Daniel Kuroski

Project 01: Explore weather trends
Data Analyst Nanodegree

Introduction

The project consists in analyzing local and global data from Udacity's database containing yearly average temperature trends.

SQL was used to fetch that data and download them as csv files. The data has been analyzed using Python programming language using IPython Notebook (Jupyter).

Progress outline

Analysis and data fetching

SQL query was used to fetch data and download two csv files:

- One containing metrics about global temperature with **year** and **average temperature**
- Other containing metrics about cities around the world with **country, city, year** and **average temperature**

used queries

Description	SQL command
Select global data	`SELECT * FROM global_data`

Select all cities data set	`SELECT *`
	`FROM city_list`
	`LEFT JOIN city_data ON city_list.city = city_data.city`

Data processing

After loading all csv's into my notebook I have:

- Listed all cities from my country to find the closest one
- Saved all city data to work later
- Cleaned all inconsistent records (null ones) and filled them with the median of the average temperature of Curitiba
    - In total there was **7** records that were null

Moving averages

As indicated on the exercise, I calculated the moving average (MA) using **7 years** to get the smooth line chart.

Python code was used for calculating the MA using Pandas built-in function:

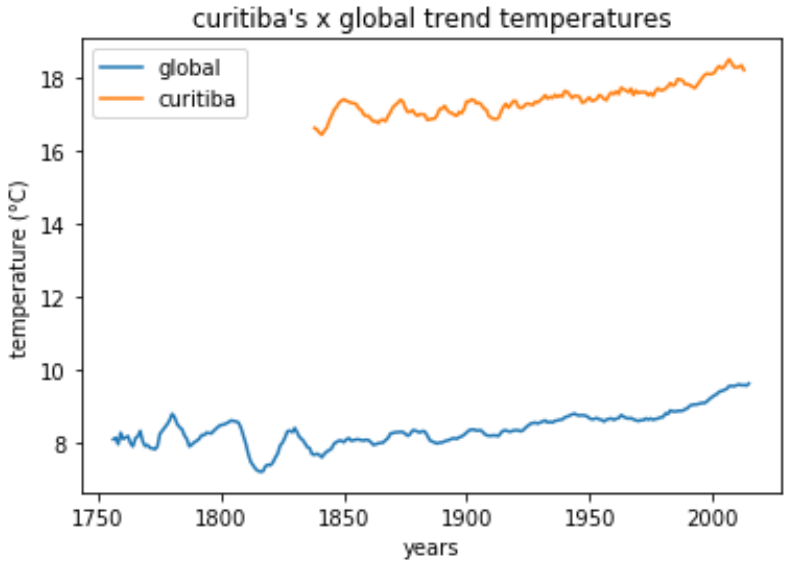
- **rolling:** provides rolling window calculations
- **mean:** return the mean of the values for the requests axis

code

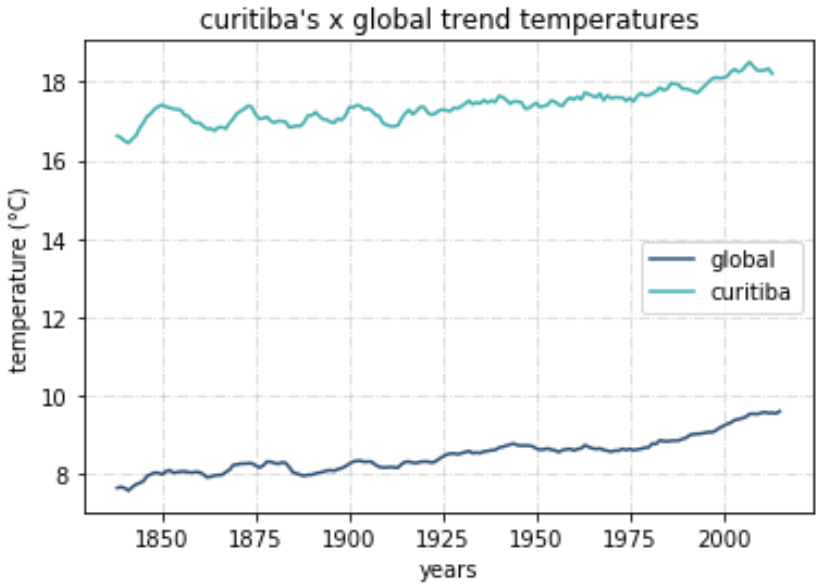
```
global_moving_average = global_data_frame["avg_temp"].rolling(7).mean()
curitiba_moving_average = curitiba_data_frame["avg_temp"].rolling(7).mean()
```

Line chart for Curitiba and Global temperature

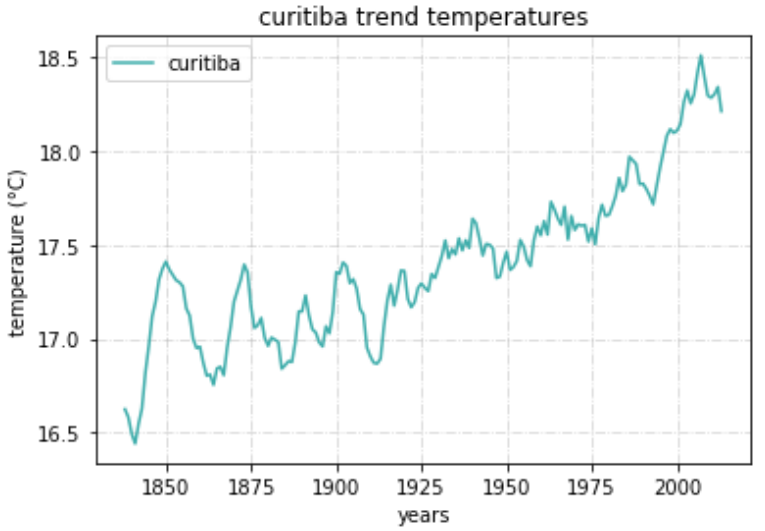
In the chart above, you can see that Curitiba doesn't have enough records compared to the global.

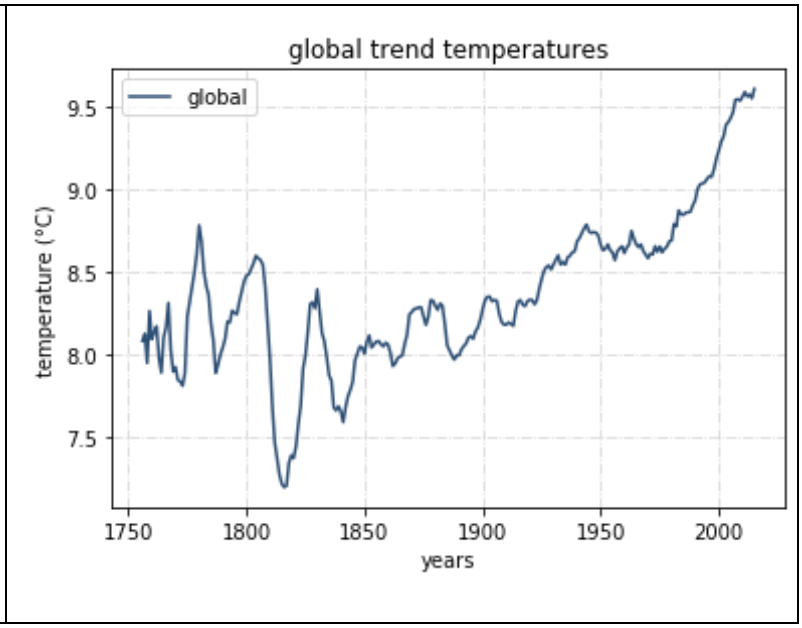


I have filtered the global data to start from the first record of Curitiba, leading to this chart:



We can see either the Curitiba and global average in separated charts:





Observations

- We can notice that the temperature is raising over the years due to climate change
- **Curitiba** city average temperature varies between **16.44 to 18.50 °C**
- **Global** average temperature varies between **7.19 to 9.60 °C**
- Comparing the global and Curitiba average temperatures then the Curitiba city is **hotter**
- Global temperature started to raise since middle of 1820 which happened to be the beginning of the industrial revolution and you can also see that in Curitiba's trending
- In 2013 Curitiba's average temperature is **0.96 °C** hotter than it's first record in 1832
- In 2015 global average temperature is **1.10 °C** hotter than it's first record in 1750
- Curitiba city is getting hotter over time

Conclusion

The evidence suggests that the global temperature is raising over the years which support the case of climate change.