

EXPLORING WEATHER TRENDS

In this project, analysis of local and Global temerature data was carried out. I compared the temperature trends of where I live (Lagos, Nigeria) to the overall Global temperatures trends.

Tools Used in this Project

- SQL Workspace provided by udacity
- Microsoft Excel
- Python in Jupyter notebook

Steps taken for my project

- The following sql queries were used to extract the city_data and the global_data and this was exported to csv.

Input

HISTORY ▾MENU ▾

SCHEMA

city

country

avg_temp

city_list ▾

global_data ▲

1

SELECT year,city,country, avg_temp

2

FROM city_data

Success!

EVALUATE

Output

70792 results

Download CSV

year	city	country	avg_temp
1849	Abidjan	Côte D'Ivoire	25.58

Input

HISTORY ▾MENU ▾

SCHEMA

avg_temp - (numeric)

city_list ▾

global_data ▲

year

avg_temp

1

SELECT year, avg_temp

2

FROM global_data

Success!

EVALUATE

Output

266 results

Download CSV

year	avg_temp
1750	8.72
1751	7.98

- On microsoft excel, I filtered out my city (Lagos, Nigeria) for comparism with the overall global data.
- I then the joined the columns of average temperature of the global data to that of my city data using the year as my primary key. Note that this was done manually i.e without the use of a formular.
- Next, I removed the rows that had no temperature values.
- I calculated the 7-years(7rows) moving average temperature for both Lagos and the world

ear	city	country	avg_temp	Local Moving Avg_temp	avg_temp p	Global Moving Avg_tem
1849	Lagos	Nigeria	25.98		7.98	
1850	Lagos	Nigeria	25.87		7.9	
1851	Lagos	Nigeria	26.1		8.18	
1856	Lagos	Nigeria	26.35		8	
1857	Lagos	Nigeria	25.45		7.76	
1858	Lagos	Nigeria	25.92		8.1	
1859	Lagos	Nigeria	26.18	=AVERAGE(D2:D8)		8.02
1860	Lagos	Nigeria	26.01	AVERAGE(number1, [number2], ...) 2		
1861	Lagos	Nigeria	25.95	25.99	7.85	8.01

- I then proceeded to Jupyter notebook to continue my analysis and visualizations

In [1]:

```
## Importing the libraries needed for the analysis
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Matplotlib is building the font cache; this may take a moment.

In [5]:

```
#Importing the csv file into the jupyter notebook environment

Weather = pd.read_csv('Weather trends.csv')
```

In [6]:

```
Weather.head()
```

Out[6]:

	year	city	country	avg_temp	Local Moving Average	avg_temp.1	Global Moving Average
0	1849	Lagos	Nigeria	25.98	NaN	7.98	NaN
1	1850	Lagos	Nigeria	25.87	NaN	7.90	NaN
2	1851	Lagos	Nigeria	26.10	NaN	8.18	NaN
3	1856	Lagos	Nigeria	26.35	NaN	8.00	NaN
4	1857	Lagos	Nigeria	25.45	NaN	7.76	NaN

In [7]:

```
#REMOVING THE NAN VALUES
df = Weather.dropna()
```

In [8]:

```
df.head()
```

Out[8]:

	year	city	country	avg_temp	Local Moving Average	avg_temp.1	Global Moving Average
6	1859	Lagos	Nigeria	26.18	25.98	8.25	8.02
7	1860	Lagos	Nigeria	26.01	25.98	7.96	8.02
8	1861	Lagos	Nigeria	25.95	25.99	7.85	8.01
9	1862	Lagos	Nigeria	25.90	25.97	7.56	7.93
10	1873	Lagos	Nigeria	26.46	25.98	8.35	7.98

In [9]:

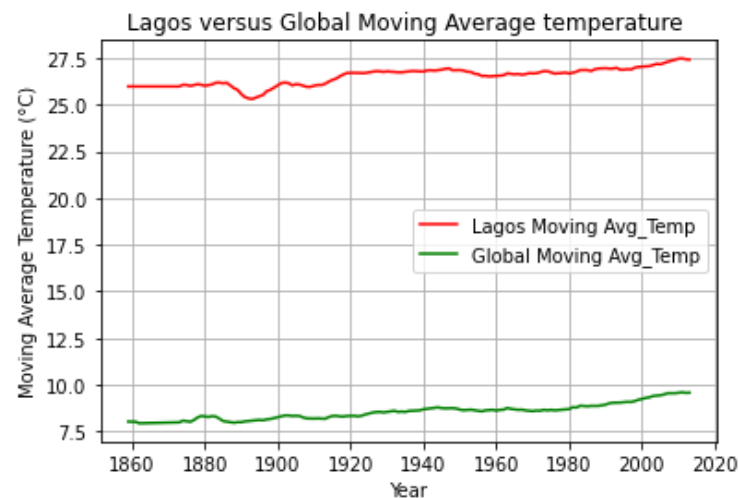
```
#ploting the line chart
x = df['year']
y = df['Local Moving Average']
z = df['Global Moving Average']
```

In [10]:

```
plt.plot(x,y,color = 'r', label = 'Lagos Moving Avg_Temp')
plt.plot(x,z,color = 'g', label = 'Global Moving Avg_Temp')
plt.legend();
plt.xlabel('Year')
plt.ylabel('Moving Average Temperature (°C)')
plt.grid()
plt.title('Lagos versus Global Moving Average temperature ')
```

Out[10]:

Text(0.5, 1.0, 'Lagos versus Global Moving Average temperature ')



From the line chart above, The following Observation was made;

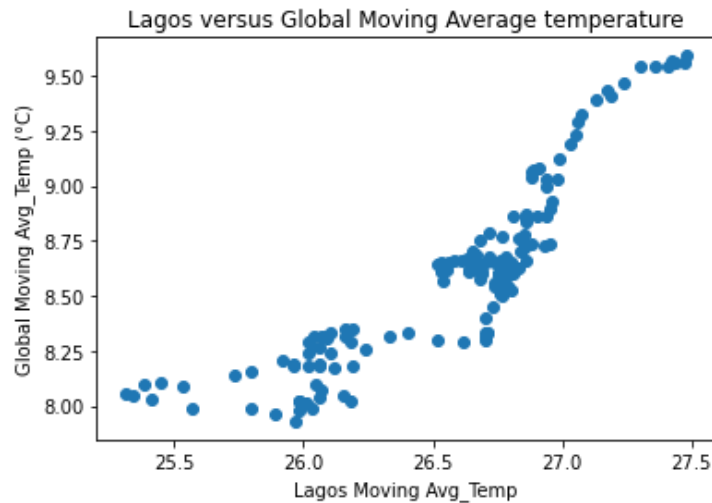
1. My City(Lagos Nigeria), temperature is hotter on average as compared to the Global average moving temperature.
2. The difference between Lagos's temperature and that of the globe fluntuates between 17-18 degrees celsius.
3. From around 1860 to 1878, the average temperatures were a bit constant for both Lagos and World. The temperature for both then increaed around 1885 and the dropped around 1890.
4. From the year 1890 till date, The temperature kept getting hotter consistently for both Lagos and the globe at large.

In [11]:

```
#Ploting yhe scatter plot
plt.scatter(y,z)
plt.xlabel('Lagos Moving Avg_Temp')
plt.ylabel('Global Moving Avg_Temp (°C)')
plt.title('Lagos versus Global Moving Average temperature')
```

Out[11]:

```
Text(0.5, 1.0, 'Lagos versus Global Moving Average temperature')
```



In [12]:

```
# To get the correlation coefficient,  
cor = np.corrcoef(y,z)  
cor
```

Out[12]:

```
array([[1.          , 0.88572823],  
       [0.88572823, 1.          ]])
```

Above is a Scattered plot to show the correlation between Lagos moving average temperature and that of the World in general.

From the plot, it is seen that they have a strong positive correlation with correlation coefficient of 0.89

Note that due to the strong positive correlation between the average temperature in Lagos and the World, using a strong prediction linear model, the temperature of Lagos can be estimated when that of the world is given.