

Explore Weather Trends

→ Database Queries

The following Steps were taken to extract the required data from the database:

1. The following queries were written to check data in the tables of the given database:

```
SELECT * FROM city_data LIMIT 10;  
SELECT * FROM city_list LIMIT 10;  
SELECT * FROM global_data LIMIT 10;
```

2. Query to extract global level average temperatures and respective years and sort it according to the year column:

```
SELECT * FROM global_data ORDER BY year;
```

3. Then in order to find the required country in the table of city_list following query was made:

```
SELECT DISTINCT country FROM city_list;
```

The name of the country was present ('India')

4. After finding the country name ('India') in the database, it was required to find the names of the cities and the nearest city. For this following query was made to get the list of cities in India in the given database.

```
SELECT city FROM city_list WHERE country = 'India';
```

The nearest city to the location was 'New Delhi'

5. Query to extract average temperatures per year for New Delhi and sort it according to the year column:

```
SELECT * FROM city_data WHERE city = 'New Delhi' ORDER BY year;
```

6. Data from steps 2 and step 5 was exported to CSV files.

→ Important Notes and Assumptions

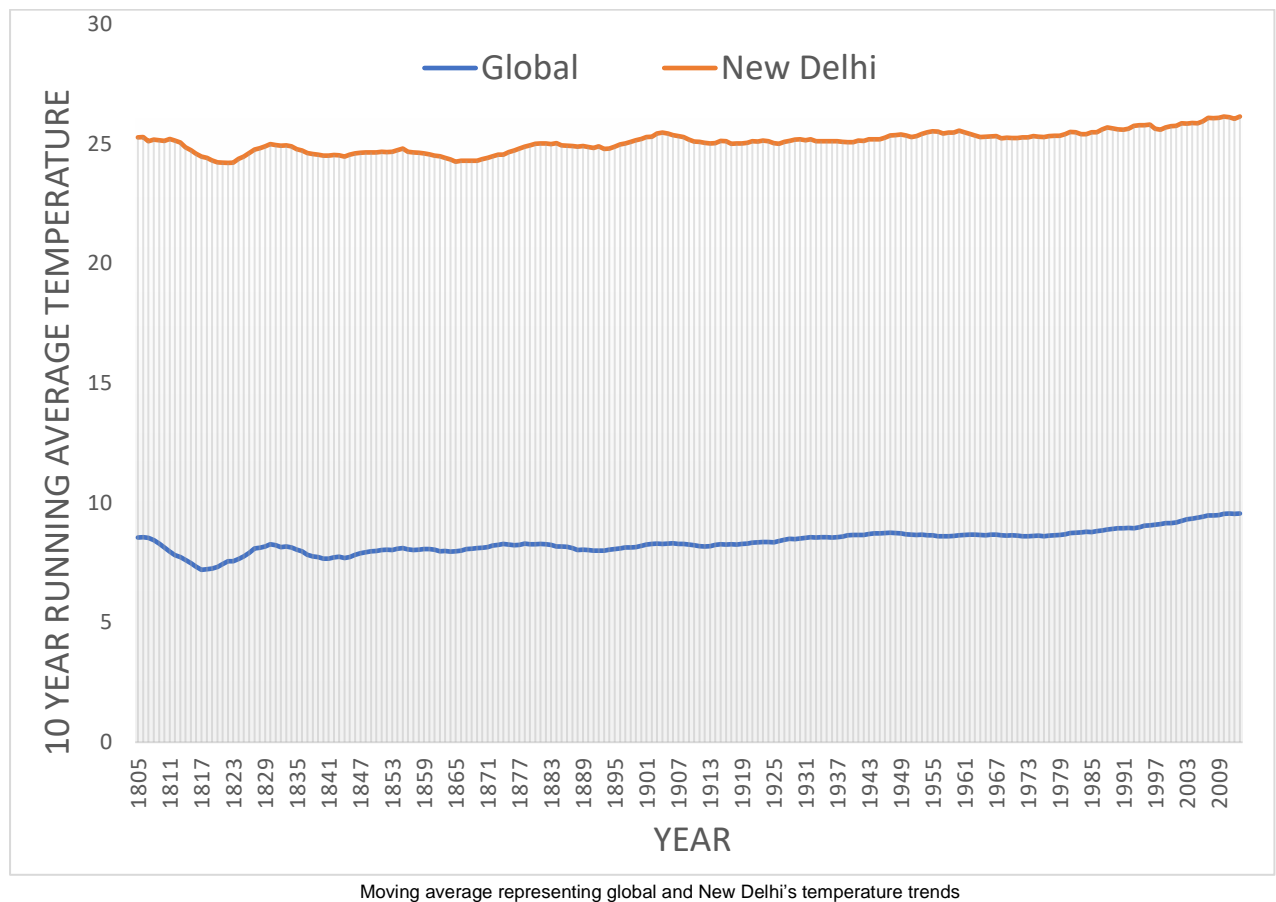
- In the given database, average temperatures were available for all the years from **1750 to 2015** for global data. But for the city of New Delhi, data was available from **1796 to 2013**.
- Also, for the city of New Delhi data was missing for the years **1808 to 1813** and **1858 to 1869**.
- Thus, the analysis is done only for years between **1796 and 2013** (data available for both cases)
- For the missing data for years **1808 to 1813**, it was assumed that the average temperature has remained constant and the same as that of the preceding year of **1807** (i.e., **24.97 °C**) and for the years **1858 to 1869**, the average temperature was the same as the preceding year of **1857** (i.e., **24.3 °C**)

→ Calculating 10-Year Moving Averages and Correlation coefficient

1. Both the CSV files were opened in **Excel**.
2. Data was edited in accordance with the assumptions mentioned in the previous section.
3. A separate column called **10-year MA** for both datasets was created. Then **AVERAGE()** function for the first 10 years was used and then dragged the column to call for the rest of the years in Excel.
4. Correlation coefficient between years and respective average temperatures was calculated for both sets of data using the **CORREL** function in excel. It turns out to be **0.838** and **0.688** for global data and New Delhi data respectively for the years between **1796 and 2013**.

→ Creating Data Visualization

Respective data were copied to a common excel file and a common **Line Chart** was made in **excel** for the **10-year moving averages** for both the sets of data.



→ Observations

The following observations were made based on the line chart and Correlation Factor:

1. The city of New Delhi is much hotter as compared to the Global temperature on average (**approx. 17°C**) and this is consistent over the time period.
2. As compared to global temperature, temperature trends of New Delhi have been more volatile for the last two centuries (especially after **1843**)
3. Between **1813** and **1840** there was a similar pattern decrease, then an increase and then again, a decrease of the pattern.
4. 10-year moving average for Global temperature reached a minimum temperature of **7.203 °C** in the year **1817** and **24.209 °C** in **1822** for New Delhi (both events occurred within a span of 5 years)
5. Maximum value of the global 10-year moving average temperature was observed in **2013** and was **9.556°C**. For New Delhi, the maximum value was **26.146°C** in **2013**. (2013 is the latest year into observation for the cases)
6. After **1890** there has been no significant dip in global temperature and a steady increase is observed till **2013**.
7. New Delhi's Temperature became quite steady between **1910** and **1950** with minor fluctuations of reading but global temperature saw a steady increase for the same time period.
8. Past century has seen an increase in temperature in both cases with a comparable pattern. Indicating that the world is getting hotter.
9. Correlation Coefficient between years and respective average temperatures per year for Global data comes out to be **0.838**. This indicates that there is a high correlation between years and global temperatures and the correlation is positive. This satisfies the conclusions drawn from the line graph.
10. Correlation Coefficient between average temperatures and respective years for New Delhi turns out to be **0.688**. This shows that there is a positive correlation between average temperature and respective years, though the correlation is not as high as Global temperatures trends.