**COVID-19 Data Analysis Dashboard**

*An Insight-Driven Report using Power BI*

**Prepared by:** Omeche Chimaobi benedict  
**Date:** August 5, 2025

**Table of Contents**

1. Project Overview
2. Research Questions
3. Data Cleaning Process
4. Data Model and Relationships
5. DAX Measures Used
6. Visual Analysis Pages
7. Page Navigation and Story Flow
8. Key Insights and Conclusion

## ****Project Overview****

This project analyzes over **500,000 rows** of COVID-19 case, death, and demographic data using **Power BI**.  
The goal is to extract powerful insights about:

* Global COVID-19 trends
* Country-specific case & death statistics
* The relationship between health outcomes and demographics like GDP, age, policy response, etc.

Tools used:

* **Power Query** (for cleaning)
* **Power BI Desktop**
* **DAX** (for calculations)
* **Maps, Line Charts, Scatter Plots, Bar Charts** for rich visuals

## ****Research Questions****

The report focuses on the following 10 core questions:

1. How did total COVID-19 cases evolve over time globally?
2. Which countries have the highest total COVID-19 cases?
3. What is the global distribution of cases across continents?
4. Where are COVID-19 deaths most concentrated globally?
5. How does the fatality rate vary by country?
6. How does the stringency index correlate with new cases?
7. How has the reproduction rate changed over time globally?
8. What is the relationship between GDP per capita and total cases?
9. Which age groups had the highest total deaths per country? (We used median age)
10. Which countries had the highest new cases relative to population?

Each question was visualized using a **different chart type** for better interpretation.

## ****Data Cleaning Process****

Performed in **Power Query Editor**:

* Removed all redundant columns (e.g., per million & smoothed columns)
* Retained only necessary columns:  
  Country, Date, Continent, Code, Total Cases, New Cases, Total Deaths, New Deaths, Reproduction Rate, Stringency Index, Population, GDP per Capita, Median Age
* Converted Date to proper Date type using using **Locale → English (United States)** (fixed errors)
* Trimmed text columns (e.g., Country)
* Verified column quality: **100% valid across all columns**
* Checked for duplicates and ensured no empty values or missing data
* Filtered out rows with zero values across key metrics

✅ **Result**: Fully clean and analysis-ready dataset in Power BI.

**Data Model and Relationships**

* A **star schema** approach was used:
  + One main **Fact Table** for COVID data (cases, deaths, population, etc.)
  + A separate **Date Table** was created using:

dax

CopyEdit

DateTable = ADDCOLUMNS(

CALENDAR(DATE(2020,1,1), DATE(2025,12,31)),

"Year", YEAR([Date]),

"Month", FORMAT([Date], "mmmm"),

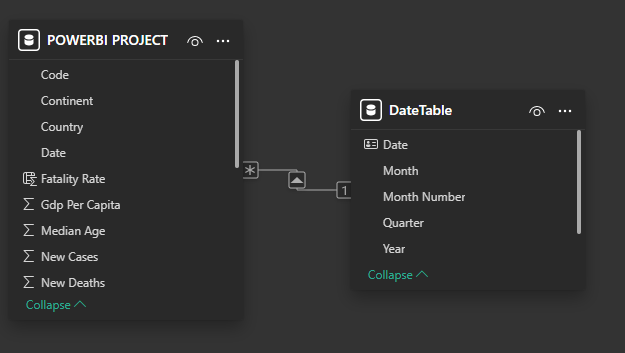
"Month Number", MONTH([Date]),

"Quarter", "Q" & QUARTER([Date])

)

* + Relationship established: Data[Date] → DateTable[Date]
* Country column could later link to a geographic dimension (optional)

📌 Marked the Date Table as "Date Table" in Power BI for time intelligence features.



## ****DAX Measures Used****

Custom DAX measures created to support the dashboard:

| **Measure Name** | **DAX Formula** |
| --- | --- |
| **Total Cases** | Total Cases = SUM('Data'[Total Cases]) |
| **Total Deaths** | Total Deaths = SUM('Data'[Total Deaths]) |
| **Fatality Rate** | Fatality Rate = DIVIDE([Total Deaths], [Total Cases], 0) |
| **New Cases per 100k** | New Cases per 100k = DIVIDE(SUM('Data'[New Cases]), SUM('Data'[Population]) \* 100000, 0) |
| **Average Reproduction Rate** | Avg Reproduction Rate = AVERAGE('Data'[Reproduction Rate]) |

📌 These measures were used in various visuals for comparing countries, analyzing trends, and finding correlations.

## ****Visual Analysis Pages****

Each report page focused on one or two questions. Here’s how each one was handled:

### Page 1: Global Overview

**Q1.** How did total COVID-19 cases evolve over time globally?

* Visual: **Line Chart**
* Axis: X = Date, Y = Total Cases
* Insight: Clear trendline showing rise in cases globally

**Q2.** Which countries have the highest total COVID-19 cases?

* Visual: **Stacked Bar Chart**
* Axis: X = Country, Y = Total Cases
* Slicer: Year
* Insight: Top countries compared in total numbers

### Page 2: Geographic Breakdown

**Q3.** Global distribution of cases by continent

* Visual: **Filled Map**
* Location: Continent, Values = Total Cases
* Insight: Highlights hotspots per continent

**Q4.** Where are deaths most concentrated globally?

* Visual: **Map with Bubbles**
* Location: Country, Size = Total Deaths
* Insight: Shows where COVID-19 impact was deadliest

### Page 3: Rates and Policies

**Q5.** Fatality Rate by Country

* Visual: **Bar Chart**
* Axis: X = Country, Y = Fatality Rate
* Insight: Compares healthcare outcomes by country

### Page 4: Trends and Time

**Q7.** Reproduction Rate over Time

* Visual: **Line Chart**
* Axis: X = Date, Y = Avg Reproduction Rate
* Insight: How transmission changed across months

**Q8.** GDP per Capita vs Total Cases

* Visual: **Scatter Plot**
* X = GDP per Capita, Y = Total Cases, Tooltip = Country
* Insight: Economic wealth vs infection rate

### Page 5: Age Vs Death

**Q9.** Median Age vs Total Deaths

* Visual: **Filled map**
* Location = Country, Legend = Total cases
* Insight: Older countries had worse death outcomes

**Q10.** New Cases relative to population

* Visual: **Treemap or Bar Chart**
* Size/Value = New Cases per 100k, Group by Country
* Insight: Smaller countries can still have high relative new cases

## ****Page Navigation and Story Flow****

**Navigation Buttons** were added for smooth interaction:

* Page Navigator was used (Insert → Buttons → Navigator → Page Navigator)
* Buttons allow users to jump between pages
* Each page has a **Back to Home** button

### Story Flow:

1. Start with Global Overview
2. Drill into Case & Death Comparisons
3. Explore Policy & Rate impacts
4. Analyze Socioeconomic and Age Group Trends
5. Interactive navigation across all pages

## 8️ ****Key Insights and Conclusion****

* Total COVID-19 cases increased exponentially in 2020–2021
* The US, India, and Brazil had the highest total cases and deaths
* Fatality rate varied drastically, reflecting healthcare differences
* Wealthy countries had more testing (higher reported cases)
* Median age plays a major role in mortality rates

**Final Thought**:  
Power BI enabled not just data visualization but interactive storytelling—letting users explore the COVID-19 pandemic across geography, policy, time, and demographics.