Christoph Uhl
1/17/2020

Exploring Weather Trends

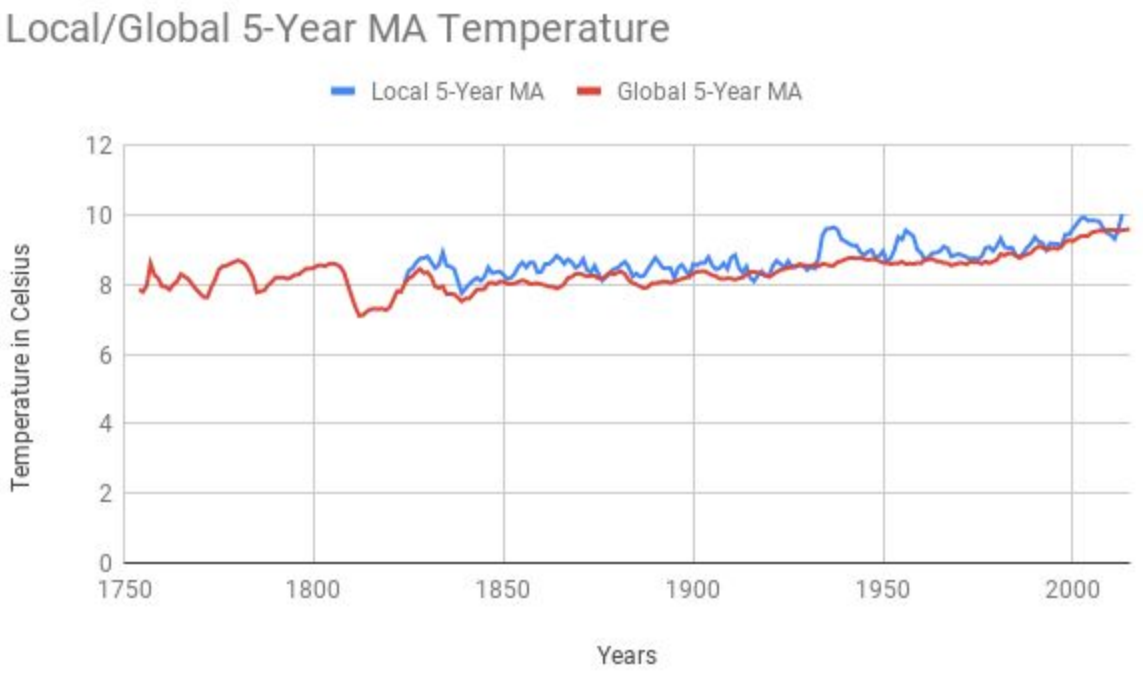
To obtain the data I used the following SQL queries for the local and global data respectively:
"SELECT * FROM city_data WHERE city = 'Denver'",
"SELECT * FROM global_data".

I then downloaded the results of those queries as csv files and imported both to Google Spreadsheets.

Looking at the imported data, I noticed that the local temperature data available had entries in the years 1775 and 1776, but after came a long period of no data. Finally in the year 1820 a temperature data point appeared. Since my goal was to plot a moving average of 5 years, I figured that I cannot use the data points of 1775/76, so I only started calculating the 5-year Moving Average from 1820, leading to the first Moving Average data point in the year 1824. The global temperature data looked entirely clean, so I had to do no data cleaning there.

The way I calculated the Moving Averages was by using Spreadsheets' 'AVERAGE()' method, and selecting the instantaneous data point of the MA we are trying to calculate along with the instantaneous data from each of the 4 years prior to it. This gives the average of the temperatures of the 5 most recent years, including the current year whose MA we are trying to calculate.

To create the graph, I had moved the year, local MA, and global MA columns next to each other and then I offset the local MA data points such that the global and local data points correctly align with the year column. Finally I selected the 3 columns and made a chart out of them. This is what I found:



My key considerations for visualizing this trend were the number of years for the Moving Average, and properly aligning the series based on the years.

As we can see, the global MA has more data to show than the local MA, this difference in data points is likely because Denver (the city used for local temperature data) was only founded in 1858. I am wondering where the two data points from 1775 and 1776 came from, which leads me to wonder where the temperature data came from in the first place.

Another observation is that Denver is consistently warmer than the World's average, excluding the 5-or-so exceptions.

Local data points seems to fluctuate much more than global ones, I would guess that this is due to the difference in scale between the entire world and a single city.

Lastly, there is a clear trend that the temperatures for both local and global 5-year Moving Averages are increasing with time. I believe this would be an argument for climate change.

This was the first Udacity project, it didn't involve much coding in Python, as using Google Spreadsheets proved a much simpler and efficient approach to make some simple analyses on the data. It was great to review my basic SQL knowledge and practice describing my work as well as analyzing a graph.