EXPLORING WEATHER TRENDS

In this project, I will analyze local and global temperature data and compare the temperature trends where I live to overall global temperature trends.

DATA EXTRACTION

To extract the data from the database I used SQL on the SQL Workspace in the Udacity Classroom. For city-level data, I chose Lagos city, Nigeria. I combined both dataset into one with the JOIN keyword. The screenshots below show the SQL query to extract the combined global data and the local data.

DATA PREPARATION

I then downloaded the datasets from the SQL Workspace in the Udacity Classroom into a csv file. I imported the libraries necessary and opened the CSV files with Pandas, the Python's library and I used the integrated development environment (IDE) Jupyterlab. See below a screen shot of the process.

```
#import libraries
import pandas as pd
import numpy as np
import seaborn as sns
from matplotlib import pyplot as plt
```

```
#read first five values of the datasets
combined_data = pd.read_csv('data.csv')
combined_data.head()
```

	year_city	city	country	avg_temp_city	year_glob	avg_temp_glob
0	1849	Lagos	Nigeria	25.98	1849	7.98
1	1850	Lagos	Nigeria	25.87	1850	7.90
2	1851	Lagos	Nigeria	26.10	1851	8.18
3	1852	Lagos	Nigeria	NaN	1852	8.10
4	1853	Lagos	Nigeria	NaN	1853	8.04

Moving averages are used to smooth out data to make it easier to observe long term trends and not get lost in daily fluctuations. If I plot the average temperature in both global and local dataset, the chart looks too volatile to interpret, like is shown below.

```
plt.plot(combined_data['year_city'],
         combined_data['avg_temp_city'],
         label='Lagos temp', color= 'm')
plt.plot(combined_data['year_glob'],
         combined_data['avg_temp_glob'],
         label='Global temp', color= 'c')
plt.legend()
plt.title('Weather Trends')
plt.xlabel('Year')
plt.ylabel('Temperature (°C)')
plt.show()
```
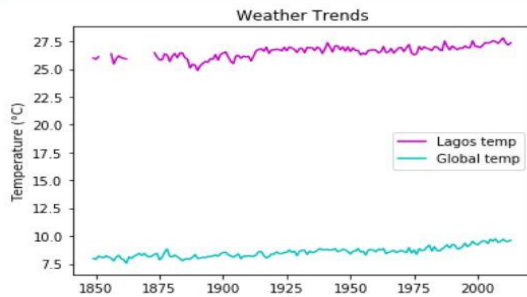

Using a moving average, I can both smooth out the daily volatility and to observe the long term trend. For calculating the moving averages I used the .rolling() method on the average temperatures column.

```
combined_data['ma_local'] = combined_data['avg_temp_city'].rolling(window=20).mean()

combined_data['ma_global'] = combined_data['avg_temp_glob'].rolling(window=20).mean()
combined_data.tail()
```

	year_city	city	country	avg_temp_city	year_glob	avg_temp_glob	ma_local	ma_global
160	2009	Lagos	Nigeria	27.53	2009	9.51	27.1420	9.3245
161	2010	Lagos	Nigeria	27.79	2010	9.70	27.1785	9.3480
162	2011	Lagos	Nigeria	27.35	2011	9.52	27.2020	9.3650
163	2012	Lagos	Nigeria	27.15	2012	9.51	27.2230	9.3985
164	2013	Lagos	Nigeria	27.36	2013	9.61	27.2460	9.4355

The line plot is shown below

```
plt.plot(combined_data['year_city'],
         combined_data['ma_local'],
         label='Lagos temp', color= 'm')
plt.plot(combined_data['year_glob'],
         combined_data['ma_global'],
         label='Global temp', color= 'c')
plt.legend()
plt.title('Weather Trends')
plt.xlabel('Year')
plt.ylabel('Temperature (°C)')
plt.show()
```
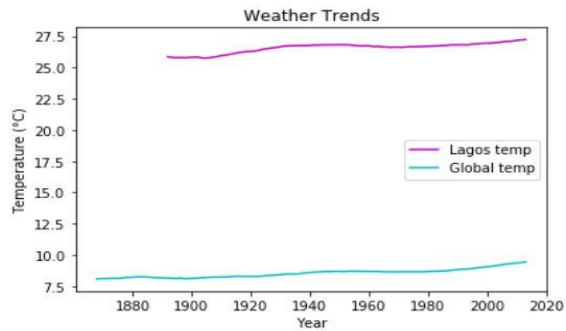

Conclusion

- Over time, it seems that the city's temperatures have change similarly than the global average.

- Over time, both types of temperature (global and local) show an upward trend, that is the world is getting hotter.

- According to the chart, it can be told that the upwards trend has been consistent over a long time ago.

- Lagos city's temperature has been consistently higher than the global average over time.