Mapping the News: COVID-19 in the NY/NJ Metro Area

by Tammy Ann Syrek-Marshall

1. Introduction
   1. Background

Epidemics and pandemics have occurred throughout human history. From the plagues of ancient Greece and Rome to the more recent Ebola Virus, these outbreaks have been responsible for the deaths of thousands if not millions.

In the 21st century, along with advances in science and medicine, we still look to the leadership of our governments to take responsibility for mitigating the effects of these epidemics as well as for keeping the public informed. This information often takes the form of local and national news reports.

On December 31, 2019 the first reports of a new strain of pneumonia afflicting dozens of people in Wuhan China made the headlines in the New York Times as well as other news outlets.

On January 7, 2020 this virus was identified as a novel, previously unknown strain.

The first death from this virus was reported on January 11, 2020.

By January 21, 2020 the first case of this Coronavirus, officially named COVID-19, was reported In Washington State.

As of March 1, 2020, while the virus continued to spread quickly to other countries, New York and New Jersey began seeing their first cases of COVID-19.

* 1. Problem and Discussion

In this report, the primary aim is to provide government officials, community leaders, as well as business and news media executives with a methodology for understanding through data visualization information dissemination during a crisis situation.

While the outbreaks are the major focus of the news, the effects of this virus can also be seen in other community related stories.

Each state in the U.S. is an entity onto its’ self, with differences in their governments and their communities. With that in mind, it might be possible to gain new insights into the effects of this epidemic on society by doing a comparative analysis of the news coverage of New Jersey to the coverage of New York.

The more information on the effects of a society-wide disruptor, the more efficient and effective the planning and problem solving can be.

1. Data
   1. Overview

By using the tools of data science and geospatial data, I propose to map out community specific news articles based on the geographical focus of their stories.

In order to create the Maps, I will use Folium library to visualize the data.

As for the data itself, the News Article data will be derived from searching databases such as ProQuest then compiling the data into an excel spreadsheet, formatted so as to be transformable into a Pandas dataframe.

Additional community data will be uploaded from FourSquare. If more information is needed, it is possible to obtain datasets from sources such as the CDC and Census. However, I expect this might not be required.

Once the data is visualized, additional information will be provided in the final product relating to each marker on the map.

As an example, a story may cover the cancellation of a public event, much like the South by Southwest (SXSW) event in Austin Texas, due to concerns over the Coronavirus. The FourSquare data can provide additional information about the venue as well as nearby venues also affected by the cancellation.