Exploring Weather Trends Project

Extracting the Data

When starting the project, I used SQL to extract the necessary data from the global_data and city_data tables (SQL below). I initially selected too much data as I only filtered on "city = 'Barcelona'", but there is also a Barcelona, Venezuela so I had to include "country = 'Spain'".

```
-- city data for Barcelona, Spain
SELECT year, avg_data
FROM city_data
WHERE country = 'Spain'
AND city = 'Barcelona'
ORDER BY 1, 2

-- global data
SELECT *
FROM global_data
ORDER BY 1, 2
```

Extracting and Importing the Data

I then exported the data as CSV and imported them into a Google Sheets document via File → Import → Append to current sheet, allowing the separator to be selected automatically.

Tidying the Data

I noticed that some of the city data was missing but it was only from years before 1750, and since the global data only started from 1750 I decided to ignore these rows. I then chose a 10-Year Moving average since decades are a natural comparison and more than that would be smoothing too much. I calculated it as shown in the tutorial with "AVERAGE(B2:B11)" in Excel.

Plotting the Data

Once I had the city data in one tab and the global data in another, I plotted both as a line graph. I then added a title, axes labels, and renamed the Avg Temp columns in both sheets to have the labels in the legend. The resulting graph is shown below:
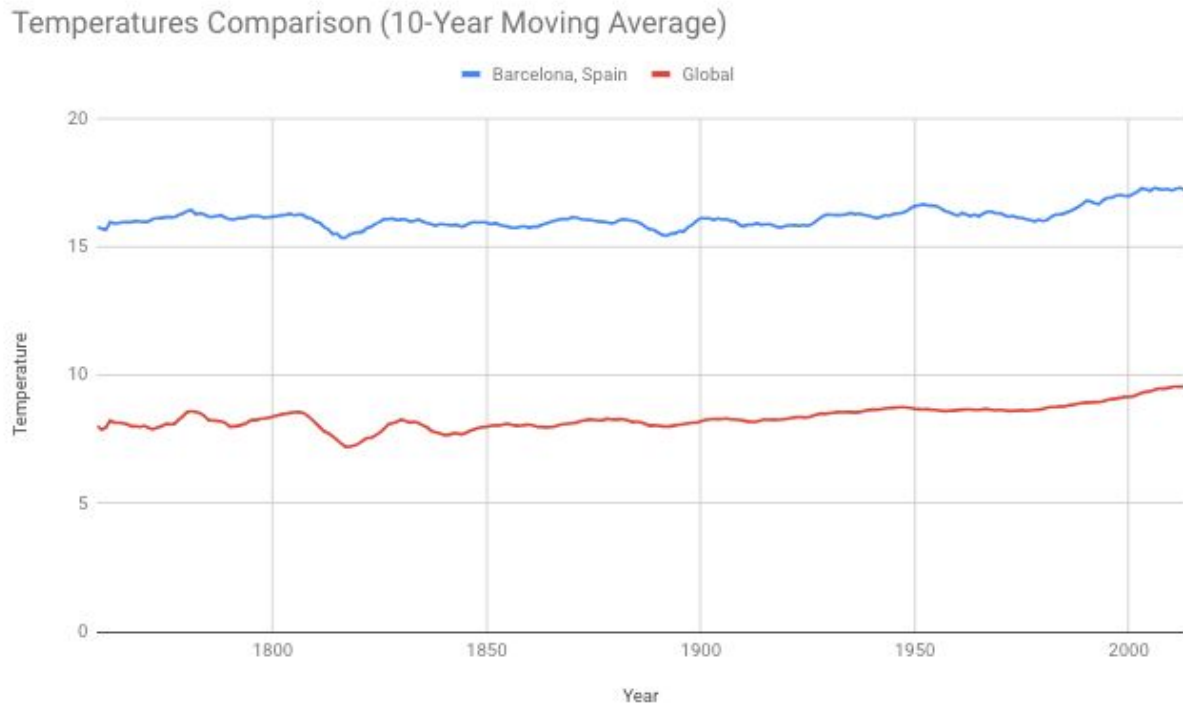
Observations

- The average temperature in Barcelona, Spain is approximately 8 degrees higher than the global average
- The difference between the temperature in Barcelona and the temperature Globally remains fairly constant over time. This is emphasised by the same drop around the year 1820.
- Barcelona's temperature seems to increase more sharply around the year 1980 than globally, suggesting one of two things: Either the global temperature increase is being dampened by other countries without this increasing trend in temperatures, or the temperature in Barcelona is increasing faster than the average.
- Barcelona and Global data have a correlation coefficient of 0.88 (using Excel's CORREL function)