

Name:	Chaitya Arun Dobariya
UID:	2021600017
Course:	Advanced Data Visualization

Experiment 3

Aim:	<p>Design Interactive Dashboards and Storytelling using Tableau / Power BI / R (Shiny) / Python (Streamlit/Flask) / D3.js to be performed on the dataset - Disease spread / Healthcare</p> <p>Create interactive dashboard - Write observations from each chart given below</p> <p>(Advanced - Word chart, Box and whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, Jitter, Line, Area, Waterfall, Donut, Treemap, Funnel</p> <p>Basic - Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, Bubble plot)</p>
-------------	--

1. Dataset

You can find the dataset [here](#).

Description

Coronaviruses are a large family of viruses which may cause illness in animals or humans. In humans, several coronaviruses are known to cause respiratory infections ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). The most recently discovered coronavirus causes coronavirus disease COVID-19 - World Health Organization

COVID-19 cases at daily level is present in covid_19_india.csv file

Statewise testing details in StatewiseTestingDetails.csv file

Fields

StatewiseTestingDetails.csv

Date: record date (Date)

State: state names (String)

TotalSamples: count of samples (Integer)

covid_19_india.csv

Sno: serial number (Integer)

Date: record date (Date)

Time: record time (String)

State/UnionTerritory: state names (String)

ConfirmedIndianNational: confirmed Indian cases (Integer)

ConfirmedForeignNational: confirmed Foreign cases (Integer)

Cured: cured count (Integer)

Deaths: death count (Integer)

Confirmed: confirmed cases count (Integer)

Importing the libraries and data:

```
import streamlit as st
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.graph_objs as go
```

```
# Load datasets
statewise_testing = pd.read_csv('StatewiseTestingDetails.csv')
covid19_india = pd.read_csv('covid_19_india.csv')

# Convert 'Date' to datetime using correct formats
statewise_testing['Date'] = pd.to_datetime(statewise_testing['Date'], format='%Y-%m-%d')
covid19_india['Date'] = pd.to_datetime(covid19_india['Date'], format='%d-%m-%Y')

# Fill NaN values with 0 for numerical columns
statewise_testing.fillna(0, inplace=True)
covid19_india.fillna(0, inplace=True)
```

StatewiseTestingDetails.csv

	A	B	C
1	Date	State	TotalSamples
2	17-04-2020	Andaman and Nicobar Islands	1403
3	24-04-2020	Andaman and Nicobar Islands	2679
4	27-04-2020	Andaman and Nicobar Islands	2848
5	01-05-2020	Andaman and Nicobar Islands	3754
6	16-05-2020	Andaman and Nicobar Islands	6677
7	19-05-2020	Andaman and Nicobar Islands	6965
8	20-05-2020	Andaman and Nicobar Islands	7082
9	21-05-2020	Andaman and Nicobar Islands	7167
10	22-05-2020	Andaman and Nicobar Islands	7263

covid_19_india.csv

	A	B	C	D	E	F	G	H	I
1	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignNational	Cured	Deaths	Confirmed
2	1	30-01-2020	6:00 PM	Kerala	1	0	0	0	1
3	2	31-01-2020	6:00 PM	Kerala	1	0	0	0	1
4	3	01-02-2020	6:00 PM	Kerala	2	0	0	0	2
5	4	02-02-2020	6:00 PM	Kerala	3	0	0	0	3
6	5	03-02-2020	6:00 PM	Kerala	3	0	0	0	3
7	6	04-02-2020	6:00 PM	Kerala	3	0	0	0	3
8	7	05-02-2020	6:00 PM	Kerala	3	0	0	0	3
9	8	06-02-2020	6:00 PM	Kerala	3	0	0	0	3
10	9	07-02-2020	6:00 PM	Kerala	3	0	0	0	3

2. Data Preprocessing

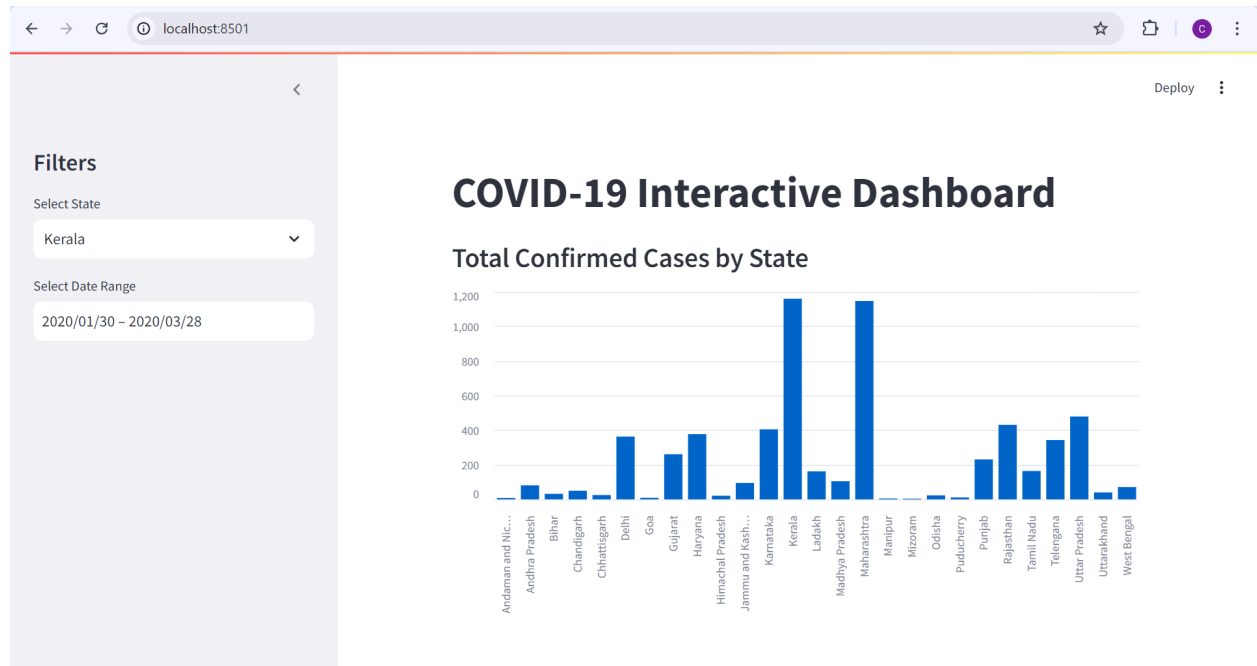
```
# Fill NaN values with 0 for numerical columns
statewise_testing.fillna(0, inplace=True)
covid19_india.fillna(0, inplace=True)

# Sidebar filters
st.sidebar.title("Filters")
selected_state = st.sidebar.selectbox("Select State", covid19_india['State/UnionTerritory'].unique())
selected_date = st.sidebar.date_input("Select Date Range", value=[covid19_india['Date'].min().date(), covid19_india['Date'].max().date()])

# Convert selected_date to datetime64
selected_date = pd.to_datetime(selected_date)

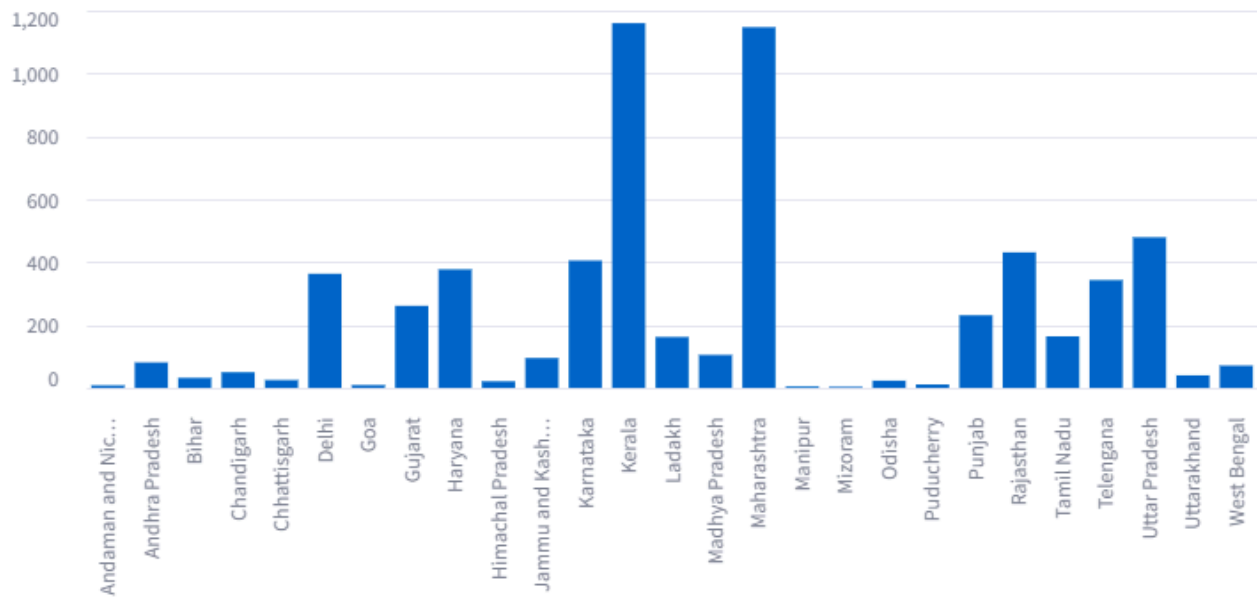
# Filter data based on selections
filtered_testing = statewise_testing[(statewise_testing['Date'] >= selected_date[0]) & (statewise_testing['Date'] <= selected_date[1])]
filtered_covid = covid19_india[(covid19_india['Date'] >= selected_date[0]) & (covid19_india['Date'] <= selected_date[1])]
```

Creating a sidebar for custom state and timeline selection



3. Charts & Plots

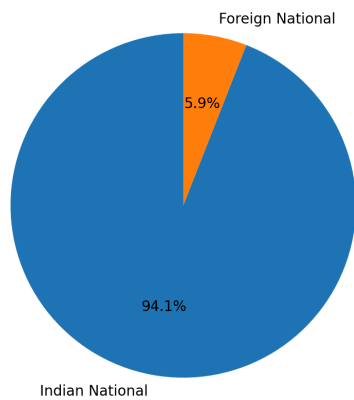
3.1 Bar Chart:



Observation:

This bar chart displays the total number of confirmed COVID-19 cases across different states in India. It provides a comparative view of the overall impact of the pandemic in each state, with states having higher case counts appearing as taller bars.

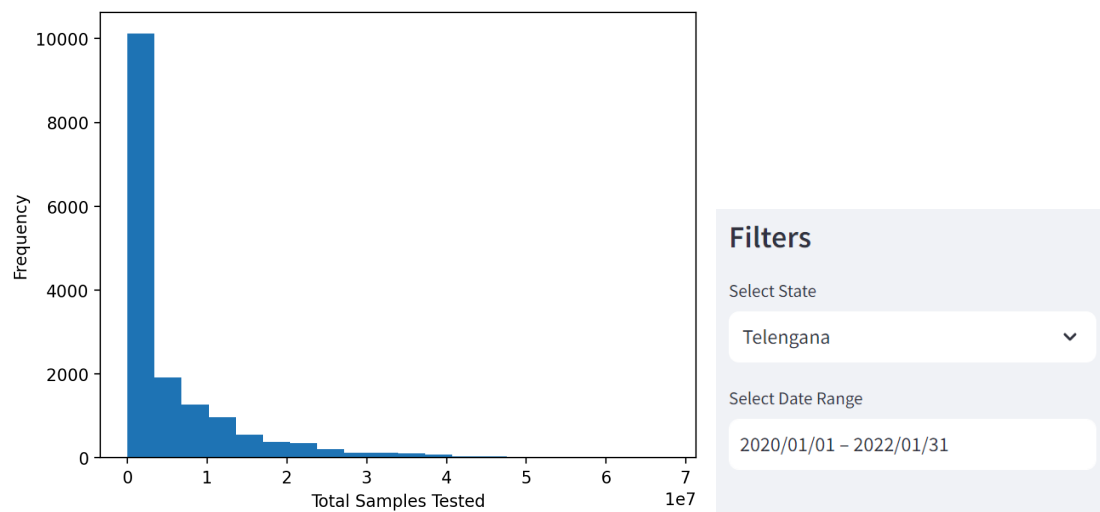
3.2 Pie Chart



Observation:

This pie chart illustrates the proportion of confirmed COVID-19 cases among Indian nationals versus foreign nationals in **Kerala**. It highlights the relative share of each group, giving insight into the demographic distribution of the confirmed cases.

3.3 Histogram



Observation:

This histogram shows the distribution of the number of total samples tested daily in **Telangana**. It provides an overview of testing activity over the selected period, highlighting the frequency of different testing volumes.

3.4 Line Chart

Confirmed Cases Over Time



Observation:

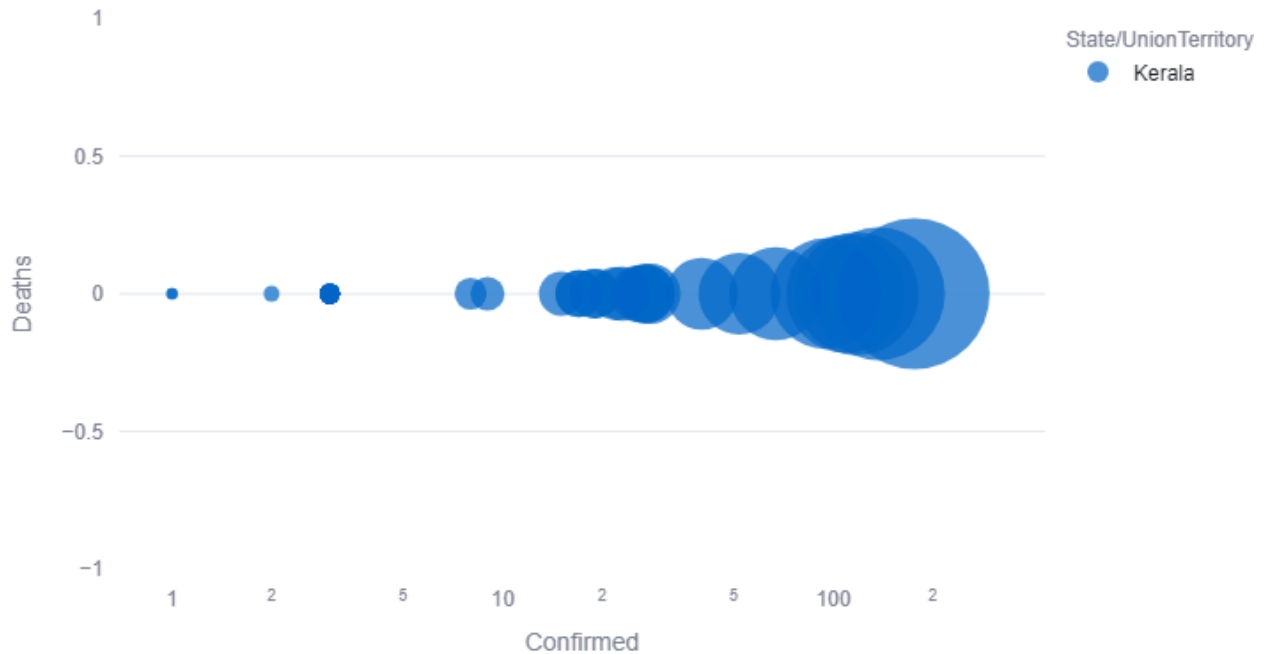
The line chart presents the trend of confirmed COVID-19 cases over time in **Telangana**. It captures how the number of cases has evolved, showing peaks and troughs in the case counts over the selected date range.

3.5 Scatter Chart

Observation:

This scatter plot visualizes the relationship between daily confirmed COVID-19 cases and the total number of tests conducted. It helps in understanding how testing volume correlates with the number of confirmed cases.

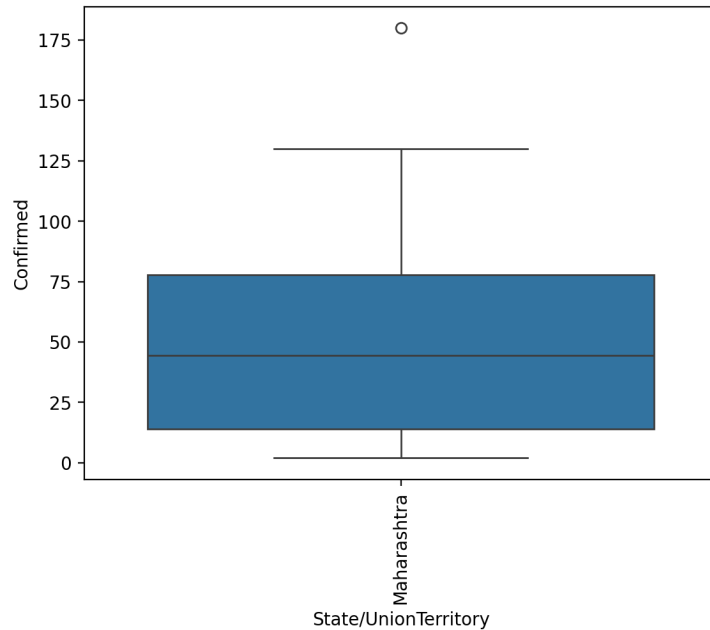
3.6 Bubble Chart



Observation:

The bubble plot displays states with bubble sizes proportional to the number of confirmed cases in **Kerala**. Each bubble represents a state, and its size indicates the volume of confirmed cases, while the position shows the number of deaths, providing a visual summary of the impact across states.

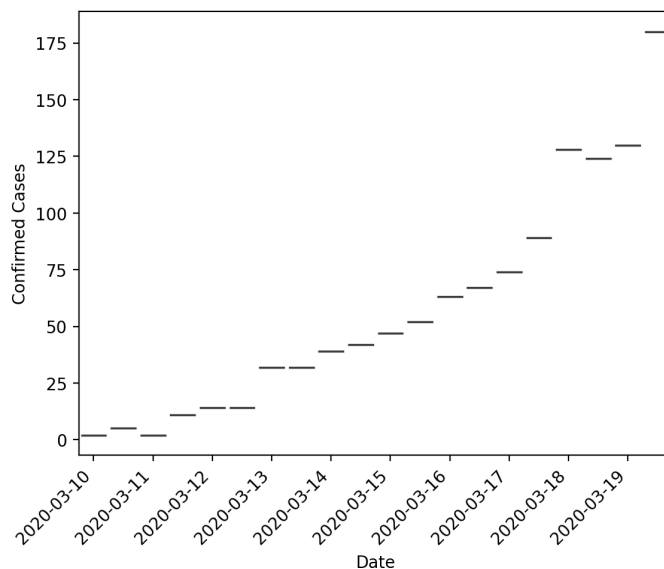
3.7 Box Plot



Observation:

This box and whisker plot shows the distribution of confirmed COVID-19 cases for the top 10 states with the highest case counts in **Maharashtra**. It reveals the spread and variability in case numbers within these states.

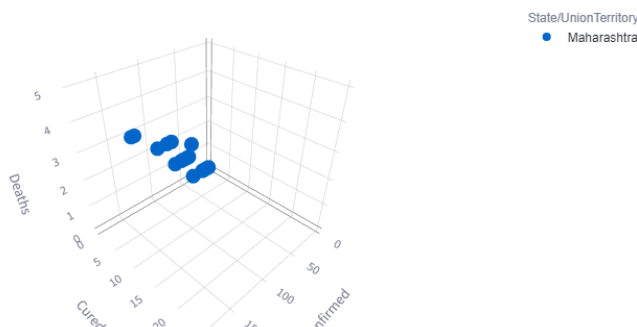
3.8 Violin Plot



Observation:

The violin plot illustrates the distribution of confirmed COVID-19 cases over time in **Maharashtra**. It provides insights into the density and distribution of cases for each date, highlighting trends and fluctuations.

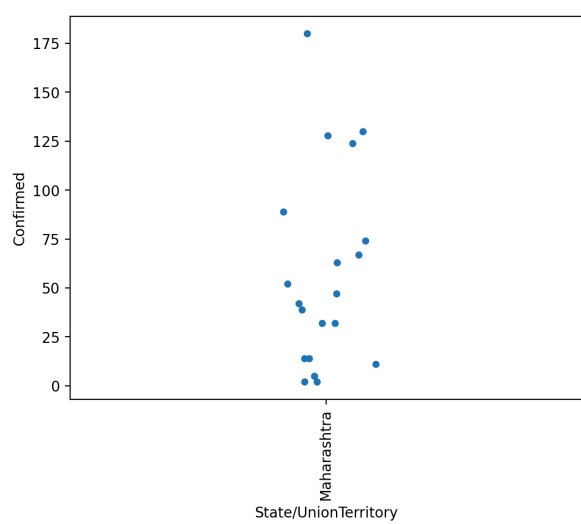
3.9 3D Scatter Plot



Observation:

The 3D scatter plot visualizes the relationship between confirmed cases, cured cases, and deaths. Each point represents a state, providing a comprehensive view of the pandemic's impact on different metrics simultaneously.

3.10 Jitter Plot

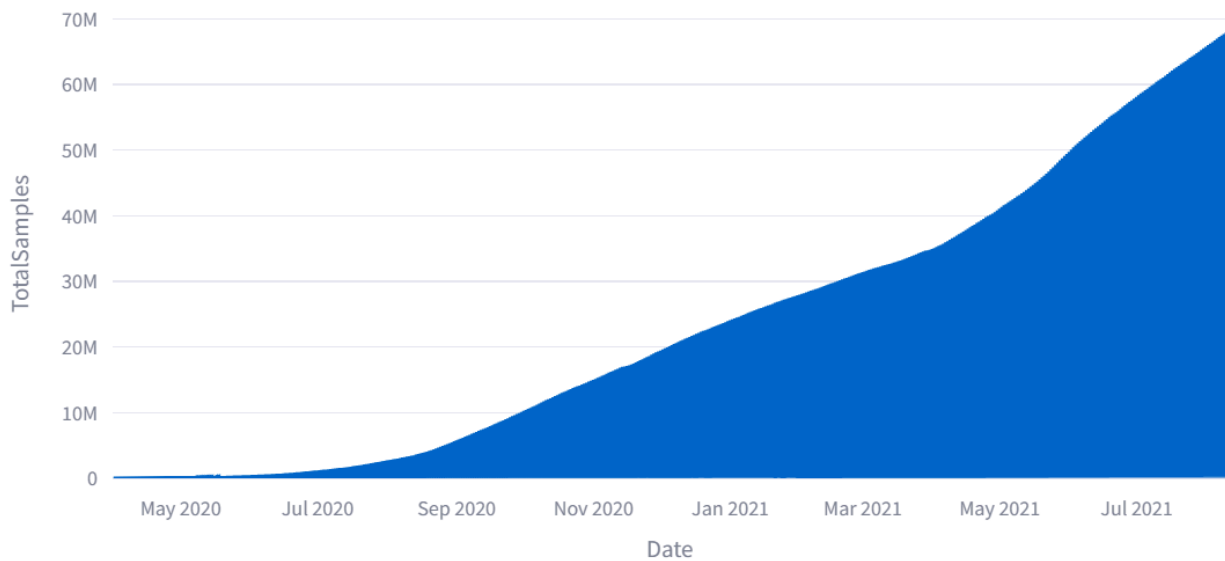


Observation:

The jitter plot displays daily confirmed COVID-19 cases across different states with jittered points. It helps in visualizing the spread of case counts across states and identifies patterns or anomalies.

3.11 Area Chart

Total Tests Conducted Over Time

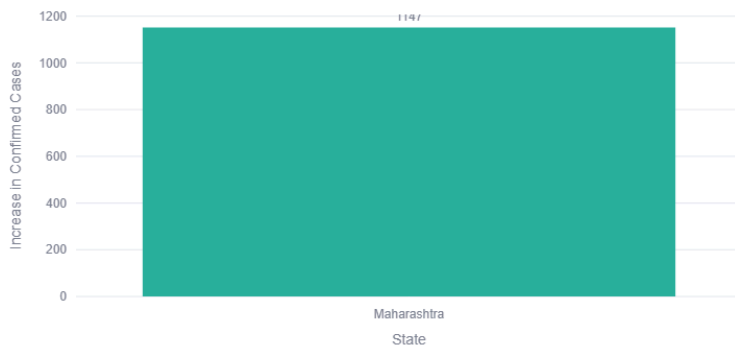


Observation:

The area chart depicts the total number of tests conducted over time. It provides a visual representation of testing efforts and their changes throughout the selected period.

3.12 Waterfall Chart

Contribution of States to Overall Increase in Cases

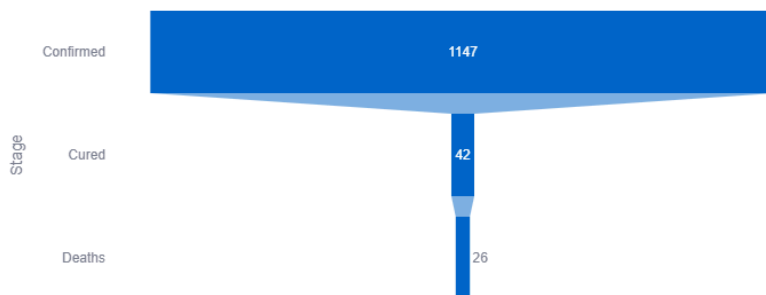


Observation:

The waterfall chart illustrates the contribution of each state to the overall increase in confirmed COVID-19 cases. It shows how case numbers have risen state by state, providing a clear view of the contribution to the national total.

3.13 Funnel Chart

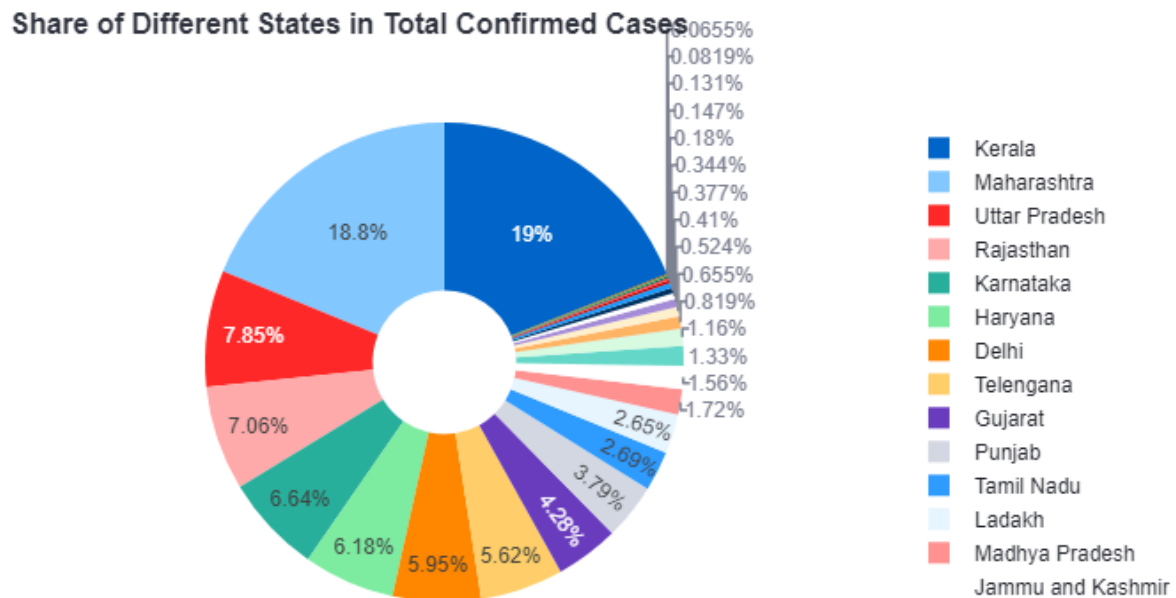
Progression of Cases from Confirmed to Cured to Deaths



Observation:

The funnel chart displays the progression of COVID-19 cases from confirmed to cured to deaths. It visually represents the stages of the pandemic and the transitions between them, highlighting the outcomes of confirmed cases.

3.14 Donut Chart

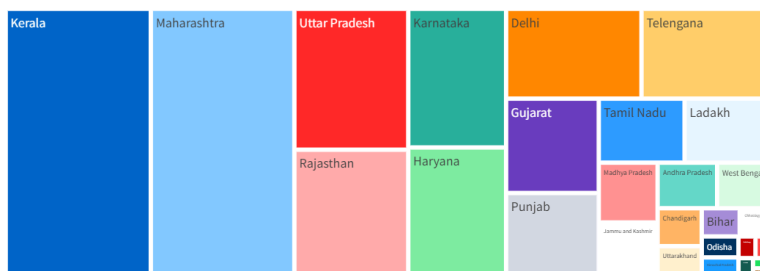


Observation:

This donut chart shows the share of total confirmed COVID-19 cases among different states. It provides a proportional view of how cases are distributed across states.

3.15 Tree Chart

Proportion of Cases by State



Observation:

The treemap visualizes the proportion of confirmed COVID-19 cases by state. Each state's area in the treemap corresponds to its share of the total cases, offering an intuitive view of case distribution across the country.

Conclusion

I created an interactive dashboard providing a comprehensive and engaging way to explore and analyze COVID-19 data. Each visualization offers unique insights into various aspects of the pandemic, from the distribution of cases across states to the relationship between testing and confirmed cases. The combination of advanced and basic chart types ensures that users can gain a deep understanding of the data, identify trends, and make informed decisions based on visualized evidence. This dashboard serves as a valuable tool for monitoring and analyzing the spread of disease and healthcare metrics, ultimately contributing to better decision-making and response strategies.