Explore Weather Trends

Overview

Analysis of the temperature trends between a local city - Bangalore, and the global averages from the 19th to 21st century.

Data and Model

Data was downloaded from Udacity SQL workspace in the csv format. The SQL queries used to query the database are:

- `Global data : SELECT * FROM global_data;`
- `Find city : SELECT * FROM city_list WHERE country='India';`
- `Bangalore data : SELECT * FROM city_data WHERE city='Bangalore';`

The downloaded data had to be modified in the following ways in order to make a correct analysis.

- The temperature averages were available from the years 1796 (Bangalore) and 1750 (Global). I considered from the year 1796 for both the lists, until the year 2013.
- Eliminated all the null tuples from the Bangalore temperature averages list.
- Bangalore spreadsheet had columns of 'country' and 'city' which had to removed and graphs were plotted only considering the 'year' and 'avg_temp' columns.

The python libraries - pandas and numpy were used find the moving averages. I have used the Simple Moving Average (SMA) formula:

$$SMA = (A1 + A2 + A3 + .. + An)/n$$

$A_n =$ the average temperature of a year

$n =$ the period size (number of years)

The pandas library was also used to get the correlation between the global and local averages.

The python library sklearn was used to fit a linear regression model to predict the local temperatures given the global temperature averages.

All the temperature measurements are in °**Celsius**.

Results

The python library matplotlib was used to plot the following graphs:

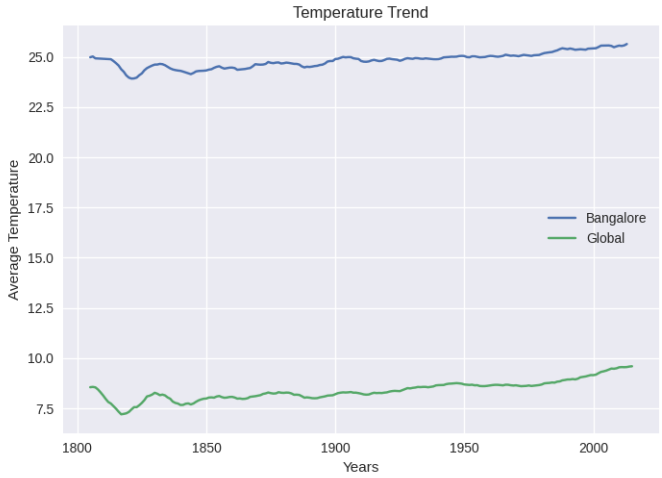

Figure 1: Linear graph describing the overall trend of temperature

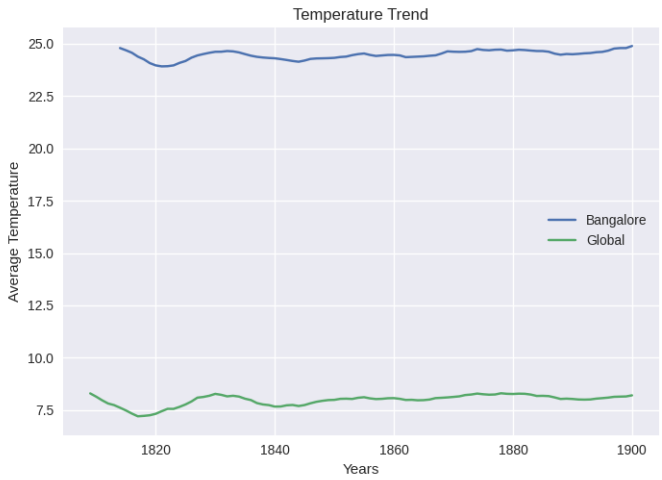

Figure 2: Linear graph describing the trend of temperature in the 19th century

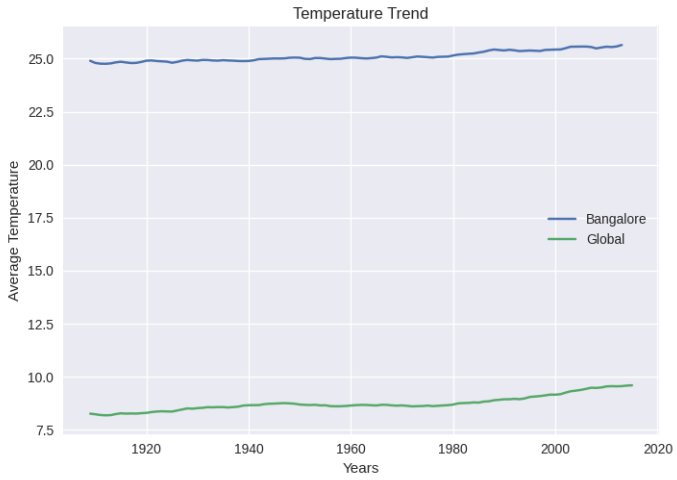

Figure 3: Linear graph describing the trend of temperature in the 20th century

As can be observed from the above graphs, the overall trend in the 200 years of data observed is almost linear and identical in both the local area and globally. Also, we see that:

- Global temperature averages:
  - Highest global temperature average = 9.83
  - Lowest global temperature average = 5.78
- Local temperature averages:
  - Highest local temperature average = 26.61
  - Lowest local temperature average = 23.3

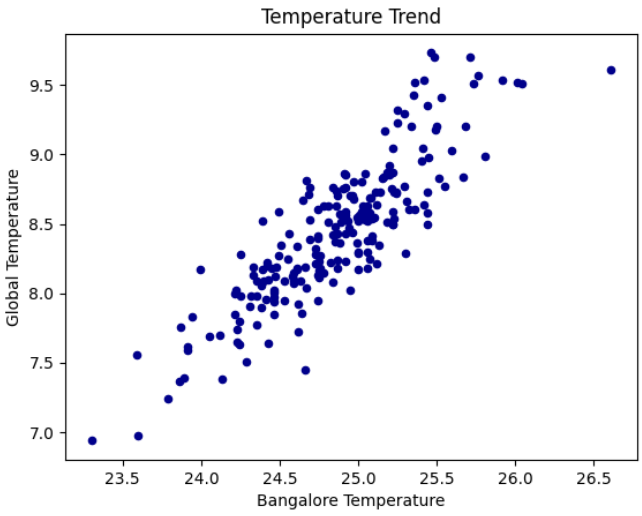

Figure 4: Scatter plot showing the relation between the local and global moving averages

The correlation coefficient is **0.86**

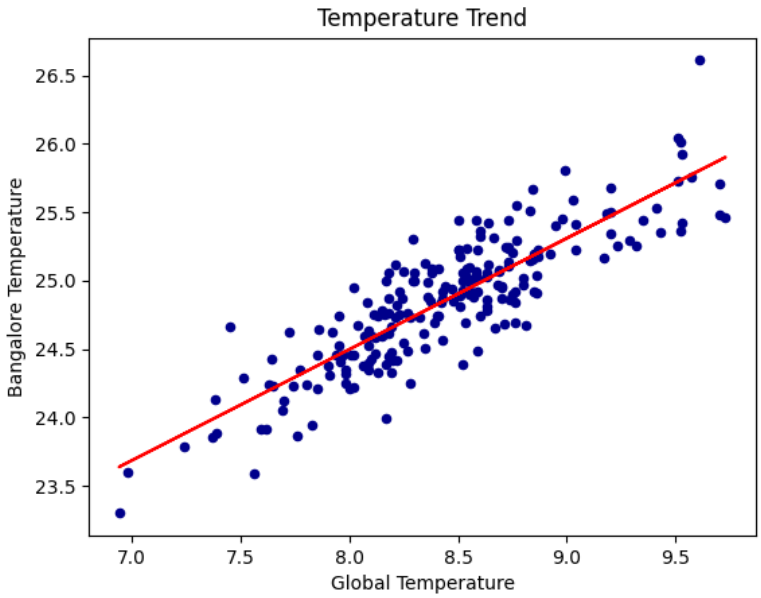

Figure 5: Linear regressor

The above linear regression model tries to predict the local (Bangalore) average temperature when the global average temperature is provided.

Conclusion

The following are the conclusions drawn from the aforementioned data:

- Bangalore is, on an average, 17°C hotter compared to the global average. This difference has been more or less constant over the years.
- There's and interesting trend that can be observed between the local and global temperatures:
  - The global and local average rise in temperatures between 1796-1896 and 1897-1996 was around 0.55°C
  - The global rise in temperature in the last 20 years = 0.88°C while the local rise in temperature in the last 20 years = 0.600°C.
    This means, the world average is increasing faster than Bangalore's average. (No wonder Bangalore is called *'The Garden City'*)
- The correlation coefficient = 0.86. This means that both the global and local temperatures are closely linked and the increase/decrease in one of them affects the other.
- As per the linear regression model, as the world average hits 11°C, Bangalore's average will cross 27°C. This could happen in between 2020-2025.