

## Project: Explore Weather Trends

### Outline:

The scope of this project is to analyse similarities and differences between global temperature trends and temperature trends of a San Jose city.

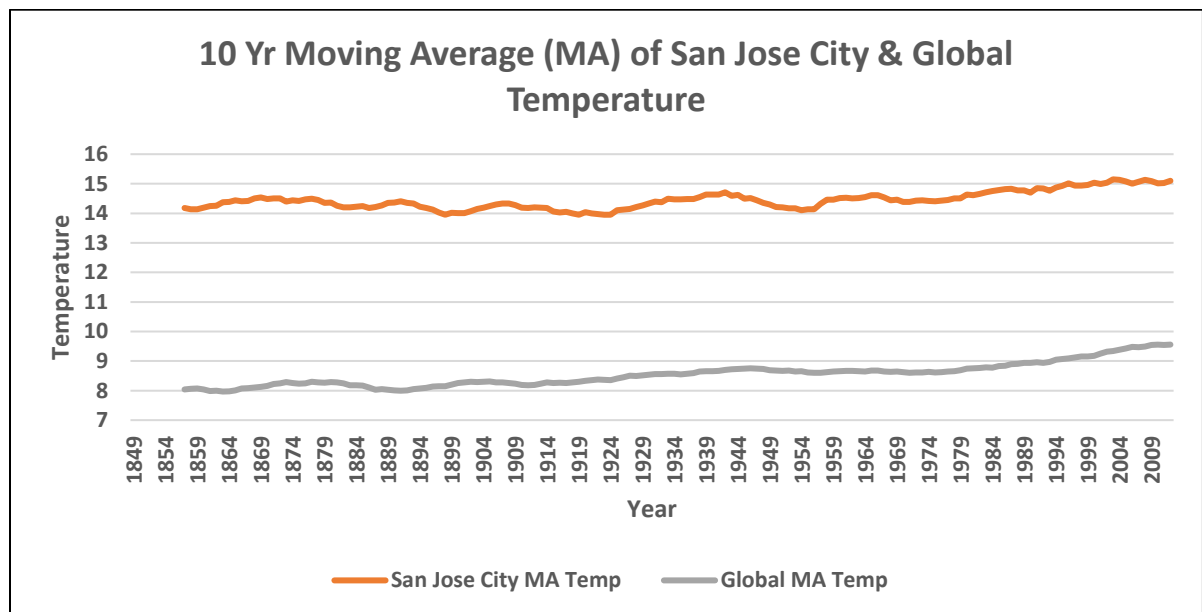
### Process to extract and analyze data:

1. The data was extracted from the database using SQL queries, exported to csv format and analysed using line charts in excel. *(Refer Annexure 1 for SQL Queries)*
2. Since the raw data is spread across 160 years, excel was used to calculate 10 year moving average to smooth the data to draw observations listed below. *(Refer Annexure 2 for 10 year moving average calculation)*
3. Excel was used to find correlation between the two weather trends. *(Refer Annexure 2 for correlation calculation in excel)*
4. Additionally, temperature data of Los Angeles city were added to the above data to understand the trend.

### Observations:

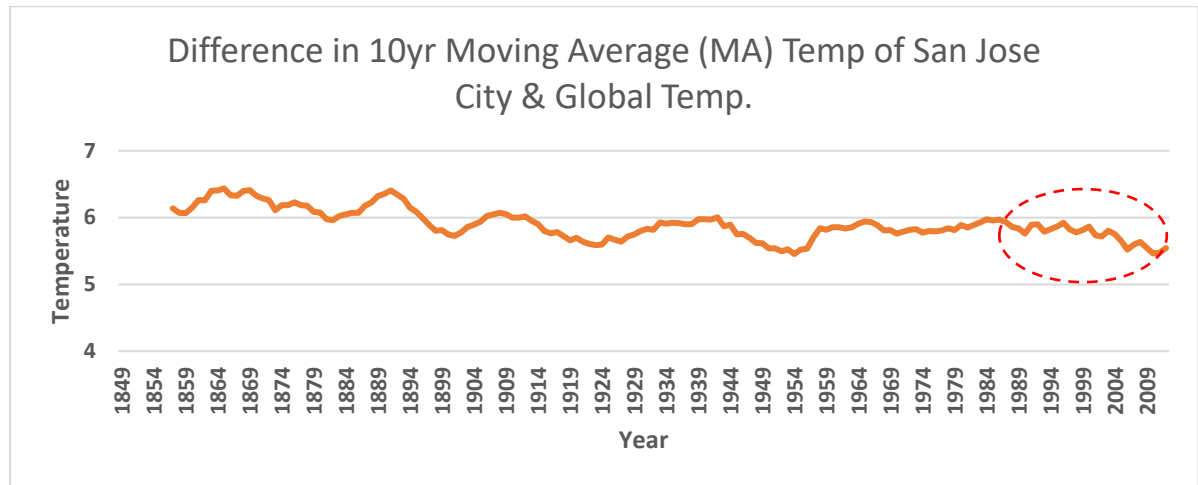
1. The city of San Jose appears hotter than the global average temperature as shown by the chart 1.1.

Chart: 1.1



2. It is interesting to note that the difference in the 10year moving average between city and global temperature tends to decrease after 1980s due to a greater increase in the global average temperature relative to San Jose (refer chart 1.2). This clearly shows that the difference is not consistent over time.

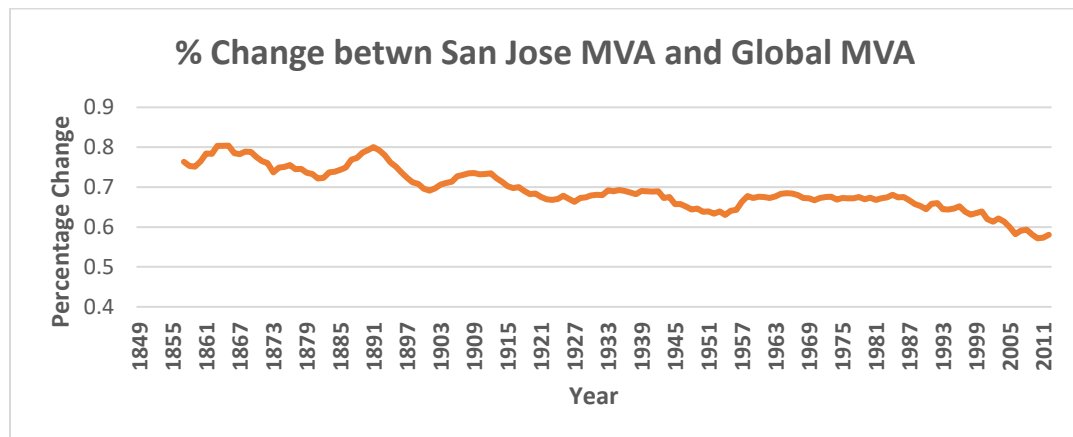
Chart 1.2



Additionally, the correlation between the San Jose city and global temperatures is 0.536 (calculated in excel using formula based on average temperatures and not moving average) implying a moderate uphill (positive) relationship between the two. In other words, a moderate uphill (positive) relationship means that if San Jose city average temperature rises so does the Global average temperature and vice a versa.

- As per the 10 year moving average chart 1.1, San Jose city temperature seems to have more fluctuations. Also, the percentage change between the global average temperature and San Jose city tends to decrease from 1980s to 2011 implying that the global average temperature has been rising (chart 1.3).

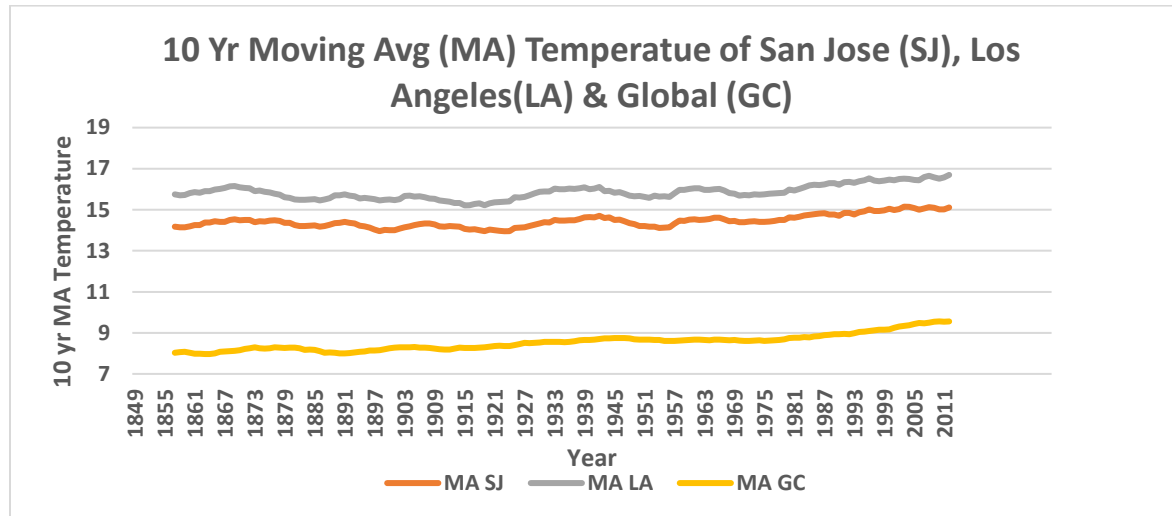
Chart: 1.3



- Looking at charts 1.1, 1.2 & 1.3 it can be concluded that global average temperatures show an upward trend and the world is getting hotter. This upward trend is observed over a data of 160 years.

5. Comparing two cities with global data: I have added the average temperatures for the city of Los Angeles (LA) to the above data to see if I can learn more about temperature trends.

Chart: 1.4



The above chart shows that LA city is hotter than San Jose and the global average temperature. Also, the correlation between average temperature of Los Angeles and San Jose is 0.8818, implying a strong positive relationship indicating that if one increases the other also increases. The global average temperature shares approximately the same correlation of 0.503 with the average temperatures of LA city portraying a moderate uphill relationship.

Note: Since the yearly average temperatures for San Jose city and San Francisco are same in the database (refer Annexure 2), San Francisco city data has not been used for the analysis.

## **Annexure 1 : SQL Queries used for extracting data**

- (1) To choose a city and country from database:

```
SELECT *  
FROM city_list
```

- (2) To view what the data looks like in city\_data table for city San Jose in United States

```
SELECT *  
FROM city_data  
WHERE city = 'San Jose' and country = 'United States'
```

- (3) SQL query to extract temperature data for city of San Jose, San Francisco, Los Angeles and Global average temperature.

```
SELECT  
cd1.year,  
cd1.avg_temp AS avg_temp_la,  
cd2.avg_temp AS avg_temp_sf,  
gd.avg_temp AS avg_temp_gd  
FROM city_data AS cd1, city_data AS cd2, global_data AS gd  
WHERE  
(cd1.year = cd2.year) AND (cd2.year = gd.year) AND  
(cd1.country = 'United States' AND cd1.city = 'Los Angeles') AND  
(cd2.country = 'United States' AND cd2.city = 'San Francisco')  
  
ORDER BY year;
```

## Annexure 2 : 10 year moving average calculation and correlation calculation

- Column B shows average temp of San Jose (SJ) City and column F shown the formula for 10 year moving average for San Jose city (MA SJ)

	A	B	C	D	E	F	G	H	I
1	year	avg_temp_SJ	avg_temp_la	avg_temp_sf	avg_temp_gd	MA SJ	MA LA	MA SF	MA GC
2	1849	14.12	15.71	14.12	7.98				
3	1850	13.8	15.28	13.8	7.9				
4	1851	14.39	15.53	14.39	8.18				
5	1852	13.81	15.61	13.81	8.1				
6	1853	14.4	16.27	14.4	8.04				
7	1854	13.98	15.74	13.98	8.21				
8	1855	14.2	15.94	14.2	8.11				
9	1856	14.1	15.52	14.1	8				
10	1857	14.78	16.19	14.78	7.76				
11	1858	14.19	15.67	14.19	8.1	=AVERAGE(B2:B11)			8.04
12	1859	13.71	15.29	13.71	8.25	14.14	15.70	14.14	8.07
13	1860	13.81	15.41	13.81	7.96	14.14	15.72	14.14	8.07
14	1861	14.88	16.51	14.88	7.85	14.19	15.82	14.19	8.04
15	1862	14.43	16.05	14.43	7.56	14.25	15.86	14.25	7.98

- Column B shows average temp of San Jose (SJ) City and column F shown the formula for next 10 year moving average for San Jose city (MA SJ)

	A	B	C	D	E	F	G	H	I
1	year	avg_temp_SJ	avg_temp_la	avg_temp_sf	avg_temp_gd	MA SJ	MA LA	MA SF	MA GC
2	1849	14.12	15.71	14.12	7.98				
3	1850	13.8	15.28	13.8	7.9				
4	1851	14.39	15.53	14.39	8.18				
5	1852	13.81	15.61	13.81	8.1				
6	1853	14.4	16.27	14.4	8.04				
7	1854	13.98	15.74	13.98	8.21				
8	1855	14.2	15.94	14.2	8.11				
9	1856	14.1	15.52	14.1	8				
10	1857	14.78	16.19	14.78	7.76				
11	1858	14.19	15.67	14.19	8.1	14.18	15.75	14.18	8.04
12	1859	13.71	15.29	13.71	8.25	=AVERAGE(B3:B12)			8.07
13	1860	13.81	15.41	13.81	7.96	14.14	15.72	14.14	8.07
14	1861	14.88	16.51	14.88	7.85	14.19	15.82	14.19	8.04
15	1862	14.43	16.05	14.43	7.56	14.25	15.86	14.25	7.98

Using the same formula 10 year moving averages have been calculated for Los Angeles, San Francisco and Global Temperatures

- Correlation calculations using excel formula (only Average Temperature data is used not Moving Average Temperature data)

	A	B	C	D	E	F	G	H	I	J	K
1	year	avg_temp_SJ	avg_temp_la	avg_temp_sf	avg_temp_gd	MA SJ	MA LA	MA SF	MA GC		
2	1849	14.12	15.71	14.12	7.98						
3	1850	13.8	15.28	13.8	7.9						
4	1851	14.39	15.53	14.39	8.18						
5	1852	13.81	15.61	13.81	8.1						
6	1853	14.4	16.27	14.4	8.04						
7	1854	13.98	15.74	13.98	8.21						
8	1855	14.2	15.94	14.2	8.11						
9	1856	14.1	15.52	14.1	8						
10	1857	14.78	16.19	14.78	7.76						
11	1858	14.19	15.67	14.19	8.1	14.18	15.75	14.18	8.04		
12	1859	13.71	15.29	13.71	8.25	14.14	15.70	14.14	8.07		
13	1860	13.81	15.41	13.81	7.96	14.14	15.72	14.14	8.07		
14	1861	14.88	16.51	14.88	7.85	14.19	15.82	14.19	8.04		

Correl SJ & LA 0.881812 =CORREL(B2:B166, C2:C166)  
 Correl LA & GC 0.503752 =CORREL(C2:C166, E2:E166)  
 Correl SJ & GC 0.536038 =CORREL(B2:B166, E2:E166)