Team Mejia_Tantsyura_Yang Project Proposal 2

Team members: Emanuel Mejia, Yixi Yang, Nadia Tantsyura

Team github: https://github.com/UC-Berkeley-I-School/Project2_Yang_Tantsyura_Mejia.git

All countries in the world were affected by the COVID-19 pandemic. However, countries dealt with the outbreak quite differently, especially in terms of lockdown stringency. Some countries were seen as strict practitioners and others lax, with a few prominent examples in each type, but there is no exact grouping for many others in the middle of the stringency spectrum. Lockdown strategy also shifted as countries balanced economic costs with public health and as vaccines became available. There is much debate on which approach is better for society, but we should be clear about how countries employed stringency over time before we attempt to evaluate the merits of the strategies.

**Our main question is: how did countries use lockdowns as a pandemic response and how did that correlate with demographic factors and COVID severity?**

We will answer this question through two parts: **what were the stringency levels of different countries overtime and how did they correlate with other factors such as demographics or COVID cases?** Our **primary dataset** is the OWID COVID dataset made available by Our World in Data. The dataset contains many variables of interest, and we do not foresee need for a supplemental dataset.

For the first part of our question, our **variables of interest** are:

- Stringency_index: Government Response Stringency Index, a composite measure based on 9 response indicators including school closures, workplace closures, and travel bans, rescaled to a value from 0 to 100 (100 = strictest response)
- Location: the countries whose data we are examining
- Date: we have data over a period of time, starting from early 2020 to now

We would like to see, on a weekly basis, the stringency levels of each country and group them into meaningful clusters based on the distribution of the stringency levels. We would track how group membership changed overtime and note any moments of drastic change. For example, perhaps many countries relaxed restrictions when summer came, or Europe and the US fluctuated while Australia remained strict. This could be interesting to see if popular stereotypes are confirmed, or if there are any surprises in which countries adopted which strategy. Furthermore, we may see that lockdowns are not black-and-white, and that there may be several broad levels of stringency, each with their adherents.

For the second part of our question, our **variables of interest** are:

- Continent: the continent for each country where we have data
- Positive_rate: The share of COVID-19 tests that are positive, given as a rolling 7-day average (this is the inverse of tests_per_case)
- People_vaccinated_per_hundred: Total number of people who received at least one vaccine dose per 100 people in the total population
- People_fully_vaccinated_per_hundred: Total number of people who received all doses prescribed by the vaccination protocol per 100 people in the total population

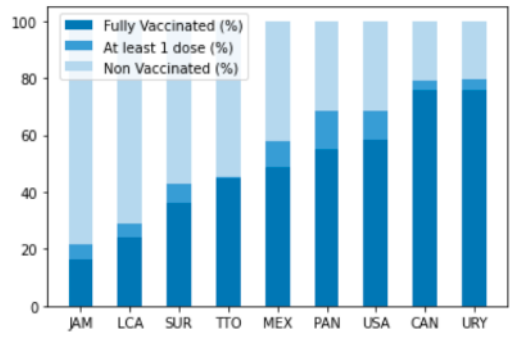
- Population_density: Number of people divided by land area, measured in square kilometers, most recent year available
- Aged_65_or_older: Share of the population that is 65 years and older, most recent year available
- Gdp_per_capita: Gross domestic product at purchasing power parity (constant 2011 international dollars), most recent year available
- Hospital_bed_per_thousands: Hospital beds per 1,000 people, most recent year available since 2010
- Human_development_index: A composite index measuring average achievement in three basic dimensions of human development—a long and healthy life, knowledge and a decent standard of living. Values for 2019, imported from http://hdr.undp.org/en/indicators/137506
- New_cases_smoothed_per_million: new confirmed cases of COVID-19 (7-day smoothed) per 1,000,000 people
- New_deaths_smoothed_per_million: new deaths attributed to COVID-19 (7-day smoothed) per 1,000,000 people

We would like to see if the stringency level of lockdowns correlated with any or all of the factors above, such as geographic location (which could be a proxy for cultural differences), health resources available to deal with a pandemic, economic resources, testing prevalence, vaccination rates, and severity of the pandemic. We could even check for correlation of such factors broken down by continent – for example, perhaps Africa has low GDP per capita but also a young population, so lockdown stringency is less severe than an equally poor country in Asia that has an older population.

Some data cleanup would be necessary before any analysis. Not all variables contained data for all countries, and some data were available weekly while others were updated daily. We will reduce the number of countries to those that contained data in most of our variables of interest, which are about 130. We will also average daily data into weekly data for better matching of time stamps across variables. Additionally, there is no testing data available for China, but we have decided to keep it for our other studies because it is too large a country to ignore.

Some sample graphs are shown below, highlighting the vaccination rates of different countries in the Americas and Europe.

*Americas*




*Europe*