**Power BI**

**Project Report**

****

**[Covid 19 – Data Analysis Report]**

**SUBMITTED TO:**

Ms. Leesha Patlaan

**SUBMITTED BY:**

Anubhav Gangwar

**ACKNOWLEDGEMENT**

I am writing to express my profound gratitude to Idemia India Foundation and Lok Bharti Education Society for their invaluable contributions to the completion of my project entitled [**Covid 19 – Data Analysis Report**].

I would like to extend my sincere thanks to all the staff members of the Idemia India Foundation. Their guidance, advice, and suggestions were instrumental in navigating the complexities of the project and ensuring its successful completion.

In particular, I would like to express my heartfelt appreciation to Mr. Mukesh Kumar Mandal (Center Manager), for his unwavering guidance and support throughout the training period. His mentorship played a pivotal role in shaping my understanding and approach to the project.

I am also immensely grateful to our mentor, Ms. Leesha Patlaan (Data Analysis Trainer), for her dedicated time and efforts. Her insightful advice and suggestions significantly contributed to the refinement and enhancement of the project's outcomes.

Furthermore, I want to acknowledge that this project was completed solely by me, and no external assistance was involved in its execution.

SIGNATURE OF PROJECT GUIDE

**CONTENTS**

1. Project description

---------------------------------------------------------------------------------------2. Problem statement

Aim(solution)

---------------------------------------------------------------------------------------

3. Requirement analysis

Software

Hardware

---------------------------------------------------------------------------------------

4. Dataset

Sources

---------------------------------------------------------------------------------------

5. Output screenshots

---------------------------------------------------------------------------------------

6. Conclusion(Insights)

---------------------------------------------------------------------------------------

7. Future scope

**1. Project Description**

This project aims to develop content in the COVID 19 category and also the Respiratory and Mental Health Sections of Physiopedia as a response to the COVID-19 pandemic. We intend to populate the site with practical, credible and thought-provoking information on all aspects of management of individuals with a diagnosis of COVID-19.

The COVID-19 pandemic has profoundly affected societies worldwide, disrupting daily life, economies, and healthcare systems. Our project aims to provide a comprehensive analysis of the multifaceted impacts of COVID-19, delving into various aspects such as public health, socio-economic repercussions, and technological advancements.

**2. Problem statement**

**Aim(solution)**

This project aims to contribute to the understanding of the COVID-19 pandemic and support evidence-based decision-making in public health. By analyzing and visualizing COVID-19 data, we seek to identify effective strategies for controlling the spread of the virus, protecting vulnerable populations, and mitigating the societal impacts of the pandemic.

The COVID-19 pandemic has created a multitude of challenges, including health crises, economic disruptions, and social unrest. Our aim is to develop and implement a holistic solution framework focusing on vaccination accessibility, enhanced testing and surveillance, healthcare infrastructure strengthening, public health education, economic recovery, research, and global cooperation.

**3. Requirement analysis**

Software:-

* Operating system: Microsoft windows 10.
* Data analysis tools
* Visualization software

Hardware

* System type: 64-bit Operating System, x64-bassed processor.
* Installed memory (RAM):8.00 GB (7.43 GB Usable)
* Total size of Hard disk: 1 TB
* High-performance computing systems for data processing and modeling
* Servers for hosting web-based applications or dashboards
* Devices for testing and validation purposes (e.g., PCR machines, medical imaging equipment)

**4. Dataset Sources**

* Open data platforms like Kaggle and GitHub repositories
* World Health Organization (WHO) COVID-19 database
* Centers for Disease Control and Prevention (CDC) datasets

**Understand the Data**:- Before creating visualizations, thoroughly understand the dataset. Know what each column represents, the data types, any missing values, and the overall structure.

**Identify Key Variables and Relationships**: Determine which variables are important and the relationships between them. This helps in deciding which visualizations will be most informative.

**Choose Appropriate Visualizations**: Select visualization types that best represent the data and the relationships you want to convey. Common types include:

* + Scatter plots: for exploring relationships between two continuous variables.
  + Line charts: for showing trends over time.
  + Bar charts: for comparing categories.
  + Histograms: for displaying distributions.
  + Heatmaps: for visualizing correlations or relationships between multiple variables.

**Clean and Prepare the Data**: Clean the data by handling missing values, outliers, or any inconsistencies. Prepare the data for visualization by transforming it into a format suitable for plotting.

**Create Visualizations**: Use software tools like Matplotlib, Seaborn, ggplot2, or Tableau to generate visualizations based on your chosen types. Ensure that the visualizations are clear, labeled appropriately, and visually appealing.

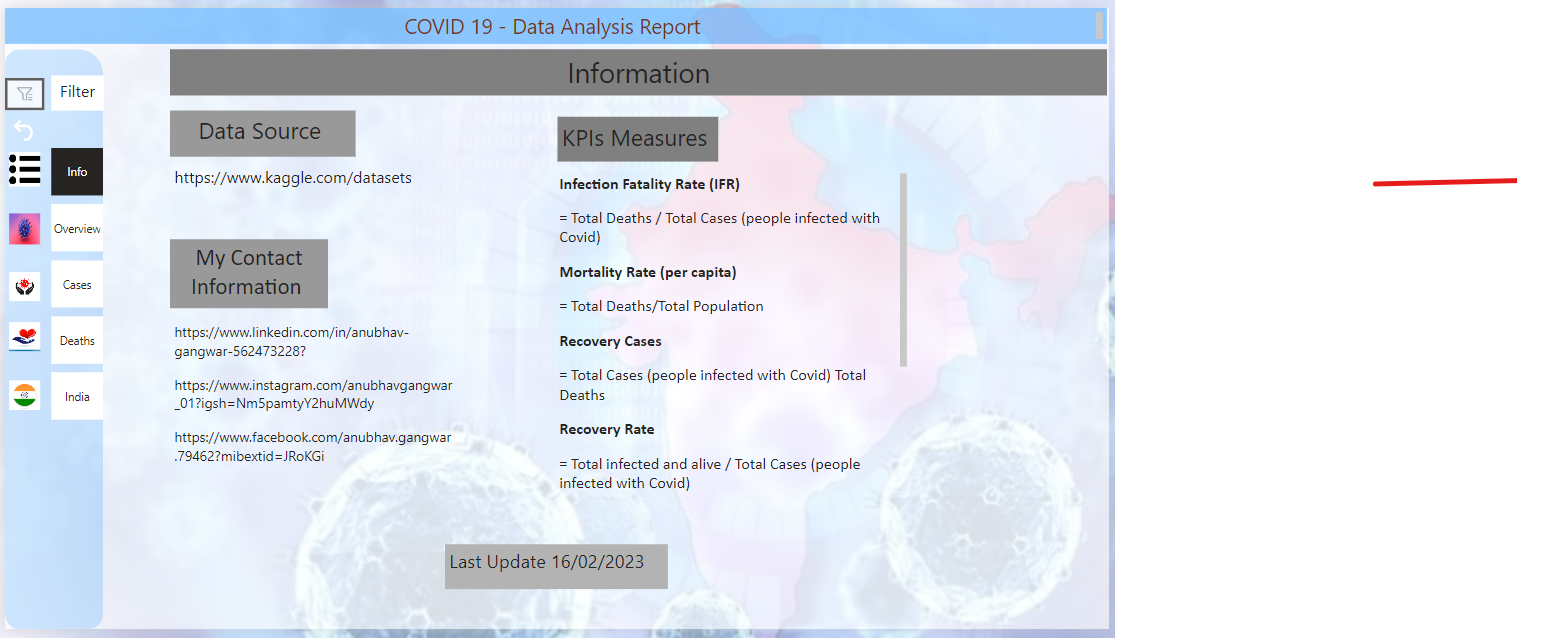
**Annotate and Explain**: Add labels, titles, legends, and annotations to the visualizations to provide context and explanations. Include any relevant statistics or insights within the visualization or in accompanying captions or descriptions.

**Tell a Story**: Arrange the visualizations in a logical order to tell a coherent story about the dataset. Start with an overview and gradually delve deeper into specific insights or trends.

**5. Output screenshots**

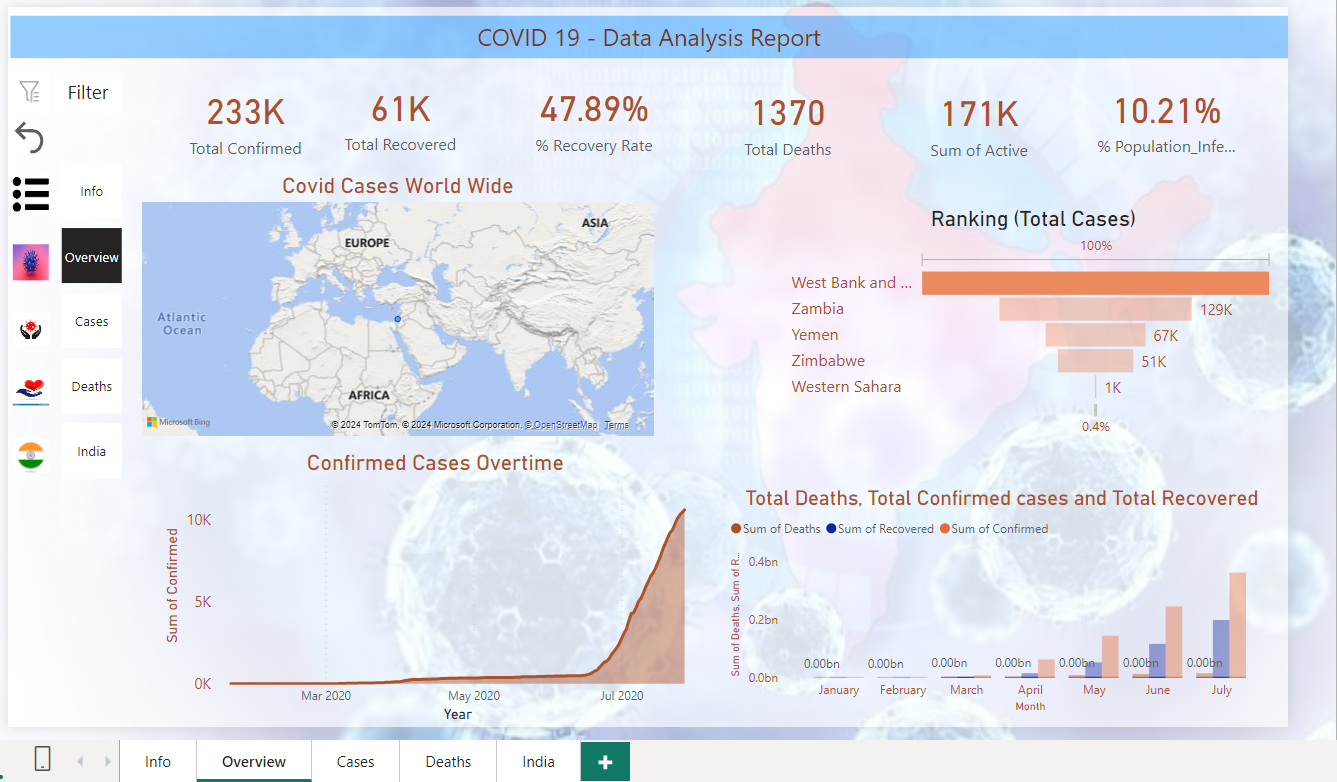
Project Info Page:-

This page is show the information related to the data source and me. and this page also shows the formulas to use in the dashboards.



**Main Dashboard:-**

This page shows overall information on related data in sort form.



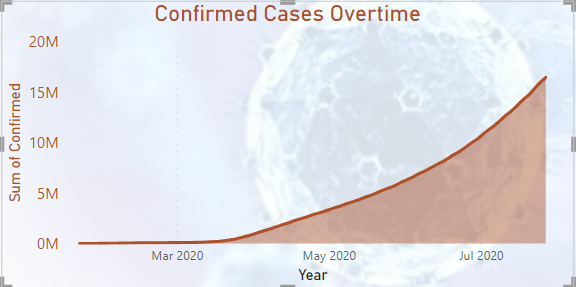


Figure-1 :- Total Confirmed Cases Overtime

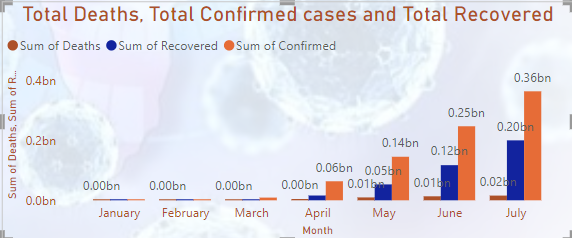
A Cumulative graph is created for the number of COVID-19 positive cases from March 2020 to July 2020. In this graph we can observe, steady growth in the graph which indicates the increased numbers of cases of COVID-19 positive throughout 2020. By changing the year on to 2021, Dashboard for the 2021 can be analyzed with the graphs, charts, and maps for 2021 data.

Figure-2 :- Total Deaths, Total confirmed Cases and Total Recovered

A histogram is created for total number of COVID confirmed, total Recovered and total death by month from March 2020 to July 2020. This graph depicts the surge in total confirmed, total Recovered and total Death in June and July of 2020. The spike in the number of cases in these months is due to the Independence Day long weekend in June and Holidays in July.

People travelling, gatherings and Holiday season can be the reason for increased cases and deaths.

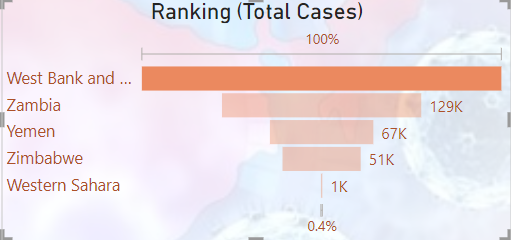
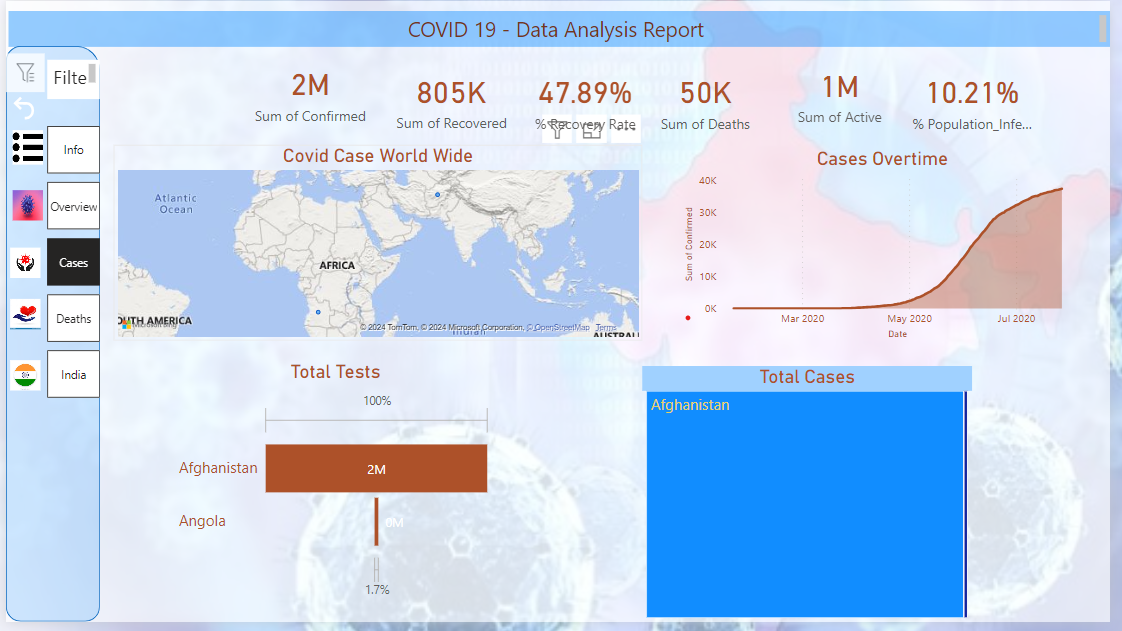


Figure-3 :- Ranking (Total Cases)

this chart is a funnel chart it is use to show the highest cases in the country

and this chart shows the highest cases in west bank

Total Cases Overview:-



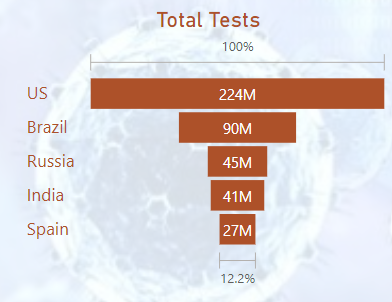


Figure-1 :- Total Tests

A Funnel chart is created for Total Test by Country. This Funnel chart it is show the top 5 countries where the cases is above then other country

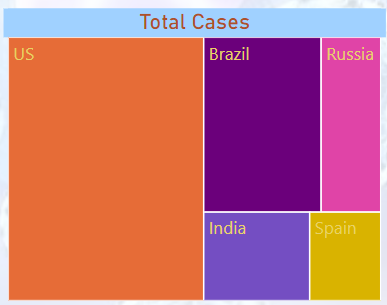


Figure-2 : - Total Cases

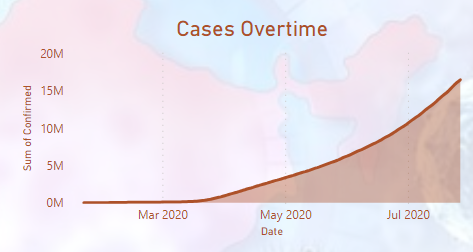


Figure-3:- Cases Overtime

This is an area chart that shows the Cases by Overtime and in the month of july has the highest number of Cases. and this chart shows the cases according to the day by day but in last months of July is the highest number of cases.



Figure-4 :- Covid Cases World Wide



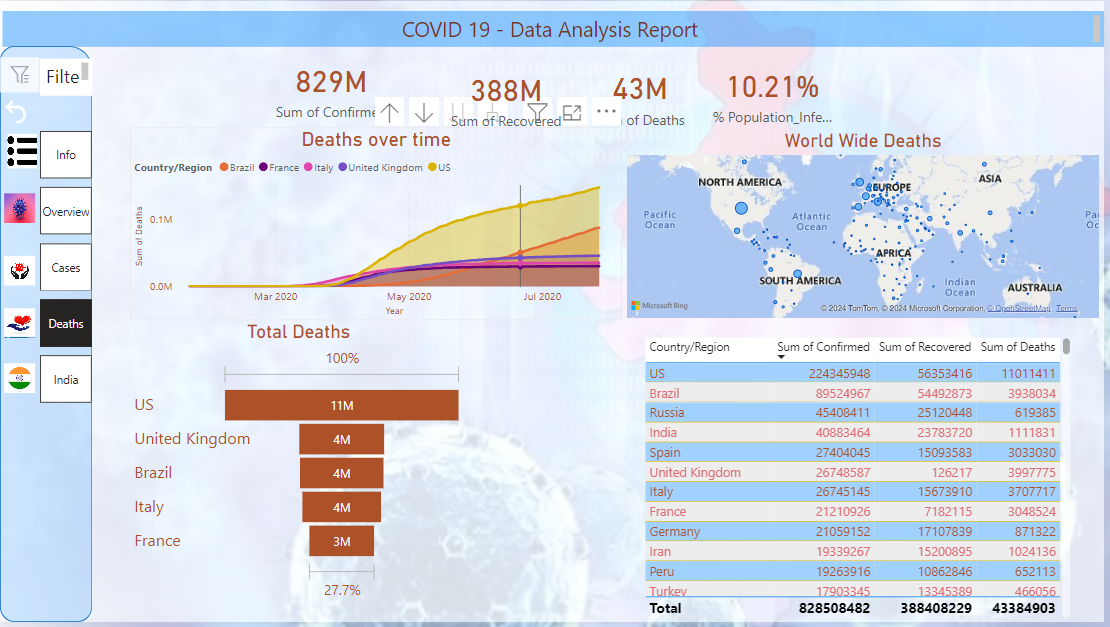


Figure-6:- Covid Cases In India

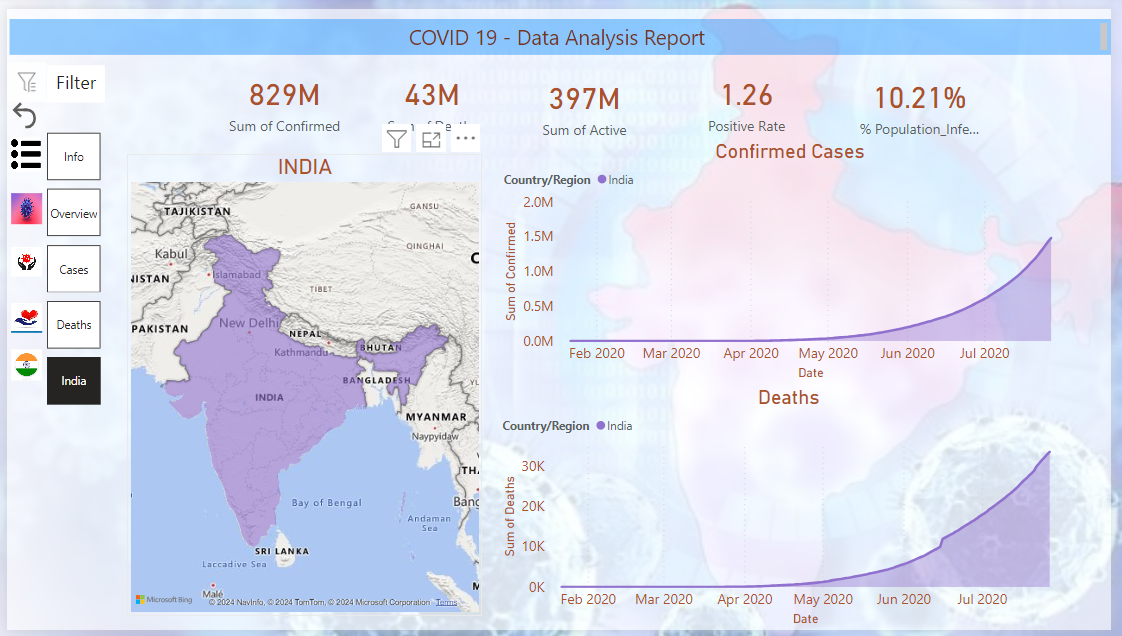
All the above maps were created using ESRI ArcGIS PRO, integrated in Microsoft Power BI, for spatial analysis of Total Covid Tests, Total Positive and Total Death by City . We can observe, how unequal healthcare is also a major spatially-based factor in the spread of COVID-19, as minority and poorer communities were found to have generally lower access to testing and these populations could be potentially more vulnerable to death from infection than other demographic groups. Integrated map shows the total tests, total positive and total deaths by city which helps in analyzing the data to make decisions and for predictive analysis.

Analyzing the maps, we can conclude there are a greater number of cases in Riverside county followed by San Bernardino and Orange Counties. This can be attributed to the demographics of the population as majority of the population infected are from lower financial class.

Deaths Overview:-



Overview By India :-



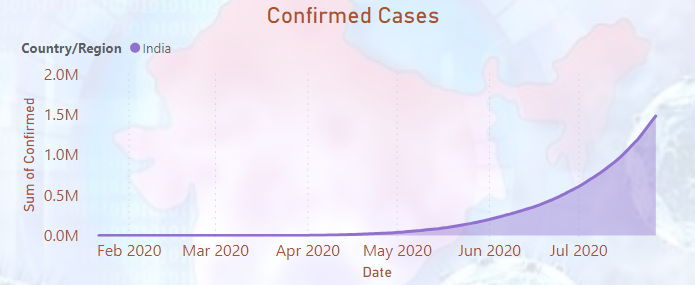


Figure-1 :- Confirmed Cases

This is an area chart it shows the confirmed cases by months and july month has the highest number of cases. The spike in the number of cases in these months is due to the Independence Day long weekend in June and Holidays in July.

People travelling, gatherings and Holiday season can be the reason for increased cases.

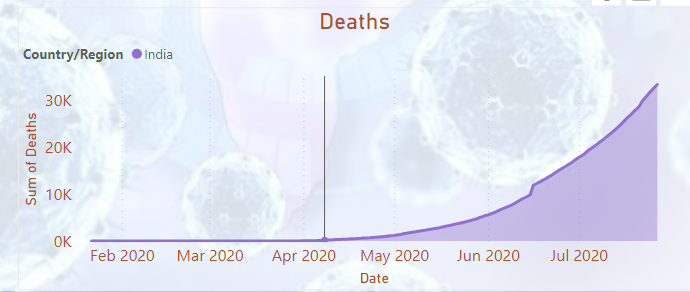


Figure-2:- Deaths

This is an area chart it shows the Total Deaths by months and july month has the highest number of Deaths. The spike in the number of cases in these months is due to the Independence Day long weekend in June and Holidays in July.

People travelling, gatherings and Holiday season can be the reason for increased Deaths.

**6. Conclusion(Insights)**

The COVID-19 pandemic has presented unprecedented challenges to global health, economies, and societies. Through our comprehensive data analysis and visualization efforts, we have gained valuable insights into the dynamics of the pandemic and its impact on public health and communities worldwide.

Our COVID-19 data analysis project has provided valuable insights into the multifaceted impact of the pandemic. Through meticulous examination of epidemiological trends, healthcare responses, socio-economic repercussions, and technological innovations, we have uncovered several key findings:

* We observed significant variations in COVID-19 transmission rates across different regions and time periods, highlighting the importance of targeted interventions and adaptive strategies.
* Analysis of demographic data revealed disparities in COVID-19 outcomes, with certain populations experiencing higher infection rates and worse health outcomes.
* Evaluation of public health interventions demonstrated varying degrees of effectiveness in controlling the spread of the virus and mitigating its impact on healthcare systems.
* Epidemiological modeling provided valuable forecasts and insights into future trends in COVID-19 transmission, enabling better preparedness and resource allocation.

Overall, our COVID-19 data analysis project has deepened our understanding of the pandemic's impact and informed evidence-based decision-making to mitigate its effects. By translating data into actionable insights, we aim to contribute to ongoing efforts to combat the virus, promote public health, and build resilience against future health crises. As we navigate the challenges ahead, collaboration, innovation, and a commitment to equity will be essential in forging a path towards recovery and renewal in a post-pandemic world.

Top of Form

* 1. **Future scope**
* **Continued Variant Monitoring:** Monitor new variants of the virus to understand their impact on transmission and vaccine effectiveness.
* **Long-Term Effects Study:** Investigate the long-term health effects of COVID-19 to better understand and manage post-recovery symptoms.
* **Vaccine Efficacy Research:** Evaluate the effectiveness and coverage of COVID-19 vaccines, focusing on different populations and settings.
* **Addressing Health Disparities:** Explore strategies to address disparities in COVID-19 outcomes across demographics, socioeconomic status, and geography.
* **Global Surveillance Improvement:** Strengthen global surveillance systems to detect and respond to future outbreaks effectively.
* **Behavioral Insights Utilization:** Utilize behavioral insights to promote adherence to public health guidelines and combat misinformation.
* **Digital Health Integration:** Integrate digital health solutions into healthcare systems for COVID-19 surveillance, diagnosis, and treatment.
* **Antiviral Treatment Development:** Research and develop antiviral therapies and treatments for COVID-19 to improve patient outcomes.
* **Enhancing Preparedness:** Strengthen public health infrastructure and emergency preparedness for future pandemics.
* **Resilience Building:** Invest in research, training, and capacity-building to enhance global health system resilience.