

# Exploring Weather Trends

using

Extracted Data from Database

in

Partial fulfilment of Data Analyst Nanodegree, Udacity



UDACITY

**By Irish Mehta**

## Table of Content

<b>Task .....</b>	<b>2</b>
<b>Softwares and Languages used.....</b>	<b>2</b>
SQL code.....	2
<b>Observations .....</b>	<b>3</b>
<b>References .....</b>	<b>4</b>

# Task

Analyzing local and global temperature data and comparing the temperature trends of where you live to overall global temperature trends. Your goal will be to create a visualization and prepare a write up describing the similarities and differences between global temperature trends and temperature trends in the closest big city to where you live

## Softwares and Languages used

1. Script Query Language (SQL)
2. Microsoft Excel

### SQL code

//from the given database, I need to first check which city is closest to my city, Mumbai

```
SELECT DISTINCT city,country
```

```
FROM city_data
```

```
WHERE country='India';
```

//Closest city whose data is available is Pune

//Now I need to combine the two temperature columns for exporting the Dataset but since the two schema have columns having the same name, I'll rename the two

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

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

//Final step is to join the two tables along the column of year

```
SELECT global_data.year, global_data.global_temp, city_data.local_temp
```

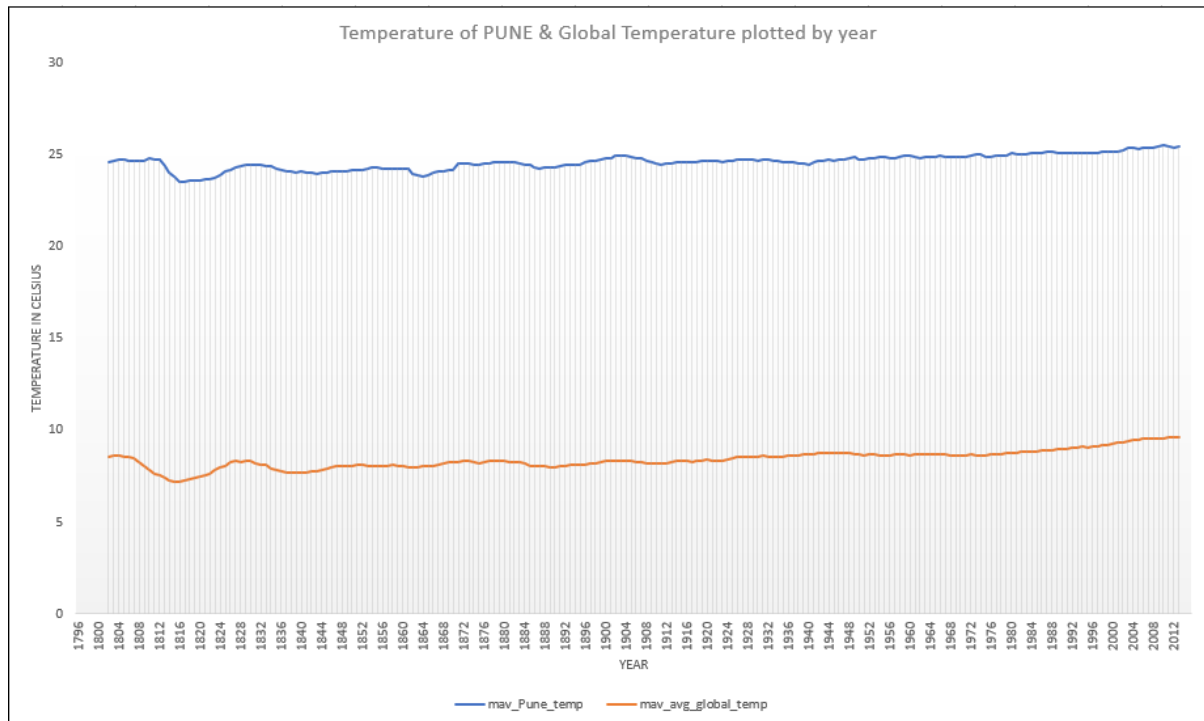
```
FROM global_data JOIN city_data
```

```
ON global_data.year = city_data.year
```

```
WHERE city='Pune';
```

Further steps required usage of Microsoft excel to calculate the moving average of the temperature. A 6-year Moving Average is used to get the following line graph:

(Note that I have changed the variable name from *global temp* to *mav\_avg\_global\_temp* when creating the Moving Average and similarly the *local\_temp* variable to *mav\_Pune\_temp*)



## Observations

1. In the year 1815, there had been a tremendous eruption of Mount Tambora (the largest eruption in history) and caused the temperature worldwide to decrease as the ash dispersed around from the eruption column. This fact is supported by the graph and there is a dip in the temperature of both Pune and the world around the decade 1810-1820.
2. The dip in the curve of Pune in India around the 1890s is mostly because of the great famine which struck the entirety of India during that time.
3. Since the year 1950, the overall temperature of the entire world has been increasing due to the onset of global warming and hence is visible from the graph of both the locations
4. The temperature of Pune is much greater than that of the world as India is situated closer to the equator and has a lot of humidity hence it has a moderate temperature throughout the year. Another major factor is that Pune is situated on the leeward side of the Sahyadri mountain ranges hence it faces severe heat for most of the year as the winds are blocked by the ranges
5. The two variables have a correlation coefficient of 0.89049 which means that there is a high degree of correlation between these two temperatures and most of the trends are common

# References

<https://en.wikipedia.org/wiki/Famine#India>

[https://en.wikipedia.org/wiki/1815\\_eruption\\_of\\_Mount\\_Tambora](https://en.wikipedia.org/wiki/1815_eruption_of_Mount_Tambora)

[https://en.wikipedia.org/wiki/Global\\_warming](https://en.wikipedia.org/wiki/Global_warming)

<https://github.com/VELLALAKAVYA>