Exploring Weather Trends – Project Submission

Zhaojunxiang (Kasey) Fu

Outline

A thorough run-through of the project and its weather data extracting will be discussed here. Because the largest city for this project is Toronto, the weather data for Toronto is compared with the global data, as well as 3 other cities which are SIMILAR to Toronto in terms of population, lifestyle, and economy.

First, SQL queries were used to extract the needed data for this investigation. On the Udacity website, SQL queries allowed me to conduct some early research for my project before I began data extracting.

First to look through the city data in order to confirm that ‘Toronto’ was included:

```
SELECT city

FROM city_list

WHERE city IN ('Toronto');
```

Next, to extract the data from Toronto and the global temperature averages, the next set of queries were used:

To Extract Toronto data

```
SELECT *

FROM city_data

WHERE city IN ('Toronto');
```

To Extract Global data

```
SELECT *

FROM global_data;
```

To Extract Data from Boston, Melbourne and Chicago for Comparison to Toronto

```
SELECT *

FROM city_data

WHERE city IN ('Boston');

SELECT *

FROM city_data

WHERE city IN ('Melbourne');

SELECT *

FROM city_data

WHERE city IN ('Chicago');
```

Here is a snapshot of the query for ‘Chicago’.

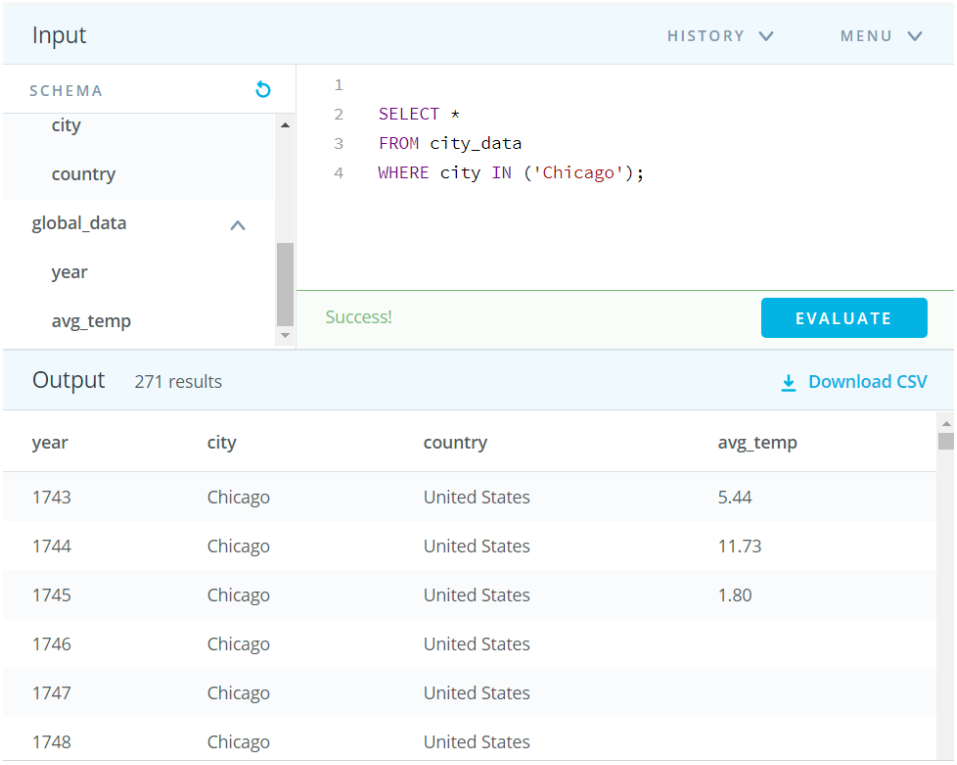

Figure 1: The SQL query to extract Chicago data.

Next, this data was extracted and viewed on different Excel files for simplicity. The data was also viewed as a CSV on text editor Atom as a second choice. To calculate the Moving Averages for every set of data extracted, I went and looked for the first 7 consecutive years where there was data provided on the average temperature. I used the Average function to calculate the average for those 7 years running, and here is an example:

=Average(B2:B8)

This Average function was applied to all sets of data (Toronto, Global, Boston, Melbourne, Chicago), to expose the moving averages for every list.

Next, due to the simplicity of connecting Excel files to Tableau, I used Tableau to visualize all the different sets of data. First, I inputted every different excel file as a data source into the Tableau

platform, and then viewed it on Tableau as a table with the 'Data Interpreter' function enabled so that any changes made to the excel files were easily changed automatically on the Tableau file.

Finally, I created different sheets to represent every different city (or global data) available. Using city as a dimension, I took 'Year' measure into the columns section, the 'Avg Temp' and 'MA Temp' into the Rows section, I was able to view two different line graphs on the Tableau screen and dashboard. The top shows the line graph with the average temperature per year, while the bottom shows the line graph with the moving average (temperature) per year, for comparison. As expected, the moving average line graph was much more smooth and 'cleaned up' compared to the top graph. Here are the graphs for Toronto, Boston, Chicago, Melbourne, and Global:

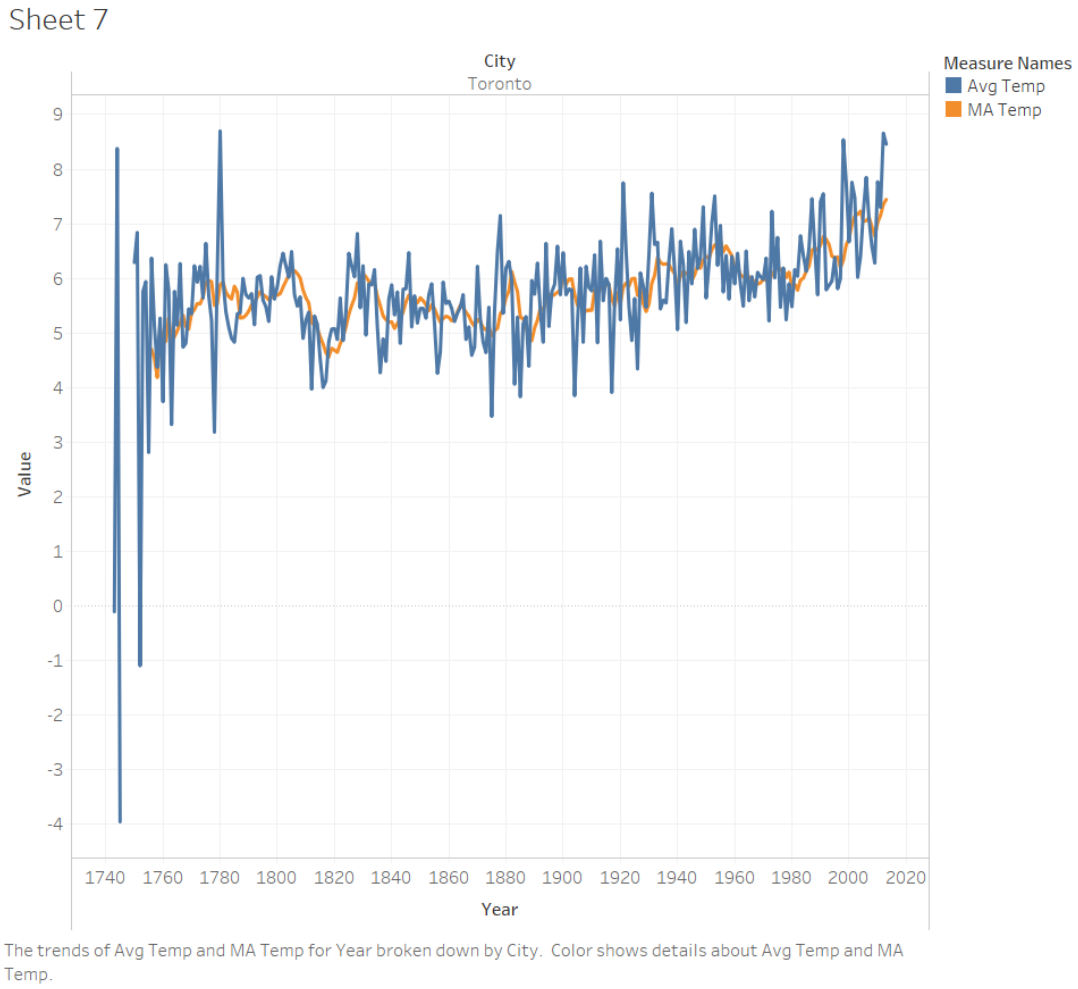

Figure 2: The average temperature and the moving average temperature for Toronto.

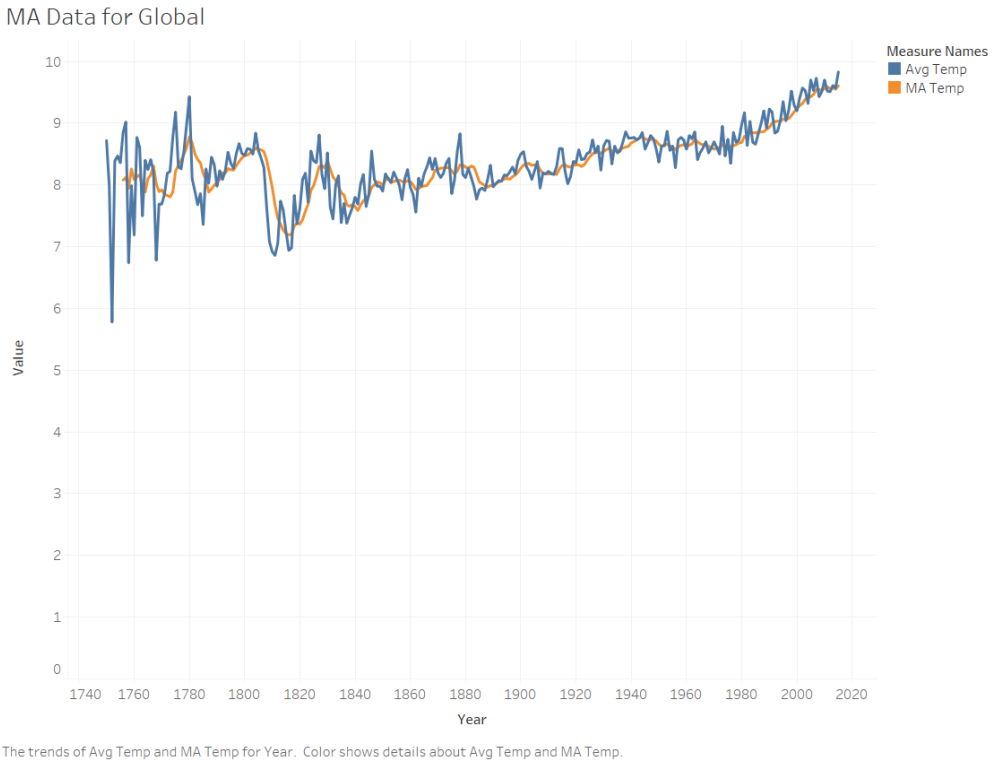

Figure 3: The average temperature and the moving average temperature for Global data.

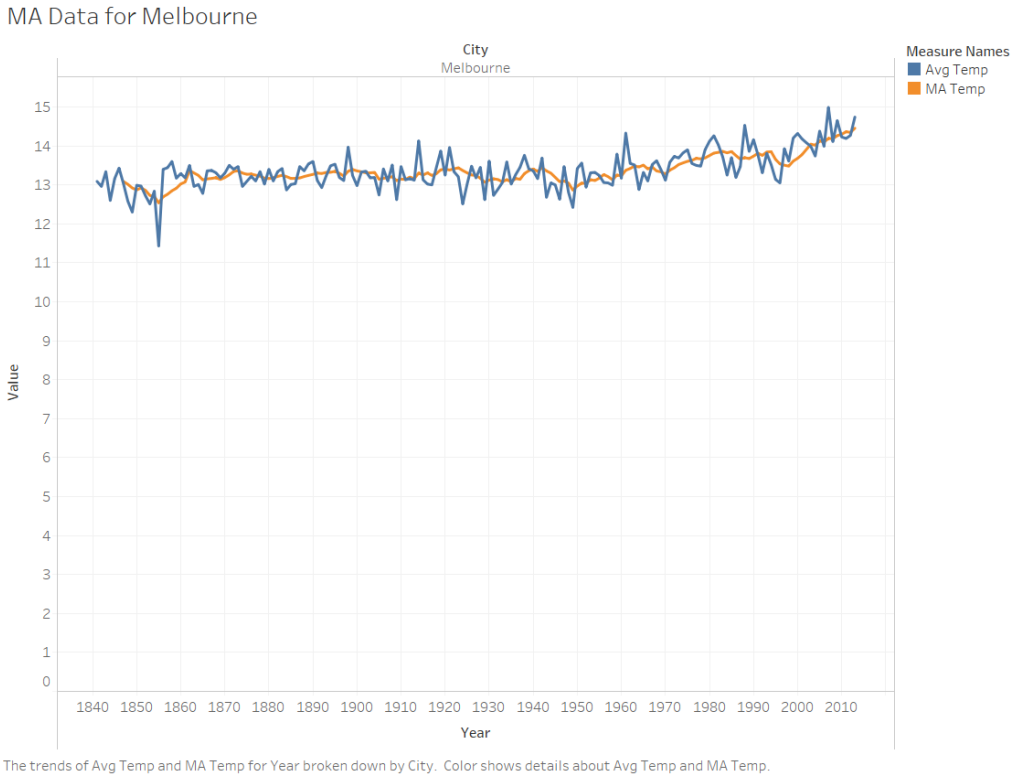

Figure 4: The average temperature and the moving average temperature for Melbourne.

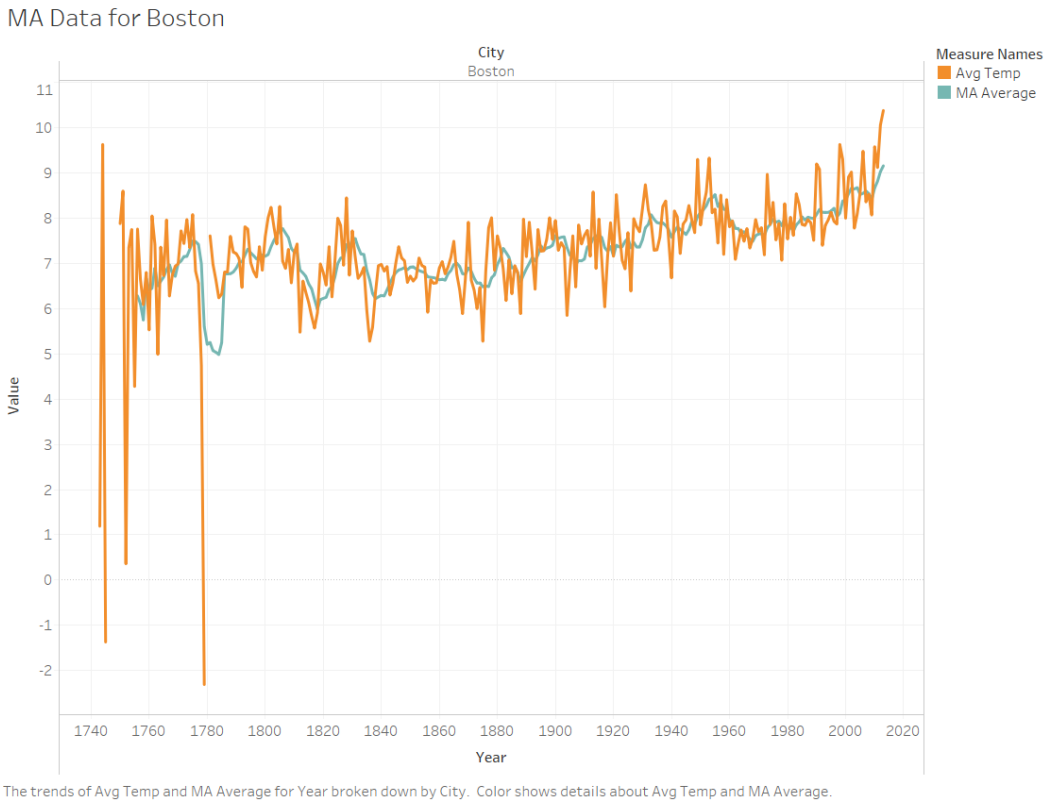

Figure 5: The average temperature and the moving average temperature for Boston.

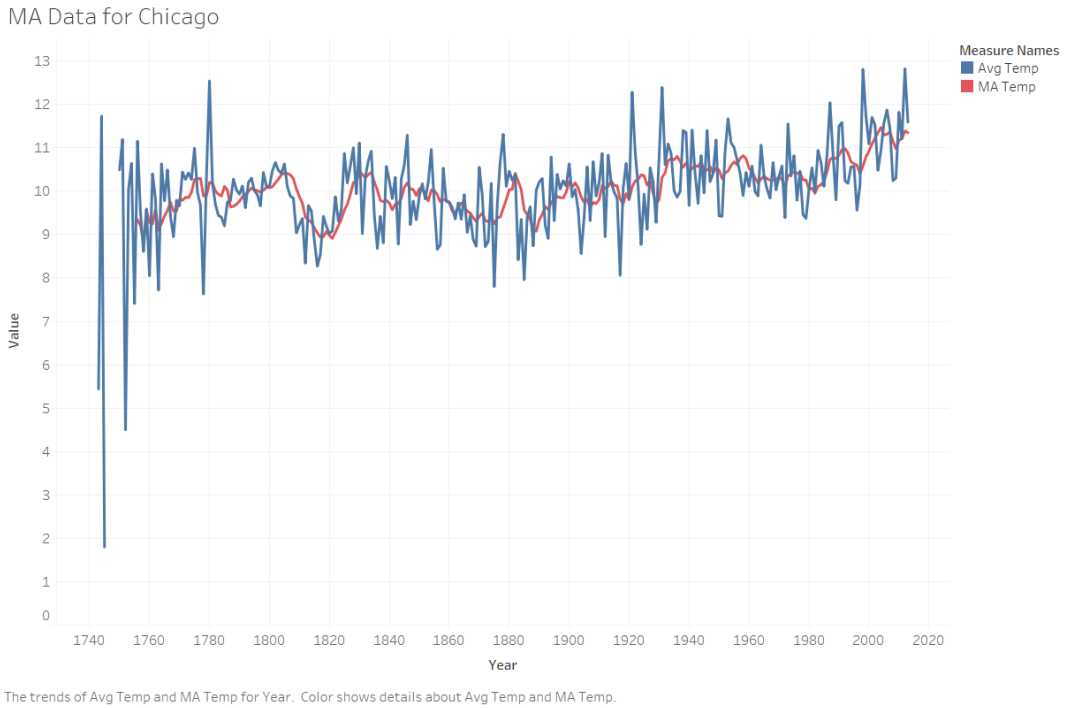

Figure 6: The average temperature and the moving average temperature for Chicago

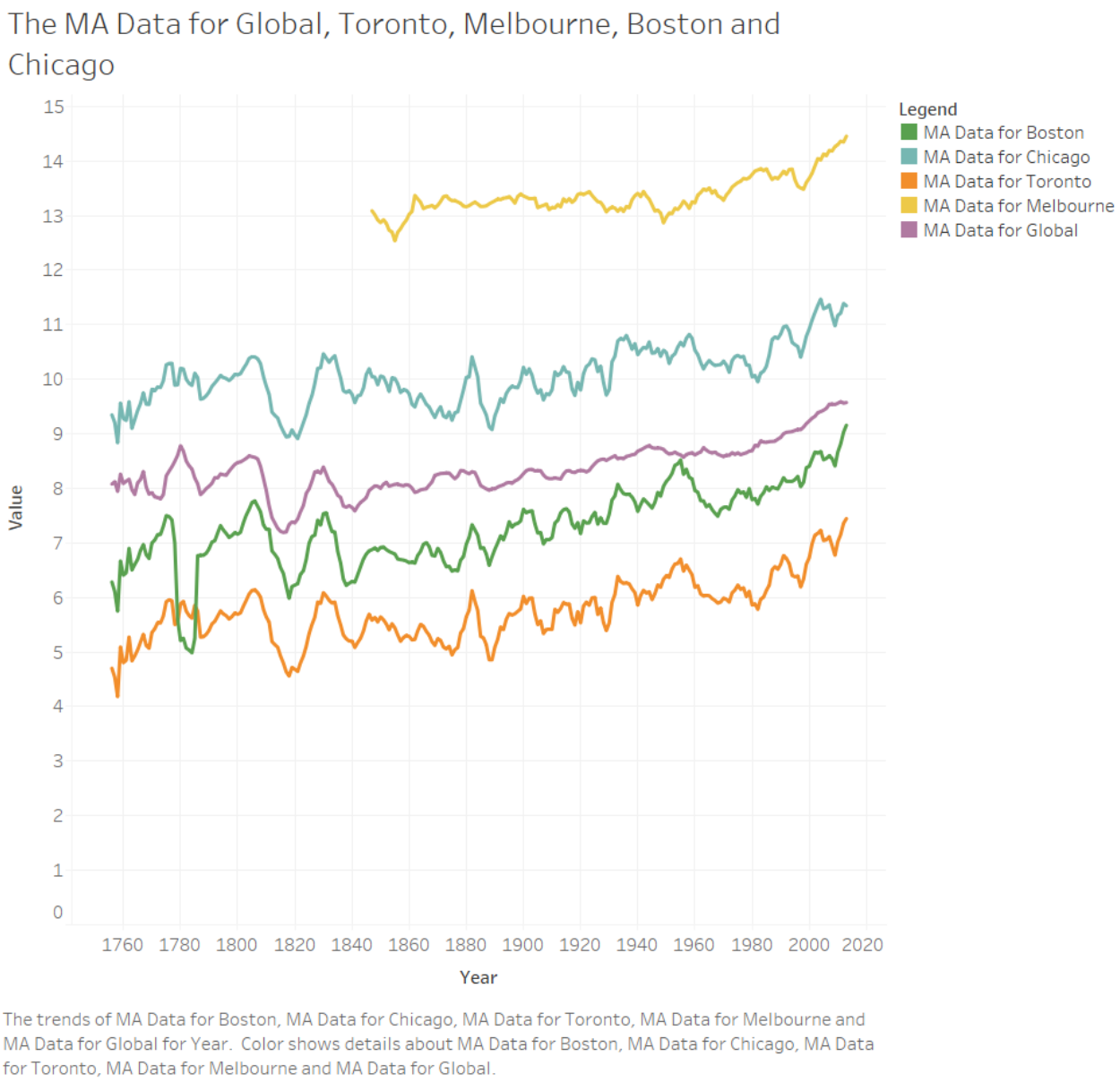

Figure 7: The Moving Averages for Global, Toronto, Melbourne, Chicago, and Boston Data.

Key Considerations

The numbers provided within the database for every city was simplistic enough to be read on both a CSV file and on an excel spreadsheet. Because the data was based on a timeline as the 'X variable' can be known as years and the 'Y variable' can be known as temperature, the perfect visualization method would be a line graph. Comparisons between every different data set on each different line graph can would also be more simplistic and efficient. Other methods of visualization such as area charts, tree maps, horizontal bars or pie charts would not produce a visualization perfect for viewing a dataset based on time.

Analysis

After visualizing these sets of graphs on Tableau, I took a look at every individual trend within every graph. My approaches to this analysis was based on the comparison of each and every graph for every different city analyzed. Instead of numbers being compared off a database, the data was visualized using the line graph method for an easier digestion for a general audience. Another key technique was to always use the global data average as a reference for every city analyzed. If a city has a greater increase in temperature throughout the centuries than the global trend, it can be considered worse when it comes to global climate change. If a city has a lower increase in temperature compared to the global trend, it can be considered better off in that regard.

While only looking at the moving average, the global data illustrates an increase of a temperature of 2 degrees Celsius. First, we will take a look at Toronto and compare its data with the global data. Toronto has rather been hit hard by global warming – more than the global moving average. Jumping from around 4.7 degrees Celsius to around 7.3 degrees Celsius. This almost shows a 3 degrees jump, as opposed to the 2 degrees jump seen within the global averages. If the average

temperatures for Toronto were to be compared to other cities around the world with a similar population and a similar economy, the results are as follows:

Chicago seems to further exemplify the global averages, as a jump from about 9.3 degrees Celsius to 11.3 degrees Celsius shows another change that reflects that of the world. At around the 1740 – 1750 mark, Chicago temperatures seem to dip and decrease significantly during those years. This data can be regarded as outliers as it would not make sense for temperatures to drop that low during that time. Although Chicago boasts a larger population and a stronger economy than that of Toronto, the latter city has taken a worse hit when it comes to rising global temperatures – particularly when compared to the global average, and Chicago's average. In addition, Toronto's line graph (Moving average) seems to boast many more spikes than Chicago. This represents the drastic (and elastic) changes in temperature within Toronto in the last couple of centuries.

Boston seems to have experienced the same rise in temperature as Toronto – almost 3 degrees Celsius has been added to the global average since the mid 18th century. However, just like with Chicago, near the beginning of the data from around 1740 – 1780, the city has witnessed large dips in temperature. This does not seem accurate and can be regarded as outliers as well. Both Toronto and Boston rank higher in temperature change compared to the global data. However, Toronto's line graph illustrates more uncertainty and less stability with its frequent spikes and changes in average temperature, which can change as much as 1.5 degrees Celsius in just around 10 years.

Melbourne is a different story. The graph depicting its moving average actually only shows an increase of one degrees Celsius – half of the global change in temperature. This would mean that Toronto and Boston have almost a temperature change that triples that of Melbourne's. Although Toronto is known to have the most uncertainty when it comes to its line graph as it illustrates many more spikes and plummets, Melbourne is the complete opposite. The line does increase slightly, but is amazingly kept straight in a span of two and a half centuries.

Conclusion

Global data has recently taken a turn for the worst in terms of global warming – which can be exemplified by different cities around the world. Toronto has a long way to go in terms of climate sustainability – the change in weather temperatures continues to outplay the change from the global average. Steps by the city to take action on this problem does not go unnoticed, as seen by Toronto's crazy spikes and changes in average temperature on its line graph. Stability seems to be one of the city's most tough challenges, as the city continues to have its average temperature increased after a cycle of spikes and plummets.