Michelle Hsieh

Professor David McDonald

DATA 512

12 December 2022

COVID-19 and Mobility Analysis – Cuyahoga County, Ohio

**Introduction**

The COVID-19 pandemic has had a profound impact on communities around the world, including Cuyahoga County in Ohio. As the spread of the virus continues to be a major concern, understanding how people are moving within the county is crucial for containing the outbreak and mitigating its effects. Mobility analysis can provide valuable insights into how the virus is spreading and how effective social distancing measures are at limiting its spread. In this paper, we will explore the importance of mobility analysis in the context of the COVID-19 pandemic in Cuyahoga County. The analysis of mobility patterns during the pandemic can help policymakers and public health officials understand how the virus is spreading and identify areas where interventions may be more necessary than others. This can help them make more informed decisions about how to contain the outbreak and mitigate its effects on the community.

Additionally, by tracking changes in mobility patterns over time, researchers can evaluate the effectiveness of different social distancing measures and determine which ones are most effective at limiting the spread of the virus. This can help guide future decision-making and help public health officials develop more effective strategies for containing and mitigating the effects of the pandemic. The mobility patterns found can also provide more insight to people’s perception of risk and this can be helpful in behavioral economics if a similar situation like COVID-19 were to arise again.

**Background/Related Work**

There has been a significant amount of research on the use of mobility data for understanding and responding to the COVID-19 pandemic. Studies have used mobility data from digital platforms such as Google and Apple to track changes in people's movements and identify potential hotspots for the virus. Other research has focused on using mobility data to evaluate the effectiveness of different social distancing measures and determine how they impact the spread of the virus. Additionally, there has been a growing interest in using machine learning and other advanced analytical techniques to better understand the relationship between mobility patterns and the spread of the virus.

In “Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges,” Hu et. al. proposed new ways to accurately calculate human mobility in order to guide researchers and policymakers in conducting data-driven evaluations and decision-making for the COVID-19 pandemic and other infectious disease outbreaks. Instead of using Google’s location tracking data, they utilized public transit systems, mobile operators, and mobile phone applications. They compared mobility dataset’s characteristics by assessing data privacy, quality, space–time coverage, high-performance data storage and processing, and accessibility. In “What human mobility data tell us about COVID-19 spread”, Alessandretti talks about how the increased accuracy of location reporting has helped pinpoint and alert people if they have potentially been exposed to someone who had COVID-19. High-resolution mobility data makes it possible to explore how contacts occurring in different types of locations (such as restaurants, gyms or hotels) contribute to disease spreading. They can further enable us to assess the role played by super-spreading individuals, who infect a high number of other people, and study how epidemics spread within and across social communities. While other research initiatives focus more on how mobility affects the rate of spread of COVID-19, I am more interested in how the mask mandate has affected people’s perception of risk and mobility changes throughout this time period. My hypotheses and research questions comprise of comparisons of mean percentage changes from the baseline across 6 different location categories during 3 different time periods:

Null Hypothesis: There is no difference in mean percentage change from the baseline for “location” before the mask mandate/during the mask mandate/after the mask mandate vs. before the mask mandate/during the mask mandate/after the mask mandate

Alternate Hypothesis: There is a difference in mean percentage change from the baseline for “location” before the mask mandate/during the mask mandate/after the mask mandate vs. before the mask mandate/during the mask mandate/after the mask mandate

The location categories consist of parks, retail and recreation, workplaces, transit stations, residential areas, and grocery stores and pharmacies.

**Methodology**

I collected data on mobility patterns in Cuyahoga County from the Google Mobility Dataset which provides anonymized and aggregated mobility data for research purposes. The preprocessing steps consisted of joining the 2020, 2021, and 2022 files together into one single CSV and filtering on Cuyahoga County yielded around 950 rows for the analysis. The data timespan is February 1, 2020 to October 15, 2022 which is when data stopped being collected.

Next, we will analyze the data using a t-test to compare the mobility patterns for various locations before, during, and after the mask mandate. This will involve calculating the means and standard deviations of the mobility data for each group and using these values to conduct a t-test to determine whether there is a significant difference between the two groups. mandate. A t-test would be appropriate since we have a large sample and the standard deviations are unknown. It is also a widely-used statistical test to analyze differences between populations and samples. I additionally plotted visualizations to show the percentage change of baseline over time for each location and overlayed the mask mandate time period to show the changepoints more easily. These visualizations helped to show the impact of the mask mandate on mobility destination changes.

Some human-centered considerations include generalizing locations as much as possible so the users’ location data would not be personally identifiable especially for residential locations. My analysis could potentially identify areas where the virus is spreading rapidly, which could lead to increased public health interventions in those areas. This could have significant implications for the people living and working in those areas, so it is important to carefully consider how my analysis might impact them.

**Findings**

In my analysis, I found that most of the locations have seen significant differences in mobility change across the different time periods. Parks has seen the most dramatic change since its one of the few locations where people can safely social distance outdoors. Especially during the beginning of the pandemic, people were either supposed to stay home or go outside while maintaining a safe social distance.

Chart, histogram

Description automatically generated

From this plot we can see that the highest peak of percentage change was close to a 400% increase from the baseline. I found that there was a significant difference in mean percentage change before the mask mandate vs. during the mask mandate as well as during the mask mandate vs. after the mask mandate. However, there was not a significant difference before the mask mandate vs. after the mask mandate. Before the mandate, the mean percentage change was 66.56%, during the mandate: 48.23%, and after the mandate: 57.51%.

Chart, histogram

Description automatically generated

For residential locations, I found that there is a significant difference among all the time periods: before the mandate (10.41%), during the mandate (8.34%), and after the mandate (4.86%). We can see that over time people have been leaving their residences and going outside more as the number of COVID-19 cases ease out.

Chart

Description automatically generated

For retail and recreation, I found that there is also a significant difference among all the time periods: before the mandate (-23.2%), during the mandate (-17.8%), and after the mandate (-10.38%). We can see that over time that people have been going to places like shopping plazas and restaurants, but mobility has not reached the baseline very often.

Chart

Description automatically generated

For workplaces, there is a significant difference for during the mandate vs. after the mandate and before the mandate vs. after the mandate. There was not a significant difference between before the mandate vs. during the mandate and this could be attributed to most workplaces not opening their offices until mid-2021. With people primarily working remotely or in a hybrid mode, the mobility changes for workplaces have been consistently low: before the mandate (-30.27%), during the mandate (-28.64%), after the mandate (-23.31%).

Chart

Description automatically generated

For grocery stores and pharmacies, there is a significant difference before the mandate vs. after the mandate and before the mandate vs. during the mandate. There was not a significant difference between during the mandate vs. after the mandate and this could be attributed to people not being as concerned about going grocery shopping as they normally would before the pandemic. Some dramatic decreases could be because of the increased usage of online grocery shopping and delivery apps such as Instacart. Since grocery stores are an essential part of people’s lives, the overall mobility has been hovering over the baseline consistently. Before the mandate: -4.07%, during the mandate: -6.82%, and after the mandate:: -6.28%.

Chart

Description automatically generated

For transit stations, there is a significant difference during vs. after the mask mandate and before vs. after the mask mandate. There is not a significant difference between before vs. during the mask mandate and this could be attributed to people not feeling safe while taking public transport and feel safer from the comfort of their own car. Most workplaces are still not open during the mask mandate so a decrease in mobility for transit stations makes sense since most people who go to transit stations are commuters. The mean percentage change for each time periods are: -24.65% for before, -24.6% for during, and -16.94% for after.

**Discussion/Implications**

We can see that for some of the locations, the mask mandate being put in place and being removed has a relationship with the mobility changes. These findings are important because we can see how people perceive risk and what locations they feel safer to go to once the mask mandate is put in place and once it is removed. Future research can build upon this study by seeing if locations with dramatic increases of mobility change are associated with higher COVID-19 infection rates. Another potential direction for future research is to explore the use of more advanced analytical techniques, such as machine learning, to better understand the relationship between mobility patterns and the spread of the virus. This could involve developing more sophisticated models that can capture the complex dynamics of the virus spread and make more accurate predictions about its future trajectory. Another potential direction for future research is to explore the use of mobility data from a broader range of sources. While this study may rely on data from digital platforms such as Google and Apple, future research could explore the use of other data sources, such as transportation data or survey data, to gain a more comprehensive understanding of mobility patterns in Cuyahoga County. Additionally, future research could focus on the development of more effective interventions for containing the spread of the virus. This could involve using mobility data to identify the most effective social distancing measures and other interventions and testing their effectiveness in real-world settings.

**Limitations**

As time passes and we move further away from the baseline period, populations might vary due to relocation or new regional and remote working options. Google’s understanding of categorized places might also change. For example, the same value today and in April 2020 might not indicate the same behavior or adherence—it might be that Google has updated information about shops and restaurants in the region or that fewer people live there now. These differences could shift the values up or down over long time periods, so we recommend using some caution when analyzing data from longer time intervals (6+ months). Additionally, the data may be subject to certain biases, such as the underrepresentation of individuals who do not use digital platforms, which could impact the accuracy of the results.

The data collected is only from users who have opted-in to sharing their location history through their Google accounts which means that the data represents a small fraction of the region population. We will still maintain the assumption that the users who have opted in are still representative of the population for the purposes of this project. Seasonality is also another unknown since we are dealing with time series data. It may be hard to extrapolate insights for parks for example since parks may have a dramatically lower percentage change during colder months. It may be hard to differentiate seasonality changes from COVID-19 phase changes.

**Conclusion**

Through this analysis, I examined the effect of COVID-19 and the mask mandate relaxation on human mobility and overall behavior. I looked into whether there was a significant difference across the three time periods (before, during, and after the mask mandate) and six locations (parks, grocery store and pharmacy, transit stations, workplaces, retail and recreation, and residences). The overall findings are that COVID-19 has had a significant impact on mobility and is still affecting people’s perception of risk and decision-making until this day. More people have been going to parks and a lot less people have been going to workplaces since there is greater flexibility in most work environments.

This study can inform the reader's understanding of human-centered data science in several ways. First, it illustrates the importance of considering human-centered considerations when collecting and analyzing data related to the COVID-19 pandemic. This study takes into account the potential privacy implications of collecting and analyzing mobility data, and it also considers the potential impacts of the study on the people of Cuyahoga County. By considering these human-centered considerations, this study demonstrates the importance of treating data related to the pandemic in an ethical and responsible manner. Additionally, this study can inform the reader's understanding of the role of data science in addressing the challenges posed by the COVID-19 pandemic. By using data to better understand the relationship between mobility patterns and the spread of the virus in Cuyahoga County, this study illustrates how data science can be used to inform policy decisions and develop effective interventions for containing the outbreak and mitigating its effects. This highlights the potential of data science to make a meaningful and positive impact on public health efforts during the pandemic.

**References**

T-test: <https://www.scribbr.com/statistics/t-test/>

Regression: <https://towardsdatascience.com/understanding-regression-using-covid-19-dataset-detailed-analysis-be7e319e3a50>

Time series analysis: <https://www.tableau.com/learn/articles/time-series-analysis>

Alessandretti, L. What human mobility data tell us about COVID-19 spread. Nat Rev Phys 4, 12–13 (2022). <https://doi.org/10.1038/s42254-021-00407-1>

Tao Hu, Siqin Wang, Bing She, Mengxi Zhang, Xiao Huang, Yunhe Cui, Jacob Khuri, Yaxin Hu, Xiaokang Fu, Xiaoyue Wang, Peixiao Wang, Xinyan Zhu, Shuming Bao, Wendy Guan & Zhenlong Li (2021) Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges, International Journal of Digital Earth, 14:9, 1126-1147, DOI: 10.1080/17538947.2021.1952324

<https://policies.google.com/terms?hl=en>

**Data Sources**

[Google COVID-19 Community Mobility Reports](https://www.google.com/covid19/mobility/)

<https://www.kaggle.com/antgoldbloom/covid19-data-from-john-hopkins-university?select=RAW_us_confirmed_cases.csv>

<https://data.cdc.gov/Policy-Surveillance/U-S-State-and-Territorial-Public-Mask-Mandates-Fro/62d6-pm5i>