

Project Report: Explore Weather Trends

Steps

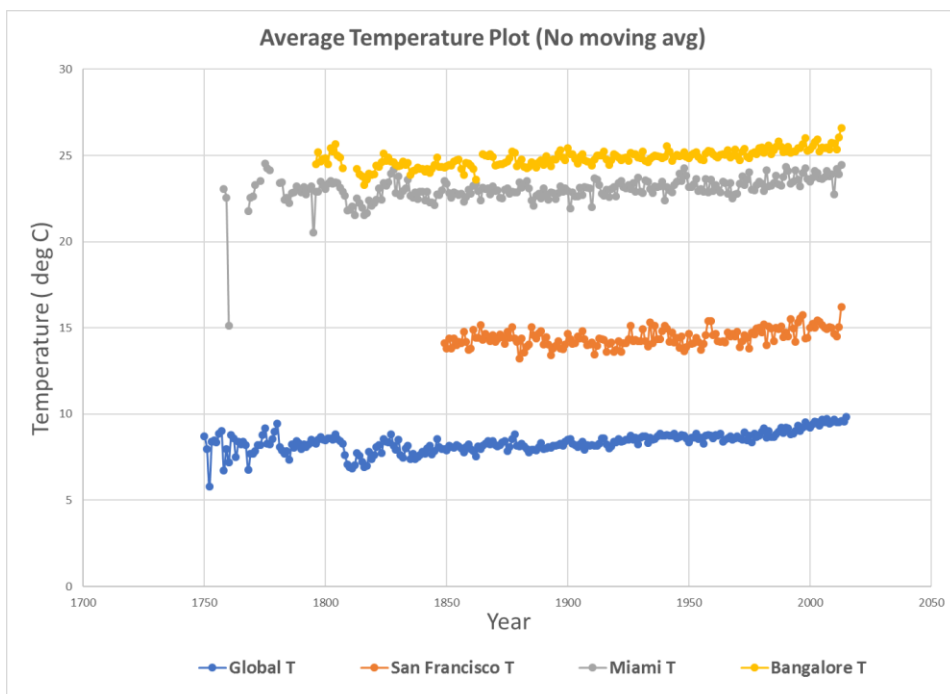
SQL Query to extract global temperature data:	<pre>SELECT * FROM global_data;</pre>
SQL Query to pull out cities in the US:	<pre>SELECT * FROM city_list ORDER BY country;</pre>

From the list of US cities, I picked San Francisco (biggest city closest to where I live) and Miami (because I'm interested in the analyzing the weather in Miami), and also picked Bangalore, a city in India.

SQL queries to extract city data:		
<pre>SELECT * FROM city_data WHERE city='San Francisco';</pre>	<pre>SELECT * FROM city_data WHERE city='Miami';</pre>	<pre>SELECT * FROM city_data WHERE city='Bangalore';</pre>

The data extracted using the queries was exported to Excel files.

Analysis in Excel: A graph of avg temp vs year was plotted to get an idea about the data:



This graph shows that the temperature data for some years is missing, and calculating the moving average may help handle missing data and help observe the trend better

Moving average for 5 year and 10 year periods were calculated. I picked the 10 year moving average for the comparison plot since the 5 year moving average had points where division by zero was encountered due to missing data. It is assumed that the 10 year moving average is reasonable imputation to fill in the gaps in data.

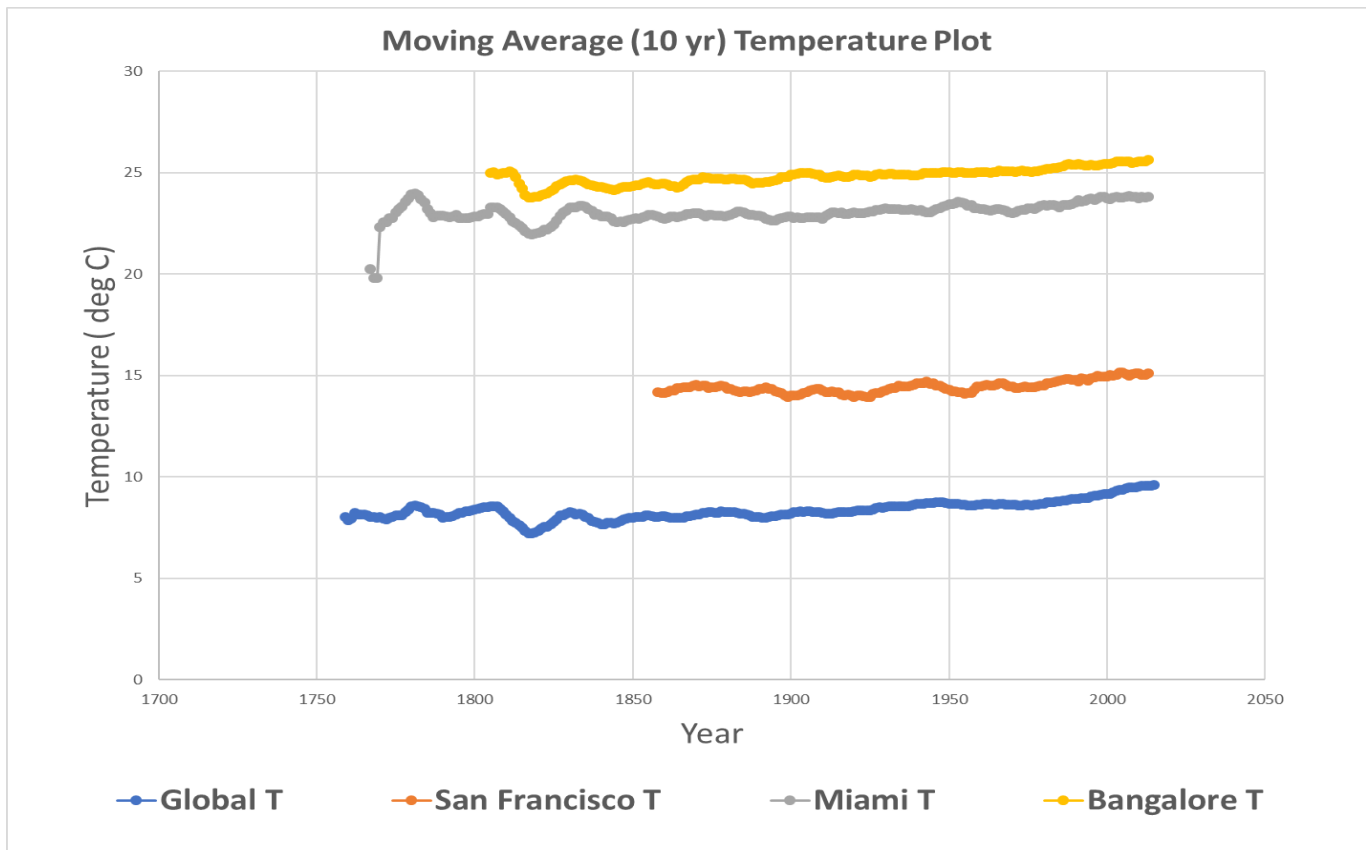
	GLOBAL				SAN FRANCISCO				MIAMI				BANGALORE			
1	year	avg_temp	5 yr MA	10 yr MA	year	avg_temp	5 yr MA	10 yr MA	year	avg_temp	5 yr MA	10 yr MA	year	avg_temp	5 yr MA	10 yr MA
2																
3	1750	8.72			1849	14.12			1758	23.05			1796	24.49		
4	1751	7.98			1850	13.8			1759	22.56			1797	25.18		
5	1752	5.78			1851	14.39			1760	15.14			1798	24.65		
6	1753	8.39			1852	13.81			1761				1799	24.81		
7	1754	8.47	7.868		1853	14.4	14.104		1762	20.25			1800	24.85	24.796	
8	1755	8.36	7.796		1854	13.98	14.076		1763	18.85			1801	24.49	24.796	
9	1756	8.85	7.97		1855	14.2	14.156		1764	15.14			1802	25.44	24.848	
10	1757	9.02	8.618		1856	14.1	14.098		1765	#DIV/0!			1803	25.22	24.962	
11	1758	6.74	8.288		1857	14.78	14.292		1766	#DIV/0!			1804	25.67	25.134	
12	1759	7.99	8.192	8.03	1858	14.19	14.177		1767	#DIV/0!	20.25		1805	25.01	25.166	24.981
13	1760	7.19	7.958	7.877	1859	13.71	14.196	14.136	1768	21.77	21.77	19.82333	1806	24.87	25.242	25.019
14	1761	8.77	7.942	7.956	1860	13.81	14.118	14.137	1769	22.57	22.17	19.82667	1807	24.25	25.004	24.926
15	1762	8.61	7.86	8.239	1861	14.88	14.274	14.186	1770	22.64	22.32667	22.32667	1808		24.95	24.95667
16	1763	7.5	8.012	8.15	1862	14.43	14.204	14.248	1771	23.28	22.565	22.565	1809		24.71	24.975
17	1764	8.4	8.094	8.143	1863	14.43	14.252	14.251	1772		22.565	22.565	1810		24.56	24.99286
18	1765	8.25	8.306	8.132	1864	15.18	14.546	14.371	1773	23.52	23.0025	22.756	1811		24.25	25.07667
19	1766	8.41	8.234	8.088	1865	14.32	14.648	14.383	1774		23.14667	22.756	1812		#DIV/0!	25.004
20	1767	8.22	8.156	8.008	1866	14.67	14.606	14.44	1775	24.52	23.77333	23.05	1813	24.23	24.23	24.806
21	1768	6.78	8.012	8.012	1867	14.46	14.612	14.408	1776	24.26	24.1	23.22286	1814	23.91	24.07	24.454
22	1769	7.69	7.87	7.982	1868	14.25	14.576	14.414	1777	24.16	24.115	23.34	1815	23.79	23.97667	24.21
23	1770	7.69	7.758	8.032	1869	14.57	14.454	14.5	1778		24.31333	23.56429	1816	23.3	23.8075	23.896
24	1771	7.85	7.646	7.94	1870	14.19	14.428	14.538	1779		24.31333	23.73	1817	23.6	23.766	23.766
25	1772	8.19	7.64	7.898	1871	14.34	14.362	14.484	1780		24.21	23.948	1818	23.94	23.708	23.795
26	1773	8.22	7.928	7.97	1872	14.63	14.396	14.504	1781	23.39	23.775	23.97	1819	23.86	23.698	23.80429
27	1774	8.77	8.144	8.007	1873	14.46	14.438	14.507	1782	23.45	23.43	23.88333	1820	23.91	23.722	23.8175
28	1775	9.18	8.442	8.1	1874	14.09	14.342	14.398	1783	22.44	23.09333	23.70333	1821	24.4	23.94	23.88222
29	1776	8.3	8.532	8.089	1875	14.76	14.456	14.442	1784	22.57	22.9625	23.54143	1822	24.33	24.088	23.927
30	1777	8.26	8.546	8.093	1876	14.44	14.476	14.419	1785	22.24	22.818	23.21571	1823	24.62	24.224	23.966
31	1778	8.54	8.61	8.269	1877	15.03	14.556	14.476	1786	22.81	22.702	23.00857	1824	25.1	24.472	24.085
32	1779	8.98	8.652	8.398	1878	14.37	14.538	14.488	1787	22.86	22.584	22.82286	1825	24.69	24.628	24.175
33	1780	9.43	8.702	8.572	1879	14.7	14.56	14.451	1788	23.21	22.738	22.87125	1826	24.88	24.724	24.333
34					1880											

A few screenshots showing the calculation of the moving average from the spreadsheet:

	F	G	H	I	J
1					
2	year	avg_temp	5 yr MA	10 yr MA	
3	1849	14.12			
4	1850	13.8			
5	1851	14.39			
6	1852	13.81			
7	1853	14.4	14.104		
8	1854	13.98	14.076		
9	1855	14.2	14.156		
10	1856	14.1	14.098		
11	1857	14.78	14.292		
12	1858	14.19	14.25	14.177	
13	1859	13.71	14.196	14.136	
14	1860	13.81	14.118	14.137	
15	1861	14.88	14.274	14.186	
16	1862	14.43	14.204	14.248	
17	1863	14.43	14.252	14.251	
18	1864	15.18	14.546	14.371	
19	1865	14.32	14.648	14.383	
20	1866	14.67	14.606	14.44	

	A	B	C	D	E	F	G	H	I
1	GLOBAL	GLOBAL				SAN FRANCISCO			
2	year	avg_temp	5 yr MA	10 yr MA		year	avg_temp	5 yr MA	10 yr MA
3	1750	8.72				1849	14.12		
4	1751	7.98				1850	13.8		
5	1752	5.78				1851	14.39		
6	1753	8.39				1852	13.81		
7	1754	8.47	7.868			1853	14.4	14.104	
8	1755	8.36	7.796			1854	13.98	14.076	
9	1756	8.85	7.97			1855	14.2	14.156	
10	1757	9.02	8.618			1856	14.1	14.098	
11	1758	6.74	8.288			1857	14.78	14.292	
12	1759	7.99	8.192	8.03		1858	14.19	14.25	14.177
13	1760	7.19	7.958	7.877		1859	13.71	14.196	14.136
14	1761	8.77	7.942	7.956		1860	13.81	14.118	14.137
15	1762	8.61	7.86	8.239		1861	14.88	14.274	14.186
16	1763	7.5	8.012	8.15		1862	14.43	14.204	14.248
17	1764	8.4	8.094	8.143		1863	14.43	14.252	14.251

A line plot showing the 10 yr moving average temperatures vs year



Observations:

- The chosen cities (San Francisco, Miami and Bangalore) are hotter compared to the global average. In order of decreasing heat, Bangalore > Miami > San Francisco > global temperature.
- Bangalore and Miami are pretty close to each other with Bangalore being only slightly hotter than Miami, consistently over the years. San Francisco is significantly cooler than Bangalore and Miami.
- Over time, the average local temperatures in San Francisco , Miami and Bangalore have gradually increased and so has the global average temperature showing that the world has been getting hotter over the years.
- This trend has been consistent over the last few hundred years and there are no steep increases. Before 1850, the fluctuations in global temperature have been more. Around 2000, the global temperature increase is more evident while 1850 – 2000 has not seen too much variation.

Correlation Coefficient

The image shows an Excel spreadsheet with the following data:

SAN FRANCISCO		GLOBAL	
year	avg_temp	year	avg_temp
1849	14.12	1849	7.98
1850	13.8	1850	7.9
1851	14.39	1851	8.18
1852	13.81	1852	8.1
1853	14.4	1853	8.04
1854	13.98	1854	8.21
1855	14.2	1855	8.11
1856	14.1	1856	8
1857	14.78	1857	7.76
1858	14.19	1858	8.1
1859	13.71	1859	8.25
1860	13.81	1860	7.96
1861	14.88	1861	7.85
1862	14.43	1862	7.56
1863	14.43	1863	8.11
1864	15.18	1864	7.98
1865	14.32	1865	8.18
1866	14.67	1866	8.29
1867	14.46	1867	8.44
1868	14.25	1868	8.25
1869	14.57	1869	8.43
1870	14.19	1870	8.2
1871	14.34	1871	8.12
1872	14.63	1872	8.19
1873	14.46	1873	8.35
1874	14.09	1874	8.43
1875	14.76	1875	7.86
1876	14.44	1876	8.08
1877	15.03	1877	8.54
1878	14.37	1878	8.83
1879	14.2	1879	8.17
1880	13.22	1880	8.12
1881	14.39	1881	8.27
1882	13.58	1882	8.12

The formula bar shows: `=CORREL(B3:B167,E3:E167)`

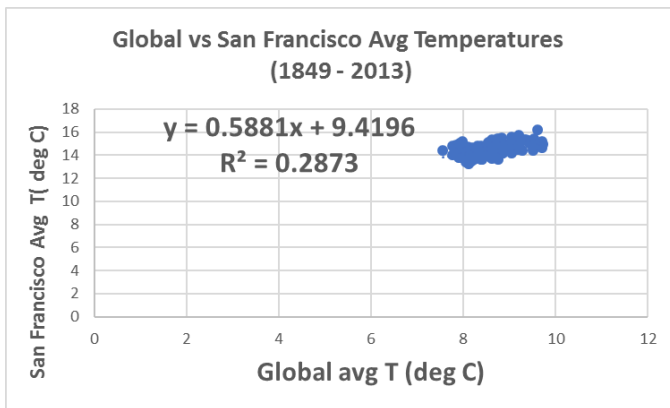
The result in cell H6 is: 0.536038

Taking the avg temperature data for San Francisco, between 1849 – 2013, we have global and local data available for those years.

Using the CORREL () and PEARSON() functions in Excel, the correlation coefficient was found to be **0.53**.

The two variables were: Average global temperature and Average local temperature in San Francisco.

Predicting City temperature based on Global temperature



Using this equation where y is the local avg temperature, and x is the global avg temperature, the city temperature can be predicted based on the global temperature.

$$y = 0.5881x + 9.4196$$

The low R^2 value suggests that the fit is not great. A linear relationship does not relate the data very well.