Covid-19 and Neighborhood Venues

Michael Yeh

August, 2020

What We Did

We took New York City's Covid-19 data, combined it with venues data from FourSquare, and looked for relationships between them.

Covid-19 data:

New York City provides public access to Covid-19 data compiled by city officials. This data is organized by zipcode and includes all of New York City. The data set includes two types of data that we used for our analyses: Covid-19 confirmed infection rate (number of infections per 100,000 resident within a zipcode neighborhood), and Covid-19 death rate (number of deahts per 100,000 resident within a zipcode neighborhood).

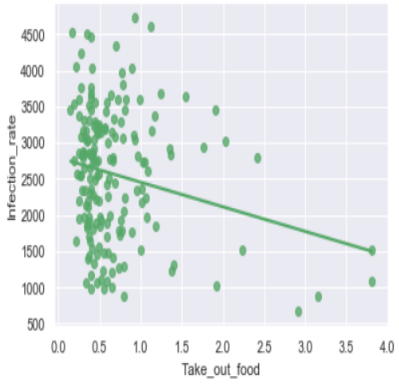
FourSquare venues data:

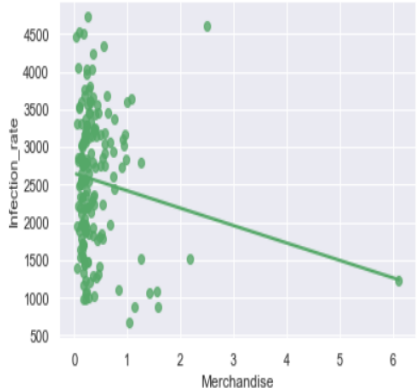
FourSquare is a web based service which provides venues data download. Venues are establishments of all kinds, such as restaurants, stores, markets, services, etc.. Venues are a reflection of the life-style within a neighborhood, and life-style may have implications on Covid-19 community spread.

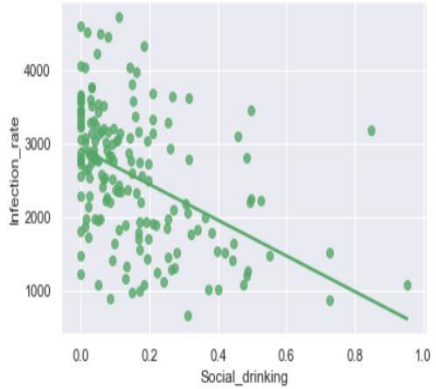
Surprising and Thought Provoking Findings (1)

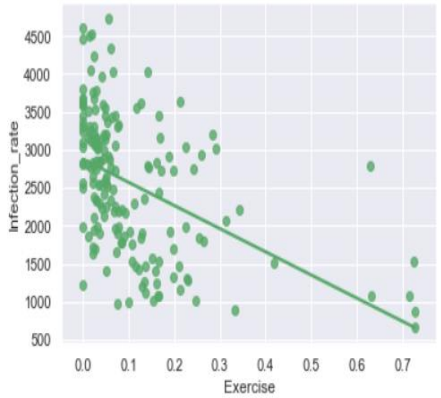
Surprise #1: More venues per resident is associated with lower infections and deaths

We transformed venues data into per-capita data (i.e, number of venues per 1000 residents). We then looked at the relationships between each category of venue versus Covid-19 infection and death rates. Take at look at the representative plots below.









We were quite surprised to see that when there are more venues per resident, of any type, the Covid-19 infection and death rates are lower! It is especially interesting that this behavior is true for Social_drinking venues, which includes bars.

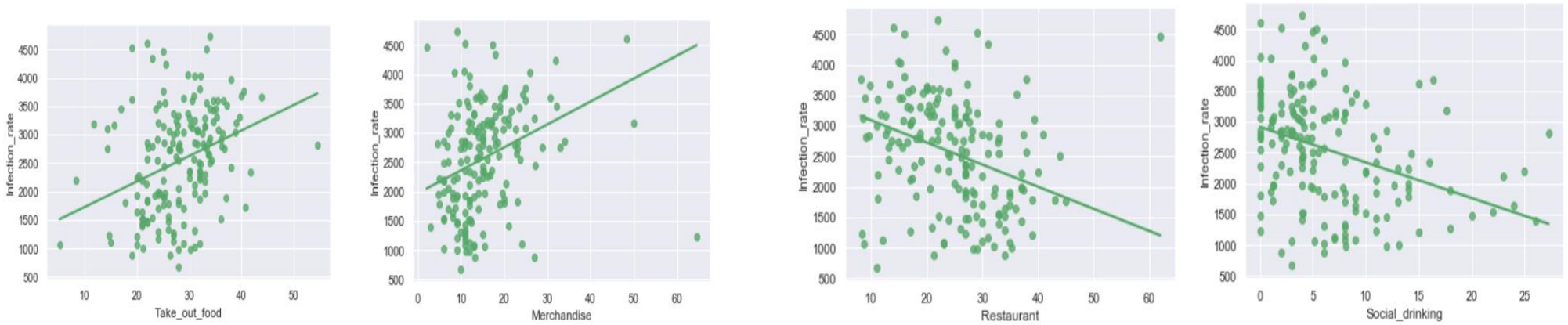
What could explain this observation? One possibility is when there are more venues, people have more choices, therefore, there is less crowding of individual venues. In other words, there is less density within each venue because when it gets crowded, people simply go to the venue half a block away that is less crowded. Less density leads to fewer infections and fewer deaths.

Public health policy makers might find this information useful in managing the lock-down and re-opening of local neighborhoods.

Surprising and Thought Provoking Findings (2)

Surprise #2: The composition of venues impacts Covid-19 infection and death rates

We transformed venues data into percentage data (i.e, number of venues in a category divided by total venues in a neighborhood). We then looked at the relationships between each category of venues versus Covid-19 infection and death rates. Take at look at the representative plots below.



On the left side are two plots of Harmful Venues (Take_out_food, Merchandise, Markets, Services, Transportation). These are establishments with high traffic, low duration (clients spend little time in them), and low cost. On the right side are "Helpful Venues" (Restaurants, Social_drinking, Exercise (indoor facilities), Entertainment). These are establishments with low traffic, high duration (clients spend a long time in them), and high cost.

The surprise this time is that there is positive association for Harmful Venues (i.e., higher percentage is associated with higher Covid-19 infection and death rates). If people spend little time in these establishments, why are the Covid-19 rates higher when a neighborhood has more of these venues relative to other venues (i.e., higher percentage)? Why are these venues positively correlated with Covid-19?

Surprising and Thought Provoking Findings (2 continued)

Surprise #2: The composition of venues impacts Covid-19 infection and death rates (continued)

Equally surprising is that the Helpful Venues are negatively associated with Covid-19 rates. Higher percentage of Helpful Venues seems to be associated with lower Covid-19 infection and death rates! Take Restaurants for example, these are sit-down culinary venues where people enjoy a meal over a long period of time, often over an hour in pre-Covid days. Doesn't the longer duration of stay in these venues expose people to higher risk of Covid-19 transmission? The data shows that when a neighborhood's venue mix (i.e., the composition of venues as measured by percentages) contains more Helpful Venues, there is lower Covid-19 infection and death rates.

What might explain these surprising observations? Perhaps the Harmful Venues are harder to keep clean due to high traffic, and also because of high traffic, there is higher likelihood that an infected person visits that establishment. Also, the type of people who frequent Harmful Venues may be prone to unhealthy life-style choices (e.g., fast food is associated with bad health). Unhealthy life-styles makes a person more susceptible to Covid-19 infection and death.

On the other hand, Helpful Venues like sit-down restaurants are easier to keep clean because of lower traffic. Furthermore, the people who visit these venues may have higher income, better health care, and are generally less likely to carry the virus into the venue. So while people spend more time in Upscale Venues, they are in a more hygienic environment with lower risk of infection.

It is also quite surprising that the Social_drinking venues are in the Upscale Venues group. Bars are included in this group. The data in this study shows that a higher percentage of Social_drinking venues is associated with lower Covid-19 rates. However, there are numerous documented cases of rapid community spread in bars across the US. So this particular type of venue (i.e., bars) must be handled with extreme care.

Conclusions

When it comes to stopping Covid-19, there is nothing better than a complete lockdown. Any contact between people is a chance for Covid-19 to spread. Therefore, the findings in this study is not to suggest that we should open up more venues. Rather, these findings can be most effectively used in a gradual re-opening or a gradual lockdown should that be necessary again.

Surprise #1 indicates that increasing the number of venues (of all types) available to each person can reduce the infection and death rates. This is because of reduced venue density. However, Surprise #2 tells us that certain types of venues (Helpful Venues) help to reduce infection rates, therefore these venues should be prioritized.

Said another way, it is important to pay attention to the mix of venues available in a neighborhood. When re-opening or gradually locking down, we should attempt to increase the percentage of negatively correlated venues (i.e., Helpful Venues) and reduce the percentage of positively correlated venues (i.e., Harmful Venues).

More analyses should be performed on the findings in this study before the concepts are put into practice. Hopefully these findings will lead to a more safe and efficient approach to re-opening the economy.