THIORO FALL July ,2019

Data Analyst Nanodegree

Project1- Explore Weather Trends

Objectif

In this project, I will analyze the local temperature of Arlington city from United States and compare them with the global temperature. Udacity provided me a database from where I have to extract, manipulate and visualize data.

Goals

1. Extraction the data from the database and exporting to CSV file
2. Create a line chart that compares my city's temperature which is Arlington and the global temperature. Plot the moving average rather than the yearly average in order to smooth out the lines, making trends more visual.
3. Analyze the charts of the average temperature of the city of Arlington and the global temperature

Tools used

1. SQL: to extract the data from the database
2. Python: to calculate the moving average and plotting line chart
3. Anaconda- Jupiter Note book: to write python code and make observations

Since there are some missing record for city of Arlington, United States, I extract global and city data from 1850.

a) Write a SQL query to extract the city level data. Export to CSV

```
SELECT *
FROM city_data
WHERE country='United States'
and city='Arlington' and year>1849;
```

b) Write a SQL query to extract the global data. Export to CSV

```
SELECT *
FROM global _data
WHERE year>1849;
```

c) Moving averages:

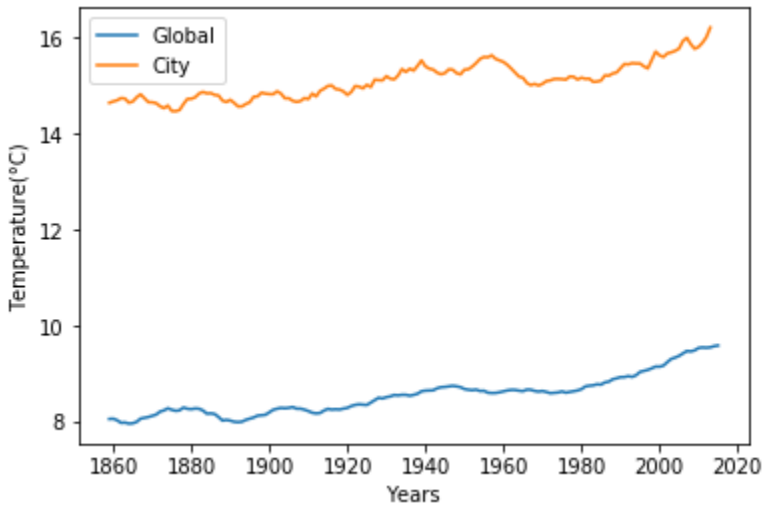
- The rolling average has been calculated for every 10 years
- Python was used to calculate the moving average which built-in function: rolling and mean
- Python code:

```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
global_data=pd.read_csv(r"C:\Users\Thioro\Desktop\global_data.csv")
city_data=pd.read_csv("C:/Users/Thioro/Desktop/city_data.csv")
#print(city_data)
glb_mv_avg = global_data['avg_temp'].rolling(10).mean()
city_mv_avg = city_data['avg_temp'].rolling(10).mean()
#print(city_mv_avg)
plt.plot(global_data['year'],glb_mv_avg, label='Global')
plt.plot(city_data['year'],city_mv_avg, label='City')
plt.legend()
plt.xlabel("Years")
plt.ylabel("Temperature(°C)")
plt.show()
```

d) Line chart with local and global temperature trends:

My key consideration will be to observe the fluctuations of the city temperature compare to the global temperature. Because of the missing records, I will visualize the trends from 1850 to 2013.

Average Temperature for Earth and Arlington city in USA



e) Observations:

- The general observation is that from 1858 to 2013 the average temperature for the Earth and Arlington city have increase progressively.
- The global average temperature went from 8°C to 10°C between 1850 to 2013; which is an increase of 2°C. However, the average temperature for Arlington city went from 14°C to 16°C within the same time frame; which is an increase of 2°C also.
- Basically, the average temperature of Arlington city is approximately twice hotter than the global average temperature.
- The reason of the high average temperature of the city of Arlington may be due to its location. Arlington city is located in Washington DC area which have a lot of pollutions. Washington Dc has been known as one of the most polluted area in the United States.

In conclusion, the average city temperature of Arlington city is hotter than the global average temperature. Climate change is affecting Arlington city. So, what will happen in 50 years, 100 years? How are people going to survive in Arlington city if the temperature continue to increase?