EXPLORE WEATHER TRENDS

UDACITY NANODEGREE DATA ANALYST -

PROJECT 1

Name: Thu Trang Bui

Date: 8 June, 2019

1. Extracting data from the database with SQL

Firstly, I took a look at the data to have an overview of data with following SQL queries:

SELECT * FROM city_list

LIMIT 5;

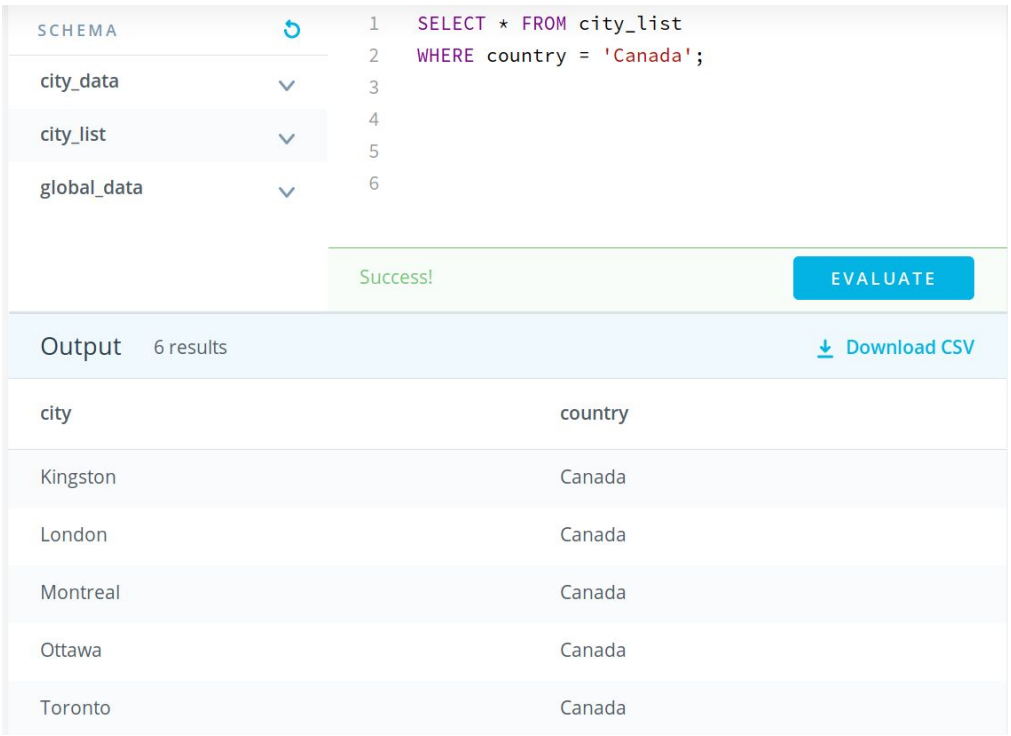
SELECT * FROM city_data

LIMIT 5;

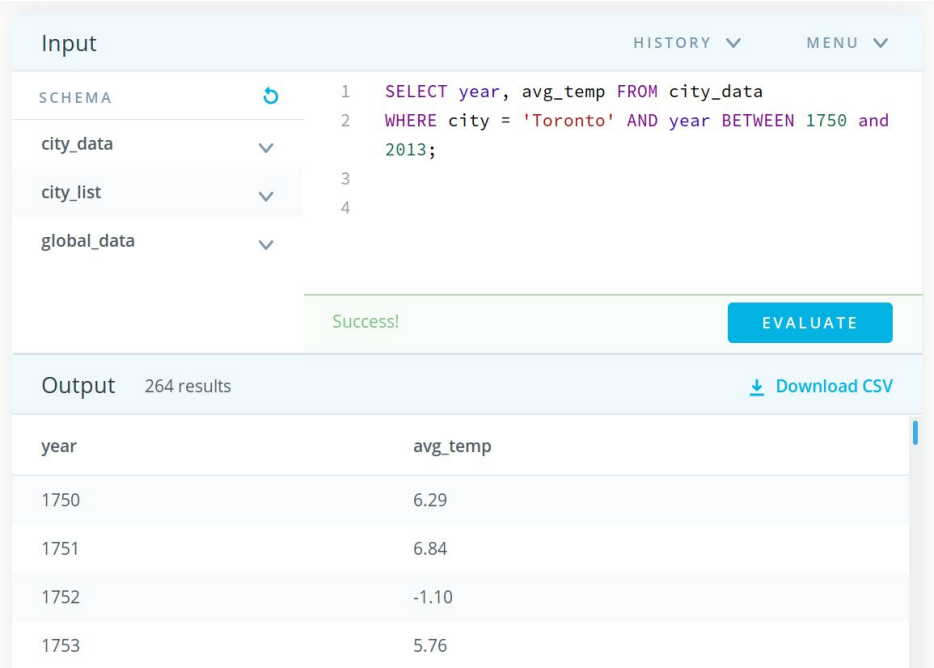
SELECT * FROM global_data

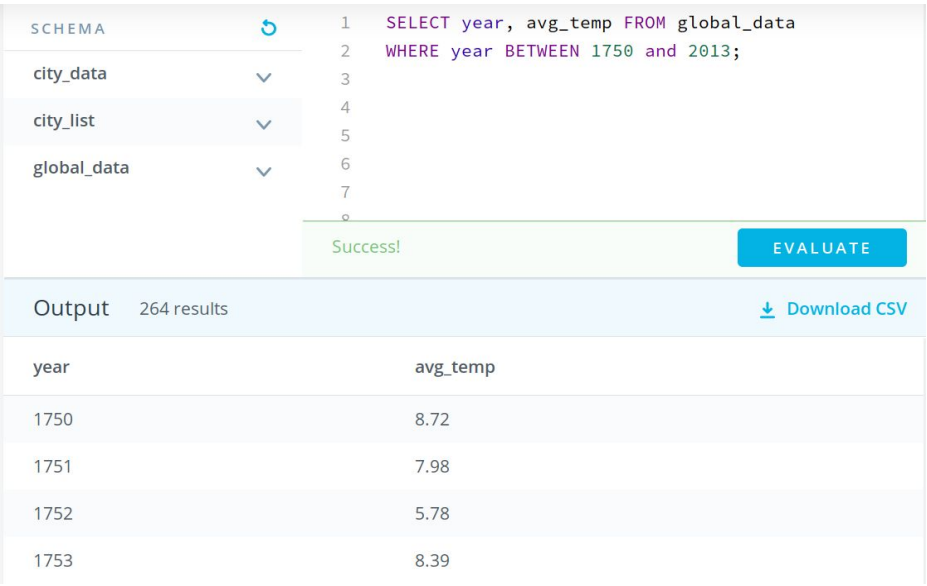
LIMIT 5;

Then I checked the data were available of which cities in Canada, the country I am living in:



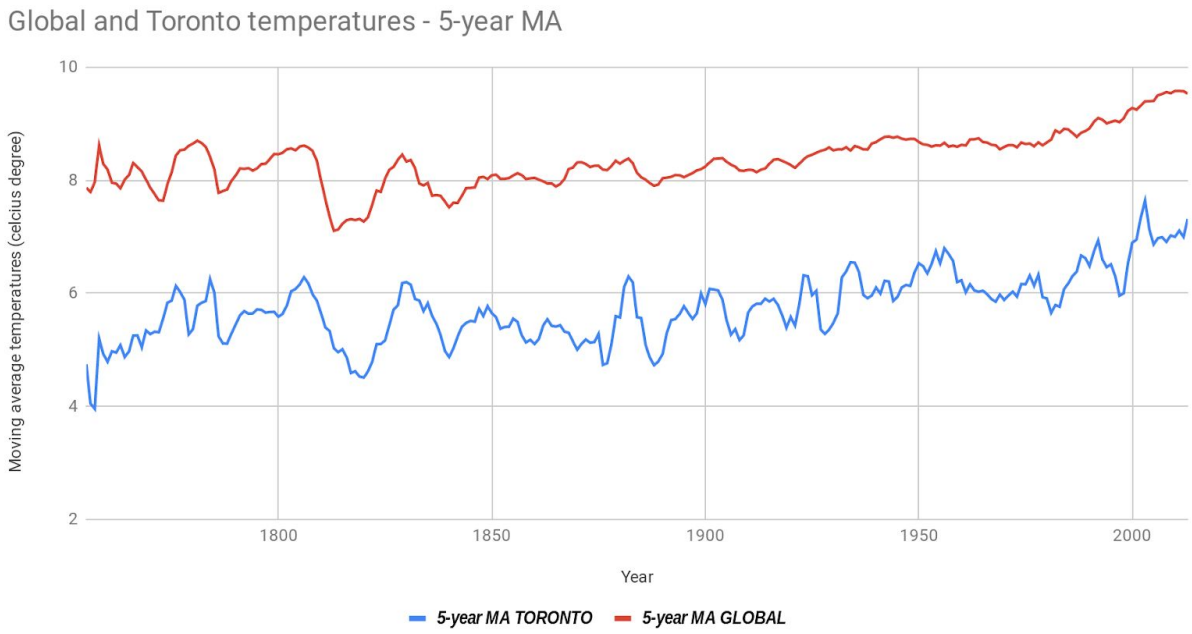
As I am living in Toronto, I choose this city to analyse. Hence, I extracted the data of Toronto. Moreover, looking at global_data file, I noticed that data of global temperature were taken from 1750 - 2015, while data of Toronto temperature were from 1743 - 2013. Thus, to be consistent in the year to yield a more accurate comparison, I selected the data of Toronto temperature and global temperature to be 1750 - 2013:





2. Data manipulation

After getting the desired data, I used Microsoft Excel to manipulate and visualise data. Since yearly temperature fluctuates too much, making it hard to define the trend, I decided to use the moving average in order to smooth out the fluctuation and easier to observe the long term trend. Then, I read some papers about calculating moving averages for temperature data to decide the time period of MA. Some papers state that 5-year moving average is suitable for temperature data in a long time, some recommend 10-year period is more preferable. Hence, I tried both methods. Firstly, I calculated 5-year moving average and got the below chart:

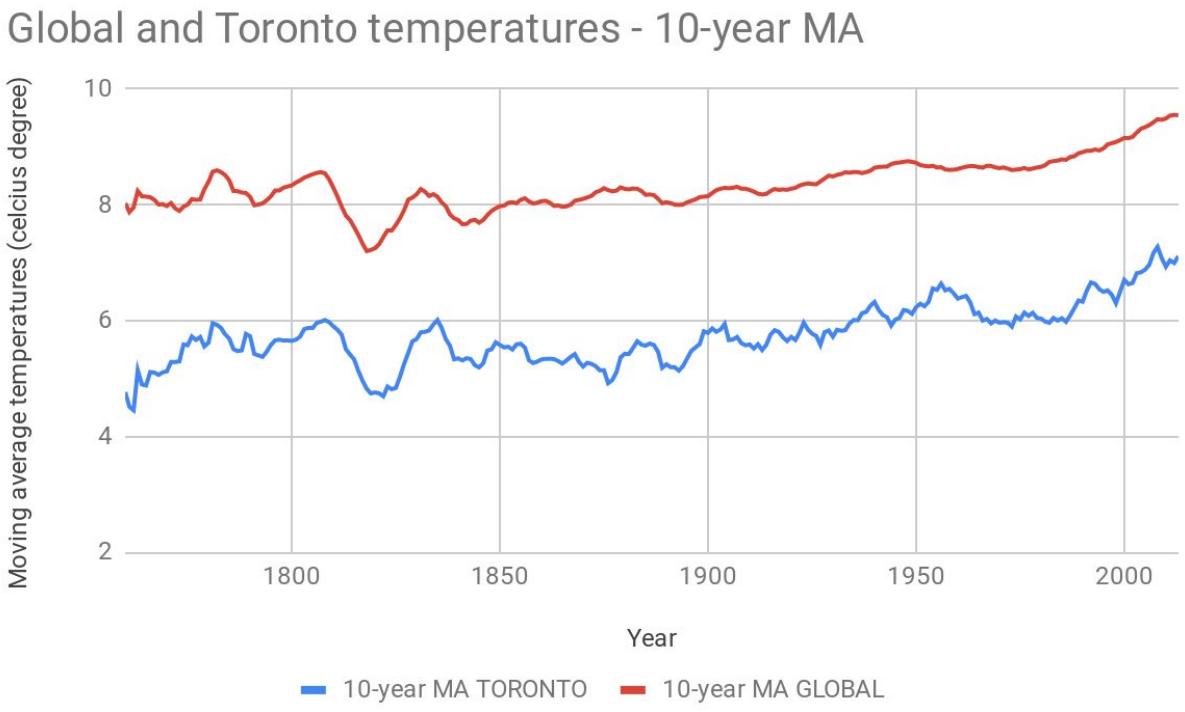


The moving average was calculated as follows:

I created a column next to the Toronto average temperature called 5-year MA Toronto, which was where the moving average field would be stored. I went down to the fifth year (1754, since the data start from the year 1750) and used the AVERAGE() function to calculate the

average temperature for the first five years. Then I copied the formula down to the end of the dataset to row 265 of the year 2013. Then I repeated the process with global average temperature.

10-year MA was calculated similarly. I created a column next to the Toronto average temperature called 10-year MA Toronto. I went down to the tenth year at row 11 (1759, since the data start from the year 1750) and used the AVERAGE() function to calculate the average temperature for the first ten years. Then I copied the formula down to the end of the dataset to row 265 of year 2013. After that, I used 10-year moving average to produce this chart:



As 10-year moving averages chart has smooth lines, making trends more observable, I decided to make observations and analysis about the data based on this chart rather than using 5-year moving average.

3. Observations about the local and global weather trend

First of all, it is clear the Toronto city has always been colder than global average temperature, since the blue line reflecting 10-year MA of Toronto is always below the red line of global average temperature.

Moreover, the difference in temperature between Toronto and the world has been very consistent over time. Looking at the chart, the distance between the two lines is quite stable, despite some points in which the gap can be smaller or bigger. Thus, it can be concluded that Toronto average temperature is lower than the global average temperature, and the gap of temperature is consistent.

Thirdly, it can be seen that the changes in Toronto's temperatures over time are in line with the changes in the global average. The two lines move in the same pattern, and the similarities in the changing pattern is stable. When the global average temperature went down during 1810 - 1820 period before going up to recover in the following 10 years, Toronto's average temperature also dropped significantly during 1810 - 1820 and went up back to the old level.

In addition, it is apparent that the global average temperature line is smoother than that of Toronto. In other words, Toronto's average temperature fluctuates more than global.

Last but not least, the overall trend of both global and Toronto temperature shows that the world in general, and Toronto in particular, is getting hotter, and this trend has been consistent from 1850 up to 2013.