

# Explore Weather Trends

Ricardo Marques

**Abstract**—In this project, the main objective is to analyze local and global temperature data and compare the temperature trends where I live to overall global temperature trends.

**Index Terms**—SQL Query, Moving Averages, Data Visualization.

## 1 INTRODUCTION

Nowadays global warm is a central topic in many debates where some believe that our planet is getting hotter most influenced by the greenhouse gases emission in our production model. Although others prefer to believe that the Earth has its own geological cycle where temperatures oscilates by warming or cooling the climate. Despites this discussion, it's interesting to analyze the global and local temperature data in order to check this trends in a timescale.

## 2 DATA ACQUISITION

The LaTeX text editor was used to write this report. In this project, Udacity provided global and local temperature data stored in an internal SQL data schema requiring some SQL query to extract this data in .csv file. Global and local temperature and data was extracted by using the following SQL query, note that I filtered Rio de Janeiro as my city in SQL data schema.

```
SELECT year, avg_temp
FROM global_data
```

```
SELECT year, country, city, avg_temp
FROM city_data
WHERE city = 'Rio De Janeiro'
```

After extraction I had 2 files, one for global temperature (global\_temp.csv) and other for local temperature (city\_temp.csv). I used Excel to read them and synchronize the time interval because global data begins in 1750 and Rio de Janeiro data begins in 1832, so I setted them to begin both in 1832.

## 3 DATA VISUALIZATION

In order to compose data visualization I used Jupyter Notebook (Anaconda Python 3.7) and loaded 2 modules Pandas and Matplotlib. I also applied the moving average to smooth the noisy shape of lines.

```
import pandas as pd
from matplotlib import pyplot as plt
```

Firstly, I plotted 3 moving averages for global temperature data to find out the best window that suited my data: 1, 5 and 10 years as shown in Figure 1. The same was done with Rio de Janeiro temperature data in Figure 2. Both plots

revealed that the 10 years moving average was enough to turn the lines smoother.

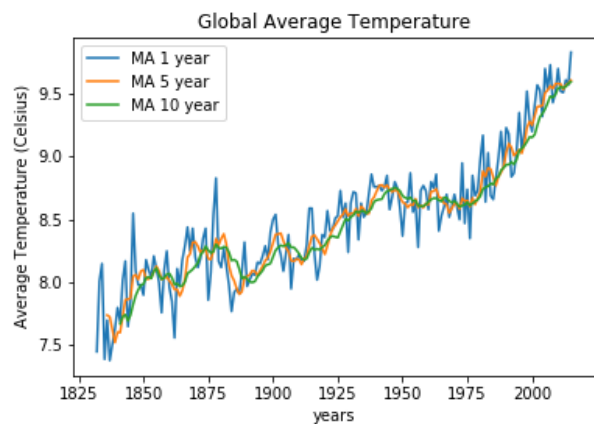


Fig. 1. Global average temperature.

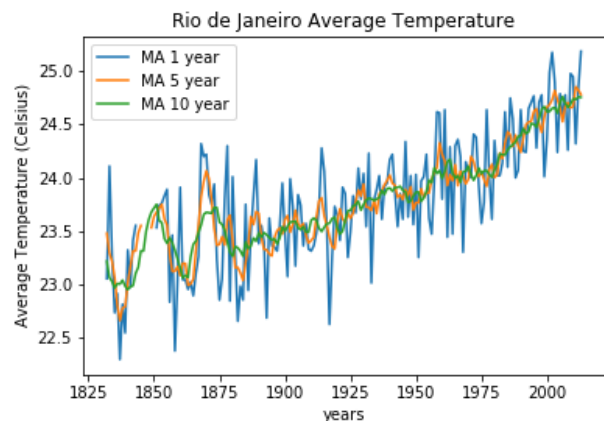


Fig. 2. Rio de janeiro average temperature.

Then I plotted both global and Rio de Janeiro temperature data in the same graph as shown in Figure 3. We can see a huge gap between lines because of the difference in the mean temperature. In order to close the gap, I normalized the data by dividing all temperature point by the first point of each series, global and Rio respectively in Figure 4. After this final polishment the data was ready to be analysed.

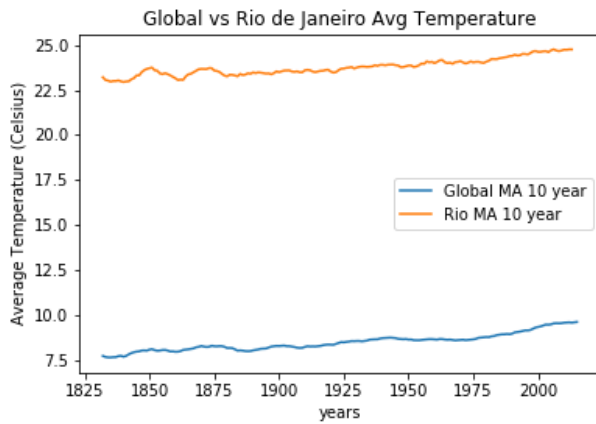


Fig. 3. Average temperature global and in Rio de Janeiro.

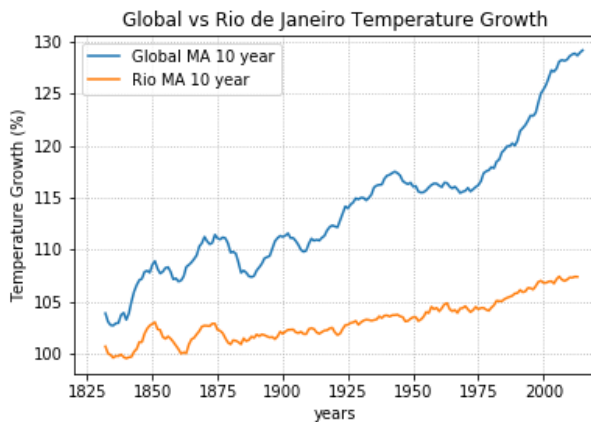


Fig. 4. Temperature growth of global and in Rio de Janeiro.

## 4 RESULTS

At first glance, Figure 3 shows that the average temperature in Rio de Janeiro was higher than the average global temperature and, as a consequence, there's a gap between lines. Figure 4 shows that the global average temperature had a higher growth rate than the Rio de Janeiro one because of its higher gradient (curve inclination) over the time. By analyzing the final point in 2013 we can check that the global average temperature rose about 25% as well as Rio one rose about 6% in the same period.

## 5 FINAL REMARKS

The aim of this work was achieved in accord to rubric supplied by Udacity, I was able to extract dataset from SQL, read the .csv files using Python. Four line graphs were plotted using moving averages strategy, and this allow us to analyze data properly and get to some conclusions with regard to global and Rio average temperature over the time.