```
In [2]: import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns
        import plotly.express as px
```

```
In [4]: df = pd.read_csv('covid 19/country_wise_latest.csv')
        df.head()
```

Out[4]:

	Country/Region	Confirmed	Deaths	Recovered	Active	New cases	New deaths	New recovered	Deaths / 100 Cases	Recovered / 100 Cases	Deaths / 100 Recovered	Confirmed last week	1 week change	1 week % increase	WHO Region
0	Afghanistan	36263	1269	25198	9796	106	10	18	3.50	69.49	5.04	35526	737	2.07	Eastern Mediterranean
1	Albania	4880	144	2745	1991	117	6	63	2.95	56.25	5.25	4171	709	17.00	Europe
2	Algeria	27973	1163	18837	7973	616	8	749	4.16	67.34	6.17	23691	4282	18.07	Africa
3	Andorra	907	52	803	52	10	0	0	5.73	88.53	6.48	884	23	2.60	Europe
4	Angola	950	41	242	667	18	1	0	4.32	25.47	16.94	749	201	26.84	Africa

```
In [4]: df.isnull().sum()

Out[4]: Country/Region            0
        Confirmed                 0
        Deaths                    0
        Recovered                 0
        Active                    0
        New cases                 0
        New deaths                0
        New recovered             0
        Deaths / 100 Cases        0
        Recovered / 100 Cases     0
        Deaths / 100 Recovered    0
        Confirmed last week       0
        1 week change             0
        1 week % increase         0
        WHO Region                0
        dtype: int64
```

```
In [6]: df.shape

Out[6]: (187, 15)
```

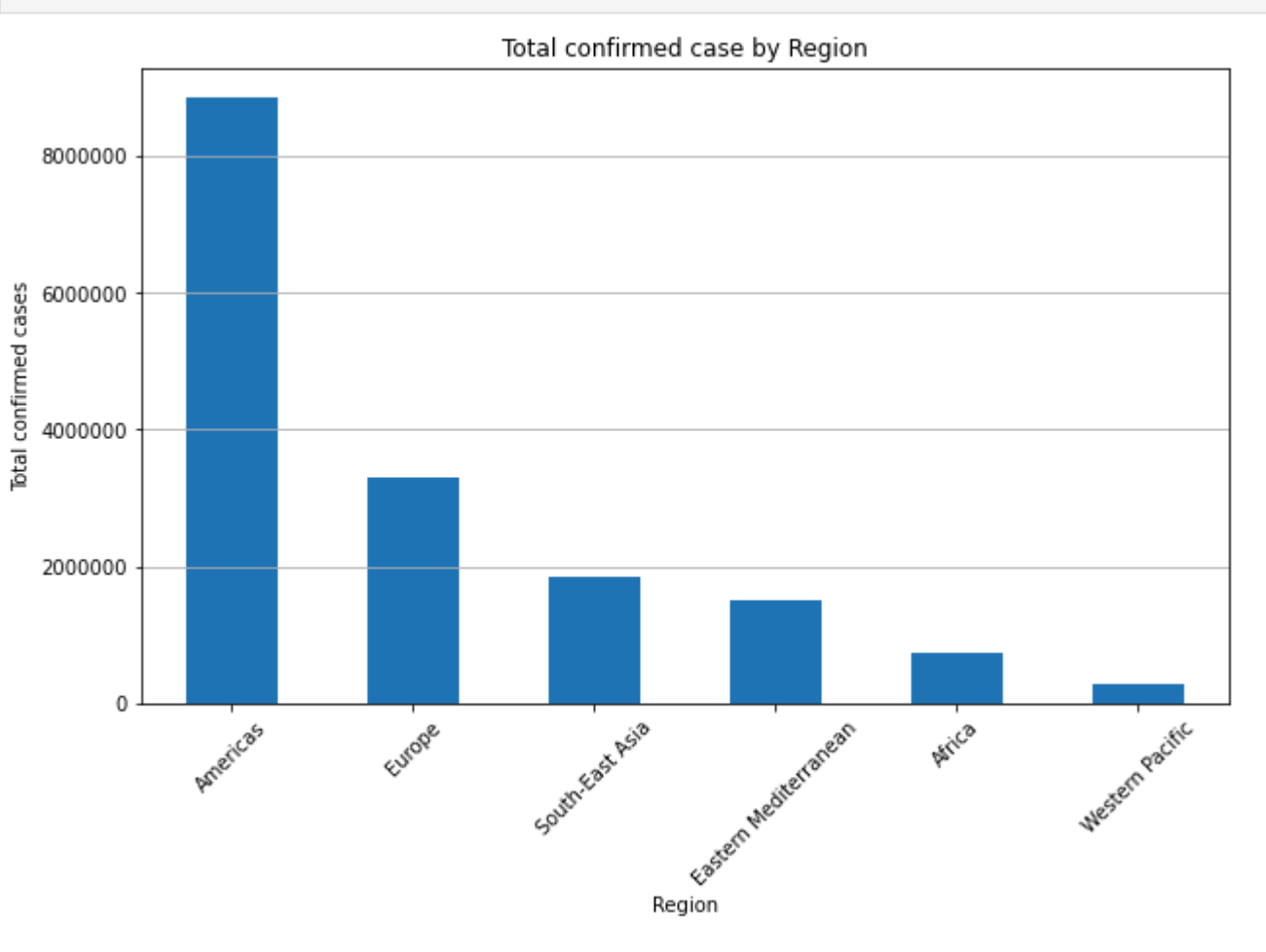
```
In [7]: df.columns

Out[7]: Index(['Country/Region', 'Confirmed', 'Deaths', 'Recovered', 'Active',
               'New cases', 'New deaths', 'New recovered', 'Deaths / 100 Cases',
               'Recovered / 100 Cases', 'Deaths / 100 Recovered',
               'Confirmed last week', '1 week change', '1 week % increase',
               'WHO Region'],
              dtype='object')
```

```
In [8]: df.info

Out[8]: <bound method DataFrame.info of          Country/Region  Confirmed  Deaths  Recovered  Active  New cases  \
        0           Afghanistan      36263    1269      25198    9796        106
        1               Albania       4880     144       2745    1991        117
        2               Algeria      27973    1163      18837    7973        616
        3               Andorra        907      52        803      52         10
        4                Angola        950      41        242     667         18
        ..                  ...        ...     ...        ...     ...        ...
        182  West Bank and Gaza      10621      78       3752    6791        152
        183      Western Sahara         10       1          8       1          0
        184               Yemen       1691     483        833     375         10
        185              Zambia       4552     140       2815    1597         71
        186            Zimbabwe       2704      36        542    2126        192

             New deaths  New recovered  Deaths / 100 Cases  Recovered / 100 Cases  \
        0            10             18                3.50                  69.49
        1             6             63                2.95                  56.25
        2             8            749                4.16                  67.34
        3             0              0                5.73                  88.53
        4             1              0                4.32                  25.47
        ..          ...            ...                 ...                    ...
        182           2              0                0.73                  35.33
        183           0              0               10.00                  80.00
        184           4             36               28.56                  49.26
        185           1            465                3.08                  61.84
        186           2             24                1.33                  20.04

             Deaths / 100 Recovered  Confirmed last week  1 week change  \
        0                      5.04                35526            737
        1                      5.25                 4171            709
        2                      6.17                23691           4282
        3                      6.48                  884             23
        4                     16.94                  749            201
        ..                      ...                  ...            ...
        182                    2.08                 8916           1705
        183                   12.50                   10              0
        184                   57.98                 1619             72
        185                    4.97                 3326           1226
        186                    6.64                 1713            991

             1 week % increase             WHO Region
        0                 2.07  Eastern Mediterranean
        1                17.00                 Europe
        2                18.07                 Africa
        3                 2.60                 Europe
        4                26.84                 Africa
        ..                 ...                    ...
        182              19.12  Eastern Mediterranean
        183               0.00                 Africa
        184               4.45  Eastern Mediterranean
        185              36.86                 Africa
        186              57.85                 Africa

        [187 rows x 15 columns]>
```

```
In [18]: df.groupby('WHO Region')['Confirmed'].sum().sort_values(ascending = False)

Out[18]: WHO Region
         Americas                 8839286
         Europe                   3299523
         South-East Asia          1835297
         Eastern Mediterranean    1490744
         Africa                    723207
         Western Pacific           292428
         Name: Confirmed, dtype: int64
```

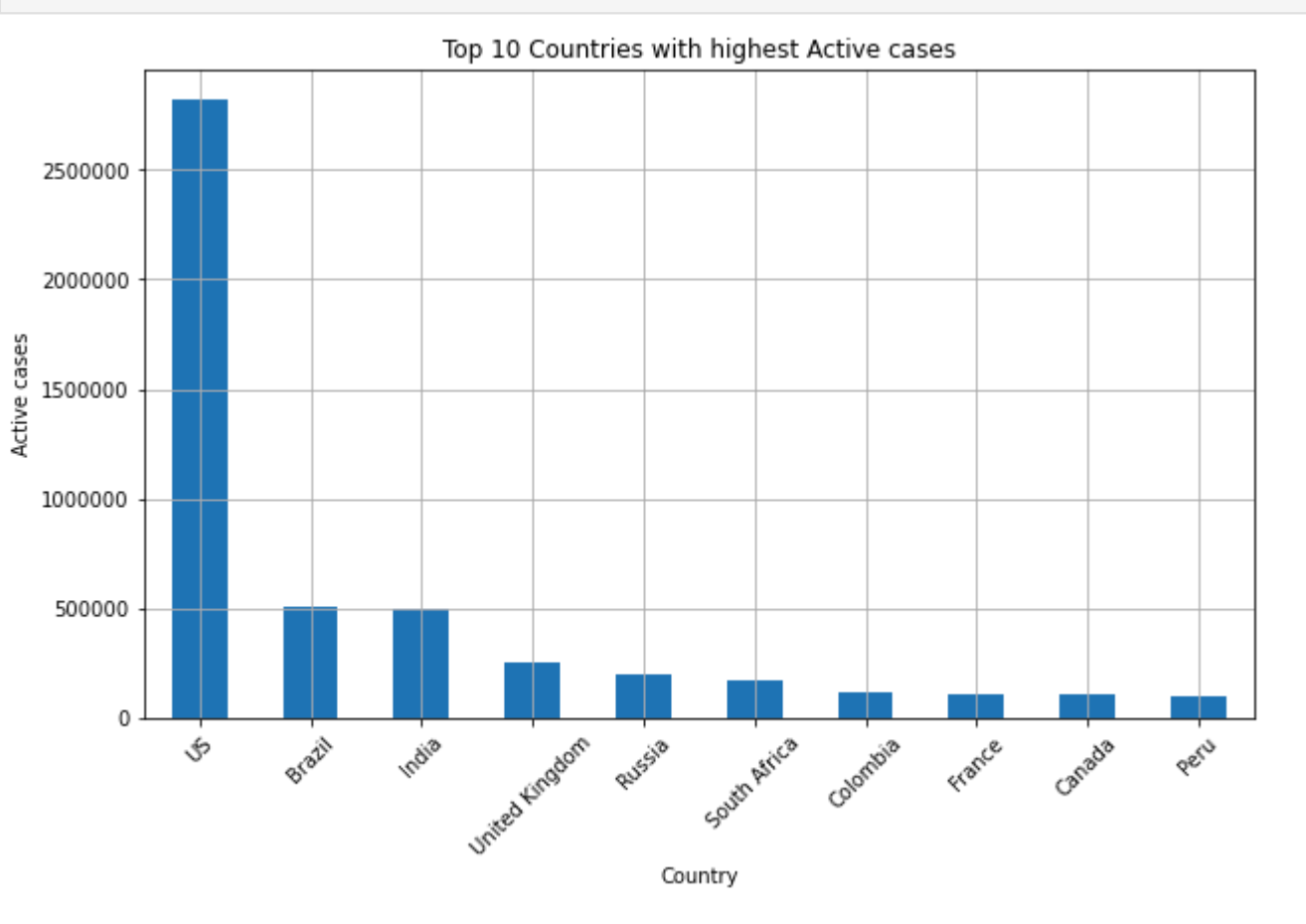
Confirmed cases by Region

```
In [41]: plt.figure(figsize = (10,6))
         df.groupby('WHO Region')['Confirmed'].sum().sort_values(ascending = False).plot(kind= 'bar')
         plt.title('Total confirmed case by Region')

         plt.ticklabel_format(style='plain' , axis = 'y')
         plt.xlabel('Region')
         plt.ylabel('Total confirmed cases')
         plt.grid(axis='y')
         plt.xticks(rotation = 45)
         plt.show()
```



Top 10 Countries with highest Active cases

```
In [46]: plt.figure(figsize = (10,6))
         df.groupby('Country/Region')['Active'].sum().sort_values(ascending = False).head(10).plot(kind= 'bar')
         plt.title('Top 10 Countries with highest Active cases')

         plt.ticklabel_format(style='plain' , axis = 'y')
         plt.xlabel('Country')
         plt.ylabel('Active cases')
         plt.grid(True)
         plt.xticks(rotation = 45)
         plt.show()
```



Correlation Matrix

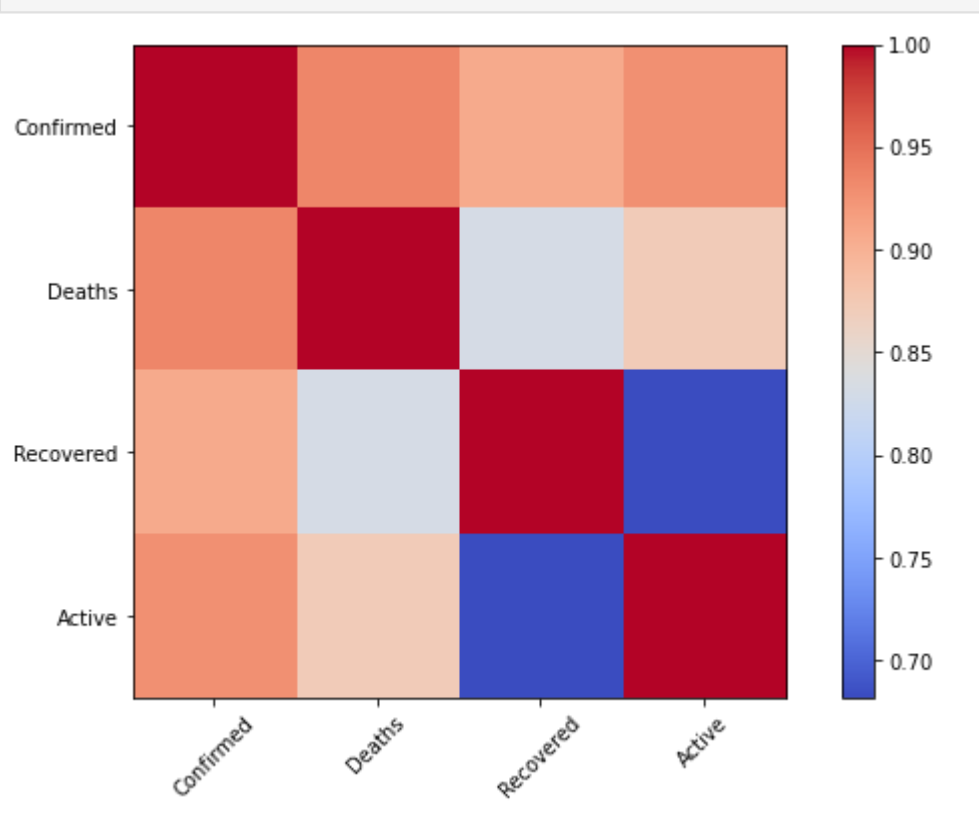
```
In [47]: correlation = df[['Confirmed' ,'Deaths' , 'Recovered' , 'Active']]
         correlation
```

Out[47]:

	Confirmed	Deaths	Recovered	Active
0	36263	1269	25198	9796
1	4880	144	2745	1991
2	27973	1163	18837	7973
3	907	52	803	52
4	950	41	242	667
...	...	...	...	...
182	10621	78	3752	6791
183	10	1	8	1
184	1691	483	833	375
185	4552	140	2815	1597
186	2704	36	542	2126

187 rows × 4 columns

```
In [60]: plt.figure(figsize = (10,6))
         correlation = df[['Confirmed' ,'Deaths' , 'Recovered' , 'Active']].corr()
         correlation

         plt.imshow(correlation, cmap = 'coolwarm' , interpolation = 'nearest')
         plt.colorbar()
         plt.xticks(range(len(correlation.columns)), correlation.columns , rotation = 45)
         plt.yticks(range(len(correlation.columns)), correlation.columns)
         plt.show()
```



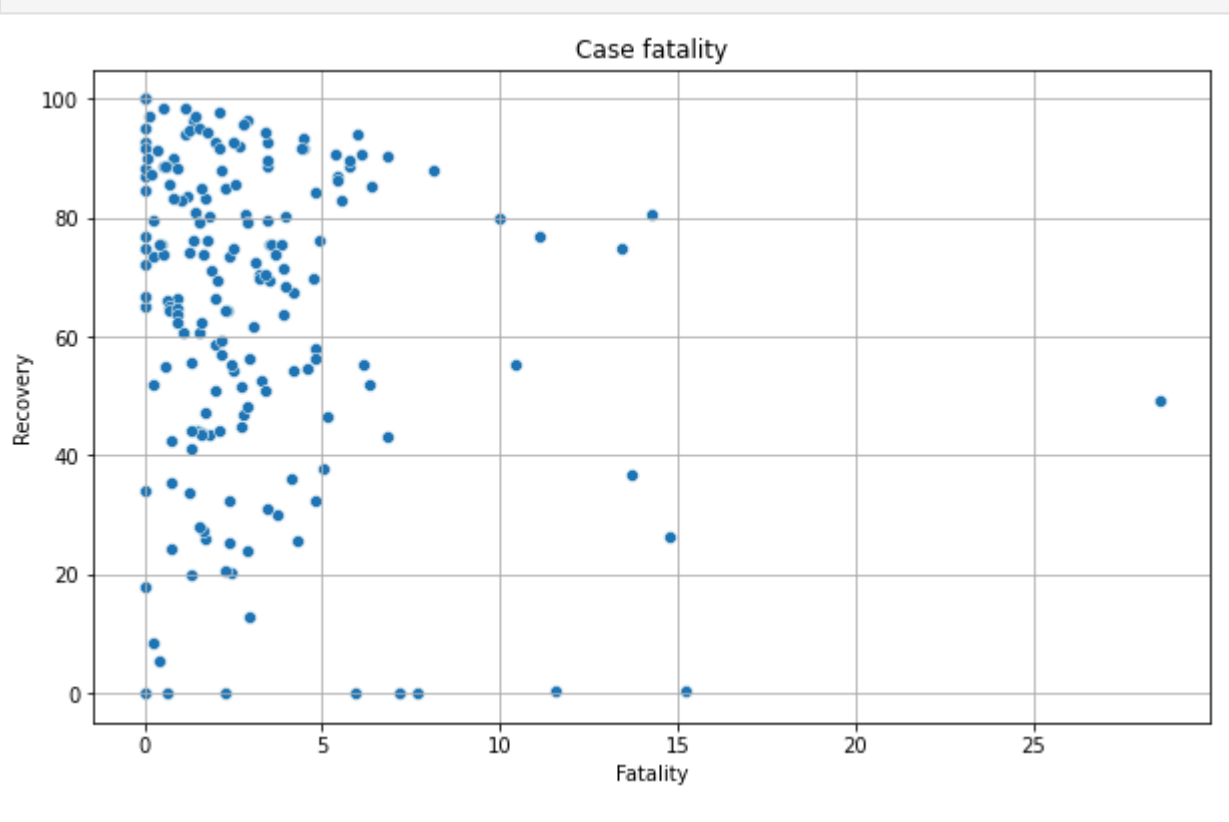
Fatality vs Recovery

```
In [63]: df['Fatality_rate'] = df['Deaths'] / df['Confirmed'] * 100
         df['Recovery_rate'] = df['Recovered'] / df['Confirmed'] * 100
         df.head()
```

Out[63]:

	Country/Region	Confirmed	Deaths	Recovered	Active	New cases	New deaths	New recovered	Deaths / 100 Cases	Recovered / 100 Cases	Deaths / 100 Recovered	Confirmed last week	1 week change	1 week % increase	WHO Region	Fatality_rate	Recovery_rate
0	Afghanistan	36263	1269	25198	9796	106	10	18	3.50	69.49	5.04	35526	737	2.07	Eastern Mediterranean	3.499435	69.486805
1	Albania	4880	144	2745	1991	117	6	63	2.95	56.25	5.25	4171	709	17.00	Europe	2.950820	56.250000
2	Algeria	27973	1163	18837	7973	616	8	749	4.16	67.34	6.17	23691	4282	18.07	Africa	4.157581	67.339935
3	Andorra	907	52	803	52	10	0	0	5.73	88.53	6.48	884	23	2.60	Europe	5.733186	88.533627
4	Angola	950	41	242	667	18	1	0	4.32	25.47	16.94	749	201	26.84	Africa	4.315789	25.473684

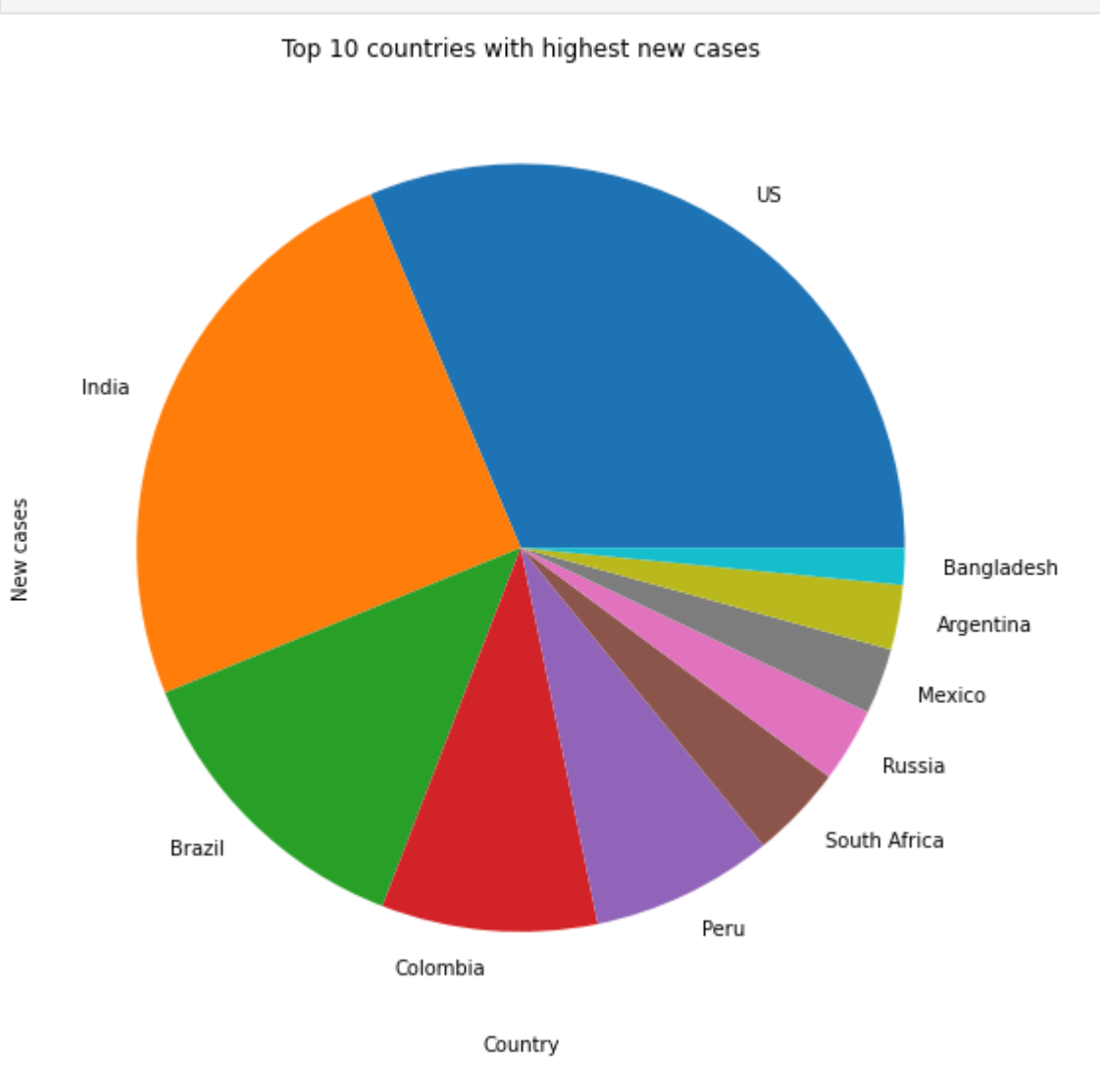
```
In [66]: plt.figure(figsize = (10,6))
         sns.scatterplot(x = 'Fatality_rate', y = 'Recovery_rate', data = df)
         plt.title('Case fatality')
         plt.xlabel('Fatality')
         plt.ylabel('Recovery')
         plt.grid(True)

         plt.show()
```



Top 10 Countries with highest new cases

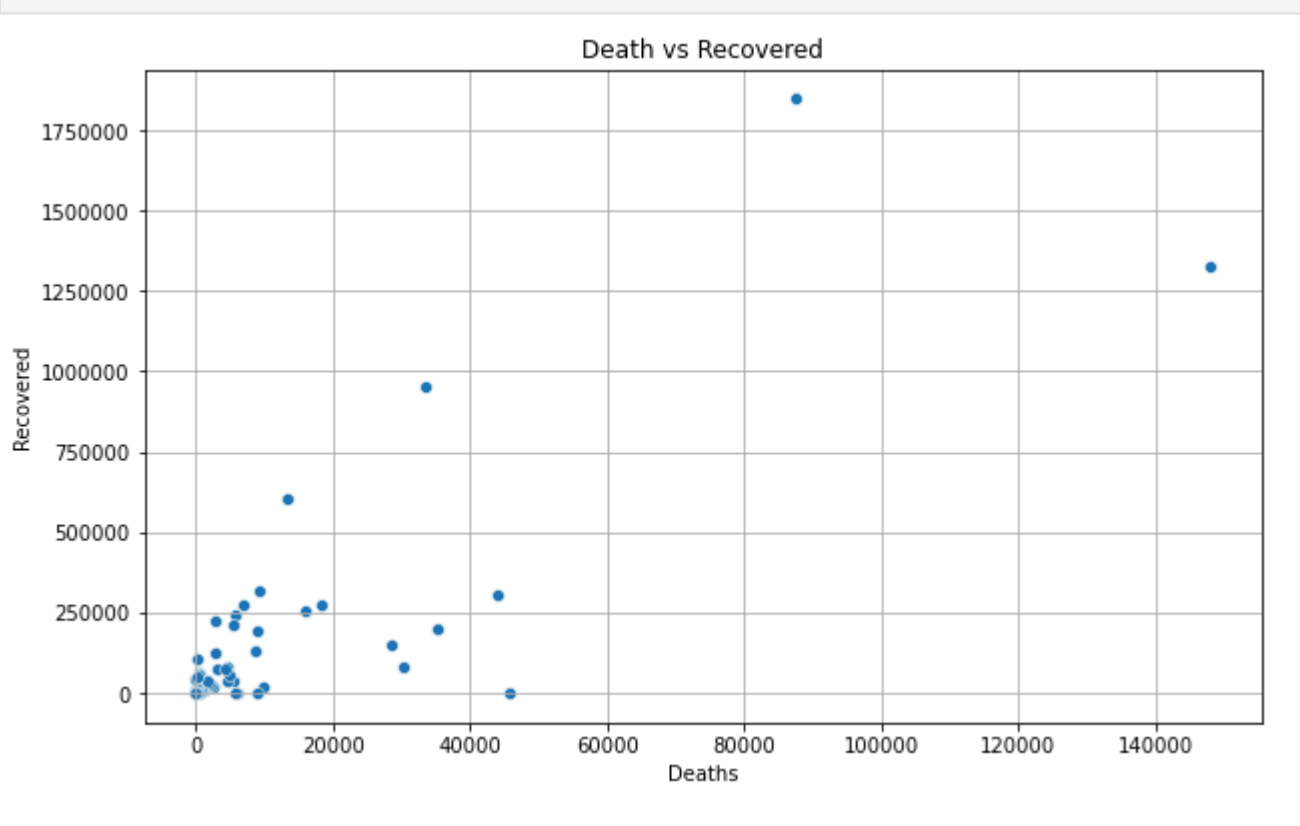
```
In [19]: df.groupby('Country/Region')['New cases'].sum().sort_values(ascending = False).head(10)

Out[19]: Country/Region
         US              56336
         India           44457
         Brazil          23284
         Colombia        16306
         Peru            13756
         South Africa     7096
         Russia           5607
         Mexico           4973
         Argentina        4890
         Bangladesh       2772
         Name: New cases, dtype: int64
```

```
In [69]: plt.figure(figsize = (15,9))
         df.groupby('Country/Region')['New cases'].sum().sort_values(ascending = False).head(10).plot(kind= 'pie')
         plt.title('Top 10 countries with highest new cases')


         plt.xlabel('Country')
         plt.ylabel('New cases')

         plt.show()
```



Deaths vs Recovered

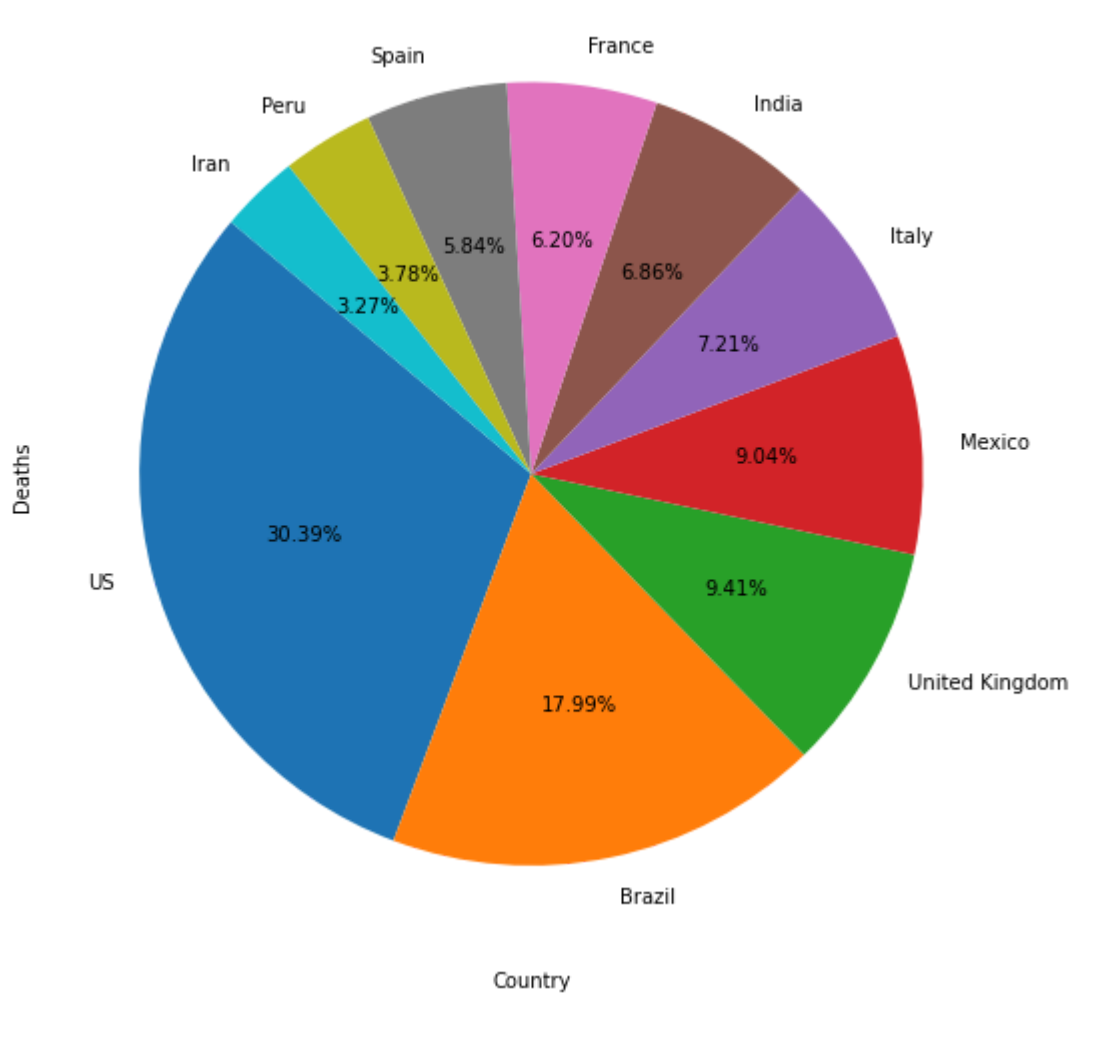
```
In [71]: plt.figure(figsize = (10,6))
         sns.scatterplot(x = 'Deaths', y = 'Recovered', data = df)
         plt.ticklabel_format(style='plain' , axis = 'y')
         plt.title('Death vs Recovered')
         plt.xlabel('Deaths')
         plt.ylabel('Recovered')
         plt.grid(True)

         plt.show()
```



Top 10 Countries with highest deaths

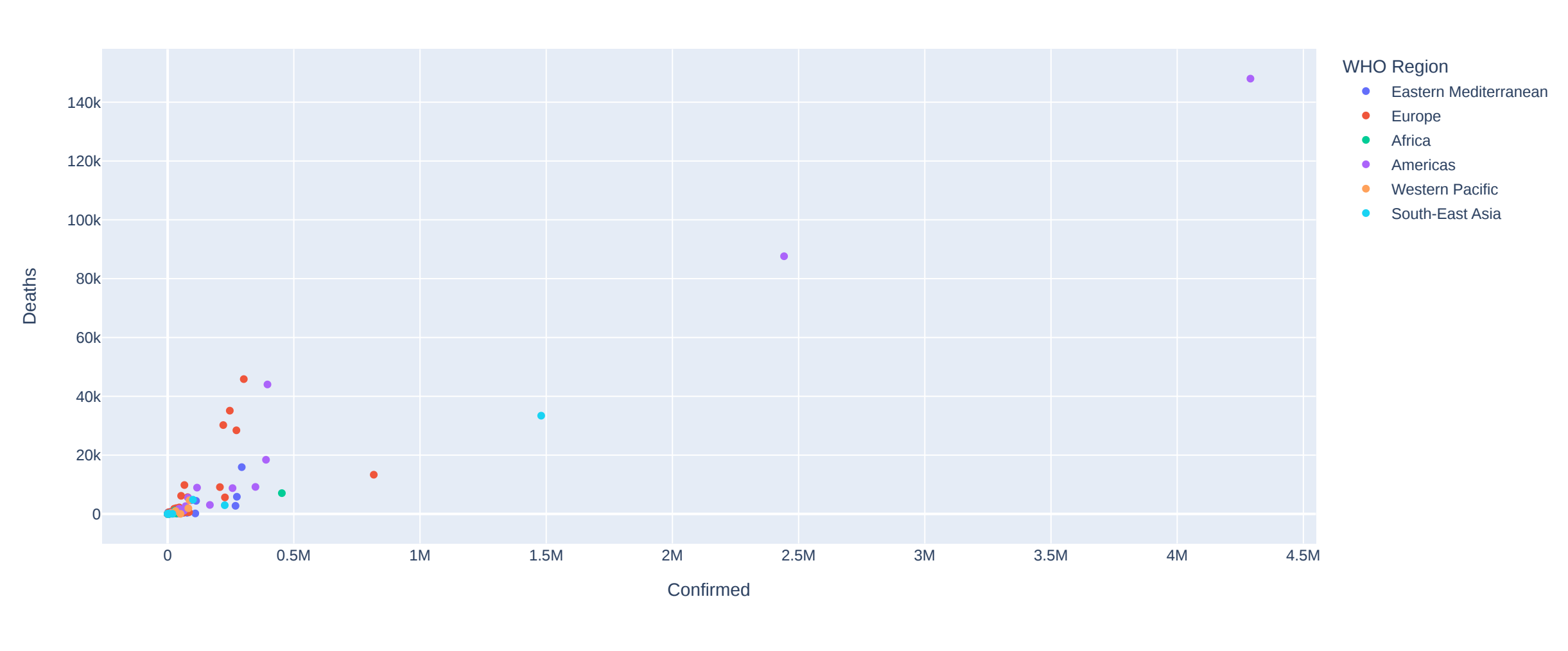
```
In [78]: df6=df.groupby('Country/Region')['Deaths'].sum().sort_values(ascending = False).head(10)
         df6.index

Out[78]: Index(['US', 'Brazil', 'United Kingdom', 'Mexico', 'Italy', 'India', 'France',
                'Spain', 'Peru', 'Iran'],
               dtype='object', name='Country/Region')
```

```
In [84]: plt.figure(figsize = (15,9))
         df6 = df.groupby('Country/Region')['Deaths'].sum().sort_values(ascending = False).head(10)
         pie_chart = plt.pie(df6, labels = df6.index, autopct = '%0.2f%%' , startangle = 140)
         plt.title('Top 10 countries with highest Deaths')
         plt.xlabel('Country')
         plt.ylabel('Deaths')
         plt.show()
```



Confirmed Cases vs. Deaths by WHO Region

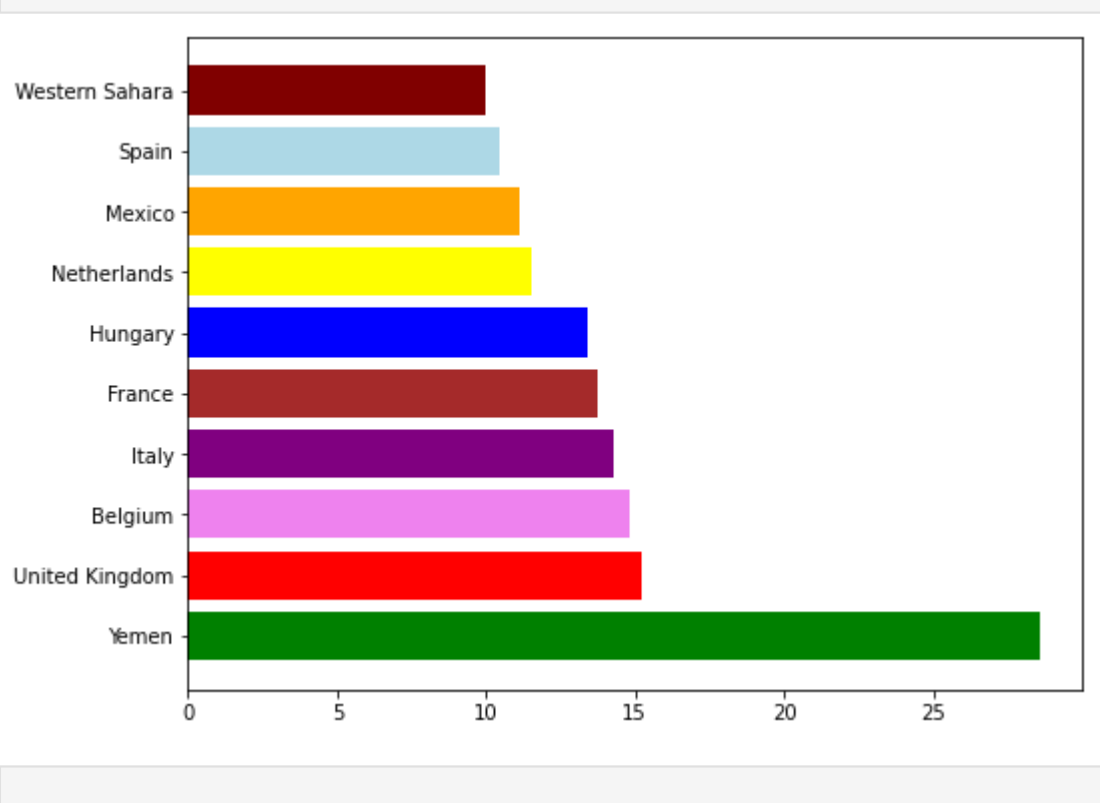
```
In [5]: px.scatter(df, x = 'Confirmed' , y = 'Deaths' , color = 'WHO Region')
```



Top 10 Countries with highest deaths per 100 cases

```
In [92]: df9 = df.groupby('Country/Region')['Deaths / 100 Cases'].sum().sort_values(ascending = False).head(10)
         colors = ['green' , 'red' , 'violet' , 'purple' , 'brown' , 'blue' , 'yellow', 'orange', 'lightblue' , 'maroon']
         df9

Out[92]: Country/Region
         Yemen             28.56
         United Kingdom    15.19
         Belgium           14.79
         Italy             14.26
         France            13.71
         Hungary           13.40
         Netherlands       11.53
         Mexico            11.13
         Spain             10.44
         Western Sahara    10.00
         Name: Deaths / 100 Cases, dtype: float64
```

```
In [94]: plt.barh(df9.index, df9, color = colors)
         plt.show()
```




```
In [ ]:
```