Exploring Weather Trends - Project 1
Chelsey Dolan

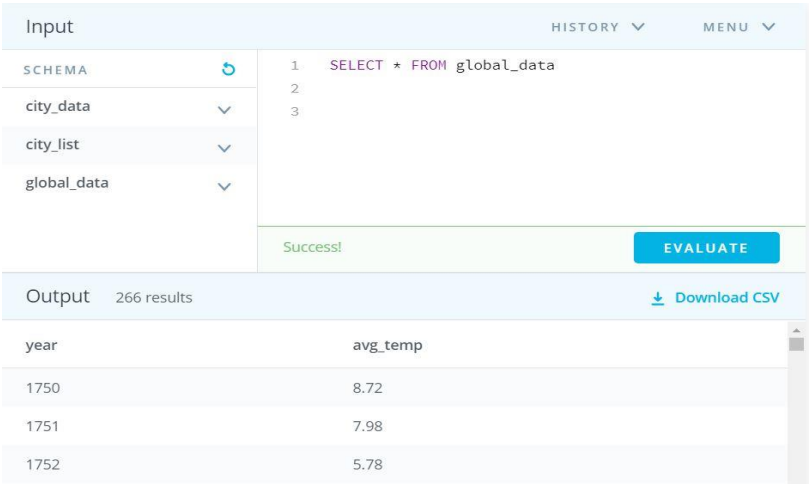
Overview

This is an analysis of local and global temperature data, comparing the temperature trends to my current city, San Diego, CA to overall global temperature trends. I chose to use Excel line charts for data visualization due to their simplicity and readability.

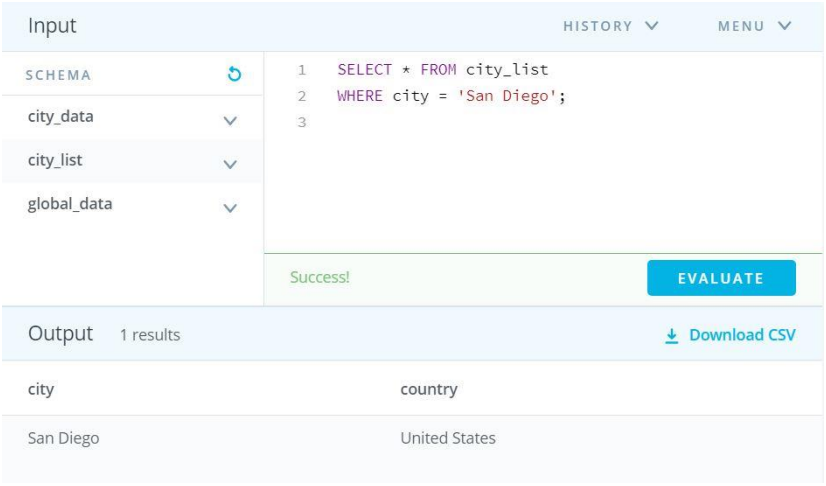
Collecting the Data

I used the SQL database in the Udacity module to first obtain my data. The following are the SQL queries I made to collect the data. All data was downloaded into separate CSV files.

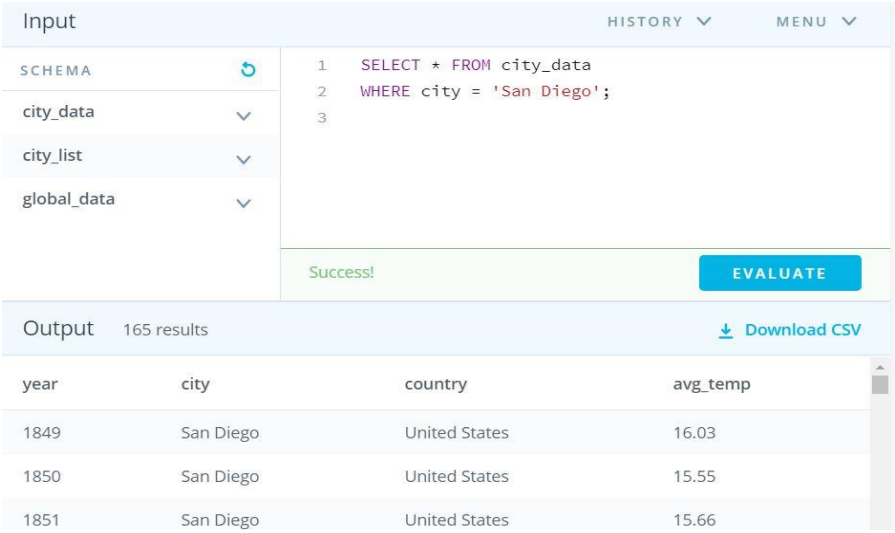
- Global Temperatures – there is data from 1750 to 2015.



- City List – success! San Diego is in the city list (otherwise I would use the nearest major city)



- City Data – the data for San Diego starts at 1849 and goes until 2013.



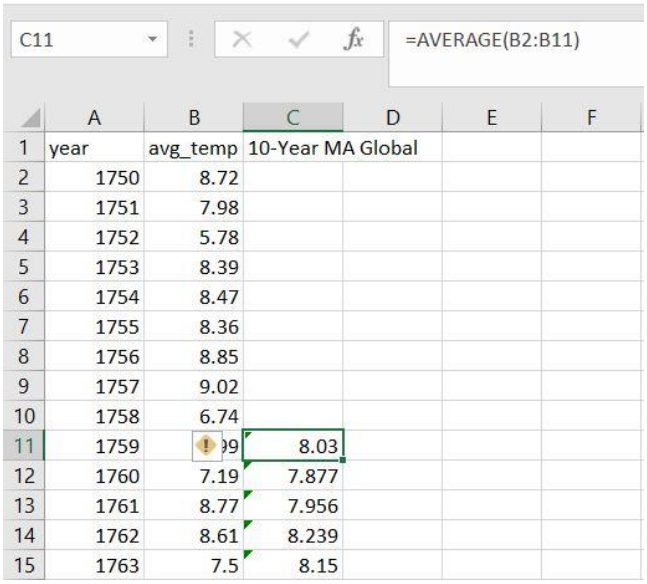
year	city	country	avg_temp
1849	San Diego	United States	16.03
1850	San Diego	United States	15.55
1851	San Diego	United States	15.66

Findings

I chose to visualize the datasets individually before visualizing them in a combined chart. While this is a personal preference (and not required for this assignment), I know that viewing the two data sets on their own may prevent assumption and bias from entering my analysis and afford the opportunity to look at the original data itself, before analysis comes into play.

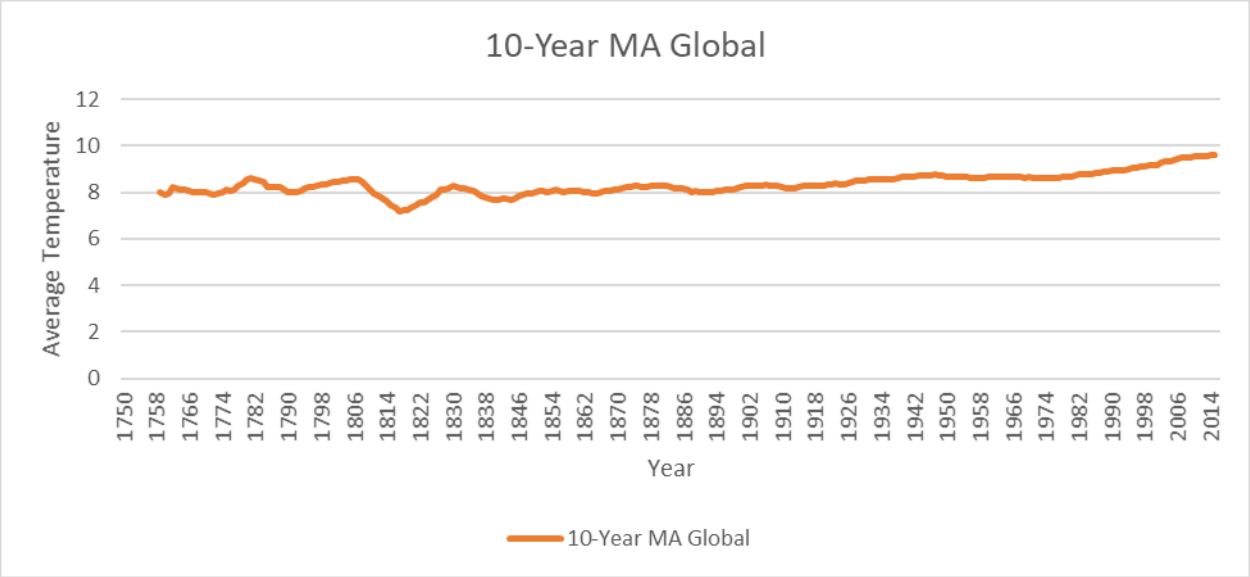
- 10-Year Moving Average – Global Temperatures in ºC

Excel spreadsheet using the AVERAGE formula for cells B2:B11 (10 years) and dragging it down the column.



	A	B	C	D	E	F
1	year	avg_temp	10-Year MA Global			
2	1750	8.72				
3	1751	7.98				
4	1752	5.78				
5	1753	8.39				
6	1754	8.47				
7	1755	8.36				
8	1756	8.85				
9	1757	9.02				
10	1758	6.74				
11	1759	9	8.03			
12	1760	7.19	7.877			
13	1761	8.77	7.956			
14	1762	8.61	8.239			
15	1763	7.5	8.15			

Once I had the CSV file downloaded onto my laptop, I immediately set to calculating the 10-year moving average. I chose 10 years because it provides a manageable way to read the data over the large period of time presented (276 years of data!). After that, I used the line chart functionality in Excel to create my line chart. I set the Year as the x-axis, and the Average Temperature as the y-axis.



This is quite a lot of data to make sense of at first glance. You can tell there is on overall upward trend in the average temperature, but there don't appear to be any major spikes in temperature, or major valleys of decreased temperature. I am not an environmental scientist, but I can tell you from looking at this data alone, the global temperature, on average, is getting warmer in general. Let's look at how San Diego, California, USA compares to the global average.

- 10-Year Moving Average – San Diego, CA Temperatures in ºC

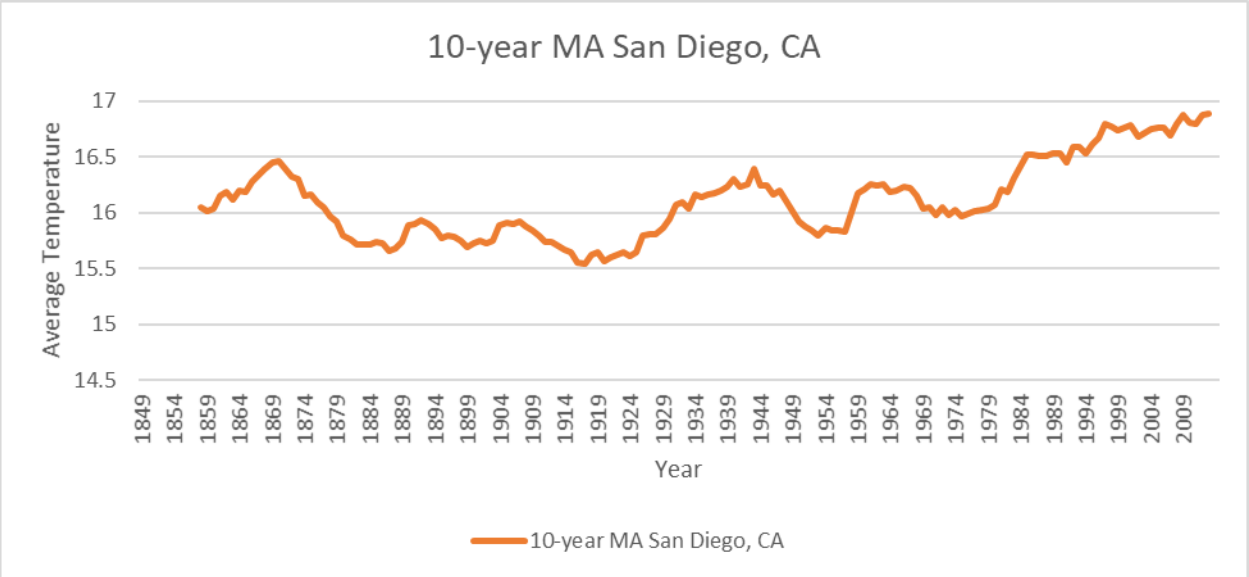
Excel spreadsheet using the AVERAGE formula for cells D2:D11 (10 years) and dragging it down the column.

The 20-year moving average was calculated using the AVERAGE formula for cells D1:21 (20 years).

E11 =AVERAGE(D2:D11)

	A	B	C	D	E	F	G	H
1	Year	city	country	Average Te	10-year M	20-YR MA		
2	1849	San Diego	United Sta	16.03				
3	1850	San Diego	United Sta	15.55				
4	1851	San Diego	United Sta	15.66				
5	1852	San Diego	United Sta	16.06				
6	1853	San Diego	United Sta	16.69				
7	1854	San Diego	United Sta	16.11				
8	1855	San Diego	United Sta	16.31				
9	1856	San Diego	United Sta	15.75				
10	1857	San Diego	United Sta	16.41				
11	1858	San Diego	United Sta	15.96	16.053			
12	1859	San Diego	United Sta	15.64	16.014			
13	1860	San Diego	United Sta	15.74	16.033			
14	1861	San Diego	United Sta	16.82	16.149			
15	1862	San Diego	United Sta	16.37	16.18			
16	1863	San Diego	United Sta	16.1	16.121			
17	1864	San Diego	United Sta	16.87	16.197			
18	1865	San Diego	United Sta	16.18	16.184			
19	1866	San Diego	United Sta	16.64	16.273			
20	1867	San Diego	United Sta	16.99	16.331			
21	1868	San Diego	United Sta	16.53	16.388	16.2205		

The San Diego dataset is much more varied than the global temperature dataset. The first observation I made after calculating the 10-year temperature average, and creating the line chart in Excel, was that the average temperature in San Diego was much higher, almost double, the global average. Yes, San Diego is close to the equator, so you'd expect the temperatures to be warmer than the global average. However, there is a sharp rise in temperatures beginning about 1984. The average temperatures rise sharply with only small dips every few years as the line rises towards 17 ºC.

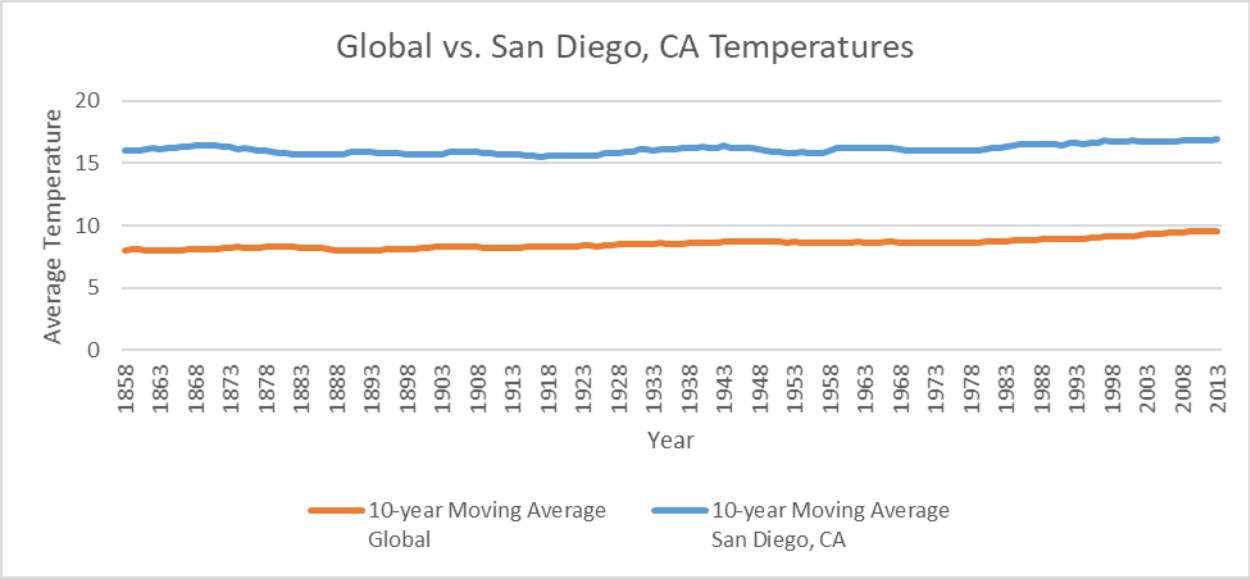


- Comparison of San Diego and Global Temperatures in ºC

  There are two major points that stand out when comparing the global temperatures with the San Diego temperatures: first, both the global and San Diego temperatures increase over time (even just slightly). Second, San Diego has a much higher average temperature than the global dataset.

  In the chart below, I only show the data that is relevant to both global and San Diego temperatures, starting with the first year the moving average was calculated for San Diego (1858). I also chose to compare the 10-year moving average for both datasets to retain continuity of scale in my data analysis. I did calculate a 20-year moving average for both global and San Diego data, but I felt the 10-year scale provided the level detail I desired to display.

  Of note, the line for San Diego looks much less varied when it's next to the global data. This is where I appreciated my choice to look at the datasets separately first because if I hadn't, I might not have seen the multiple variations in the San Diego dataset.



Observations

- San Diego is hotter than the global average by almost double. The overall average of global temperatures is 8.37 ºC, while the average temperature in San Diego is 16.13 ºC. San Diego is consistently hotter than the global average over time, never coming close to the global data.
- The changes in San Diego's temperatures over time are much more varied than the global temperatures. As discussed previously, there are many peaks and valleys of data, while the global data is less varied over time. This makes it difficult to see the overall increase in temperature for San Diego, especially when you observe the first peak in temperatures around 1869 and following decline.
- The overall trend for both global and San Diego temperatures is that they are getting hotter. The world is getting warmer, and San Diego is too (despite already being warm).
- The global data is much more consistent than the San Diego data. It raises many questions about how the data was collected in the first place – not to say it was collected, or even calculated, wrong, but I am curious to peel back the layers on the data collection efforts for both the global and city datasets. Readers may also question the value of presenting the global data because so many cities from around the world play a factor into this average. There are a multitude of cities that are far colder than the global average, and places much hotter than San Diego that are possibly skewing the dataset. Therefore, it is valuable to analyze individual cities, even multiple cities within the same country first, to get a better feel for the comparison between global and city average temperatures from a regional perspective.

Summary

In conclusion, this was a fantastic exercise in harnessing the simple power of data visualization. A line chart of the moving average in Excel provides an effective way to take a large dataset and break it down into a more manageable display for the reader. While the temperature trends do show that both global and San Diego, California, United States temperatures are getting warmer, readers can use this knowledge and data in meaningful ways as it pertains to their own city, and global temperatures.