Team J Cubed: Jiahao Huang, James Lo, Jessica Wong
Emails: jiahao303@gmail.com, goobyme@ucla.edu, jessicacwong@g.ucla.edu

Components of Social Distancing

The COVID-19 pandemic has forced legislators across the United States to enforce social distancing measures. The extent of social distancing restrictions varies from state to state, but additionally, each state often experiences marked differences in effectiveness from county to county. Our team wanted to explore the different components of social distancing and which demographic factors may affect each component. To do so, we used the BigQuery data warehouse to query data from the Google Community Mobility Reports, which monitor changes from baseline in movement across six categories: retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential areas. We compared counties' changes from baseline with demographic information using the Bureau of Economic Analysis' GDP and Income by County dataset, as well as the United States Census Bureau's American Community Survey. Using county data from all three sources, we investigated whether factors such as per capita income, population, and number of families with young children (weighted by population) were associated with trends in any of the six types of movement changes.

We first examined relationships between demographic factors and movement trends for counties in California. Per capita income appeared to be associated with an increase in movement among residential areas, as well as reductions in movement for grocery stores and pharmacies, retail and recreation areas, workplaces, and transit stations. Per capita income did not display a correlation with movement trends for parks; park visit changes have varied greatly between counties, likely because some people prefer to avoid public areas while others seek exercise at parks rather than gyms, and this phenomenon does not appear to be different across socioeconomic statuses. We hypothesized that high per capita income for a county is associated with greater movement among residential areas and less movement among all other locations (except parks), as wealthier counties are likely to have more residents in white collar jobs where remote work is possible, while less wealthy counties are likely to have more residents in blue collar jobs where remote work is not feasible, restricting the extent of their social distancing. We used logarithmic models to represent the relationships between per capita income and the six movement categories, with the highest R-squared value appearing in the model relating change in residential movement and income per capita in California counties as 0.62926. The positive association from the model aligns with our hypothesis. A logarithmic transformation is likely appropriate because per capita income has a range spanning more than one order of magnitude, from around $30,000 to over $130,000—the four counties in the upper bound have very high costs of living, softening the marginal effect of income on movement.

We then investigated the relationships on a national level, finding positive associations between per capita income and change in residential movement and negative associations between movement for all other locations (except parks, which displayed a lack of significant correlation likely for a similar reason). Logarithmic transformations were again useful since the range of per capita income spanned more than one order of magnitude. For national county data, we also explored the population and the proportion of families with young children as possible influential factors within the context of the relationship between earnings and change in movement. For all movement categories except parks, county population seems to be associated with larger changes in movement. Additionally, it appears that counties with some of the lowest weighted amounts of families with young children tend to experience less changes in movement, although there do not appear to be significant differences between counties with relatively moderate versus relatively high weighted amounts of families with young children.