UDACITY
Data Analyst Nanodegree Program

PROJECT DETAILS: **Exploring Weather Trends**

STEPS FOR PREPARATION OF DATA:

Tool used – **SQL** (Structured Query Language)

- ➢ **STEP1**: I am from INDIA, I wanted to choose New Delhi for analysis.

    To see which cities are available for "INDIA" in the given dataset

```
SELECT *

FROM city_list

WHERE country LIKE 'India'
```

- ➢ **STEP 2**: I found from the SCHEMA that city_data and global_data both have same column named 'Avg_temp', so to rename it used following sql query-

```
ALTER TABLE city_data RENAME COLUMN avg_temp to CAT;

ALTER TABLE global_data RENAME COLUMN avg_temp to GAT;
```

- ➢ **STEP 3**: Extraction of Data

```
SELECT global_data.year, global_data.GAT, city_data.CAT

FROM global_data JOIN city_data        -- Joining the two tables

ON global_data.year = city_data.year    -- reference for joining

WHERE city LIKE 'New Delhi';
```

- ➢ **STEP 4**: Downloading the extracted Data and saving it as a csv file.

delhi.csv

<u>STEPS FOR DATA MANIPULATION:</u> CALCULATION OF MOVING AVERAGE

Tool Used : Python, Pandas and Numpy libraries

➢ **STEP 1:** Importing important libraries

```python
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt        -for plotting the chart
```

➢ **STEP 2:** Importing the extracted dataset

```
csv_path = 'D:/bertelsmann udacity/delhi.csv'    -directory where file is saved

df=pd.read_csv(csv_path)                          -creating a data_frame
```

➢ **STEP 3:** Calculating the Moving Average

```python
def moving_avg(mA_range, data_input):

        output = data_input.rolling(window = mA_range, on ="cat").mean()

        return output

        chart_moving_avg = moving_avg(20, data)        -function calling
```

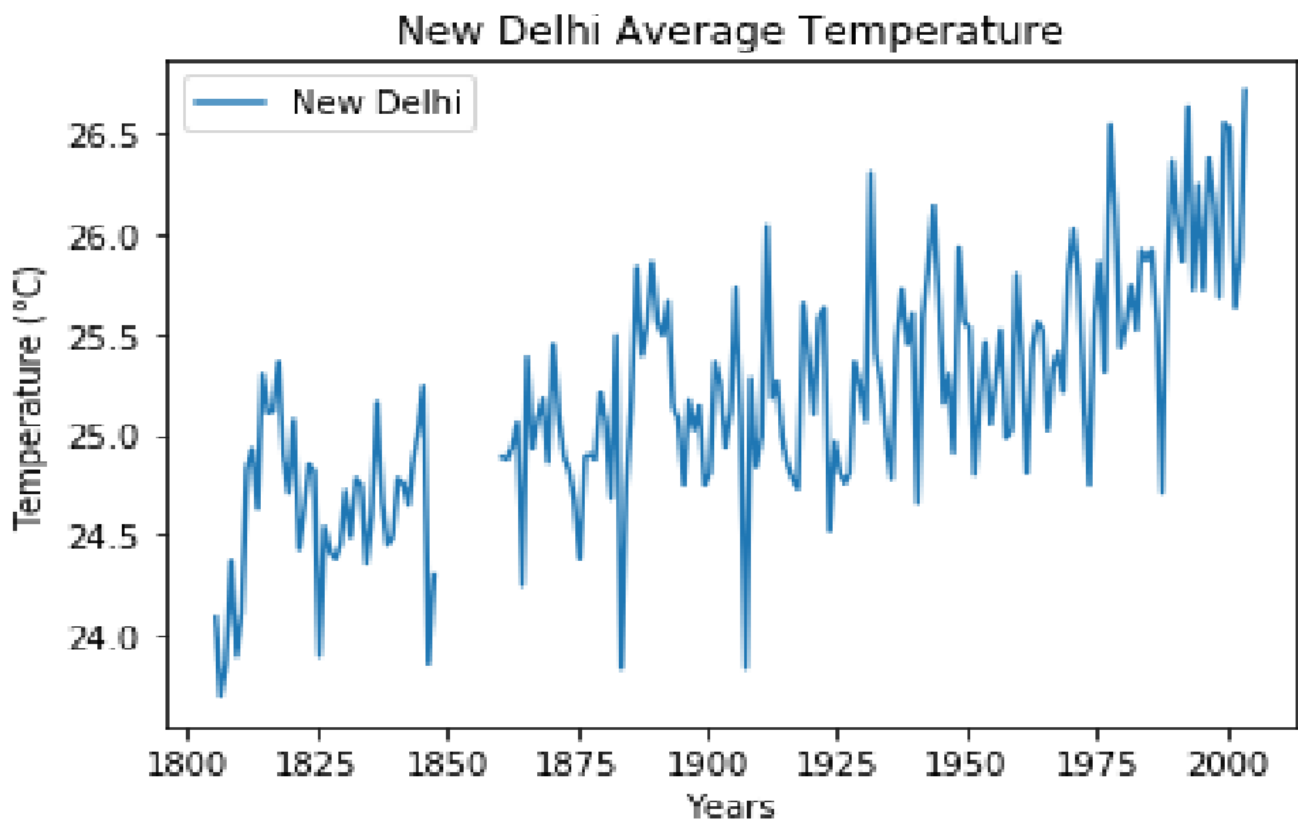$$\bar{a}_{SM} = \frac{x_n + x_{n-1} + \ldots + x_{M-(n-1)}}{M}$$

$$\bar{a}_{SM} = \frac{1}{M} \sum_{i=0}^{n-1} x_{M-i}$$

STEPS FOR DATA VISUALIZATION:
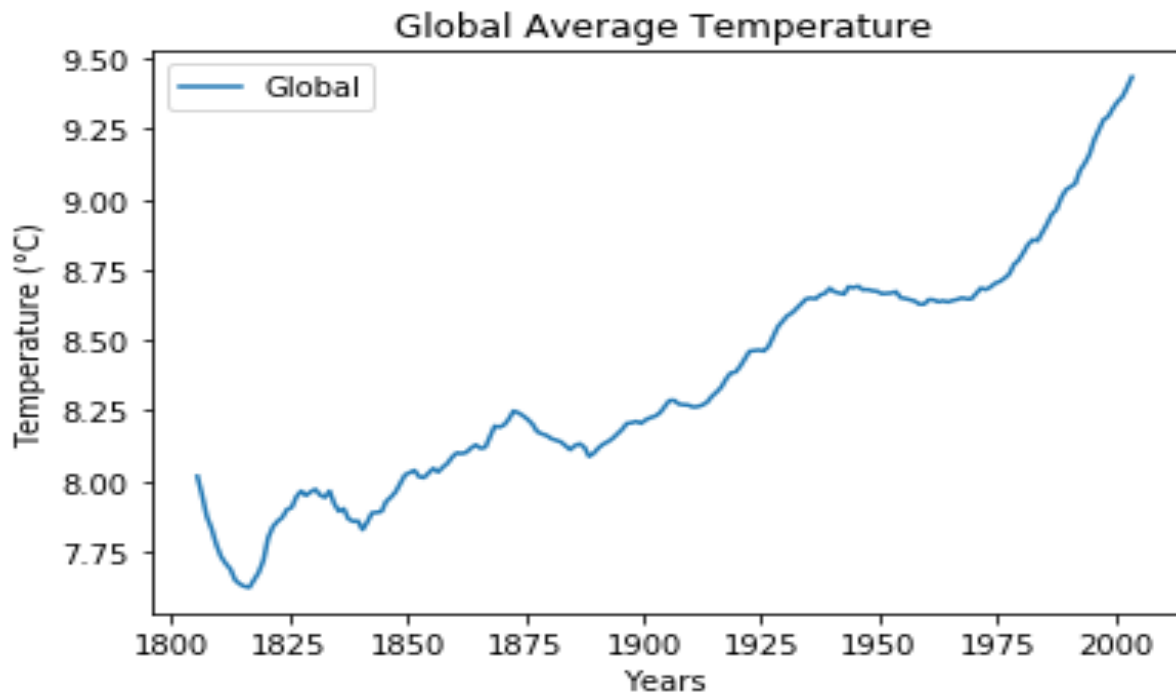
Tool Used: Python, Matplotlib library

➤ **STEP 1:** Plotting the graph of New Delhi's average temperature

```
plt.plot(chart_moving_avg['year'], chart_moving_avg['cat'], label='New Delhi')

plt.legend()

plt.xlabel("Years")

plt.ylabel("Temperature (°C)")

plt.title("New Delhi Average Temperature")

plt.show()
```
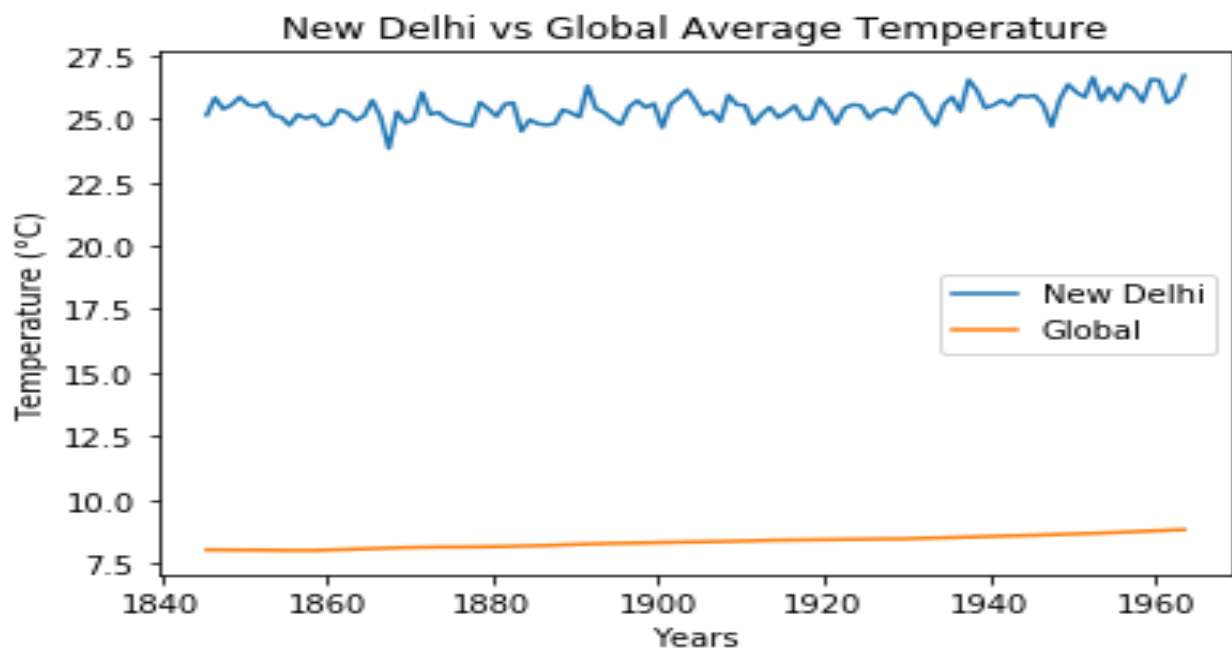

➤ **STEP 2:** Plotting the graph of Global average temperature

```
plt.plot(chart_moving_avg['year'], chart_moving_avg['gat'], label='Global')

plt.legend()

plt.xlabel("Years")

plt.ylabel("Temperature (°C)")

plt.title("Global Average Temperature")

plt.show()
```

Global Average Temperature

> **STEP 3:** Plotting the graph of New Delhi vs Global average temperature

```
plt.plot(chart_moving_avg['year'], chart_moving_avg['cat'], label='New Delhi')

plt.plot(chart_moving_avg['year'], chart_moving_avg['gat'], label='Global')

plt.legend()

plt.xlabel("Years")

plt.ylabel("Temperature (°C)")

plt.title("New Delhi vs Global Average Temperature")

plt.show()
```


New Delhi vs Global Average Temperature

DATA INTERPRETATION

The following is the key observations made on the basis of the chart made by plotting the Moving Averages of the Temperatures recorded in the selected city – New Delhi and the global average temperature over the period – (1796 – 2013 years)

1. The selected city – New Delhi is pretty much hotter than the rest of the globe, having an average temperature of 25.16627*C. While the Global average temperature being 8.403532*C. With a difference of 16.7672*C.

2. It is observed that the difference in temperatures between New Delhi and the rest of the world has been consistent over the past 100 years (1900-2000).

3.
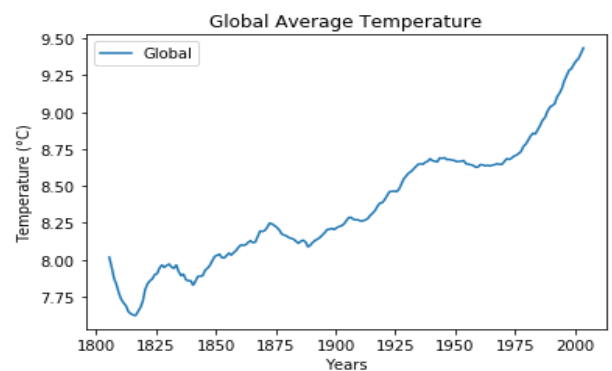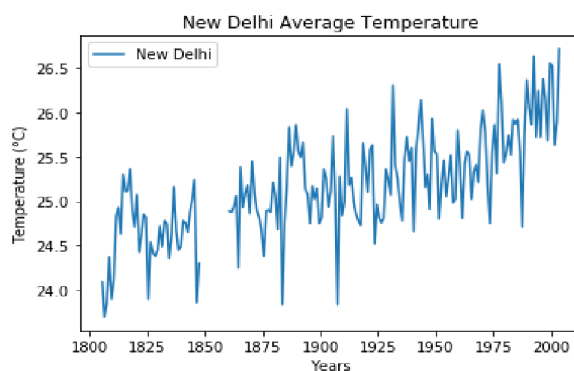
- (1808-1812 & 1858-1869) missing data of the city's average temperature
- 1835-city's temp fall by 0.92 and globally it falls by 0.76 with respect to the previous year's recorded temperature
- 1875- there is a global decrement in temperature by 0.57*C but on the other hand New Delhi's temperature increases by 1.13*C.
- In the early years(1800-1900) the trend of weather could not be exactly determined whether the world is becoming hotter or cooler due to in-consistent weather conditions.
- In the latter years(1950-2013) the global temperature has risen from (8.37 – 9.61) accounting to a total of 1.24*C , Whereas New Delhi's has risen from (24.66 - 26.71) that is about 2.05*C. Hence there exist a trend which indicates that the world is getting hotter and New Delhi is getting even more hotter.
- In the latter years(1950-2013) the temperature has risen exponentially both globaly and locally, this mainly is due to the industrial revolution that occurred in Europe, North America and Asia. It changed the agriculture based economies to large-scale machine based economy. Hence running large machine to meet the demands the industries produced more and more heat, eventually increasing the temperature both locally and globally.

- The correlation co-efficient of the global and local average temperature is nearly 0.762654 , which indicates that with every positive increase in global temperature there is an increase in the local temperature of New Delhi.
- The local temperature of New Delhi is on average 16.7672*C greater than the global temperature. So an estimate can be made looking on the trend that for the past 100years the local as well as global temperature is rising consistently.

- Selecting 'New York' as local city: -
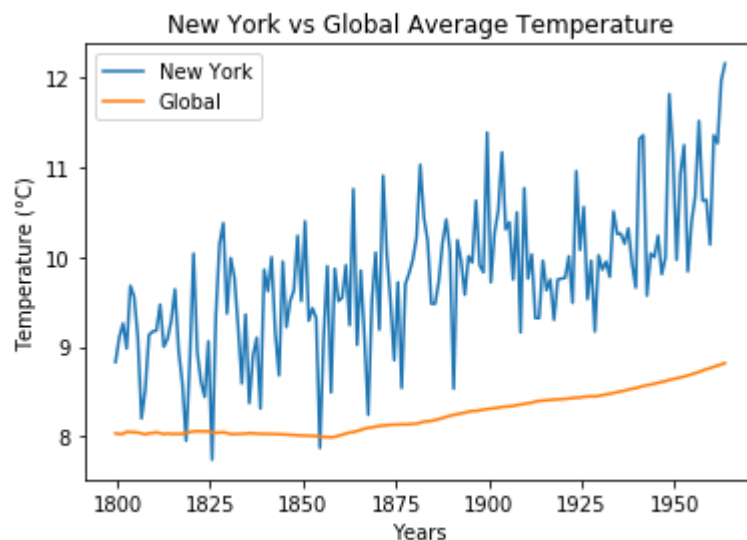
New York city has an average temperature slightly greater than the global average temperature. The temperature has risen from 9.5*C to 12*C in the years(1950-2013), whereas the overall global temperature has risen from 8.5*C to 8.9*C. This indicates that New York city is getting much hotter in compared to the rest of the world.

- Selecting 'Miami' as local city: -
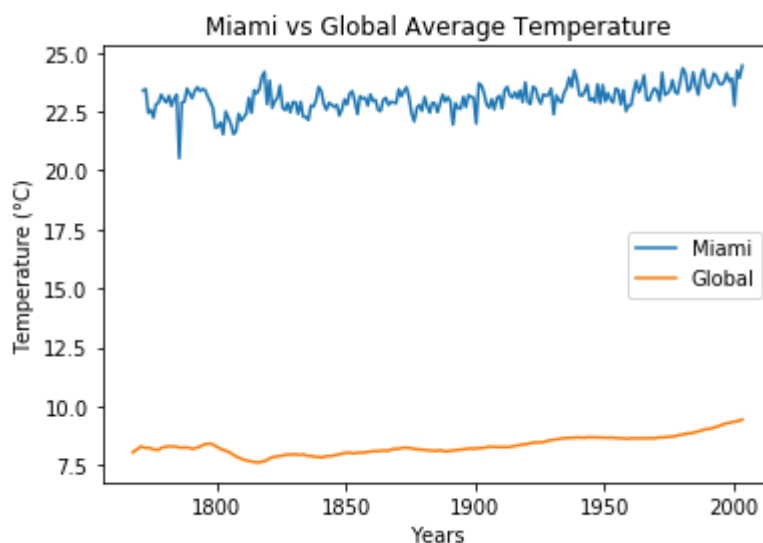
The average temperature of Miami is quite high in comparison to the global average temperature. It is consistent with the global temperature rise.