

Name: Rishu Singh

Date: 21st May 2019

Place: New Delhi, INDIA

UDACITY

Data Analysis Nanodegree

Project 01:



Exploring weather Trends

Overview:

In this project, I have analyzed local temperature of New Delhi, India in accordance with the global temperature data and compared. I had been provided with a database on Udacity portal from where I have to extract, manipulate and visualize the data as in the following goals.

Goals:

1. Extraction of data from the database and export to CSV file
2. Making a line chart visualization based on extracted data
3. Observation based on chart

Tools Used:

1. SQL: To extract the data from the database.
2. Python: For calculating moving average and plotting line chart.
3. ANACONDA - Jupyter Notebook: For writing python code and making observations.
4. Excel: Having a look at the data and writing project.

STEP 1 -

Extraction of Data from provided Database I have done the following activity in order to make a relevant dataset. I have learnt the SQL basics from lessons provided before this project.

I have also done an introductory course on SQL and relational database from which I have used some concepts.

1. To see which cities are available for "India" in the given dataset:

```
SELECT * FROM city_list WHERE country LIKE 'India'
```

2. I know that I can make a relevant dataset by joining the two tables. But, I found from the SCHEMA that both city_data and global_data contains same column named 'avg_temp'.

So I have changed the names of the columns respectively in order to have distinct columns.

```
ALTER TABLE city_data RENAME COLUMN avg_temp to CAT;  
-- CAT = City Average Temp.
```

```
ALTER TABLE global_data RENAME COLUMN avg_temp to GAT;  
-- GAT = Global Average Temp.
```

3. Now I have written following code in order to join the two tables and have the relevant data:

```
SELECT global_data.year, global_data.GAT, city_data.CAT  
FROM global_data JOIN city_data  
ON global_data.year = city_data.year  
WHERE city LIKE 'New Delhi';
```

Now I have got an option of downloading the file as CSV format.

Moving Averages:

- To observe the trends in temperature I calculated moving average(MA).
- I used 10 years Moving Average to get the smooth line chart.

Excel commands for Moving Averages:

| Moving Average | Excel Commands |
|-----------------------------|------------------|
| For 10 years Moving Average | =AVERAGE(B2:B11) |

This is how the excel sheet looks like:

| | A | B | C | D | E | F |
|----|------|------|--------|-------|----------|---|
| 1 | year | gat | mA_gat | cat | mA_cat | |
| 2 | 1796 | 8.27 | | 25.03 | | |
| 3 | 1797 | 8.51 | | 26.71 | | |
| 4 | 1798 | 8.67 | | 24.29 | | |
| 5 | 1799 | 8.51 | | 25.28 | | |
| 6 | 1800 | 8.48 | | 25.21 | | |
| 7 | 1801 | 8.59 | | 24.22 | | |
| 8 | 1802 | 8.58 | | 25.63 | | |
| 9 | 1803 | 8.5 | | 25.38 | | |
| 10 | 1804 | 8.84 | | 25.68 | | |
| 11 | 1805 | 8.56 | 8.551 | 25.3 | 25.273 | |
| 12 | 1806 | 8.43 | 8.567 | 25.22 | 25.292 | |
| 13 | 1807 | 8.28 | 8.544 | 24.97 | 25.118 | |
| 14 | 1808 | 7.63 | 8.44 | | 25.21 | |
| 15 | 1809 | 7.08 | 8.297 | | 25.20125 | |
| 16 | 1810 | 6.92 | 8.141 | | 25.2 | |
| 17 | 1811 | 6.86 | 7.968 | | 25.36333 | |
| 18 | 1812 | 7.05 | 7.815 | | 25.31 | |
| 19 | 1813 | 7.74 | 7.739 | 24.56 | 25.146 | |
| 20 | 1814 | 7.59 | 7.614 | 23.73 | 24.756 | |
| 21 | 1815 | 7.24 | 7.482 | 24.09 | 24.514 | |
| 22 | 1816 | 6.94 | 7.333 | 23.7 | 24.21 | |

STEP 2 – Python Code for Making Line Chart

So I have used some python libraries here, I have written these codes on Jupyter Notebook.

Importing the important Libraries

```
import numpy as np
```

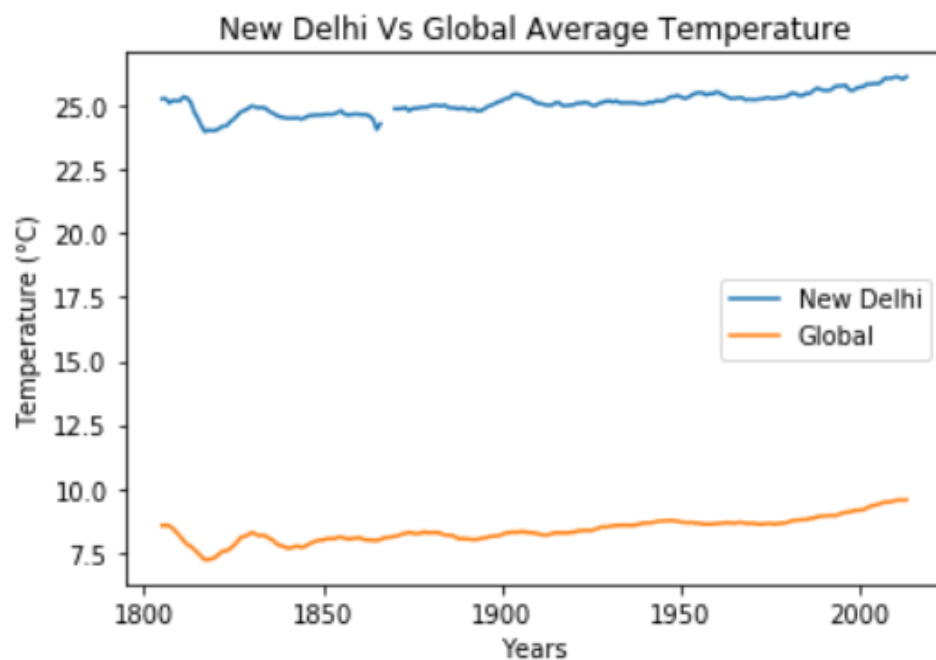
```
import pandas as pd
```

```
from matplotlib import pyplot as plt
```

Importing the extracted Data Set

```
data = pd.read_csv("results.csv")
```

Line Chart for New Delhi and Global Temperature:



Observation:

- **Global** average temperature varies between **7.20 to 9.56** Degree Celsius but **New Delhi** city average temperature is varies between **24.03 to 26.14** Degree Celsius.
- If comparing the Global average temperature and New Delhi average temperature, then the **New Delhi city is hotter** than Global average temperature.
- Change in temperature over time:

| Year | Change in Global average temperature | Change in New Delhi average temperature | Increasing/Decreasing over time |
|-------------|--------------------------------------|-----------------------------------------|---------------------------------|
| 1796 - 1866 | 8.27 – 8.04 | 25.03 – 24.30 | Decreasing |
| 1880 - 1950 | 8.27 – 8.68 | 25.01 – 25.36 | Increasing |
| 1951 - 2013 | 8.67 – 9.56 | 25.28 – 26.14 | Increasing |

- According to the graph and above table the difference between Global average temperature and New Delhi average temperature is been **consistent over time**.
- New Delhi and Global average temperature have similar kind of trends. During early years, both trends seems to have ups and downs then approx. around **1982** the moving average temperature starts to **increase** at a steady rate.
- According to the graph the **world is getting hotter** because from 1834 to 2013 Temperature increases.

Final Conclusion:

Global Temperature \propto New Delhi Temperature

The World is Getting Hotter.