

PROJECT 1: Exploring Weather Trends

UDACITY

Data Analysis Nanodegree

Deepak Nandipati

Date Created: May 1, 2019

City: Toronto, Ontario. Canada

Introduction

The purpose of this project is to examine the trends of average temperature between global and local city temperatures. As I reside in Toronto, the temperatures of this city was extracted and compared directly with global temperatures.

Tools used

- **Jupyter Notebook:** Text editor used to observe changes made with Python
- **SQL:** Data Wrangling and extraction of raw data
- **Python:** Executing calculations for moving averages and plotting graphs
 - Packages used: *Numpy, Pandas, Matplotlib*

Outline of Project

The project used tools: SQL and Jupyter Notebook utilizing python language to deduce the data and make appropriate observations. Before any data can be observed it first needed to be extracted from raw data, then cleaned, and exported to CSV where the python will now utilize the CSV file to make calculations and plot data.

SQL:

Step 1: Identifying if 'Toronto' exists in database

- SQL query was used to first determine if my local city showed on the list of cities. For that I used the code:

```
SELECT *  
FROM city_list
```

```
WHERE country LIKE 'Canada'
```

- The output listed Toronto as one of the values; next step can now be progressed

Step 2: Changing column names

- The column name for average temperature is same for both `global_data` and `city_data`; I changed it in order to prevent any confusion later on during the calculations; This step is optional

```
ALTER TABLE city_data
    RENAME COLUMN avg_temp TO city_avg_temp;
ALTER TABLE global_data
    RENAME COLUMN avg_temp TO global_avg_temp;
```

Step 3: Select necessary data and extract to CSV

- I selected the information needed to analyze data, such as the average temperatures of both global and Toronto along with years. I merged the tables and formed one final table with needed data and exported CSV file

```
SELECT global_data.year, global_data.global_avg_temp,
city_data.city_avg_temp
FROM global_data

JOIN city_data
ON global_data.year = city_data.year
WHERE city LIKE 'Toronto';
```

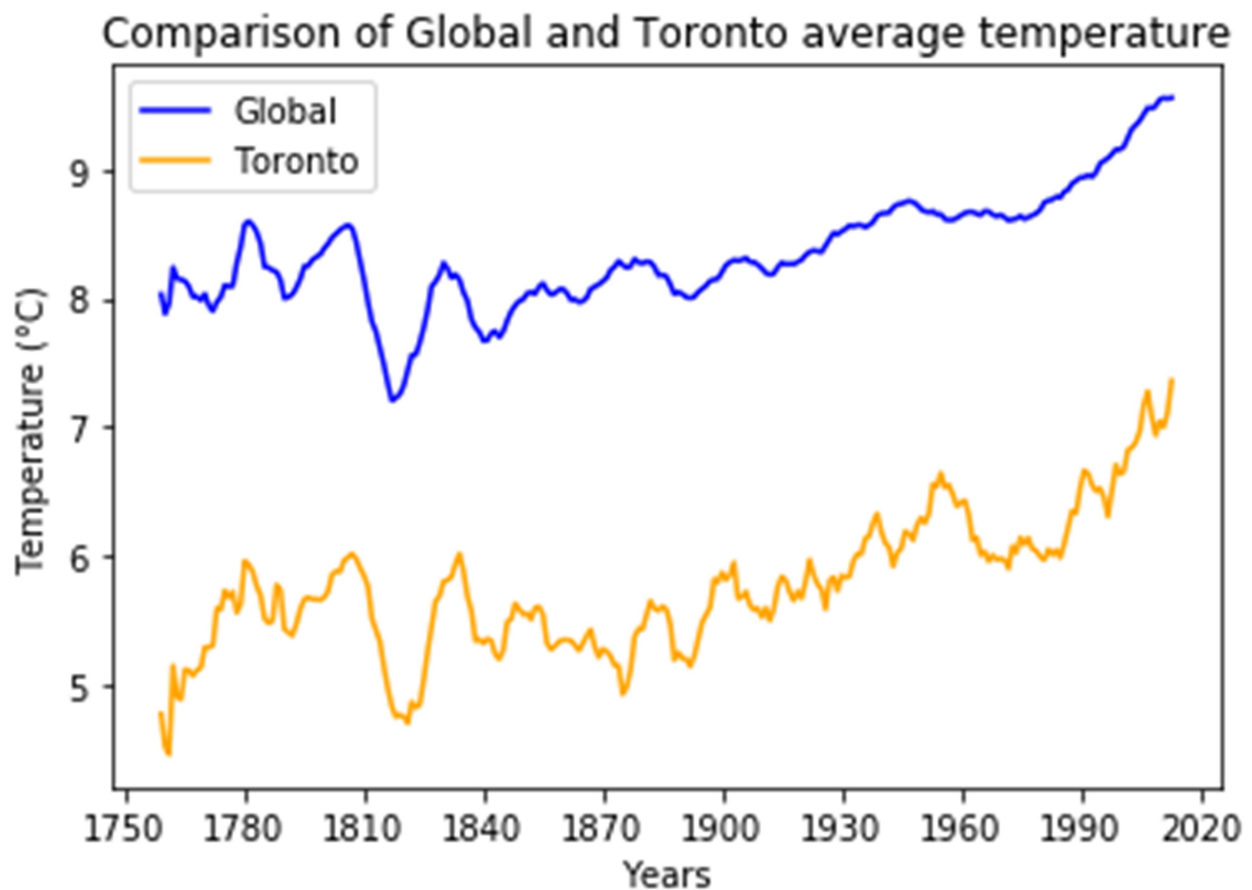
Python:

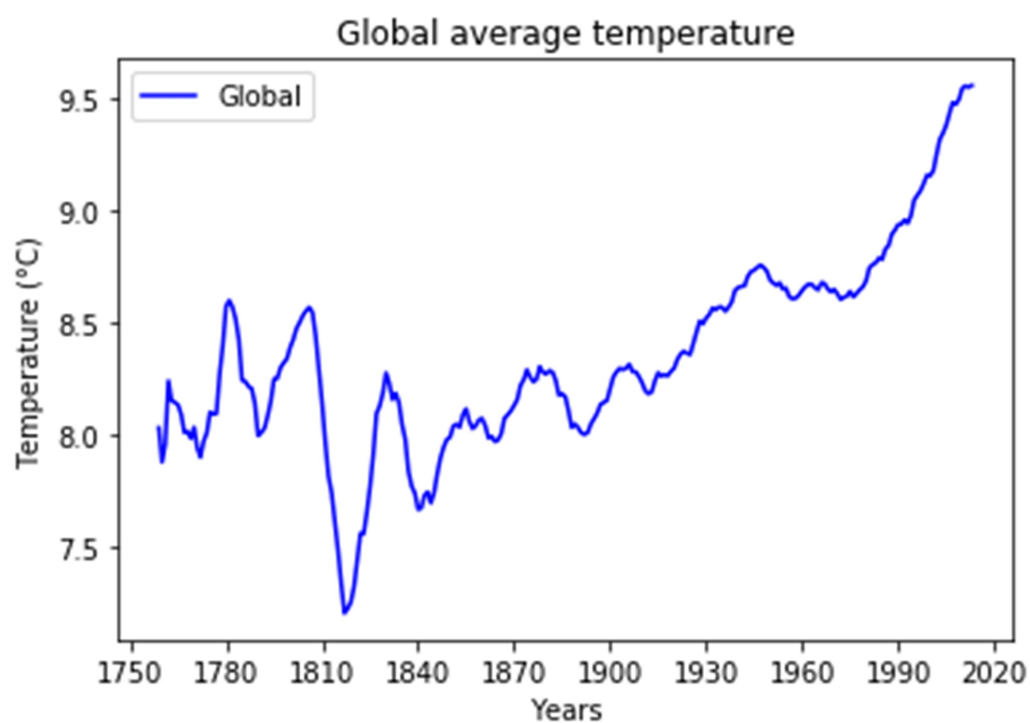
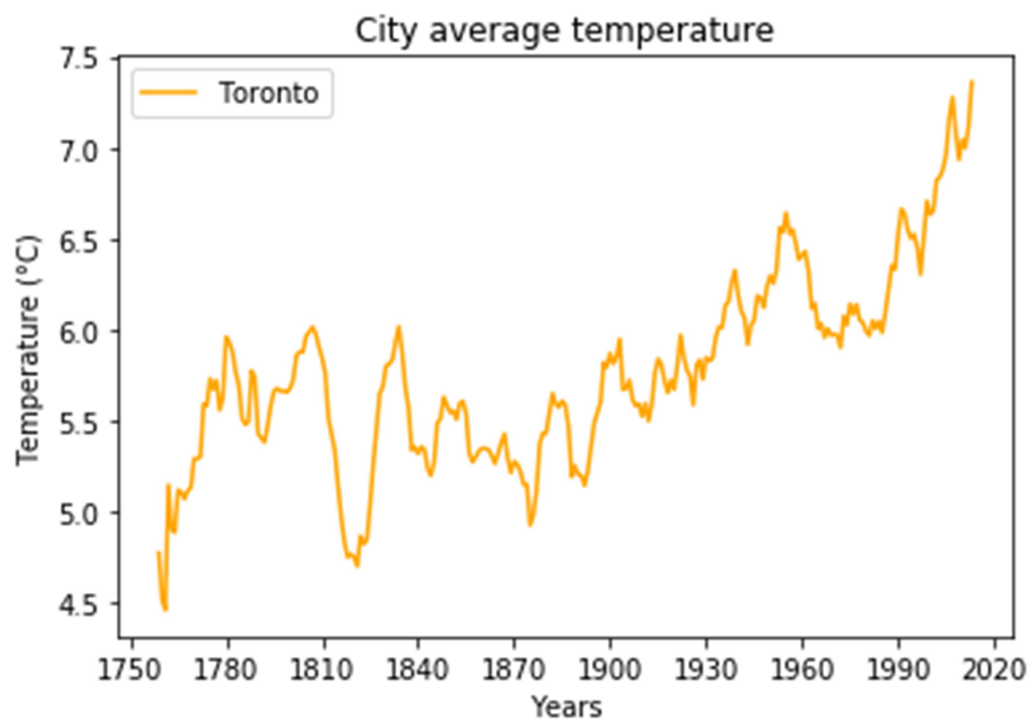
Moving average calculations

- Moving averages was needed in order to smoothen the data across the graph; this process can be conducted in various ways but I used the *pandas* function
 - ‘rolling’ along with the ‘mean’ function
 - `Pandas.DataFrame.rolling.mean()`
- The moving average or rolling average calculated for every 10 years

- The reason for the 10 years rolling average was chosen because the earliest data is of 1750 and with the latest being 2013. With the spread of years, the window function needed to be higher to smooth the graph, but not too high as the trends in temperature can be missed. With trial and error, the window function #10 was chosen.

Line Graphs





Key Observations

- When you look at individual graphs, it is evident that both the global and city temperatures are both increasing and the future trends suggest that it will continue to increase
- Combined line graphs provide clearer indication that the Toronto average temperature is lower than the global temperature; it was calculated that the average city temperature of Toronto during the given periods of 1750 to 2013 is of 5.7 degrees Celsius, while global is of 8.4 degrees Celsius. This indicates that there is only a 2.6 degrees differences between the Toronto average temperature and global average temperature
- Global and city increase in temperature is more evident after the 1900s and has been more rapidly increasing since. Several reasons for that but most obvious is due to the 1900s was right after the industrial revolution and the period of both world wars where a lot of industrialization has occurred.
- Both global and city average temperature decreased suddenly during the 1810 to 1840 period. The cause is unknown but this difference is easily identical between both. They both decreased approximately 1 degree Celsius
- The increase in temperature can be determined by observing the latest years of temperature in comparison to the earlier temperature records. Below you can see the first 5 and last 5 values of temperatures. One can notice the increases in temperatures for both global and city averages. It is projected to further increase.

	year	global_avg_temp	city_avg_temp
0	1750	8.72	6.29
1	1751	7.98	6.84
2	1752	5.78	-1.10
3	1753	8.39	5.76
4	1754	8.47	5.94

	year	global_avg_temp	city_avg_temp
259	2009	9.51	6.28
260	2010	9.70	7.77
261	2011	9.52	7.30
262	2012	9.51	8.66
263	2013	9.61	8.46

Conclusion

In summary the trend appears to be very similar for global and Toronto average temperatures, where both are projected to increase with time. The differences in temperature between global and my city is not drastic, but only different by less than 5 degrees Celsius. More information is needed to further decipher temperature fluctuations such as the period of 1810 to 1840.