**Assignment 1 – PS6 - [COVID-19 DATA ANALYSIS] - [Weightage 12%]**

1. **Problem Statement**

Coronavirus disease 2019 (COVID‑19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It was first identified in December 2019 in Wuhan, Hubei, China, and has resulted in an ongoing pandemic. As of 17 September 2020, more than 29.8 million cases have been reported across 188 countries and territories with more than 940,000 deaths; more than 20.3 million people have recovered. As the number of cases were increasing day by day, Indian Government has taken multiple measures to handle the situation which includes the lockdown for the entire nation .However the primary purpose of the lockdown was to slow down the spread of the disease and give time to governments, both Centre and states, to be better prepared on all fronts to deal with the fallout of rising infections and a drastically slowing economy. Assuming the trajectory of new cases starts showing a flattening curve, the costs of extending the lockdown will far exceed the benefits. Now there are discussions about partial reopening of Colleges /Schools as well.

With the Government planning to reopen colleges, The principal of the Engineering Collegewhere you studied wanted to give an awareness webinar to **the students** before they start attending classes. Knowing that you are a Data Science Consultant, he thought that you would be the best candidate for the purpose.

With the given context, you need to create a dashboard using TABLEAU. (Use the concepts learned in the class).

For Data Sets: Refer  [this](https://api.covid19india.org/documentation/csv/) .Pull the details as required. [Clearly mention which all Data sets you used] <https://www.covid19india.org/about>

**Data preparation has to be done**. If required, feel free to make proper assumptions about the data and the analysis process to be followed. Don’t forget to mention the assumptions at the right places.

**The objectives include**

**Demonstrate the VISUALISATION CONTEXT**

1. **KNOW YOUR AUDIENCE(First question is answered for you)**
2. **List the primary groups or individuals to whom you’ll be communicating.**

The students of Engineering College

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

To a student who is careless in terms of health and safety and following the safety guidelines and precautionary measures.

1. **What does your audience care about?**

The students care about their health and safety and they want to know the precautionary measures and safety guidelines that they need to follow to prevent the spread of COVID-19.

1. **What action does your audience need to take?**

The students need to follow the preventive measures and take appropriate precautions which include

* + - 1. Physical or social distancing, hand washing and keeping unwashed hands away from the face
      2. Use of face mask to minimize the risk of transmission
      3. Clean your hands with sanitizer before entering the class
      4. Cover your nose and mouth with your bent elbow or a tissue when you cough or sneeze.

1. **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?**

**What is at stake?**

Students, Staff, Administration and Teachers/ Faculty life and health is at stake.

**What is the benefit if the audience acts in the way you want them to?**

If the students take the preventive measures and follow the precautions then there will be a very less chance that they will be affected by COVID-19.

**What are the risks if they don’t?**

If the students does not follow the precautionary measures and the safety guidelines, then there is a very high risk of being effected by the COVID and not putting only there life and health at stake but also the life and health of their class fellows, friends and teachers will be at stake and the college may gets closed again.

## WHAT?

* **What are you trying to communicate?**

The awareness, precautionary measures and the safety guidelines to prevent and control the spread of COVID19.

* **What questions are you trying to answer/display in your visualizations?**

From the dataset link given in the description, I have used the three datasets and their details are given as following

**Cases Time Series Dataset**

Time series Dataset of Confirmed, Recovered and Deceased cases in India

**Vaccine Doses State wise**

Number of vaccine doses administered state wise

**Raw Dataset**

Raw Data from October 10 to onward

In the visualization number one, i have drawn an analysis of the number of confirmed COVID-19 cases in India, how many recovered after suffering from COVID and those deceased due to COVID.

In the visualization number two, the details regarding the vaccination being carried out state wise and the total vaccination done so far is given.

In the visualization number three, the details about the COVID cases district wise is given

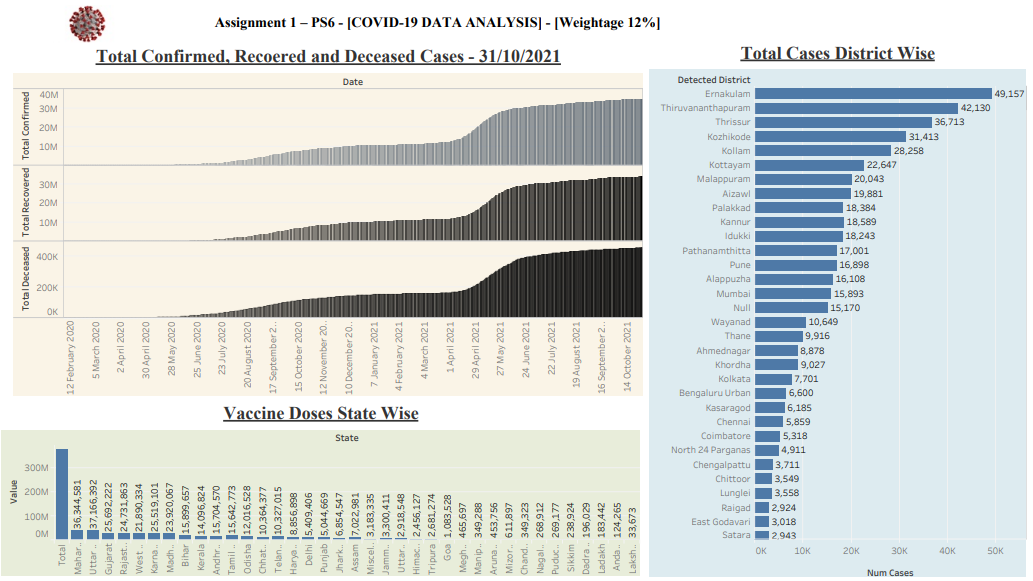


Figure COVID 19 Dashboard

* **Write these as specific questions. You need to come up with 3 questions at least, each of which will be answered using one Viz.**
* Question 1:

Provide details about confirmed, recovered and deceased cases in India

* Question 2:

Provide details about the Vaccination done state wise, to cope with COVID-19 crisis

* Question 3:

Provide details about COVID cases district wise.

* **Data preparation needed to answer the specific queries must be done.**

1. **Present the BIG IDEA**.

The program was successful at raising awareness, precautionary measures and safety guidelines students need to follow. Following the guidelines will prevent the spread of COVID-19 and there will be a very less chance that they will be affected by COVID-19. By not following the safety guidelines and precautionary measures, will not only put their life and health at stake but also the life and health of their class fellows, friends and teachers will be at stake and as a result of it college may gets closed again and it will affect their studies.

## HOW?

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

In the chart number one, i mapped total confirmed, total recovered and total deceased cases in a single graph, which precisely shows the number of people suffered from COVID, the number of people recovered from COVID and the number of people died from COVID at any date, year or month.

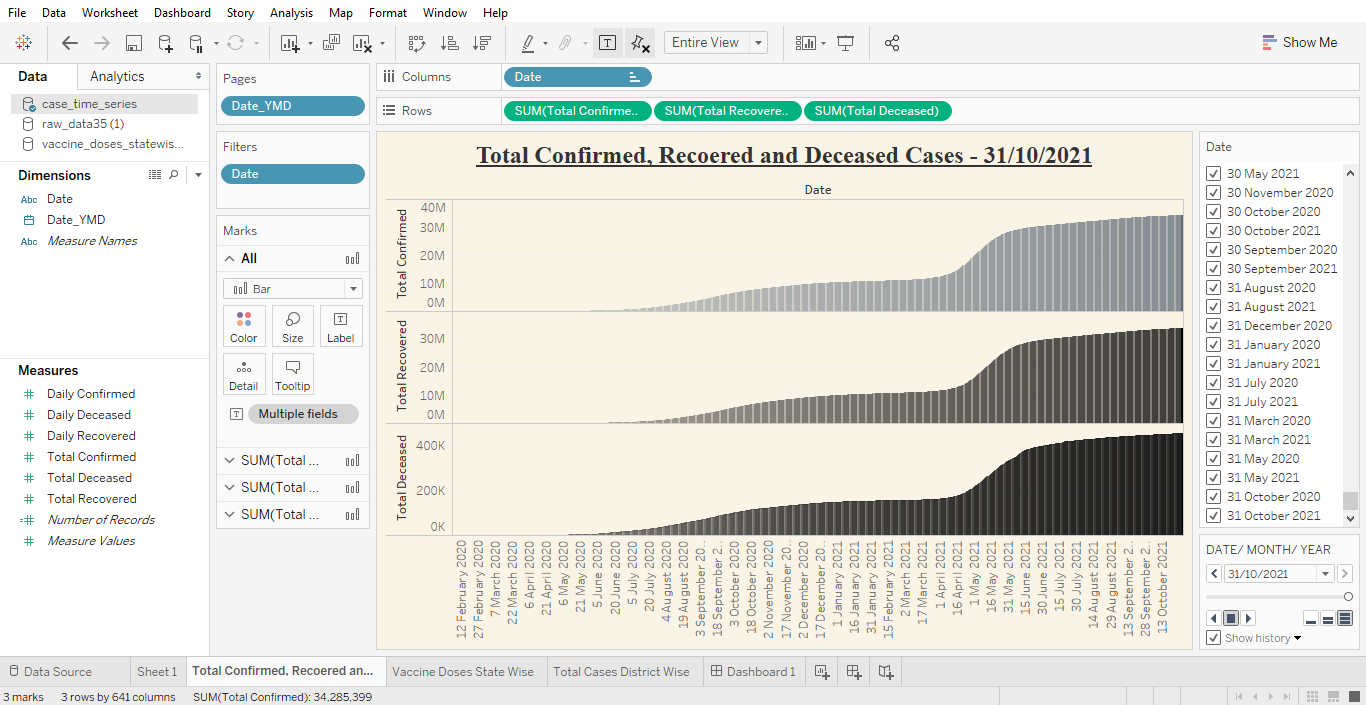


Figure Total Confirmed. Recovered and Deceased Cases

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

In the chart number two, i provided details of the vaccination done state wise. The purpose of selecting this is to explain the students, how vaccination is an effective solution to cope up with the increasing COVID cases, and as a result of rapid vaccination, how the number of people being effected by COVID are decreasing.

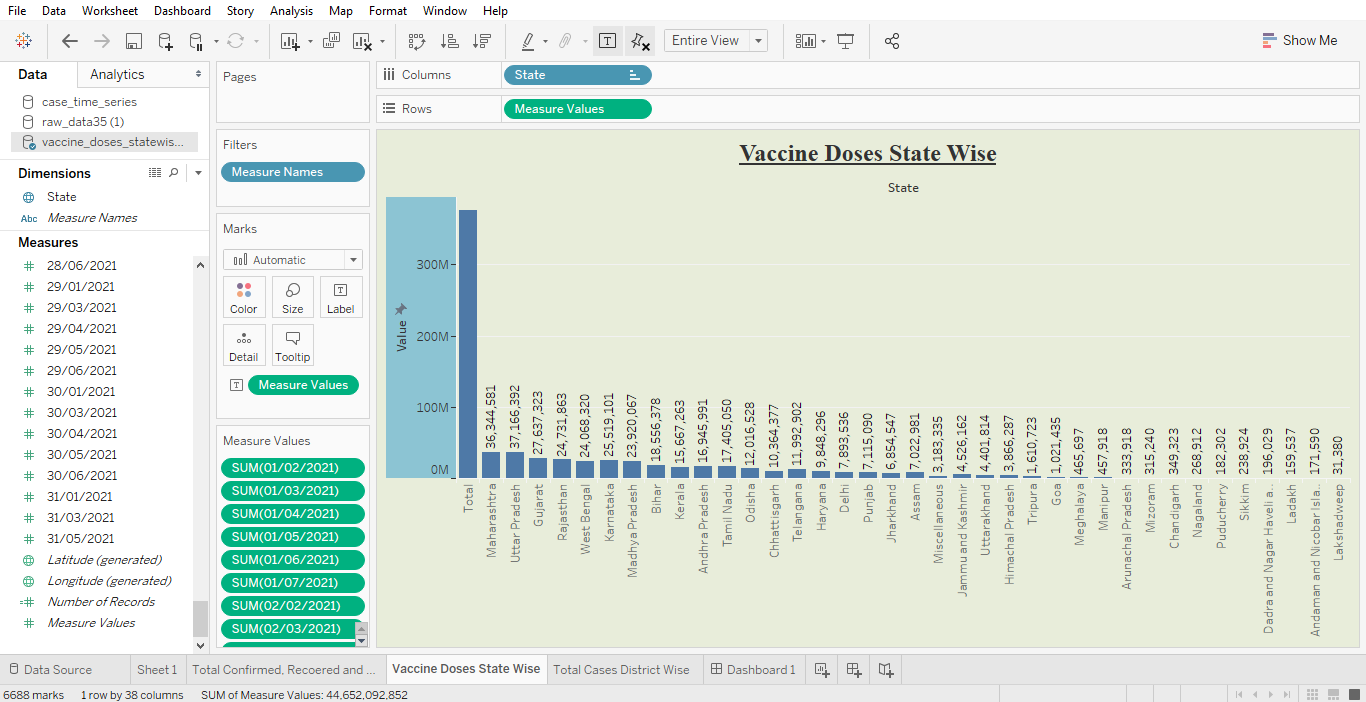


Figure Vaccine Doses State Wise

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

In the chart number three, i provided details of the COVID cases district wise. The purpose of selecting this viz is to explain the students, which districts are more affected due to COVID, so they can refrain travelling to such districts

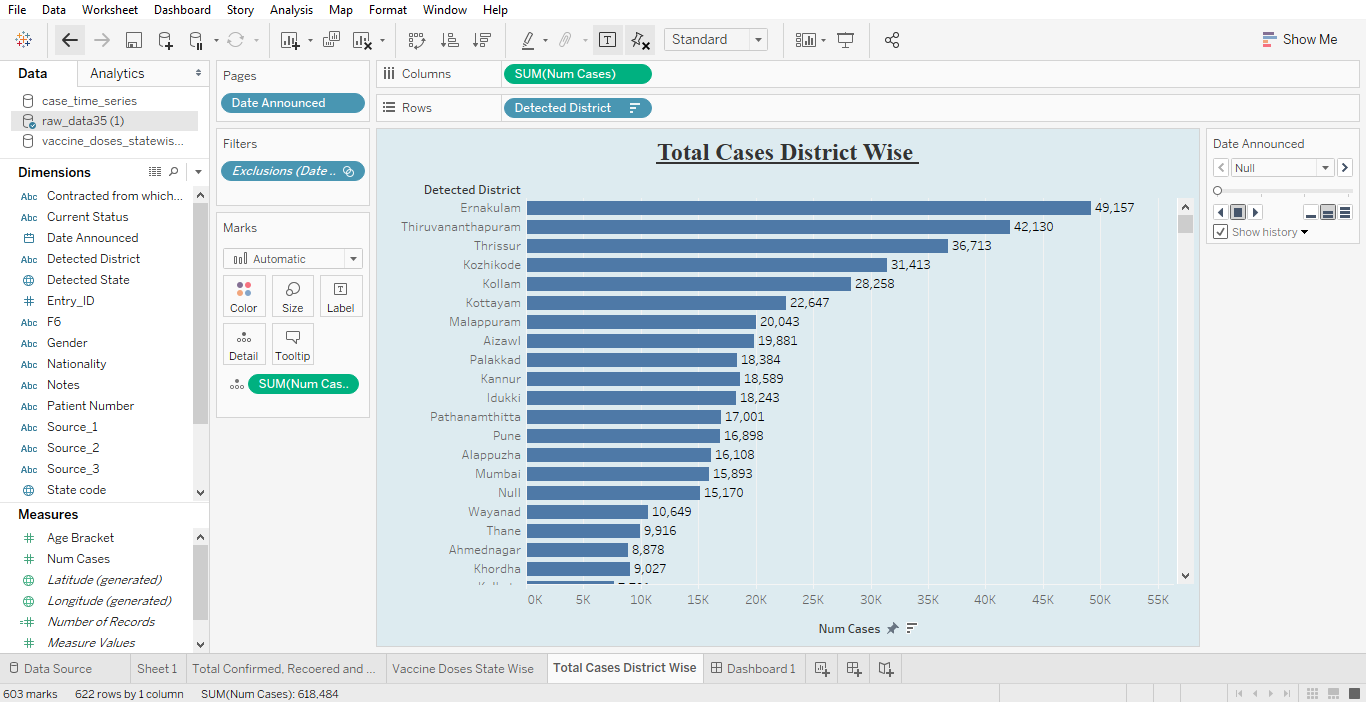


Figure Total Case District Wise

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

Chart Number 1:

In the Chart Number 1 the Gestalt principles employed are:

1. Symmetry & Order
2. Similarity
3. Proximity
4. Figure/ Ground

Chart Number 2:

In the Chart Number 2 the Gestalt principles employed are:

1. Symmetry & Order
2. Figure/ Ground
3. Similarity

Chart Number 3:

In the Chart Number 3 the Gestalt principles employed are:

1. Symmetry & Order
2. Figure/ Ground
3. Similarity
4. Proximity
5. **Create your dashboard**

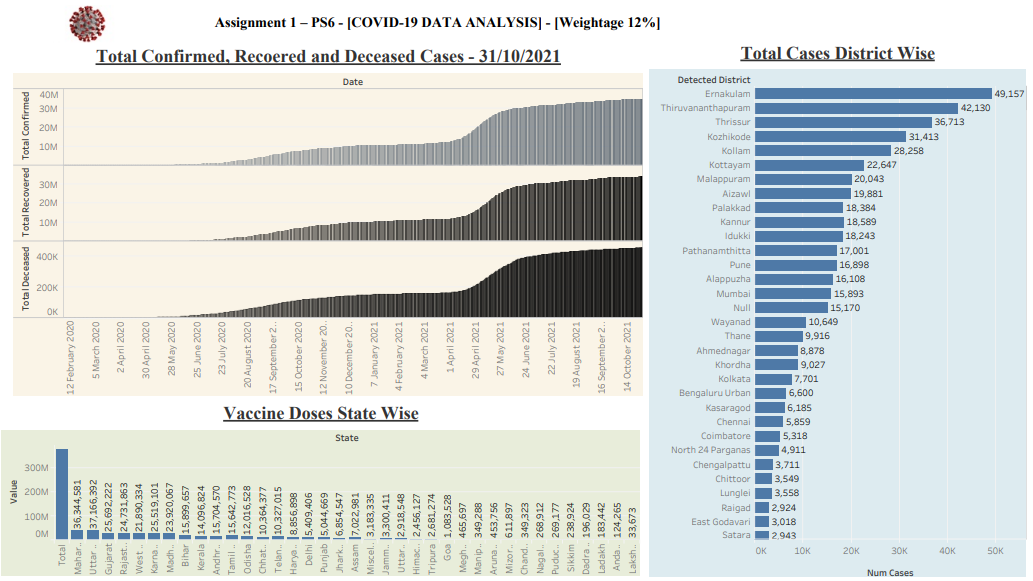
****

Figure Complete Dashboard

## Deliverables

Zipped file containing

1. The word doc with answers to question 1 through 4 above.
2. The tableau workbook(.twbx) with 3 Sheets(Each sheet should have 1 visual) and 1 Dashboard comprising all three visuals.
3. The source file after preprocessing(if any).

The file name should be the respective group name.

## Deadline

* The strict deadline for submission of the assignment is **<December 26, 2021>EoD.**
* Late submissions won’t be evaluated.

## How to submit

* This is a group assignment.
* All members of the group will work on the same problem statement.
* Each group should zip the deliverables and upload in CANVAS in respective locations under ASSIGNMENT Tab.
* Assignment submitted via means other than through CANVAS will not be graded

## Evaluation

|  |  |  |
| --- | --- | --- |
| **Sl. No** | **Criteria** | **Description** |
| 1 | Know your audience.  (10%) | * Getting to know our audience and understanding their needs and what drives them is an important early part of the process for successfully communicating with data. |
| 2. | WHAT?  Effective Formulation of Contextual Questions  (25%) | * Identify what needs to be communicated very clearly and frame the questions accordingly justifying the context. |
| 3 | BIG IDEA (10%) | * The Big Idea can help us get clear and succinct on the main message we want to get across to our audience |
| 4 | Choice of appropriate visuals  (25%) | * Identify the appropriate visuals for communicating the message |
| 5 | No clutter in the visuals  (10%) | * The visuals presented should not have any unwanted elements that reduces the understanding of data |
| 6 | Audience attention  (10%) | * The visuals presented should have the right kind of visual cues that helps the audience to focus the attention wherever required. |
| 7 | Dashboard in Tableau  (10%) | * Use Principles of Effective Dashboard Design to come up with an interesting Dashboard |

**ALL GROUP MEMBERS WILL BE CREDITED THE SAME MARKS. ITS INDIVIDUAL’S RESPONSIBILITY TO ENSURE HIS/HER PARTICIPATION AS WELL AS TEAM’S RESPONSIBILITY TO ENSURE EVERYONE’S PARTICIPATION.**