

# **Project: Exploring Weather Trends**

By: Abhinav Jha

## 1. What tools did you use for each step? (Python, SQL, Excel, etc)

a. I used two SQL queries to extract my data from the database:

i. Global Data:

```
SELECT *
FROM global_data
```

ii. Local Data: (for San Jose, CA)

```
SELECT year, city, avg_temp
FROM city_data
WHERE city = 'San Jose'
```

b. I exported the data to CSV files using [Download CSV](#) the link

c. I used MS Excel to convert the CSV file to XSLX file, calculate the moving average and create the charts.

## 2. How did you calculate the moving average?

I tried 7, 10, 20-year moving averages to see which average is better to smooth out data. To calculate the moving average in MS Excel, I used the AVERAGE function (the same approach as in the lesson) as shown below:

	A	B	C	D
1	year	city	Loc_avg_temp	Loc_7Mov_avg
2	1849	San Jose	14.12	
3	1850	San Jose	13.8	
4	1851	San Jose	14.39	
5	1852	San Jose	13.81	
6	1853	San Jose	14.4	
7	1854	San Jose	13.98	
8	1855	San Jose	14.2	=AVERAGE(C2:C8)
9	1856	San Jose	14.1	AVERAGE(number1,

Figure 1: 7-year Moving Average

	A	B	C	D	E
1	year	city	Loc_avg_temp	Loc_7Mov_avg	Loc_10Mov_avg
2	1849	San Jose	14.12		
3	1850	San Jose	13.8		
4	1851	San Jose	14.39		
5	1852	San Jose	13.81		
6	1853	San Jose	14.4		
7	1854	San Jose	13.98		
8	1855	San Jose	14.2	14.10	
9	1856	San Jose	14.1	14.10	
10	1857	San Jose	14.78	14.24	
11	1858	San Jose	14.19	14.21	=AVERAGE(C2:C11)
12	1859	San Jose	13.71	14.19	AVERAGE(number1, num
13	1860	San Jose	13.81	14.11	14.14
14	1861	San Jose	14.88	14.24	14.19

Figure 2: 10-year Moving Average

1	year	city	Loc_avg_temp	Loc_7Mov_avg	Loc_10Mov_avg	Loc_20Mov_avg	Loc
2	1849	San Jose	14.12				
3	1850	San Jose	13.8				
4	1851	San Jose	14.39				
5	1852	San Jose	13.81				
6	1853	San Jose	14.4				
7	1854	San Jose	13.98				
8	1855	San Jose	14.2	14.10			
9	1856	San Jose	14.1	14.10			
10	1857	San Jose	14.78	14.24			
11	1858	San Jose	14.19	14.21	14.18		
12	1859	San Jose	13.71	14.19	14.14		
13	1860	San Jose	13.81	14.11	14.14		
14	1861	San Jose	14.88	14.24	14.19		
15	1862	San Jose	14.43	14.27	14.25		
16	1863	San Jose	14.43	14.32	14.25		
17	1864	San Jose	15.18	14.38	14.37		
18	1865	San Jose	14.32	14.39	14.38		
19	1866	San Jose	14.67	14.53	14.44		
20	1867	San Jose	14.46	14.62	14.41		
21	1868	San Jose	14.25	14.53	14.41	=AVERAGE(C2:C21)	
22	1869	San Jose	14.57	14.55	14.50	=AVERAGE(number1, [range])	

Figure 3: 20-year Moving Average

### 3. What were your key considerations when deciding how to visualize the trends?

The key consideration was to determine the timeframe for data visualization; Looking at the local temperature data for San Jose, the data covers the period between **1849 to 2013**, where in the global temperature data covers the period between **1750 and 2015**. Therefore, the analysis was performed for the range between **1849 to 2013**. To make sure local and global temperature data is mapped correctly, I used VLOOKUP to retrieve the global temperature data worksheet into the local data worksheet.

Another consideration was to adjust the starting point for each chart as follows:

- 7-year moving average starting point: 1855 (1849 + 7) See figure 6
- 10-year moving average starting point: 1858 (1849 + 10) See figure 5
- 20-year moving average starting point: 1868 (1849 + 20) See figure 4

To help assess the data variance and frequency of change between global and local temperature levels, I calculated the following:

- The Global & Local annual change percentage:

1	year	Loc_avg_temp	Loc_inc_%	Glo_avg
2	1849	14.12	0.000%	
3	1850	13.8	=IFERROR((B3-B2)/B2,0)	
4	1851	14.39	IFERROR(value, value_if_error)	
5	1852	13.81	-4.031%	

G	K	L
Glo_avg_temp	Glo_inc_%	Differenc
7.98	0.000%	
7.9	=IFERROR((G3-G2)/G2,0)	
8.18	IFERROR(value, value_if_error)	

- The Local/Global temp. average difference:

1	year	Loc_avg_temp	Glo_avg_temp	Difference
2	1849	14.12	7.98	=B2-G2
3	1850	13.8	7.9	5.90
4	1851	14.39	8.18	6.21
5	1852	13.81	8.1	5.71
6	1853	14.4	8.04	6.36
7	1854	13.98	8.71	5.27

Also, I used Pivot table to calculate the Max, Min, Average, Standard Deviation, High/Low (%) change as follows:

3	Average of Loc_avg_temp	Average of Glo_avg_temp	14	Average of Loc_Inc_prct	Average of Glo_Inc_prct
4	14.45	8.55	15	0.152%	0.146%
5			16		
6	Min of Loc_avg_temp	Min of Glo_avg_temp	17		
7	13.22	7.56	18	Max of Loc_Inc_prct	Max of Glo_Inc_prct
8			19	9.907%	7.275%
9	Max of Loc_avg_temp	Max of Glo_avg_temp	20		
10	16.23	9.73	21	Min of Loc_Inc_prct	Min of Glo_Inc_prct
11			22	-8.698%	-7.475%
12	StdDev of Loc_avg_temp	StdDev of Glo_avg_temp	23		
13	0.505	0.460	24	Max of Difference	Min of Difference2
			25	7.20	4.86

All the calculations above were summarized in table 1 and table 2 below

### Observations:

- The San Jose is hotter than the global temperature (please refer to Min, Max and Avg. columns in the table below)
- The local (San Jose) and global temperature levels are both increasing.
- The global moving average experiences less fluctuations than the local moving average in San Jose.
- The global temperature levels have a smaller variance than the local temperature changes.
- To determine the slope, we used the Linear TREND function for the local and global temperature data, we got the following the following equations:
  - o Local temperature:  $y = 0.0049x + 14.041$
  - o Global temperature:  $y = 0.0033x + 7.8644$

By comparing the two slopes (Slope 1 = 0.0049) & (Slope 2 = 0.0033), we note the local trend is increasing more rapidly than the global trend.

- The highest difference between local and global temperature is 7.20 °. This was recorded in year 1864; where the lowest difference between local and global temperature is 4.84 °. this was recorded in year 1998 (see figure 7)

	Min	Max	Avg.	SD	highest Inc. (%)	Lowest Dec. (%)	Avg. Change
San Jose	13.22 °	16.23 °	14.5 °	0.505	9.907 %	- 8.698 %	0.152 %
Global	7.56 °	9.73 °	8.55 °	0.460	7.275%	- 7.475 %	0.146 %

Table 1: Global Vs. Local (Summary 1)

Highest Difference	Lowest Difference
7.20 °	4.86 °

Table 2: Highest & Lowest Average Difference

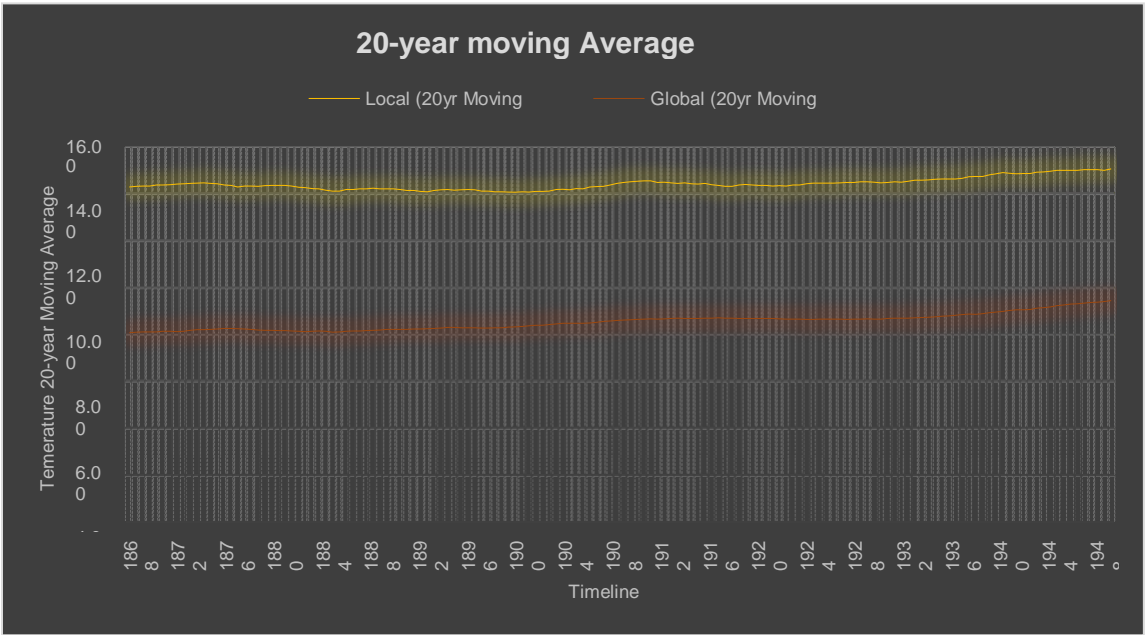


Figure 4: 20-year Moving Average Comparison

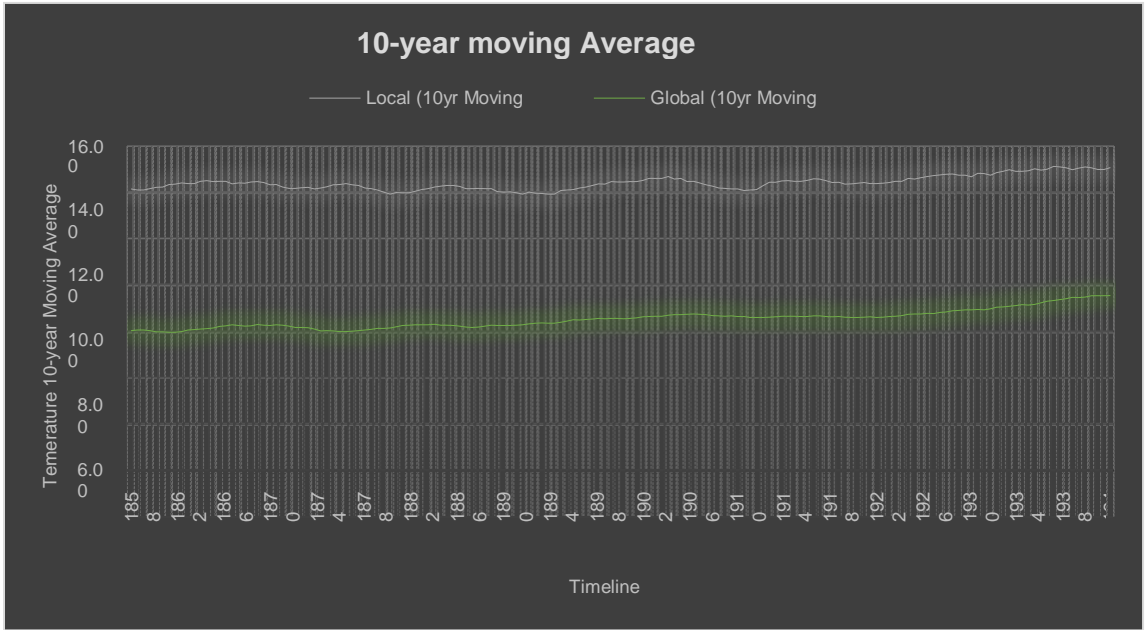


Figure 5: 10-year Moving Average Comparison

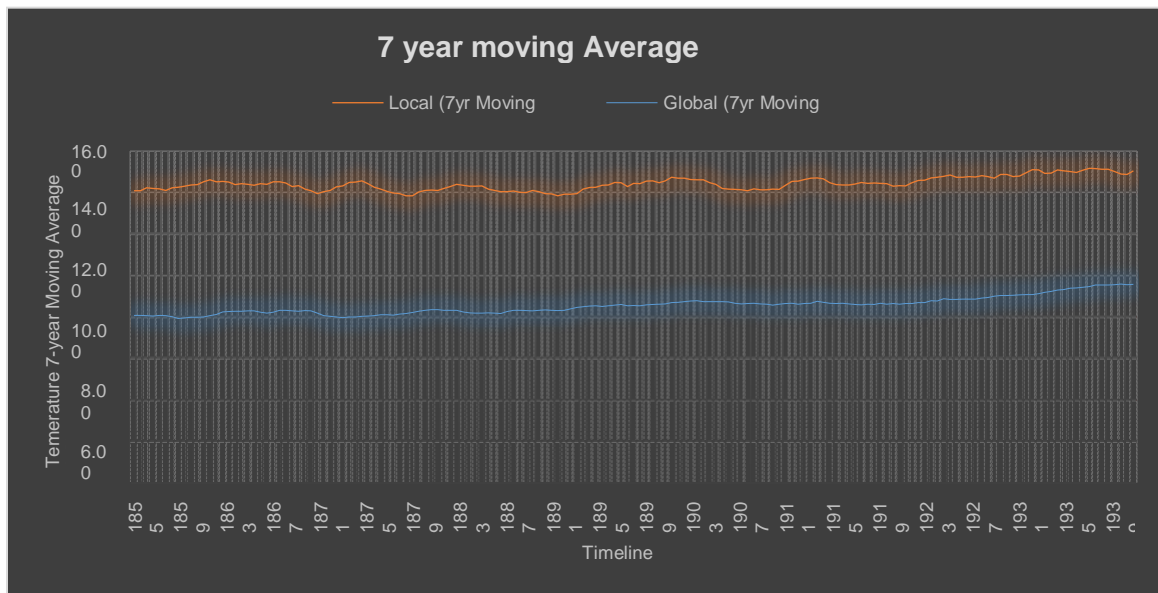


Figure 6: 7-year Moving Average Comparison

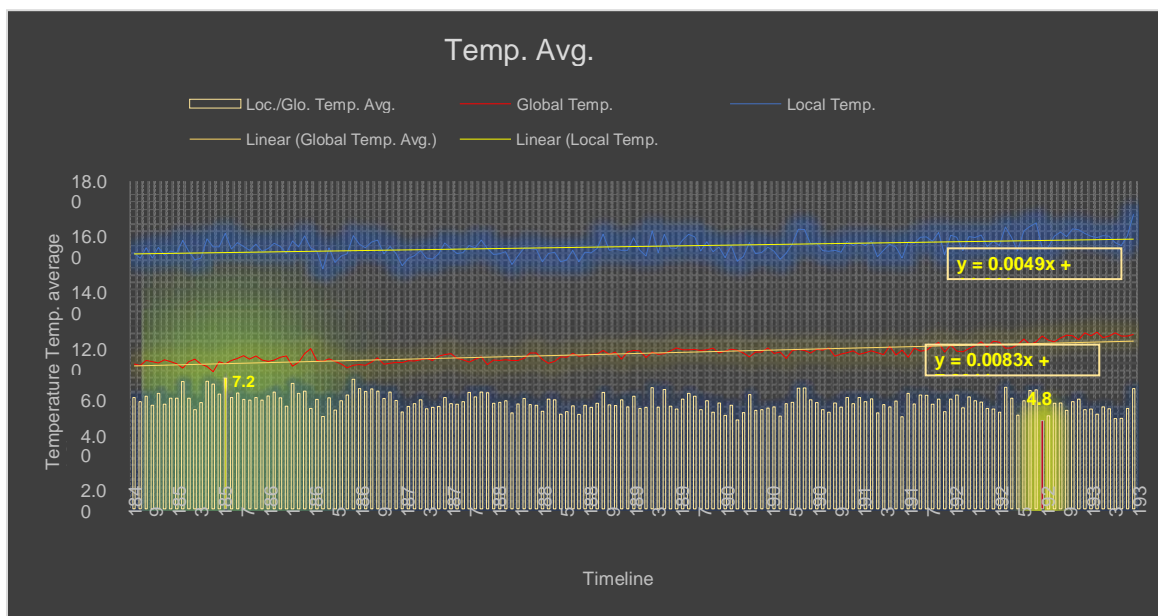


Figure 7: Temp. Average Comparison & Maximum Difference