Andres Crucetta Nieto
Udacity Data Analyst Nanodegree
Project 1: Weather Trends

What tools did you use for each step? (Python, SQL, Excel, etc)

- For the first step I used SQL. This proved to be a good tool to extract the data and manipulate it appropriately
- For the second step, I decided to use Google Sheets. At first, I thought of using Python, however, due to how simple the exercise was, I decided sheets would be an easier way to portray the data.
- For the third step, I used Tableau and SQL. With SQL I managed to extract the data and with Tableau I was able to calculate the Moving Average and visualize different cities at the same time.

How did you calculate the moving average?

I calculated the moving average for step 2 by taking the average of the previous 7 years for each row starting in 1756. I didn't portray the values from 1750-1755 since I assumed these would be contained in the Moving Average.

I calculated the moving average for step 3 by using Tableau's table calculation for moving Average. I modified it within the interface so that it would include the previous 7 values for each city.

What were your key considerations when deciding how to visualize the trends?

My main considerations were to have the global trends as well as the Chicago trends within the same graph. Doing so would allow the user to detect variations in the trends more easily. Another key consideration was the amount of data and making sure that I balanced the volatility of the graph with the moving average function (or a similar method)

STEPS

Step #1: Extracting the Data

For this step I used SQL, the code is below:

```
SELECT city_data.year, city_data.city, city_data.avg_temp, global_data.avg_temp global_temp
FROM city_data
JOIN global_data
ON city_data.year = global_data.year
AND city_data.city IN('Chicago','New York','Los Angeles','Berlin');
```

Step #2: Calculating the Moving Average

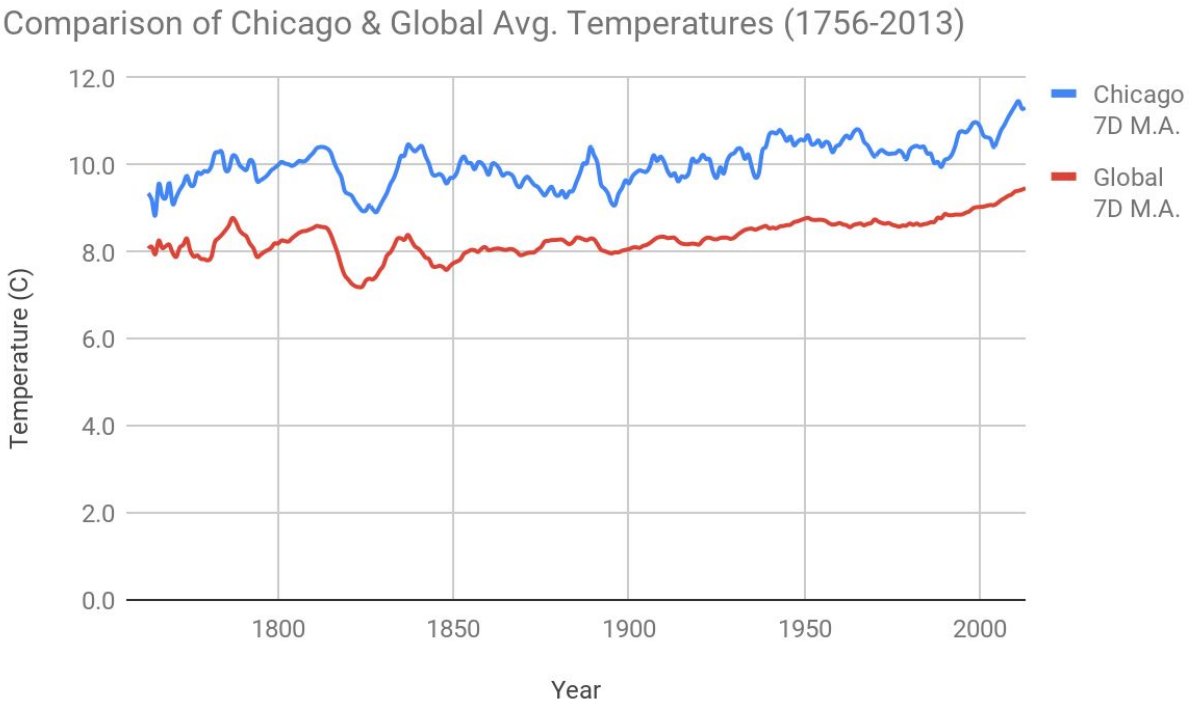


Figure 1: Chicago vs. Global Avg. Temperatures (1756 - 2013)

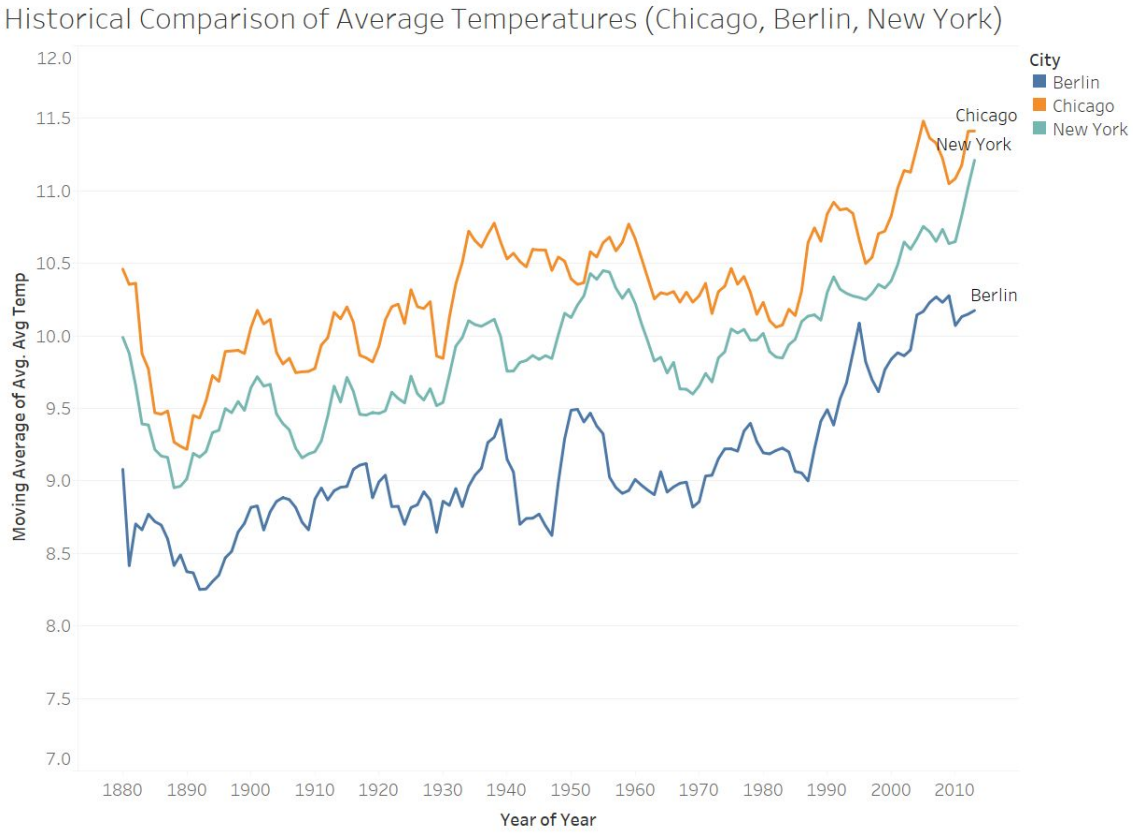
Observations:

1. **The world is getting warmer:** The general trend shows a slight increase as we get closer to the 2000's. I would assume this is a result from CO2 pollution and countries like China going through industrialization.
2. **Peak in temperatures in Chicago during the 1880's:** During the 1880's, there seemed to be an increase in temperatures in Chicago that wasn't present in Global

temperatures. I would be interested in finding out whether this had to do with the industrialization of the city and the levels of pollution.

3. **Peaks & Troughs in Chicago:** Chicago seemed to have a higher variability of temperatures throughout the years. My assumption would be that global temperatures don't portray that volatility due to the vast sample size in which it consists of.
4. **Chicago and Global temperatures have gone under similar trends:** Overall, Chicago has followed world temperatures throughout the years. Even though there are some slight differences, in the long run, they hold a steady increase pattern.

Step #3: Comparing my Favorite Cities with Tableau



Observations:

- All three cities increased their temperatures over the last 100+ years
- Something happened in New York around the 1950's that made it tie Chicago's temperature.
- Chicago had a large increase in temperature on the 2000's as compared with Berlin and New York.