Project 1 - Explore Weather Trends

Preparing the data

For this project, I used SQL from the Udacity platform to extract the data as a CSV, and then imported both files into a Google Sheets document, in the following steps:

1. I imported each CSV file into separate sheets inside the same document to have a better view of each file.
2. After this, I calculated the *moving average of 7 years* using the function `AVERAGE` from the Google Sheets, selecting the last 7 observations (years) and calculating it. Then, I accepted the suggestion to autofill the rest of the column with the same function, calculating for all years until the end.
3. As my city (*Curitiba, Brazil*) had 7 years without any observation between 1844 and 1850, the last year had an error as dividing by zero is not allowed (the sum of 7 zeroes values are zero, and is used in the average calculation). To correct this instead of deleting the year and having a blank point in the trendline, I copied the value from the previous year and replicated it to this problematic year (1850), as the differences between consecutive years are very small, so this should not affect the trendline.
4. For plotting the trendline, I created a new sheet to merge both tables (my city and global data), as this was easier for plotting on Google Sheets. To see the big picture, I kept all the years from the global data (from 1750 to 2015) and left the blank observations from my city (it has data only from 1837 to 2013). This way, we can see the global trend over the time, as well as the trend from my city during the applicable period. Also, to help visualize the trend of both series, I put a trendline behind the curves, making it easier to compare both series.

SQL Queries

These are the queries that I used to extract the data from the database:

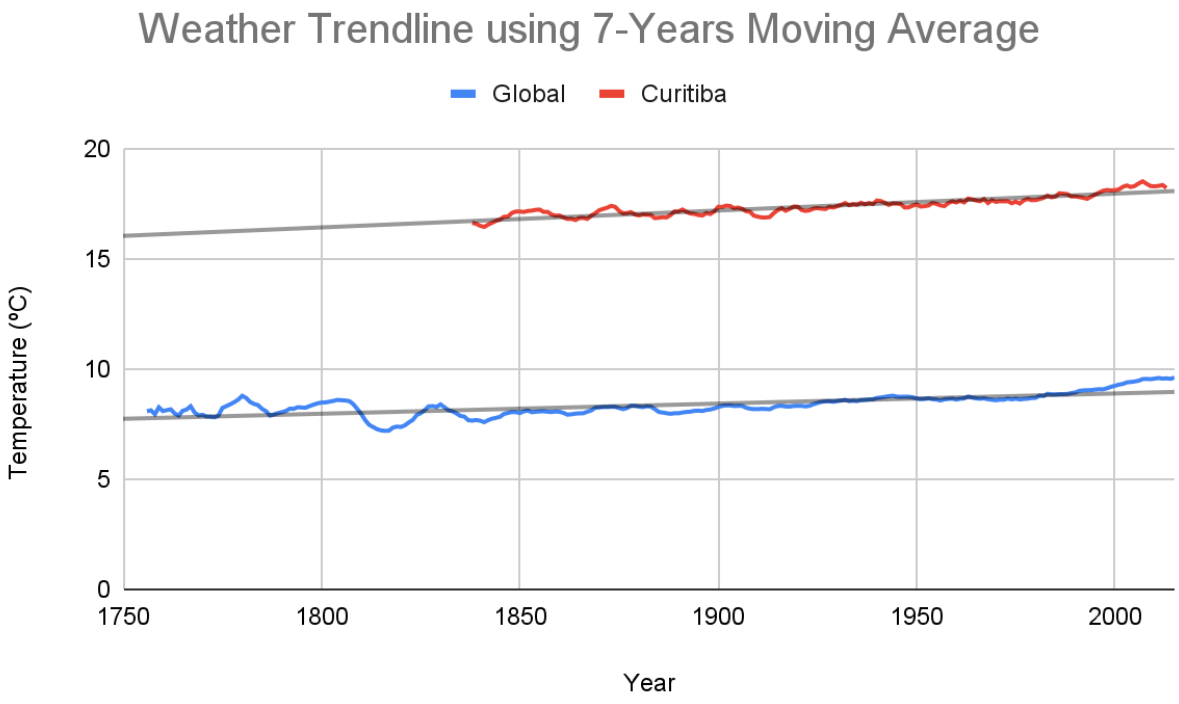
My City Data

```
SELECT *
FROM city_data
WHERE city = 'Curitiba' AND country = 'Brazil'
ORDER BY year;
```

Global Data

```
SELECT *
FROM global_data
ORDER BY year;
```

Line Chart



Observations

The first characteristic to note is that my city is almost 10 ºC hotter than the global average, and the difference was constant during the years. One interesting thing

is that the peaks and valleys in Curitiba's series follow the global ones, but they usually are higher or deeper (more intense), showing a bigger variation between the years than globally.

Also, both the temperature of my city and global temperature have been increasing over the years. This can be seen looking at both trendlines with a positive inclination. In the global series, the first 10 years had averages lower than 8 ºC, and the last decade had averages higher than 9,4 ºC. In Curitiba, the first decade had averages lower than 16,9 ºC, and the last 10 years had averages higher than 18,2 ºC.

The last observation is that the global temperature is increasing over the last 250 years, but not consistently. Between 1750 and 1850 the global temperature was varying a lot, with a high peak around 1775 and a low valley around 1830, with almost a difference of 2 ºC between them in only 60 years. After 1830, the temperature stopped to vary so much between the years and started to follow a slightly crescent trend, looking almost like a line with a small positive inclination.