**Exploring Weather Trends**

1. Extraction of Data

In this step data was extracted from database where three tables were present namely city\_list, city\_data and global\_data.

* city\_list - This contains a list of cities and countries in the database. Look through them in order to find the city nearest to you.

SCHEMA: Attribute (city (character) , country (character))

* city\_data - This contains the average temperatures for each city by year (ºC).

SCHEMA: Attribute (year (integer), city (character), country (character), avg\_temp(numeric))

* global\_data - This contains the average global temperatures by year (ºC).

SCHEMA: Attribute (year (integer), avg\_temp(numeric))

In this step to extract the required data from database SQL Query was performed of the database which is given below:

select city\_data.year,city\_data.city,city\_data.country,city\_data.avg\_temp , global\_data.avg\_temp as global\_avg\_temp from city\_data inner join global\_data on city\_data.year=global\_data.year where city\_data.city like 'New Delhi'

After this the output generated was downloaded in a csv format and Attribute named were changed for convenience , below is the output generated:

A screenshot of a cell phone

Description automatically generated

1. Processing the Data

As it is clearly visible that the dataset having missing value also and since it has a total of 218 rows and in that city\_avg\_temp\_na attribute have 17 values missing which is 7.79 % of that attribute which is less and normally 25% to 30% missing values are accepted and It is know that the **city\_avg\_temp\_na** attribute is an important attribute and is of numeric type, so we will calculate the mean value and insert it in the place of missing values.

Formula for calculating mean:

=ROUND(AVERAGE(IF(ISNUMBER(D2:D219),D2:D219)),2)

Formula for replacing missing value with mean

=IF(ISNUMBER(D2),D2,$K$1) 🡪 here $K$1 is the cell where mean value is stored.

Given Below is the dataset after performing said operation.

A screenshot of a cell phone

Description automatically generated

Mean was calculated which came around to be 25.17, so all missing values were replaced with it and a new column **city\_avg\_temp** was created.

1. Calculating Moving Average

In this step moving average was calculated in order to smooth out the data, so moving average for a interval of 5 years was done after which the data set looks like this:

Formula for Calculating moving average: =AVERAGE(F2:F6) 🡪 for City Moving Average

=AVERAGE(E2:E6) 🡪 for Global Moving Average

A screenshot of a cell phone

Description automatically generated

Also a Moving Average for a interval of 7 years was calculated to check which gives a smooth visualization

,the data set looks like this:

Formula for Calculating moving average: =AVERAGE(F2:F8) 🡪 for City Moving Average

=AVERAGE(E2:E8) 🡪 for Global Moving Average

A close up of text on a white background

Description automatically generated

Key Consideration for visualization:

1). Plotting year on X-axis

2). Plotting Temperature in Y-axis

3). Calculate the moving average of city and global temperature to smooth the trend line.

1. Visualization of Data

In this step a Line graph was plotted in order to extract out the insights, given below is the line chart.

Here we can see that the 7 year moving average gives a smooth trend line. So for inferences we will use 7 Year moving average visualization. And also chart without moving average is plotted to see the difference and understand the graph better.

Here it can be seen that the plot is not smooth at all.

1. Observations

**Question.** Is your city hotter or cooler on average compared to the global average? Has the difference been consistent over time?

**Answer**: Comparing to the global average the city I live in is hotter than the global average and over the years the difference has somewhat been consistent.

**Question.** How do the changes in your city’s temperatures over time compare to the changes in the global average?

**Answer**: Between 1802 to 1817 the global temperature saw a dip of 2℃ whereas the temperature in my city saw a very little dip and between 1817 to 1820 when global temperature was rising and the temperature in my city was declining and after that the temperature has been fluctuating continuously in my city and is not as steady as global temperature.

**Question.** What does the overall trend look like? Is the world getting hotter or cooler? Has the trend been consistent over the last few hundred years?

**Answer**: If we will look at the overall trend then we can infer that the global temperature has been fluctuating less whereas the temperature in my city is fluctuating continuously and plummet during 1802-1803 when global temperature was steady and also during year 1900 and 1936-37 the temperature deviated whereas the global temperature is steady. Looking at the global and city temperature trend we can easily say that the world is getting hotter because the temperature increases by at least 2℃ after hitting the lowest average. Looking at last few hundred years we can say that between 1850-1950 the trend in my city saw a fluctuation whereas the global temperature was steady.

**Question.** How the temperature in 2012 is relative to 1802 ?

**Answer**: Looking at the visualization it can be infer that both the global and city temperature has seen a increase of average of 1℃. And during the period of 1814-1817 the temperature were lowest for both city and global.