

Name: Hanaa Ballahmar
Date: 6th September, 2019
Country: Saudi Arabia

UDACITY - Data Analyst Nanodgree

Overview

In this project, I will analyze the local temperature of Makkah in Saudi Arabia and compared it with the global temperature data. I have been taken the data from Udacity portal to meet the criteria of this project.

The used tools

- **SQL:** To extract the data from the database.
- **Excel:** To look up and understand the attributes of the data to write a report of a project.
- **Python:** To calculate the moving average and plotting a line chart.
- **ANACONDA - Jupyter Notebook:** To write python code and making observations.

Extraction of Data

Data were retrieved from the database using SQL - Structured Query Language. First, retrieve cities in Saudi Arabia to search which is the closest to where I live from city_list table by using this SQL command:

```
1  select * from city_list Where country = 'Saudi  
    Arabia'
```

The result of previous select statement given two city Makkah and Riyadh. Second, retrieve the temperature of Makkah data from city_data table by using this SQL command:

```
1  select * from city_data Where country = 'Saudi  
    Arabia' and city = 'Mecca'
```

Last, retrieve the temperature of global data from global_data table by using this SQL command:

```
1  select * from global_data
```

Name: Hanaa Ballahmar

Date: 6th September, 2019

Country: Saudi Arabia

There are two tables to analyze data which make confused due to redundancy of data. Therefore, we joined it's into one table to analyze without redundancy of data. We found the SCHEMA for both city_data and global_data table contains the same column named 'avg_temp'. So, I have changed the names of the columns respectively in order to have distinct columns and retrieve the final data from a database from multiple tables by using this SQL command:

```
1  alter table city_data rename column avg_temp to
    local_avg_temp;

1  alter table global_data rename column avg_temp to
    global_avg_temp;

1  select global_data.year,
    global_data.global_avg_temp,
    city_data.local_avg_temp
2  from global_data,city_data where (global_data.year
    = city_data.year) and
3  (city_data.city = 'Mecca' and city_data.country =
    'Saudi Arabia')
```

Moving Averages:

The rolling Average has been calculated to smooth out data and to make it easier to observe the trends for the short term when it is shown in the charts. The Moving Average has been calculated for every 15 years to every single data but the first 15 years are not included. There were some missing data in local temperature. I filled that data by calculating the mean of the values of the whole column and then filling that mean to missing cells. I have been used Python to calculate the Moving Average Using built-in Functions such as (“rolling”, “mean”).

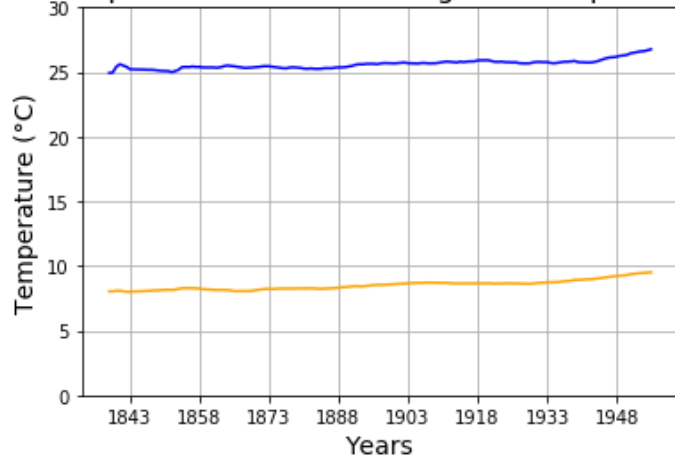
Name: Hanaa Ballahmar

Date: 6th September, 2019

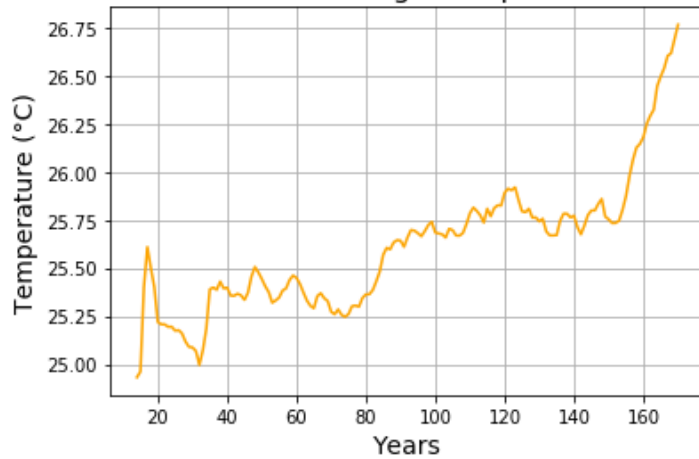
Country: Saudi Arabia

Line Chart for the data:

Temperature in Makkak vs global temperature



Makkak Average Temperature



Global Average Temperature



Name: Hanaa Ballahmar

Date: 6th September, 2019

Country: Saudi Arabia

Observations

- My main concern to visualize the weather trends is observing an increase or a decrease in the moving average temperature.
- The global temperature trends represented in blue color whereas Makkah temperature was represented in orange color.
- Since 1875, I found that the average global and Makkah temperature is increasing over time due to climate change.
- The difference between year 1843 and ~2013 in temperature is more than 1.5 degrees in the Global average chart (increasing).
- The difference between year 1843 and ~2013 is more than 2.4 degrees in Makkah City average chart (increasing).
- The overall trends either global or Makkah are increasing gradually day by day. The world is getting hotter based on the trends above.