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

```
df = pd.read_csv("covid_case.csv")

# Display basic information
print("Dataset Shape:", df.shape)
print("\nColumn Names:", list(df.columns))
print("\nTotal Records:", len(df))
```

```
Dataset Shape: (1200, 6)

Column Names: ['Date', 'Country', 'Confirmed', 'Deaths', 'Recovered', 'Active']

Total Records: 1200
```

```
print("\nMissing values per column:")
print(df.isnull().sum())
```

```
Missing values per column:
Date         0
Country      0
Confirmed    0
Deaths       0
Recovered    0
Active       0
dtype: int64
```

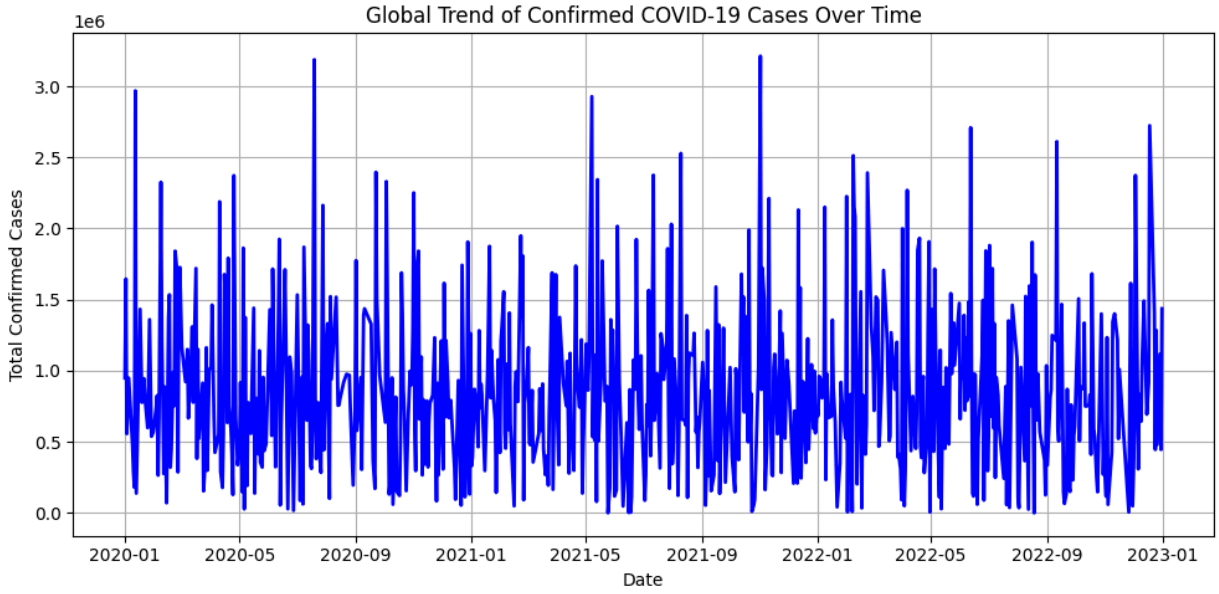
```
df = df.fillna(0)
```

```
print("\nStatistical Summary:")
print(df.describe())
```

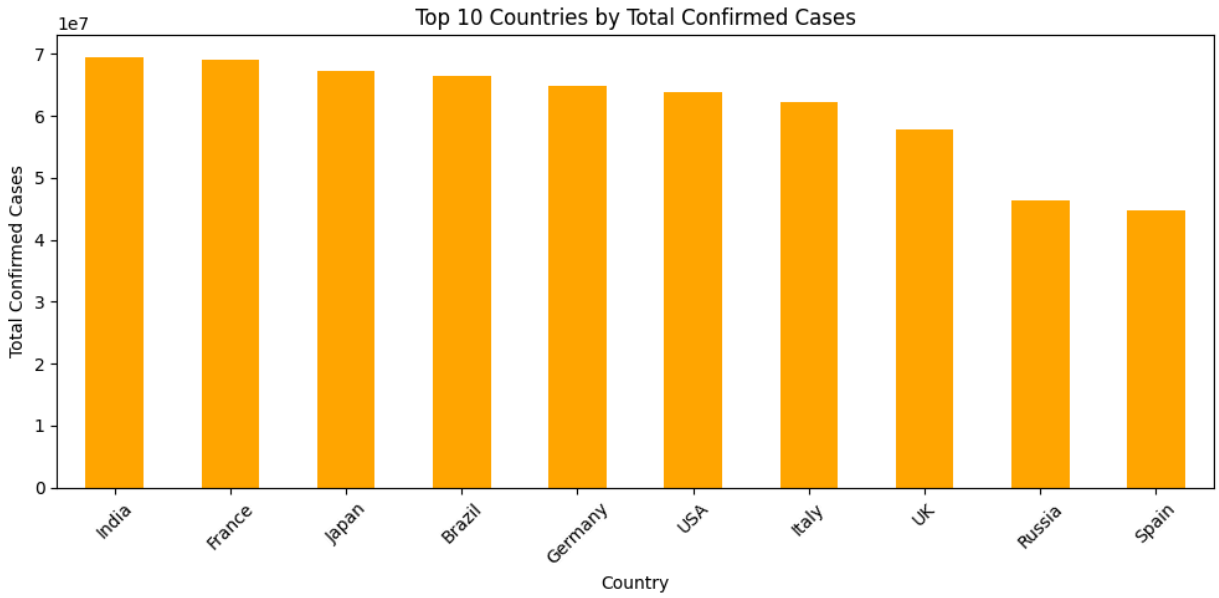
```
Statistical Summary:
           Confirmed        Deaths      Recovered         Active
count    1200.000000   1200.000000    1200.000000    1200.000000
mean   510035.990000   8778.575000  394193.676667  107063.738333
std    284907.522467   6406.087455  226903.461987   84097.944365
min      1215.000000      8.000000     831.000000     129.000000
25%    263086.500000   3637.750000  200101.000000   36993.250000
50%    514039.000000   7442.000000  388718.500000   84334.500000
75%    754526.500000  12686.500000  576461.500000  161058.250000
max    999653.000000  29119.000000  913816.000000  365420.000000
```

```
df['Date'] = pd.to_datetime(df['Date'], errors='coerce')
```

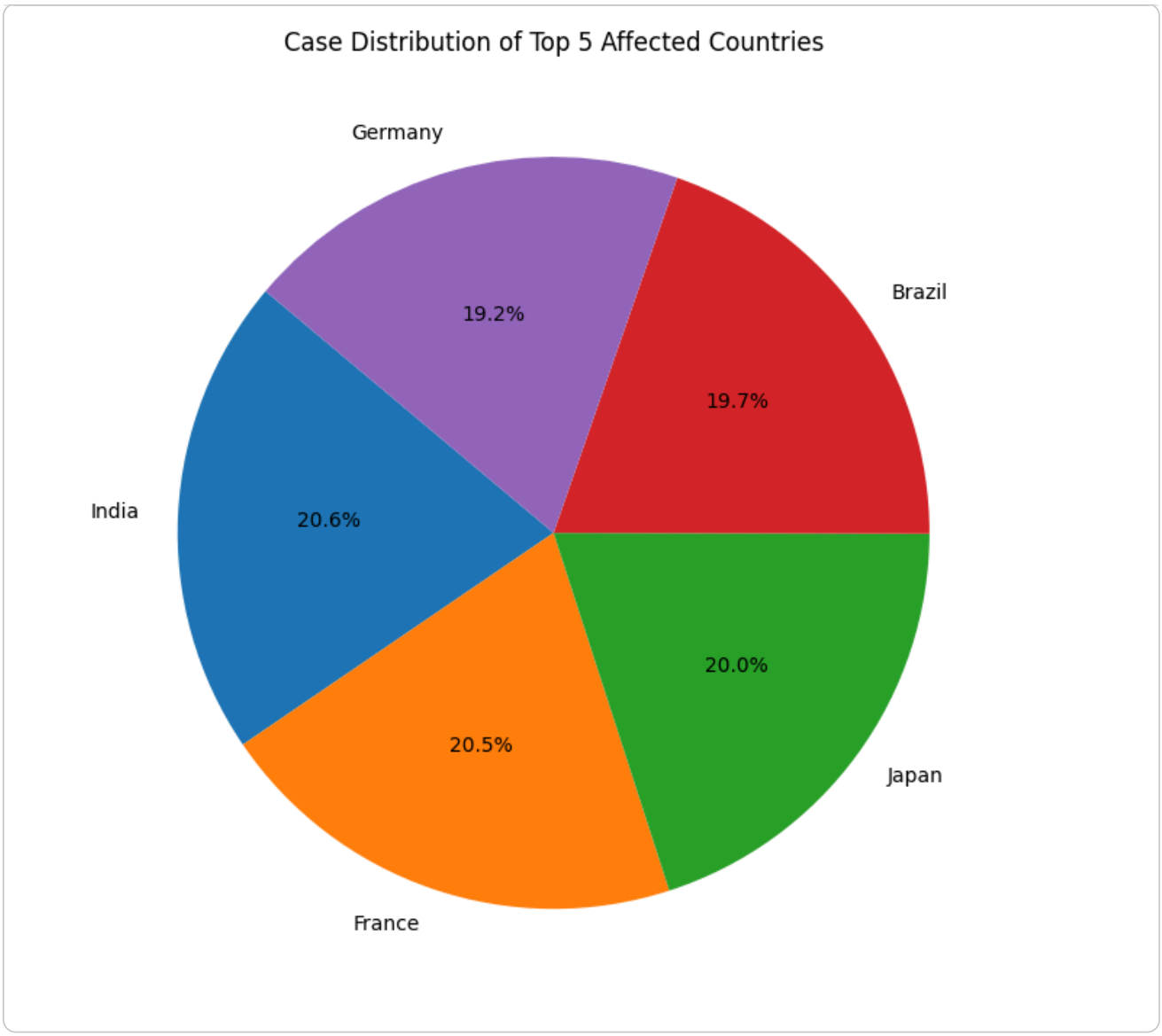
```
global_cases = df.groupby('Date')['Confirmed'].sum()
plt.figure(figsize=(10,5))
plt.plot(global_cases.index, global_cases.values, color='blue', linewidth=2)
plt.title('Global Trend of Confirmed COVID-19 Cases Over Time')
plt.xlabel('Date')
plt.ylabel('Total Confirmed Cases')
plt.grid(True)
plt.tight_layout()
plt.show()
```



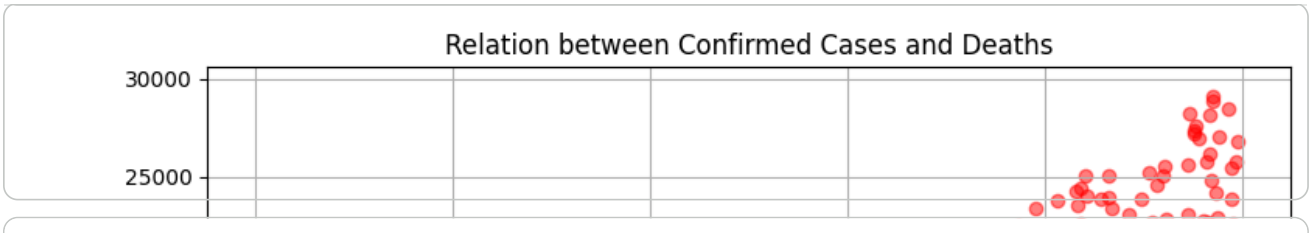

```
top10 = df.groupby('Country')['Confirmed'].sum().sort_values(ascending=False).head(10)
plt.figure(figsize=(10,5))
top10.plot(kind='bar', color='orange')
plt.title('Top 10 Countries by Total Confirmed Cases')
plt.xlabel('Country')
plt.ylabel('Total Confirmed Cases')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



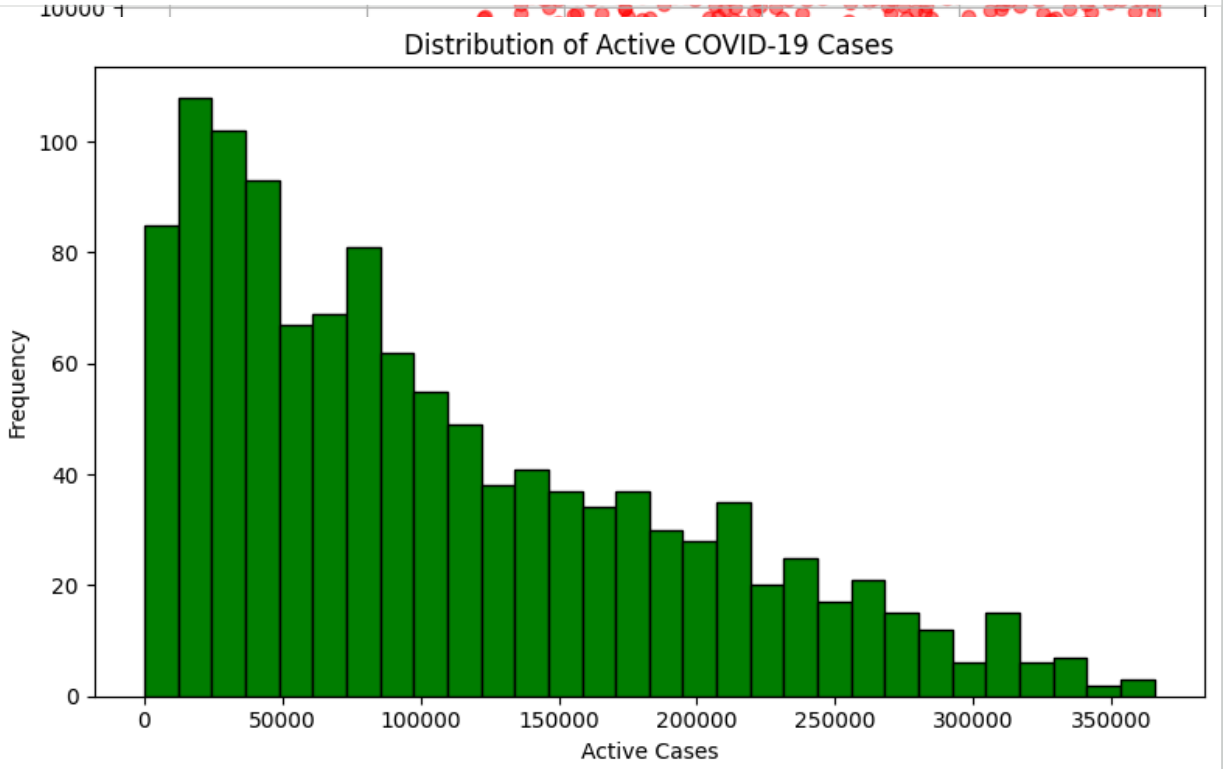

```
top5 = df.groupby('Country')['Confirmed'].sum().sort_values(ascending=False).head(5)
plt.figure(figsize=(7,7))
plt.pie(top5.values, labels=top5.index, autopct='%1.1f%%', startangle=140)
plt.title('Case Distribution of Top 5 Affected Countries')
plt.tight_layout()
plt.show()
```




```
plt.figure(figsize=(8,5))
plt.scatter(df['Confirmed'], df['Deaths'], alpha=0.5, color='red')
plt.title('Relation between Confirmed Cases and Deaths')
plt.xlabel('Confirmed Cases')
plt.ylabel('Deaths')
plt.grid(True)
plt.tight_layout()
plt.show()
```




```
plt.figure(figsize=(8,5))
plt.hist(df['Active'], bins=30, color='green', edgecolor='black')
plt.title('Distribution of Active COVID-19 Cases')
plt.xlabel('Active Cases')
plt.ylabel('Frequency')
plt.tight_layout()
plt.show()
```




```
print("\nAnalysis and visualization completed successfully.")
```

```
Analysis and visualization completed successfully.
```