



# Udacity Data Analyst Project 01:

## Exploring Weather Trends

Name: Niladri Ghosh

Date: 20<sup>th</sup> May, 2019

Country: India

### ABSTRACT:

SQL Query was used to download (CSV) file that contains yearly average temperature of the cities 'Ranchi', 'Patna', 'Bangalore' and the global temperature.

Later the data has been analyzed using Python Programming Language using Jupyter Notebook Environment.

### INTRODUCTION:

In this project, I will be analyzing the three local cities and the global temperature data and compare the temperature trends for my own city where I live: 'Ranchi' in addition I am going to add two cities according to their air pollution ratings of 2018 - 'Patna' being the highest polluted whereas 'Bangalore' as the least one, to the overall global temperature trends. It would be fun to check out weather the pollution contributes to the rise of average temperature of a city.

Aim of this project is to create a visualization and describe the changing temperature of each cities with respect to the passing time. Let's see if the graph deviates from the least to the highest polluted cities in India. At last I'll be preparing a brief report weather we are facing a global temperature crisis.

Cities which I chose:

1. Ranchi (own city) - 142 (PM10)
2. Patna (highly polluted) - 119.7 (PM2.5)
3. Bangalore (least polluted) - 71 mg/cubic meter (RSPM)

# PROGRESS OUTLINE:

## 1. SQL Query was used to download four CSV files that contains:

- Global temperature data such as ("Year", "Average Temperature") has been imported by using the query below:

```
SELECT * FROM global_data;
```

- "Ranchi" (my own city) data such as ("Year", "Average Temperature") and has been imported by using the query below:

```
SELECT * FROM city_data  
WHERE city = 'Ranchi' AND country = 'India';
```

- "Patna" (most polluted city) data such as ("Year", "Average Temperature") has been imported by using the query below:

```
SELECT * FROM city_data  
WHERE city = 'Patna' AND country = 'India';
```

- "Bangalore" [now Bengaluru] (least polluted city) data such as ("Year", "Average Temperature") has been imported by using the query below:

```
SELECT * FROM city_data  
WHERE city = 'Bangalore' AND country = 'India';
```

## 2. Importing necessary libraries in python.

```
import pandas as pd # for working with data!  
import matplotlib.pyplot as plt # for visualizing the data!  
import numpy as np # for calculating the moving average!
```

## 3. Importing the data.

```
global_temp = pd.read_csv('global_data.csv') # importing 'global temperature data'  
ranchi_temp = pd.read_csv('ranchi_data.csv') # importing '~Ranchi temperature data' : data for Ranchi over multiple year  
patna_temp = pd.read_csv('patna_data.csv') # importing '~Patna temperature data' : data for Patna over multiple year  
bangalore_temp = pd.read_csv('bangalore_data.csv') # importing '~Bangalore temperature data' : data for Bangalore over mu
```

#### 4. Moving Averages for the Global data and Local data ('Ranchi', 'Patna' & 'Bangalore') to smoothen out charts while plotting.

- Rolling Average has been calculated to smooth out data and to make it easier to observe the trends when it be shown in Charts.
- The Rolling Average has been calculated for every 10 years to each single data but the first 10 can't be predicted due to lack of data i.e. lack of data for moving averages.
- Python was used for calculating the Moving Average Using Built-in Functions such as ("rolling", "mean").
- Python Code provided below:

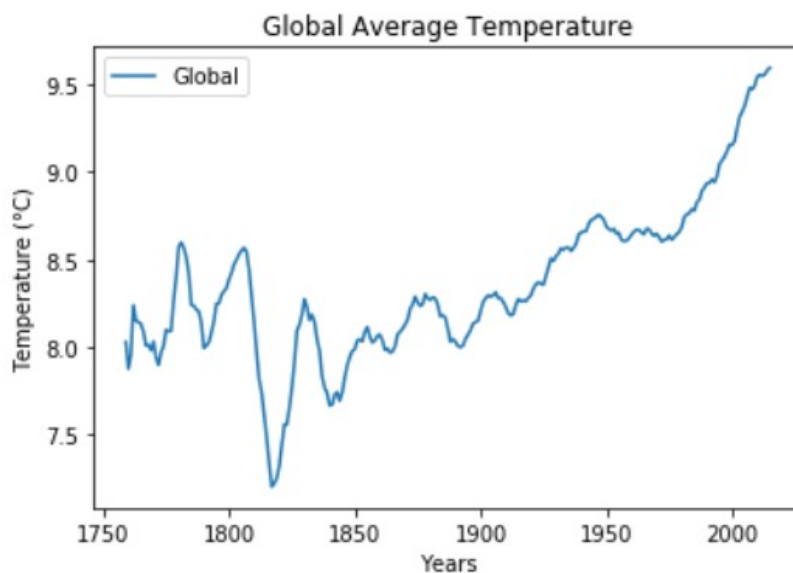
```
global_mv_avg = global_temp['avg_temp'].rolling(10).mean()
ranchi_mv_avg = ranchi_temp['avg_temp'].rolling(10).mean()
patna_mv_avg = patna_temp['avg_temp'].rolling(10).mean()
bangalore_mv_avg = bangalore_temp['avg_temp'].rolling(10).mean()
```

#### 5. Line Chart for data.

- **Individual Plotting with code**

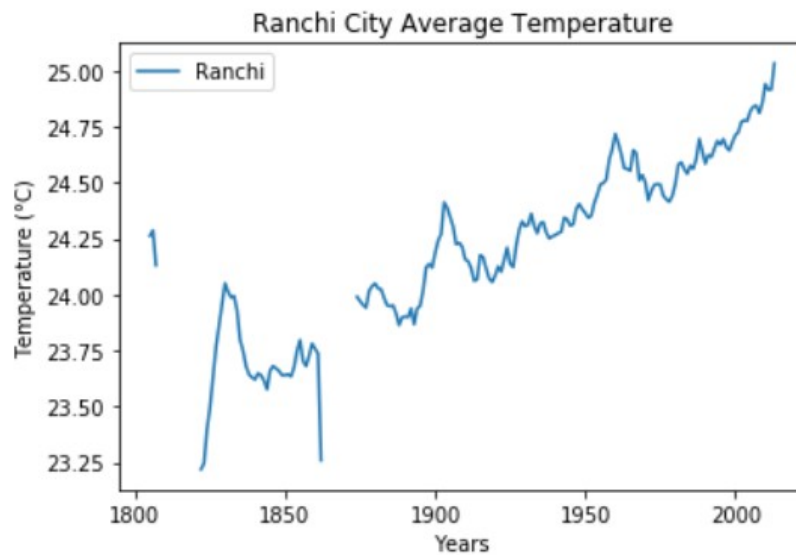
GLOBAL AVERAGE TEMPERATURE

```
#Global Average Temperature
plt.plot(global_temp['year'], global_mv_avg, label='Global')
plt.legend()
plt.xlabel("Years")
plt.ylabel("Temperature (°C)")
plt.title("Global Average Temperature")
plt.show()
```



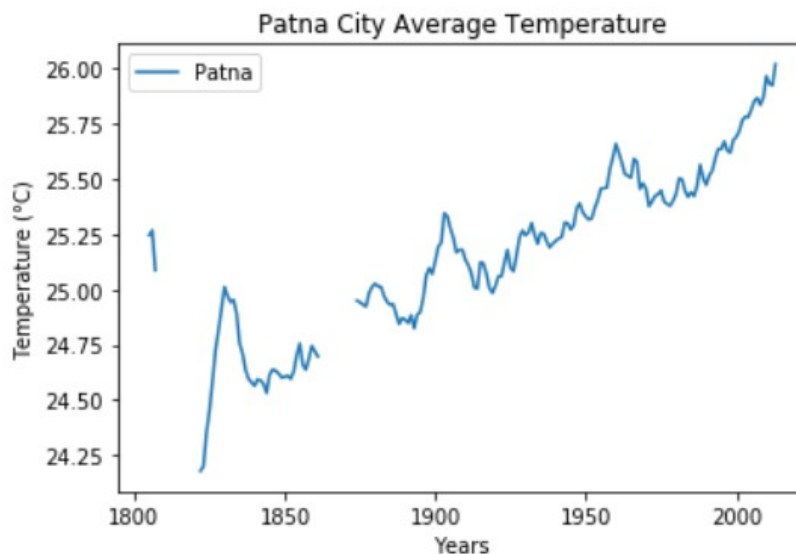
## RANCHI AVERAGE TEMPERATURE

```
#Ranchi Average Temperature  
plt.plot(ranchi_temp['year'], ranchi_mv_avg, label='Ranchi')  
plt.legend()  
plt.xlabel("Years")  
plt.ylabel("Temperature (°C)")  
plt.title("Ranchi City Average Temperature")  
plt.show()
```



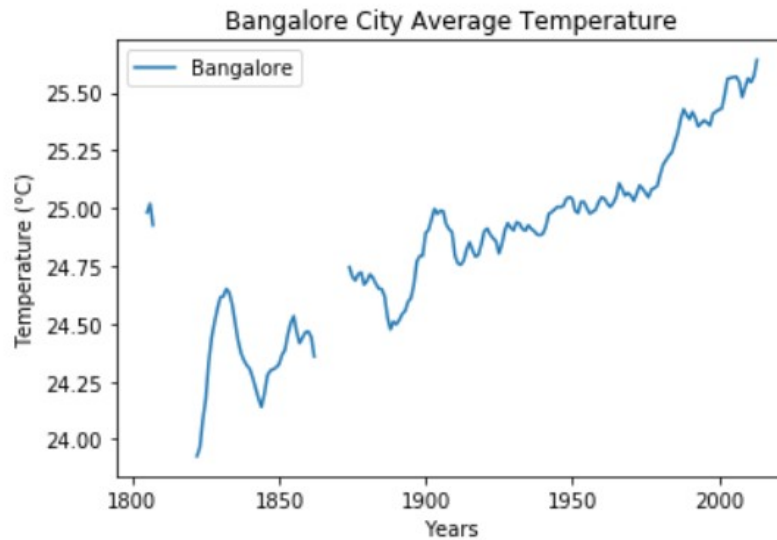
## PATNA AVERAGE TEMPERATURE

```
#Patna Average Temperature  
plt.plot(patna_temp['year'], patna_mv_avg, label='Patna')  
plt.legend()  
plt.xlabel("Years")  
plt.ylabel("Temperature (°C)")  
plt.title("Patna City Average Temperature")  
plt.show()
```



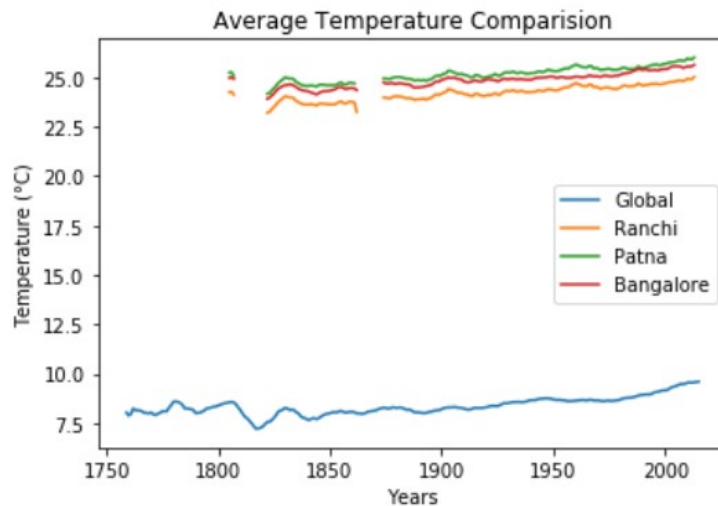
## BANGALORE AVERAGE TEMPERATURE

```
#Bangalore Average Temperature
plt.plot(bangalore_temp['year'],bangalore_mv_avg,label='Bangalore')
plt.legend()
plt.xlabel("Years")
plt.ylabel("Temperature (°C)")
plt.title("Bangalore City Average Temperature")
plt.show()
```



- Comparison of each city's temperature with global temperature data

```
#Comparision
plt.plot(global_temp['year'],global_mv_avg,label='Global')
plt.plot(ranchi_temp['year'],ranchi_mv_avg,label='Ranchi')
plt.plot(patna_temp['year'],patna_mv_avg,label='Patna')
plt.plot(bangalore_temp['year'],bangalore_mv_avg,label='Bangalore')
plt.legend()
plt.xlabel("Years")
plt.ylabel("Temperature (°C)")
plt.title("Average Temperature Comparision")
plt.show()
```



## OBSERVATIONS:

- Thoroughly going through each and every line we can definitely conclude that the average temperature is rising over the century due to the climate change.
- The average temperature all over the world has started rising exponentially due the rapid growth in industrialization.
- The global temperature dropped around 1820 since then it has been gradually increasing ever since.
- The difference in temperature for globe between 1820 and 2013 is more than 2° C.
- Few of the data are missing for country India during 1810 to 1820 and 1860 to 1870.
- The difference in temperature for Ranchi between 1820 to 2013 is around 1.7° C.
- The difference in temperature for Patna between 1820 to 2013 is around 1.9° C.
- The difference in temperature for Bangalore between 1820 to 2013 is around 1.7° C.
- The increase in temperature for Patna (highly polluted) is around 1.9° C where as in Bangalore and Ranchi (comparably less polluted) is around 1.7° C.
- The difference in rate of increase of temperature is almost 0.2° C which in fact is a towering difference.
- The globe with all the cities is getting warmed up day by day.

## CONCLUSIONS:

Hence from the observations above, the global temperature is raising over the years which support the case of climate change. As mentioned earlier the level of pollution in an area effects the average temperature of that city can be proved by the difference in the increase in temperature of cities Patna and Ranchi, Bangalore which is around 0.2° C which justifies our statement.

Climate change is the environmental challenge of this generation, and it is imperative that we act before it's too late.

John Delaney