



# Exploring Weather Trends : Columbus, OH

Author: Amber Simmons, PhD

Date: 2023-05-17

Photo by [Oz Seyrek](#) on [Unsplash](#)

## Purpose Statement

In this project, I will analyze local and global temperature data and compare the temperature trends my local city (Columbus, OH) to the overall global temperature trends. I will seek to answer the following questions:

- How does the temperature in Columbus compare to the global average, and has this trend been consistent over time?
- How have changes in Columbus's temperatures over time compared to changes in the global average?
- What does the overall temperature trend look like? Has the trend been consistent over the last few hundred years?
- Are the temperature averages generally correlated, and what does that relationship look like?

## Methods

The following section describes the steps taken to extract and wrangle the local and global temperature data.

## Data Extraction

The following SQL queries were used to extract the data from the database provided:

City-level Data:

```
SELECT *
FROM city_data
WHERE country = 'United States'
      AND city = 'Columbus';
```

Global Data:

```
SELECT *
FROM global_data;
```

The results of the above queries were exported to CSVs.

## Data Wrangling

First, the query-export CSVs were imported into RStudio. The year columns were then typecast into the Date type, and the annual average temperature columns were converted to a numeric type. Because a simple way to prepare data for plotting comparisons is to join the two data sets together, the city-level data was assigned a group variable where each value is 'Columbus', and the global data was assigned a group variable with value 'Global'. The moving averages for each set were calculated as follows using the *zoo* library:

```
cbus_data_full <- cbus_data %>%
mutate('5-Year MA' = zoo::rollmean(avg_temp, k = 5, fill = NA, align = 'right'),
      '10-Year MA' = zoo::rollmean(avg_temp, k = 10, fill = NA, align = 'right'),
      '50-Year MA' = zoo::rollmean(avg_temp, k = 50, fill = NA, align = 'right')
)

global_data_full <- global_temps %>%
mutate('5-Year MA' = zoo::rollmean(avg_temp, k = 5, fill = NA, align = 'right'),
      '10-Year MA' = zoo::rollmean(avg_temp, k = 10, fill = NA, align = 'right'),
      '50-Year MA' = zoo::rollmean(avg_temp, k = 50, fill = NA, align = 'right')
)
```

For exploratory purposes, 5, 10, and 50 year moving averages were calculated. This method was verified against the example dataset provided in the course. Finally, the two datasets were combined into one set using *dplyr::bind\_rows*. Ultimately, the 10-year moving average was selected for visualization; the 5-year moving average provided such a detailed view that the overall trends were difficult to assess, and the 50-year moving average cast such a wide aggregation that the subtle changes between decades were lost.

## Observations

1. *Columbus historically has been hotter than the global average.* According to the figure below, Columbus is consistently 5-6 degrees warmer than the global average (according to 10-year global averages).
2. *Changes in average temperature year-over-year in Columbus tend to align with changes to global temperature.* For example, the dip in average global temperature between 1800-1820 is also reflected in the average Columbus temperatures.

3. *Similar to the global average temperature since ~1800, Columbus has been steadily increasing in average temperature year over year.* The world has been getting consistently warmer overall since about 1820, as seen in the plot below. While the Columbus average temperatures tend to vary year over year more than the global averages (Columbus shows more gains/losses in average temps when looking at small windows of time), but the overall trend shows a similar steady increase since about 1800.
4. *Columbus 10-year moving average temperatures are statistically correlated with the 10-year moving average global temperatures.* The Columbus and global average temperatures have a correlation coefficient of 0.78, indicating that the temperatures are both positively and strongly related.

