

**Udacity-data analyst nanodegree**

**Project 1**

**Topic:- Explore weather trends**

**-By SOHINI CHAUDHURI**

PROBLEM STATEMENT:- In this project, you will analyze local and global temperature data and compare the temperature trends where you live to overall global temperature trends.

INSTRUCTIONS:-

Your goal will be to create a visualization and prepare a write up describing the similarities and differences between global temperature trends and temperature trends in the closest big city to where you live.

**Steps:-**

* **Extract the data** from the database and write a SQL query to extract the city level data and the global data and export to CSV.
* **Open up the CSV** in whatever tool (I will use Python).
* **Create a line chart** that compares your city’s temperatures with the global temperatures. Make sure to plot the moving average.
* **Make observations** about the similarities and differences between the world averages and your city’s averages, as well as overall trends.

OBJECTIVE:-

* **Extraction of the data from the database provided.**
* **Chart visualization based on extracted data.**
* **The observations gathered based on the chart made.**

TOOLS USED:-

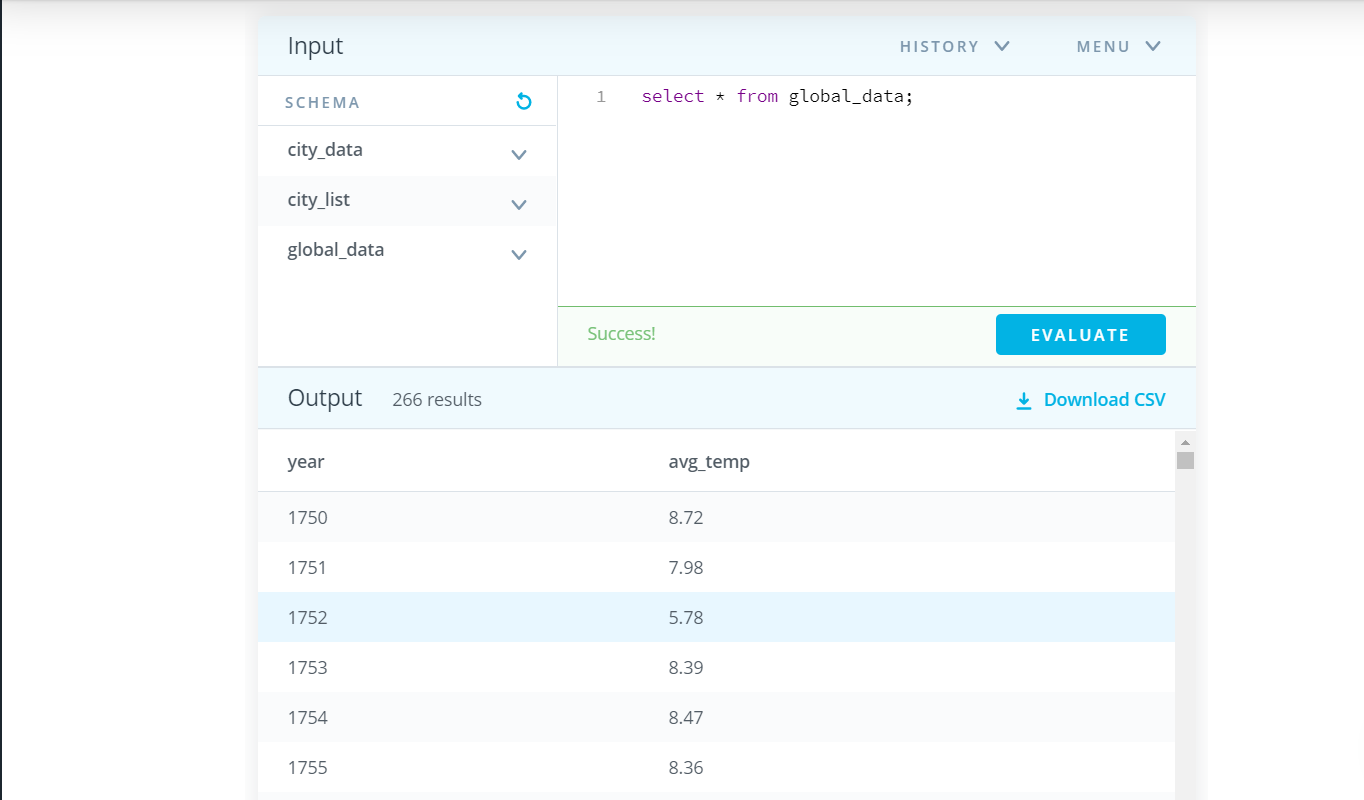
* **SQL: For the extraction of the data from the database.**
* **Python(Using Jupyter Notebook):**

**Pandas: For manipulation of the data and filtering them to plot the graphs.**

**Matplotlib: For plotting the graphs**

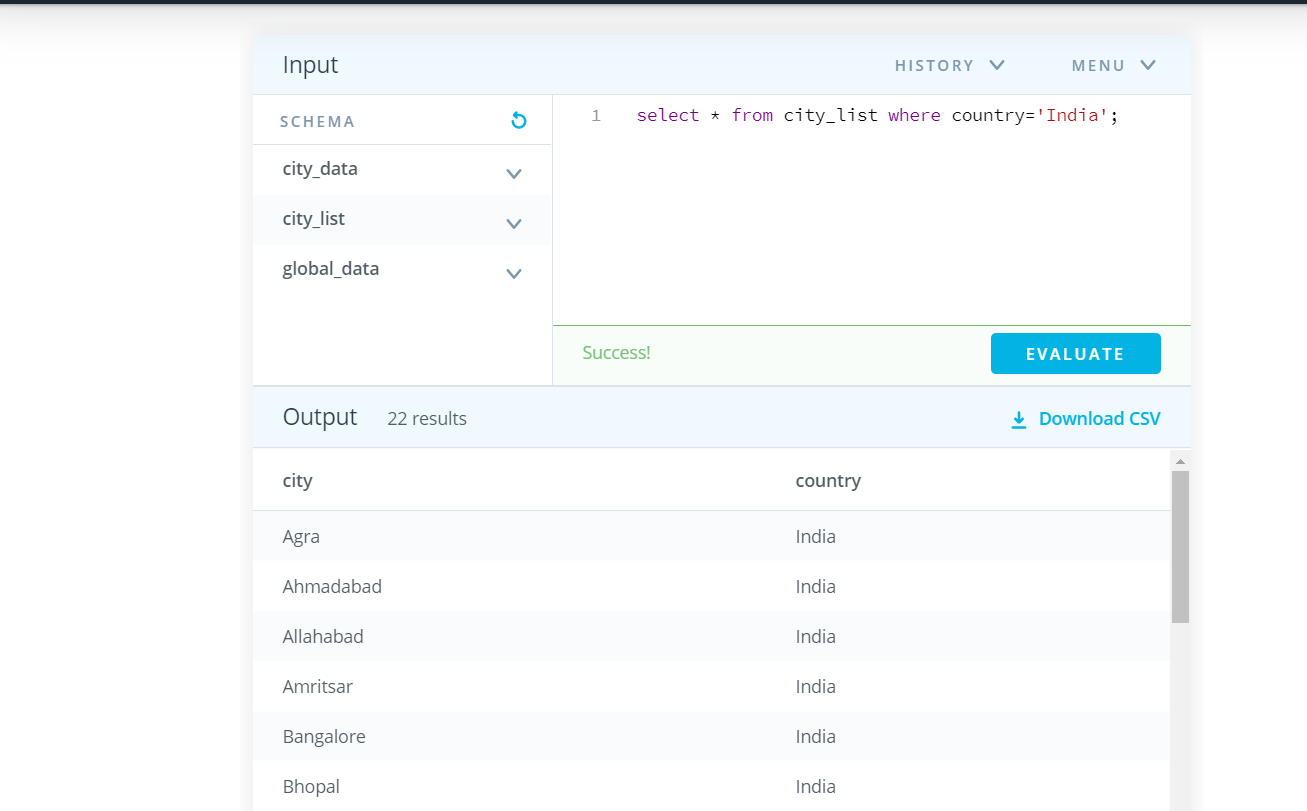
PROCEDURE:-

Step 1:- Extraction of the global data from the database provided to us



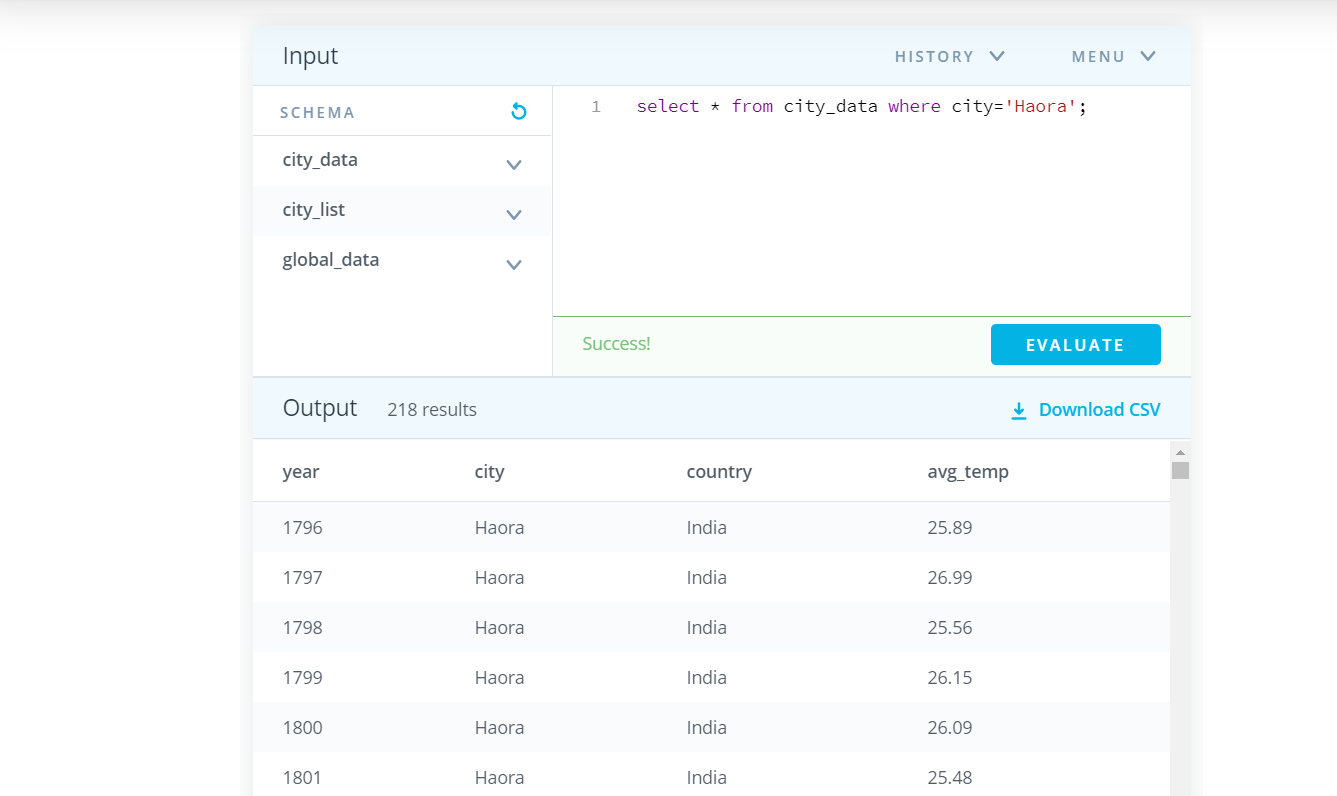
It produced 266 results for the global data.

Step 2:- Extraction of the country data from the database provided to us



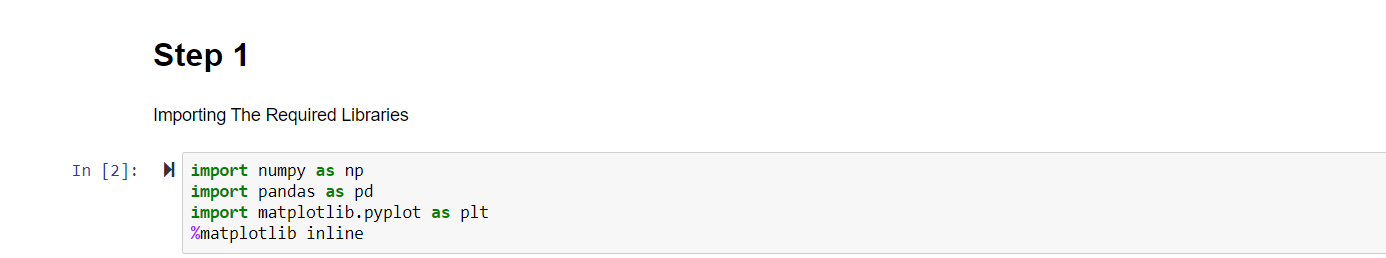
It produced 22 results for the country data.

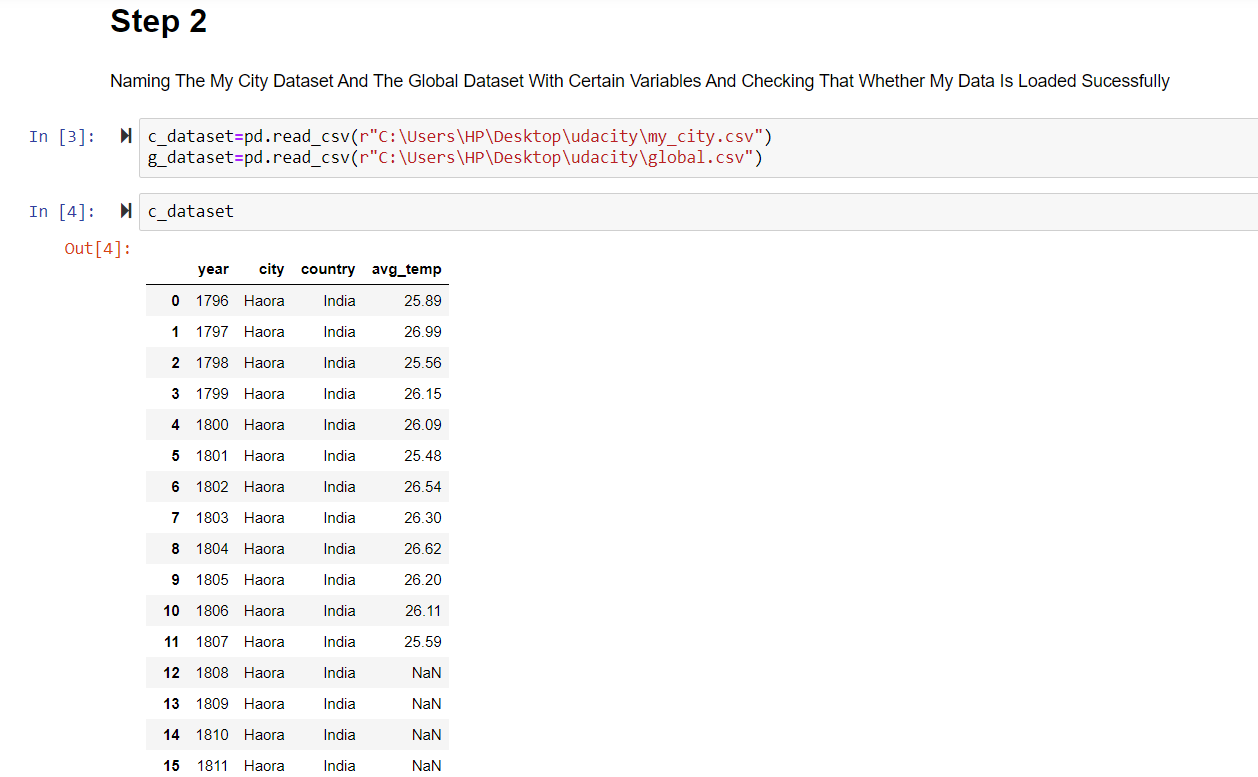
Step 3:- Extraction of the city data from the database provided to us



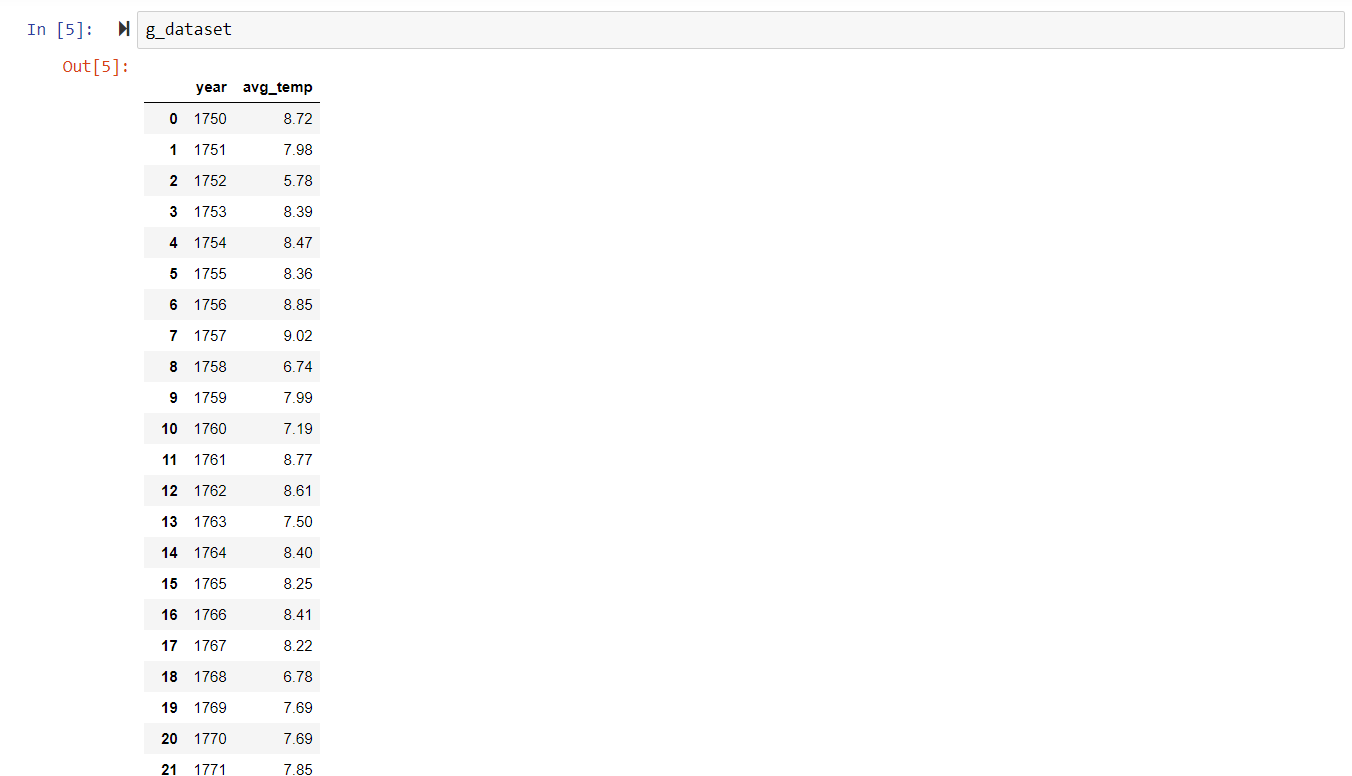
It produced 218 results for the city data.

Step 4:- Processing Of The Data Using Python Libraries With The Help Of Jupyter Notebook

STEPS ARE AS FOLLOWS:- 



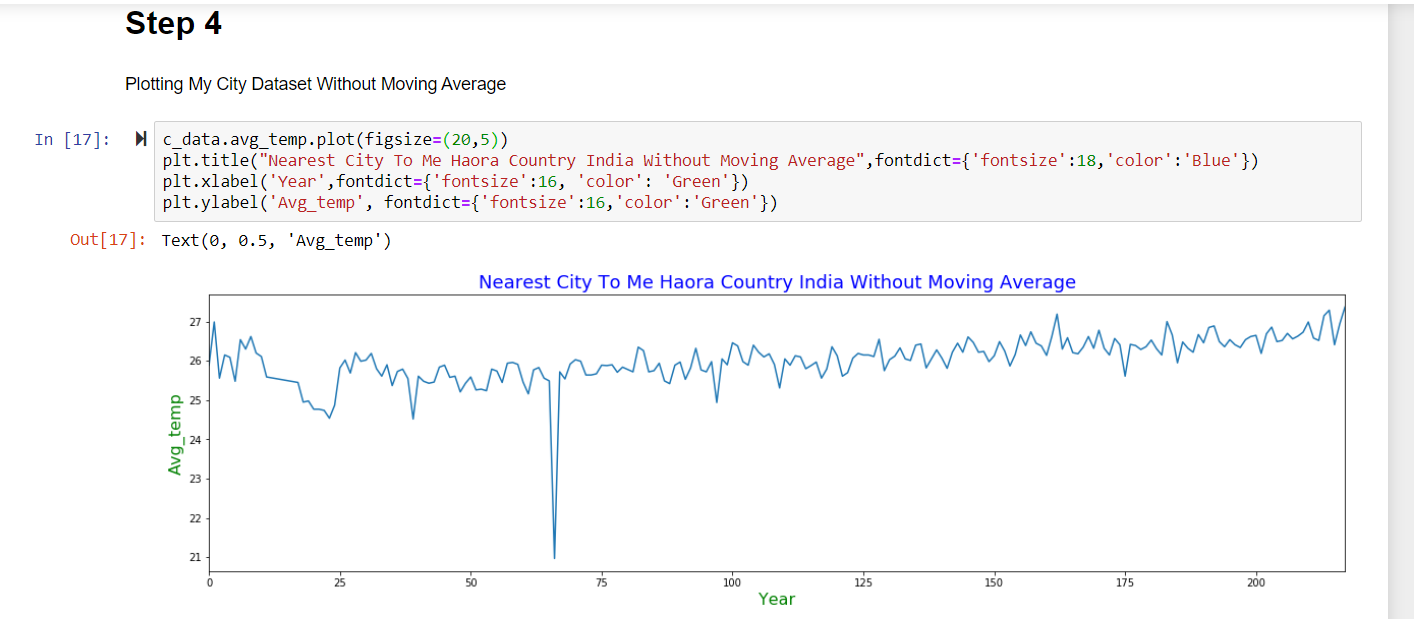
218 rows × 4 columns were produced.

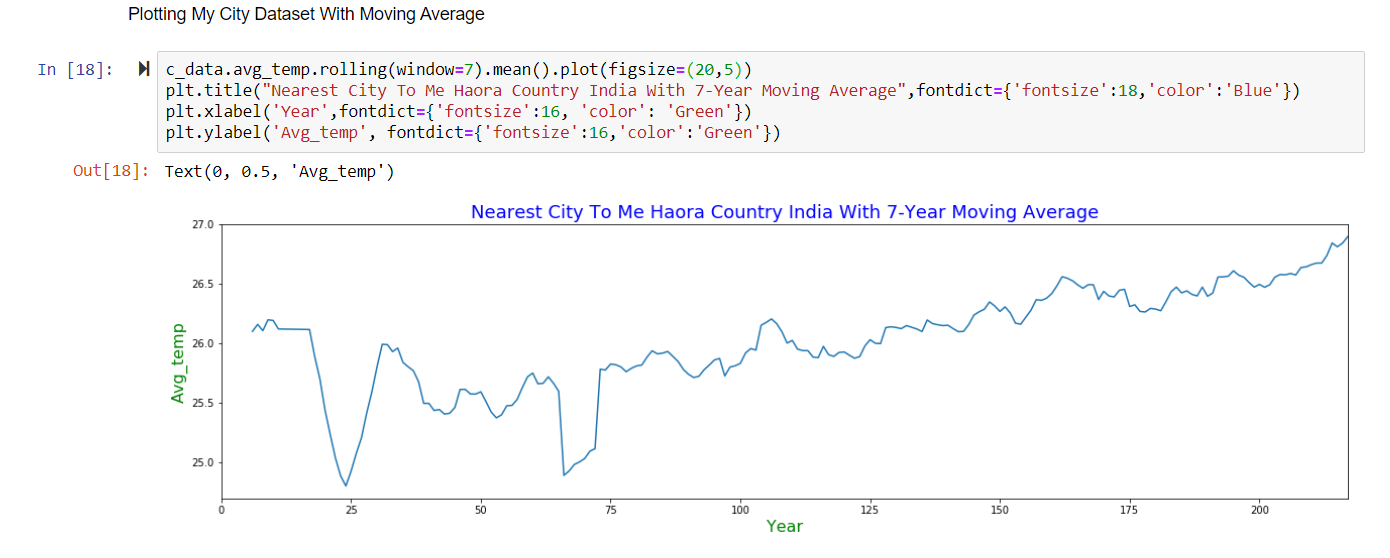


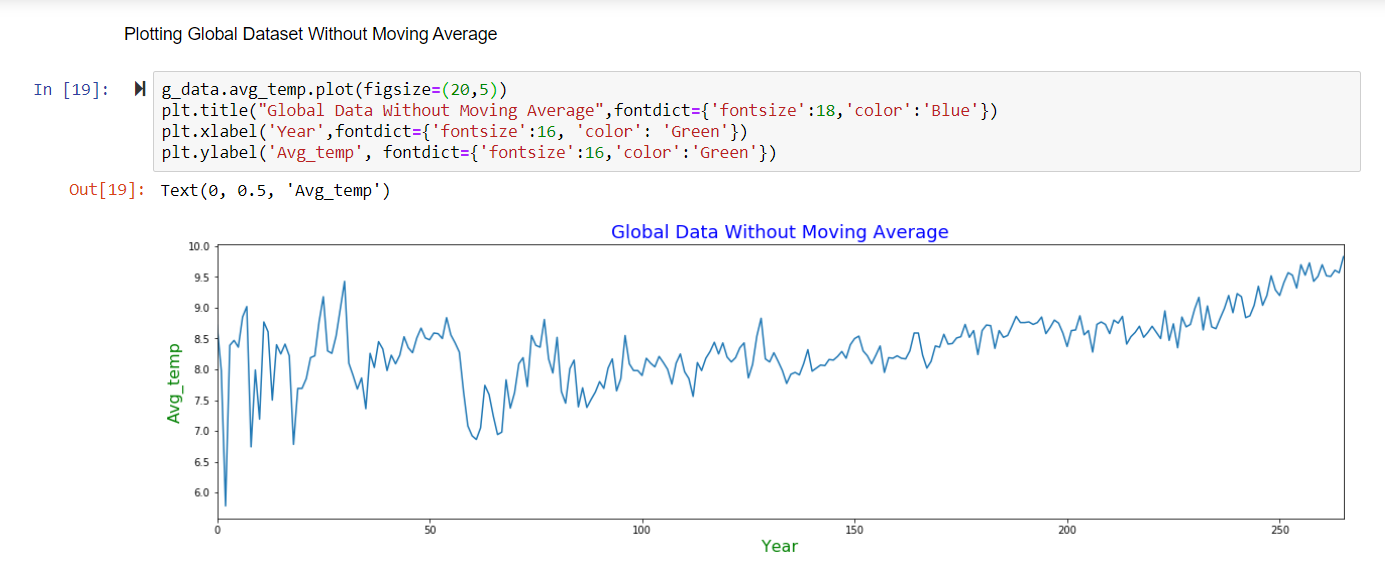
266 rows × 2 columns were produced.

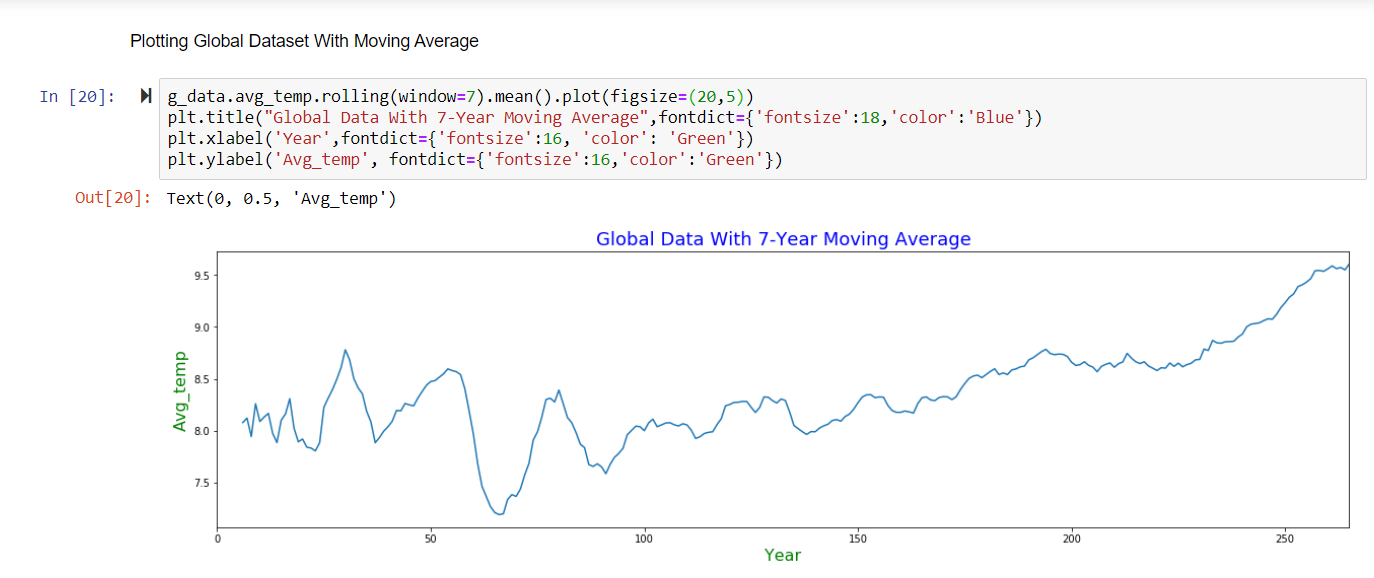


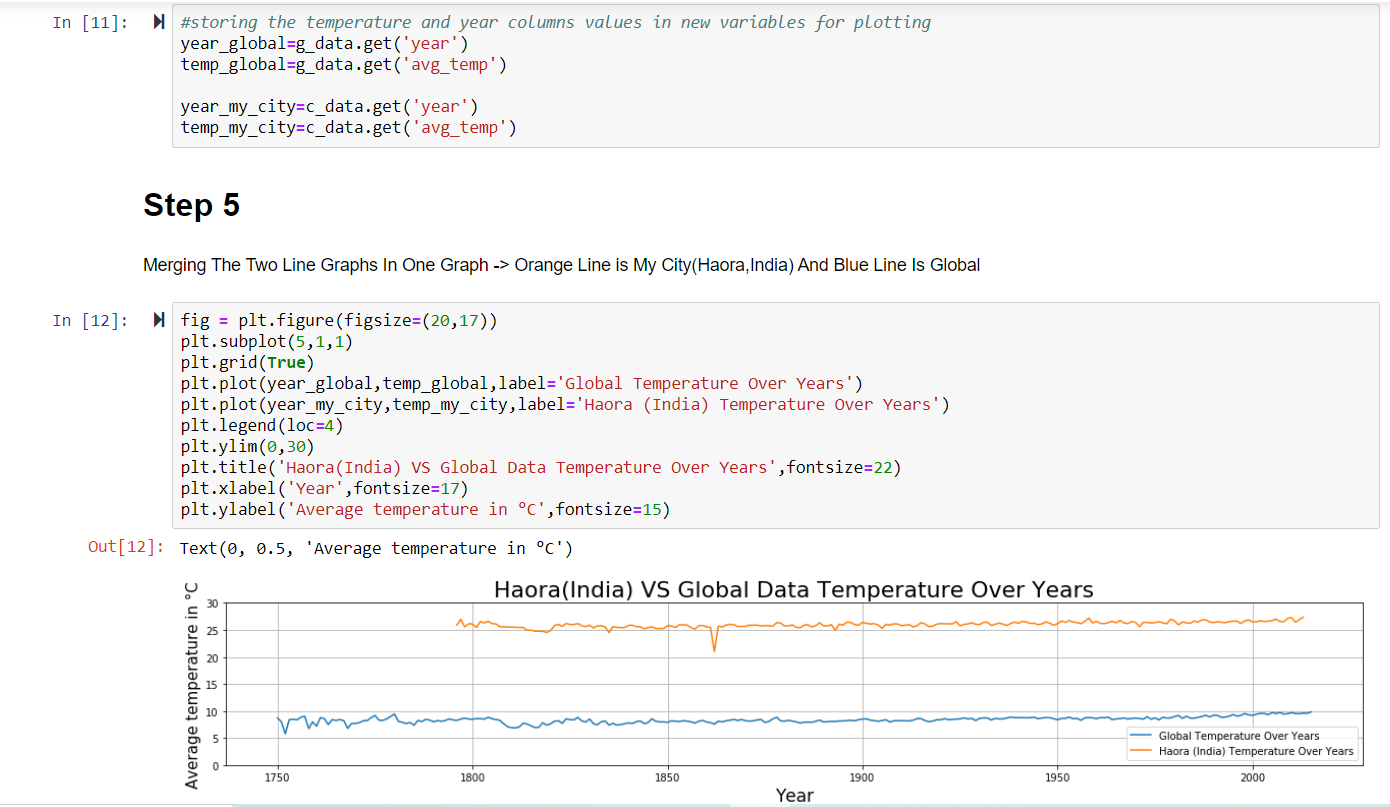
Plotting graphs without and with moving average to observe the difference and benefits





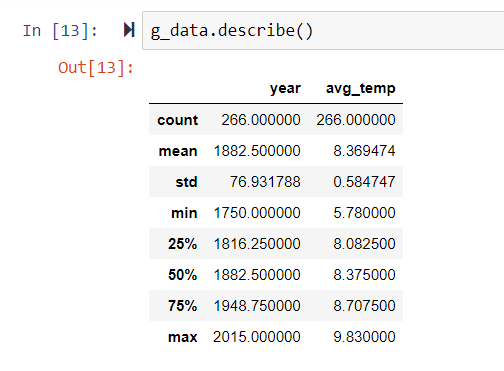
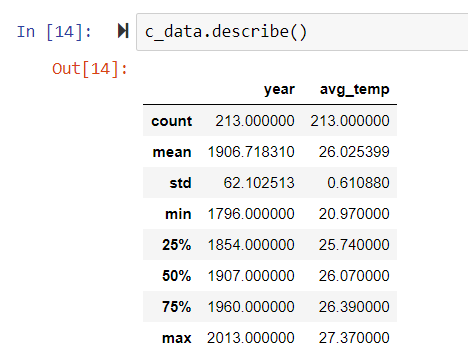




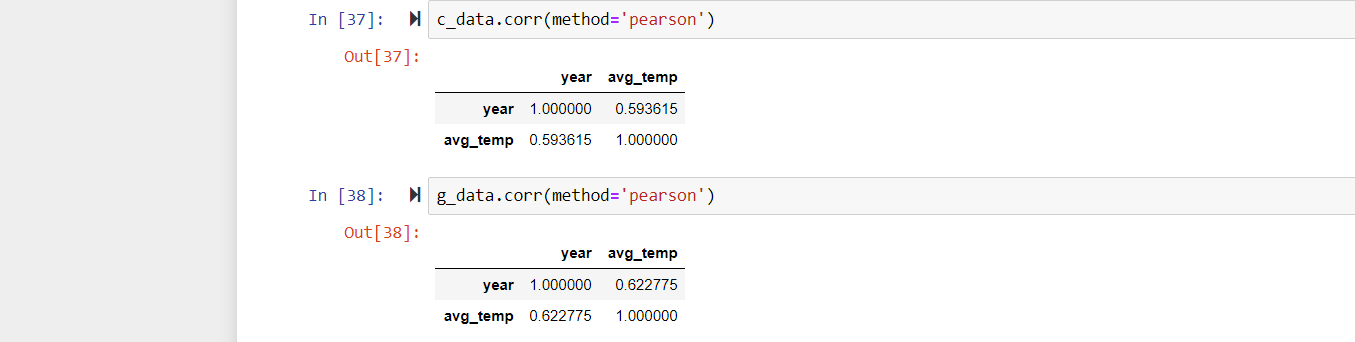


OBSERVATIONS:-

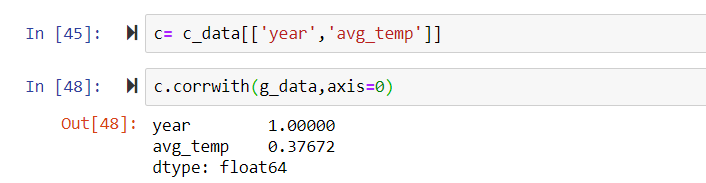
1. The following significant differences between the City Avg\_Temp (Haora in my case) and the Global Avg\_Temp

* The city Haora(in India) is much hotter than the global temperature.
* From 1750 to some years before of 1800 there is change in the global data but after that remained almost constant. However, the city\_data for the city Haora(in India) shows that the temperature has been constant but it fell towards 200 C in the years between 1850 and 1900 and then again increased towards 250 C and maintained it’s consistency almost.
* Thus, the city quite oscillation when compared to the global data.
* The maximum and minimum average temperature of the global is much greater than that of the city data.
* The following are the correlation among the columns of city\_data and the global\_data:-



* The following id the correlation among the columns of city\_data and the global\_data:-



* **The average temperature of city\_data is 8.369 and that of global\_data is 26.025 .**

CONCLUSION:-

* **Although the two line graphs are at different level of temperatures but from the year 1900 the pattern followed by them is almost similar if viewed on large scale.**

**-Thank You**