

1 Global Analysis of COVID-19 Related Deaths

Tyler Barton, Kayla Benitez, Patrick Chen, Victoria Christensen

1.0.0.1 Author contributions

1. Tyler Barton prepared the data description as well as created Figure 1 and wrote the results and discussion that followed.
2. Kayla Benitez studied the topic and prepared the background and worked on the abstract and methods paragraph.
3. Patrick Chen studied the basic properties of the dataset and created Figure 2 and wrote the results that followed.
4. Victoria Christensen proposed a few methods to explore questions about the dataset and worked on the abstract and aims and compiled our work into one file.
5. We all worked on fixing the plotting codes together and proposing interesting questions to explore.

1.0.0.2 Abstract

The COVID-19 pandemic has resulted in the deaths of millions of people around the globe and we were interested in seeing the difference in COVID-19 death rates between continents over time. We set out to explore how gdp per capita impacted the cumulative number of deaths per million for each continent and we expected to find a positive relationship between the two. We found that as gdp per capita increases, total deaths per million decreases except in North America and Africa. We also looked out how the COVID-19 death rate changed over time among the different continents by analyzing the data on a monthly basis and we expected to see a decline with the introduction of vaccines and new treatments. We found that at the start of the pandemic COVID-19 death rates were high but then began to decline at the end of 2020 and have stayed below 3% ever since.

1.1 Introduction

1.1.1 Background

We will be analyzing COVID-19 data for various countries to see the global impact the pandemic has had. According to the World Health Organization, Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus. Where most people infected with the virus will experience mild to moderate respiratory illness and recover without requiring special treatment. However, some will become seriously ill and require medical attention. Vaccines, testing and other variables aim to decrease the transmission of the COVID-19 virus and reduce cases and the severity of cases that lead to ICU & hospitalizations or even death. Our dataset comes from data maintained by [Our World in Data](https://github.com/owid/covid-19-data/blob/master/public/data/README.md#confirmed-cases) (<https://github.com/owid/covid-19-data/blob/master/public/data/README.md#confirmed-cases>) (OWID) that has been updated daily throughout the duration of the COVID-19 pandemic. The dataset includes multiple countries and has various variables mentioned below.

The metrics of our dataset includes:

- Continents
- Locations
- Vaccinations
- Tests & Positivity
- Confirmed Cases
- Confirmed Deaths
- GDP per Capita
- Total deaths per million

To explore the data, we can keep track of:

- Monthly number of confirmed cases
- Cumulative number of confirmed deaths
- Death rates
- GDP comparison

The COVID-19 pandemic is still ongoing and with this data, we wish to explore the impact of the pandemic and how countries have been able to mitigate cases. We will analyze for any trends, or relationships, such as whether total deaths or death rates have decreased over time and the effect that tests, vaccinations, policy responses, and other variables of interest have on confirmed cases, and confirmed deaths.

1.1.2 Aims

This project aimed to explore the difference in the total number and rates of COVID-19 related deaths between the different continents around the globe. Using visualizations and smoothing curves, we explored how gdp per capita impacted the number of COVID-19 related deaths. We had expected countries with lower gdp per capita to have higher COVID-19 related deaths because those countries probably had less resources to fund vaccinations and testing sites. We had expected countries with higher gdp per capita to have lower COVID-19 related deaths. We found some evidence for several of the continents that did not support our initial predictions. We also examined the differences in the COVID-19 related death rates between continents over the course of the pandemic. As vaccinations and testing became more widespread, we expected the death rates to drop which did occur in every continent.

1.2 Materials and methods

1.2.1 Datasets

Name	Variable description	Type	Units of Measurement
date	Date of observation	String	Calender Year
continent	Continent of the geographical location	String	Continent
location	Geographical location	String	Country
gdp_per_capita	Gross domestic product divided by population	Numeric	Local current currency
total_cases	Total confirmed cases of COVID-19	Numeric	positive covid tests
total_deaths	Total deaths attributed to COVID-19	Numeric	Mortality number from covid
total_deaths_per_million	Total deaths attributed to COVID-19 per 1,000,000 people	Numeric	Deaths per 1,000,000 people count

Hannah Ritchie, Edouard Mathieu, Lucas Rod  s-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Diana Beltekian and Max Roser (2020) - "Coronavirus Pandemic (COVID-19)". Published online at OurWorldInData.org. Retrieved from: '<https://ourworldindata.org/coronavirus>' (<https://ourworldindata.org/coronavirus>) [Online Resource].

1.2.1.1 Sample and measurement information

All 225 locations sampled received updated data every day. The data was collected with the help of thousands of researchers and the information of COVID-19 tests worldwide. The COVID-19 related death and cases data is collected with the help of health clinics and organizations around the globe. There are missing values because many countries did not update the values every single day or some countries did not record those values at all.

The relevant population is the globe as it is a sample of 225 locations that reported their data as they wanted to get a wide range of data. The sampling mechanism is through the use of the SDG-Tracker which is paired with Our World in Data and thousands of health organizations to gather the data. This is a convenience sample because it is not a random sample, the data comes from countries who voluntarily submitted their data. The sampling frame is a subset of the population where the sampling frame consists of the countries reporting and the population consists of the entire globe. Because data is dependent on voluntary reporting the scope of inference can not be specified.

1.2.2 Methods

Our methods consisted of focusing on a sample of the original COVID-19 dataset, filtering out by dates and variables. The frame we focused on was within the timeline of March 2020 until February 2022. The variables utilized were 'date','continent', 'location', 'gdp_per_capita', 'total_cases', 'total_deaths' and 'total_deaths_per_million'. We explored the data by initially using a scatter plot of gdp per capita vs total deaths per million. The total_cases were differentiated by the size of the dots and the continents were graphed by the different colored dots. The scatter was plotted with a smooth trend line to determine any patterns or trends in the data. This plot would determine if the economic situations of the countries affected the total_deaths experienced. Our exploratory analysis also consisted of plotting the continents versus the death rate for COVID-19; where the death rate was calculated by dividing the total_cases by total_deaths. We utilized this plot to determine which continents experienced the highest death rates compared to the global data.

1.3 Results

Figure 1: Compares gdp per capita to total deaths due to COVID-19 per million colored by continent. The size of the dots represent the total cases reported in the various continents

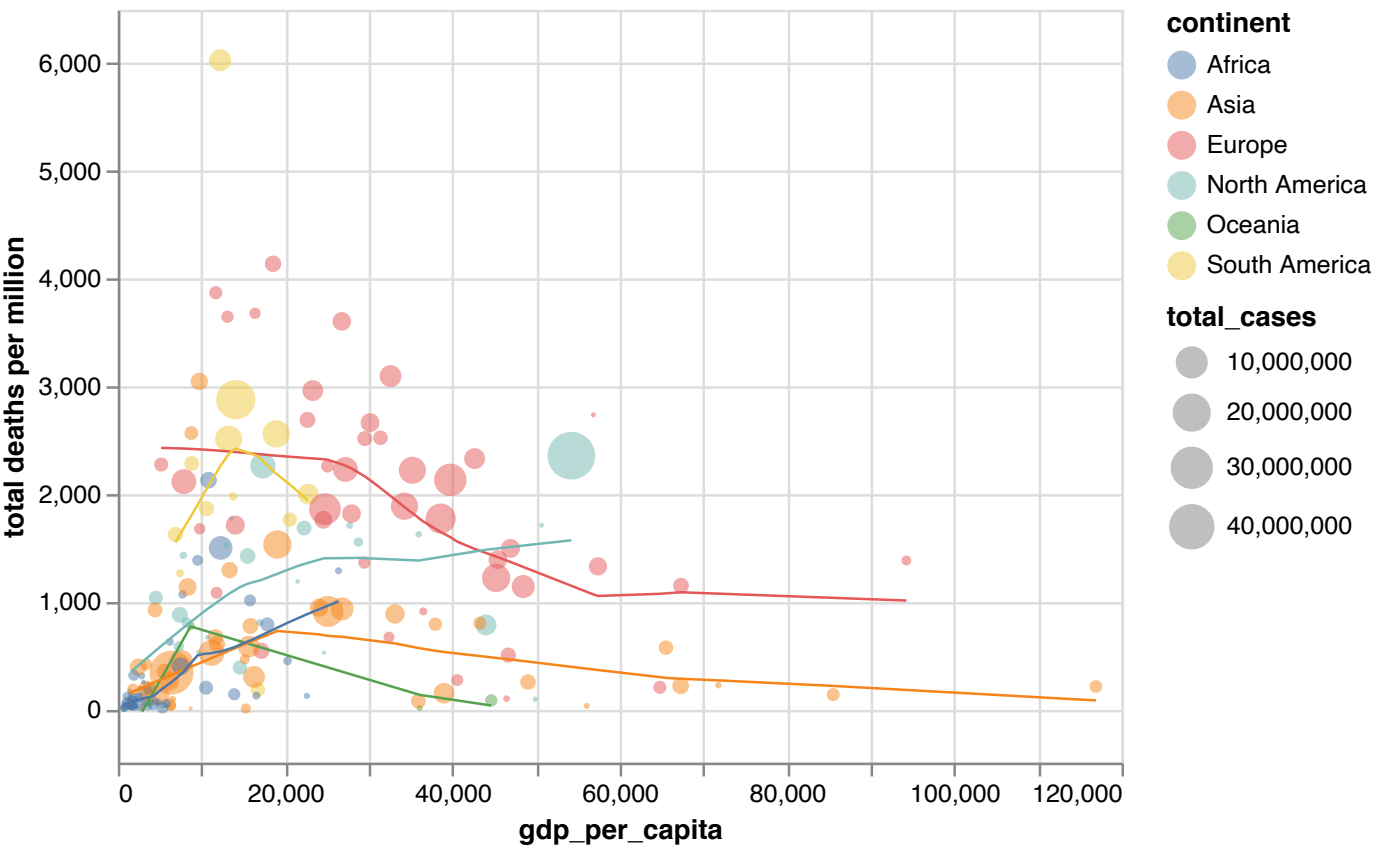
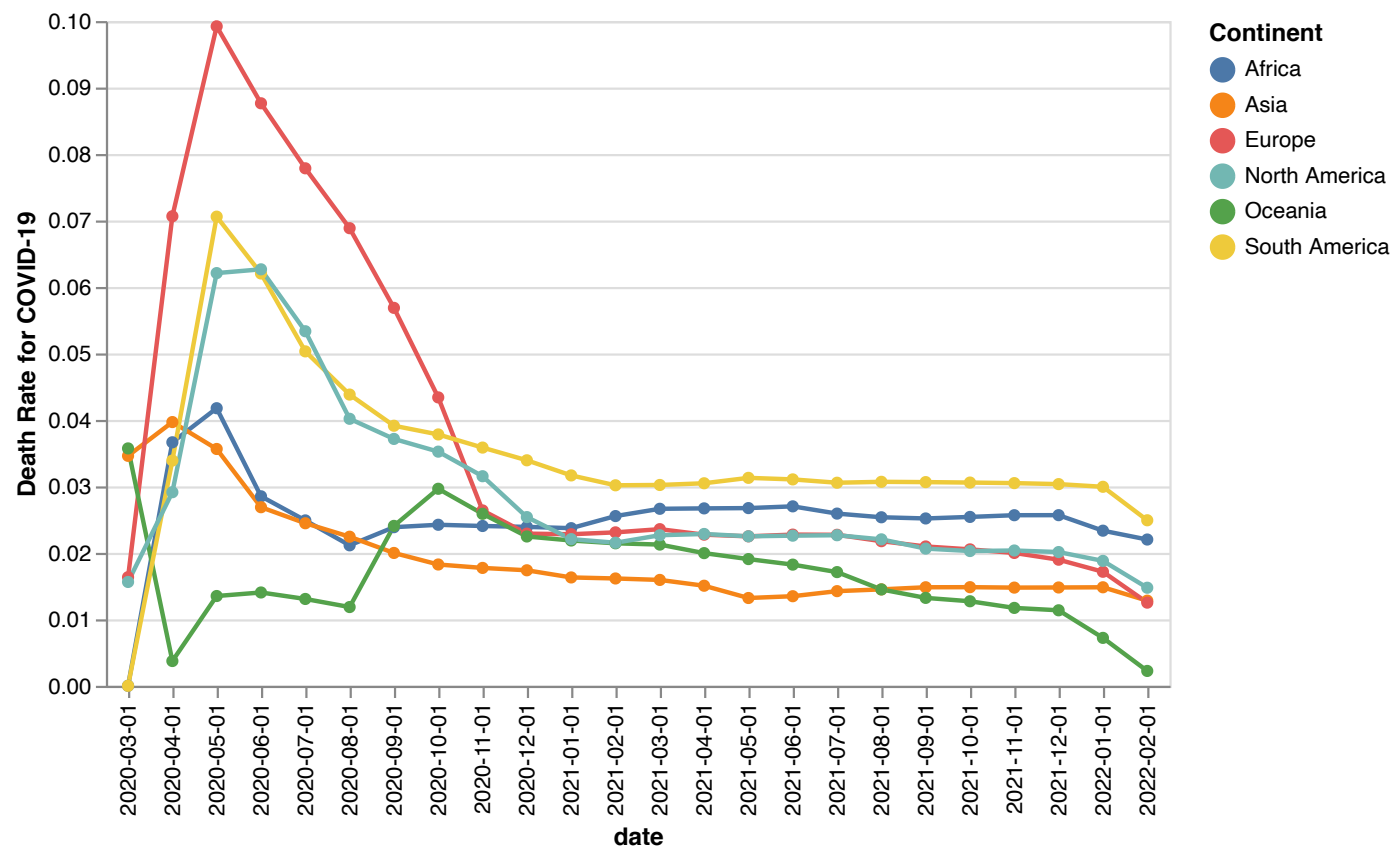


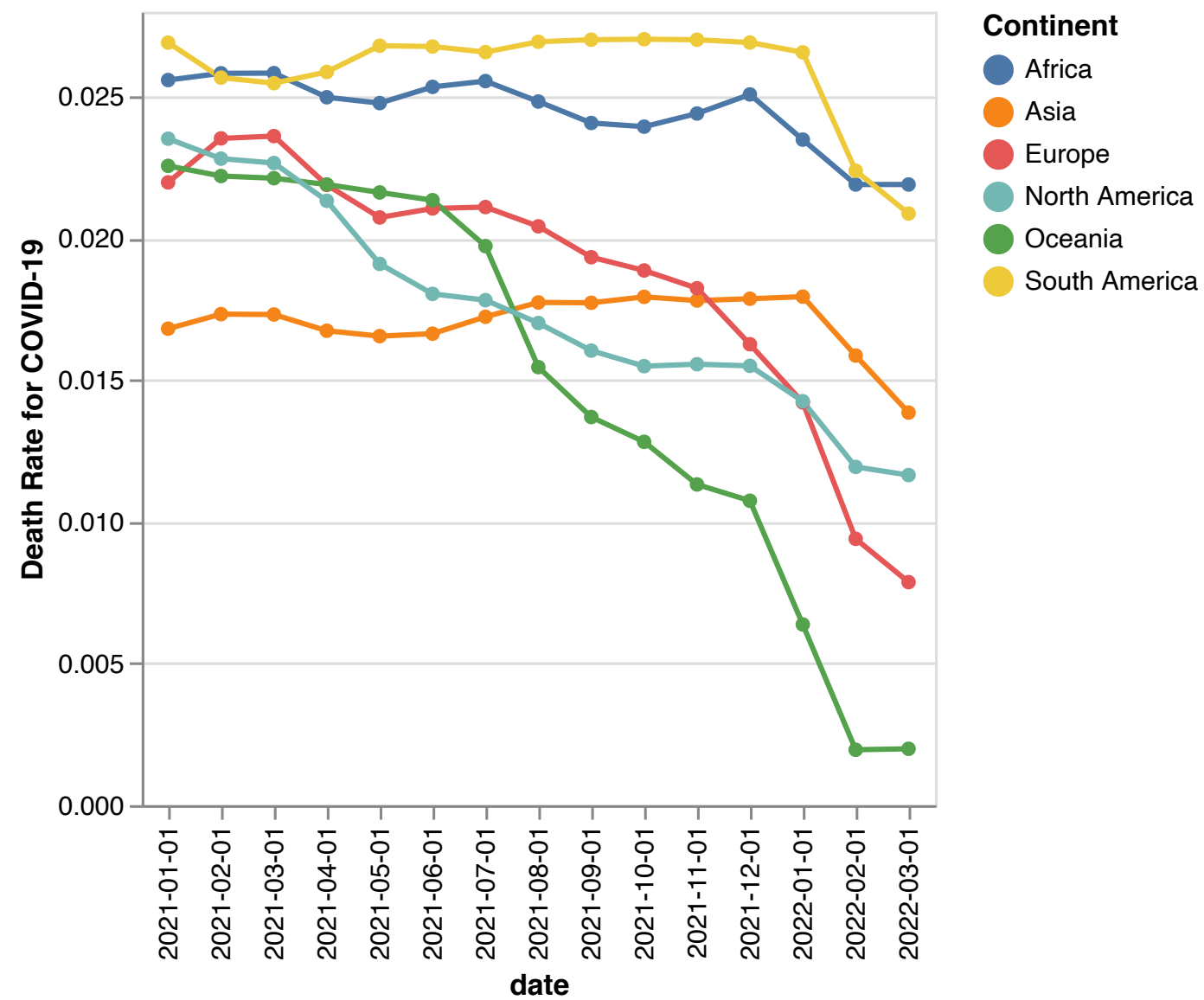
Figure 1 shows that as gdp per capita increases, total deaths per million tend to go down except for North America and Africa. The trend lines show the average gdp per capita by total deaths per million for each continent. The countries with the highest gdp per capita have some of the lowest death rates because of COVID-19 out of the countries in the dataset.

Figure 2: Death Rate for COVID-19 on each continent besides Antarctica from March 2020 to February 2022.



Overall, each continent had a higher death rate for COVID-19 in 2020 than 2021. From the plot, it shows the peak for COVID-19 globally was around May and June 2020. All continents had their highest death rate except Oceania. Europe had the overall highest death rate (almost 10%) and South America had a 7% death rate for COVID-19 during the summer of 2020. Oceania had the highest death rate at 3% around October 2020. After the year of 2020, every continent had a stable death rate in the range from 1% to 3%.

Figure 3: Death Rate for COVID-19 on each continent besides Antarctica from January 2021 to March 2022



Basically, this is a zoom-in version of Figure 2. Since January 2021, every continent has had a lower than 3% death rate. Starting from December 2021, the death rate dropped globally. Until March 2020, Africa had 2.2%, South America had 2.1%, Asia had 1.4%, North America had 1.2%, Europe had 0.8%, and Oceania had 0.2% death rate due to COVID-19.

1.4 Discussion

Looking at Figure 1, the positive trend lines in Figure 1 for North America and Africa can be due to various reasons such as more testing, inadequate funding even in countries with money, people in the upper class who are anti-mask or anti-vaccine, or funding incentives to mark more deaths as COVID-19 deaths even when the person could have died from another cause. Moreover, using Figure 1 we can see how countries in Europe, Asia, Oceania, and South America with more gdp per capita have less deaths which could be because of more funding in healthcare, sanitation, testing, and vaccines. There exists an outlier in South America (Peru) that has 6031.551 deaths per million people with a GDP of 12,236.706. This means that Peru has the highest death rate due to COVID-19 compared to the other countries.

Looking at Figure 2, we can see that the death rates spiked in most of the continents right before the stay at home orders were put in place. We noticed that Oceania never had a drastic spike because they mitigated cases by shutting their borders down and having strict COVID-19 protocols. In January 2021, the COVID-19 death rates decreased to levels below 3% and this might be due to mask mandates, more knowledge of effective drugs such as Remdesivir, and the development of vaccines. Remdesivir was approved by the FDA in late October of 2020 and according to their website it is known to help alleviate symptoms and potentially reduce chances of dying from the virus. Overall we see that the death rates decreased in all continents due to the COVID-19 preventative measures and treatments.

We would have liked to explore more deeply into how certain countries dealt with the pandemic such that we could research countries that did well during the pandemic and countries that did not and compare how the death rate, total cases, and total vaccines shifted over time. Furthermore, we could have seen which countries had the highest and the lowest death rates and compared how the variables changed over time for these countries. With Figure 1, we did a general view of this, but going into more detail on the outliers and finding out why the trends happen would be a good topic to explore more in.