Global COVID-19 Data Analysis: Insights & Strategy

Introduction

"I am here today to delve into a comprehensive analysis of global COVID-19 data, drawing insights from extensive datasets that track the pandemic's progression, vaccination efforts, and various demographic impacts.

My objective is to transform this vast amount of information into clear, actionable intelligence that can inform our strategic planning and help us better understand the critical factors shaping public health and economic landscapes worldwide. Let's explore the trends and patterns that emerge from this data. This dataset was obtained randomly from Kaggle."

By Austine Obasuyi

Theme: Key Questions to Answer and Analysis Technique

Trend Analysis:
How do cases/deaths vary over time by country/region. Are there rising or declining trends. Time-series analysis (line plots by date)
Hotspot Identification:
Which countries/regions report the highest case or death counts? Sorting & ranking, heat maps:
Testing & Detection:
What is the testing rate per population? Positivity rate? Ratio calculations (tests/cases), bar plots:
Vaccination Progress:
What is the vaccination rate globally/regionally? Cumulative % vaccinated, maps:
Mortality & Recovery:
What is the case fatality rate and recovery rate? Derived metrics: deaths/cases, recovered/cases:
Comparative Insights:
How do metrics vary by region, income level, or healthcare capacity? Cross-sectional analysis with filters/aggregates

About Dataset

Daily updated dataset of all Coronavirus (COVID-19) Cases in all countries in the world

Confirmed cases and deaths: this data is collected from the World Health Organization Coronavirus Dashboard. The cases & deaths dataset is updated daily.

Note 1: Time/date stamps reflect when the data was last updated by WHO. Due to the time required to process and validate the incoming data.

Note 2: Counts and corrections made after these times will be carried forward to the next reporting cycle for that specific region. Delayed reporting for any specific country, territory or area may result in pooled counts for multiple days being presented, with a retrospective update to counts on previous days to accurately reflect trends. Significant data errors detected or reported to WHO may be corrected at more frequent intervals.

Hospitalizations and intensive care unit (ICU) admissions: data is collected from official sources and collated by World in Data.

Testing for COVID-19: this data is collected by the Our World in Data team from official reports; you can find further details in the post on COVID-19 testing, including checklist of questions to understand testing data, information on geographical and temporal coverage, and detailed country-by-country source information. On 23 June 2022 on kaggle.

Vaccinations against COVID-19: this data is collected by the Our World in Data team from official reports.

Other variables: this data is collected from a variety of sources (United Nations, World Bank, Global Burden of Disease, Blavatnik School of Government, etc.). More information is available in our codebook.

Analysis Plan

Theme	Key Questions to Answer	Analysis Technique
Trend Analysis	How do cases/deaths vary over time by country/region? Are there rising or declining trends?	Time-series analysis (line plots by date)
Hotspot Identification	Which countries/regions report the highest case or death counts?	Sorting & ranking, heat maps
Testing & Detection	What is the testing rate per population? Positivity rate?	Ratio calculations (tests/cases), bar plots
Vaccination Progress	What is the vaccination rate globally/regionally?	Cumulative % vaccinated, maps
Mortality & Recovery	What is the case fatality rate and recovery rate?	Derived metrics: deaths/cases, recovered/cases
Comparative Insights	How do metrics vary by region, income level, or healthcare capacity?	Cross-sectional analysis with filters/aggregates

Trend Analysis How cases/deaths vary over time by region

Kpi

Continent	70 Years Older	Aged 65 Older	Median Age	Cardiovasc Death Rate	Fatality Rate	Total Tests	Total Cases	Total Deaths
Africa	212,817	331,753	1,999,287	25,970,587	10,599,771	22,213,357,604	14,616,886,098	306,955,815
Asia	338,205	530,271	2,536,786	24,101,501	17,868,743	576,014,519,305	252,167,317,226	1,790,404,880
Europe	768,870	1,187,098	2,818,274	15,804,485	12,711,177	609,317,604,566	236,756,684,151	2,361,106,039
North America	281,651	416,201	1,535,677	8,577,785	10,653,200	381,789,868,848	127,073,670,231	1,880,636,077
Oceania	98,315	161,096	601,064	8,420,251	16,386,100	24,123,978,485	11,598,053,992	23,933,802
South America	109,074	166,323	649,472	3,763,732	1,334,510	61,969,482,409	73,484,570,403	1,646,924,372

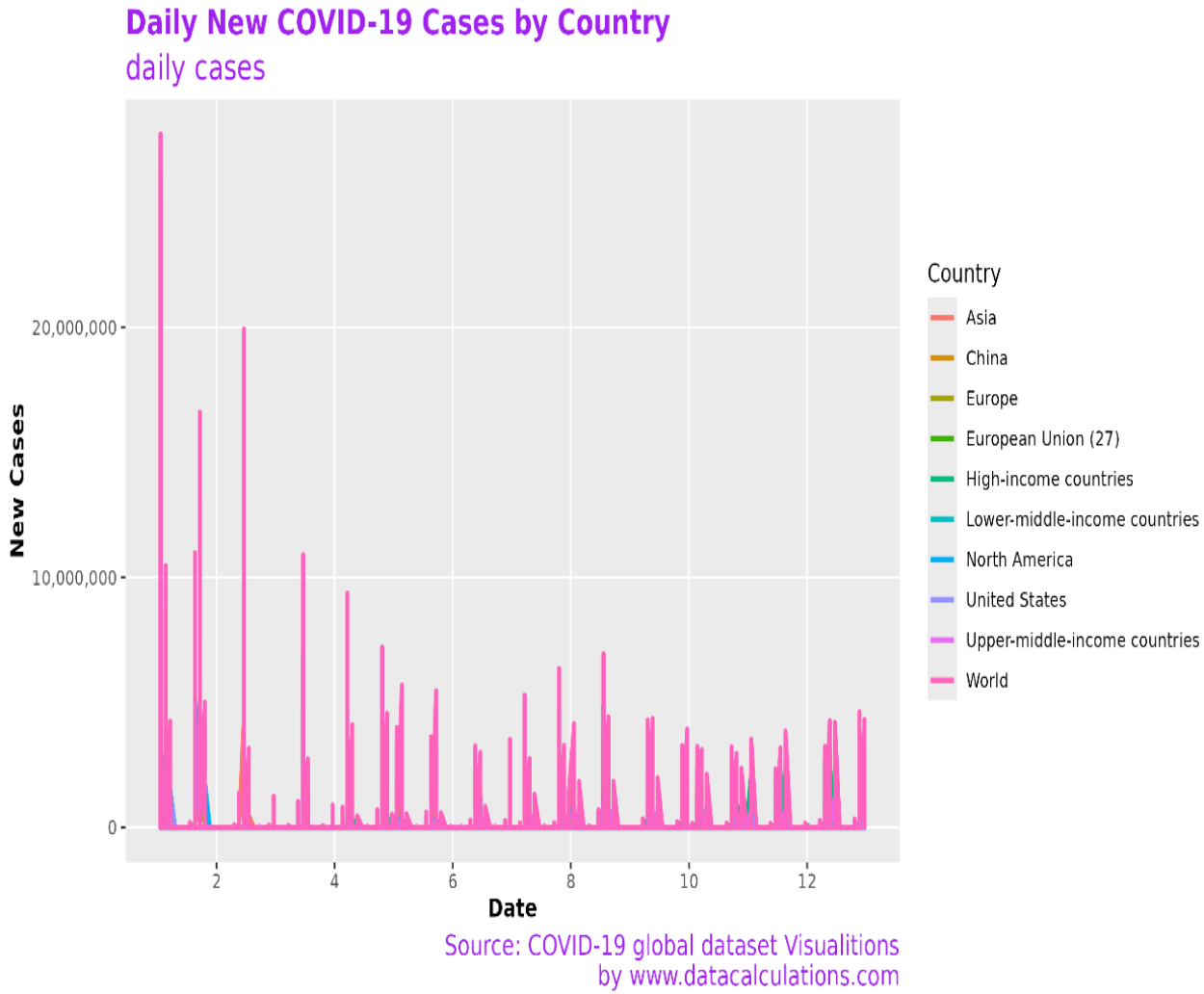
Measure Names
- 70 Years Older
- Aged 65 Older
- Median Age
- Cardiovasc Death Rate
- Fatality Rate
- Total Tests
- Total Cases
- Total Deaths

Daily New COVID-19 Cases by Country

This line plot illustrates the daily fluctuations in new COVID-19 cases over time, comparing the trends across several selected countries. It effectively visualizes the various waves and peaks of the pandemic for each nation, highlighting their individual trajectories of case surges and declines.

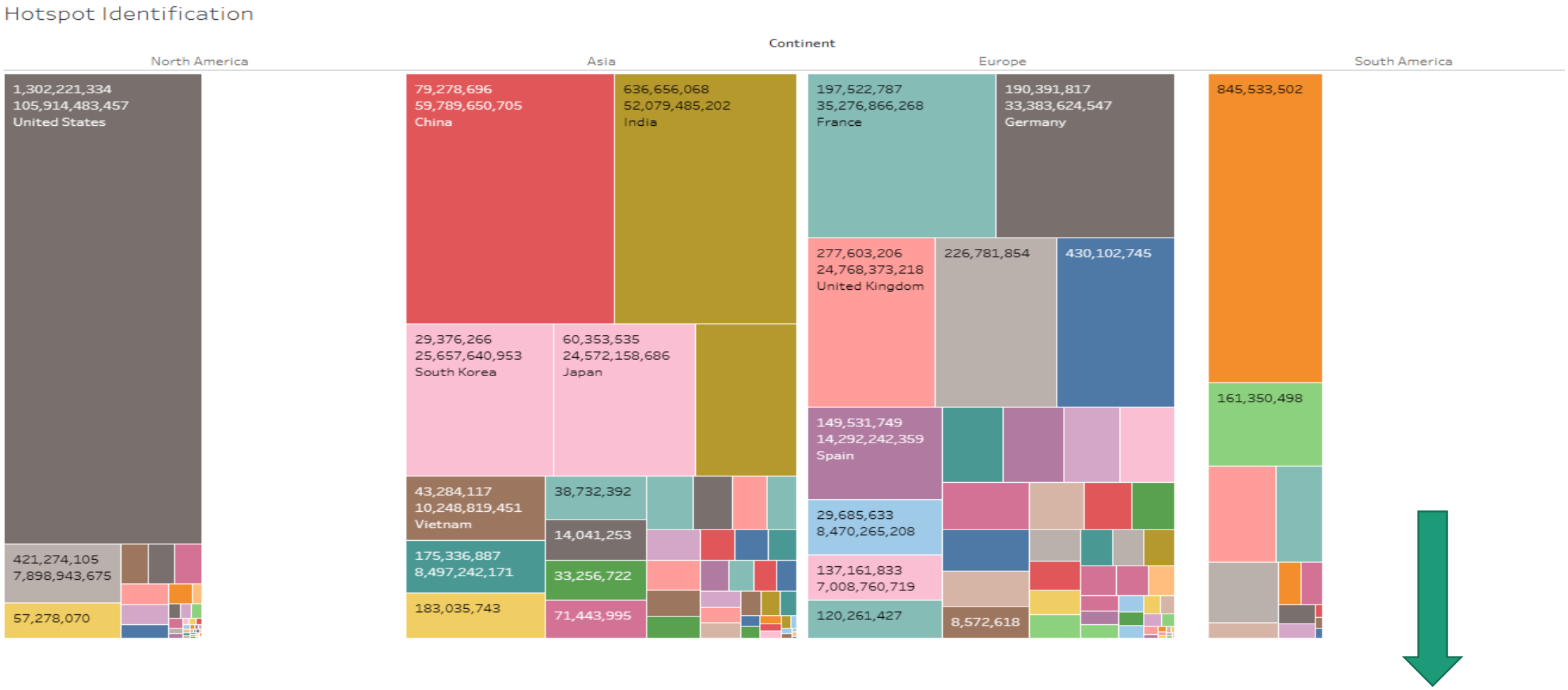
Figure 2 : Daily New COVID-19 Cases by Country.

This chart tracks the daily reported new COVID-19 cases for a selection of countries over time, revealing the intensity and timing of their respective pandemic waves.



HOTSPOT IDENTIFICATION.

Countries/regions with the highest case or death report, heat maps



Hotspot Identification (Tree map)

This treemap visualization, titled "Hotspot Identification," effectively displays the distribution of a large numeric metric (likely population or a significant count) across various countries, categorized by continent. The size of each rectangular segment is proportional to the metric's value, allowing for quick identification of the largest "hotspots" within each continent and globally. It provides a hierarchical view, starting with continents and then breaking down into individual countries, showcasing their relative contributions to the overall metric.

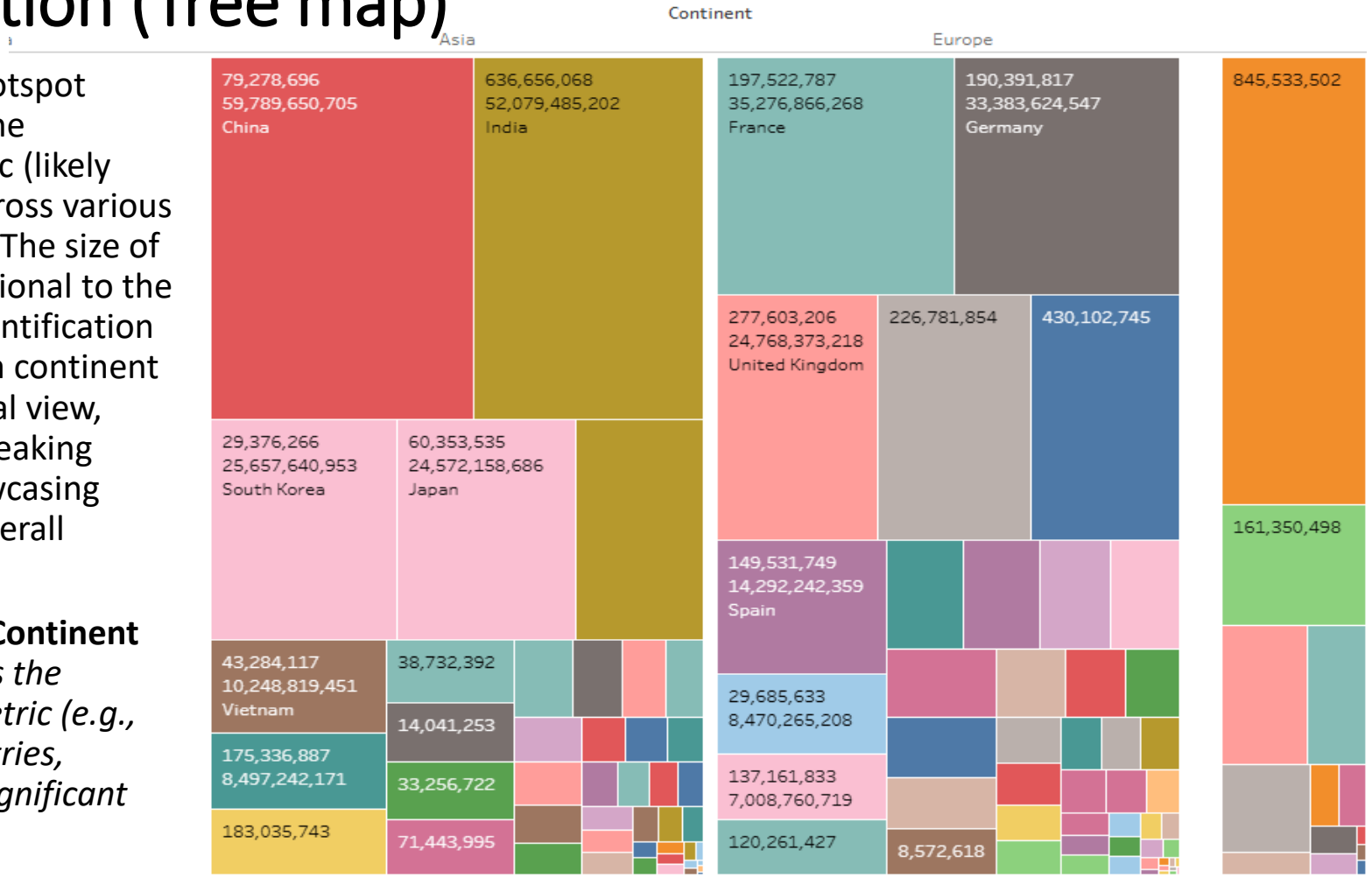
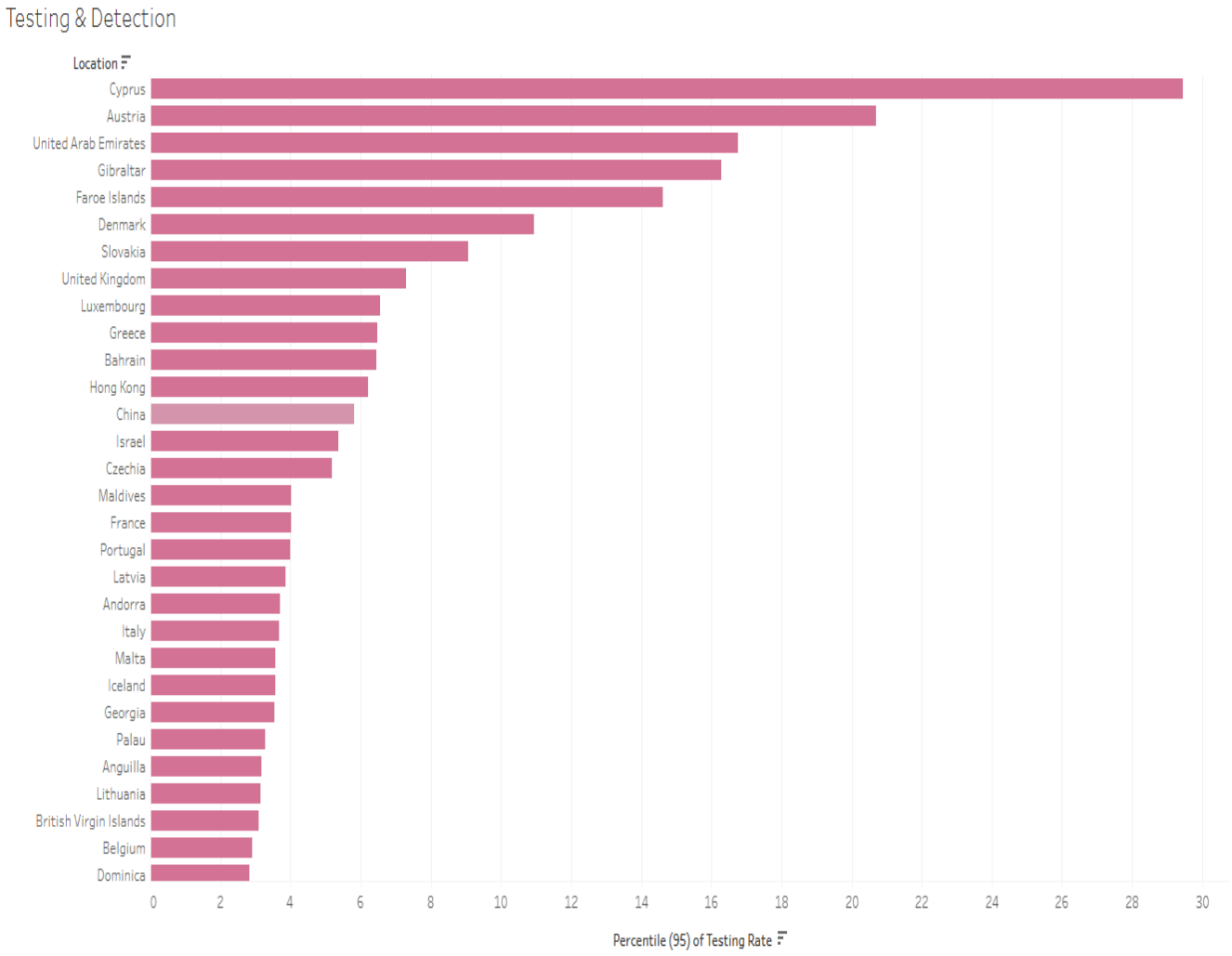


Figure 3: Hotspot Identification by Continent and Country *This treemap illustrates the proportional distribution of a key metric (e.g., population, case count) across countries, grouped by continent, to highlight significant areas or "hotspots" at a glance.*

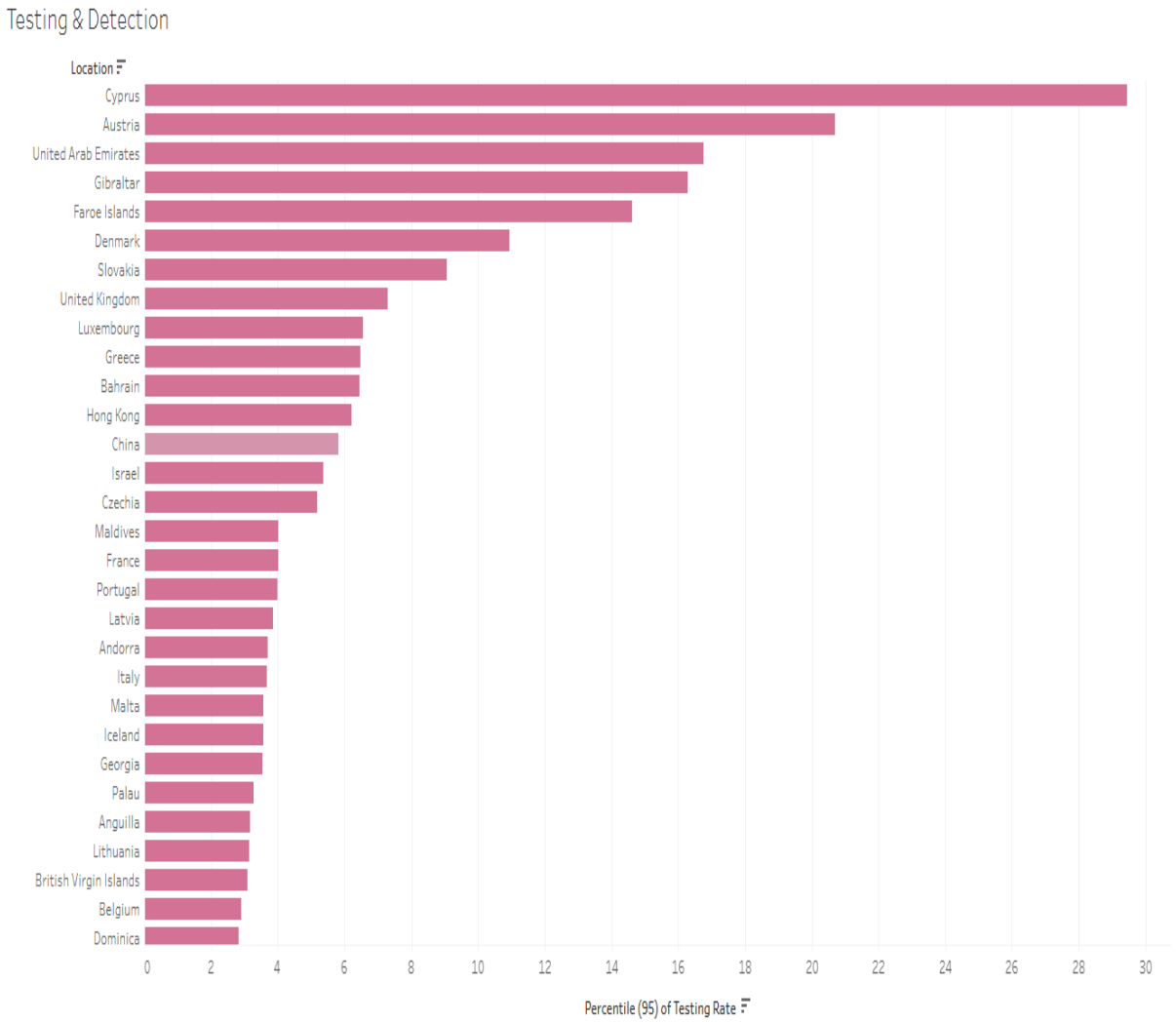
Testing & Detection (Testing Rate)

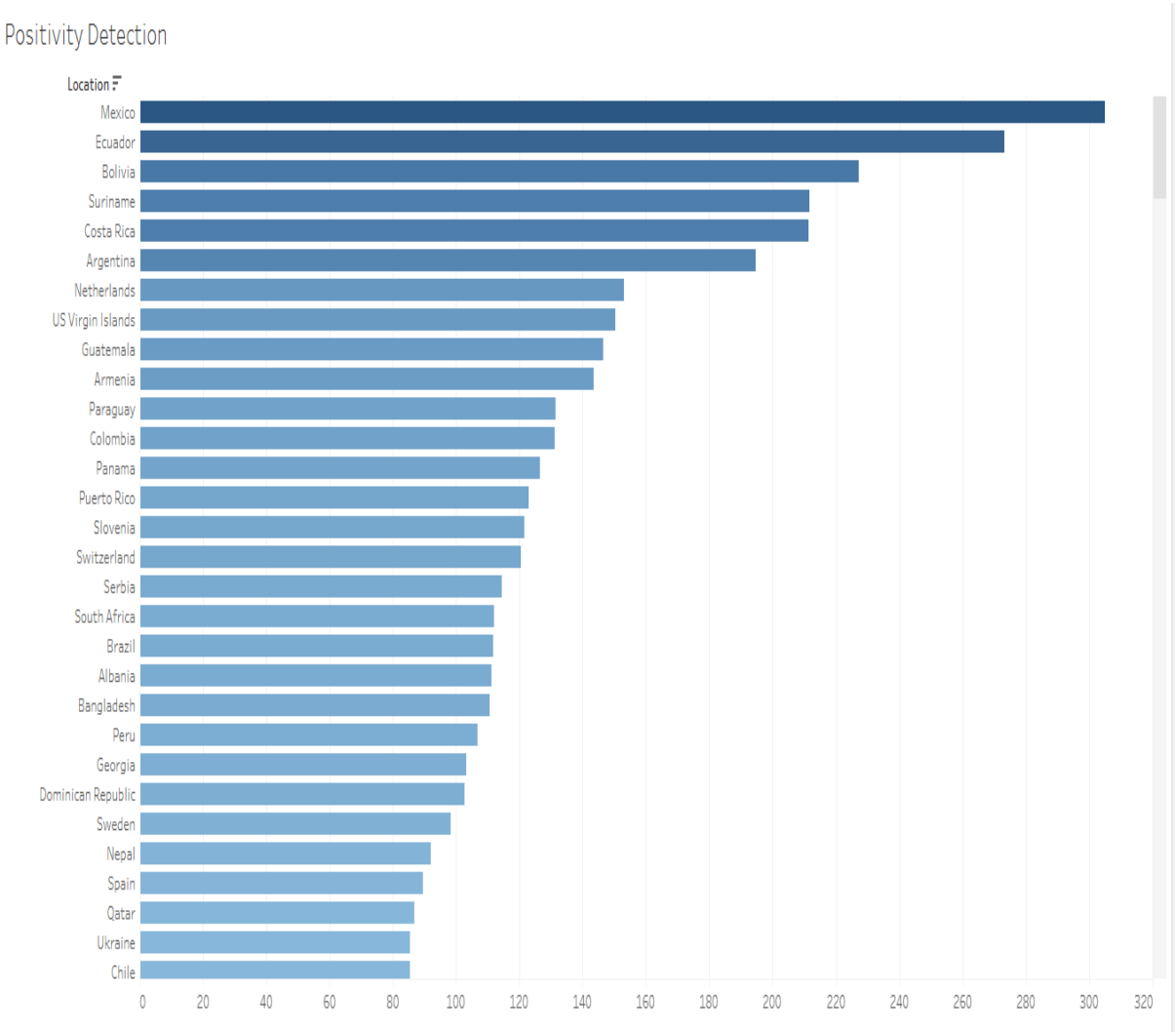
- This horizontal bar chart presents the 95th percentile of the COVID-19 testing rate across various countries, providing a snapshot of their peak testing capacities relative to their populations. The chart allows for a clear comparison of how different locations performed in terms of testing intensity, with longer bars indicating higher testing rates.
- **Figure 4: 95th Percentile of COVID-19 Testing Rate by Country** *This bar chart illustrates the peak testing rate (95th percentile) achieved by various countries, offering insight into their testing efforts during the pandemic.*

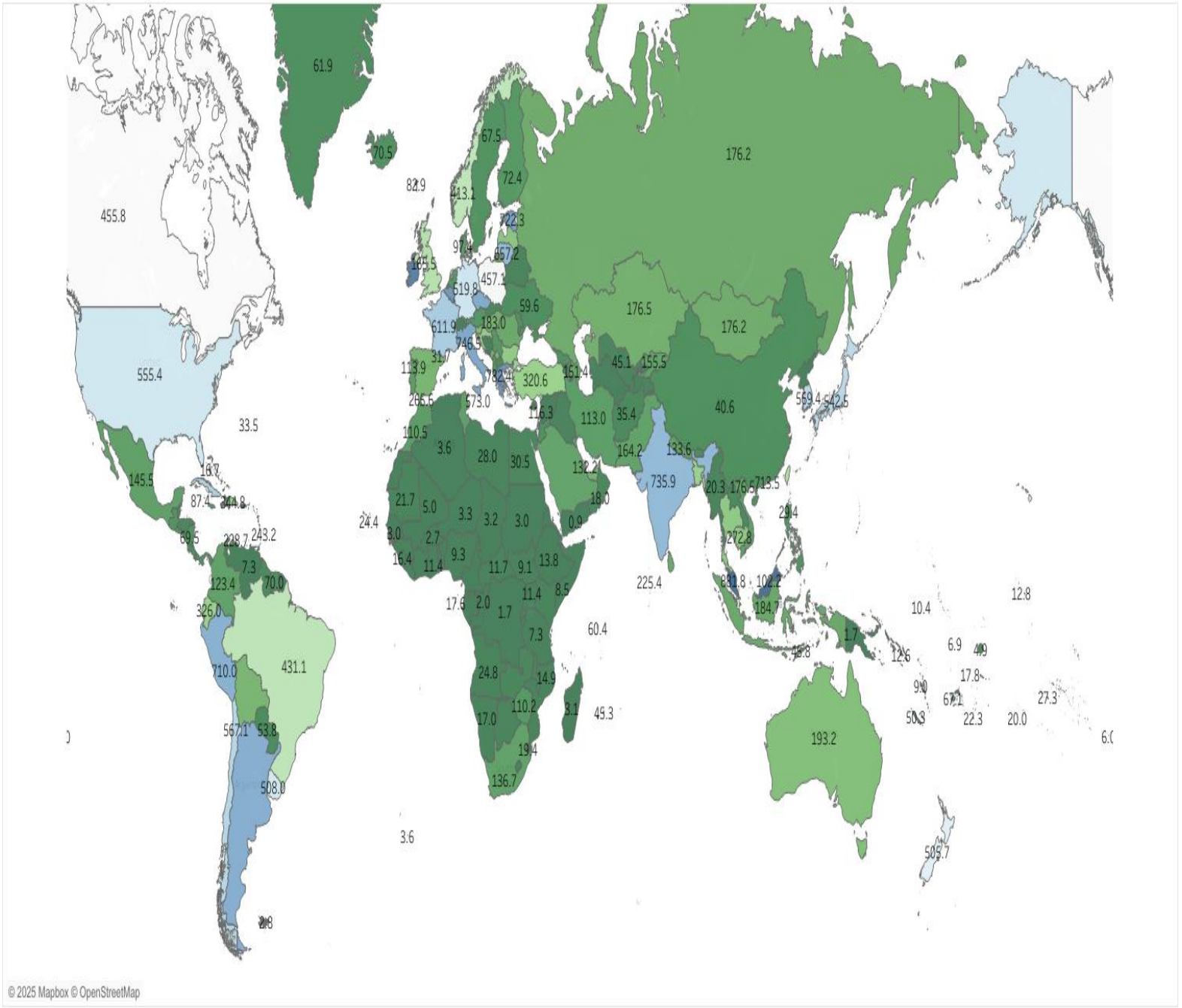


Testing & Detection.

Testing rate per population / Positivity rate & Ratio calculations (bar plots)







Global Vaccination Progress (Choropleth Map)

This choropleth map visually represents the cumulative percentage of the population vaccinated across different countries. The color intensity of each country on the map indicates its vaccination progress, allowing for a quick understanding of regional and global disparities in vaccination rates. Darker shades likely represent higher vaccination percentages, while lighter shades or gray areas may indicate lower rates or unavailable data.

Figure 5: Global Cumulative Percentage Vaccinated *This choropleth map displays the cumulative percentage of the population that has been vaccinated in various countries, illustrating global and regional vaccination progress through color-coded intensity.*

Vaccination Progress.

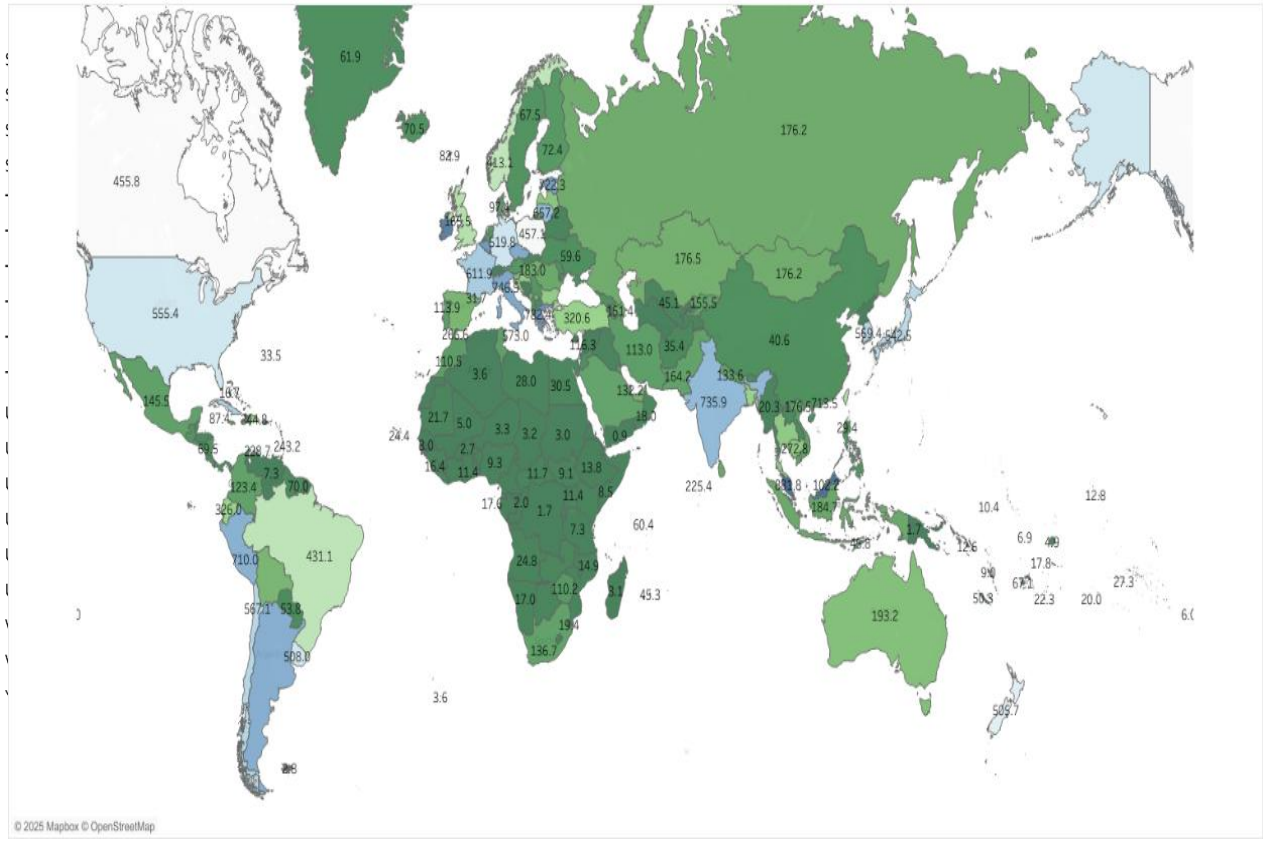
Vaccination rate globally/regionally Cumulative % vaccinated

lew COVID-19 Cases by Country

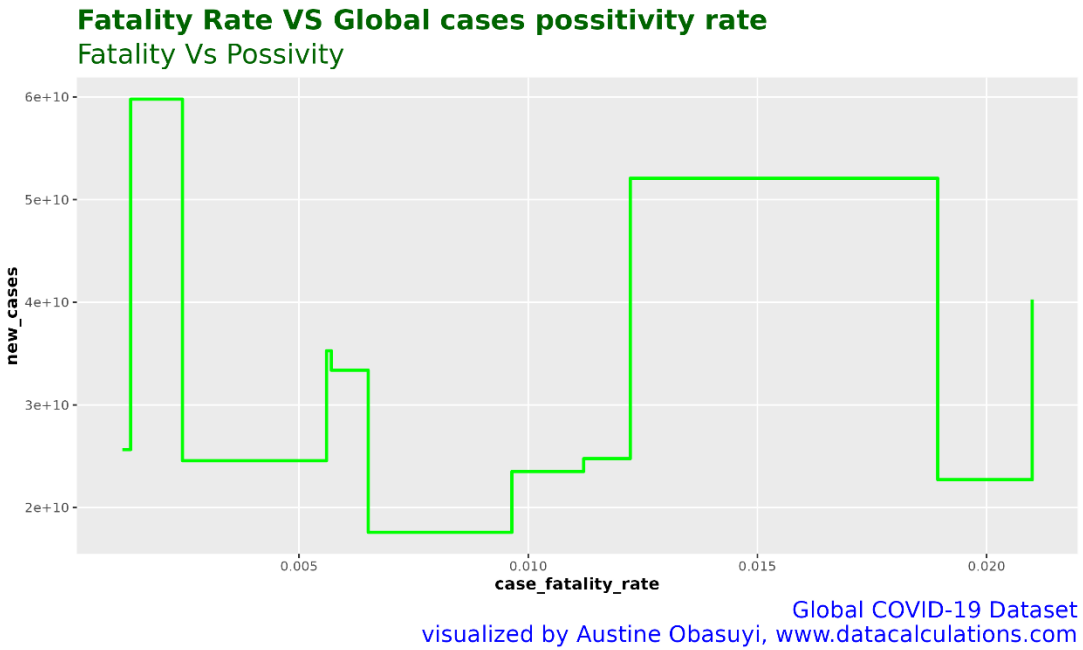
ses

ıanistan	British Virgin Islands	Finland	Isle of Man	Martinique	Portugal
ınia	Brunei	France	Israel	Mexico	Puerto Rico
orra	Bulgaria	Georgia	Italy	Moldova	Qatar
uilla	Cambodia	Germany	Jamaica	Monaco	Romania
gua and Barbuda	Canada	Gibraltar	Japan	Mongolia	Russia
enia	Cayman Islands	Greece	Jersey	Montenegro	Saint Barthelemy
ɔa	China	Greenland	Jordan	Montserrat	Saint Kitts and Nevis
tria	Costa Rica	Grenada	Kazakhstan	Myanmar	Saint Lucia
rbaijan	Croatia	Guadeloupe	Kosovo	Nepal	Saint Martin (French part)
amas	Cuba	Guatemala	Kuwait	Netherlands	Saint Pierre and Miquelon
rain	Curacao	Guernsey	Kyrgyzstan	Nicaragua	Saint Vincent and the Grenadines
gladesh	Cyprus	Haiti	Laos	North Korea	San Marino
ɔados	Czechia	Honduras	Latvia	North Macedonia	Saudi Arabia
ırus	Denmark	Hungary	Lebanon	Norway	Serbia
ium	Dominica	Iceland	Liechtenstein	Oman	Singapore
ze	Dominican Republic	India	Lithuania	Pakistan	Sint Maarten (Dutch part)
nuda	East Timor	Indonesia	Luxembourg	Palestine	Slovakia
tan	El Salvador	Iran	Malaysia	Panama	Slovenia
aire Sint Eustatius and Saba	Estonia	Iraq	Maldives	Philippines	South Korea
nia and Herzegovina	Faroe Islands	Ireland	Malta	Poland	Spain

Choropleth Map (Color-coded by vaccination progress)



Top 10 Countries by Total Deaths



Fatality Rate vs. Global Cases Positivity Rate

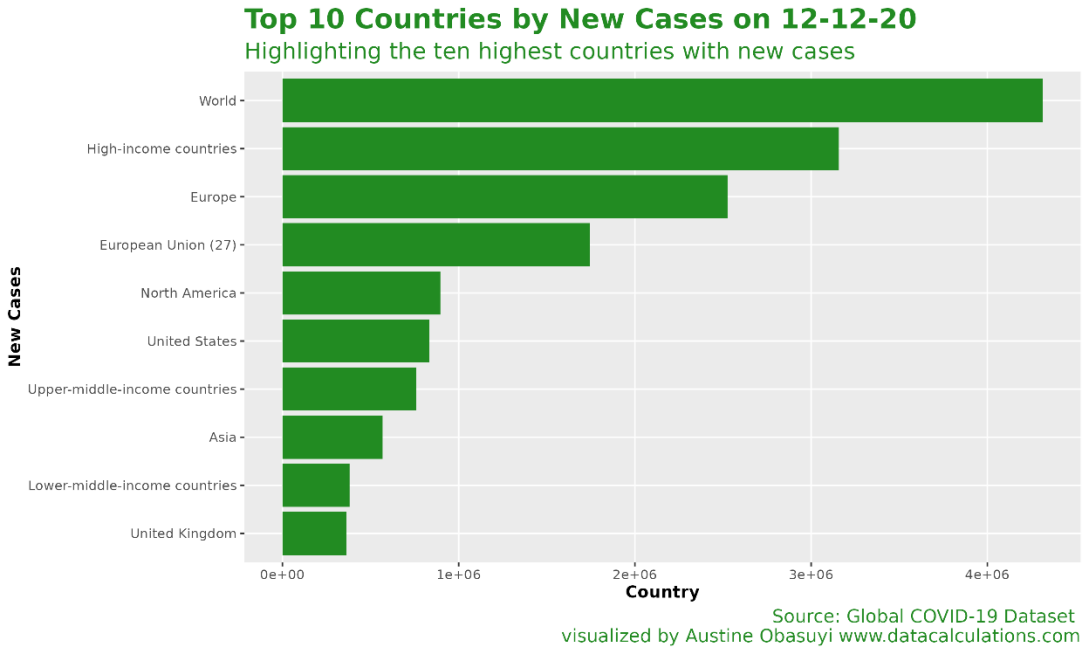
This line plot, titled "Fatality Rate VS Global cases positivity rate," illustrates the relationship between the case fatality rate and new cases, potentially reflecting the global cases positivity rate. The visualization shows how the number of new cases (new_cases) fluctuates across different levels of case_fatality_rate, indicating potential thresholds or ranges where new cases either stabilize or change significantly in relation to the fatality rate.

Figure 6(B): Fatality Rate vs. Global Cases Positivity Rate *This chart explores the interplay between the case fatality rate and the volume of new COVID-19 cases, highlighting how new case counts vary across different fatality rate percentages.*

This bar chart provides a clear comparison of the cumulative COVID-19 fatalities among the ten countries with the highest recorded total deaths. It offers an immediate visual ranking of the nations most impacted in terms of absolute lives lost to the pandemic.

Figure 6(A): Top 10 Countries by Total COVID-19 Deaths.

This bar chart presents the cumulative number of COVID-19 deaths, ranking the top ten most affected countries by total fatalities.



Mortality & Recovery

Mortality & Recovery Rates

This visualization (a line chart) illustrates the dynamics of COVID-19 mortality and recovery. It presents the case fatality rate (deaths per confirmed cases) and the recovery rate (recovered cases per confirmed cases) over time or across different regions. This allows for an understanding of the severity of the disease and the effectiveness of medical interventions, showing how these crucial metrics evolve.

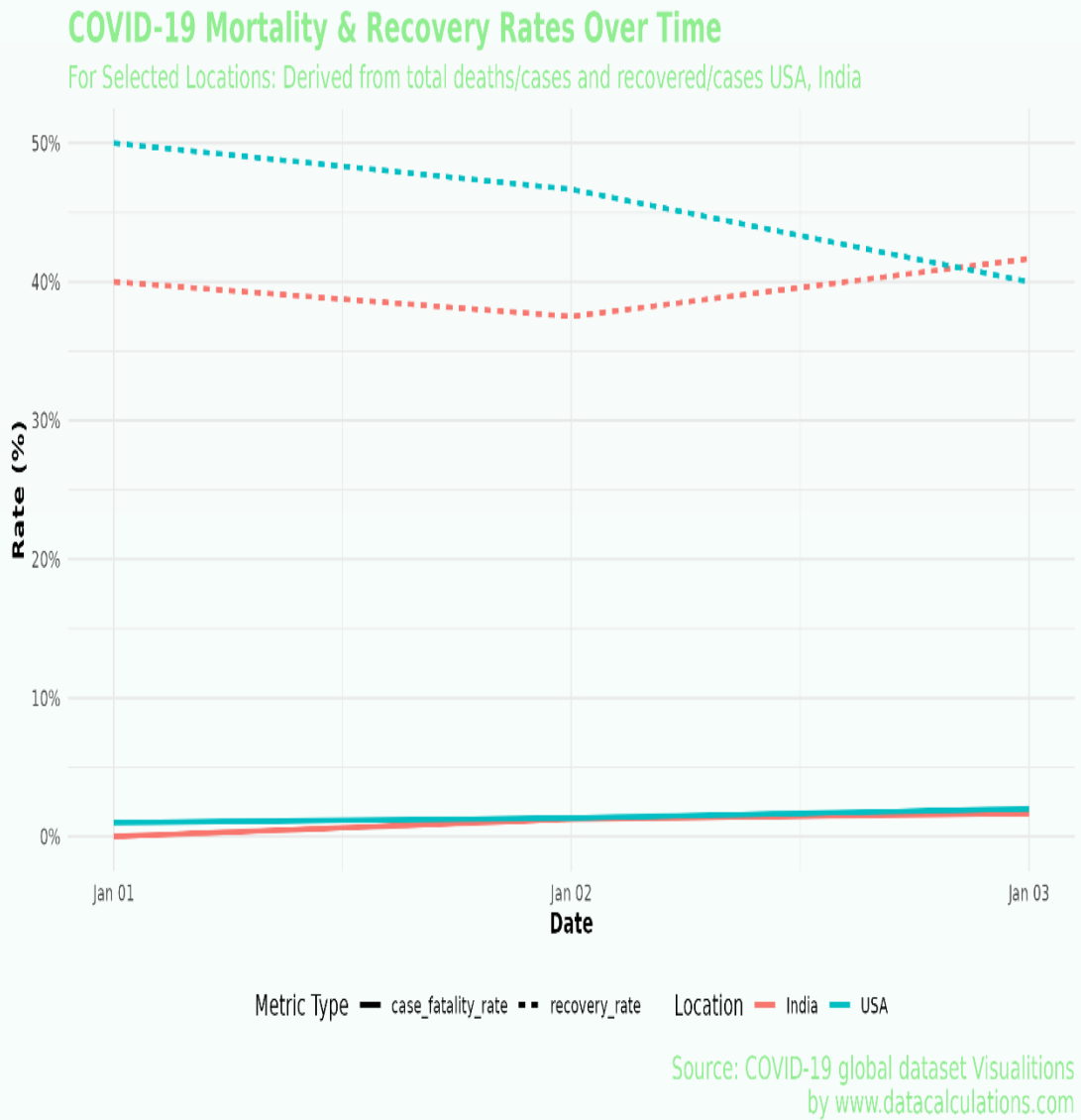


Figure 7: Case Fatality Rate and Recovery Rate
This chart tracks the case fatality rate (deaths per case) and recovery rate (recovered per case), providing insights into the severity and progression of COVID-19 outcomes.

Comparative Insights (Metrics by Region/Income/ Healthcare)

This visualization (a set of grouped bar charts) provides comparative insights into how key COVID-19 metrics (such as case rates, death rates, or vaccination rates) vary when analyzed by different categorical dimensions like geographical region, economic income level, or healthcare capacity. It aims to highlight disparities and correlations, revealing which factors might influence pandemic outcomes across diverse populations and systems.

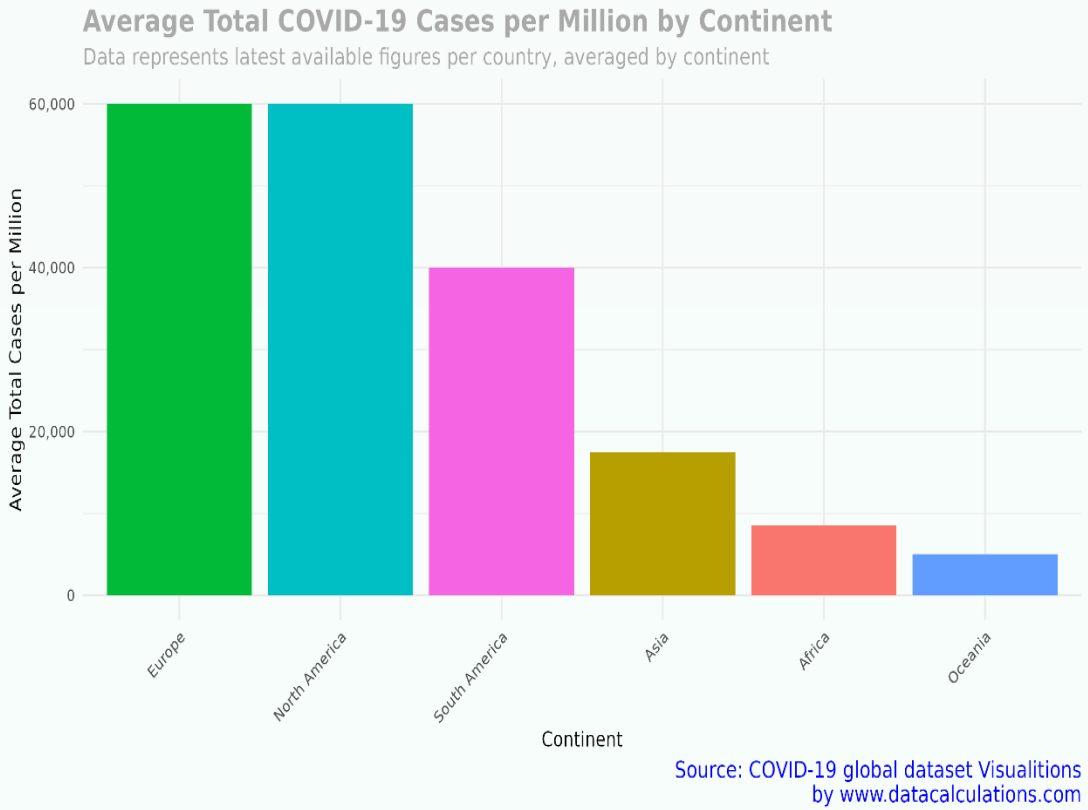
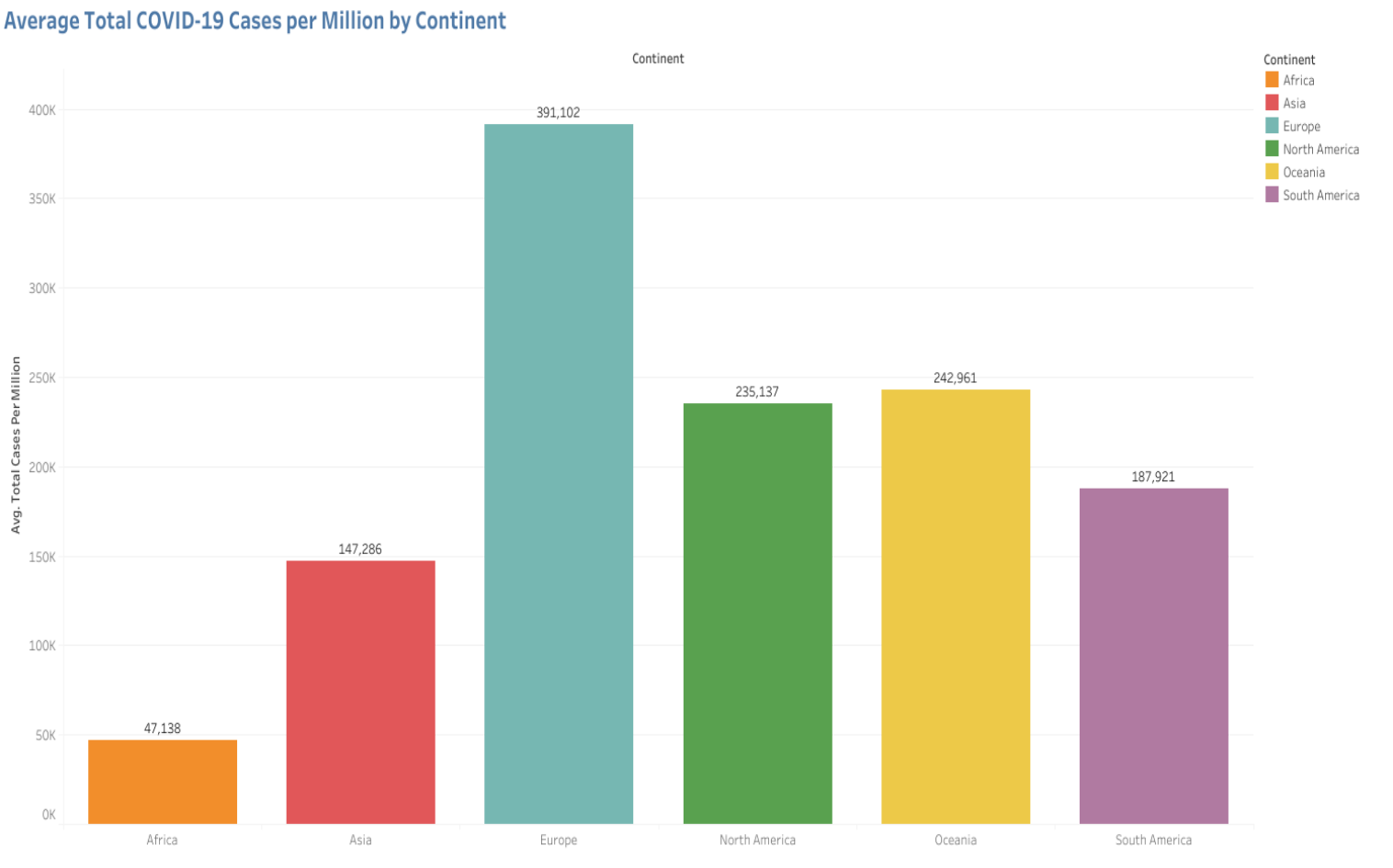


Figure 8: COVID-19 Metric Variations by Region, Income Level, and Healthcare Capacity. *This chart set illustrates the differential impact and progression of COVID-19 metrics across various regions, income levels, and healthcare capacities, revealing key comparative insights into the pandemic's global landscape.*

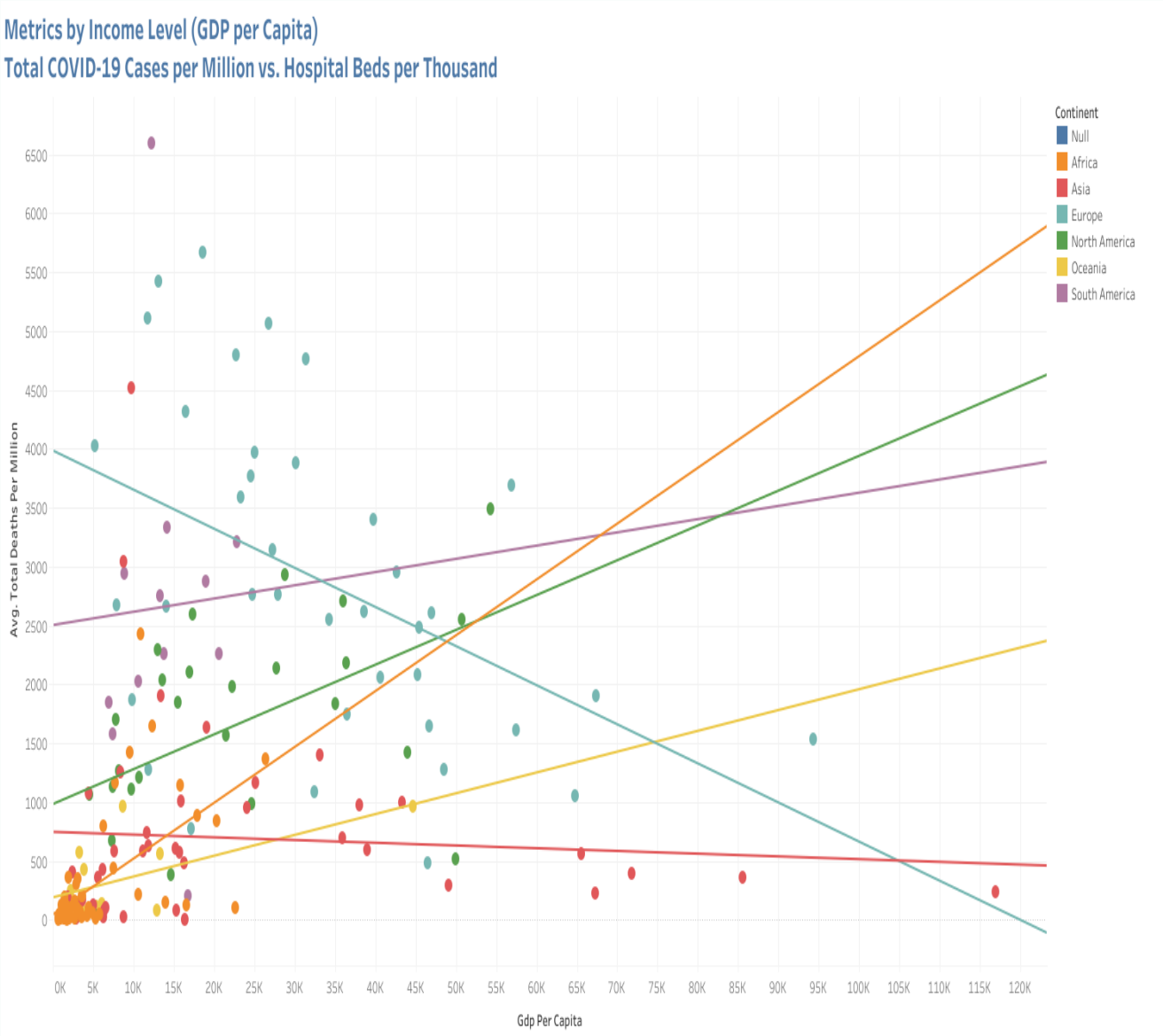
Average Total COVID-19 Cases per Million by Continent



Average Total COVID-19 Cases per Million by Continent

This bar chart illustrates the average total COVID-19 cases per million people across different continents. It provides a clear comparative view of the pandemic's intensity relative to population size in each major geographical region, highlighting which continents have experienced higher average case burdens.

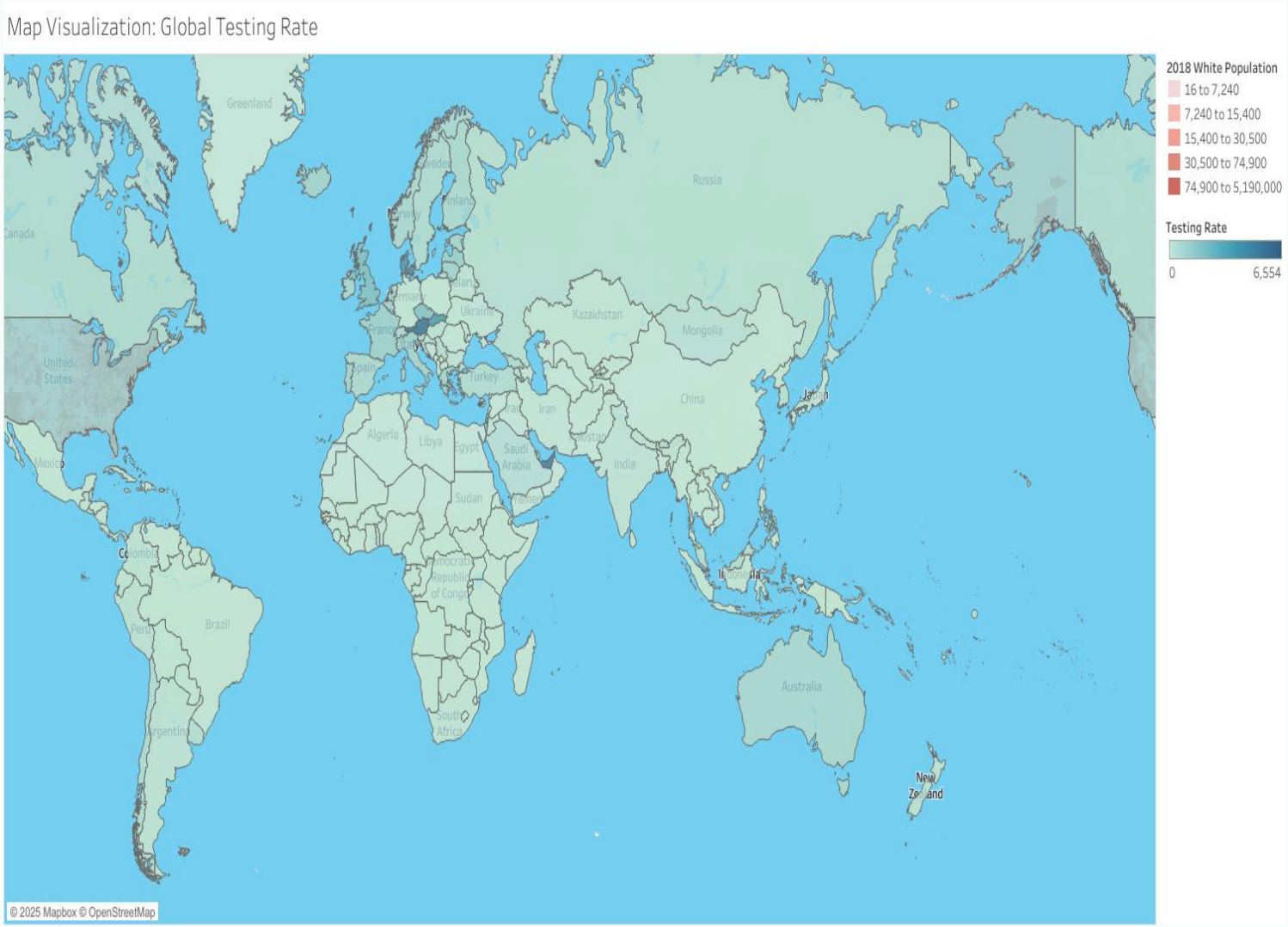
Figure 9: Average Total COVID-19 Cases per Million by Continent *This chart compares the average total COVID-19 cases per million across the world's continents, offering a standardized view of the pandemic's relative impact by region.*



Metrics by Income Level (GDP per Capita)
This scatter plot, titled "Metrics by Income Level (GDP per Capita)," explores the relationship between a country's GDP per capita (as an indicator of income level) and its average total COVID-19 cases per million, also showing hospital beds per thousand. Each point represents a country, colored by continent, and trend lines are included for each continent. This visualization helps to identify potential correlations between economic development, healthcare infrastructure, and the scale of the pandemic's impact on cases, suggesting how these factors might influence observed metrics.

Figure 10: Total COVID-19 Cases per Million vs. GDP Per Capita and Hospital Beds per Thousand *This scatter plot examines the interplay between economic income (GDP per capita), healthcare capacity (hospital beds), and the total COVID-19 cases per million across countries, segmented by continent, with trend lines indicating general relationships.*

Map Visualization: Global Testing Rate



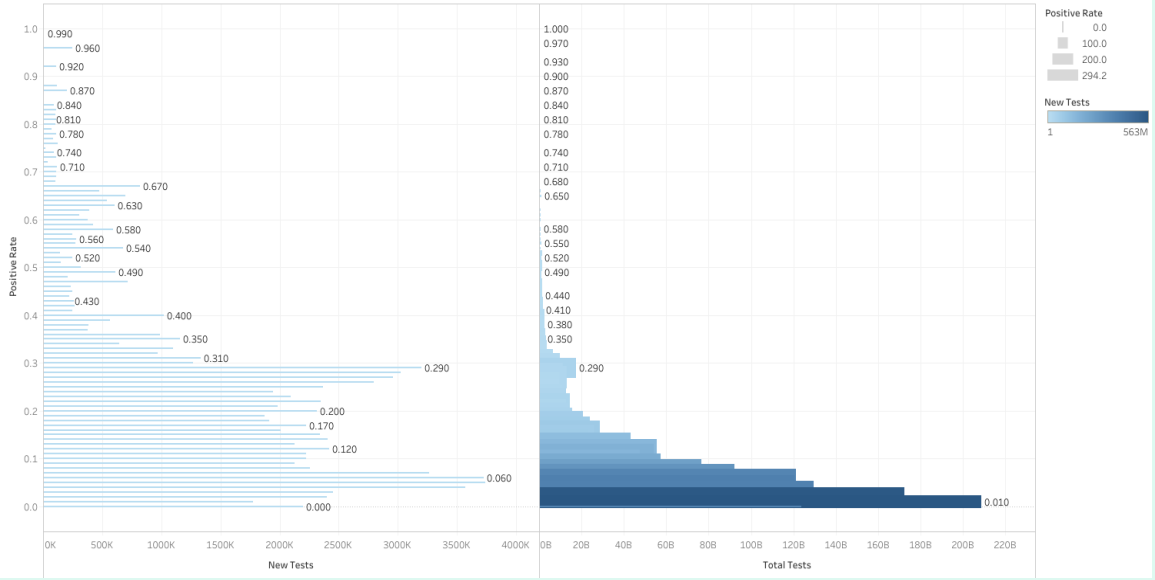
Global Testing Rate (Map)
Global Testing Rate," visually represents the COVID-19 testing rate across different countries worldwide. The shading of each country on the map indicates its testing rate, providing a geographical overview of testing efforts. This visualization helps to identify regions with higher or lower testing intensities, which can influence the reported case numbers and overall understanding of the pandemic's spread.

Figure 11: Global COVID-19 Testing Rate by Country *This map illustrates the global distribution of COVID-19 testing rates across countries, highlighting geographical variations in testing intensity.*

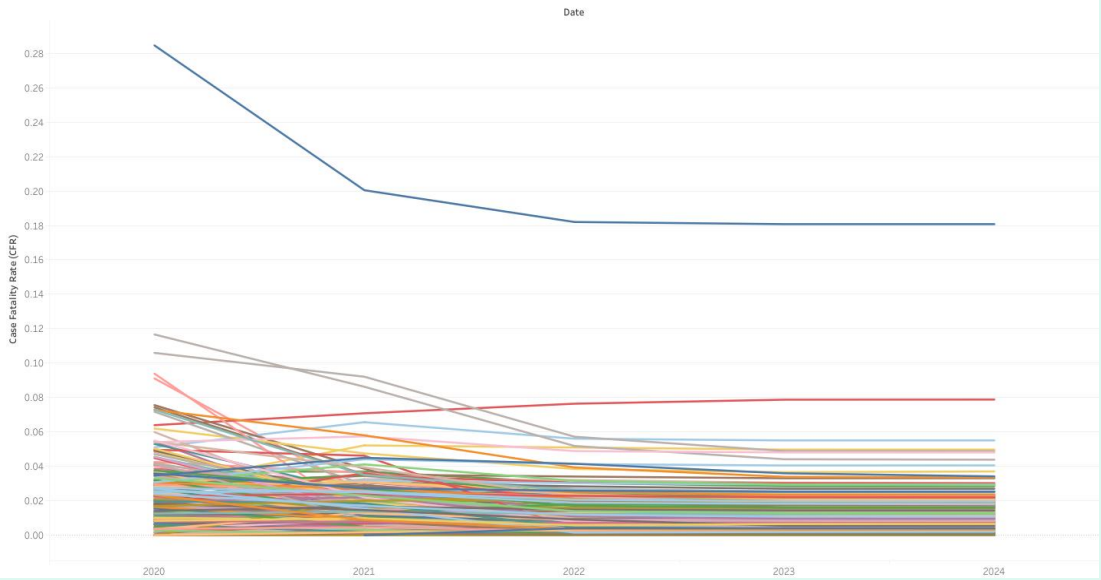
A tibble: 11 × 7

location	new_cases	new_deaths	new_tests	new_vaccinations	case_fatality_rate	posibility_rate
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
China	59789650705	79278696	9464000000	1.712764e+12	0.001325960	6.3175878
India	52079485202	636656068	296477930381	2.099624e+12	0.012224700	0.1756606
Brazil	40274470410	845533502	12343376260	2.019023e+11	0.020994280	3.2628407
France	35276866268	197522787	83214454871	1.029527e+11	0.005599216	0.4239271
Germany	33383624547	190391817	6628848380	1.096179e+11	0.005703150	5.0361123
South Korea	25657640953	29376266	20722710943	5.449662e+10	0.001144932	1.2381411
United Kingdom	24768373218	277603206	144400410838	6.060255e+10	0.011207971	0.1715256
Japan	24572158686	60353535	13429500083	1.836412e+11	0.002456176	1.8297151
Italy	23516696786	226781854	63349522489	1.122985e+11	0.009643440	0.3712214
Russia	22717049874	430102745	57999569801	5.145820e+10	0.018933037	0.3916762
Turkey	17595151920	114536586	47891178076	6.314574e+10	0.006509554	0.3673986

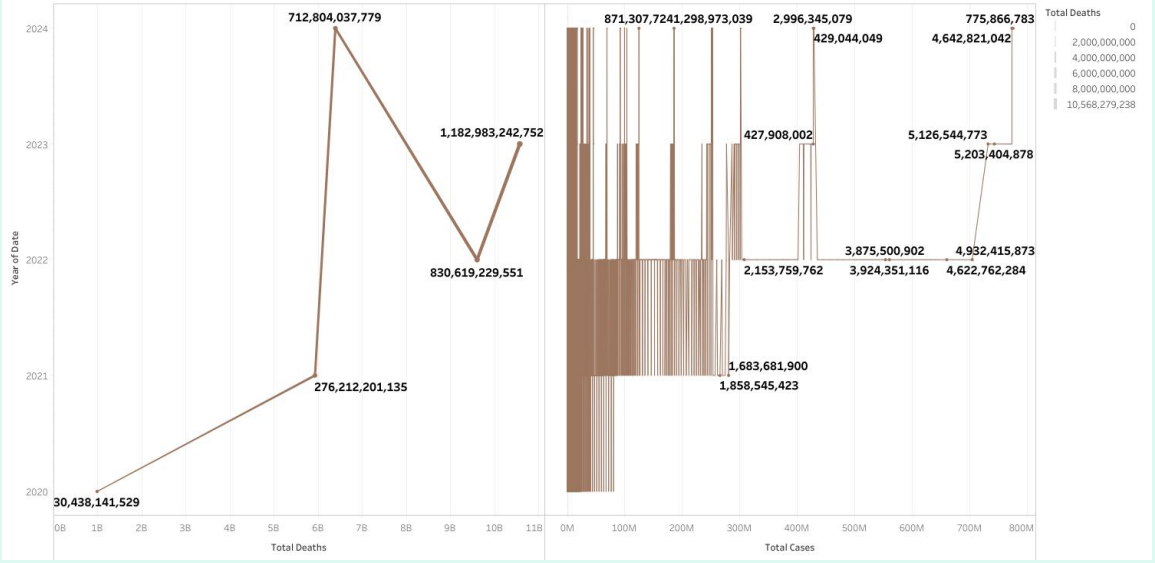
Test & Positivity



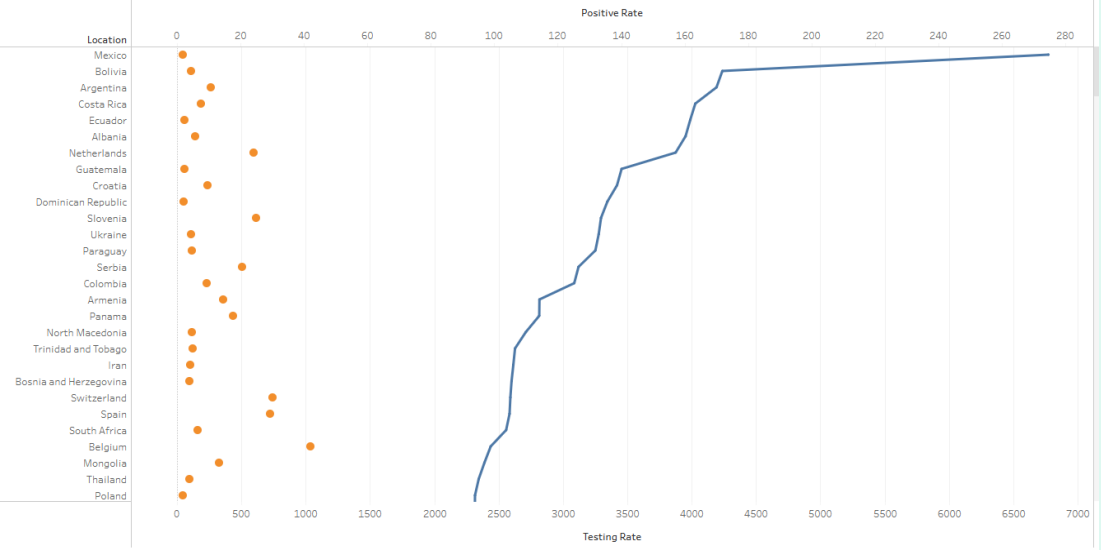
Mortality & Recovery



Trend Analysis



Dual-Axis Chart: Testing Rate vs. Positivity Rate



First page | Second page | Third page

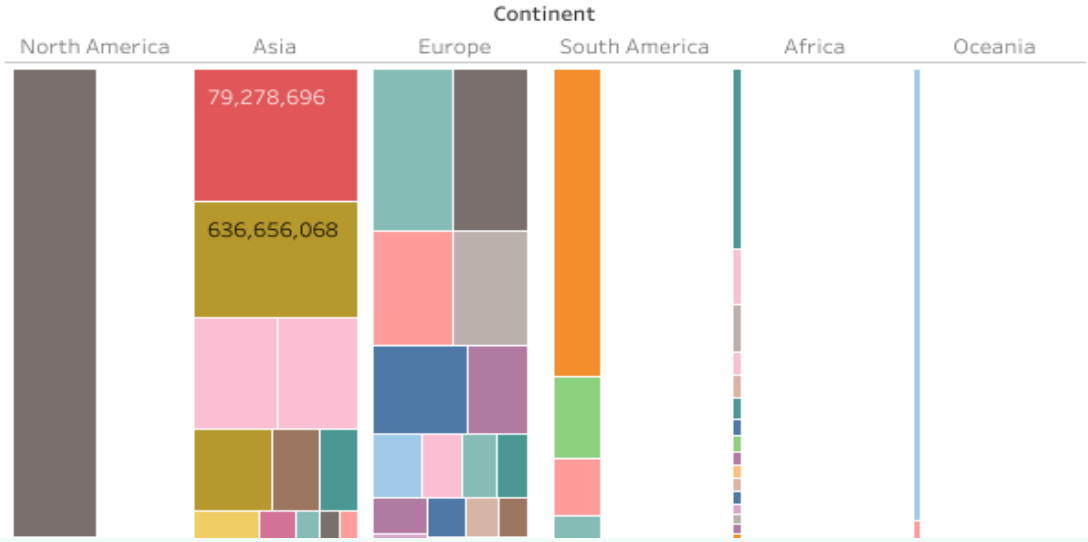
GLOBAL COVID-19 DASHBOARD

KPi

Continent	70 Years Older	Aged 65 Older	Median Age	Cardiovasc Death Rate	Fatality Rate	Total Tests	Total Cases	Total Deaths
Africa	212,817	331,753	1,999,287	25,970,587	10,599,771	22,213,357,604	14,616,886,098	306,955,815
Asia	338,205	530,271	2,536,786	24,101,500	17,868,743	##########	##########	1,790,404,880
Europe	768,870	1,187,098	2,818,274	15,804,485	12,711,177	##########	##########	2,361,106,039
North Ame..	281,651	416,201	1,535,677	8,577,785	10,653,200	##########	##########	1,880,636,077
Oceania	98,315	161,096	601,064	8,420,251	16,386,100	24,123,978,485	11,598,053,992	23,933,802
South Ame..	109,074	166,323	649,472	3,763,732	1,334,510	61,969,482,409	73,484,570,403	1,646,924,372

Dual-Axis Chart: Testing Rate vs. Positive Rate

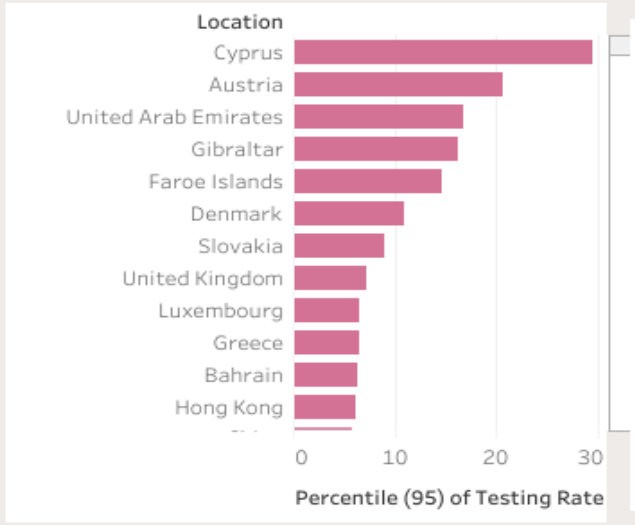
Hotspot Identification



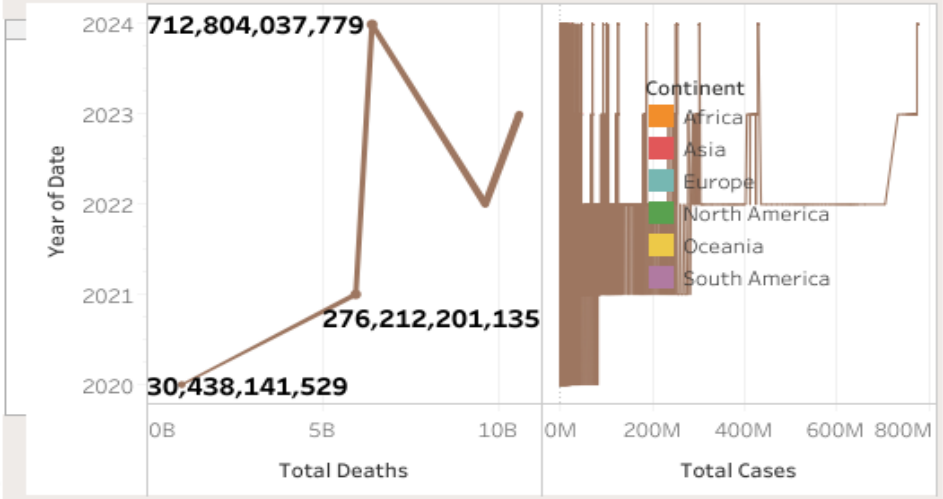
Map Visualization: Global Testing Rate

First page | Second page | Third page

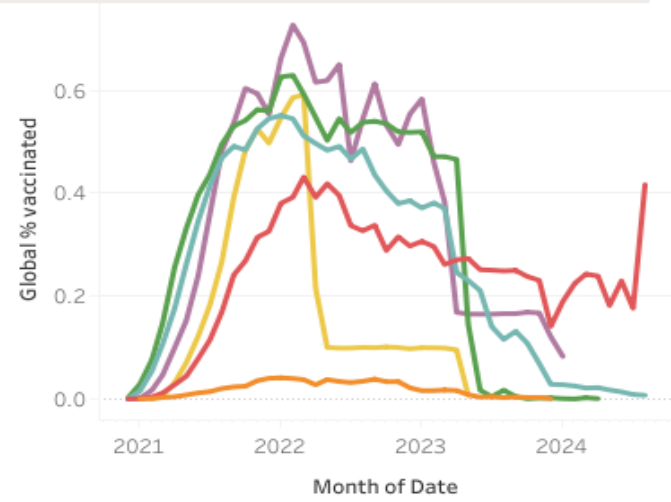
Testing & Detection



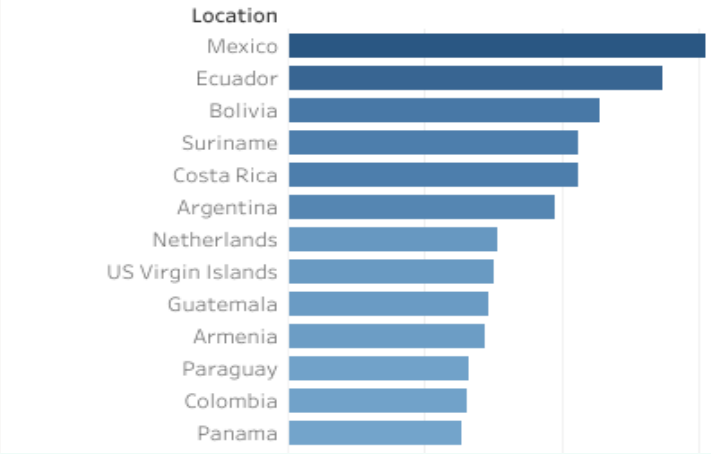
Trend Analysis



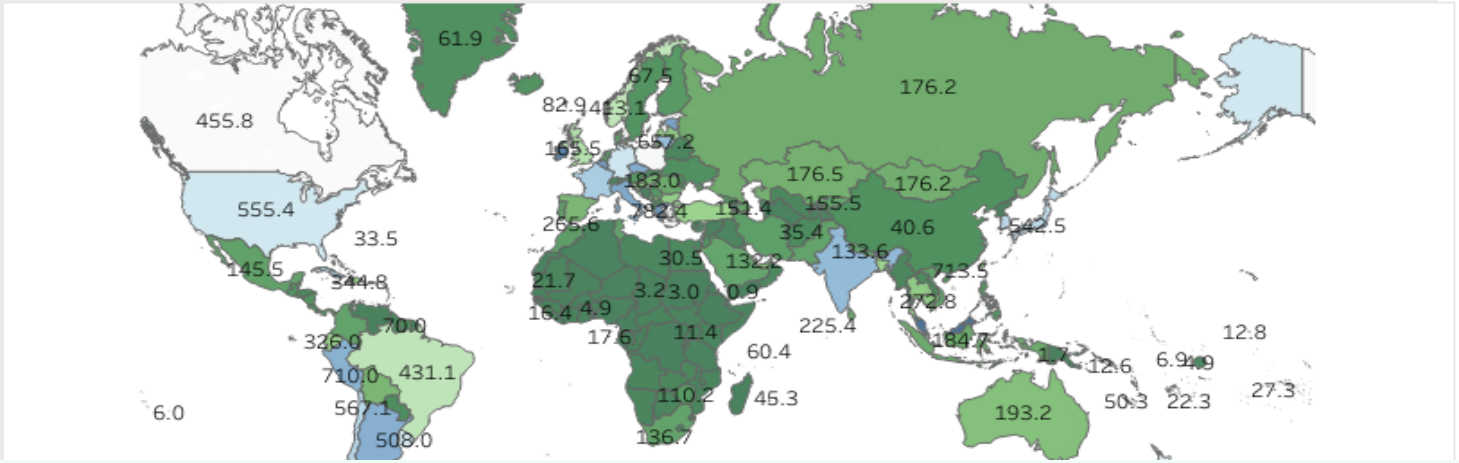
Time Series by Region



Positivity Detection

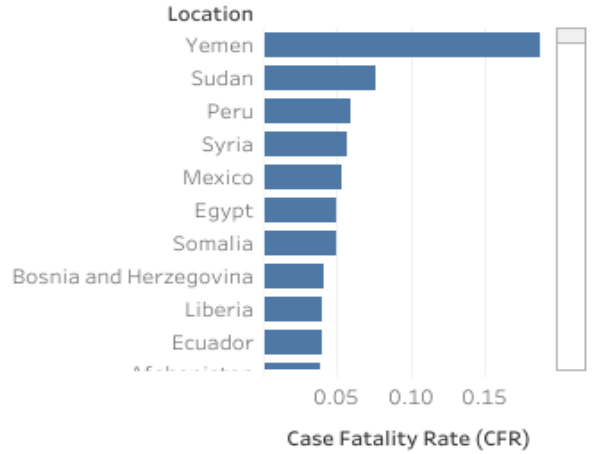


Choropleth Map (Color-coded by vaccination progress)

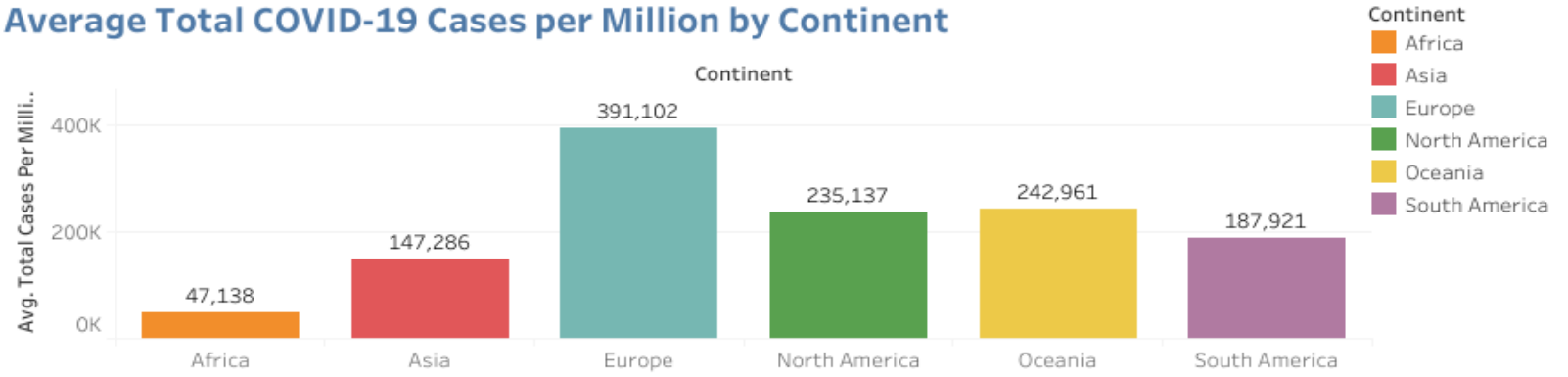


DATASET SOURCE: GLOBAL COVID-19 DATASET,
VISUALIZED BY AUSTINE OBASUYI
WWW.DATACALCULATIONS.COM..

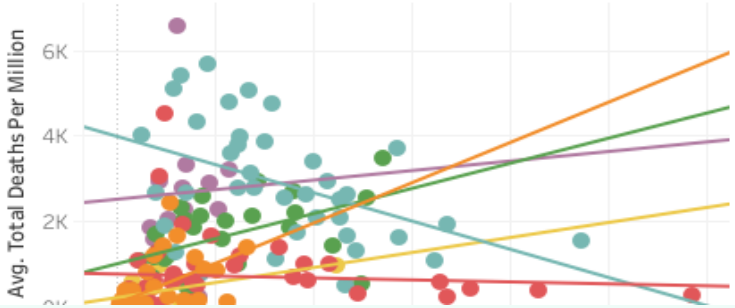
Recommended for Snapshot Comparison



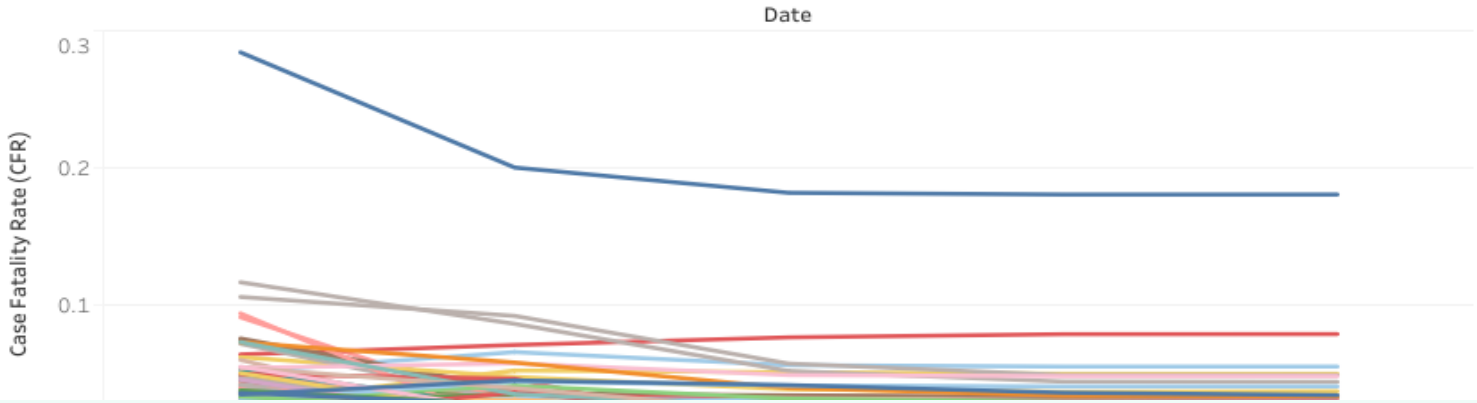
Metrics by Region (Continent).
Average Total COVID-19 Cases per Million by Continent



Metrics by Income Level (GDP per Capita)
Total COVID-19 Cases per Million vs.
Hospital Beds per Thousand



Mortality & Recovery



Executive Summary:
Global COVID-19 Data Analysis

- **Purpose:** This project aims to identify global COVID-19 patterns, uncover critical risk zones, and provide actionable insights to inform strategic planning and decision-making for stakeholders.

- **Tools:** I Austine leveraged a robust toolkit, including **R** for advanced data wrangling, statistical analysis, and custom visualization using ggplot2, and **Tableau** for creating interactive dashboards and geographical insights.

- **Objective of the Analysis:** The primary objective of this analysis was to transform a vast and complex global COVID-19 dataset into clear, actionable intelligence. I sought to understand the pandemic's progression, assess the impact of various interventions (like testing and vaccination), and identify the underlying factors—such as socio-economic conditions and healthcare capacity—that have shaped its outcomes worldwide. This comprehensive understanding is crucial for informing future strategies and resource allocation.

- **Dataset Description:** The core of this analysis is a comprehensive global COVID-19 dataset, randomly obtained from Kaggle. This dataset tracks crucial metrics related to the pandemic's progression, including daily new cases, total cases, total deaths, testing volumes, and vaccination efforts. It also incorporates various demographic, economic (e.g., GDP per capita), and healthcare capacity (e.g., hospital beds per thousand) indicators, enabling a multi-dimensional examination of the pandemic's impact.

- **Methodology:** My methodology involved a rigorous multi-stage approach:
- **Data Wrangling & Cleaning:** Initial steps focused on cleaning and preparing the raw data, including handling missing values, standardizing date formats, and converting data types to ensure accuracy and consistency.

- **Derived Metric Calculation:** Key performance indicators such as Case Fatality Rate (CFR - deaths per cases), testing rates per population, and cases per million were calculated to provide standardized comparative metrics.

- **Exploratory Data Analysis (EDA):** I performed extensive EDA using R to identify initial trends, outliers, and relationships within the data.

- **Visualization & Dashboarding:**
  - **R (ggplot2):** Utilized for creating detailed line plots (e.g., Daily New Cases, Fatality Rate vs. Positivity) and scatter plots (e.g., Cases/Deaths vs. GDP/Hospital Beds) to illustrate trends and correlations.
  - **Tableau:** Employed for developing interactive dashboards, including bar charts (e.g., Top 10 Deaths, Average Cases by Continent, Testing Rate Percentiles), treemaps (e.g., Hotspot Identification), and choropleth maps (e.g., Global Vaccination Rate, Global Testing Rate) to provide geographical and comparative insights.
- **Comparative Analysis:** Insights were derived by comparing metrics across continents, income levels, and healthcare capacities, allowing for a nuanced understanding of varying pandemic experiences.
- **Key Insights:** I revealed several critical insights into the global COVID-19 pandemic:
- **Diverse Pandemic Trajectories & Hotspots:** Visualizations of daily new cases demonstrated varied pandemic waves and peaks across countries, highlighting the dynamic nature of outbreaks. Treemaps and bar charts of total deaths effectively identified geographical "hotspots" and nations with the highest absolute burdens (e.g., **Figure 2: Top 10 Countries by Total COVID-19 Deaths**, **Figure 3: Hotspot Identification by Continent and Country**).
- **Significant Testing Gaps:** I also analyze testing rates (e.g., **Figure 4: 95th Percentile of COVID-19 Testing Rate by Country**, **Figure 11: Global COVID-19 Testing Rate by Country**) revealed substantial disparities in testing capacities and efforts globally. Regions with lower testing rates may have experienced underreported cases, impacting the accuracy of observed infection rates. Addressing these gaps is crucial for effective pandemic management.

- **Uneven Vaccination Progress:** The global vaccination map (**Figure 5: Global Cumulative Percentage Vaccinated**) clearly illustrated an uneven distribution of vaccination progress worldwide. While some regions achieved high coverage, others lagged significantly, impacting global herd immunity and highlighting the need for equitable vaccine distribution strategies.

- **Mortality Patterns & Influencing Factors:** The relationship between case fatality rates and new cases (**Figure 6: Fatality Rate vs. Global Cases Positivity Rate**) provided insights into the severity of the disease and potential reporting biases. The Case Fatality Rate (CFR) itself was a critical indicator of healthcare system strain and disease impact.

- **Socio-Economic & Healthcare Correlations:** Comparative analyses (**Figure 9: Average Total COVID-19 Cases per Million by Continent**, **Figure 10: Total COVID-19 Cases per Million vs. GDP Per Capita and Hospital Beds per Thousand**) indicated that a nation's economic status (GDP per capita) and healthcare infrastructure (hospital beds per thousand) often correlate with its ability to manage the pandemic, influencing observed case and death rates. Continents like Europe showed higher average cases per million, while the relationship with GDP and hospital beds varied, suggesting a complex interplay of factors.

- **Recommendations:** Based on these insights, I recommend the following strategic actions:
- **Targeted Public Health Interventions:** Tailor public health strategies to specific regional and national contexts, leveraging insights from hotspot identification and varying pandemic trajectories.
- **Strengthen Global Testing Infrastructure:** Invest in and support regions with identified testing gaps to improve early detection, contact tracing, and containment efforts, thereby enhancing the accuracy of global pandemic data.
- **Advocate for Equitable Vaccine Distribution:** Prioritize initiatives that promote fair and timely access to vaccines globally, recognizing that widespread and equitable vaccination is key to mitigating future waves and achieving global recovery.
- **Data-Driven Policy Formulation:** Continuously monitor and analyze evolving COVID-19 data to inform agile public health policies, optimize resource allocation, and support economic recovery plans.
- **Invest in Healthcare Resilience:** Emphasize the importance of robust healthcare systems and infrastructure, particularly in regions with lower healthcare capacity, to better withstand future health crises.
- **Enhance Preparedness for Future Pandemics:** Utilize these lessons learned to develop comprehensive strategies for future pandemic preparedness, focusing on early warning systems, rapid response mechanisms, and resilient supply chains.

THANKS

Mr Austine Obasuyi

milltech@datacalculations.com

sirmilltech@gmail.com

+90-538-779-3556

MY WEBSITE

https://github.com/sirmilltech

https://www.datacalculations.com/

https://datascienceportfol.io/sirmilltech

https://www.linkedin.com/in/austine-o-obasuyi-209269352/

DATA CALCULATIONS(Milltech)