

# Explore Weather Trends

Made By: **Tariq Issa**

---

## Overview

In this project, Require To analyze the local temperature for my country, Unfortunately my Real country Does not Exist in the city data tables So My choice was **damascus, Syria**.  
I used the database from Udacity portal, as will shown in the following goals.

## Goals

1. Extraction of data from the database and export to CSV file
2. Making a chart visualization based on extracted data using plots
3. Observation based on chart as requirements

## Tools Used:

1. SQL: To extract the data from the database
2. Python: For calculating moving average and plotting line chart ( small callable function )
3. ANACONDA - Jupyter Notebook: For writing python code and making observations(python 2)

**NOTE:** In this project I used some functionality from some special sites and i will refers to them

### Phase 1 - Extraction of Data from provided Database

To see which cities are available for "Syria" in the given dataset:

```
SELECT *  
FROM city_list  
WHERE country = 'Syria' ;
```

**\*\*** I made some changes of the database, Depends on these two tables (city\_data, global\_data) , i notice that there is a same column name 'avg\_temp' for both of ,, so I joined these two tables together depends on the 'avg\_temp' And I rename each of these tables columns As shown :

A- **ALTER TABLE city\_data RENAME COLUMN avg\_temp to city\_avg\_temp;**  
--> city\_avg\_temp = City Average Temp.

B- **ALTER TABLE global\_data RENAME COLUMN avg\_temp to global\_avg\_temp;**  
--> global\_avg\_temp = Global Average Temp.

C- Now I wrote the following code in order to join the two tables and have the relevant data:

-- join two table

```
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 = 'Damascus';
```

D- I downloaded the file as CSV format. Downloaded as "Syria\_joined\_table.csv".

## Phase 2 - Python Code for Making Line Chart

I wrote this code on Jupyter Notebook.

# Importing the important Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

# Importing the Data Set(csv file )

```
data = pd.read_csv('Syria_joined_table.csv')
```

defined a function for the calculation of moving averages in order to get smooth graph.

# function that calculates the MOVING AVERAGE

```
def moving_avg(mA_range, data_input):
    output = data_input.rolling(window = mA_range, center = False, on = 'city_avg_temp').mean().dropna()
    return output
```

# Function Calling with the range of Moving Average

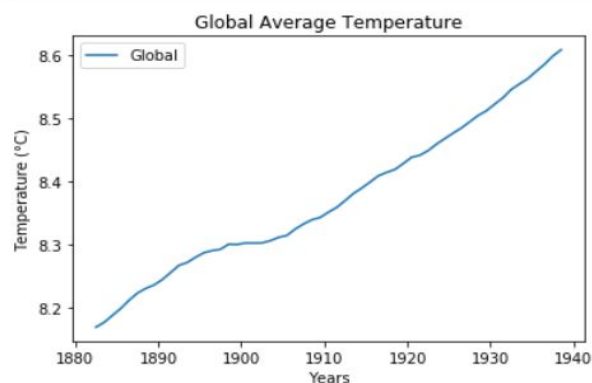
```
mA_value = 150
```

```
chart_moving_avg = moving_avg(mA_value, data)
```

# Plotting the graph: Global Temperature

```
plt.plot(chart_moving_avg['year'], chart_moving_avg['global_avg_temp'], label='Global')
plt.legend()
plt.xlabel('Years')
plt.ylabel(u 'Temperature (°C)' ) # u : for declare the (°C)as unicode
plt.title("Global Average Temperature")
plt.show()
```

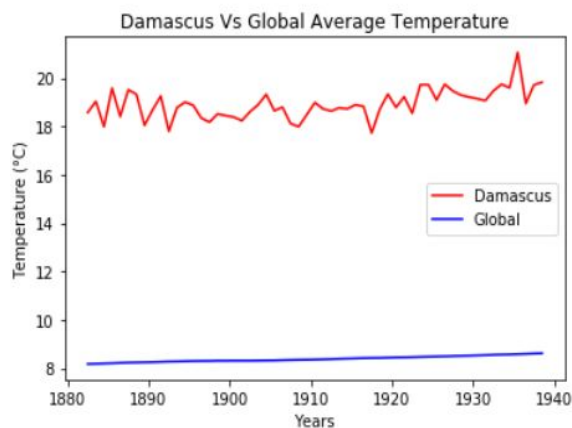
```
In [177]: plt.plot(chart_moving_avg['year'], chart_moving_avg['global_avg_temp'], label='Global')
plt.legend()
plt.xlabel('Years')
plt.ylabel(u'Temperature (°C)' )
plt.title("Global Average Temperature")
plt.show()
```



I analysed the Global Average temperatures And MY country temperatures each one is separate at the same year  
As shown :

```
# Drawing the graph: damascus and Global Temperature
plt.plot(chart_moving_avg['year'], chart_moving_avg['city_avg_temp'], label='Damascus',c='r')
plt.plot(chart_moving_avg['year'], chart_moving_avg['global_avg_temp'], label='Global',c='b')
plt.legend()
plt.xlabel("Years")
plt.ylabel(u"Temperature (°C)")
plt.title("Damascus Vs Global Average Temperature")
plt.show()
```

```
In [190]: # Drawing the graph: damascus and Global Temperature
plt.plot(chart_moving_avg['year'], chart_moving_avg['city_avg_temp'], label='Damascus',c='r')
plt.plot(chart_moving_avg['year'], chart_moving_avg['global_avg_temp'], label='Global',c='b')
plt.legend()
plt.xlabel("Years")
plt.ylabel(u"Temperature (°C)")
plt.title("Damascus Vs Global Average Temperature")
plt.show()
```



```
data.tail(10)
```

	year	global_avg_temp	city_avg_temp
196	2004	9.32	19.22
197	2005	9.70	19.15
198	2006	9.53	19.07
199	2007	9.73	19.47
200	2008	9.43	19.75
201	2009	9.51	19.60
202	2010	9.70	21.07
203	2011	9.52	18.95
204	2012	9.51	19.72
205	2013	9.61	19.83

```
data.head(10)
```

	year	global_avg_temp	city_avg_temp
0	1808	7.63	14.91
1	1809	7.08	17.04
2	1810	6.92	17.09
3	1811	6.86	17.12
4	1812	7.05	17.07
5	1813	7.74	17.62
6	1814	7.59	17.68
7	1815	7.24	17.42
8	1816	6.94	16.97
9	1817	6.98	17.15

As shown **Above**

## RESULTS - :

I have observed that, if I choose a **short range** for moving average, like 10 or 50, I will get messy line. Also the range of the “Years” on x-axis becomes **short**.

If I use a **larger** moving average, say 150, I will get relatively smooth graph and range of x-axis is **longer**.

**Observation** from the Line Charts:

1. The chart of Damascus Vs Global Temperature: Very big difference between the average temperature of Damascus and the Global temperature.
2. Since I have got a slight inclined straight line for global data. I have separately plotted the graph of global data. (The first graph)
3. From the first graph: I noticed that global temp. is increasing quite constantly with years by 0.1 degree centigrade.

To verify my findings I have used : **data.tail (10)** which displayed me the 10 rows from the last and I have seen the same insight as in the graph.

\*\*\* Coming back to the second chart, I observe that Damascus have temperatures **greater** than the global average.

\*\*\* If I draw a tangent line touching the two troughs or crests of line of Damascus, I see that there is a **consistent** change between this line and the line of global average over time.

To verify the temperatures of recent years, like 2010 or 2012, I used data.tail (10) which showed me the same insight from the table.

\*\*\* The Damascus city seems to be **hotter** than any other cool place in the world.

\*\*\* From the graph, Noticing that the temperature of the world is on constant rise.

This is all of my observations for the data used in this project.

Key Considerations:

1. Unit of Temperature: Centigrade, on Y-axis
2. Years shown on X-axis
3. Different colors of lines for city and global average
4. Use of Matplotlib library for visualization
5. Applied moving average on City data in order to get relatively smooth line
6. Defined a function for easy code
7. Saved all of the codes in .ipynb files (Jupyter Notebook) for later reference and regenerations or revisions.

## REFERENCES:

Change Column Name: <https://www.1keydata.com/sql/alter-table-rename-column.html>

Joining the tables: <http://www.dofactory.com/sql/join>

Calculation command for Moving Average in Python:

<http://www.learn datasci.com/python-finance-part-3-moving-average-trading-strategy/>

Parameters for .rolling()

<https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.rolling.html>