

Project: Explore Weather Trends

Submitted by: Amir Khanlou

For

Data Analyst Nanodegree Program

March 23, 2020

Step 1: Data Extraction

For this the following SQL queries used to extract the data we need.

- 1) First we check which cities in Canada are listed in the database

```
SELECT *  
FROM city_list  
WHERE country LIKE 'Canada'
```

The results are 6 cities as per this:

Kingston	Canada
London	Canada
Montreal	Canada
Ottawa	Canada
Toronto	Canada
Victoria	Canada

The nearest city to Vancouver where I live is Victoria thus Victoria is selected

- 2) Looking at the city data. For this two queries are used to extract the data as CSV

Query 1: for just having a look at the data

```
SELECT *  
FROM city_data  
LIMIT 10
```

Query 2: for getting all the data of only Victoria and saving it as CSV. The observation was that there are 186 lines of data available

```
SELECT *  
FROM city_data  
WHERE city LIKE 'Victoria'
```

- 3) Getting the global data. For this one query is used to extract the data as CSV. The observation was that there are 266 lines of data available

```
SELECT *  
FROM global_data
```

Step 2: Data Manipulation in the Spreadsheet

The two downloaded CSV files, one being the temperature trend of the city of Victoria and the other the global temperature trend were imported to a single Excel spreadsheet file for further processing.

I decided to make two more columns for each data set, first being the moving average of the yearly temperature trend averaged over 5 years span and the second averaged over 10 years span. The line charts of them are shown below.

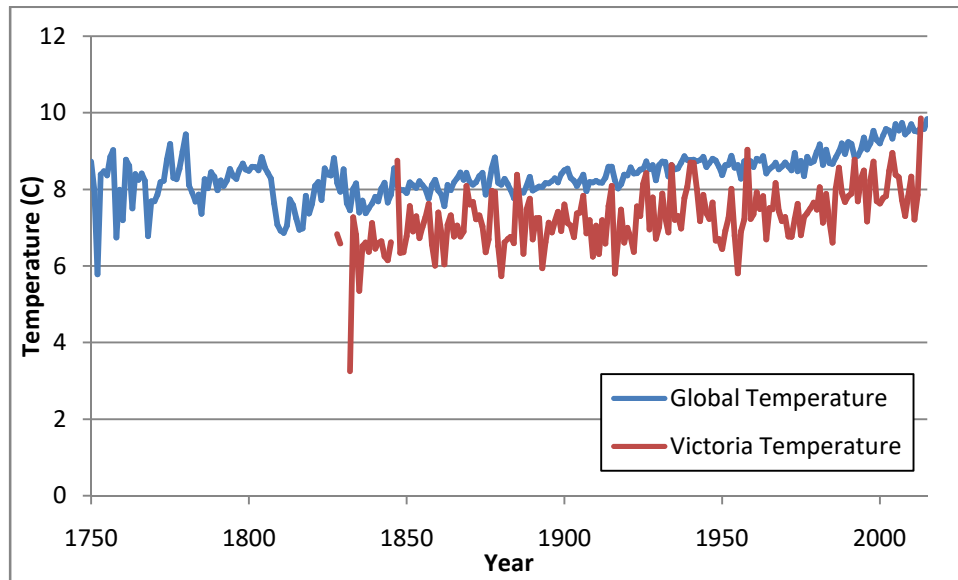


Figure 1. Yearly global and Victoria temperature

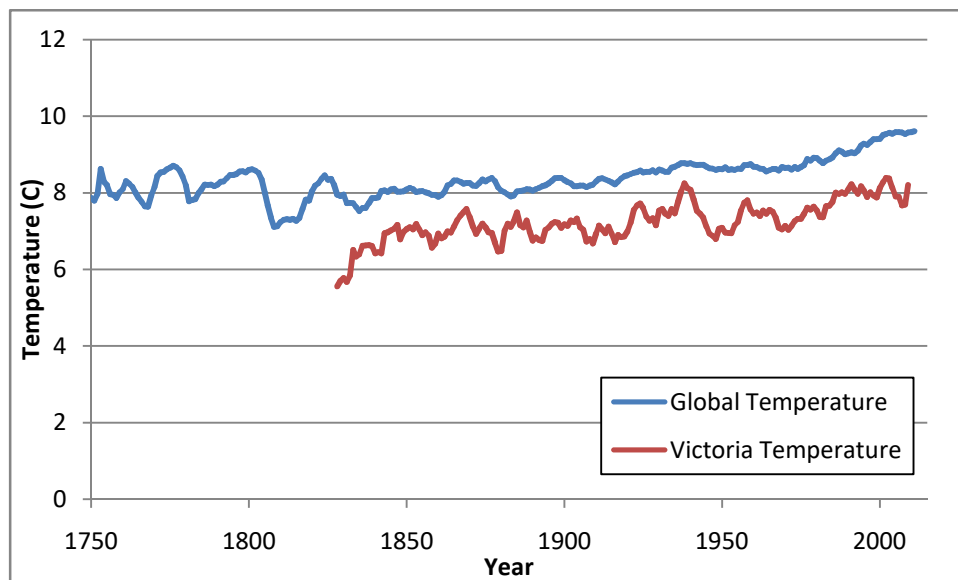


Figure 2. Yearly global and Victoria temperature with moving average over 5 years span

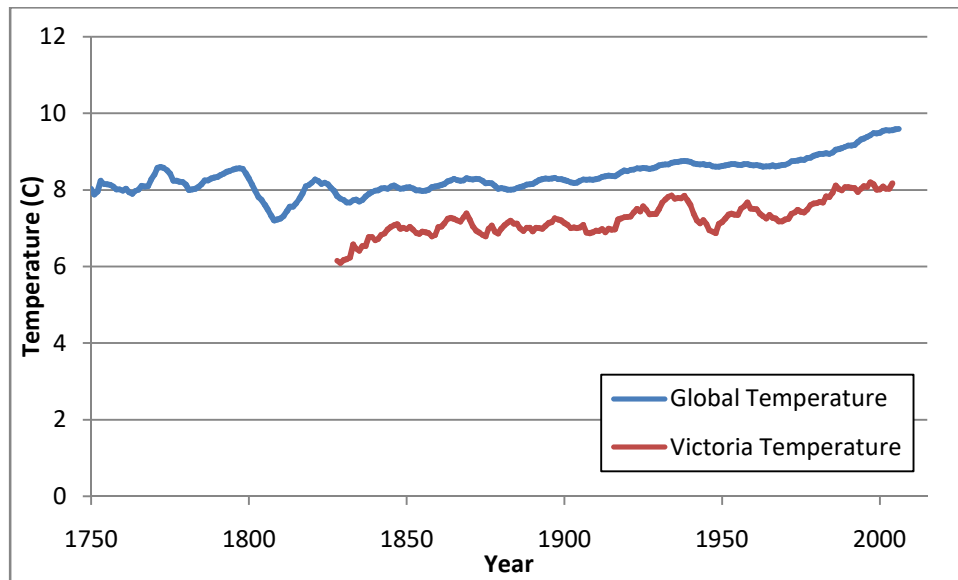


Figure 3. Yearly global and Victoria temperature with moving average over 10 years span

Figure 1 uses the raw data as it was extracted from the database. This figure shows significant fluctuations and although one can conclude the upward trend as the year's progress, the fluctuation of the data makes the chart less desirable to observe and make comparisons. Because of this, the moving average technique was used to make the chart smoother and more readable. First I experimented with averaging the data over the span of 5 years and the result is shown on Figure 2. This chart looks much smoother than the original chart; however as a second experiment, I tried averaging the data over 10 years period which is shown on Figure 3. The Figure 3 appears to be fairly clean and observable therefore this chart was selected for the subsequent analyses.

Step 3: Observations

- 1) Observation 1: Looking at Figure 3, it seems Victoria is overall colder than the global average temperature. The Victoria data is missing some of the earlier data preceding year 1828. For this reason a window of data where both data overlap is selected between years 1850 to 2010 for making comparisons between the two datasets which would equate to 160 years of observations. The average global temperature is 8.5C for this period while Victoria's average temperature is 7.3C, 1.2C colder than global average temperature. While this difference is averaged over the entire 160 years, just by looking at the line chart one can observe that this temperature difference seems very consistent over the entire overlap period.
- 2) Observation 2: Both data sets show a general upward trend meaning the world is getting hotter. Interestingly the rise of the temperature accelerates around year 1840 when the industrial revolution reached its maturity (https://en.wikipedia.org/wiki/Industrial_Revolution) and continued to rise afterward.
- 3) Observation 3: In order to quantify how much the temperature has risen in the selected period, the maximum and minimum temperature of each data set was selected and the difference is calculated as per the table below. This table shows the global temperature is risen by 1.58 degree over the period of 160 years. This number is 1.48 degree for the city of Victoria.

	Unit	Global	Victoria
Max Temp	C	9.54	8.20
Min Temp	C	7.97	6.72
Delta Temp	C	1.58	1.48
Average Temp	C	8.50	7.30

- 4) Observation 4: In order to better quantify the trend of temperature rise and compare the two dataset, figure 3 is redrawn with only overlapping data. This helps adding trend line to these curves using regression methods to extract further numerical values. For this, linear trend line is selected to fit straight lines to these two curves. Although this may not be ideal given the nonlinearity of the dataset, for the purpose of this study and simplicity, we deem this to be an acceptable compromise. Figure 4, shows these curves with the associated trend lines. What is important here is the slope of the trend lines. The global temperature curve follows the slope factor of 0.0075 while Victoria's curve shows 0.0061 for the slope. From this one can conclude that the global temperature is rising more rapidly than Victoria's temperature given this methodology.

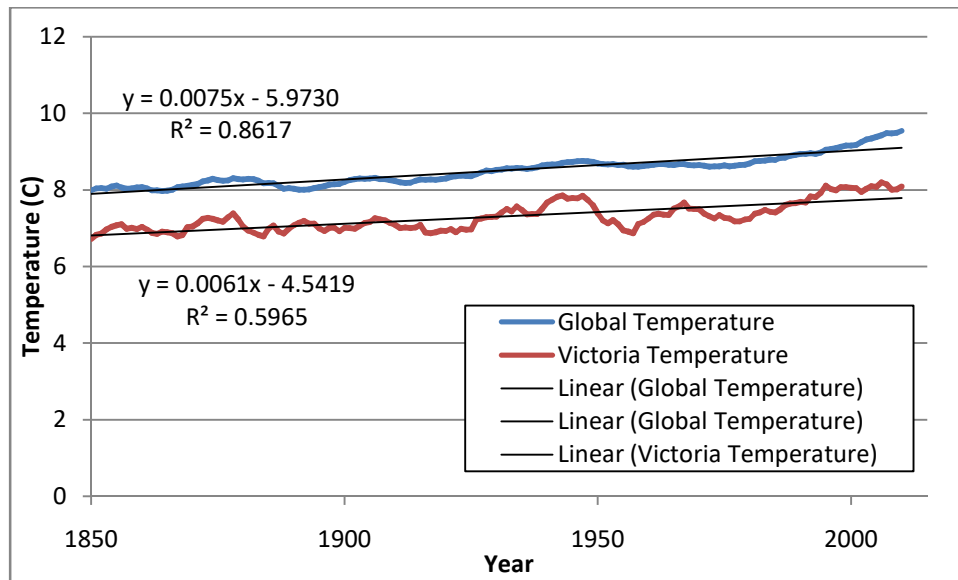


Figure 4. Yearly global and Victoria temperature with moving average over 10 years span for 160 years overlap with linear regression fit