

# **Data Analyst Nanodegree**

## **Exploring Weather Trends**

**By Tianfang Xie**

## Step 1. Extract the data from the database

Tools needed to be used : SQL

I live in the Indiana state of USA, so I used the SQL database to find the name of the nearest big city by inputting “**SELECT \* FROM city\_list WHERE country = 'United States'**”.

Then I find the nearest big city named “Chicago”.

The screenshot shows a web-based SQL query editor. On the left, under the 'Input' tab, there is a 'SCHEMA' section with a list of tables: 'city\_data', 'city\_list', 'city', 'country', and 'global\_data'. The 'city\_list' table is selected. In the center, the SQL query is entered: `1 SELECT * FROM city_list WHERE country = 'United States'`. Below the query, a green bar indicates 'Success!'. To the right of the success message is a blue 'EVALUATE' button. Below the input section, the 'Output' section shows '52 results'. A 'Download CSV' link is available. The output is a table with two columns: city names and 'United States'. The city 'Chicago' is highlighted in blue. At the bottom, there are 'Menu' and 'Expand' buttons.

City	Country
Birmingham	United States
Boston	United States
Charlotte	United States
Chicago	United States
Colorado Springs	United States
Columbus	United States
Dallas	United States
Denver	United States
Detroit	United States

Then I extracted the temperature data of Chicago from the database by inputting “**SELECT \* FROM city\_data WHERE country = 'United States' AND city = 'Chicago'**”

Input

HISTORY ▾MENU ▾

SCHEMA ↺

city\_data ^

year

city

country

avg\_temp ▾

1

SELECT \* FROM city\_data WHERE country = 'United States' AND city = 'Chicago'

Success!

EVALUATE

Output 271 results

Download CSV

year	city	country	avg_temp
1743	Chicago	United States	5.44
1744	Chicago	United States	11.73
1745	Chicago	United States	1.80
1746	Chicago	United States	
1747	Chicago	United States	
1748	Chicago	United States	
1749	Chicago	United States	
1750	Chicago	United States	10.49

^ Menu

Expand

Then I clicked the “Download CSV” button to download the csv file.

Similarly, I found the temperature of global by inputting “SELECT \* FROM global\_data” and then clicked the “Download CSV” button.

Input

HISTORY ▾MENU ▾

SCHEMA ↻

city\_data ▾

city\_list ▾

global\_data ▲

year

avg\_temp ▾

1 SELECT \* FROM global\_data

Success!

EVALUATE

Output 266 results

Download CSV

year	avg_temp
1750	8.72
1751	7.98
1752	5.78
1753	8.39
1754	8.47
1755	8.36
1756	8.85
1757	9.02

^ Menu ↗ Expand

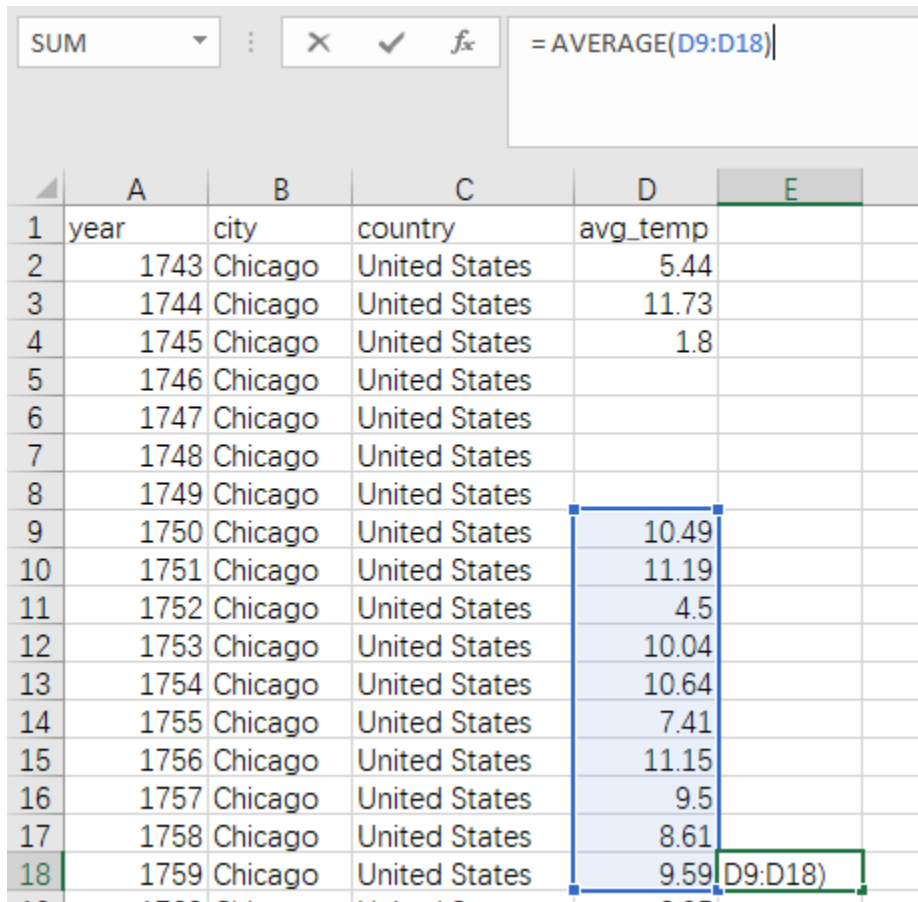
## Step 2. Open the csv file and calculate the moving average temperature

Tools needed to be used : Excel

I used the software Excel to open the csv files and calculated the moving average temperature by inputting “= AVERAGE(D9:D18)”.

The key considerations when deciding how to visualize the trends are:

1. Chicago lacks temperature data from 1746 to 1749, and the global temperature data begins from 1950, thus I need to begin the comparison from year 1750.
2. I used decade (10 years) as the choice of years to calculate the moving average temperature, this kind of choice fits people's custom.



The screenshot shows an Excel spreadsheet with the following data:

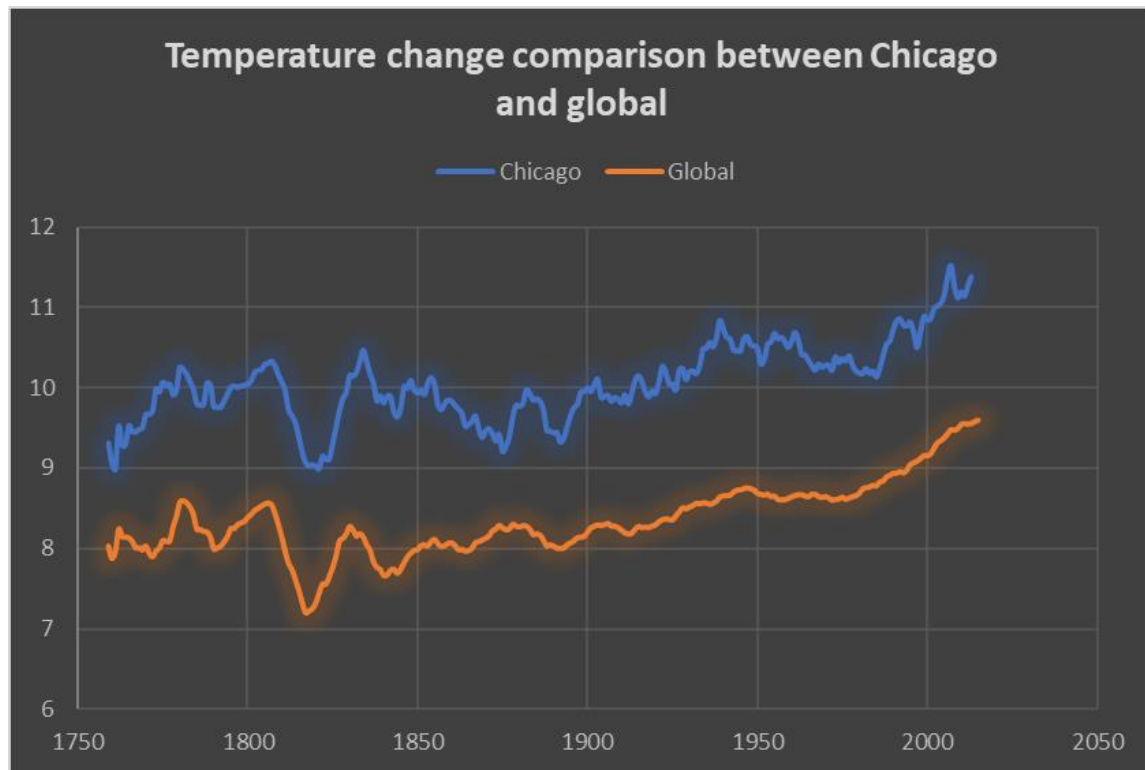
	A	B	C	D	E
1	year	city	country	avg_temp	
2	1743	Chicago	United States	5.44	
3	1744	Chicago	United States	11.73	
4	1745	Chicago	United States	1.8	
5	1746	Chicago	United States		
6	1747	Chicago	United States		
7	1748	Chicago	United States		
8	1749	Chicago	United States		
9	1750	Chicago	United States	10.49	
10	1751	Chicago	United States	11.19	
11	1752	Chicago	United States	4.5	
12	1753	Chicago	United States	10.04	
13	1754	Chicago	United States	10.64	
14	1755	Chicago	United States	7.41	
15	1756	Chicago	United States	11.15	
16	1757	Chicago	United States	9.5	
17	1758	Chicago	United States	8.61	
18	1759	Chicago	United States	9.59	

The formula bar at the top shows: `= AVERAGE(D9:D18)`

### Step 3. Create a line chart to compare Chicago's temperatures with the global temperatures

Tools needed to be used : Excel

I used the Excel to generate the line chart to visualize the temperature comparison.



#### Step 4. Make observations about the similarities and differences

Tools needed to be used : Excel

Observations:

1. The global average temperature is between 7.203 to 9.594 Celsius, and the Chicago's average temperature is between 8.989 to 11.517 Celsius, the Chicago city is hotter on average compared to the global average.
2. The difference between the global average and the Chicago's average is not consistent over time. For example, during the period from 1850 to 1875, the Chicago's average temperature was decreasing and the global average temperature was increasing.
3. During the past centuries, especially from 1850 to present, the changes in Chicago city's temperatures over time compare to the changes in the global average have the same trend. The temperatures of Chicago and global share the trend of increase.
4. For the overall trend, the world is getting hotter. The trend was consisted over the past century from 1900 to present. And the global temperature also have several significantly decrease trend during the period from 1800 to 1850.