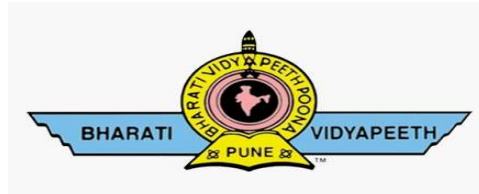**BHARATI VIDYAPEETH'S**

COLLEGE OF ENGINEERING FOR WOMEN

KATRAJ-DHANKAWADI, PUNE-43.

INFORMATION TECHNOLOGY DEPARTMENT
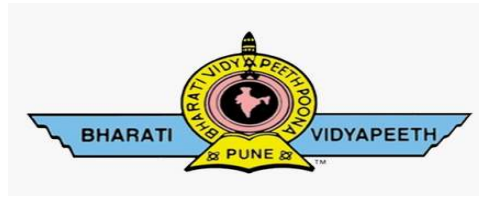
A PROJECT REPORT

ON

"Covid-19 Dataset Analysis"

Project Guided By

Prof. A.D. Khairkar

Submitted By

Miss. Srushti Mule (Roll no. 3342)

Miss. Apurva Nangare (Roll no. 3343)

Miss. Chaitali Nigade (Roll no. 3344)

Miss. Manasvi Pudale (Roll no. 3352)

Class: TE IT

Academic Year:2023-24

BHARATI VIDYAPEETH'S COLLEGE OF ENGINEERING FOR WOMEN

KATRAJ-DHANKAWADI, PUNE-43.

INFORMATION TECHNOLOGY DEPARTMENT

CERTIFICATE

This is to certify that the Project titled

"Covid-19 Dataset Analysis"

was prepared and presented by

Miss. Srushti Mule (Roll no. 3342)

Miss. Apurva Nangare (Roll no. 3343)

Miss. Chaitali Nigade (Roll no. 3344)

Miss. Manasvi Pudale (Roll no. 3352)

Class: TE IT

Under the guidance of

Prof. A. D. Khairkar

Towards partial fulfillment of

TE IT course under

Savitribai Phule Pune University

(Year 2023-24)

Prof. A.D. Khairkar   **Prof. Dr. D. A. Godse**

Subject Teacher   **HOD IT**

ACKNOWLEDGEMENT

To complete any type of project work is team work. It involves all the technical/non-technical expertise from various sources. The contribution from the experts in the form of knows-how and other technical supports is of its vital importance. We are indebted to our inspiring guide **Prof. A.D. Khairkar** who has extended all valuable guideline, help and constant encouragement through the various difficult stages for the onslaught of the project. We are also very thankful to our **HOD Prof. Dr. D. A. Godse, Prof. Dr. K. B. Naik** for their valuable guidance.

We have great pleasure in offering our honorable Principal **Prof. Dr. P. V. Jadhav.** Last but not least, we would like to thanks all the direct and indirect help provided by the staff and our entire classmate for successful completion of this project.

Date:-

Miss. Srushti Mule (Roll no. 3342)

Miss. Apurva Nangare(Roll no. 3343)

Miss. Chaitali Nigade (Roll no. 3344)

Miss. Manasvi Pudale (Roll no. 3352)

T.E. (I.T)

Covid-19 Dataset Analysis

ABSTRACT

This project presents a comprehensive **Analysis of COVID-19 Datasets**, focusing on confirmed cases, deaths, and recovered cases. Leveraging diverse datasets sourced from reputable repositories, our study employs advanced statistical techniques and machine learning algorithms to uncover key patterns and trends in the pandemic's progression.

Through temporal analysis, we identify critical periods in infection rates and potential contributing factors. Geospatial mapping reveals regional disparities in disease prevalence and healthcare resource allocation. Additionally, demographic profiling sheds light on differential susceptibility and outcomes among various population cohorts.

Furthermore, predictive modeling aids in forecasting disease trajectories, facilitating timely interventions and resource allocation strategies. Social media sentiment analysis offers insights into public perceptions, influencing adherence to preventive measures. Overall, our findings contribute empirical evidence and actionable insights to the discourse on COVID-19, informing public health policy formulation and future research directions.

Covid-19 Dataset Analysis

CONTENTS

Covid-19 Dataset Analysis

INTRODUCTION

The COVID-19 pandemic has posed unprecedented challenges to public health systems globally, necessitating rigorous data-driven approaches to understand its dynamics and inform effective response strategies. In this project, we undertake a comprehensive analysis of COVID-19 datasets, focusing on vaccinated patients, confirmed cases, deaths, and recovered patients. Additionally, we delve into area-wise and gender-wise analyses to gain deeper insights into the pandemic's impact, particularly within the context of India.

With the advent of vaccination campaigns worldwide, understanding the efficacy and impact of vaccines on COVID-19 patients is paramount. By scrutinizing data pertaining to vaccinated individuals, we aim to assess vaccine effectiveness, identify breakthrough infections, and discern any emerging trends or patterns among this subset of patients. Such analysis is crucial for optimizing vaccination strategies and informing public health policies aimed at curbing the spread of the virus.

Furthermore, our examination of confirmed cases, deaths, and recovered patients provides a holistic view of the pandemic's progression and its toll on affected populations. Through statistical analyses and visualization techniques, we seek to elucidate temporal trends, geographical hotspots, and demographic disparities in disease burden. This not only aids in resource allocation and healthcare planning but also informs targeted interventions to mitigate transmission and reduce mortality rates.

Moreover, our area-wise analysis zooms in on regional variations in COVID-19 incidence, healthcare infrastructure, and socio-economic factors that influence disease dynamics. By identifying high-risk areas and vulnerable populations, we can tailor intervention strategies to address localized needs effectively. Similarly, our gender-wise analysis sheds light on potential disparities in infection rates, severity of illness, and healthcare access, thus advocating for gender-sensitive approaches in pandemic response efforts.

Covid-19 Dataset Analysis

In essence, this project endeavors to harness the power of data analytics to unravel the complexities of the COVID-19 pandemic, offering valuable insights that can guide evidence-based decision-making and bolster public health resilience, particularly in the Indian context.

Covid-19 Dataset Analysis

DESCRIPTION OF DATASET

The dataset comprises 7840 rows and 10 columns, each representing a unique observation related to the COVID-19 pandemic within India. The dataset captures crucial information pertaining to various aspects of the pandemic, including epidemiological data, vaccination coverage, and demographic characteristics.

1. Area: This column delineates the geographical area or region within India where the observation was recorded, providing insights into regional variations in COVID-19 incidence and response efforts.

2. Date and Time: These columns indicate the date and time of each observation, facilitating temporal analysis to track the progression of the pandemic over time and identify trends or patterns.

3. State/Union Territories: This column specifies the administrative division within India where the observation was recorded, enabling analysis at the state or union territory level to understand regional disparities in disease burden and healthcare infrastructure.

4. Cured Cases, Deaths, Confirmed Cases: These columns document the number of individuals who have recovered from COVID-19, the number of fatalities attributed to the virus, and the total number of confirmed cases, respectively. These metrics are fundamental for assessing disease severity, mortality rates, and disease transmission dynamics.

5. Covaxin, Covishield, Sputnik: These columns represent the number of individuals vaccinated with specific COVID-19 vaccines, namely Covaxin, Covishield, and Sputnik. Tracking vaccine coverage is essential for evaluating vaccination campaigns' effectiveness and identifying vaccination disparities across regions.

6. Male, Female, Transgender: These columns provide demographic breakdowns of vaccinated individuals based on gender, facilitating gender-specific analyses to understand potential disparities in vaccine uptake and COVID-19 outcomes.
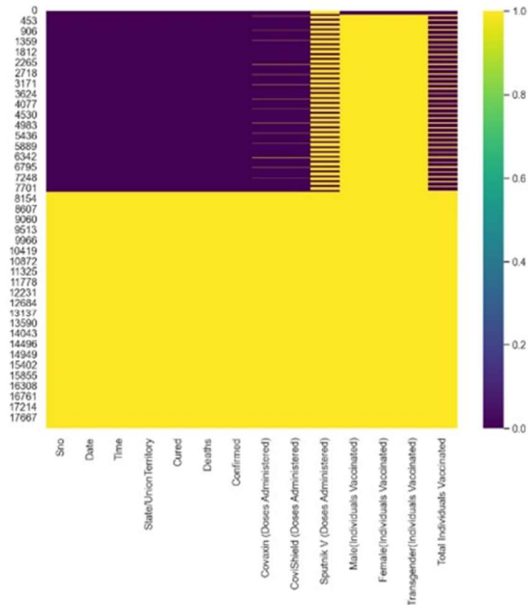
7. Total Individuals Vaccinated: This column aggregates the total number of individuals vaccinated, irrespective of the vaccine type administered. It serves as a key indicator of vaccination progress and coverage rates within the population.
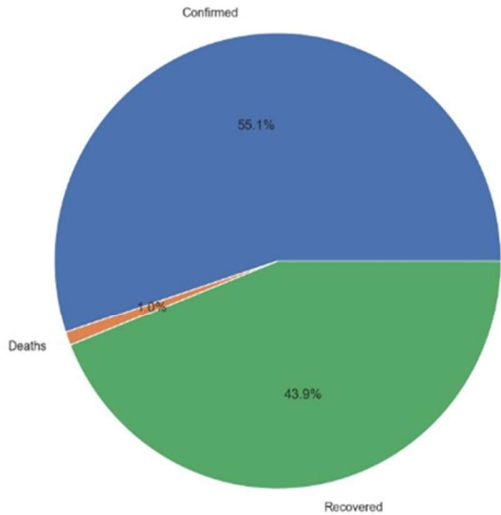
GUI SCREENSHOTS

COVID-19 Dashboard

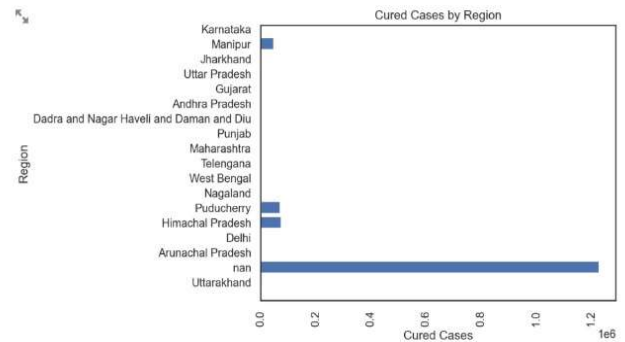## Missing Values Heatmap

Pie Chart: COVID-19 Data
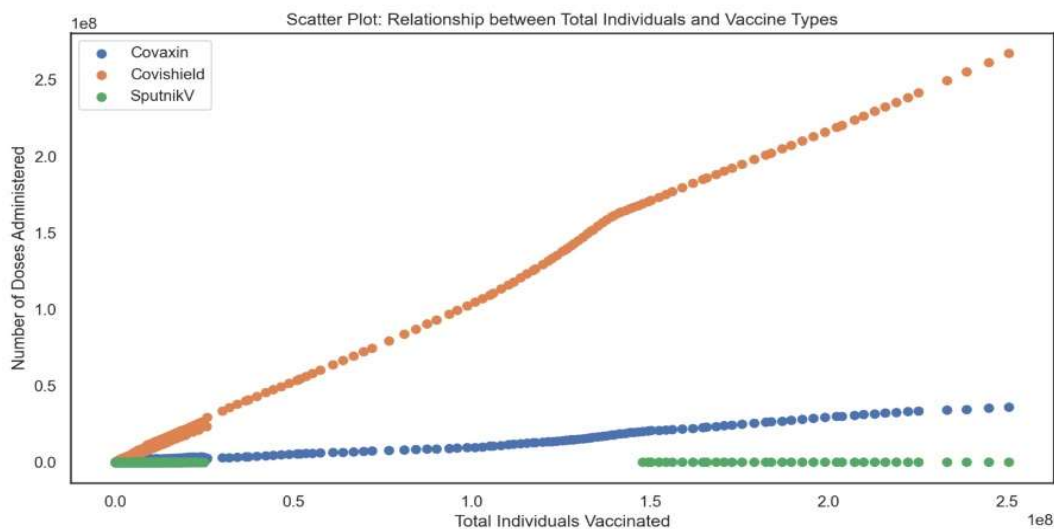
Covid-19 Dataset Analysis

Confirmed Cases by Region
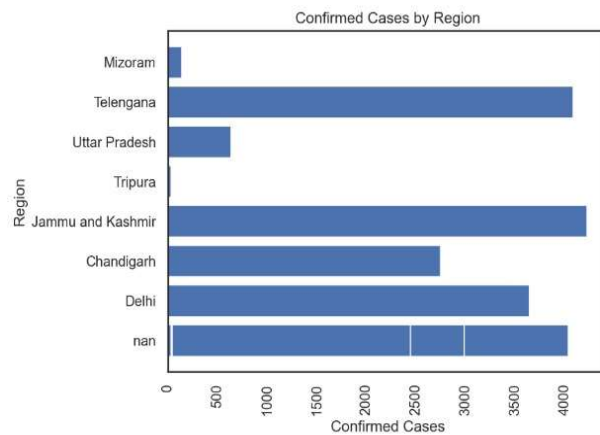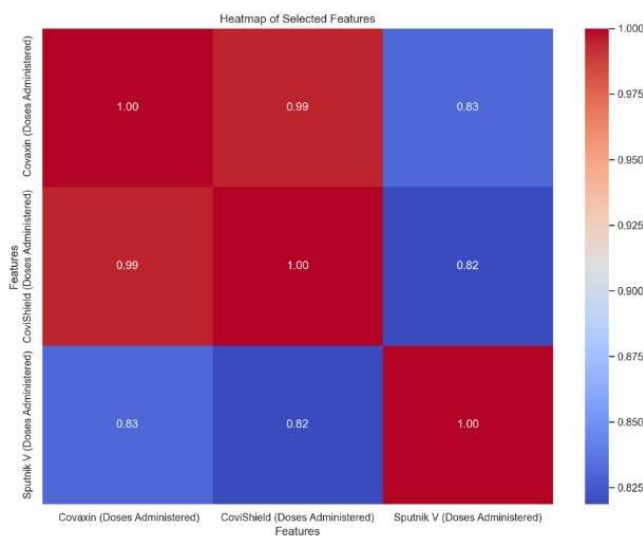

Cured Cases by Region

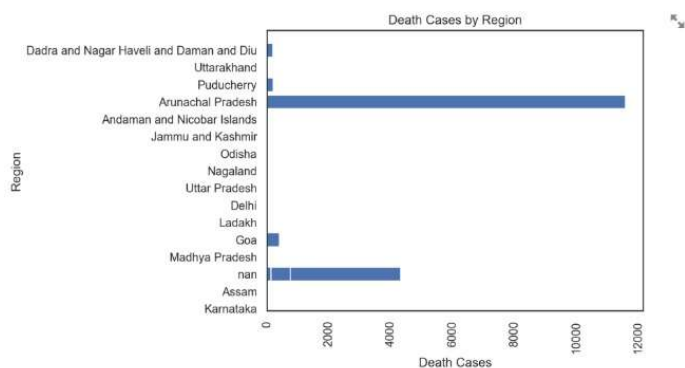# Scatter Plot: Relationship between Total Individuals and Vaccine Types

Covid-19 Dataset Analysis

Heatmap of Selected Features

Death Cases by Region

CONCLUSION

The entire project has been developed and deployed as per the requirements stated by user, it is found to be bug free as per the testing standard that are implemented.

Computer does maximum work with in minimum time. Because it is used in every field so that it provides comfort and suitability to everyone. Providing maximum facilities and comfort to customers to customers is main goal of the firm. To achieve this goal, other modern facilities relating to computer should have to be provided.