Data Analyst Nanodegree Program Project: Exploring Weather Trends

Outline – Executive Summary

In this document, I explore weather data provided by Udacity in the form of an SQL database. ***First****, I extracted the data using SQL and converted my query into a CSV file.* ***Second****, I used Google Sheets to manipulate and filter the data. I calculated the moving average by using the* `=AVERAGE()` *function in Google Sheets across the years, every 3 years.* ***Third****, I created a* ***Line Chart*** *in Google Sheets with two trend lines (for each moving average).* ***Finally****, I made four observations from the visualization.*

The steps are:

1. EXTRACTING WEATHER DATA
2. MANIPULATING WEATHER DATA
3. VISUALIZING WEATHER DATA
4. OBSERVATIONS.

EXTRACTING DATA

To extract the data, I used the following SQL query using the Udacity platforms SQL editor. My intention is to extrapolate global average temperature, global year, London's average city temperature, and city year. I made a condition where the results displayed years equal to 1900 or greater as I wish to explore recent weather trends.

The query is...

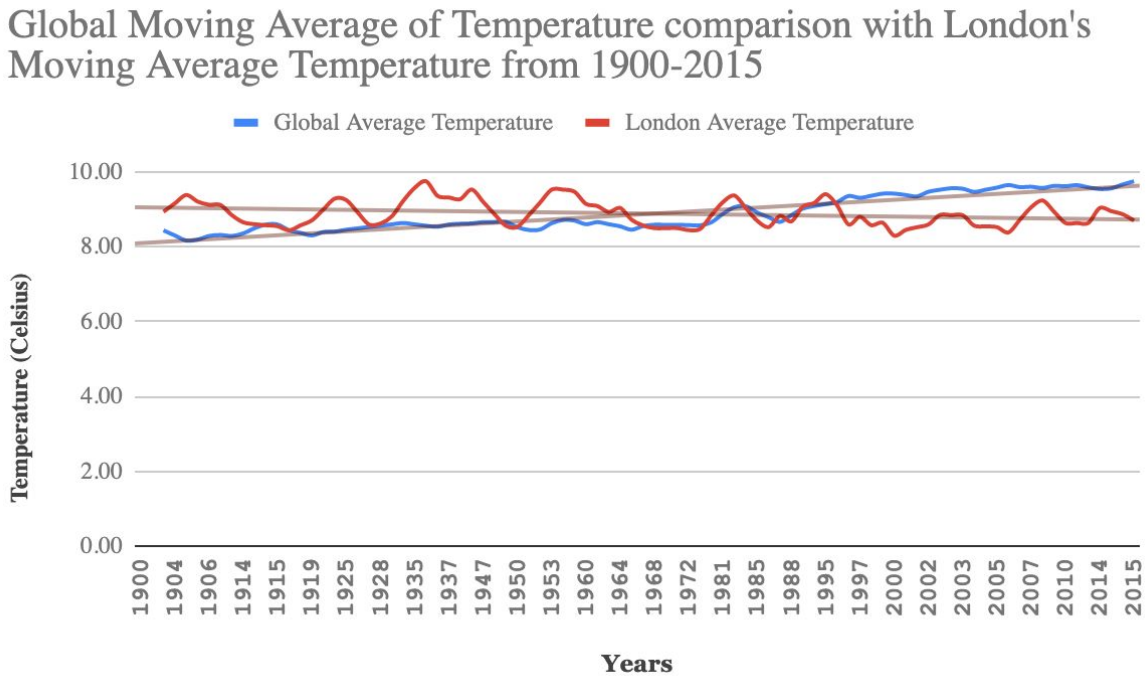
```
SELECT DISTINCT city_data.city,
                city_data.year,
                city_data.avg_temp   avg_city_temp,
                global_data.year     global_year,
                global_data.avg_temp global_avg_temp
FROM   city_data
       JOIN city_list
         ON city_data.city = city_list.city
       JOIN global_data
         ON city_data.avg_temp = global_data.avg_temp
WHERE  city_data.city = 'London'
       AND city_data.year >= 1900
       AND global_data.year >= 1900
```

MANIPULATING WEATHER DATA

After extracting the SQL query into a CSV file successfully, I used Google Sheets to perform my calculations. (Click here to view the Google Sheet). First, I separated the global columns from the city columns as they were displaying results differently when filters were placed on them. After that, I sorted the numbers by ascending order to get both global and city data to match years. Next, I calculated the 3-year moving average using the `=AVERAGE()` function in a new column. Then, the next step was to visualize the calculated data.

VISUALIZING WEATHER DATA

I created a **Line Chart** of the global moving average and the London moving average temperatures. I added two trendlines to create a distinction between the two and to see where each one was heading over time. I used colors that are clear and legible, and I added descriptive axes titles and a legend. The X-axis represents the years, and the Y-axis represents temperature in Celsius. See visualization:



OBSERVATIONS

1. Global Average temperatures have *increased* across the past century.
2. London's average temperature is increasing and decreasing at rates similar to the global average.
3. London's Global Average temperatures were higher than the global average in the years 1972-1982, and until 1995, the average temperature in London was almost always higher.
4. Global Average temperatures are trending higher whereas London's average temperature seems to be trending lower or in parallel to the Global Average which indicates that London's temperature may increase.
    a. In June of 2019, the BBC reported:

        "*The Met Office said Heathrow and Northolt in west London had reached **34C** (93.2F) making it one of the warmest June days for about 40 years."*

Although the line chart data stops at 2015, we see that this trend has steadily increased and every year the average temperature increases. (See Source section)

SOURCE

1. **UK weather: Hottest day of the year as temperatures soar**
   https://www.bbc.com/news/uk-48810170