

Amit Kumar Shankar
Project 1

SQL query to extract data

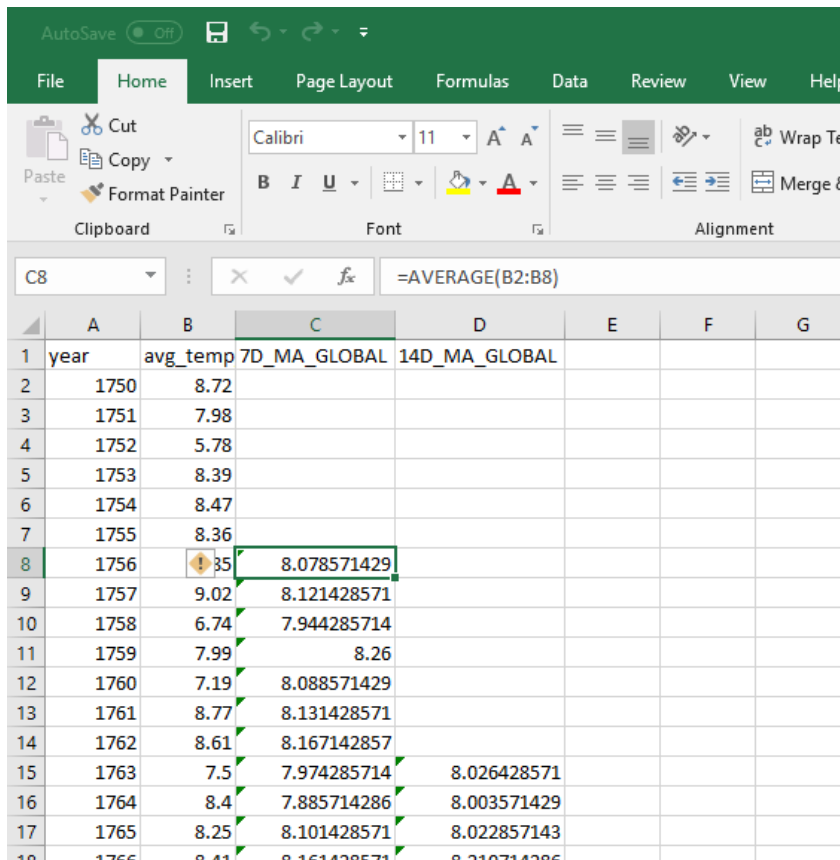
```
SELECT *  
FROM city_list  
WHERE country in ('United States')  
--nearest city is San Francisco
```

```
SELECT year, avg_temp  
FROM city_data  
WHERE country in ('United States') AND city in ('San Francisco')
```

```
SELECT *  
FROM global_data
```

Excel Documentation

Save the two csv files on desktop and opened both the files in excel. Created a new sheet in a new excel file and copied and pasted both the csv files in on sheet and created the line graphs. The next two images show the excel formulas used to calculate the 7 day and the 14 day moving averages.



	A	B	C	D	E	F	G
1	year	avg_temp	7D_MA_GLOBAL	14D_MA_GLOBAL			
2	1750	8.72					
3	1751	7.98					
4	1752	5.78					
5	1753	8.39					
6	1754	8.47					
7	1755	8.36					
8	1756	8.35	8.078571429				
9	1757	9.02	8.121428571				
10	1758	6.74	7.944285714				
11	1759	7.99	8.26				
12	1760	7.19	8.088571429				
13	1761	8.77	8.131428571				
14	1762	8.61	8.167142857				
15	1763	7.5	7.974285714	8.026428571			
16	1764	8.4	7.885714286	8.003571429			
17	1765	8.25	8.101428571	8.022857143			
18	1766	8.41	8.161428571	8.110714286			

AutoSave Off

File Home Insert Page Layout Formulas Data Review View

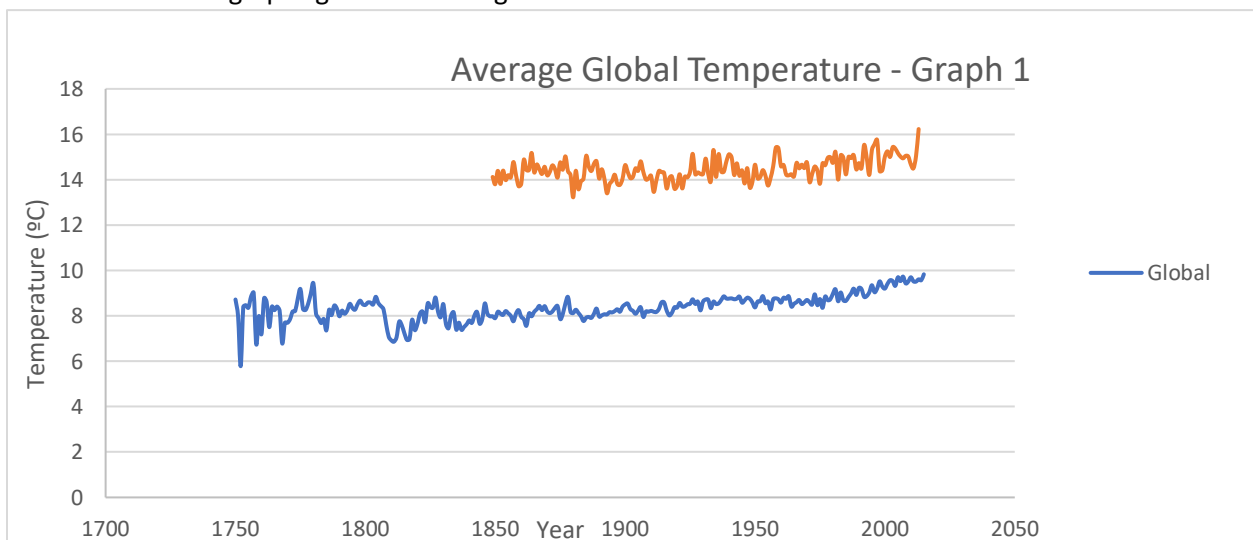
Cut Copy Paste Format Painter

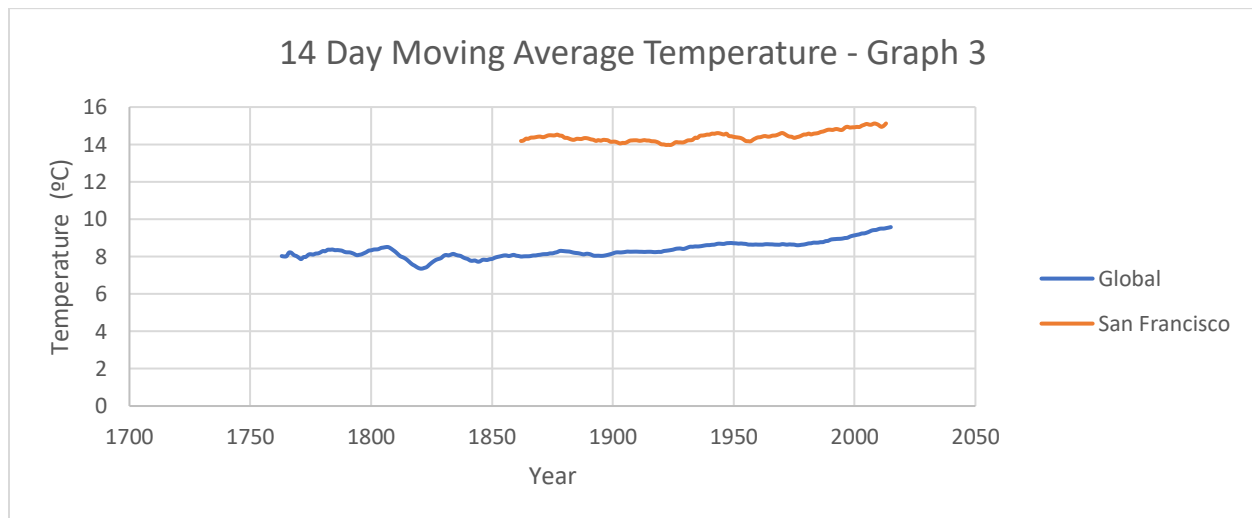
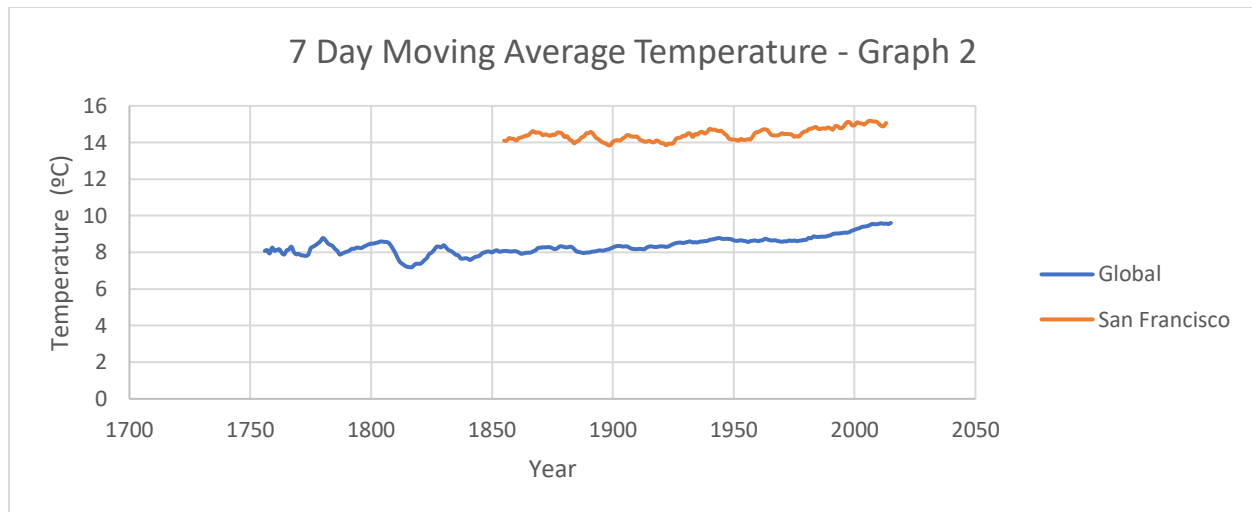
Clipboard Font Alignment

D15 \times \checkmark f_x =AVERAGE(B2:B15)

	A	B	C	D	E	F
1	year	avg_temp	7D_MA_GLOBAL	14D_MA_GLOBAL		
2	1750	8.72				
3	1751	7.98				
4	1752	5.78				
5	1753	8.39				
6	1754	8.47				
7	1755	8.36				
8	1756	8.85	8.078571429			
9	1757	9.02	8.121428571			
10	1758	6.74	7.944285714			
11	1759	7.99	8.26			
12	1760	7.19	8.088571429			
13	1761	8.77	8.131428571			
14	1762	8.61	8.167142857			
15	1763	7.5	7.974285714	8.026428571		
16	1764	8.4	7.885714286	8.003571429		
17	1765	8.25	8.101428571	8.022857143		

Below are the line graphs generated using the data.





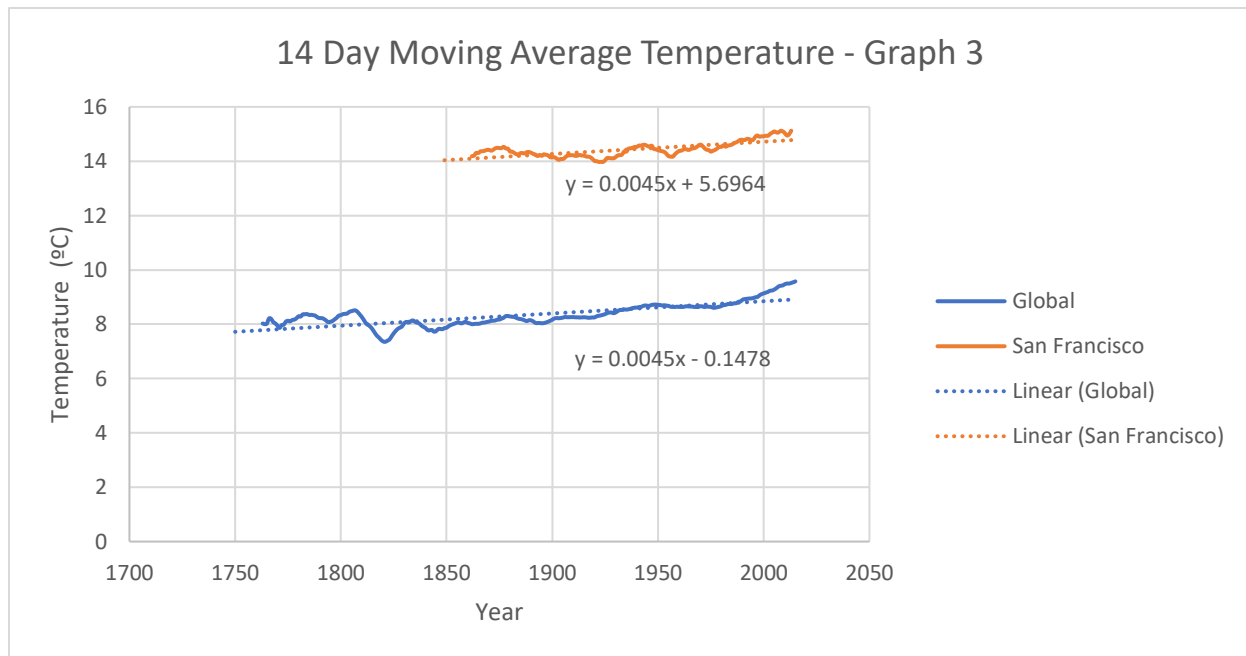
Graph 1 shows the actual temperature values against time. Because of the variations in the temperature and large number of data points, it makes it hard to see the overall trends. Therefore, the 7 day and 14 day moving average line graphs were created to identify the trends.

Observations

1. Both Global and San Francisco temperatures are increasing.
2. Temperature in San Francisco is higher (hotter) than the Global temperature.
3. The Global temperature is between 8-10 degrees while the temperature in San Francisco is between 14-16degrees. Overall, the temperature in San Francisco is approximately 4 degrees higher than the Global temperature.
4. The San Francisco and Global temperature patterns are very similar. Both increase or remain constant together. It could also be that the San Francisco temperature lags while the global temperature leads.

The screen below shows the correlation coefficient between Global and San Francisco temperature to be 0.54. The shows that there is a positive correlation between the two temperatures. This also means that as one temperature goes up, the other goes up as well.

Looking at the Slope for the two graphs:



It's interesting to see that the slopes of both the 14 day MA graphs are positive and same (0.0045) but with a different y intercept. This means that the linear steepness of the graphs are same (moving at the same rate) but from different starting points (the y intercept). The positive slope confirms an upward trend.