Battling Climate Change: Implications on the Canadian Climate Policies in Recent Years*

Owen Huang

27 April 2022

Abstract

In this paper, we have fitted a multiple linear regression model to predict the annual national greenhouse gas (GHG) emission in Canada and determined that investments in business, machinery equipment, and exports are the economic factors that are contributing the most to Canada's annual GHG emissions. With the current data presented, we are not able to foretell the future trend of Canada's GHG emissions as we cannot quantify how the economy is going to behave after the pandemic has ended. However, there remains hope as the Government of Canada is actively trying to suppressing the production of GHG, and the innovation of a renewable and green source and energy is imminent. Fulfilling the Paris Agreement (reduce GHG emission levels in 2030 to 30% below the 2005 levels) is the first step toward mitigating the devastating effects of climate change and providing the future generation with a safe world to live in, not a corrupted one.

Contents

*Code and data are available at: https://github.com/Ovven999/Climate_Change

1 Introduction

Our planet is burning. Greenhouse gases are the by-products of energy creation, which humanity will only produce more as we demand more energy for transportation, manufacturing goods, building infrastructures, and much more. Burning fossil fuels such as oil, coal, and natural gas is the primary method for creating energy which makes up approximately 85% of the world's total energy output. The consequence of burning fossil fuels to create energy is the release of greenhouse gases into the atmosphere such as carbon dioxide, sulfur, methane, etc. Greenhouse gases are good at trapping heat within the planet and capturing and retaining heat that radiates from outer space. Thus, a reasonable amount of greenhouse gas emissions will not have a large effect on the Earth's climate, but make the planet more desirable for lives. However, over the past decades of greenhouse gas accumulation, the greenhouse gases that are released into our atmosphere is increasing at an accelerating speed since the industrial revolution. With an increasing amount of greenhouse gases we are pouring into the atmosphere, too much heat is trapped between the Earth's crust and the atmosphere, resulting in the rise of global temperature and it will make the planet warmer and warmer each year we step into the future. Such phenomenon is known as global warming, and its devastating effects are becoming more apparent in recent years. Icecap melting, sea level rising, regional droughts, and the possible sixth mass extinction are all direct and indirectly related to the rising global temperature.

Realizing this issue, 196 countries had come together in 2015 and signed the Paris Agreement to reduce greenhouse gas emissions and fight climate change. Countries need to seek ways to reduce the amount of greenhouse gas emissions, as well as reduce the number of greenhouse gases in the atmosphere before we reach a point of no turning back. The goal of Canada was to reduce GHG emissions by 30% below 2005 levels by 2030. As a result Greenhouse, Gas Pollution Pricing Act (GGPPA) was introduced in 2018 which targets pricing carbon pollution, climate resilience, and clean innovation, technology, and jobs.

This paper sets out to investigate the effectiveness of the actions performed against climate change. We are interested to find out which components of the economical activities are contributing the most to greenhouse gas emissions. In doing so, we can pinpoint the potential sectors of the economy that the policymakers impose policies on to reduce the overall greenhouse gas emissions. We will also briefly discuss government interventions such as carbon pricing, promoting zero-emission vehicles, and energy efficiency retrofits for buildings. We are interested to know, according to the current performance of the government policies, is Canada able to obtain the goal of reducing greenhouse gas emissions by 30% below the 2005 levels.

The most recent data on Canada's greenhouse gas emissions ended in 2019 before the COVID-19 pandemic struck. As a result, we have also taken into account the potential effects of the pandemic in the **Discussion** section, and we determined it would decrease the greenhouse gas emissions in the short run and will have an overall positive effect upon reaching the final target. The Pan-Canadian Framework on Clean Growth and Climate Change took place in 2018, but soon after its settlement, it is interrupted by the pandemic in late 2019. Since we do not have access to the total greenhouse gas emissions in Canada after 2020, we hypothesize and fitted a generalized linear mixed model greenhouse gas emissions against various factors such as various components of GDP, population, and year to figure out which economic activities are the primary driving force of greenhouse gas emissions.

After we obtained the result, we conclude that with the evidence we are presented with, investments in business, machinery and equipment, and exports are the primary economic factors that are contributing the most to Canada's annual GHG emissions. This can give insights to policymakers to weigh the tradeoffs between stimulating investments and limiting GHG emissions. Since the pandemic struck the globe in 2020, the most recent data on Canada's GDP and GHG emission levels in 2021 and 2022 has not been revealed, thus we cannot make meaningful predictions on the future trend of the economy and GHG emissions. However, we suggest to keep implementing the current climate change policies such as carbon pricing, promoting zero-emission vehicles, and energy efficiency retrofits for buildings to limit the growth of GHG emission levels, and shift time and resources to research and development branches for innovations in renewable, replenishable, and green energy sources. This would be a more sustainable solution in the long run.

In **Data** section, we will give detailed information on the data set we will use in this analysis, including their sources and the variables that we are interested in. We will also perform an Exploratory Data Analysis (EDA) to determine the nature of the data set to verify the assumptions that we need to fit an appropriate

model. In **Result** section we will present the resulting figures and models and discuss their implications on the research question. In **Model** section under results, we will explain the statistical method we will use to answer our research question, along with the general implications and application. Lastly, in the **Discussion** section, we will elaborate on our results in real-world settings and explain what the results we obtained from the analysis tell us about the future greenhouse gas emission levels, the economic implications, and the takeaway for the reader of this paper.

2 Data

Throguhout the analysis of this report, we used *R* (R Core Team (2020)), the *tidyverse* (Wickham et al. (2019)) package, the *kableExtra* (Zhu (2021)) package, the *dplyr* (Wickham et al. (2022)) package, the *ggplot*2 (Wickham (2016)) package,the *ggpubr* (Kassambara (2020)) package, the *knitr* (Xie (2021)) package, the *ggrepel* (Slowikowski (2021)) package, the *readr* package, the *readxl* (Wickham and Bryan (2019)) package, and the *lme*4 package for the data analyzing process and graph plotting.

The data we have collected is from 3 major sources. First, we obtained the nation-wise greenhouse gas emissions from the Integrated Carbon Observation System (ICOS). ICOS provides standardized open data about various climate measurements. It is consisted of over 140 measurement stations worldwide and works with numerous scientists and universities around the globe to provide the most accurate data and analysis on our climate (Global Carbon Project (2021)). Below is a table describing variables from data we collected from ICOS after cleaning.

Table 1: GHG Emissions, by country

Variable	Descriptions
Year	Years from 1959 to 2020
Country	Countries containing Australia, Brazil, Canada, USA, China, etc.(total of 12 countries)
GHG Emissions	the greenhouse gas emissions in million tonnes per year

Data on Canada's annual greenhouse gas emission, annual GHG emission by different economic sectors, and annual GHG emission by different provinces are obtained from the Government of Canada website (Government of Canada (2022a)). Below is a table describing the data about GHG emission levels in different Canadian provinces and territories that we collected from the Government of Canada after cleaning.

Table 2: Canada's GHG Emissions, by Province or Territory

Variable	Descriptions
Year	3 specified years: 1990, 2005, 2020
Province or Territory	All the Canadian provinces and territories (ON, BC, etc.)
GHG Emissions	the greenhouse gas emissions in megatons of CO2 equivalent

Below is a table describing the data about GHG emission levels in different Canadian economic sectors that we collected from the Government of Canada after cleaning.

Table 3: Canada's GHG Emissions, by Province or Territory

Variable	Descriptions
Year	Years from 1990 to 2020
Total	The total GHG emissions in the specified year, in Mt CO2 eq
Sector	oil and gas, transports, buildings, electricity, heavy industry, agriculture, waste and others
GHG Emission	the greenhouse gas emissions in megatons of CO2 equivalent, for the sector

Canada's GDP per capita (The World Bank (2022b)) and population (The World Bank (2022a)) are obtained from the World Bank national accounts data and OECD National Accounts data files. Data on Canada's GDP by expenditure is obtained from Statistics Canada (Statistics Canada (2021)). We have organized the data sets into one as follows. Notice that GDP expenditure has many more sub-categories, here we are only listing a few important ones.

Table 4: Canada's GDP, population, and GHG emissions, by year

Variable	Descriptions
Year	Years from 1990 to 2020
GDP per capita	Country's GDP divided by its population
GDP	THe Gross Domestic Product in Canada
Population	The Canada's total population
GHG emissions	The total GHG emissions in the specified year, in Mt CO2 eq
Final Consumption Expenditure	Include consumption by households, institutions, and government
Gross Fixed Capital Formation	Country's total investment
Exports	Exports to other countries
Imports	Imports from other countries

3 Results

3.1 Annual Greenhouse Gas Emission, by Country

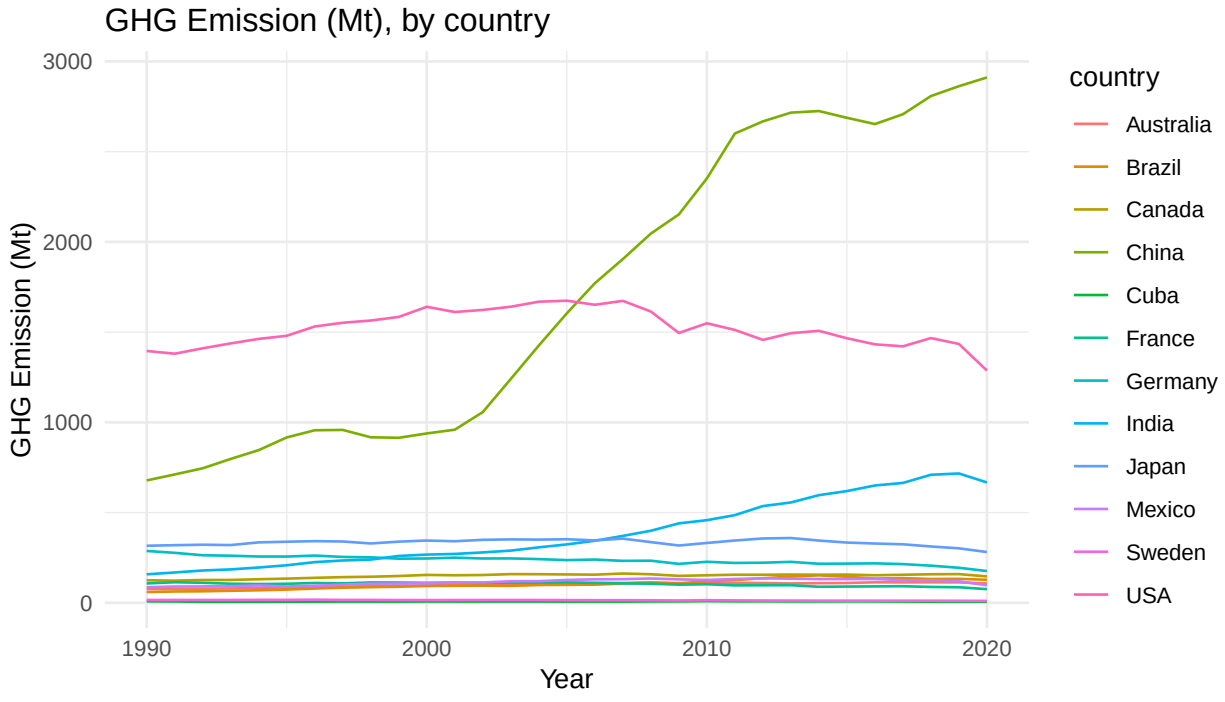

Figure 1: GHG Emission (Mt), by country

We want to get a sense of the greenhouse gas emission situation in a global sense, so it helps to figure out which countries are currently producing the most amount of greenhouse gases. From figure 1, we observed that before 2005, the USA was producing the most of the greenhouse gases in the world. China's greenhouse

gas emission amount drastically increase in the early 2000s and quickly exceeded the US in 2005 and became the country that emits the most greenhouse gases annually. As of 2020, China emits almost twice as much GHG as the US. Moreover, we have observed a trend in decreasing GHG emissions for the US since 2005, and the GHG emission from China shows no sign of slowing down. This is expected as China is one of the largest production countries that are producing goods that are sent off to the rest of the world. The United States, however, although they possess most of the leading commercial and technological companies, is producing less GHG. This is because the US is setting their production plants in other countries like China and India, which not only shifts GHG emissions out of the US, but firms can also take the advantage of cheaper labor and resource costs.

3.2 Canada's Greenhouse Gas Emission and Economic Indicators

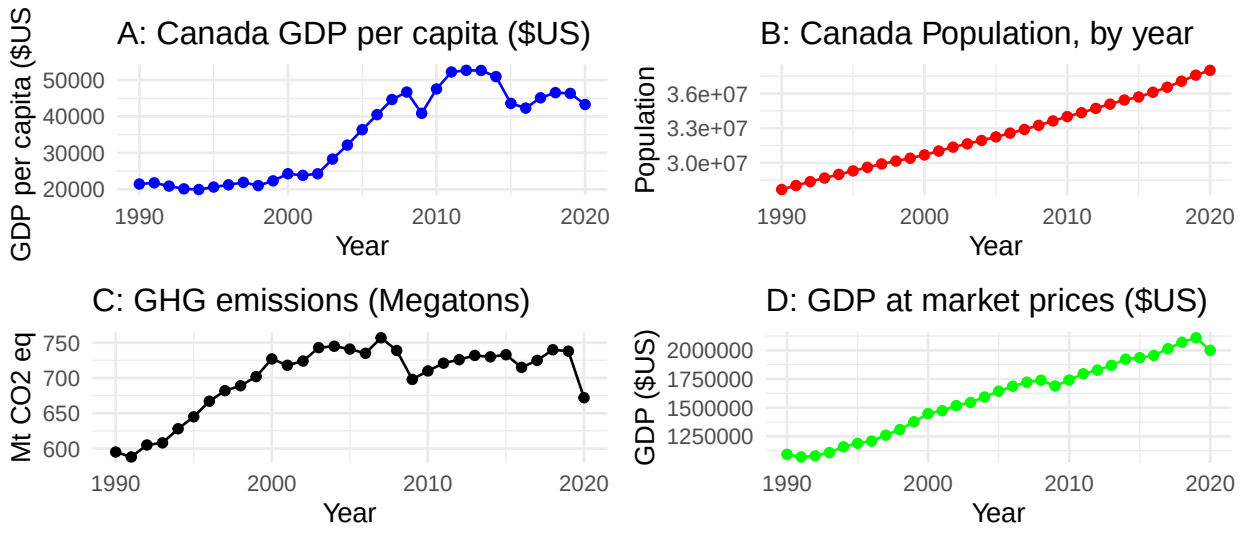

Figure 2: Canada's GHG emissions, GDP, and population

This is an overview of the current situation of Canada's greenhouse gas emissions in relation to its population and GDP. From figure 2D we can see that the Canadian economy has been growing consistently in the past few years. However, it is hard to say that it is going to continue to grow in the upcoming years. We can see that there is a significant decrease in GDP from 2019 to 2020, and this is the result of the COVID-19 pandemic in which many businesses were forced to shut down due to the quarantine policy imposed to combat the virus. As the situation continued to escalate throughout 2020 and 2021, although we do not have access to specific numbers, we are expecting the GDP to continue to decrease, and thus production will decrease further. This recession may not be good for the economy from an economic point of view, but it surely is contributing to mitigating climate change in the form of reducing greenhouse gas emissions.

From Figures 2A and 2C, we can see that Canada's greenhouse gas emissions follow largely a similar pattern with the GDP per capita. When the economy is doing well, there will be lots of production occurring to match up with the large demand on the market, and the increasing production will lead to production plants and business buildings emitting more greenhouse gases. Also, when the economy is doing well, there will be a lot more demand for transportation and that will increase the number of fossil emissions to the atmosphere. A similar argument can be made for when the economy is in a recession.

Figure 2B presents Canada's population. Just as the population of a country is tightly connected to its economy, it also has a significant effect on the greenhouse gas emissions of a country. More people means more demand and consumption for goods and services. As the population of Canada continues to grow, we would expect fresh blood to be injected into the country, which would help with the long-term growth of the economy. But as we have mentioned before, there will be a trade-off between the well-being of the economy

and the greenhouse gas emissions in the country. We need to seek a way that promotes economical growth, all while preserving or reducing greenhouse gas emissions.

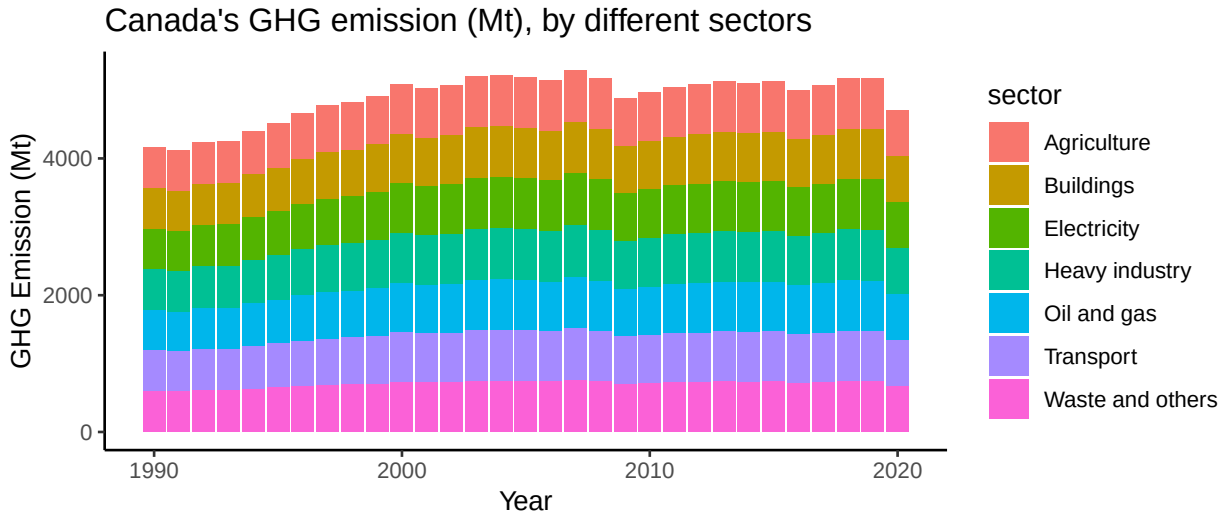

Figure 3: Canada's GHG emission (Mt CO2 eq), by different sectors

Understanding which economic sectors are producing the most amount of greenhouse gases is crucial for policymakers when they are taking action against climate change and mitigating greenhouse gas emissions. From 1990 to 2020, we have observed the largest increase in GHG emissions from the oil and gas sector, which increase from 102.6 Mt CO_2 eq 1990 to 178.8 Mt CO_2 eq in 2020. The second-largest increase was from the transportation sector which increased from 120.5 Mt CO_2 eq in 1990, to 159.2 Mt CO_2 eq in 2020. These increases accompany the growth of the economy and are for the most part inevitable.

However, there is an offset effect in the electricity and heavy industry sectors. Greenhouse gas emissions in the electricity sector decreased from 94.7 Mt CO_2 eq in 2019 to 56.2 Mt CO_2 eq in 2020, while emissions in heavy industry decreased from 97.2 Mt CO_2 eq to 71.8 Mt CO_2 eq. Comparing the emission level of 2020 to 2005, we can see that of the overall 69 Mt CO_2 eq decrease in GHG emission, most of it comes from the electricity sector which accounts for 61 Mt CO_2 eq GHG emission reduction.

This figure shows that Alberta is producing the most amount of greenhouse gases, followed by Ontario, Quebec, and Saskatchewan. Emissions in most provinces are below the 2005 emission level except for Alberta. In 1990, Ontario was emitting the most amount of GHG due to its large-scale manufacturing industries. However, Alberta quickly surpassed Ontario to be the province that emits the most GHG due to its uprising activities in the oil and gas industry. According to the Government of Canada, the decrease in GHG emissions in Ontario was driven by the shutdown of coal-fired electricity generation plants. Quebec's decrease was the result of the decrease in the residential and heavy metal industry. Saskatchewan's decrease was due to the emission reduction in the oil and gas industry (Government of Canada (2022b)). The Government of Canada also provided an insight into the greenhouse gas emission levels during the first year of the pandemic and stated that all provinces and territories experienced a 5% to 16% decrease.

3.3 Model

We are going to fit a multiple linear regression model to further analyze the topic of interest: what components of economical activities are the factors that are driving greenhouse gas emissions.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_i + \epsilon$$

Where Y is the response variable we are interested in predicting, X_i is the predictors that are used to predict the response variable Y. β_0 is the y-intercept, if there are categorical variables as the predictor,

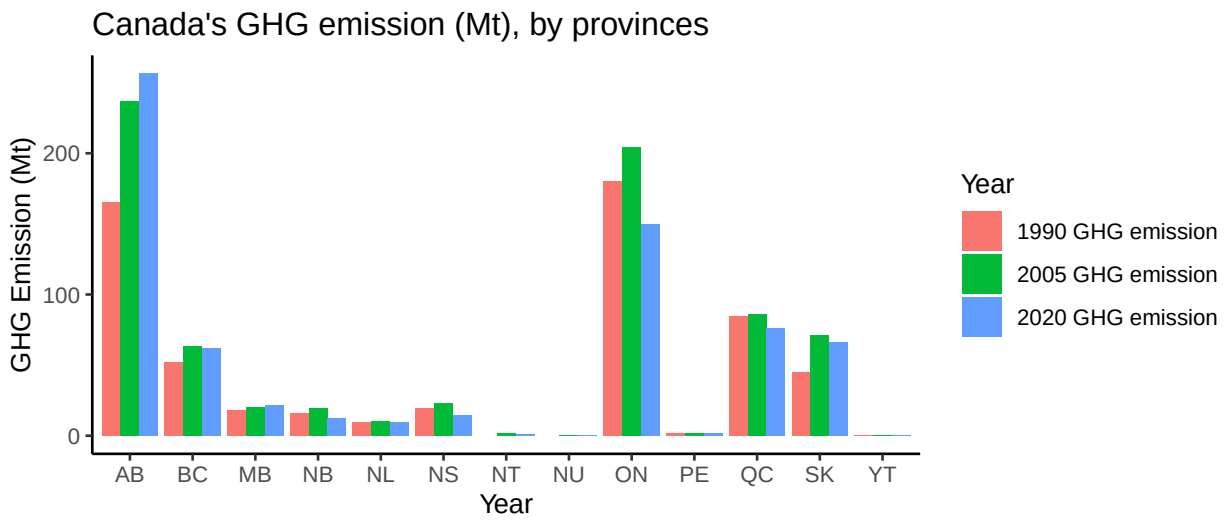

Figure 4: Canada's GHG emission (Mt CO2 eq), by provinces

then β_0 represent the value for a particular reference group, otherwise, β_0 represents the y-intercept where all predictors X_i are equal to 0. For the first model, we fit Canada's annual greenhouse gas emissions as the response variable and then fit the 4 major GDP categories as the predictor, including final consumer expenditure, gross fixed capital formation (GFCF or investment), investments in inventories, and exports to other counties and we included a year as a reference to time. Here we are excluding imports from other counties because these goods and services were produced outside of Canada, so it does not possess a significant portion of the total greenhouse gas emissions.

The summary of the linear model shows that GFCF (investment) and exports have strong statistical significance with a p-value <0.001. P-value is a test statistic that represents the probability of this data set occurring while assuming the correctness of the null hypothesis. For the linear regression model, the null hypothesis is that $\beta_i = 0$. Since we have obtained a very small p-value for all predictors, we conclude that we have strong evidence against the null hypothesis that $\beta_i = 0$ and there is a relationship between the predictor variables and the response variable. This makes sense because the money invested by companies directed impacts the goods and services they produce, without technological innovation, the increasing output will likely increase the greenhouse gas emission from buildings and factories. Exports are contributing positively to greenhouse gas emissions due to a similar reason.

Furthermore, we understand that not all investments go into producing products and greenhouse gases, so we fitted a new model with GFCF split into sub-categories of investment such as business GFCF, residential structures, non-residential structure, machinery and equipment, intellectual property products, non-profit institutions serving households' GFCF, general governments GFCF, along with year and exports. Then we run bi-directional step-wise AIC on the model to reduce the model to obtain the final model below. AIC stands for Akaike Information Criterion, which is a mathematical expression that describes how well the model fits a particular data set and is used to compare the goodness of fit between different models. Bi-directional step-wise AIC begins with the full model with all the possible predictors, then it is going to remove or add a predictor recursively until we have obtained the model with the smallest AIC value.

$$Total\ GHG\ emissions = 11390.74 - 5.51(Year) + 8.26 \times 10^{-4}(GFCF) + 9.26 \times 10^{-4}(Business\ GFCF) +$$

$$1.38 \times 10^{-3}(Machinery\ and\ equipment) + 5.20 \times 10^{-4}(Exports)$$

Here we will examine whether or not the final model violates any of the assumptions of a linear regression model. A violation would limit our ability to interpret the model. In the Y versus Y-hat plot, Y represents

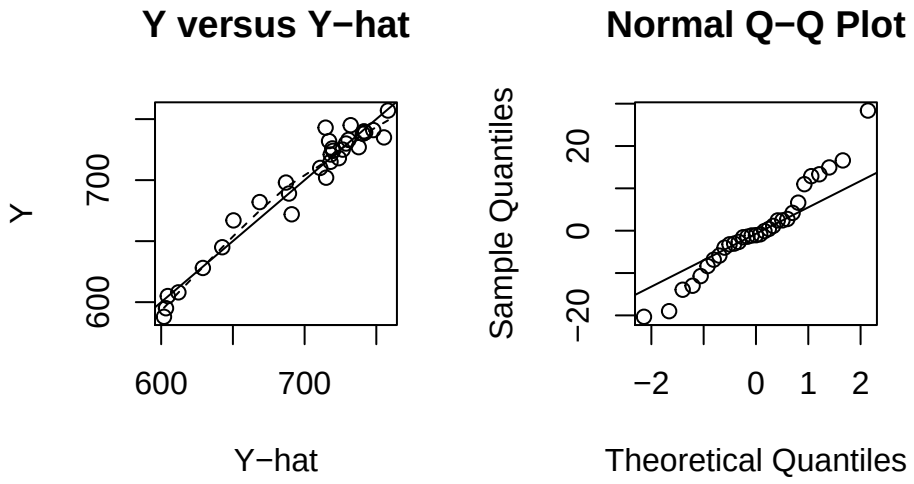

Figure 5: Checking for multiple linear regression assumptions: Y vs Y_hat and QQ plot

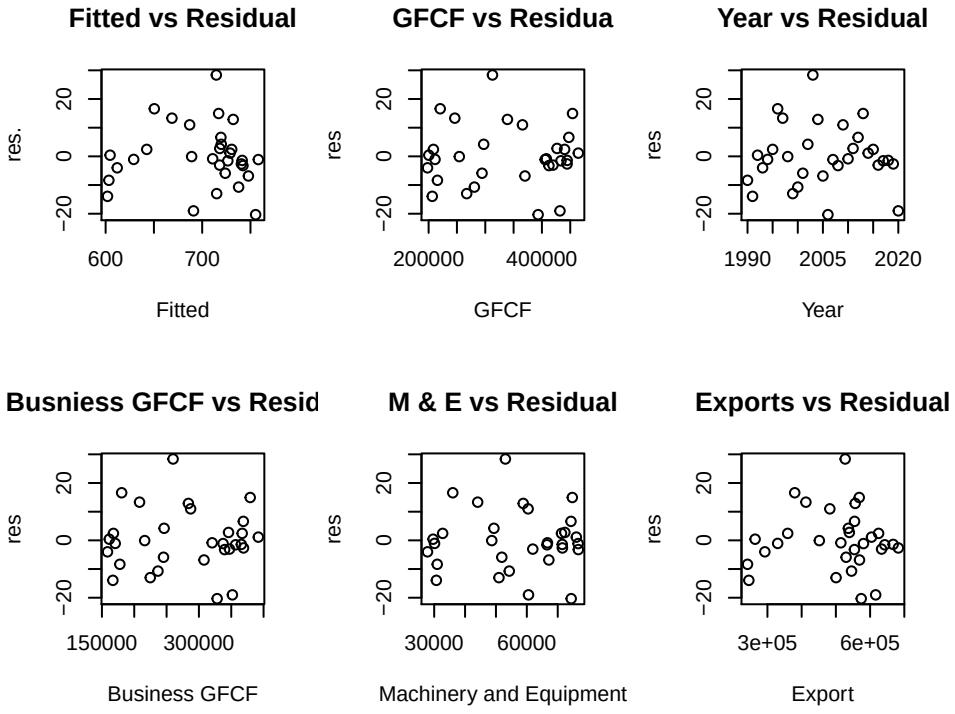

Figure 6: Check for the residual plots of the multiple linear regression model

the actual values and Y-hat represents the predicted value. We can see that when plotting them against each other, they follow largely a straight line at the 45-degree angle, suggesting that the relationship is indeed linear. The Normal Q-Q plot is used to verify the normality assumption. We see that the data set deviates towards the tails of the distribution suggesting that it is not a perfectly normal distribution.

We do not observe any patterns in the residual plot and that all points are randomly scattered around the space. This verifies that the model and data are linear, have constant variance, and are uncorrelated. Thus, we conclude that the model is a valid model that satisfies the assumptions of the multiple linear regression model. The final model is as follows:

$$Total\ GHG\ emissions = 11390.74 - 5.51(Year) + 8.26 \times 10^{-4}(GFCF) + 9.26 \times 10^{-4}(Business\ GFCF) +$$

$$1.38 \times 10^{-3}(Machinery\ and\ equipment) + 5.20 \times 10^{-4}(Exports)$$

β_0 means that if all other predictors are 0, the total GHG emissions in Canada for the year 0 would be 11390.74 megatons, which is meaningless because a year of 0 does not make sense. $\beta_1 = -5.51$ shows that there is a negative relationship between year and GHG emissions, which means that if we keep all else fixed and increase year by 1, we would expect the total greenhouse gas emission for that year to decrease by 5.51 megatons. $\beta_2 = 8.26 \times 10^{-4}$ means that if we keep all else fixed and increase GFCF (investment) by 1, we would expect the total greenhouse gas emission for that year to increase by 8.26×10^{-4} megatons. Similarly, 1 unit increase in Business GFCF would increase total GHG emissions by 9.24×10^{-4} megatons, 1 unit increase in machinery and equipment would increase GHG emission by 1.38×10^{-3} megatons, and lastly, 1 unit increase in exports would increase emission levels by 5.30×10^{-4} megatons.

4 Discussion

4.1 Canada'g GHG Emission Levels and the Economy

As settled in the Paris Agreement, the target emission by the year 2030 will be 30% less than the emission level of 2005. The greenhouse gas emission level for 2005 settles ta 741 Megatons of CO_2 equivalent (Mt CO_2 eq). The emission level for 2019 is at 738 MT CO_2 eq, while the emission level for 2020 is down to 672 Mt CO_2 eq. However, the significant decrease in GHG emissions between 2019 to 2020 is unlikely to sustain when the pandemic is over and the economy is rebooted. We would expect a kickback in greenhouse gas emission levels that exceed the 2005 levels, as the country strives for recovering from the pandemic.

As of 2020, the GHG emission level is 9.3% lower than that of the 2005 levels, which is still far away from the target of 30%. From figure 3, we discovered that the driving force of the reduction of GHG emissions was in the electricity and heavy industry sector of the economy. The major reduction that occurred from 2019 to 2020 was the introduction of the COVID-19 pandemic, which slowed down the industry and the economy, and posted a significant reduction in trade and mobility such as travel by air and land. These targets contributed significantly to the reduction of GHG emissions as its effects are visible in the transportation sector, where a 14% decrease (185.5 Mt CO_2 eq in 2019 to 159.2 Mt CO_2 eq in 2020) was observed. Needless to say, the amount of greenhouse gas emissions is positively related to the well-being of the economy. When the economy is performing well, we will have an increasing demand for goods and services, which would drive up production quantity to meet the demand. The manufacturing plants, office buildings, and power plants all rely on energy inputs such as electricity and natural gas. The increased amount of economic activities would increase the mobility of its participants, resulting in more people traveling around both by air and land. While burning fossil fuel remains to be one of the primary sources of energy, the increased demand during the economic boom will most definitely increase the amount of GHG emissions. Does that mean we have to give up the growth of our economy to reduce the amount of greenhouse gas emissions? We did make a promise on the Paris Agreement after all.

Due to the effect of the pandemic, the GHG emission levels are 9.3% below the 2005 emission levels. We have seen a significant decrease in the amount of GHG emissions in the electricity and heavy industry sector, and we are seeing little to no decreases in other sectors since 2005, such as oil and gas, transportation, buildings,

agriculture, and waste and others. The most direct solution to reducing GHG emissions in other sectors is to promote green energy. The Government of Canada has plans to promote the zero-emission vehicles as part of the actions taken against climate change, and already has plans to increase the electric-vehicle rebates. We can expect the GHG emission levels to continue to go down as the pandemic continues in 2021 and onward, but this does not help with sustainable economic growth. The oil and gas sector and transportation sector will continue to be two of the major economic sectors that contribute the most to the GHG emissions, thus, it is crucial to find sustainable and green energy to slow down the growth of the GHG emissions in these two sectors.

4.2 Suggestions for policymakers

From the multiple linear regression model fitted from the previous part, we observed that the four major components of GDP expenditure: consumption, gross fixed capital formation (GFCF), investment in inventories, and exports (imports are excluded because they are not produced in Canada), GFCF and exports have strong statistical significance when fitted against total greenhouse gas emissions in Canada. And finally, we have determined that money spends on GFCF, business GFCF, machinery and equipment, and exports are the most significant contributing factors to GHG emissions over the years. As we have mentioned before GHG emissions and GDP are positively related, so an increase in spending on any of the components mentioned above would result in an increase in GHG emissions. Investments in business and product inventories would increase the energy consumed when operating business buildings and manufacturing plants. And money spent on purchasing more machinery and equipment would also increase the energy consumed. The increase in the number of exports would also mean that we are producing more goods and services domestically. These are all the reasons that each factor is positively affecting GHG emission levels individually. We want to try to avoid giving up too many of these market growth opportunities while decreasing GHG emission levels.

Investments are crucial for economic growth as they can directly influence interests rate, and carelessly imposing monetary policies on the market could result in unexpected behaviors of investors and backfires on the economy and the goal of reducing greenhouse gas emissions. The carbon tax introduced in recent years is a good way to limit GHG emissions to a specified amount, but it could also prevent the potential growth of firms, markets, and the economy. The carbon tax was introduced in 2019 at $20 per ton of carbon dioxide produced, and the Canadian government planned to raise that tax amount by $10 each year. Thus, instead of trying to decrease investments to decrease GHG emissions, we should be spending more money on the research and development sectors of the economy. Increase policies that would benefit innovative firms that are actively trying to reduce their GHG outputs or actively switching to greener energy sources. AS we cannot slow down the growth of the economy too much in the effort of trying to reach the goal specified in the Paris Agreement, policymakers should focus on the subsidizing the R&D sectors, especially researchers, firms, universities, and laboratories that are focusing on the development of renewable energy and products. Accelerating the process by which these world-changing inventions come to life is better than limiting GHG emissions in the short run, all while holding back the growth of the economy.

4.3 The world

Canada meeting its target is not enough to slow down the current state of global climate change, but it certainly would help. What is important is to find a way to slow down the rate at which China and the US are emitting GHG. The two of them are emitting more GHG each year than the rest of the world combined. But too sudden of an effort in trying to decrease GHG emissions would result in economic malfunctions within their countries and would radiate to the rest of the world, much like how countries are behaving during the COVID-19 pandemic. Similar models and arguments can also be applied to other countries where we can post policies such as carbon tax and the cap and trade system to limit GHG emissions in the short run to buy them time to develop new and greener sources of energy.

Moreover, the COVID pandemic's contribution to GHG emissions does not limit to less production and traveling, but also the changing lifestyle of individuals and operating strategies of business give us insight into how the future workplace would look like and how it helps to reduce GHG emissions from transportation and business buildings. As the pandemic continued for the past 2 years, individuals and firms have adapted

to remote working environments where people can work from home and business buildings are operating at a minimal cost. This not only reduces GHG emissions from mandatory transportation as people routinely travel from home to workplace and then back home but also reduces the GHG emissions emitted by operating business buildings. Working remotely not only saves the time and resources of traveling but also allows for more flexible schedules that contribute to the well-being of the individuals. When the economy worldwide is finally back on track in the future once the pandemic is over, a policymaker can try to promote these types of remote working environments among traffic-heavy cities, and this would help with reducing GHG emissions by a lot while buying time for a new energy source to be introduced.

4.4 Weaknesses and next steps

Also, all data sets are obtained from credible authorities, but they still may post inaccuracies such as underestimating the total greenhouse gas emission levels because GHG emissions are hard to measure and quantify. There are more ways to combat climate change and reduce GHG emissions that are not being mentioned in this report, methods may include planting trees, reducing deforestation, and desertification. Moreover, the amount of GHG emissions reported in 2020 coincide with the first year of the COVID-19 pandemic, which had a significant impact on various aspects of the economy, especially on the transportation and energy sectors. Thus, estimates including 2020 emission levels need to be interpreted with caution as they may not reflect the true trend of GHG emissions. AS the pandemic incoming to an end hopefully in the near future, it is unknown how the reopening is going to affect the economy and the greenhouse gas emission level. Thus, the model may not accurately predict future behavior based solely on the data from 1990 to 2020.

For the next steps, we would like to have access to the more recent data on GDP and total greenhouse gas emission levels during the pandemic lockdown. That way we will be able to provide more meaningful insight into how the market and GHG emission level is going to behave after the pandemic has passed.

Appendix

A Datasheet (Gebru et al. 2021)

Motivation

1. *For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.*
   - The data for GHG emissions for different countries is collected from the Integrated Carbon Observation System (ICOS). ICOS provides standardized open data about various climate measurements. It is consisted of over 140 measurement stations worldwide, and working with numerous scientist and universities around the globe to provide the most accurate data and analysis on our climate. Canada's population and GDP per capita is obtained from the world bank. World bank targets to help regions and countries to sustainable growth and strives to end poverty. Statistic Canada provide open data for analysis and educational uses.
2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
   - Datasets from Integrated Carbon Observation System (ICOS) were created by a group of scientist such as Pierre Friedlingstein1,2, Matthew W. Jones3, Michael O'Sullivan, etc.. The World Bank, which is consisted of 5 institutions: The International Bank for Reconstruction and Development, The International Development Association, The International Finance Corporation, The Multilateral Investment Guarantee Agency, and The International Centre for Settlement of Investment Disputes. Statistic Canada, which is the national statistical office.
3. *Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.*
   - Statistic Canada's dataset were funded by the Government of Canada
4. *Any other comments?*
   - N/A

Composition

1. *What do the instances that comprise the dataset represent (for example, documents, photos, people, countries)? Are there multiple types of instances (for example, movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.*
   - The data sets are comprised of different countries and the Canada's national data. Canada's nation data is comprised of GDP expenditure components, and Canada's provinces and territories
2. *How many instances are there in total (of each type, if appropriate)?*
   - GDP expenditure components: 27
   - province and territories: 13
   - countries: 12
   - Total: 52
3. *Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (for example, geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (for example, to cover a more diverse range of instances, because instances were withheld or unavailable).*
   - The data set contains all the possible instances
4. *What data does each instance consist of? "Raw" data (for example, unprocessed text or images) or features? In either case, please provide a description.*
   - Each instances contains 1 categorical variable that provides the description of the instance. GDP expenditure components are followed by 2 continuous variables representing the year and total expenses. province and territories and countires are followed by 2 continuous variables representing the year and GHG emission levels.
5. *Is there a label or target associated with each instance? If so, please provide a description.*
   - No

6. *Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (for example, because it was unavailable). This does not include intentionally removed information, but might include, for example, redacted text.*
   - No
7. *Are relationships between individual instances made explicit (for example, users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.*
   - NO, there is no relationship
8. *Are there recommended data splits (for example, training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.*
   - No there is not
9. *Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.*
   - There are no errors, sources of noise, or redundancies in the dataset
10. *Is the dataset self-contained, or does it link to or otherwise rely on external resources (for example, websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (that is, including the external resources as they existed at the time the dataset was created); c) are there any restrictions (for example, licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.*
    - It is self-contained
11. *Does the dataset contain data that might be considered confidential (for example, data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.*
    - All dataset used in this report are open data and not confidential
12. *Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.*
    - There are no sensitive data
13. *Does the dataset identify any sub-populations (for example, by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset.*
    - No it does not
14. *Is it possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset? If so, please describe how.*
    - No
15. *Does the dataset contain data that might be considered sensitive in any way (for example, data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? If so, please provide a description.*
    - No
16. *Any other comments?*
    - N/A

Collection process

1. *How was the data associated with each instance acquired? Was the data directly observable (for example, raw text, movie ratings), reported by subjects (for example, survey responses), or indirectly inferred/derived from other data (for example, part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.*
   - Data from Integrated Carbon Observation System (ICOS) were directly collected from measurement stations worldwide. Data from Government of Canada and Statistic Canada are directly collected and calculated by government statistical agencies.
2. *What mechanisms or procedures were used to collect the data (for example, hardware apparatuses or sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?*

- Data are collected from measurement stations, and Canadian government statistical and economical agencies.

3. *If the dataset is a sample from a larger set, what was the sampling strategy (for example, deterministic, probabilistic with specific sampling probabilities)?*
   - N/A
4. *Who was involved in the data collection process (for example, students, crowdworkers, contractors) and how were they compensated (for example, how much were crowdworkers paid)?*
   - Scientists worldwide are involved in the creation of ICOS datasets. Compensation for which were not disclosed.
5. *Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (for example, recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.*
   - ICOS datasets are updated each year for the purpose of creating reports such as the Global Carbon Budget 2021 report. These datasets were last modified in 2022. Canada's GDP expenditure data were last modifed in 2021, and greenhouse gas emission data was last modifed in 2022.
6. *Were any ethical review processes conducted (for example, by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.*
   - There are no mention of ethical review process being conducted
7. *Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?*
   - Data were collected from ICOS https://essd.copernicus.org/articles/14/1917/2022/, Government of Canada https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/greenhouse-gas-emissions.html#agriculture, and Statistic Canada https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610022201.
8. *Were the individuals in question notified about the data collection? If so, please describe (or show with screenshots or other information) how notice was provided, and provide a link or other access point to, or otherwise reproduce, the exact language of the notification itself.*
   - N/A
9. *Did the individuals in question consent to the collection and use of their data? If so, please describe (or show with screenshots or other information) how consent was requested and provided, and provide a link or other access point to, or otherwise reproduce, the exact language to which the individuals consented.*
   - N/A
10. *If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? If so, please provide a description, as well as a link or other access point to the mechanism (if appropriate).*
    - N/A
11. *Has an analysis of the potential impact of the dataset and its use on data subjects (for example, a data protection impact analysis) been conducted? If so, please provide a description of this analysis, including the outcomes, as well as a link or other access point to any supporting documentation.*
    - The analysis of potentoal impact of the dataset and its use on data subjects were not conducted
12. *Any other comments?*
    - N?A

Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.*
   - GHG emission levels were converted into megatons of CO2 equivalent. Which is calculated by the GHG's activity data multiply by its emission factor. For example, methane's ability to trap heat inside the atmosphere is 25 times that of carbon dioxide, so methane's emission factor would be 25.
2. *Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (for example, to support unanticipated future uses)? If so, please provide a link or other access point to the "raw" data.*
   - the raw data are saved in input/data/GDP_per_cap.xls, inputs/data/Global_Carbon_Budget_2021v1.0.xlsx,

inputs/data/National_Carbon_Emissions_2021v1.0.xlsx, /inputs/data/pop.xls, inputs/data/ghg-emissions-regional-en.csv, inputs/data/ghg-emissions-sector-en.csv.

3. *Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.*
   - R studio: https://www.r-project.org/
4. *Any other comments?*
   - N/A

Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
   - Data from ICOS was used to create the report on Global Carbon Budget 2021. Data from the government of Canada is used to create the report on Environment and Climate Change Canada (2022) Canadian Environmental Sustainability Indicators: Greenhouse gas emissions.
2. *Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.*
   - https://github.com/Ovven999/Climate_Change
3. *What (other) tasks could the dataset be used for?*
   - Evaluating the current state of climate change for different countries.
4. *Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (for example, stereotyping, quality of service issues) or other risks or harms (for example, legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?*
   - The cleaned data set should not be interpreted outside the context of this report. Data set cannot be used for commercial uses without the consent of the original creator.
5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*
   - Data set cannot be used for commercial uses without the consent of the original creator.
6. *Any other comments?*
   - N/A

Distribution

1. *Will the dataset be distributed to third parties outside of the entity (for example, company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.*
   - The dataset will not be distributed to third parties outside of the entity on behalf of which the dataset was created.
2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*
   - The dataset will be distributed through gitbub
3. *When will the dataset be distributed?*
   - The dataset will be distributed on April 27, 2022
4. *Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/ or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.*
   - N/A
5. *Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.*
   - No thrid party have imposed IP-based or other restrictions on the data associated with the instances.
6. *Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.*

- No

7. *Any other comments?*
   - N/A

Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
   - Owen Huang
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*
   - o.huang@mail.utoronto.ca
3. *Is there an erratum? If so, please provide a link or other access point.*
   - No
4. *Will the dataset be updated (for example, to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (for example, mailing list, GitHub)?*
   - The dataset will not be updated
5. *If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (for example, were the individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.*
   - The dataset does not relate to people
6. *Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.*
   - No it will not
7. *If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.*
   - The repository can be clone from the github link https://github.com/Ovven999/Climate_Change. The contributions will not be validated/verified, since I am not responsible for any extended/augmented/built on/contributed datasets other than my own.
8. *Any other comments?*
   - N/A

B GHG emission levels for different economic sectors in Canada

Table 5: GHG emission levels for different economic sectors in Canada, in Mt CO2 eq

Year	Oil and gas	Transport	Buildings	Electricity	Heavy industry	Agriculture	Waste and others
1990	102.6	120.5	71.2	94.7	97.2	51.7	56.7
1991	102.2	114.3	70.6	96.1	97.1	52.3	55.3
1992	110.8	115.2	72.4	102.5	94.6	54.7	54.9
1993	117.5	116.7	76.0	93.1	94.1	56.5	53.7
1994	121.5	121.3	76.3	95.1	99.7	59.1	55.1
1995	127.6	122.1	77.0	98.2	100.4	62.2	57.9
1996	135.4	125.7	83.3	98.2	103.2	63.4	57.5
1997	136.5	131.5	80.8	109.4	102.6	64.0	57.5
1998	140.9	137.3	72.4	122.1	97.6	63.8	54.5
1999	149.9	143.0	76.3	119.1	94.7	63.5	55.2
2000	155.1	144.9	82.8	129.0	94.1	64.1	56.9
2001	156.6	146.6	79.5	129.2	88.3	62.9	55.2
2002	161.5	147.7	83.8	123.6	89.0	62.9	55.7
2003	165.9	151.7	89.0	127.3	88.3	65.1	55.9
2004	168.2	156.1	87.6	119.0	92.2	65.9	56.2
2005	171.3	160.1	83.7	117.5	87.2	66.4	55.0
2006	178.0	161.0	78.5	111.6	87.0	65.0	53.5
2007	183.0	164.4	84.3	119.6	86.0	65.0	54.5
2008	179.9	164.5	84.2	108.8	84.5	64.2	52.9
2009	176.9	161.5	82.6	93.7	71.5	61.9	49.5
2010	181.4	167.2	79.4	94.6	74.6	61.9	50.5
2011	187.2	168.3	84.7	86.6	80.4	62.4	51.3
2012	194.2	170.5	83.2	83.3	80.2	63.7	50.5
2013	198.9	174.0	84.3	79.7	78.6	65.4	51.3
2014	204.5	171.5	84.9	76.3	79.4	63.8	49.3
2015	204.8	172.1	83.8	79.7	77.8	64.6	49.8
2016	194.4	173.5	82.2	74.3	76.3	64.9	49.5
2017	196.5	178.9	86.6	72.6	75.5	64.3	50.5
2018	205.0	184.1	92.9	62.8	77.5	66.3	51.5
2019	203.5	185.5	92.0	61.8	77.4	66.7	51.5
2020	178.8	159.2	87.8	56.2	71.8	68.7	50.0

C Canada's GHG emission levels by province and territories

Table 6: Canada's GHG emission levels by province or territory, in Mt CO2 eq

Province or Territory	1990 GHG Emissions	2005 GHG Emissions	2020 GHG Emissions
Newfoundland and Labrador (NL)	9.6	10.5	9.5
Prince Edward Island (PE)	1.8	1.9	1.6
Nova Scotia (NS)	19.5	23	14.6
New Brunswick (NB)	16.2	19.8	12.4
Quebec (QC)	84.5	86.3	76.2
Ontario (ON)	180	204.4	149.6
Manitoba (MB)	18.3	20.5	21.7
Saskatchewan (SK)	45.1	71.3	65.9
Alberta (AB)	165.6	237.1	256.5
British Columbia (BC)	51.7	63.6	61.7
Yukon (YT)	0.6	0.6	0.6
Northwest Territories (NT)	1.8[A]	1.7	1.4
Nunavut (NU)[A]	n/a	0.6	0.6

References

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92.

Global Carbon Project. 2021. "Supplemental Data of Global Carbon Project 2021." Global Carbon Project. https://doi.org/10.18160/GCP-2021.

Government of Canada. 2022a. "Environment and Climate Change Canada (2022) Canadian Environmental Sustainability Indicators: Greenhouse Gas Emissions." https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/greenhouse-gas-emissions.html.

———. 2022b. "Environment and Climate Change Canada (2022) Canadian Environmental Sustainability Indicators: Greenhouse Gas Emissions." https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/greenhouse-gas-emissions.html.

Kassambara, Alboukadel. 2020. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. https://CRAN.R-project.org/package=ggpubr.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Slowikowski, Kamil. 2021. *Ggrepel: Automatically Position Non-Overlapping Text Labels with 'Ggplot2'*. https://github.com/slowkow/ggrepel.

Statistics Canada. 2021. "Table 36-10-0222-01 Gross Domestic Product, Expenditure-Based, Provincial and Territorial, Annual (X 1,000,000)." https://doi.org/https://doi.org/10.25318/3610022201-eng.

The World Bank. 2022a. "Population, Total - Canada."

———. 2022b.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Jennifer Bryan. 2019. *Readxl: Read Excel Files*. https://CRAN.R-project.org/package=readxl.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. https://yihui.org/knitr/.

Zhu, Hao. 2021. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.