Project 1 – Exploring Weather Trends

Analysis Methodologies

SQL Queries

City List

```
SELECT *
FROM city_list
ORDER BY country, city;
```

Notes regarding query:

1. Actual city is Vancouver, BC, Canada, therefore, closest city is Victoria, Canada

City Data

```
SELECT year, avg_temp
FROM city_data
WHERE country = 'Canada'
AND city = 'Victoria'
ORDER BY year;
```

Notes regarding query:

1. It was noted that data from city commenced at 1828 so this was selected as a filter for the Global data
2. Based on encouragement from the course rubric, the above query was also completed for all listed Canadian cities (Kingston, London, Montreal, Ottawa, Toronto, Victoria) and the code was adjusted in the following way to pull all cities:
   ```
   SELECT city, year, avg_temp
   FROM city_data
   WHERE country = 'Canada'
   ORDER BY city, year;
   ```

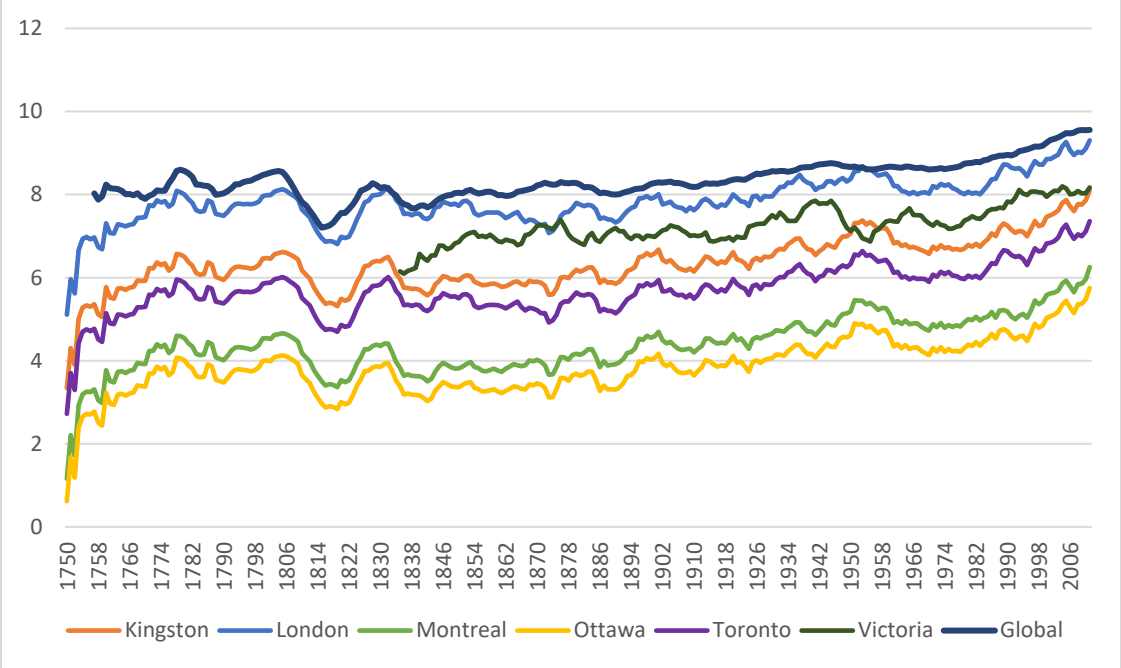
Global Data

```
SELECT *
FROM Global_data
ORDER BY year;
```

Excel and Other Data Steps and Rationale

Data Examination and Clean Up

1. Excel was selected to examine the data because this is the program that is most familiar to the author.
2. The data from the amended query for the City Data was filtered by city using Excel's filter options to copy each of the city information into a new spreadsheet. (This was considered more efficient than exporting six different sets of data from City Data) The data from the Global Data was also copied to the new spreadsheet. Data was aligned to ensure that each row represented the same year for all cities.
3. It was determined that a decade would be used as the interval to create the moving average (MA) as this is a common interval for comparing the passages of time. An additional column was created to calculate the MA for each of the six cities and the Global trends.
4. The full data was graphed for all cities, but it was noted that there was substantial variability in this data, with sharp increases seen in the first few years that were not replicated elsewhere in the data. It was further noted that there were multiple missing years of data in the first decade.



Therefore, the data visualized in the line chart above was amended to show data only from the first year of availability for the Global data (1759), (and then the first available year for Victoria - 1837) as this was considered a more accurate representation of the trends. The final year of data visualization was set at the last year that all cities had available data, 2013. The inclusion of zero as a reference in the vertical scale is a personal preference of the author.

5. The graph was formatted to ensure ease in reading axis labels and distinguishing between the various cities, as well as highlighting the Global trend information.

Comparison between Canadian Cities

6. It was noted that the variation in MA for all Canadian cities listed followed a similar pattern of variability, with the exception of Victoria. This was explored further. The locations of all cities were plotted on a map of Canada (using Google Maps as a reference) and it was identified that Kingston, London, Montreal, Ottawa and Toronto are all found within a relatively similar geographic location, on the Eastern side of Canada, whereas Victoria is found on the Western coast of Canada.
7. To confirm the differences in variations, correlations were calculated between the various Canadian cities, using the CORREL formula, and the significance of these correlations was calculated using http://vassarstats.net/textbook/ch4apx.html.

Comparison to Global Temperatures

8. It was noted that Canadian cities had a lower MA compared to the Global MA and a t-test assuming unequal variances, using the Excel Data Analysis add-in, following the instructions at https://www.rwu.edu/sites/default/files/downloads/fcas/mns/running_a_t-test_in_excel.pdf, was performed to compare them. Averages were compared for where all cities had data (1837 to 2013).
9. Latitudes of the observed Canadian cities were obtained via Google search.

Predicting City Moving Averages

10. The MA of each Canadian city was also correlated to the Global MA using the CORREL formula for all years that had data across all cities, 1837 to 2013.
11. Annual temperature information for each of the Canadian cities was obtained from https://victoria.weatherstats.ca/metrics/temperature.html (by changing the city in the address for all six cities). This information was added to the original data set to calculate the subsequent MA for each of the cities.
12. The formula LINEST was used to calculate the slope, y-intercepts, accounted variance and standard error of predicted y of the regression line for each Canadian city compared to the Global changes, following the instructions provided at https://www.youtube.com/watch?v=6wbcPbYbq6M.

	Kingston	London	Montreal	Ottawa	Toronto	Victoria
slope	1.167	0.9940	1.248	1.232	1.078	0.9285
intercept	-3.351	-0.4157	-5.991	-6.396	-3.240	-0.6090
r^2	0.8275	0.7942	0.8385	0.8372	0.8110	0.8274
se(y)	0.2352	0.2235	0.2419	0.2400	0.2298	0.1828
α	0.05	0.05	0.05	0.05	0.05	0.05
df	253	253	253	253	253	175
crit. value	1.969	1.969	1.969	1.969	1.969	1.974
me(y)	0.4631	0.4400	0.4763	0.4725	0.4525	0.3608

Table 1: Regression model statistics for predicting city moving average from global moving average.

13. The city MA for 2014 and 2015 was compared to the predicted MA based on the regression formula.

14. Confidence intervals of 95% for the predicted results were also calculated by referring to the Prediction Intervals information at http://www.stat.yale.edu/Courses/1997-98/101/linregin.htm and http://stattrek.com/regression/slope-confidence-interval.aspx?Tutorial=AP using http://stattrek.com/online-calculator/t-distribution.aspx as a t-distribution calculator.
15. Figures were converted to graph format using the Insert Combo Chart Excel action to assist in understanding the data. The standard format was amended to assist in demarking the confidence intervals and the scale was adjusted to assist in identifying data that exceeded the upper confidence intervals.
16. It is recognized that adding to a data set is often not done when completing dataset analysis. As result, the additional MAs were not included in the data visualization and only used to test the predictive capability of the regression models. It is also recognized that the predictive capabilities of the regression models have limitations because it is not possible to confirm consistent methodology for calculating the annual temperature of the cities.

Evaluating Temperature Increase

17. Visual inspection of the data revealed that both Globally and for all Canadian cities, the greatest increase in temperature was from approximately 1980 onward. Due to the high correlation between the Global data and that of the Canadian cities, only Global data was used to examine this observation. The Excel Data Analysis add-in was used to conduct t-tests assuming unequal variance. The years 1981 to 2015 were compared to 1946 to 1980, and 1946 to 1980 was compared to 1759 to 1793. This allowed a comparison between the 35 years with the greatest temperature change to the previous 35 years, with further comparison to the first available 35 years of data, to provide insight into changes in temperature over time.

Statistical Results

Comparisons between Canadian Cities

Correlations of moving average temperature (MA) were calculated between each Canadian city with the following results. All correlations had a p-value of less than 1×10^{-6}.

	Kingston	London	Montreal	Ottawa	Toronto	Victoria
Kingston	----					
London	0.9723	----				
Montreal	0.9926	0.9450	----			
Ottawa	0.9968	0.9601	0.9977	----		
Toronto	0.9871	0.9949	0.9673	0.9796	----	
Victoria	0.7654	0.7430	0.7792	0.7759	0.7456	----

Table 2: Correlation of moving average temperature between Canadian cities.

Closest correlations were between the five cities of Kingston, London, Montreal, Ottawa and Toronto. Correlations between these cities ranged from r = 0.9450 between London and Montreal to r = 0.9968 between Ottawa and Kingston. These five cities were all correlated lower with Victoria, ranging from r = 0.7430 with London to r = 0.7792 with Montreal.

Comparison to Global Temperatures

The mean MA across all Canadian cities was compared to the mean MA of the Global temperatures for 1837 to 2013 using a t-test assuming unequal variances.

MA Temperature	Mean	Difference	df	t	p
Canadian Cities	5.525	-2.938	342	-69.52	$<1 \times 10^{-6}$
Global	8.462				

Table 3: Comparison of means for Canadian cities and Global moving average temperatures from 1837 to 2013.

The average Global MA was found to be 2.938°C higher than the average MA for the Canadian cities with a p-value of less than 1×10^{-6}.

Predicting City Moving Averages

Correlations between the Canadian cities' MA and the Global MA were calculated for the years 1837 to 2013. All city MAs were highly correlated with the Global MAs and had a p-value of less than 1×10^{-6}. Correlations ranged from 0.8912, explaining 79% of the variance for London to 0.9157, explaining 84% of the variance with Montreal.

MA	Kingston	London	Montreal	Ottawa	Toronto	Victoria
Global correlation	0.9097	0.8912	0.9157	0.9150	0.9006	0.9096
Variance explained	83%	79%	84%	84%	81%	83%

Table 4: Correlation and explained variance between Canadian cities' the Global moving average temperatures from 1837 to 2013.

Regression lines were calculated for the correlations between the cities and the Global MA. Upper (UCI) and lower (LCI) onfidence intervals at $\alpha = 0.05$ were then calculated for the predicted y-value. These values were compared to the actual city MAs for 2014 and 2015.

	Kingston	London	Montreal	Ottawa	Toronto	Victoria
Predicted MA Temp 2014	7.812	9.097	5.952	5.397	7.077	8.277
LCI 2014	7.349	8.657	5.475	4.925	6.624	7.916
UCI 2014	8.275	9.537	6.428	5.870	7.529	8.638
Actual MA Temp 2014	8.222	9.380	6.452	5.895	7.469	8.345
Predicted MA Temp 2015	8.115	9.355	6.276	5.718	7.357	8.518
LCI 2015	7.652	8.915	5.800	5.245	6.905	8.158
UCI 2015	8.578	9.795	6.752	6.190	7.810	8.879
Actual MA Temp 2015	8.264	9.421	6.567	5.982	7.607	8.637

Table 5: Upper and lower confidence intervals for $\alpha = 0.05$ for predicting city moving average temperatures from global moving average temperatures for 2014 and 2015.

Evaluating Temperature Increase

Mean Global MA of 35-year intervals were calculated for 1759 to 1793, 1946 to 1980 and 1981 to 2015. Their respective means were 8.150°C, 8.656°C and 9.154°C. Two t-tests assuming unequal variance were conducted to compare to means between 1759 to 1793 and 1946 to 1980, and between 1946 to 1980 and 1981 to 2015.

Global MA Temp.	Mean	Difference	df	t	p
1759 – 1793	8.150	-0.5064	37	-14.95	$<1 \times 10^{-6}$
1946 – 1980	8.656	-0.4978	35	-10.09	$<1 \times 10^{-6}$
1981 – 2015	9.154				

Table 6: T-tests comparing mean Global moving average temperatures between 1759 to 1793 and 1946 to 1980 and between 1946 to 1980 and 1981 to 2015.

The difference in mean Global MA between 1946 to 1980 and 1759 to 1793 was 0.5064°C. The difference in mean Global MA between 1981 to 2015 and 1946 to 1980 was 0.4978°C. The p-value for both differences was less than 1×10^{-6}.

Comparing Canadian Temperature Changes to Global Changes

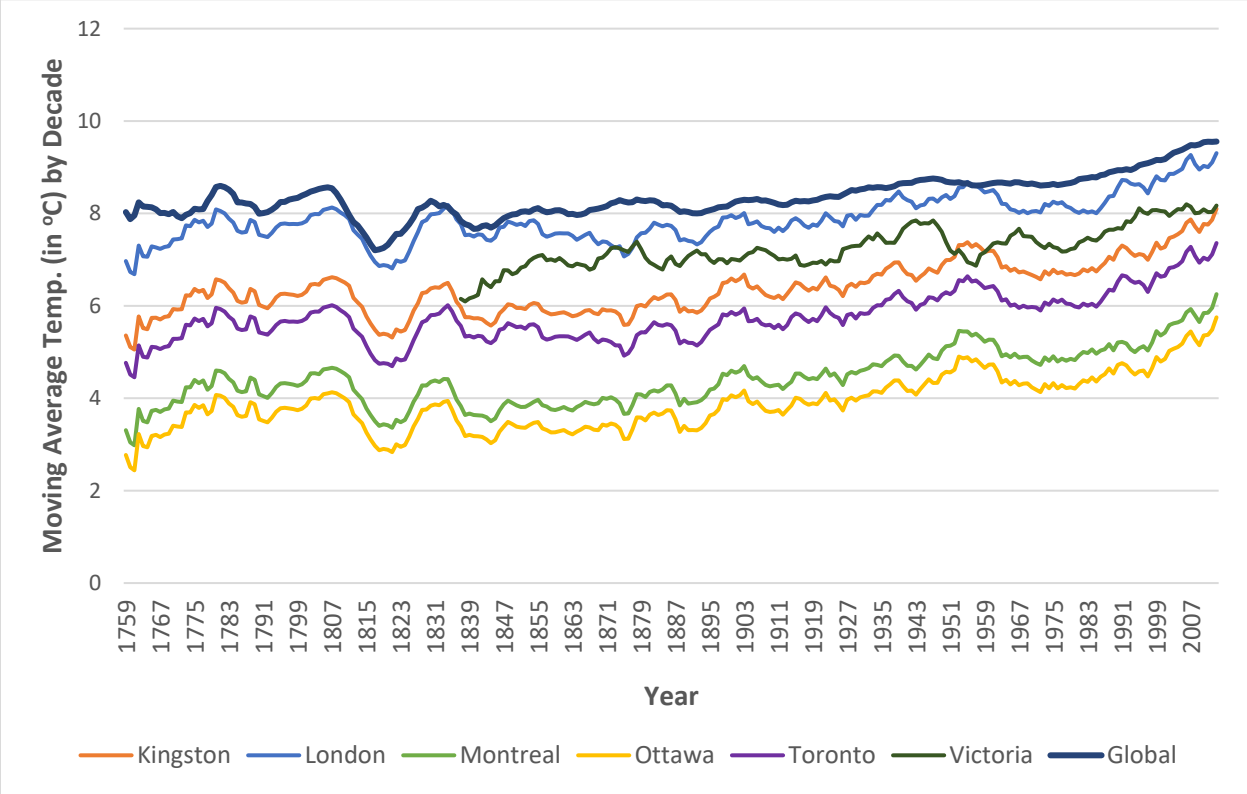


Figure 1: Moving average temperatures for six Canadian cities and the globe between 1759 and 2013.

Geography Influences Temperature Variations between Cities

After noticing that the variations in temperature for Kingston, London, Montreal, Ottawa and Toronto were more similar than those for Victoria, their locations were plotted on a map. As shown in Figure 2, Kingston, London, Montreal, Ottawa and Toronto are all located on the eastern side of Canada, in relative proximity to the St Lawrence River or Lake Ontario. Victoria, is located on Vancouver Island, just off the west coast of Canada.

Conclusion: This suggests that geographic proximity and location influence temperature variations found within Canadian cities.

Figure 2: Map of six Canadian cities. (Map and location information retrieved from Google Maps February 14, 2018.)

Canadian Cities are Colder than the Global Average

Influences of geographic location are also seen in the comparison of the mean temperature for Canadian cities to the global temperature. The average temperatures of Canadian cities were found to be lower than that of the average global temperatures. It is noted that all Canadian cities are found substantially above the equator, with latitudes ranging from 43°N to 48°N.

Conclusion: As such, it should be expected that the temperatures of Canadian cities are lower than global temperatures as they are not off-set by temperatures from warmer climates that are encompassed within the global average.

Global Temperatures Strongly Correlate and Predict City Temperatures

While geographic proximity and location may influence similarities in city temperatures, the high correlations between each of the Canadian cities' temperatures to the global temperatures suggests that city temperatures are also influenced by global temperatures. This is the case even if there are varied correlations in temperature between the cities; the temperature of each city remains highly correlated to the global temperature.

The influence of global temperatures on city temperatures can also be seen in the ability to predict city temperatures from global temperatures with reasonable accuracy. As can be seen in Figures 3 and 4, of the 12 predicted temperatures, only Montreal and Ottawa were just above 95% confidence intervals in 2014.

Conclusion: These results indicate that as global temperatures rise, cities will see corresponding changes in their average temperatures.

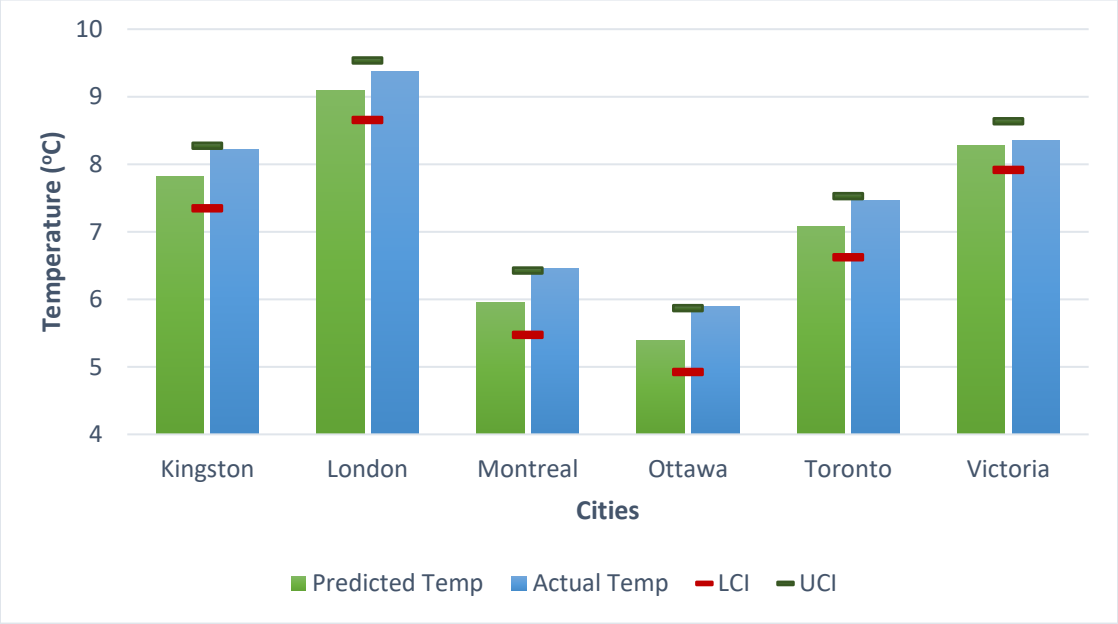


Figure 3: Predicted moving average temperatures for six Canadian cities from Global moving average temperatures compared to actual moving average temperatures for 2014.

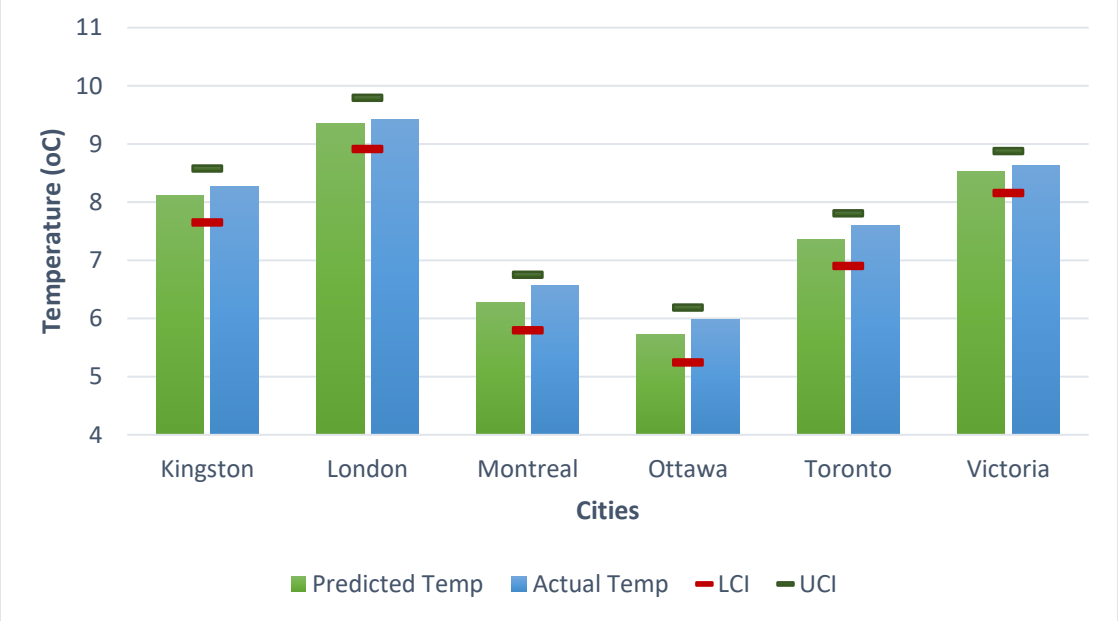


Figure 4: Predicted moving average temperatures for six Canadian cities from Global moving average temperatures compared to actual moving average temperatures for 2014.

Global Temperatures Have Seen a Sharp Increase Since 1980

As can be seen in Figure 5, the difference in global temperatures for the first 35 years of available data compared to the 35 years before 1980 are incredibly similar to the difference in global temperatures for the 35 years before and after 1980, with an approximate difference of 0.5°C. That is, the rate of change in global temperature over a 150-year interval is essentially the same as the change in temperature for the most recent 35-year interval.

Conclusion: These results suggest that the rate of global temperature increase has changed dramatically in the last 35 years, and, based on the previous observations, we can expect to see similar changes within the temperatures of Canadian cities.

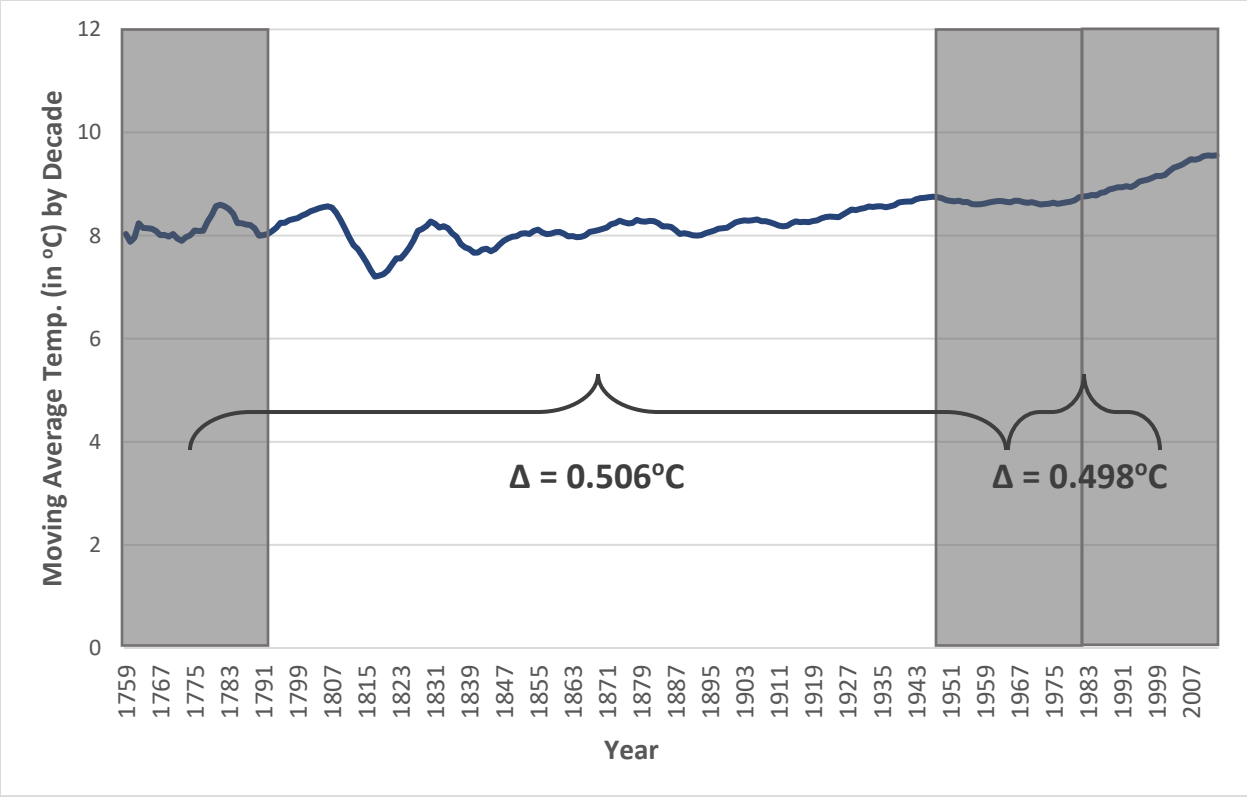


Figure 5: Global moving average temperatures from 1759 to 2015.