To: Udacity Data Science

From: Varsha Raj

Re: Weather Trends

Document Purpose

A simple data analysis was performed using R programming tool to analyze the trends in local and global average temperature data. Temperature data were imported from the Udacity Database using SQL query. This analysis used local temperature data from Philadelphia, PA and Bangalore, India.

Temperature Data

Global data has temperature records for 1750-2013 periods. Bangalore temperature data have records dating back to 1796 up to 2013 while for Philadelphia, records go all the way dating back to 1746 up to 2015. Some missing data were observed in Philadelphia for the years 1746 through 1749 and 1780. Temperature data for Bangalore had missing observations for the years 1808 through 1812 and 1863-1864. To maintain the integrity of the data and to utilize the maximum amount of data available, temperature data for period 1750--2013 was used in the analysis.

Method

The following query was to download city and wide and global data:

SELECT * FROM city_data

SELECT * FROM global_data

Using the dplyr package in R, further filtering was done to get local data of interest (please see weather_trends.R)

Using the available temperature data, line plots were created for the two cities, Philadelphia (1750-2013) and Bangalore (1796-2013) respectively and were compared against the line plots from the global data (1750--2013). A 5-year moving average was used to make the trends more observable. Any missing data were removed when calculating moving averages. Since the analysis uses nearly 200 years of data, ignoring a couple of missing data does not affect the trends observed. Figure 1 below shows the average temperature data for the two cities compared against the global temperatures. The “rollmean” function found in R tool was used to calculate the moving average.

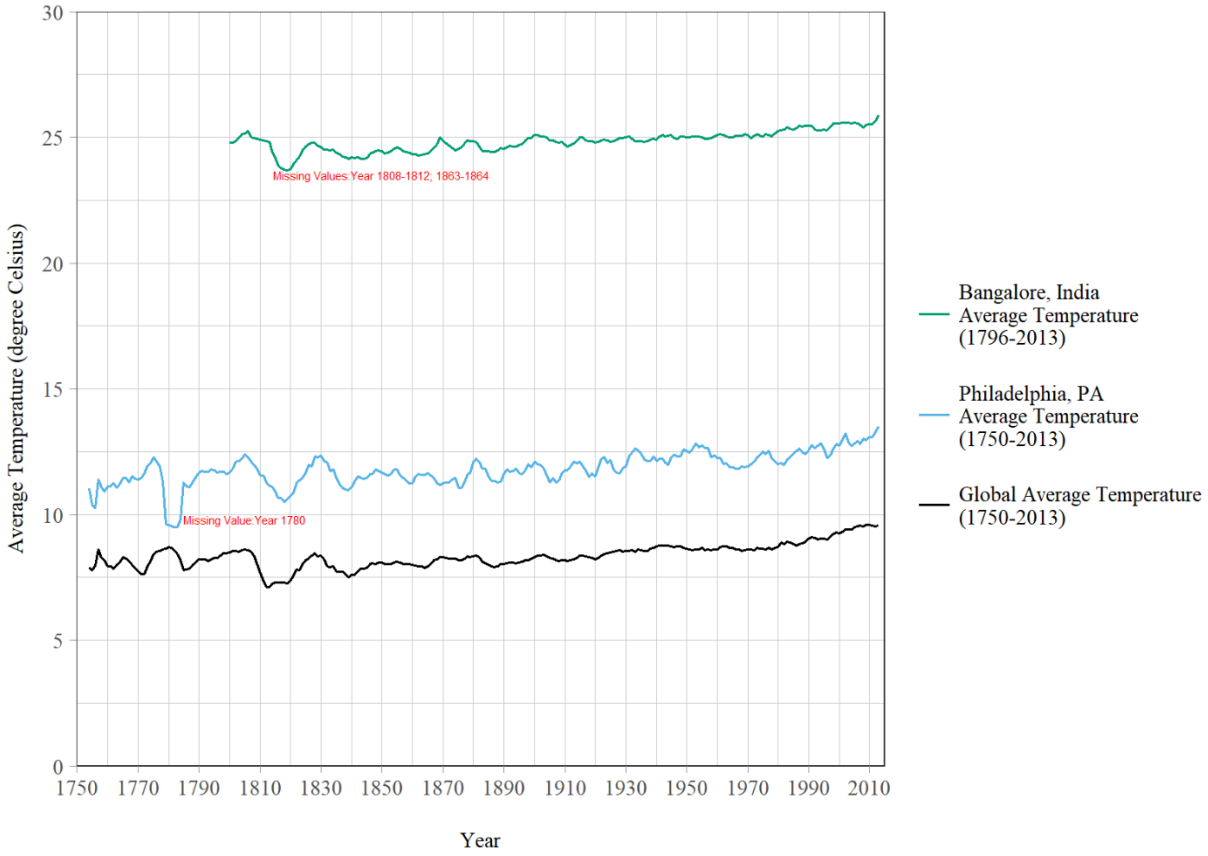

Figure 1: Temperature trends for Philadelphia, PA and Bangalore, India compared to global averages

Discussions

The sudden depression in the Philadelphia line plot during the year 1780 is due to the missing data for that year. Missing data for the years 1808-1812 and 1863-64 explains the depression observed in the Bangalore line plot.

Philadelphia city on average is hotter than the global average temperature. Depending on the months, Philadelphia gets hotter or colder and this could explain some of the fluctuations in the plots when compared to the global temperatures. The temperature in Philadelphia in the year 1952 dropped to around 5 degree-Celsius resulting in a fairly lower 5-year average during the year 1954. This is indicated by a depression in the early 1950's in the Philadelphia line plot while the global averages show a peak rise during that period. From the mid 1700's, Philadelphia experiences a rising limb while global average temperatures have combinations of highs and lows until 1785. Both Philadelphia and global averages show rising limb in the late 1700's and these characteristics can be correlated to the onset of industrial revolution that began somewhere in the 1760's. It is also interesting to note that between the years 1800--1850, both Philadelphia and global average temperatures experience similar combinations of gradual recessing limb and rising limb. Starting from the late 1700's, Philadelphia shows similar temperature patterns as the global averages.

Bangalore shows a much higher than average temperatures because of its proximity to the equator. Bangalore because of its elevation has in fact a much moderate climate (warm summers and cool winters) when compared to other neighboring cities like Mysore that experience more dry heat during summers.

Both globally and locally, trend is the increasing temperatures. One of the biggest contributor to the increasing average temperatures around the world is urbanization and greenhouse gas emissions. Bangalore has shown a remarkable increase in temperature starting from around 1980when compared to Philadelphia and global averages. The biggest factor to this is the tech boom that began around that time resulting in air pollution because of construction and increase in traffic.

Both Philadelphia and Bangalore show an increase in average temperatures in the last 5 years when compared to previous years.

A linear regression test (average temperature as a function of year) was also performed to calculate slopes to estimates degree-Celsius increase each year. The “lm” function found in R tool was used to perform this analysis. The analysis indicated that globally, the average temperature is rising by 0.46% (0.0046 degree-Celsius) each year. Philadelphia and Bangalore also show a positive trend as observed in the plots with temperatures increasing by 0.006 degree-Celsius and 0.005 degree-Celsius each year respectively.