


Tools


- Python 3
- Jupyter Notebook


Data Extraction

SQL Query to get global data

Input		HISTORY ▾	MENU ▾
SCHEMA		1 <code>SELECT * FROM global_data</code>	
city_data	▾		
city_list	▾		
global_data	▾		
		Success!	EVALUATE

SQL Query to get the list of cities in my country

Input		HISTORY ▾	MENU ▾
SCHEMA		1 <code>SELECT city FROM city_list</code> 2 <code>WHERE country = 'Algeria'</code>	
city_data	▾		
city_list	▾		
global_data	▾		
		Success!	EVALUATE

Output	1 results	 Download CSV
city		
	Algiers	

There is only one city available which is Algiers, actually, I live in Algiers

SQL Query to get local data

Input

HISTORY ▾

MENU ▾

SCHEMA

city_data

year

city

country

avg_temp

1

SELECT year, avg_temp FROM city_data

2

WHERE country = 'Algeria' AND city = 'Algiers'

Success!

EVALUATE

Output

261 results

Download CSV

CSV files opening

```
#read the data from csv files
import pandas as pd
local_data = pd.read_csv('C:\\Users\\pc\\Desktop\\workspace\\data_analysis\\project1\\local_data.csv')
global_data = pd.read_csv('C:\\Users\\pc\\Desktop\\workspace\\data_analysis\\project1\\global_data.csv')
```

Data description

```
global_data.describe()
```

	year	avg_temp
count	266.000000	266.000000
mean	1882.500000	8.369474
std	76.931788	0.584747
min	1750.000000	5.780000
25%	1816.250000	8.082500
50%	1882.500000	8.375000
75%	1948.750000	8.707500
max	2015.000000	9.830000

```
local_data.describe()
```

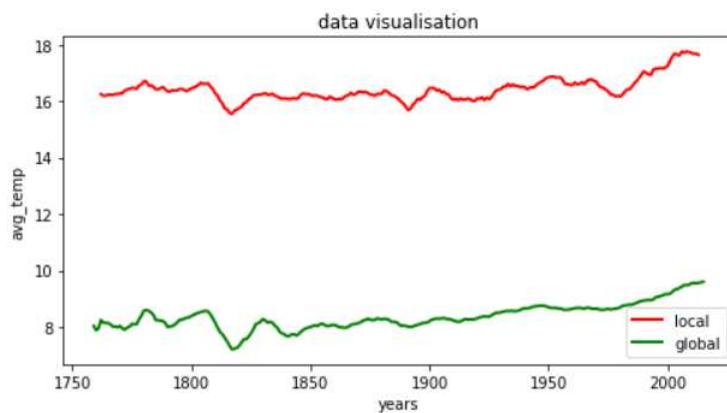
	year	avg_temp
count	261.00000	261.000000
mean	1883.00000	16.433218
std	75.48841	0.592456
min	1753.00000	15.160000
25%	1818.00000	16.050000
50%	1883.00000	16.350000
75%	1948.00000	16.780000
max	2013.00000	18.130000

Calculating moving averages

```
local_data['ma_10'] = local_data['avg_temp'].rolling(10).mean()  
global_data['ma_10'] = global_data['avg_temp'].rolling(10).mean()
```

Creating line chart

```
plt.figure(figsize=(8,4))  
plt.title("data visualisation")  
plt.xlabel('years')  
plt.ylabel('avg_temp')  
plt.plot('year', 'ma_10', data=local_data, marker='', color='red', linewidth=2, label='local')  
plt.plot('year', 'ma_10', data=global_data, marker='', color='green', linewidth=2, label='global')  
plt.legend()  
plt.show()
```



Observations

- According to the local temperature average my city is hotter compared to the global temperature average.
- The global average temperature is increasing over the years, the world is getting hotter
- Similarly the local average temperature is increasing over the years

- The difference between local average temperature and global average temperature is consistent over the years (slight changes)

```
diff = local_data['avg_temp'] - global_data['avg_temp']  
diff.hist()
```

<matplotlib.axes._subplots.AxesSubplot at 0x2f1ad3b7c88>

