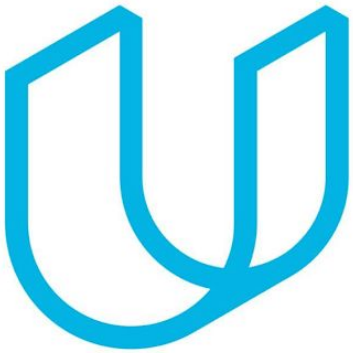
Data Analyst Nanodegree



UDACITY

EXPLORING WEATHER TRENDS

- ABHISHEK ANGNE

Exploring Weather Trends between Dublin, New York and the Global temperatures

Nature is the most important part of our lives and taking care of it is one of our essential needs for a longer survival. This project analyzes weather trends and looks at how the earth has been heating up over a span of approximately 250 years.

Goal: To compare rising temperature averages for New York (United States) and Dublin (Ireland) to the global averages and to each other to understand weather trends.

Story: I've always wanted to travel to the city of New York, which is why New York is a part of my analysis. In the year 2019, I moved to Dublin to study MSc. in Data Analytics. Through this analysis, I want to look at the weather conditions of these two technological hubs in the world of Analytics. Also, this nano-degree by Udacity is a step further in refreshing and shaping my skills as an Analyst amidst the quarantine due to COVID-19 and a big thank you to Udacity for providing the opportunity.

Data Extraction and Transformation:
SQL queries:

From city_list	**SELECT * FROM city_list WHERE country = 'Ireland' AND city = 'Dublin' SELECT * FROM city_list WHERE country = 'United States' AND city = 'New York'**
From global_data	**SELECT * FROM global_data;**
From city_data	**SELECT * FROM city_data WHERE city = 'Dublin'; SELECT * FROM city_data WHERE city = 'New York'';**

Using the aforementioned queries, I downloaded the .csv formats of the required data.
Treating Missing Values: Removed years 1743-1749 from the city_data as the global_data database has years ranging from 1750. Therefore, the years selected for the analysis were from 1750-2013. The values were also treated by averaging but 1750 was chosen as the first year due to data availability. Creating synthetic data and then calculating its moving average and then relative averages could well affect the overall calculation which is why only 1750-2013

were chosen.

Insights: A quick overview

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

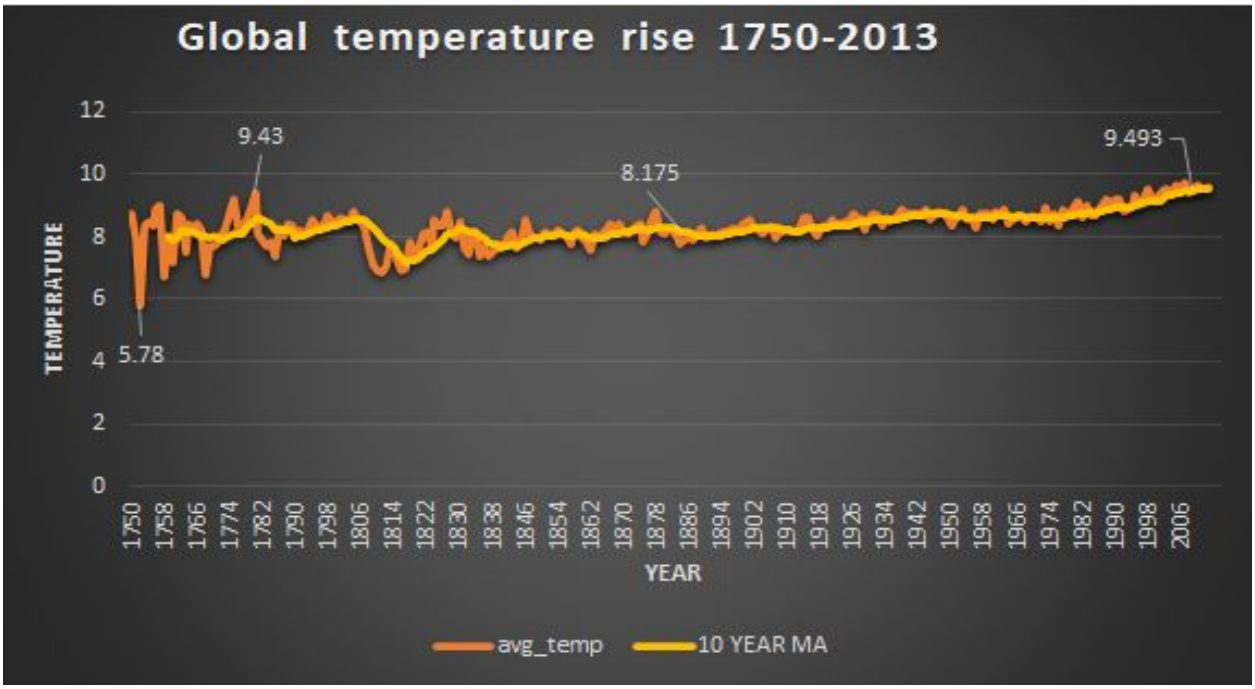
A: SQL and Excel were used for analyzing the weather data. SQL was used to pull data using queries from the 3 files namely, city_data, global_data, and city_listl. I used Google Docs for documenting the entire project.

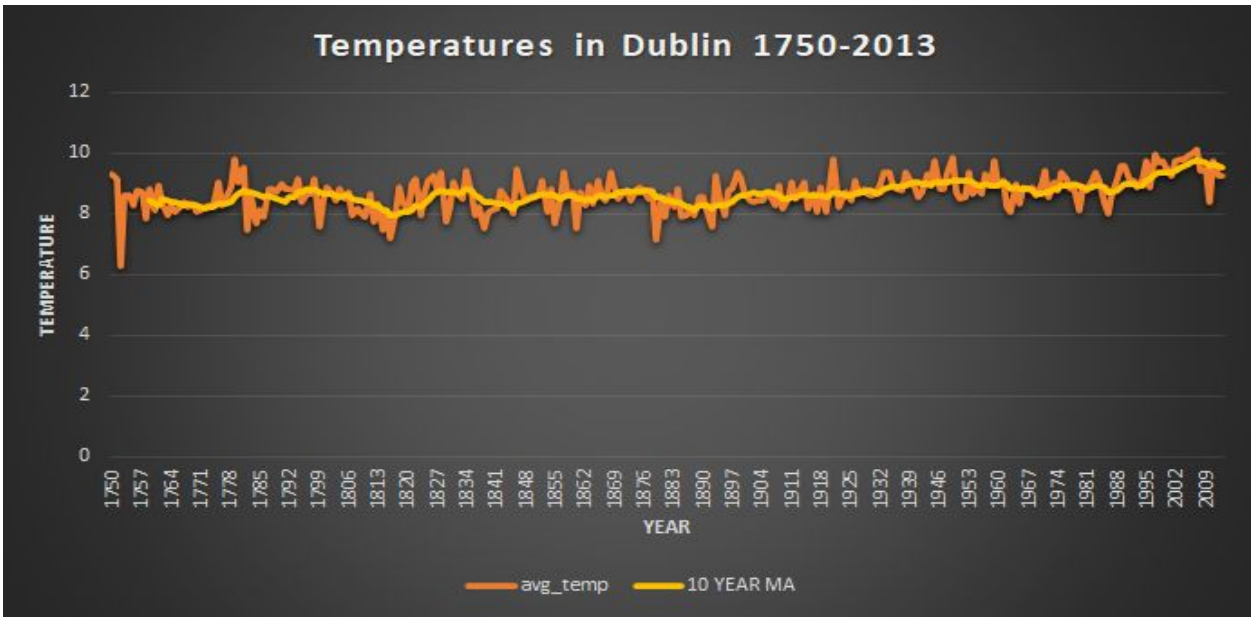
Q) How did you calculate the moving average?

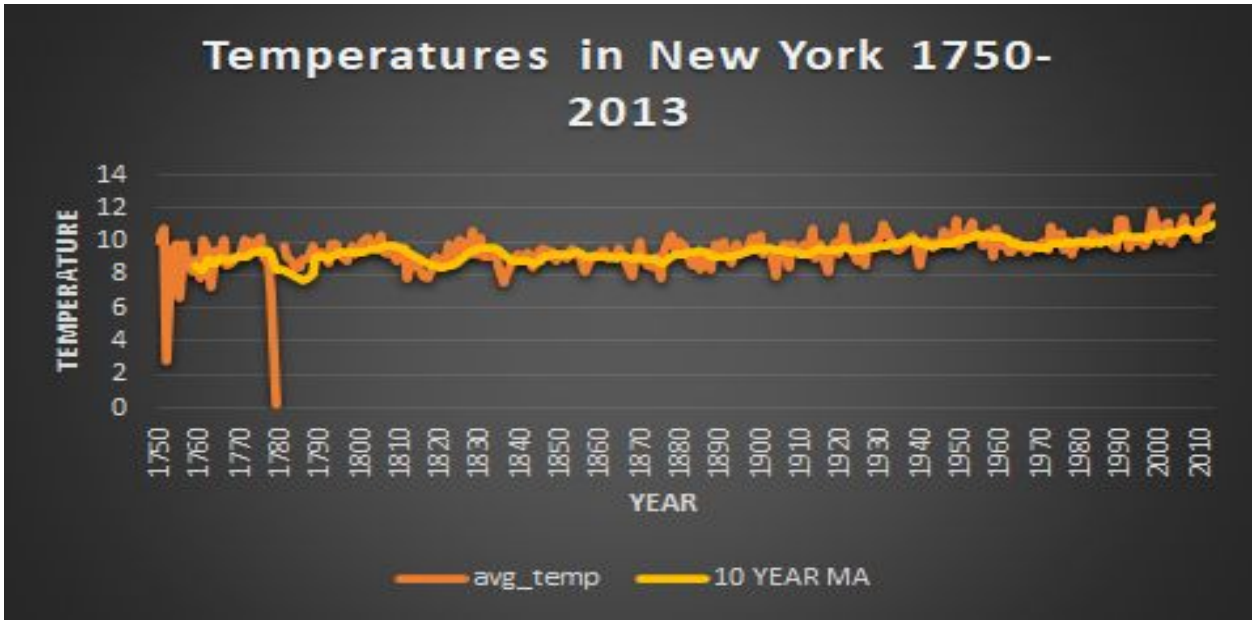
A: The moving average was calculated by using the average function in Excel. AVERAGE(celli:cellj) wherein i and j are cell numbers. I have used 10-year moving averages for the analysis.

Is your city hotter or cooler on average compared to the global average? Has the difference been consistent over time?

The city of Dublin has an average temperature of 8.67 Degree Celsius. The city of New York has an average temperature of 9.46 Degree Celsius and the global average temperature is 8.34 Degree Celsius. Therefore, it can be easily understood that the cities of Dublin and New York have higher averages than the global average. It could be because of both being economic boosters of their respective countries.





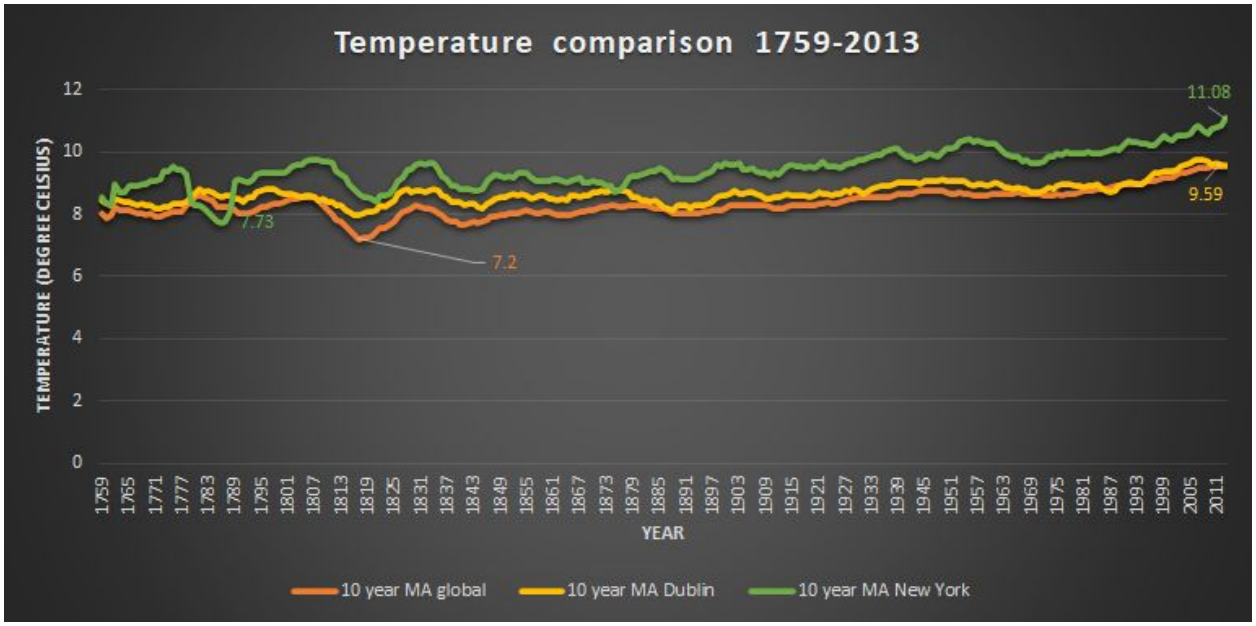


“How do the changes in your city’s temperatures over time compare to the changes in the global average?”

1) Here, 10 year MA’s are considered and 6 year gaps are visualized. New York’s all time high is 11.08 in the 6 year period post 20011, 2013 to be specific. Their all-time low is in the year 1786, i.e. 7.73. This is one of the few times in the overall time period considered that New York has an average temperature substantially lower than the Dublin and Global averages.

2) The global all time low is 7.20 degrees in the year 1817. Dublin’s weather in the same year was 7.95 degrees and New York’s was 8.66 degrees.

3) 1779 to 1788 are the only years wherein New York's average weather has been lower than both the global average and the averages in Dublin.
4) Conclusively, New York being a global center for tourism, commerce, economy and much more, does attract more levels of pollution, thus, resulting in an average weather warmer than the global average.



Recommendation: I would recommend observing ozone friendly strategies to be followed one hour per week or month and contribute in individual and group strategies and awareness to address the very obvious and ever looming threat of global warming.

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

Line graphs were used in order to analyze time-series data and line graphs are perfect in depicting data related to time and years.
Color scheme: Colourful Palette 3 on MS Excel
a) I used a color scheme which had warm tones like orange and chrome yellow to depict that the planet is heating up over the years.
b) My key considerations were to investigate whether the selected cities were getting warmer or not. Also, I tried to understand which out of the two were warmer.
c) Dublin has had the highest temperatures in 2007 with the yearly temperature being at 10.11 and the 10 year MA at 9.75. However, in 2007, the global temperature was 9.73 and a 10 year MA of 9.48.
d) New York, on the other hand, had its highest average temperature in the year 2013 of 12.16 degrees and a 10 year MA of 11.076. In 2013, the global average temperature was 9.61 with the 10 year MA being 9.55 degrees. Dublin had a moving average of 9.53 which was less than the global moving average.

Conclusion:

Global greenhouse gas emissions have led to continuous growth in the global average temperatures.

REFERENCES:

https://www.investopedia.com/ask/answers/122314/what-are-main-advantages-using-moving-averages-ma.asp

Research about CO2 Emission under the Current Energy Structure in Tianjin (https://ieeexplore.ieee.org/document/5660365)