Data Analyst Nanodegree

Explore Weather Trends

April-26,2020, Project 0

Bayan Bin Manea

Introduction

In this project, I have worked with the temperature database from Udacity workspace, to extract global and city temperatures data, with the aim of analyzing, visualizing, observing, and comparing my findings.

Tools Used:

1. **SQL:** to extract data from the database.
2. **Google Sheet:** To calculate the moving averages for global and city data, and to plot a line chart.

Step 1: Extract data from the database

1. Extract data from the global_data table

```
SELECT * FROM global_data
```

2. Check which cities are available in Saudi Arabia in city_list table

```
SELECT * FROM city_list

WHERE country = 'Saudi Arabia'
```

3. Extract Riyadh city data from city_data table

```
SELECT * FROM city_data

WHERE city ='Riyadh'
```

Step 2: Working with CSV files

Working with Riyadh.csv file I noticed that there were two missing values in the year column so I decided to research the problem on the web and the knowledge section, I found out that the best way to deal with missing values is to either drop the value or calculate the mean. I choose to calculate the mean for the missing values.

year	city	country	avg_temp	Average
1843	Riyadh	Saudi Arabia	24.74	25.21414201
1844	Riyadh	Saudi Arabia	15.45	
1845	Riyadh	Saudi Arabia	20.82	
1846	Riyadh	Saudi Arabia		
1847	Riyadh	Saudi Arabia		
1848	Riyadh	Saudi Arabia	24.56	
1849	Riyadh	Saudi Arabia	24.8	
1850	Riyadh	Saudi Arabia	24.34	

I combined both Riyadh.csv and Global.csv into a single sheet and changed the headers into meaningful names to make plotting the line chart easier.

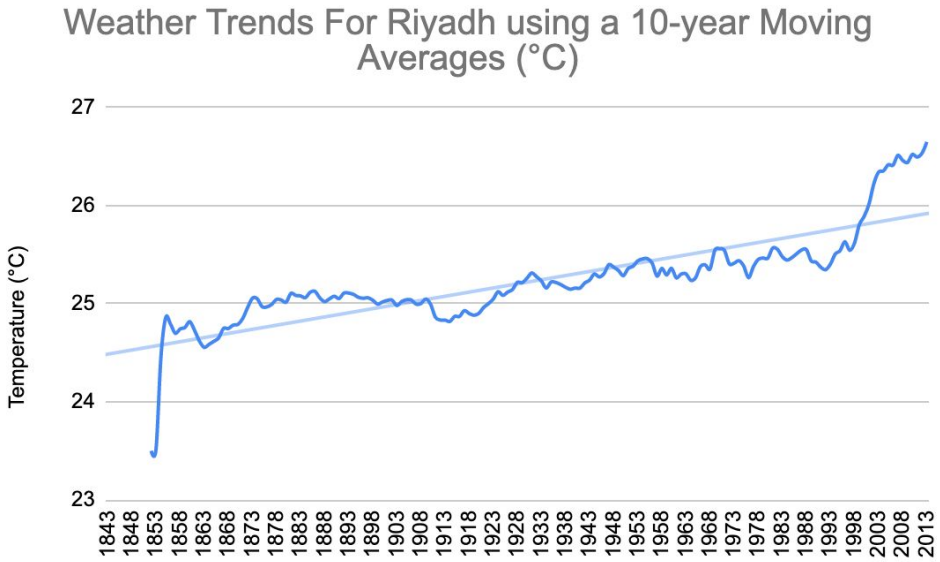
city	Riyadh_years	country	Riyadh_avg_temp	Average	Riyadh	G_year	Global_avg_temp	Global
Riyadh	1843	Saudi Arabia	24.74	25.21409357		1750	8.72	
Riyadh	1844	Saudi Arabia	15.45			1751	7.98	
Riyadh	1845	Saudi Arabia	20.82			1752	5.78	
Riyadh	1846	Saudi Arabia	25.21			1753	8.39	
Riyadh	1847	Saudi Arabia	25.21			1754	8.47	
Riyadh	1848	Saudi Arabia	24.56			1755	8.36	
Riyadh	1849	Saudi Arabia	24.8			1756	8.85	
Riyadh	1850	Saudi Arabia	24.34			1757	9.02	
Riyadh	1851	Saudi Arabia	25.03			1758	6.74	

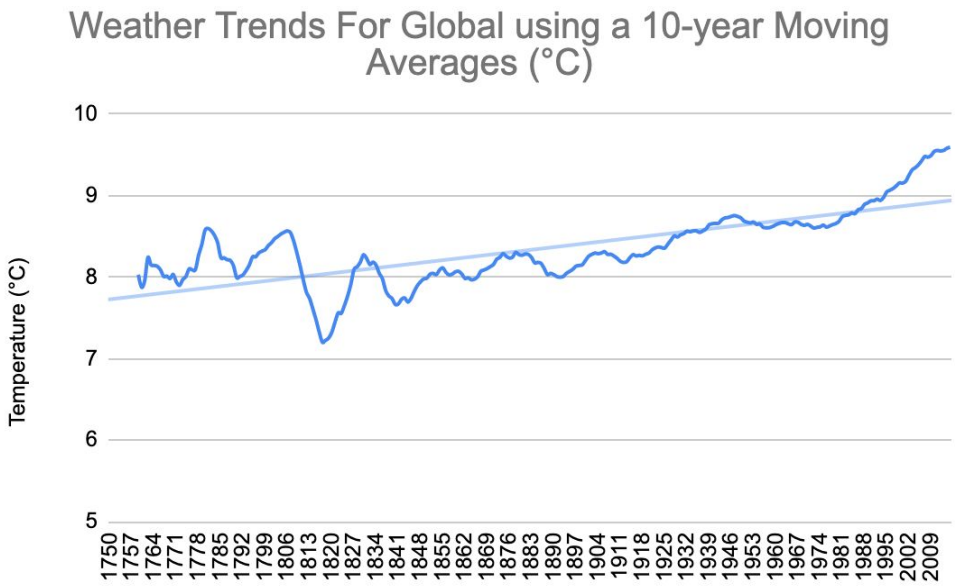
Step 3: Calculating the Moving Average for both Global and Riyadh temperatures

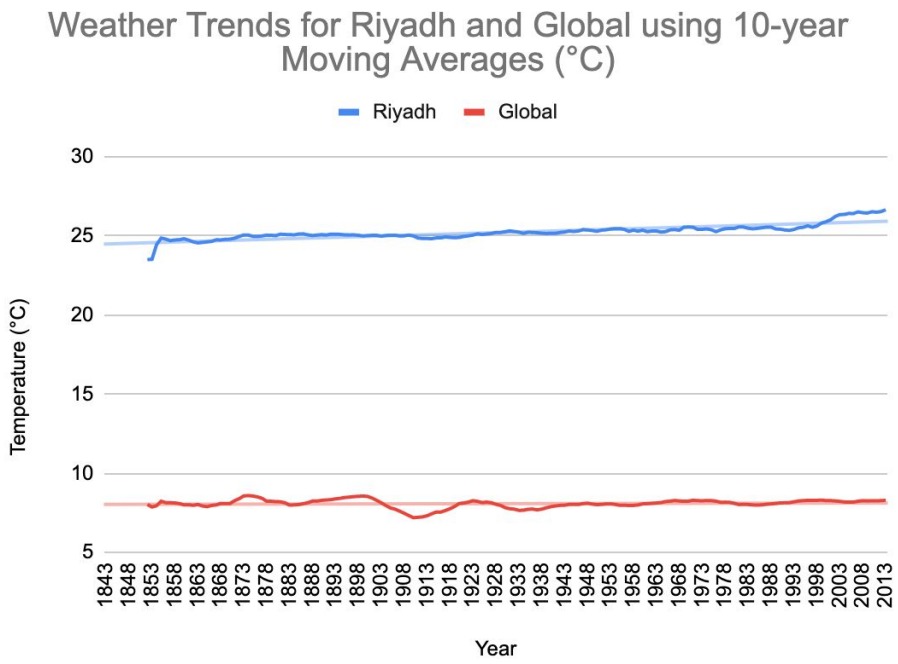
I have done 10 years moving average, I did it by using the Average formula in Google Sheets.

Step 4: Plotting a Line Chart

I have plotted a Line chart for Riyadh and Global data separately to get a clear understanding of the temperatures and to observe the difference between Riyadh temperatures and the Global temperatures.







Observations & Findings

- Overall, the chart shows that temperature degrees are increasing over the years globally and in Riyadh as a result of climate change.
- Temperatures in Riyadh start with their lowest level in 1852 with 23.5° degrees then continue to rise up gradually.
- Global temperatures start at their lowest level in 1852 with 8.03° degrees then slightly vary over the years.
- The global temperature fluctuates between 7° to 9° degrees and continues to rise up slowly.
- The lowest global temperature declined to 7.2° degrees in 1913, then rose up to 8.2° in 1992.
- Temperatures in Riyadh rose up gradually from 25° - 26° In the years 2000 to 2013 with an increased rate of approximately 0.1 degrees.

- The Riyadh average temperatures ranged between 25° - 26° degrees.
- There is a major difference between global temperatures and Riyadh temperatures, global weather temperatures are considered to be colder than Riyadh temperatures.

References:

1. How to handle missing data(statistics)? What imputation techniques do you recommend or follow