

Dataset Exploring and Exporting

Since the original dataset is in an exported sql format, I imported such file into MySQL database by using MySQL Workbench, and it turns out there are 3 data tables: global_data, city_list and city_data. After a few queries to explore the dataset, 3 SQL queries to generate the 3 CSV files (as below) have been implemented:

```
USE temperatures;
```

```
SELECT year, city, avg_temp FROM city_data  
WHERE city = "toronto" AND avg_temp IS NOT NULL;
```

```
SELECT city, country FROM city_list;
```

```
SELECT year, avg_temp FROM global_data WHERE avg_temp IS NOT NULL;
```

Line Charting the Weather Trends

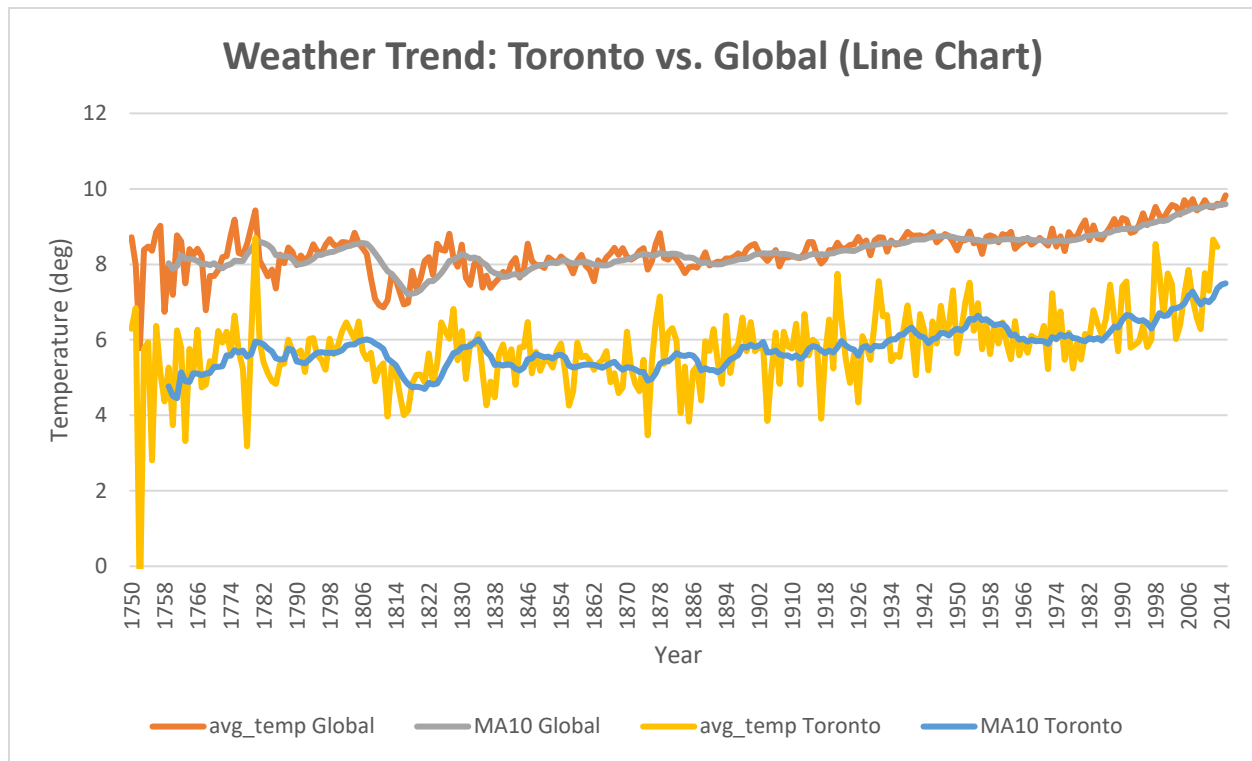
In order to explore the Weather Trends visually, 2 CSV files: city_data_toronto.csv and global_data.csv, have been imported into Excel application. The Line Chart or X-Y Scatter Plot of Excel empowers us to visualize the data and weather trends in an easy way.

After merging the 2 CSV files (said above) into an Excel spreadsheet, I also generated 2 columns: “MA10 Global” and “MA10 Toronto” as per the instructions of the assignment (Table 1 below).

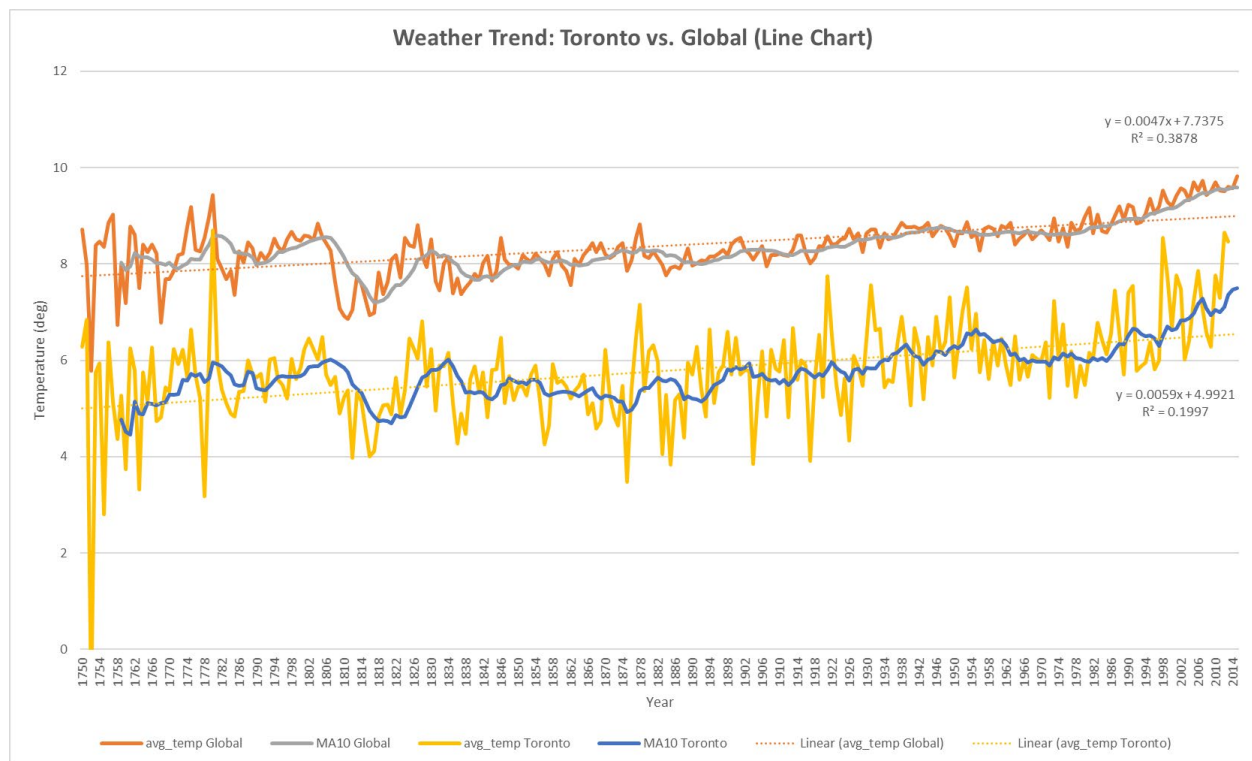
year	avg_temp Global	MA10 Global	avg_temp Toronto	MA10 Toronto
1750	8.72		6.29	
1751	7.98		6.84	
1752	5.78		-1.1	
1753	8.39		5.76	
1754	8.47		5.94	
1755	8.36		2.81	
1756	8.85		6.37	
1757	9.02		5.13	
1758	6.74		4.37	
1759	7.99	8.03	5.27	4.77
1760	7.19	7.88	3.74	4.51
1761	8.77	7.96	6.25	4.45
1762	8.61	8.24	5.79	5.14
1763	7.5	8.15	3.32	4.90
1764	8.4	8.14	5.76	4.88
1765	8.25	8.13	5.14	5.11
1766	8.41	8.09	6.27	5.10
1767	8.22	8.01	4.74	5.07
1768	6.78	8.01	4.81	5.11

Table 1 – Data Table of Avg_temp and the MA10's

The Line Chart can be drawn as below (Figure 1).



We can draw the Trend lines with regressions across the 2 datasets: Global and Toronto (Figure 2).



Observations

(Any 4 observations are good enough)

1. The average temperature trend of Toronto is moving along with the global trend. Both have been trending upwards over the period of 1750-2014. This might verify the Global Warming effect.
2. Both the local and global temperatures have experienced a greater fluctuation in 1816's ("big drop") is known as the "Year Without a Summer" because of severe climate abnormalities that caused average global temperatures to decrease by 0.4–0.7 °C. Summer temperatures in Europe were the coldest on record between the years of 1766–2000 (refer to Wikipedia for more details).
3. Toronto temperature seems to experience much more pronounced peaks and troughs in its data than overall global data. That's due to the "smooth-out" effect of the global temperature data.
4. The 10-year moving average for global temperatures is a lot smoother (less fluctuations) compared to the 10-year moving average for local temperatures of Toronto.
5. The MA10 temperatures of Toronto is ~ 2.2 degrees lower than the global temperatures most likely due to Toronto locating too "Northern" at latitude of 43.65°N.
6. The trendlines in Figure 2 reveals Toronto has a slightly steeper slope at 0.0059 versus global's 0.0047. Moreover, Global data has a higher R^2 , indicating Global temperatures are less volatile due to the "smooth-out" effect, hence smaller variance and higher R^2 .