Explore Weather Trends

18 March 2020

Udacity Data Analyst Nanodegree

Project-1

Tanmay Jain

New Delhi,

India

Overview

In the project, I have analyzed the local temperature of my city, New Delhi and compared it with the global temperature with the help of the data provided by Udacity.

Tools Used

- Python
- SQL
- Jupyter Notebook
- Google Docs

Steps Taken

1. An initial query to find my city in the city_list table.

```
SELECT * FROM city_list
WHERE country='India'
AND city LIKE '%elhi'
```

2. Extract city_data for my city.

```
SELECT * FROM city_data
WHERE city = 'New Delhi'
```

3. Alter the city data table and global data table to change their avg_temp columns to CAT (City Average Temperature) and GAT (Global Average Temperature) for ease of use.

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

4. Create another table that consists of columns necessary for comparing temperatures and download it in .csv format using the given button.

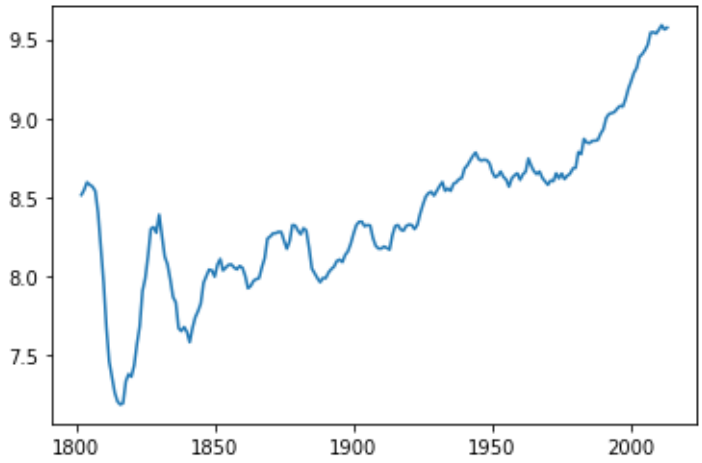
```
SELECT g.year, g.gat, c.cat FROM city_data c JOIN
 global_data g
ON c.year=g.year
WHERE c.city='New Delhi'
```

5. Reading the resultant csv in Jupyter Notebook using the Pandas Library.

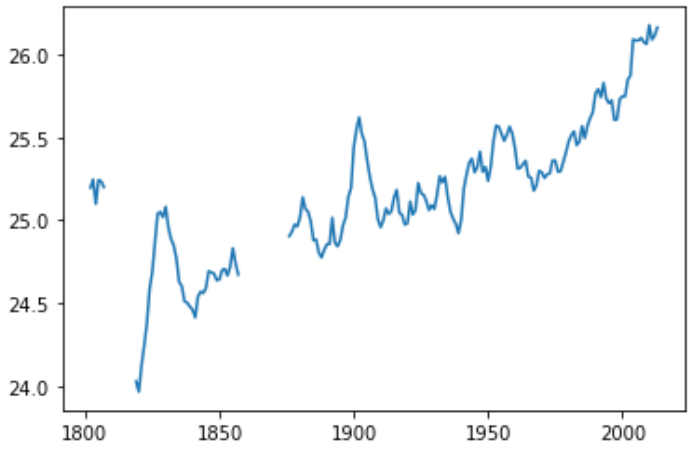
```
1 import pandas as pd
2 df=pd.read_csv('results.csv')
3 df.head()
```

	year	gat	cat
0	1796	8.27	25.03
1	1797	8.51	26.71
2	1798	8.67	24.29
3	1799	8.51	25.28
4	1800	8.48	25.21

6. Plotting the global average temperature with the help of matplotlib.pyplot library with a moving average of 7 years.

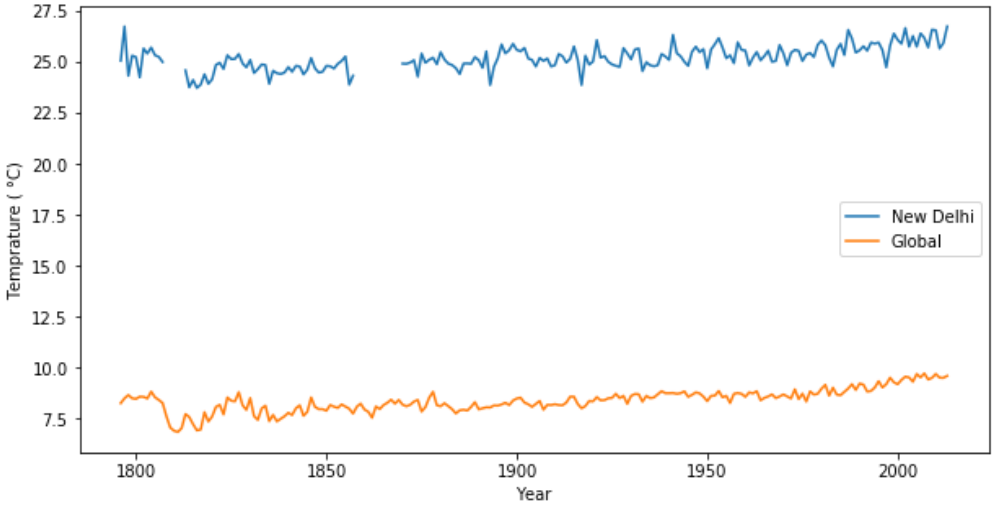


7. Plotting the average temperature of New Delhi with a moving average of 7 years.



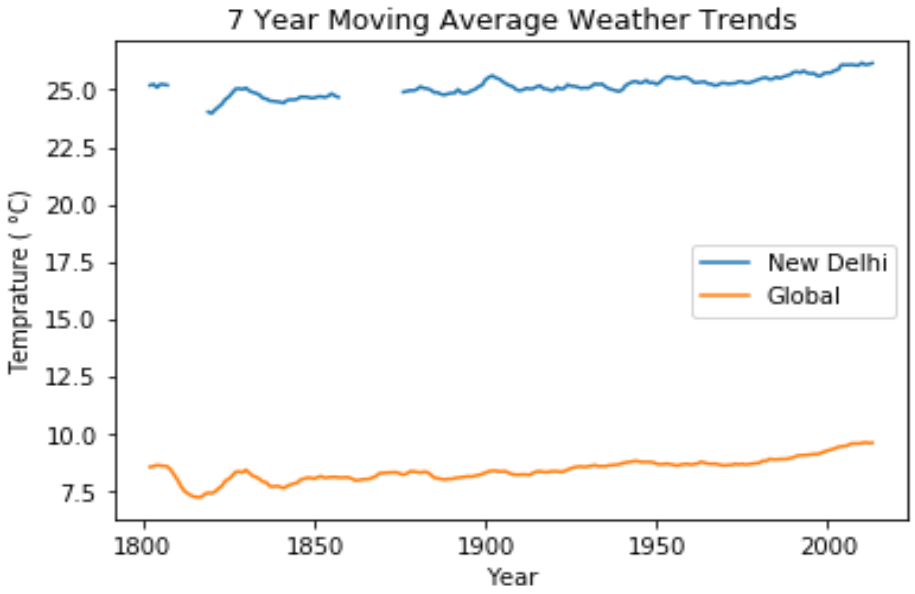
8. Plotting the data of New Delhi and Global Average temperature together.

```
plt.plot(df.year,df.cat)
plt.plot(df.year,df.gat)
plt.legend(['New Delhi', 'Global'])
plt.xlabel('Year')
plt.ylabel('Temprature ( °C)')
plt.show()
```



9. To remove these irregularities, we use the Moving Average of the given average temperatures. Below, we use the 7 years moving average of the given temperatures and plot them.

```
plt.plot(df.year,df.cat.rolling(window=7).mean())
plt.plot(df.year,df.gat.rolling(window=7).mean())
plt.title('7 Year Moving Average Weather Trends')
plt.legend(['New Delhi', 'Global'])
plt.xlabel('Year')
plt.ylabel('Temprature ( °C)')
plt.show()
```



10. Describing the dataset.

	gat	cat
count	218.000000	201.000000
mean	8.403532	25.166269
std	0.548662	0.594003
min	6.860000	23.700000
25%	8.092500	24.800000
50%	8.415000	25.140000
75%	8.727500	25.550000
max	9.730000	26.710000

Observing the Trends

The data clearly shows the difference between the average global temperature and New Delhi's temperature. From the above visualisations and trends, we can conclude the following:

- The mean temperature of New Delhi is more than three times that of global mean temperature.
- The temperature in New Delhi is such that it never goes below 23 °C whereas the average global temperature never went above 10 °C.
- The average Global temperature and New Delhi's temperature iis inclining steeply for the past year which shows the weather changes.
- The global temperature and New Delhi's temperature are proportionally related to each other. As one increases, the other increases too.