Exploring Weather Trends

This is Project #1 from Udacity Data Analyst nanodegree. In this project, we analyze local and global temperature data and compare the temperature trends in Cairo, Egypt to the global temperature trends.

Data Extraction

The provided database has three tables: *city_list*, *city_data*, and *global_data*. city_list contains a list of cities along with their corresponding countries. city_data contains yearly average temperatures of a number of cities. Lastly, the global_data table contains a list of global yearly average temperatures.

To extract the relevant temperature data, we used SQL queries. Then, we saved the result sets into spreadsheets.

To get a list of cities, we ran this SQL query:

*SELECT * FROM city_list*

To get Cairo average temperatures (as per the assignment), we ran this SQL query:

*SELECT * FROM city_data WHERE city='Cairo'*

To get a list of average global temperatures, we ran this SQL query:

*SELECT * FROM global_data*

Note: As shown in the query below, an INNER JOIN between city_data and global_data, using the common column (i.e., year), could be used to substitute the last two queries with one single query, and generate a combined dataset at once.

```
SELECT c.year, c.country, c.city, c.avg_temp AS cairo_avg_temp,
g.avg_temp AS global_avg_temp
FROM city_data AS c
INNER JOIN global_data AS g
ON c.year = g.year
WHERE c.city = 'Cairo'
```

Processing Tools

As explained above, SQL was used to extract temperature data from the provided database. Microsoft Excel was used to process our data. Microsoft Excel offers great tools for data analysis including excellent data presentation, as well as a set of very useful tools and functions. The extracted two spreadsheets are then merged into a single file to calculate the moving averages and generate our chart.

Calculating Moving Averages

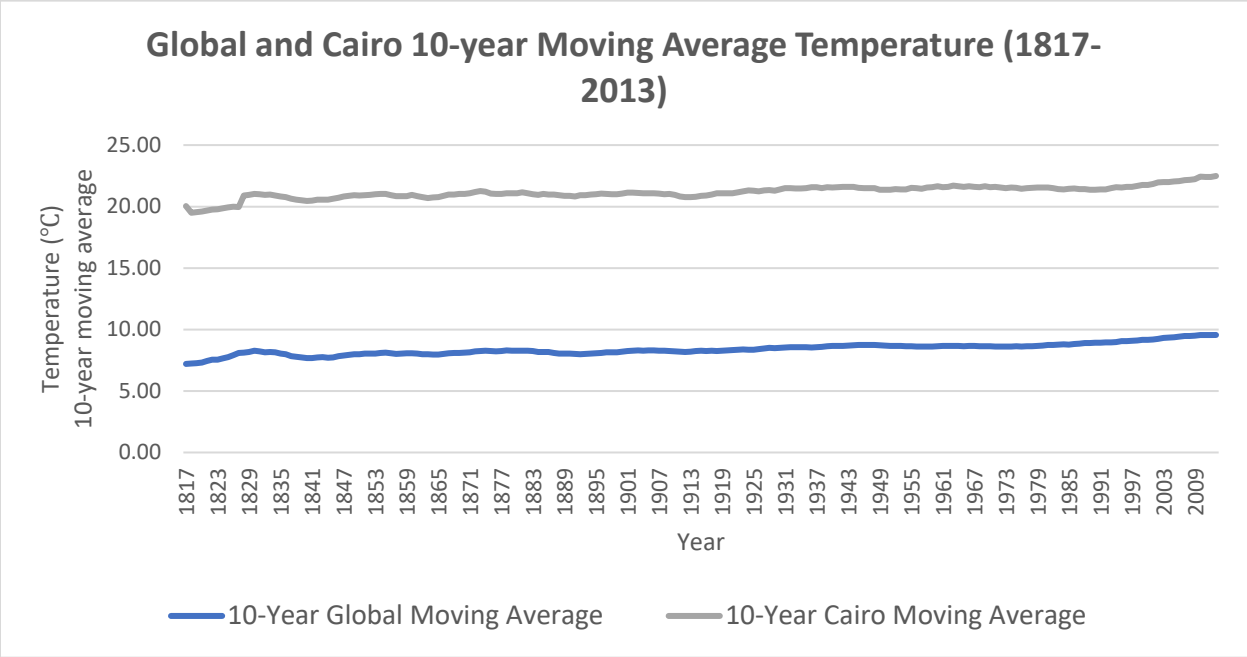
We started with 5-year moving averages to smooth our data, but then we settled for 10-year moving averages to improve the readability of our graph. The Average function in Microsoft Excel was used to calculate the moving city average temperatures, as well as the moving global average temperatures.

Below is a screenshot of part of the final dataset.

year	city_avg_temp	global_avg_temp	10_year_city_ma	10_year_global_ma
1808	17.11	7.63		
1809	19.87	7.08		
1810	19.93	6.92		
1811	20.00	6.86		
1812	19.93	7.05		
1813	20.51	7.74		
1814	20.43	7.59		
1815	20.30	7.24		
1816	20.51	6.94		
1817	21.88	6.98	20.05	7.20
1818	11.60	7.83	19.50	7.22
1819	20.31	7.37	19.54	7.25
1820	20.58	7.62	19.61	7.32
1821	20.63	8.09	19.67	7.45
1822	20.72	8.19	19.75	7.56
1823	20.71	7.72	19.77	7.56
1824	21.44	8.55	19.87	7.65
1825	21.00	8.39	19.94	7.77
1826	20.94	8.36	19.98	7.91
1827	21.63	8.81	19.96	8.09
1828	20.99	8.17	20.90	8.13
1829	20.91	7.94	20.96	8.18

Data Chart and Key Observations

Now we have prepared our data and calculated the moving averages for both Cairo temperature and the global temperatures, we can generate our chart.



From this chart, we can make some observations:

1. Cairo is way hotter than the average global temperatures, this has been consistent over the years.
2. The local temperatures seem to increase when the global temperatures increase, and drop when the global temperatures drop.
3. Although the trend has been generally consistent over the years, there are some anomalies in our data. For example, Cairo temperature dropped from 21.88 ℃ in 1817 to 11.60 ℃ in 1818. Maybe 1818 was an exceptional year in Cairo, or maybe it is just bad data.
4. There were periods where both local and global temperatures dropped a bit. For example, up until year 1826, the average global temperatures were below 8 ℃, then it hit 8 ℃ up until year 1836, after which it drops again below 8 ℃ up until year 1951.
5. From the numbers presented in our dataset, we can see that earth temperatures are increasing and it is getting hotter over time; both local temperatures and global temperatures have increased over 2 ℃ in the evaluated period.