

Explore Weather Trends

Data Analyst Nanodegree Project #1

Date: 24 March, 2018
Revision: A
Prepared by: Khaled Khaled

1. Extract Data from Database

To start, I first need to find the city which is closest to where I live. In order to do that, I wrote an SQL query to retrieve the cities in the United Arab Emirates:

Input		HISTORY ▾	MENU ▾
SCHEMA	↻	<pre>1 SELECT * 2 FROM city_list 3 WHERE country = 'United Arab Emirates'</pre>	
city_data	▾		
city_list	▾		
global_data	▾		
		Success!	EVALUATE
Output		1 results	Download CSV
city		country	
Abu Dhabi		United Arab Emirates	

Figure 1: SQL Query (Local City)

Accordingly, one city appears for the country of the United Arab Emirates which is **Abu Dhabi**. Coincidentally, this is the city which I currently reside in and hence will be chosen as the local city to be compared with global data.

1.1 Extract Abu Dhabi Data

Now it is time to extract all the temperature data for the city of Abu Dhabi using SQL Query as follows:

Input		HISTORY ▾	MENU ▾
SCHEMA	↻	<pre>1 SELECT year, avg_temp 2 FROM city_data 3 WHERE city = 'Abu Dhabi'</pre>	
city_data	▾		
city_list	▾		
global_data	▾		
		Success!	EVALUATE
Output		171 results	Download CSV
year		avg_temp	
1843		26.04	
1844		26.26	
1845			
...			

Figure 2: SQL Query (Abu Dhabi Data)

This will return 171 results in total from the year 1843-2013.

1.2 Extract Global Data

Similarly, we use SQL Query to extract global data as follows:

The screenshot shows a web-based SQL query interface. At the top, there's a header with 'Input', 'HISTORY', and 'MENU' dropdowns. Below this, a 'SCHEMA' section lists databases: 'city_data', 'city_list', and 'global_data', each with a dropdown arrow. To the right, the SQL query is displayed:

```
1 SELECT *  
2 FROM global_data
```

. Below the query, a green 'Success!' message is shown next to an 'EVALUATE' button. The 'Output' section indicates '266 results' and includes a 'Download CSV' link. Below this, a table is shown with two columns: 'year' and 'avg_temp'. The table contains three rows of data: 1750 with avg_temp 8.72, 1751 with avg_temp 7.98, and a third row with a blank 'year' cell and a blank 'avg_temp' cell.

year	avg_temp
1750	8.72
1751	7.98

Figure 3: SQL Query (Global Data)

This on the other hand returns 266 results from the year 1750-2015.

2. Data Selection & Manipulation

2.1 Data Selection

All data from previous exercise has been extracted as CSV files and imported into Excel for further evaluation. As mentioned in the previous section, there were more results in the global database as data was available for a larger range of years.

Furthermore, while extracting the 171 results from the Abu Dhabi Data, it seems that the results have returned null values as shown in figure 2 above.

In order to provide a more accurate and useful dataset for comparison, I proceeded with choosing the range of common years only (1843-2013). Furthermore, dataset for range of years 1843-1860 was excluded due to large sum of null values for these years. Hence, comparison shall be made for the data in years ranging from **1861-2013**.

2.2 Data Manipulation

Due to fluctuations in yearly averages, it is best to evaluate the data considering the moving average to provide smoother results during data visualization.

The moving average has been prepared on the same excel spreadsheet on a 7-year basis. This is done by calculating the average temperature for the first 7 years (1861-1867). The same is then repeated from years 1862-1868 and so forth.

The same procedure is performed for both the local city data and global data to obtain data which will be used for data visualization.

3. Data Visualization

I am now able to plot a line chart to show a comparison between the local city average temperature and global average temperature. This is done by plotting the moving average temperatures on the y-axis and the year range on the x-axis. After inputting the corresponding values on excel, the following line chart is generated.

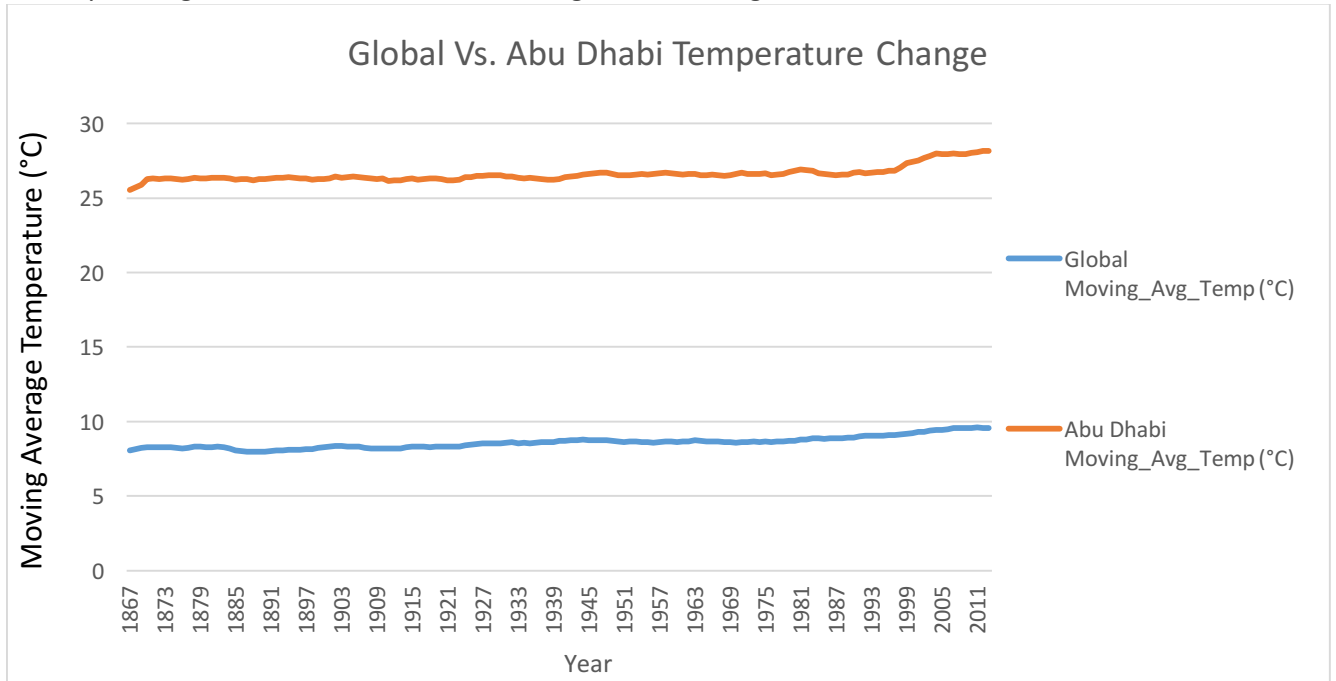


Figure 4: Global Vs. Abu Dhabi Temperature Range (7 Years Moving Average)

Another chart is plotted to show the global average temperature range across all years as shown in figure 5 below. This elaborates a clearer trend in the global average temperature over a wide-spread range.

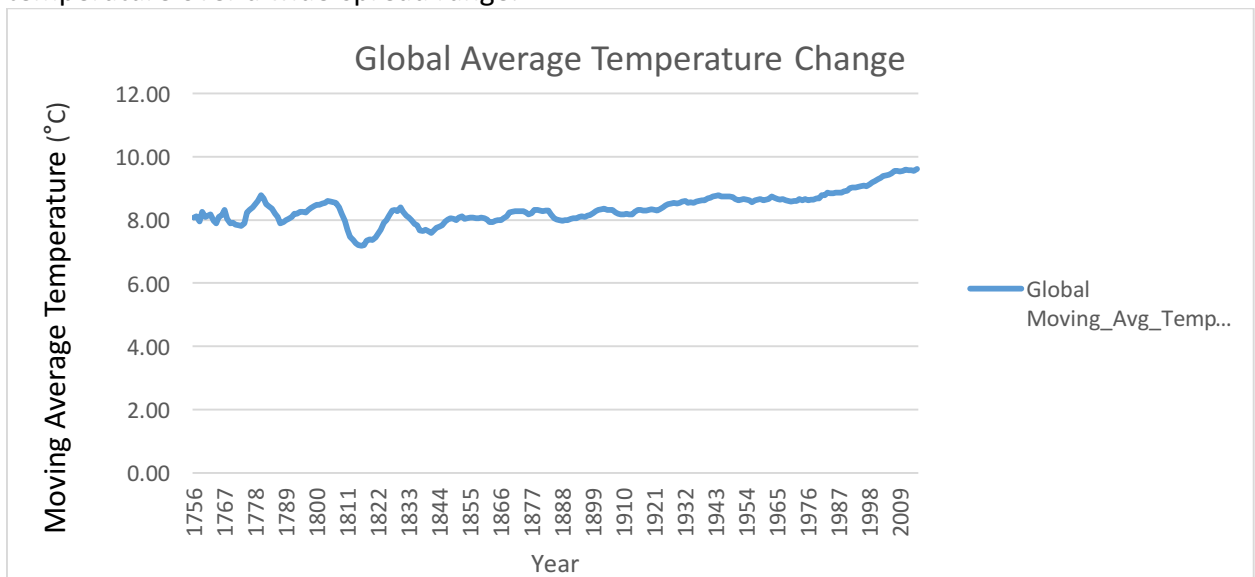


Figure 5: Global Average Temperature Range (7 Years Moving Average)

4. Observations

According to line charts, the following observations may be deduced:

- 1) Abu Dhabi's weather is much warmer than the global average considering that the temperature has always been greater in the past couple hundred years.
- 2) In both cases, we can see that the average temperature is gradually increasing throughout the entire time-frame.
- 3) A significant rise in the yearly average temperature can be observed for Abu Dhabi in the past couple of decades starting from the year 1994 upwards.
- 4) The yearly average temperature seems to be increasing abnormally on a global scale in the last 3-4 decades. The same may be noticed when looking at the wide-range chart in figure 5 where it is evident that the global average temperature had fluctuated around the 8 °C mark up until approximately 1890. From then on, we notice an uptrend in the temperature rise.