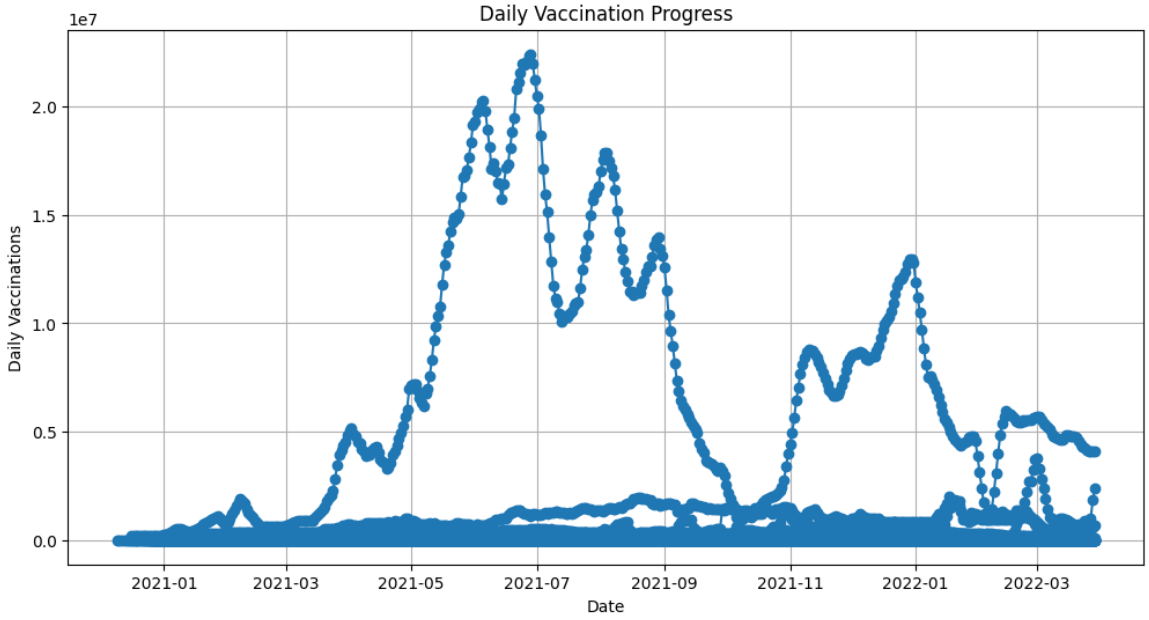
In [1]:

```
# Step 1: Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt

# Step 2: Load the dataset
# Make sure you have downloaded the dataset from the provided Kagg
# and specify the correct file path when reading it.
data = pd.read_csv('country_vaccinations.csv')  # Replace with you

# Step 3: Explore the dataset
# You can start by checking the first few rows of the dataset to u
print(data.head())



# Check for missing values
missing_values = data.isnull().sum()
print("Missing Values:\n", missing_values)

# Fill missing values (if necessary)
# Example: data['column_name'].fillna(value, inplace=True)

# Convert date columns to datetime
data['date'] = pd.to_datetime(data['date'])



plt.figure(figsize=(12, 6))
plt.plot(data['date'], data['daily_vaccinations'], marker='o', lin
plt.title('Daily Vaccination Progress')
plt.xlabel('Date')
plt.ylabel('Daily Vaccinations')
plt.grid()
plt.show()
```

```
       country iso_code        date  total_vaccinations  people_vaccinated  \
0  Afghanistan      AFG  2021-02-22                 0.0                 0.0
1  Afghanistan      AFG  2021-02-23                 NaN                 NaN
2  Afghanistan      AFG  2021-02-24                 NaN                 NaN
3  Afghanistan      AFG  2021-02-25                 NaN                 NaN
4  Afghanistan      AFG  2021-02-26                 NaN                 NaN

   people_fully_vaccinated  daily_vaccinations_raw  daily_vaccinations  \
0                      NaN                     NaN                 NaN
1                      NaN                     NaN              1367.0
2                      NaN                     NaN              1367.0
3                      NaN                     NaN              1367.0
4                      NaN                     NaN              1367.0

   total_vaccinations_per_hundred  people_vaccinated_per_hundred  \
0                             0.0                            0.0
1                             NaN                            NaN
```

```
2                                   NaN                                   NaN
3                                   NaN                                   NaN
4                                   NaN                                   NaN

   people_fully_vaccinated_per_hundred  daily_vaccinations_per_million  \
0                                  NaN                             NaN
1                                  NaN                            34.0
2                                  NaN                            34.0
3                                  NaN                            34.0
4                                  NaN                            34.0

                                        vaccines  \
0  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
1  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
2  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
3  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
4  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...

                 source_name            source_website
0  World Health Organization  https://covid19.who.int/
1  World Health Organization  https://covid19.who.int/
2  World Health Organization  https://covid19.who.int/
3  World Health Organization  https://covid19.who.int/
4  World Health Organization  https://covid19.who.int/
Missing Values:
 country                                0
iso_code                                0
date                                    0
total_vaccinations                  13859
people_vaccinated                   14665
people_fully_vaccinated             15532
daily_vaccinations_raw              16487
daily_vaccinations                     89
total_vaccinations_per_hundred      13859
people_vaccinated_per_hundred       14665
people_fully_vaccinated_per_hundred 15532
daily_vaccinations_per_million         89
vaccines                                0
source_name                             0
source_website                          0
dtype: int64
```



In [2]:

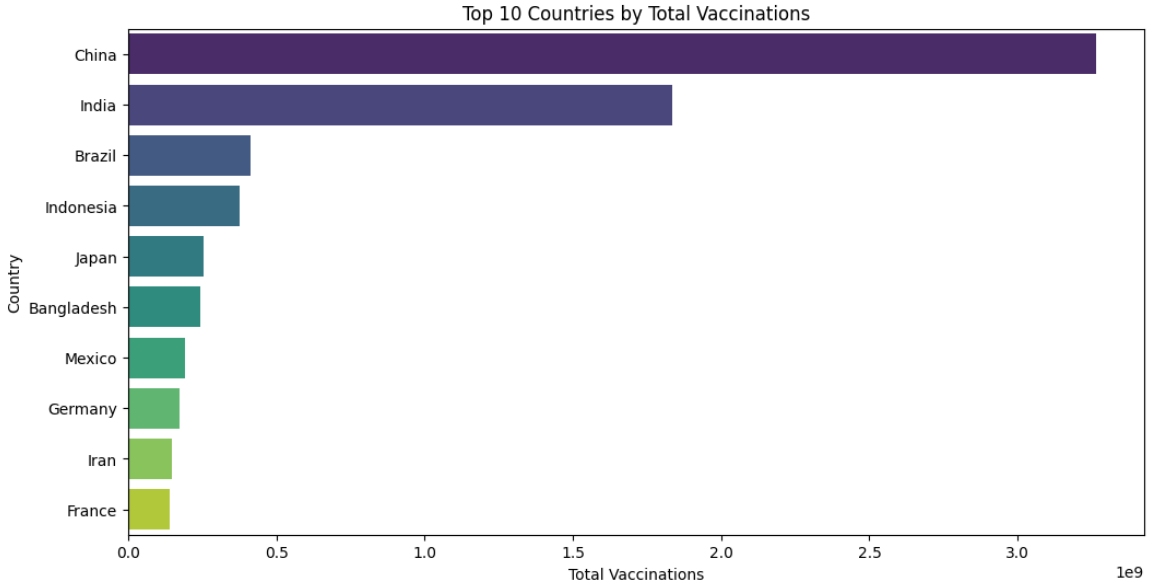
```
import seaborn as sns

# Load the dataset
data = pd.read_csv("country_vaccinations.csv")

# Display basic information about the dataset
```

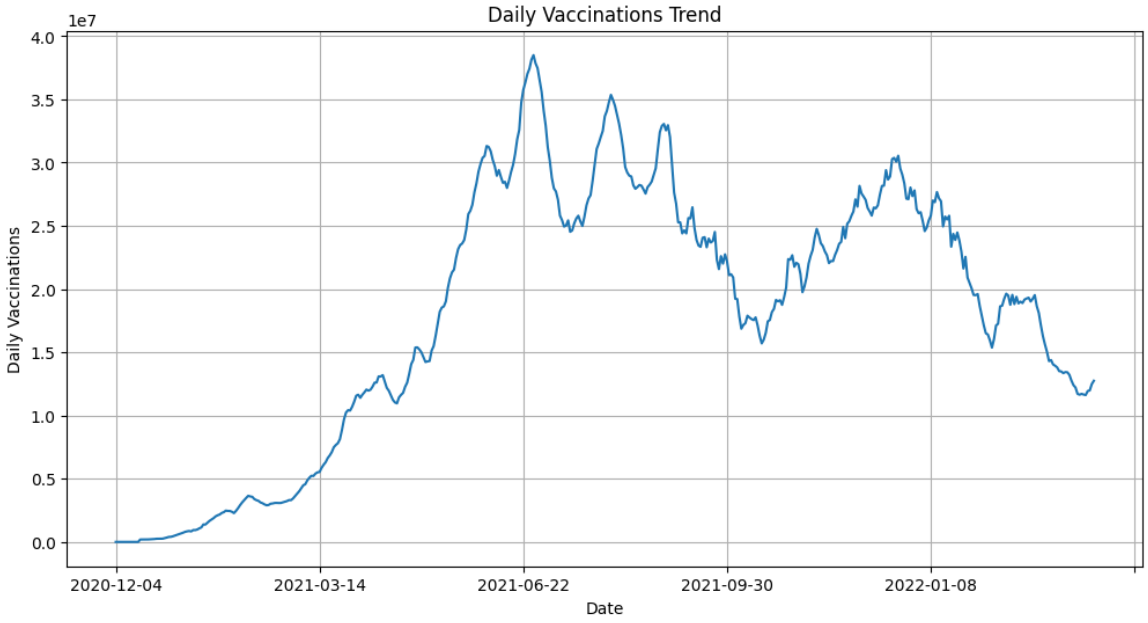
```
print("Basic Information About the Dataset:")
print(data.info())

# Display the first few rows of the dataset
print("\nFirst 5 Rows of the Dataset:")
print(data.head())

# Exploratory Data Analysis
# Summary statistics
print("\nSummary Statistics:")
print(data.describe())

# Check for missing values
print("\nMissing Values:")
print(data.isnull().sum())

# Statistical Analysis
# Calculate the total vaccinations per country
total_vaccinations_by_country = data.groupby('country')['total_vac
print("\nTotal Vaccinations by Country:")
print(total_vaccinations_by_country.head(10))

# Visualizations
# Plot total vaccinations by country
plt.figure(figsize=(12, 6))
sns.barplot(x=total_vaccinations_by_country.head(10), y=total_vacc
plt.title('Top 10 Countries by Total Vaccinations')
plt.xlabel('Total Vaccinations')
plt.ylabel('Country')
plt.show()

# Plot daily vaccinations trend
daily_vaccinations = data.groupby('date')['daily_vaccinations'].su
plt.figure(figsize=(12, 6))
daily_vaccinations.plot()
plt.title('Daily Vaccinations Trend')
plt.xlabel('Date')
plt.ylabel('Daily Vaccinations')
plt.grid(True)
plt.show()
```

```
Basic Information About the Dataset:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 57860 entries, 0 to 57859
Data columns (total 15 columns):
 #   Column                               Non-Null Count  Dtype
---  ------                               --------------  -----
 0   country                              57860 non-null  object
 1   iso_code                             57860 non-null  object
 2   date                                 57860 non-null  object
 3   total_vaccinations                   28539 non-null  float64
 4   people_vaccinated                    27516 non-null  float64
```

```
 5   people_fully_vaccinated              25712 non-null  float64
 6   daily_vaccinations_raw               22833 non-null  float64
 7   daily_vaccinations                   57640 non-null  float64
 8   total_vaccinations_per_hundred       28539 non-null  float64
 9   people_vaccinated_per_hundred        27516 non-null  float64
 10  people_fully_vaccinated_per_hundred  25712 non-null  float64
 11  daily_vaccinations_per_million       57640 non-null  float64
 12  vaccines                             57859 non-null  object
 13  source_name                          57859 non-null  object
 14  source_website                       57859 non-null  object
dtypes: float64(9), object(6)
memory usage: 6.6+ MB
None

First 5 Rows of the Dataset:
       country iso_code        date  total_vaccinations  people_vaccinated  \
0  Afghanistan      AFG  2021-02-22                 0.0                0.0
1  Afghanistan      AFG  2021-02-23                 NaN                NaN
2  Afghanistan      AFG  2021-02-24                 NaN                NaN
3  Afghanistan      AFG  2021-02-25                 NaN                NaN
4  Afghanistan      AFG  2021-02-26                 NaN                NaN

   people_fully_vaccinated  daily_vaccinations_raw  daily_vaccinations  \
0                      NaN                     NaN                 NaN
1                      NaN                     NaN              1367.0
2                      NaN                     NaN              1367.0
3                      NaN                     NaN              1367.0
4                      NaN                     NaN              1367.0

   total_vaccinations_per_hundred  people_vaccinated_per_hundred  \
0                             0.0                            0.0
1                             NaN                            NaN
2                             NaN                            NaN
3                             NaN                            NaN
4                             NaN                            NaN

   people_fully_vaccinated_per_hundred  daily_vaccinations_per_million  \
0                                  NaN                             NaN
1                                  NaN                            34.0
2                                  NaN                            34.0
3                                  NaN                            34.0
4                                  NaN                            34.0

                                      vaccines  \
0  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
1  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
2  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
3  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...
4  Johnson&Johnson, Oxford/AstraZeneca, Pfizer/Bi...

                source_name            source_website
0  World Health Organization  https://covid19.who.int/
1  World Health Organization  https://covid19.who.int/
2  World Health Organization  https://covid19.who.int/
3  World Health Organization  https://covid19.who.int/
4  World Health Organization  https://covid19.who.int/

Summary Statistics:
       total_vaccinations  people_vaccinated  people_fully_vaccinated  \
count        2.853900e+04       2.751600e+04             2.571200e+04
mean         5.439970e+07       1.845563e+07             1.431103e+07
std          2.724200e+08       8.271212e+07             6.635533e+07
min          0.000000e+00       0.000000e+00             1.000000e+00
25%          3.946855e+05       2.783510e+05             1.747380e+05
50%          2.600550e+06       1.620988e+06             1.196778e+06
75%          1.476710e+07       7.269298e+06             6.123212e+06
max          3.263129e+09       1.275541e+09             1.240777e+09

       daily_vaccinations_raw  daily_vaccinations  \
count            2.283300e+04        5.764000e+04
mean             3.307485e+05        1.504113e+05
std              1.482829e+06        9.232620e+05
min              0.000000e+00        0.000000e+00
25%              3.317000e+03        7.610000e+02
```

```
50%              2.006400e+04        5.739000e+03
75%              1.111570e+05        3.663500e+04
max              2.474100e+07        2.242429e+07

       total_vaccinations_per_hundred  people_vaccinated_per_hundred  \
count                    28539.000000                   27516.000000
mean                        80.300755                      41.196622
std                         68.197795                      29.429686
min                          0.000000                       0.000000
25%                         15.910000                      11.477500
50%                         68.220000                      41.910000
75%                        132.430000                      67.772500
max                        345.370000                     124.760000

       people_fully_vaccinated_per_hundred  daily_vaccinations_per_million
count                         25712.000000                    57640.000000
mean                             35.639925                     3228.340475
std                              28.526565                     4149.646875
min                               0.000000                        0.000000
25%                               7.190000                      620.000000
50%                              31.570000                     1955.000000
75%                              62.262500                     4547.500000
max                             122.370000                   117497.000000

Missing Values:
country                                  0
iso_code                                 0
date                                     0
total_vaccinations                   29321
people_vaccinated                    30344
people_fully_vaccinated              32148
daily_vaccinations_raw               35027
daily_vaccinations                     220
total_vaccinations_per_hundred       29321
people_vaccinated_per_hundred        30344
people_fully_vaccinated_per_hundred  32148
daily_vaccinations_per_million         220
vaccines                                 1
source_name                              1
source_website                           1
dtype: int64

Total Vaccinations by Country:
country
China         3.263129e+09
India         1.834501e+09
Brazil        4.135596e+08
Indonesia     3.771089e+08
Japan         2.543456e+08
Bangladesh    2.436427e+08
Mexico        1.919079e+08
Germany       1.719400e+08
Iran          1.467926e+08
France        1.416662e+08
Name: total_vaccinations, dtype: float64
```


Top 10 Countries by Total Vaccinations
China
India
Brazil
Indonesia
Japan
Bangladesh
Mexico
Germany
Iran
France
Country
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Total Vaccinations
1e9


Daily Vaccinations Trend
1e7
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Daily Vaccinations
2020-12-04
2021-03-14
2021-06-22
2021-09-30
2022-01-08
Date