

Outline of Steps

1. Extracted the data from the SQL Workspace using the following queries;

Reason	SQL Query
To scan/assess the data	<pre>select city from city_list</pre>
To get data for one specific city “Islamabad, Pakistan”	<pre>select * from city_data where city = 'Islamabad'</pre>
To get global data	<pre>select * from global_data</pre>
<i>Downloaded the csv files for the above.</i>	

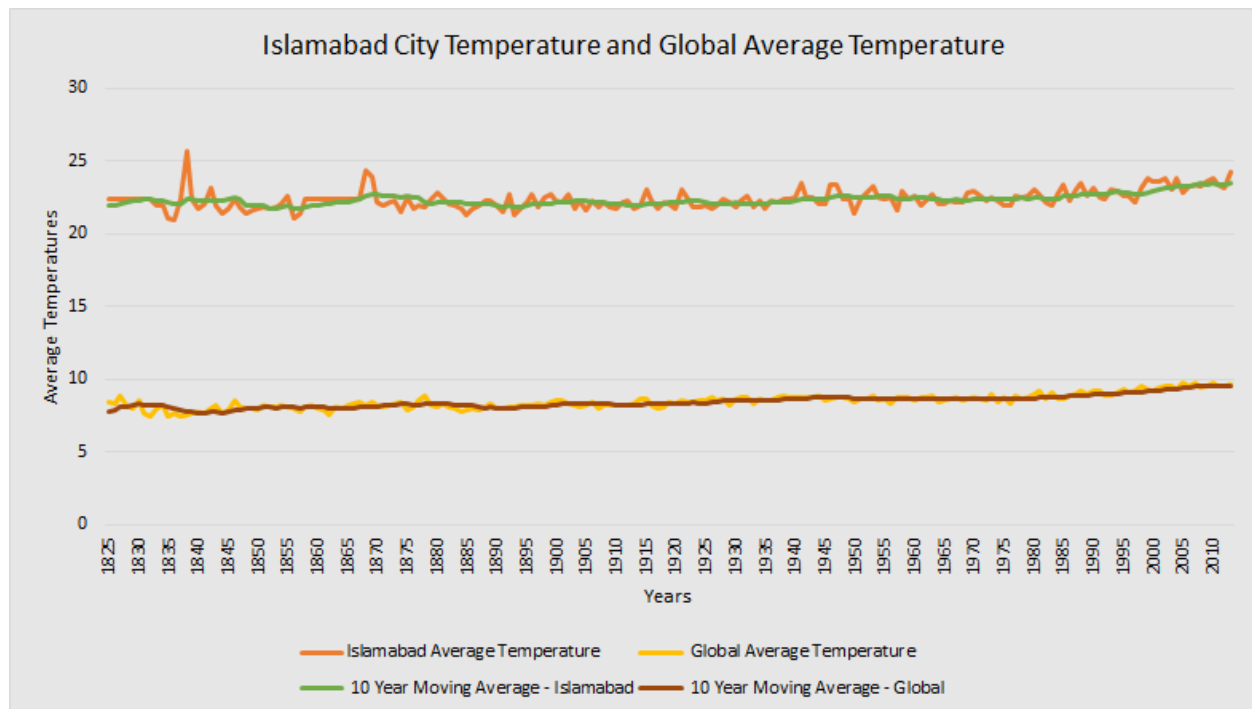
2. Used Excel to process and analyze the data using the following steps;

- Put the city and global data together on one sheet.
- Since there was a discrepancy in years for when the data was recorded, i.e. the global data was from 1750 and Islamabad’s data was from 1816, used the time period which was common for both starting from 1816.
- Calculated the moving averages for both sets of data, using a 10-year gap. A smaller gap of 5 years was giving a slightly volatile reading. For missing values, the average of the Islamabad city temperatures was taken and imputed.
- **Calculating Moving Averages:**
Since the average temperatures, when visualized was giving volatile readings, along with some gaps for missing values; 10-year moving averages were used. Two separate columns were created for 10-year moving averages of global and Islamabad city data, respectively. By going down to the 10th value and using the AVERAGE() function, the average temperatures for the first 10 years of temperatures for Islamabad city was calculated. The same formula was copied to the end of the list using the drag down method. The process was repeated for the second column of global temperatures.
- Following this, pivot tables were used to extract the relevant data and better summarize and analyze it.
- Line Charts were then used to visualize the data after which the data, axes, legends were formatted.
- The CORREL function was used to gauge the correlation between the two datasets. After this the variance and standard deviations were calculated for the average temperatures and the moving averages.

Key Considerations

- The data should be visualized in a continuous, non-volatile manner.
- It should be easily comprehensible to the one viewing it.
- A clear trend should be visible over the course of the years for both datasets.
- Should be presented in a manner that makes gleaning meaningful insights easier.

Line Chart



The visual shows the average temperatures for Islamabad and the 10-year moving averages for the same.

Summary Statistics

Statistics	Islamabad City Average Temperature - 10-Year Moving Average	Global Average Temperature - 10 Year Moving Average	Islamabad Average City Temperature	Average Global Temperature
Variance	0.135003411	0.181384786	0.427139974	0.242419977
Standard Deviation	0.367428103	0.425892928	0.653559465	0.492361632
Mean	22.32988889	8.438846561	22.36560847	8.478042328
Median	22.277	8.303	22.34	8.44
Mode	22.149	8.184	22.34	8.18
Correlation	0.803281148		0.528040226	

Observations

- From the visual, it can be seen that there is a direct positive correlation between the change in average global temperatures and the average temperature of Islamabad. An overall trend of increasing average temperatures is seen in both, indicating the world has gotten hotter on average.
- The moving average variances show very little changes in the temperature from the average, while the variances in actual recorded temperatures is significantly higher for Islamabad and moderately higher for the world. Using the variance for actual temperatures, it is indicated that the temperatures in Islamabad have varied relatively more vis-à-vis the average temperature for the world over the course of almost two centuries.

- The standard deviation for the average temperatures of Islamabad indicates that the recorded temperatures are more spread out, while the global temperatures are less spread out over the mean temperature. There is likely to be some discrepancy in the variance and standard deviation for Islamabad due to the missing values for which the mean was imputed, making the variance and standard deviation lower than the actual.
- The correlation coefficient is calculated to be 0.53 which shows a moderately high degree of positive correlation. This implies that the temperatures in Islamabad are likely to be higher/lower if the global temperatures are on the rise/on the decline.
- However, the average temperature of Islamabad is considerably higher than the global average; the former being in the range of 21-24C and the latter being in the range of 7.6-9.6C. This makes the city relatively hotter than the global average and this trend is seen to be consistent over time.