My SQL Queries

Input HISTORY MENU

SCHEMA

city_data

city_list

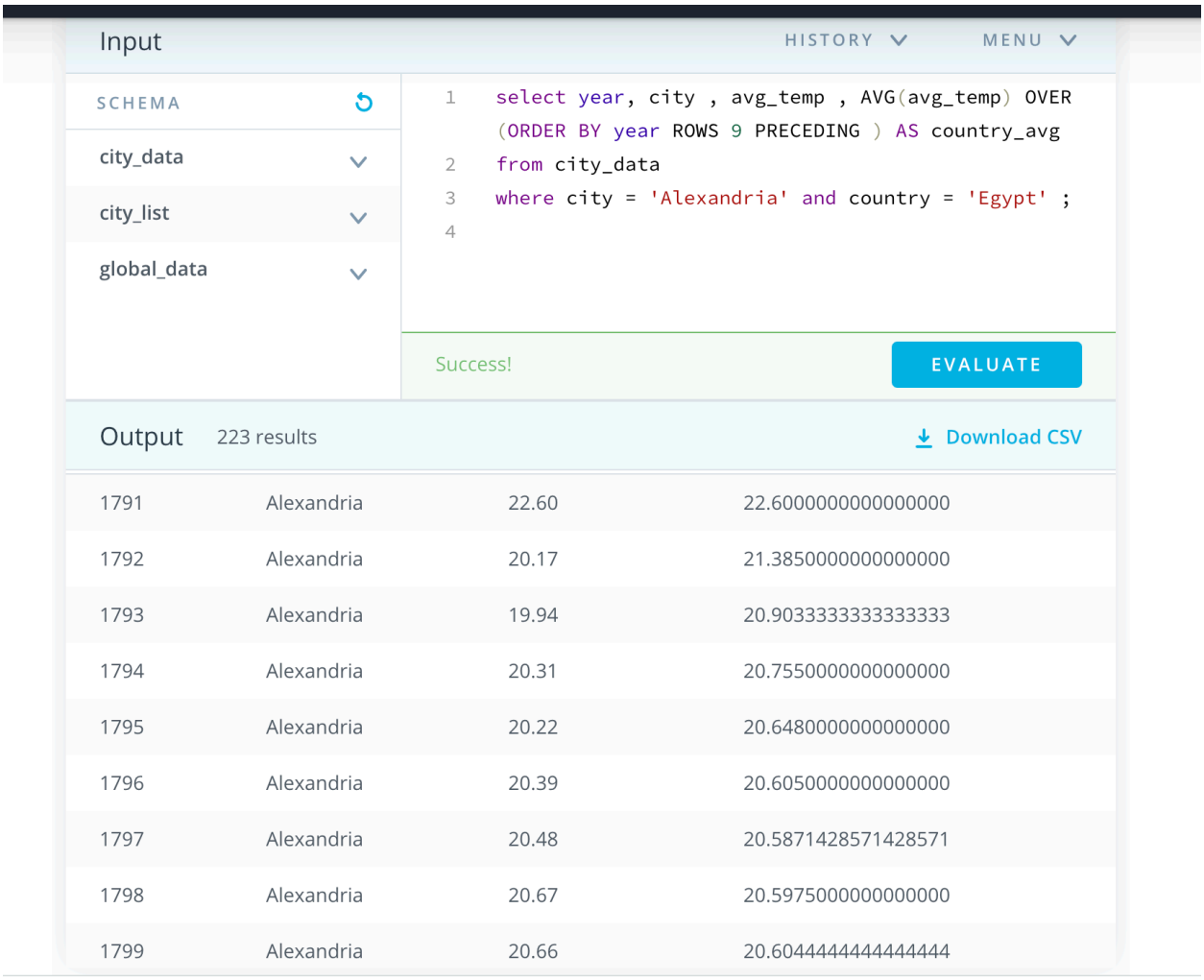
global_data

```
1  select * , AVG(avg_temp) OVER (ORDER BY year ASC
   ROWS 9 PRECEDING ) AS global_avg
2  from global_data
3  where year between 1791 and 2013 ;
4
5
```

Success! EVALUATE

Output 223 results Download CSV

1801	8.59	8.4230000000000000
1802	8.58	8.4720000000000000
1803	8.50	8.4990000000000000
1804	8.84	8.5300000000000000
1805	8.56	8.5510000000000000
1806	8.43	8.5670000000000000
1807	8.28	8.5440000000000000
1808	7.63	8.4400000000000000


Input
HISTORY
MENU
SCHEMA
city_data
city_list
global_data
1 select year, city , avg_temp , AVG(avg_temp) OVER
(ORDER BY year ROWS 9 PRECEDING) AS country_avg
2 from city_data
3 where city = 'Alexandria' and country = 'Egypt' ;
4
Success!
EVALUATE
Output
223 results
Download CSV
1791 Alexandria 22.60 22.6000000000000000
1792 Alexandria 20.17 21.3850000000000000
1793 Alexandria 19.94 20.9033333333333333
1794 Alexandria 20.31 20.7550000000000000
1795 Alexandria 20.22 20.6480000000000000
1796 Alexandria 20.39 20.6050000000000000
1797 Alexandria 20.48 20.5871428571428571
1798 Alexandria 20.67 20.5975000000000000
1799 Alexandria 20.66 20.6044444444444444

The project

In [7]:
```
import pandas as pd
url1='/Users/ahmedyasser/Downloads/results.csv'
```

In [8]:
```
df1= pd.read_csv(url1)
```

In [9]:
```
df1
```

Out[9]:

	year	avg_temp	global_avg
0	1791	8.23	8.230000
1	1792	8.09	8.160000
2	1793	8.23	8.183333
3	1794	8.53	8.270000
4	1795	8.35	8.286000
...	...	...	...
218	2009	9.51	9.493000
219	2010	9.70	9.543000
220	2011	9.52	9.554000
221	2012	9.51	9.548000
222	2013	9.61	9.556000

223 rows × 3 columns

In [10]:
```
url2='/Users/ahmedyasser/Downloads/results (1).csv'
```

In [11]:
```
df2= pd.read_csv(url2)
```

In [12]:
```
df2
```

Out[12]:

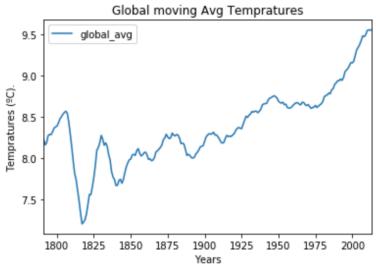
	year	city	avg_temp	country_avg
0	1791	Alexandria	22.60	22.600000
1	1792	Alexandria	20.17	21.385000
2	1793	Alexandria	19.94	20.903333
3	1794	Alexandria	20.31	20.755000
4	1795	Alexandria	20.22	20.648000
...	...	...	...	...
218	2009	Alexandria	21.67	21.308000
219	2010	Alexandria	22.46	21.480000
220	2011	Alexandria	21.18	21.453000
221	2012	Alexandria	21.55	21.462000
222	2013	Alexandria	21.44	21.484000

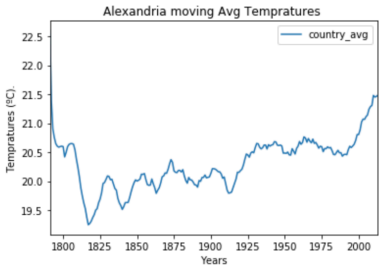
223 rows × 4 columns

In [15]:
```
import matplotlib as mpl
import matplotlib.pyplot as plt
```

In [17]:
```
df1.plot(kind='line', x='year', y='global_avg')
plt.title('Global moving Avg Tempratures')
plt.xlabel('Years')
plt.ylabel('Tempratures (ºC).')
df2.plot(kind='line',x='year',y='country_avg')
plt.title('Alexandria moving Avg Tempratures ')
plt.xlabel('Years')
plt.ylabel('Tempratures (ºC).')
plt.show
```

Out[17]:
```
<function matplotlib.pyplot.show(*args, **kw)>
```





Observation

1- Alexandria is a hot city compared to the global average

2- As the Global temprature average tends to increase after year 1800 alexandria's temprature tends to decrease

3- The Global and Alexandria's temprature moving average tend to increase in 1900

4- The Average global temperature is getting so much higher while also alexandria's temprature is getting higher but not with the same Rate compared to the average global temperature .

Out lines

1- I extracted the data using sql queries and use python to analyze and visualize them used them on jupyter notebook .

2- I calculated the moving average using an aggregate function as in this sql code for the city

(select year, city , avg_temp , AVG(avg_temp) OVER (ORDER BY year ROWS 9 PRECEDING) AS country_avg from city_data where city = 'Alexandria' and country = 'Egypt' ;)

and this sql code for the global average

(select * , AVG(avg_temp) OVER (ORDER BY year ASC ROWS 9 PRECEDING) AS global_avg from global_data where year between 1791 and 2013 ;)

3- the key consideration is to see how much the temprature is changed globaly compared to Alexandria through the years

Ahmed Yasser - Ahmadyasser1999@gmail.com