

UDACITY

Data Analyst Nanodegree Program

Project 01: Exploring Weather Trends.

Prepared by: Apurva Verma

GOAL- My goal will be to create visualization and prepare a write up describing the similarities and differences between global temperature trends and temperature trends in the city where I live.

INTRODUCTION:

I have been provided the temperature database for the global temperature and a database with a list of city temperature across the world from the udacity portal where I have extracted the data.

Extract data from the database-

To retrieve the temperature of the closest city I live in I wrote an SQL command. SQL Query was used to download (CSV) file that contains yearly average temperature of the City 'New York' and the global temperature.

Queries used to extract the databases are:

- 1.To select the global temperature database.
- 2.To check which countries and cities are available in the database.
- 3.To select data from the City database

```
select year,avg_temp from global_data;
select city, country from city_list;
select year,avg_temp from city_data where
'New York' and country='United States';
```

Moving Averages:

- Rolling Average has been calculated to smooth out data and to make it easier to observe the trends when it be shown in Charts.
- Moving Average has been calculated for every 10 years to each single data but the first 10!
- IPython was used for calculating the Moving Average Using built-in Functions such as (“rolling”, “mean”).
- Code has been explained underneath:

```
glb_avg = global_data.rolling(window=20).mean()  
glb_avg
```

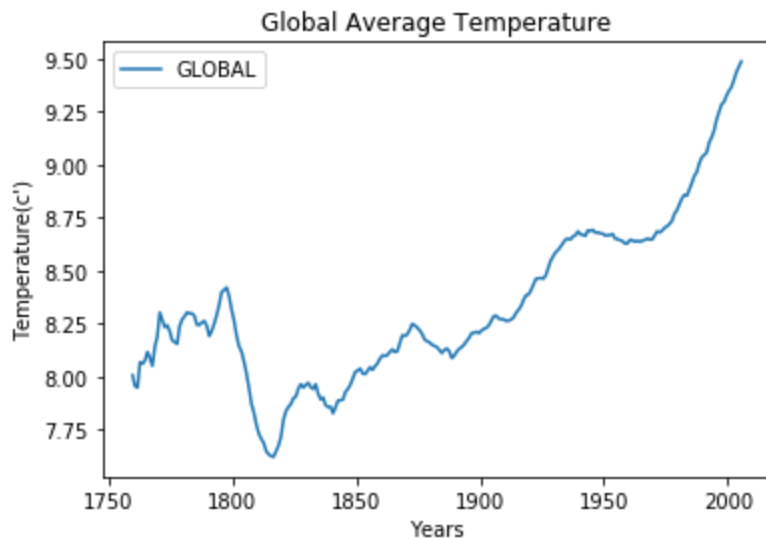
```
city_avg= city_data.rolling(window=20).mean()  
city_avg
```

Line Chart for the data:

The data has been analyzed using Python Programming Language and IPython Notebook (Jupyter). I have plotted a chart for glb_avg and city_avg separately to observe difference between Global Average Temperature and the temperature of the city New York.

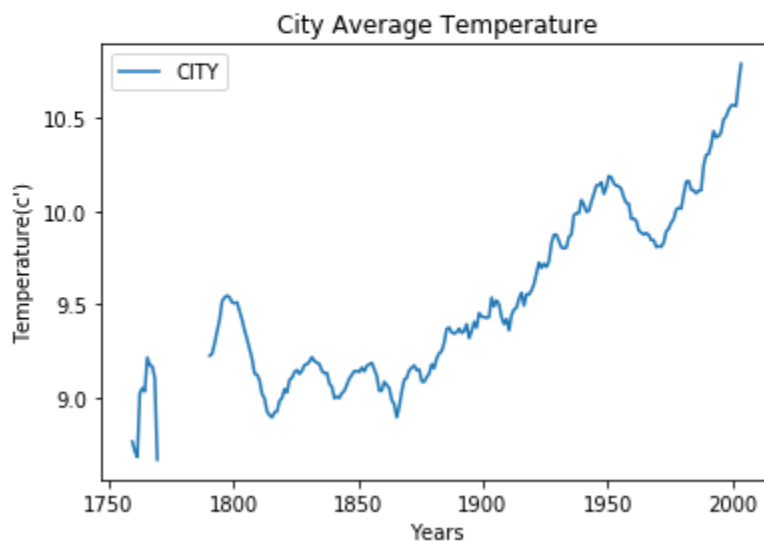
For Global Temperature:

```
glb_avg.plot(x='year',y='avg_temp', label='GLOBAL')  
plt.xlabel("Years")  
plt.ylabel("Temperature(c')")  
plt.title("Global Average Temperature")  
plt.show()
```



For City Temperature:

```
city_avg.plot(x='year',y='avg_temp', label='CITY')  
plt.xlabel("Years")  
plt.ylabel("Temperature(c')")  
plt.title("City Average Temperature")  
plt.show()
```



I observed that the column called avg_temp is same in both city_data and global_data. To change the column name of one dataframe I used 'rename' function and then joined both the dataframes. I changed the column name of city_avg dataframe and renamed it as city_avg_temp. Then I used the append function to join both the dataframes to continue the analysis further.

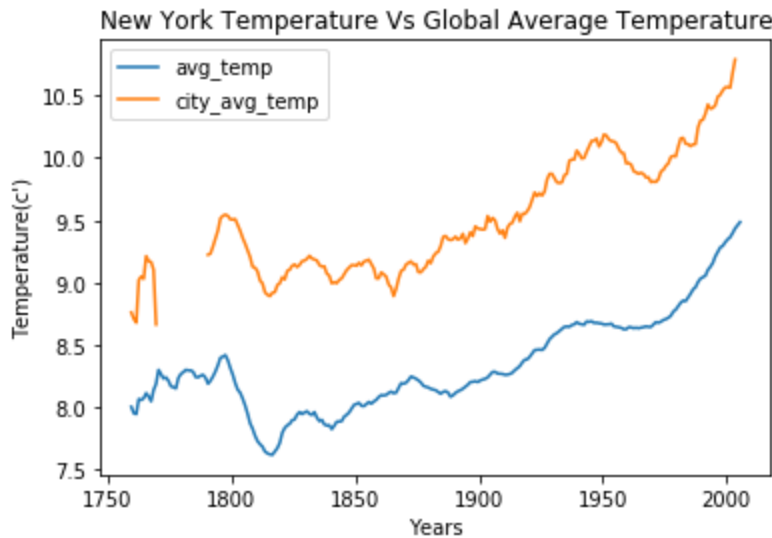
```
city_avg.rename(columns={'avg_temp':'city_avg_temp'},inplace=True)  
city_avg
```

```
df=city_avg.append(glb_avg)  
df
```

Data Visualization:

I am now able to plot a line chart to show a comparison between the city average temperature and global average temperature. This is done by plotting the moving average temperatures on the y-axis and the year range on the x-axis. After inputting the corresponding values, the following line chart is generated as given below.

```
df.plot(x='year',y=['avg_temp','city_avg_temp'])  
plt.xlabel("Years")  
plt.ylabel("Temperature(c')")  
plt.title("New York Temperature Vs Global Average Temperature")  
plt.show()
```



The following observations may be deduced:

- 1) New York's weather is much warmer than the global average considering that the temperature has always been greater in the past couple hundred years.
- 2) In both cases, we can see that the average temperature is gradually increasing throughout the entire time-frame.
- 3) The difference between year 1800 and ~2010 in temperature is more than 2 degrees in the Global average chart (increasing).
- 4) The difference between year 1800 and ~2010 is more than 2.5 degrees in the New York City average chart (increasing).

CONCLUSION:

There is an evidence suggests that the global temperature is monotonically increasing over the years which support the case of climate change. The final conclusion of this project is New York is hotter than global temperature and temperature is increasing day by day due to changes in the climate.

