UDACITY DATA ANALYST NANODEGREE PROGRAM

EXPLORING WEATHER TRENDS PROJECT

BY

ADETOLA HABEEB ADELASE

17 OCTOBER 2021

In this project, I analyzed local and global temperature data and compared the temperature trends of the nearest city where I live (London) to/with the overall global temperature trends. The objective(s) of this project is to create a visualization and note the similarities and differences between global temperature trends and temperature trends of the nearest city where I live (London). The various steps and process in achieving this are highlighted below.

- **EXTRACTING THE DATA**

- I wrote a SQL query to extract the **global temperature data** and exported to CSV on my local PC. This is shown in the figure 1 below:

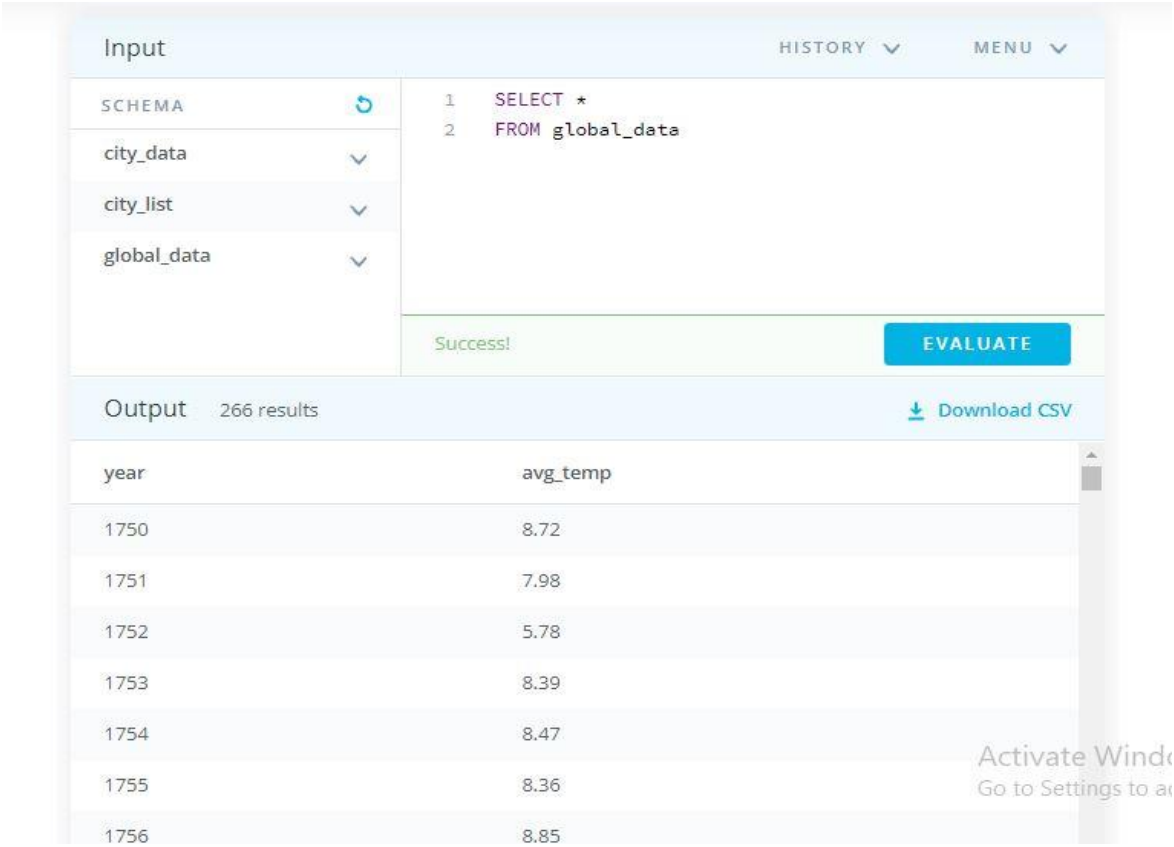

Figure 1 : SQL query and result. (Global Temperature Data)

The query returned an output of 266 results. Average Temperature ranging from the year 1750 to 2015.

- For the **local temperature data**, I ran an SQL query to see the closest city near me which is London and I noticed there was two cities named London, one in United Kingdom and the other one being in Canada as shown in figure 2.

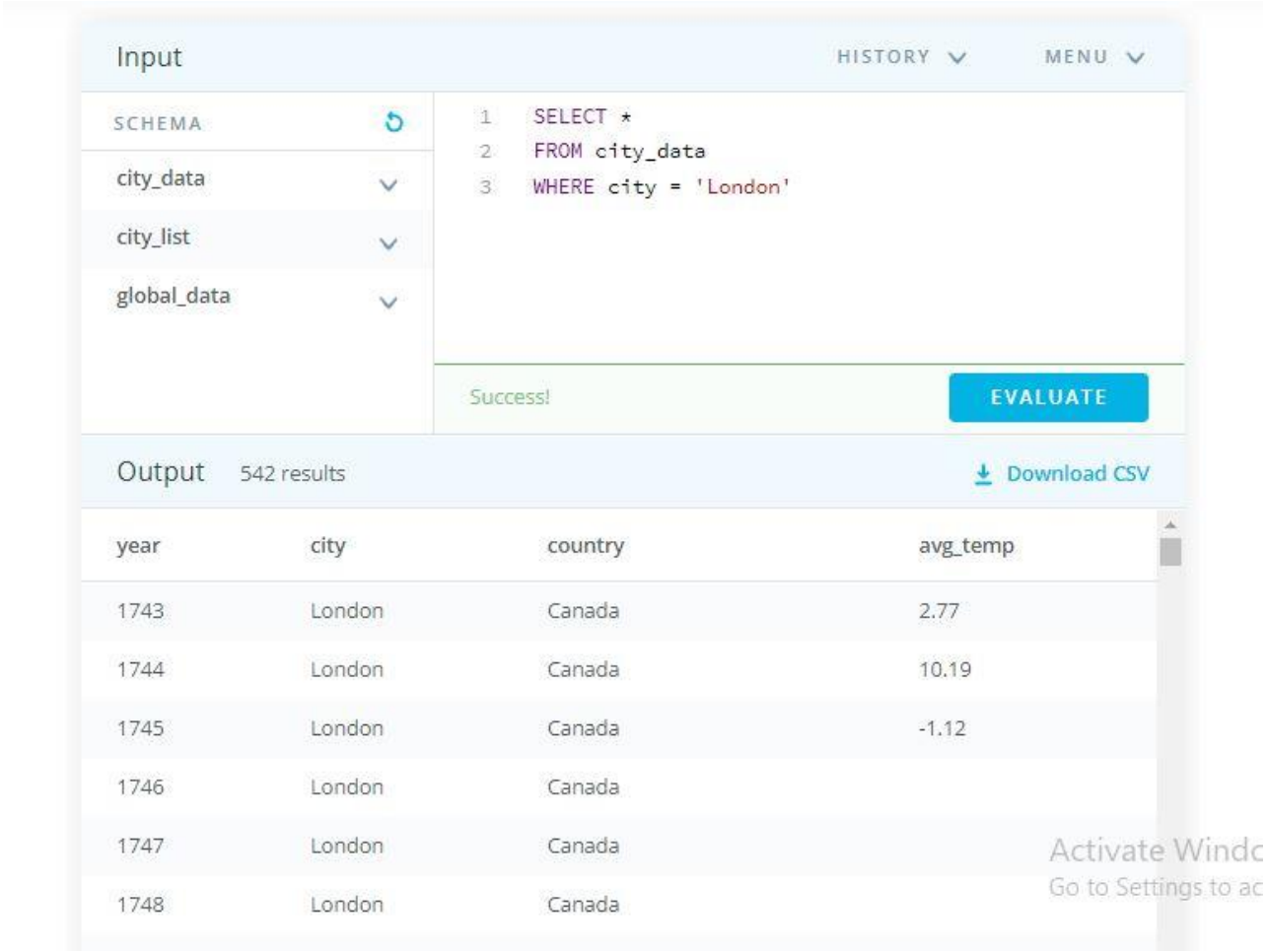

Figure 2 : SQL query and result showing London city in Canada. (Local Temperature Data)

The query returned an output of 542 results. Ranging from the year 1743 to 2015 for the two countries Canada and United Kingdom.

- After the above process, I ran a SQL query to extract all the temperature data only for the city of London (United Kingdom) and exported the data into CSV on my local PC, which is shown below in figure 3:

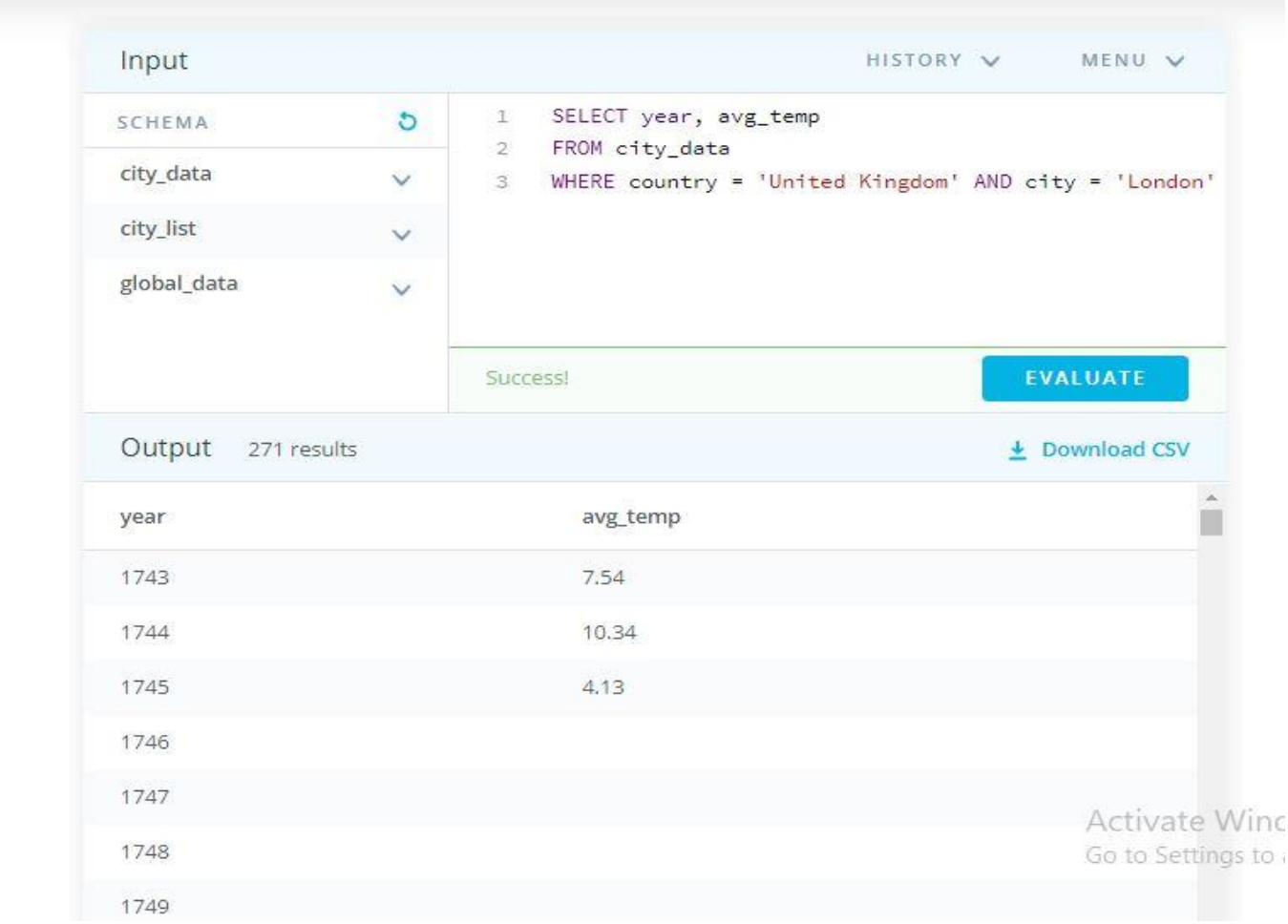

Figure 3 : SQL query and result. (London(UK) Local Temperature Data)

The query returned an output of 271 results. Average Temperature values ranging from the year 1743 to 2015 with some returning null values.

- **DATA EXTRATION AND ANALYSIS**

The local and global temperature data exported to my local PC was extracted and analyzed using Jupyter Notebook with the aid of Python packages such as Pandas, NumPy and Matplotlib.

During the data analysis I noticed there were more data ranging in more years in the global temperature data compared to the local temperature data and the there were null values in the local data temperature extracted as shown in figure 3.

- **DATA MANIPULATION AND VISUALIZATION**

Since I would be comparing my local city's temperatures (London) with the global temperatures. I plotted the ***moving average*** of the data rather than the yearly averages to smooth out the data and line and make the trend more observable and to not get lost in daily fluctuation.

Moving average, also called a rolling or running average, is used to analyze the time-series data by calculating averages of different subsets of the complete dataset. Since it involves taking the average of the dataset over time, it is also called a moving mean (MM) or rolling mean.

The moving average in the project was calculated using the Pandas package function *rolling().mean()*. ***rolling(window_size*)** function is used to get data containing each window and **mean()** is to find the average of each window. The parameter window_size was set to 20. Which means that our moving average runs over 20 rows — in this case, 20 years.

The whole process was also carried out using Jupyter Notebook with the Python packages (Pandas and NumPy) mentioned above. I calculated the moving average for both the local and global data and plotted the line chart (which are shown below) for the last 100, 50 10 years and all the years in the data for better comparison on both data. I plotted the year range on the X-axis and moving average temperatures on the Y-axis and.

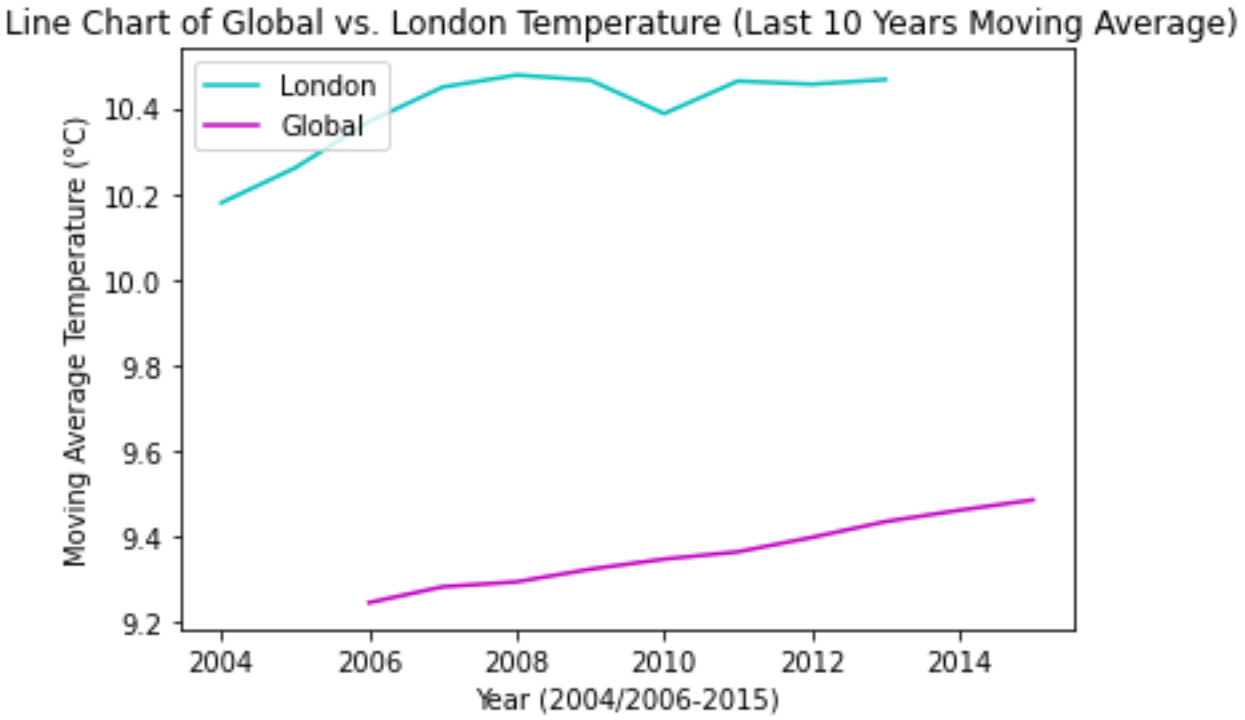

Figure 4: Last 10 years moving average chart. (London vs Global)

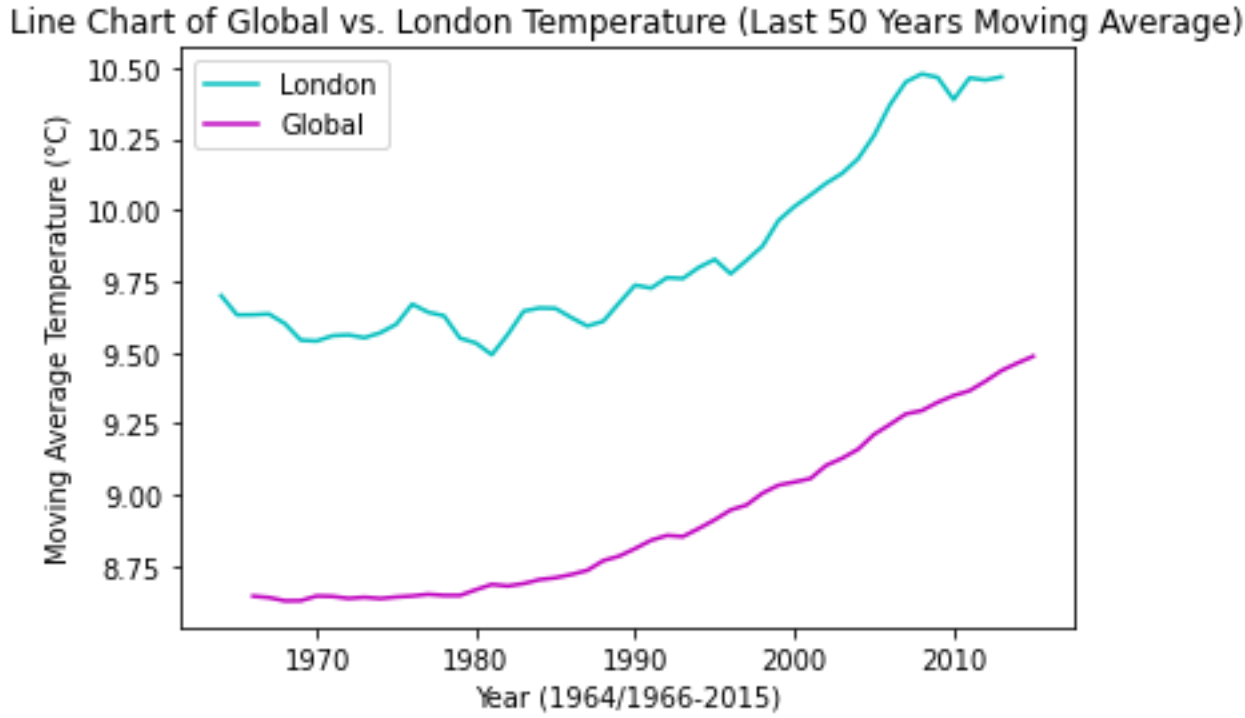

Figure 5: : Last 50 years moving average chart. (London vs Global)

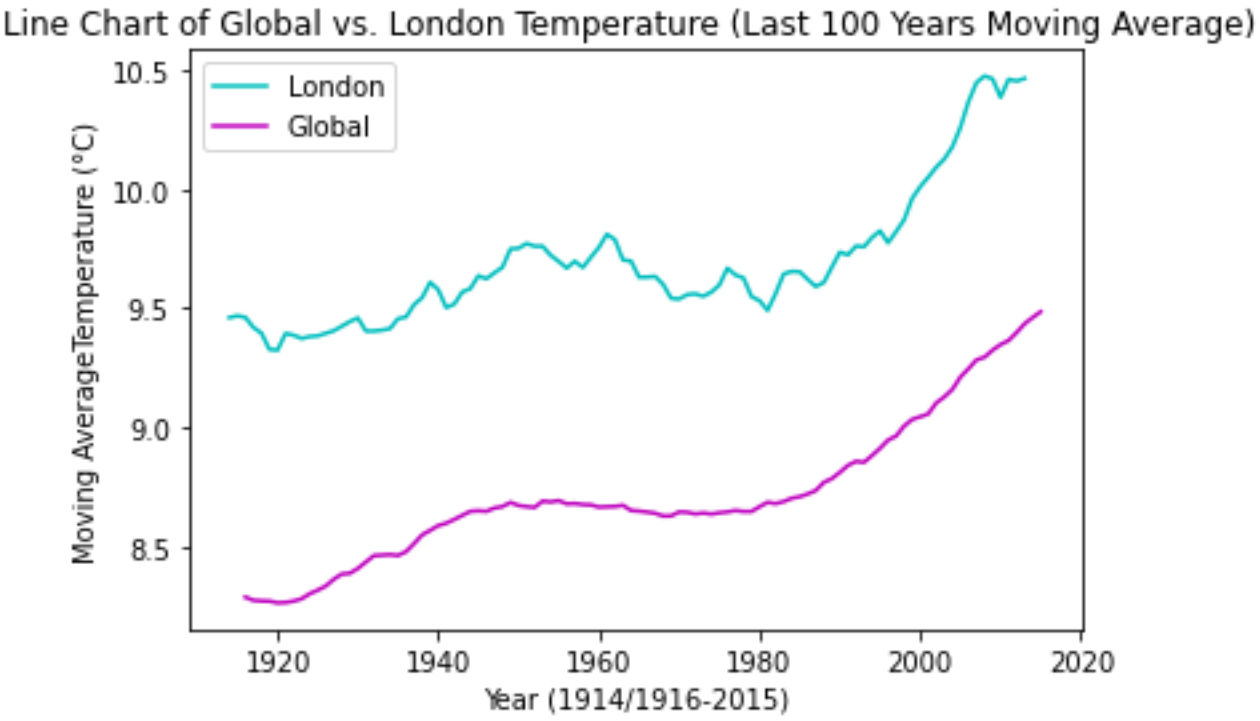

Figure 6: : Last 100 years moving average chart. (London vs Global)

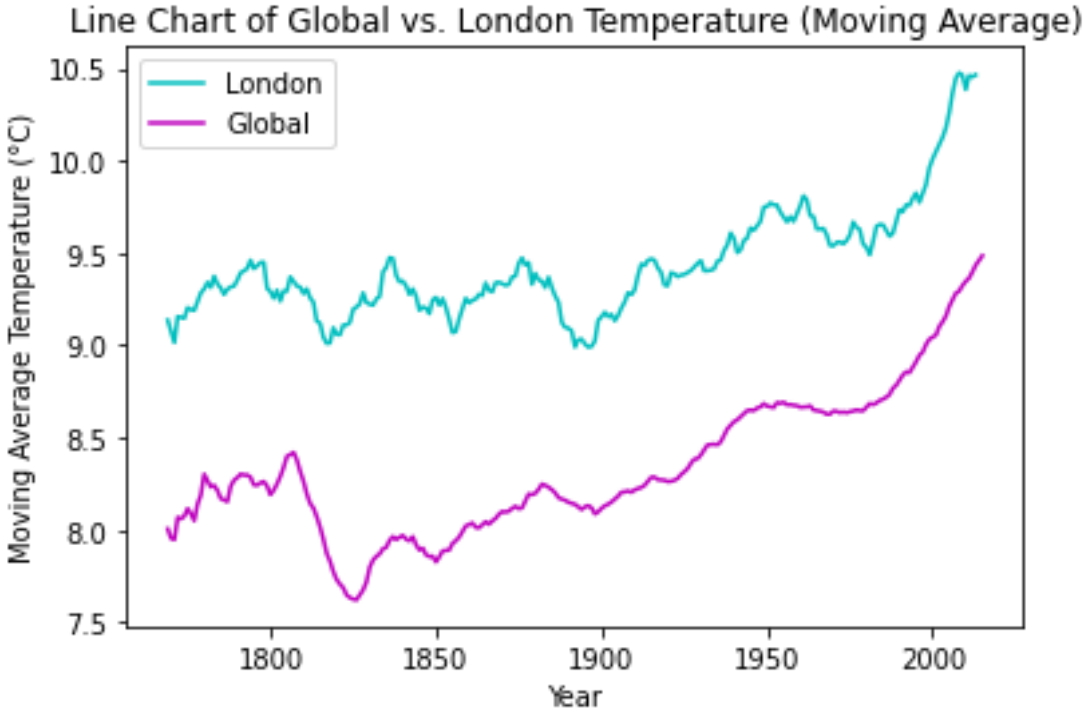

Figure 7: : Moving average chart across all the years in the data. (London vs Global)

- **OBSERVATION**

From the above line charts, here are some of my observations:

1. With London's average temperature average being 9.4359 compared to global temperature's average of 8.3695. London is hotter on average compared to the global average. The difference has been consistent over the years but the in last couple of years (1990 -2015), global average temperature made a huge jump.

2. Over the years, the changes in London's average temperature have been a bit inconsistent, as it fluctuates a lot before the steady increase in recent years (2000 -2015) compared to global average which as be gradually increasing over the years aside from the early year's (1820 – 1905) fluctuation.

3. From the average trend of the global and London temperature, the observation is the world is getting hotter and it's going to get hotter as rise in temperature as be consistent for the past 50 years.

4. There was a massive drop in the average of early years of the global temperature which is eerily similar to the sudden increase in recent years.