



Data Analyst Nanodegree

Explore Weather Trends

Name: Amr Mohamed Hosny

September 2017

Pulling the data from the database:

1) Check the nearest city

```
SELECT *  
FROM city_list;
```

Nearest City: city = Cairo , country = Egypt

2) Pull global_data

```
SELECT *  
FROM global_data  
ORDER BY year;
```

3) Pull Cairo's data

```
SELECT year, avg_temp  
FROM city_data  
WHERE city = 'Cairo'  
AND country = 'Egypt'  
ORDER BY year;
```

Manipulating csv files :

- Python is used to calculate the moving average and generate the line plot.
- Since all datapoints are one year apart then the moving average at year y is

$$\sum_{i=startyear}^{i=y} temp[i] / (1 + y - startyear)$$

```
import csv
import matplotlib.pyplot as plt

def calculate_moving_average(years, temps):
    if not years:
        return []
    moving_average = [temps[0]]
    summation_temp = temps[0]
    start_year = years[0]
    for idx in range(1, len(years)):
        summation_temp += temps[idx]
        moving_average.append(
            summation_temp / (1 + years[idx] - start_year))
    return moving_average

if __name__ == '__main__':

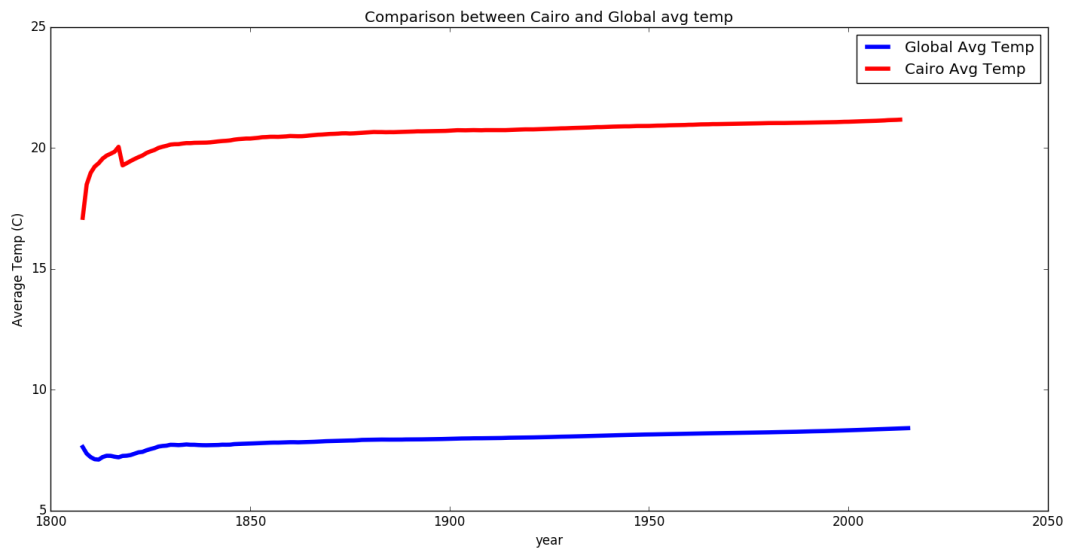
    cairo_temp = []
    cairo_years = []
    with open('cairo_data.csv') as cairo_data_file:
        reader = csv.DictReader(cairo_data_file)
        for row in reader:
            cairo_temp.append(float(row['avg_temp']))
            cairo_years.append(int(row['year']))

    global_temp = []
    global_years = []
    with open('global_data.csv') as global_data_file:
        reader = csv.DictReader(global_data_file)
        for row in reader:
            # Ignore years that came before the start year of Cairo's data
            cur_year = int(row['year'])
            if cur_year < int(cairo_years[0]):
                continue
            global_temp.append(float(row['avg_temp']))
            global_years.append(int(row['year']))

    global_temp_moving_avg = calculate_moving_average(
        global_years, global_temp)
    cairo_temp_moving_avg = calculate_moving_average(
        cairo_years, cairo_temp)
```

```
global_handle, = plt.plot(
    global_years, global_temp_moving_avg,
    label = 'Global Avg Temp', color = 'blue',
    linewidth = 4)
cairo_handle, = plt.plot(
    cairo_years, cairo_temp_moving_avg,
    label = 'Cairo Avg Temp', color = 'red',
    linewidth = 4)
plt.legend(handles = [global_handle, cairo_handle])
plt.xlabel('year')
plt.ylabel('Average Temp (C)')
plt.ylim([5,25])
plt.title("Comparison between Cairo and Global avg temp")
plt.show()
```

Data Interpretation:



- Temperature in Cairo, Egypt is for sure higher than the global average temperature.
- Temperatures tend to fluctuate at the start of the 1800s.
- There is a significant increase in the average temperatures for both Cairo, Egypt and the Globe. The world tend to be hotter as time progresses.
- During the last 200 years, the difference between Cairo's average temp and the global average temp lies in range [+11,+13] which is nearly constant.