

Explore Weather Trends



DATA ANALYST NANODEGREE



Mohamed Gamal Elbayoumi
21-May 2020

Overview

I have analyzed, in this project, local temperature of Cairo, Egypt in accordance with the global temperature data and compared. I had been provided with a database on Udacity portal.

Goals

- ❑ Selecting city and country from the database “ city_list ”.
- ❑ Extracting the City level data from the database “ city_data ” and export to CSVfile.
- ❑ Extracting the global temperature from the database “ global_data ” and export to CSV file.

Tools

- ❑ Python : For calculating moving average and plotting line chart
- ❑ SQL : To extract the data from the database
- ❑ Google Sheet : To calculate Moving Averages of global and city temperatures

Step 1 : Data Extraction

- check available countries and cities

```
SELECT *  
FROM city_list WHERE Country='Egypt'  
AND City='Cairo' ;
```

- select data from the City database

```
SELECT avg_temp,year,city,country  
FROM city_data  
WHERE city='Cairo' ;
```

- I observed from the SCHEMA that both city_data and global_data contain the same column named 'avg_temp'. So I have changed the names of the columns respectively in order to have distinct columns.

```
ALTER TABLE city_data  
RENAME COLUMN avg_temp to city_avg_temp;  
ALTER TABLE global_data  
RENAME COLUMN avg_temp to global_avg_temp;
```

- I have joined the two tables using JOIN also called as INNER JOIN as avg_temp is the same in both the tables.

```
SELECT global_data.year,  
global_data.global_avg_temp,city_data.city_avg_temp  
FROM global_data JOIN city_data
```

```
ON global_data.year = city_data.year  
WHERE city LIKE 'Cairo';
```

I downloaded the file as “results.csv”.

Step 2 : Plot Line Chart

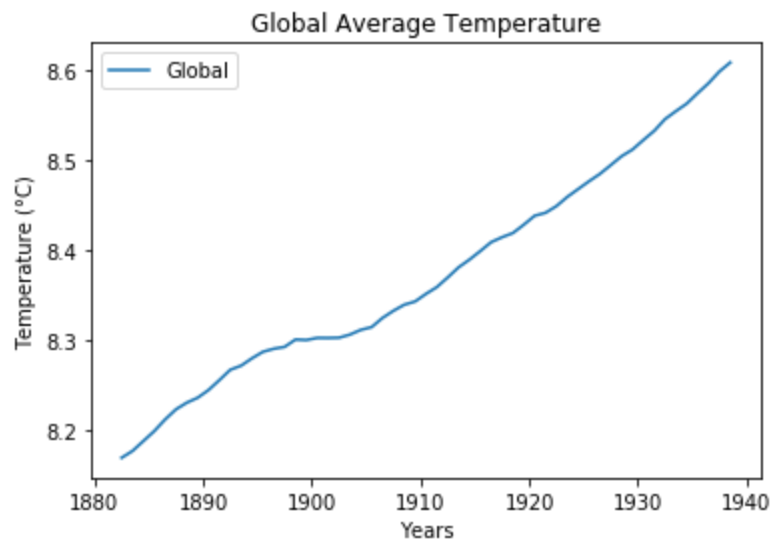
Codes

```
# Libraries  
import numpy as np  
import pandas as pd  
from matplotlib import pyplot as plt  
  
#Read Data Set  
data = pd.read_csv("results.csv")  
  
# function that calculates the MOVING AVERAGE  
def moving_avg(mA_range, data_input):  
    output = data_input.rolling(window = mA_range, on  
    ="cat").mean().dropna()  
    return output  
  
# Function Calling with the range of Moving Average  
mA_value = 150  
chart_moving_avg = moving_avg(mA_value, data)  
  
# Drawing the graph: Global Temperature
```

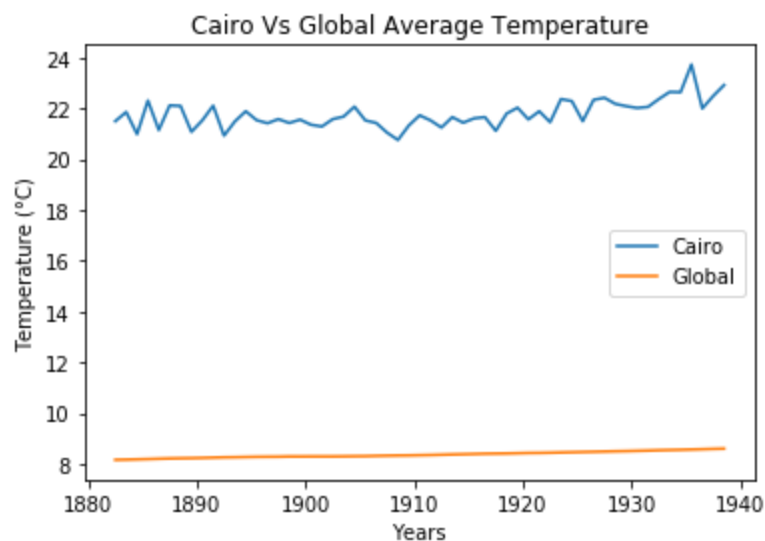
```

plt.plot(chart_moving_avg ['year'], chart_moving_avg ['gat'], label = 'Global')
plt.legend()
plt.xlabel ("Years")
plt.ylabel ("Temperature (°C)")
plt.title ("Global Average Temperature")
plt.show ()

```



Now combined with Cairo data,



Observations:

- ❖ Global Average Temperature for 10 yr MA varies between 8.5°C to 9.5°C
- ❖ The Chart of Cairo Vs Global has a very big difference in the temperatures.
- ❖ Cairo Average Temperatures is hotter than global average temperature.
- ❖ From the first Graph, I observed global temperature is increasing smoothly from about 8 to 8.5 over the provided period.
- ❖ Cairo average temperatures are ups and downs during the early years, later during
- ❖ Both the temperatures increased due to increase in temperature.

References:

<https://www.statisticshowto.com/moving-average/>

