In [1]:
```
import pandas as pd
```

In [2]:
```
data =  pd.read_csv('D:\\DATA ANALYTICS\\4. covid_19_data (1).csv')
```

In [3]:
```
data
```

Out[3]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252
1	4/29/2020	NaN	Albania	766	30	455
2	4/29/2020	NaN	Algeria	3848	444	1702
3	4/29/2020	NaN	Andorra	743	42	423
4	4/29/2020	NaN	Angola	27	2	7
...	...	...	...	...	...	...
316	4/29/2020	Wyoming	US	545	7	0
317	4/29/2020	Xinjiang	Mainland China	76	3	73
318	4/29/2020	Yukon	Canada	11	0	0
319	4/29/2020	Yunnan	Mainland China	185	2	181
320	4/29/2020	Zhejiang	Mainland China	1268	1	1263

321 rows × 6 columns

In [5]:
```
data.count()
```

Out[5]:
```
Date         321
State        140
Region       321
Confirmed    321
Deaths       321
Recovered    321
dtype: int64
```

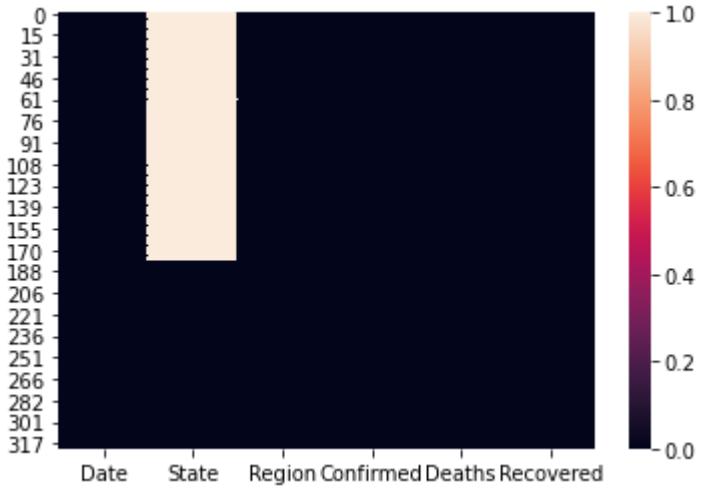
In [6]:
```
data.isnull().sum()
```

Out[6]:
```
Date           0
State        181
Region         0
Confirmed      0
Deaths         0
Recovered      0
dtype: int64
```

In [8]:
```
import seaborn as sns
import matplotlib.pyplot as plt
```

In [12]:
```
sns.heatmap(data.isnull())
```

Out[12]: <AxesSubplot:>



In [13]:
```
data.groupby('Region')['Confirmed','Recovered'].sum()
```

```
<ipython-input-13-20fd7b835859>:1: FutureWarning: Indexing with multiple keys (implicitl
y converted to a tuple of keys) will be deprecated, use a list instead.
  data.groupby('Region')['Confirmed','Recovered'].sum()
```

Out[13]:

Region	Confirmed	Recovered
Afghanistan	1939	252
Albania	766	455
Algeria	3848	1702
Andorra	743	423
Angola	27	7
...	...	...
Venezuela	331	142
Vietnam	270	222
West Bank and Gaza	344	71
Zambia	97	54
Zimbabwe	32	5

180 rows × 2 columns

In [15]:
```
data = data[~(data.Confirmed < 10) ]
```

In [16]:
```
data
```

Out[16]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252
1	4/29/2020	NaN	Albania	766	30	455
2	4/29/2020	NaN	Algeria	3848	444	1702
3	4/29/2020	NaN	Andorra	743	42	423
4	4/29/2020	NaN	Angola	27	2	7
...	...	...	...	...	...	...
316	4/29/2020	Wyoming	US	545	7	0
317	4/29/2020	Xinjiang	Mainland China	76	3	73
318	4/29/2020	Yukon	Canada	11	0	0
319	4/29/2020	Yunnan	Mainland China	185	2	181
320	4/29/2020	Zhejiang	Mainland China	1268	1	1263

304 rows × 6 columns

In [17]:
```
data.groupby('Region')['Confirmed','Recovered'].sum()
```

```
<ipython-input-17-20fd7b835859>:1: FutureWarning: Indexing with multiple keys (implicitl
y converted to a tuple of keys) will be deprecated, use a list instead.
  data.groupby('Region')['Confirmed','Recovered'].sum()
```

Out[17]:

Region	Confirmed	Recovered
Afghanistan	1939	252
Albania	766	455
Algeria	3848	1702
Andorra	743	423
Angola	27	7
...	...	...
Venezuela	331	142
Vietnam	270	222
West Bank and Gaza	344	71
Zambia	97	54
Zimbabwe	32	5

180 rows × 2 columns

In [21]:
```
data
```

Out[21]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252
1	4/29/2020	NaN	Albania	766	30	455
2	4/29/2020	NaN	Algeria	3848	444	1702
3	4/29/2020	NaN	Andorra	743	42	423
4	4/29/2020	NaN	Angola	27	2	7
...	...	...	...	...	...	...
316	4/29/2020	Wyoming	US	545	7	0
317	4/29/2020	Xinjiang	Mainland China	76	3	73
318	4/29/2020	Yukon	Canada	11	0	0
319	4/29/2020	Yunnan	Mainland China	185	2	181
320	4/29/2020	Zhejiang	Mainland China	1268	1	1263

304 rows × 6 columns

In [30]:
```
data.groupby('Region')['Deaths'].sum().sort_values(ascending = True)
```

Out[30]:
```
Region
Cambodia                      0
Seychelles                    0
Saint Lucia                   0
Central African Republic      0
Saint Kitts and Nevis         0
                           ...
France                    24121
Spain                     24275
UK                        26165
Italy                     27682
US                        60967
Name: Deaths, Length: 180, dtype: int64
```

In [38]:
```
data[data.Region == 'India']
```

Out[38]:

	Date	State	Region	Confirmed	Deaths	Recovered
74	4/29/2020	NaN	India	33062	1079	8437

In [41]:
```
data.sort_values(by = ['Confirmed'])
```

Out[41]:

	Date	State	Region	Confirmed	Deaths	Recovered
156	4/29/2020	NaN	Suriname	10	1	8
70	4/29/2020	NaN	Holy See	10	0	2
59	4/29/2020	NaN	Gambia	10	1	8
318	4/29/2020	Yukon	Canada	11	0	0

	Date	State	Region	Confirmed	Deaths	Recovered
217	4/29/2020	Greenland	Denmark	11	0	11
...	...	...	...	...	...	...
57	4/29/2020	NaN	France	165093	24087	48228
168	4/29/2020	NaN	UK	165221	26097	0
80	4/29/2020	NaN	Italy	203591	27682	71252
153	4/29/2020	NaN	Spain	236899	24275	132929
265	4/29/2020	New York	US	299691	23477	0

304 rows × 6 columns

In [ ]: