**Make observations** about the similarities and differences between the world averages and your city’s averages, as well as overall trends. Here are some questions to get you started.

* Is your city hotter or cooler on average compared to the global average? Has the difference been consistent over time?
* “How do the changes in your city’s temperatures over time compare to the changes in the global average?”
* What does the overall trend look like? Is the world getting hotter or cooler? Has the trend been consistent over the last few hundred years?

Your submission should be a PDF that includes:

* **An outline** of steps taken to prepare the data to be visualized in the chart, such as:
  + What tools did you use for each step? (Python, SQL, Excel, etc)
    - SQL was used, as instructed, on the website to extract the data. Python was used for calculations and visualizations. Packages used are pandas, numpy, matplotlib, scipy.
  + How did you calculate the moving average?
    - The MA is calculated by using a user-defined *n* window size for each year that averages (calculates the mean) values *n* number of years back.
  + What were your key considerations when deciding how to visualize the trends?
    - I want to visualize mainly two things: comparison of *trend* between the two measurements together with correlation. To visualize a trend, a trend line will be a good addition. Therefore, a line plot is sufficient, but also using a simple regression to fit the trend line. Comparisons of the slope coeffiecient for each model for each dataset can tell us the how much the temperature is rising in average.
    - I want an easy way to see the *distributions* of the two datasets to see if my city is warmer of colder than the global average. A histogram with sufficient bin size can tell us that (a line plot may also be obvious, but a histogram is good at presenting this).

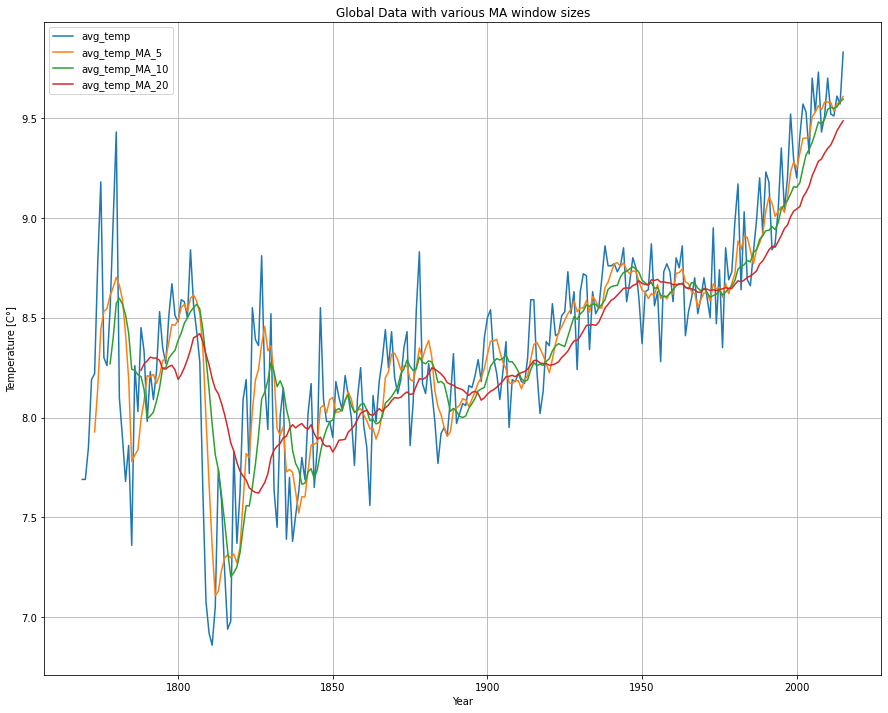
**Line chart** with local and global temperature trends

Figure 1: Global dataset with MA for different window sizes, for comparison.

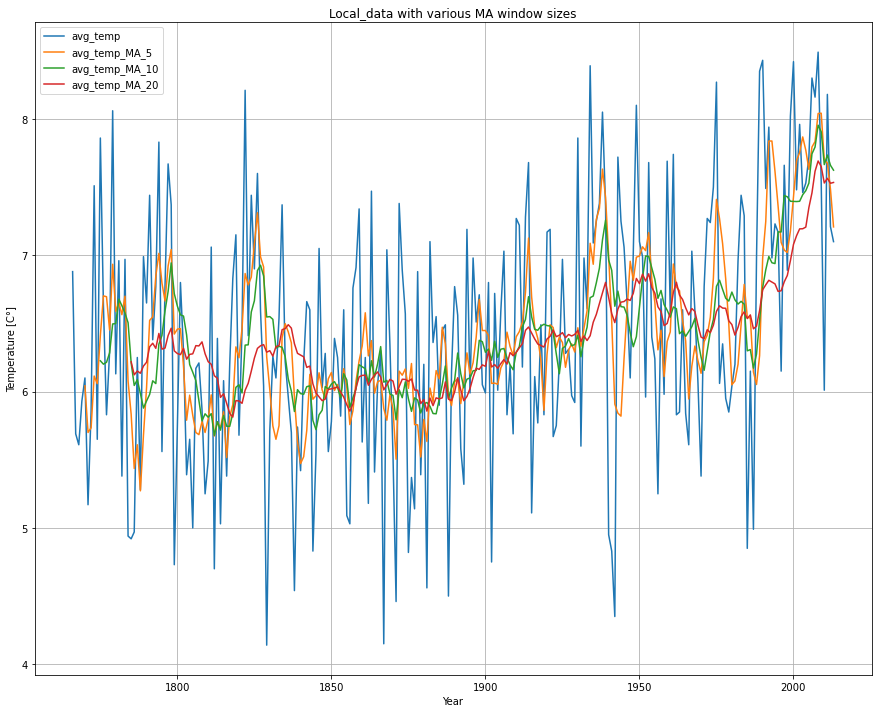


Figure 2: Local dataset with MA for different window sizes, for comparison.

Initially to “remove” the noise in the data, we compare different implementations of a moving average presented in Figure 1 & Figure 2 for each dataset. The higher values on the window size, the smoother the line. Though, when increasing the window size, we also see that the measurement is lagged. When moving further in the report, window size of 20 is used.

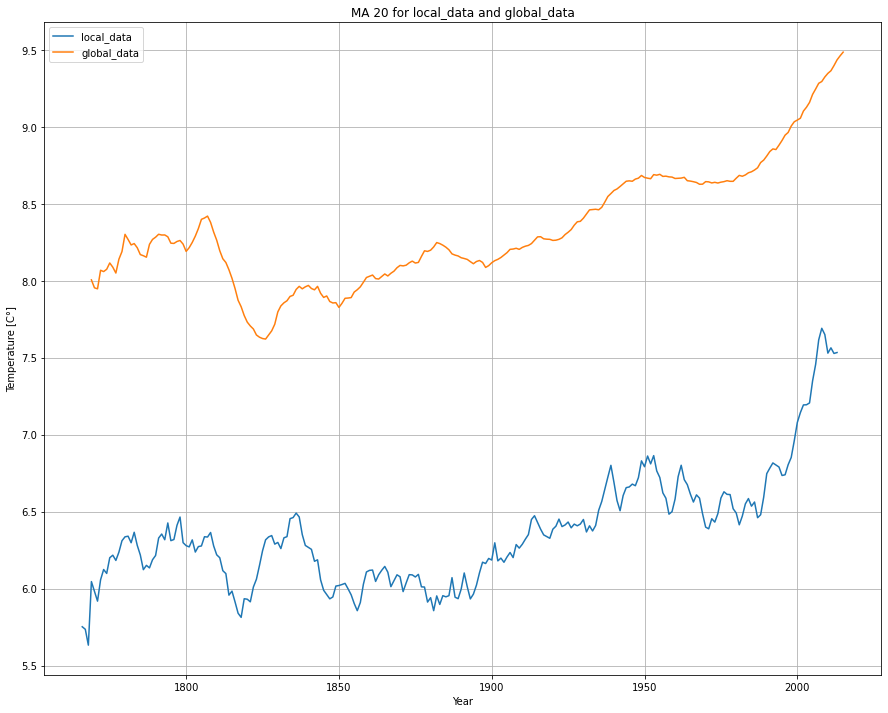


Figure 3: Moving average of both datasets where window size = 20 for MA.

Figure 3 presents the moving average for both the datasets where window size = 20. It is obvious that the global average temperature is higher than the temperature in Stockholm through the years. Noticeable are also the dips and peaks throughout the history that seem to have a high correlation.

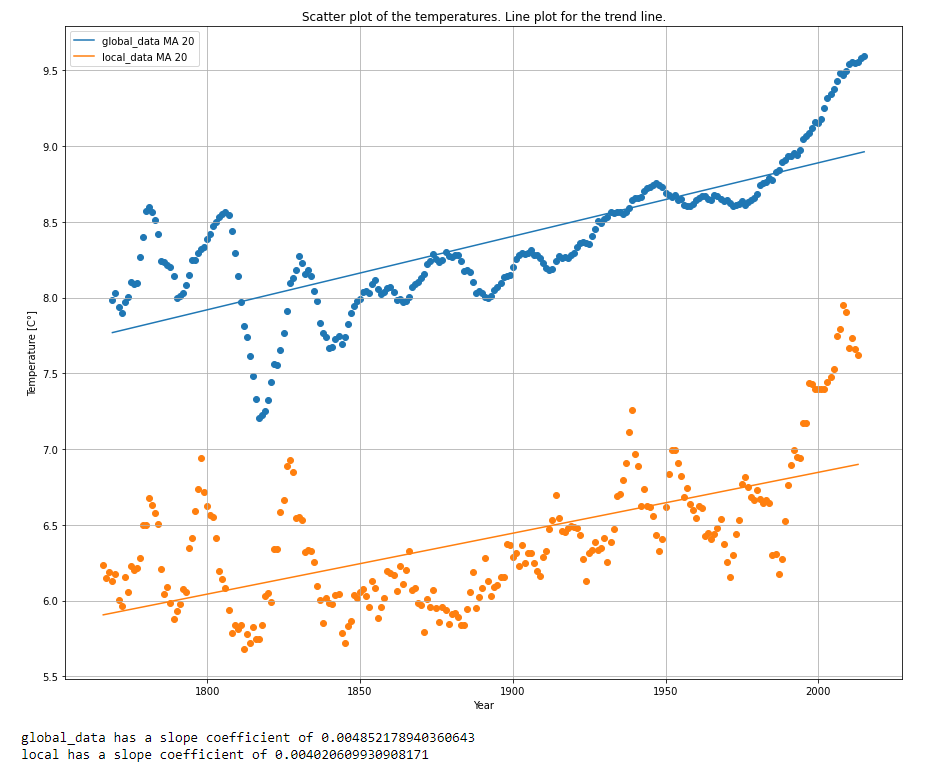


Figure 4: Line/Scatter plot of both datasets with a linear regression implemented. The slope coeffiecient is presented under the plot.

* At least **four observations** about the similarities and/or differences in the trends

Select \*

FROM city\_data

WHERE 'Stockholm' in (city\_data.city)

Select \*

FROM global\_data