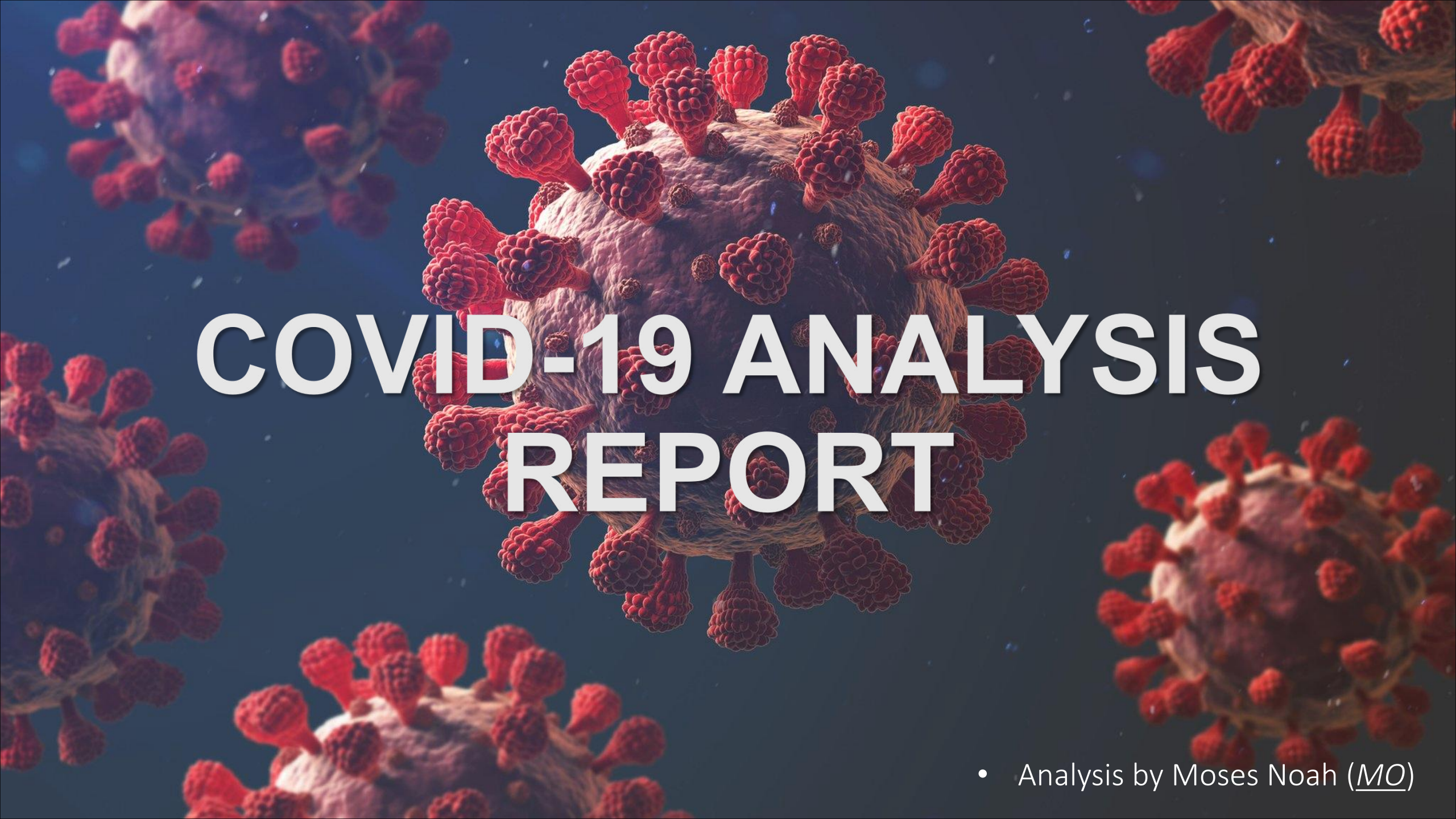


COVID-19 ANALYSIS REPORT

- Analysis by Moses Noah (*MO*)

INTRODUCTION

COVID-19

COVID-19, caused by the SARS-CoV-2 virus, emerged in late 2019 leading to a global pandemic that impacted health, economies, and daily life worldwide. With a range of symptoms and severe outcomes, it prompted unprecedented public health responses, including social distancing, lockdowns, and vaccination drives.

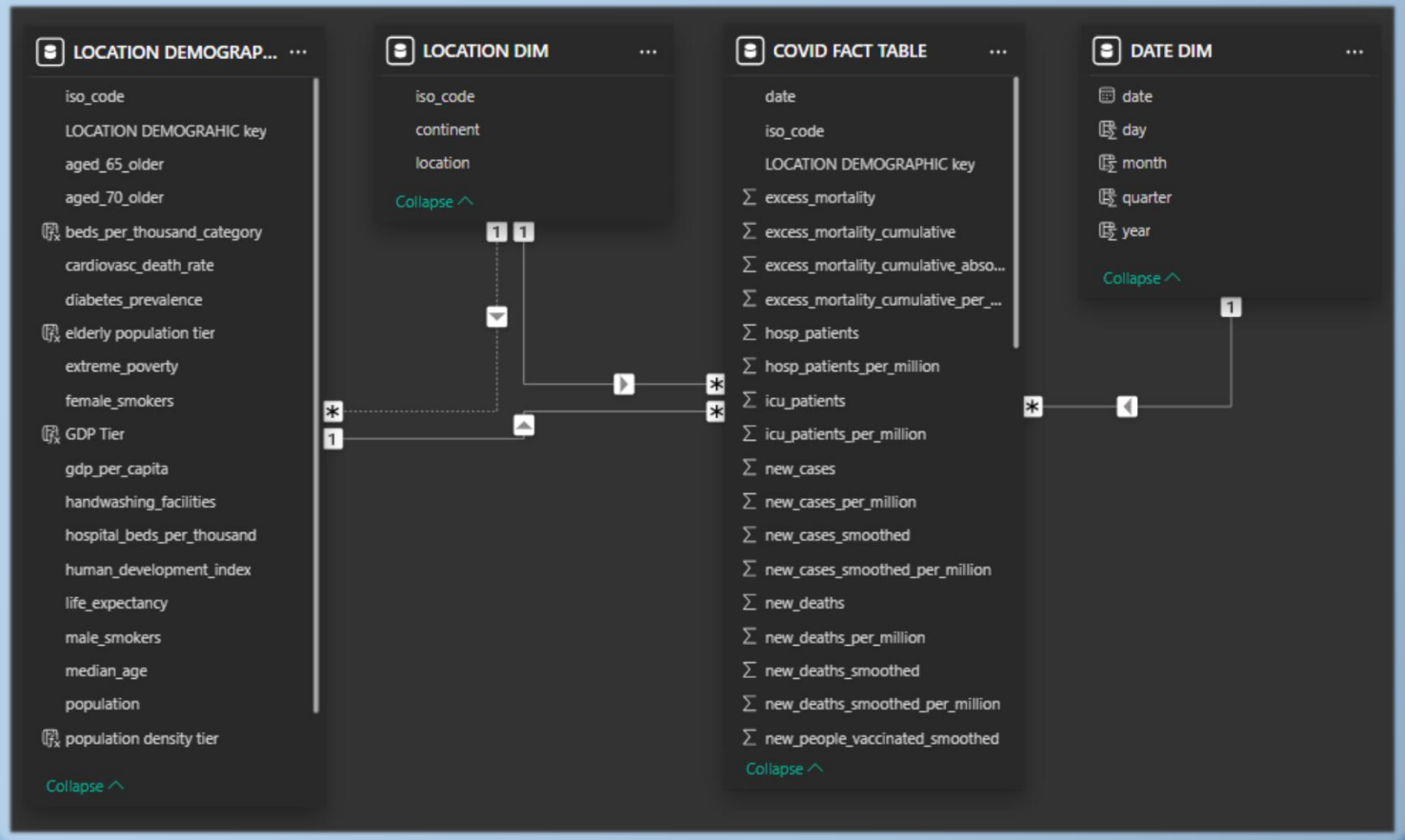
Analyzing the spread, response measures, and impacts of COVID-19 offers insight into managing infectious diseases and understanding factors that shaped the pandemic's trajectory.

Data Overview

Data Source: Our World In Data(owid-covid-data)

Data Model

The data is structured using a hybrid of Star Schema model and Snowflake Schema model (due to further normalization of the Location Dimension), normalized into four main tables: a central **COVID Fact Table** and three supporting Dimension Tables—**Location Demographics Dimension**, **Location Dimension**, and **Date Dimension**. This design allows efficient, structured analysis of COVID-19 trends by linking key demographic and temporal information to case data.



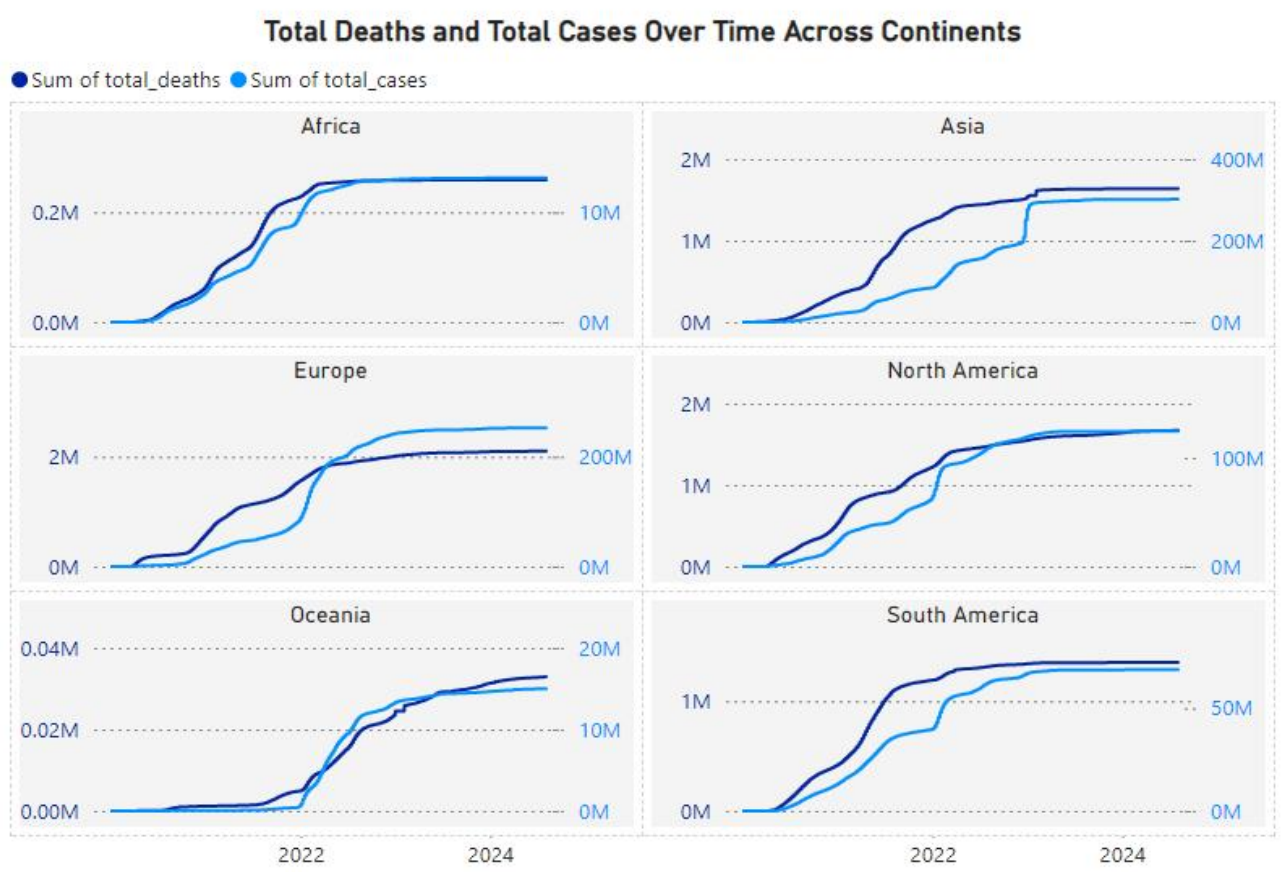
Data Cleaning: Cleaning the dataset was a fairly straightforward process since most of the data was pre-prepared for analysis. Key steps involved converting data types to ensure accuracy and standardizing decimal points, replacing commas with periods where needed. Missing values were managed appropriately, 'null' was used for non-numerical fields while '0' filled gaps in purely numerical columns. These steps ensured consistency, accuracy, and readiness for the insights I aimed to extract.

Data Transformation: For the data transformation phase, I used DAX functions to create new columns across all tables, enhancing analysis capabilities. Key transformations included categorizing GDP, elderly population, and population density into defined tiers, which allows for more focused comparisons across socio-economic levels. I also organized the data by creating a stringency index category to classify government response levels and added breakdowns for time by days, months, quarters, and years in the Date Dimension. These transformations enable streamlined, multifaceted analyses within the model.

Research Questions

1. How have total cases and total deaths evolved over time across different continents?
2. How did the trend of new cases differ between countries with stringent lockdown policies and those with more relaxed measures?
3. What is the trend of deaths in countries with limited healthcare capacity compared to those with greater healthcare capacity?
4. How does population density affect the spread of COVID-19?
5. What impact does GDP per capita have on COVID-19 vaccination rates?
6. What impact does GDP per capita have on COVID-19 mortality rates?
7. How does the stringency index correlate with the reduction in new cases over time?
8. Did countries with more stringent government response have fewer deaths than those with less stringent measures?
9. What is the relationship between the number of people fully vaccinated and the decline in COVID-19-related deaths?
10. Do countries with higher tests rates have lower COVID-19 death rates?

How Have Total Cases and Total Deaths Evolved Over Time Across Different Continents?



These trends suggest that continents with higher healthcare capacity, early vaccination access, or strict public health measures managed to reduce the rate of deaths relative to cases over time. The varied response effectiveness highlights the impact of regional factors on COVID-19 outcomes, with population density, health policy, and infrastructure playing key roles.

Key Observations

Asia: Asia experienced the highest cumulative cases, peaking at around 300+ million. However, deaths in Asia remained moderately lower despite it's high population indicating a potential impact of early mitigation measures or healthcare capacity.

Europe: Europe experienced the highest total deaths of all continents at slightly above 2 million but notably the trend of deaths flatlined early into 2022 while cases kept rising (early 2022 was the begining of a new wave) suggesting that Europe might have benefited from vaccinations or improved treatment options during later waves.

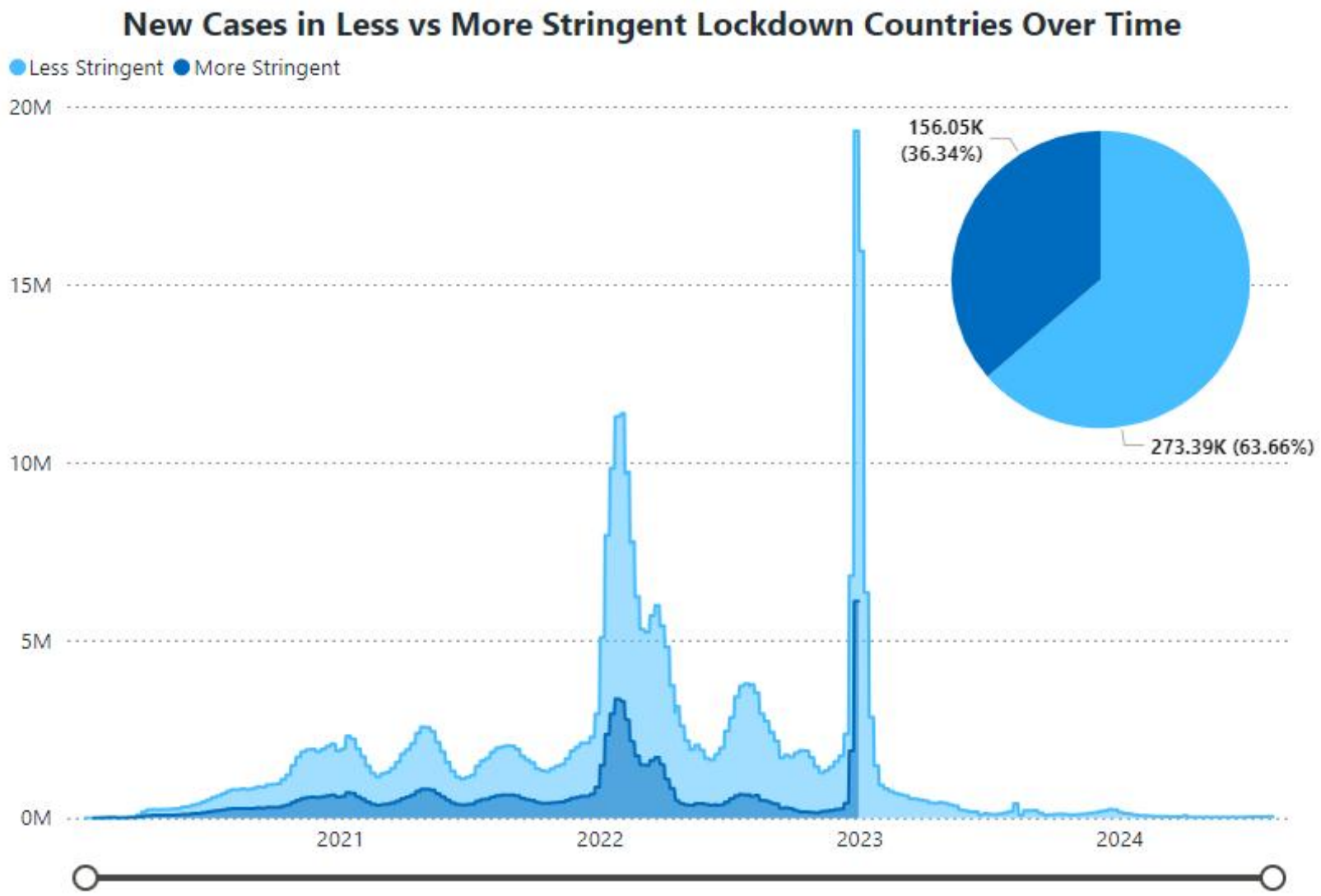
North America: North America saw steady increases in both cases and deaths, with total cases reaching over 100 million. Deaths remained proportionally high, likely reflecting the early pandemic waves that heavily impacted the continent before widespread vaccination.

Africa: Africa saw a rapid rise in both deatgs and cases, Africa also saw a relatively low death rate despite increasing cases, which could be due to factors like a younger population or less urban density.

Oceania: Oceania shows the lowest figures in both cases and deaths, reflecting the continent's strict border control measures and delayed exposure to the virus.

South America: South America experienced rapid increase in deaths and cases (though total deaths is quite lower than other high populace continents), this could reflect challenges in healthcare infrastructure during major waves of infection.

How Did The Trend of New Cases Differ Between Countries With Stringent Lockdown Policies and Those With More Relaxed Measures?



Key Observations

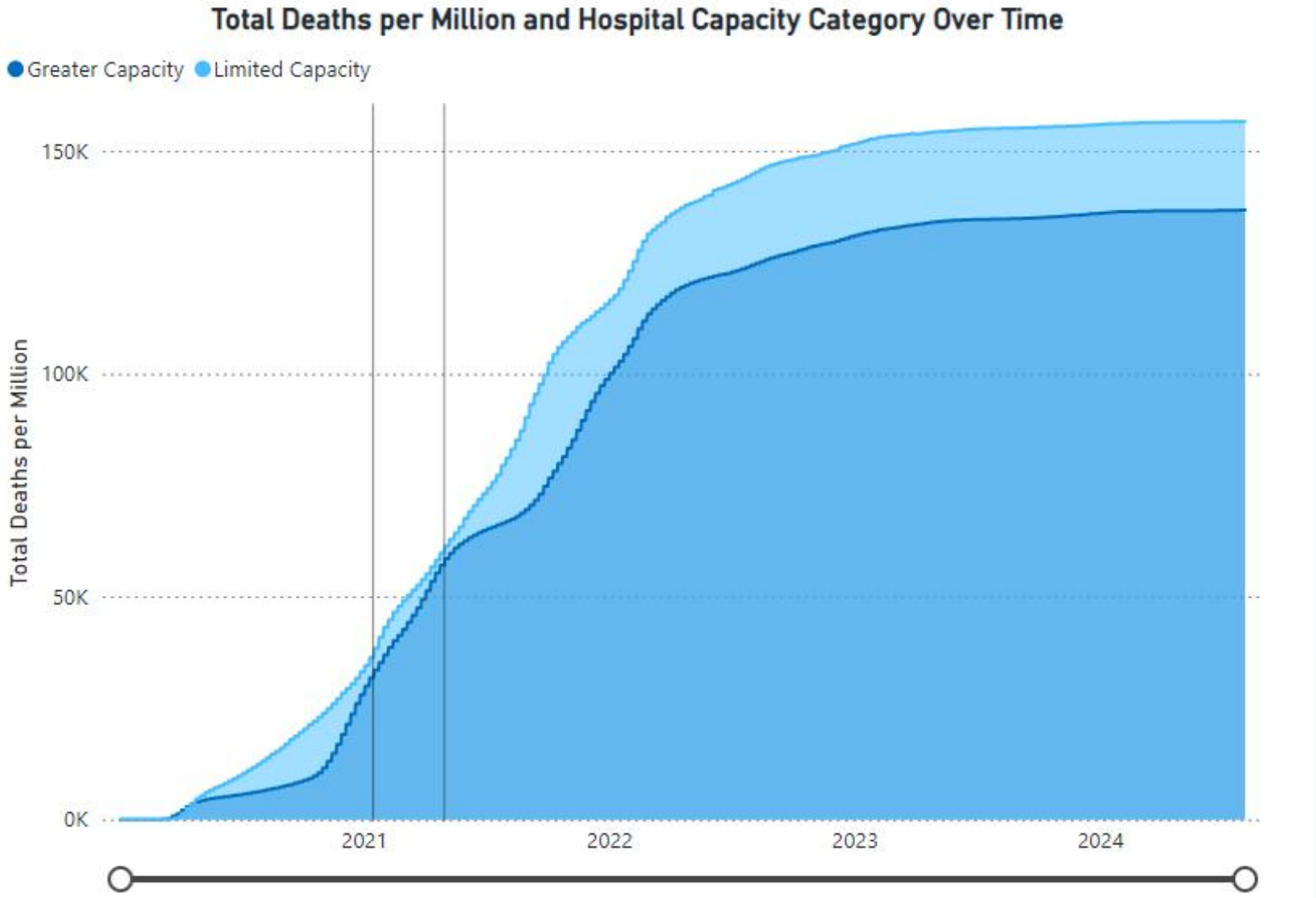
Countries with more stringent lockdown policies saw moderately lesser new cases compared to countries with less stringent policies.

As indicated by the pie chart, there are more instances with less stringent lockdown policies than those with more stringent lockdown policies, this can explain why there are more cases with less stringent lockdown policies.

The chart also indicates several time periods when there was surges in cases around the world (this is depicted by the several bumps in the chart along the x-axis), this is also called *covid-19 waves*

This categorization is realized by dividing the stringency index into two demographics namely Less Stringent and More Stringent, this division is gotten by simply divding the highest stringency index (32.7) in two as there are no officially specified stringency demogrphics.

What is The Trend of Deaths in Countries With Limited Healthcare Capacity Compared to Those With Greater Healthcare Capacity?



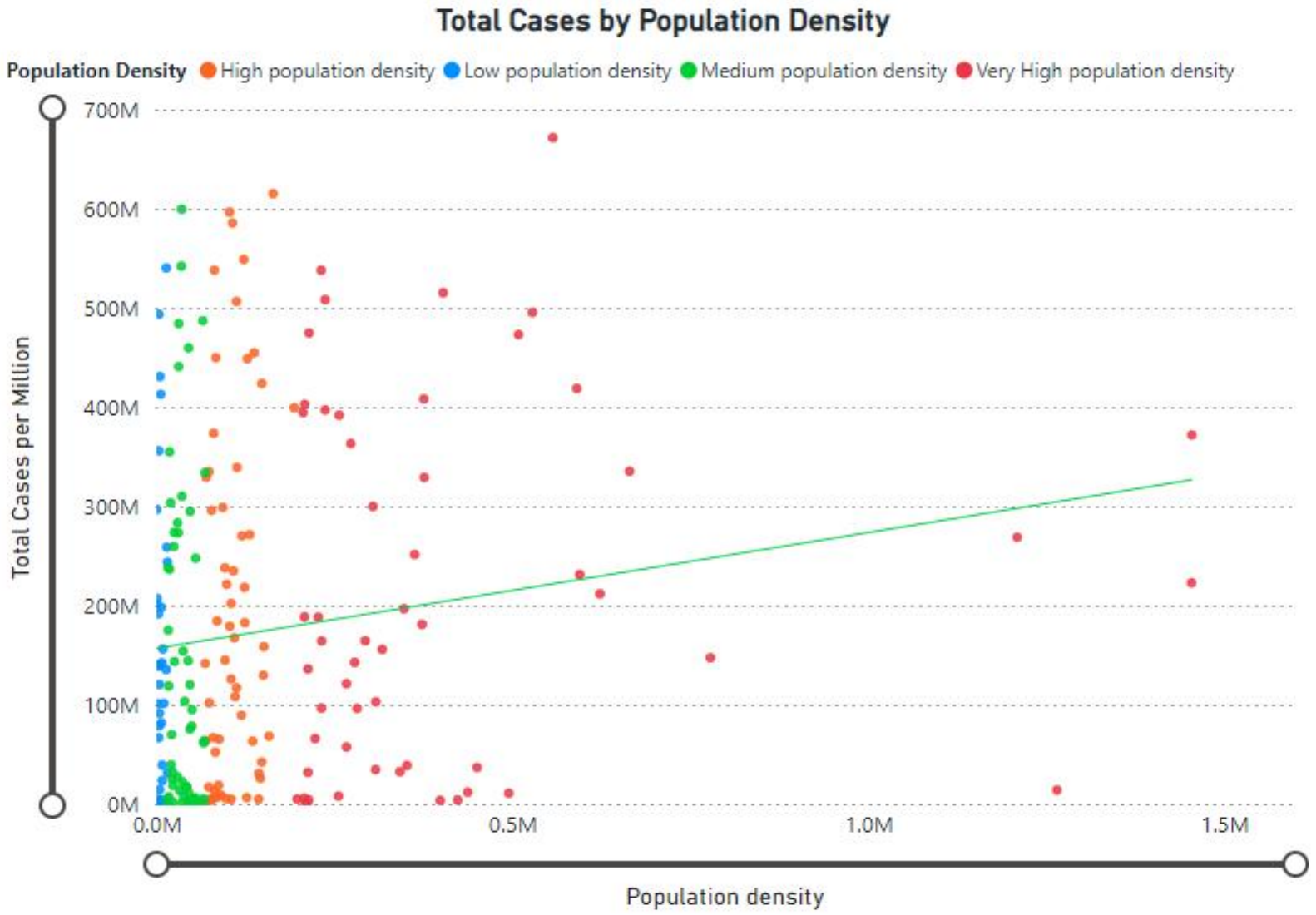
Key Observations

According to the visual to the left, countries with limited healthcare capacity saw a steeper increase in deaths during the early months of covid-19 outbreak compared to countries with greater healthcare capacity, this gap will go on to be narrowed during the *covid-19 wave* around early-2021 and also around mid-2021.

Countries with greater healthcare capacity are countries with strong economies/first world countries that invests a lot in healthcare compared to countries with less strength in economy.

After the surge of mid-2021 the gap between deaths in greater healthcare capacity and limited healthcare capacity countries would go on to quiclkly widen towards the wave of late 2021, this can be explained as the world was a year into covid-19 already and the countries with strong economies had invested much into cure and precaution of covid contamination already.

How Does Population Density Affect The Spread Of COVID-19?



Key Observations

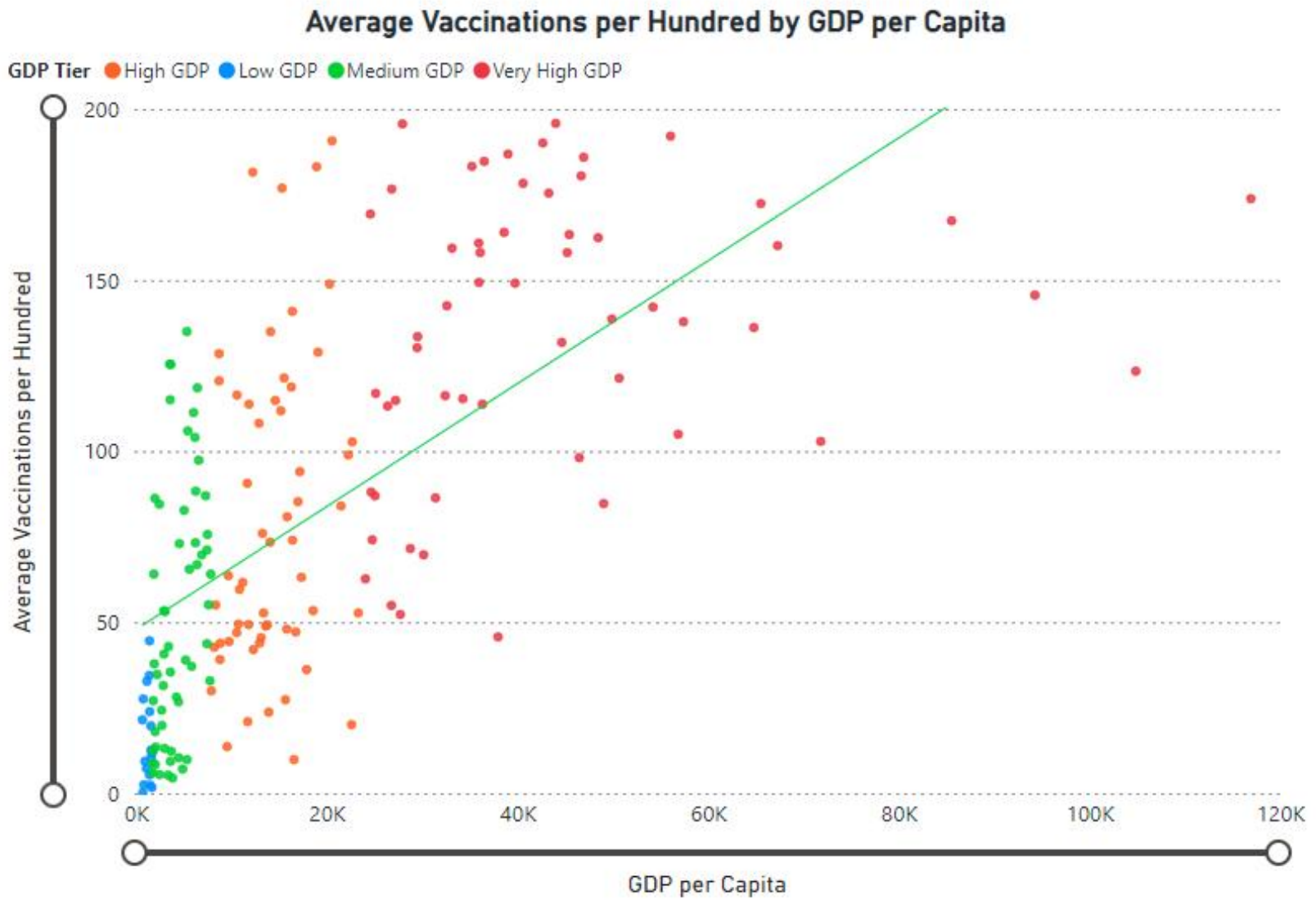
Using a scatter plot, we can derive a correlation between the two variables.

The green line (representing a positive correlation) indicates that as population density rises, the rate of cases also rise, this can be explained simply by the definnition of population density as it refers to the amount of people located in a specified space.

The more people we have in a given amount of space we the more the likelyhood of contamination which can be described by the number of cases per million.

outliers on both axis have been modertately exempted so as to reduce skewing of the visual

What Impact Does GDP per Capita Have On COVID-19 Vaccination Rates?



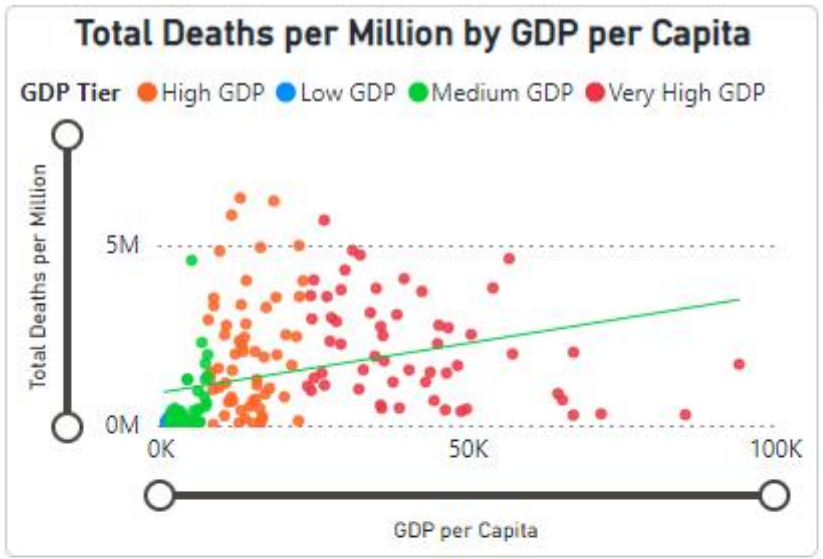
Key Observations

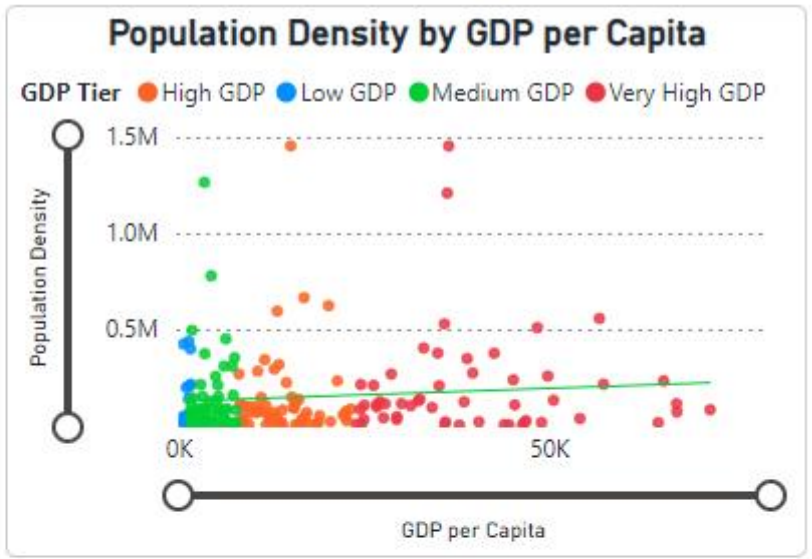
Using a scatter plot again, we can derive a correlation between the two variables.

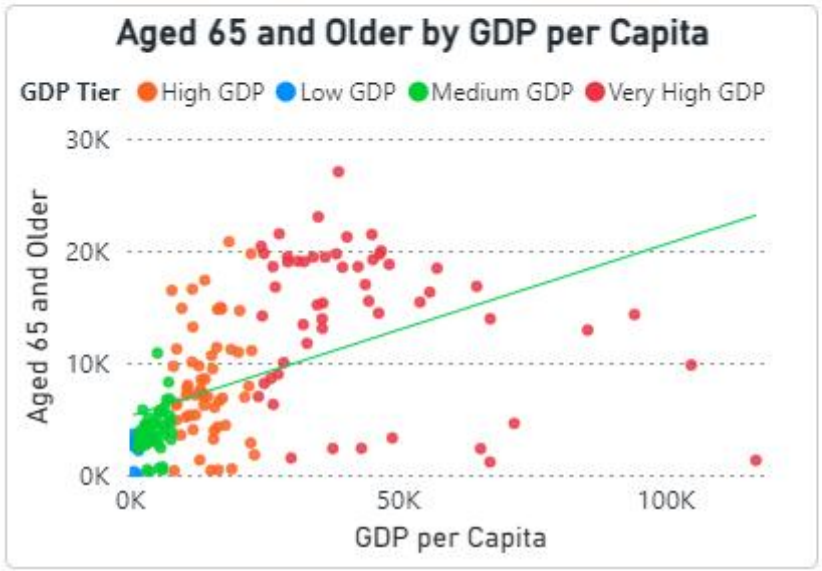
Also here, the green line (representing a positive correlation) indicates that as GDP per capita rises, the average rate of vaccination also rise, this can be explained by the fact that increase in GDP per capita describes a stronger economy and potentially a richer country showing that such coiuntries must have invested more into obtaining and distributing covid-19 vaccines and earlier too.

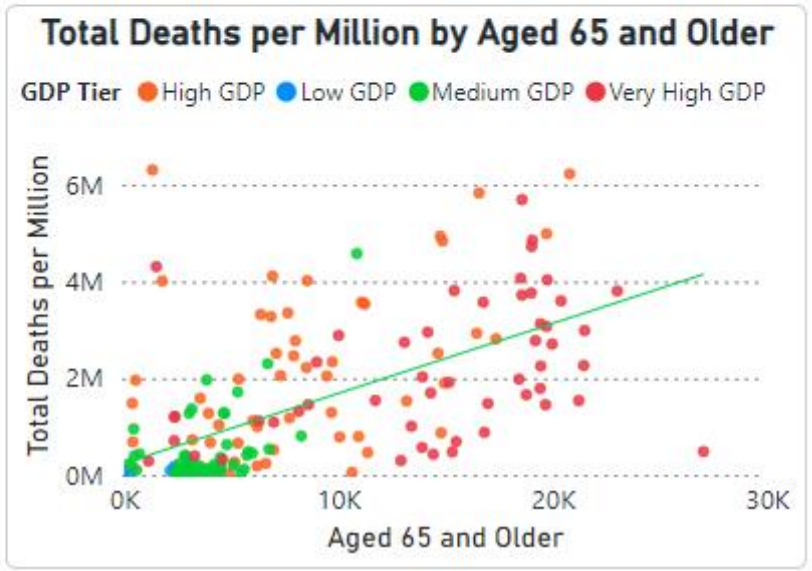
outliers on both axis have been modertately exempted so as to reduce skewing of the visual

What Impact Does GDP per Capita Have On COVID-19 Mortality Rates?









Key Observations

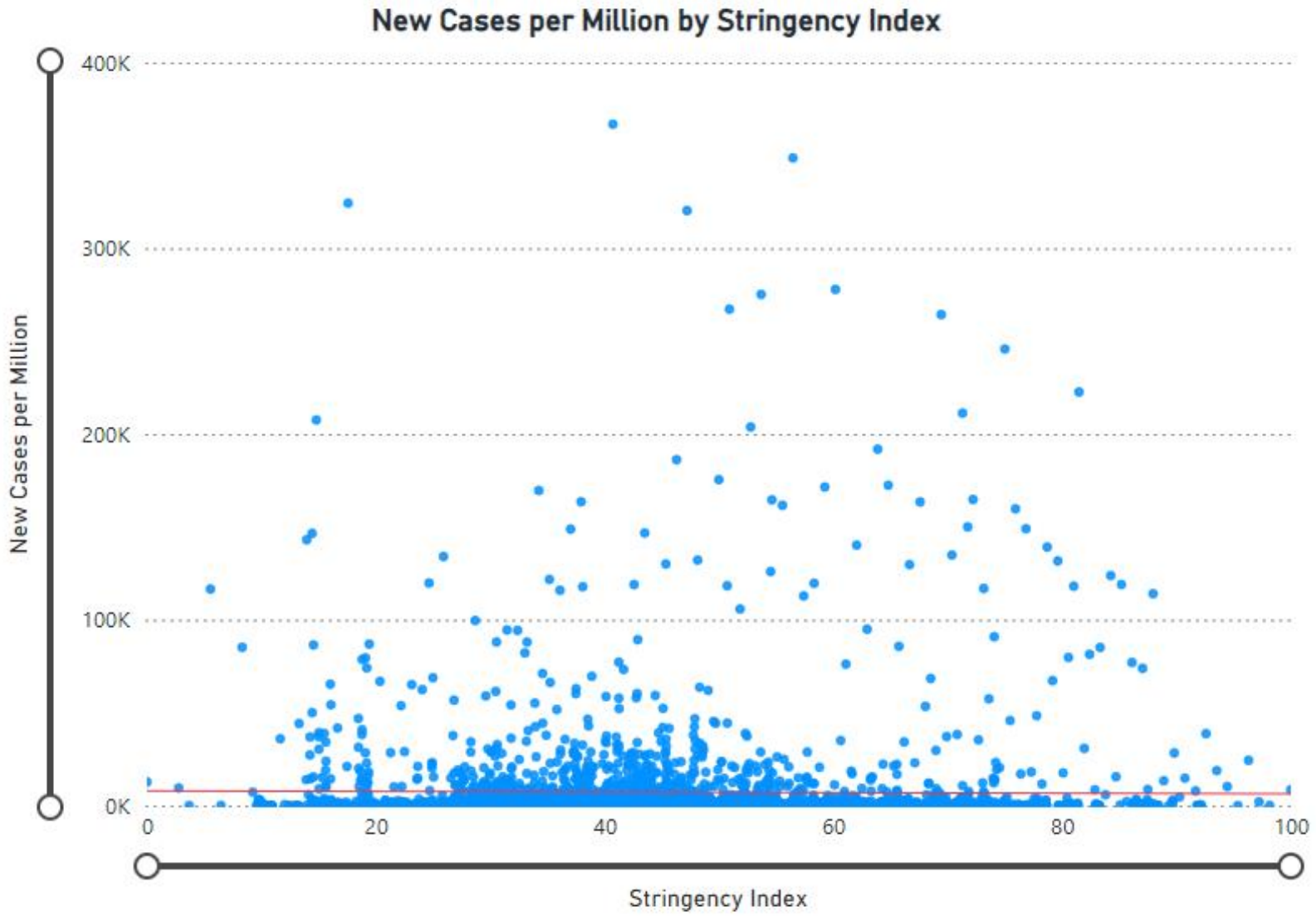
The scatter plot at the top left shows a positive correlation between GDP per Capita and deaths per million, while this might seem counter-intuitive as we already established that countries with higher GDP per Capita (potentaillly rich countries) have much invested into counter measures against COVID-19, the positive correlation can be explained with some key demographic about countries with higher GDP per Capita and a fact about the COVID-19 virus.

Firstly, countries with higher GDP per Capita tend to have higher population density (demonstrated by the top right scatter plot). High population density increases the likelyhocod of COVID-19 contamination/spread but this does'nt fully describe the increase in death rates.

Countries with Higher GDP per Capita also have more old age population (bottom left scatter plot) which is the age group where COVID-19 is the most lethal (bottom right scatter plot).

outliers on both axis have been modertately exempted so as to reduce skewing of the visuals

How Does The Stringency Index Correlate With The Reduction In New Cases Over Time?



Key Observations

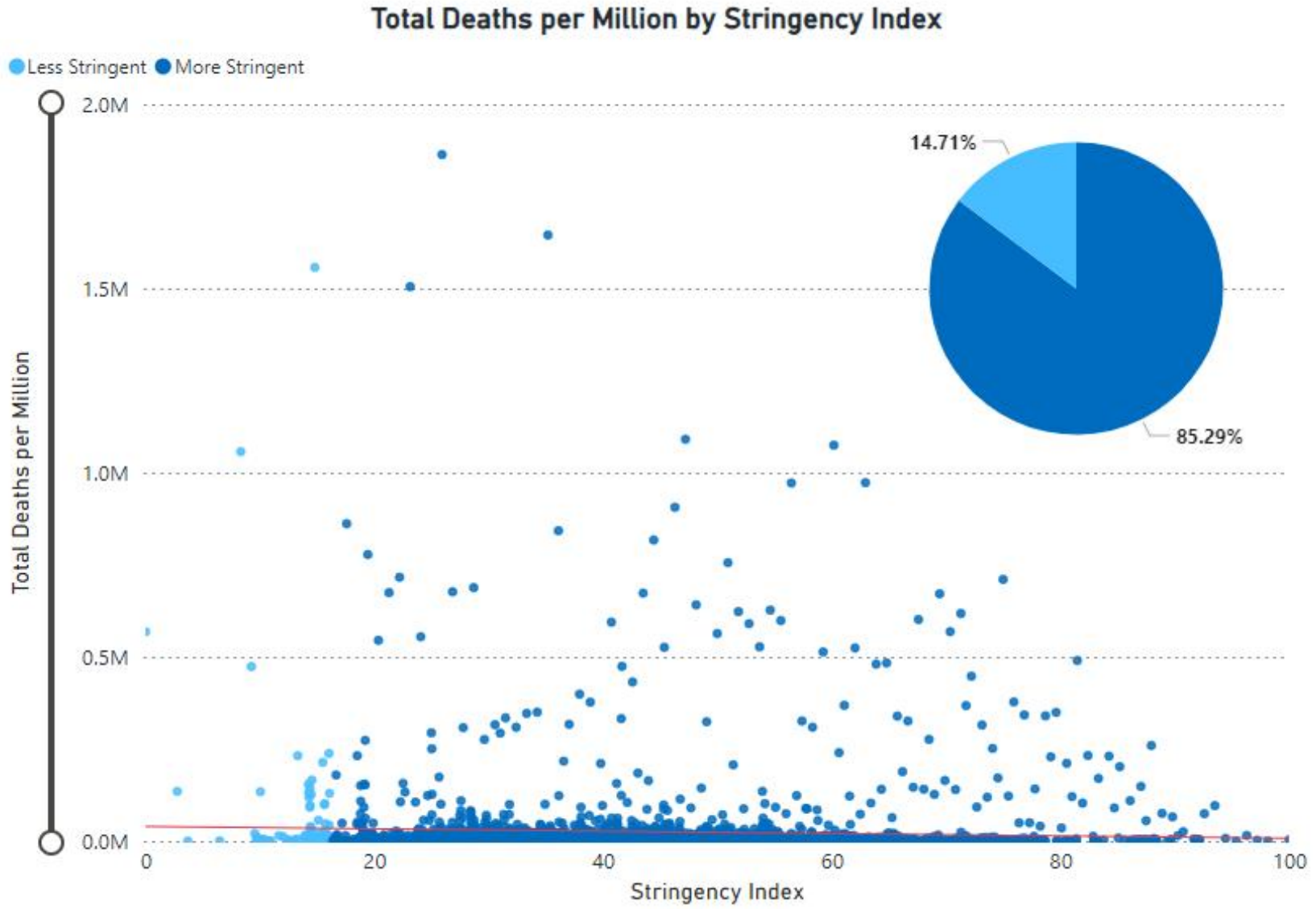
According to the scatter plot, increase in stringency index suggests a decrease in the rate of COVID-19 cases, this is determined with the trend line (red) signaling a negative correlation of the two variables.

It is worth noting that the correlation between both variables is quite minimal as the trend line isn't much steeped downwards. This lack of a strong correlation is in line with what was showcased in previous slides which is that countries with higher GDP per Capita (potentialy rich countries with densly populated areas) tend to have a higher chace of COVID-19 contamination/spread which deminishes the country's efforts to counter the spread of COVID-19.

It is also worth noting that countries with higher GDP per Capita tend to have more stringency index.

outliers on both axis have been modertately exempted so as to reduce skewing of the visual

Did Countries With More Stringent Government Responses Have Fewer Deaths Than Those With Less Stringent Measures?



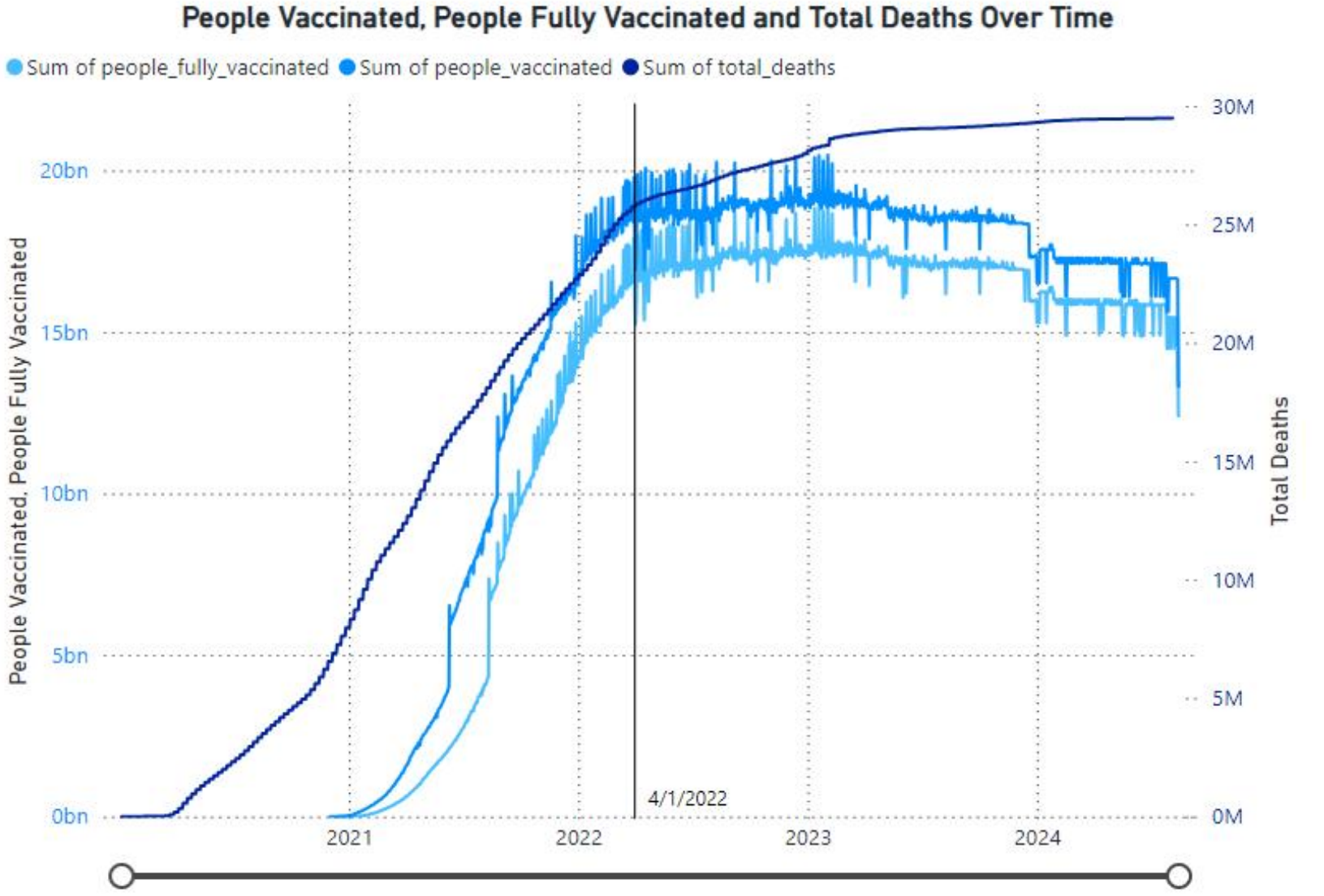
Key Observations

Just like in the previous slide, there is a negative correlation between both variables and the correlation is also weak here but nonetheless, there is still the fact that an increases in government stringency leads to a decrease in the rate of COVID-19 deaths as demonstrated by the scatter plot.

Most probably due to the effectiveness of the COVID-19 virus and the catastrophies experienced during the early days of COVID-19 going world wide, most governments around the world opted for stringent lockdown policies so as to reduce COVID-19 related casualties, this is described with the pie chart.

outliers on both axis have been modertately exempted so as to reduce skewing of the visual

What Is The Relationship Between The Number Of People Fully Vaccinated and The Decline In COVID-19 Related Deaths?

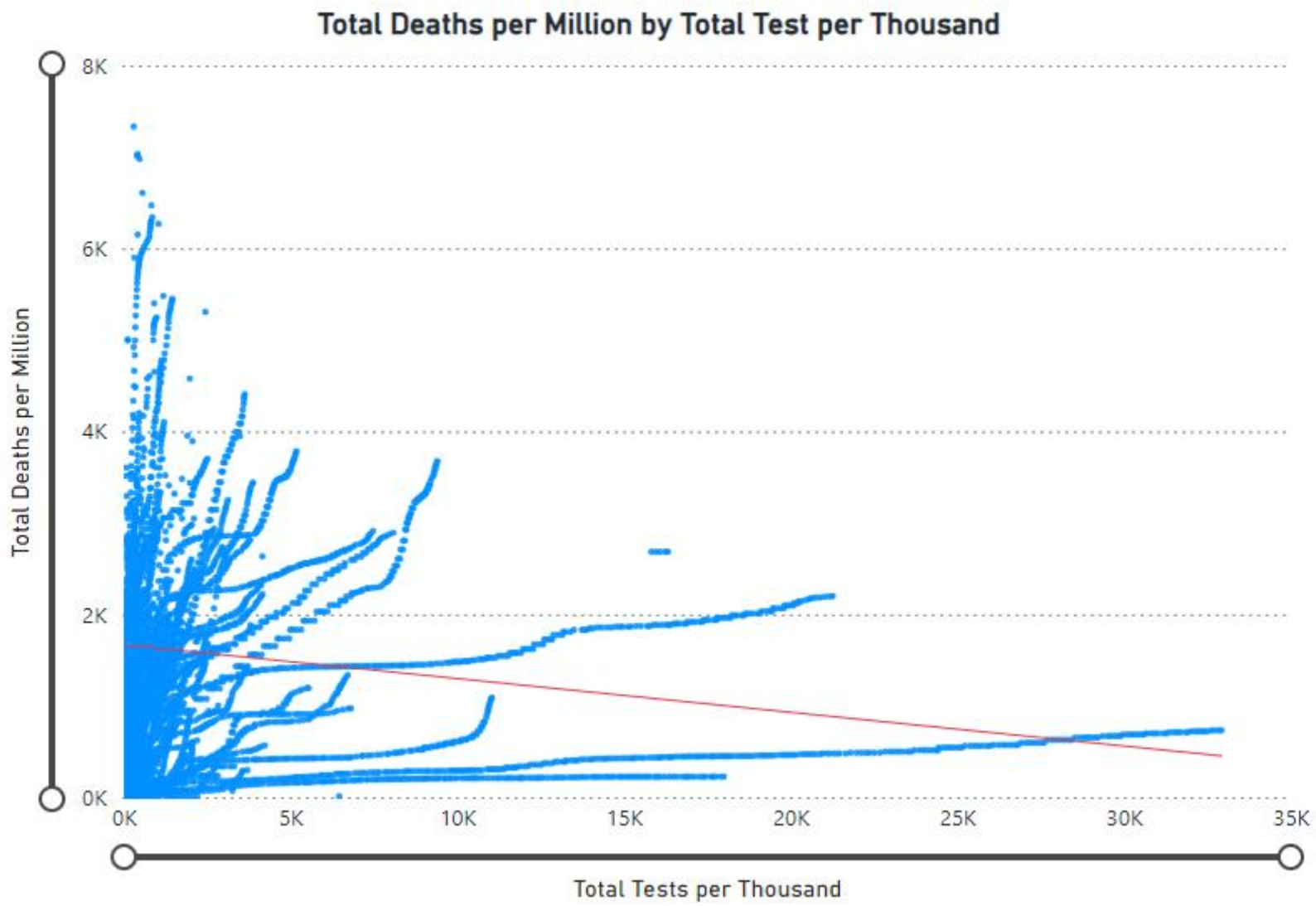


Key Observations

According to the chart, the rate of COVID-19 deaths rose rather quickly and it took untiil late 2020 before the use (recorded) of COVID-19 vaccines.

It is clearly shown that the world adopted the COVID-19 vaccines very quickly as the rise in people vaccinated and people fully vaccinated is quite steep, so much so that people vaccinated and people fully vaccinated reached its peak in just about a year (as indicated by the reference line) which is relatively the same time when the trend of total deaths noticeably starts to flatline.

Do countries with higher tests rates have lower COVID-19 death rates?



Key Observations

As demonstrated by the scatter plot to the left we can see that there is a negative correlation between both variables explaining that as the rate of tests increase, the rate of deaths decreases.

This can be explained with the fact that early awareness to an ilness prompts earlier actions to combat potential negatives of the illness.

It is worth noting that majority of the data points are located near the origin point along the x axis while they are more distributed along the y axis as the y axis describes the distribution of death rate, it makes sense because COVID-19 deaths saw a rather steep rise as established in the previous slide.

outliers on both axis have been modertately exempted so as to reduce skewing of the visual

Conclusion & Recommendations

Conclusion:

The analysis shows distinct trends in COVID-19 cases, deaths, and vaccination impacts across various demographics and government response levels. Countries with higher healthcare capacity and vaccination rates generally saw lower mortality rates, while stricter lockdown measures correlated with slower case growth. Economic and demographic factors, like GDP per capita and population density, influenced COVID-19 outcomes, underscoring the need for tailored health responses.

Recommendations:

1. Targeted Health Resources: Increase healthcare capacity in regions with limited hospital beds to improve pandemic resilience.
2. Vaccination Programs: Prioritize vaccination drives in low-GDP areas to reduce mortality and improve immunity rates.
3. Responsive Policy Design: Implement adaptable, data-driven lockdown measures that align with local COVID-19 spread patterns and healthcare capacity.