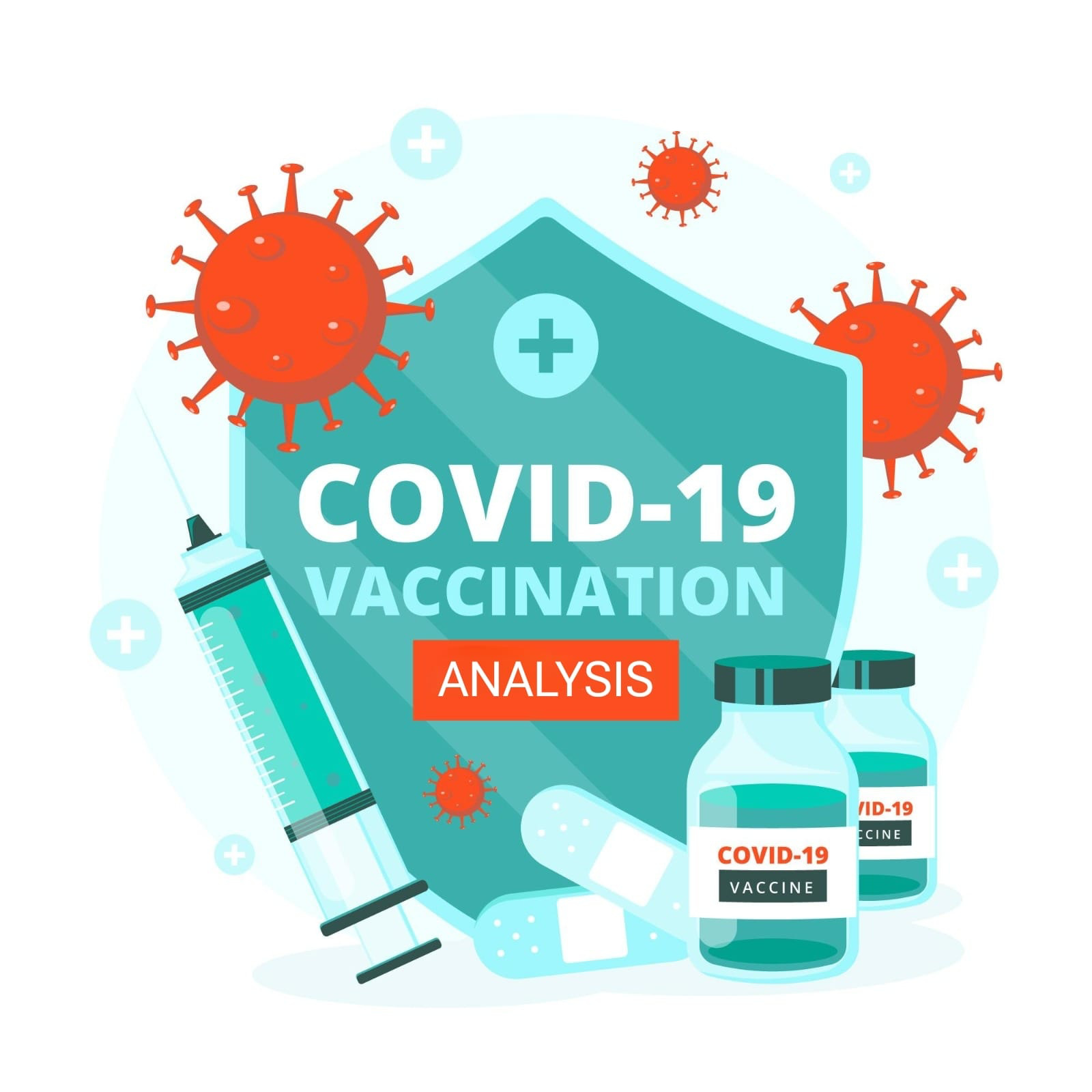
**COVID Vaccines Analysis Project**

**Project Title:** **COVID Vaccines Analysis Project**



**TEAM MEMBER:**

|  |  |  |
| --- | --- | --- |
| **NAME** | **REG NO** | **EMAIL ID** |
| **JEYASURIYA S** | **312821243010** | **21aids11@act.edu.in** |

**Problem Statement:**

Conduct an in-depth analysis of Covid-19 vaccine data, including efficacy, distribution, and adverse effects, to provide insights that aid policymakers and health organizations in optimizing vaccine deployment strategies.

**Abstract:**

The COVID-19 pandemic has posed an unprecedented global challenge, necessitating a rapid and coordinated response to mitigate its impact. Vaccination campaigns have been a cornerstone of this response, with vaccines developed and distributed at an unprecedented pace. To ensure the effectiveness of these vaccination efforts, it is essential to conduct a comprehensive analysis of COVID-19 vaccine data, encompassing aspects such as efficacy, distribution, and adverse effects.

This project embarks on an in-depth exploration of COVID-19 vaccine data, leveraging a diverse dataset from reputable sources. Through a structured approach encompassing data collection, preprocessing, exploratory data analysis, statistical analysis, and visualization, the project aims to uncover critical insights that can guide policymakers and health organizations in optimizing vaccine deployment strategies.

The project's overarching goal is to contribute to the ongoing fight against COVID-19 by providing evidence-based recommendations for vaccine distribution and management. By analyzing vaccine efficacy across different populations, investigating the prevalence and impact of adverse effects, and visualizing vaccination progress, this project seeks to inform decision-makers and support the global effort to combat the pandemic.

In a world striving for a return to normalcy, data-driven insights derived from this analysis hold the potential to facilitate more efficient and equitable vaccine distribution, ultimately saving lives and curbing the spread of COVID-19.

**Dataset Link**

Kaggle is a widely recognized and influential platform in the field of data science and machine learning. It has made significant contributions to the success of our document and the COVID Vaccines Analysis project

**COVID-19 World Vaccination Progress Dataset:**

**[ https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress ]**

**Project Steps**

**Phase 1: Problem Definition and Design Thinking**

**Problem Definition**: The problem is to conduct an in-depth analysis of Covid-19 vaccine data, focusing on vaccine efficacy, distribution, and adverse effects. The goal is to provide insights that aid policymakers and health organizations in optimizing vaccine deployment strategies. This project involves data collection, data preprocessing, exploratory data analysis, statistical analysis, and visualization.

**Design Thinking:**

1. **Data Collection:** Collect Covid-19 vaccine data from reputable sources like health organizations, government databases, and research publications.

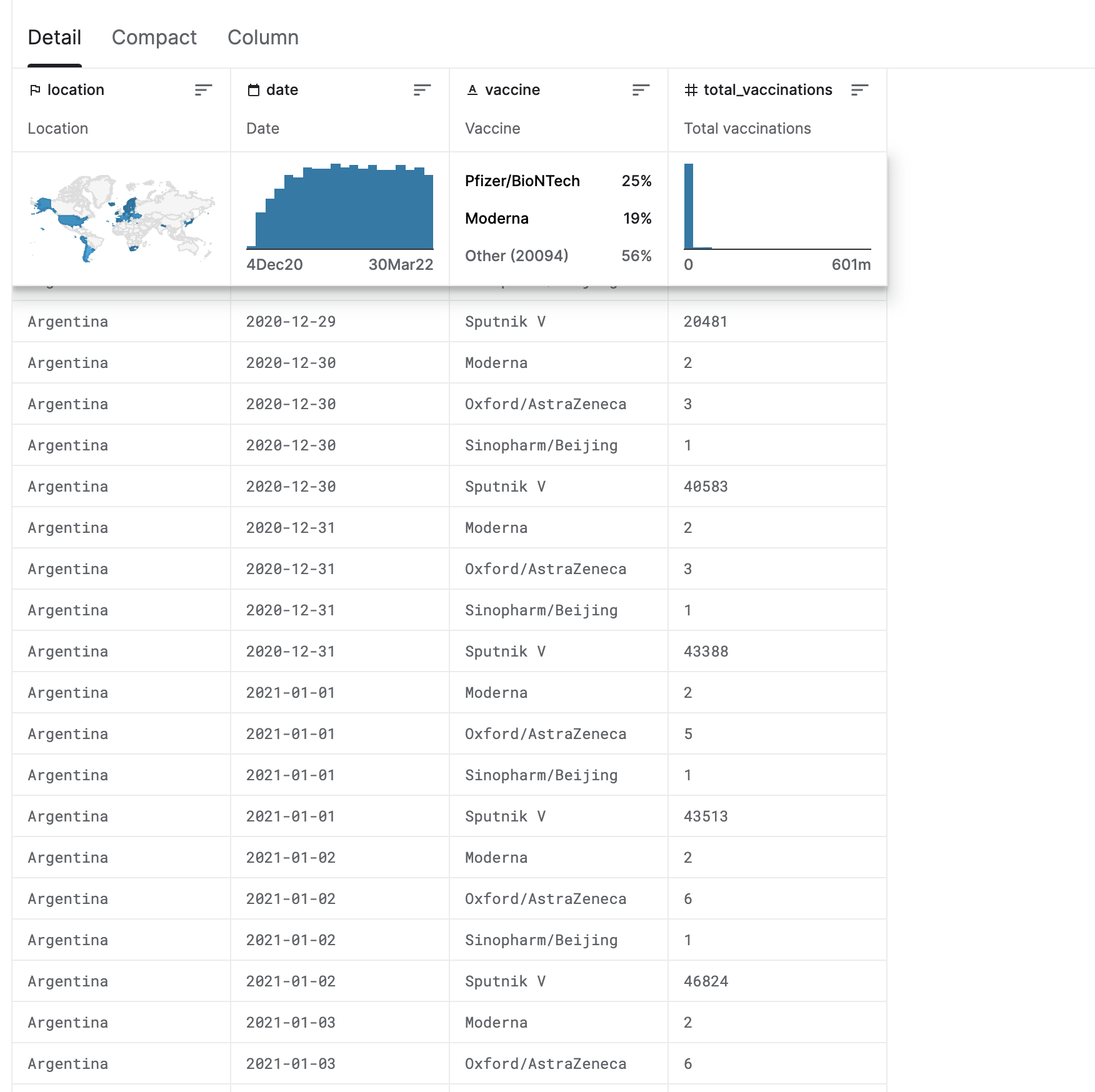
2. **Data Preprocessing:** Clean and preprocess the data, handle missing values, and convert categorical features into numerical representations.

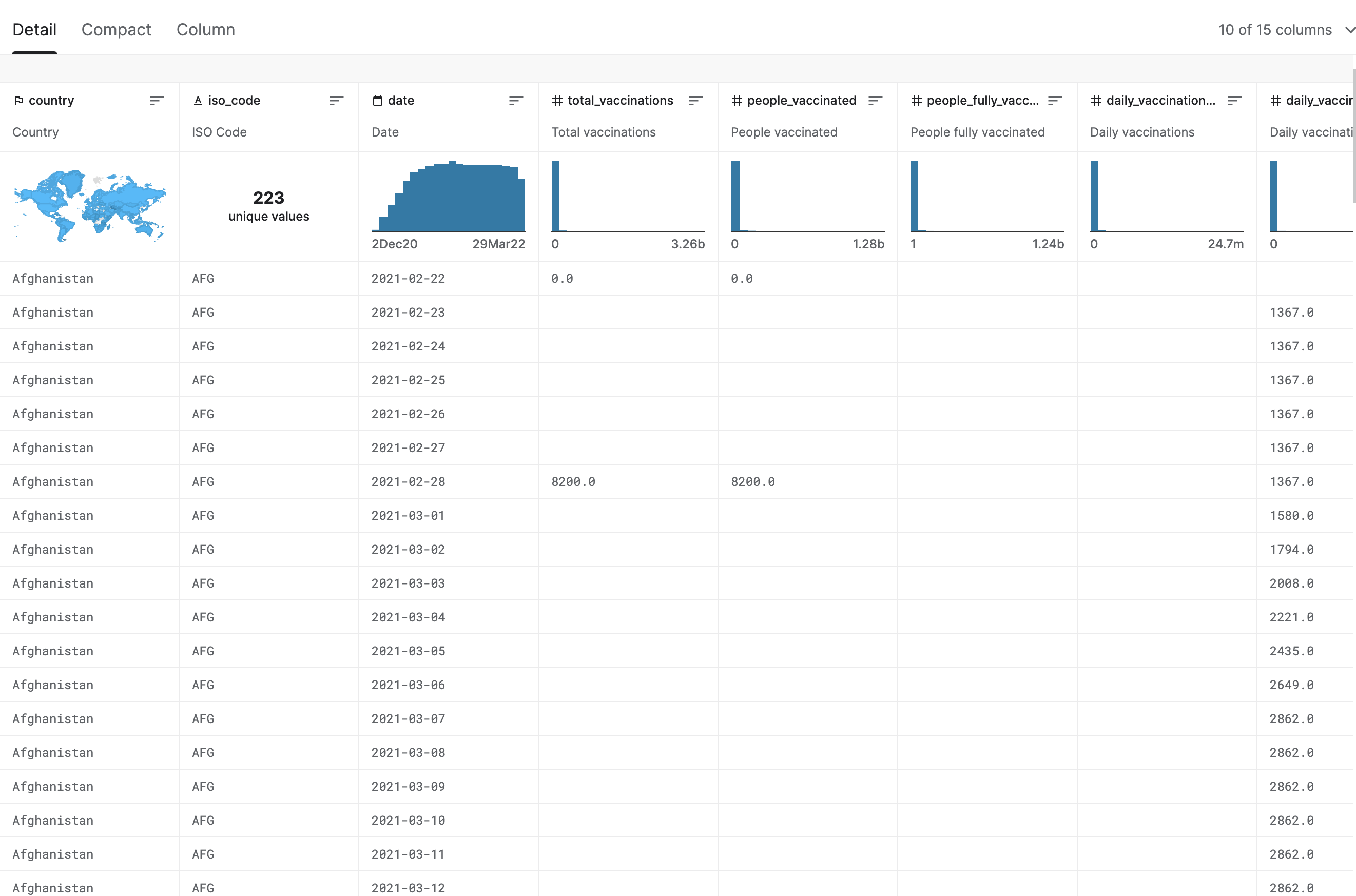
3. **Exploratory Data Analysis:** Explore the data to understand its characteristics, identify trends, and outliers.

4. **Statistical Analysis:** Perform statistical tests to analyze vaccine efficacy, adverse effects, and distribution across different populations.

5. **Visualization:** Create visualizations (e.g., bar plots, line charts, heatmaps) to present key findings and insights.

6. **Insights and Recommendations:** Provide actionable insights and recommendations based on the analysis to assist policymakers and health organizations.





## **About Dataset**

### **Context**

Data is collected daily from **[Our World in Data](https://ourworldindata.org/" \t "/Users/aravind/Desktop/x/_blank)** GitHub repository for [covid-19](https://github.com/owid/covid-19-data" \t "/Users/aravind/Desktop/x/_blank), merged and uploaded. Country level vaccination data is gathered and assembled in one single file. Then, this data file is merged with locations data file to include vaccination sources information. A second file, with manufacturers information, is included.

### **Content**

The data (country vaccinations) contains the following information:

* ****Country****- this is the country for which the vaccination information is provided;
* ****Country ISO Code**** - ISO code for the country;
* ****Date**** - date for the data entry; for some of the dates we have only the daily vaccinations, for others, only the (cumulative) total;
* ****Total number of vaccinations**** - this is the absolute number of total immunizations in the country;
* ****Total number of people vaccinated**** - a person, depending on the immunization scheme, will receive one or more (typically 2) vaccines; at a certain moment, the number of vaccination might be larger than the number of people;
* ****Total number of people fully vaccinated**** - this is the number of people that received the entire set of immunization according to the immunization scheme (typically 2); at a certain moment in time, there might be a certain number of people that received one vaccine and another number (smaller) of people that received all vaccines in the scheme;
* ****Daily vaccinations (raw)**** - for a certain data entry, the number of vaccination for that date/country;
* ****Daily vaccinations**** - for a certain data entry, the number of vaccination for that date/country;
* ****Total vaccinations per hundred**** - ratio (in percent) between vaccination number and total population up to the date in the country;
* ****Total number of people vaccinated per hundred**** - ratio (in percent) between population immunized and total population up to the date in the country;
* ****Total number of people fully vaccinated per hundred**** - ratio (in percent) between population fully immunized and total population up to the date in the country;
* ****Number of vaccinations per day**** - number of daily vaccination for that day and country;
* ****Daily vaccinations per million**** - ratio (in ppm) between vaccination number and total population for the current date in the country;
* ****Vaccines used in the country**** - total number of vaccines used in the country (up to date);
* ****Source name**** - source of the information (national authority, international organization, local organization etc.);
* ****Source website**** - website of the source of information;

There is a second file added recently (country vaccinations by manufacturer), with the following columns:

* ****Location**** - country;
* ****Date**** - date;
* ****Vaccine**** - vaccine type;
* ****Total number of vaccinations**** - total number of vaccinations / current time and vaccine type.

**Conclusion:**

In conclusion, our COVID Vaccines Analysis project, fueled by the rich dataset from Kaggle and the collaborative spirit of the data science community, has undertaken a comprehensive exploration of COVID-19 vaccine data. Through rigorous data collection, preprocessing, exploratory analysis, statistical examination, and impactful visualization, we have strived to uncover vital insights that can aid policymakers and health organizations in optimizing vaccine deployment strategies.

As we navigate through the complexities of the ongoing pandemic, data-driven insights are indispensable in the fight against COVID-19. Our project's findings and recommendations, rooted in evidence and analysis, stand as a valuable contribution to the global effort to combat the virus. By leveraging the power of data science and the resources offered by Kaggle, we aim to facilitate more efficient and equitable vaccine distribution, ultimately working towards a safer and healthier world for all.

This project highlights the transformative potential of data analysis and the collaborative nature of platforms like Kaggle in addressing critical real-world challenges. It underscores the importance of data-driven decision-making and reaffirms the belief that, together, we can make a meaningful impact in our collective battle against the COVID-19 pandemic.