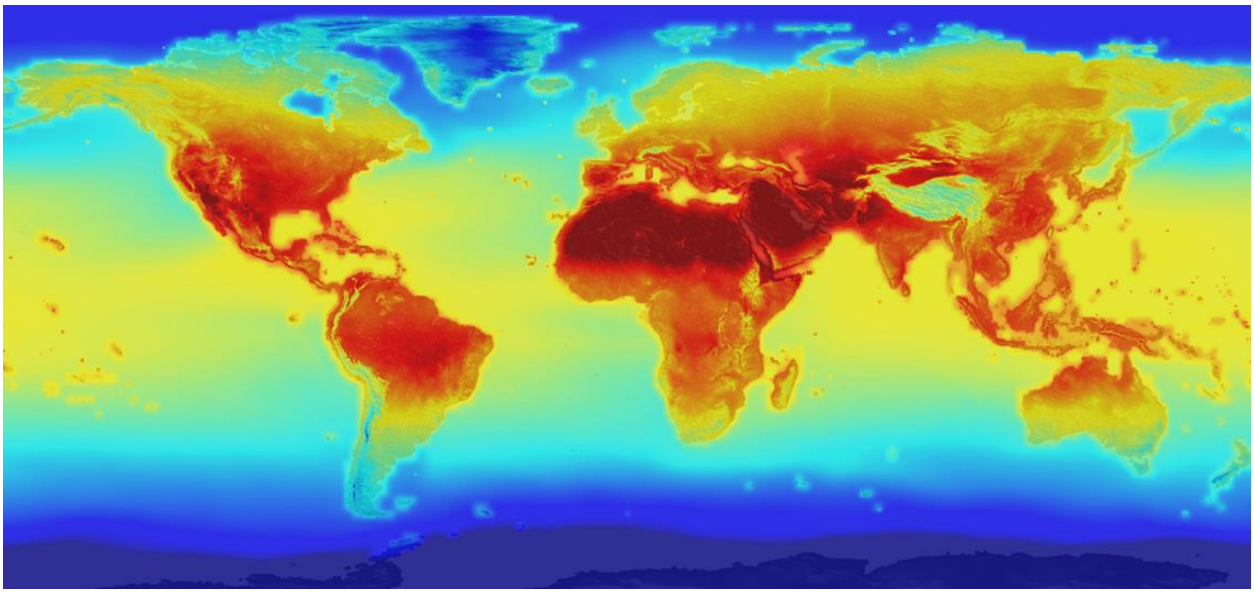
Ryan Sikhrangkur

Udacity Data Analyst Nanodegree

Exploring Weather Trends



For this presentation I shall be analyzing the weather trends of my local major city of New York, United States, and through this data discuss my observations of these local trends as well as how they correlate with global weather trends.

Step 1: Retrieving the Data

My first step was to retrieve the data I needed through the Database Schema and identify from the list of cities which would be most applicable to me. After viewing the columns within each table, I filtered the number of cities available by writing a query which would return all listed within my country. (Figure 1.1)

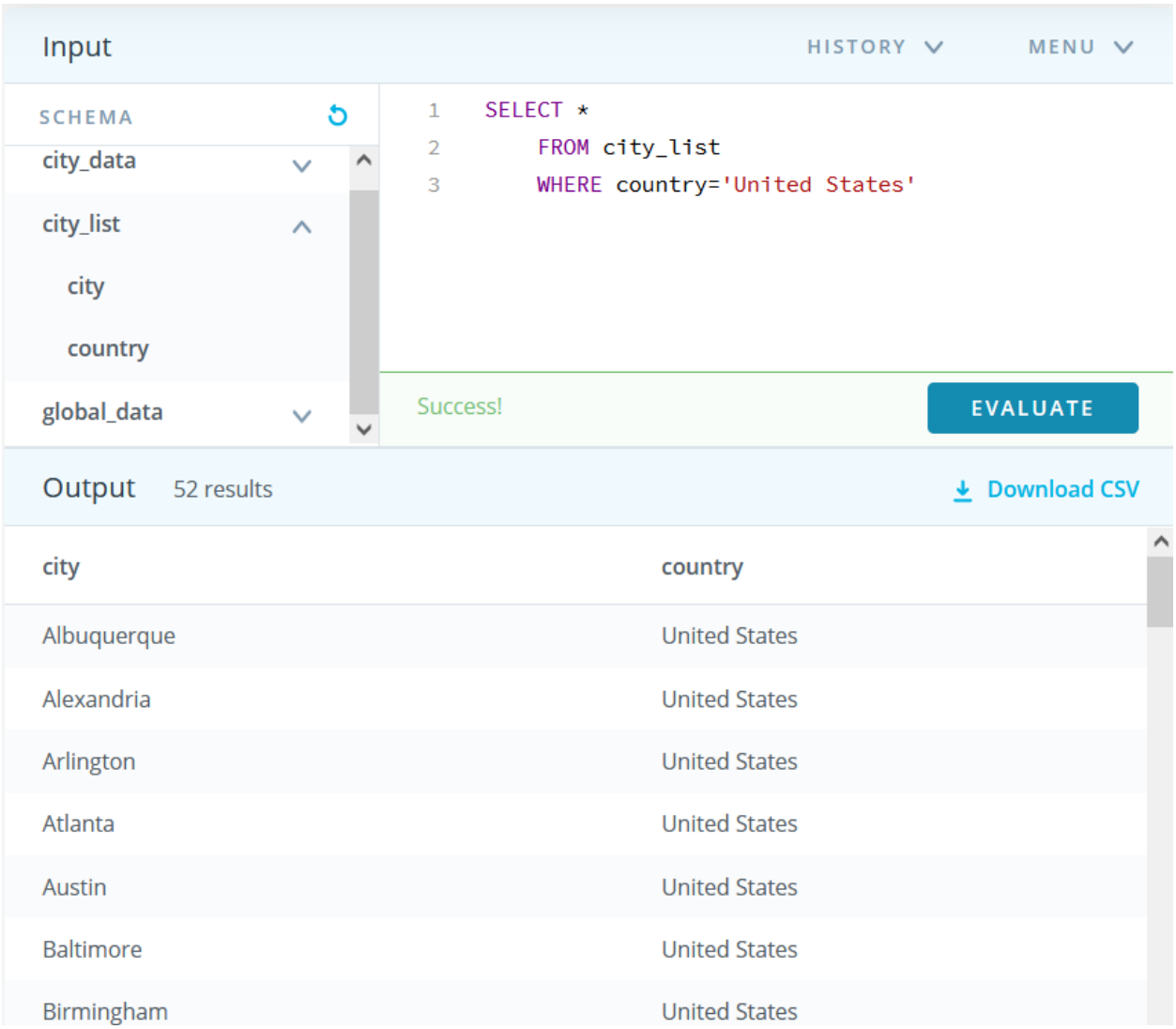


Figure 1.1: SQL Query city_list

Once I browsed the list and confirmed the most accurate reading for my local weather trends would be New York, I wrote a query which would compile a basic list from the data necessary for my local area, preorganized by column which I would find easiest to read once downloaded in .CSV format. (Figure 1.2)

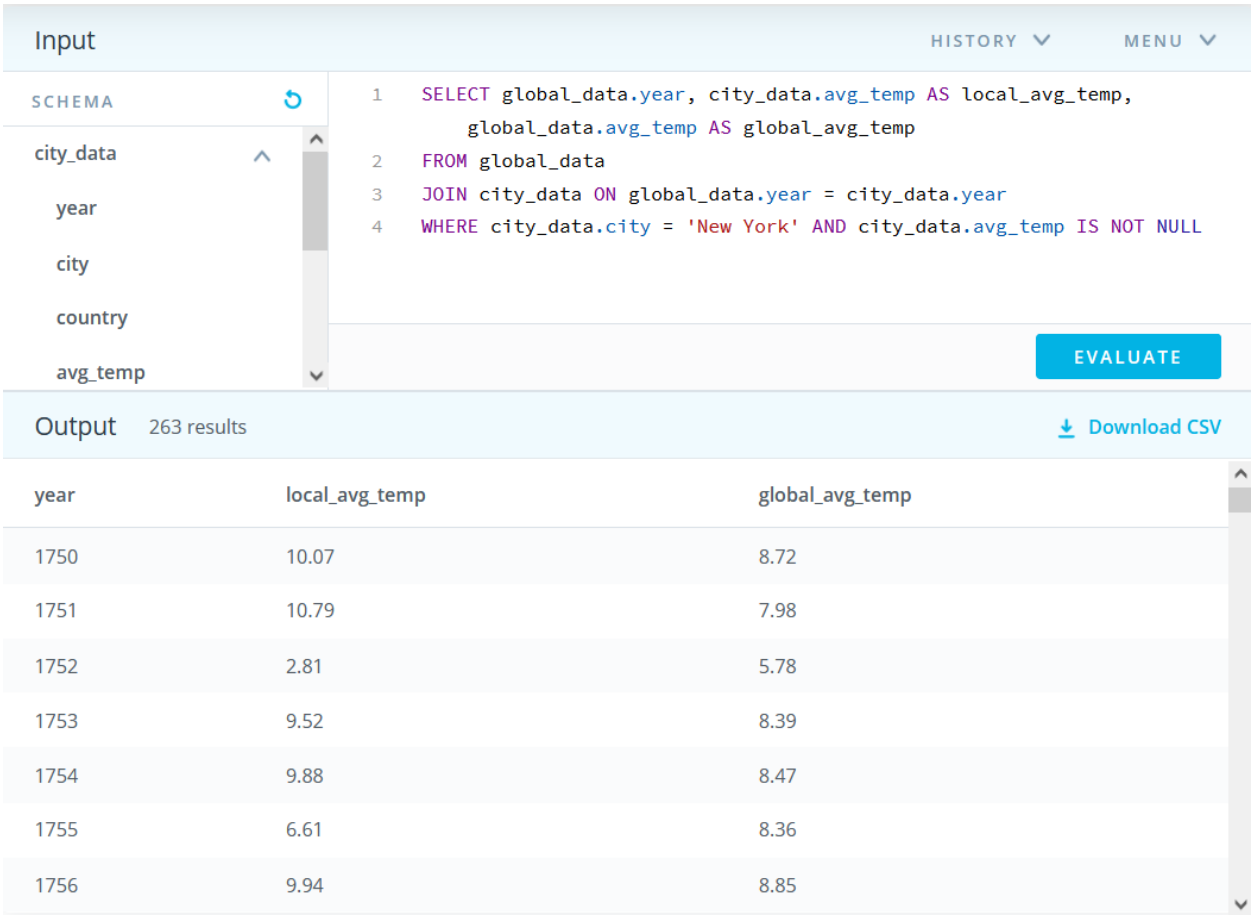


Figure 1.2: SQL Query city_data and global_data

The query request was made with the data organized so it would be easiest for me to read as a .CSV/Excel file, with the JOIN and IS NOT NULL conditions limiting the scope of data to only the years recorded for both local and global temperature readings.

Step 2: Reading the Data

With the data successfully queried and saved, I read it using Microsoft Excel. There I relabeled the column categories and calculated the moving average for both lists of temperature in a 10-year span using Excel's =AVERAGE(start : end) command. I decided 10 years to be a reasonable length of time to perform the moving averages for legibility purposes. (Figure 1.3)

	A	B	C	D	E
1	Year	New York Avg. Temp	10-Year Moving Avg. (NY)	Global Avg. Temp	10-Year Moving Avg. (Global)
2	1750	10.07		8.72	
3	1751	10.79		7.98	
4	1752	2.81		5.78	
5	1753	9.52		8.39	
6	1754	9.88		8.47	
7	1755	6.61		8.36	
8	1756	9.94		8.85	
9	1757	8.89		9.02	
10	1758	8.15		6.74	
11	1759	9.01	8.567	7.99	8.03
12	1760	7.73	8.333	7.19	7.877
13	1761	10.18	8.272	8.77	7.956
14	1762	9.55	8.946	8.61	8.239
15	1763	7.23	8.717	7.5	8.15
16	1764	9.55	8.684	8.4	8.143
17	1765	8.96	8.919	8.25	8.132
18	1766	10.09	8.934	8.41	8.088
19	1767	8.52	8.897	8.22	8.008
20	1768	8.67	8.949	6.78	8.012
21	1769	9.1	8.958	7.69	7.982
22	1770	9.04	9.089	7.69	8.032
23	1771	10.18	9.089	7.85	7.94
24	1772	9.64	9.098	8.19	7.898
25	1773	10.08	9.383	8.22	7.97
26	1774	9.52	9.38	8.77	8.007
27	1775	10.24	9.508	9.18	8.1
28	1776	9.13	9.412	8.3	8.089
29	1777	8.77	9.437	8.26	8.093

Figure 1.3: Record of Average Temperatures

Here I would like to make an additional note: The data retrieved had recorded yearly temperatures as early as 1750 until 2013. In that expansive list of data, I did identify that New York did not have a temperature reading for the year 1780. For accuracy I returned for an additional SQL query for that year's global reading and considered the missing New York reading to be negligible, for which I recorded a 9-year moving average from 1780 to 1789 rounded to three decimals. (Figure 1.4)

1780	N/A	8.3
1781	9.79	8.257
1782	9.15	8.202
1783	8.81	8.061
1784	8.4	7.937
1785	8.49	7.742
1786	8.98	7.726
1787	8.97	7.748
1788	9.77	8.068
1789	9.42	9.087

Figure 1.4: 9-Year Moving Average

Step 3: Presenting the Data

With the moving averages completed, I created a line chart to compare for any noticeable correlation. (Figure 1.5)

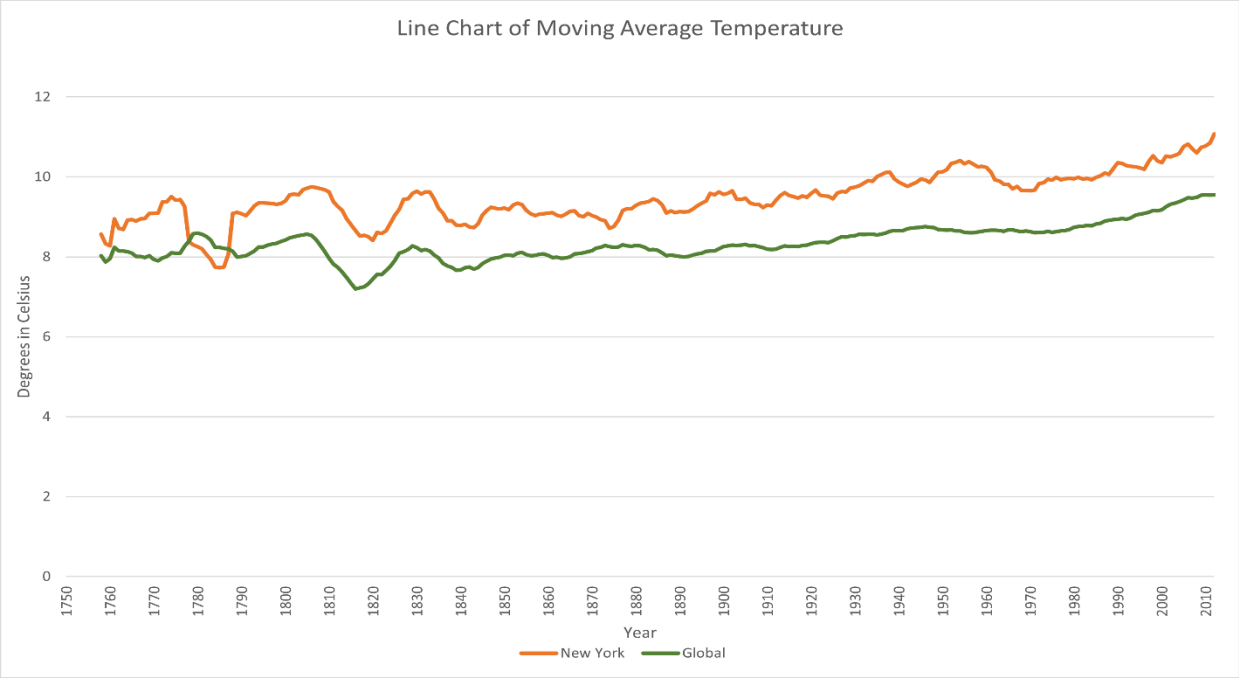


Figure 1.5: New York vs. Global Moving Average

Viewing the data in this way, I can make a few observations:

1. **1.)** The temperature in New York is warmer than the global average, with an overall average temperature of 9.465°C compared to the global average of 8.344°C. Additionally, both averages show an increasing temperature as we reach more recent years at a consistent rate.
2. **2.)** New York's temperature was lower than the global moving average from 1779 until 1788. While there is some concern that for all but the first year New York is missing the data from 1780 and therefore leaves a margin of error, it is most likely due to the area's record-low temperature reading of 0.25°C in 1779. Unfortunately, I could not find any outside sources regarding that timeframe.
3. **3.)** Utilizing the =VARP() function in Excel, we see that New York's moving average is more volatile with a variance of 0.364, almost twice the Global variance of 0.194. However, with exception of the timeframe from 1776 through 1789, the overall trend in temperature change is very similar. This is supported with a correlation coefficient (=CORREL(NY_Avg, GBL_Avg) on Excel) of 0.789.
4. **4.)** As a more personal observation, I was surprised to see little change in the average temperature readings for New York after the year 2001. Following the incident on September 11, and how conditions were for the city and how it has affected the health of those who have been in the area, I assumed an impact on its environment would show a more drastic change in the temperature readings for a few years at the very least.

This concludes my presentation of the weather trends of New York, and how they correlate to the global temperature readings.