

https://www.udacity.com/

|  |
| --- |
| 703-675-3323 |
| Smart Phone |
| jonlipke@gmail.com |
| Envelope |
| Jonathon K. Lipke |
| User |

Explore Weather Trends

In this project, you will analyze local and global temperature data and compare the temperature trends where you live to overall global temperature trends.

3/29/2019

Jonathon K. Lipke

Data Analyst Nanodegree Program

Contents

[Data Analyst Nanodegree Program 1](#_Toc4749716)

[I. Executive Summary 3](#_Toc4749717)

[II. Average Temperature of Alexandria, VA and the Global Average 3](#_Toc4749718)

[III. 3 Year Moving Average Analysis 5](#_Toc4749719)

[IV. 10 Year Moving Average Analysis 6](#_Toc4749720)

[V. Methodology 7](#_Toc4749721)

[**Normalizing the data** 7](#_Toc4749722)

[**SQL Script** 7](#_Toc4749723)

[**Excel modification and Data visualization** 8](#_Toc4749724)

[Figure 1. Average Temperature of City and Global Average 4](#_Toc4749736)

[Figure 2. Outliers from Average Temperature Data 4](#_Toc4749737)

[Figure 3. 3 Year Moving Average 5](#_Toc4749738)

[Figure 4. Outliers from 3 Year Moving Average 6](#_Toc4749739)

[Figure 5. 10 Year Moving Average 6](#_Toc4749740)

[Figure 6. Outliers from 10 Year Moving Average 7](#_Toc4749741)

## Executive Summary

The purpose of the data report is to analyze the weather trends of the city the author resides in compared to the global average. The author, yours truly, currently resides in the city of Alexandria, which is in the state of Virginia (VA) just outside of the U.S. Capital, Washington D.C. According to the project instructions, the average weather for Alexandria, VA and the global average was extracted from 1750 to 2013. In addition, the three- and ten-year moving averages and scatter plots were developed for this project to provide insights into the weather trends. Please select [link](https://app.powerbigov.us/view?r=eyJrIjoiYmZiYzEzYzctYjNkNS00MWYyLWFiNjItZTM1ODI1MGJlMDExIiwidCI6ImMxMmE5ZjI3LTUwNWQtNGZjNi05YWZhLWEwZmQ2NWQ5ZTk4NCJ9) to view public interactive dashboard of the weather trends project.

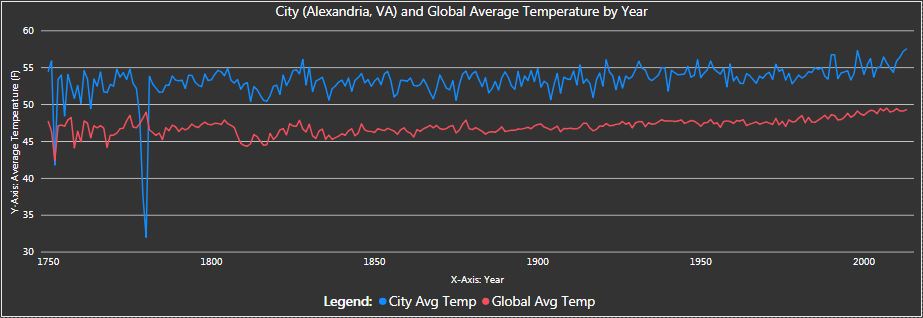
## Average Temperature of Alexandria, VA and the Global Average

As depicted above, the average weather data was extracted for the city of Alexandria, VA and for the global average. The data shown in figure 1 below illustrates that the temperature in the city of Alexandria is approximately 6.3 degrees Fahrenheit warmer than the global average. Both the city and the global average exhibit a slight linear trend as the compound annual growth rates are 0.02% and 0.01% for the city average and global average, respectively. In addition, both lines exhibit seasonality.

The weather for the city of Alexandria drops exponentially around the year 1781. Initial research was not easy to obtain, but according to the America Revolutionary historians those years experienced some of the harshest winters ever which had a direct impact on the Revolutionary war. Based on George Washington’s historical papers, the ground was still frozen in mid April.[[1]](#footnote-1)

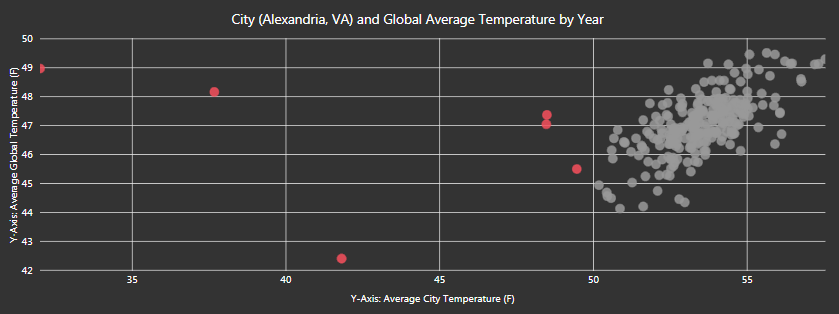
The weather for both data sets also decrease around the years 1816 to 1818. Additional research indicated that the year of 1816 was known as the “year without a summer.” The global average temperature decreased suddenly because of a volcano eruption on Mt. Tambora, a 13,000-foot-high volcano on the island of Sumbawa, near Bali, Indonesia. Between the volcano eruption and cooling sun, the Northern Hemisphere experienced unusual low temperatures with erratic weather patterns.[[2]](#footnote-2)

Figure 1. Average Temperature of City and Global Average



In addition to the line chart, a scatter plot was developed to show the outliers in the data set. The average temperature data sets have outliers for years 1780, 1779, 1752, 1778, 1755, and 1763. These dates are shown as outliers because the average temperature is less than the average temperature compared to the rest of the data. For example, the years 1779 and 1780 has average city temperatures of 32o and 37o,respectively when the preceding and following temperatures average around 50o.

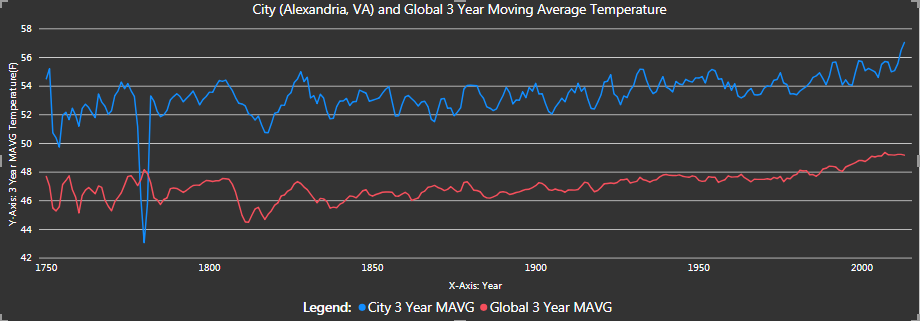
Figure 2. Outliers from Average Temperature Data



## 3 Year Moving Average Analysis

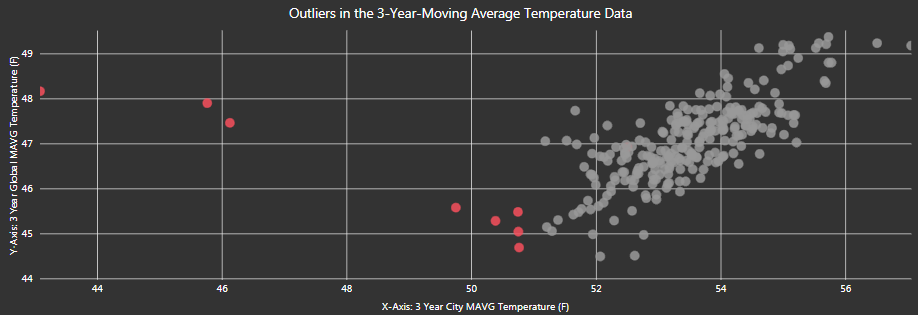
The development of the three-year moving average of both the city and global average lines begin to smooth out and continue to exhibit a linear growth trend. The lines exhibit the same decreases during 1781 for the city and during 1816 for both the city and the global average.

Figure 3. 3 Year Moving Average



In addition to the line chart, a scatter plot was developed to show the outliers in the 3-year-moving-average data set. The average temperature data sets have outliers for years 1780, 1781, 1779, 1754, 1752, 1753, 1818, and 1817. When compared to the average temperature outlier scatter plot, the new outliers in the 3-year-moving-average data set are years 1818 and 1817. The 3-year-moving average temperatures are 50.75o and 50.76o for the years 1818 and 1817, respectively. The data points before and after the years of 1818 and 1817 are greater than 51o. As a direct result of the 3-year-moving average smoothing, the spacing between the outliers and non-outliers is getting smaller. The same behavior is shown for the remaining outlier years for the 3-year-moving-average temperature data set.

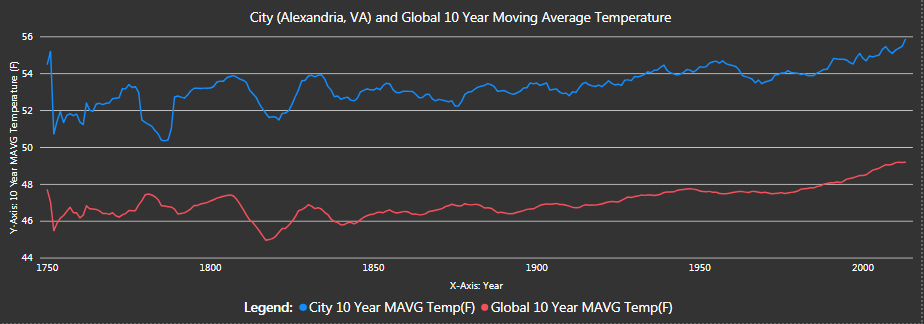
Figure 4. Outliers from 3 Year Moving Average



## 10 Year Moving Average Analysis

The development of the ten-year moving average of both the city and global average lines appear to exhibit smoother lines as compared to the three-year moving average and continue to exhibit a linear growth trend. The sudden decrease during 1781 for the city has shortened exponentially, still showing a drop, but not nearly as significant as the two preceding line charts. The decreases in 1816 for both the city and the global average has smoothed out a bit and both lines begin to exhibit similarities proceeding the decrease of 1816.

Figure 5. 10 Year Moving Average



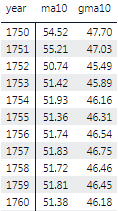
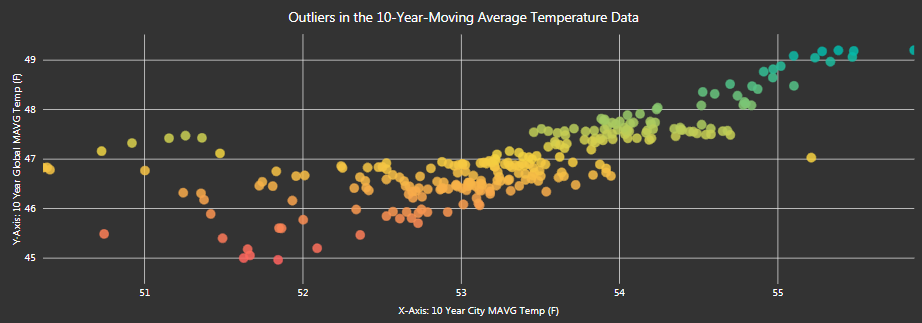


Figure 6 below illustrates the direct impact of a 10-year-moving-average as the data is much smoother when compared to the 3-year-moving-average and the average data sets. Instead of using a pre-attentive color scheme of red and gray, figure 6 is using a diverging color scale to illustrate the temperatures from low to high (red is low, yellow is middle, and green is high). This data set is exhibiting less prevalent outliers when compared to the other data sets. However, the years 1751 and 1752 are indicated in the figure as they can be considered outliers. As shown in the call out box, the spacing between data points is small.

Figure 6. Outliers from 10 Year Moving Average



1752

1751

## Methodology

The approach for extracting and normalizing the data consisted of identifying null values, developing the SQL for data extraction, converting the Celsius temperature to Fahrenheit, and developing a data visualization dashboard in Microsoft Power BI.

**Normalizing the data**

The global data set started in 1750 and the city data set starts in 1743. The data set was modified to have both data sets start in 1750.

**SQL Script**

The SQL script developed extracted the data based on the following criteria:

* Left join with the City\_Data and the Global\_Data
* Filtered the data for years greater than 1749
* Filtered the data for cities equaling Alexandria
* Filtered the data for countries equaling United States
* Created columns for the 3-year moving average for the city and global
* Created columns for the 10-year moving average for the city and global
* See script below

SELECT City\_data.year, City\_data.city, City\_data.country, City\_data.avg\_temp, Global\_data.avg\_temp AS GAVG, AVG(City\_data.avg\_temp) OVER (ORDER BY City\_data.year ASC ROWS Between 2 PRECEDING AND CURRENT ROW) AS MA3, AVG(City\_data.avg\_temp) OVER (ORDER BY City\_data.year ASC ROWS Between 9 PRECEDING AND CURRENT ROW) AS MA10, AVG(Global\_data.avg\_temp) OVER (ORDER BY Global\_data.year ASC ROWS Between 2 PRECEDING AND CURRENT ROW) AS GMA3, AVG(Global\_data.avg\_temp) OVER (ORDER BY Global\_data.year ASC ROWS Between 9 PRECEDING AND CURRENT ROW) AS GMA10 FROM City\_data LEFT JOIN Global\_data ON City\_data.year = Global\_data.year WHERE City\_data.year > 1749 AND City\_data.city ='Alexandria' AND City\_data.country = 'United States' ORDER BY City\_data.year;

**Excel modification and Data visualization**

Once the data was extracted, the “Convert” function was leveraged to convert the Celsius degrees to Fahrenheit degrees as follows: CONVERT(Cell Reference,"C","F"). The CSV extract file was then converted to excel and a new worksheet was developed for the Fahrenheit degrees (one worksheet for Celsius and one worksheet for Fahrenheit).

After that the modified excel worksheet was uploaded into a Microsoft Power BI PBIX file and a data visualization dashboard was developed. The author leveraged his Power BI Pro site to push the PBIX file into a public shareable [link](https://app.powerbigov.us/view?r=eyJrIjoiYmZiYzEzYzctYjNkNS00MWYyLWFiNjItZTM1ODI1MGJlMDExIiwidCI6ImMxMmE5ZjI3LTUwNWQtNGZjNi05YWZhLWEwZmQ2NWQ5ZTk4NCJ9).

1. <http://gwpapers.virginia.edu/general-washington-records-the-weather/> [↑](#footnote-ref-1)
2. [https://www.almanac.com/extra/year-without-summer#](https://www.almanac.com/extra/year-without-summer) [↑](#footnote-ref-2)