CSC110 Project Report: Effects of Greenhouse Gas Emissions in Canada

Dana Alshekerchi, Nehchal Kalsi, Rachel Kim, Kathy Lee

Monday, December 14, 2020

Introduction: Problem Description and Research Question

Greenhouse gas (GHG) emissions are a significant driver of climate change. Ultimately, it is the greenhouse effect that is causing global warming on Earth: "scientists attribute the global warming trend observed since the mid-20th century to a human expansion of the 'greenhouse effect'" [1]. This effect can arise both from humans directly emitting GHGs into the atmosphere and from indirect human activities that hinder the Earth's natural ability to remove GHGs from the atmosphere, thus exacerbating the issue (an example of the latter is deforestation). Our project will focus on the former.

It is also important to note that the negative repercussions of climate change are distributed in an unequal manner, both globally and within Canada. This inequality is caused both by geographical factors and the infrastructure already in place to adequately deal with the changing climate; for our project, we plan on focusing on the former. Notably, in Canada, Northern areas are affected by melting ice and rising temperatures, coastal areas (such as British Columbia and the Maritimes) are affected by extreme rainfall and floods, the prairie regions are affected by increasing droughts due to rising temperatures, and Alberta is affected through both the increasing number and severity of wildfires [2]. Many of these effects are caused, at least in part, by rising temperatures (both average temperatures and extreme temperatures).

Through our project, we aim to better understand the relationship between GHG emissions and their effects on Canada and how certain regions are more acutely affected than others. We also aim to understand how dire the effects of climate change could be in our future. Thus, our research question is: **How have GHG emissions affected temperatures in various regions across Canada?**

Dataset Description

For our project, we used two main data sources to produce the results, as well as two data sources to create the maps:

Our first dataset, Canada's Official Greenhouse Gas Inventory, is published by Environment and Climate Change Canada. [3] Formatted as a csv file, this set consists of greenhouse gas emissions data across all Canadian provinces and territories, spanning from 1990 to 2018. The specific columns that were used were the years (column 1), the province names (column 2), and the CO2 equivalent of the GHG emissions (second last column). Only the rows that recorded the total CO2 emissions for that province or territory (as opposed to a certain sector in the industry) were used.

The second dataset, Adjusted and Homogenized Canadian Climate Data – Daily Temperature and Precipitation (AHCCD – daily TP), is published by Environment and Climate Change Canada. [4] The data of daily mean temperatures are available as a ZIP file containing 780 text files, each of which represents a climate-reporting station in Canada. In each `txt` file, the average daily temperature of the station is recorded. The first two rows provide station information, and the following rows have with 33 columns, displaying the year, month, and days (from 1 to 31). While the start year varies per file, the recordings all span to 2019. For our program, we extracted the station name and province/territory name, years, and average daily temperatures for use.

There were two additional datasets used for the maps. The first was a file containing the coordinates for all the borders for each of the provinces and territories in Canada in a json file, in a specific format called a geojson. [5] The files were created by Kan Nishida. The second was also a geojson file, called Adjusted and Homogenized Canadian

Climate Data (AHCCD) Stations, containing the coordinates for all of climate stations in Canada. [6] The source of this file is the Government of Canada.

Computational Overview

The greenhouse gases dataset provides values for different greenhouse gases such as CO2 (carbon dioxide), CH4 (methane), N2O (nitrous oxide), HFCs (hydrofluorocarbons) and such throughout Canada, from 1990 to 2018, with the exception of Nunavut which was founded in 1999. We chose to work with the CO2 equivalent of the GHG emissions.

The program reports our results in a visual manner in two different ways. First, it creates a graph for the selected station, displaying both the CO2 equivalent levels of GHGs and change in average temperature from year to year. This allows the user to view the correlation between the two variables. The program also creates three maps: one that shows the raw CO2 equivalent output from each province in a given year, one that shows the difference in CO2 equivalent output from each province for a given year compared to 1990, and one that shows the difference in daily mean temperatures for each weather station for a given year compared to 1990.

Data Cleaning and Reading

There were four main files to be processed: the Greenhouse Gas Emissions file, the AHCCD – daily TP (daily mean temperatures) files, the Canada province and territory borders geojson file, and the Canada climate stations coordinates geojson file.

The Greenhouse Gas Emissions file was processed for two different applications: one for the graphs (in `read_ghg_emissions` in `data_reading.py`) and one for the maps (in `read_ghg_emissions_data_reading.py`). The end result for the processing for the graphs is a nested list, with the inner list containing the year (`int`), province (`str`), and CO2 equivalent of greenhouse gas emissions (`float`) in that order. The end result for the processing for the maps is a dictionary. The keys of the returned dictionary are years from 1990 to 2018 and the list associated with each key contains the CO2 equivalent emissions for each of the provinces in alphabetical order. Only the rows containing the total sum of all the CO2 emissions for a province or territory (rather than for a certain sector in the industry) were included in the output. The csv library was used for this function.

The AHCCD – daily TP files were also processed in two different ways: one for the graphs (in `read_daily_mean_temps_all_files`) and one for the maps (in `read_daily_mean_temps_all_files_for_maps`). Both return dictionaries, with the keys as tuples and the values as another dictionary. The tuple for the first dictionary contains the name of the weather station and the province or territory it is located in, respectively. The tuple in the second dictionary contains the id for the station, the name of the weather station, and the province or territory it is located it. In the inner dictionary, the key of this inner dictionary is the year and the value of this inner dictionary is a list of all the temperatures for every day of that year.

There were many issues to account for when processing the daily mean temperatures. The first was the presence of letters beside temperature values, which were removed. The stations that did not contain temperatures from January 1990 to December 2018 had empty dictionaries as the values. Additionally, every month has 31 weather entries, regardless of whether or not the month actually had 31 days. The entries for the non-existent days (such as February 31) were all in the form of `-9999.9`; however, the entries for the days where the climate station failed to record any data also had `-9999.9` as the entries. These two different types had to be distinguished. The former was removed in the processing, while for the latter, the average of the nearest recorded temperature before and the nearest recorded temperature after was taken. The os library was used to process each of the 780 text files at once.

Evidently, because of these issues, the program took quite a long time to run. Thus, the processed data was saved in a json file; the output of `read_daily_mean_temps_all_files` was saved in `data.json` and the output of `read_daily_mean_temps_all_files_for_maps` was saved in `data_for_maps_since_1990.json`. In the subsequent steps of the program, the data was called from these files rather than running the processing program again.

Graphic User Interface

Using the tkinter library, we built a graphic user interface to serve as the first visual component of our program. [7] This library provided us with the option to insert elements such as labels, buttons, and comboboxes, which allows users to interact with our algorithm. Moreover, this file reads and applies the processed data (json file) to generate our interactive visual components.

When the program first runs, a window opens featuring two sections, one to get the map and the other to get the graph. In the map section, the user is prompted to input a year between 1991 and 2018. After that, the user presses the button in order to create the three maps on separate browsers. Such is made possible with the `map_open` function, which checks for valid user input, takes in the data, and runs `plot_emissions_map` as well as `plot_temperatures_map` from `maps.py`. In the second section, the user is prompted to find choose the location by using the combobox. Since there are many stations, the user is able to filter through them based on the province/territory as well as based on their input in the search bar which is unlocked after the province/territory is chosen. [8]

At the bottom of the screen there are two buttons. The first button called "Instructions" opens a new window, where the user can read in more detail about what the main interface is displaying. The second button called "Creators" has a drawn picture of the creators. [9, 10, 11]

Graphing

For plotting purposes, two different data sets are plotted and overlayed.

The first dataset is the greenhouse gases emission data for all provinces. The code extracts the carbon dioxide data for the given province, and then plots carbon dioxide levels in metric kilotons over the years using a `plotly` line graph.

The second dataset is the mean temperature data for different stations. Our program first calculates the average temperature per year for the given station. Then, the temperature anomaly is calculated for each given year, by subtracting the average temperature for one year by the average temperatures over the span of all given years (within function `temp_anomaly`). Using the `plotly` module, the temperature anomalies are then plotted onto a line graph.

And then, our program combines the two line plots for the above-mentioned datasets, making a single interactive line plot with years on the x-axis and primary and secondary y-axes representing carbon dioxide levels and temperature anomalies respectively. [12, 13, 14]

Map

The `plotly.express` library was used to create the maps. The program generates three maps: one that shows the raw CO2 equivalent output from each province in a given year, one that shows the difference in CO2 equivalent output from each province for a given year compared to 1990, and one that shows the difference in daily mean temperatures for each weather station for a given year compared to 1990. The first two maps are created by plot_emissions_map and they are choropleth maps. The third map is created by plot_temperatures_map and it is a scatter map. The colours on the maps for the raw emissions data and the temperature differences were determined on a logarithmic scale, as this allowed the differences to be more easily seen, whereas the emissions difference was graphed on a linear scale. [16]

Instructions to Obtain Data Sets and Run Program

a. Install Python Libraries

In order to be able to run our program, all of the Python libraries listed under `requirements.txt` must be first installed. Please ensure that plotly is updated to the latest version (4.14)! No other preparation is deemed necessary.

b. Download Datasets

The raw datasets used in our final submission can be obtained from the following links:

- Greenhouse gases data from the official website of Government of Canada:
`http://donnees.ec.gc.ca/data/substances/monitor/canada-s-official-greenhouse-gas-inventory/GHG_IPCC_Can_Prov_Terr.csv`

- ZIP file containing temperature data for 780 different Canadian stations:
`ftp://ccrp.tor.ec.gc.ca/pub/EC_data/AHCCD_daily/Homog_daily_mean_temp_v2019.zip`

- geojson file containing the data for the borders of each province and territory in Canada:
`https://blog.exploratory.io/making-maps-for-canadas-provisions-and-census-divisions\-in-r-c189b88ccd8`

- geojson file containing the data for the locations of each of the climate stations in Canada:
`https://geo.weather.gc.ca/geomet-beta/features/collections/ahccd-stations/items`

As stated in our Computational Overview, we cleaned and formatted the raw data into a usable form. The processed datasets can be found at UTSend. Storing all program files under the same folder (that is, the Python files and the dataset files should be in the same folder) is vital for experiencing the program's expected functionality. The following is the claim ID and passcode to access files on UTSend:

- Link to UTSend: `https://send.utoronto.ca/pickup.php`
- Claim ID: sAsdR44Nu3T2BT9i
- Claim Passcode: EZZuKAbE8YeyeJau

Please also ensure that `title_image.png`, `leaf.ico` and `creator_image.png` from MarkUs is stored in the same folder as the `.py` files.

c. Running `main.py`

Upon running the file `main.py`, the user is prompted with a yellow window (Fig 1). The window consists of the title of our project: "Effects of Greenhouse Gas Emissions in Canada" and a couple drop-down lists (or combo-boxes), buttons, one search field, and one input field for year. There are two options for the user to investigate:

To start, there is a View Map option. Under this section, one has to input a valid year (between 1991 and 2018) in the textbox. To the right of the textbox is a button labelled `Map`. Then, on clicking the `Map` button, along with a valid year, three different kinds of maps will be generated on the user's browser tab: one that shows the raw CO2 equivalent output from each province in a given year, another that shows the difference in CO2 equivalent output from each province for a given year compared to 1990, and the final scatter map which shows the difference in daily mean temperatures for each weather station for a given year compared to 1990 (Fig 2). The map shown does not depend on any of the other inputs from combo-boxes, rather only on the input year. That is, the map stays the same regardless of any province or station chosen.

Additionally, there is the View Graph option. Here, the user must first choose a Canadian province or territory from the "Province/Territory" drop-down menu. The search field provides assistance in filtering out the station names containing letters typed into the field, which is executed once the `Search` button is clicked. The stations in that particular province or territory are displayed in the "Station" drop-down list, filtered from the search field input. As soon as one of the stations from the combo-box is chosen, an overlayed line graph of carbon-dioxide levels and temperatures over the years of the province/territory and that particular station respectively is opened in a new browser tab (Fig 3).

Apart from this, there are a couple other buttons that do not contribute towards the running of the program, but are included for additional information and instructions. For instance, the "Instructions" button provides the user with instructions on how to use the program (Fig 4). The "Creators" button contains some general information about the program creator (Fig 5). [15]

Note: the figure numbers refer to the labels of actual figures in the "Figures" section of this paper. Clicking numbers will take you to the respective labels.

Changes to Project Proposal

Everything, no matter how small or big, requires planning before execution. At times, during execution, plans might change depending on situations, and so did our project.

Initially, we had intended to look at the inter-relation of greenhouse gas emissions and change in temperature, and rain and snowfall patterns at a global level. To start, our final submission only examines the greenhouse gas data in terms of CO2 equivalents and the temperature trend within Canada, over the span of 30 years from 1990 to 2018. As a consequence, we used a new dataset as per our requirements. While the C02 emissions dataset from our proposal comprised of values on a global scale, our new source considers only values in Canada. For our map implementation, we have also manipulated the two geojson files as mentioned in Dataset Description, which contain coordinate for the Canadian provinces and territories, as well as coordinates for all weather stations across Canada.

We also decided to shift our focus of the effects of greenhouse gases solely on temperature. Omitting our proposed idea of evaluating snowfall and precipitation, this allows our program to fixate on temperatures and the topic of global warming. To limit the scope of the project, we initially intended selecting one station each in Maritimes, Northern Areas, Alberta and British Columbia as part of our study. With our final work, we have decided to examine stations individually, rather than selecting several from a region or grouping stations. This in turn allows us to compare each station temperature with the CO2 emissions within the overall province/territory.

Moreover, our initial plan was to implement a predictive program trained on the datasets which would allow us to predict changes in temperature, and rainfall and snowfall patterns in regions depending on their amount of greenhouse gas emissions. Instead, our new design investigates the carbon dioxide equivalent of greenhouse gases-temperature association for 512 stations in Canada, and draws conclusion about how temperatures in these regions have been affected over the 1990-2018 period.

To display the relation between greenhouse gas emissions and temperature over time specific to a region, we have implemented a graph for users to view. Users are prompted to select a province/territory and station, which then allows the graph to display.

Additionally, our final submission implements a graphic user interface. This prompts two options for the user: to enter a year to view the map, as well as select the province/territory and search for station location to view the CO2 emissions/temperature graph.

Following our feedback, we have reduced the number of Python libraries used. While our proposed plan included the implementation of NumPy, pandas, statsmodels, scikit-learn, and matplotlib, our final work applied tkinter and plotly modules.

Discussion

Our initial question was **How have GHG emissions affected temperatures in various regions across Canada?** The visualizations from our program have allowed this question to be answered to a certain extent.

For the maps, it is most interesting to note the year input of 2018. The “CO2 Emissions Raw Data Map” shows that the provinces contribute significantly more than the territories to greenhouse gas (GHG) emissions, with Alberta and Ontario contributing the most. The “CO2 Emissions Difference” maps shows that nearly all the provinces and territories increased their net CO2 production from 1990 to 2018, with Alberta facing the steepest increase. To answer the question, these maps must be compared with the “Daily Mean Temperatures Difference” map to see if the provinces and territories that contribute most to GHG emissions also face the largest temperature increases. From this comparison, it is clear that this is not the case. Although there were only a few viable stations in Northern Canada that could be analyzed, all of them are yellow, indicating an increase in average temperature of 1 to 10 degrees between 1990 and 2018. Many areas in central and eastern Canada, however, did not face such drastic changes in average temperature, despite the far greater GHG emissions from these provinces. Western Canada, specifically British Columbia, have a cluster of orange and yellow dots, indicating that they also faced greater temperature increases compared to the rest of the provinces despite contributing similarly to CO2 output. From the maps, we can conclude that climate change affects different regions to different extents, and that this is not proportional to the contributions to climate change (that is, GHG emissions) that a region makes. Those who contribute least (that is, the northern regions of Canada) face the greatest burdens in the form of temperature increases. Climate change is not only an environmental detriment, but also exacerbates the inequalities that already exist in today's world. This highlights the urgency and the dire need to address the issue. To add, through the graph implementation, we are able to scale down our comparison of temperature and emissions by each station, over the span of several years. The graphs examine the changes individually by station. Overall, we can observe that for years with greater values in CO2 equivalent-GHG emissions, there is a trend of increasing temperatures.

There were many limitations throughout the course of the project. Most notably, the daily mean temperatures datasets were quite messy, as many of the temperatures had to be estimated and thus had some measure of inaccuracy. Furthermore, when creating the maps, the only weather stations that could be plotted were the ones that had both their coordinates in the Canada Climate Stations geojson file as well as temperature data from 1990 to 2018. These restrictions on the types of data that we could use resulted in many weather stations being removed, especially in northern areas. It would have been helpful to see more data points from this region, since it seems to be one that is heavily impacted by climate change. Another limitation was with plotly.express, which did not have a default map for Canada. To counter this, a geojson file with the borders for each of the provinces and territories was found. With regards to the task of graphing, our values relating to temperature is produced by calculating the anomaly. While this is more suitable for our purposes (with comparing to emissions), it requires additional operations in contrast to examining the yearly average temperatures. To add, while the provided dataset recorded daily temperatures, we were required to calculate the average per year, over a considerable number of years. An additional limitation we came across concerns the data for territories within the temperature dataset. Some regions, Nunavut and Northwest Territories, for instance, were merged together in the set.

As next steps, we could try to find more data points in the north by decreasing the time range that is considered to a smaller range. It would also be helpful to find a way to combine the emissions and the temperatures maps into one map, as this would allow for the visual comparison between emissions and temperature data to be made more easily. Additionally, for further exploration, we can extrapolate trends for both emissions and temperature within the station on the graphs. Perhaps in prompting the user to enter a year beyond 2018, we may display the estimation of yearly values on our graph based on our current data.

References

[1] NASA. *The Causes of Climate Change.* (2020, August 18). Retrieved November 05, 2020, from https://climate.nasa.gov/causes/.

[2] Climate Action Network. *Impacts in Canada.* (n.d.). Retrieved November 05, 2020, from https://climateactionnetwork.ca/issues/impacts-and-adaptation/learning-cente/impacts-in-canada/.

[3] Environment and Climate Change Canada. (2020, April 21). *Canada's Official Greenhouse Gas Inventory.* Retrieved December 9, 2020, from https://open.canada.ca/data/en/dataset/779c7bcf-4982-47eb-af1b-a33618a05e5b

[4] Environment and Climate Change Canada. (2020, February 19). *Adjusted and Homogenized Canadian Climate Data - Daily Temperature and Precipitation (AHCCD – daily T&P).* Retrieved November 06, 2020, from https://geo.weather.gc.ca/geomet-beta/features/collections/ahccd-stations/items.

[5] Nishida, K. (2017, January 02). *Making Maps for Canada's Provisions and Census Divisions in R.* Retrieved December 14, 2020, from https://blog.exploratory.io/making-maps-for-canadas-provisions-and-census-divisions-in-r-c189b88ccd8a

[6] *Adjusted And Homogenized Canadian Climate Data (AHCCD) Stations: Items.* (n.d.). Retrieved December 13, 2020, from https://geo.weather.gc.ca/geomet-beta/features/collections/ahccd-stations/items

[7] *Python GUI's With TKinter.* (2016, January 18). Retrieved December 14, 2020, from https://www.youtube.com/playlist?list=PLCC34OHNcOtoC6GglhF3ncJ5rLwQrLGnV

[8] User1404. (2019, April 12). Is autocomplete search feature available in tkinter combo box? Retrieved December 14, 2020, from https://stackoverflow.com/questions/55649709/is-autocomplete-search-feature-available-in-tkinter-combo-box

[9] Frankot. (2015, July 11). Tkinter Label image without border. Retrieved December 14, 2020, from https://stackoverflow.com/questions/31360480/tkinter-label-image-without-border

[10] Hughes, J. [james hughes] (2015, June 06). How do I change the text size in a label widget, python tkinter. Retrieved December 14, 2020, from https://stackoverflow.com/questions/30685308/how-do-i-change-the-text-size-in-a-label-widget-python-tkinter

[11] Hu, J. (2019, March 15). *Set Tkinter Background Color.* Retrieved December 14, 2020, from https://www.delftstack.com/howto/python-tkinter/how-to-set-tkinter-backgroud-color/

[12] *Multiple Axes.* (n.d.). Retrieved December 12, 2020, from https://plotly.com/python/multiple-axes/

[13] *Python Figure Reference: layout.* (n.d.). Retrieved December 14, 2020, from https://plotly.com/python/reference/layout/#layout-paper_bgcolor

[14] *Python Figure Reference: Single-Page.* (n.d.). Retrieved December 14, 2020, from https://plotly.com/python/reference/#scatter-line-color

[15] *Inserting Images.* (n.d.). Retrieved December 14, 2020, from https://www.overleaf.com/learn/latex/Inserting_Images

[16] Singh, N. (2020, July 19). *Plotting Choropleth Maps using Python (Plotly).* Retrieved December 14, 2020, from https://www.youtube.com/watch?v=aJmaw3QKMvk

Figures

Different outputs of our program

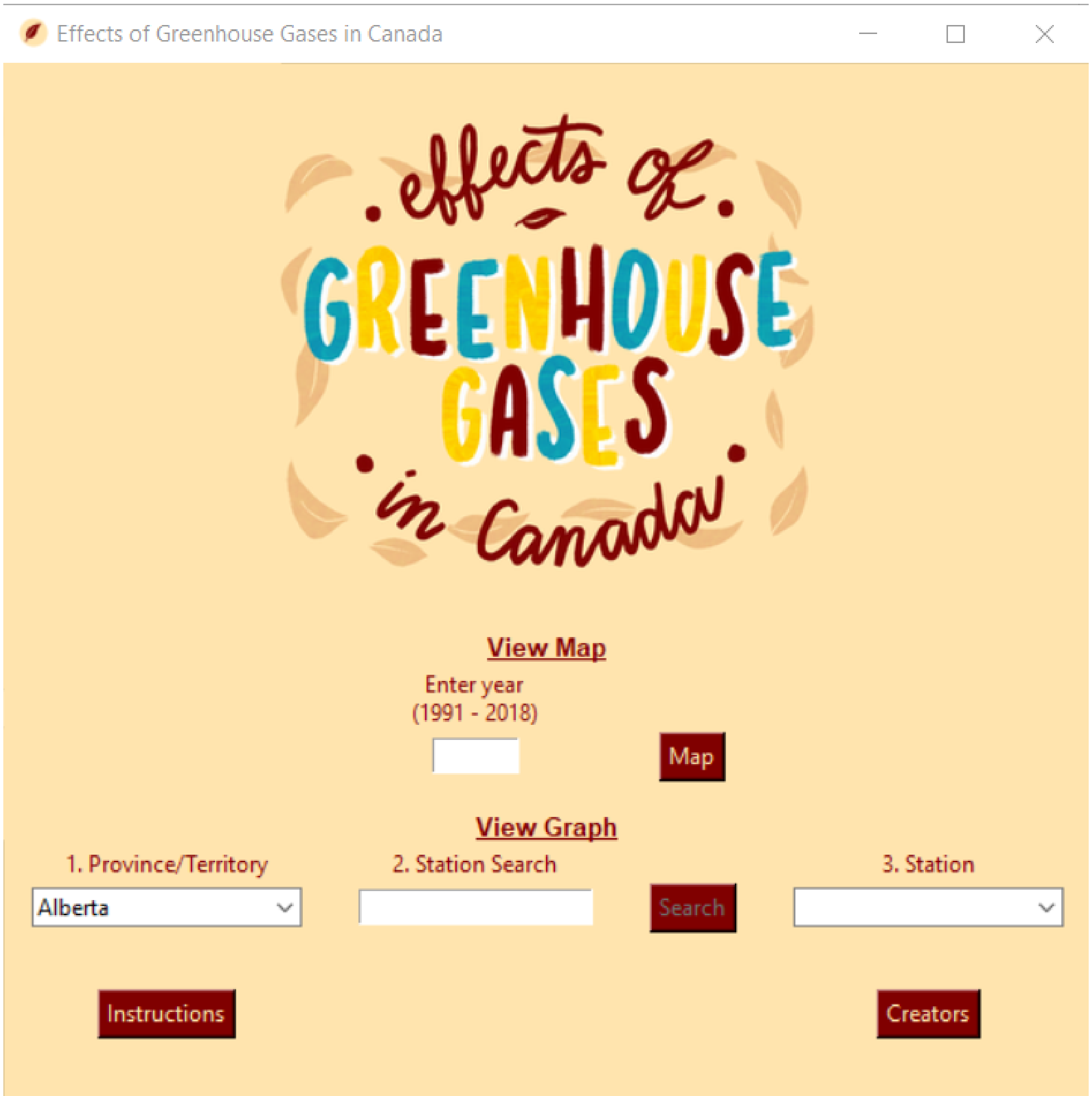
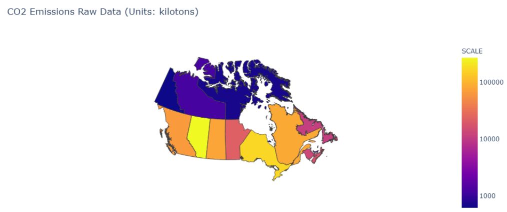
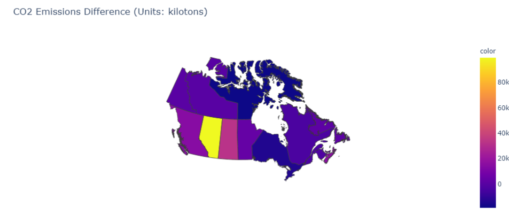


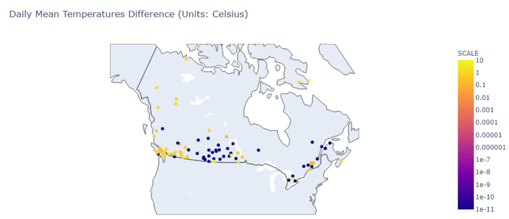
Figure 1: Window



[a]



[b]



[c]

Figure 2: Map for [a] Raw Emissions Data [b] Emissions Difference [c] Daily Temperature Difference with Year Input 2018

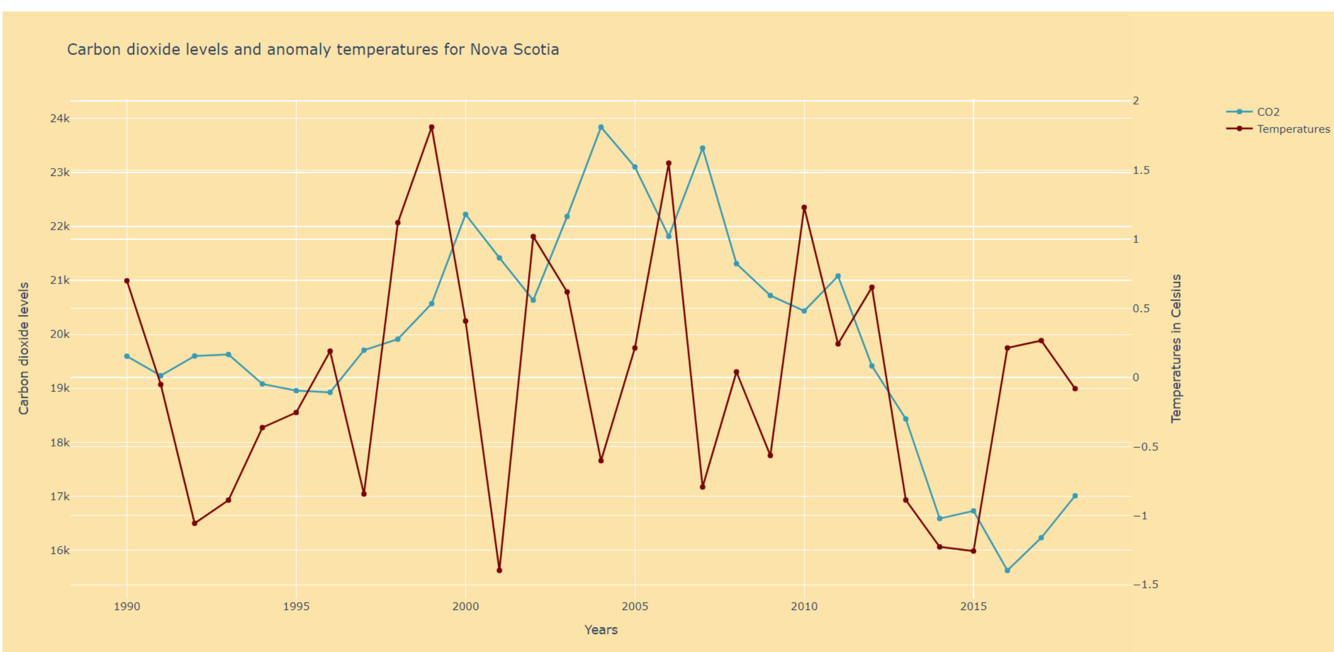


Figure 3: Line graphs for carbon dioxide and temperature

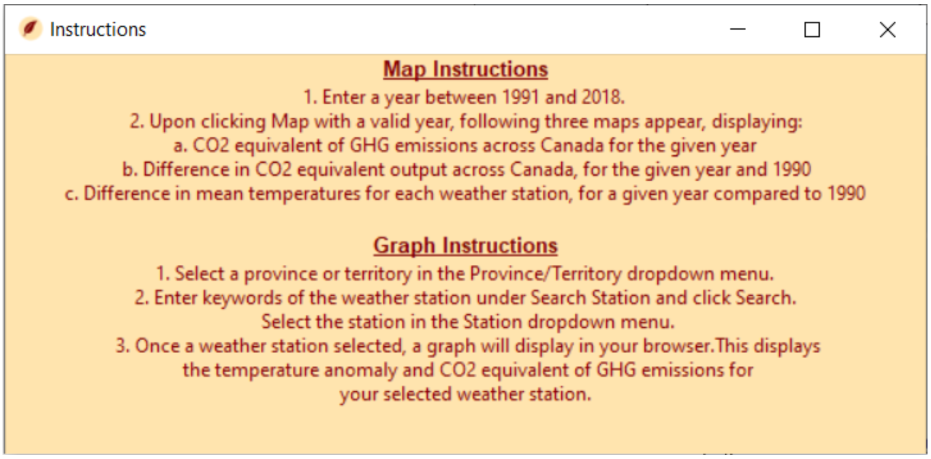


Figure 4: Instructions window

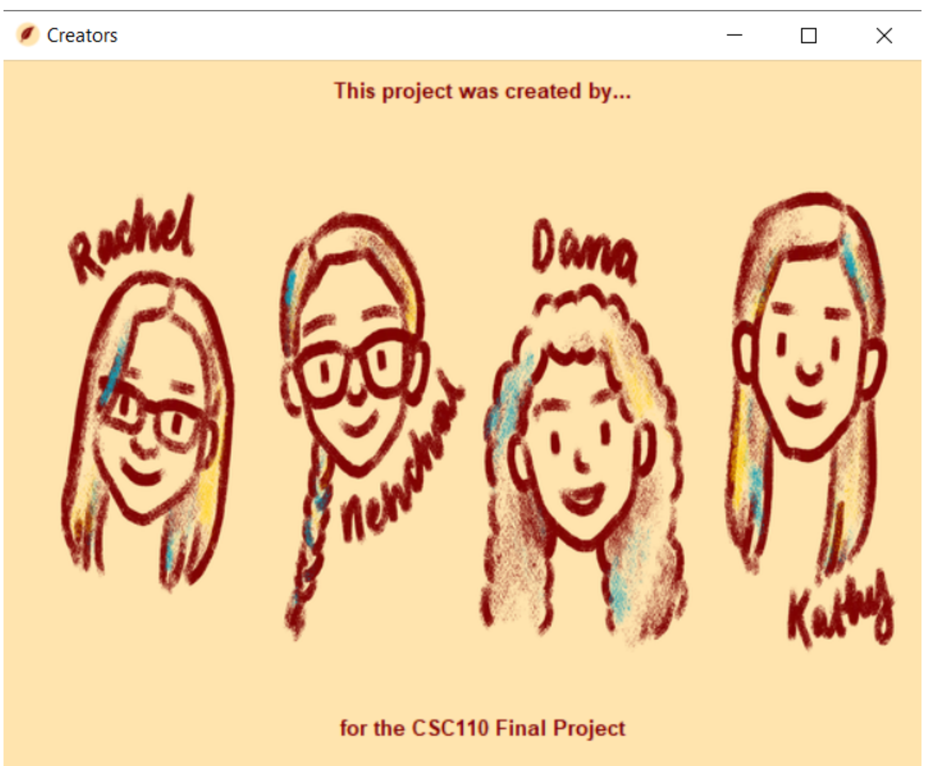


Figure 5: Creators window