Project : Exploring Weather Trends

Outline of the steps followed:

- **Extracting the data**
- **Cleaning the data**
- **Data Exploration**
- **Data Visualization**
- **Observations**

Extracting the data

I extracted data for 2 cities - New York in United States (where I currently live) and Pune in India (where I belong to) for exploring and comparing the global weather trends.

Following SQL queries were used to extract data from schema:

- **City data for New York, United States**

```
select * from city_data where country ='United States' and city ='New York';
```

- **City data for Pune, India**

```
select * from city_data where country ='India' and city ='Pune';
```

- **Global data**

```
select * from global data;
```

Data Cleaning

After extracting data in CSV format, Data Analysis, Exploration and Visualization is done in Microsoft Excel. Now, it is observed that Pune temperatures are available from 1796, New York temperatures from 1743 and Global from 1750. So, for comparing all temperatures in the same time frame, we will keep data from year 1796 onwards and remove rest years from analysis.

Also, few of the temperatures are missing for Pune city, so to fill in those gaps, we will use the average temperature for Pune throughout the years. Therefore, the missing values for temperatures in the Pune city data were replaced by 24.6 degrees for further analysis.

Data Exploration

Now, for better visualization, we use 10 year moving average of temperatures rather than yearly averages. Moving averages will smooth out data to make it easier to observe long term trends and not get lost in daily fluctuations.

Here is the glimpse of the data with moving averages calculated for New York, Pune and Global temperatures.

year	city	country	avg_temp	New York	Pune avg t	Pune temp MA	global temp	Global temp MA
1796	New York	United Sta	8.97		24.39		8.27	
1797	New York	United Sta	8.85		25.17		8.51	
1798	New York	United Sta	9.71		24.05		8.67	
1799	New York	United Sta	9.19		24.68		8.51	
1800	New York	United Sta	9.58		24.67		8.48	
1801	New York	United Sta	10		23.94		8.59	
1802	New York	United Sta	10.29		25.18		8.58	
1803	New York	United Sta	10.07		24.95		8.5	
1804	New York	United Sta	9.69		25.33		8.84	
1805	New York	United Sta	10.43	=AVERAGE(D2:D11)		24.716	8.56	8.551
1806	New York	United Sta	9.38	9.719	24.68	24.745	8.43	8.567
1807	New York	United Sta	9.17	9.751	24	24.628	8.28	8.544

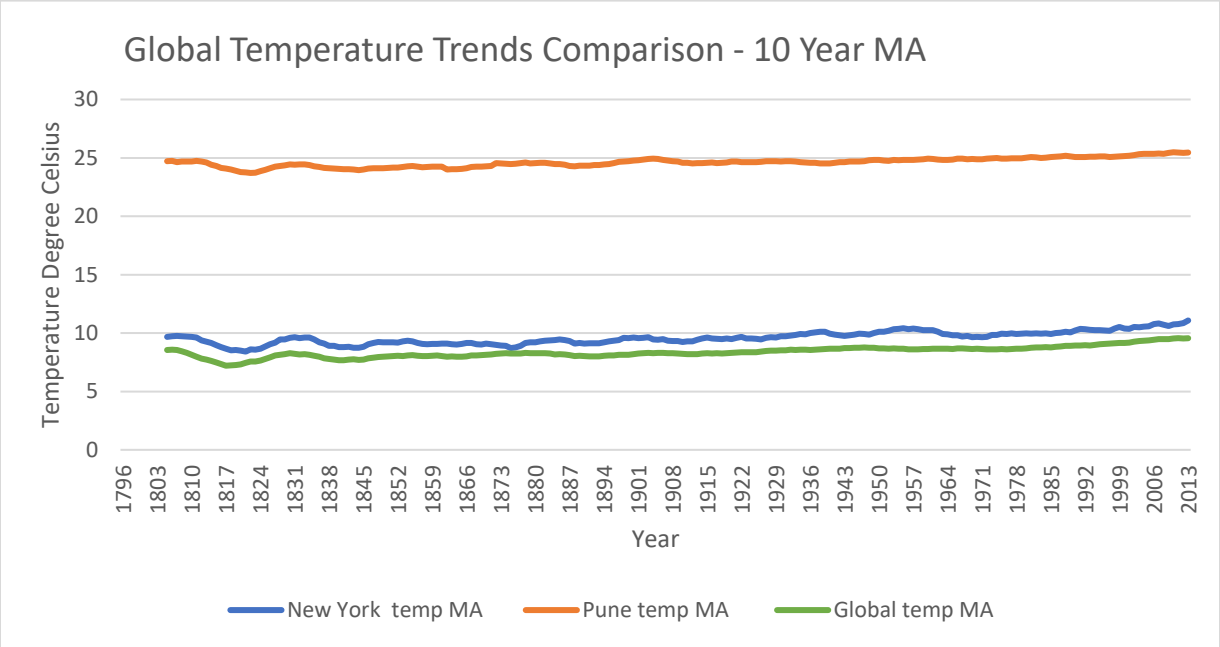
The correlation coefficients are calculated for Global vs New York and Global vs Pune temperatures moving averages.

Correlation coefficient for Global Vs. New York = 0.93

Correlation coefficient for Global Vs. Pune = 0.91

Data Visualization:

For better visualization of the trends, a Line chart is plotted with Years on the X axis and the Moving Average Temperatures for 3 categories – Global, New York and Pune on the Y axis.



Observations:

- The minimum global average temperature observed is 7.2 degrees (around 1820) and maximum is 9.55 degrees(2013)
- There is an upward trend for global average temperatures as well as Pune and New York average temperatures as observed from past 200 years
- For about a decade from 1817, it is observed that there was a dip in temperature globally as well as for Pune and New York
- The climate of New York is continental with hot summers and cold winters and its average temperature is around 10 degrees which is closer to the average global temperature
- But as Pune has a tropical climate, it has an average temperature around 25 degrees and is way hotter than New York and Global average
- The Global temperature ranges from 5 to 10 degrees on average from the 19th to 21th century but there is an observed increase at the beginning of the 21st century which shows that the climate is changing rapidly now
- There exists a high positive correlation between both Global and New York temperature moving averages (Coefficient of 0.93) and Global vs Pune temperature moving averages (Coefficient of 0.91)