Udacity Data Analytics

Project 1: Exploring Weather Trends

Summary

In this project, I will analyze local (Moscow, Russia and Munich, Germany) and global temperature data and compare the temperature trends in Moscow and Munich to overall global temperature trends.

Workflow

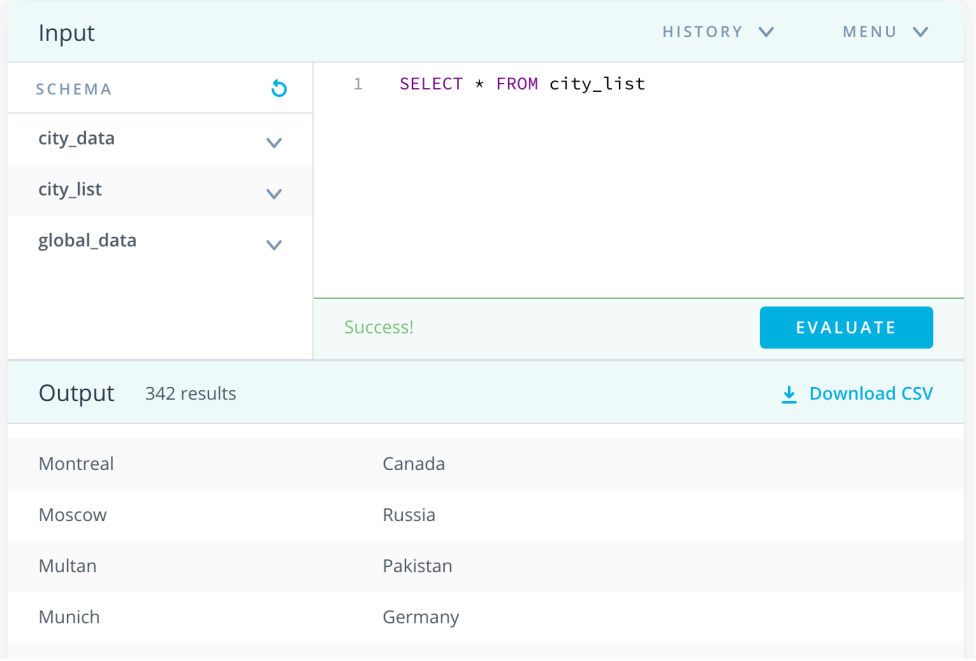
1) Selecting data for Moscow, Munich and the world
2) Downloading the resulting datasets into csv files.
3) Opening the csv files, cleaning the data.
4) Applying the moving average to transform the data.
5) Plotting a line graph based on the moving average data.
6) Making observations based on the line graph.

Tools used

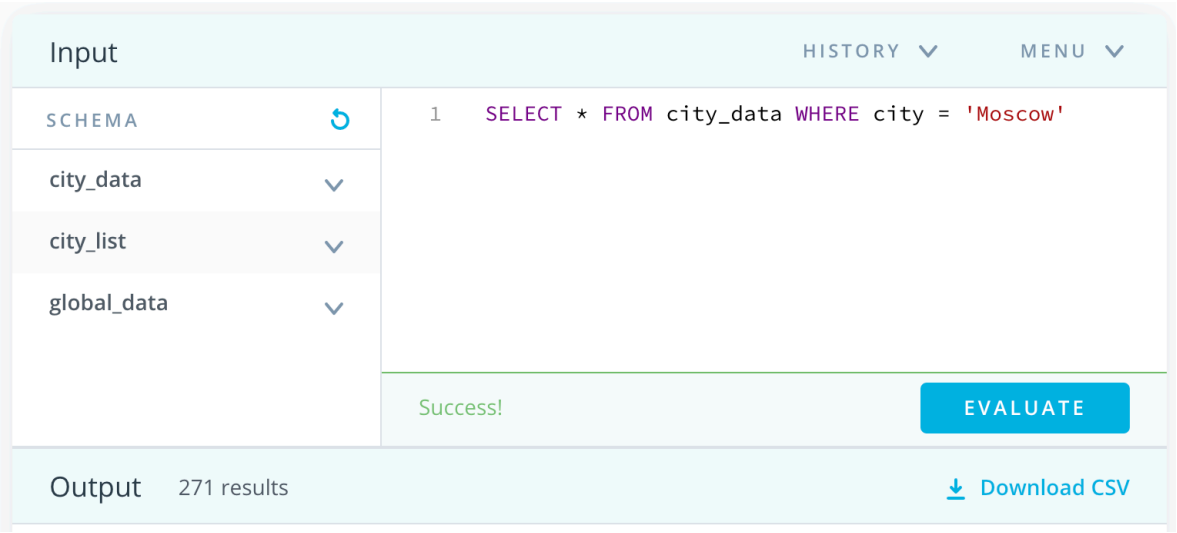
- SQL to extract data from the Udacity datasets
- Excel to prepare the data and plot the line graph

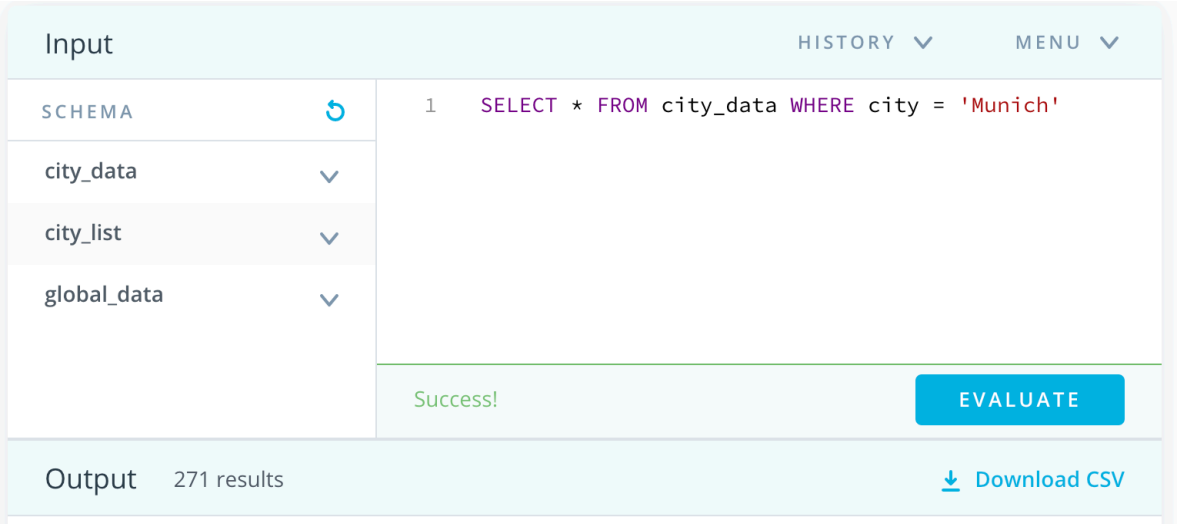
Selecting and downloading data

Using SQL-queries, make sure that the dataset contains information for Moscow and Munich. It does contain this information, as shown on the screen shot below:



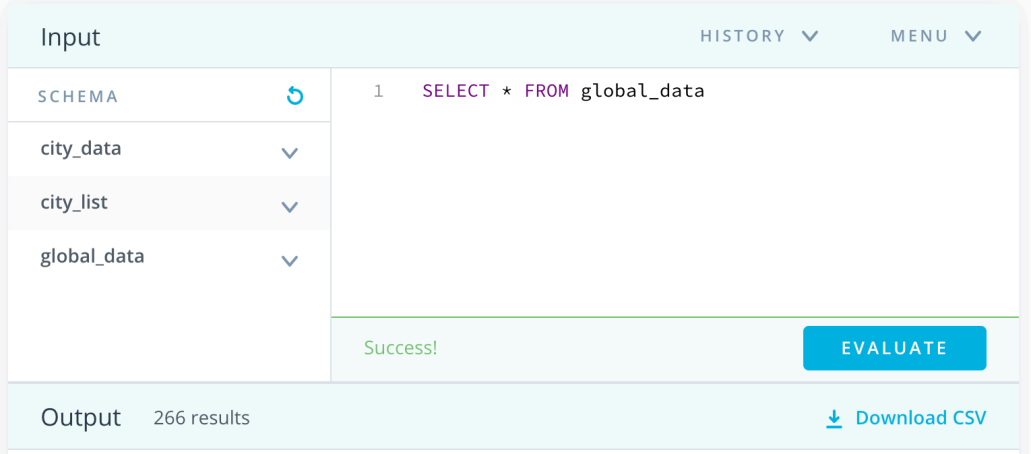
We see that both Moscow and Munich are in the dataset. So now, using SQL queries, we will select a subset for Moscow and download it to moscow.csv file and select a subset for Munich and download it to munich.csv file.





The datasets for Moscow and for Munich also contain 271 entries.

And, finally, we extract all global weather data and download it to global.csv file:

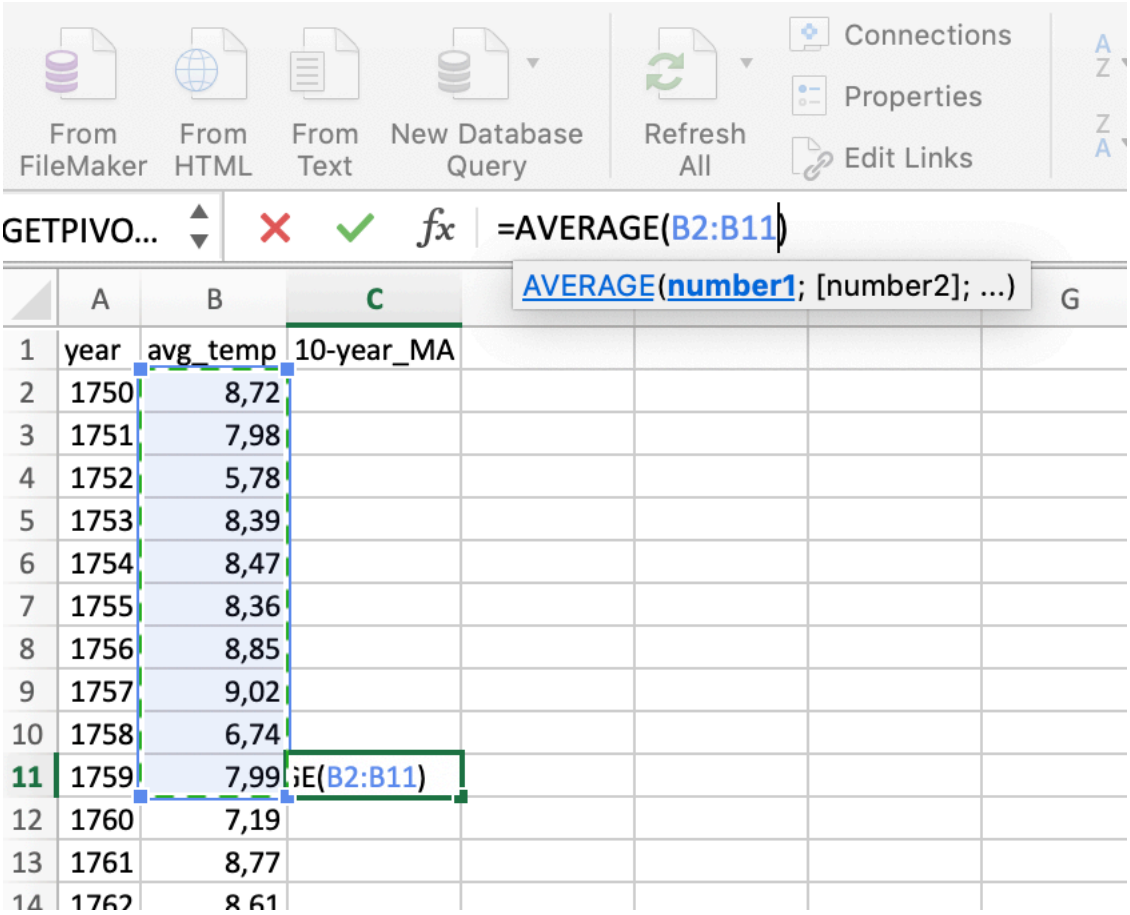


Opening and cleaning data

Having downloaded the datasets for Moscow, Munich and the world, I opened the three csv files in Excel using the option Data-From Text – by choosing the files and setting necessary parameters (comma as the delimiter and point as the separator for the number format).

Applying the moving average

Then I added the 10-year moving average column to all three datasets and calculated the moving average as shown on the screen shot below:

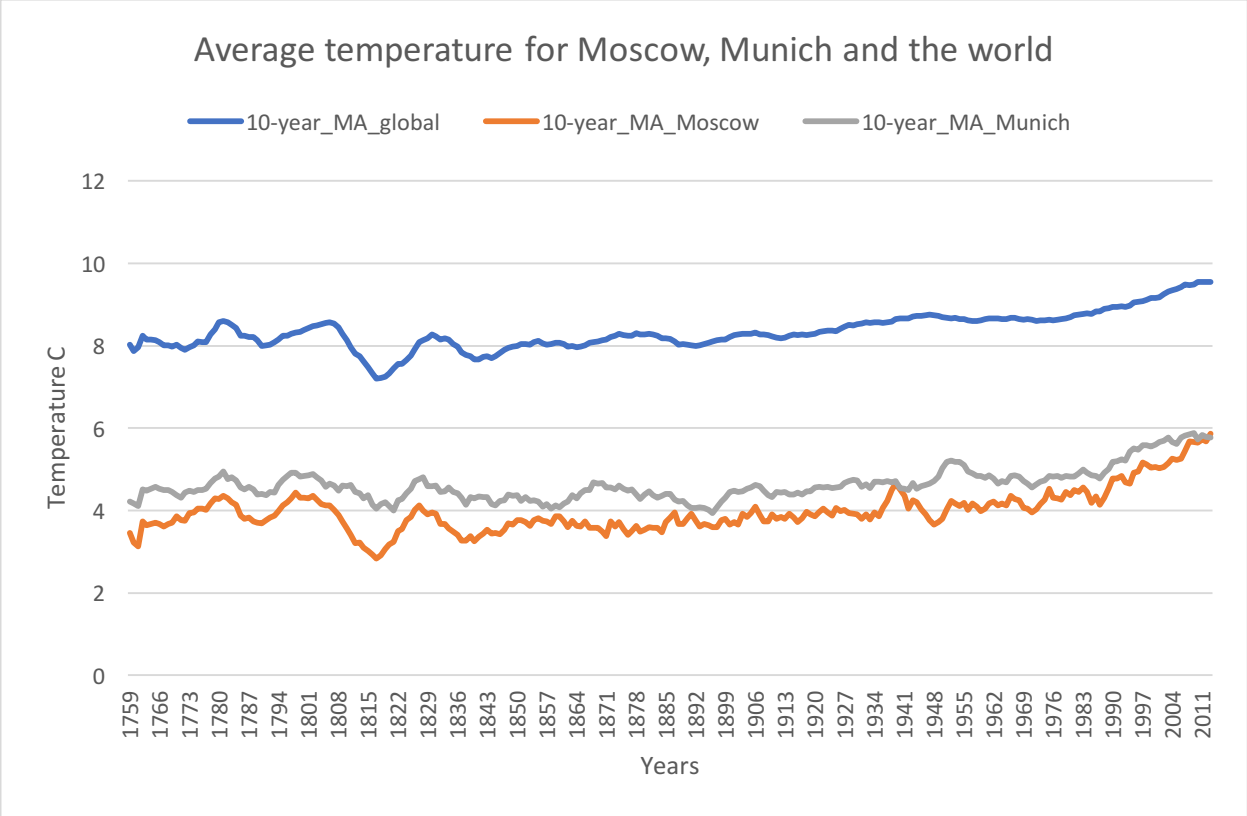


After calculating 10-year moving averages for the two cities and the world, I moved the 10-year_MA columns into one sheet to make it easier working with the data.
I noticed that the datasets did not have the same number of entries and I deleted those years for which there were no entries in at least one of the datasets.
These were years 1750-1758 and 2014 and 2015.

year	10-year_MA_global	10-year_MA_Moscow	10-year_MA_Munich
1750			
1751			
1752			
1753			
1754			
1755			
1756			3,436
1757			3,734
1758			3,508
1759	8,03	3,46	4,212
1760	7,877	3,217	4,174
1761	7,956	3,124	4,114
1762	8,239	3,742	4,51
1763	8,15	3,646	4,474
1764	8,143	3,674	4,523
1765	8,132	3,705	4,57
1766	8,088	3,677	4,534
1767	8,008	3,616	4,489

Plotting the line graph

Having cleaned the data, I used it as the basis to plot a line graph as follows:



Making observations

Based on the line graph I can make the following observations:

1) The average weather increased by approximately 1.3 C^o over the period of from 1760 until 2013, from approximately 8.0 C^o in the 1760s to 9.5C^o in 2010s.
2) The average temperature for Moscow and Munich does not differ that much as one might think these datasets were not available.
   The average temperature for Moscow is only 0.7 C^o lower than in Munich.
3) The average temperatures both for Moscow and Munich are considerably lower (by approximately 4.4 Co for Moscow and 3.7Co for Munich).
4) There was a period in 1816-1818 when the global weather went down by approximately 1.5C^o. Both Munich and Moscow followed this trend.
5) The moving average graph for global temperature is smoother than the moving average graph for Munich and even more so for Moscow.
6) Both Munich and Moscow graphs follow on average the same trends as the global temperature graph.
7) There were two periods when Moscow average temperature was almost equal to or even higher than that in Munich: in 1939 and in 2013.