Kaggle Task Analysis: Transmission, Incubation & Environmental Sustainability

Yasha Ronquillo

ICS 360

Debasis Bhattacharya

May 3, 2020

**Analysis**

**Temperature, Air Travel & Transmission - Anand Venkataraman**

In his notebook, Venkataraman suggests that temperature, relative humidity, and air travel are factors of the coronavirus cases. Venkataraman used a set of maps to show his case. His most notable illustrations are two maps showcasing the globe based on its reported cases, and the other based on the countries’ global average land temperatures. Both maps have red and green zones, with red (the top half of the globe) are causes of concern and the green (bottom half) are “safer.” He also noted the spread of the virus on March 1, 2020 in relation to air routes from Beijing. This data insinuates that air travel is a factor of the virus spread. This is already a known fact as direct contact increases the transmission of the virus.

After noting the top nine countries with the most COVID-19 cases (as of March 29, 2020)-- US, Italy, China, Spain, Germany, France, Iran, UK, Switzerland, and Netherlands-- it was noted that these countries are in the northern hemisphere. Most notably, such countries have average global temperatures within 5-10°C. On the other hand, Based on his “Progression of spread of COVID-19” visual, the virus also tended to appear first within those regions as of February 29, 2020. Meanwhile, countries that are in the southern hemisphere--specifically those in more tropical or dry climates-- came into the scene later.

With the data provided, it can be presumed that temperature may have an effect on the transmission or lifespan of the virus. Venkataraman concludes that the virus is able to thrive in more temperate zones, such as the United States and China. On the other hand, he states that the virus is “not able to thrive in warmer climates or surroundings,” hence the delay or the lower number of cases in such countries (Venkataraman, 2020).

**Demographics & observation for pandemic escalation - Aman Kumar**

Another task submission was from Kumar, who looked into the different demographic factors that would have impacted the transmission and the spread of the coronavirus. In his notebook, Kumar looked into factors such as temperature, wind speed, precipitation, fog presence, population, population density, median age, sex ratio, hospital beds, and life expentancy.

With his data Kumar found, similar to Venkataraman, that rising temperatures tend to slow down the number of cases, having a negative correlation. He furhter analyzed this with a box-and-whisker scatter plot comparing temperature(x-axis) and the number of reported cases(y-axis). It is clearly seen in the plot that there is a unimodal curve, where the median is around 6.5°C. Conducting a hypothesis test and obtaining a p-value of 0.0089, Kumar concluded that he accepted his null hypothesis that temperature is the same over the population data and that temperature alone is not the sole reason for COVID-19. In terms of demographics, median age also has an effect on the cases, where the higher the age, the more cases there are--an already-known fact as older adults are more susceptible to get infected by the virus. Kumar also looked into gender and case correlation and found that males 60+years had a greater chance of death than females of the same age. This is so as life expectancy is also affected with COVID-19, even more so with men with a weak correlation. Kumar noted that this is so as “males are more subjected to moving out of homes even in quarantine zones.” He also found that smokers percentage in population, obesity, and diabetes affect COVID-19 (Kumar, 2020).

**Four possible ways to fight COVID-19 - Yuan So**

After presenting the effects of medical and health service capacity, level of economic development, national infection rate, and policy responses as factors for COVID-19, So presented four possible ways to combat the virus. The first call to action was to strengthen the medical response. In the data that compared fatality rate with cumulative number of deaths, it was shown that Germany and Switzerland were on the lower range of both spectrums, with Germany having an “abnormally low mortality rate.” This is so because such countries already have strong medical capabilities in place. Other countries do not have this same readiness however, as their medical resources are outnumbered or have been exceeded by the number of infected individuals in the population.

So then proposed for excellent national health awareness and to take a large number testing. Japan is not doing enough tests and may lead to more outbreaks than recorded. So observes that Japan’s current data may be underreported. So then said to use time differences to effectively use medical resources. Seeing that large countries such as the United States, China, Italy, France, and Germany are unable to prevent the exponential rate of infection after 14 days after the 500th case of their countries, cities must make sure to value that 14-day time. With 14 days of isolation, the virus can be potentially contained.

Finally isolation should be made mandatory according to So. Looking into China’s medical resources per capita, cities like Hubei were able to contain the virus within one month of isolation.

**References**

Kumar, Aman(2020, April). Demographics & observation for pandemic escalation. Retrieved from https://www.kaggle.com/aestheteaman01/demographics-observation-for-pandemic-escalation

So, Yuan (2020, April). Four possible ways to fight COVID-19. Retrieved from https://www.kaggle.com/yuanso/four-possible-ways-to-fight-covid-19

Venkataraman, Anand (2020, March). COVID-19 - Temperature, Air Travel & Transmission. Retrieved from https://www.kaggle.com/sixteenpython/covid-19-temperature-air-travel-transmission/notebook