Țîbuliac Andrei – Udacity Data Analyst Nanodegree

Project 1 - Explore Weather Trends

Introduction

In my first Data Analyst project, I used in the first step the SQL query to extract the data on a “csv” file of the cities and global data. I extract in a “csv” file the temperatures data of the city Bucharest from Romania. I used Microsoft Office Excel to create an data chart of the moving averages of the city and global data.

Process Method

My Weather Trend Project consisted of two steps:

1. **The SQL part.**

   The SQL part of the project where I selected the Bucharest, Romania city from where I gatered the data.

   I first extract the city list and countries database in order to easily find the city near me using this query:

   *SELECT * FROM city_list*;

   I then extract the global data that contain the average temperatures by year using the following query:

   *SELECT * FROM global_data;*

   In the last step I extract the average temperatures for the Bucharest city from the 1743 until 2013 using several operators like “SELECT”, “FROM”, “WHERE”, “LIKE” in the following query:

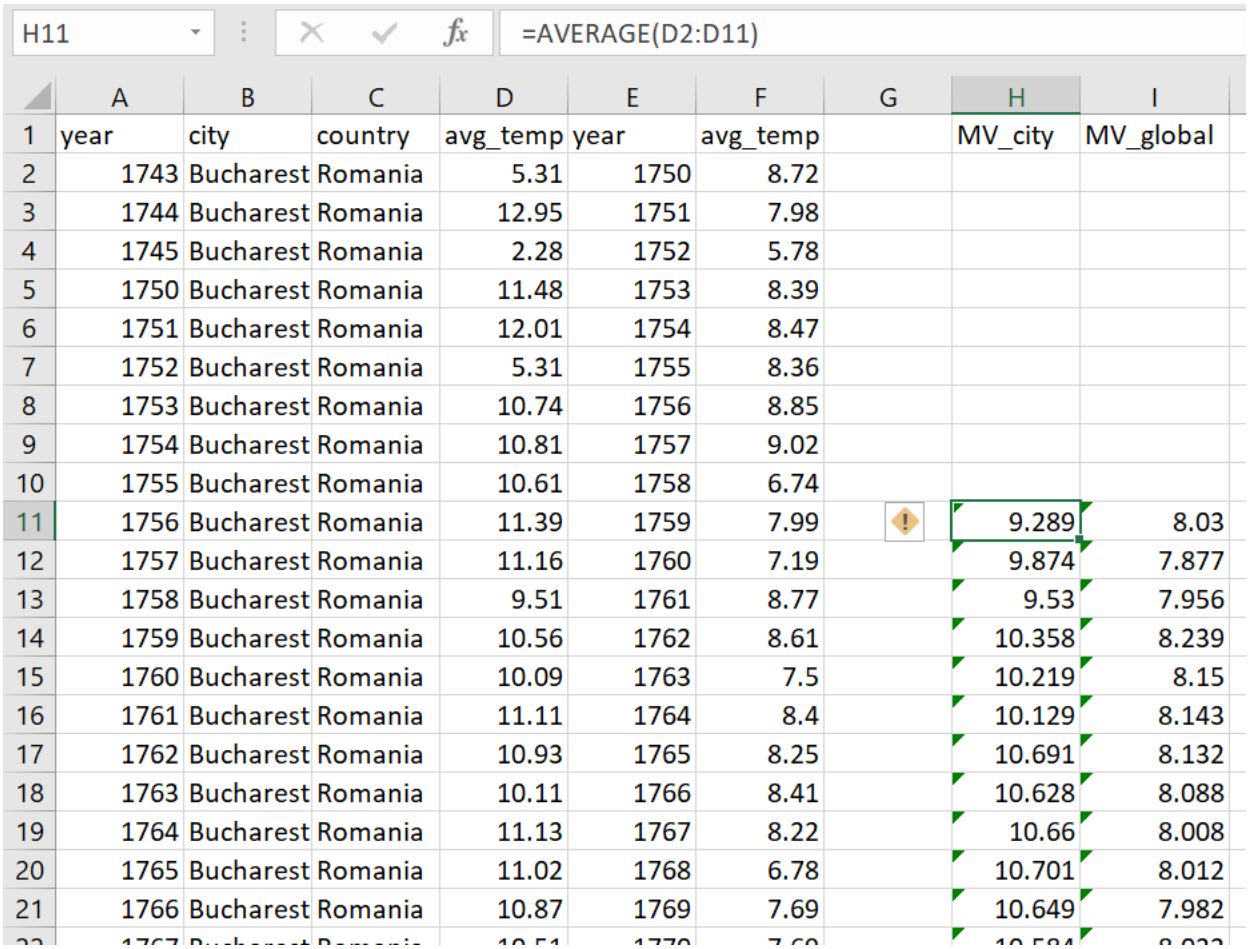
   *SELECT **

FROM city_data

WHERE city_data.city LIKE 'Bucharest' AND city_data.country LIKE 'Romania';

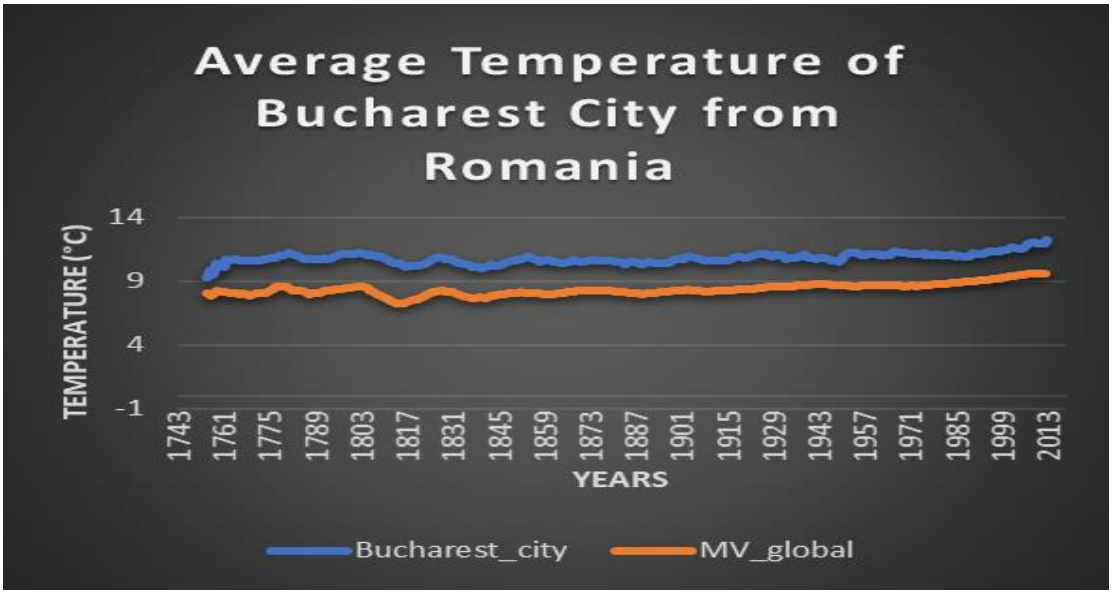
2. The Excel Method

After I have obtained the Bucharest city and global data in two separated "csv" file I created a new file in which I combined those two databases in order to calculate the Moving Averages of the Bucharest city and global temperatures. For the moving averages I used a baseline of 10 years using a *"=AVERAGE(D2:D11)"* variance formula like in the photo below:



	A	B	C	D	E	F	G	H	I
1	year	city	country	avg_temp	year	avg_temp		MV_city	MV_global
2	1743	Bucharest	Romania	5.31	1750	8.72			
3	1744	Bucharest	Romania	12.95	1751	7.98			
4	1745	Bucharest	Romania	2.28	1752	5.78			
5	1750	Bucharest	Romania	11.48	1753	8.39			
6	1751	Bucharest	Romania	12.01	1754	8.47			
7	1752	Bucharest	Romania	5.31	1755	8.36			
8	1753	Bucharest	Romania	10.74	1756	8.85			
9	1754	Bucharest	Romania	10.81	1757	9.02			
10	1755	Bucharest	Romania	10.61	1758	6.74			
11	1756	Bucharest	Romania	11.39	1759	7.99		9.289	8.03
12	1757	Bucharest	Romania	11.16	1760	7.19		9.874	7.877
13	1758	Bucharest	Romania	9.51	1761	8.77		9.53	7.956
14	1759	Bucharest	Romania	10.56	1762	8.61		10.358	8.239
15	1760	Bucharest	Romania	10.09	1763	7.5		10.219	8.15
16	1761	Bucharest	Romania	11.11	1764	8.4		10.129	8.143
17	1762	Bucharest	Romania	10.93	1765	8.25		10.691	8.132
18	1763	Bucharest	Romania	10.11	1766	8.41		10.628	8.088
19	1764	Bucharest	Romania	11.13	1767	8.22		10.66	8.008
20	1765	Bucharest	Romania	11.02	1768	6.78		10.701	8.012
21	1766	Bucharest	Romania	10.87	1769	7.69		10.649	7.982

After figuring out the Moving Averages I was able to produce a graph chart of lines that better represent the cities and global temperatures in correlation to 1743 to 2015 years using an black background to better observe the lines.



Analysis and conclusions

1. They both have a positive slope which is gradually increase from 1743, which means the global and Bucharest temperature gradually increase, both being linked to a potential global warming evolution.
2. The Bucharest temperatures are general higher than global average. This is due to continental climate of the Romania country.
3. There is a suddenly decrease of temperature from 1803 that stabilize approximately in the year 1831.

Bonus:

I also search the internet for information regarding the graph charts that can be done in Python and by so I was able to do the same project in Python with the same results, using a 12 years period, using an orange color for global line and a green color for Bucharest line.

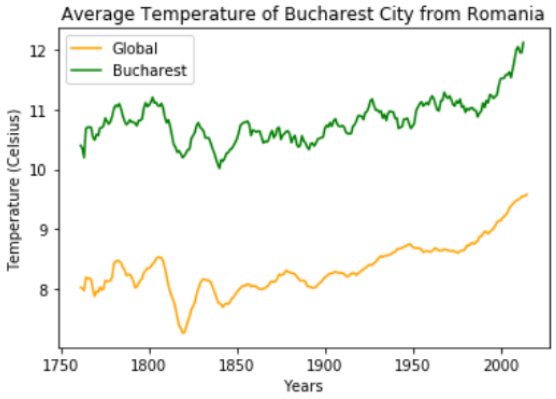
jupyter weathertrends Last Checkpoint: a day ago (unsaved changes)

File Edit View Insert Cell Kernel Widgets Help

Run Code

In [6]:
```
tempglobal = pd.read_csv('global_data.csv')
Moving_Averages_global=tempglobal['avg_temp'].rolling(12).mean()
tempcity = pd.read_csv('Romania_Bucharest_city_data.csv')
Moving_Averages_city=tempcity['avg_temp'].rolling(12).mean()
```

In [7]:
```
plt.plot(tempglobal['year'],Moving_Averages_global,label='Global',color="Orange")
plt.plot(tempcity['year'],Moving_Averages_city,label='Bucharest',color="Green")
plt.title("Average Temperature of Bucharest City from Romania")
plt.xlabel("Years")
plt.ylabel("Temperature (Celsius)")
plt.legend()
plt.show()
```



I done this project using Jupyter Notebook. My opinion is that Python display data in a more nicely matter giving me the possibility to observe temperature changes more easier.