COVID-19 Coronavirus data set.

```
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
                                       Other names ISO 3166-1 alpha-3
               Country
CODE \
           Afghanistan
                                       Afghanistan
AFG
               Albania
                                           Albania
1
ALB
               Algeria
                                           Algeria
DZA
3
               Andorra
                                           Andorra
AND
                Angola
                                            Angola
4
AG0
. .
220
    Wallis and Futuna Wallis and Futuna Islands
WLF
221
        Western Sahara
                                    Western Sahara
ESHÂ
222
                 Yemen
                                              Yemen
YEM
                Zambia
223
                                             Zambia
ZMB
224
              Zimbabwe
                                          Zimbabwe
ZWE
     Population Continent Total Cases Total Deaths Tot Cases//IM
pop
       40462186
                     Asia
                                 177827
                                                  7671
0
4395
                                                  3492
        2872296
                   Europe
                                 273870
95349
       45236699
                   Africa
                                 265691
                                                  6874
5873
          77481
                    Europe
                                  40024
                                                   153
516565
       34654212
                   Africa
                                  99194
                                                  1900
2862
```

220	10894	0ceania	454	7		
41674						
221 16	623031	Africa	10	1		
222 381	30975258	Asia	11806	2143		
223 16442	19284482	Africa	317076	3967		
224 16174	15241601	Africa	246525	5446		
	Dootha	/1M Doot	h manaantana			
0 1 2 3 4	ot Deaths	190 190 1216 152 1975 55	1.275058 2.587216 0.382271 1.915438			
220 221 222 223 224		643 2 69 206 357	1.541850 10.000000 18.151787 1.251120 2.209107			
[225 r	ows x 10	columns]				
df.hea	nd()					
Contin	Country ment \	Other names	ISO 3166-1 alpha	-3 CODE	Population	
	hanistan	Afghanistan		AFG	40462186	
1	Albania	Albania		ALB	2872296	
Europe 2 Africa	Algeria	Algeria		DZA	45236699	
3 Europe	Andorra	Andorra		AND	77481	
4 Africa	Angola	Angola		AG0	34654212	
Tot 0 1 2 3 4	273870 273870 265691 40024 99194	Total Deaths 7671 3492 6874 153 1900	95 5 516	395 349 873	Deaths/1M pop 190 1216 152 1975 55	

```
Death percentage
0
           4.313743
1
           1.275058
2
           2.587216
3
           0.382271
4
           1.915438
df.tail()
                                       Other names ISO 3166-1 alpha-3
               Country
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
                   Africa
223
       19284482
                                 317076
                                                  3967
16442
224
                   Africa
       15241601
                                 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
                               224 non-null
     Other names
                                                object
     ISO 3166-1 alpha-3 CODE
 2
                               225 non-null
                                                object
 3
     Population
                               225 non-null
                                                int64
 4
     Continent
                               225 non-null
                                                obiect
 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()
                       Total Cases
                                    Total Deaths
         Population
                                                   Tot Cases//1M pop
count
       2.250000e+02
                      2.250000e+02
                                    2.250000e+02
                                                           225.000000
       3.507321e+07
                      2.184781e+06
                                    2.744813e+04
                                                        136900.373333
mean
std
       1.392418e+08
                      7.275938e+06
                                    9.689177e+04
                                                        145060.340289
       8.050000e+02
                      1.000000e+00
                                    0.000000e+00
                                                             9.000000
min
25%
       5.665570e+05
                      2.407100e+04
                                    1.890000e+02
                                                         11384.000000
50%
       5.827911e+06
                      1.639360e+05
                                    1.965000e+03
                                                         88987.000000
       2.190585e+07
                      1.092547e+06
75%
                                    1.366000e+04
                                                        223335.000000
max
       1.439324e+09
                      8.183905e+07
                                    1.008222e+06
                                                        696044.000000
       Tot Deaths/1M pop
                           Death percentage
              225.000000
count
                                 225.000000
             1096.715556
                                    1.444125
mean
                                    1.741728
std
             1195.715543
                0.000000
                                    0.000000
min
25%
              123.000000
                                    0.511291
50%
              708.000000
                                    1.036905
75%
             1795.000000
                                    1.977017
             6286.000000
                                  18.151787
max
df.isnull().sum()
Country
                            0
                            1
Other names
                            0
ISO 3166-1 alpha-3 CODE
Population
                            0
                            0
Continent
                            0
Total Cases
                            0
Total Deaths
                            0
Tot Cases//1M pop
Tot Deaths/1M pop
                            0
Death percentage
dtype: int64
df.dropna()
```

CODE	\	Country	0	ther names IS	60 3166-1 alpha-3
0	-	anistan	А	fghanistan	
AFG 1	,	Albania		Albania	
ALB 2	1	Algeria		Algeria	
DZA		_		_	
3 AND	, , ,	Andorra		Andorra	
4 AGO		Angola		Angola	
220 W WLF	<i>l</i> allis and	Futuna Wa	llis and Futu	na Islands	
221 ESHÂ	Western	Sahara	West	ern Sahara	
222 YEM		Yemen		Yemen	
223 ZMB		Zambia		Zambia	
224	Zi	imbabwe		Zimbabwe	
ZWE					
pop \	opulation	Continent	Total Cases	Total Deaths	s Tot Cases//1M
0	40462186	Asia	177827	7671	
				, 0, -	-
4395 1	2872296	Europe	273870	3492	
1 95349 2	2872296 45236699				2
1 95349 2 5873 3	45236699 77481	Europe	273870	3492	<u>?</u> I
1 95349 2 5873	45236699 77481	Europe Africa	273870 265691	3492 6874	<u>?</u> I 3
1 95349 2 5873 3 516565 4 2862	45236699 77481	Europe Africa Europe	273870 265691 40024 99194	3492 6874 153	<u>2</u> I 3
1 95349 2 5873 3 516565 4 2862	45236699 77481 34654212	Europe Africa Europe Africa	273870 265691 40024 99194	3492 6874 153 1906	2
1 95349 2 5873 3 516565 4 2862	45236699 77481	Europe Africa Europe Africa Oceania	273870 265691 40024 99194	3492 6874 153 1906	
1 95349 2 5873 3 516565 4 2862 220 41674 221	45236699 77481 34654212	Europe Africa Europe Africa	273870 265691 40024 99194	3492 6874 153 1906	
1 95349 2 5873 3 516565 4 2862 220 41674 221 16 222	45236699 77481 34654212 10894	Europe Africa Europe Africa Oceania	273870 265691 40024 99194 454	3492 6874 153 1906	
1 95349 2 5873 3 516565 4 2862 220 41674 221 16 222 381 223	45236699 77481 34654212 10894 623031	Europe Africa Europe Africa Oceania Africa	273870 265691 40024 99194 454	3492 6874 153 1906 	
1 95349 2 5873 3 516565 4 2862 220 41674 221 16 222 381	45236699 77481 34654212 10894 623031 30975258	Europe Africa Europe Africa Oceania Africa Asia	273870 265691 40024 99194 454 10 11806	3492 6874 153 1906 	

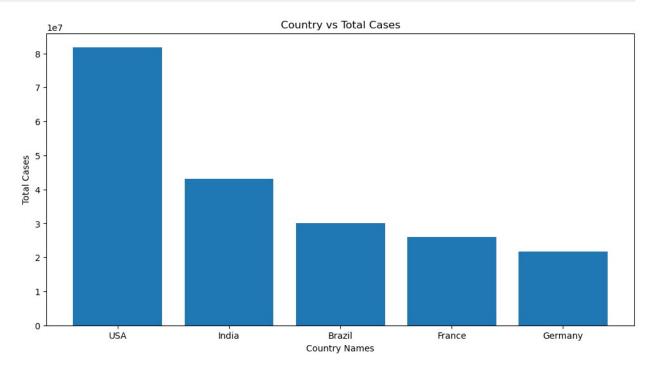
```
Tot Deaths/1M pop
                        Death percentage
0
                   190
                                 4.313743
1
                  1216
                                 1.275058
2
                                 2.587216
                   152
3
                  1975
                                 0.382271
4
                    55
                                 1.915438
220
                                 1.541850
                   643
221
                     2
                                10.000000
222
                    69
                                18.151787
223
                                 1.251120
                   206
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	1		Other names	ISO 3166-1 alpha-3
CODE 0	\ Afghanistan	1		Afghanistan	
AFG 1	Albania	1		Albania	
ALB 2	Algeria	ì		Algeria	
DZA 3	Andorra	1		Andorra	
AND 4 AGO	Angola	l		Angola	
131 FSM	Micronesia	a Micrones	ia (Federated	States of)	
132 MDA	Moldova	ì	Republic	of Moldova	
133 MC0	Monaco)		Monaco	
134 MNG	Mongolia	1		Mongolia	
135 MNE	Montenegro			Montenegro	
		Continent	Total Cases	Total Death	s Tot Cases//1M
pop 0 4395	40462186	Asia	177827	767	1
1 95349	2872296	Europe	273870	349	2
2 5873	45236699	Africa	265691	687	4
3 51656	77481 55	Europe	40024	15	3
4 2862	34654212	Africa	99194	190	0
131 9	117134	Oceania	1		0
132 12798		Europe	514199	1144	
133 27289		Europe	10842	5	
134 13902		Asia	468610	217	
135 3714	628205 17	Europe	233326	270	5

```
Tot Deaths/1M pop
                          Death percentage
                                  4.313743
0
                    190
1
                   1216
                                  1.275058
2
                                  2.587216
                    152
3
                   1975
                                  0.382271
4
                     55
                                  1.915438
                                  0.000000
131
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
                                         Other names Iso Code
                Country
                                                                Population
0
           Afghanistan
                                         Afghanistan
                                                           AFG
                                                                  40462186
                Albania
                                             Albania
                                                           ALB
                                                                    2872296
1
2
                Algeria
                                             Algeria
                                                           DZA
                                                                  45236699
3
                Andorra
                                             Andorra
                                                           AND
                                                                      77481
                                                                  34654212
                 Angola
                                              Angola
                                                           AG0
     Wallis and Futuna
                         Wallis and Futuna Islands
220
                                                           WLF
                                                                      10894
221
        Western Sahara
                                     Western Sahara
                                                         ESHÂ
                                                                     623031
222
                                                           YEM
                                                                  30975258
                  Yemen
                                               Yemen
223
                 Zambia
                                              Zambia
                                                           ZMB
                                                                   19284482
               Zimbabwe
                                            Zimbabwe
224
                                                           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
                      40024
                                        153
                                                         516565
       Europe
```

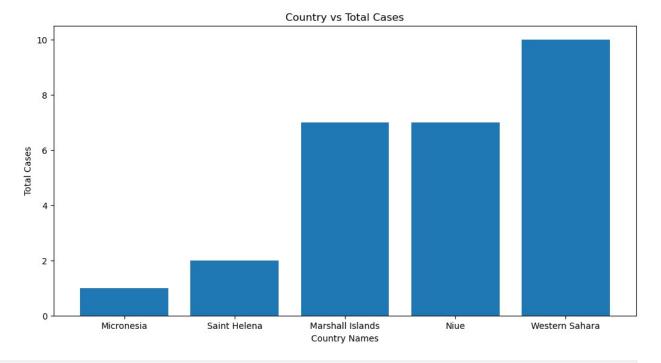
```
4
       Africa
                      99194
                                      1900
                                                          2862
220
      Oceania
                        454
                                         7
                                                         41674
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
                      2
221
                                 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()
                            Other names Iso Code
     Country
                                                    Population \
214
         USA
              United States of America
                                              USA
                                                     334400597
92
       India
                                   India
                                              IND
                                                   1403754381
26
      Brazil
                                  Brazil
                                              BRA
                                                     215204501
                                              FRA
70
      France
                                  France
                                                      65526369
76
     Germany
                                 Germany
                                              DEU
                                                      84252947
                            Continent Total Cases
                                                      Total Deaths \
214
                     Northern America
                                           81839052
                                                           1008222
92
                                  Asia
                                           43029044
                                                            521388
     Latin America and the Caribbean
26
                                           29999816
                                                            660269
70
                                                            142506
                                Europe
                                           25997852
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
                                       3068
                                                      2.200910
                 139401
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")
```

```
plt.ylabel("Total Cases")
plt.title("Country vs Total Cases")
#plt.xticks(rotation = 90)
plt.show()
```



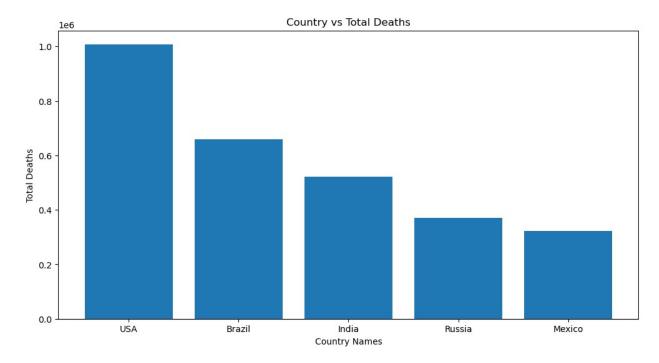
<pre>Country = df.nsmallest(5, 'Total Cases') Country.head()</pre>					
Country			Other names	Iso_Code	
Population \		/ F		ECM	
131 Micronesia 117134	Micronesia	(Federa	ted States of)	FSM	
168 Saint Helena			Saint Helena	SHN	
6109				2	
125 Marshall Islands		Ма	rshall Islands	MHL	
59889			N 1.	NTII	
148 Niue 1645			Niue	NIU	
221 Western Sahara			Western Sahara	ESHÂ	
623031					
6 1 7 1 6	.	D 11	T C //1M	,	
Continent Total C 131 Oceania	ases lotal 1	Deaths 0	lot Cases//IM	pop \	
168 Africa	2	0		327	
125 Oceania	7	0		117	
148 Oceania	7	0	4	4255	
221 Africa	10	1		16	

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



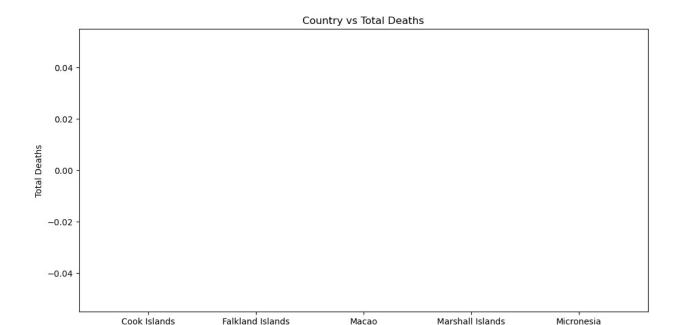
```
Country = df.nlargest(5, 'Total Deaths')
Country.head()
                          Other names Iso Code
                                                 Population \
    Country
214
        USA
             United States of America
                                            USA
                                                  334400597
26
     Brazil
                                            BRA
                                Brazil
                                                  215204501
92
     India
                                 India
                                            IND
                                                 1403754381
165
     Russia
                   Russian Federation
                                            RUS
                                                  146044010
130
     Mexico
                               Mexico
                                            MEX
                                                  131303955
                                                    Total Deaths \
                           Continent Total Cases
214
                    Northern America
                                          81839052
                                                         1008222
26
     Latin America and the Caribbean
                                          29999816
                                                          660269
92
                                 Asia
                                          43029044
                                                          521388
```

```
165
                                           17896866
                                                           369708
                               Europe
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
                                                     2.065770
                                      2531
                  43147
                                                     5.705041
130
                                      2462
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 = df.nsmallest(5, 'Total Deaths')
Country.head()
                                                       Other names
              Country
Iso Code \
46
         Cook Islands
                                                      Cook Islands
COK
     Falkland Islands
                                       Falkland Islands (Malvinas)
67
FLK
118
                Macao China, Macao Special Administrative Region
MAC
```

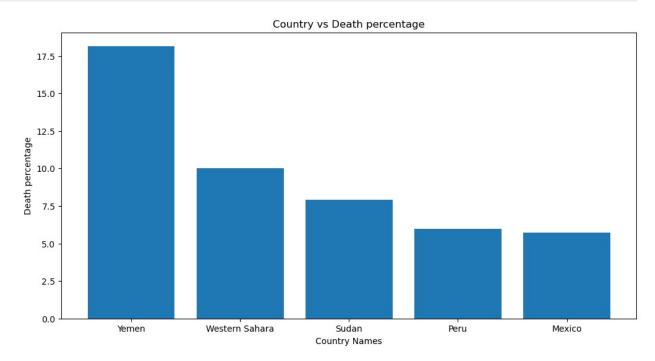
```
125
     Marshall Islands
                                                  Marshall Islands
MHL
131
           Micronesia
                                 Micronesia (Federated States of)
FSM
     Population
                                        Continent Total Cases Total
Deaths
       /
          17592
                                          Oceania
                                                          2118
46
0
67
           3657 Latin America and the Caribbean
                                                           123
0
118
         664828
                                                            82
                                             Asia
0
125
                                                             7
          59889
                                          Oceania
131
         117134
                                          Oceania
                                                             1
0
     Tot Cases//1M pop Tot Deaths/1M pop
                                            Death percentage
46
                120396
                                                         0.0
                                         0
                                                         0.0
67
                 33634
                                         0
118
                   123
                                                         0.0
125
                   117
                                         0
                                                         0.0
131
                                         0
                                                         0.0
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()
```



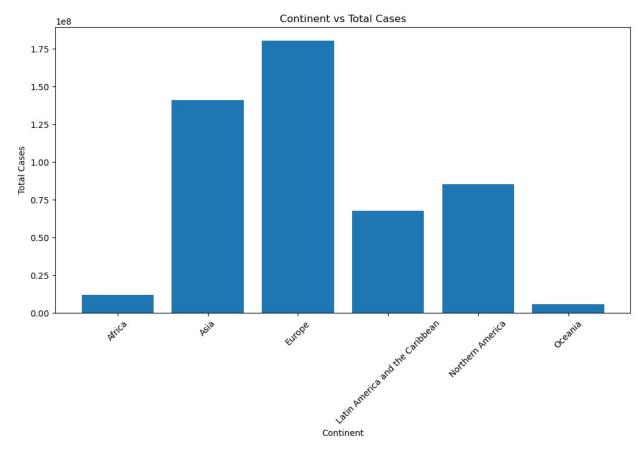
Country Names

```
Country = df.nlargest(5, 'Death percentage')
Country.head()
                         Other names Iso Code
                                                Population \
            Country
222
                               Yemen
                                           YEM
                                                  30975258
              Yemen
221
     Western Sahara
                      Western Sahara
                                         ESHÂ
                                                     623031
193
              Sudan
                               Sudan
                                                  45640385
                                           SDN
158
               Peru
                                Peru
                                           PER
                                                  33775745
130
                                           MEX
             Mexico
                              Mexico
                                                 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
     Latin America and the Caribbean
130
                                            5665376
                                                            323212
     Tot Cases//1M pop
                         Tot Deaths/1M pop
                                             Death percentage
222
                                                     18.151787
                    381
                                         69
221
                                          2
                     16
                                                     10.000000
193
                   1357
                                        108
                                                      7.920265
158
                                                      5.983499
                 105062
                                       6286
130
                 43147
                                       2462
                                                      5.705041
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()
```



```
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()
```

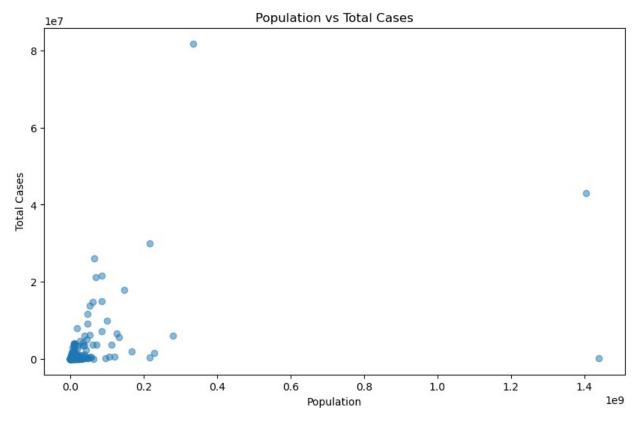


```
plt.figure(figsize=(10, 6))

# 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()
```



```
# 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\shaqu\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')
```

