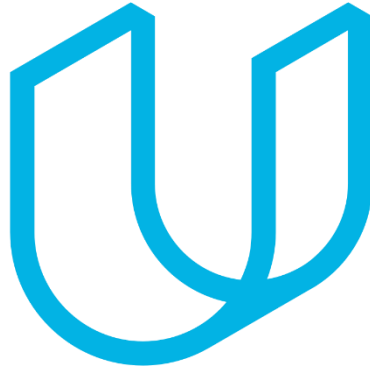


UDACITY DATA ANALYST NANODEGREE



UDACITY

Vasileios Garyfallos

February 2020

PROJECT 1: EXPLORE WEATHER TRENDS

OUTLINE OF STEPS TAKEN:

1. Using SQL, I extracted the local and global data and downloaded them as .csv files.

SQL queries used:

local data

```
SELECT * FROM city_data  
WHERE city='Berlin' AND country='Germany '
```

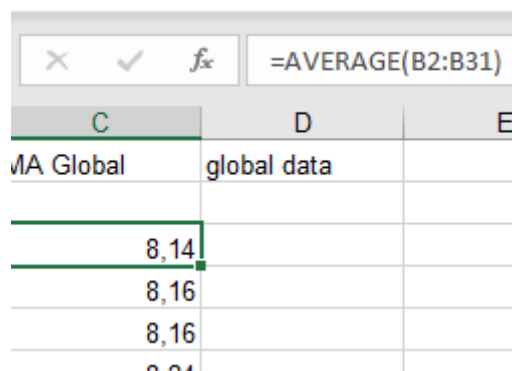
global data

```
SELECT * FROM global_data
```

2. Using MS Excel, I calculated the MA for both global and local data. I calculated the MA using sample for 30 consecutive years and by using the following formula:

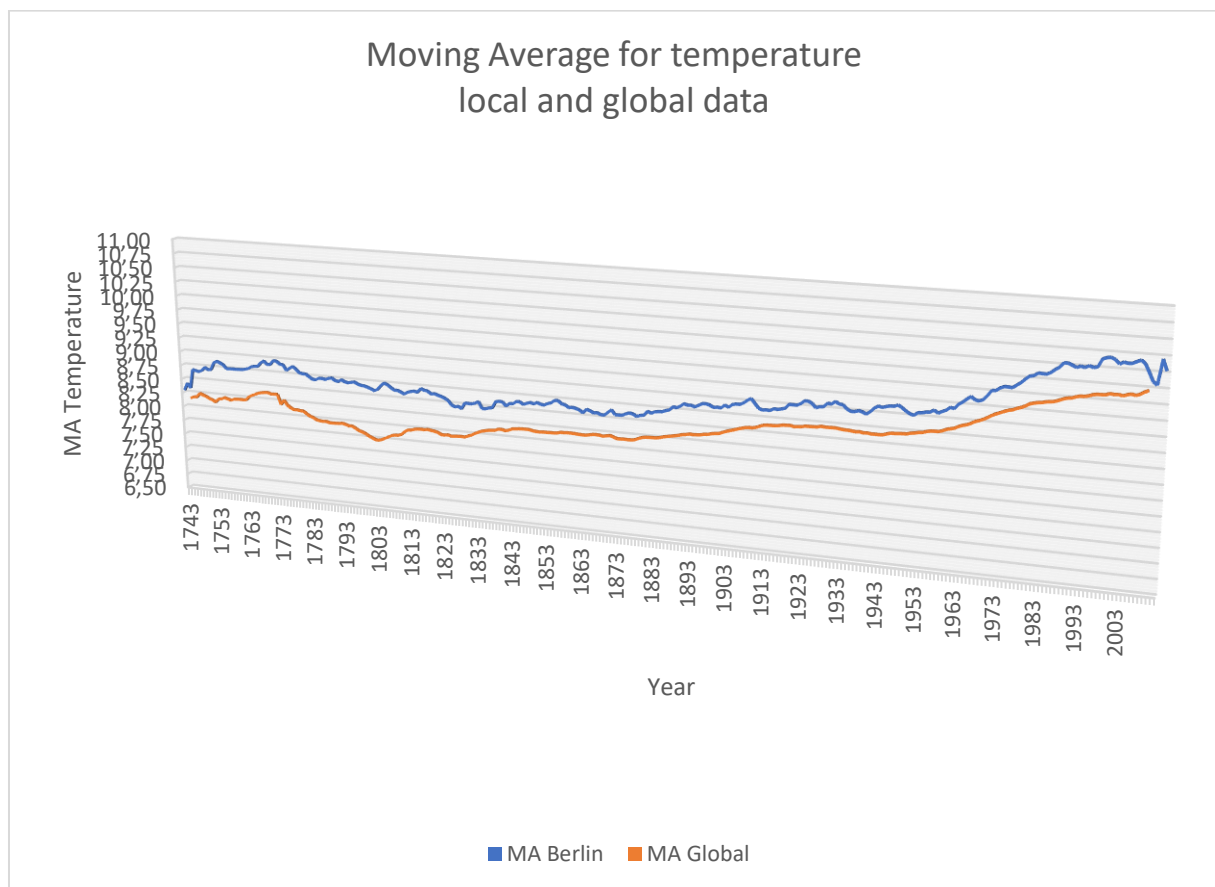
$$(AVGy1 + AVGy2 + AVG...yX + AVGy30) / 30$$

Using sample of 30 years, I was able to smooth out the lines of the chart in order to give more valuable output, filtering out the fluctuations.



C	D	E
MA Global	global data	
8,14		
8,16		
8,16		
8,14		

3. I created a double line chart in MS Excel in order to plot the moving average.



OBSERVATIONS:

1. The global average temperature since 1743 shows an upward trend due to global warming. There are small periods of cooling (1781 – 1811).
2. Berlin's average temperature over the observed timeline is generally higher than the global by approx. 0,5 degrees and is timeline-consistent, except for the period 2000 – 2005, in which Berlin's temperature has been rapidly decreased, and thus the two lines tend to intersect in this period. (Pearson strong positive linear correlation $r = 0,95$)
3. Berlin's temperature fluctuations tend to be higher than the global temperature fluctuations for the entirety of the observed timeline. (Berlin's $\sigma = 0,50$ whereas global $\sigma = 0,67$)
4. The coolest year globally was 1752 with an annual average temperature of 5,78 degrees. The hottest year globally was 2015 with an annual average temperature of 9,83 degrees.