

# Data Analysis Nanodegree (*Project : 1*)

UDACITY

## Exploring Weather Trend

### Introduction :

Using SQL , data was downloaded as CSV file which contains name of city and average temperature along with years and also global temperature.

After downloading data as CSV file , data has being analysed using **Python** programming language using **Jupyter Notebook**.

### Progress Outline :

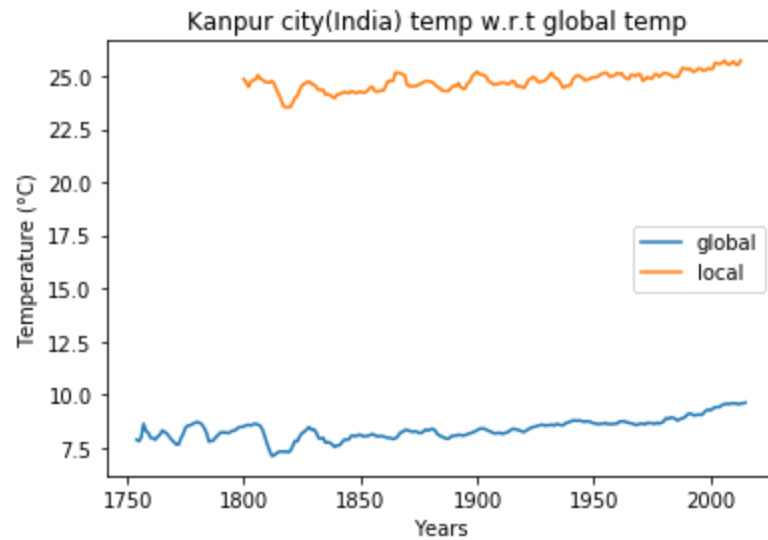
#### 2 files were downloaded using SQL :

- 1) Data about global temperature using :  
`SELECT * FROM global_data`
- 2) Data about local city temperature using:  
`SELECT*FROM city_data WHERE city = 'Kanpur' AND country='India'`

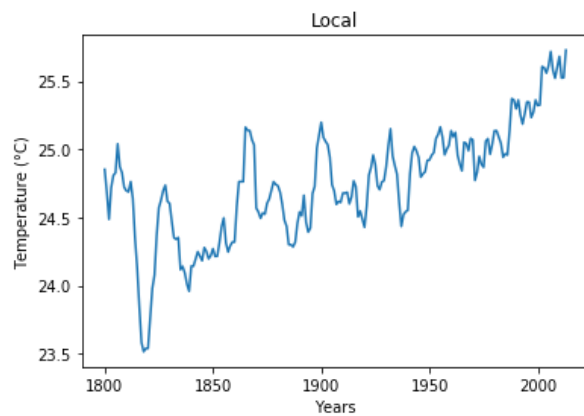
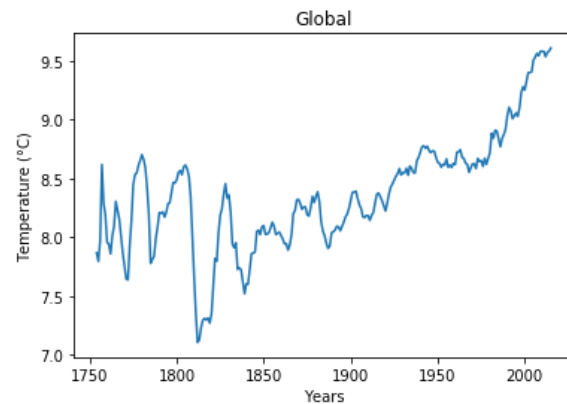
#### Moving Average :

Rolling Average of every 5 years is calculated.

#### Line Chart :



**Separate Chart for Global and Local Average temperature :**



**OBSERVATIONS :**

- 1)The average global temperature is less than Local city (Kanpur's (India)) temperature.
- 2)Average temperature both globally and of particular city is increasing w.r.t years .
- 3)Since 1980 temperature is being constantly rising.Reason may be global warming.
- 4)There is a drop in temperature between 1800 and 1850 , both globally and locally.
- 5)Kanpur city is hotter as compared globally.

### Exploring weather trends

```
In [2]: # Importing necessary Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
In [3]: # Loading the dataset
city_df=pd.read_csv('city.csv')
global_df=pd.read_csv('global_data.csv')
```

```
In [4]: # Checking for null values
city_df.isnull().sum()
```

```
Out[4]: year      0
city      0
country    0
avg_temp   12
dtype: int64
```

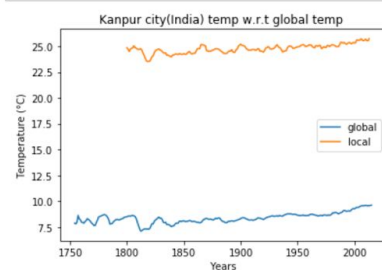
```
In [5]: # Checking for null values
global_df.isnull().sum()
```

```
Out[5]: year      0
avg_temp    0
dtype: int64
```

```
In [6]: # Filling nan values with mean values
mean=city_df['avg_temp'].mean()
city_df['avg_temp'].fillna(mean,inplace=True)
```

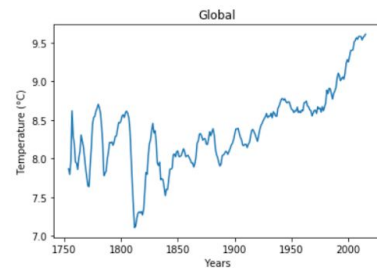
```
In [7]: # Doing moving average for city and global data
global_moving_avg=global_df['avg_temp'].rolling(5).mean()
city_moving_avg=city_df['avg_temp'].rolling(5).mean()
```

```
In [8]: # Plotting the data
plt.plot(global_df['year'],global_moving_avg,label='global');
plt.plot(city_df['year'],city_moving_avg,label='local');
plt.xlabel("Years");
plt.ylabel("Temperature (°C)");
plt.legend();
plt.title("Kanpur city(India) temp w.r.t global temp ");
```



```
In [9]: plt.plot(global_df['year'],global_moving_avg)
plt.xlabel("Years");
plt.ylabel("Temperature (°C)");
plt.title("Global")
```

Out[9]: Text(0.5, 1.0, 'Global')



```
In [11]: plt.plot(city_df['year'],city_moving_avg)
plt.xlabel("Years");
plt.ylabel("Temperature (°C)");
plt.title("Local")
```

Out[11]: Text(0.5, 1.0, 'Local')

