

Explore Weather Trends

Data Analyst Nanodegree Project #1

Date : 18-05-2020

Prepared by: Akash Yadav

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 India :

The screenshot shows a web-based SQL query interface. The 'Input' section on the left lists a schema with tables: city_data, city_list, and global_data. The main query area contains the following SQL code:

```
SELECT *  
FROM city_data  
WHERE city = 'Hyderabad' AND country = 'India';
```

Below the query, a green 'Success!' message and a blue 'EVALUATE' button are visible. The 'Output' section shows '218 results' and a 'Download CSV' link. The output is displayed as a table with the following columns: year, city, country, and avg_temp. The first row of data is:

| year | city | country | avg_temp |
|------|-----------|---------|----------|
| 1796 | Hyderabad | India | 26.53 |

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.

This will return 218 results in total from the year 1796-2013.

1.2 Extract Global Data

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

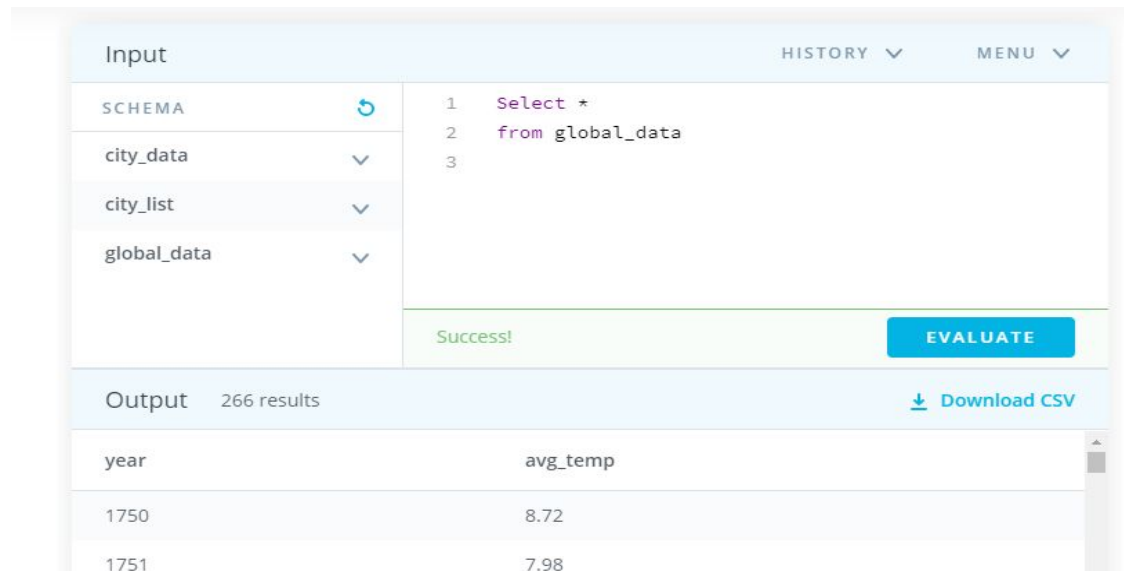


Figure 2: SQL Query (Global Data)

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 218 results from the Hyderabad , 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 (1796-2013). Furthermore, dataset for range of years 1796-1860 was excluded due to the large sum of null values for these years. Hence, comparison shall be made for the data in years ranging from 1796-2013.

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 Jupyter notebook .This is done by calculating the average temperature .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 jupyter notebook , the following line chart is generated.

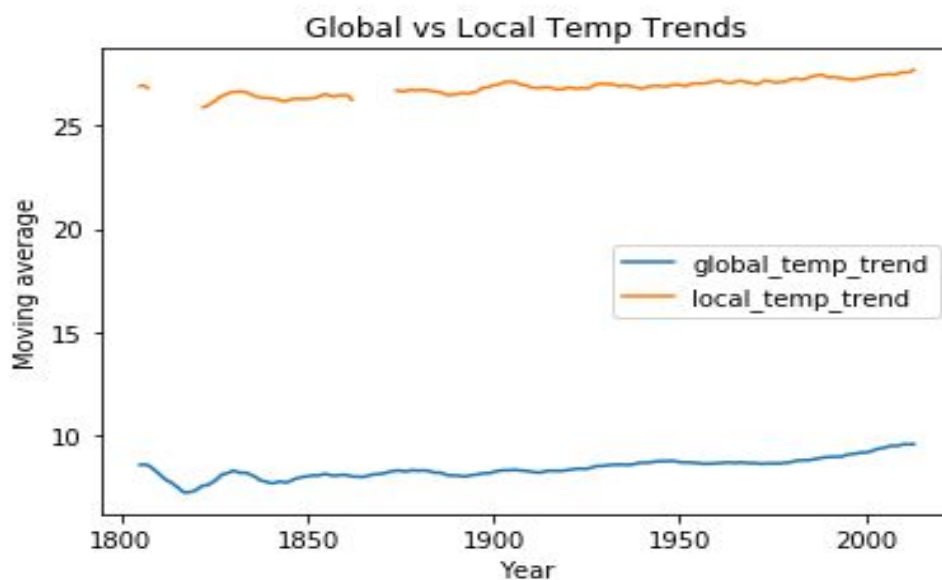


Figure 3: global _temp_trend vs local_temp_trend

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.

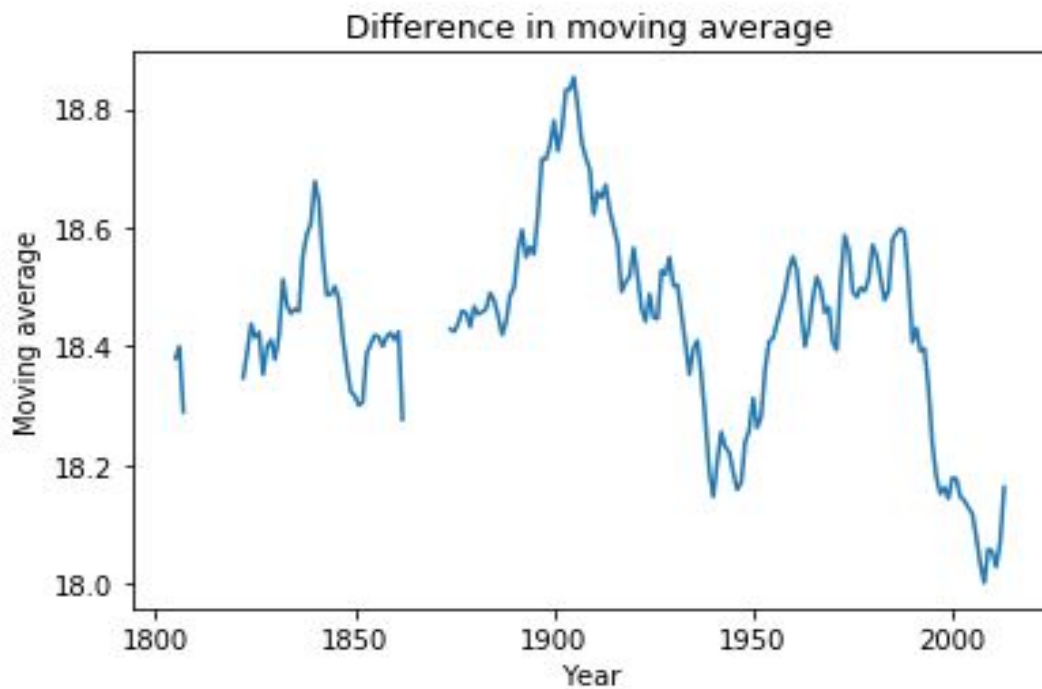


Figure 4: Global Average Temperature Range

Observations

1) On average, the temperature for local (Hyderabad) is higher compared to the global, which means the city I live in is hotter compared to the global.

a) And this trend is consistent over time which can be seen for years 1800s - 2000s

2) At first glance from the green line in above visual, which indicates the difference in moving average for local and global temperature shows that the rate of change of temperature for Hyderabad compared with global is almost the same for all years. But when you dig in, to see what's happening with difference in moving averages closely. It fluctuates between 18°C and 19°C . Observing closely reveals that from the 1900s there was a drop in the rate of change of temperature of Hyderabad compared with global.

3) Looking at the overall trend it looks like the world is getting hotter because the curves keep on rising as years progress.

a) And this trend has been consistent for the last few hundred years.

4) Looking at the visual it is found that most of the countries are cooler as compared to where I live, that is because the region I live in comes under the tropical region which occupies a small region on Earth as compared.