Weather Trend Exploration

Outline:

The purpose of this exercise was to import data from an SQL server and perform some basic analysis of the global and local weather trends.

Tools used in the process:

- SQL query - extracting data from the SQL server
- Python – Calculation of rolling averages and visualisation.

The query to extract the data looked as follows:

```
SELECT * FROM global_data;

SELECT * FROM city_data
WHERE country = 'Poland' AND city = 'Warsaw';
```

In order to calculate the moving averages, I imported the pandas library [1] and used the rolling [2] module in order to perform rolling window calculations.

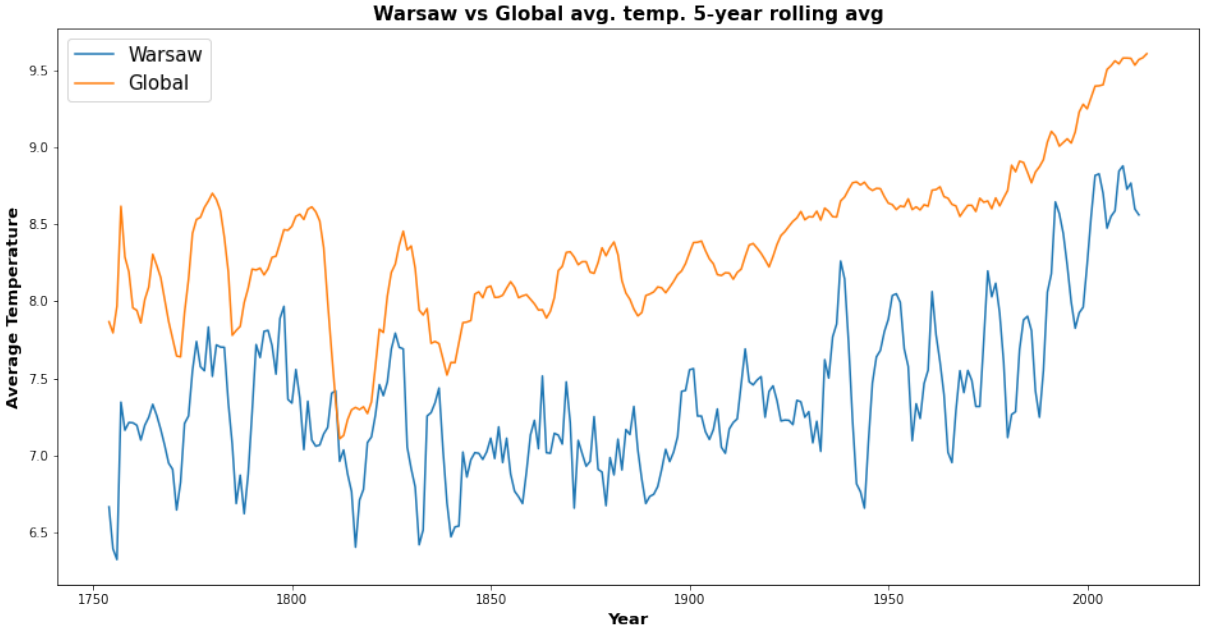
The rolling averages were calculated in the following way:

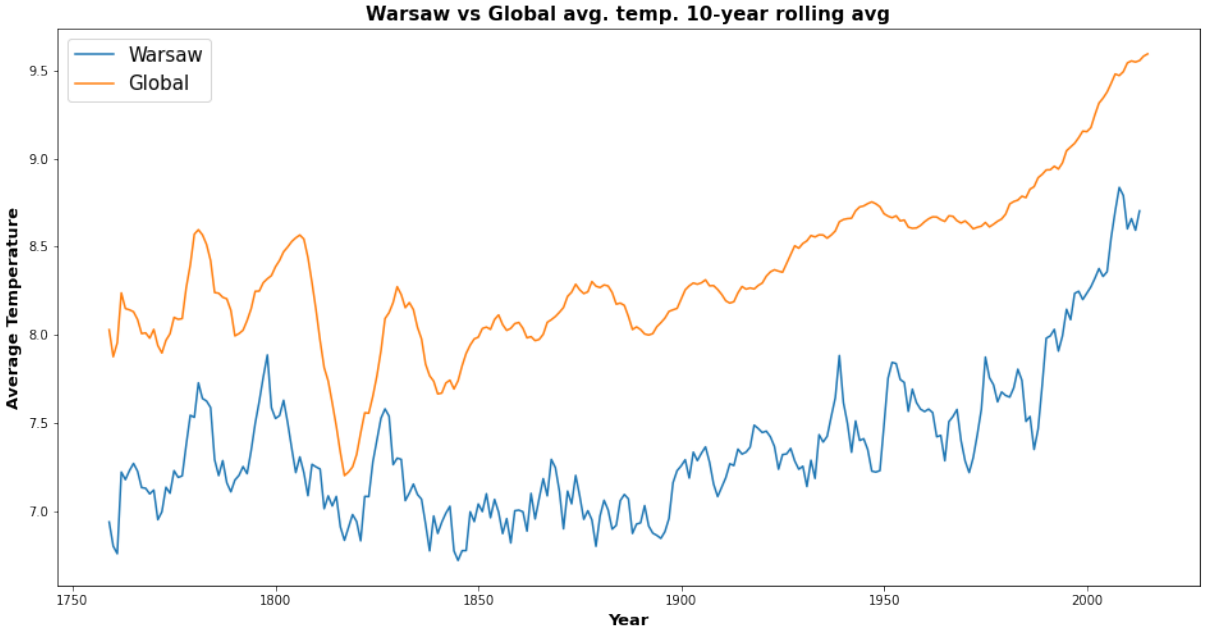
```
df_city["5y_mean"] = df_city.avg_temp.rolling(window=5).mean()
df_city["10y_mean"] = df_city.avg_temp.rolling(window=10).mean()
df_city["15y_mean"] = df_city.avg_temp.rolling(window=15).mean()
```

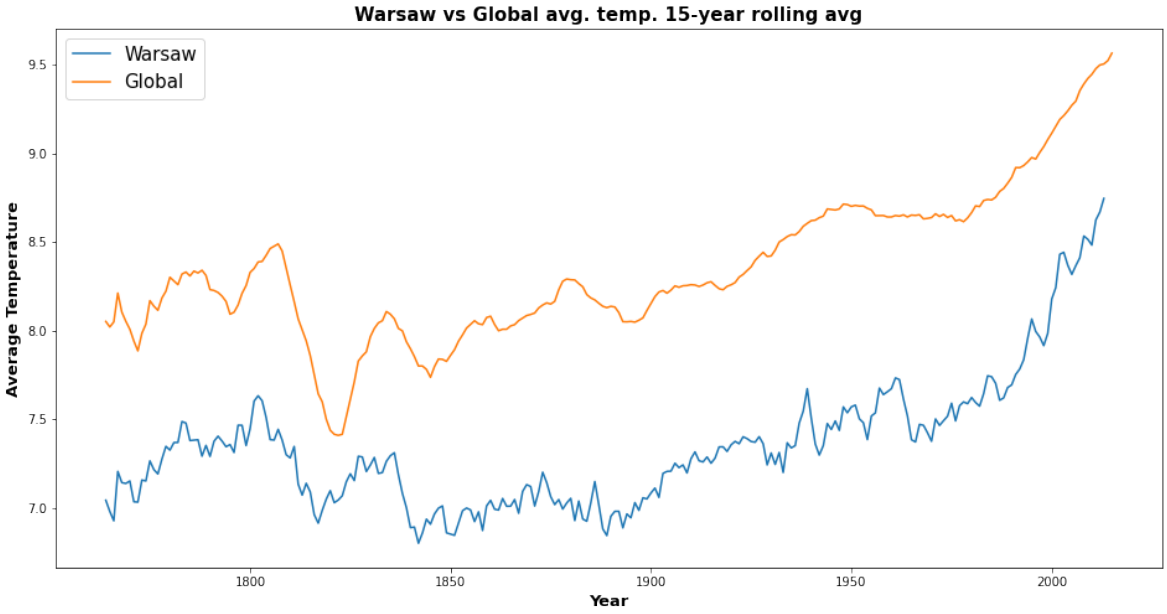
Considerations regarding trend visualisation:

Plotting yearly averages yielded extremely noisy plots, that were essentially unreadable.
By calculating the 5,10 and 15-year rolling averages, I managed to smooth out the plots, making them readable.
However, choosing the window size has its consequences, too small of a window leads to little noise being removed.
While too big of a window results in loss of data and insights, as some trends may disappear on too large of a scale.
Exploring with 3 different window sizes allowed me to choose the most optimal one.
The graphs presented below, are visualisations of these 3 different rolling average windows.

[1] https://pandas.pydata.org/
[2] https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.rolling.html


Warsaw vs Global avg. temp. 5-year rolling avg
Warsaw
Global
Average Temperature
Year
9.5
9.0
8.5
8.0
7.5
7.0
6.5
1750
1800
1850
1900
1950
2000


Warsaw vs Global avg. temp. 10-year rolling avg
Warsaw
Global
Average Temperature
Year
9.5
9.0
8.5
8.0
7.5
7.0
1750
1800
1850
1900
1950
2000



As mentioned earlier, the larger the window size, the smoother the plot.
I believe that the second graph, with the 10-year rolling average, provides most insight into the local and global weather trends.

Observations:

1) In both cases, the Global and Local weather data, the average temperature is rising from 1845 onwards. Suggesting that the climate is getting warmer, which might be linked to the Industrial Revolution which took place in the 19th century.

2) Throughout the entire period, the Local average temperature is consistently lower than the global average temperature. That is most likely due to Poland's geographical location.

3) The local average temperature is more or less similar to the global average temperature. The gaps seems to be closing from the year 2000 onwards.

4) The changes in the local average weather have a higher magnitude comparing to the global average weather. That however, is probably due to the global weather being an aggregate of all of the countries in the world (Including the Local one)