COVID-19 Coronavirus data set.

```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

df  = pd.read_csv(r"C:\Users\shagu\Downloads\COVID-19
Coronavirus.csv")
df
```

	Country	Other names	ISO 3166-1 alpha-3 CODE
0	Afghanistan	Afghanistan	AFG
1	Albania	Albania	ALB
2	Algeria	Algeria	DZA
3	Andorra	Andorra	AND
4	Angola	Angola	AGO
..	...	...	...
220	Wallis and Futuna	Wallis and Futuna Islands	WLF
221	Western Sahara	Western Sahara	ESHÂ
222	Yemen	Yemen	YEM
223	Zambia	Zambia	ZMB
224	Zimbabwe	Zimbabwe	ZWE

	Population	Continent	Total Cases	Total Deaths	Tot Cases//1M pop
0	40462186	Asia	177827	7671	4395
1	2872296	Europe	273870	3492	95349
2	45236699	Africa	265691	6874	5873
3	77481	Europe	40024	153	516565
4	34654212	Africa	99194	1900	2862

```
..         ...     ...          ...              ...              ..
.
220       10894   Oceania      454               7
41674
221       623031  Africa       10                1
16
222       30975258  Asia       11806             2143
381
223       19284482  Africa     317076            3967
16442
224       15241601  Africa     246525            5446
16174

     Tot Deaths/1M pop  Death percentage
0                  190          4.313743
1                 1216          1.275058
2                  152          2.587216
3                 1975          0.382271
4                   55          1.915438
..                 ...               ...
220                643          1.541850
221                  2         10.000000
222                 69         18.151787
223                206          1.251120
224                357          2.209107

[225 rows x 10 columns]

df.head()

        Country  Other names ISO 3166-1 alpha-3 CODE  Population
Continent  \
0  Afghanistan  Afghanistan                     AFG    40462186
Asia
1       Albania      Albania                     ALB     2872296
Europe
2       Algeria      Algeria                     DZA    45236699
Africa
3       Andorra      Andorra                     AND       77481
Europe
4        Angola       Angola                     AGO    34654212
Africa

   Total Cases  Total Deaths  Tot Cases//1M pop  Tot Deaths/1M pop  \
0       177827          7671               4395                190
1       273870          3492              95349               1216
2       265691          6874               5873                152
3        40024           153             516565               1975
4        99194          1900               2862                 55
```

```
    Death percentage
0           4.313743
1           1.275058
2           2.587216
3           0.382271
4           1.915438
```

df.tail()

```
               Country             Other names ISO 3166-1 alpha-3
CODE  \
220  Wallis and Futuna  Wallis and Futuna Islands
WLF
221      Western Sahara              Western Sahara
ESHÂ
222               Yemen                       Yemen
YEM
223              Zambia                      Zambia
ZMB
224            Zimbabwe                    Zimbabwe
ZWE

     Population Continent  Total Cases  Total Deaths  Tot Cases//1M
pop  \
220       10894    Oceania          454             7
41674
221      623031     Africa           10             1
16
222    30975258       Asia        11806          2143
381
223    19284482     Africa       317076          3967
16442
224    15241601     Africa       246525          5446
16174

     Tot Deaths/1M pop  Death percentage
220                643          1.541850
221                  2         10.000000
222                 69         18.151787
223                206          1.251120
224                357          2.209107
```

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 225 entries, 0 to 224
Data columns (total 10 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Country               225 non-null    object
```

```
 1    Other names              224 non-null    object
 2    ISO 3166-1 alpha-3 CODE  225 non-null    object
 3    Population               225 non-null    int64
 4    Continent                225 non-null    object
 5    Total Cases              225 non-null    int64
 6    Total Deaths             225 non-null    int64
 7    Tot Cases//1M pop        225 non-null    int64
 8    Tot Deaths/1M pop        225 non-null    int64
 9    Death percentage         225 non-null    float64
dtypes: float64(1), int64(5), object(4)
memory usage: 17.7+ KB
```

df.describe()

```
        Population    Total Cases    Total Deaths    Tot Cases//1M pop  \
count   2.250000e+02  2.250000e+02   2.250000e+02           225.000000
mean    3.507321e+07  2.184781e+06   2.744813e+04        136900.373333
std     1.392418e+08  7.275938e+06   9.689177e+04        145060.340289
min     8.050000e+02  1.000000e+00   0.000000e+00             9.000000
25%     5.665570e+05  2.407100e+04   1.890000e+02         11384.000000
50%     5.827911e+06  1.639360e+05   1.965000e+03         88987.000000
75%     2.190585e+07  1.092547e+06   1.366000e+04        223335.000000
max     1.439324e+09  8.183905e+07   1.008222e+06        696044.000000
```

```
        Tot Deaths/1M pop  Death percentage
count          225.000000        225.000000
mean          1096.715556          1.444125
std           1195.715543          1.741728
min              0.000000          0.000000
25%            123.000000          0.511291
50%            708.000000          1.036905
75%           1795.000000          1.977017
max           6286.000000         18.151787
```

df.isnull().sum()

```
Country                   0
Other names               1
ISO 3166-1 alpha-3 CODE   0
Population                0
Continent                 0
Total Cases               0
Total Deaths              0
Tot Cases//1M pop         0
Tot Deaths/1M pop         0
Death percentage          0
dtype: int64
```

df.dropna()

```
                 Country             Other names ISO 3166-1 alpha-3
CODE    \
0           Afghanistan             Afghanistan
AFG
1               Albania                 Albania
ALB
2               Algeria                 Algeria
DZA
3               Andorra                 Andorra
AND
4                Angola                  Angola
AGO
..                  ...                     ...
...
220  Wallis and Futuna  Wallis and Futuna Islands
WLF
221      Western Sahara          Western Sahara
ESHÂ
222                Yemen                   Yemen
YEM
223                Zambia                  Zambia
ZMB
224              Zimbabwe                Zimbabwe
ZWE

     Population Continent  Total Cases  Total Deaths  Tot Cases//1M
pop   \
0      40462186      Asia       177827          7671
4395
1       2872296    Europe       273870          3492
95349
2      45236699    Africa       265691          6874
5873
3         77481    Europe        40024           153
516565
4      34654212    Africa        99194          1900
2862
..          ...       ...          ...           ...
.
220       10894   Oceania          454             7
41674
221      623031    Africa           10             1
16
222    30975258      Asia        11806          2143
381
223    19284482    Africa       317076          3967
16442
224    15241601    Africa       246525          5446
16174
```

```
     Tot Deaths/1M pop  Death percentage
0                  190          4.313743
1                 1216          1.275058
2                  152          2.587216
3                 1975          0.382271
4                   55          1.915438
..                 ...               ...
220                643          1.541850
221                  2         10.000000
222                 69         18.151787
223                206          1.251120
224                357          2.209107

[224 rows x 10 columns]
```

```
df.shape
```

```
(225, 10)
```

```
df1 = df[df.isna().any(axis=1)]
```

```
df1
```

```
        Country Other names ISO 3166-1 alpha-3 CODE   Population
Continent  \
135  Montenegro         NaN                      MNE       628205
Europe

     Total Cases  Total Deaths  Tot Cases//1M pop  Tot Deaths/1M
pop  \
135       233326          2705             371417               4306


     Death percentage
135          1.159322
```

```
newdf = df['Other names'].fillna('Montenegro', inplace = True)
newdf
```

```
df1 = df[df.isna().any(axis=1)]
df1
```

```
Empty DataFrame
Columns: [Country, Other names, ISO 3166-1 alpha-3 CODE, Population,
Continent, Total Cases, Total Deaths, Tot Cases//1M pop, Tot Deaths/1M
pop, Death percentage]
Index: []
```

```
df.head(136)
```

```
            Country                     Other names ISO 3166-1 alpha-3
CODE   \
0     Afghanistan                       Afghanistan
AFG
1         Albania                           Albania
ALB
2         Algeria                           Algeria
DZA
3          Andorra                          Andorra
AND
4           Angola                           Angola
AGO
..           ...                               ...
...
131   Micronesia  Micronesia (Federated States of)
FSM
132      Moldova               Republic of Moldova
MDA
133       Monaco                            Monaco
MCO
134      Mongolia                          Mongolia
MNG
135   Montenegro                        Montenegro
MNE

     Population Continent  Total Cases  Total Deaths  Tot Cases//1M
pop  \
0      40462186      Asia       177827          7671
4395
1       2872296    Europe       273870          3492
95349
2      45236699    Africa       265691          6874
5873
3         77481    Europe        40024           153
516565
4      34654212    Africa        99194          1900
2862
..          ...       ...          ...           ...           ..
.
131      117134   Oceania            1             0
9
132     4017550    Europe       514199         11446
127988
133       39729    Europe        10842            54
272899
134     3370682      Asia       468610          2177
139025
135      628205    Europe       233326          2705
371417
```

```
     Tot Deaths/1M pop  Death percentage
0                  190          4.313743
1                 1216          1.275058
2                  152          2.587216
3                 1975          0.382271
4                   55          1.915438
..                 ...               ...
131                  0          0.000000
132               2849          2.225986
133               1359          0.498063
134                646          0.464565
135               4306          1.159322

[136 rows x 10 columns]
df.rename(columns={"ISO 3166-1 alpha-3 CODE": "Iso_Code"},
inplace=True)

df
```

	Country	Other names	Iso_Code	Population
0	Afghanistan	Afghanistan	AFG	40462186
1	Albania	Albania	ALB	2872296
2	Algeria	Algeria	DZA	45236699
3	Andorra	Andorra	AND	77481
4	Angola	Angola	AGO	34654212
..	...	...	...	...
220	Wallis and Futuna	Wallis and Futuna Islands	WLF	10894
221	Western Sahara	Western Sahara	ESHÂ	623031
222	Yemen	Yemen	YEM	30975258
223	Zambia	Zambia	ZMB	19284482
224	Zimbabwe	Zimbabwe	ZWE	15241601

	Continent	Total Cases	Total Deaths	Tot Cases//1M pop
0	Asia	177827	7671	4395
1	Europe	273870	3492	95349
2	Africa	265691	6874	5873
3	Europe	40024	153	516565

```
4      Africa          99194          1900          2862
..        ...            ...           ...           ...
220   Oceania           454             7          41674
221    Africa            10             1             16
222      Asia         11806          2143           381
223    Africa        317076          3967          16442
224    Africa        246525          5446          16174

     Tot Deaths/1M pop  Death percentage
0                  190          4.313743
1                 1216          1.275058
2                  152          2.587216
3                 1975          0.382271
4                   55          1.915438
..                 ...               ...
220                643          1.541850
221                  2         10.000000
222                 69         18.151787
223                206          1.251120
224                357          2.209107

[225 rows x 10 columns]

Country = df.nlargest(5, 'Total Cases')
Country.head()

      Country                 Other names  Iso_Code   Population  \
214       USA   United States of America       USA    334400597
92      India                      India       IND   1403754381
26     Brazil                     Brazil       BRA    215204501
70     France                     France       FRA     65526369
76    Germany                    Germany       DEU     84252947

                             Continent   Total Cases   Total Deaths  \
214                  Northern America      81839052        1008222
92                            Asia      43029044         521388
26    Latin America and the Caribbean      29999816         660269
70                           Europe      25997852         142506
76                           Europe      21646375         130563

     Tot Cases//1M pop  Tot Deaths/1M pop  Death percentage
214             244734               3015          1.231957
92               30653                371          1.211712
26              139401               3068          2.200910
70              396754               2175          0.548145
76              256921               1550          0.603163

plt.figure(figsize =(12,6))
plt.bar(x=Country["Country"],height=Country["Total Cases"])
plt.xlabel("Country Names")
```
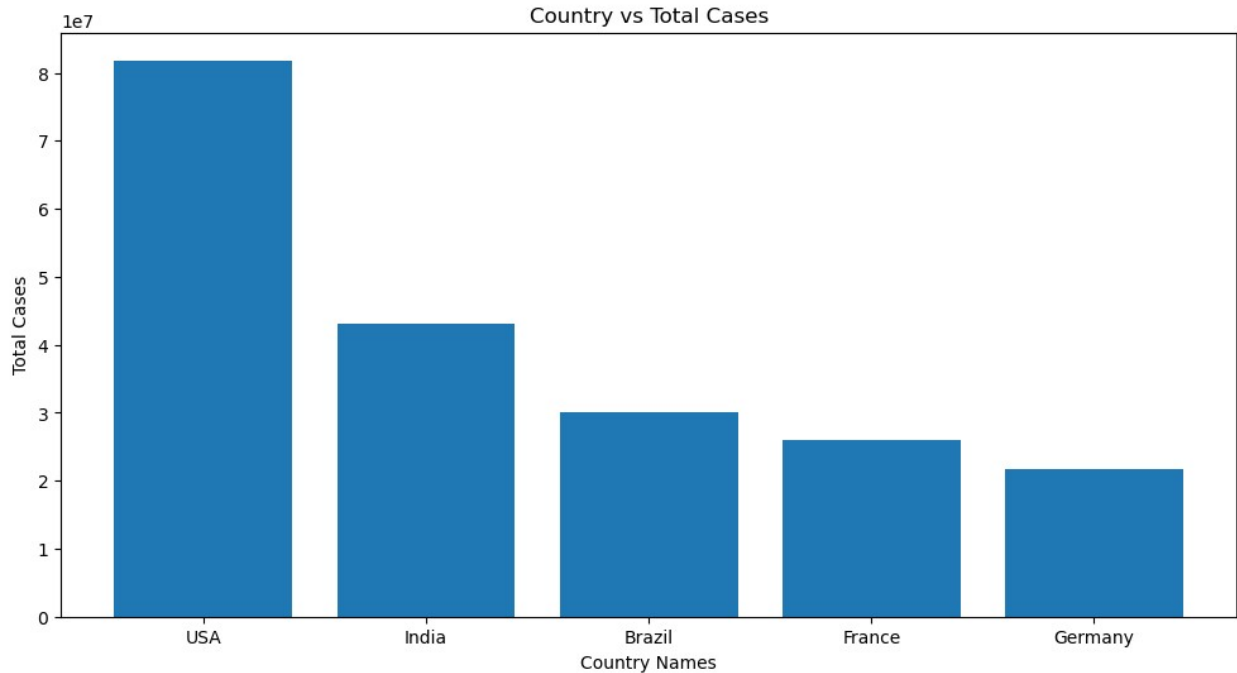
```python
plt.ylabel("Total Cases")
plt.title("Country vs Total Cases")
#plt.xticks(rotation = 90)
plt.show()
```



```python
Country = df.nsmallest(5, 'Total Cases')
Country.head()
```

```
            Country                 Other names Iso_Code
Population  \
131        Micronesia  Micronesia (Federated States of)       FSM
117134
168     Saint Helena                      Saint Helena       SHN
6109
125   Marshall Islands                  Marshall Islands       MHL
59889
148             Niue                              Niue       NIU
1645
221    Western Sahara                    Western Sahara       ESHÂ
623031

    Continent  Total Cases  Total Deaths  Tot Cases//1M pop  \
131   Oceania            1             0                  9
168    Africa            2             0                327
125   Oceania            7             0                117
148   Oceania            7             0               4255
221    Africa           10             1                 16
```

```
     Tot Deaths/1M pop  Death percentage
131                  0               0.0
168                  0               0.0
125                  0               0.0
148                  0               0.0
221                  2              10.0
```

```python
plt.figure(figsize =(12,6))
plt.bar(x=Country["Country"],height=Country["Total Cases"])
plt.xlabel("Country Names")
plt.ylabel("Total Cases")
plt.title("Country vs Total Cases")
#plt.xticks(rotation = 90)
plt.show()
```
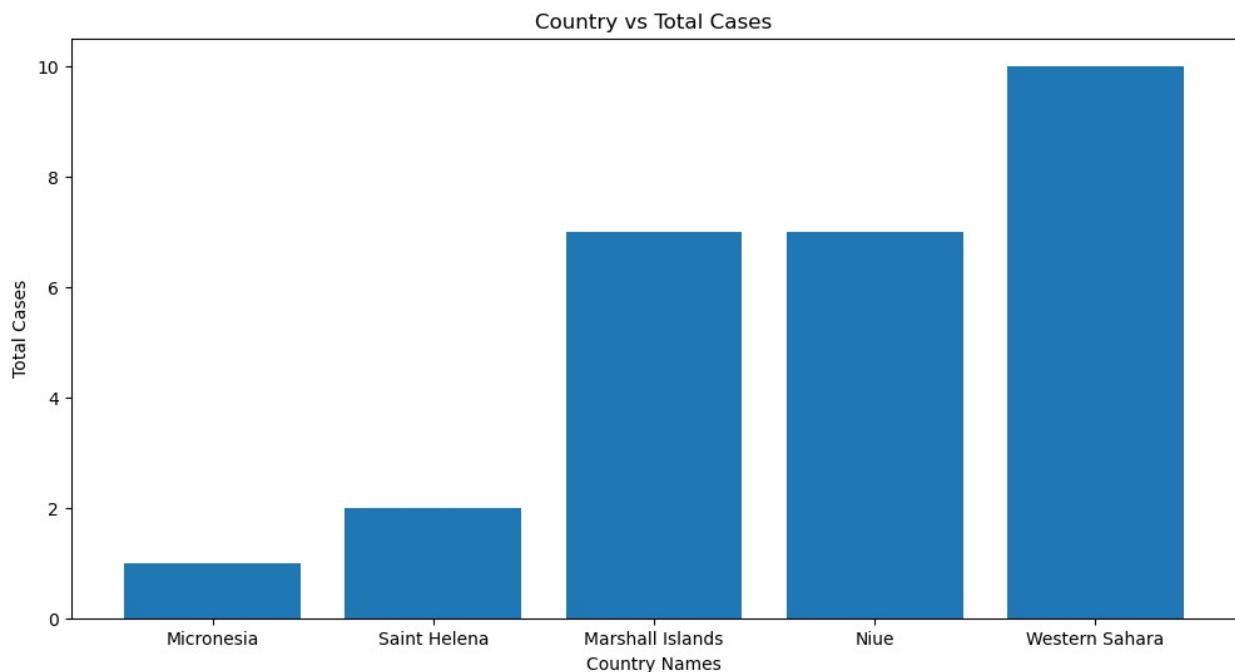


```python
Country = df.nlargest(5, 'Total Deaths')
Country.head()
```

```
     Country               Other names Iso_Code  Population  \
214      USA  United States of America      USA   334400597
26   Brazil                     Brazil      BRA   215204501
92    India                      India      IND  1403754381
165  Russia         Russian Federation      RUS   146044010
130  Mexico                     Mexico      MEX   131303955


                            Continent  Total Cases  Total Deaths  \
214                 Northern America     81839052       1008222
26    Latin America and the Caribbean     29999816        660269
92                              Asia     43029044        521388
```

```
165                             Europe       17896866         369708
130   Latin America and the Caribbean         5665376         323212


      Tot Cases//1M pop  Tot Deaths/1M pop  Death percentage
214             244734               3015          1.231957
26              139401               3068          2.200910
92               30653                371          1.211712
165             122544               2531          2.065770
130              43147               2462          5.705041
```

```python
plt.figure(figsize =(12,6))
plt.bar(x=Country["Country"],height=Country["Total Deaths"])
plt.xlabel("Country Names")
plt.ylabel("Total Deaths")
plt.title("Country vs Total Deaths")
#plt.xticks(rotation = 90)
plt.show()
```
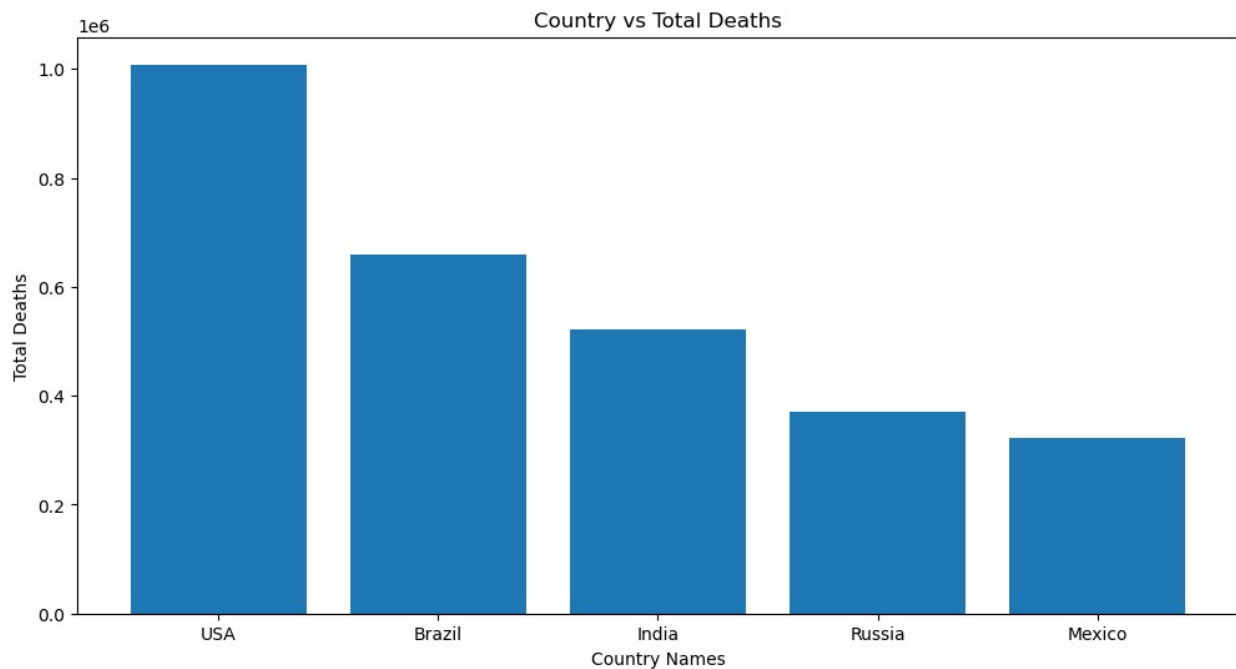


```python
Country = df.nsmallest(5, 'Total Deaths')
Country.head()
```

```
              Country                              Other names
Iso_Code  \
46        Cook Islands                            Cook Islands
COK
67    Falkland Islands            Falkland Islands (Malvinas)
FLK
118              Macao   China, Macao Special Administrative Region
MAC
```

```
125   Marshall Islands                           Marshall Islands
MHL
131        Micronesia          Micronesia (Federated States of)
FSM

      Population                          Continent  Total Cases  Total
Deaths  \
46         17592                           Oceania         2118
0
67          3657  Latin America and the Caribbean          123
0
118       664828                              Asia           82
0
125        59889                           Oceania            7
0
131       117134                           Oceania            1
0

      Tot Cases//1M pop  Tot Deaths/1M pop  Death percentage
46               120396                  0               0.0
67                33634                  0               0.0
118                 123                  0               0.0
125                 117                  0               0.0
131                   9                  0               0.0
```

```python
plt.figure(figsize =(12,6))
plt.bar(x=Country["Country"],height=Country["Total Deaths"])
plt.xlabel("Country Names")
plt.ylabel("Total Deaths")
plt.title("Country vs Total Deaths")
#plt.xticks(rotation = 90)
plt.show()
```

Country vs Total Deaths
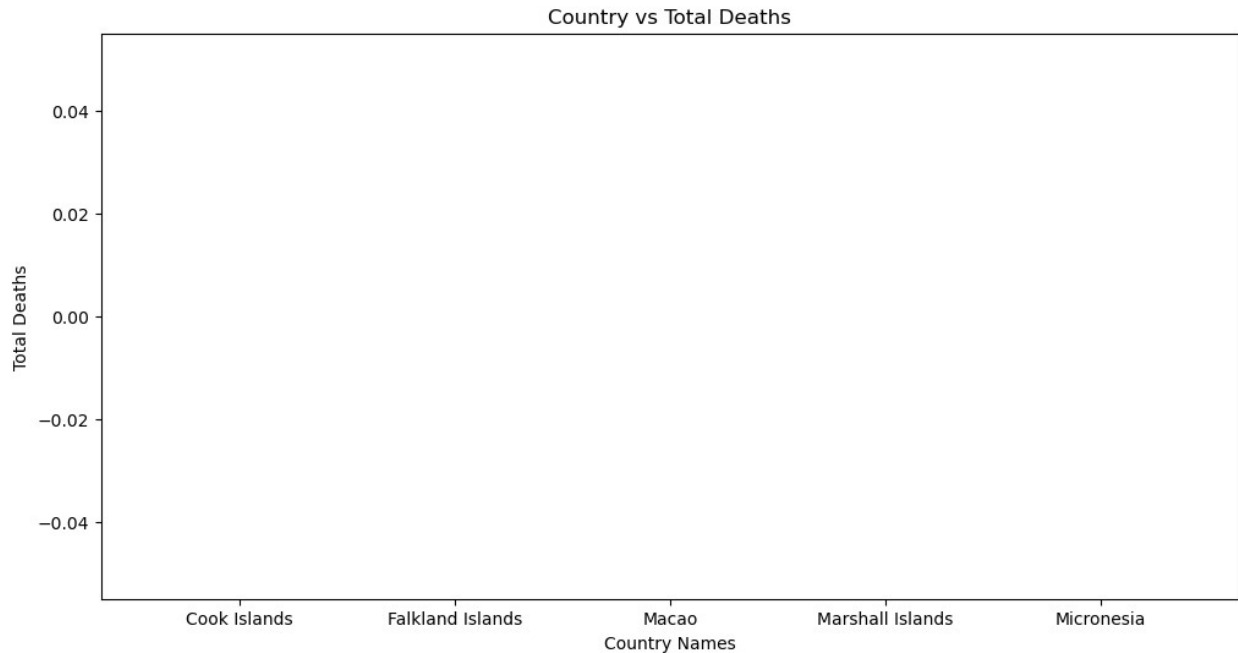

```python
Country = df.nlargest(5, 'Death percentage')
Country.head()
```

	Country	Other names	Iso_Code	Population	
222	Yemen	Yemen	YEM	30975258	
221	Western Sahara	Western Sahara	ESHÂ	623031	
193	Sudan	Sudan	SDN	45640385	
158	Peru	Peru	PER	33775745	
130	Mexico	Mexico	MEX	131303955	

	Continent	Total Cases	Total Deaths	
222	Asia	11806	2143	
221	Africa	10	1	
193	Africa	61955	4907	
158	Latin America and the Caribbean	3548559	212328	
130	Latin America and the Caribbean	5665376	323212	

	Tot Cases//1M pop	Tot Deaths/1M pop	Death percentage
222	381	69	18.151787
221	16	2	10.000000
193	1357	108	7.920265
158	105062	6286	5.983499
130	43147	2462	5.705041

```python
plt.figure(figsize =(12,6))
plt.bar(x=Country["Country"],height=Country["Death percentage"])
plt.xlabel("Country Names")
plt.ylabel("Death percentage")
plt.title("Country vs Death percentage")
```

```
#plt.xticks(rotation = 90)
plt.show()
```


Country vs Death percentage
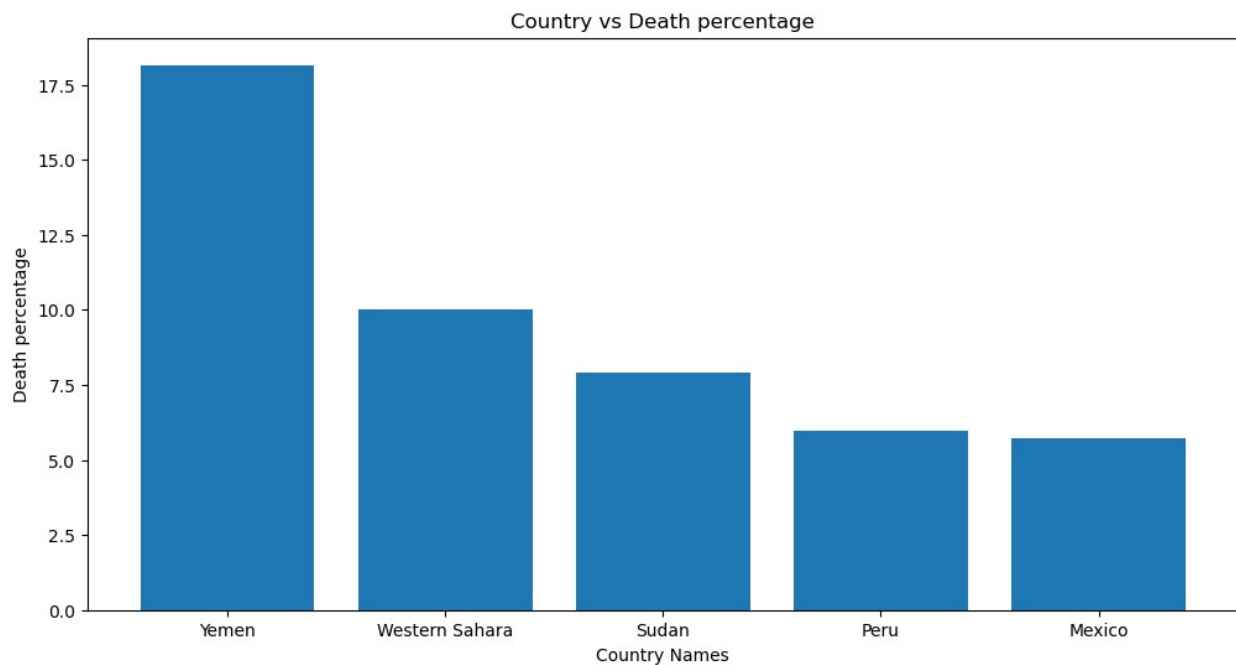
```
df2 = df.groupby('Continent')['Total Cases'].sum().reset_index()
df2
```

	Continent	Total Cases
0	Africa	11764207
1	Asia	140957179
2	Europe	180332483
3	Latin America and the Caribbean	67509231
4	Northern America	85364770
5	Oceania	5647957

```
plt.figure(figsize =(12,6))
plt.bar(x=df2['Continent'],height=df2['Total Cases'])
plt.xlabel("Continent")
plt.ylabel("Total Cases")
plt.title("Continent vs Total Cases")
plt.xticks(rotation = 45)
plt.show()
```
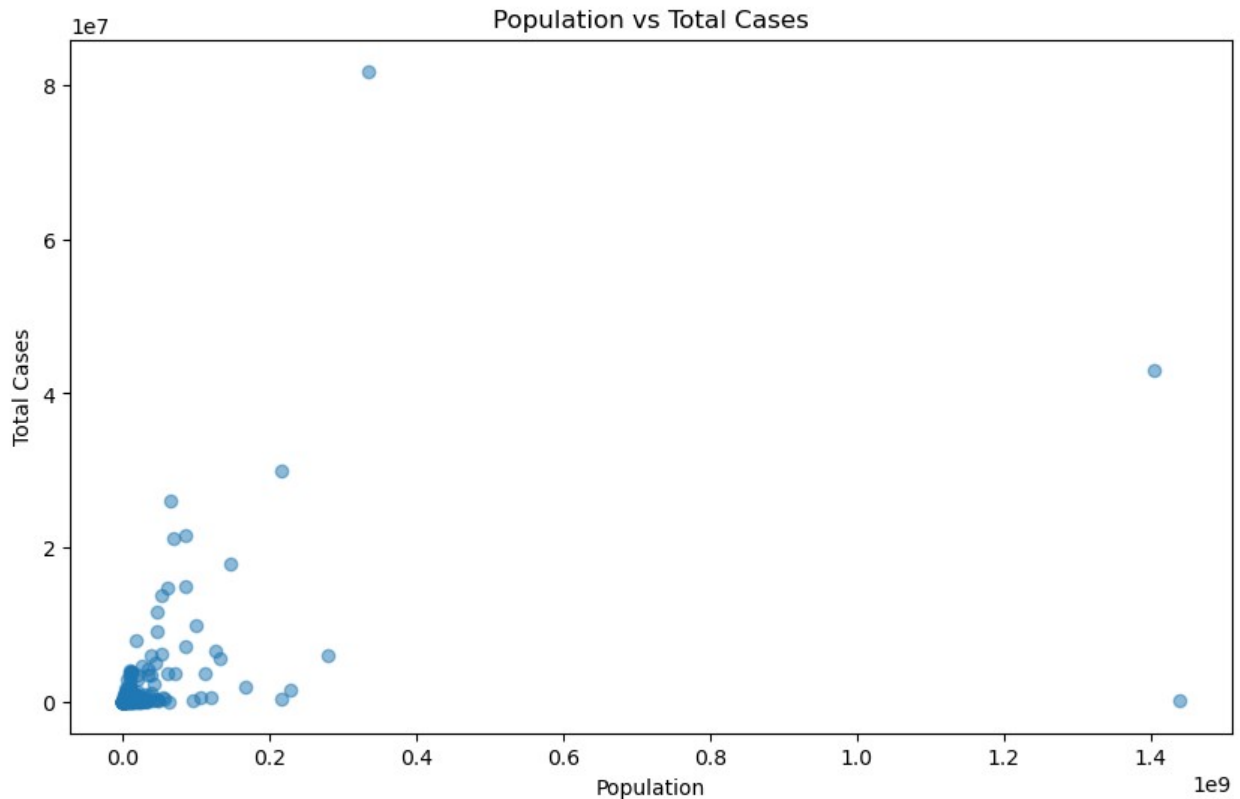
Continent vs Total Cases
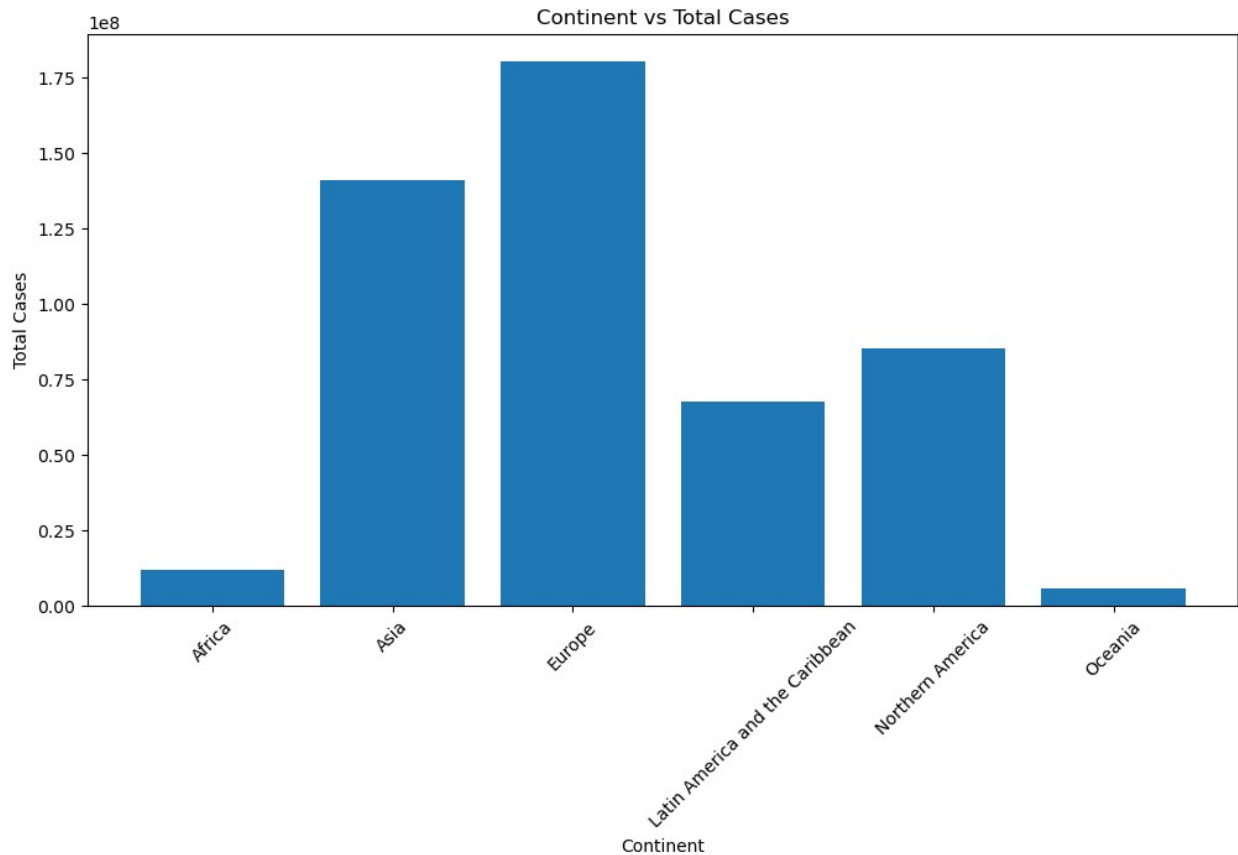
```
plt.figure(figsize=(10, 6))

# Create a scatter plot
plt.scatter(df["Population"], df["Total Cases"], alpha=0.5)

# Set labels and title
plt.xlabel("Population")
plt.ylabel("Total Cases")
plt.title("Population vs Total Cases")

# Show the plot
plt.show()
```

Population vs Total Cases

```python
# Calculate the Spearman rank correlation matrix
correlation_matrix = df.corr(method='spearman')

# Set up the figure size
plt.figure(figsize=(10, 8))

# Create the heatmap
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm", center=0)

# Set title
plt.title("Spearman Rank Correlation Heatmap")

# Show the plot
plt.show()
```

```
C:\Users\shagu\AppData\Local\Temp\ipykernel_8884\2144558676.py:2:
FutureWarning: The default value of numeric_only in DataFrame.corr is
deprecated. In a future version, it will default to False. Select only
valid columns or specify the value of numeric_only to silence this
warning.
  correlation_matrix = df.corr(method='spearman')
```

Spearman Rank Correlation Heatmap