INF1340   
Sheldon Chen, Lo Humeniuk, Benjamin Kelly, Will Power-Jenkins  
19 December 2021

# Canadian Crop Yield & Climate Analysis

## Introduction

This program allows a user to view historic environmental data spanning from 1940 up to 2019, and to observe whether there is a relationship between yearly crop yields and precipitation and mean annual climate, gathered by province. Within the graphic user interface, a user has the option to select a province, type of crop, and temperature of precipitation data to view a histogram showing the relationships within this data.

The data available for this program was robust, with many weather stations collecting daily, or hourly, temperatures. In order to analyze a large set of data in a short amount of time, we elected to use mean annual temperature data within our program as opposed to looking at highs and lows or extremes. We recognize there are implications in this, and that it may not give a true glimpse at how crop yields and temperature relate over time or account for the fact that climate change doesn’t necessarily mean that temperatures are getting warmer every season of every year consistently (in fact, the warming of the Arctic has been linked to colder winters further south).[[1]](#footnote-0) Still, we feel it still gives a look at general trends over time and may serve as a starting point for those interested in using data science to look at environmental factors and correlations with agricultural yields.

## Build Status

Data was not collected from all selected weather stations each year, and some provinces began collecting data later than others; as a result, there are some gaps in data. Likewise, some provinces did not have some crop yield data readily available; as a consequence, some of the graphs and maps generated may seem skewed as they reflect this missing data. Currently, the program does not scrape new data as it is collected, though this is a possible modification that could be made. In the ‘How to Execute’ section, a link to our collaborative coding document is included; code was maintained here which shows how environmental data was imported and cleaned from a dataset hosted on Kaggle[[2]](#footnote-1) and from Climate Change Canada[[3]](#footnote-2); those interested in modifying their dataset or looking at different weather stations may find this code useful.

Within the ‘How to Execute’ and ‘Features’ sections, we offer additional options for how this program might be built upon, modified, or extended to suit different needs.

## Requirements

Those interested in running this program must be able to install or run the following:

\* numpy (np)

\* pandas (pd)

\* StatsCan

\* tkinter (tk)

\* matplotlib.pyplot (plt)

\* sklearn

\* statsmodels (sm)

\* statsmodels.formula (smf)

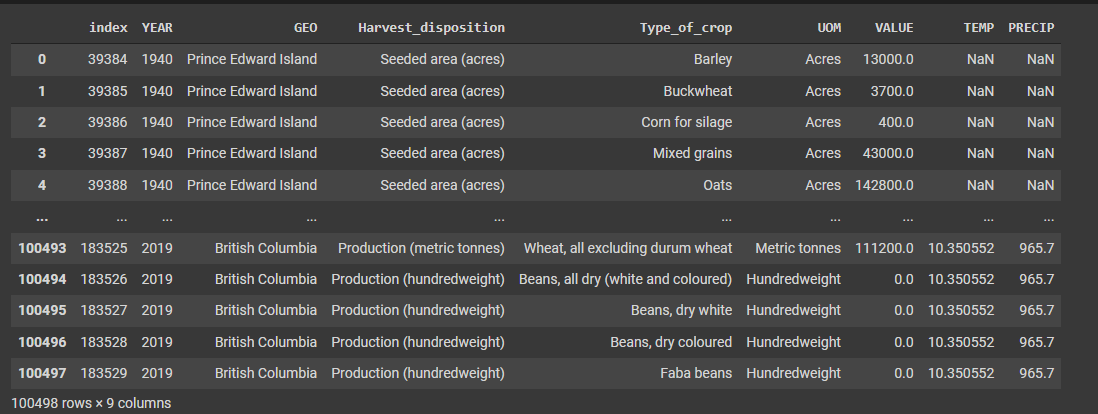
## How to Execute the Program

Files used are stored at <https://github.com/BenKelly-Data/Canadian-Agricutural-Yields-vs-Climate>.

To execute, run the .py file. If .pip install does not work on your system, use conda install - c conda-forge stats\_can. The program may encounter a long runtime of up to 10 minutes due to data cleaning.

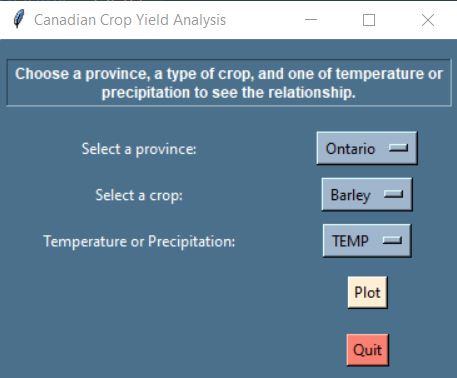
If the .py has install issues you can also try our Google Collab code: <https://colab.research.google.com/drive/1sEcezR13fLWKICQGDxlkSFKFCUAswJd1?usp=sharing>

Within this program, data is first imported and cleaned; as mentioned, we used Statistics Canada’s API for crop yield data, and 2 other sources for environmental data (see notes 2 and 3).



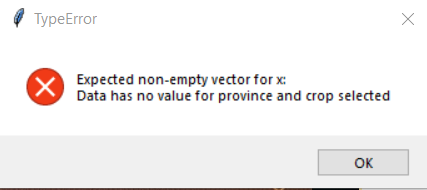
*Figure 1: How this data looks when aggregated into one table*

The GUI we developed has a main menu with three drop-down options for the user: type of crop, province, and either temperature or precipitation.

**

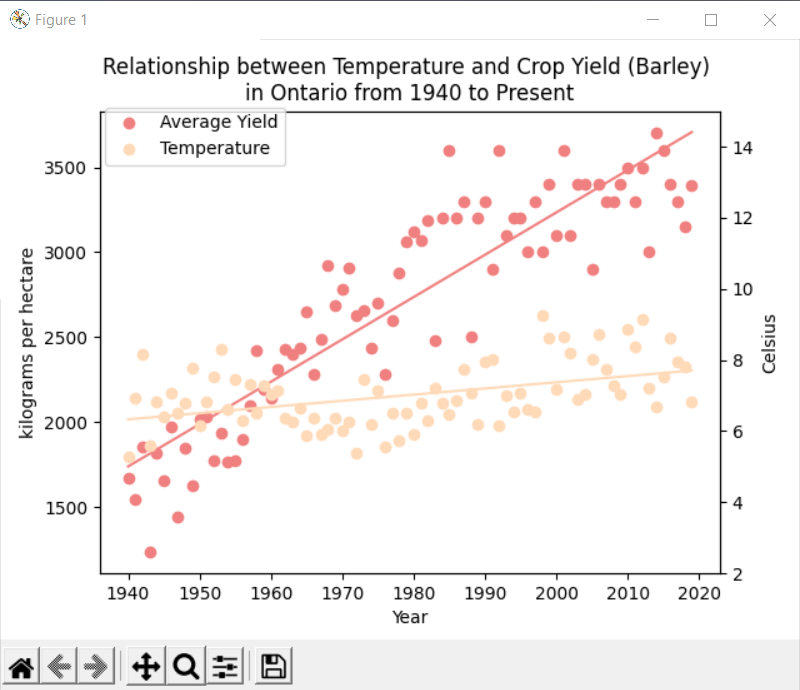
*Figure 2: Main menu showing instructions and buttons for province, crop, temperature/precipitation, plot, and quit.*

If a user selects a ‘bad value’ (i.e., attempts to view data that is not available), they will receive an error message explaining why the graph cannot be plotted.



*Figure 3: Error message noting: “Expected non-empty vector for x: Data has no value for province and crop selected.*

Once a selection is made, the user has two buttons they can choose from: Plot or Quit.

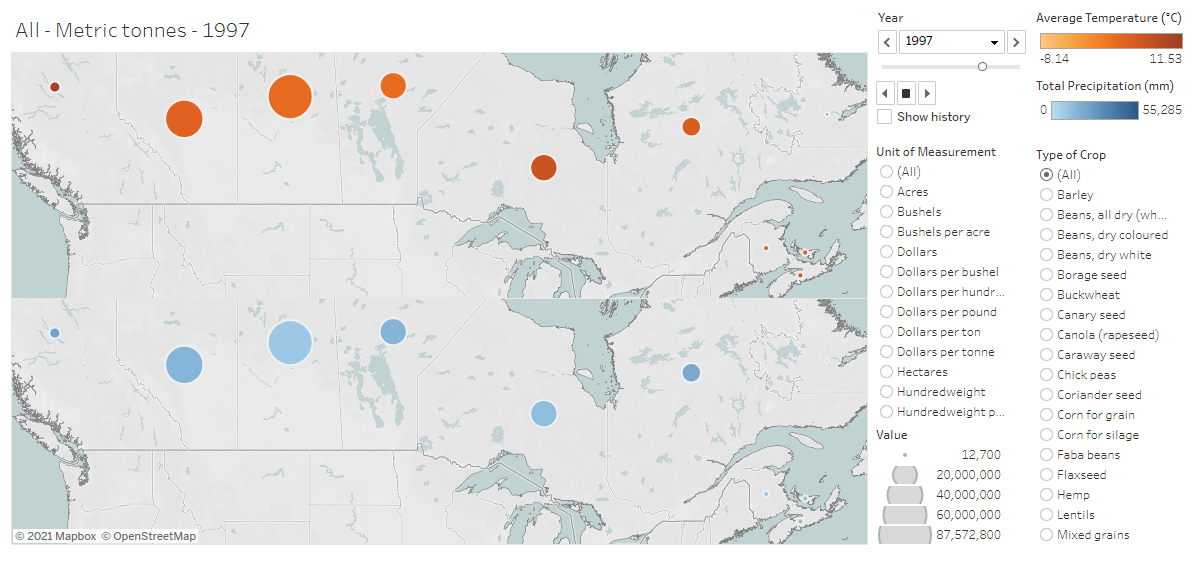
The Plot button will produce a graph such as this:

*Figure 4: Example graph showing data when Ontario (province), Barley (crop type) and Temperature are selected.*

## Features

### Map

We have plotted all of the data onto a map in Tableau based on the datasets and code created. A user can click on plot points within each province to view average annual temperature, total annual precipitation, and value (modified by unit of measurement as the actual amount), and can toggle to view this data by year, crop type, and unit of measurement.



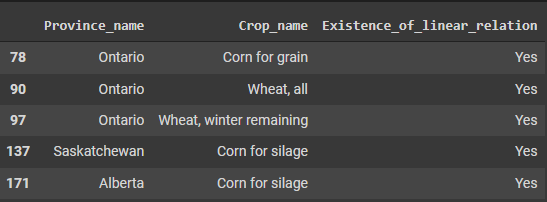
*Figure 5: Display of mapped data for metric tonnes of all crop types in 1997.*

Here is the link to interact with the map: <https://public.tableau.com/shared/DMJBHZYXR?:display_count=n&:origin=viz_share_link>

### 

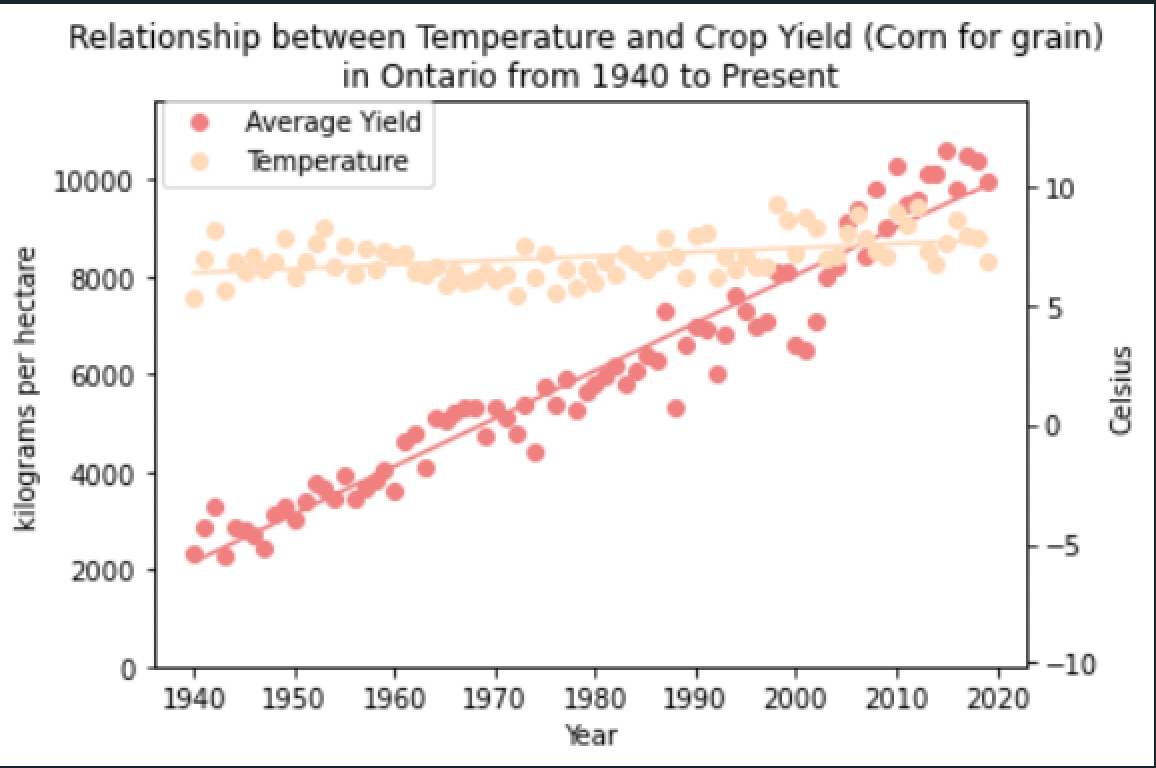
### Linear Regression

In our analysis of these data sets, we also ran linear regressions; the following are what we found to have significant relationships.



*Figure 6: Chart showing provinces and crop names that have positive linear relationships.*

Within our analysis, both the temperature *and* the precipitation must have a correlation with crop yield for it to indicate a positive relationship. Though we have not included this within our program or GUI, a feature could be added to show which variable (temperature or precipitation) crop yield is correlated with (and in that instance, a feature could show whether it has a correlation with one, or both variables).

**

*Figure 7: A positive correlation between corn for grain (crop type), year and temperature in Ontario (province).*

This information can be extrapolated on, and code could optionally extend such that future crop yields could be predicted for those instances of positive correlations. Please refer to the code in our collaborative coding document that shows how this could be done.

1. See McGrath, M. (2021, September 2). Climate change: Arctic warming linked to colder winters. *BBC News.* https://www.bbc.com/news/science-environment-58425526 [↑](#footnote-ref-0)
2. Data downloaded from: Turner, A. (2019). 80 years of Canadian climate data. *Kaggle.* <https://www.kaggle.com/aturner374/eighty-years-of-canadian-climate-data> [↑](#footnote-ref-1)
3. Data downloaded from: Government of Canada. (2021). Historical data. https://climate.weather.gc.ca/historical\_data/search\_historic\_data\_e.html [↑](#footnote-ref-2)