 1960-2018	World Bank
API_SP.POP.TOTL.DS2_en_csv_v2_936048.csv	Total population, 1960-2018	UN Population Division
API_SP.DYN.LE00.IN_DS2_en_csv_v2_935946.csv	Life expectancy at birth, 1960-2018	UN Population Division
API_SP.URB.TOTL_DS2_en_csv_v2_936795.csv	Total urban population, 1960-2018	UN Population Division
API_SL.UEM.TOTL.ZS_DS2_en_csv_v2_935940.csv	Unemployment rate, 1991-2019	International Labour Organisation