Explore Weather Trends

OCTOBER 1

Udacity DAND Project
Authored by: Mahmoud Abdelgawad

Introduction

Problem Overview

We need to analyze the trends in global temperatures and compare that with the trends in local temperatures where I live.

What are the requirements?

Multiple steps need to be done to reach the final verdict:

1. Extract the data from the database
   We need to extract the global and local data from the database and present them in a suitable way to be dealt with. We'll need to export them as CSV files.
2. Open the CSV files
   We need to have a look at the data, we can use any program that can deal with CSV files, like Microsoft Excel, or open it using a programming language.
3. Draw the charts
   After we have the data ready to be analyzed, we'll need to get a better look at the data over time. For that we'll create a line chart illustrating the trends in both global and local temperatures.
4. Make observations
   Now we have our date ready and charted, we need to extract useful information from that and draw conclusions based on the charts.

Tools Used

- SQL – To extract the required data from the tables and export them to CSV
- Microsoft Excel – To open the CSV files and analyze the data by computing moving averages and drawing the charts

Practical Steps

We'll now start listing and explaining the steps taken to have our final output and the results.

Extracting the Data

The data are found in 3 SQL tables, the first contains the global average temperatures, the second contains a list of cities and their countries, and the third contains the average temperature data for each city.
Using simple SQL commands, we're easily able to extract the required data:

1. To extract the global temperature data, I used this command:

```
SELECT *
FROM global_data;
```

2. To search for a near local city, I used this command:

```
SELECT *
FROM city_data
WHERE country LIKE 'Egypt';
```

I found 2 cities within Egypt, so I chose Cairo, since it is my residence city.
The problem I faced here though, is that not all the range in the global data was available for my city, so I used this command to get the same period for both:

```
SELECT c.year,
  c.city,
  c.country,
  c.avg_temp AS local_avg_temp,
  g.avg_temp AS global_avg_temp
FROM city_data c
  JOIN global_data g ON g.year = c.year
WHERE c.city = 'Cairo' AND c.country = 'Egypt';
```

Exporting to CSV

After getting the query result, we use the option to export the result as a CSV file and save it locally.

Drawing the Charts

Opening the CSV file using Microsoft Excel gives us the data in the same view as we queried. To draw the charts, we first need to calculate the moving averages, and I chose to use a 10-year moving average calculation for this data.

I drew the charts for each average temperature separately at first, then drew the chart for both moving averages together to get a sense of the correlation between the two.

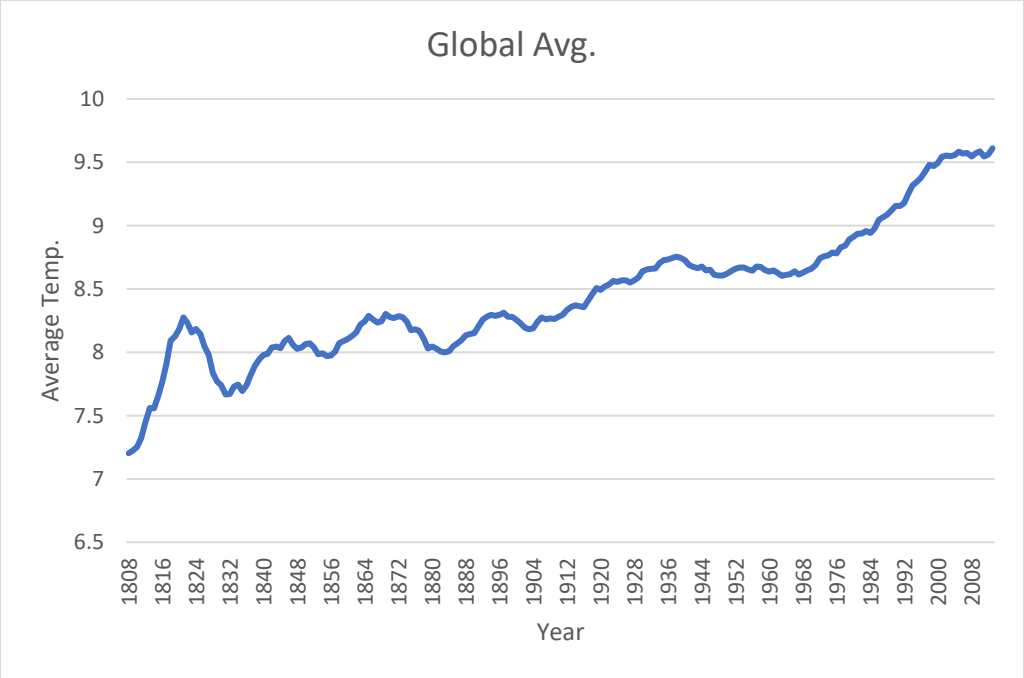

Figure 1. Global Moving Average Temperature

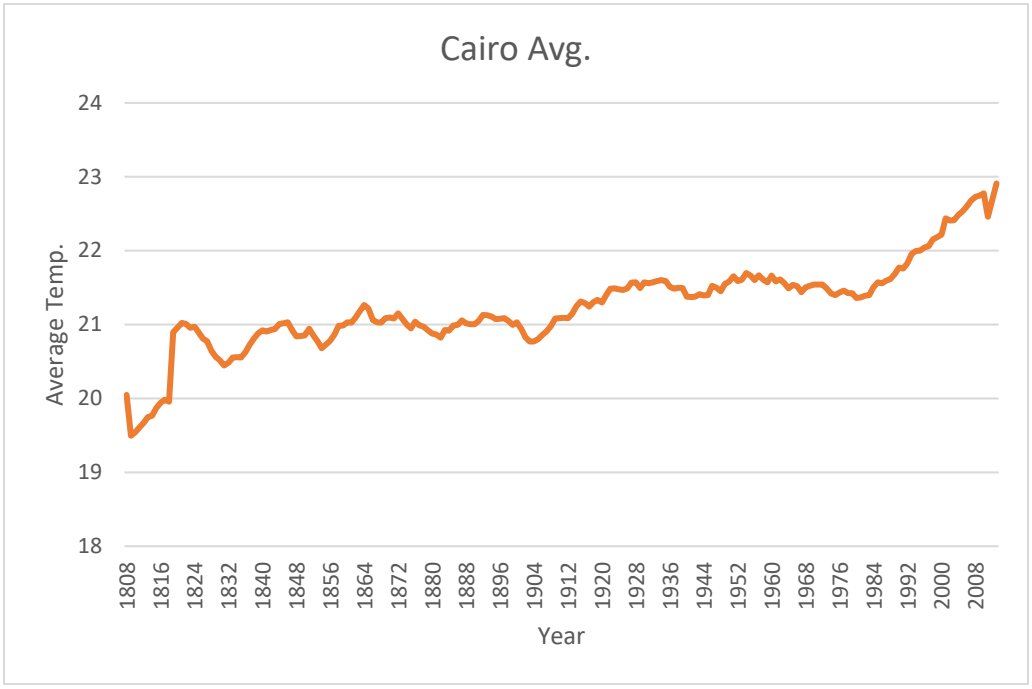

Figure 2. Cairo Moving Average Temperatures

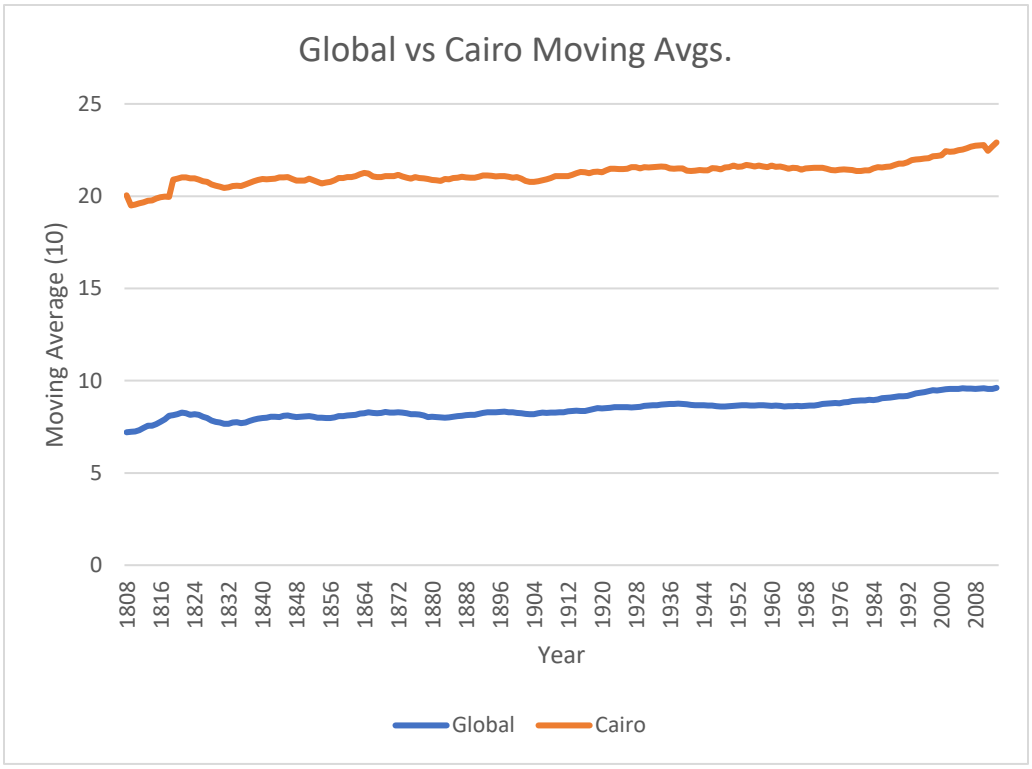

Figure 3. Global vs Cairo Moving Average Temperatures

Observations

From the individual averages we can observe the following:

- The global average temperatures range between ***7.203*** and ***9.585***.
- The local averages range between ***19.496*** and ***22.775***.
- It is apparent that the temperature, overall, increases steadily.
- At first the averages keep bouncing between decreasing and increasing, but since the **1980**s they have been increasing steadily.

From the combined graph, we can see the following observations:

- The average local temperature is much higher than the global average temperature.
- The decrease and/or increase in temperature is correlated, both global and local averages decrease or increase by almost same rate.

By calculating the correlation coefficient, using Excel's built in ***CORREL*** function, we can confirm the correlation between the two temperature averages.

The correlation coefficient between the averages is not a true indicator, since the averages are changing at non-constant rate, it equals ***0.58267***.
The true confirmation comes when computing the correlation coefficient for the moving averages, it gives a coefficient that equals ***0.946357***. We can see then that the moving averages are strongly correlated confirming the observation we see on the chart.

Expanded Results

Given the data we have, it's relevant to look at other nearby local cities to get more insight and compare the different results.
I've included the results for Alexandria, Egypt, in the chart below, and the observations are in line with what we got above.

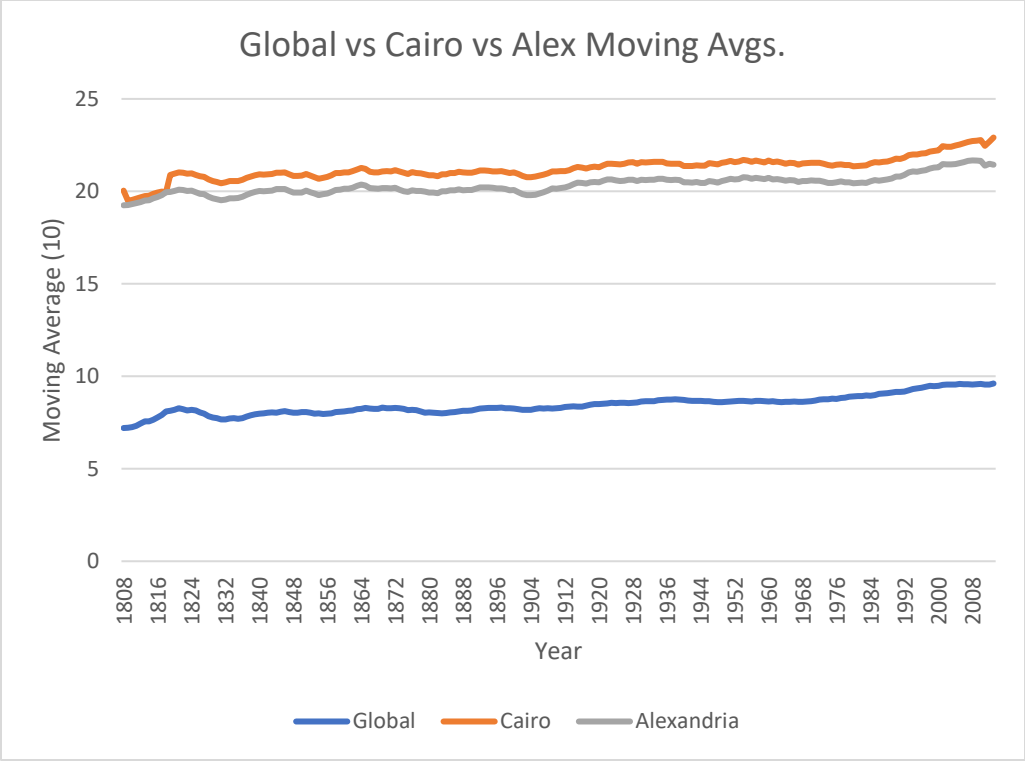

Figure 4. Global vs Cairo vs Alexandria Moving Average Temperatures

From the above chart, we can see the same observations apply to Alexandria. The average temperatures keep rising over the years, albeit the averages are less than Cairo's since Alexandria is a coastal city on the Mediterranean which helps keep its temperatures down.

Calculating the correlation coefficients also give almost the same results. The correlation coefficient between Global and Alexandria's moving averages is ***0.963164***, and between Cairo and Alexandria is ***0.968099***.

Conclusion

The average global -and local- temperatures have been rising since the **1830**s, which is not coincidentally after the industrial revolution. Recently, since the **1980**s the increase rate has become higher, all pointing to global warming being a serious problem that will soon endanger the life on the planet.