# Covid-19 Statewise Vaccination Analysis

KLE Technological University | Creating Value, Leveraging Knowledge
Dr. M. S. Sheshgiri Campus, Belagavi

Team B-01

Department of Computer Science and Engineering

## Background

*Coronaviruses*, a diverse group of viruses. In humans, these viruses can cause a range of *respiratory infections*, from mild colds to more serious conditions like ***Middle East Respiratory Syndrome (MERS)*** and ***Severe Acute Respiratory Syndrome (SARS)***. The newest addition to this family of viruses is the ***one responsible for the COVID-19*** disease, as identified by the World Health Organization

## Problem statement

- ***Problem Statement:*** Analyze the distribution and administration of COVID-19 vaccines across different states in India
- ***Objectives:*** We aim to uncover regional disparities, assess vaccination rates, understand demographic trends, and predict future vaccine needs. This will provide actionable insights for vaccine deployment strategies and combat the pandemic effectively.

## Knowing The Dataset

- the dataset has 24 attributes [2 catagorical and 22 Numeric ]
- There are 7845 instances, ranges from 0 to 7844.
- Source URL : https://www.kaggle.com/datasets/sudalairajkumar/covid19-in-india (Appeared on kaggle)
- The data was recorded from 16th Jan 2021 - 12th Aug 2021
- Memory Usage = 1.4+ MB

## Hypothesis

- How is the distribution of vaccine types represented across all states based on the total doses administered?
- How is the distribution of Covaxin, CoviShield and Sputnik V vaccine doses administered across all states of India ?
- How does vaccination coverage vary across different age groups?
- How does vaccination coverage vary across different gender groups?
- What is the state-wise distribution of different age group individuals vaccinated?
- Which state has the highest of total individuals vaccinated?
- Is there a significant difference between the number of first doses and second doses administered within each state?
- Is there a significant difference between the Total Doses Administered and Total Individuals Vaccinated within each state?
- What is the loss of vaccine doses administered?
- How are Doses Administered by Gender and Age Group ?
- What is the Distribution of Doses Administered by Vaccine Type and Gender ?
- How many sessions were conducted in the whole vaccination process ?
- Is there any relation between adverse events following immunization (AEFI) and the total number of vaccine doses administered ?

## Data Preprocesssing

- **Handling Missing Values:** - Dropped 3 columns due to 97.96 percent missing values. - Dropped first 200 rows due to inconsistent data.
- **Dealing with Duplicates:** - No duplicate values found in the dataset.
- **Encoding Categorical Variables:** - Applied Label Encoding on "Updated On" and "State" columns.
- **Outlier Detection and Handling:** - Used box plot and Capping in IQR method to detect and cap outliers. Significant outliers in several columns left untreated.
- **Skewness and Normalization:** - Checked skewness of numerical features. - Normalization or scaling performed for certain analysis or modeling techniques.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

- **Correlation Analysis:** - Identified relationships between variables and potential multicollinearity.

## Exploratory Data Analytics

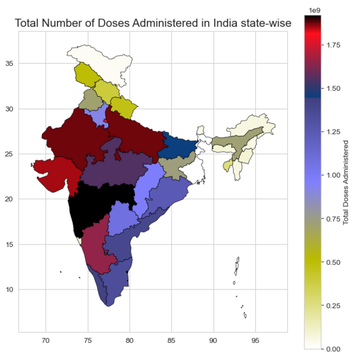


Figure 1. Total number of doses administrated state-wise in India

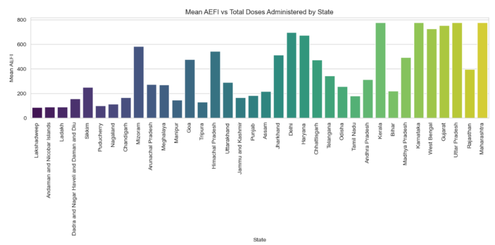


Figure 2. Mean AEFI compared with total doses administrated in all states

## Exploratory Data Analysis

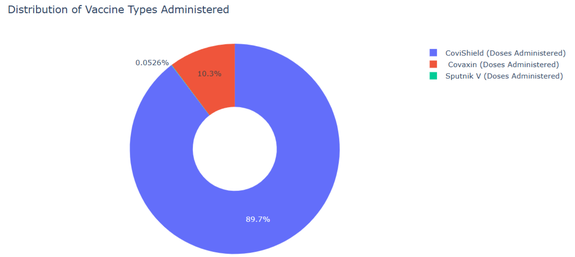


Figure 3. Distribution of vaccine types across total doses administrated

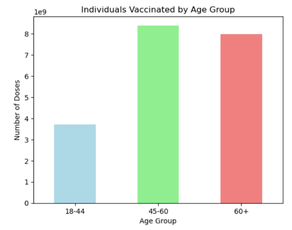


Figure 4. Individuals vaccinated by age Group

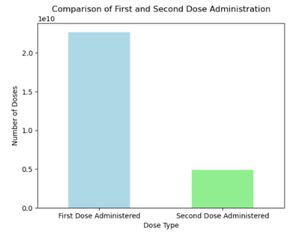


Figure 5. Comparison of first and second dose administration

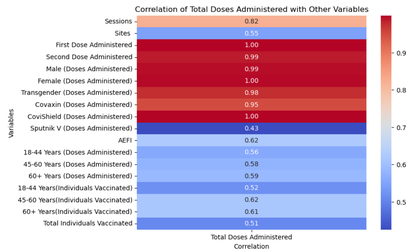


Figure 6. Total doses administrated is an important factor in predicting the trends

## Exploratory Data Analytics

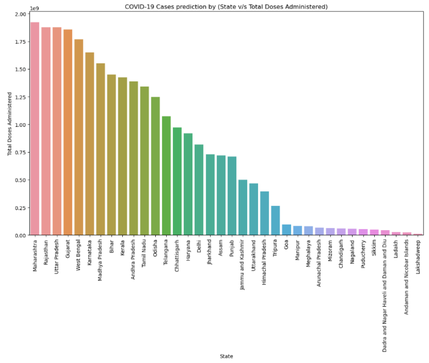


Figure 7. COVID-19 cases prediction by (State v/s Total doses administered)

## Conclusion

- CoviShield was the most administered vaccine, followed by Covaxin and Sputnik-v.
- Maharashtra had the highest number of vaccine doses administered, while some states had significantly lower vaccination rates.
- The age group 45-60 received the highest number of vaccine dose
- Males had a higher vaccination rate compared to females and transgenders
- A significant difference is observed between the number of first doses and second doses administered within each state.
- The study provided insights into vaccine wastage and suggested ways to optimize vaccine distribution to minimize loss

## References

[1] Jian Pei Jiawei Han, Micheline Kamber. Data mining : Concepts and techniques. *The Morgan Kaufmann Series in Data Management Systems*, 3, 2011.

## Under The Guidance of :

Dr. Prema T. Akkasaligar

## Team Members



| Name | SRN |
|---|---|
| Akhilesh Joshi | 02FE22BCS013 |
| DilipSingh Rajpurohit | 02FE22BCS035 |
| Sana Mulla | 02FE22BCS099 |
| Seeba Doddmani | 02FE22BCS116 |