Udacity NAND

Project One

Explore Weather Trends

James S Wood

Udacity Data Analyst Nanodegree

Exploring Weather Trends Project

The aim of the project is to analyze and visualize local and global temperature data. I have considered (Raleigh NC) as my local city and compared it with overall global temperature trends. This report focuses on providing four interesting insights/observations about the change in temperature trends. The weather data (local and global) have been visualized using Excel, where SQL query is used to extract the data from Udacity website.

Introduction

The change in weather trends has always been an interesting topic among scientist, politicians, environmentalists, and others. The goal of the project is to compare and analyze the similarities and dissimilarities between the local and global temperature data. The observations can be drawn by visualizing the temperature data. The report consists of following sections: I. Data extraction from database, II. Data manipulation methods and tools, III. Data visualization, IV. Data interpretation.

Data Extraction

The data is provided by the Udacity as a part of the nanodegree program. The data needs to be extracted through workspace which is connected to the database using SQL query. There are three tables in the database:

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.

city_data - This contains the average temperatures for each city by year.

global_data - This contains the average global temperatures by year.

The SQL query used to extract the data: (A) Local city data (Raleigh) and (B) global data respectively:

A. Write a SQL query to extract the city level data. Export to CSV.

```
SELECT *
FROM city_data
WHERE CITY= 'Raleigh';
```

B. Write a SQL query to extract the global data. Export to CSV.

```
SELECT *
FROM global_data;
```

Data Manipulation

I have used MS Excel for data manipulation. I have created columns with local_avg_temp and global_avg_temp for storing the moving average data. The calculation for moving average is same as explained in projects “moving average section”. I have calculated moving average for 10 years both for global and local data.

Data Visualization

To visualize the data, I have used the line charts provided within the Excel. I have analyzed data by plotting the 3 charts, (1) Global moving average of temperature over a period of 10 years (2) Local moving average of temperature over a period of 10 years (3) Comparison between local and global moving averages. Observations were made and explained in the data interpretation section.

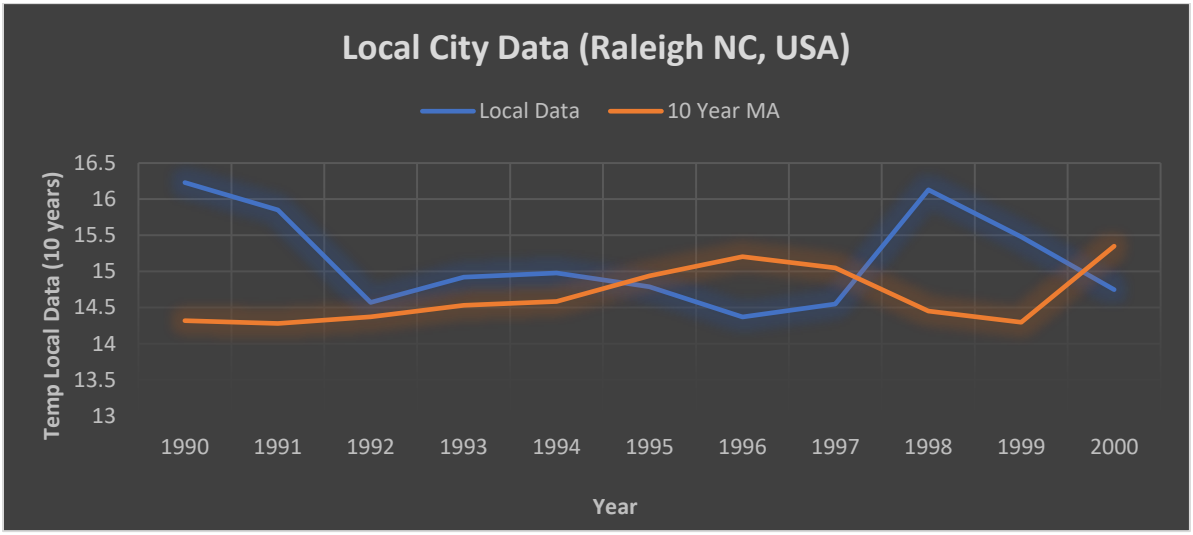

Figure 1. Local city (Raleigh) temperature data taken over 10 year moving average

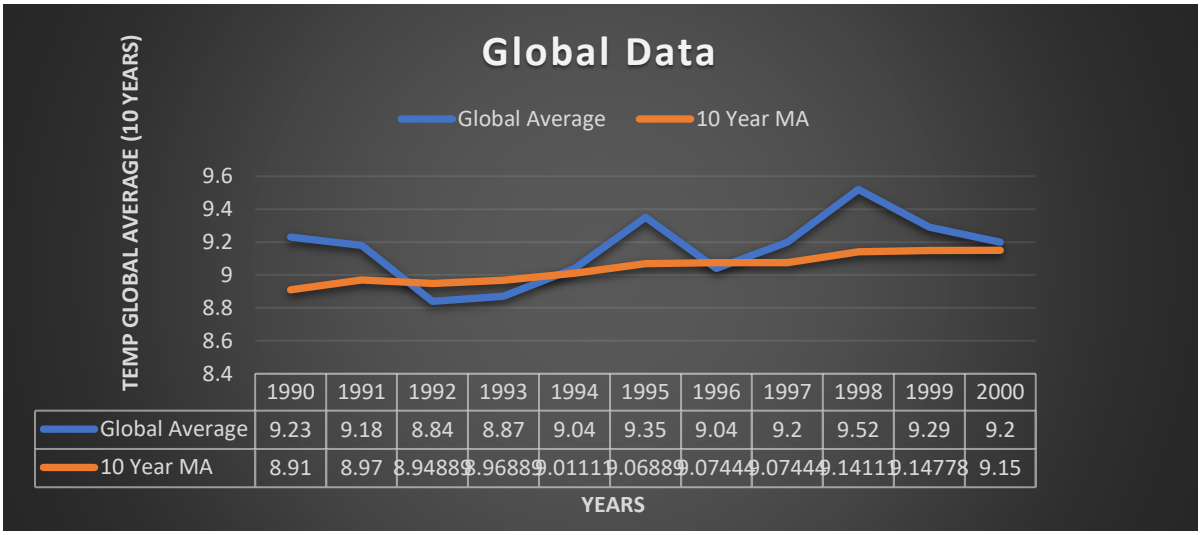

Figure 2. Global temperature data taken over 10 year moving average

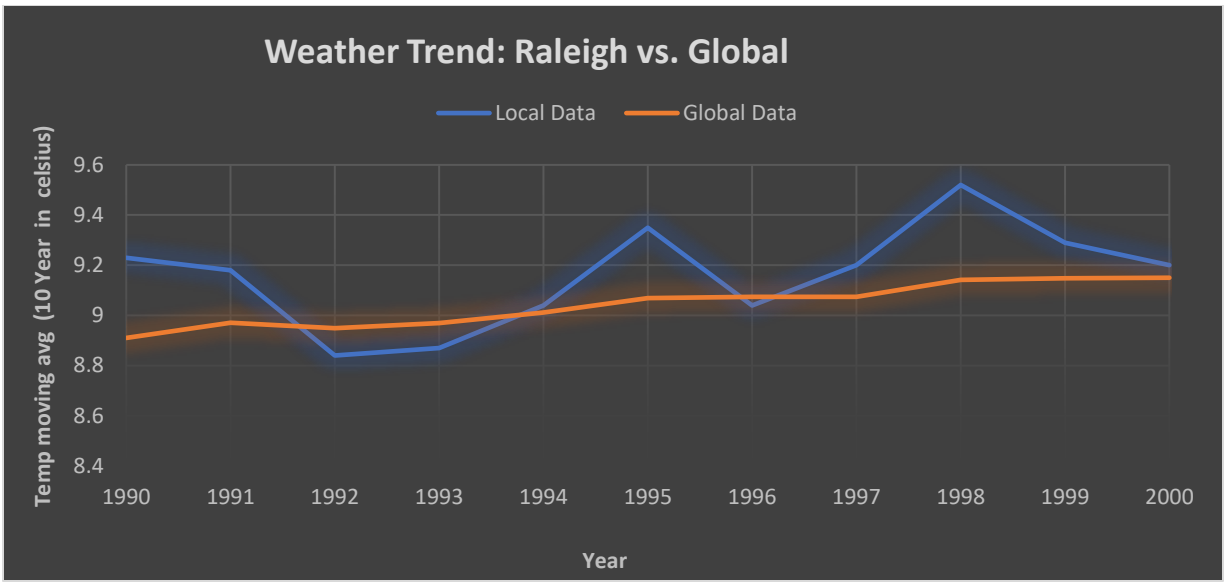

Figure 3. Comparison of Raleigh and Global temperature trend

Data Interpretation

There following observations were drawn after analyzing the data from line charts:

(1) Raleigh NC is hotter on average compared to global average temperature trend and the difference is consistent over time.

(2) Global and Raleigh NC average temperatures have seen a consistent incremental change around the starting of 1990, before that there were fluctuations in average temperature.

(3) The overall trend shows increase in temperature in both global and local level. This means the world is getting hotter. This trend is consistent over the last one hundred years.

(4) There is a significant increase in temperature can be observed in global data.