- Project 1 –
Exploration of weather trends over time: World vs. Denmark

by Sebastian Sbirna

This project demonstrates several found trends in the fluctuations of the average weather temperature of the world in the time interval between the years 1750 to 2013, compared with the local yearly average temperature in Copenhagen, Denmark, where I am currently living in.

Firstly, the data was collected from an online SQL database using the two different commands. The first command, $SELECT * from\ city\_data\ WHERE\ (country = 'Denmark'\ OR\ city = 'Copenhagen');$ filters the data from the *city_data* database to extract only the fields which interest us, namely the ones related to '*Copenhagen*' or '*Denmark*'. The second command retrieves all the global yearly temperature values from the *global_data* dataset, using the command $SELECT * from\ global\_data;$

After these steps, we have obtained two CSV files with the data we need. For each of these data sets (the global, respectively the local data temperature values), we will calculate a moving average over 10 years, in order to smoothen out small abrupt trend changes. The moving average temperature for a year is calculated as taking the sum of the values of average temperature of each of the past 10 years (with the 10th year being included as the current year for which we are calculating the value), after which this sum is divided by the number of years over which we are calculating the average (in this case, 10).

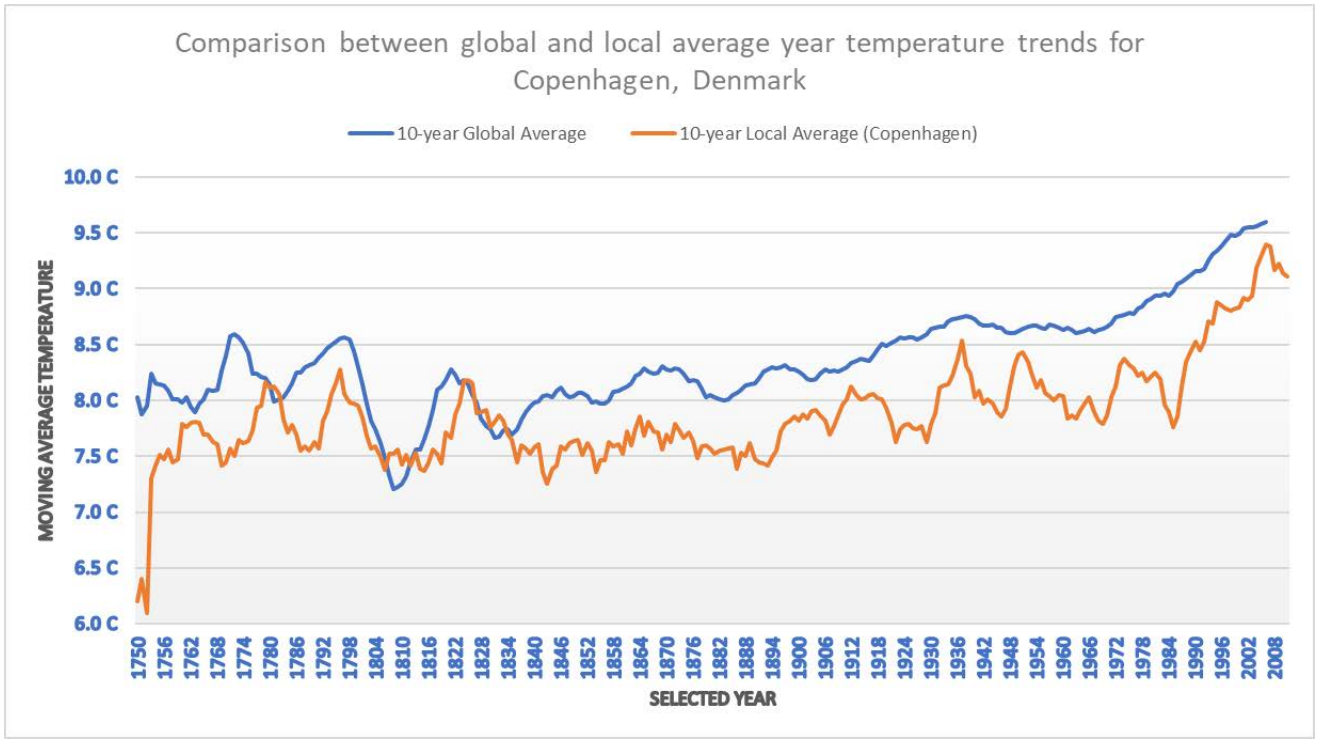

Figure 1. *Line chart graph concomitantly showing global and local temperature moving average trends*

After the moving average for all the possible years in both data sets are calculated, we prepare and present a line chart graph (figure 1), which was made to clearly expose trends in the moving average. Since the two data sets have data starting and ending in different years (1750 – 2015 for *global_data* and 1743 – 2013 for the Copenhagen-filtered values from *city_data*), we have chosen to intersect these two time-range intervals, in order to be sure to always have data from both sets to compare.

From this line chart, we can see the following trends and observations:

- The falling and rising temperature trends in Copenhagen almost always follow closely the trends of the world, with a slight time delay of 4-10 years (*verified by checking the curvatures of the lines throughout the whole time period, however the trends are very visible in: global downfall of 1771, respectively 1780 for Copenhagen; abrupt rise of global temperature around 1808, respectively 1818 for Copenhagen; constant rising since the 1990s, noticed as well in Copenhagen)*. This means that the global weather status greatly impacts Copenhagen's weather temperature as well.
- Even though the global temperature has continued to steadily increase since 1835, the local temperature has suffered a steady cooling, and in the years 1836-1847, Copenhagen's temperature trend has deviated from the global one, resulting in a difference between the two average temperatures of about 0.5°C ever since 1836, although this gap has been fluctuating over the years, varying from as much as about 1.2°C (in *1985*) to about 0.2°C (in *1937* or *2006*). The creation of this gap between the trends has resulted in Copenhagen not being close to the general average city temperature ever since 1836, however the cause of this gap creation is unknown.
- It can be noticed from the graph that, with high degree of generality, Copenhagen's weather is colder than the global average. This is because in only very few years has the local weather reached or surpassed the global average, with such exceptions being in the years 1780-1782 and 1825-1834. This results in the assumption that Copenhagen is and has been, for a long time, one of the colder cities of the world.
- While the global weather trend has been a smooth increase in temperature ever since 1904 (*with some downfall and stagnation between 1942 and 1968*), Copenhagen's average temperature is much more uneven, with constant upturns and downfalls every 10-15 years. This is most likely due to the fact that the global temperature is calculated by taking the average of many major cities around the world, which smoothens out the average data greatly, while Copenhagen's average comes only from one city's data: itself.

There are more observations that could be done from this graph, and the trends presented above have been some of the patterns that could be found from analyzing the moving average value differences between global and local average temperature data.

Explore Weather Trends

REVIEW | HISTORY

Meets Specifications

Congratulations!

EXCELLENT JOB ON THIS PROJECT! YOU HAVE MET ALL THE REQUIREMENTS OF THE RUBRIC.

You demonstrated your ability to retrieve data from a SQL Database and derive interesting, accurate results from the output of your query. You were further able to manipulate this data using external software and create a meaningful visualization to demonstrate your observed results. This is a tremendously important skill and will prove useful throughout your career in data analytics.

Before you move on to your next lessons, take pride in the effort you've put into this project. I hope you found this exercise both challenging and rewarding. Keep up the exceptional work and effort here, and I look forward to seeing you rock those future submissions!

Analysis

- **The SQL query used to extract the data is included.**
- **The query runs without error and pulls the intended data.**

Great work here in extracting the data for Copenhagen and comparing that to global temperatures. Your queries were spot on!

If you're interested in bolstering your SQL mastery with more questions and puzzles, here are a couple websites I often enjoy to looking for extra coding practice for SQL (not affiliated with Udacity):

- https://www.hackerrank.com/domains/sql/select
- https://leetcode.com/problemset/database/
- https://lagunita.stanford.edu/courses/DB/SQL/SelfPaced/courseware/ch-sql/seq-vid-introduction_to_sql/

You'll get a chance to practice increasingly difficult questions and learn how to interact with multiple tables at once. As an example, here is another way to get the data that you want for both Copenhagen and Global temperatures while excluding the empty years in one table output!

```
SELECT city_data.year,
    city_data.avg_temp as city_temp,
    global_data.avg_temp as global_temp
FROM city_data, global_data
WHERE city_data.year = global_data.year
    AND NOT city_data.avg_temp is NULL
    AND city_data.city = 'Copenhagen'
```

Moving averages are calculated to be used in the line chart.

Excellent work here in calculating the 10 year moving average for both Copenhagen and Global temperatures. There's no specific rule of thumb when trying to find the correct number for moving averages, although something like 2 and 3 years would be too short to succinctly smooth out the data and something like 50 years would certainly be too large given the context of our dataset. From that perspective your choice of moving

average works well given what we are trying to achieve with the output. The gap between these two lines is very

apparent here.

I noticed that that there were some missing data from Copenhagen. Here's a link to a blog that details a number of techniques we can use when dealing with missing data. I encourage you to check it out in your free time!

https://www.iriseekhout.com/missing-data/missing-data-methods/

✓

- **A line chart is included in the submission.**
- **The chart and its axes have titles, and there's a clear legend (if applicable).**

The line chart included in your submission looks fantastic! The chart contains a clearly represented title that explains the details of the presented line graph. It also includes a well place legend, intuitive axis labels and clear tick labels. I especially appreciate that you rotated the year values on the x-axis so that they don't overlap one another. This attention to detail really goes a long way to help communicate your results to an audience.

Here is a handy link describing some best practices when creating graphs!

https://www.fusioncharts.com/charting-best-practices/

✓

- **The student includes four observations about their provided data visualization.**
- **The four observations are accurate.**

Great work in looking at the output of your graph and making four observations from them. Your observations are accurate, and can be clearly reflected from the output of your visualization. Well done! Global temperatures are indeed getting warmer, but hopefully with your new skills in data wrangling, you can help better inform and advise others on the importance of climate change and conservation!

DOWNLOAD PROJECT

RETURN TO PATH