Project - Exploring weather trends

Global temperature has been a hot topic over years. In this project, I analyzed the local and global temperature data to compare the temperature trends where I live to overall temperature trends.

> Firstly, we need to gather the required dataset of the city and the world from the Database provided by Udacity. For this, I used SQL commands that is the SELECT Query to get the datset and then I downloaded it from the source.

Query for selecting city's data

- select * from city_data where city='Patna' and country='India';

Query for slecting global data

- select * from global_data;

In [12]:

```
#import the packages
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

> Gather City's Data using python's library Pandas

In [13]:

```
df_city=pd.read_csv('C:/Users/somya/Desktop/explore_weather_trends/city.csv')
```

In [14]:

```
df_city.head(20)
```

Out[14]:

	year	city	country	avg_temp
0	1796	Patna	India	24.99
1	1797	Patna	India	26.49
2	1798	Patna	India	24.27
3	1799	Patna	India	25.25
4	1800	Patna	India	25.20
5	1801	Patna	India	24.19
6	1802	Patna	India	25.64
7	1803	Patna	India	25.40
8	1804	Patna	India	25.72
9	1805	Patna	India	25.30
10	1806	Patna	India	25.21
11	1807	Patna	India	24.69
12	1808	Patna	India	NaN
13	1809	Patna	India	NaN
14	1810	Patna	India	NaN
15	1811	Patna	India	NaN
16	1812	Patna	India	NaN
17	1813	Patna	India	24.55
18	1814	Patna	India	23.80
19	1815	Patna	India	24.08

There are some missing values in the dataset. So, we need to remove those missing value rows to avoid any discontinuity in graph.

In [15]:

```
df_city.dropna(axis=0,inplace=True)
```

In [16]:

```
df_city.head(20)
```

Out[16]:

	year	city	country	avg_temp
0	1796	Patna	India	24.99
1	1797	Patna	India	26.49
2	1798	Patna	India	24.27
3	1799	Patna	India	25.25
4	1800	Patna	India	25.20
5	1801	Patna	India	24.19
6	1802	Patna	India	25.64
7	1803	Patna	India	25.40
8	1804	Patna	India	25.72
9	1805	Patna	India	25.30
10	1806	Patna	India	25.21
11	1807	Patna	India	24.69
17	1813	Patna	India	24.55
18	1814	Patna	India	23.80
19	1815	Patna	India	24.08
20	1816	Patna	India	23.81
21	1817	Patna	India	23.87
22	1818	Patna	India	24.00
23	1819	Patna	India	23.74
24	1820	Patna	India	24.02

Now, we will calculate the moving average of avg_temp using python's mean function for 10 observations per window.

In [17]:

```
moving_avg_city = df_city['avg_temp'].rolling(window=10).mean()
```

In [18]:

```
moving_avg_city
```

Out[18]:

```
0         NaN
1         NaN
2         NaN
3         NaN
4         NaN
5         NaN
6         NaN
7         NaN
8         NaN
9      25.245
10     25.267
11     25.087
17     25.115
18     24.970
19     24.858
20     24.820
21     24.643
22     24.503
23     24.305
24     24.177
25     24.144
26     24.179
27     24.200
28     24.351
29     24.452
30     24.583
31     24.725
32     24.815
33     24.912
34     25.011
        ...
188    25.418
189    25.440
190    25.422
191    25.470
192    25.564
193    25.503
194    25.472
195    25.515
196    25.536
197    25.591
198    25.635
199    25.634
200    25.669
201    25.628
202    25.618
203    25.675
204    25.689
205    25.716
206    25.764
207    25.781
208    25.779
209    25.814
210    25.852
```

```
211    25.865
212    25.834
213    25.868
214    25.964
215    25.930
216    25.922
217    26.018
Name: avg_temp, Length: 210, dtype: float64
```

Gathering Global Dataset

In [19]:

```
df_global=pd.read_csv('C:/Users/somya/Desktop/explore_weather_trends/Global.csv')
```

In [20]:

```
df_global.head()
```

Out[20]:

	year	avg_temp
0	1750	8.72
1	1751	7.98
2	1752	5.78
3	1753	8.39
4	1754	8.47

Finding the moving Average of global avg_temp

In [21]:

```
moving_avg_global = df_global['avg_temp'].rolling(window=10).mean()
```

In [22]:

```
moving_avg_global
```

Out[22]:

```
0        NaN
1        NaN
2        NaN
3        NaN
4        NaN
5        NaN
6        NaN
7        NaN
8        NaN
9      8.030
10     7.877
11     7.956
12     8.239
13     8.150
14     8.143
15     8.132
16     8.088
17     8.008
18     8.012
19     7.982
20     8.032
21     7.940
22     7.898
23     7.970
24     8.007
25     8.100
26     8.089
27     8.093
28     8.269
29     8.398
       ...
236    8.827
237    8.841
238    8.892
239    8.911
240    8.936
241    8.937
242    8.957
243    8.941
244    8.976
245    9.045
246    9.066
247    9.087
248    9.119
249    9.156
250    9.153
251    9.176
252    9.249
253    9.315
254    9.343
255    9.378
256    9.427
257    9.480
258    9.471
```

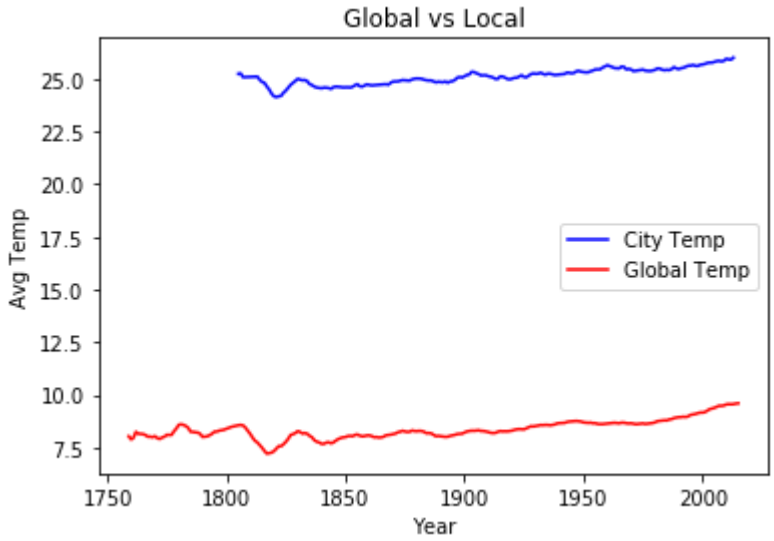
```
259    9.493
260    9.543
261    9.554
262    9.548
263    9.556
264    9.581
265    9.594
Name: avg_temp, Length: 266, dtype: float64
```

> Plotting the line chart using matplotlib.

- First, we will plot both the moving averages together to compare differences between local and global temperature.

In [31]:

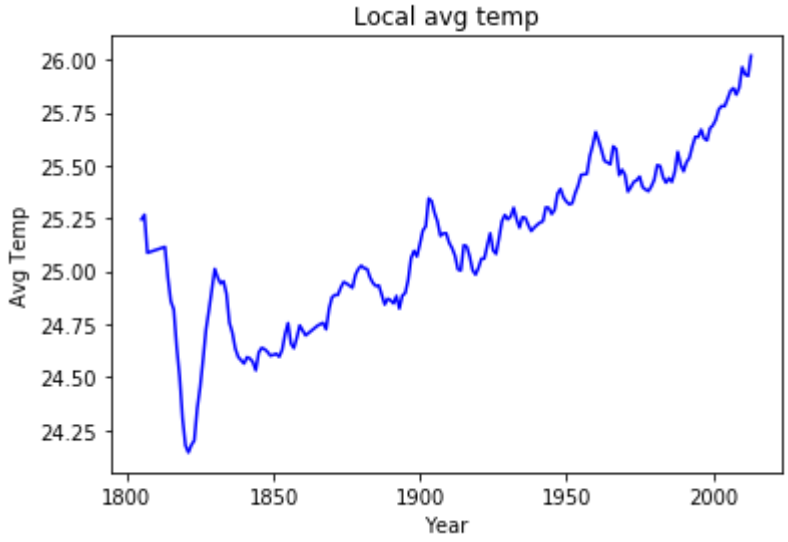
```
plt.plot(df_city['year'],moving_avg_city, label='City Temp', color='blue')
plt.plot(df_global['year'],moving_avg_global, label='Global Temp', color='red');
plt.xlabel('Year')
plt.ylabel('Avg Temp')
plt.title('Global vs Local')
plt.legend();
```



- Now, we will plot individual line charts to observe local and global trends over time.

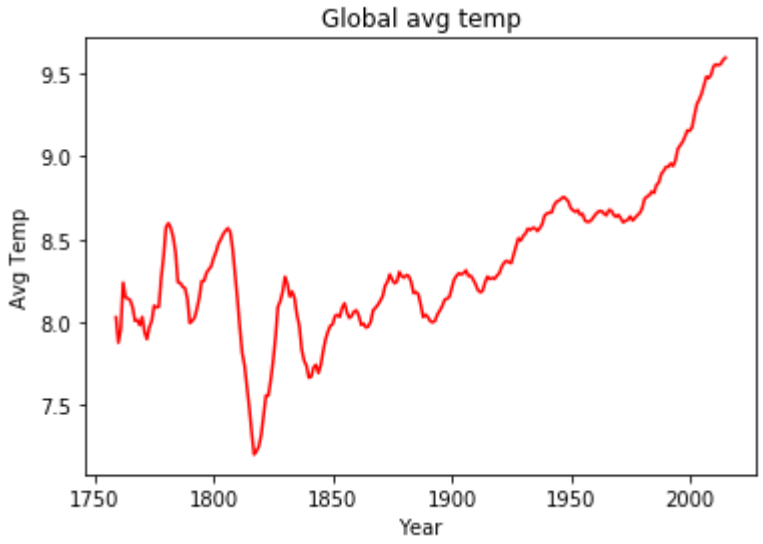
In [33]:

```
plt.plot(df_city['year'],moving_avg_city, label='City Temp', color='blue')
plt.xlabel('Year')
plt.ylabel('Avg Temp')
plt.title('Local avg temp');
```



In [34]:

```
plt.plot(df_global['year'],moving_avg_global, label='Global Temp', color='red');
plt.xlabel('Year')
plt.ylabel('Avg Temp')
plt.title('Global avg temp');
```



Conclusions:

- The average temperature of my city ranges from 23 degree celcius to 25 degree celcius whereas the average temperature of world ranges from 4 degree celcius to 10 degree celcius. My city is hotter with respect to world.
- We can observe from the above graph that the difference is consistent over years.
- Between 1800 and 1850,for 2-3 years there is a drastic decrease in the global temperature as well as the city temperature and then sudden increase in the temperature of city and world. After 1850, there were small fluctuations in temperature.After 1975, world's temperature is increasing exponentially,whereas city's temperature still fluctuates i.e. it decreases or increases.
- The overall trends looks like the world is getting hotter over years.Over the last 100 years, the temperature is increasing due to climate change.
- Thus we conclude that, we need to control climate changes otherwise the world will become more hotter in the coming years as the temperature is increasing exponentially.

References:

https://towardsdatascience.com/implementing-moving-averages-in-python-1ad28e636f9d (https://towardsdatascience.com/implementing-moving-averages-in-python-1ad28e636f9d)

https://pythonbasics.org/matplotlib-line-chart/ (https://pythonbasics.org/matplotlib-line-chart/)

In [ ]: