

COVID VACCINES ANALYSIS

1 . Data Collection: Gather data from reliable sources such as government health agencies, research organizations, or datasets related to COVID vaccine information.

2 . Data Loading: Import the collected data into your analysis environment. This could be done using tools like Python (Pandas), R, or data analysis software.

3 . Data Inspection: Examine the dataset to understand its structure, size, and the variables it contains. Use descriptive statistics and summary functions to get an initial overview of the data.

4 . Handling Missing Values: Identify and address missing data by either imputing missing values or removing rows or columns with excessive missing data.

5 . Data Type Conversion: Ensure that data types for each variable are appropriate. For example, date fields should be in date format, and numerical data should be in the correct numeric format.

6 . Data Cleaning: Clean the data by addressing outliers, inconsistencies, and errors in the dataset. This may involve removing duplicates, correcting data entry mistakes, and standardizing values.

7 . Feature Engineering: Create new features or transform existing ones to extract meaningful information. For COVID vaccine data, this might involve calculating vaccination rates, infection rates, or other relevant metrics.

8 . Data Aggregation: If the data is collected at a granular level, consider aggregating it to different time intervals (daily, weekly, monthly) or geographic regions (country, state) as needed for your analysis.

9 . Data Transformation: Normalize or scale data if required. Transformation methods could include standardization, log transformation, or other techniques to make the data suitable for analysis.

10 . Data Visualization: Create visualizations such as bar charts, line plots, heatmaps, or scatter plots to represent the data. Visualization can help identify trends, patterns, and relationships in the data.

11 . Data Export: Save the cleaned and analyzed data for future reference or share it with others. Common formats for export include CSV, Excel, or specific data analysis software formats.

```
import pandas as pd
import matplotlib.pyplot as plt

# Load your COVID vaccine dataset
# (replace 'your_data.csv' with your actual dataset)
data = pd.read_csv('your_data.csv')

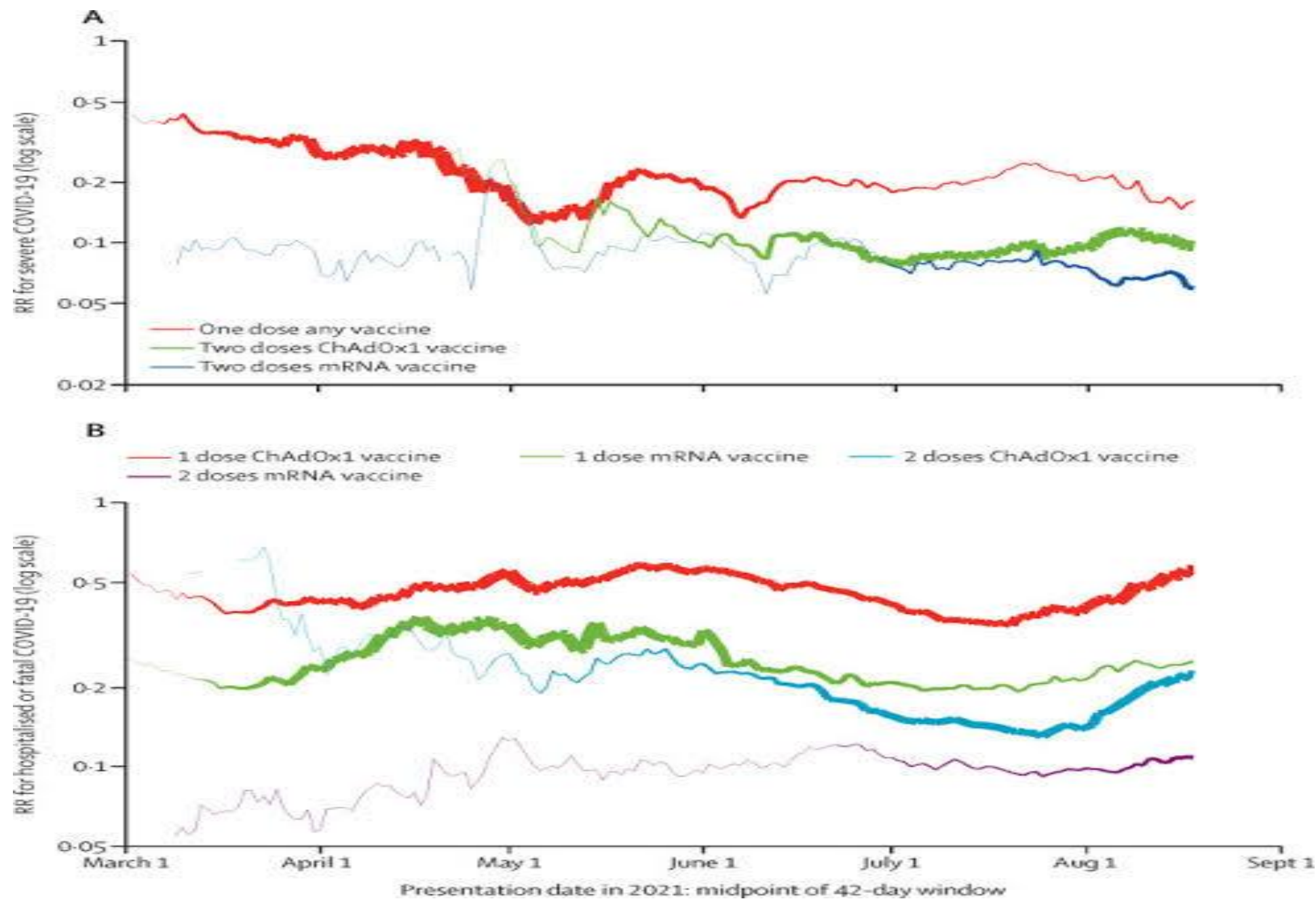
# Perform data analysis
total_vaccinations = data['total_vaccinations'].sum()
total_people_vaccinated = data['people_vaccinated'].sum()
total_people_fully_vaccinated = data['people_fully_vaccinated'].sum()
```

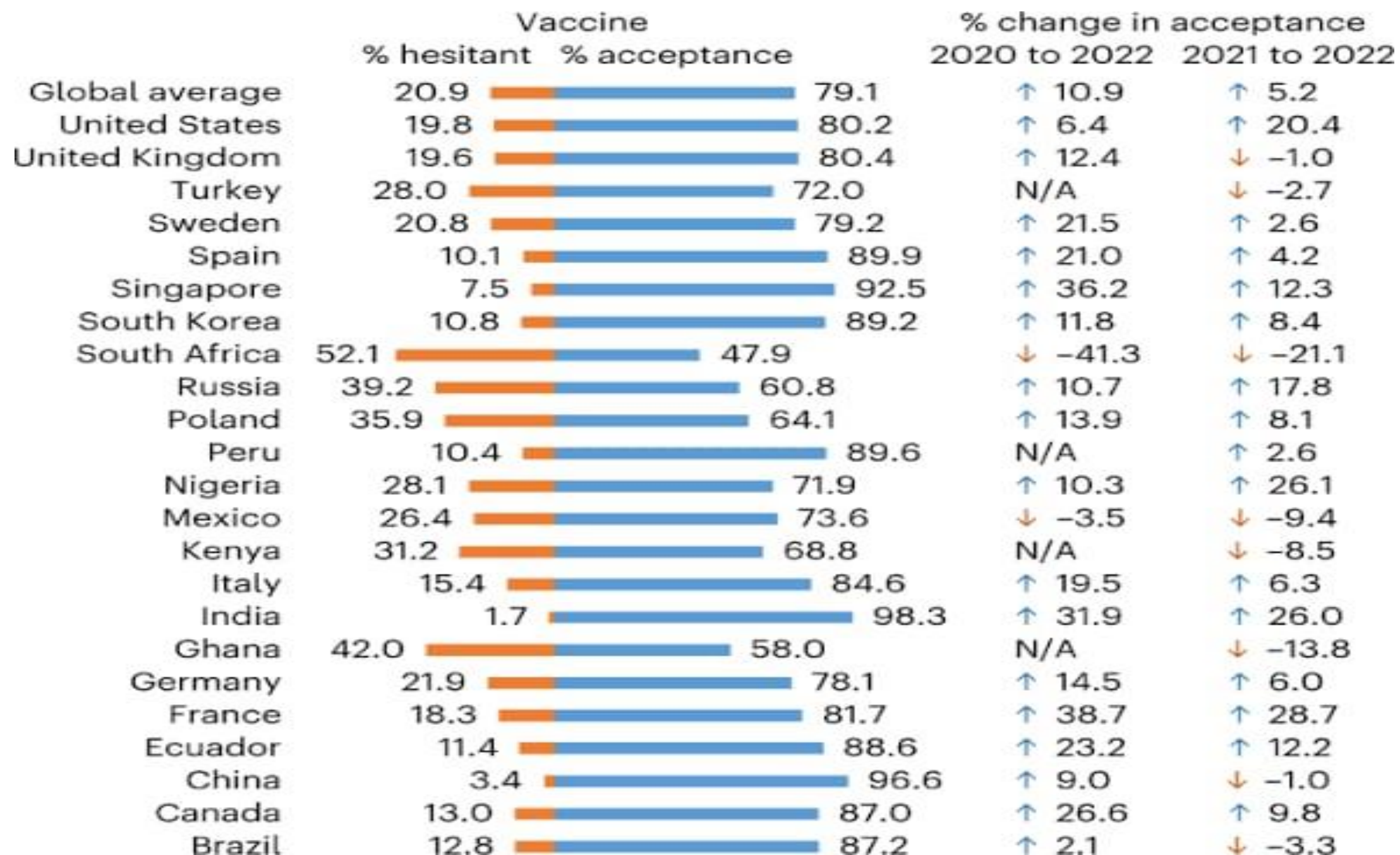
```
# Print the analysis resultsprint(f'Total  
Vaccinations:  
{total_vaccinations}')print(f'Total People  
Vaccinated:  
{total_people_vaccinated}')print(f'Total  
People Fully Vaccinated:  
{total_people_fully_vaccinated}')
```

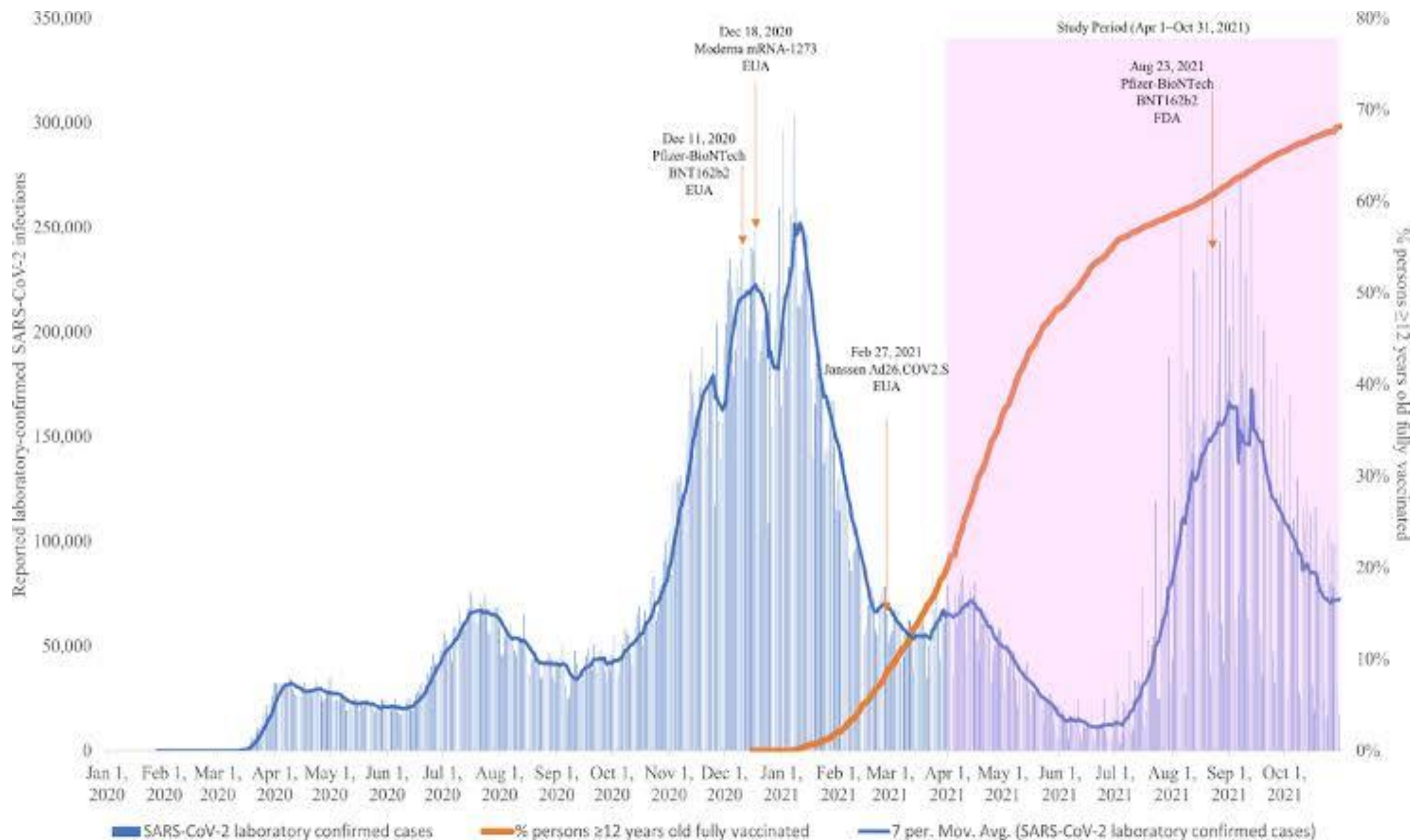
```
# Create a simple bar chartdata['date'] =  
pd.to_datetime(data['date'])data =  
data.set_index('date')
```

```
plt.figure(figsize=(10, 6))plt.plot(data.index,  
data['total_vaccinations'], label='Total  
Vaccinations')plt.plot(data.index,  
data['people_vaccinated'], label='People  
Vaccinated')plt.plot(data.index,  
data['people_fully_vaccinated'], label='People  
Fully  
Vaccinated')plt.xlabel('Date')plt.ylabel('Count')plt.ti  
tle('COVID Vaccine Data  
Analysis')plt.legend()plt.grid(True)
```

```
# Save and display the  
chartplt.savefig('covid_vaccine_analysis.  
png')plt.show()
```







Daily COVID-19 vaccine doses administered

7-day rolling average. All doses, including boosters, are counted individually.

Our World
in Data



Data source: Official data collated by Our World in Data

CC BY