Exploring Weather Trends

By Vikas Kumar Choudhary

Introduction

In this project we will analyze the local and global temperature data and compare the trends of Delhi, India temperature and Global Temperature. We will try to find some insights like :

- Trends in Delhi and Global Temperature
- Is there any relation between Delhi and Global Temperature
- which has more variability in temperature and many more...

Dataset Structure

Weather Data : This dataset is provided by udacity for the analysis. This dataset is extracted from udacity sql workspace using the following query:

```
select c.year, c.avg_temp as city_temp, g.avg_temp as global_temp
from city_data c
join global_data g
on c.year=g.year
and c.city='Delhi'
order by 1;
```

This dataset has 3 columns

- `year` : Year in which temperature is recorded
- `city_temp` : Temp of Delhi City in the given year
- `global_temp` : Global Temp in the given year

There is one data quality issue: city data has some empty fields

Steps performed to prepare the data for analysis

1. Data is extracted from udacity sql workspace using sql query.
2. Data is loaded in Excel
3. Calculated the average of city_temp and fill all the null cells in city_temp with that averag value.
4. Then calculated the moving average for both city data and global data.
5. Then created a line plot between city_temp moving average and global_temp moving average.

Calculation of moving average :

- Go down till the 5th row of city_temp and use the AVERAGE() function to calculate the average temperature for the first years in a new column.
- Then drag the same formula till the end of data.
- Repeat the above 2 steps and finally we have running average for city and global data.

Key considerations to decide how to visualize the trends:

- As we have 2 variables to plot so I decided to go with Bivariate plots, so I choose to plot `Line Plot` and `Scatter Plot`
- For line plot I choose to go with running averages as it will smooth out data to make it easier to observe long term trends and not get lost in yearly fluctuations.

Analysis

Line Plot

Temp Trend for Delhi and Global

Scatter Plot

Delhi Temperature vs Global Temperature
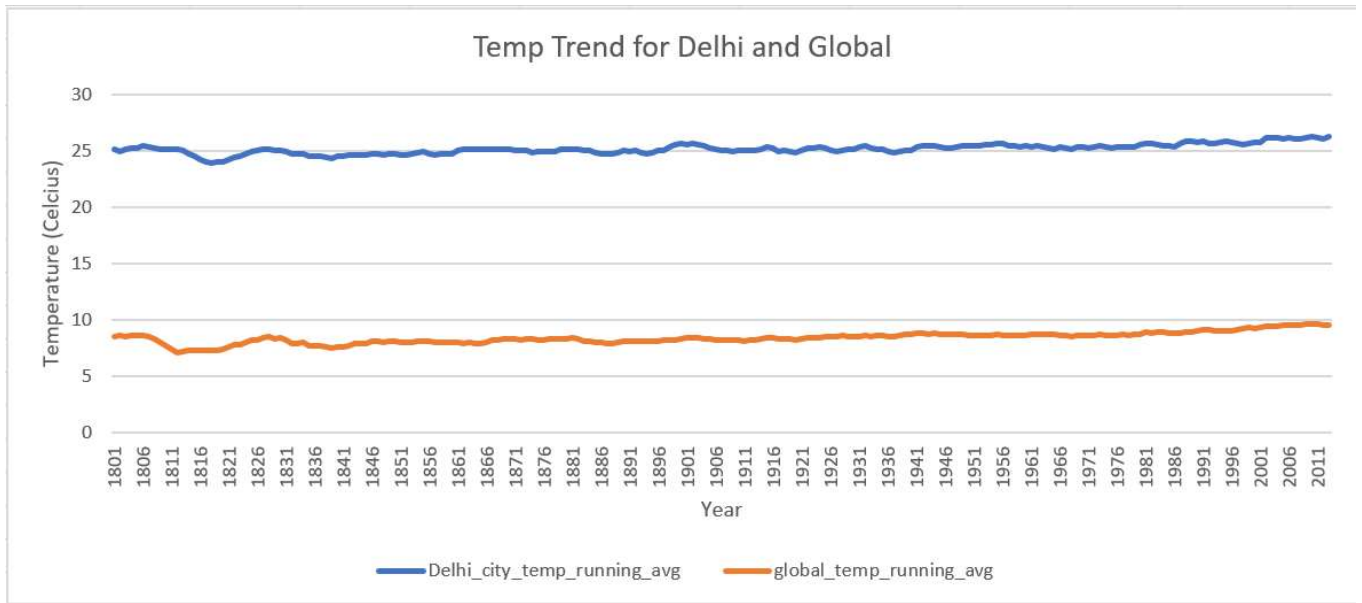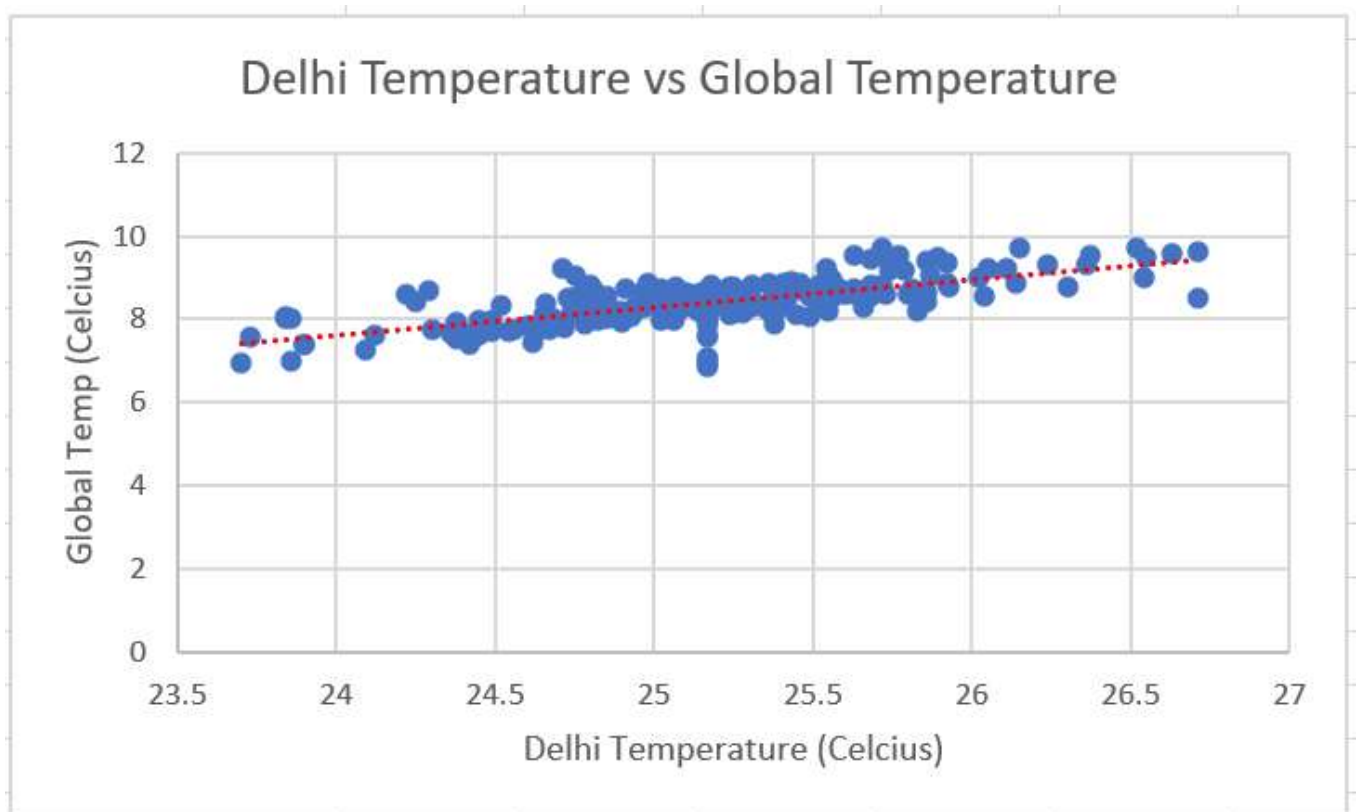
Statistical Analysis

- `Correlation Coefficient` between city data and global data : **0.870898275**
- `Standard Deviation` of city data : **0.433327217**
- `Standard Deviation` of global data : **0.502285232**

Insights

Insight 1

- Delhi Data is highly correlated with Global Data

Insight 2

- Delhi is hotter compared to Global Temperature

Insight 3

- Both City and Global weather shows increasing trend in last few years

Insight 4

- The world is getting hotter year by year

Insight 5

- Global Weather has more variability compared to city

Insight 6

- City Temp also goes down when global temp goes down

Insight 7

- Delhi City gets more hotter compared globally as difference in max and min temp for city is 3.01 degree celcius while for global the difference is 2.87 only.

Conclusion

Both the city data and global data are highly correlated to each other. And the trend shows that the world is getting hotter year by year. There can be many reasons for this like

- Pollution
- Global Warming etc.

The world need to take action otherwise it will be unbearable to live on earth after few years if the same trend goes on in future too...