KLE Technological University | Creating Value, Leveraging Knowledge

Dr. M. S. Sheshgiri Campus, Belagavi

Covid-19 Statewise vaccination Analysis

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**

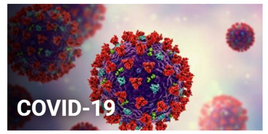

Figure 1. Coronavirus

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 to optimize vaccine deployment strategies and combat the pandemic effectively.

Knowing The Dataset

- the dataset has 24 attributes [2 catagorical and 1 numeric]
- There are 7845 instances, ranges from 0 to 7844.
- Source URL : DATASET (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 ?

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 analyses or modeling techniques.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

- **Correlation Analysis:** - Identified relationships between variables and potential multicollinearity.
  - **Data Visualization:** - Created various visualizations (Donut Chart,Bar plot,Stacked bar plot,Multiple histogram,Gauge plot,Choropleth map,Scatter plot) to understand the distribution and relationships between features.
  - **Ridge Regression:** - Applied Ridge Regression, a type of linear regression with a regularization term to prevent overfitting by penalizing large coefficients.

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\} \tag{2}$$

Exploratory Data Analysis

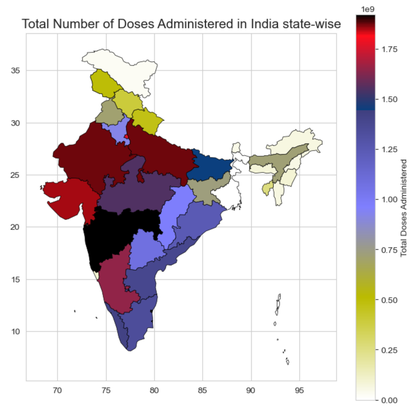

Figure 2. Coverage of vaccination process all over the regions of India

Exploratory Data Analysis

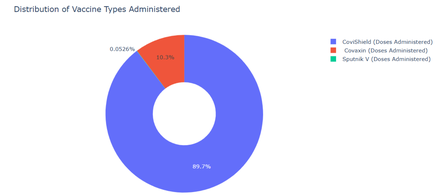

Figure 3. CoviShield Vaccine has been administered in more number - 89.7 percent

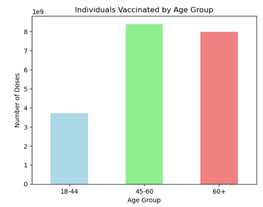

Figure 4. The Individuals of age group 45-60 have been vaccinated in good numbers

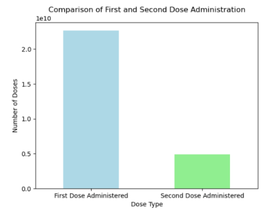

Figure 5. More than half of the people who were administrated with 1st dose , Didnt recieve the second dose

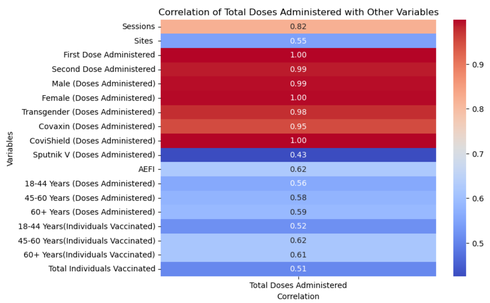

Figure 6. Total Doses Administrated is an important factor in predicting the trends in vaccination process

Exploratory Data Analytics

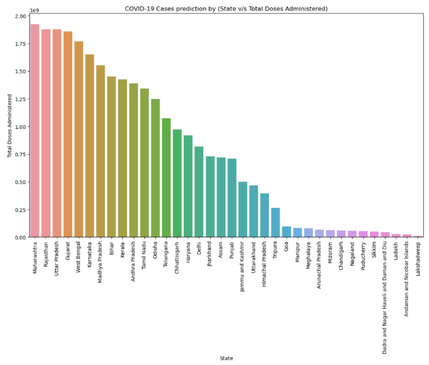

Figure 7. State with most COVID-19 Cases prediction by (State v/s Total Doses Administered)' is Maharashtra. due to high pouplation density and less care in precautions.

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 was 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

Donec quis posuere ligula. Nunc feugiat elit a mi malesuada consequat. Sed imperdiet augue ac nibh aliquet tristique. Aenean eu torto

Under The Guidance of :

Dr. Prema T. Akkasaligar

Team Members

Name	SRN
Akhilesh Joshi	02FE22BCS013
DilipSingh Rajpurohit	02FE22BCS035
Sana Mulla	02FE22BCS099
Seeba Dodmani	02FE22BCS116