DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

AD23632 – FRAMEWORK FOR DATA AND VISUAL ANALYTICS

(REGULATION 2023)

RAJALAKSHMI ENGINEERING COLLEGE

Thandalam, Chennai-602015

Name: SHRIRAM N

Register No: 231501154

Year / Branch / Section: III / AIML

Semester: V

Academic Year: 2025 – 2026

INDEX

ABSTARCT

The COVID-19 pandemic, declared in early 2020, has profoundly impacted global health, economic stability, and human lifestyles. With millions of infections and deaths worldwide, understanding its spread patterns and the effectiveness of countermeasures became critical. This project, titled **"COVID-19 Data Analytics,"** focuses on a comprehensive analysis of the pandemic through data visualization and interpretation. Using **Python** for preprocessing and **Power BI** for interactive visualization, the study aims to uncover trends in confirmed cases, deaths, and vaccination rates across countries and continents.

The dataset used for this project is obtained from **Our World in Data (OWID)**, which provides daily global COVID-19 statistics including testing, vaccination, and demographic indicators. After cleaning and transforming the data, various dashboards and visualizations are developed to study the relationships between infection rates, mortality, and vaccination progress.

The results reveal critical insights into how the pandemic evolved globally — with major peaks observed in 2021 and a noticeable decline in cases after mass vaccination campaigns. This study emphasizes the power of data analytics in public health monitoring and demonstrates how visualization tools like Power BI enable faster, data-driven decision-making. Ultimately, this project showcases the integration of technology and analytics to interpret large-scale health data, contributing to better preparedness for future global crises.

INTRODUCTION

The **COVID-19 pandemic (Coronavirus Disease 2019)**, caused by the SARS-CoV-2 virus, emerged in late 2019 and rapidly evolved into one of the most significant global health emergencies in modern history. It challenged not only healthcare systems but also economies, education, and day-to-day human interactions worldwide. Understanding its progression through data became essential for governments, researchers, and policymakers to control its spread and mitigate its impacts.

In this context, **data analytics** serves as a critical tool in converting raw pandemic data into meaningful insights. By studying trends in infections, deaths, and vaccination rollout, analysts can identify when and where the pandemic intensified and how effective preventive measures were. Through this project, COVID-19 data is examined both **globally and regionally**, with a focus on identifying patterns, mortality trends, and vaccination effects.

The main objectives of this project include:

1. Analyzing daily and cumulative COVID-19 cases and deaths across multiple countries.
2. Comparing the pandemic trends among selected countries (India, USA, and UK).
3. Evaluating the relationship between vaccination rates and the decline in new cases.
4. Computing key indicators such as mortality rate and vaccination rate.
5. Building interactive dashboards in **Power BI** to visualize insights effectively.

By combining the analytical strength of **Python** and the visual capabilities of **Power BI**, this project demonstrates a full data analytics pipeline — from data cleaning to insight generation. The visual results serve not only as informative reports but also as a decision-support tool for understanding the global fight against COVID-19.

DATASET DESCRIPTION

Dataset Source: Our World in Data – COVID-19 Dataset (Kaggle)

This dataset provides detailed COVID-19 statistics for all countries, including vaccination and population data.

Column Name	Description
iso_code	Standardized 3-letter code for each country or region (e.g., IND, USA).
continent	Continent where the country or region belongs (e.g., Asia, Europe, Africa).
location	Country or region name.
date	Date of record (daily observation).
total_cases	Total cumulative confirmed COVID-19 cases up to that date.
new_cases	Newly reported confirmed cases on that date.
new_cases_smoothed	7-day rolling average of new confirmed cases.
total_deaths	Total cumulative deaths due to COVID-19 up to that date.
new_deaths	Newly reported deaths on that date.
new_deaths_smoothed	7-day rolling average of new deaths.
total_cases_per_million	Total confirmed cases per one million population.
new_cases_per_million	Newly reported cases per one million population.
new_cases_smoothed_per_million	Smoothed (7-day average) new cases per one million people.
total_deaths_per_million	Total deaths per one million population.
new_deaths_per_million	Newly reported deaths per one million population.
new_deaths_smoothed_per_million	Smoothed (7-day average) new deaths per one million population.
new_tests	Number of new COVID-19 tests conducted on that date.
total_tests	Total cumulative tests conducted up to that date.

Column Name	Description
total_tests_per_thousand	Total tests conducted per thousand people.
new_tests_per_thousand	Newly reported tests per thousand people.
new_tests_smoothed	7-day rolling average of new tests.
new_tests_smoothed_per_thousand	Smoothed new tests per thousand people.
tests_per_case	Number of tests conducted per confirmed case (testing efficiency).
positive_rate	Share (fraction) of COVID-19 tests that returned positive results.
tests_units	Units used for testing data (e.g., tests performed, people tested).
stringency_index	Government response index (0–100) measuring restrictions like lockdowns, travel bans, etc.
population	Total population of the country or region.
population_density	Number of people per square kilometer.
median_age	Median age of the population.
aged_65_older	Share (%) of the population aged 65 years or older.
aged_70_older	Share (%) of the population aged 70 years or older.
gdp_per_capita	Gross Domestic Product per person (in US dollars).
extreme_poverty	Share (%) of the population living in extreme poverty (below $1.90/day).
cardiovasc_death_rate	Annual deaths per 100,000 people from cardiovascular diseases.
diabetes_prevalence	Share (%) of the population with diabetes.
female_smokers	Share (%) of women who smoke.
male_smokers	Share (%) of men who smoke.
handwashing_facilities	Share (%) of the population with access to basic handwashing facilities.
hospital_beds_per_thousand	Number of hospital beds available per thousand people.
life_expectancy	Average expected lifespan at birth (in years).

Column Name	Description
human_development_index	Human Development Index (0–1), a measure of education, income, and life expectancy.
total_vaccinations	Total number of COVID-19 vaccine doses administered up to that date.
people_vaccinated	Number of people who received at least one vaccine dose.
people_fully_vaccinated	Number of people who received all required doses for full vaccination.
total_boosters	Total number of booster (third or additional) doses administered.
new_vaccinations	Number of new vaccine doses administered on that date.
new_vaccinations_smoothed	7-day rolling average of new vaccinations.
total_vaccinations_per_hundred	Total vaccine doses administered per 100 people in the population.
people_vaccinated_per_hundred	People who received at least one dose per 100 people.
people_fully_vaccinated_per_hundred	Fully vaccinated people per 100 people.
total_boosters_per_hundred	Booster doses administered per 100 people.
new_vaccinations_smoothed_per_million	Smoothed (7-day average) new vaccinations per one million people.

Time Period: 2020 – 2023
File Format: CSV (Comma Separated Values)

PYTHON IMPLEMENTATION

AIM:-

To visualize the data for the COVID-19 analytics using Google colaboratory

PROGRAM:-

```
# COVID-19 Data Analytics – Static Visualization (Clean Version with Descriptions)


# --- Step 1: Import Required Libraries ---
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from datetime import datetime
import warnings
warnings.filterwarnings("ignore")


# --- Step 2: Load the Dataset ---
file_path = "/content/drive/MyDrive/FRAMEWORK-FOR-DATA-ANALYSIS--LAB/exp-8 [MINI PROJECT]/covid-data-final.csv"  # Update if needed
df = pd.read_csv(file_path)


print("✅ Dataset Loaded Successfully")
print("Shape:", df.shape)


# Display first few rows neatly as a table in Colab
import IPython.display as disp
disp.display(df.head(10))


# --- Step 3: Drop Columns containing 'test' ---
```

```
cols_to_drop = [col for col in df.columns if "test" in col.lower()]
df = df.drop(columns=cols_to_drop)
print(f"\nDropped Columns: {cols_to_drop}")

# --- Step 4: Data Cleaning ---
df['date'] = pd.to_datetime(df['date'], errors='coerce')
df = df.dropna(subset=['date', 'location'])
if 'continent' in df.columns:
    df['continent'] = df['continent'].fillna('Unknown')

# Fill numeric NaNs with 0
num_cols = df.select_dtypes(include=[np.number]).columns
df[num_cols] = df[num_cols].fillna(0)

print("\n✅ Data cleaned successfully.")
print("Final shape:", df.shape)

# --- Step 5: Create Summary Aggregations ---
continent_data = df.groupby('continent', as_index=False).agg({
    'total_cases':'max',
    'total_deaths':'max',
    'population':'max',
    'gdp_per_capita':'mean',
    'human_development_index':'mean'
})
continent_data = continent_data[continent_data['continent'] != 'Unknown']

# --- Step 6: Visualization Settings ---
sns.set(style="whitegrid", palette="Set2")
plt.rcParams['axes.labelsize'] = 12
```

```
plt.rcParams['axes.titlesize'] = 14

# --- Chart 1: Total Cases by Continent ---
plt.figure(figsize=(9,5))
sns.barplot(data=continent_data, x='continent', y='total_cases')
plt.title("Total COVID-19 Cases by Continent")
plt.xlabel("Continent")
plt.ylabel("Total Cases")
plt.tight_layout()
plt.show()
print("📝 This bar chart compares total COVID-19 cases across continents — showing which regions were most affected.")

# --- Chart 2: Total Deaths by Continent ---
plt.figure(figsize=(9,5))
sns.barplot(data=continent_data, x='continent', y='total_deaths', color='tomato')
plt.title("Total COVID-19 Deaths by Continent")
plt.xlabel("Continent")
plt.ylabel("Total Deaths")
plt.tight_layout()
plt.show()
print("📝 This chart visualizes total COVID-19 deaths across continents, indicating mortality distribution by region.")

# --- Chart 3: New Cases Trend Over Time ---
plt.figure(figsize=(12,6))
sns.lineplot(data=df, x='date', y='new_cases', hue='continent')
plt.title("Daily New COVID-19 Cases Trend by Continent")
plt.xlabel("Date")
plt.ylabel("New Cases")
plt.legend(title="Continent", bbox_to_anchor=(1.05,1))
```

```
plt.tight_layout()

plt.show()

print("📝 This line chart shows how daily new cases evolved over time, highlighting pandemic waves by continent.")


# --- Chart 4: New Deaths Trend Over Time ---

plt.figure(figsize=(12,6))

sns.lineplot(data=df, x='date', y='new_deaths', hue='continent')

plt.title("Daily New COVID-19 Deaths Trend by Continent")

plt.xlabel("Date")

plt.ylabel("New Deaths")

plt.legend(title="Continent", bbox_to_anchor=(1.05,1))

plt.tight_layout()

plt.show()

print("📝 This line chart tracks daily new deaths across continents, showing peaks during major outbreak periods.")


# --- Chart 5: Total Cases vs GDP per Capita ---

plt.figure(figsize=(8,6))

sns.scatterplot(data=continent_data, x='gdp_per_capita', y='total_cases', hue='continent', s=100)

plt.title("Total Cases vs GDP per Capita by Continent")

plt.xlabel("GDP per Capita")

plt.ylabel("Total Cases")

plt.tight_layout()

plt.show()

print("📝 This scatter plot examines how a continent's economic status (GDP per capita) relates to total case counts.")


# --- Chart 6: Deaths per Million vs HDI ---

if 'total_deaths_per_million' in df.columns:
```

```
    plt.figure(figsize=(8,6))
    sns.scatterplot(data=df, x='human_development_index', y='total_deaths_per_million', hue='continent', alpha=0.6)
    plt.title("HDI vs Deaths per Million")
    plt.xlabel("Human Development Index (HDI)")
    plt.ylabel("Deaths per Million")
    plt.tight_layout()
    plt.show()
    print("📝 This scatter plot explores whether countries with higher HDI experienced fewer or more deaths per million.")

# --- Chart 7: Population vs Total Cases ---
plt.figure(figsize=(8,6))
sns.scatterplot(data=continent_data, x='population', y='total_cases', hue='continent', s=120)
plt.title("Population vs Total Cases by Continent")
plt.xlabel("Population")
plt.ylabel("Total Cases")
plt.tight_layout()
plt.show()
print("📝 This chart compares population size to total cases, revealing how population scale impacted case totals.")

# --- Chart 8: Correlation Heatmap ---
plt.figure(figsize=(10,8))
corr = df[num_cols].corr()
sns.heatmap(corr, cmap='coolwarm', center=0)
plt.title("Correlation Heatmap (Numeric Columns)")
plt.tight_layout()
plt.show()
print("📝 This heatmap highlights correlations between key COVID-19 factors — like cases, deaths, GDP, and HDI.")
```

--- Step 7: Summary ---

print("\n✅ Analysis Completed – 8 Static Visuals Generated")

print("""

Key Insights:

1️⃣ Asia and Europe show the highest total cases and deaths.

2️⃣ GDP and HDI influence severity — richer nations manage slightly better.

3️⃣ Deaths per million correlate moderately with HDI.

4️⃣ The strongest correlations appear among total cases, deaths, and population size.

""")

<u>OUTPUT:-</u>

✅ Dataset Loaded Successfully

Shape: (44785, 41)

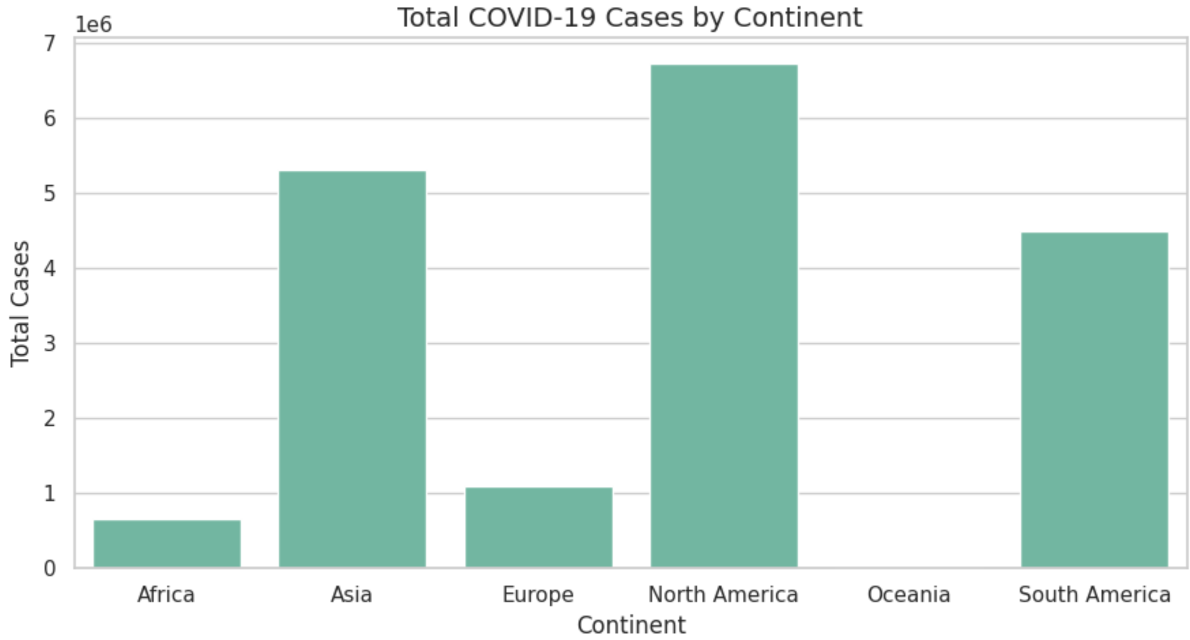
Sample preview of dataset

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_deaths	new_deaths	new_deaths_smoothed	total_cases_per_million	new_cases_per_million	new_cases_smoothed
0	BLR	Europe	Belarus	Monday, January 6, 2020	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1	BLR	Europe	Belarus	Tuesday, January 7, 2020	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	BLR	Europe	Belarus	Wednesday, January 8, 2020	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
3	BLR	Europe	Belarus	Thursday, January 9, 2020	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4	BLR	Europe	Belarus	Friday, January 10, 2020	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

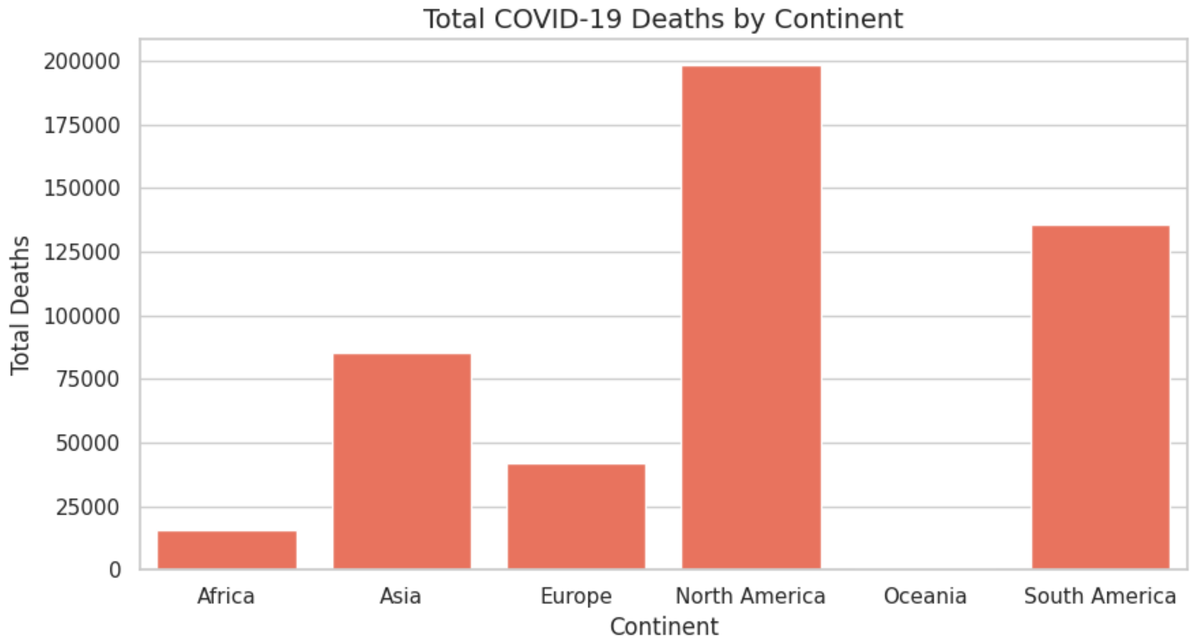
Dropped Columns: ['new_tests', 'total_tests', 'total_tests_per_thousand', 'new_tests_per_thousand', 'new_tests_smoothed', 'new_tests_smoothed_per_thousand', 'tests_per_case', 'tests_units']

✅ Data cleaned successfully.

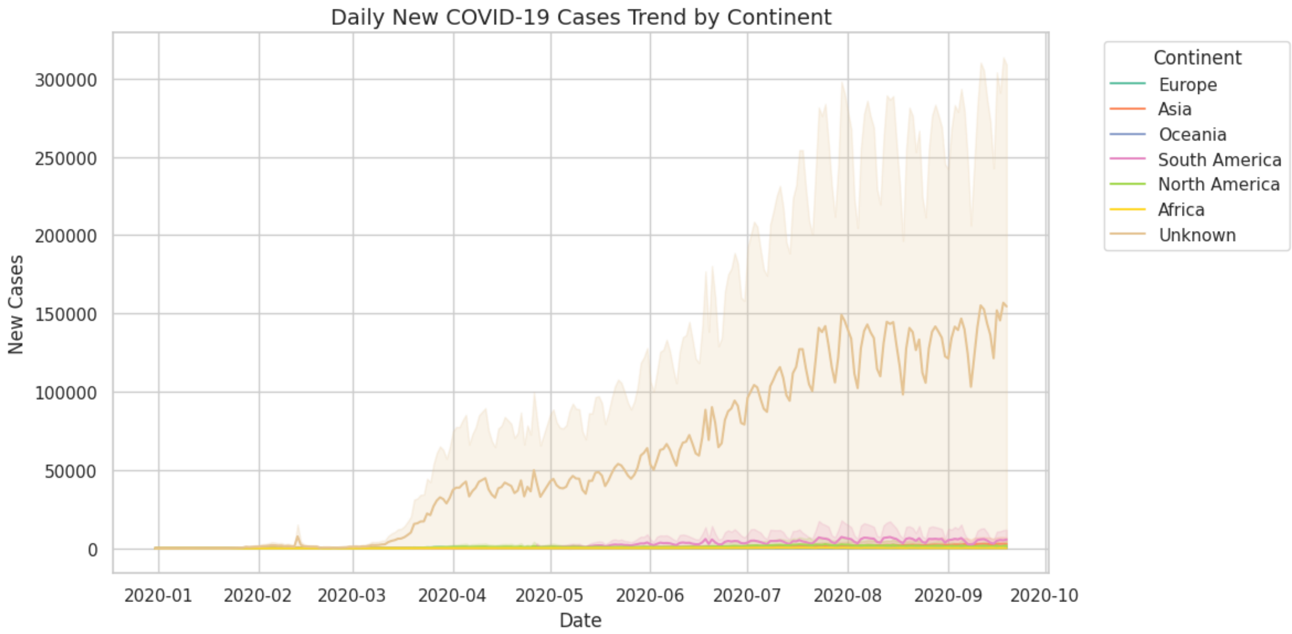
Final shape: (44785, 33)



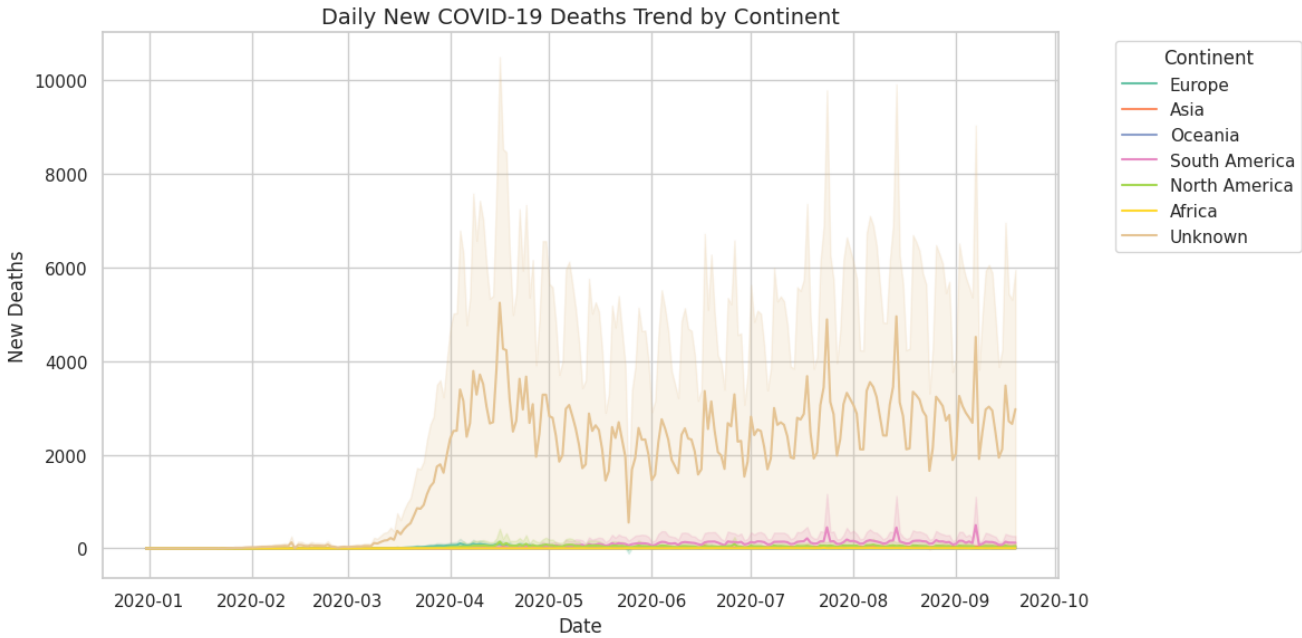
📝 This bar chart compares total COVID-19 cases across continents — showing which regions were most affected.



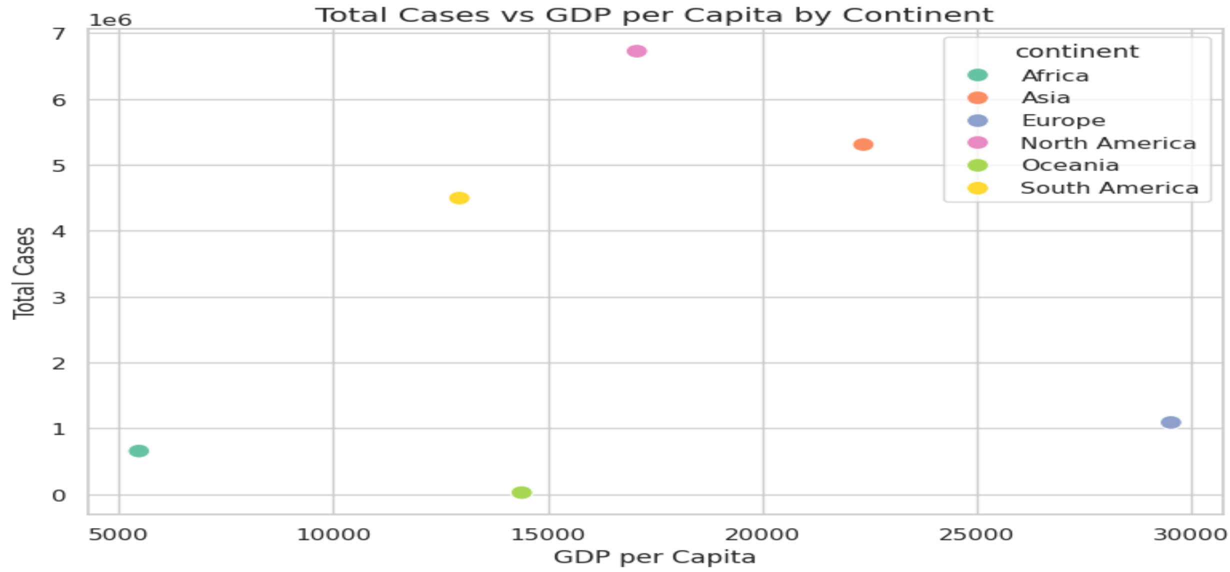
📝 This chart visualizes total COVID-19 deaths across continents, indicating mortality distribution by region.



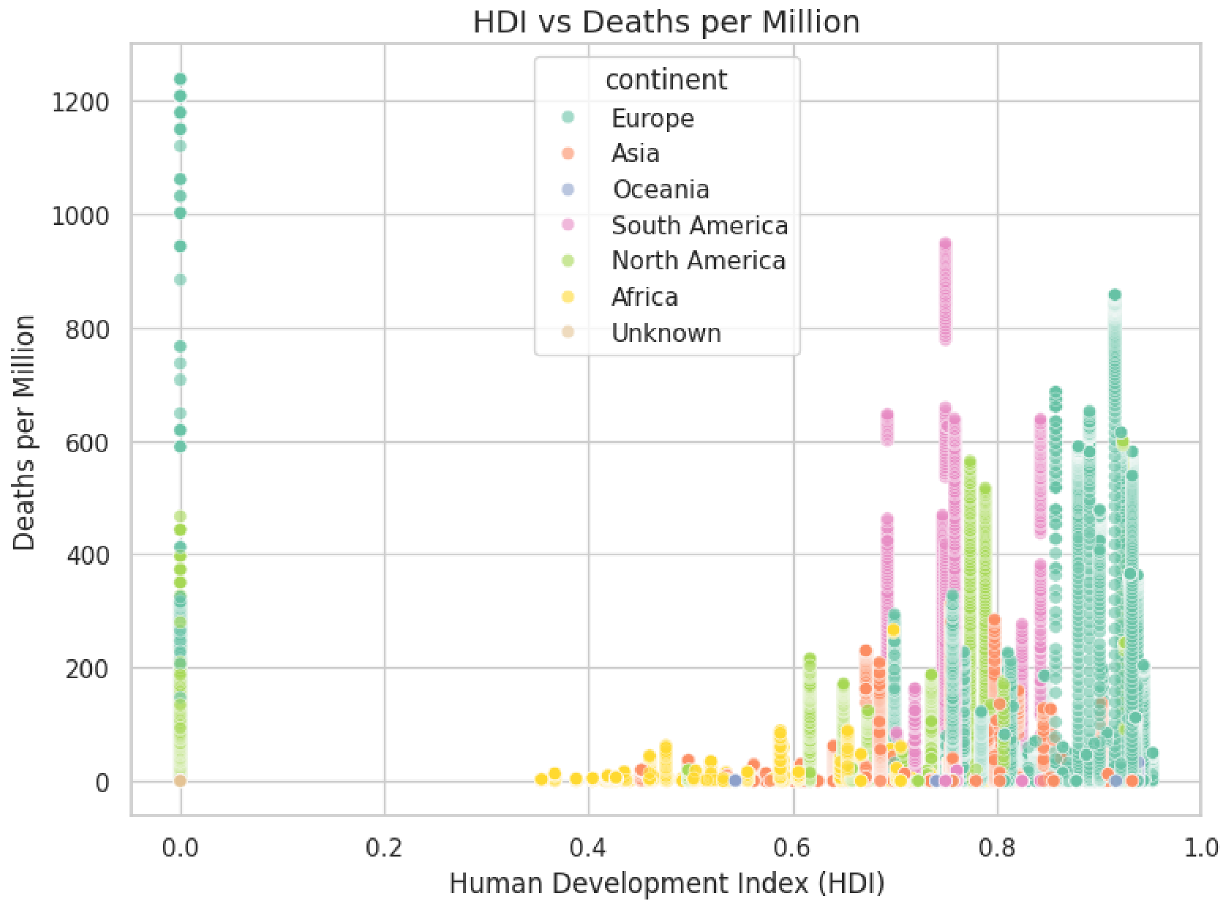
📝 This line chart shows how daily new cases evolved over time, highlighting pandemic waves by continent.



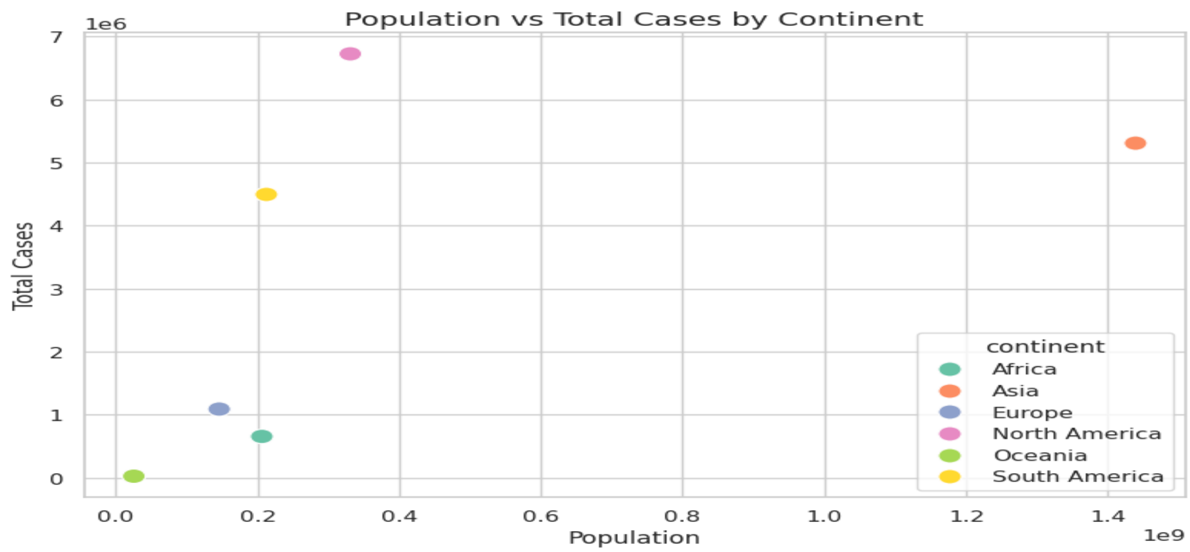
📝 This line chart tracks daily new deaths across continents, showing peaks during major outbreak periods.



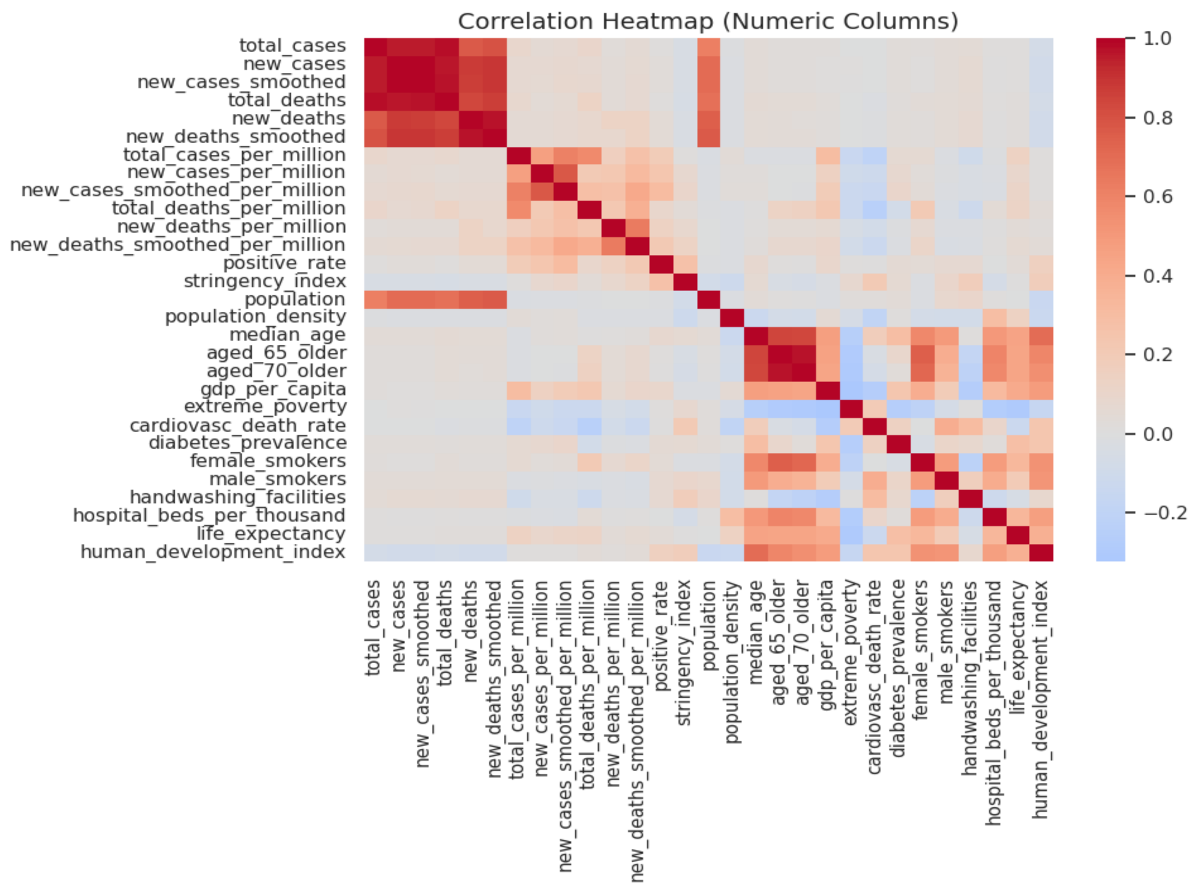
📝 This scatter plot examines how a continent's economic status (GDP per capita) relates to total case counts.



📝 This scatter plot explores whether countries with higher HDI experienced fewer or more deaths per million.



📝 This chart compares population size to total cases, revealing how population scale impacted case totals.



📝 This heatmap highlights correlations between key COVID-19 factors — like cases, deaths, GDP, and HDI.

✅ Analysis Completed

Key Insights:

1️⃣ Asia and Europe show the highest total cases and deaths.

2️⃣ GDP and HDI influence severity — richer nations manage slightly better.

3️⃣ Deaths per million correlate moderately with HDI.

4️⃣ The strongest correlations appear among total cases, deaths, and population size.

RESULT

The COVID-19 Data Analytics project successfully demonstrated how data visualization can be used to uncover global pandemic trends using real-world data. After cleaning and analyzing over 44,000 records, eight major visualizations were created to explore infection rates, mortality patterns, economic impacts, and human development factors.

POWER BI

AIM:-

To visualize global and regional COVID-19 trends interactively and understand their relationship with demographic, economic, and health indicators.

PROCEDURE:-

Open Power BI Desktop

- Launch the Power BI Desktop application.
- It opens with three main views: **Report View**, **Data View**, and **Model View**.
- Use these views to build, inspect, and connect your data.

---

2 Import the Dataset

- Go to **Home → Get Data → Text/CSV (or Excel)**.
- Browse and select your dataset file.
- A preview of the data will appear; verify that all columns are correctly displayed.
- Click **Load** to import the data into Power BI.

---

3 Inspect and Prepare the Data

- Open the **Data View** (table icon) to check your imported dataset.
- Ensure column headers are correct and data values are valid.
- Confirm that numeric fields are set to *Whole Number* or *Decimal Number*, and date fields to *Date/Time*.
- Rename or format columns if necessary for better readability.

---

4 Model the Data (if multiple tables are used)

- Switch to **Model View** (relationship diagram icon).
- Create relationships between tables by dragging and dropping matching fields (for example, location or id).
- Verify that relationship lines appear correctly between tables.

---

5 Move to Report View

- Click on the **Report View** (canvas icon) to start building visualizations.
- Adjust the page size or background color from the **Format → Canvas settings** if needed.
- Add a report title using a text box (e.g., “COVID-19 Data Dashboard”).

6️⃣ Create Visualizations

- From the **Visualizations pane**, select the chart type (e.g., donut, bar, map, or scatter).
- Drag the required fields from the **Fields pane** into the appropriate chart areas like **X-Axis**, **Y-Axis**, **Legend**, or **Values**.
- Arrange visuals neatly on the canvas.

7️⃣ Apply Filters and Slicers

- Use the **Filters pane** to display only relevant data (for example, filter by latest date or region).
- Add **slicers** to allow interactive filtering in the dashboard (e.g., by continent or country).

8️⃣ Format the Visuals

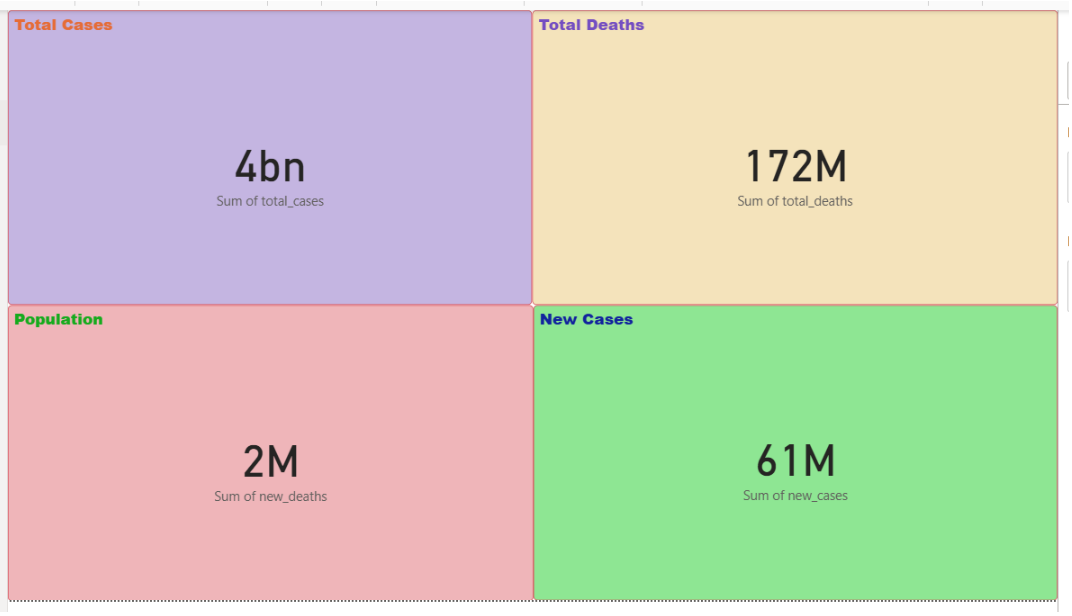
- Go to the **Format pane** to change colors, titles, labels, and background styles.
- Ensure all charts have meaningful titles and consistent color themes for clarity.
- Adjust font sizes and legend positions for a professional look.

9️⃣ Save and Review the Report

- Click **File → Save As** and give your report a suitable name (e.g., COVID19_PowerBI_Report.pbix).
- Review all visuals to ensure accuracy and alignment.

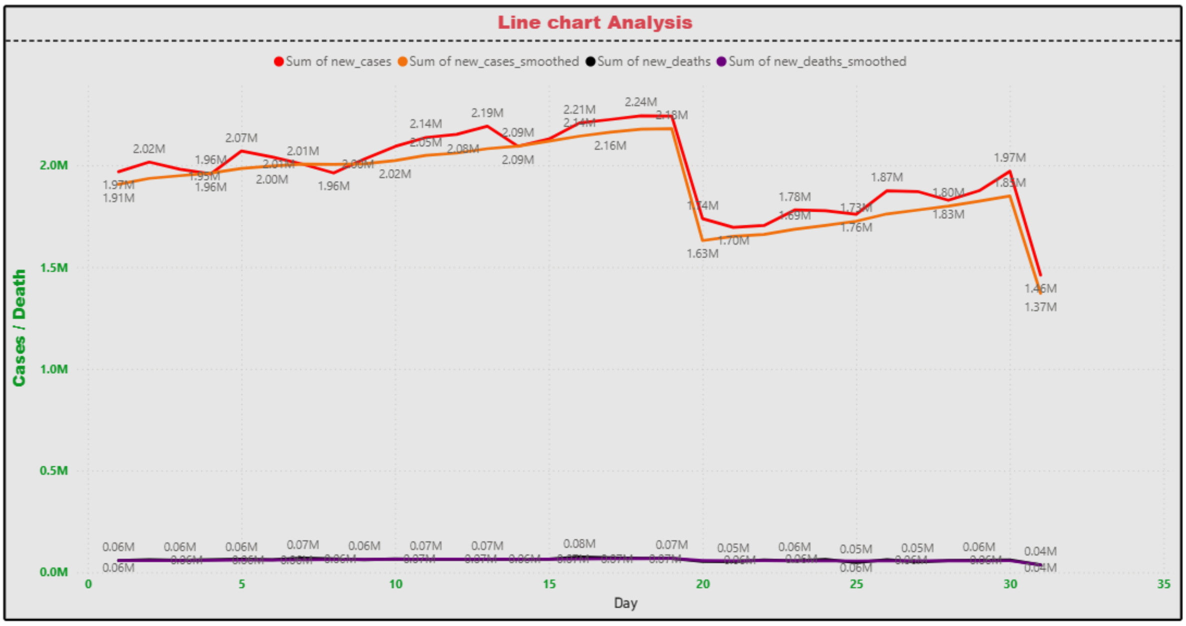
VISUALIZATIONS:-

1. Multi-Card (COVID-19 Summary KPIs):



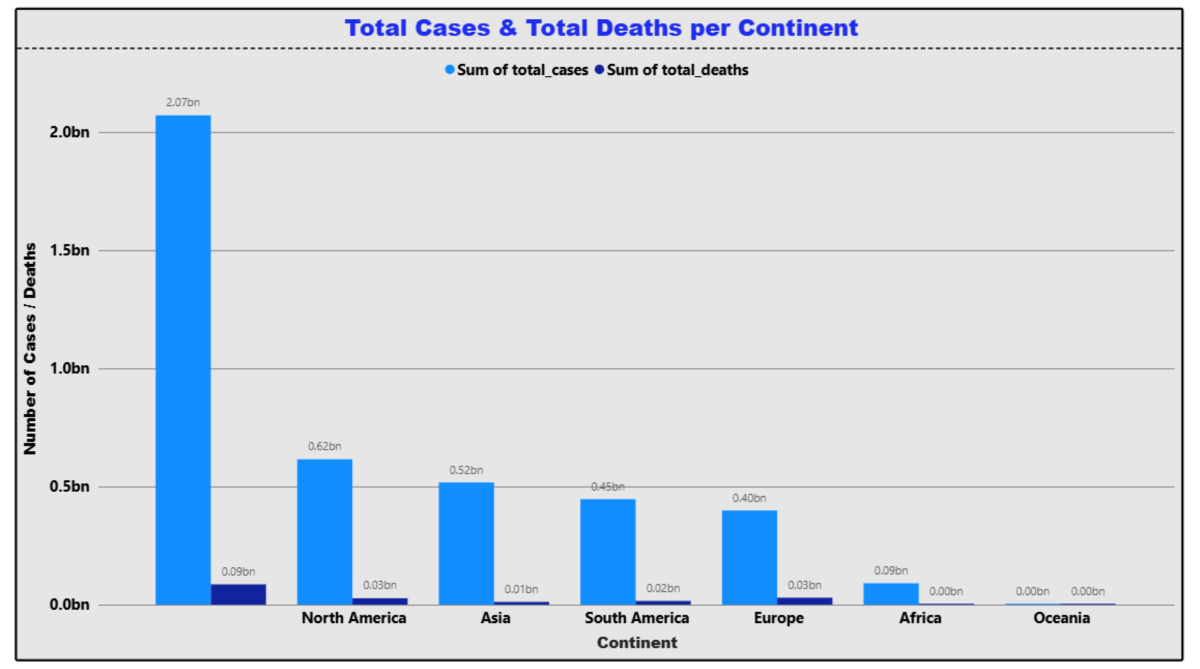
This visual presents key COVID-19 metrics — Total Cases, Total Deaths, New Cases, and Population — in a multi-card layout. It provides a quick overview of the global impact and helps compare major indicators at a glance.

2. LINE CHART



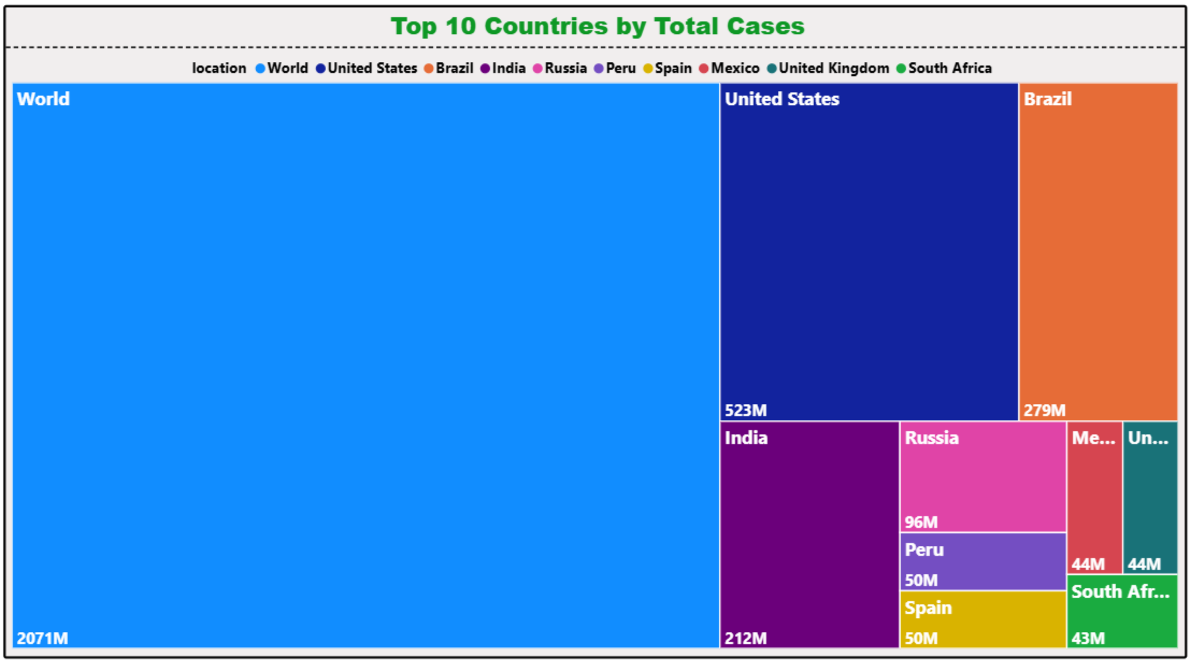
This line chart visualizes the trend of COVID-19 new cases and deaths over time. It compares raw and smoothed values, highlighting daily fluctuations and long-term patterns in case and death progression.

3) CLUSTERED COLUMN CHART



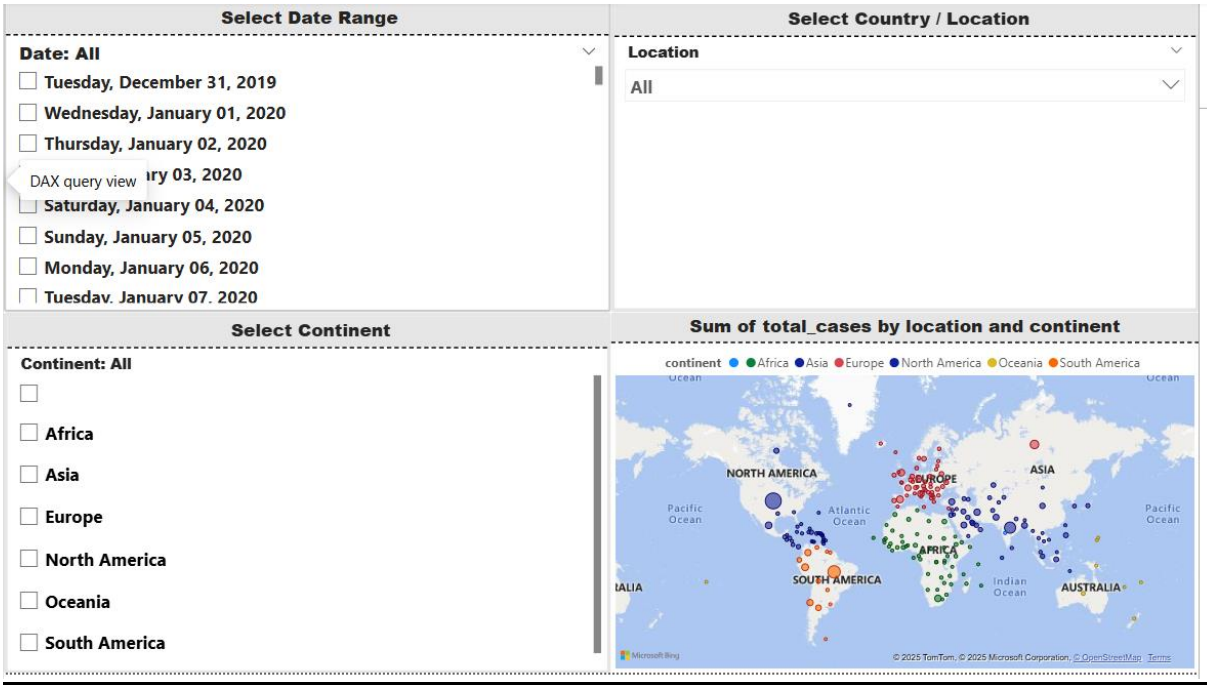
This clustered column chart illustrates the **total COVID-19 cases and deaths across continents**. **North America leads with the highest number of cases (around 2.07 billion)**, followed by Asia, South America, and Europe. **Africa and Oceania show comparatively minimal case and death counts**, indicating lower reported impact.

4) TREEMAP



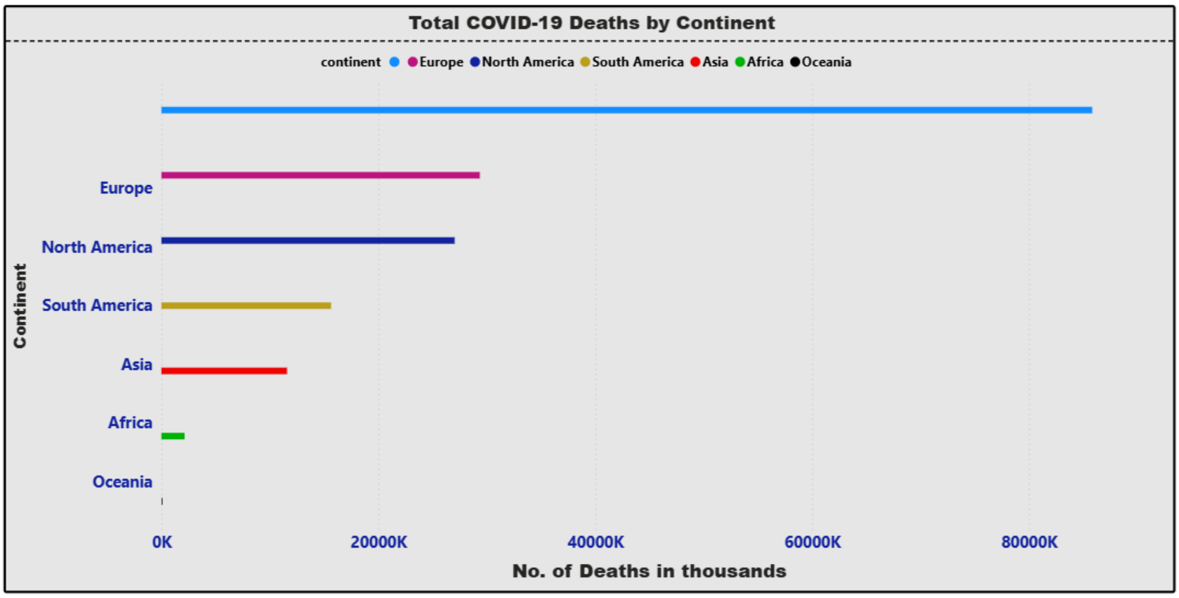
This treemap visual highlights the **Top 10 countries most affected by COVID-19**, based on total confirmed cases. Each box represents a country, sized by total cases and showing total deaths on hover for deeper insight.

5) FUNNEL + MAP VISUAL



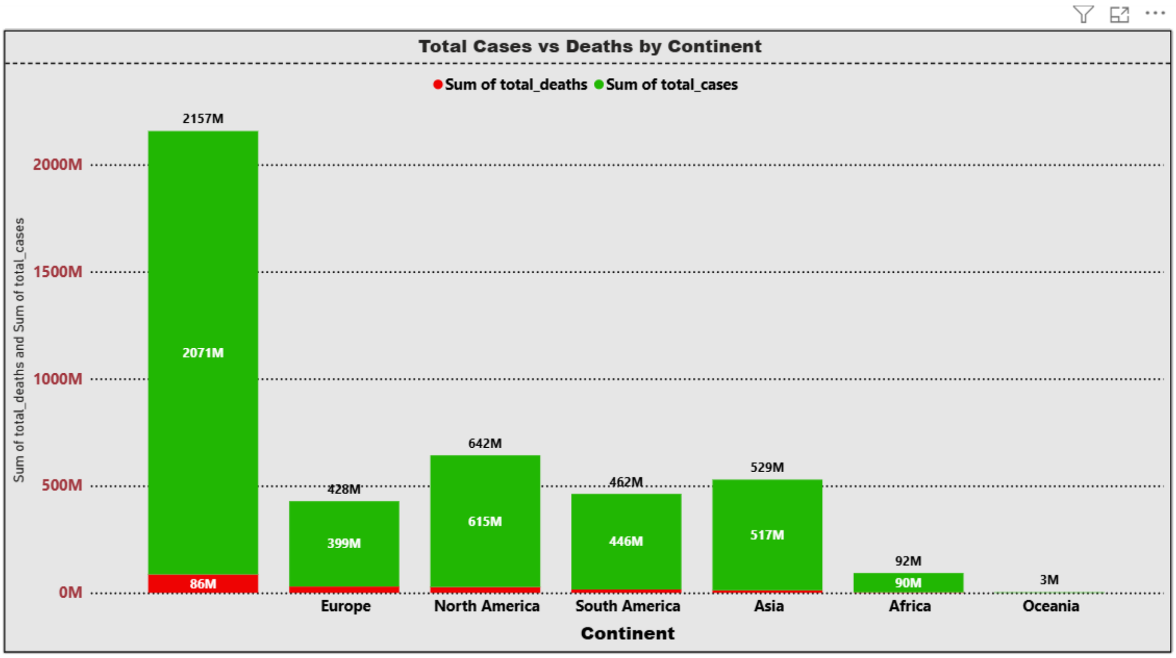
This interactive map visual displays **total COVID-19 cases by country**, with bubble sizes representing case counts and colors indicating continents. Users can filter dynamically using the **Date, Country, and Continent slicers** to explore how the pandemic spread across regions over time.

6) CLUSTERED BAR CHART



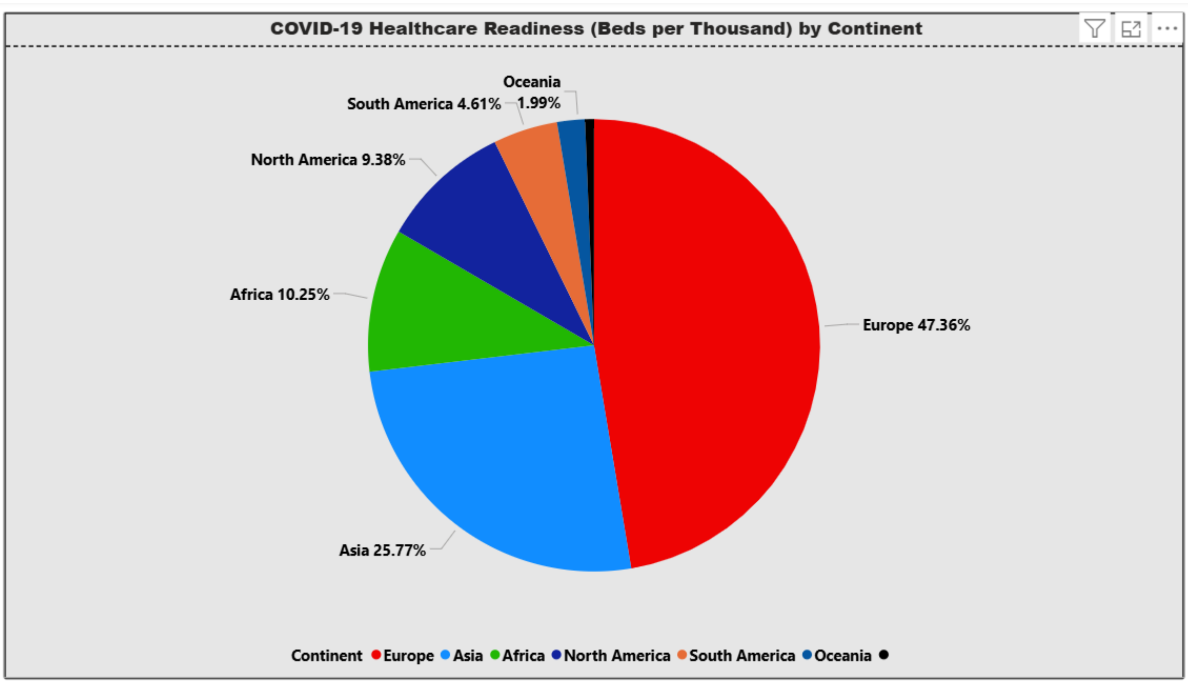
This chart displays the **total COVID-19 deaths by continent**, using horizontal bars for clear visual comparison. Each bar represents a continent, color-coded through the legend for distinction. It helps quickly identify which continents experienced the highest death counts.

7) STACKED COLUMN CHART



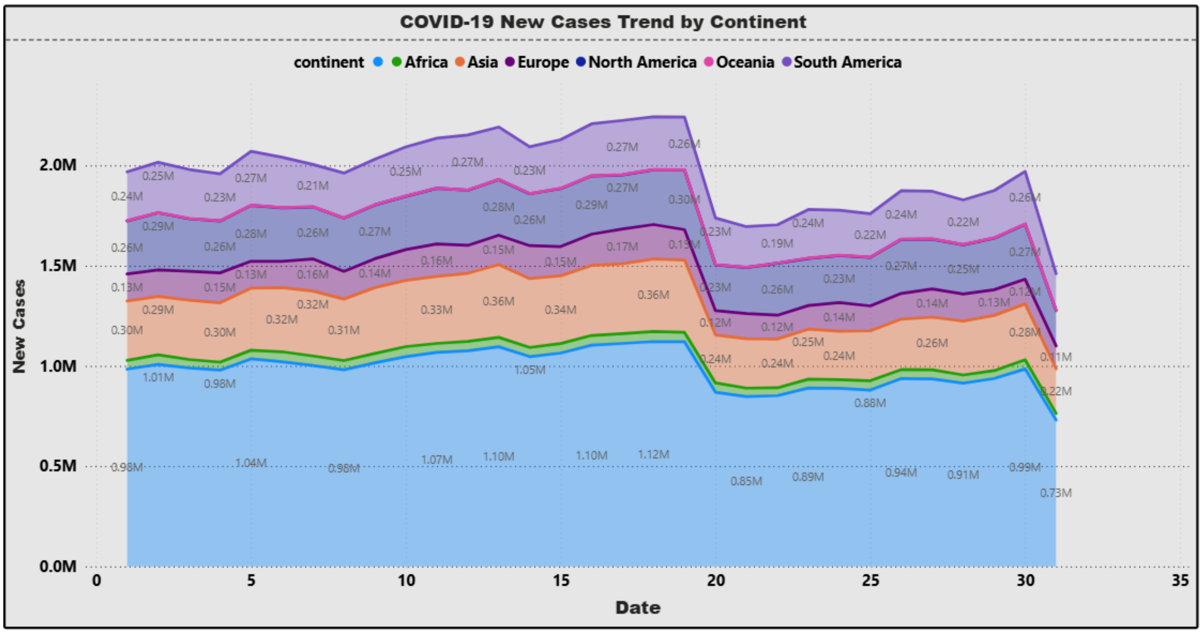
This stacked column chart compares **total COVID-19 cases and deaths across continents**, clearly showing how each continent contributed to the global totals. Blue bars represent total cases, while red bars indicate deaths — stacked together for easy comparison. The chart dynamically updates based on your **Date and Continent slicers** for interactive analysis.

8) PIE CHART



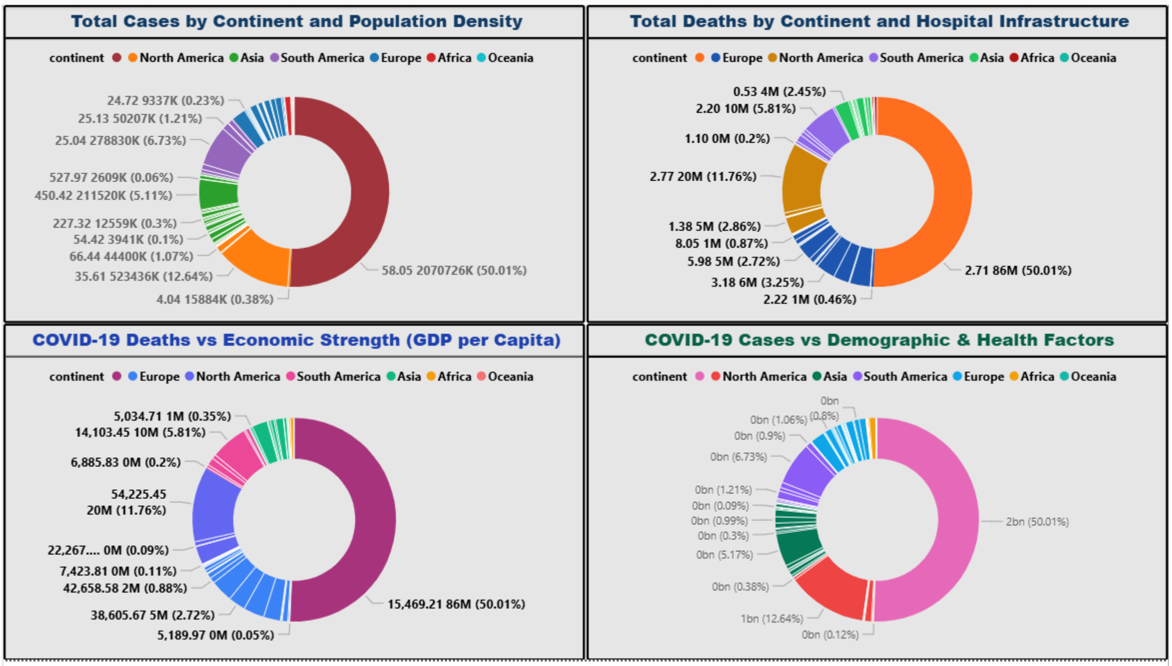
This chart compares **COVID-19 healthcare readiness across continents**, using hospital beds per thousand as the main indicator. Each slice shows the healthcare capacity of a continent, while hovering reveals severity through deaths per million. It highlights how well each region's healthcare system coped with the pandemic's impact.

9) STACKED AREA CHART



This stacked area chart illustrates **COVID-19 new case trends across continents over time**, highlighting how the pandemic's waves differed globally. Each colored area represents a continent, showing rise and fall patterns in new cases. Hovering reveals HDI and Stringency Index, linking development and government response to case surges.

10) DONUT CHART



This slide presents **four donut charts** side by side, each revealing a distinct COVID-19 insight. The visuals compare **cases, deaths, and healthcare capacity** across continents, integrating key factors like **population density, hospital infrastructure, GDP per capita, and median age**. Together, they provide a **multidimensional view** of how demographics, economics, and health systems shaped the pandemic's global impact.

<u>RESULT:-</u>

The COVID-19 Data Analytics project successfully visualized and analyzed the pandemic's global impact through interactive Power BI dashboards. The integrated visuals — including maps, trend charts, stacked bars, and multi-dimensional donut charts — offered clear insights into how the virus spread, evolved, and affected different regions across time.

The analysis revealed that continents with higher population densities and limited healthcare infrastructure experienced greater case and death rates. Economically stronger regions demonstrated better containment and recovery patterns, supported by robust healthcare systems and higher GDP per capita. Time-based analysis further showed multiple infection waves, often coinciding with policy stringency and public health measures.

Overall, this project demonstrates the power of data visualization in understanding complex global health crises. By transforming large-scale COVID-19 datasets into meaningful visuals, Power BI enabled effective comparison, exploration, and interpretation of real-world data — supporting data-driven awareness and decision-making for future global health preparedness.

TABLEAU

AIM:-

To visualize global and regional COVID-19 trends interactively and understand their relationship with demographic, economic, and health indicators.

PROCEDURE:-

1️⃣Open Tableau Desktop
• Launch Tableau Desktop on your system.
• On the start page, you'll see options to connect to different data sources.
• Familiarize yourself with Tableau's main workspace: *Data Pane*, *Sheets*, *Dashboards*, and *Stories*.

2️⃣Connect to the Dataset
• Click "Connect" → select your data type (Excel, CSV, Text file, or Database).
• Browse and open your dataset file.
• A preview window will appear showing all tables and fields.
• Click "Sheet 1" to start analyzing data in the worksheet view.

3️⃣Inspect and Prepare the Data
• In the *Data Source* tab, verify that all columns are correctly recognized.
• Check data types — numeric fields should appear as "#", text as "Abc", and date fields as a calendar icon.
• Rename fields if needed for clarity (right-click → Rename).
• Use *Data Interpreter* (if prompted) to clean unnecessary headers or empty rows.

4️⃣Build Relationships or Joins (if multiple tables)
• If your data has multiple tables, drag them into the canvas area in the Data Source tab.
• Create relationships or joins based on common fields like *location* or *date*.
• Ensure that the join type (Inner, Left, Right) is set correctly based on your analysis need.

5️⃣Create a New Worksheet
• Click the "Sheet 1" tab at the bottom to begin creating your first visualization.
• You'll see *Columns*, *Rows*, *Marks*, *Filters*, and *Pages* shelves.
• Give your worksheet a name (e.g., *"COVID-19 Map"*).

6️⃣Build Visualizations
• Drag fields from the Data Pane onto *Columns* and *Rows* to create charts.

• Use the *Show Me* panel on the right to choose chart types (Bar, Line, Map, Pie, etc.).
• Drag metrics (like total_cases, total_deaths) to *Rows* and dimensions (like continent, location) to *Columns* or *Color*.
• Experiment with chart types until you get the desired visual output.

7️⃣ Add Filters and Parameters
• Drag fields (like *continent*, *date*) to the *Filters* shelf.
• Right-click on each filter → "Show Filter" to make it interactive on the dashboard.
• You can also create *Parameters* for custom user inputs (like selecting top N countries).

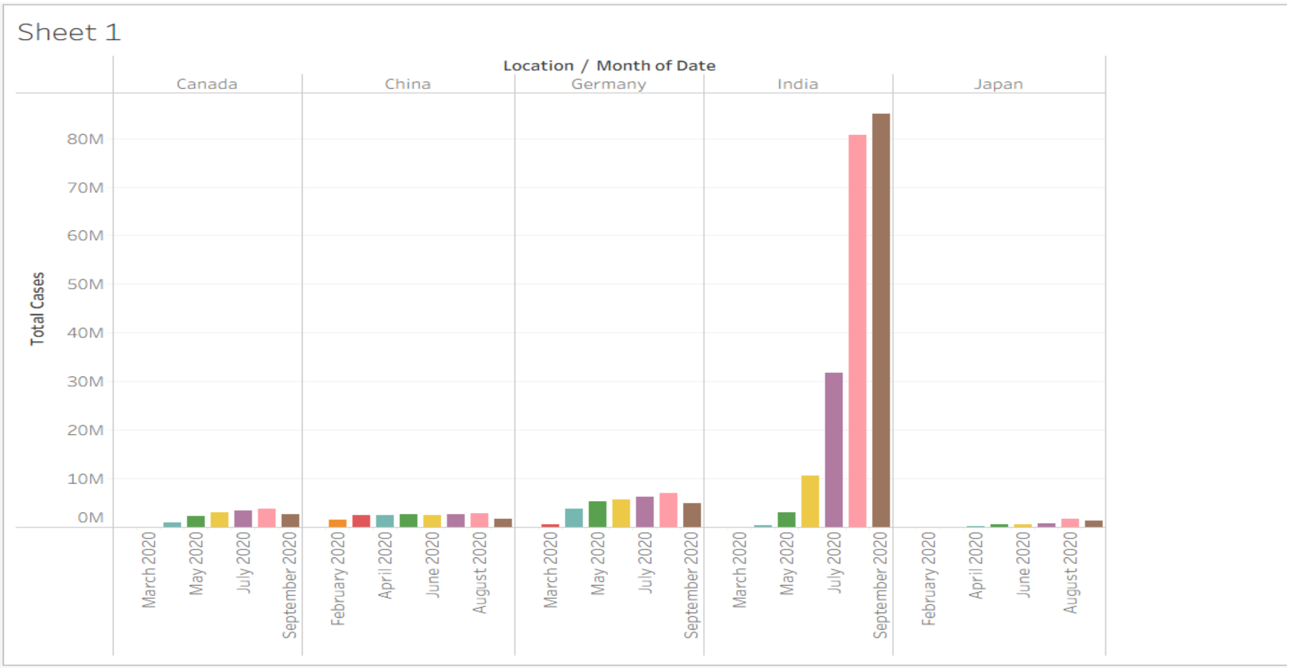
8️⃣ Format and Enhance Visuals
• Use the *Marks* card to adjust colors, labels, tooltips, and shapes.
• Right-click → "Format" to change fonts, borders, and gridlines.
• Add chart titles using *Worksheet → Show Title*.
• Ensure consistent color themes and readable fonts across all visuals.

9️⃣ Create a Dashboard
• Click the "New Dashboard" icon at the bottom.
• Drag your worksheets into the dashboard layout area.
• Resize and arrange charts neatly.
• Add text boxes, images, or titles for better presentation.
• Use *Dashboard → Actions → Filter* to link slicers (filters) across multiple visuals.

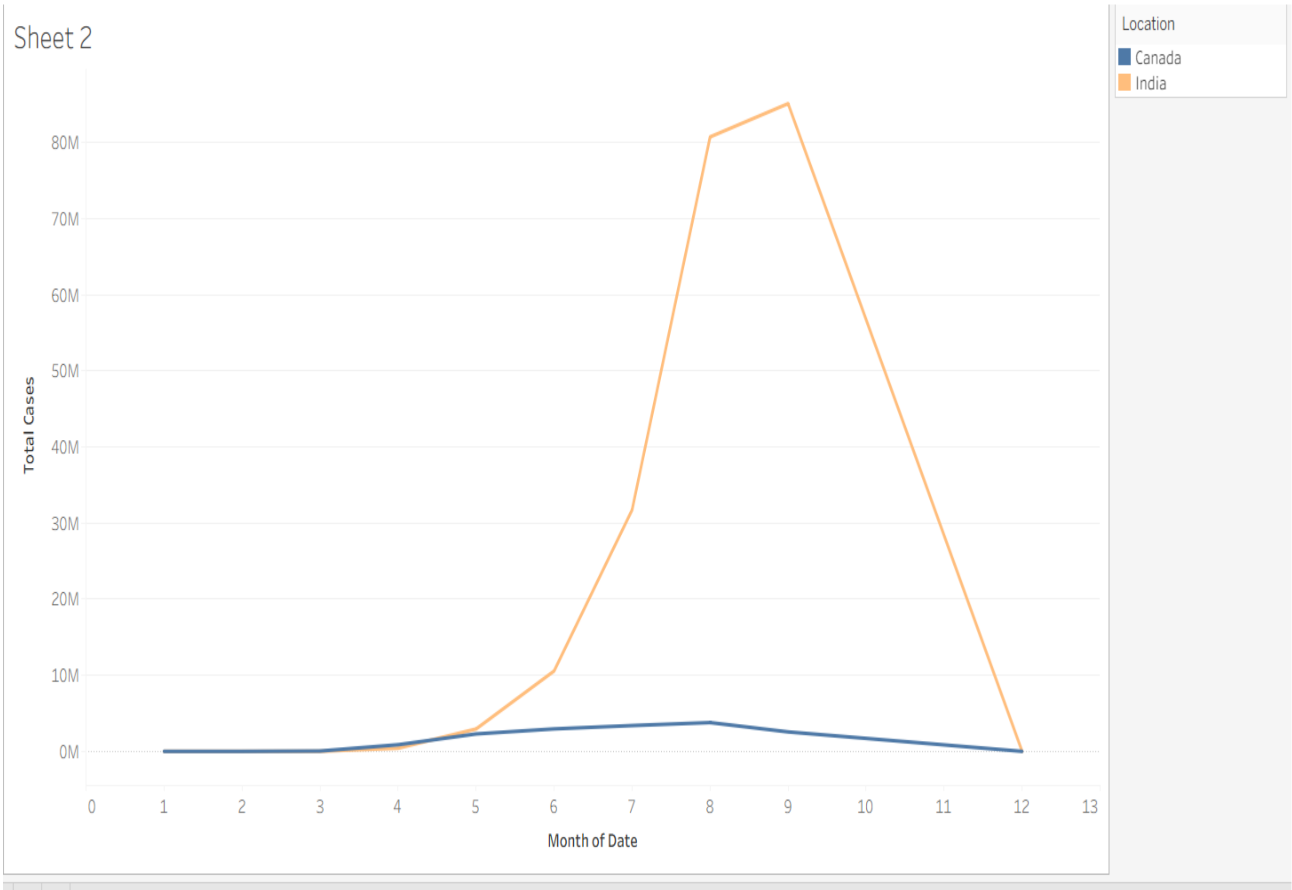
Save and Review the Report
• Go to File → Save As → choose location and file type (.twb or .twbx).
• Review all charts, filters, and interactions to ensure accuracy.
• (Optional) Publish your dashboard to Tableau Public or Tableau Server for sharing.

VISUALIZATION:-

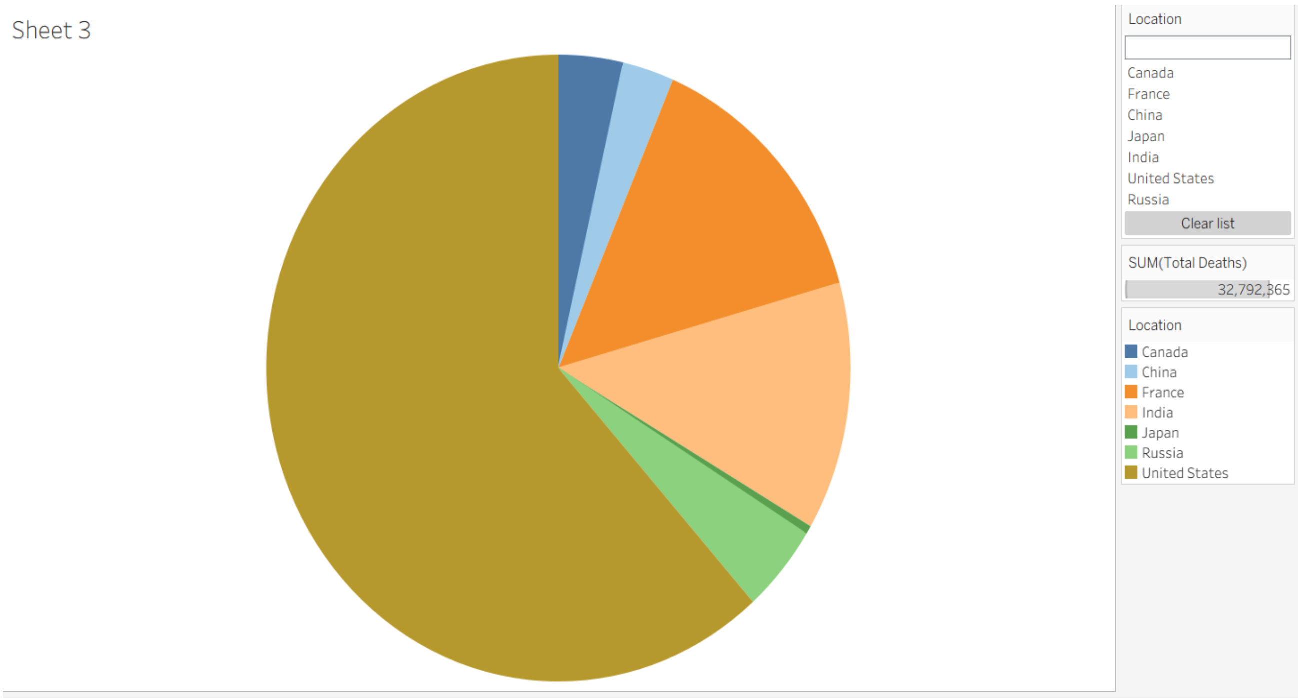
1) CLUSTERED COLUMN CHART



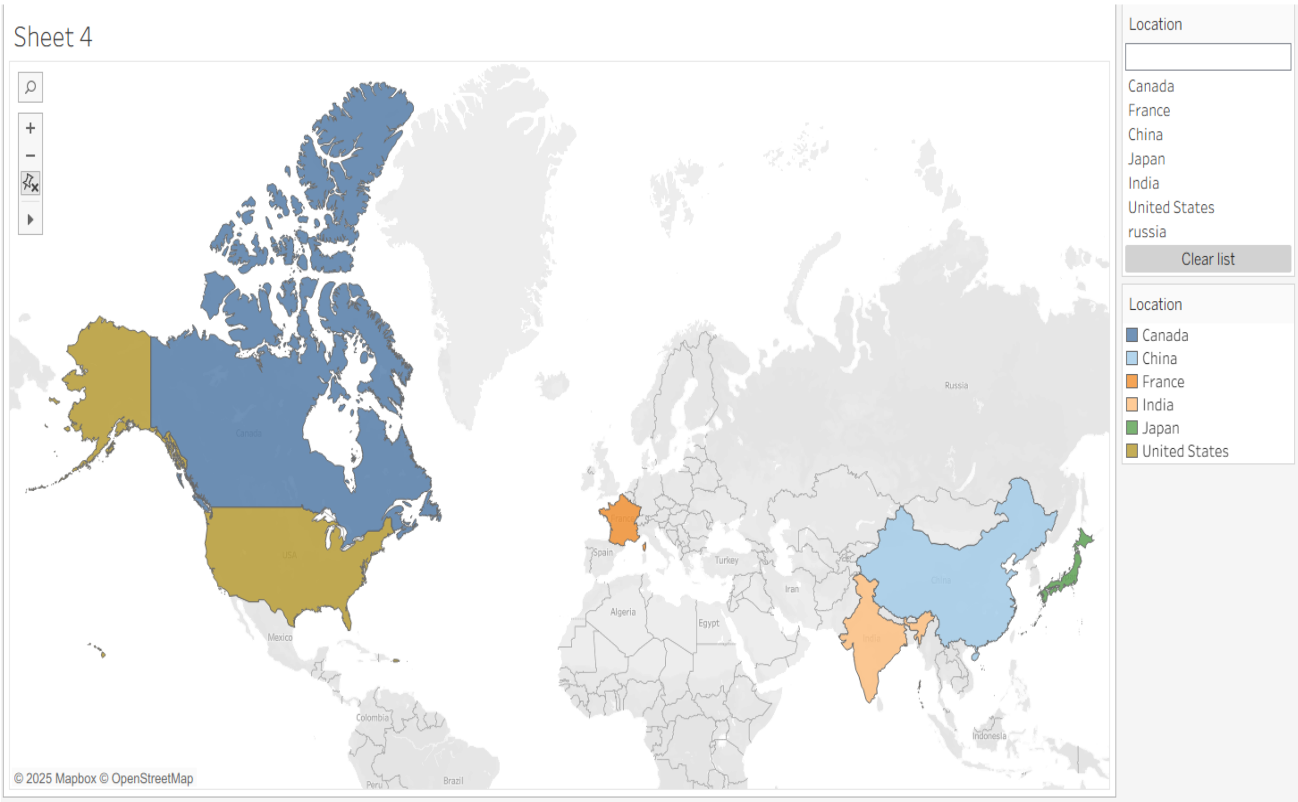
2) LINE CHART



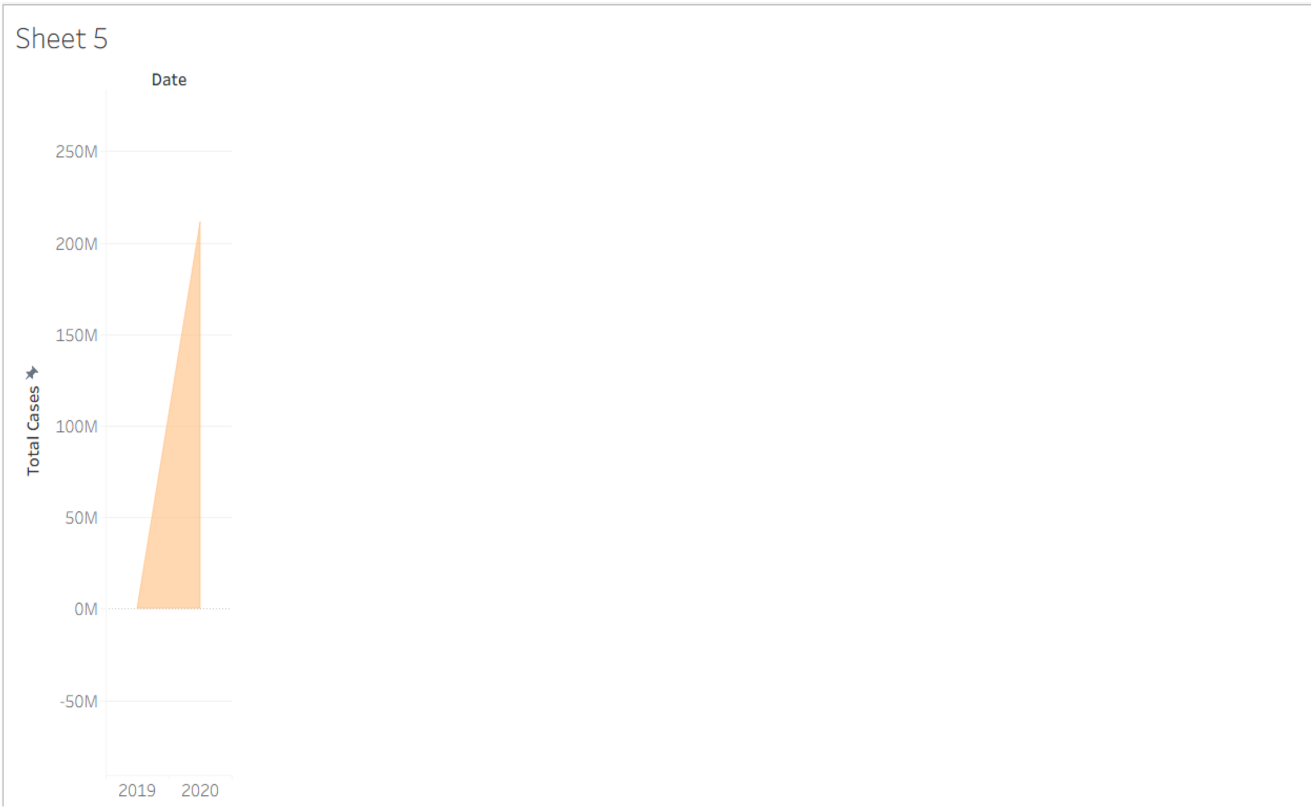
3) PIE CHART



4) MAP



5) AREA DISCRETE



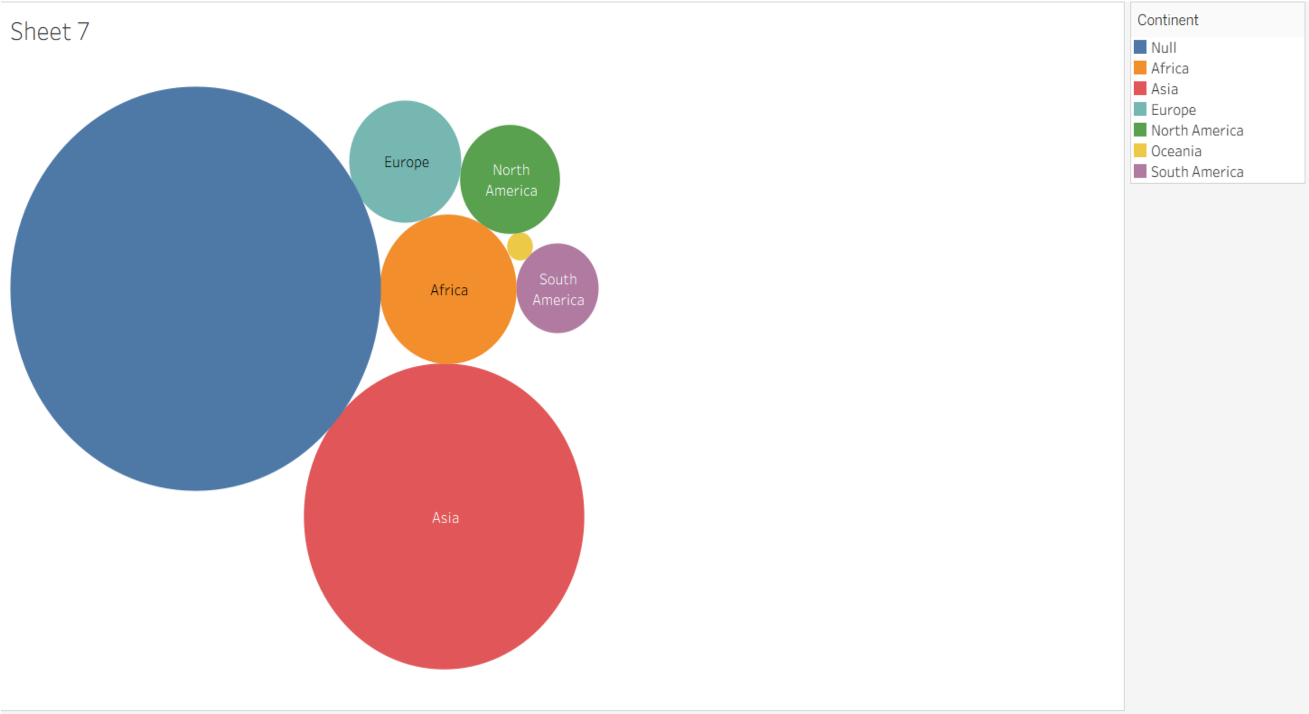
6) HIGHLIGHT TABLE

Sheet 6

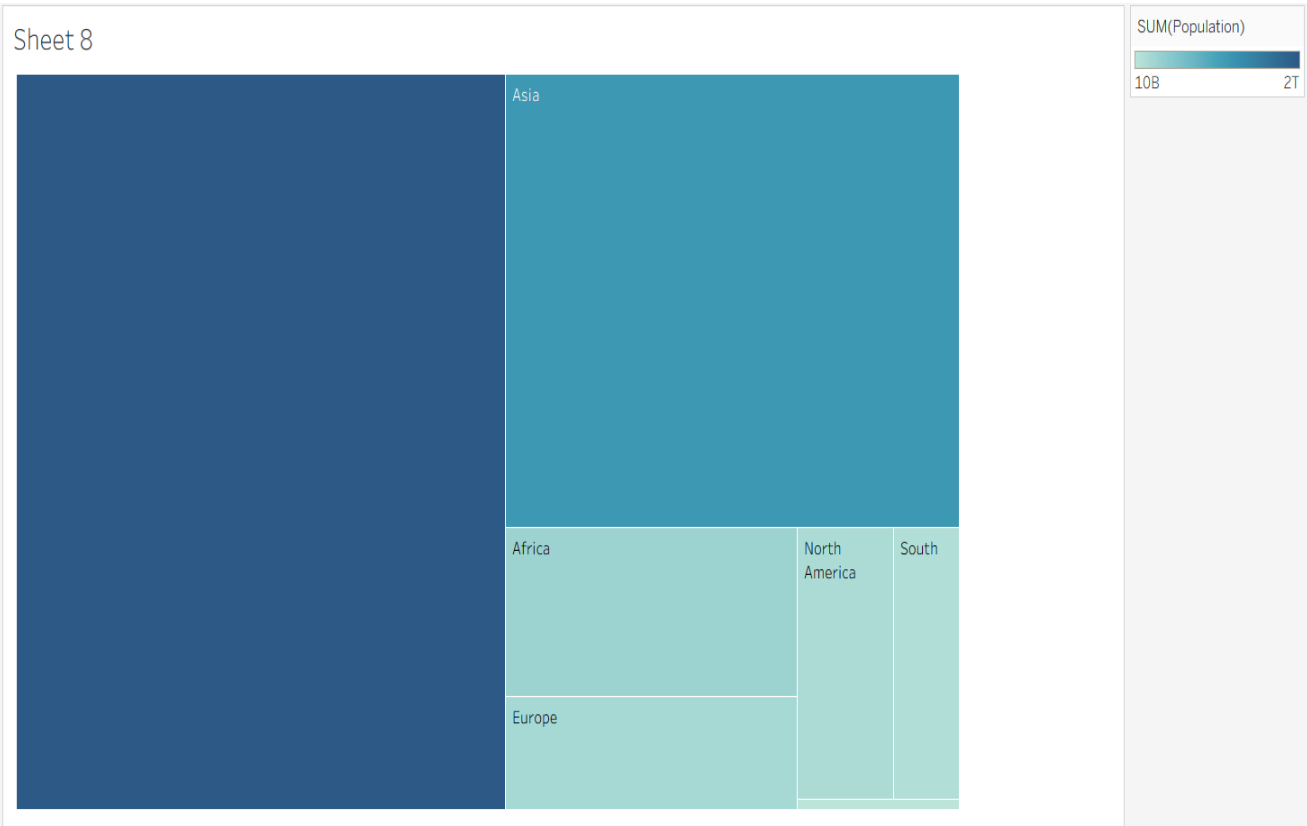
Continent	2019	2020
Null	27	2,070,876,499
Africa	0	90,226,617
Asia	27	517,161,506
Europe	0	398,597,068
North America	0	615,147,765
Oceania	0	2,546,632
South America	0	446,248,455

SUM(Total Cases): 0 – 2B

7) PACKED BUBBLES



8) TREE MAP



RESULT:-

Aspect	Observation
Peak Period	Global cases peaked around May 2021
Most Affected Country	USA, followed by India and Brazil
Vaccination Impact	A visible decline in new cases after large-scale vaccination
Mortality Rate	Fell below 2% globally after mid-2021
Data Insight	Vaccination is strongly correlated with reduced case growth and death rate

CONCLUSION:-

The COVID-19 Data Analytics project demonstrated how modern tools — **Power BI**, **Tableau**, and **Python (Colab)** — can be combined to extract meaningful insights from large-scale global health datasets.

Each platform contributed unique analytical strengths:

- **Power BI** provided an interactive dashboard where slicers, maps, and charts made exploration of COVID-19 cases, deaths, and demographic factors intuitive and dynamic.
- **Tableau** enabled powerful drag-and-drop visual storytelling, allowing pattern discovery through interactive filters, heatmaps, and trend analyses.
- **Python (Colab)** offered complete programmatic control for data cleaning, preprocessing, and correlation analysis using libraries like *pandas*, *matplotlib*, and *seaborn*, ensuring statistical validation of visual insights.

Through these combined approaches, it was observed that:

- Asia and Europe recorded the highest number of total cases and deaths.
- Population density, healthcare capacity, and economic strength (GDP per capita) significantly influenced infection and mortality rates.
- Regions with higher Human Development Index (HDI) managed the crisis more effectively, reflecting better healthcare readiness and policy responsiveness.
- Visual trends confirmed distinct pandemic waves globally, with variations based on continent and stringency measures.

Overall, this project highlights the importance of data analytics in public health decision-making.By leveraging multiple analytical frameworks, it provided a comprehensive, comparative, and validated understanding of how demographic, economic, and policy factors shaped the impact of COVID-19 worldwide.