EXPLORING WEATHER TRENDS

SEPTEMBER-15,2018

UDACITY-Data Analyst Nanodegree

ANSARI ARHAM ABRAR

TERM-1,PROJECT-1,Explore Weather Trends.

PUNE.INDIA.

OVERVIEW

In this project I have analysed local temperature of Pune,India with the global temperature data and compared it.I had been provided entire database from Udacity portal from where I have to extract data and manipulate data and visualise it.

Goals

1. To extract data from database and export it to CSV file.
2. To make a chart visualisation based on extracted data.
3. Observation based on charts.

Tools Used

1. SQL: It is used to extract data from Udacity's database.
2. EXCEL: To analyse data and visualise it.

STEP 1: Extraction of data from the Udacity's database

1. To see which cities are available for "India" in the given dataset:

   ```
   SELECT * FROM city_list WHERE country LIKE 'India'
   ```

2. I can make a relevant dataset by joing the two tables.But I found that both the Schema contains both city_data and global_data contains same column 'avg_temp'.So I have change the column name of both the table from 'avg_temp' of global_data to 'GAT' and from 'avg_temp' of city_data to 'CAT'.In order to have a distinct columns.

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

   CITY AVERAGE TEMPREATURE(CAT)

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

   GLOBAL AVERAGE TEMPREATURE(GAT)

3. Now I have written code in order to join the two tables:

   ```
   SELECT global_data.year,global_data.GAT,city_data.CAT
          FROM global_data   JOIN city_data    --join tables
          ON global_data.year = city_data.year   --joining refrences
          WHERE city LIKE 'Pune';
   ```

After evaluating the query in workspace now the file is downloadable in

Comma Seprated Value(CSV) format as results.csv.

STEP 2: ANALYSING FILE IN EXCEL SHEET

After CSV file is downloaded it comes to analysing .

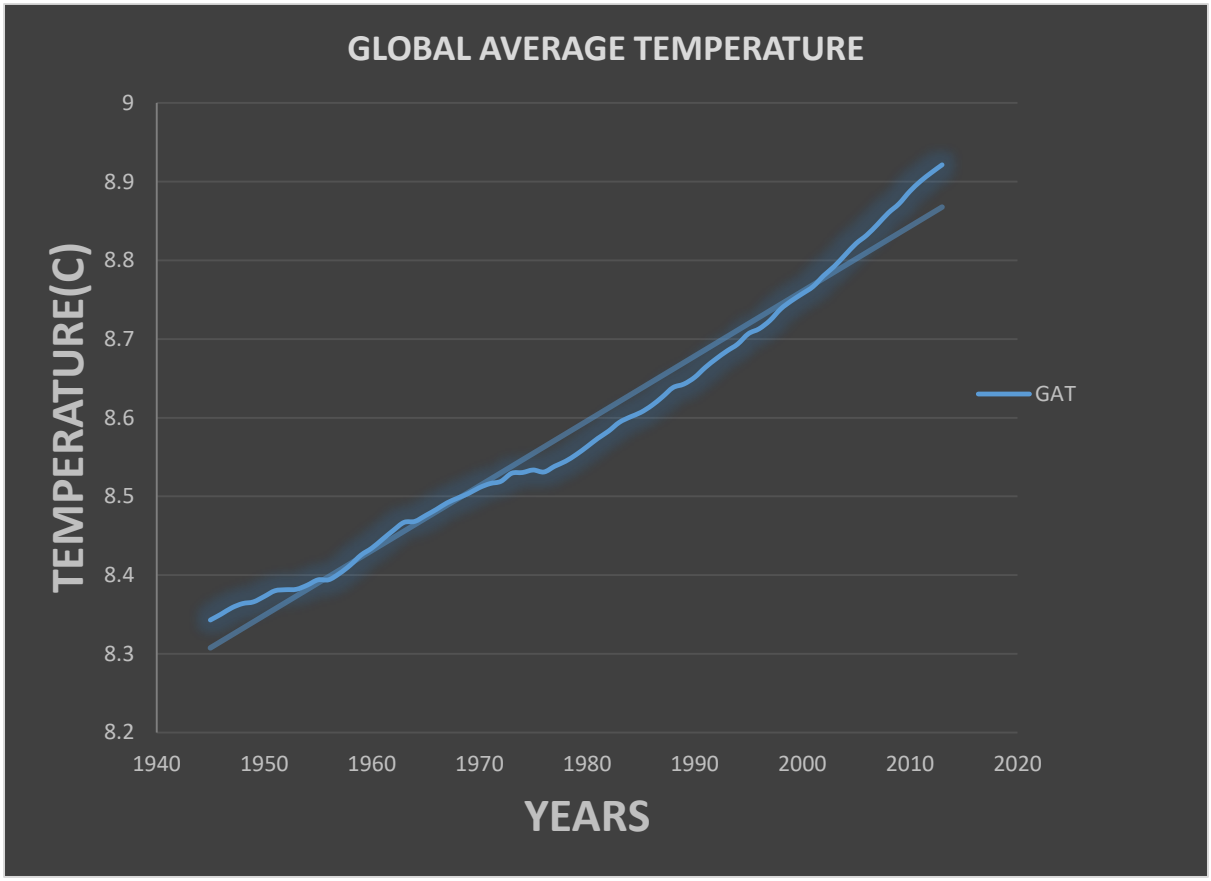
I imported data analysis library in EXCEL through Add-ins and then calculated 75 years moving average for global temperature and 4 years Pune temperature in order to get smooth curve.

Firstly, I analysed global temperature separately to check and distinguish it from combined data of global temperatures and local temperatures.

Calculating moving average for global average temperature and city average temperature:

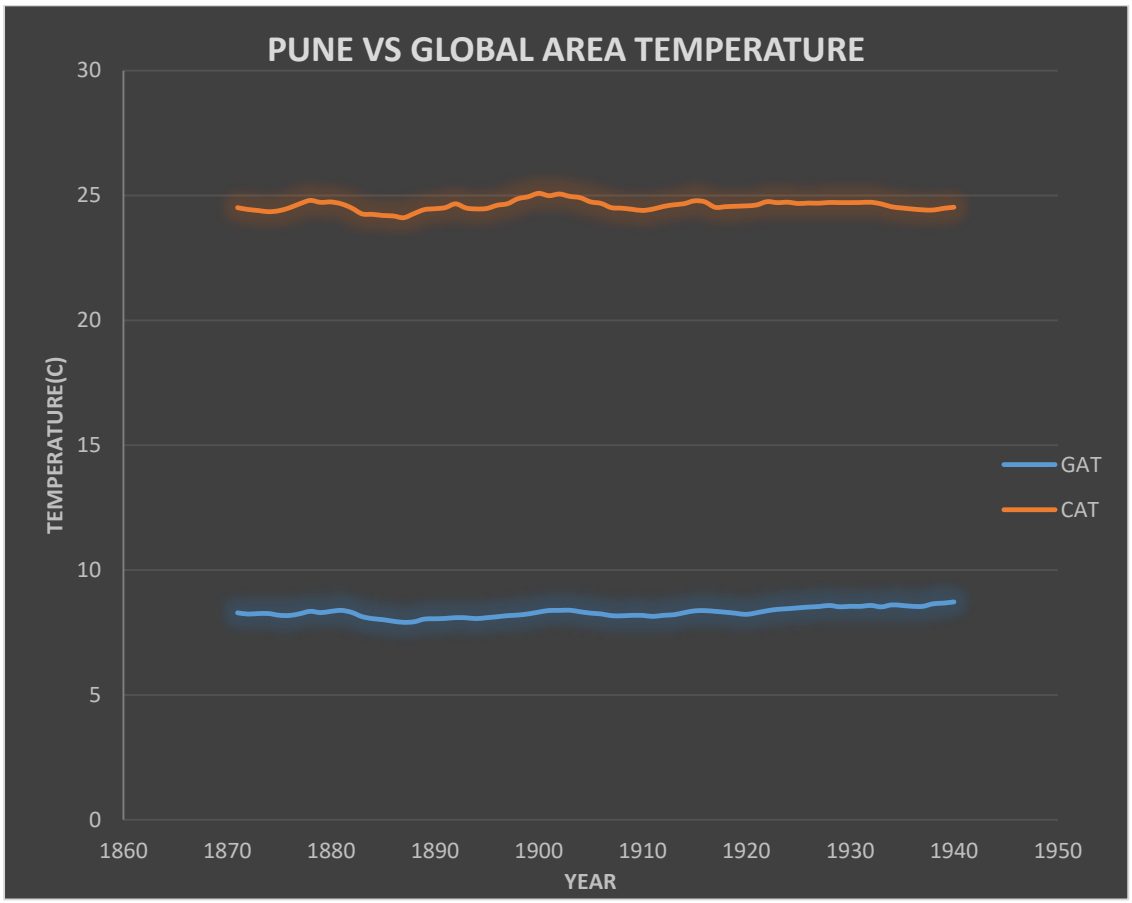
- First I have added Data analysis ToolPak from Add-Ins.
- Data analysis ToolPak consist of in-built library for calculating moving average in Excel.
- After selecting range for input,output and interval,moving average is calculated

Following Line graph shows global average temperature :



Following graph shows 75years moving average of Global average temperature

Then,I have calculated 4 year moving average for Global vs Local area temperature and plotted it in the line graph shown:

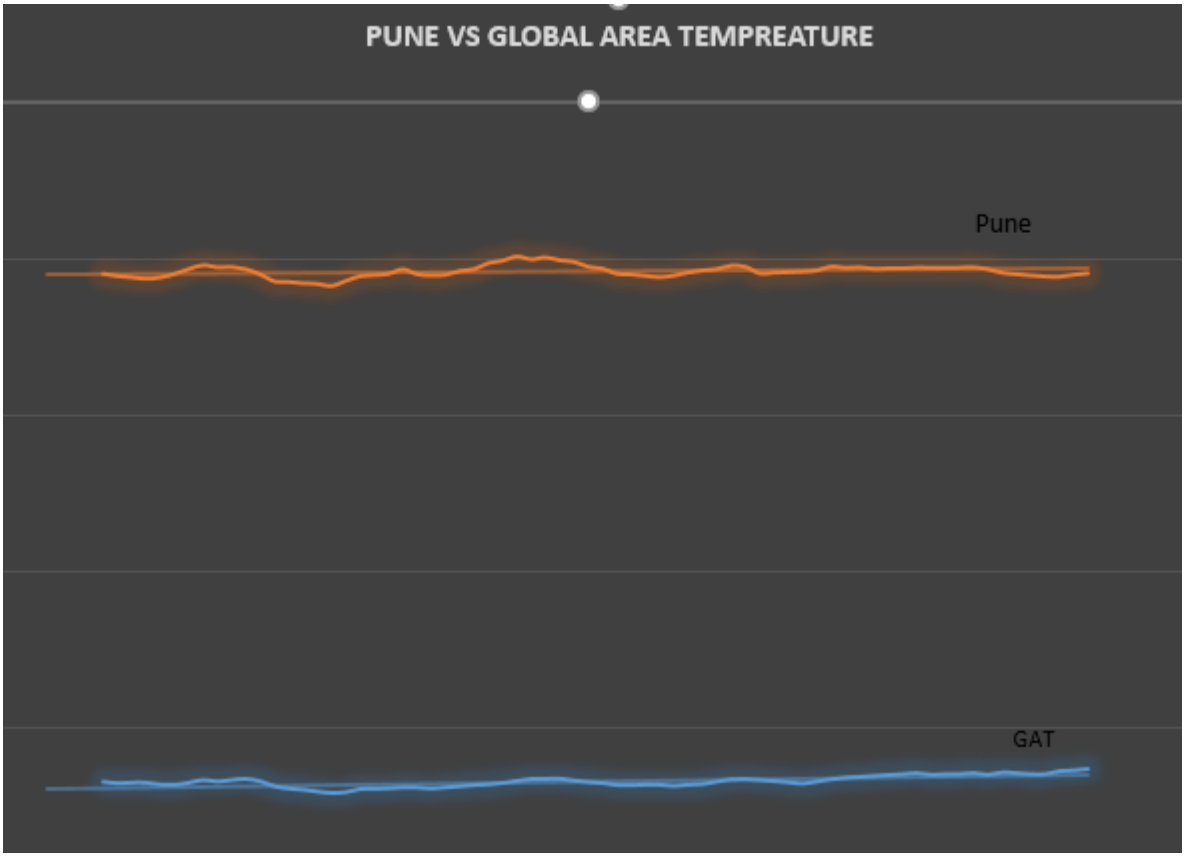


RESULTS: Observation

I observed that,if I choose a short range for moving average i.e 10 or 50.I will get messy line in the graph .If I use a larger moving average i.e 150 or more then that I will get a smooth graph which has longer range.

Observation from a line chart

1. The chart of Pune vs Global Temperature shows quite big difference between average temperature of Pune and that of world.
2. Since,I have got a slight inclined straight line for global data,I have separately plotted the graph of global data.
3. From the graph of global temperature data it is clear that temperature is increasing quite constantly with years.
4. It is observable in the second chart that temperature of Pune is quite high with respect to global average temperature.
5. If I draw a tangent line between crest and trough of Pune and Global temperature, there is a consistent change between this line and the line of global average over time.



6.It seems to be Pune is lot hotter as compared to any other cool places in the world.

7.Observing the graph global average temperature of the world shows a constant rise (Observed by the linear trendline).

Key Considerations:

1. Unit of temperature :Centigrade,on Y-axis.
2. Years shown on X-axis.
3. Calculated 4 years moving average on City average temperature and 75 years moving average to get global average temperature to get a smooth line.
4. Different colours of line for City and Global average temperature.
5. Linear trendline for better changing observations.