State Measures on Summer Travel

W203 - Summer 2021
Authors: Mai La, Matt Pribadi, Matt Lyons

Defining Research Question

The 2019 novel coronavirus (COVID-19) pandemic had a massive impact on many aspects of life around the world. Although many industries were impacted, the tourism industry, in particular, was devastated. Because each state within the U.S. took a different approach to public health mandates during the pandemic, each state's tourism industry was impacted to a different degree. As part of the tourism bureau, our team is interested in understanding how to restore our tourism industry to pre-pandemic levels. In order to do this, our team proposes a research question examining **COVID-19 control measures taken in a state effect on the number of long-distance trips taken in June 2021.** In particular, by focusing on long distance travel in summer 2021, we can develop a better understanding of the important contributors to revenues associated with travel including hotels, restaurants, airfare. This time period was chosen because of many reasons: the simultaneous start of summer vacation trips, the lifting of several mask mandates among the states, and the push in vaccination rates by the federal government.

Variables & Data Sources

In regards to the outcome variables, we derive the number of long-distance trips per capita for each state by selecting how many trips above 100 miles are taken, standardized by the population of that state. In regards to each states' policies, we are examining several key predictors, specifically around mask mandates, vaccination rates, and closure and reopening data for different industries. For all of these variables, we are controlling for the number of COVID-19 cases and work from home capabilities within each state.

A. **Trips by Distance**: from the Bureau of Transportation. This data set is produced from collecting mobile device location data, listing the number of trips aggregated by distance taken by residents in different states.
B. **State Population**: from the US Census Bureau Survey, the data set contains population and other demographic information for all US states
C. **Mask Mandates**: from the COVID-19 US State Policy database (CUSP), with information about the end date of face mask mandate in all states.
D. **Opening of restaurants and recreation business**: from the COVID-19 US State Policy database (CUSP), with information about the reopening date for businesses (i.e. bars, casinos, theaters, etc.) in all states.
E. **Vaccination Rate**: from the CDC Data on Vaccinations, include the percentage of population completed COVID-19 vaccination in all states.
F. **Number of Covid-19 cases**: from the New York Times Covid-19 Data, include the number of cases and deaths in all states.

G. **Work From Home Data**: from the Bureau of Transportation Statistics. These data include income levels and a measurement of how many households included at least one adult who began shifting at least in part to telework.

Plan of Action

Several of the variables above will require transformations, specifically those which include dates. The outcome variable has daily values, which may be too granular to be useful. We will standardize the total trips taken on the state population, so that it will represent the sum of trips of distances >100 miles taken per capita in a given state over the course of the selected time period. For the data states' policies (e.g. mask mandates, recreational facilities closing and reopenings), we will calculate the number of days that the policy is in effect within June 2021 and divide it by 30 days to standardize it. For example, if the state of Vermont lifted their mask mandate on June 14th, then that variable would be 14 divided by 30, which is 0.47; however, if the mandate were lifted on any date before June 2021, the variable would be 0. As another example, if bars were closed on a given week but were opened the following week in June, the total number of days the bars were open would be divided by 30 resulting in a number between 0 and 1. For states with multiple closings/reopenings, we will follow the same logic.

After cleaning and structuring the data, we will build three regression models:

The first model will examine the relationship between our output variable, the number of trips longer than 100 miles taken per capita and our input variables including mask mandates, vaccination rates (%), and open/closed status of various recreational facilities.

The second model will use all of these variables and add per capita COVID-19 cases as a control variable. By adding this variable, we expect we will have a better picture of the effect of the policies themselves, independent of whether these policies are having an impact on COVID cases.

Lastly, and more experimentally, we will build a third model including several other variables: the CUSP (opening and closing) data, the Census Bureau demographic data, and the BTS Travel Behavior data on income level and work-from-home increases. We would like to examine (1) whether we may be missing potentially powerful predictors, (2) determine whether these effects might disappear when certain demographic variables are controlled for, and (3) try to better assess collinearity among our variables.

We will include all of these parameters in a regression table and make a final recommendation of which model best suits our goals. Our goal is to determine what combination of policies was best aligned with higher amounts of travel so that these policies can be encouraged among the remaining states. We believe that by implementing these effective policies nationwide, the tourism industry can recover more quickly and once again continue to flourish.