



Data Analyst Nanodegree

Project Name : Explore Weather Trends



VS



Abhishek Kumar Shukla
Varanasi, India.
May-3, 2020.

Overview :

In this project I have analyzed Global average temperature with Varanasi, India. The dataset I used has been taken from Udacity Portal in which global average temperature and city temperature average data is provided. I extracted the data in .csv format and then I imported into Google Sheets to analyze.

Goals:

- Get the data by writing SQL query and then extract it in .csv format from Udacity portal to analyze.
- Open the CSV file in any tool for making charts. I used Google Sheets.
- Calculate moving average for drawing smooth graphs. I used a 10 year moving average starting from 1913.
- Create a line chart.
- Make observations based on the chart.

Questions to be asked:

- How do the changes in Varanasi temperature over time compare to the changes in the global average?
- Is Varanasi hotter or cooler when compared to global temperature?
- What is the correlation between both?
- What does the overall trend look like?

Tools Used:

- 1) PostgreSQL^{*} : To extract data from the database.
- 2) Google Sheets : To calculate moving average and plot the line chart.
- 3) Google Docs : To store SQL query.

Project Walkthrough :

Step 1 : Extract data from the provided database

1. Get a list of Cities in India from the database.

```
1  SELECT * FROM city_data
2      where Country LIKE 'India';
3
```

2. City data and global data is given in different tables and need to combine for further study. The column name for temperature is the same in both the tabel to I have to rename it to merge.

```
1  ALTER TABLE city_data
2      RENAME COLUMN avg_temp to city_avg;
3
```

```
1  ALTER TABLE global_data
2      RENAME COLUMN avg_temp to global_avg;
3
```

3. Merge both tables using INNER JOIN.

```
1  SELECT c.city_avg , c.year , g.global_avg
2      FROM city_data c
3      JOIN global_data g
4      ON c.year = g.year
5      WHERE city = 'Varanasi'
6
```

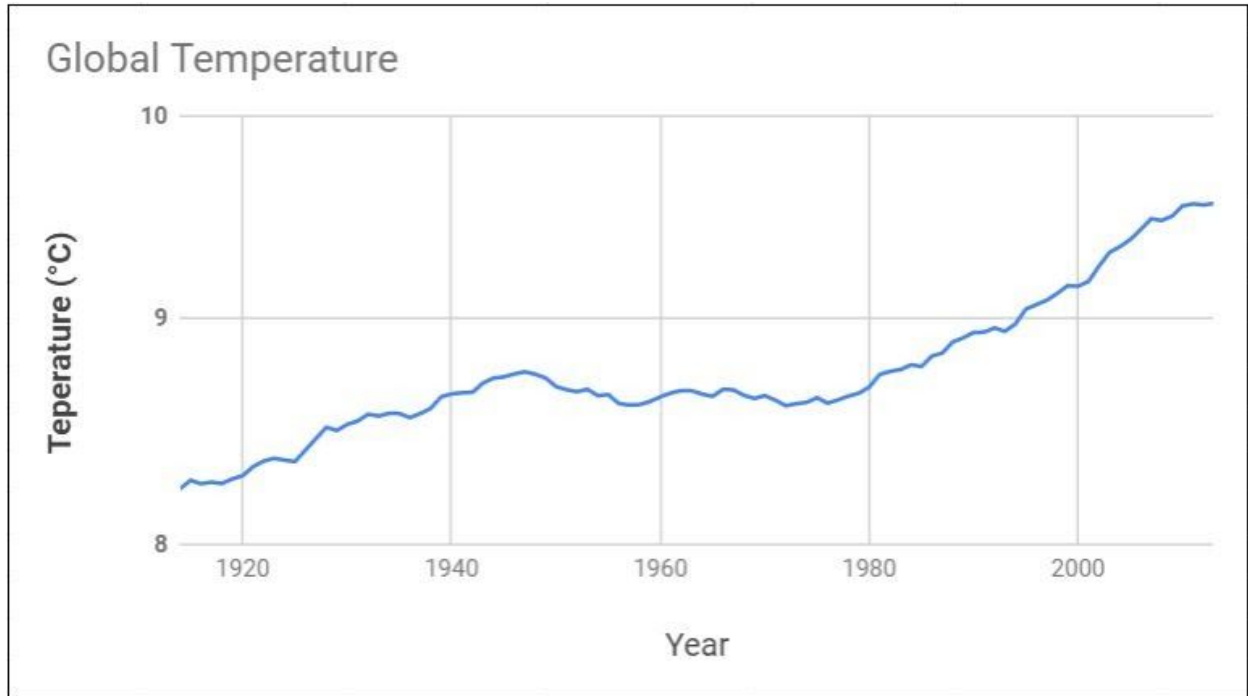
I have got the option “Download CSV” , using that I downloaded the data in CSV format.

Step 2 : Calculate Moving average.

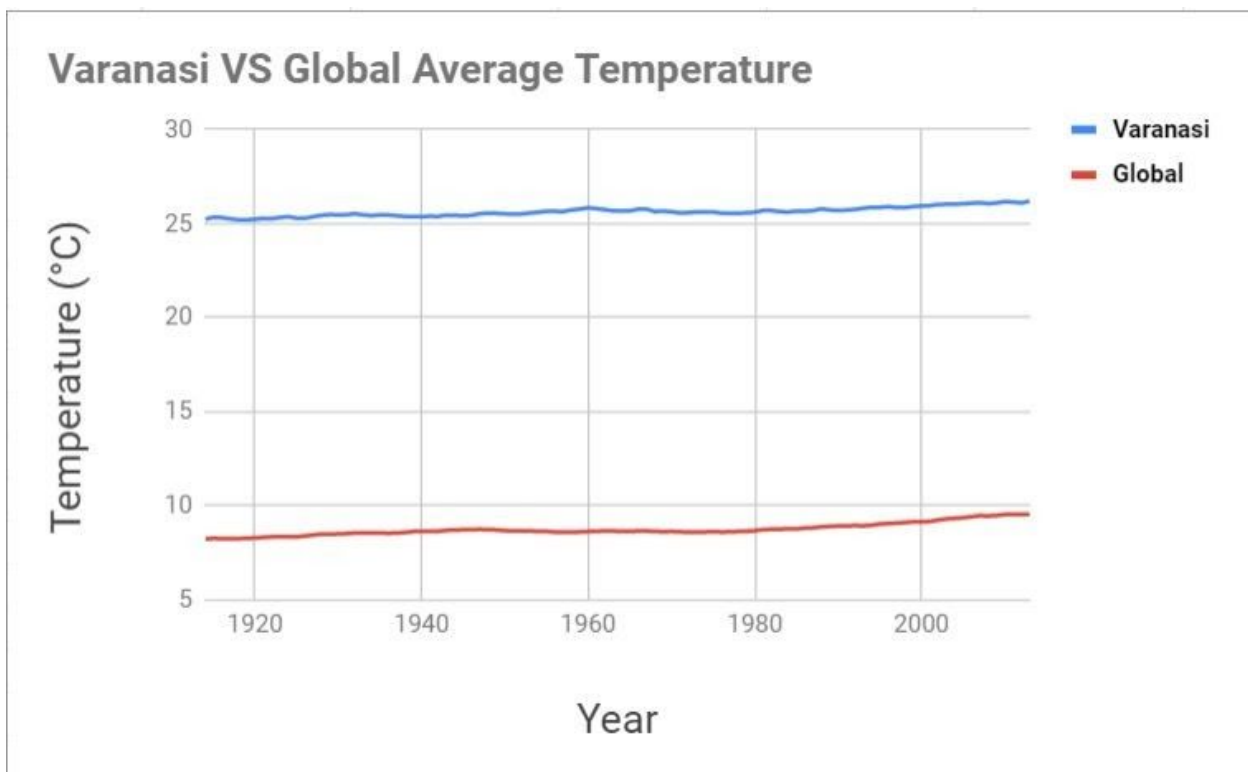
1. Moving average is calculated for a smooth chart.
2. The 10 year moving average is calculated starting from 1913.

Step 3 : Draw line chart.

I have plotted a line chart for global data first to make a clear comparison and then I combined the line for global and Varanasi data in a single chart to make a clear observation.



Global Temperature Data



Comparison between Varanasi and Global Data.

Observations :

1. To get a smooth graph we have to calculate moving average with a big interval otherwise the chart will be messy. I used a 10 year moving average for a smooth graph.
2. Varanasi is very hot compared to the global average.
3. Considering the first chart we have a clear idea that the global temperature is increasing at a very high rate due to industrialization and advent of new technologies.
4. Cutting down trees is the major factor in such a hike in the temperature.
5. As the total carbon emissions are increasing at a very high rate from 1900 and hence temperature rise can be seen.
6. I observed that 50 year MA gives a very smooth chart when compared to 10 year MA.
7. The global temperature seem to decrease after 1950 and then sudden increase can be seen.
8. I check the temperature of India's capital Delhi. It is way hotter than Varanasi due to industries and the crowd. Population density in Varanasi is 2,399 inhabitants per square mile whereas in Delhi it is 25,259. It is a major factor why Delhi is hotter than Varanasi.
9. I observed that there is a **positive correlation** between Varanasi and Global Temperature. The correlation is **0.9218723429** (close to 1) which means highly correlated. If the global temperature rises so does the temperature of Varanasi and vice-versa.

Key Considerations :

1. Use of CORREL() function of Google Docs to calculate correlation between temperatures.
2. Applied moving average to get a smooth graph.

References :

- <https://www.investopedia.com/ask/answers/032515/what-does-it-mean-if-correlation-coefficient-positive-negative-or-zero.asp>
- <https://economictimes.indiatimes.com/news/environment/global-warming/how-rise-in-earths-average-global-temperature-is-affecting-our-planet/articleshow/72039042.cms?from=mdr>
- <https://logopond.com/pdcreative/showcase/detail/102924>
- <https://www.youtube.com/watch?v=onu0h890QAE>