Exploring Weather trend

May 29, 2020

1 Introduction

Global warming is a hot topic today. On average, it is believed that Earth's climate system is warmer now than it was about eighty years ago. This implies that regions that are warm are getting warmer, while colder regions are becoming warm. Experts (including data scientists and analyst) see the trend is accelerating; they demonstrate global warming by analyzing temperature data collected by weather stations on land and in the ocean, and by measurements of various effects of the warming.

This project provides a proof of global warming. Using the weather trend data on a database in Udacity portal, we have these objectives:

- to analyze temperature trend in Victoria, the capital city of Canada province, British Columbia and compare with the global temperature trend,

- create a data visualization and finally,

- write up our findings.

2 Methodolody

The computing languages used in this project include

1. SQL for extracting data from Udacity database

2. Pandas in Jupyter Notebook, for viewing and cleaning my data.

3. Microsoft excel for calculations and creating chart.

4. LaTeX for writing this report.

2.1 Organizing the project

The first thing to do is make sure that all our files relating to this project are organized in a single folder to ensure efficient running of all codes. For example, to insert picture in this LaTeX file, we will need to have the picture stored in the same folder where the Latex file is. Hence since this project is for the Udacity Nanodegree program, I have a folder on my desktop called Udacity_program, and in this folder I created another folder called Weather_trend. This Weather_trend will store all files concerning the weather trend project.

2.2 Extracting data from Udacity database

For this project, three tables in the database were provided with their corresponding instructions:

- city_list - This contains a list of cities and countries in the database. Look through them in order to find the city nearest to you.

- city_data - This contains the average temperatures for each city by year (°C).

- global_data - This contains the average global temperatures by year (°C).

Therefore using the workspace connected to Udacity database schema, I wrote the following SQL queries to extract my data

1. I live in Canada; so the following code will show me the list of cities in Canada from which temperature data was collected from.

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

Out of the six cities given, the city closest to where I live is Victoria, the capital city of the Canada province, British Columbia.

2. The following code extracts temperature data from Victoria city and join it with global data to obtain a single dataset.

```
SELECT
        city_data.city AS "city_name",
        city_data.year AS "city_year",
        city_data.avg_temp AS "city_average_temp.",
        global_data.avg_temp AS "global_average_temp."
    FROM
        city_data
    JOIN
        global_data
    ON
        city_data.year = global_data.year
    WHERE
        city_data.city='Victoria' AND city_data.country='Canada'
```

Our final data has 186 rows and 4 columns. The column names include city_name, city_year, city_average_temp, and global_average_temp respectively. From the download link on the SQL database, I then export the result of this query to .csv.

2.3 A look at the data/ data cleaning

I opened up the downloaded csv file in Jupyter Notebook from my personal computer and performed the following operations:

- I wrote the following code to have a quick look at my data

```
import pandas as pd                    #importing pandas

df=pd.read_csv('results-9.csv')        #we assign the csv file to the variable df
df.head()                              #prints few lines of my data

df.isnull().sum()                      #checks for missing values in data
```

- Result from the last command shows that there were three missing values in the Victoria temperature data. To fix the problem of missing data I wrote the following code:

```
#compute mean of Victoria temperature data and assign it to variable means
means = df['city_average_temp.'].mean()

#fill missing values with (means) the calculated mean
df['city_average_temp.'].fillna(means,inplace=True)
```

which fills the missing values with the mean of the Victoria temperature data.

- I confirmed that the missing values were filled using the code
  df.isnull().**sum**().**any**()

- Finally I saved the changes to result_edited.csv using the code
  df.to_csv('result_edited.csv', index=False)

3 Data Visualization

I opened up the edited .csv file in Microsoft excel spreadsheet.

3.1 Creating chart for average temperature

With the data opened on Microsoft spread sheet, I created a line chart with the following instruction

- from the table, highlight city_year, city_average_temp and global_average_temp columns respectively.

- From Insert menu, choose the XY(Scatter) > scatter with straight lines chart option and the chart is created.

- Then from 'Chart design > Add Chart Element ¿ Axis Titles, I labeled the horizontal and vertical axis respectively. Finally I gave the chart a title

- I saved chart as ave in .PDF format. This chart was appended into this .tex document using the following code
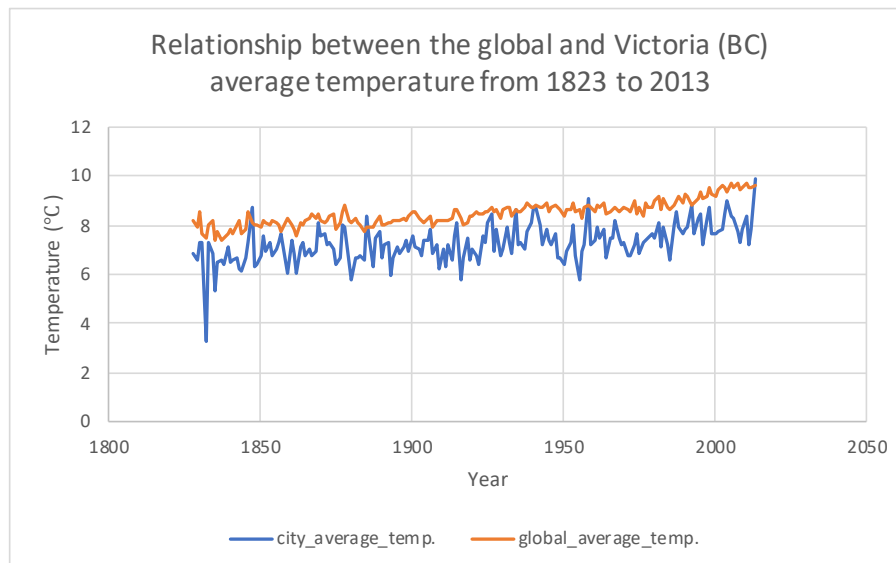

Figure 1:   Comparing the global and Victoria (Canada) average temperature from the year 1823 to 2013.

```
\begin{figure}[h!]
\begin{center}
\includegraphics[width=0.75\textwidth]{ave}
\caption{ Comparing the global and Victoria (Canada) average temperature from
the year 1823 to 2013. }
\label{setups}
\end{center}
\end{figure}
```

Fig 1 compares the temperature trend between Victoria city and the global temperature trend. The fluctuations make it difficult to perform our analysis or give a proper report on the data set. A convenient tool is the moving average as we will see in the next section.

3.2 Moving averages

Moving averages are used to smooth out data to make it easier to observe long term trends and not get lost in daily fluctuations. I calculated the 14-year moving average for each temperature data as follows

- I created two new columns, one titled 14-Day MA_city where the 14-year moving averages for Victoria city would be stored, and the other titled 14-Day MA_global, where the global moving average would be stored.

- From the 14-Day MA_city column, I scrolled down to the fourteenth (14) year (which is 1841) and used the AVERAGE() function to calculate the average temperature for the first fourteen years of temperature data (from 1823 - 1841). Then I dragged the cell all the way down to 2013 to fill the entire column with the formula. I repeated this process for the 14-Day MA_global column.

3.3 Creating chart for the moving averages

Following the same procedure as in 3.1, this time around with the city_year, 14-Day MA_city and 14-Day MA_global columns respectively highlighted, the following chart was created
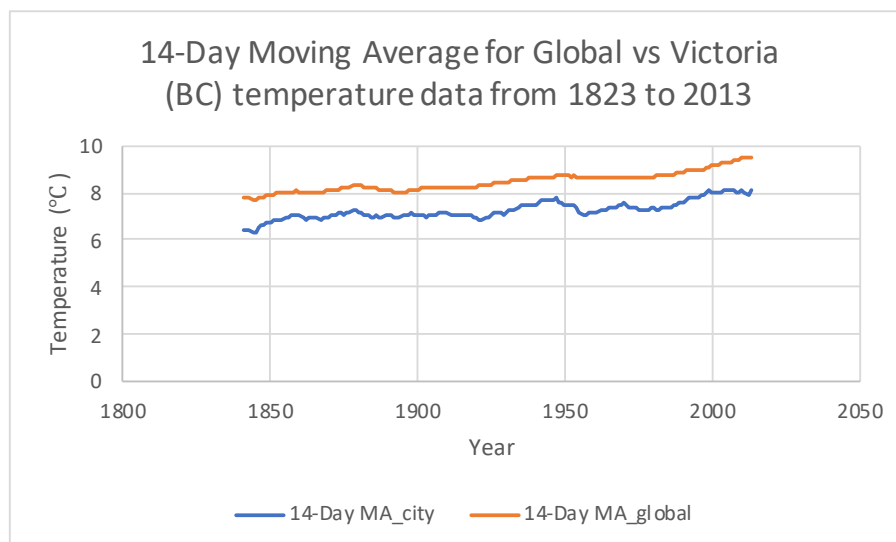

Figure 2: Comparing the 14-day moving average for global and Victoria (Canada) temperature trends. See inset for colour scheme

4 Observation

- The positive increase in the line chart (Figure 2) between the global average temperature and the Victoria city over years, is indeed an indication that the world's climate is gradually warming up.

- Figure 2 shows that the Earth was 1.74 degrees hotter in 2013 than it was in 1823, while Victoria city was 1.65 hotter in 2013 than it was in 1823.

- Canada is cold compared to many countries in the world and it is believed that its average temperature have warmed up by approximately 1.3 degrees which is about twice the global rate []. Our data and chart shows that average temperature warming in Victoria city still fall below the

global average temperature. We can also observe from the chart that the last and most recent 13 years (2000-2013), temperature across Victoria was steady with slight increase (about 0.06 degrees) across years, while the global average temperature continues to rise by 0.4 degrees.

- We can also observe from the chart that around the years (1887 -1922), temperature across Victoria and the Earth was steady with slight increase across years.

- The coldest year in Victoria and the Earth were observed in 1845 and 1844 respectively while the hottest years were 2005 and 2013 for Victoria city and the Earth respectively.