Udacity Nanodegree

Project 1- Explore Weather trends.

Tools used : Python 3(Spider),SQL

Objective : The objective of this report is to determine the effect of global warming in various parts of world.

Procedure :- First thing I did was importing the dataset from udacity database. SQL(Structured Query language. This data then used in python 3(Spider) with libraries like numpy ,panda and Numpy

SQL(Structured Query language):-The data was available in data base . First thing was what type of columns the dataset contains. There were three data sets were available Such as global data , city list and city data.Then there is something common relation between the data is city name ,global sheet avg temp ,city data avg temp .So to if I can combine these columns from three sheets then the required data would be ready.Then I decided to alter to column's name avg_temp in two data set as mention below:

"ALTER TABLE global_data RENAME COLUMN avg_temp to global_avg_temp;

ALTER TABLE city_data RENAME COLUMN avg_temp to city_avg_temp;"

Then I used inner join statement in sql data base because I will need the data of two different data set together .Therefore the data is imported by using below code after alteration :-

"SELECT global_data.year, global_data.global_avg_temp, city_avg_temp

FROM global_data INNER JOIN city_data /*For joining the two table */

ON global_data.year=city_data.year where city like 'Delhi'"

Python 3 (Spyder):- The next step is to import the data to python library .There to calculate the moving averages there is function available in pandas library called Pandas.rolling.This function used to reduce noise in data.The below code I mention is used for rolling function :-

```
"def simplifiedRollingMean(windowRolling, dataset_input):

    dataset_output = dataset_input.rolling(window = windowRolling, center=False, on = "year").mean().dropna()

return dataset_output"
```

<mark>Note Dropna used to drop empty values or nun entries in data set</mark>

Conclusion :-

To determine the conculusion I used 4 cities data . Firstly I tried with 7,14,50 moving averages but I find using 10 moving averages will give me good result.The cities which I used is Delhi,Tornto,Rome,Berlin.
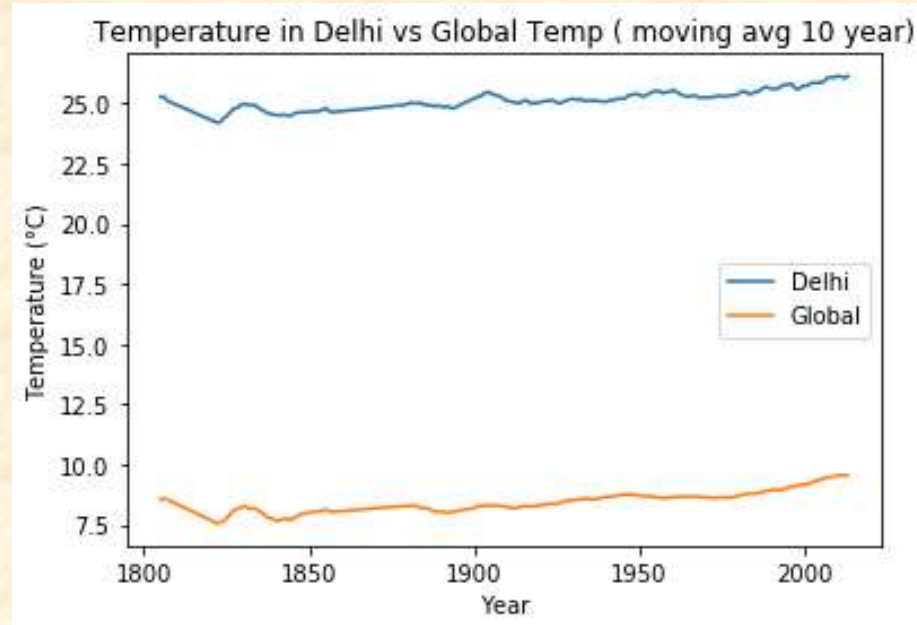

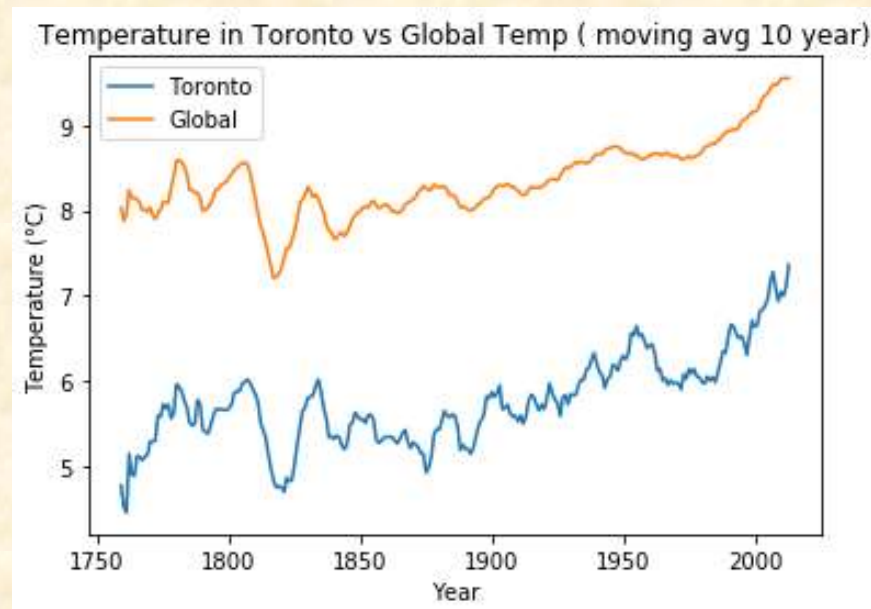Fig 1:- Moving averages Delhi vs Global Temperatures

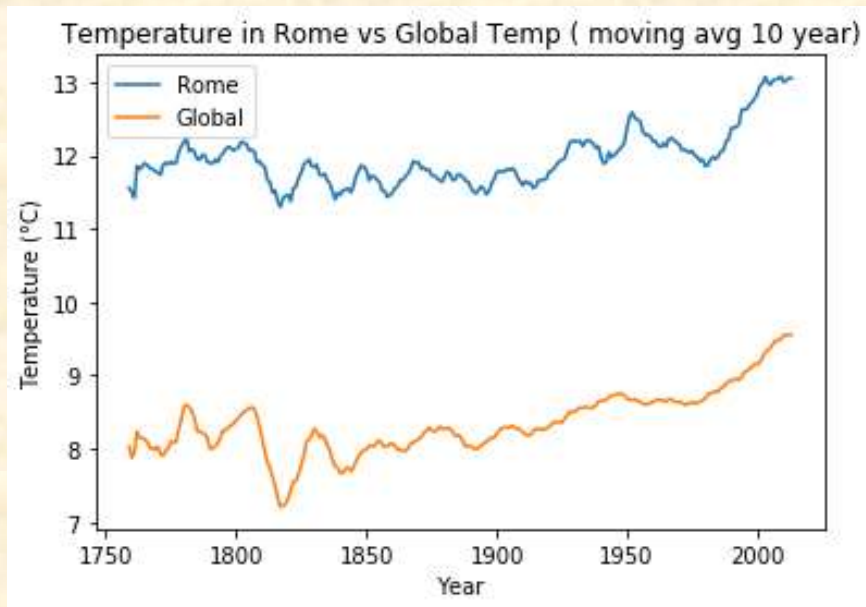Fig 2 Moving averages Tornto Vs Global Temp
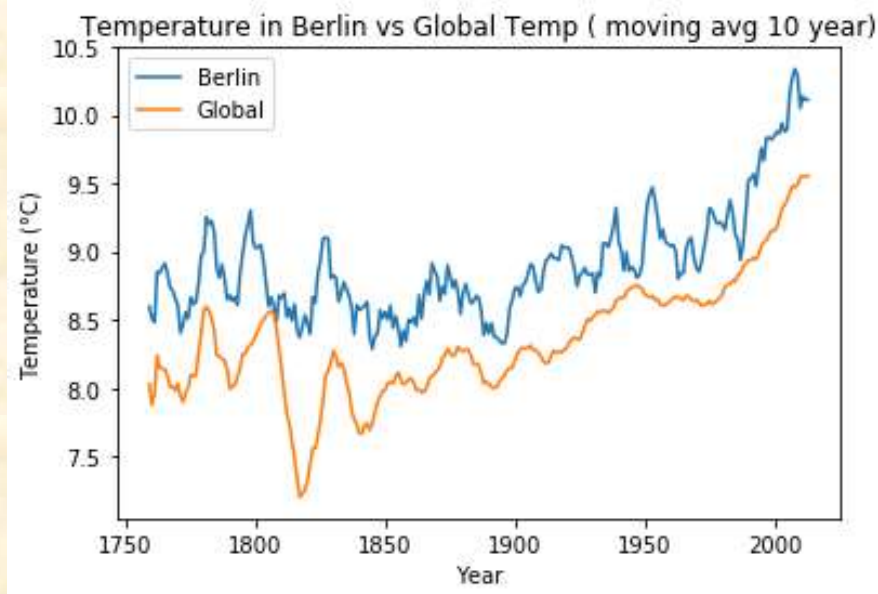
Fig 3- Moving averages Rome vs Global

Fig 3- Moving averages Berlin vs Global

From all the figures it can be already identified that delhi is quite hotter than the rest of the globe.For the delhi it seems the temperature increases after 1950s as global temperature tends to increases.The average temperature of delhi increases as global temperature increases.For the cties like berlin ,Toronto and rome their temperature also increases from 1950s. These cities are also hotter than global average temperature.Global warming is serious issue and is constantly increasing. The global temperature is increasing .

For code viewing visit my git hub profile.