Exploring Weather Trends

1. Extraction of Data

In this step data was extracted from database where three tables were present namely city_list, city_data and global_data.

- city_list - This contains a list of cities and countries in the database. Look through them in order to find the city nearest to you.
  SCHEMA: Attribute (city (character) , country (character))

- city_data - This contains the average temperatures for each city by year (ºC).
  SCHEMA: Attribute (year (integer), city (character), country (character), avg_temp(numeric))

- global_data - This contains the average global temperatures by year (ºC).
  SCHEMA: Attribute (year (integer), avg_temp(numeric))

In this step to extract the required data from database SQL Query was performed of the database which is given below:

```
select city_data.year,city_data.city,city_data.country,city_data.avg_temp , global_data.avg_temp
as global_avg_temp from city_data inner join global_data on city_data.year=global_data.year where
city_data.city like 'New Delhi'
```

After this the output generated was downloaded in a csv format and Attribute named were changed for convenience , below is the output generated:

	A	B	C	D	E
1	year	city	country	city_avg_temp_na	global_avg_temp
2	1796	New Delhi	India	25.03	8.27
3	1797	New Delhi	India	26.71	8.51
4	1798	New Delhi	India	24.29	8.67
5	1799	New Delhi	India	25.28	8.51
6	1800	New Delhi	India	25.21	8.48
7	1801	New Delhi	India	24.22	8.59
8	1802	New Delhi	India	25.63	8.58
9	1803	New Delhi	India	25.38	8.5
10	1804	New Delhi	India	25.68	8.84
11	1805	New Delhi	India	25.3	8.56
12	1806	New Delhi	India	25.22	8.43
13	1807	New Delhi	India	24.97	8.28
14	1808	New Delhi	India		7.63
15	1809	New Delhi	India		7.08
16	1810	New Delhi	India		6.92
17	1811	New Delhi	India		6.86
18	1812	New Delhi	India		7.05
19	1813	New Delhi	India	24.56	7.74
20	1814	New Delhi	India	23.73	7.59
21	1815	New Delhi	India	24.09	7.24
22	1816	New Delhi	India	23.7	6.94
23	1817	New Delhi	India	23.86	6.98
24	1818	New Delhi	India	24.37	7.83
25	1819	New Delhi	India	23.9	7.37
26	1820	New Delhi	India	24.13	7.62

2. Processing the Data

As it is clearly visible that the dataset having missing value also and since it has a total of 218 rows and in that city_avg_temp_na attribute have 17 values missing which is 7.79 % of that attribute which is less and normally 25% to 30% missing values are accepted and It is know that the **city_avg_temp_na** attribute is

an important attribute and is of numeric type, so we will calculate the mean value and insert it in the place of missing values.

Formula for calculating mean:

=ROUND(AVERAGE(IF(ISNUMBER(D2:D219),D2:D219)),2)

Formula for replacing missing value with mean

=IF(ISNUMBER(D2),D2,K1) → here K1 is the cell where mean value is stored.

Given Below is the dataset after performing said operation.

	A	B	C	D	E	F
1	year	city	country	city_avg_temp_na	global_avg_temp	city_avg_temp
2	1796	New Delhi	India	25.03	8.27	25.03
3	1797	New Delhi	India	26.71	8.51	26.71
4	1798	New Delhi	India	24.29	8.67	24.29
5	1799	New Delhi	India	25.28	8.51	25.28
6	1800	New Delhi	India	25.21	8.48	25.21
7	1801	New Delhi	India	24.22	8.59	24.22
8	1802	New Delhi	India	25.63	8.58	25.63
9	1803	New Delhi	India	25.38	8.5	25.38
10	1804	New Delhi	India	25.68	8.84	25.68
11	1805	New Delhi	India	25.3	8.56	25.3
12	1806	New Delhi	India	25.22	8.43	25.22
13	1807	New Delhi	India	24.97	8.28	24.97
14	1808	New Delhi	India		7.63	25.17
15	1809	New Delhi	India		7.08	25.17
16	1810	New Delhi	India		6.92	25.17
17	1811	New Delhi	India		6.86	25.17
18	1812	New Delhi	India		7.05	25.17
19	1813	New Delhi	India	24.56	7.74	24.56
20	1814	New Delhi	India	23.73	7.59	23.73
21	1815	New Delhi	India	24.09	7.24	24.09
22	1816	New Delhi	India	23.7	6.94	23.7
23	1817	New Delhi	India	23.86	6.98	23.86
24	1818	New Delhi	India	24.37	7.83	24.37
25	1819	New Delhi	India	23.9	7.37	23.9
26	1820	New Delhi	India	24.13	7.63	24.13

Mean was calculated which came around to be 25.17, so all missing values were replaced with it and a new column city_avg_temp was created.

3. Calculating Moving Average

In this step moving average was calculated in order to smooth out the data, so moving average for a interval of 5 years was done after which the data set looks like this:

Formula for Calculating moving average: =AVERAGE(F2:F6) → for City Moving Average

=AVERAGE(E2:E6) → for Global Moving Average

	A	B	C	D	E	F	G	H	I	J	K	L
1	year	city	country	city_avg_temp_na	global_avg_temp	city_avg_temp	city_mov_avg_5	global_mov_avg_5		Average	25.17	
2	1796	New Delhi	India	25.03	8.27	25.03						
3	1797	New Delhi	India	26.71	8.51	26.71						
4	1798	New Delhi	India	24.29	8.67	24.29						
5	1799	New Delhi	India	25.28	8.51	25.28						
6	1800	New Delhi	India	25.21	8.48	25.21	25.304	8.488				
7	1801	New Delhi	India	24.22	8.59	24.22	25.142	8.552				
8	1802	New Delhi	India	25.63	8.58	25.63	24.926	8.566				
9	1803	New Delhi	India	25.38	8.5	25.38	25.144	8.532				
10	1804	New Delhi	India	25.68	8.84	25.68	25.224	8.598				
11	1805	New Delhi	India	25.3	8.56	25.3	25.242	8.614				
12	1806	New Delhi	India	25.22	8.43	25.22	25.442	8.582				
13	1807	New Delhi	India	24.97	8.28	24.97	25.31	8.522				
14	1808	New Delhi	India		7.63	25.17	25.268	8.348				
15	1809	New Delhi	India		7.08	25.17	25.166	7.996				
16	1810	New Delhi	India		6.92	25.17	25.14	7.668				
17	1811	New Delhi	India		6.86	25.17	25.13	7.354				
18	1812	New Delhi	India		7.05	25.17	25.17	7.108				
19	1813	New Delhi	India	24.56	7.74	24.56	25.048	7.13				
20	1814	New Delhi	India	23.73	7.59	23.73	24.76	7.232				
21	1815	New Delhi	India	24.09	7.24	24.09	24.544	7.296				
22	1816	New Delhi	India	23.7	6.94	23.7	24.25	7.312				
23	1817	New Delhi	India	23.86	6.98	23.86	23.988	7.298				
24	1818	New Delhi	India	24.37	7.83	24.37	23.95	7.316				
25	1819	New Delhi	India	23.9	7.37	23.9	23.984	7.272				
26	1820	New Delhi	India	24.13	7.63	24.13	23.99	7.248				

Also a Moving Average for a interval of 7 years was calculated to check which gives a smooth visualization

,the data set looks like this:

Formula for Calculating moving average: =AVERAGE(F2:F8) → for City Moving Average

=AVERAGE(E2:E8) → for Global Moving Average

|	A	B	C	D	E	F	G	H	I	J	K
1	year	city	country	city_avg_temp_na	global_avg_temp	city_avg_temp	city_mov_avg_7	global_mov_avg_7		Average	25.17
2	1796	New Delhi	India	25.03	8.27	25.03					
3	1797	New Delhi	India	26.71	8.51	26.71					
4	1798	New Delhi	India	24.29	8.67	24.29					
5	1799	New Delhi	India	25.28	8.51	25.28					
6	1800	New Delhi	India	25.21	8.48	25.21					
7	1801	New Delhi	India	24.22	8.59	24.22					
8	1802	New Delhi	India	25.63	8.58	25.63	25.19571429	8.515714286			
9	1803	New Delhi	India	25.38	8.5	25.38	25.24571429	8.548571429			
10	1804	New Delhi	India	25.68	8.84	25.68	25.09857143	8.595714286			
11	1805	New Delhi	India	25.3	8.56	25.3	25.24285714	8.58			
12	1806	New Delhi	India	25.22	8.43	25.22	25.23428571	8.568571429			
13	1807	New Delhi	India	24.97	8.28	24.97	25.2	8.54			
14	1808	New Delhi	India		7.63	25.17	25.33571429	8.402857143			
15	1809	New Delhi	India		7.08	25.17	25.27	8.188571429			
16	1810	New Delhi	India		6.92	25.17	25.24	7.962857143			
17	1811	New Delhi	India		6.86	25.17	25.16714286	7.68			
18	1812	New Delhi	India		7.05	25.17	25.14857143	7.464285714			
19	1813	New Delhi	India	24.56	7.74	24.56	25.05428571	7.365714286			
20	1814	New Delhi	India	23.73	7.59	23.73	24.87714286	7.267142857			
21	1815	New Delhi	India	24.09	7.24	24.09	24.72285714	7.211428571			
22	1816	New Delhi	India	23.7	6.94	23.7	24.51285714	7.191428571			
23	1817	New Delhi	India	23.86	6.98	23.86	24.32571429	7.2			
24	1818	New Delhi	India	24.37	7.83	24.37	24.21142857	7.338571429			
25	1819	New Delhi	India	23.9	7.37	23.9	24.03	7.384285714			

Key Consideration for visualization:

1). Plotting year on X-axis

2). Plotting Temperature in Y-axis

3). Calculate the moving average of city and global temperature to smooth the trend line.

4. Visualization of Data

In this step a Line graph was plotted in order to extract out the insights, given below is the line chart.

City vs Global Avg. Temp. 7 Year Moving Average
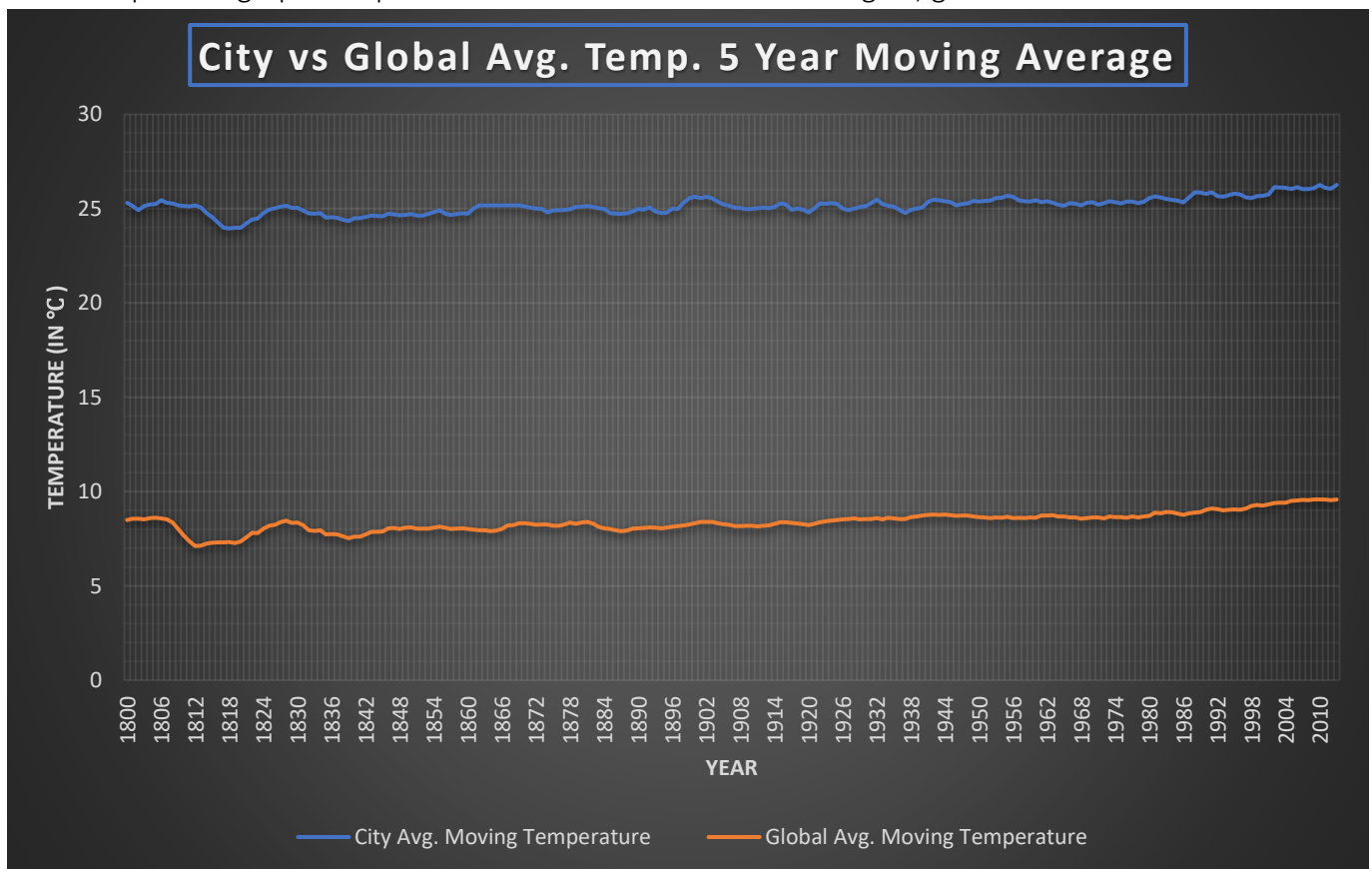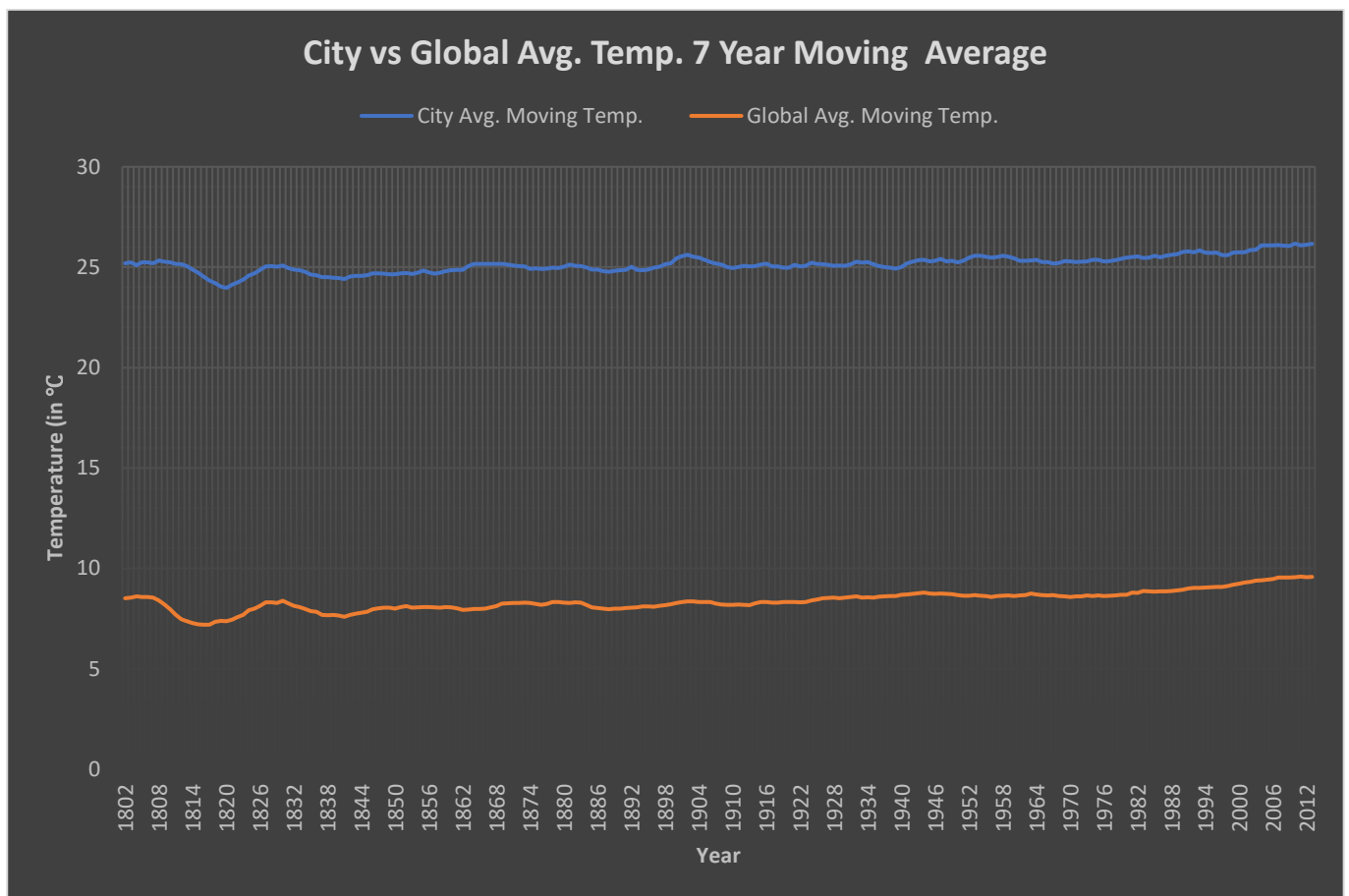
Here we can see that the 7 year moving average gives a smooth trend line. So for inferences we will use 7 Year moving average visualization. And also chart without moving average is plotted to see the difference and understand the graph better.

City Avg. Temp vs Global Avg. Temp.
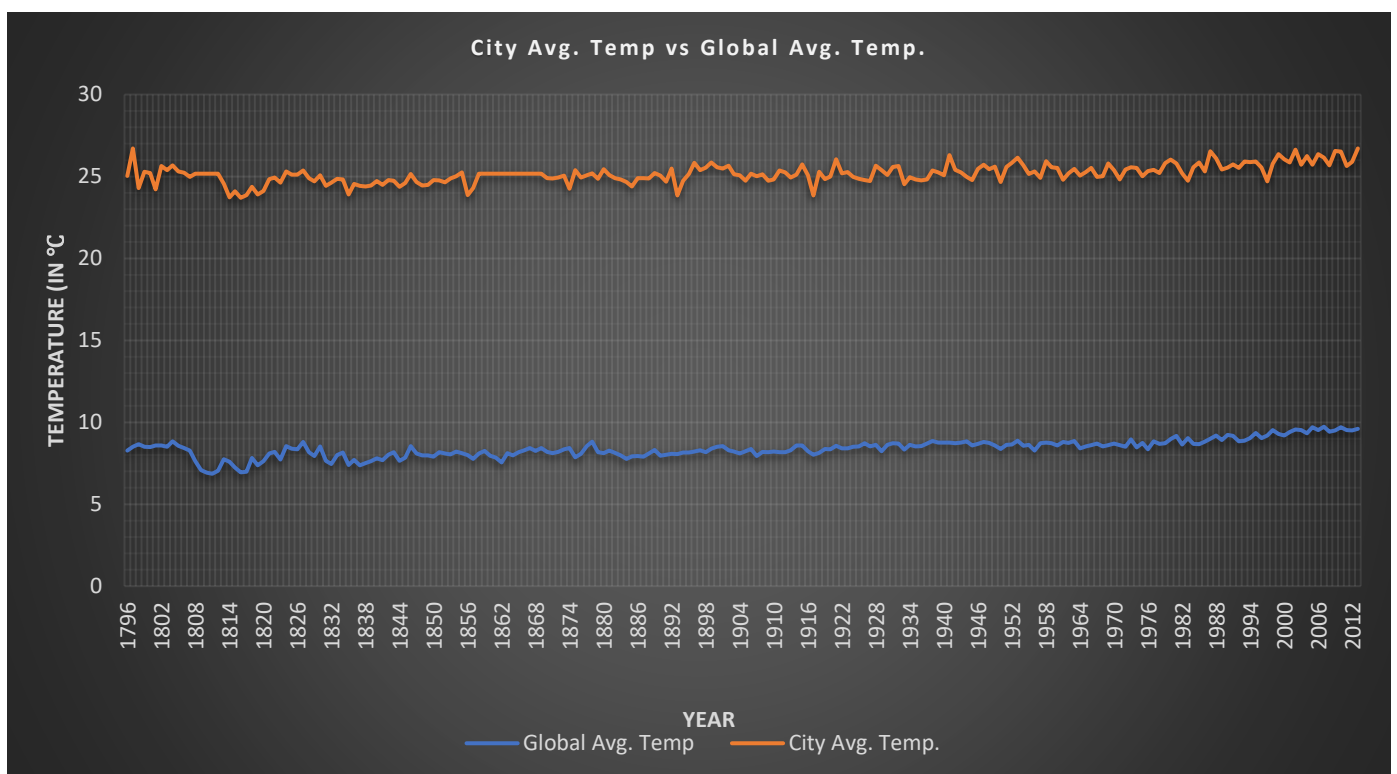
Here it can be seen that the plot is not smooth at all.

5. Observations

Question. Is your city hotter or cooler on average compared to the global average? Has the difference been consistent over time?

Answer: Comparing to the global average the city I live in is hotter than the global average and over the years the difference has somewhat been consistent.

Question. How do the changes in your city's temperatures over time compare to the changes in the global average?

Answer: Between 1802 to 1817 the global temperature saw a dip of 2°C whereas the temperature in my city saw a very little dip and between 1817 to 1820 when global temperature was rising and the temperature in my city was declining and after that the temperature has been fluctuating continuously in my city and is not as steady as global temperature.

Question. What does the overall trend look like? Is the world getting hotter or cooler? Has the trend been consistent over the last few hundred years?

Answer: If we will look at the overall trend then we can infer that the global temperature has been fluctuating less whereas the temperature in my city is fluctuating continuously and plummet during 1802-1803 when global temperature was steady and also during year 1900 and 1936-37 the temperature deviated whereas the global temperature is steady. Looking at the global and city temperature trend we can easily say that the world is getting hotter because the temperature increases by at least 2°C after hitting the lowest average. Looking at last few hundred years we can say that between 1850-1950 the trend in my city saw a fluctuation whereas the global temperature was steady.

Question. How the temperature in 2012 is relative to 1802 ?

Answer: Looking at the visualization it can be infer that both the global and city temperature has seen a increase of average of 1°C. And during the period of 1814-1817 the temperature were lowest for both city and global.