

Data Visualization & Interpretation

[COVID-19 DATA ANALYSIS]

Demonstrate the VISUALISATION CONTEXT

1. KNOW YOUR AUDIENCE

a. List the primary groups or individuals to whom you'll be communicating.

Ans. The primary groups or individuals to whom I will be communicating would be **the students of Engineering College.**

b. If you had to narrow that to a single person, who would that be?

Ans. If I had to narrow that to a single person, it would **be the student of the Engineering College.**

c. What does your audience care about?

Ans. The audience(students) care about the **spread of COVID-19 in their state, health and well-being, educational progress** and **future prospects.**

d. What action does your audience need to take?

Ans. The action that the audience (students) needs to take is to **understand the risks** and **preventive measures related to COVID-19** and **follow the guidelines** provided by the college and health authorities.

e. What is at stake? What is the benefit if the audience acts in the way you want them to? What are the risks if they don't?

Ans. The stakes involve the **students and their family's health and safety**, as well as the successful continuation of their education. If the audience (students) acts in the desired way by following the guidelines and precautions, they can protect themselves and their peers from COVID-19, reduce the risk of transmission, and ensure a safe learning environment.

The benefits include a reduced chance of contracting the virus, minimizing the disruption to their education, and contributing to the overall containment of the pandemic.

On the other hand, if the students don't adhere to the guidelines and precautions, there is a higher risk of COVID-19 transmission within the college community, potential

outbreaks, and subsequent negative impacts on their health, education, and the wider community.

2. WHAT?

- What are you trying to communicate? What questions are you trying to answer/display in your visualizations? Write these as specific questions. You need to come up with 3 questions at least, each of which will be answered using one Viz.
- Data preparation needed to answer the specific queries must be done.

Ans. We are trying to communicate the impact of Covid-19 and how the situation has evolved over time.

- a. How has the COVID-19 situation evolved over time in the state where the college is located?
- b. What is the trend of total confirmed COVID-19 cases in the state?
- c. How does the number of recovered cases compare to the number of confirmed and deceased cases?

3. Present the BIG IDEA.

- It should: (1) articulate your point of view, (2) convey what's at stake, and (3) be a complete (and single!) sentence.

Ans.

This dashboard helps in visualizing the pandemic's impact on the state, the importance of following safety measures and helps students stay informed about the situation.

4. HOW?

a. Chart 1: What type of viz did you create? Why did you select the viz that you did?

Ans. Line Chart depicting the number of cases over time in each state.

Line chart is the best to show the trends in confirmed cases, recoveries, and deaths over time, helping students understand the progression of the pandemic.

b. Chart 2: What type of viz did you create? Why did you select the viz that you did?

Ans. Horizontal Bar Chart depicting comparison of status of COVID-19 cases in the state with other states.

Bar Chart is the best for comparing number of cases between different states and understand the proportion of active, recovered, and deceased cases.

c. Chart 3: What type of viz did you create? Why did you select the viz that you did?

Ans. Area Chart depicting relationship between testing trends and confirmed cases.

Area chart is the best to show relationship relationship between testing trends and the number of confirmed cases, helping students understand the importance of testing.

d. For each of the Visualisation, identify at least 3 Gestalt principles employed.

Ans.

Gestalt Principles Employed:

Chart 1:

Law of Figure/Ground: The more the contrast between the figure and the ground, the easier it will be to distinguish between the two types of objects.

Law of Focal Point: Distinctive objects create a focal point, drawing user attention.

Law of Similarity: Using the same color for related data series (e.g., confirmed, recovered, deceased).

Chart 2:

Law of Prägnanz: Keeping the chart Simple.

Law of Similarity: Using consistent colors to represent the same categories across all states.

Law of Continuity: Keeping it in order. Aligning elements linearly to facilitate comparison of different items that are in a related grouping.

Chart 3:

Law of Proximity: Data points for testing and confirmed cases are close to each other, making the comparison intuitive.

Law of Prägnanz: Keeping the chart Simple.

Law of Similarity: Using the same color for related data series (testing and confirmed cases).

e. For each of the Visualisation, mention how you strategically used pre-attentive attributes to draw the audience's attention.

Ans.

Chart 1:

- i. **Color:** Different colors are assigned to the lines representing "Confirmed," "Recovered," and "Deceased" cases. This immediately draws attention to the different categories of data and helps the audience distinguish between them.
- ii. **Lines:** The lines themselves are pre-attentive attributes that guide the audience's focus along the trends over time. The upward or downward direction of the lines conveys the growth or decline of each category.

Chart 2:

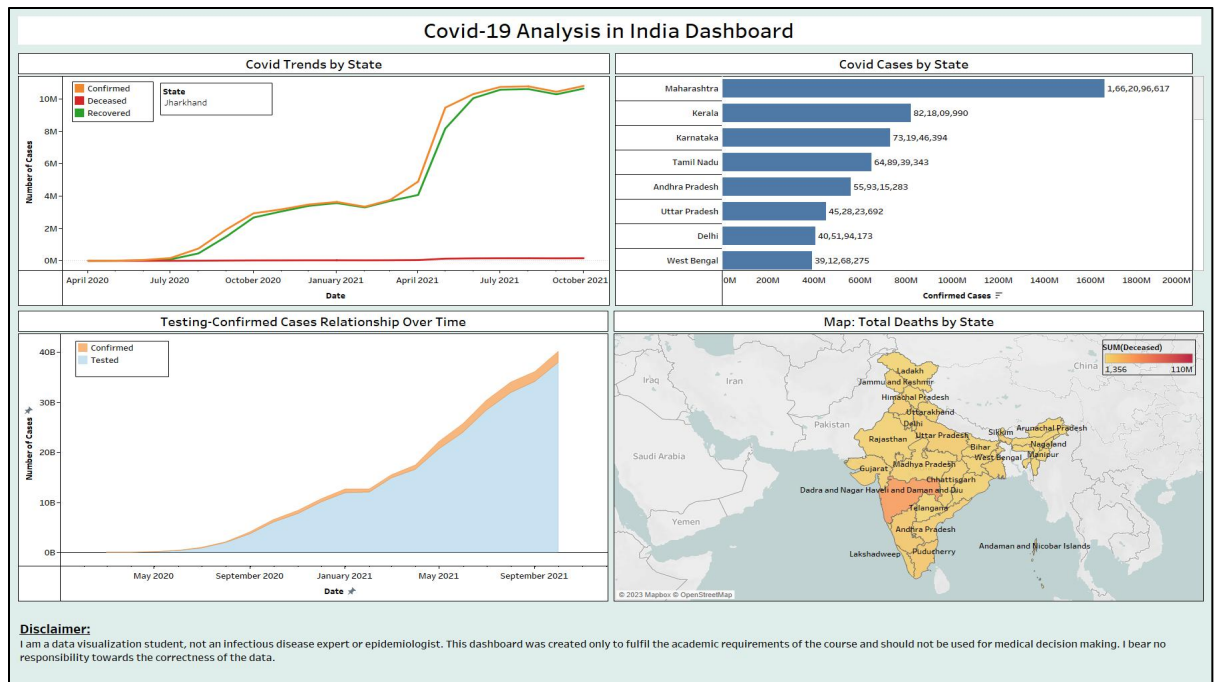
- i. **Color:** Each bar is color-coded consistently across the states. This enables quick identification of different categories and facilitates easy comparison between states.
- ii. **Spatial Position:** Each bars are equally spaced, which allows easy comparison between states.

Chart 3:

- i. **Color:** Different colors are used for Number of testing cases and Confirmed cases, which helps for easy visualization.
- ii. **Intensity:** Using color intensity to highlight the number of confirmed cases makes it stand out for the audience.

5. Create your dashboard

Along with the 3 charts in the dashboard, we have added one more Map in the dashboard depicting the total deaths by state.



Assumptions:

1. Data Preparation: The "tested" column in the dataset "states.xlsx" contains null values. Since the number of tested cases is an important parameter for analyzing the COVID-19 situation, we have decided to drop rows with null values in the "tested" column.

In our case, 763 out of 21675 rows have missing values in the 'Tested' column which corresponds to approximately 3.5% of our data. Since the percentage of missing values is small (e.g., less than 5-10%), dropping those rows might not have a significant impact on your analysis.

2. This assignment is only to fulfil the academic requirements of the course and should not be used for medical decision making. We bear no responsibility towards the correctness of the data.

3. There might be delays in testing and reporting, leading to potential discrepancies between the actual occurrence of cases and their reporting dates. This could affect the accuracy of short-term trends.