

Udacity Data Analyst Nanodegree - Project 1 Submission

- **Step 1: Query database to pull relevant annual temperature data for**
 - 1) Global, and
 - 2) Local city (Toronto, Canada)

*Note - for effective comparison, I only included data for years that had both Toronto and Global average temperature available. (1750-2013).

Query:

```
SELECT c.year, c.city, c.avg_temp as local, g.year, g.avg_temp as global
FROM city_data c
RIGHT JOIN global_data g
ON c.year = g.year
WHERE city = 'Toronto'
```

- **Step 2: Manipulate Data (used Google Sheets)**

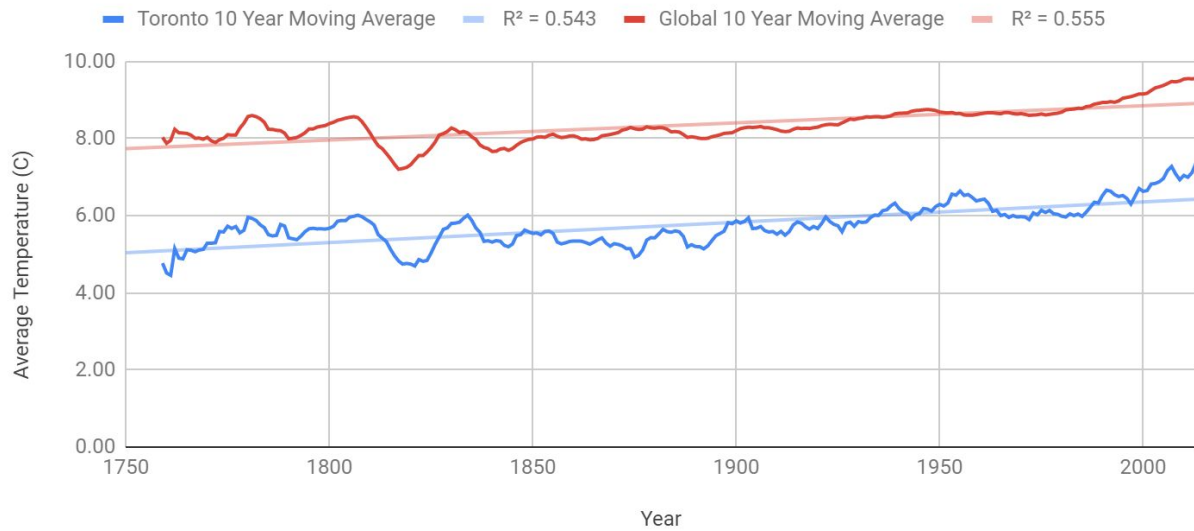
Calculated 10-year moving averages for both Toronto and Global temperatures in order to reduce the prevalence of outlier years and create a visualization that more clearly depicts the trends over time.

To calculate moving averages for each year, I used the mean of temperature values from the past 10 years. So 1760's data point would be the mean of the temperatures recorded from 1750-1759.

- **Step 3: Visualize Data Trends**

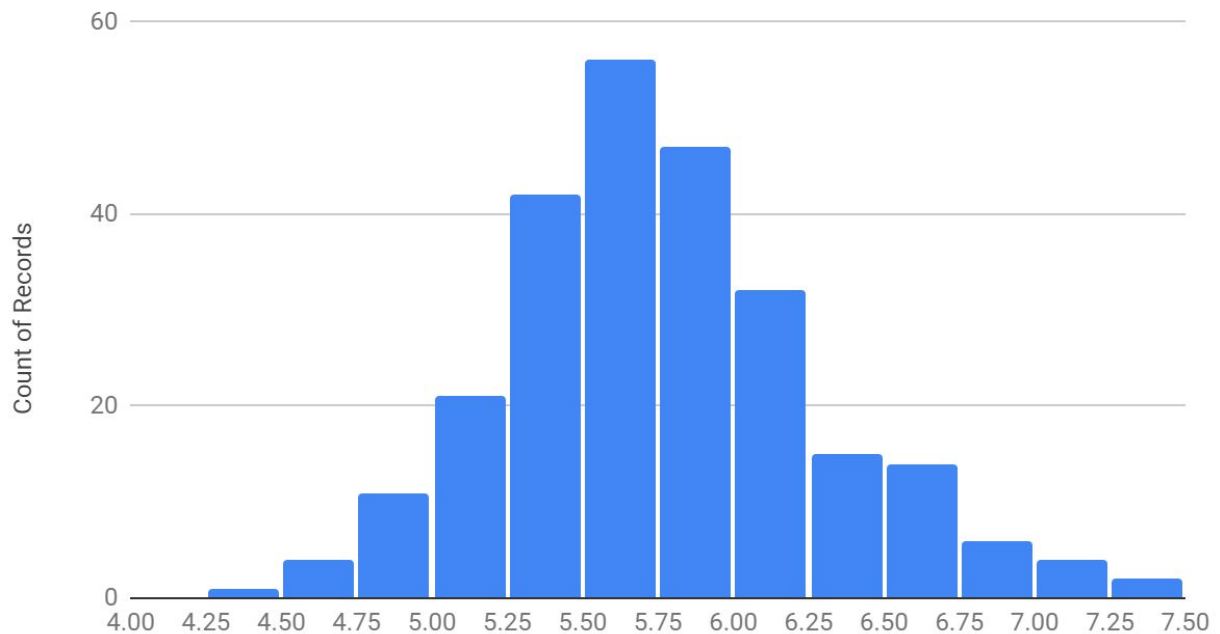
Below is a line chart that compares the changes in Toronto and Global average temperatures over time. It includes a trendline and R-squared value to provide the viewer with more information about the overall trend for each.

Comparing Toronto and Global Average Temperatures over Time

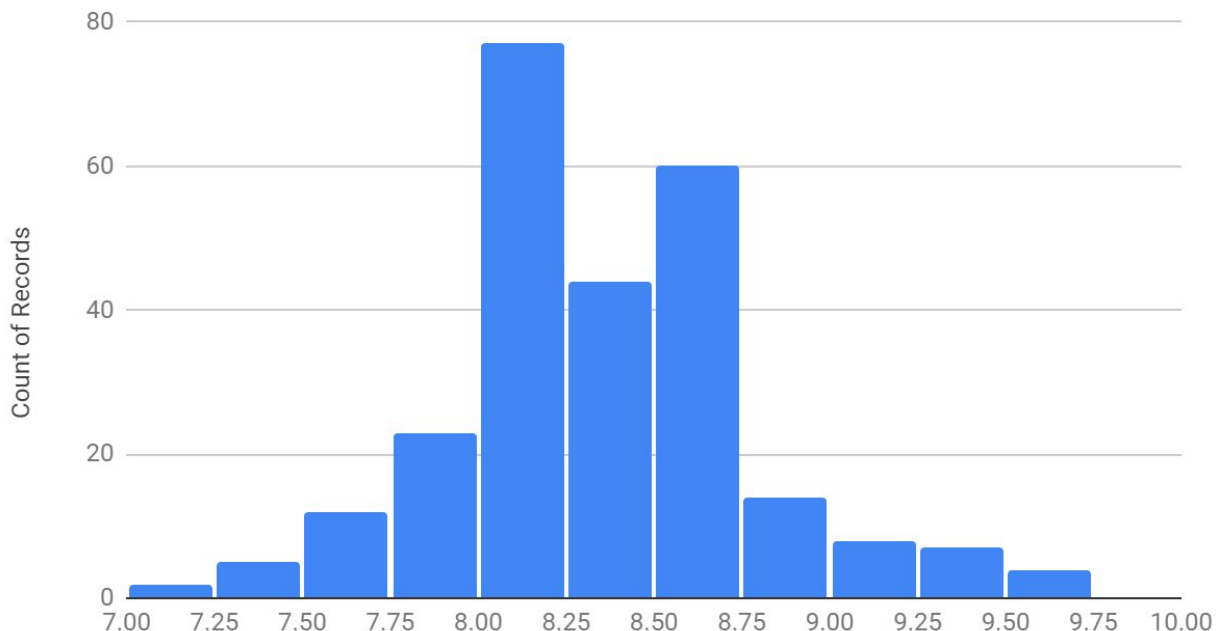


I also included histograms to roughly assess how the temperature data is distributed for both Toronto and Global.

Distribution of Toronto Average Temperature



Distribution of Global Average Temperature



- **Step 4: Interpretation of the Data**

Chart and Data Observations

- i) Toronto is consistently 2-3 degrees Celsius cooler than the global average temperature. This has been generally consistent throughout the timeframe data was collected for.
- ii) On the whole, both Toronto and Global temperatures have gotten slightly warmer over time. They follow similar patterns of cooler temperatures in the 1820s and increasingly warmer temperatures since the late 1990s. In the past 25 years alone, the average temperature for both Toronto and Global has increased by nearly 1 degree Celsius.
- iii) On average, the Global temperature has increased slightly faster than Toronto's. This can be validated through the R-squared values of 0.555 globally versus 0.543 for Toronto.
- iv) Both Toronto and Global annual temperature datasets follow a roughly normal distribution, with very similar mean and median values. However, Toronto has a slightly higher standard deviation, meaning the data is more spread out from the mean. This can also be seen visually, as the Toronto (blue) line has a lot more local peaks and valleys when compared to the smoother Global (red) line.

Region	Toronto	Global
Mean	5.76	8.34
Median	5.72	8.27
Standard Deviation	0.53	0.44