SQL query used to extract data for both Lagos City & Toronto

SELECT * FROM city_data WHERE city = 'Lagos'

SELECT * FROM city_data WHERE city = 'Toronto'

In [1]:
```
# importing libraries to be used in my project
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:
```
city_data_lagos = pd.read_csv('./city_data_lagos.csv', usecols = ['year', 'avg_temp'])
city_data_lagos
```

Out[2]:

	year	avg_temp
0	1849	25.98
1	1850	25.87
2	1851	26.10
3	1852	NaN
4	1853	NaN
...	...	...
160	2009	27.53
161	2010	27.79
162	2011	27.35
163	2012	27.15
164	2013	27.36

165 rows × 2 columns

In [3]:
```
# Replacing NaN values with the mean avg temp for Lagos city
city_data_lagos.fillna(city_data_lagos.mean(), inplace = True)
city_data_lagos.isnull().sum()
```

Out[3]:
```
year        0
avg_temp    0
dtype: int64
```

Calculating a 10 year moving average

In [7]:
```
# Calculating and creating a new column for 10 year moving average for the city of Lagos
city_data_lagos['moving_average_lagos'] = city_data_lagos['avg_temp'].rolling(10).mean()
```

In [8]:
```
# New df that includes a new column for 10 year moving average for Lagos city
city_data_lagos.head(20)
```

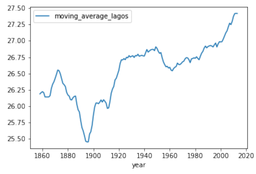
Out[8]:

	year	avg_temp	moving_average	moving_average_lagos
0	1849	25.980000	NaN	NaN
1	1850	25.870000	NaN	NaN
2	1851	26.100000	NaN	NaN
3	1852	26.553113	NaN	NaN
4	1853	26.553113	NaN	NaN
5	1854	26.553113	NaN	NaN
6	1855	26.553113	NaN	NaN
7	1856	26.350000	NaN	NaN
8	1857	25.450000	NaN	NaN
9	1858	25.920000	26.188245	26.188245
10	1859	26.180000	26.208245	26.208245
11	1860	26.010000	26.222245	26.222245
12	1861	25.950000	26.207245	26.207245
13	1862	25.900000	26.141934	26.141934
14	1863	26.553113	26.141934	26.141934
15	1864	26.553113	26.141934	26.141934
16	1865	26.553113	26.141934	26.141934
17	1866	26.553113	26.162245	26.162245
18	1867	26.553113	26.272556	26.272556
19	1868	26.553113	26.335868	26.335868

Plotting a line chart with the calculated moving average

In [16]:
```
# Plotting a line chart for visualization
city_data_lagos.plot.line(x='year', y = 'moving_average_lagos')
```

Out[16]:
```
<AxesSubplot:xlabel='year'>
```



Repeating the above step for the city of Toronto

In [10]:
```
city_data_toronto = pd.read_csv('./city_data_toronto.csv', usecols = ['year', 'avg_temp'])
city_data_toronto
```

Out[10]:

	year	avg_temp
0	1743	-0.11
1	1744	8.38
2	1745	-3.96
3	1746	NaN
4	1747	NaN
...	...	...
266	2009	6.28
267	2010	7.77
268	2011	7.30
269	2012	8.66
270	2013	8.46

271 rows × 2 columns

In [11]:
```
city_data_toronto.fillna(city_data_toronto.mean(), inplace = True)
city_data_toronto.isnull().sum()
```

Out[11]:
```
year        0
avg_temp    0
dtype: int64
```

In [12]:
```
city_data_toronto['moving_average_toronto'] = city_data_toronto['avg_temp'].rolling(10).mean()
```

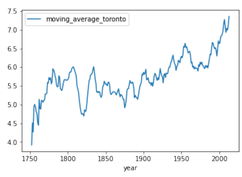
In [13]:
```
city_data_toronto.head(20)
```

Out[13]:

	year	avg_temp	moving_average_toronto
0	1743	-0.110000	NaN
1	1744	8.380000	NaN
2	1745	-3.960000	NaN
3	1746	5.721985	NaN
4	1747	5.721985	NaN
5	1748	5.721985	NaN
6	1749	5.721985	NaN
7	1750	6.290000	NaN
8	1751	6.840000	NaN
9	1752	-1.100000	3.922794
10	1753	5.760000	4.509794
11	1754	5.940000	4.265794
12	1755	2.810000	4.942794
13	1756	6.370000	5.007596
14	1757	5.130000	4.948397
15	1758	4.370000	4.813199
16	1759	5.270000	4.768000
17	1760	3.740000	4.513000
18	1761	6.250000	4.454000
19	1762	5.790000	5.143000

In [15]:
```
city_data_toronto.plot.line(x='year', y = 'moving_average_toronto')
```

Out[15]:
```
<AxesSubplot:xlabel='year'>
```



My Obeservation

1. The 10 year moving average data for Lagos city shows that it was hotter than Toronto in the year 2000 reaching a temperature of over 26 degrees while that of Toronto was between 5 - 6 degrees
2. Even though Toronto seems to be colder that than Lagos, it has more variations in temperature change compared to Lagos with the given available data
3. The hottest temperate for Lagos was 27.50 degrees between year 2000 - 2020 while that of Toronto was 7.5 beyond year 2000
4. From the given data, Lagos city has its lowest temperature a litle below 25.50 degrees which occurred between the year 1880 - 1900 while the lowest for Toronto was a little below 4 degrees which happened around 1750

In [ ]:
```
```