An Analysis of US Social Distancing Amidst the Covid-19 Pandemic through Google Mobility Reports

Undergraduates: Jack Lichtenstein (jsl61@duke.edu) and Shannon Houser (smh116@duke.edu)
Duke University DataFest Competition 2020

Duke UNIVERSITY

Introduction

Google recently released downloadable mobility reports based on cell phone location data that highlight the percent change in visits to popular places such as:

- Retail/Recreation
- Grocery/Pharmacy
- Parks
- Transit Stations
- Workplaces
- Residential

Each mobility trend is compared to a baseline, defined as the median value, for the corresponding day of the week, during the 5-week period Jan 3–Feb 6, 2020. The values released by Google are the percent change for each state compared to this baseline.

As we are not epidemiologists, our goal is simply to study and understand social distancing trends in the US during the COVID-19 pandemic. We are not interested in predicting or modelling how the virus has spread or will spread in the future, but rather how different parts of the US have reacted to the virus. We are interested in exploring the factors that are most strongly correlated to social distancing measures.

Objectives

Understand Mobility Trends During Covid-19: Which areas of the country are taking social distancing most seriously?

- Calculate a *Social Distancing Score* for each state by normalizing their mobility trend across each of the six categories of places
- Compare and contrast social distancing trends across states
- Visualize the differences in mobility in the specific categories as the pandemic progresses
- Figure out which states were reactive and which were proactive in distancing

Determine the Factors that Drive Social Distancing: What is driving certain states or areas of the country into social distancing? Why are some parts of the country doing it better than others? What drives communities to distance themselves?

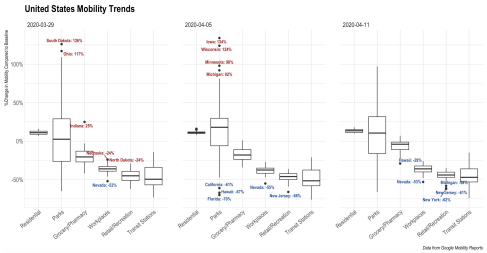
- Are political affiliations of state governors at play?
- Are there regional differences in distancing?
- How do the numbers of cases or cases per capita correlate to social distancing measures?

Exploring Social Distancing in the US

Basic Exploration:

Let's get a sense of these data with some basic exploratory data analysis. Notice that

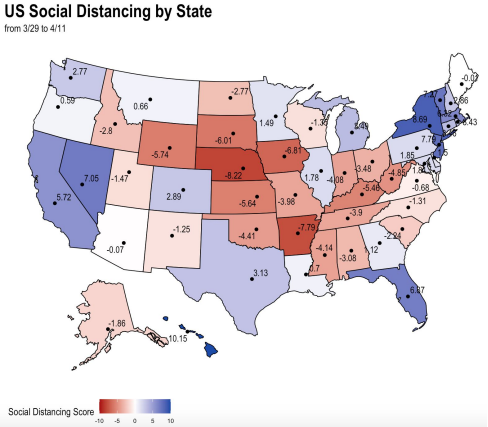
	Date	Retail/Recreation	Grocery/Pharmacy	Parks	Transit Stations	Workplaces	Residential
	2020-03-29	−46.02%	−20.12%	8.64%	−46.14%	−35.94%	11.30%
	2020-04-05	−46.58% ▼	−17.38% ▲	16.40% ▲	−48.52% ▼	−37.76% ▼	11.38% ▲
	2020-04-11	−44.56% ▲	−6.02% ▲	8.76% ▼	−45.22% ▲	−36.44% ▲	13.34% ▲
Avg.	—	−45.72%	−14.51%	11.27%	−46.63%	−36.71%	12.01%



people are primarily halting their activity in retail/recreation, workplaces and transit stations, while increasing their time spent at home and at parks. Also note that while many shied away from groceries early on, they have increased shopping more recently.

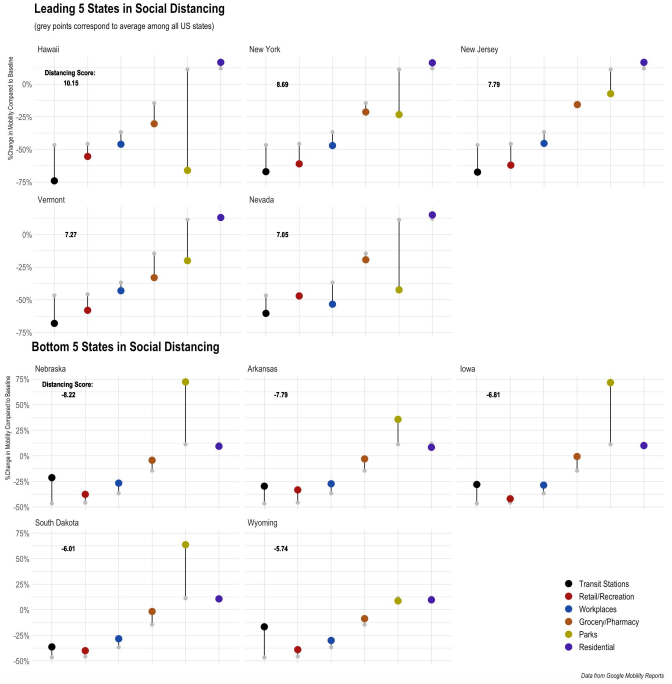
Measuring Social Distancing:

Using the mobility trends for each of the six categories of places, we calculated a composite measure for rating a state's social distancing. Note that higher values (blue) mean better social distancing. Unsurprisingly, social distancing trends appear to be regionally specific, with the north east, west coast, and Gulf Coast distancing the most, while the midwest and deep south are distancing the least.



Best/Worst States

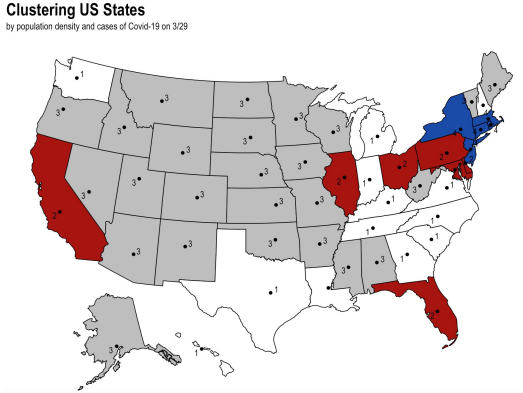
With our new metric, we can now easily rank each state based on how seriously they have taken social distancing.



Understanding Social Distancing Trends in the US

Methods:

In order to more deeply understand our data, we must note that Covid-19 has spread through the US on different timelines. That is, the first case in the US (Washington) was recorded 56 days before the first case hit West Virginia. Moreover, US states have varying population densities that alter the way in which the virus spreads. To control for these two variables, we used k-means clustering to cluster the US into four clusters based on a state's population density and confirmed cases on 3/29 (first date of google mobility data). Much like the social distancing scores, these clusters are fairly regional, notably with cluster 4 consisting of all northeastern states. By controlling for different starting points and population densities, we can examine social distancing measures between comparable states.



Findings:

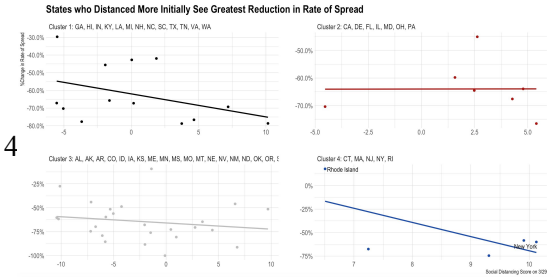
Here are some summary stats of the clusters. In all four clusters, the number of cases per capita increased, while the rate of spread declined. Notice that cluster 4, which distanced the most, experienced the least percent reduction in rate of spread and greatest percent increase in cases/capita. While we do not have mobility data for dates before 3/29, our hypothesis is that this is a consequence of having the largest starting point in cases, causing their distancing, despite being better than all other clusters, to be less effective since the virus had already been so widespread. However, by controlling for cluster, we can see that those who distance more initially see the greatest reduction in the rate of spread. Notice that Rhode Island, who had the lowest distancing score among cluster 4, was the only state (among all clusters) to experience an increase in the rate of spread. Meanwhile, states like New York (cluster 4 were able to successfully distance and decrease the spread.

Cluster	Initial Cases	Initial SDS[1]	Final SDS[1]	Rate of Spread[2]	Cases/Capita[2]
1	2.00K	0.35	−0.07 ▼	−62.4% ▼	315.7% ▲
2	3.20K	2.37	1.23 ▼	−64.0% ▼	389.7% ▲
3	514.36	−2.24	−2.11 ▲	−64.4% ▼	264.0% ▲
4	15.95K	8.62	3.91 ▼	−48.5% ▼	433.5% ▲

[1] SDS denotes Social Distancing Score

[2] Denotes %change from 3/29 to 4/11

All values are averages within each cluster. Initial/Final refers to data from 3/29 and 4/11, respectively.



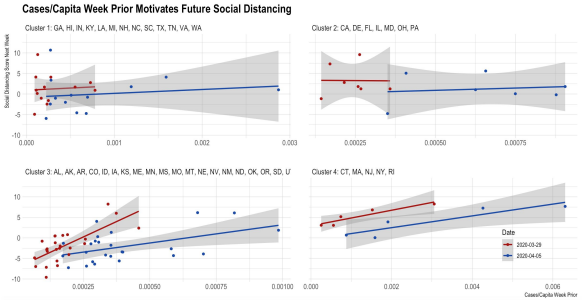
What Influences the Magnitude of a State's Social Distancing?

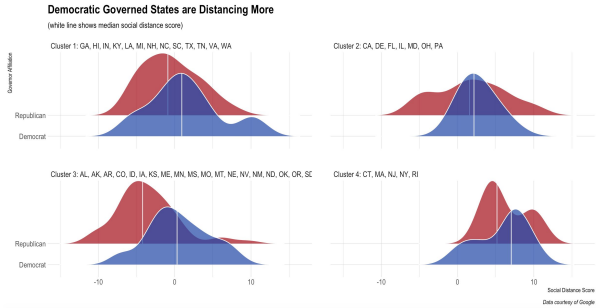
Driving Question:

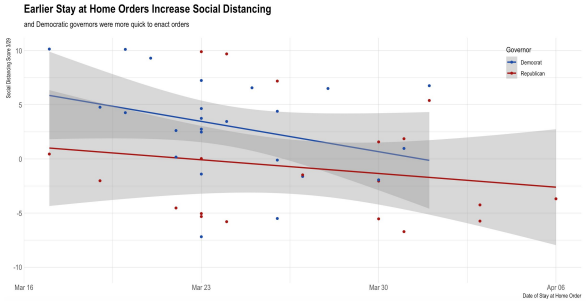
We have seen that social distancing has been effective in decreasing the spread of Covid-19, and that distancing preemptively (before many cases have hit the area) and seriously yields better results. But what is driving certain states or areas of the country into social distancing and why are some parts of the country doing it better than others?

Findings:

A primary motivating factor in social distancing are the cases/capita of the area. There is a positive correlation between the cases/capita of a state a week prior and the amount they distance the following week. Intuitively, it makes sense that the more pressure a community feels (in terms of the proximity of the virus), the more they are willing to take drastic distancing measures. Notably, most states decreased the amount they distanced over our dataset. It seems that as social distancing works to slow the rate of spread, communities may feel less inclined to distance in the future. As shown in the following graphic, the governor affiliation of a state seems to be a major factor in determining the way in which a state distances. There is no question that states with a Democratic governor have distanced more, regardless of cluster. Digging deeper, this has a lot to do with stay at home orders. Enacting these orders earlier correlates to higher social distancing scores, and Democratic governors have given these orders sooner on average than their Republican counterparts. The graph to the left excludes the six Republican governed states who have yet to declare stay at home orders as of 4/20 (all Democratic governors have issued this order).







"COVID-19 Community Mobility Report." *Google*, Google, 2020, www.google.com/covid19/mobility/.

"State Data and Policy Actions to Address Coronavirus." *The Henry J. Kaiser Family Foundation*, 17 Apr. 2020, www.kff.org/health-costs/issue-brief/state-data-and-policy-actions-to-address-coronavirus/#statelevledata.

"Coronavirus Locations: COVID-19 Map by County and State". USAFacts. usafacts.org/visualizations/coronavirus-covid-19-spread-map/