# Exploring Weather Trends

Udacity

## Purpose

For this data analysis project, I have compared the temperature trends for Dallas, Texas to the global average temperature.

## Data Preparation

The data for this project was housed in a Udacity workspace that contained two relevant tables: global_data and city_data. The global_data table contains the average temperature across the globe by year, dating back to 1750. The city_data table contains the average temperature by city for each year. The table did not contain the same year ranges for each city; however, this project only focused on the city of Dallas, Texas, which had a year range from 1820 to 2013.

SQL was used to extract the data from each table in Udacity into two CSV files of the same names. For the city_data table, the extract pulled all the data in the table, rather than filtering to just Dallas.

The data from both CSVs was then moved into an Excel workbook for further analysis. Excel was used for this data analysis due to the small amount of data and ease of creating visualizations in Excel. Each CSV was copied into its own tab, and given the same name as the table from which it was extracted. An additional tab was then created in the workbook titled Project Work – data, which contains three columns: Year, Global 7-Year Average, and Dallas 7-Year Average:

- **Year** – the year for which the temperature data was gathered
- **Global 7-Year Average** – the moving average pulled from the global_data tab that averages the previous seven years of temperature data
- **Dallas 7-Year Average** – the moving average pulled from the city_data tab that averages the previous seven years of temperature data
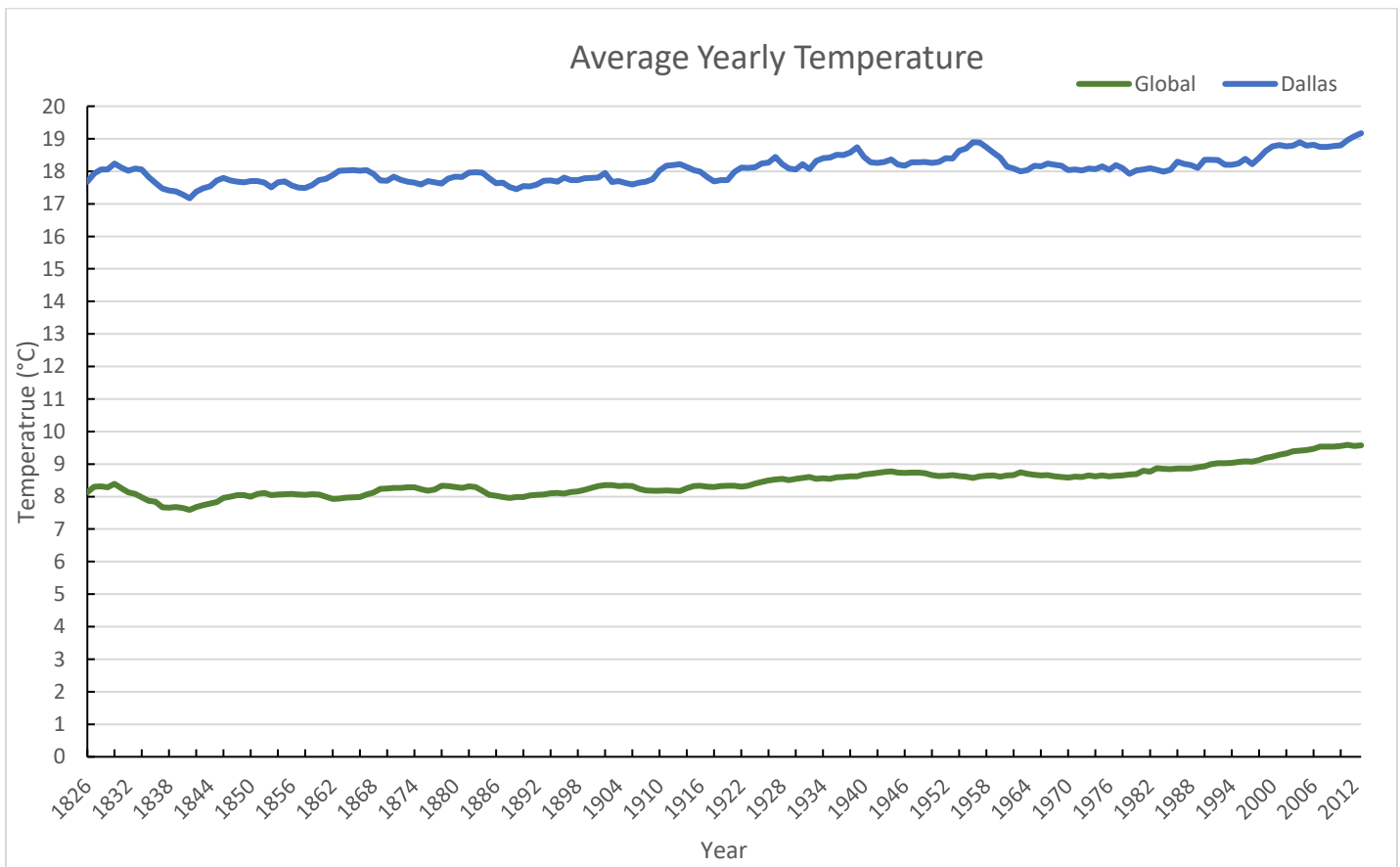


** Despite having global data from 1750 onward, a year range of 1826 to 2013 was used due to a lack of data for Dallas.

**Observations**

- **Dallas is 2x as hot as the global average**

Compared to the global average, Dallas is consistently twice as hot as the average for the rest of the world. This is likely due to Dallas's proximity to the equator and the global_data data source factoring in several extremely cold climates in Russia, Canada, and the Scandanavian countries. These countries consistently have temperatures averaging in the negatives, which would severely weight the global average.

- **Dallas is more volatile**

The moving 7-year averages for Dallas appear to be much more volatile than the global averages. This volatility is best explained by the fact that the global_data contains much more data than the Dallas data source. This larger amount of data helps to smooth out volatility.

- **Average temperature has increased**

Dating back to 1826, where our analysis begins, the temperatures in both Dallas and across the globe have increased by approximately 1.5°C. Although controversial, this overall increase in temperature is best attributed to global warming.

- **Temperature rising quicker as of recent**

For both the Dallas and global data sets, it appears that the average temperature for both has started rising quicker and more consistently since 1994. This is a hard correlation to explain and would be another interesting project to dig into what occurred in both America and worldwide during the 1970s and 1980s that may have led to an increase in temperature throughout the 90s.