



# COVID-19 VACCINES ANALYSIS

Phase – 2: INNOVATION

211521243084 – Karthikeyan S R  
Phase - 2 Documentation Submission

# Contents

Problem Framing:	2
Efficient Data Pre-processing:	2
Insightful Exploratory Data Analysis (EDA):	2
Interactive Data Visualization:	3
1. Automated Data Pre-processing:	3
2. Insightful Exploratory Data Analysis (EDA):	3
3. Interactive Data Visualization:	3
4. Time Series Analysis:	4
5. Clustering for Insights:	4

---

## INTRODUCTION

---

This document explores innovative approaches to conduct a foundational analysis of a Covid-19 dataset. The aim is to provide beginner-friendly strategies in data preprocessing, exploratory data analysis (EDA), and data visualization.

---

## DESIGN THINKING AND INNOVATION

---

### Problem Framing:

Analyzing Covid-19 data effectively is vital for informed decision-making and public health strategies. The challenge lies in making this analysis accessible and insightful, especially for individuals new to data analysis.

---

## DESIGN OBJECTIVES:

---

### Efficient Data Pre-processing:

Develop automated techniques to handle missing data and optimize the dataset for analysis.

### Insightful Exploratory Data Analysis (EDA):

Utilize simplified EDA with visualization and summary statistics to grasp the dataset's characteristics.

### Interactive Data Visualization:

Implement interactive visualizations to present critical Covid-19 statistics in an engaging and informative manner.

---

## INNOVATING THE APPROACH

---

### 1. Automated Data Pre-processing:

**Innovation:** Utilizing automated data pre-processing techniques.

**Approach:** Automated imputation of missing values and data cleaning to ensure data readiness for analysis.

### 2. Insightful Exploratory Data Analysis (EDA):

**Innovation:** Simplified yet effective EDA with visualization.

**Approach:** Employing basic visualization techniques to better understand the age distribution and other relevant aspects of the dataset.

### 3. Interactive Data Visualization:

**Innovation:** Enhancing data visualization for engagement.

**Approach:** Creating interactive and visually appealing charts to present Covid-19 cases by country, making the data more accessible and engaging.

#### 4. Time Series Analysis:

**Innovation:** Leveraging time series analysis for temporal insights.

**Approach:** Analyzing the dataset over time to identify trends, seasonality, or unusual patterns in the spread of Covid-19.

#### 5. Clustering for Insights:

**Innovation:** Utilizing clustering to uncover patterns and group similar data points.

**Approach:** Applying clustering algorithms to group regions or countries based on Covid-19 spread characteristics, aiding in targeted interventions and resource allocation.

---

## CONCLUSION

---

By incorporating these innovative approaches, we can do the analysis on the given covid-19 dataset. These techniques are the basic approaches to analyse a given dataset.

We conclude this phase by giving this innovative approach to our problem statement based on the Design thinking I provided earlier.

I am sure that these innovative approaches will prove to be useful for our analysis.

