

## Project 1 – Explore Weather Trends

The **first step** involved using SQL to extract (as .csv file) the yearly average temperatures for Berlin and global average temperatures.

For Berlin, the following query was used:

```
SELECT *  
FROM city_data  
WHERE city = 'Berlin'
```

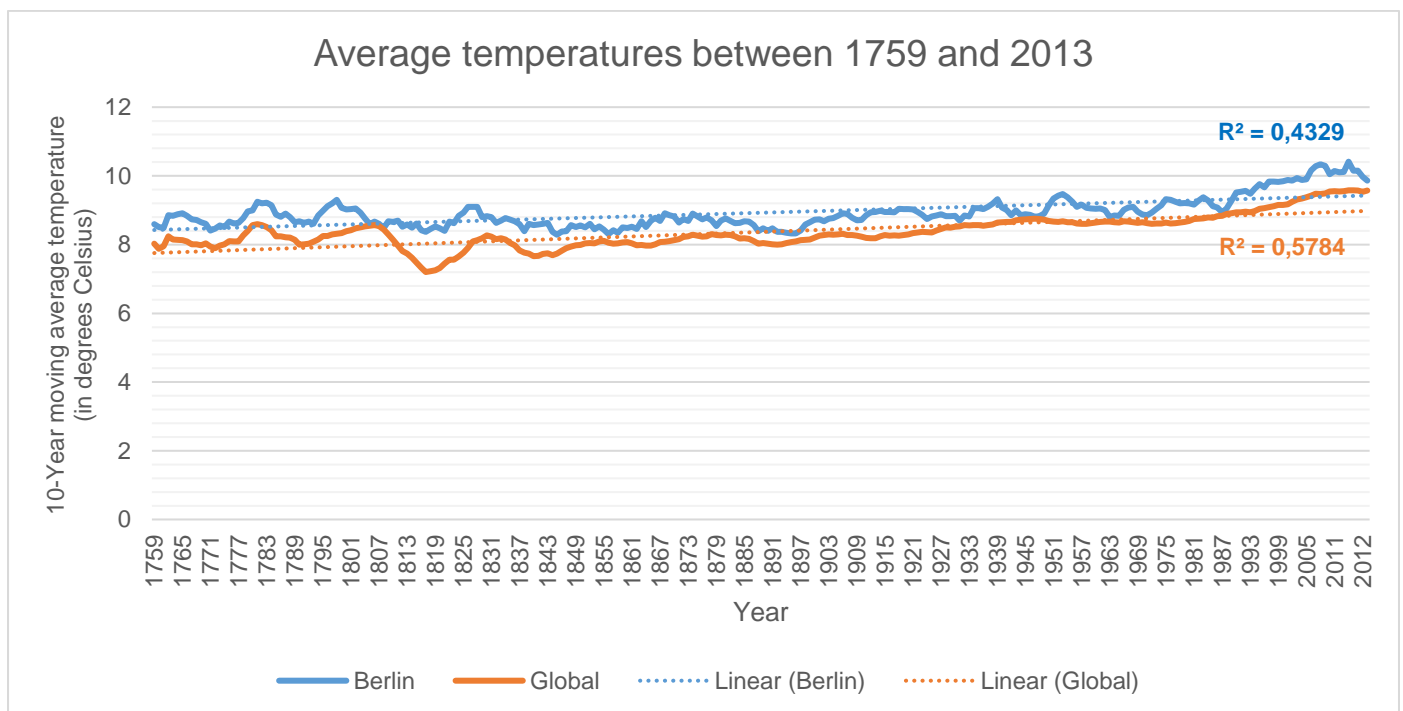
For Global, the following query was used:

```
SELECT *  
FROM global_data
```

The **second step** was to use Excel to prepare the data. The following steps were taken:

1. Integrate both datasets by putting the yearly values side by side.
2. Calculate the 10-year moving average temperatures for both Berlin and global. This was done by selecting 10 rows and calculating that average. The resulting value was the average for the last year in the 10-year period (e.g. the average between 1950-1959 was the value for the year 1959).
3. Remove from the analysed range all the years without values in both Berlin and global in order to avoid missing values in the visualisation, while maintaining the longest possible period of years.
4. Add a linear trendline to observe the trend in both Berlin and global and add their  $R^2$  to show the reliability of the trendlines.

The resulting line graph was the following:



### OBSERVATIONS

1. **Berlin has a higher average temperature than the global average.** Over the observed period, Berlin's average temperature was 8.92°C while that of the global average was 8.36°C.
2. **The temperature in Berlin shows more noise than the global average.** As possible to observe in the graph lines, Berlin shows more variance in temperature due to more “ups and downs”, while the global average tends to be steadier and more constant.
3. **Both in Berlin and globally the average temperatures have been increasing.** Climate change deniers be aware: the trendlines' upward slope indicates that average temperatures have been getting higher over time.
4. **Comparing averages over a short period of time can be misleading.** Even though the general trend is for the average temperature to increase, there are periods where the average temperatures decrease year on year (e.g. clearly seen in the beginning of the 19<sup>th</sup> Century). This means one cannot accurately infer insights about long-term changes by using short-term values.