EXPLORE WEATHER TRENDS

- SQL Queries to extract the data from the database are:

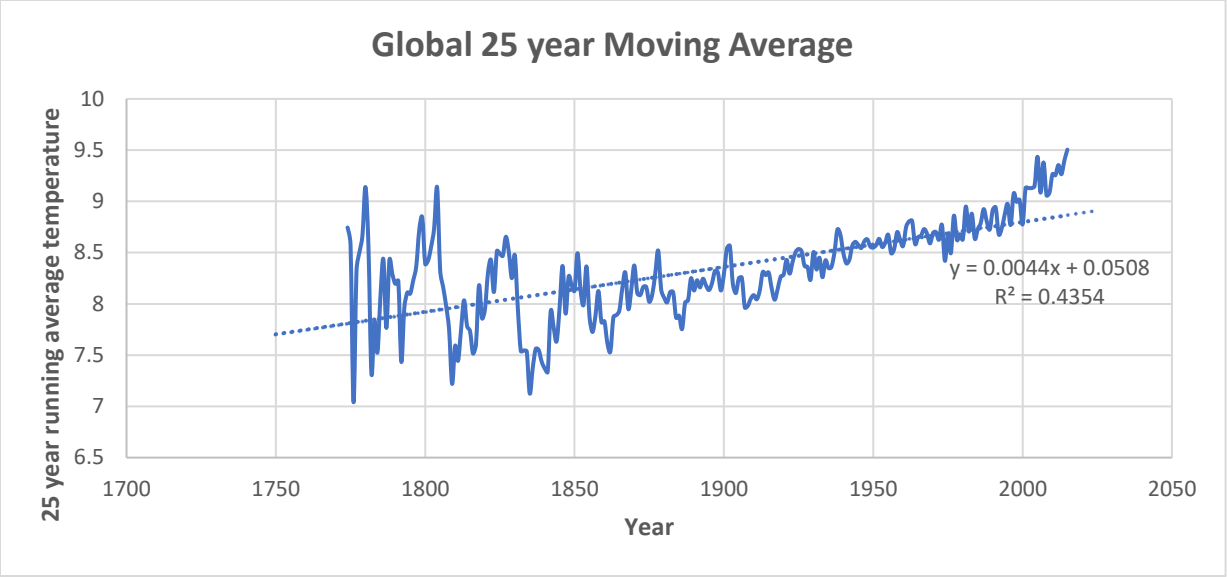
1. **To extract whole data from city_data:** SELECT *
   FROM city_data;
2. **To extract whole data from city_list:** SELECT *
   FROM city_list;
3. **To extract whole data from global_data:** SELECT *
   FROM global_data;
4. **To extract data related to near by city i.e. Delhi:** SELECT *
   FROM city_data
   WHERE city='Delhi';

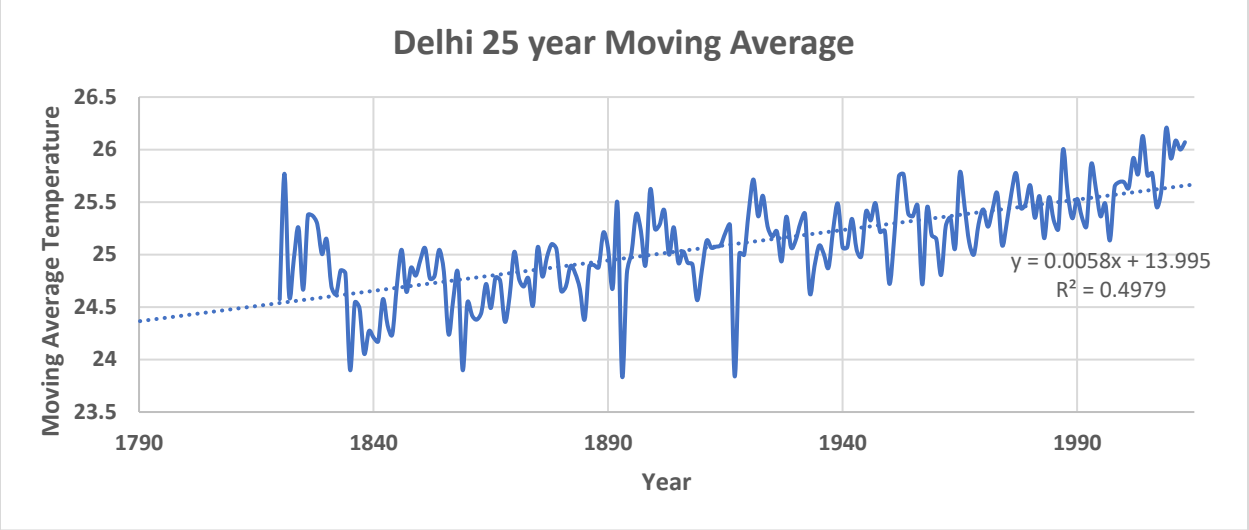
- Comparison of Delhi Temperature and Global Temperature:

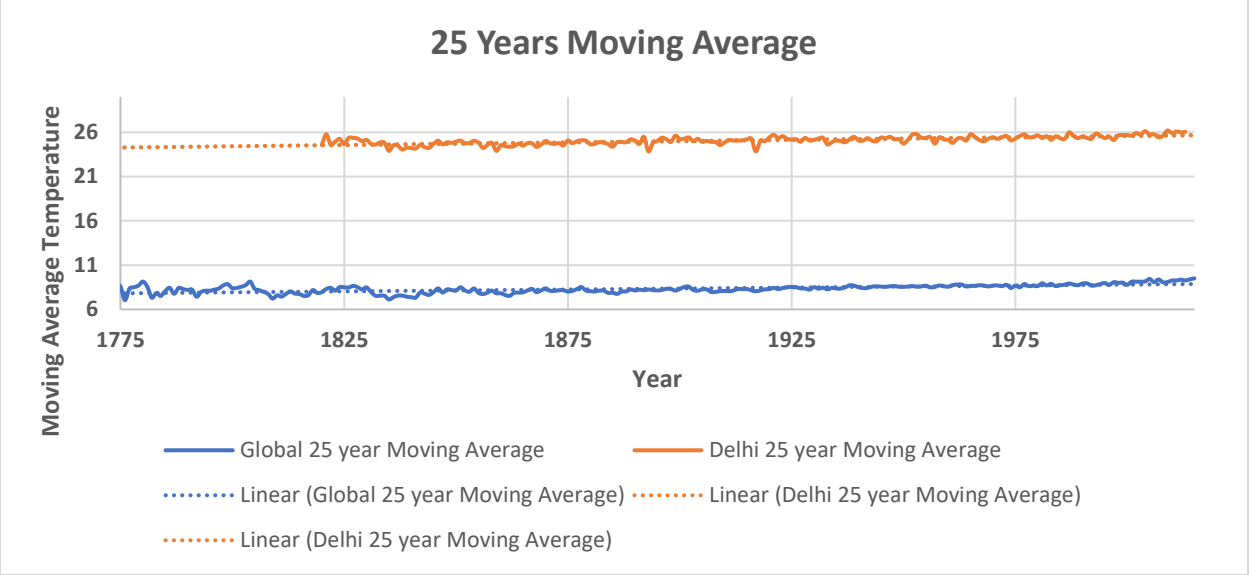
  Both the temperatures were taken in 25 year moving average form. In both cases averages of 25 years are taken from 1750 to 2015 in global temperature and 1772 to 2013 in case of Delhi.

  Observations:

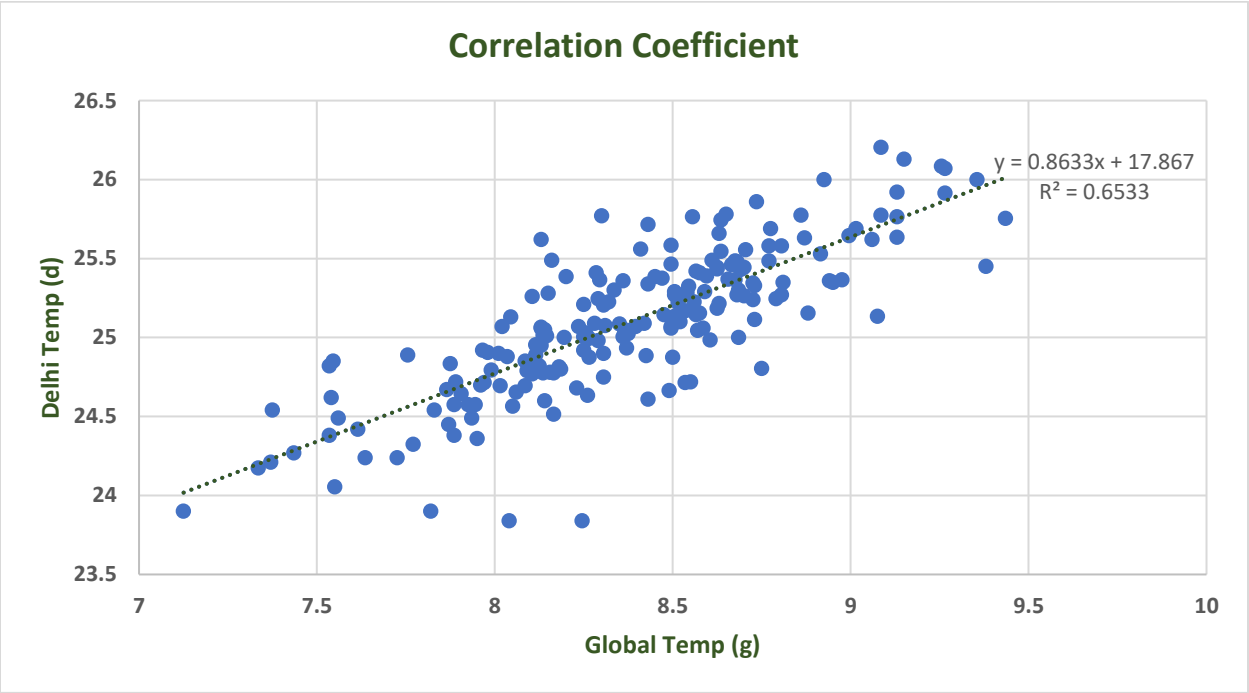
  1. Delhi is hotter compared to the global temperature. It is about 20 ^{0}C than the average global temperature.
  2. Whole world is becoming hotter by years due to global warming. Average global temperature is changed to 8.75 0C in 1774 to 9.27 0C in 2013. For Delhi its, 24.58 0C in 1820 to 26.07 0C in 2013.
  3. In the graph it can be seen that variation around the mean of average temperature during the years in 18th century was more while in 21st century the graph is steadier and is growing upwards. That means the weather is more variant during the 18th century.
  4. Delhi graph is more variant than global. Hence it can be seen that if we consider only a small area, seasons affects more to that as compared to whole planet.


Global 25 year Moving Average
25 year running average temperature
10
9.5
9
8.5
8
7.5
7
6.5
1700
1750
1800
1850
1900
1950
2000
2050
Year
y = 0.0044x + 0.0508
R² = 0.4354


Delhi 25 year Moving Average
Moving Average Temperature
26.5
26
25.5
25
24.5
24
23.5
1790
1840
1890
1940
1990
Year
y = 0.0058x + 13.995
R² = 0.4979



Correlation Coefficient: Correlation coefficient is a statistical measure used to predict the change in one variable with the changes made in another variable. In positive coefficient, both the variables increases or decreases in tandem. In negative coefficient, if one increases other decreases and vice-versa. Its value is between +1 and -1. More it is near the zero, less is the correlation. It is maximum at both the ones.



Correlation coefficient (R) = $\dfrac{n\sum dg - (\sum d)(\sum g)}{\sqrt{\{n\sum d^2 - (\sum d)^2\}}\sqrt{\{n\sum g^2 - (\sum g)^2\}}}$

Where d= 25 year moving average of Delhi temperature.

g= 25 year moving average of global temperature.
n= number of years whose data is considered

Here , **R= 0.8082**

i.e. There is strong correlation between the global and Delhi temperature values and Delhi temperature can be predicted if we know the global data.
Using the linear relation: **d = a + bg;**
Where, $b = \frac{n(\sum dg) - \sum d \sum g}{n\sum g^2 - (\sum g)^2} = 0.863$

$a = \bar{d} - b\,\bar{g} = 17.869$
putting the values of a and b in linear relation equation, we get **d= 17.869+0.863g**.

Now this relation can be used to estimate the value of Delhi temperature if we know the global temperature.
For Example, if g=8.5, then d= 17.869+0.863* 8.5 = 25.20 ^{0}C.