

Explore Weather Trends Project

Here are the steps to prepare data:

1. First, select the biggest and nearest city near me, use sql:

```
SELECT * FROM city_list;
```

Click evaluate to get the city_list in csv file. Open csv file and use filter to filter the country as “United States”, then I can get the biggest, nearest city to me is “Boston”. Be careful, here Boston starts with capital letter.

The screenshot shows a data exploration tool interface. On the left, the 'Input' section displays a schema with columns: city_data, city_list, city, country, and global_data. A SQL query 'SELECT * FROM city_list;' is entered in the '1' slot. Below the query, a 'Success!' message and an 'EVALUATE' button are visible. The 'Output' section shows '342 results' and a 'Download CSV' link. A table of results is displayed with columns 'city' and 'country'. On the right, a spreadsheet view shows a list of cities and their countries, with 'Boston' highlighted in yellow.

city	country
Abidjan	Côte D'Ivoire
Abu Dhabi	United Arab Emirates
Abuja	Nigeria
Accra	Ghana
Adana	Turkey
Adelaide	Australia
Agra	India
Ahmadabad	India
Albuquerque	United States
Alexandria	United States
Arlington	United States
Atlanta	United States
Austin	United States
Baltimore	United States
Birmingham	United States
Boston	United States
Charlotte	United States
Chicago	United States
Colorado Sprin	United States
Columbus	United States
Dallas	United States
Denver	United States
Detroit	United States
El Paso	United States
Fort Worth	United States
Fresno	United States
Houston	United States
Indianapolis	United States
Jacksonville	United States
Kansas City	United States
Las Vegas	United States
Long Beach	United States
Los Angeles	United States

2. Then, use city_data by filter “Boston” as the city:

```
SELECT * FROM city_data
WHERE city='Boston';
```

Click evaluate and download csv to get the city_data in csv file.

The screenshot shows the same data exploration tool interface. The 'Input' section now displays a two-step SQL query: '1 SELECT * FROM city_data' and '2 WHERE city='Boston';'. The 'Output' section shows '271 results' and a 'Download CSV' link. A table of results is displayed with columns: year, city, country, and avg_temp.

year	city	country	avg_temp
1743	Boston	United States	1.19
1744	Boston	United States	9.63
1745	Boston	United States	-1.37
1746	Boston	United States	
1747	Boston	United States	
1748	Boston	United States	
1749	Boston	United States	

Use the same method to download global_data csv file:

```
SELECT * FROM global_data;
```

The Database Schema

There are three tables in the database:

- city_list - This contains a list of cities and countries in the database. Look through them in order to find the city nearest to you.
- city_data - This contains the average temperatures for each city by year (°C).
- global_data - This contains the average global temperatures by year (°C).

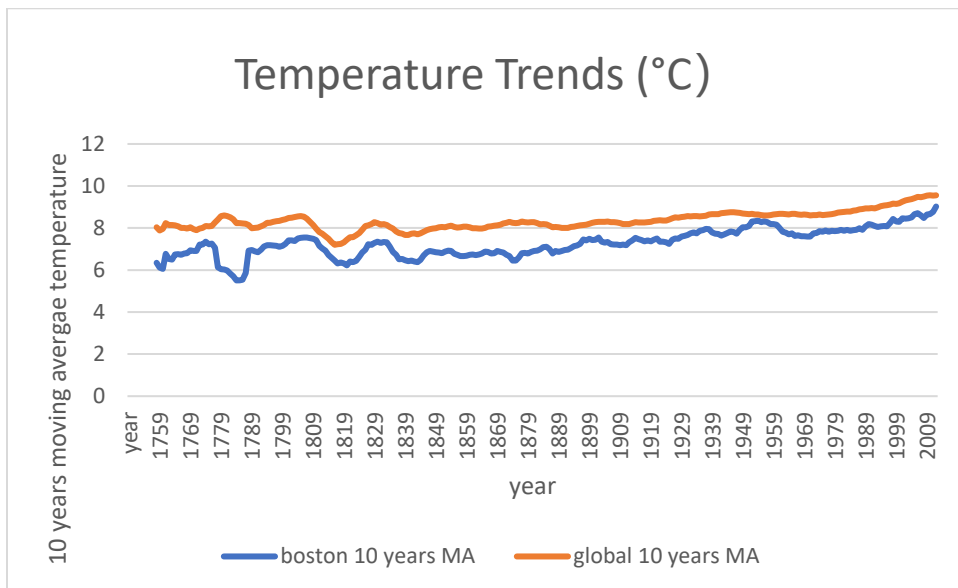
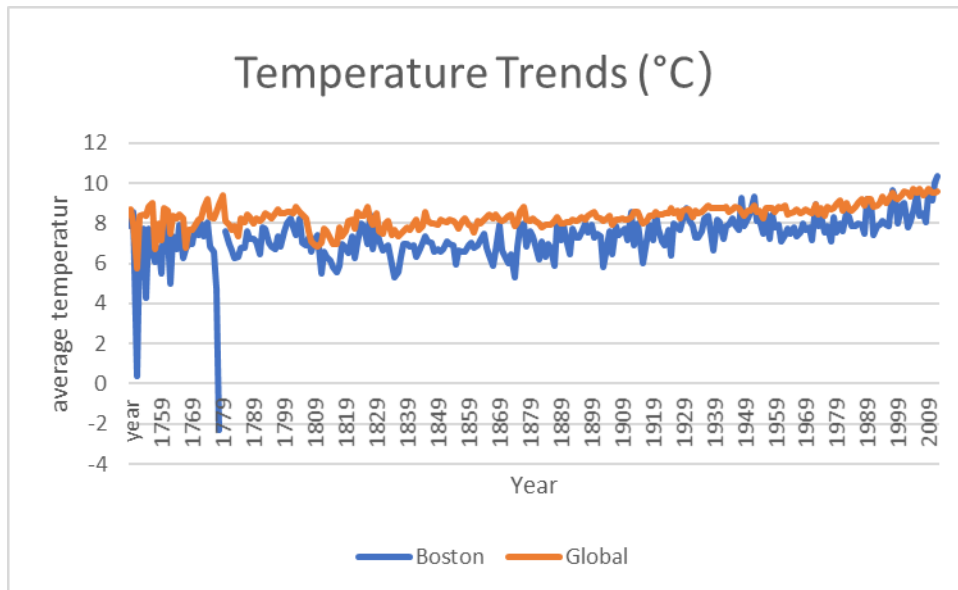
The screenshot shows a database query interface. In the 'Input' section, the query 'SELECT * FROM global_data;' is entered. Below the query, a 'Success!' message is displayed. In the 'Output' section, which shows '266 results', a table is displayed with two columns: 'year' and 'avg_temp'. The table contains data for years 1750 through 1756, with average temperatures ranging from 5.78 to 8.85.

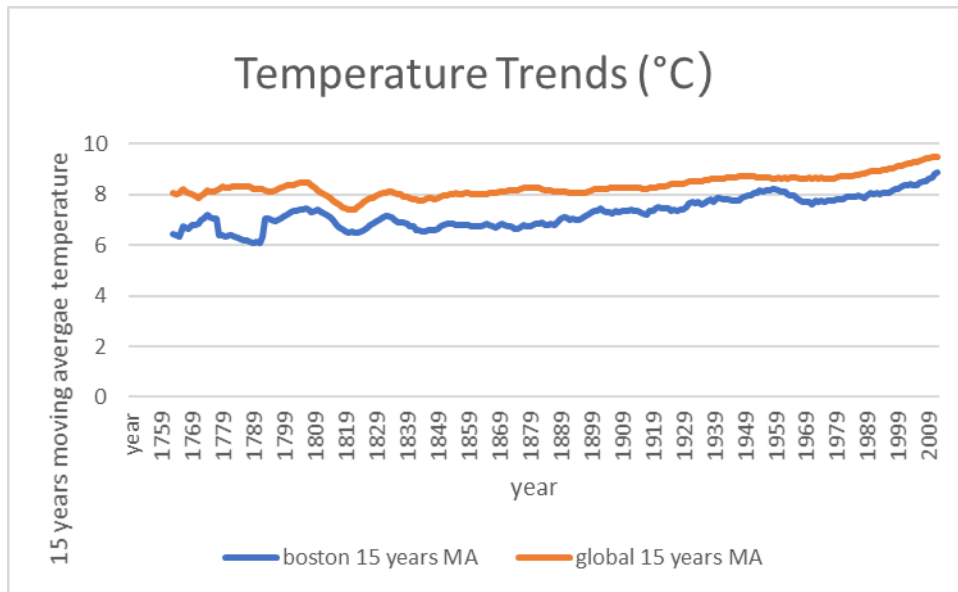
year	avg_temp
1750	8.72
1751	7.98
1752	5.78
1753	8.39
1754	8.47
1755	8.36
1756	8.85

After we get the data in csv file, we need to decide to use which tools to calculate the moving averages and make it visualize the trends. In this step, I use both excel and tableau to get the line chart.

Excel steps:

1. Open new excel and save it as "Explore Weather Trends"
2. Data->Get Data->From file-> city_data_boston.csv-> import table as table to the new worksheet
3. Repeat the step 2 to import the global_data.csv
4. After we get data in the new file, we can use them without destroy the original data. Now, join two tables from two worksheets to one table. I use inner join by years, because either left join or full outer join will make a lot of null values, null values will affect the visualize step. To inner join two tables, Data-> Get Data-> Combine queries-> Merge-> select city_data_boston and global_data and let them inner join by the matching year.
5. Then, use AVERAGE to calculate the 10 years and 15 years moving averages for both city_data_boston and global_data.
6. Use Insert-> Line Graphs to create the graphs for temperature trends, remember to add titles, axis titles and legends to make the graphs easy to read.
7. Here are the temperature trends graphs by every year, temperature trends graphs by 10 year moving average, and temperature trends graphs by 15 year moving average.

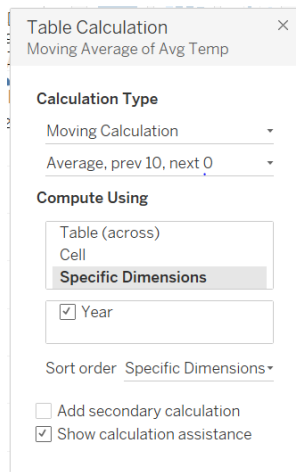




8. The graph for each year is too dense, and it cannot see the long-term trends in the graph. The 10 years moving average graph can give us a better view for the long term. However, 15 years moving average graph is smoother at some point. After comparison, 15 years moving average graph will be used for my result.

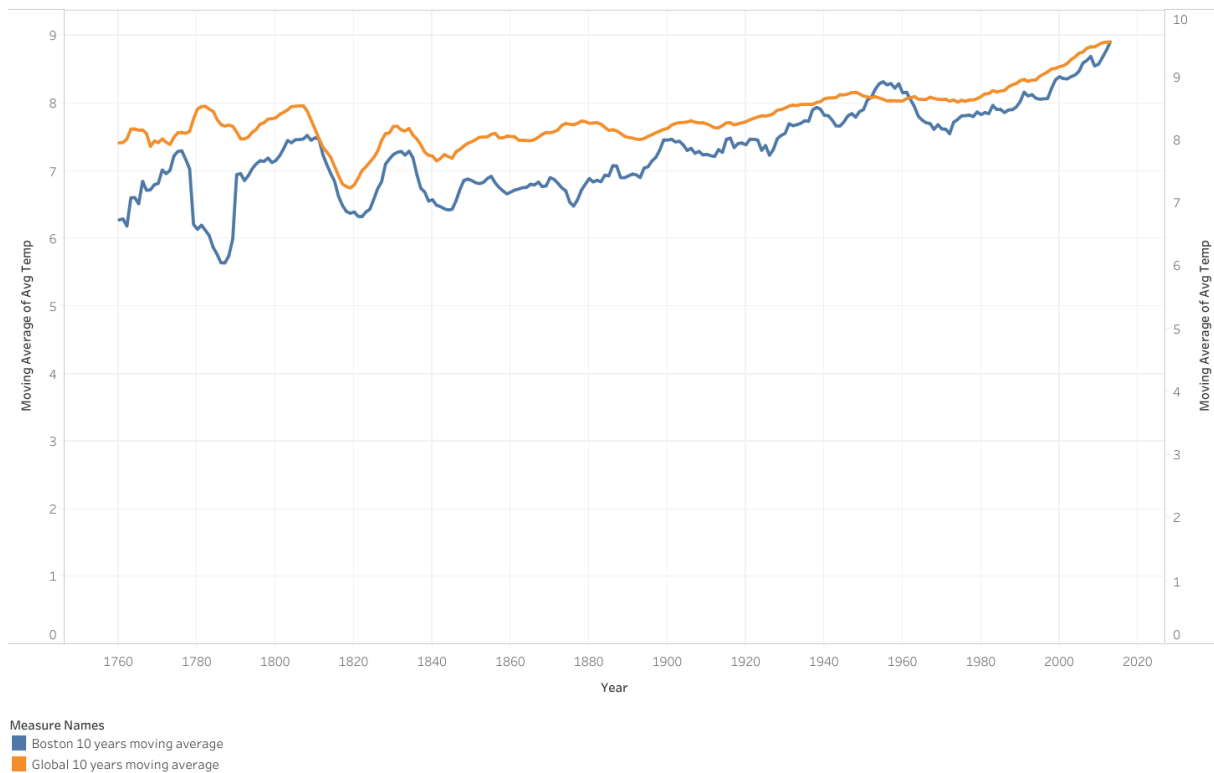
Tableau steps:

1. Tableau can join table and visualize the data in much easier way than the excel. When I use excel join, excel takes long time and almost crashed at first time. Tableau also can get the graphs much easier. First, open city_data_boston and create inner join with global data in the Data Source.
2. Go to a sheet, select year in the column.
3. Select boston_avg_temp and global_avg temp in the Rows, select dual lines in the show me. Now we can get a basic temperature trends graph with average temperature per year.
4. We already know the yearly graphs cannot tell the trends of the long-term. We need to add the calculation for the data, tableau have a really easy option called "Table Calculation", which can directly add the table calculation for the graph.
5. We do the table calculation and select moving calculation, select average and set 10 years for the value. I will also select "Null if there's not enough value", this can make my graph more accurate.



6. Repeat the table calculation steps for both Boston and global data, the 10 year graph is good to go. To make it more readable, I will create a Dashboard for the graph so that I can edit title, legends and axis title.

Temperature Trends (°C)

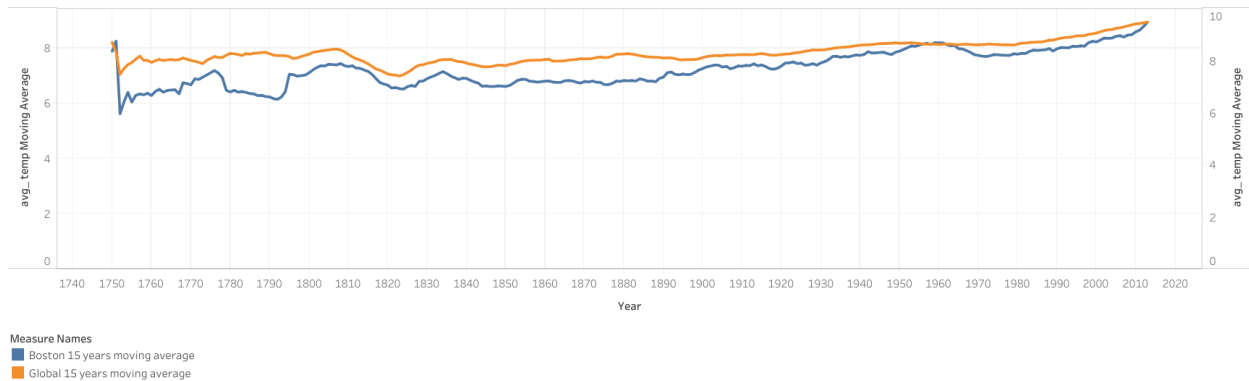


Click the link can go to my tableau public to view the graph:

https://public.tableau.com/views/temperaturetrends/Dashboard10yearMA?:language=en&:display_cou nt=y&:origin=viz_share_link

7. Since the graphs is not smooth, I will repeat step 5 and 6 to create other graph for the 15-year moving average of Avg Temp. Here is the 15 years moving average graph, which is perfect for the it visualize the trends.

Temperature Trends (°C)



Click the link can go to my tableau public to view the graph:

<https://public.tableau.com/profile/sookie.yuan#!/vizhome/temperatretrends/Dashboard15yearMA>

Observations:

1. My city Boston is cooler on average compared to the global average, except year 1951,1962, and 2013. In these three years, the average temperature in Boston and global is the same. Therefore, we can find Boston is always cooler on average compared to the global average.
2. From 1750 to 2013, we can see the average temperature in both Boston and global increase. However, the temperature difference of the Boston is over 3.6 °C, the global temperature difference is around 2.3 °C.
3. The overall trend is increases over the time, and the increase trend and decrease trends are almost same for Boston and global temperature.
4. From two line, we can also find Boston was cooler during the 1776-1795. However, in next two hundred years, Boston's weather trend is same as the global. This shows a long-term global warming, and it will affect all the place. The effect of global warming in cold places will be more pronounced.