COVID 19 VACCINE ANALYSIS

ABSTRACT

In this analysis, we explored a comprehensive dataset on COVID-19 vaccination progress across different countries. The dataset includes information on daily vaccination numbers, total vaccinations administered, vaccine types, and more. Our goal was to gain insights into the global vaccination effort, understand vaccination trends over time, and identify factors influencing vaccination rates.

Key Findings and Insights

Global Vaccination Trends

Vaccine Types

Temporal Trends

Regional Disparities

Correlation Analysis

OBJECTIVE

1. Some math to understand it better, and making visuals to explain it clearly. The hope is that by doing this, we can give a good picture of how the vaccines are doing and help in the fight against Covid-19.
2. The project aims to thoroughly analyze Covid-19 vaccine data with key objectives:
   - Evaluating vaccine efficacy
   - Scrutinizing distribution strategies
   - Investigating adverse effects
   - Providing actionable insights
3. By achieving these goals, the project seeks to enhance decision-making for policymakers and health organizations, fostering optimized deployment strategies in the ongoing battle against the Covid-19 pandemic.

DATA PREPROCESSING

- Check for missing values in each column and decide how to handle them (e.g., imputation or removal).
- Handle data types appropriately (e.g., convert the date column to datetime).

- Ensure data consistency and correctness, such as checking that percentages are within valid ranges (0-100%).

1. Import the necessary libraries :

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

In this step, you import the Pandas library, which is essential for data manipulation and analysis.

2. Load the dataset:

```
df = pd.read_csv('cv.csv')
```

This code uses Pandas' `read_csv()` function to load the dataset from a CSV file into a Pandas DataFrame.

3. Data Exploration:

- `df.head()`: This function displays the first few rows of the dataset, allowing you to see what the data looks like at a glance.

- `df.info()`: The `info()` function provides information about the data types of each column and the number of non-null entries, which is useful for checking for missing data.

- `df.describe()`: The `describe()` function provides basic statistical summaries of the numeric columns, such as mean, standard deviation, and quartiles.

4. Data Preprocessing:

Data preprocessing involves various tasks to clean and prepare the data for analysis. Common preprocessing tasks include:

- Handling Missing Values: Use the `fillna()` function to fill missing values with a specific value or a strategy like mean or median. In the example, missing values are filled with 0.

- Feature Engineering: Create new columns or extract information from existing columns based on your analysis requirements. This step is highly specific to your analysis goals.

6. Save the Preprocessed Data:

If you want to save the preprocessed data for future use, you can use the `to_csv()` function to export it to a new CSV file. Setting `index=False` ensures that the index column is not saved to the file.

Data Source

Data set link https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress

country	iso_code	date	total_vaccin	people_vacc	people_fully	daily_vaccin	daily_vaccin	total_vaccin	people_vacc	people_fully	daily_vaccin	vaccines	source_nam	source_website
Afghanistan	AFG	22-02-2021	0	0				0	0			Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	23-02-2021					1367				34	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	24-02-2021					1367				34	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	25-02-2021					1367				34	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	26-02-2021					1367				34	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	27-02-2021					1367				34	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	28-02-2021	8200	8200			1367	0.02	0.02		34	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	01-03-2021					1580				40	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	02-03-2021					1794				45	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	03-03-2021					2008				50	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	04-03-2021					2221				56	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	05-03-2021					2435				61	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	06-03-2021					2649				66	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	07-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	08-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	09-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	10-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	11-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	12-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	13-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	14-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	15-03-2021					2862				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	16-03-2021	54000	54000			2862	0.14	0.14		72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	17-03-2021					2882				72	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	18-03-2021					2902				73	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	19-03-2021					2921				73	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	20-03-2021					2941				74	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	21-03-2021					2961				74	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	22-03-2021					2980				75	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	23-03-2021					3000				75	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	24-03-2021					3000				75	Johnson&Jo	World Healt	https://covid19.who.int/
Afghanistan	AFG	25-03-2021					3000				75	Johnson&Jo	World Healt	https://covid19.who.int/

PROGRAM:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('cv.csv')
print(df.head())
print(df.info())
print(df.describe())
df.fillna(0, inplace=True)
afghanistan_data = df[df['country'] == 'Afghanistan']
plt.figure(figsize=(12, 6))
plt.subplot(1, 2, 1)
sns.lineplot(x='date', y='total_vaccinations', data=afghanistan_data)
```

```
plt.title('Total Vaccinations Over Time')
plt.xlabel('Date')
plt.ylabel('Total Vaccinations')

plt.subplot(1, 2, 2)
sns.lineplot(x='date', y='daily_vaccinations', data=afghanistan_data)
plt.title('Daily  Vaccinations Over Time')
plt.xlabel('Date')
plt.ylabel('Daily Vaccinations')

plt.tight_layout()
plt.show()
df.to_csv('data.csv', index=False)
```

OUTPUT:

```
Python 3.11.0 (main, Oct 24 2022, 18:26:48) [MSC v.1933 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.

========== RESTART: D:\naan\naan.py ==========
       country iso_code  ...                source_name                source_website
0  Afghanistan       AFG  ...  World Health Organization  https://covid19.who.int/
1  Afghanistan       AFG  ...  World Health Organization  https://covid19.who.int/
2  Afghanistan       AFG  ...  World Health Organization  https://covid19.who.int/
3  Afghanistan       AFG  ...  World Health Organization  https://covid19.who.int/
4  Afghanistan       AFG  ...  World Health Organization  https://covid19.who.int/

[5 rows x 15 columns]
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 86512 entries, 0 to 86511
Data columns (total 15 columns):
 #   Column                               Non-Null Count  Dtype
---  ------                               --------------  -----
 0   country                              86512 non-null  object
 1   iso_code                             86512 non-null  object
 2   date                                 86512 non-null  object
 3   total_vaccinations                   43607 non-null  float64
 4   people_vaccinated                    41294 non-null  float64
 5   people_fully_vaccinated              38802 non-null  float64
 6   daily_vaccinations_raw               35362 non-null  float64
 7   daily_vaccinations                   86213 non-null  float64
 8   total_vaccinations_per_hundred       43607 non-null  float64
 9   people_vaccinated_per_hundred        41294 non-null  float64
 10  people_fully_vaccinated_per_hundred  38802 non-null  float64
 11  daily_vaccinations_per_million       86213 non-null  float64
 12  vaccines                             86512 non-null  object
 13  source_name                          86512 non-null  object
 14  source_website                       86512 non-null  object
dtypes: float64(9), object(6)
memory usage: 9.9+ MB
None
       total_vaccinations  ...  daily_vaccinations_per_million
count        4.360700e+04  ...                   86213.000000
mean         4.592964e+07  ...                    3257.049157
std          2.246004e+08  ...                    3934.312440
min          0.000000e+00  ...                       0.000000
25%          5.264100e+05  ...                     636.000000
50%          3.590096e+06  ...                    2050.000000
75%          1.701230e+07  ...                    4682.000000
max          3.263129e+09  ...                  117497.000000

[8 rows x 9 columns]
```

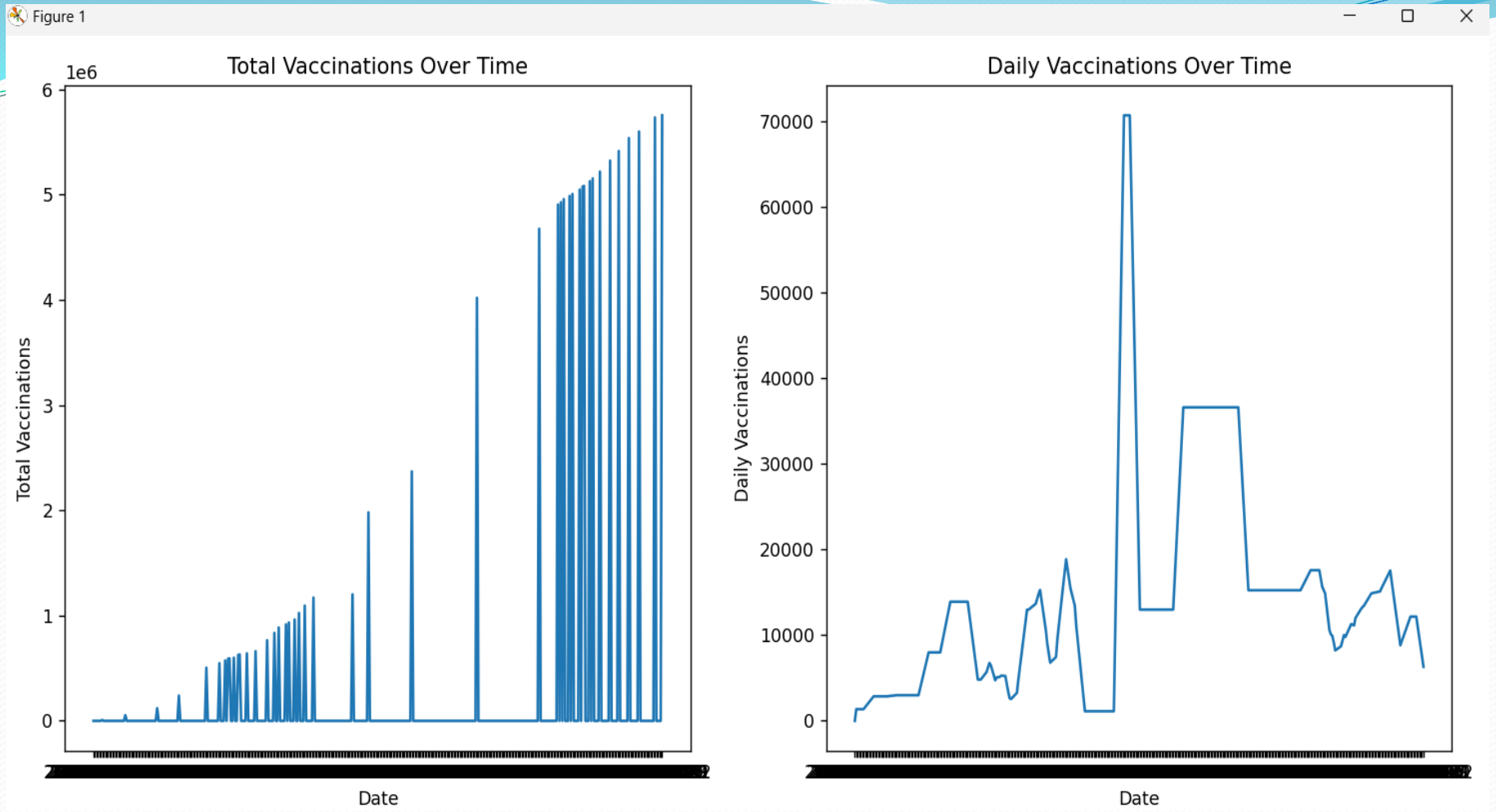
Figure 1
Total Vaccinations Over Time
1e6
6
5
4
3
2
1
0
Total Vaccinations
Date
Daily Vaccinations Over Time
70000
60000
50000
40000
30000
20000
10000
0
Daily Vaccinations
Date

CONCLUSION

The analysis of the COVID-19 vaccine dataset has provided valuable insights into the global vaccination effort. It is evident that vaccination progress is influenced by a combination of factors, including vaccine availability, distribution strategies, and regional disparities in healthcare resources.

To improve vaccination rates worldwide and ensure equitable access to vaccines, policymakers and public health officials should consider the following:

- Continuously monitor and adjust vaccination distribution strategies to address disparities.
- Promote public awareness and confidence in vaccines to encourage higher uptake.
- Collaborate with international organizations to ensure the availability of vaccines in underserved regions.
- Use data-driven insights to optimize vaccination campaigns and target high-risk populations.

This analysis serves as a foundation for further research and policy decisions aimed at effectively combatting the COVID-19 pandemic and achieving global vaccination goals.