Weather Trends

Abdelrahman Torki

Steps:

- First, I extracted the data from the online workspace using the following SQL commands:
    - ‘SELECT * FROM global_data;’ to extract the global data.
    - ‘SELECT * FROM city_data WHERE city = 'Cairo';’ to select local data.
    - ‘SELECT * FROM city_data;’ to get all of the city data.
    - ‘SELECT * FROM city_list;’, but the city_list is some kind of redundant data.
- Second, I tried to use the exponential moving average (EMA) instead of simple moving average as exponential moving average gives more weight on recent observations thereby alleviating the lagging indicator in simple moving average (SMA), but when I compared the two plots, there wasn’t much difference and I ended using SMA [figure1]. I also preferred a 10 year moving average over 7 year as the curve was smoother more linear and consequently better correlation coefficient. I used Excel’s Correl() function to calculate the correlation coefficient for linear regression.
- Then, I tried to use contrast dim colors in visualizing data.

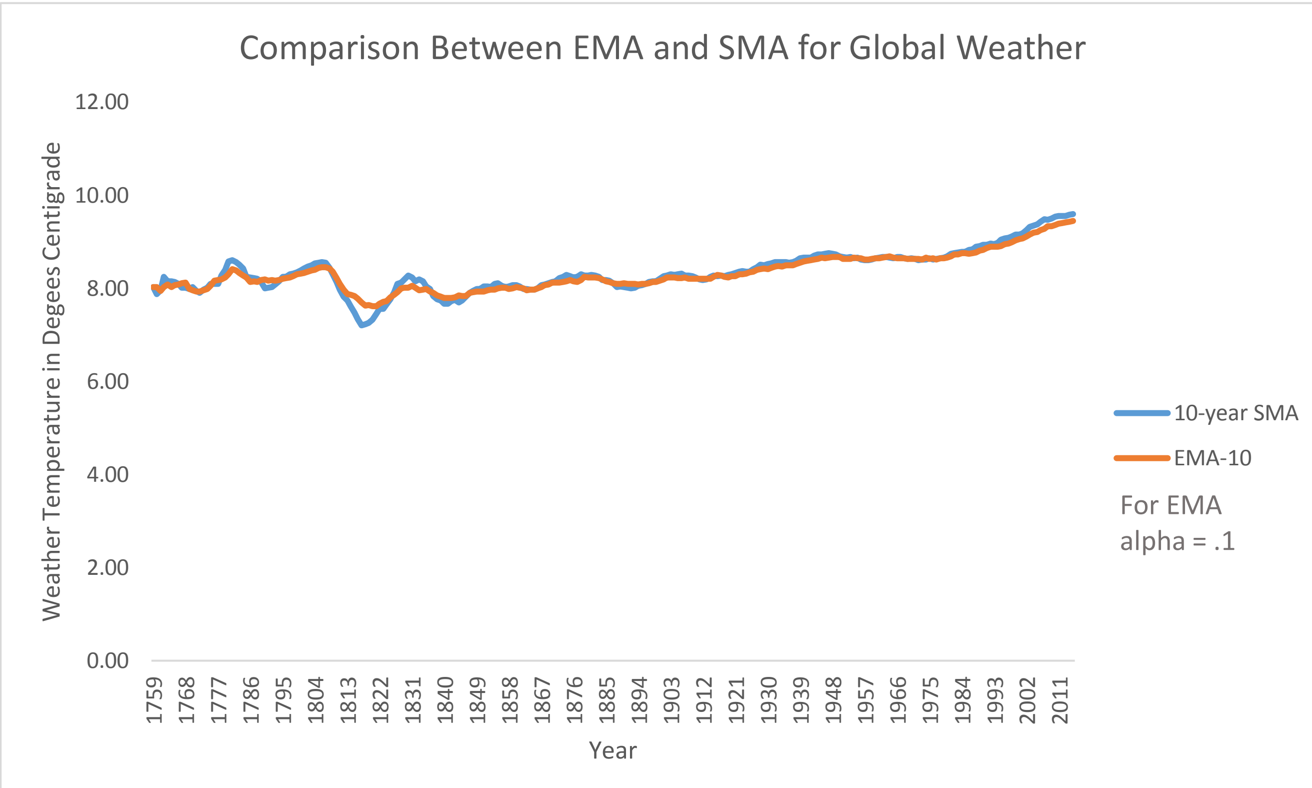

Figure 1 comparison between EMA for alpha = .1 and SMA for global_data where SMA using 10-year for rolling average and EMA uses a 10 rolling year as a first entry in the averaging process. Formula used for EMA:
St = alpha * Yt + (1 – alpha) * St
where Yt is the previous closing entry (the previous average global temperature) and St is the current exponential moving average calculated entry.

Global Trends:

- First, I studies global data alone from year 1750 up to 2015. Using a decade rolling average to smooth the graph.
- As we see in the below graph the globe gets hotter especially from the nineties to reach its climax in 2015 (9.95 degrees).
- The correlation coefficient for linear regression is .751 closer to 1 than 0. Considering the correlation coefficient and that moving averages are lagging indicators, it's highly probable that the word gets hotter with every upcoming year according to the trend line.
- In 1816, there is a great drop in global weather temperature and I when I searched, I found that this year was called 'Year Without Summer' because of severe climate abnormalities happened in this year whose cause are believed to be the massive 1815 volcanic eruption of Mount Tambora.

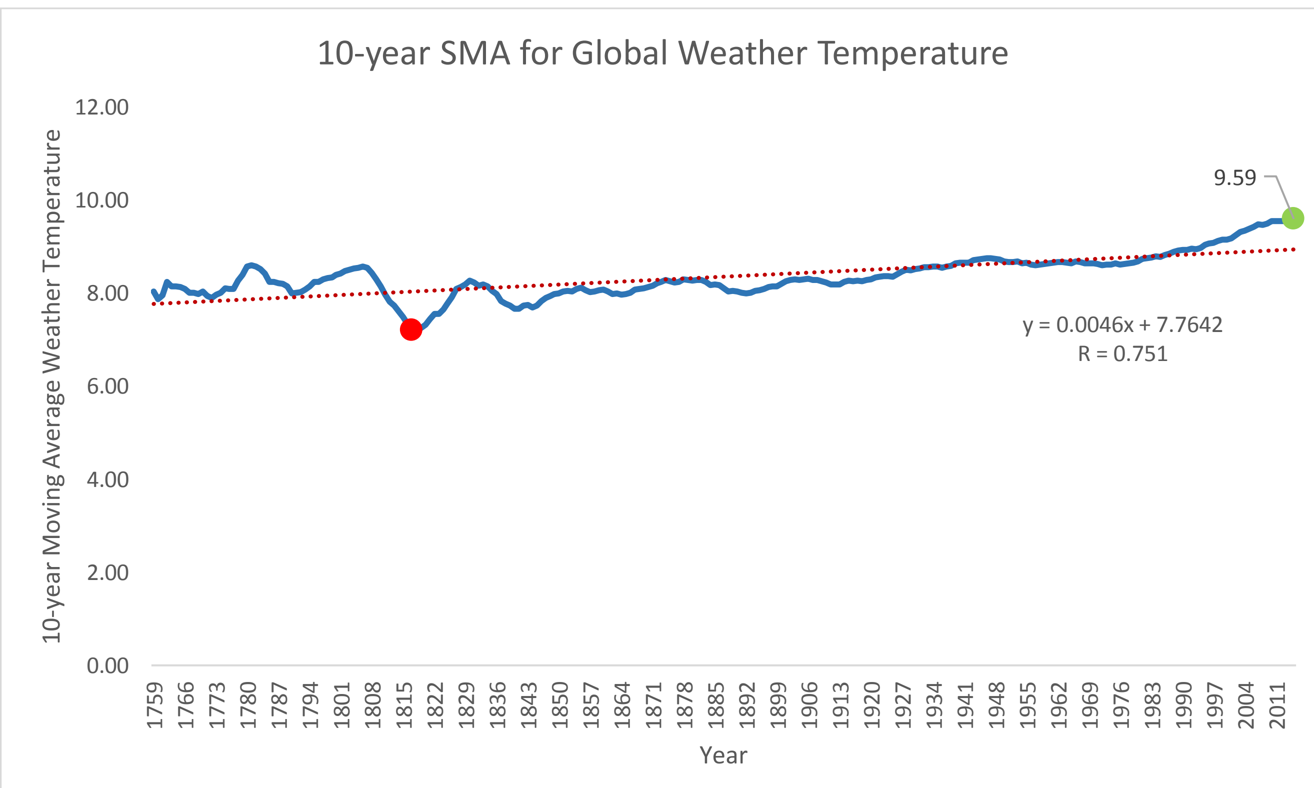

Figure 2 global rolling average temperature from 1759 to 2015

Local Trends (Cairo):

- The rate of increase in temperature in Cairo (.0079 degree/year) is almost double the rate of global increase (.0046 degree/year), but since the period of global data are more expanded, I compared the global weather and local ones over the same common years (figure 4) and found that the rates are almost the same (.0079 for Cairo and .0076 for global data). Hence, Cairo’s temperature increasing rate is a little bit faster than the globe in general.
- In years 1816 to 1818, there is a drop in temperature in Cairo that resembles the global data drop.
- The average temperature for Cairo is generally hotter than the average global temperature and that is consistent over time as illustrated in figure 4.
- There is an increase tendency in temperature in Cairo from nineties as in global data reaching its climax at 2013 and probably both 2014 and 2015 in Cairo (which we don’t have data of) are hotter based on global data in (2014, 2015) and the trend line of local data.
- The correlation coefficient is high for both global data (.925) and local data (.866) between (1817-2013) conforming the conclusion that in upcoming years both local and global weather will get hotter.

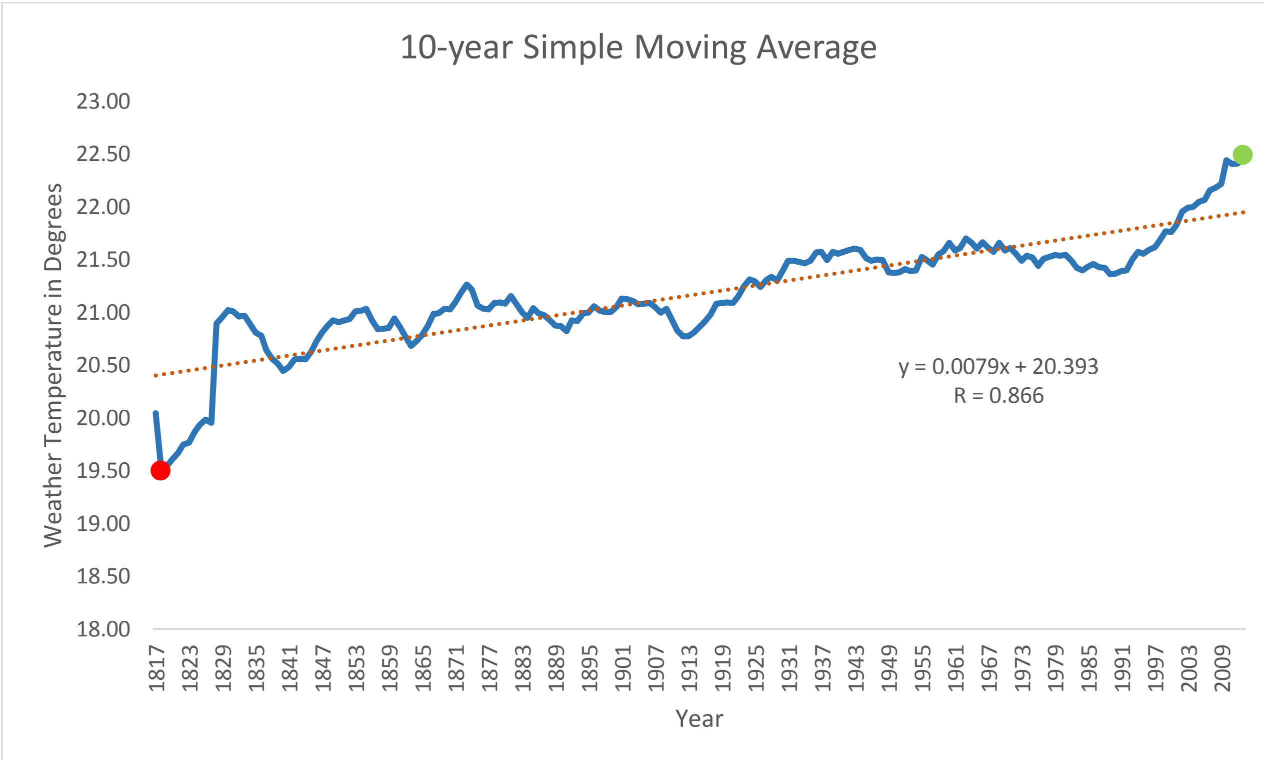

Figure 3 Cairo rolling average temperature from 1817 to 2013.

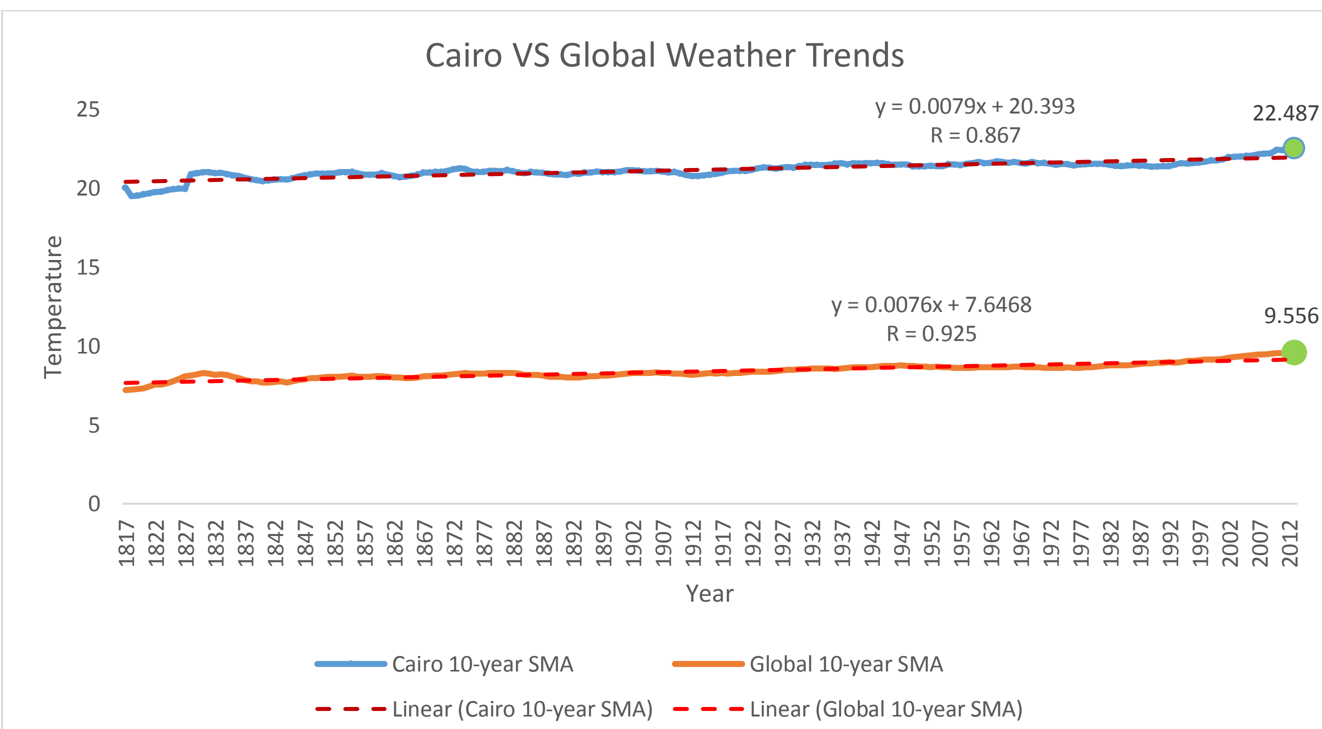

Figure 4 Cairo vs global temperature trends.

Estimating local data from global data:

Local data can't be estimated from global_data dataset. As every point in global_data is an average. For example, at 1990 global average temperature is 9.23 which is calculated be averaging all temperatures of the globe including Cairo's data which could be much higher than or much lower (outlier) or close to the global average and there's not enough data to estimate Cairo's temperature unless you know all the temperatures of other cities included in calculating global temperature for this year. We can't even say whether the local trend is the same as the global trend or not as this trend may be constituted of those possibilities:

- Cities that all of them have closing trend to global trend.
- Cities of opposite trends and those close to the global trend outweigh those of opposite trends.
- Most of the cities of opposite trends, but some few cities are outlies and their extreme trends oppose the later. Hence, the global trend doesn't actually represent the trend of the majority.

On the other hand, if we used city_data dataset. We may get geographical data for every city (longitude and latitude) and apply a regression algorithm on those cities close to Cairo and hence give good estimate of Cairo's temperature, but not very rough estimation though.

Finally here's a graph of weather trends for my favorite cities:

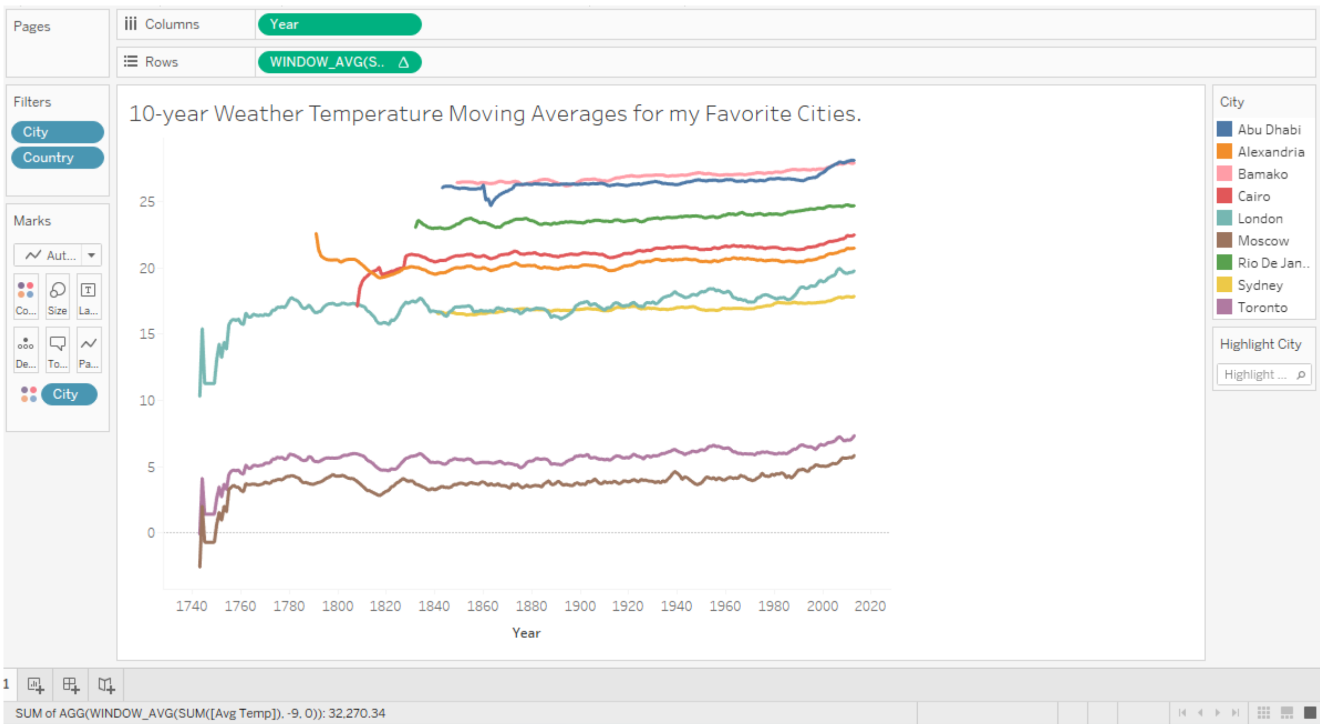

Figure 5 This graphs are generated using Tableau Public