Weather Trends Project Outline

I first used the following SQL queries to extract the data needed to analyze global weather trends versus weather trends local to my hometown. I was able to conclude that Detroit was the nearest city within the given data set.

```
/* Query to find my closest city */
SELECT c.country, c.city
FROM city_list c
WHERE c.country = 'United States'
ORDER BY c.city


/* Query to pull temperature data for Detroit */
SELECT c.year, c.city, c.avg_temp
FROM city_data c
WHERE c.city = 'Detroit'


/* Query to pull global temperature data */
SELECT g.year, g.avg_temp
FROM global_data g
```

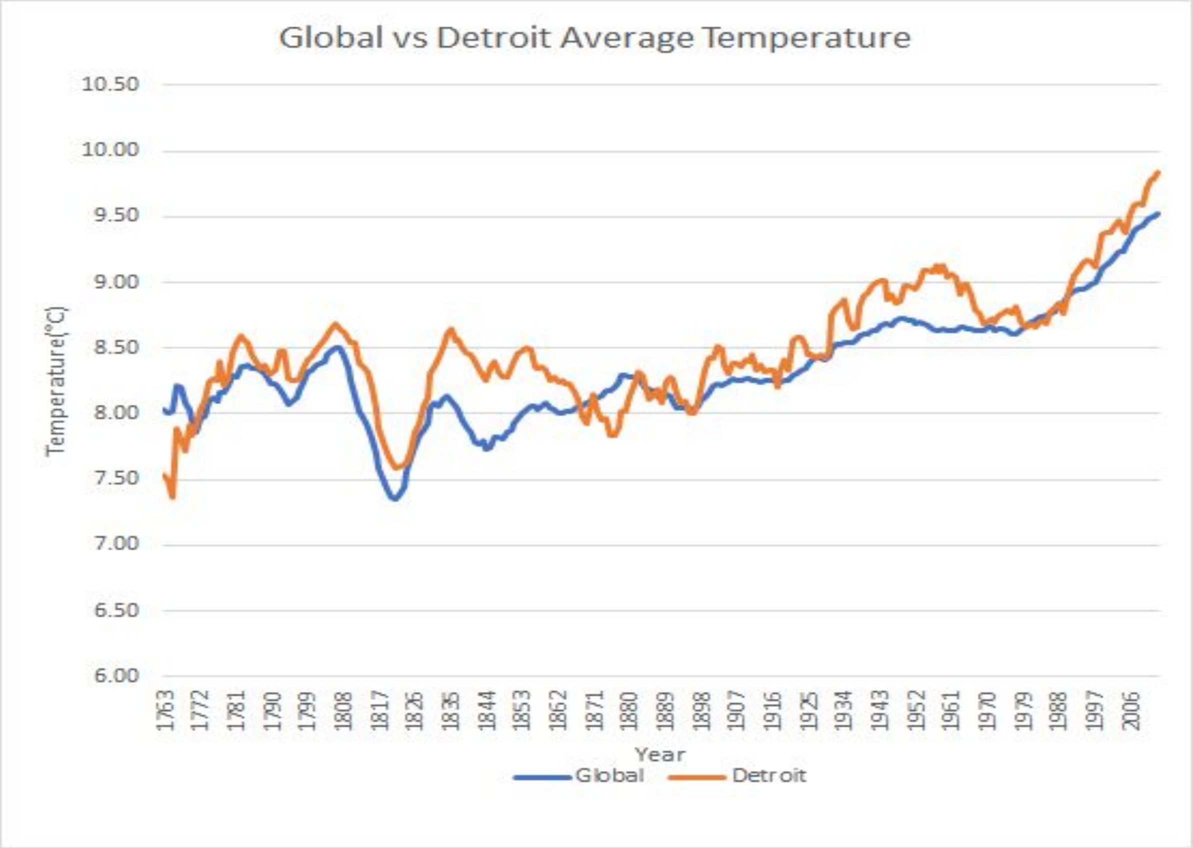
Weather Trends Outline (cont.)

I extracted both the global and Detroit weather trends as separate csv files to be analyzed in Excel. The first thing I noticed when looking at the data is that the first year that the average temperature was recorded for Detroit is 1743. Since the global data does not start until 1750, I chose to omit 1743-1749 weather trends for Detroit. I also chose not to use global weather trends for 2014 and 2015 because there are no temperature averages for Detroit in the given data for those years. So, to start off with, all of the data was narrowed down to temperatures recorded from 1750 to 2013.

To calculate the moving average, I used the average formula to find the average temperature of the first 7 years from the global data, filled the formula down the rest of the column, and then did the same for the Detroit temperatures. I also followed the same process to calculate the average temperature for every 14 years. While the 7-year average was a vast improvement as far as smoothing out the line graph and it is a more accurate visual representation of the data because it shows how little difference there has been between Detroit temperatures and the global average, the 14-year average is better at showing us where the few differences lie.

I also converted the original data into Fahrenheit for those of us that live in countries that do not rely on the metric system. I wanted to clearly understand how much of a difference there was between the global and Detroit data, and it was much more clear once I saw the 14-year moving average in Fahrenheit and in a line graph.

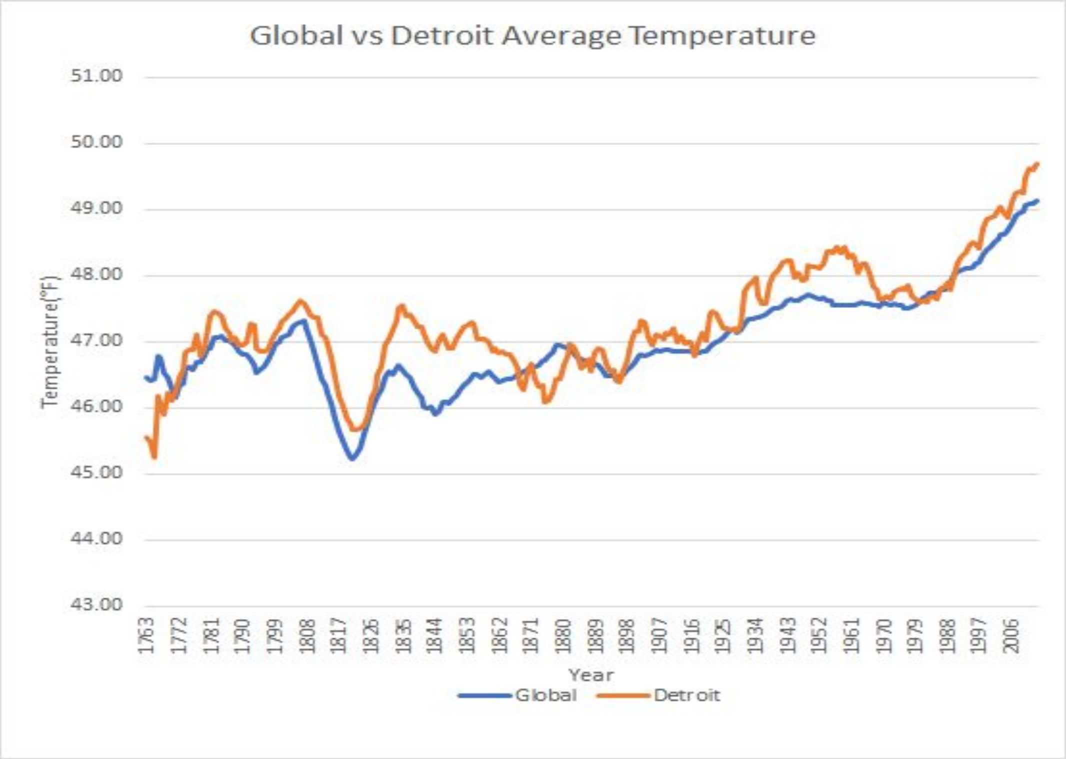
14-year Moving Average (ºC)



Here we have the data based on a 14-year moving average. To reiterate, this graph is useful for prominently showing us where there have been differences between global and Detroit temperatures. In particular, the largest gap seems to be from 1830 to 1870.

However, for anyone used to a non-metric system, the differences might seem larger than they really are, and even though I initially converted individual temperatures into Fahrenheit, it wasn't until I saw the full line graph in Fahrenheit that I intuitively understood how small the differences are between the Detroit and global data.

14-year Moving Average (ºF)

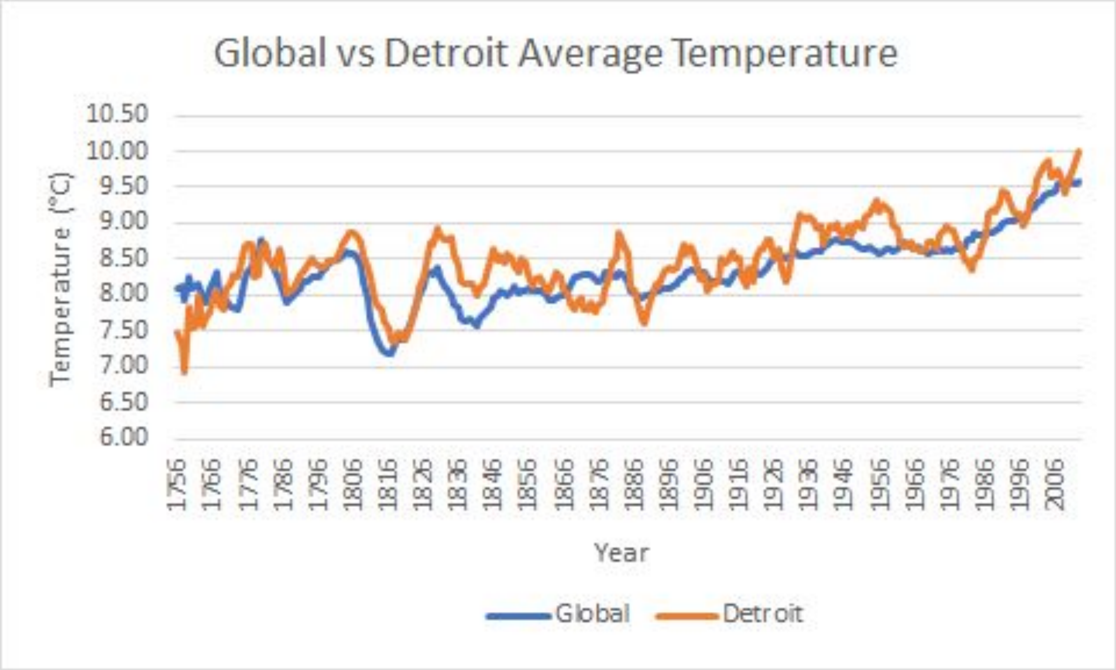


In the previous graph, one of the largest gaps between each line is approximately 7.73 °C for the global temperature and 8.25 °C for Detroit. While it certainly doesn't sound like a large variation, converting the data to Fahrenheit makes it easier to understand for those not used to the metric system.

Unless someone routinely works with both systems or data on a regular basis, there is an inherent understanding of how small a 1°F difference is compared to just over a 0.5°C discrepancy for the average person.

However, a one degree rise in global temperature overall is concerning, so I contemplated whether or not a difference of one degree was significant in this case. I will address this concern a bit later in the last graph.

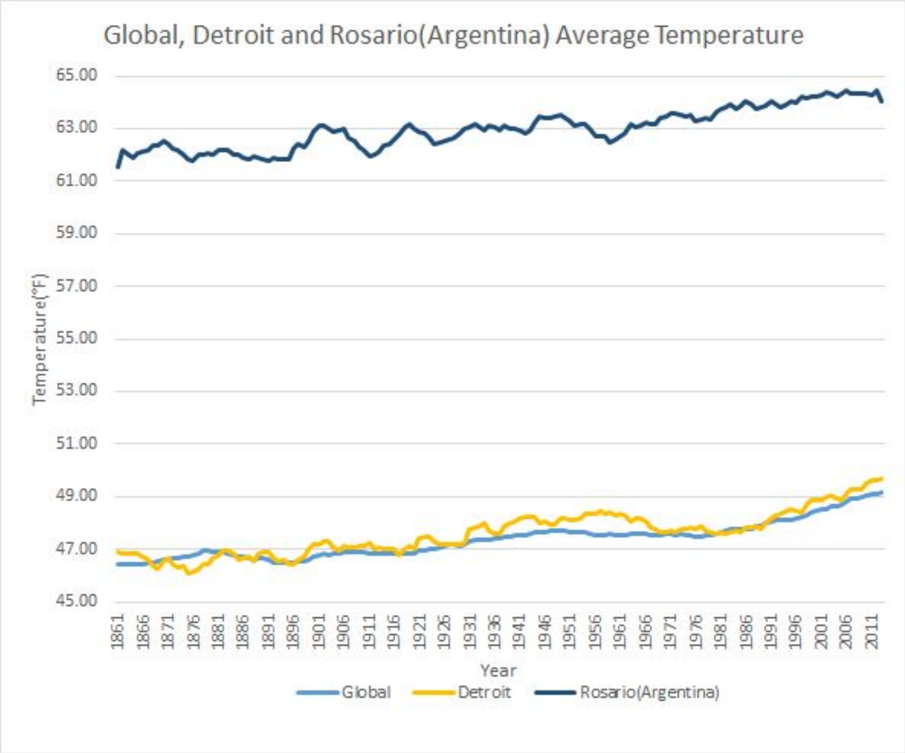
7-year Moving Average (°C)



Here it can be seen how using a 7-year moving average more accurately shows how little deviation there is between the Detroit and global temperature data. Take note that I made sure the bounds are the same as the graph for the 14-year moving average.

However, when looking at any of these graphs I don't believe that many people with an understanding of statistics will be surprised to hear that the correlation coefficient for the global data is 0.61 and for Detroit is 0.41. This informs us that both data sets are positively correlated to the year, with a moderate correlation, which is apparent regardless of which graph we are viewing. The Detroit data is slightly less moderately correlated, while the global is slightly more moderately correlated. It confirms, yet again, that these data sets are quite similar.

14-year Moving Average (ºF)



Finally, we will come back to whether or not one of the few gaps between the Detroit and global temperatures of one degree (°F) is actually a valid concern.

While even a one degree change in global temperature upon itself is something to worry about, the gap between Detroit and global temperature is interesting to note, but not a concern. Otherwise, the large gap between the temperatures in Rosario, Argentina (one of the hottest locations in South America) and the global data would be detrimental.

What is worrisome is the rise in temperature for all data sets. Detroit temperatures have risen almost 4.5°F, and global temperatures have risen about 2.6°F. Also, since Rosario's data didn't begin until 1855, I recalculated the correlation coefficients for Detroit and the global data starting in 1855. While the Detroit temperatures became more moderately correlated to the year with a 0.55 positive correlation, the global data now has a strong positive correlation with a 0.86, and Rosario starts off with a strong correlation at 0.68.