**COVID-19 and the Weather: A data visualization**

Nick P Palacio

Department of Computer Science, University of Nebraska at Omaha

Capstone Project Proposal

Dr. Brian Dorn, Dr. Rex Cammack, Dr. Brian Ricks

**COVID-19 and the Weather: A data visualization**

On January 9th 2020 the World Health Organization (WHO) announced that a coronavirus related pneumonia had been spreading in Wuhan, China. The US confirmed its first coronavirus case on January 21st (*A Timeline of COVID-19 Developments in 2020*, n.d.). By March 11th 2020 the WHO had declared the COVID-19 pandemic.

Since then the US has experienced several waves of increased infection rates that have varied in severity across the country. According to the CDC, COVID-19 can spread from human to human via respiratory droplets in the air. Specifically, the virus is known to spread more easily indoors where there is less air ventilation (CDC, 2020).

One could hypothesize that as the weather changes to less favorable conditions outdoors more people do more things indoors, increasing the spread of COVID-19. My proposed project aims to let users explore the relationship between the spread of COVID-19 and the weather in the US.

My project is a map based web application that would allow users to explore the relationship between COVID-19 infection rates and the weather. This visualization would be useful to the public and policy makers alike. Policy makers would be able to better understand the role of weather in infection rates which would help them make the right policy decisions to control the spread of the disease. Helping the public understand this same relationship would better inform them and their daily decisions that help prevent the spread of COVID-19.

**Related Work**

Influenza is another respiratory illness that is spread via respiratory droplets in the air. The link between influenza and the weather is well established (Huang et al., 2017; Roussel et al., 2016). Weather variables such as temperature, humidity and daily variation of both have been found to have a significant effect on influenza infection rates(Roussel et al., 2016). While the coronavirus is certainly not the same thing as the flu it does spread in a very similar manner. This makes the relationship between COVID-19 transmission and the weather worth exploring.

There are existing data visualizations for COVID-19 that help people understand things like infection rates, hospitalizations and policy timelines. The John Hopkins University of Medicine has several of these. They have visualizations that show charts with timelines to see how infection rates have changed over time. They also have map based visualizations that allow users to explore data for their area of interest (*COVID-19 United States Cases by County*, n.d.). Their US Map data visualization is similar to what my proposed project would be with the addition of weather data.

The COVID Tracking Project by The Atlantic has several COVID-19 data visualizations. They show you data points like case counts by state, hospitalizations and tests. Most of their visualizations are timeseries charts with one spatial visualization that lets you view the data on a map of the US. Their visualizations help keep the public informed about what is going on in their area and in the US as a whole.

**Methods**

**Data Sources**

For my COVID-19 data source I will be using one of the datasets generated and maintained by the New York Times hosted on GitHub (The New York Times, 2021). Specifically, I will be using the us-counties.csv dataset. This dataset contains a full history of cumulative COVID-19 cases and deaths by county by day in the US.  
For my weather data I will be using an API from Weather Source. Weather Source is a technology company that provides a suite of products that help businesses leverage weather and climate data. On March 16th, 2020 Weather Source opened their API for free to any researchers exploring the relationship between weather and the COVID-19 pandemic. Their Weather History API exposes many different weather data points that can be queried with a date range along with latitude and longitude, or zip code. Data can be returned in an hourly or daily format. For my purposes I will be retrieving average temperature, average relative humidity and average absolute humidity in a daily format.

**References**

A Timeline of COVID-19 Developments in 2020. (n.d.). AJMC. Retrieved March 3, 2021, from <https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020>

CDC. (2020, October 28). COVID-19 and Your Health. Centers for Disease Control and Prevention. <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html>

COVID-19 United States Cases by County. (n.d.). Johns Hopkins Coronavirus Resource Center. Retrieved March 3, 2021, from <https://coronavirus.jhu.edu/us-map>

Huang, X., Mengersen, K., Milinovich, G., & Hu, W. (2017). Effect of Weather Variability on Seasonal Influenza Among Different Age Groups in Queensland, Australia: A Bayesian Spatiotemporal Analysis. The Journal of Infectious Diseases, 215(11), 1695–1701. https://doi.org/10.1093/infdis/jix181

Roussel, M., Pontier, D., Cohen, J.-M., Lina, B., & Fouchet, D. (2016). Quantifying the role of weather on seasonal influenza. BMC Public Health, 16, 441. <https://doi.org/10.1186/s12889-016-3114-x>

The New York Times. (2021). Coronavirus (Covid-19) Data in the United States. Retrieved March 7, 2021, from <https://github.com/nytimes/covid-19-data>.