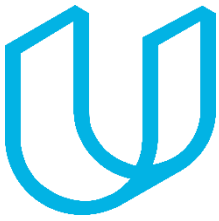

Udacity

Data Analytics Nanodegree

Project 01:

Exploring Weather Trends

Riyadh, Saudi Arabia Vs. Global

Student

Munirah Salah Almithn

September 9, 2019

Tools and Techniques being Used

1- Structure query language (SQL)

SQL is a language used in this project to retrieve the required records form Udacity's database.

2- Microsoft Excel

Excel is a useful tool used in this project to calculate the moving average and create data visualization for making observations.

3- Microsoft Word

Word is used for writing the report.

Data Extraction

This project aims to analyze and explore global and local temperature trends. Therefore, it is important to collect useful and relevant data that will provide the best results.

Database schema:

Table 1:Database schema

Table Name	Description
city_list	This table contains a list of cities and countries in the database.
city_data	This table contains the average temperatures for each city by year (ºC).
global_data	This table contains the average global temperatures by year (ºC).

The following SQL queries have been applied to retrieve data from provided database.

- **To retrieve global temperature:**

```
SELECT *

FROM global_data
```

- **List all cities in my country to choose a nearby city:**

```
SELECT *

FROM city_list

WHERE country = 'Saudi Arabia'
```

- **To retrieve local temperature (Riyadh, Saudi Arabia):**

```
SELECT year,avg_temp

FROM city_data

WHERE country = 'Saudi Arabia' AND city = 'Riyadh'
```

Key Considerations

1- Data cleaning

Typically, most of the collected raw data contain fields that are missing. Thus, data pre-processing is essential to clean and organize these data for further analysis. The local dataset has two missing values in Avg_Temp attribute as shown in Figure [1]. To have accurate results, the missing data was handled by filling the cells with the minimal average value which is 15.45.

	A	B
1	year	Avg_Temp
2	1843	24.74
3	1844	15.45
4	1845	20.82
5	1846	
6	1847	
7	1848	24.56
8	1849	24.8
9	1850	24.34
10	1851	25.03

Figure 1: Missing Values

2- Calculate moving average

The moving average of 11-yeare was calculated using AVERAGE() formula in excel in order to get a smooth graph.

Data Visualization

The Line Chart is used in this project as a candidate visualization tool because it clearly shows trends over time (years, months, or days). Figure 1 shows global temperature and Figure 2 shows Riyadh temperature.

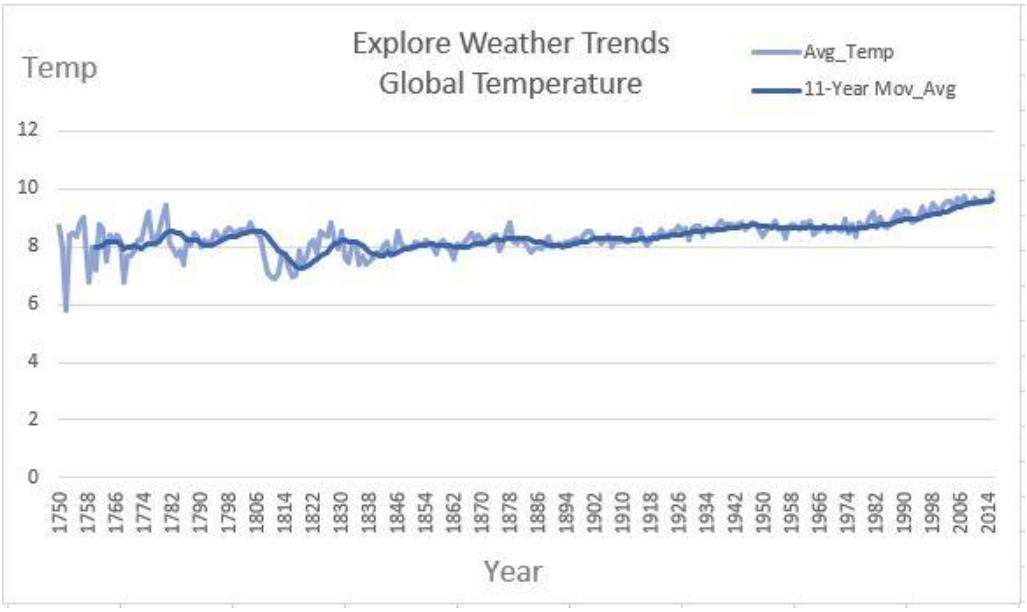

Figure 2: Global Temperature

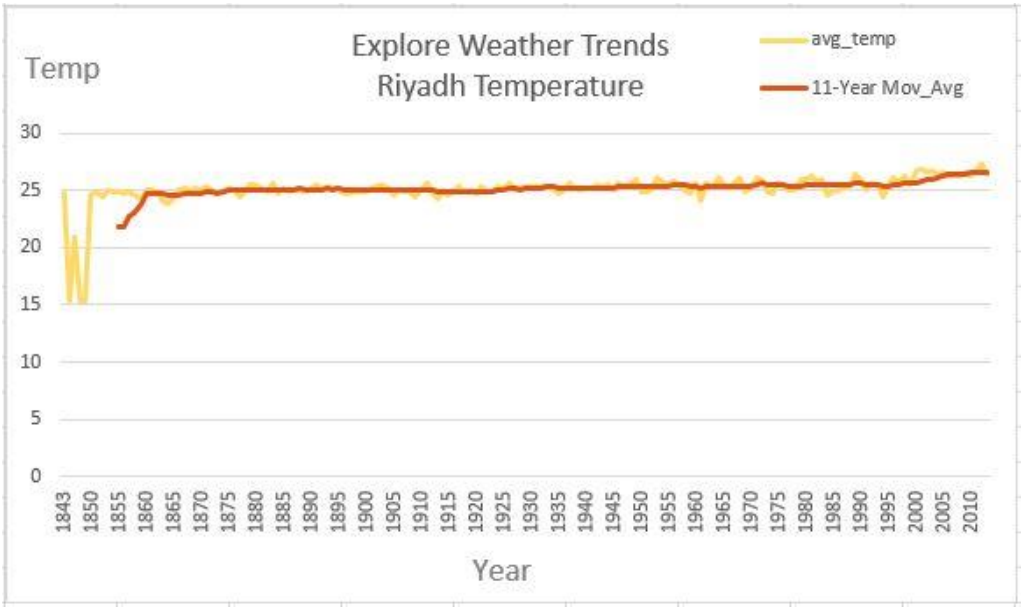

Figure 3: Riyadh Temperature

Observations

1- Riyadh moving average temperature varies between 21.85 ºC to 26.52 ºC while global moving average temperature varies between 7.23 ºC to 9.60ºC. This indicates that Riyadh temperature is much hotter compared to global temperature.

2- The moving average temperature of Riyadh shows a consistent change from 1860 to 1997. In 1998, the temperature of Riyadh is continuously starting to increase.

3- The moving average temperature of global shows inconsistent change from 1860 to 1850. The temperature gets consistent change till 1921 (in range between 7.9 - 8.3). After that, the temperature of global is starting to increase over the years.

4- Generally, the moving average temperature of Riyadh and global is consistently getting hotter over time.