

DATA ANALYST NANODEGREE



Explore Weather Trends



Udacity : Data Analyst Nanodegree

Ashish Kumar Singh

Project - 1 , Explore Weather Trends
Bangalore, India

Overview:

I have been provided the temperature database from the portal from where I have extracted the data related to global temperature and my city temperature. I analyzed the temperature around the global with the city I live by extracting the data from the database.

Goals :

1. Selecting city and country from the database “city_list”.
2. Extracting the City level data from the database “city_data” and export to CSV file.
3. Extracting the global temperature from the database “global_data” and export to CSV file.

Tools Used :

1. **SQL** : To extract the data from the database
2. **Google Sheets** :
 - > To calculate Moving Averages of global and city temperatures
 - > To plot Line Chart

STEP : 1 Extraction of data from the database

1. To check which countries and cities are available in the database.

```
SELECT *  
FROM city_list WHERE Country='India'  
AND City='Bangalore';
```

2. To select data from the City database

```
SELECT avg_temp,year,city,country  
FROM city_data  
WHERE city='Bangalore';
```

3. I observed there is a column called `avg_temp` which is same in both `city_data` and `global_data`. I want to change the schema so I joined both the tables and changed the column names in both the databases.

```
ALTER TABLE city_data
    RENAME COLUMN avg_temp to CAT;
ALTER TABLE global_data
    RENAME COLUMN avg_temp to GAT;
```

4. I have joined the two tables using JOIN also called as INNER JOIN as `avg_temp` is same in both the tables.

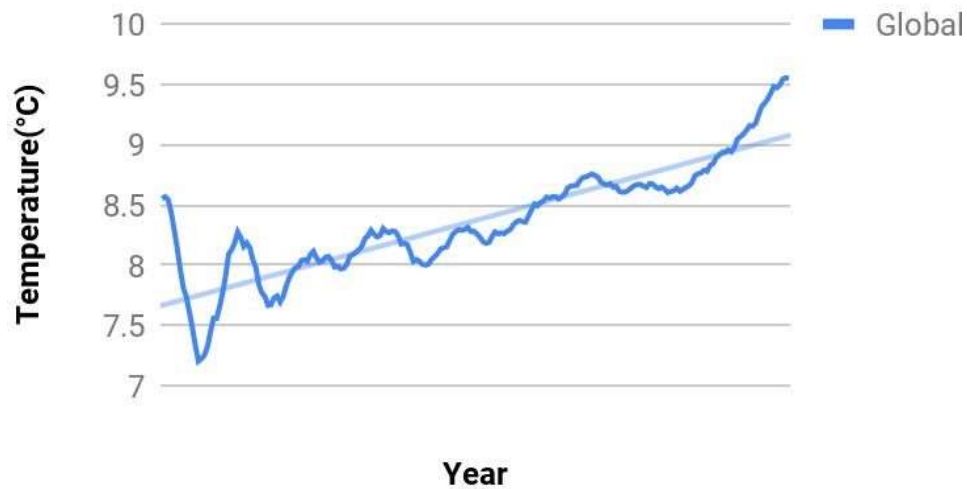
```
SELECT city_data.CAT, global_data.GAT,global_data.year
FROM global_data
JOIN city_data
ON global_data.year = City_data.year
WHERE city='Bangalore' Country='India';
```

Now, I have got an option to download a CSV file . I downloaded file as “results.csv”.

Moving Averages :

1. To smooth the data and to observe trends in the temperature.
2. I have done 10 year Moving Average to get the smooth line chart.
3. I used a command `=AVERAGE(A2:A11)` to see the Moving Average Value for 10 years.

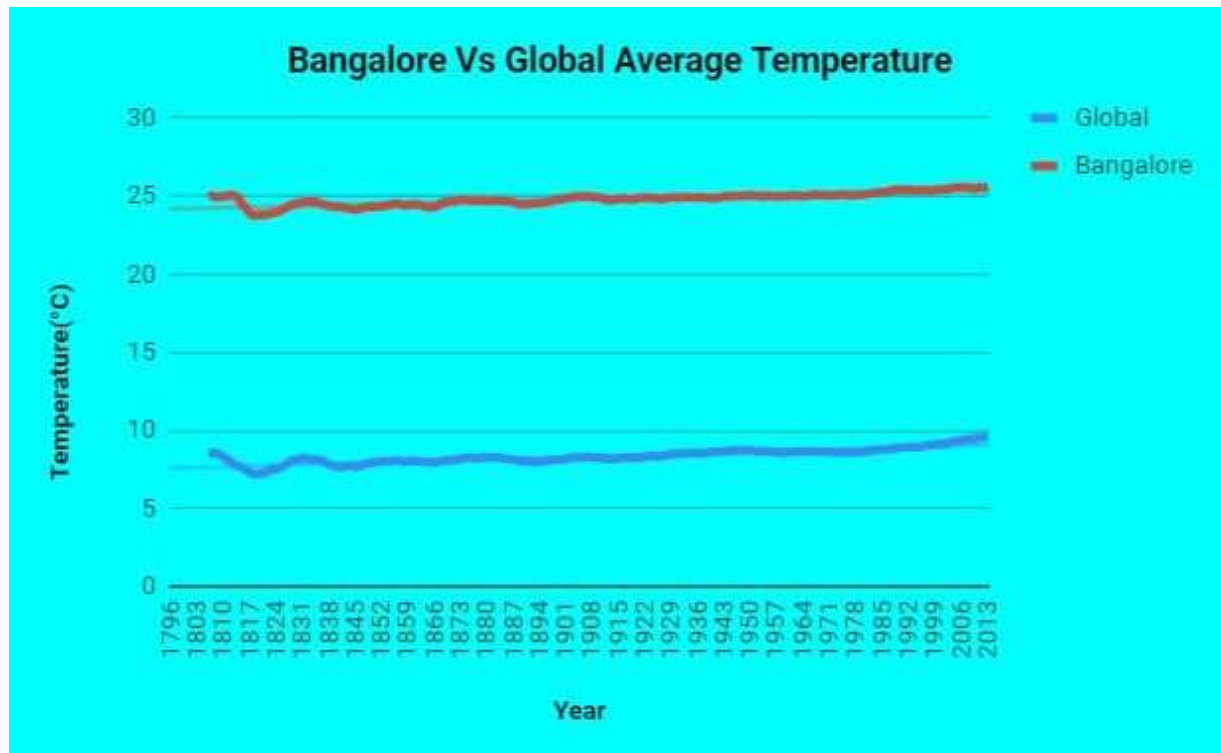
Global Temperature



I have plotted Line chart for global data separately to observe difference between Global Average Temperature and the city Bangalore.

Now I have combined both Global Average Temperature and Bangalore with 10 year MA.

Here is a Line Chart of GAT and Bangalore Average Temperatures for 10 Year MA.



OBSERVATIONS :

1. Global Average Temperature for 10 yr MA varies between 8.5°C to 9.5°C
2. Bangalore city Average Temperature for 10 yr MA varies between 24.98°C to 25.57°C
3. The Chart of Bangalore Vs Global has very big difference in the temperatures.
4. If comparison is made between Global and Bangalore Average Temperatures Bangalore is hotter than global average temperature .
5. From the first Graph, I observed global temperature is increasing from 8.5 to 9.5. From the second graph I observed the temperature of both global and Bangalore average temperatures are ups and downs during the early years, later during 1996 to 2013 both the temperatures increased due to increase in temperature.
6. The final conclusion of this project is Bangalore is hotter than global temperature and temperature is increasing day by day due to changes in the climate.

7. Further I have also checked with the other cities in India and also plotted charts for 50 Year MV.
8. I observed that for 50 MA the chart is smoother when compared to 10 year or 5 year MV.

Considerations:

X-Axis : Years

Y-Axis : Temperature($^{\circ}\text{C}$)

Legends are plotted in the charts

Different colors of lines for global and city temperatures.

References:

<http://www.statisticshowto.com/moving-average/>