Exploring Weather Trends

APRIL 28

UDACITY - Data Analyst Nanodegree
Project by: Sarthak Sahu
Bangalore, India

Introduction:

In this project, I have analyzed the temperature of Bangalore and Global temperature and compared them. I have extracted data from Udacity database and manipulated according to my goals.

Goals:

1. Extract data from the Udacity database (.csv file)
2. Create a chart Visualization (Line Chart)
3. Observation based on chart

Tools Used:

- SQL: To Extract data from database
- Microsoft Excel: To view extracted data
- Python: To Manipulate Data
- Anaconda – Jupyter Notebook: To Implement the Manipulated Data (Python Code)

STEP – 1 Extraction of Data from Database:

The following steps are taken for the extraction of data from database, the tool used here is SQL.

1. Find the cities available for 'INDIA' in the database:

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

2. I got the list and I've chosen 'Bangalore' for this project; I have provided with the SCHEMA of the tables. The table of 'city' and 'global' have same column name (i.e. 'avg_temp').
So, I renamed the column and join the table Using 'JOIN' function().

```sql
SELECT c.year,c.avg_temp AS city_avg_temp,g.avg_temp AS
global_avg_temp
FROM city_data c
JOIN global_data g ON c.year = g.year
WHERE c.city = 'Bangalore'
AND5c.avg_temp IS NOT NULL;
```

3. After getting the Output, I have downloaded the data as "results.csv".

STEP – 2 Visualization of Data:

The following steps are taken for the Visualization of Data, the tool used here is Python and Jupyter Notebook.

1. # Importing Libraries
   ```python
   import matplotlib.pyplot as plt    # For Plotting Line Chart
   import pandas as pd                # For Importing the .csv files
   ```

2. # Importing Datasets
   ```python
   dataset = pd.read_csv('results.csv')
   ```

3. For Calculating Moving Average and Rolling function, I have taken reference which I've mentioned below.
   ```python
   # Calculating Moving Average
   moving_avg_global = dataset.rolling(window = 150, center = False, on =
   'global_avg_temp').mean()                    # global moving avg
   moving_avg_city = dataset.rolling(window = 150,center = False,on = 'city_avg_temp').mean()
                                                # city moving avg
   ```

4. # Visualizing the Plot of Global and City Temperature
   ```python
   plt.figure(figsize =(10,5))
   plt.grid(True)
   plt.plot(moving_avg_global['year'],moving_avg_global['global_avg_temp'],label='Global')
   plt.plot(moving_avg_city['year'],moving_avg_city['city_avg_temp'],label='Bangalore')
   plt.legend()
   plt.xlabel('Years')
   plt.ylabel('Temperature(°C)')
   plt.show()
   ```
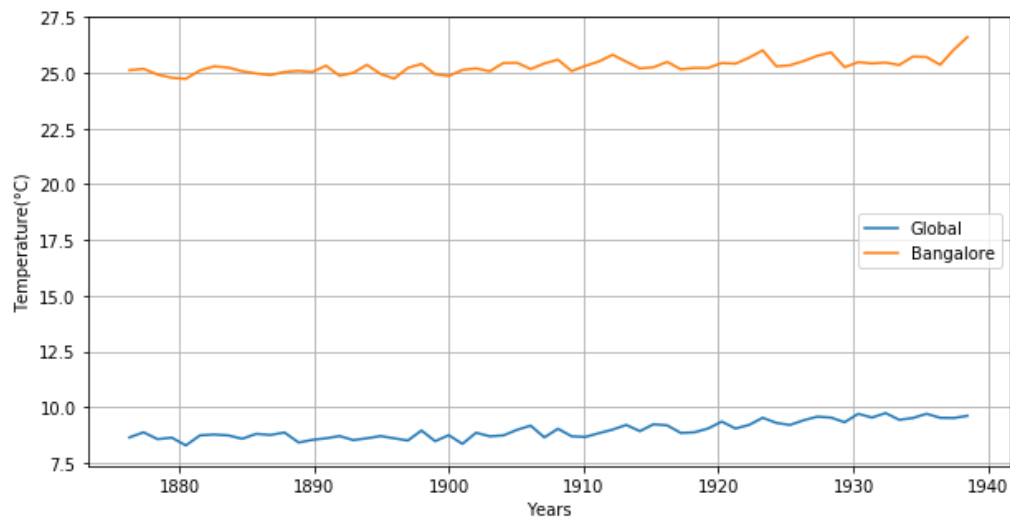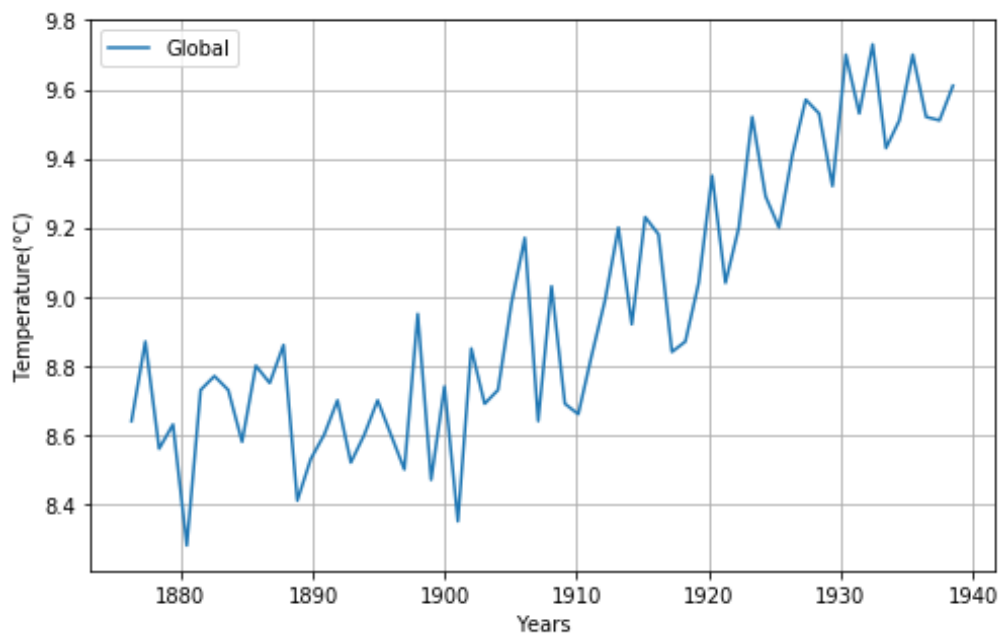
5. # Visualizing The Plot of Global Temperature
   plt.figure(figsize =(8,5))
   plt.grid(True)
   plt.plot(moving_avg_global['year'],moving_avg_global['global_avg_temp'],label='Global')
   plt.legend()
   plt.xlabel('Years')
   plt.ylabel('Temperature(°C)')
   plt.show()

6. # Visualizing the Plot of City Temperature
    ```python
    plt.figure(figsize =(8,5))
    plt.grid(True)
    plt.plot(moving_avg_city['year'],moving_avg_city['city_avg_temp'],label='Bangalore')
    plt.legend()
    plt.xlabel('Years')
    plt.ylabel('Temperature(°C)')
    plt.show()
    ```
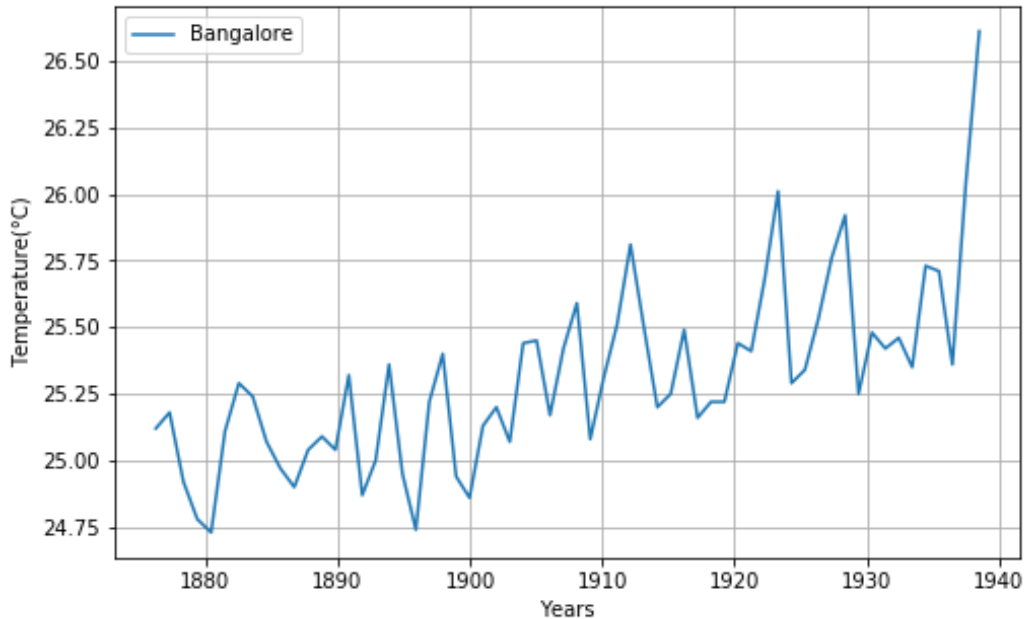

Analysis and Conclusion:

1. The Global average temperature varies from 8.4 °C to 9.6 °C While Bangalore average temperature varies from 24.98 °C to 25.57 °C which is a big difference.

2. From the chart, I have also Observed a very big difference between the average temperature of Bangalore and that of World.

3. I also notice that the global temperature is rising 0.1 °C (approx.) every year which is constant.

4. Due to Industrialization, the global temperature has started increasing at a higher rate.

5. As Bangalore is found near the Equator, Hence the Global average temperature is very less as compared to equatorial region.

Considerations:

- X -Axis: Years
- Y-Axis: Temperature (°C)
- Different color lines for global and city average temperature.
- Used matplotlib.pyplot for plotting line chart.
- Used pandas for importing the dataset.
- All the Codes are saved in Jupyter Notebook (.ipynb file) for reference and improvements.

Reference:

1. Cover Image: http://500px.com/photo/32208933
2. SQL Query: Udacity Chat Hub
3. Moving Average: https://www.learndatasci.com/tutorials/python-finance-part-3-moving-average-trading-strategy/
4. Rolling Function Parameters: https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.rolling.html