



Data Analyst for Enterprise Nanodegree Program

Exploring Weather Trends

By

Mariam Massadeh

10/4/2019

Contents

PROJECT SPECIFICATION	3
Analysis	3
Suggestions to Make Your Project Stand Out!	3
Used Tool	4
Extract the data from the database	4
SQL query to extract the city level data	4
SQL query to extract the global data	4
Chart	5
Line chart with local trend	5
Global temperature trend	6
Moving Average	6
Damascus MA	7
Global MA	8
General Average	8
Correlation Coefficient	8
My favorite cities	8
Observations	10

PROJECT SPECIFICATION

Analysis

CRITERIA	MEETS SPECIFICATIONS
Student is able to extract data from a database using SQL.	<ul style="list-style-type: none">• The SQL query used to extract the data is included.• The query runs without error and pulls the intended data.
Student is able to manipulate data in a spreadsheet or similar tool.	<ul style="list-style-type: none">• Moving averages are calculated to be used in the line chart.
Student is able to create a clear data visualization.	<ul style="list-style-type: none">• A line chart is included in the submission.• The chart and its axes have titles, and there's a clear legend (if applicable).
Student is able to interpret a data visualization.	<ul style="list-style-type: none">• The student includes four observations about their provided data visualization.• The four observations are accurate.

Suggestions to Make Your Project Stand Out!

Think about other ways to compare and find insights from this data beyond interpreting the chart. Here are a few ideas:

- What's the correlation coefficient?
- Can you estimate the average temperature in your city based on the average global temperature?
- Multiple cities - Add your favorite cities from around the globe to your visualization. What do you learn about them?

Used Tool

I used the **SQL** to extract the required data from the database and I used **Microsoft Excel** for the sheet I extracted to calculate the moving average, generating the line chart and getting observations taking in consider the temperature's trend either decreasing or increasing.

Extract the data from the database

SQL query to extract the city level data

Initially I have to check the available list to determine which is the closest country to me that have an available information.

```
SELECT * FROM city_list;
```

So I didn't find any data about my country Jordan so I selected Syria, and I retrieved the two required columns (206 results).

```
SELECT year,avg_temp FROM city_data WHERE country='Syria' and city='Damascus';
```

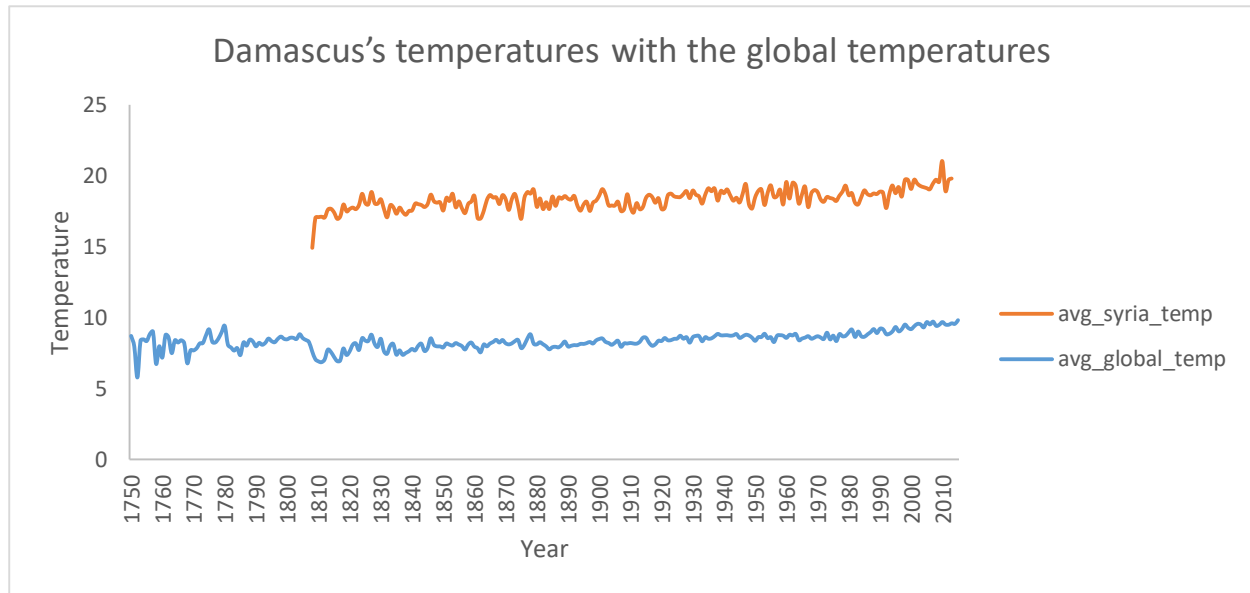
SQL query to extract the global data

I retrieved the two required columns (266 results).

```
SELECT * FROM global_data;
```

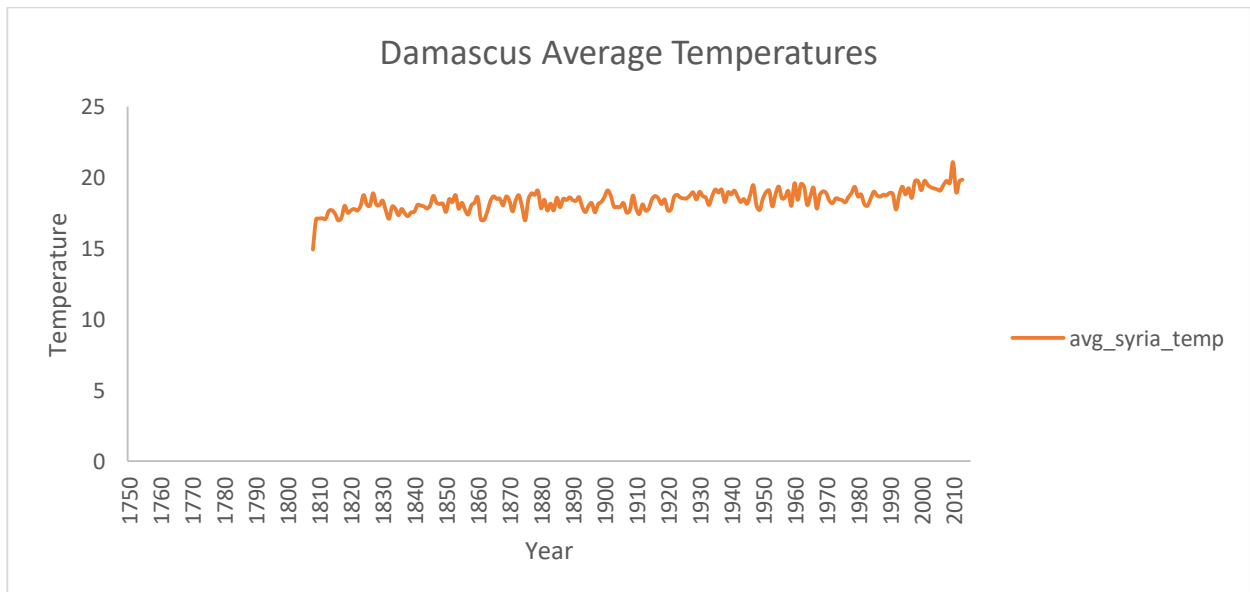
Chart

Now for the line chart that compares my city's temperatures with the global temperatures:

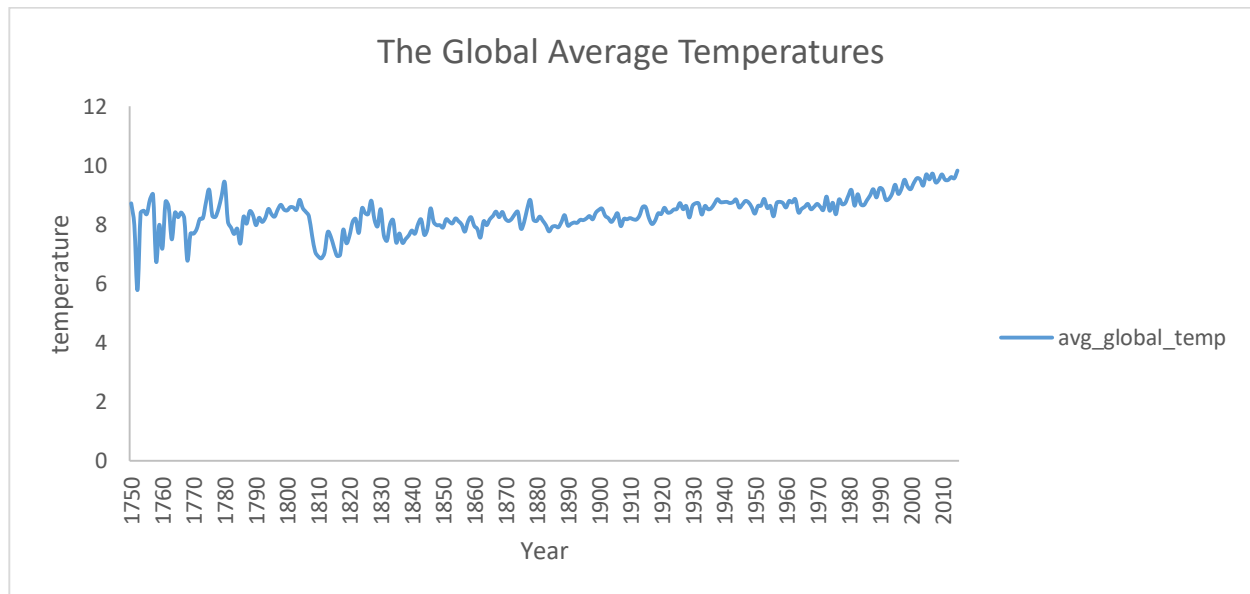


Regarding to this line chart we can see that the temperature moves upward, so the global temperature directly leads **Damascus's** temperatures to increase.

Line chart with local trend



Global temperature trend



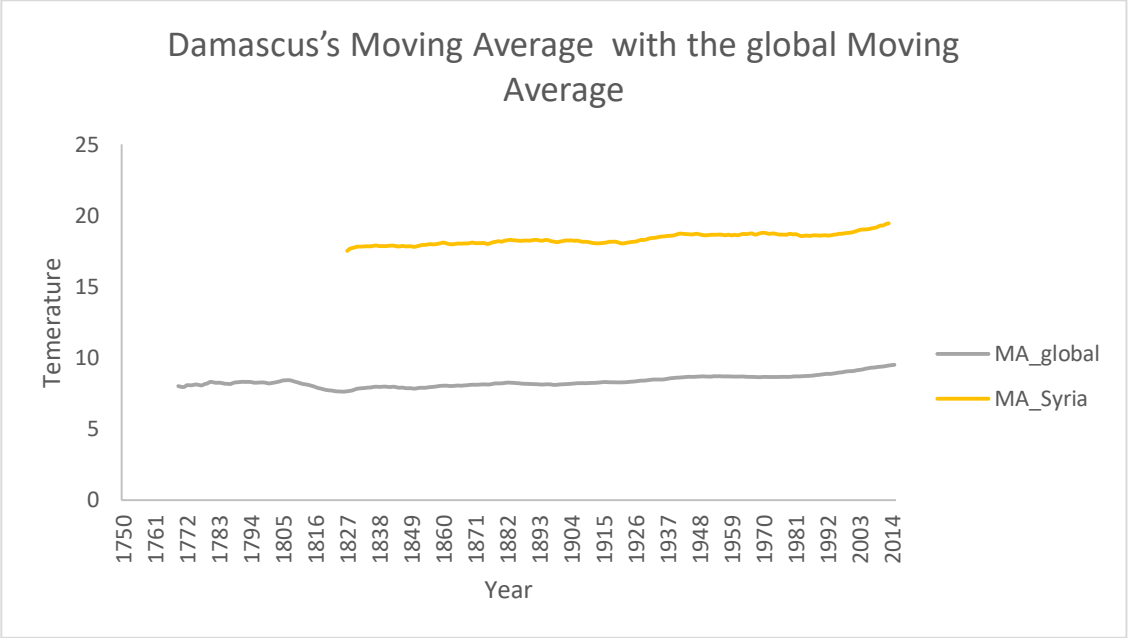
Moving Average

Moving averages are used to smooth out data to make it easier to observe long term trends and not get lost in daily fluctuations; Using a moving average, you can both smooth out the daily volatility and allow you to observe the long term trend.

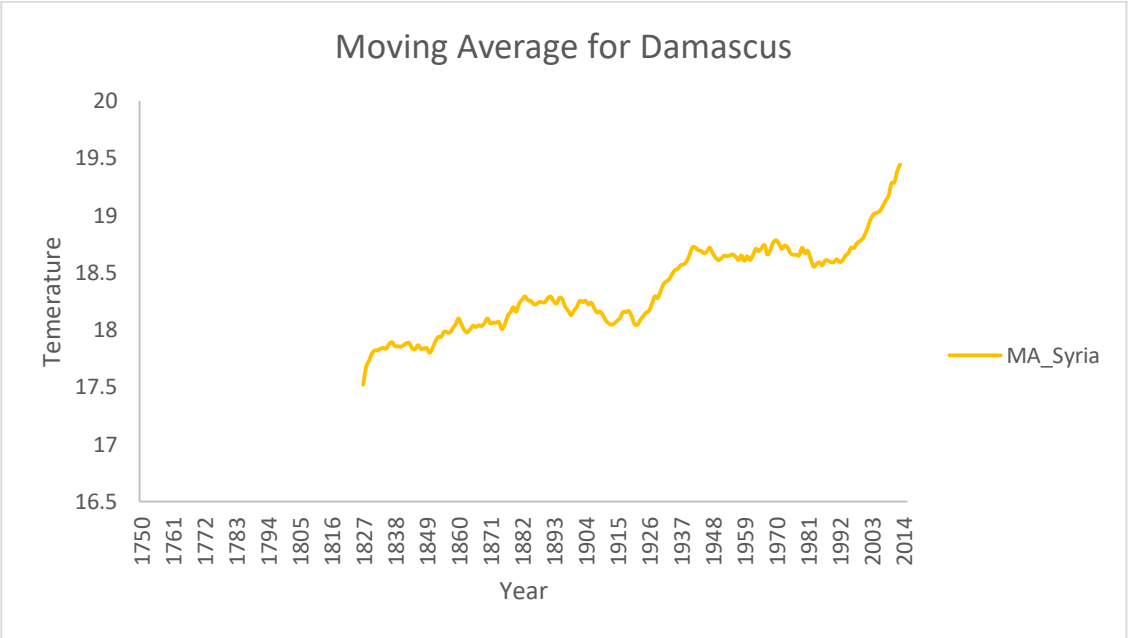
I created a new column called **MA**, which is where the moving average field will be stored. Go down by **20 results** and I used the excel formula

```
=AVERAGE(cell2:cell21)
```

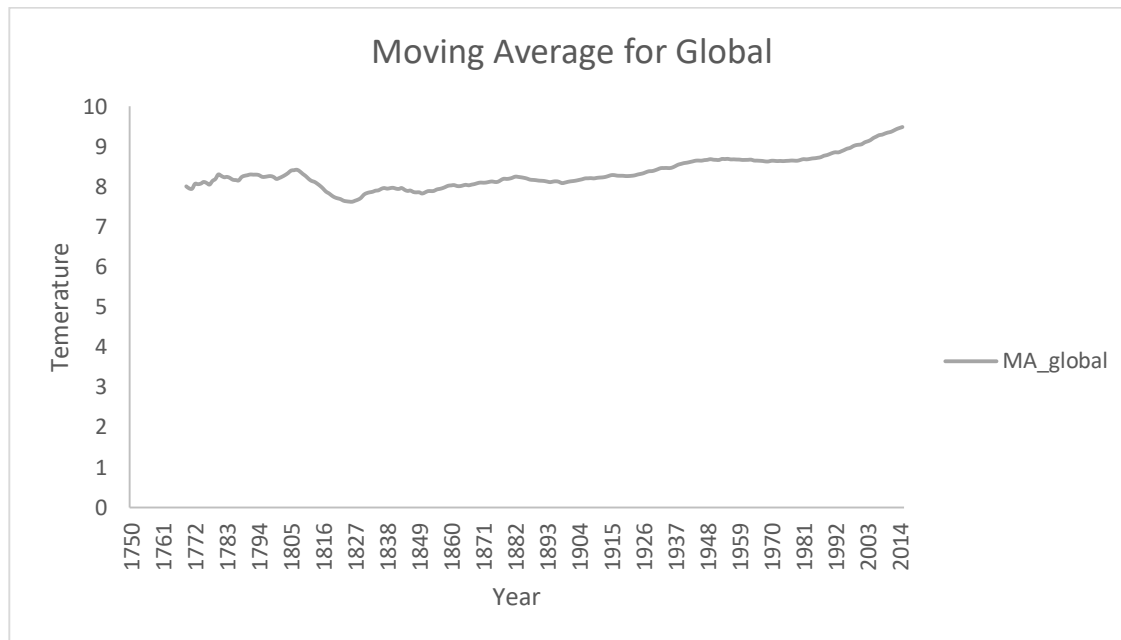
to calculate the average temperature for both local and global, then via **CTRL+C**, **CTRL+V** and **CTRL+D** to all the rest of cells.



Damascus MA



Global MA



General Average

Via the **AVERAGE formula in Excel**, the Global average temperature equal to **8.369474** while **Damascus** average temperature equal to **18.35**, so **Damascus** is hotter than the **Global**.

Correlation Coefficient

For correlation coefficient I selected data for the period 1808-2013, via the Excel formula

```
=CORREL(B60:B265,C60:C265)
```

the result was **0.773409** (highly correlated) which means that the relation between the **Global** and **Damascus** average temperatures have a **positive correlation** so the points are located close to the straight line.

My favorite cities

Let me try with other cities:

- **Valencia** in **Spain** as a city from in **Europe**.

```
SELECT year,avg_temp FROM city_data WHERE country='Spain' and city='Valencia';
```


The correlation of **Valencia** and **Global** is 0.7219 (highly correlated) which means that the relation between the **Global** and **Valencia** average temperatures have a **positive correlation** so the points are located close to the straight line.

- **Tijuana** in **Mexico** as a city from the **Latin America**.

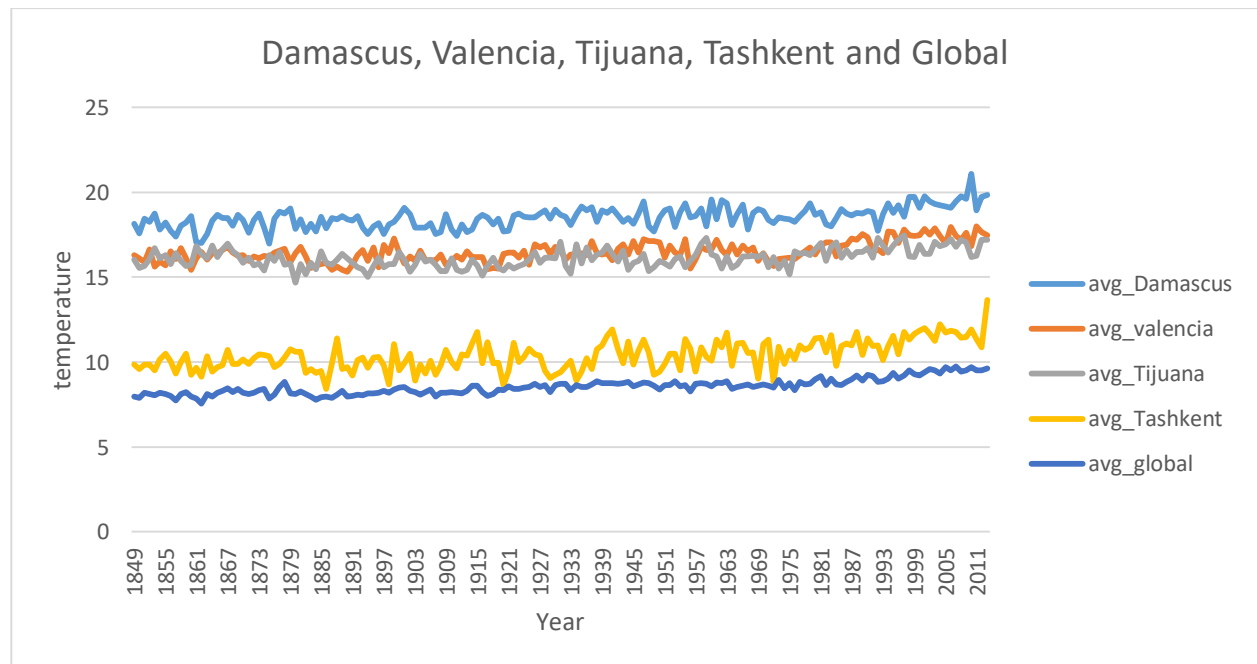
```
SELECT year,avg_temp FROM city_data WHERE country='Mexico' and city='Tijuana';
```

The correlation of **Tijuana** and **Global** is 0.505094 (moderately correlated) which means that the relation between the **Global** and **Tijuana** average temperatures have a **positive correlation** so the points are located close to the straight line.

- **Tashkent** in **Uzbekistan** as a city from **Asia**

```
SELECT year,avg_temp FROM city_data WHERE country='Uzbekistan' and city='Tashkent';
```

The correlation of **Tashkent** and **Global** is 0.721578 (highly correlated) which means that the relation between the **Global** and **Tashkent** average temperatures have a **positive correlation** so the points are located close to the straight line.



Observations

- Global average temperature range is **(5.78-9.83)** while **Damascus** city average temperature range is **(14.91-21.07)**, I calculated these via **MAX** and **MIN** formulas in Excel.
- Regarding to these ranges both increase over time, in general Global increases by **4.05 degrees**, while **Damascus** increases by **6.16 degrees**.
- The Global average temperature equal to **8.369474** while **Damascus** average temperature equal to **18.35**, then **Damascus** is hotter than the Global average temperature.
- **Damascus** city and the world are getting hotter year by year depending the correlation coefficient and also regarding to the diagrams; they have the same trend to be consistent over time which support the case of climate change.