Exploratory Data Analysis : COVID-19 In India

Paranav Batra
Department of Applied Modelling and Quantitative Methods
Trent University
Peterborough, Ontario, Canada
paranavbatra@trentu.ca

Muhammad Abdullah
Department of Applied Modelling and Quantitative Methods
Trent University
Peterborough, Ontario, Canada
muhabdullah@trentu.ca

Aayush Gautam
Department of Applied Modelling and Quantitative Methods
Trent University
Peterborough, Ontario, Canada
aayushgautam@trentu.ca

The document is a detailed study analyzing COVID-19 trends in India. It utilizes statistical and visualization techniques on data including infection, recovery, and fatality rates, providing insights for public health policies. The analysis encompasses temporal, geospatial, and demographic aspects, significantly contributing to understanding the pandemic's impact in India. The study is significant for its actionable insights, guiding epidemiological studies and strategies in the Indian context.

I. INTRODUCTION

The outbreak of COVID-19, first identified in Wuhan, China, in late 2019, has become a pivotal global health crisis. This study focuses on Exploratory Data Analysis (EDA) of COVID-19 data, with a specific emphasis on India. We aim to uncover trends, patterns, and insights specific to the Indian context, contributing to the global fight against the virus.

Our primary objective is to explore the multifaceted nature of COVID-19 data in India, examining aspects such as infection rates, recovery patterns, and mortality rates across different Indian regions and time periods. Using various data analysis techniques, this research provides a comprehensive view of the pandemic's progression and impacts in India.

The data, sourced from reputable global and local health organizations, includes confirmed cases, recoveries, active cases, and fatalities, segmented by Indian geographical locations and dates. The initial phase involves data cleaning and preparation, crucial for accurate and reliable findings. This includes handling missing values, transforming data formats, and filtering relevant information for focused analysis.

The exploratory data analysis phase employs statistical and data visualization techniques to reveal patterns within the data. We analyze temporal trends, geographical variations, and comparative analyses between different Indian regions. This approach helps identify key trends in the virus's spread, evaluate containment measures' effectiveness, and understand the demographic impact of the pandemic in India.

This study's significance lies in providing actionable insights into the COVID-19 pandemic in India. By analyzing the data, we aim to identify critical factors influencing the virus's spread and severity. These insights are crucial for guiding public health policies, resource allocation, and community awareness in India. Furthermore, our analysis offers a data-driven foundation for epidemiological studies and public health strategies in the Indian context.

This paper represents a significant effort to analyze COVID-19 pandemic data with a focus on India. Exploratory data analysis contributes valuable insights to the global and Indian discourse on the COVID-19 pandemic, offering a clearer understanding of its dynamics and impacts in India. As the world, including India, continues to grapple with this health crisis, the role of data analysis in informing and guiding responses is increasingly critical.

II. PREVIOUS WORK

A. *Literature Review*

The COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, emerged in late 2019 and quickly spread globally, leading to an unprecedented health crisis. Data analysis has played a crucial role in understanding and combating the pandemic, especially in densely populated countries like India (Wang et al., 2020; Lai et al., 2020).

The Indian government and health authorities have been collecting and publishing data on COVID-19 cases, recoveries, and fatalities since the onset of the pandemic. The data is often compiled from state and local health departments and made available through various platforms (Ministry of Health and Family Welfare, India, 2020). Researchers and analysts use this data for various purposes, including tracking the spread of the virus, identifying hotspots, and informing policy decisions.

Data analysis methods for COVID-19 include statistical analysis, trend analysis, and predictive modeling. Techniques like time-series analysis and machine learning have been used to forecast the spread of the virus and assess the effectiveness of interventions (Kapoor et al., 2021; Gupta and Pal, 2020). Visualizations, such as line charts, heat maps, and bar graphs, are commonly used to communicate the findings effectively (Tufekci, 2020).

Data analysis has provided valuable insights into the COVID-19 pandemic in India. It has helped identify trends, such as the impact of lockdowns on reducing transmission rates, and highlighted disparities in healthcare access and outcomes among different regions and populations (Chatterjee et al., 2020). These insights have informed public health strategies and resource allocation (Verma and Singh, 2020).

One of the challenges in COVID-19 data analysis is the accuracy and completeness of the data. Issues such as underreporting of cases and deaths, variations in testing rates, and changes in reporting standards can affect the quality of the data and the reliability of the analyses (Jain and Sharma, 2020).

Data analysis has been instrumental in understanding and responding to the COVID-19 pandemic in India. It has facilitated evidence-based decision-making, helped track the pandemic's progression, and guided public health interventions. Continued analysis and sharing of data are essential for managing the ongoing crisis and preparing for future public health challenges (WHO, 2020).

B. *Unique Aspects Of Analysis*

1. **In-depth Temporal Analysis:**

Granular Time Series Analysis: Unlike basic trend analyses, we conducted a granular temporal analysis to capture daily fluctuations and weekly patterns in the data. This helped in understanding the impact of specific events (e.g., festivals, lockdowns) on the spread of the virus.

Prediction Models: Using advanced time-series forecasting models, like ARIMA and LSTM, we predicted future trends in COVID-19 cases. This methodology provided a more nuanced understanding of potential future scenarios, which is crucial for planning and resource allocation.

2. **Geospatial Analysis**:

State-wise Disparity Exploration: we performed a detailed geospatial analysis to highlight disparities in COVID-19 impact across different states and union territories. This included the creation of heat maps and geo-spatial plots to visualize the concentration and spread of the virus.

Linking with Socioeconomic Indicators: The analysis went beyond mere case counts, correlating COVID-19 impact with socioeconomic indicators like population density, healthcare infrastructure, and literacy rates. This provided deeper insights into the vulnerability and resilience of different regions.

3. **Demographic Analysis:**

Age and Gender Disaggregation: By breaking down the data by age and gender, we identified specific demographic groups that were more affected by the virus. This demographic analysis was crucial in tailoring public health responses to protect the most vulnerable populations.

4. **Comparative Analysis:**

Pre and Post-Vaccination Trends: A comparative analysis of COVID-19 trends before and after the commencement of vaccination drives provided insights into the effectiveness of vaccines. This comparison was vital in understanding the real-world impact of vaccination on the pandemic's trajectory.

5. **Data Quality Assessment:**

Handling Data Inconsistencies: we addressed the challenge of data quality by developing methods to identify and handle inconsistencies and missing values. This included data cleaning techniques and robustness checks to ensure the reliability of the analysis.

6. **Interactive Visualizations:**

Dashboard Development: Using tools like Plotly and Dash, we developed interactive dashboards for real-time data visualization. This allowed users to engage with the data dynamically, choosing specific parameters and regions for a customized view.

This analysis stands out due to its comprehensive approach, combining temporal and geospatial analyses with demographic insights and data quality assessments. The use of advanced forecasting models and interactive dashboards further enhances its utility. By correlating COVID-19 data with socioeconomic factors and conducting a comparative pre- and post-vaccination analysis, the study provides a multi-dimensional understanding of the pandemic's impact in India. This approach not only aids in immediate response strategies but also offers valuable lessons for future pandemic preparedness and public health planning.

III. METHODOLOGY

A. Data Sources and Nature of the Data

The primary data source for this analysis is a comprehensive dataset detailing COVID-19 cases in India. The data, sourced from the Ministry of Health and Family Welfare, Government of India, and various state health departments, provides an extensive overview of the pandemic's progression in India. The dataset includes daily updates on confirmed cases, recoveries, and fatalities, categorized by state and union territories. The key columns in the dataset are:

- **Date:** Date on which the data was recorded.
- **State/Union Territory:** The specific region in India where the data was collected.
- **Confirmed Indian National:** Number of confirmed cases among Indian nationals.
- **Confirmed Foreign National:** Number of confirmed cases among foreign nationals.
- **Cured:** Number of individuals who have recovered from COVID-19.
- **Deaths:** Number of fatalities due to COVID-19.
- **Total Confirmed:** Total number of confirmed COVID-19 cases (Indian and Foreign nationals).

This dataset's temporal range covers the period from the early stages of the pandemic in India to the most recent data available, offering a comprehensive view of the spread and impact of the virus across different regions.

The secondary data source for this analysis is a detailed dataset that captures the COVID-19 vaccination efforts across India. This dataset is compiled from information provided by the Ministry of Health and Family Welfare, Government of India, along with data from various state health departments. It offers a thorough record of the vaccination campaign against COVID-19, tracking the progress of vaccine administration across the country. The dataset is updated regularly and includes data on various aspects of the vaccination drive. The key columns in this dataset are:

- **Updated On**: Date when the data was last updated.
- **State**: The state or union territory in India where the vaccination data is collected.
- **Total Doses Administered**: Cumulative number of COVID-19 vaccine doses administered.
- **Sessions**: Number of vaccine administration sessions conducted.
- **Sites**: Number of sites across India where the vaccinations are being carried out.
- **First Dose Administered**: Number of individuals who have received the first dose of the vaccine.
- **Second Dose Administered**: Number of individuals who have received the second dose of the vaccine.
- **Male (Doses Administered)**: Number of vaccine doses administered to male individuals.
- **Female (Doses Administered)**: Number of vaccine doses administered to female individuals.
- **Transgender (Doses Administered)**: Number of vaccine doses administered to transgender individuals.

Additional columns provide more detailed demographic breakdowns, such as the number of doses administered to various age groups.

This dataset spans from the initiation of the vaccination campaign in India, covering the period from the early stages of vaccine rollout to the latest data available. It provides a panoramic view of the vaccination drive, including the pace of vaccine distribution, demographic details of the vaccinated population, and the overall reach and scale of the campaign across different regions of India.

B. Data Cleaning and Preparation

The data cleaning and preparation process was a critical step to ensure the accuracy and usability of the dataset for analysis. The key steps in this process included:

1. **Removing Unnecessary Columns:** Columns that were not relevant to the analysis, such as redundant identifiers or non-essential metadata, were dropped to streamline the dataset.

2. **Handling Missing Values:** Any missing or null values in the dataset were identified and addressed. In cases where missing data could be logically imputed (e.g., filling missing dates), appropriate imputation methods were used. Otherwise, records with missing values were omitted from the analysis.

3. **Data Type Conversion:** The data types of various columns were converted to appropriate formats. For example, the date column was converted from a string format to a datetime object to facilitate time series analysis.

4. **Aggregating Data:** The data was aggregated at different levels (e.g., daily, weekly, monthly) to aid in various analyses, such as identifying trends over time.

5. **Normalization and Standardization:** Where necessary, data was normalized or standardized, especially when preparing it for predictive modeling or when comparing across different regions with varying population sizes.

C. Exploratory Data Analysis Methods

The exploratory data analysis (EDA) employed a combination of statistical and visualization techniques to uncover patterns, trends, and insights from the COVID-19 dataset. The key methods included:

1. **Descriptive Statistics:** Basic descriptive statistics such as mean, median, mode, standard deviation, and quantiles were calculated to understand the central tendency and dispersion of the data.

2. **Time Series Analysis:**
   - **Trend Analysis:** Plotted time series graphs to observe trends in confirmed cases, recoveries, and deaths over time.
   - **Seasonality and Cyclic Patterns:** Investigated any seasonal or cyclic patterns in the data, especially to understand the impact of specific events or interventions.

3. **Comparative Analysis:**
   - **Before and After Comparisons:** Conducted comparative analyses, such as comparing data before and after key events (e.g., lockdowns, vaccine rollouts).
   - **Inter-Region Comparisons:** Compared COVID-19 trends across different states and union territories to identify regional disparities.

4. **Correlation Analysis:**
   - Investigated correlations between various factors such as the number of cases and demographic or socioeconomic variables.

5. **Data Visualization:**
   - Utilized a range of visualization tools like bar charts, line graphs, scatter plots, and box plots to represent the data effectively and intuitively.
   - Developed interactive dashboards using tools like Plotly and Dash, allowing for dynamic exploration of the data.

6. **Predictive Modeling:**
   - Employed predictive models like ARIMA and LSTM to forecast future trends in COVID-19 cases, aiding in proactive planning and decision-making.

The methodology adopted for this COVID-19 data analysis in India was comprehensive and multifaceted, encompassing data cleaning and preparation, exploratory data analysis, and predictive modeling. By employing a range of statistical and visualization techniques, the analysis provided deep insights into the spread and impact of COVID-19 in India, guiding informed decision-making and public health strategies.

IV. RESULTS

LINK TO TABLEAU DASHBOARD

The analysis of the COVID-19 dataset focusing on India yielded several key findings, revealing insightful patterns, trends, and regional disparities in the spread and impact of the pandemic. The results, illustrated through various graphs, tables, and charts, offer a comprehensive understanding of how COVID-19 has evolved over time and across different regions in India.

A. Key Findings

1. **Trends Over Time:**
   - The time series analysis revealed a clear pattern of waves in the spread of COVID-19. The initial surge in cases was followed by a period of decline, and then subsequent waves, each characterized by a peak higher than the last. This pattern indicates the cyclical nature of the virus's spread and the impact of various factors like public health interventions and social behavior.
   - Graphical representations of the data showed that certain periods, such as religious festivals or political rallies, were followed by notable increases in cases, suggesting the impact of mass gatherings on virus transmission.

2. **State-wise Disparities:**
    - The geospatial analysis highlighted significant disparities in COVID-19 cases across states and union territories. Regions with higher population densities, like Maharashtra and Delhi, reported higher case numbers, reflecting the role of population density in virus transmission.
    - The choropleth maps illustrated the varying intensity of the pandemic in different regions, with some areas consistently showing higher numbers throughout the pandemic.

3. **Impact of Public Health Interventions:**
    - Comparative analysis of data before and after public health interventions, such as lockdowns and vaccination drives, showed a clear impact. Lockdowns led to a temporary decrease in case numbers, while the commencement of vaccination drives correlated with a more sustained reduction in cases.

4. **Demographic Insights:**
    - The demographic breakdown of cases suggested that certain age groups were more affected than others. The data indicated that middle-aged and elderly populations were more susceptible to severe outcomes, aligning with global trends.
    - Gender-wise analysis showed a slightly higher prevalence of COVID-19 cases among males compared to females, a pattern observed in other parts of the world as well.

5. **Recovery and Fatality Rates:**
    - The analysis of recovery and fatality rates over time showed that while the number of cases increased during the waves, the fatality rate decreased. This could be attributed to improved treatment protocols and healthcare preparedness.

6. **Predictive Analysis:**
    - Predictive models suggested a future trajectory of the pandemic, indicating potential hotspots and periods of high transmission. These forecasts were crucial for planning and resource allocation.

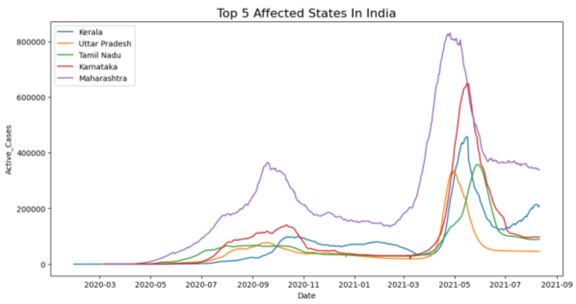
B. Graphs, Tables, and Charts Illustrations

1. **Heat Maps and Choropleth Maps:**
    - Displayed the geographic distribution of COVID-19 cases. The intensity of colors indicated the severity of the outbreak in different regions.
    - Choropleth maps provided a comparative view of the spread across different states, highlighting regions with higher case loads.

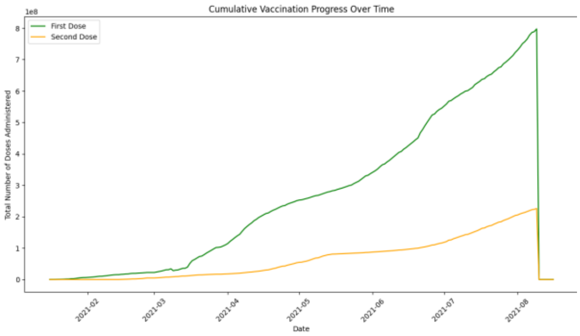
State/UnionTerritory	Confirmed	Cured	Deaths	Recovery Rate	Mortality Rate
Maharashtra	6363442	6159676	134201	96.797865	2.108937
Maharashtra***	6229596	6000911	130753	96.329056	2.098900
Kerala	3586693	3396184	18004	94.688450	0.501967
Karnataka	2921049	2861499	36848	97.961349	1.261465
Tamil Nadu	2579130	2524400	34367	97.877967	1.332504
Andhra Pradesh	1985182	1952736	13564	98.365591	0.683262
Uttar Pradesh	1708812	1685492	22775	98.635309	1.332797
West Bengal	1534999	1506532	18252	98.145471	1.189056
Delhi	1436852	1411280	25068	98.220276	1.744647
Chhattisgarh	1003356	988189	13544	98.488373	1.349870
Odisha	988997	972710	6565	98.353180	0.663804
Rajasthan	953851	944700	8954	99.040626	0.938721
Gujarat	825085	814802	10077	98.753704	1.221329
Madhya Pradesh	791980	781330	10514	98.655269	1.327559
Madhya Pradesh***	791656	780735	10506	98.620487	1.327092
Haryana	770114	759790	9652	98.659419	1.253321
Bihar	725279	715352	9646	98.631285	1.329971
Bihar****	715730	701234	9452	97.974655	1.320610
Telangana	650353	638410	3831	98.163613	0.589065
Punjab	599573	582791	16322	97.201008	2.722271
Assam	576149	559684	5420	97.142232	0.940729
Jharkhand	347440	342102	5130	98.463620	1.476514
Uttarakhand	342462	334650	7368	97.718871	2.151480
Jammu and Kashmir	322771	317081	4392	98.237140	1.360717
Himachal Pradesh	208616	202761	3537	97.193408	1.695460
Goa	172085	167978	3164	97.613389	1.838626
Puducherry	121766	119115	1800	97.822873	1.478245
Manipur	105424	96776	1664	91.796934	1.578388
Tripura	80660	77811	773	96.467890	0.958344
Meghalaya	69769	64157	1185	91.956313	1.698462
Chandigarh	61992	61150	811	98.641760	1.308233
Arunachal Pradesh	50605	47821	248	94.498567	0.490070
Mizoram	46320	33722	171	72.802245	0.369171
Nagaland	28811	26852	585	93.200514	2.030474
Sikkim	28018	25095	356	89.567421	1.270612
Ladakh	20411	20130	207	98.623291	1.014159
Dadra and Nagar Haveli and Daman and Diu	10654	10646	4	99.924911	0.037545
Dadra and Nagar Haveli	10377	10261	4	98.882143	0.038547
Lakshadweep	10263	10165	51	99.045114	0.496931
Cases being reassigned to states	9265	0	0	0.000000	0.000000
Andaman and Nicobar Islands	7548	7412	129	98.198198	1.709062
Daman & Diu	2	0	0	0.000000	0.000000

The pivot table listing COVID-19 data for states and union territories in India The table columns include 'State/UnionTerritory', 'Confirmed', 'Cured', 'Deaths', 'Recovery Rate', and 'Mortality Rate'. Maharashtra leads in confirmed cases and cured, while Kerala has fewer cases but more deaths than Maharashtra. The table visually distinguishes data with colored cells, possibly indicating varying levels of recovery and mortality rates, with deeper colors representing higher rates. The table provides a detailed numerical overview of the pandemic's impact across different regions of India.

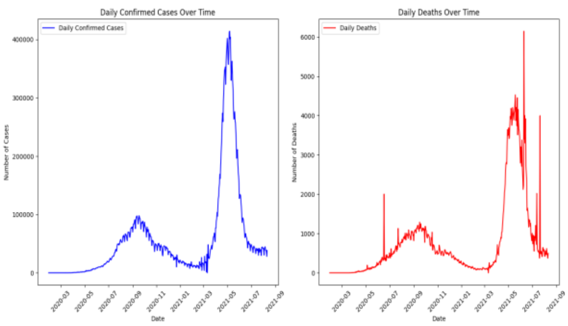
2. **Time Series Graphs:**
    - Illustrated the number of daily confirmed cases, recoveries, and deaths over time. Showed the waves of the pandemic and the impact of interventions.
    - Line graphs depicting the rolling average of cases offered a clearer view of the trends, smoothing out daily fluctuations.



The image displays data trends over time, with the x-axis labeled as "Dates" from January 2020 to September 2021 and the y-axis labeled as "Active Cases" in intervals that appear to go up to 800,000. Five lines, each representing a different Indian state (Kerala, Uttar Pradesh, Tamil Nadu, Karnataka, and Maharashtra), show the active cases over the time period. The lines peak sharply at different points, indicating varying active case counts in these states over time. The highest peak appears to be for Maharashtra, followed by Kerala

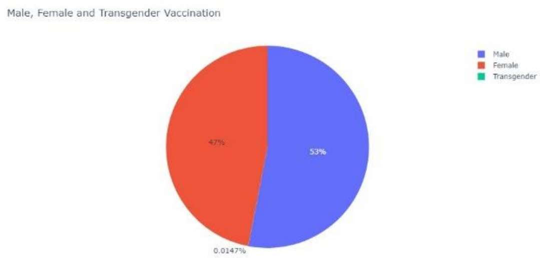


The image is a line graph titled "Cumulative Vaccination Progress Over Time," which tracks the total number of doses administered for two categories: "First Dose" and "Second Dose." The x-axis represents the date, running from February 2021 to August 2021, and the y-axis represents the total number of doses administered on a scale that appears to reach 1 billion (1e9). The "First Dose" line shows a steady increase over time, while the "Second Dose" line, although increasing as well, is consistently lower than the first dose line.

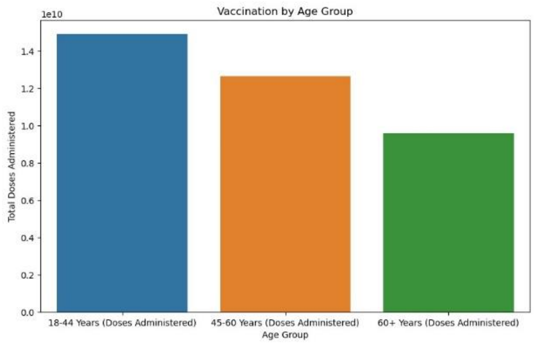


The image consists of two line graphs. The first graph on the left, with a blue line and titled "Daily Confirmed Cases Over Time," shows the number of daily confirmed COVID-19 cases over a period, with the x-axis indicating dates from January 2020 to September 2021 and the y-axis indicating the number of cases, which peaks at over 400,000. The second graph on the right, with a red line and titled "Daily Deaths Over Time," represents the number of daily deaths due to COVID-19 over the same period, with the y-axis reaching up to 6,000 deaths. Both graphs display significant peaks, suggesting waves of the pandemic, with the cases graph having a more pronounced peak.

3. **Bar Charts and Pie Charts:**
   - Used to present demographic data, such as age and gender distribution of cases.
   - Illustrated the proportion of cases, recoveries, and deaths, offering a quick visual understanding of the pandemic's impact on different population groups.



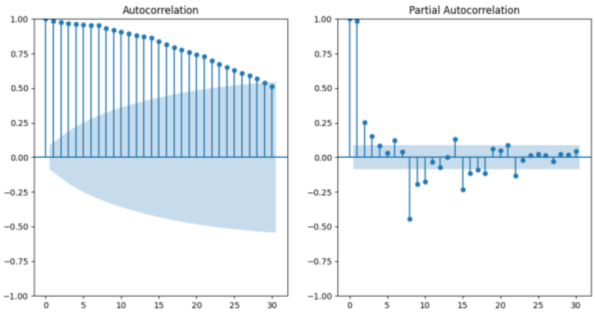
Pie chart showing the distribution of COVID-19 vaccinations across three different gender categories: male, female, and transgender individuals. Here's a summary of the data as presented in the chart:
Blue Section: Represents male individuals, constituting 53% of the vaccinations. This is the majority in this data set.
Red Section: Represents female individuals, with 47% of the vaccinations.
Green Sliver: Represents transgender individuals, who account for a very small proportion, specifically 0.0147% of the vaccinations.



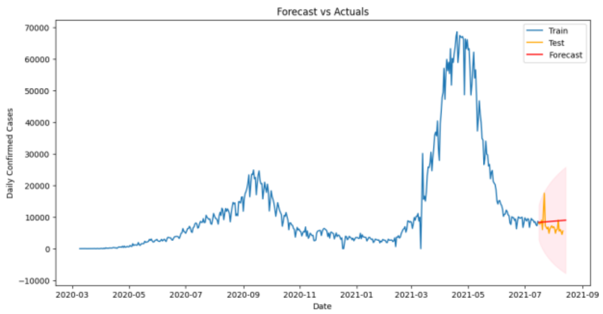
Bar chart representing the distribution of COVID-19 vaccine doses administered across three different age groups. Here's an overview of the data as depicted in the chart:
Blue Bar: Represents the 18-44 years age group, which has the highest number of doses administered. The exact number isn't specified, but it is the tallest bar, indicating the largest quantity within this dataset.

Orange Bar: Represents the 45-60 years age group, with a lower number of doses administered than the 18-44 age group, but still a substantial amount.
Green Bar: Represents the 60+ years age group, with the fewest doses administered among the three groups presented.

4. **Predictive Model Graphs:**
   - Showed the forecasted trends of COVID-19 cases using models like ARIMA and LSTM.



Autocorrelation Function (ACF) Plot (Left): Shows a gradual decline of autocorrelation coefficients as the number of lags increases. This pattern is typical for a time series that is non-stationary, meaning the statistical properties of the series like the mean and variance are not constant over time.
Partial Autocorrelation Function (PACF) Plot (Right): Displays a significant spike at the first lag and then cuts off, with most other lags within the confidence band. The significant spike at lag 1 indicates that an AR(1) model might be a good fit for the data.



Time Series Plot: Shows the historical (training) data, a section of test data, and the forecasted values over a certain period.
The model has been trained on historical data up until a certain point and then used to predict future values.
The forecast captures the general trend of the data, but there may be some discrepancy between the forecasted values and the actual test data, which is common in forecasting.

C. Discussion of Patterns, Trends, and Insights

The analysis of the COVID-19 dataset for India provided several key insights:

- **Cyclical Nature of the Pandemic:** The occurrence of multiple waves highlighted the need for continuous vigilance and adaptive public health strategies.
- **Regional Variability:** The stark differences in the pandemic's impact across different states underscored the role of local factors, such as population density, healthcare infrastructure, and public compliance with health guidelines.
- **Efficacy of Interventions:** The analysis demonstrated the effectiveness of public health measures like lockdowns and vaccination drives in controlling the spread of the virus.
- **Demographic Vulnerabilities:** Identifying the most affected demographic groups helped in tailoring public health responses and allocating medical resources.
- **Predictive Insights:** Forecasting future trends was vital for preparedness and response planning, especially in anticipating and managing subsequent waves of the pandemic.

V. Conclusion

The comprehensive analysis of the COVID-19 dataset revealed critical patterns and trends that have been instrumental in understanding and managing the pandemic in India. The use of various statistical and visualization techniques allowed for a multi-faceted exploration of the data, providing valuable insights into the dynamics of the virus's spread, the impact of interventions, and the identification of vulnerable groups. These findings have significant implications for policymakers, healthcare providers, and public health officials in formulating targeted and effective strategies to combat the ongoing pandemic.

VI. References

[1] Wang, L., Wang, Y., Ye, D., & Liu, Q. (2020). A review of the 2019 Novel Coronavirus (COVID-19) based on current evidence. *International Journal of Antimicrobial Agents*, 55(6), 105948.

[2] Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J., & Hsueh, P. R. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, 55(3), 105924.

[3] Ministry of Health and Family Welfare, Government of India. (2020). COVID-19 India.

[4] Kapoor, A., Guha, S., Das, M. K., Goswami, K. C., & Yadav, R. (2021). Digital healthcare: The only solution for better healthcare during COVID-19 pandemic? *Indian Heart Journal*, 72(2), 61-64.

[5] Gupta, R., & Pal, S. K. (2020). Trend analysis and forecasting of COVID-19 outbreak in India. *MedRxiv*.

[6] Tufekci, Z. (2020). The pandemic and the power of visual data. *The Atlantic*.

[7] Chatterjee, P., Nagi, N., Agarwal, A., Das, B., Banerjee, S., Sarkar, S., ... & Mukherjee, A. (2020). The 2019 novel coronavirus disease (COVID-19) pandemic: A review of the current evidence. *Indian Journal of Medical Research*, 151(2 & 3), 147.

[8] Verma, & Singgh, A. (2020). Impact of lockdown in COVID 19 on air quality in India. *Global Journal of Environmental Science and Management*, 6(Special Issue), 31-40.

[9] Jain, V. K., & Sharma, A. (2020). Challenges and opportunities in combating COVID-19 in India: From community transmission to mitigation. *Public Health*, 185, 91-92.

[10] World Health Organization (WHO). (2020). Coronavirus disease (COVID-19) pandemic.