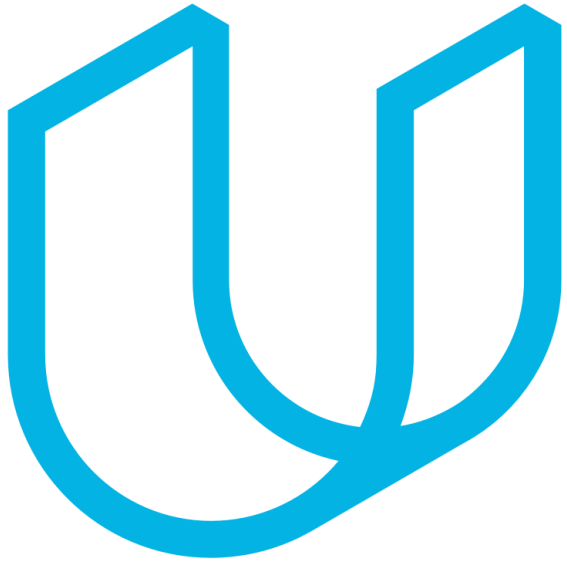
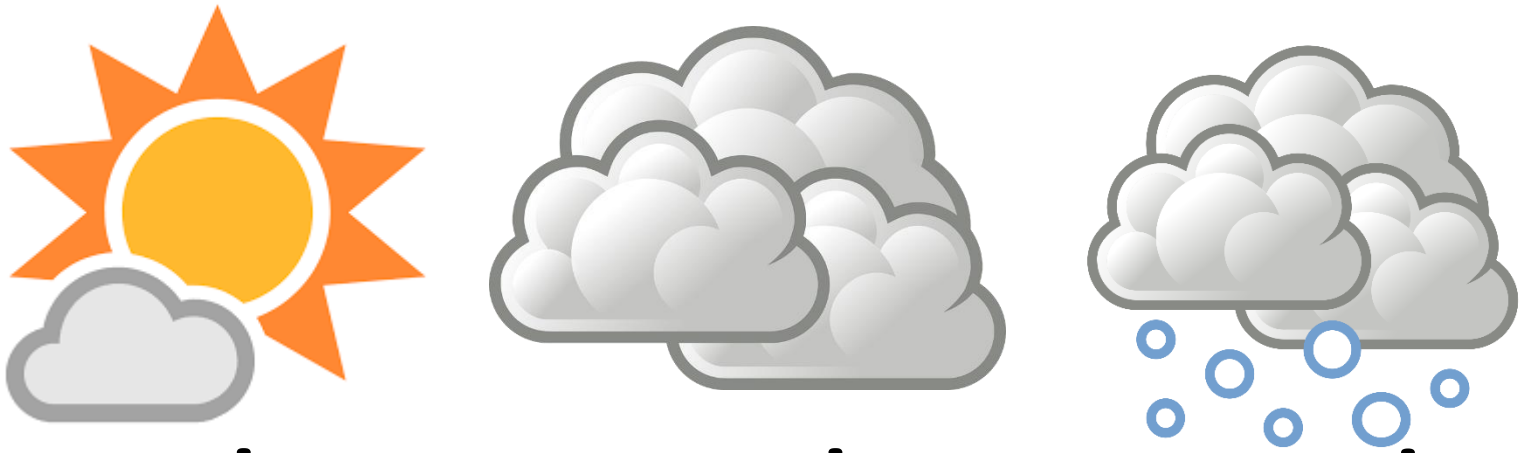
Data Analyst Nanodegree: Project 1



UDACITY



Explore weather Trends

By Bibhash Kalita

Analysis:

CITY	SQL Commands for Extracting Data
Patna	SELECT year,avg_temp FROM city_data WHERE city like ‘Patna' AND country like ‘India’;
Paris	SELECT year,avg_temp FROM city_data WHERE city like ‘Paris' AND country like ‘France’;
Helsinki	SELECT year,avg_temp FROM city_data WHERE city like ‘Helsinki' AND country like ‘Finland’;

For Global Data	SELECT * FROM global_data

Approach To get the desired data:

- I Copied the global data to all the three CSV files of cities to get the desired data for analysis.
- I Used Microsoft Excel for the analysis side by side I used google speed sheets as well.

Moving Averages 

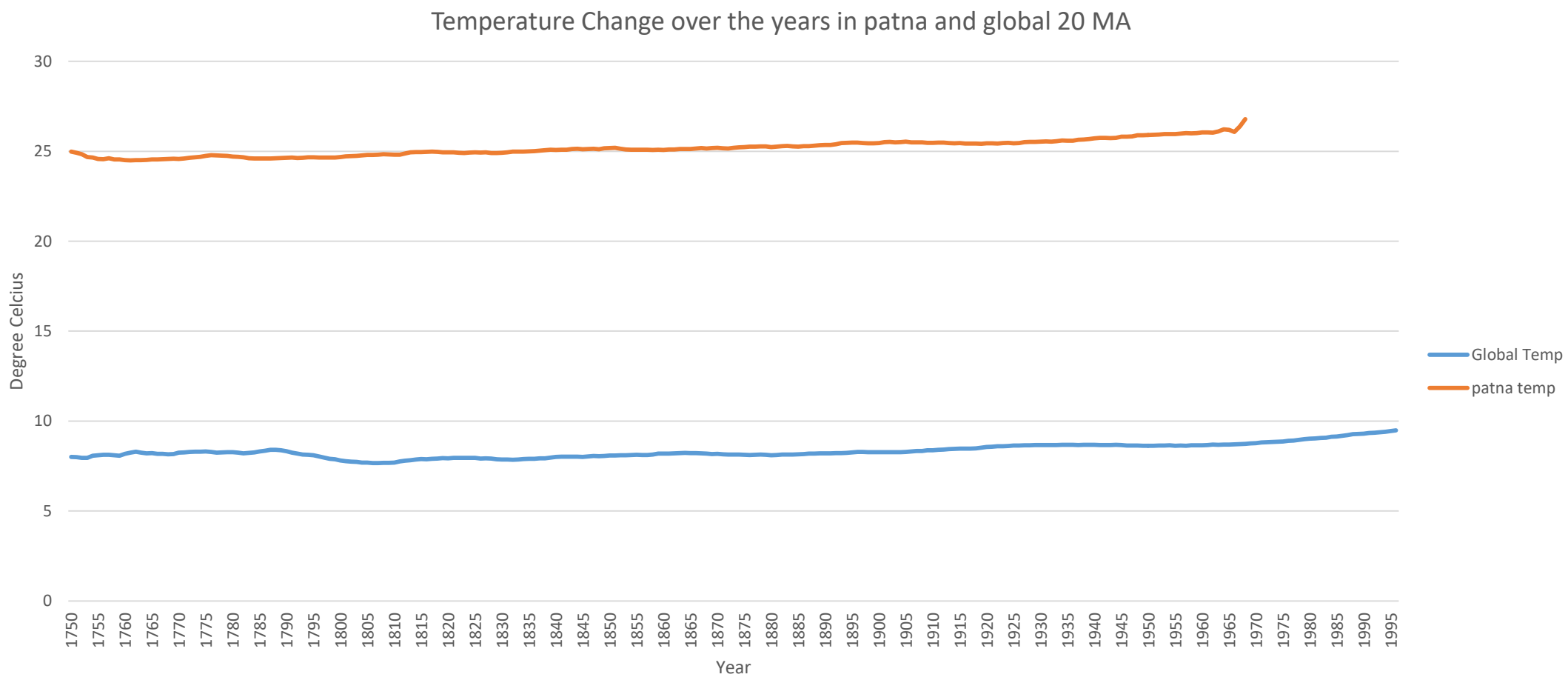
- To observe the trends in temperature I calculated moving average(MA).
- First I tried with 7 years MA but the line chart was not smooth so next I tried with 15 years MA , again I was not satisfied, then tried with 30 years MA but the line became way smooth .
- Finally after doing a lot of hit and trial I was satisfied with 20 years MA
- So I took 20 years MA in all my line charts

Excel commands for Moving Averages:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

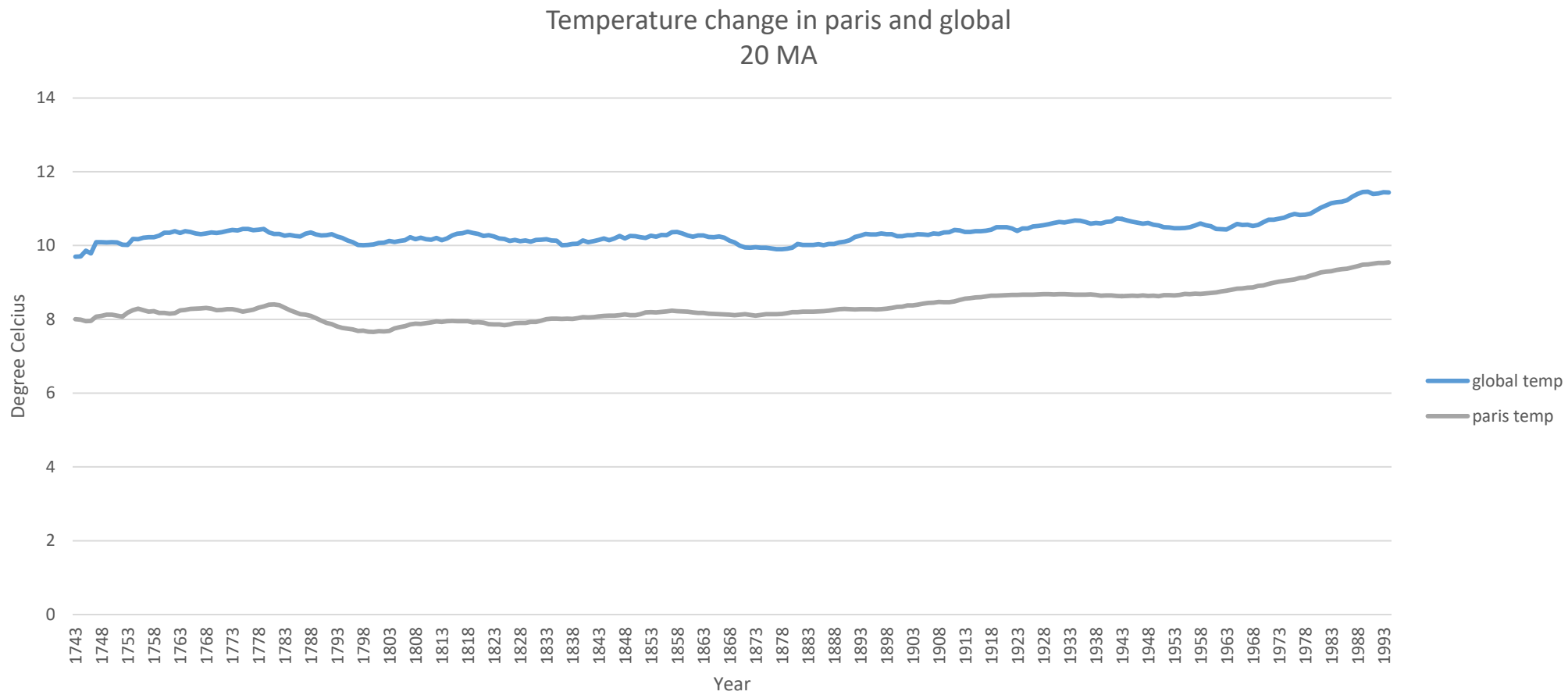
Moving Average	Excel Commands
For 7 years Moving Average:	=Average(B2:B8)
For 15 years Moving Average:	=Average(B2:B16)
For 30 years Moving Average:	=Average(B2:B31)
For 20 years Moving Average:	=Average(B2:B21)

Line Chart for Patna and Global Temperature:

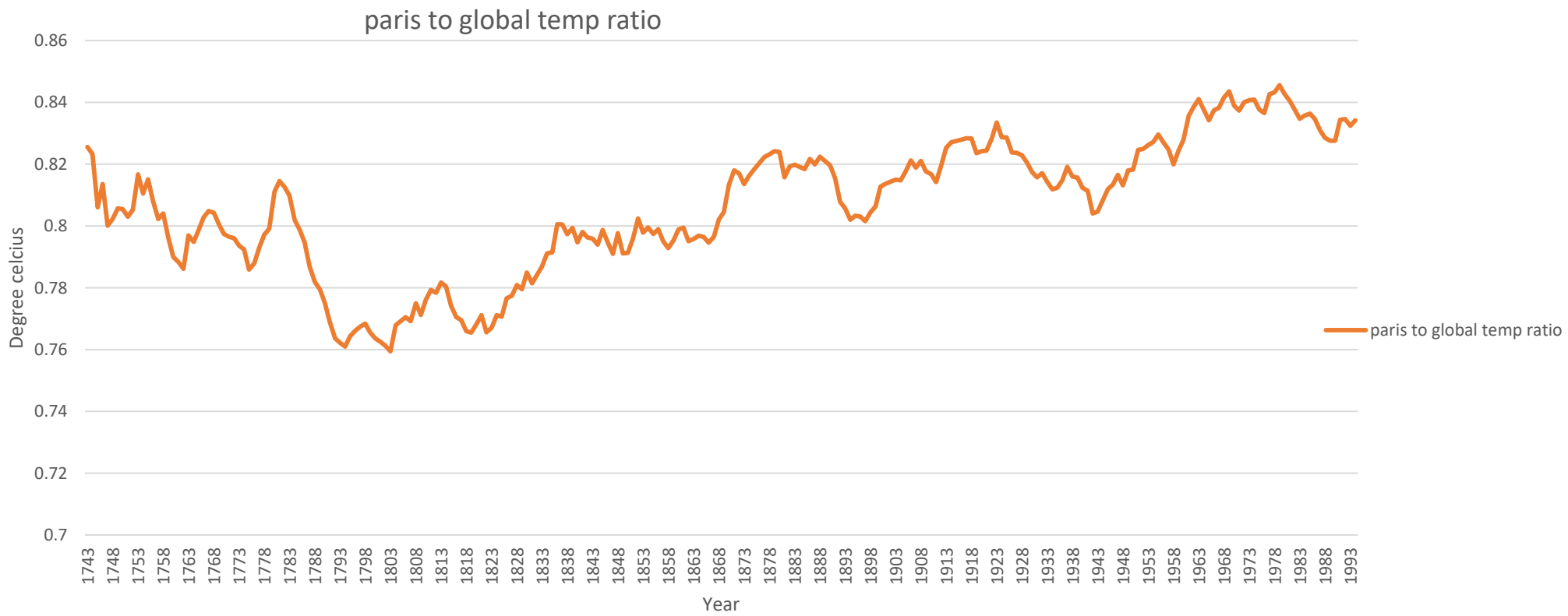


Line Chart for Patna to global Temperature Ratio:

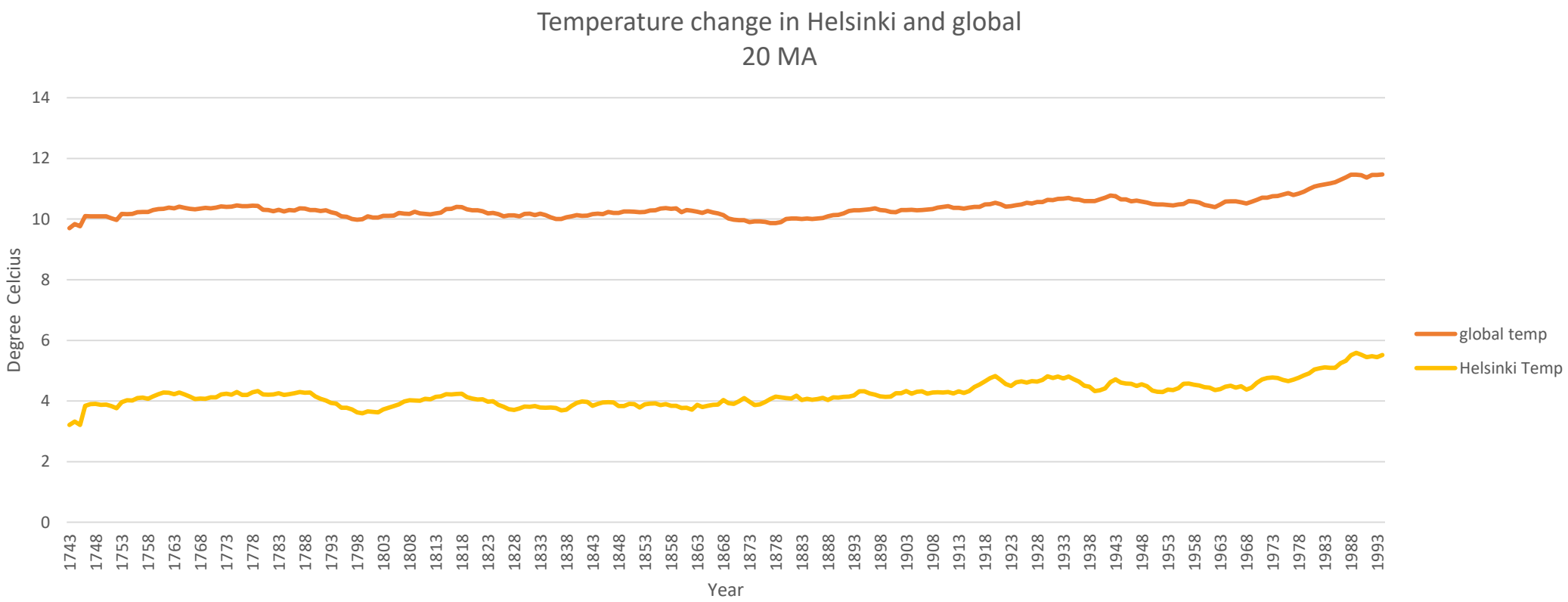
Line Chart for Paris and Global Temperature:



Line Chart for Paris to global Temperature Ratio:



Line Chart for Helsinki and Global Temperature:



Line Chart for Helsinki to global Temperature Ratio:

Correlation Coefficient:

$$r = \frac{Covariance\ (x,y)}{S.D.(x)S.D.(y)}$$

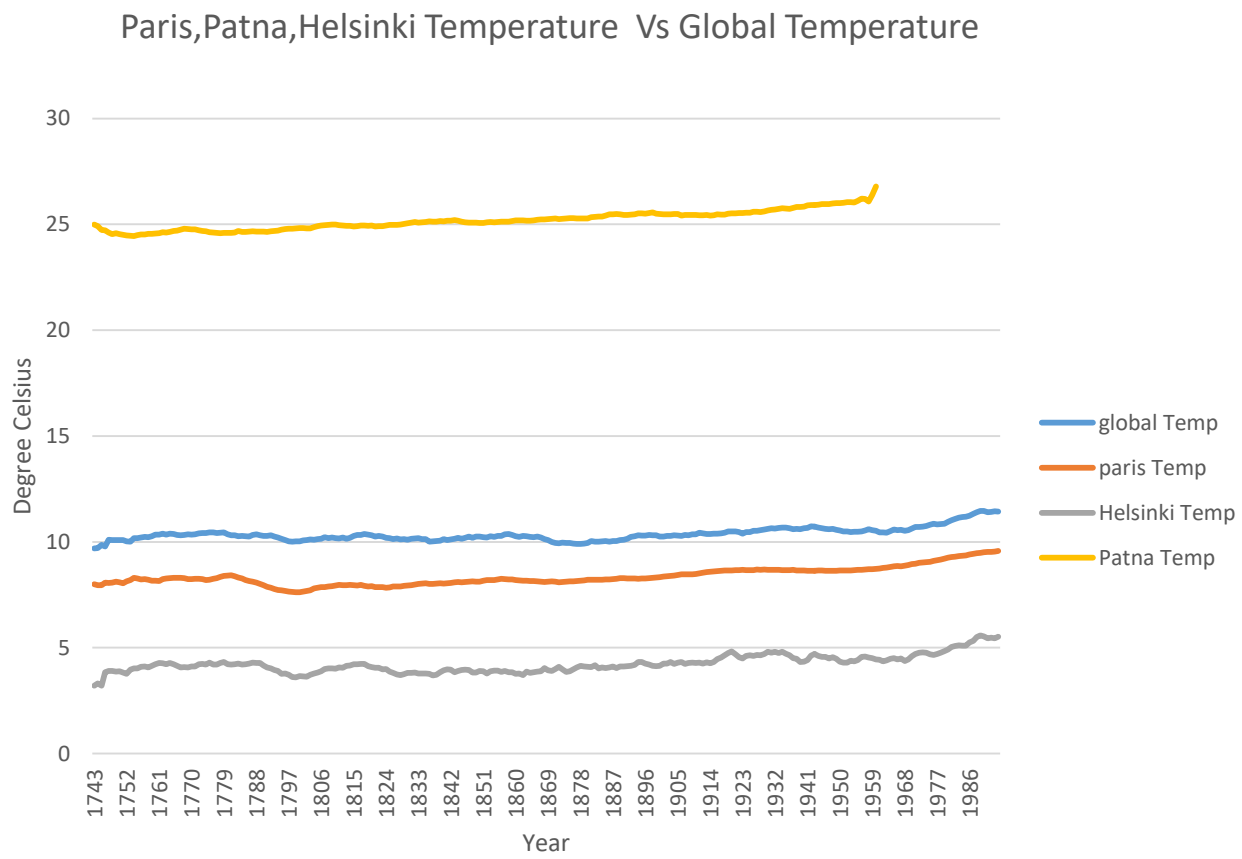
- Paris - 0.886919 ~ 0.89 a strong positive relationship
- Patna- 0.711665 ~ 0.71 a strong positive relationship
- Helsinki- 0.921642 ~ 0.92 a strong positive relationship
- Correlation shows if the city temperature is increasing the global temperature is increasing as well
- Since correlation is a symmetrical relationship, so we only observe a relationship between city temperature and global temperature

Excel Command to find Correlation Coefficient	**=PEARSON(B21:B272,K21:k272)**

Findings:

- The global temperature has increased by 1 degree Celsius from 8.72 to 9.83
- The temperature of Patna has increased by 2 degree Celsius over the years
- Since correlation coefficient of Patna is around 0.89 we can say that both global as well as Patna temperatures are increasing over the period of time
- Now when I related all three cities I saw that for Paris there was a decrease in temp by 1 degree but eventually with time it increased by 1 degree , but from this we can say that Patna is way hotter then Paris , Helsinki on the other hand showed an increase in temp by 2 Degree Celsius
- So we can say that Patna is way hotter than Paris and Helsinki
- Patna is 3 times hotter than global temperature
- Paris is interesting because its colder than global temperature by 2 degree Celsius
- Helsinki is 3 times colder than global temp which is again an interesting finding

Conclusion Line Chart:



From This line chart we can see that eventually the graph is moving upwards which means the global temperature is rising which is directly proportional to increase in temperatures of all the three cities.

Global Temperature $\propto$ Paris, Patna, Helsinki Temperature

Final Conclusion:

The World is Getting Hotter

Resources Used:

- Google – For Pictures
- Books Used :- Statistics in a nutshell to Find about Correlation Coefficient


END