

In [1]:

```
import pandas as pd
```

In [7]:

```
data = pd.read_csv("Covid Data Analysis.csv")  
data
```

Out[7]:

	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 [9]:

```
data.count()
```

Out[9]:

```
Date      321  
State     140  
Region    321  
Confirmed 321  
Deaths    321  
Recovered 321  
dtype: int64
```

In [11]:

```
data.isnull().sum()
```

Out[11]:

```
Date      0  
State     181  
Region    0  
Confirmed 0  
Deaths    0  
Recovered 0  
dtype: int64
```

In [13]:

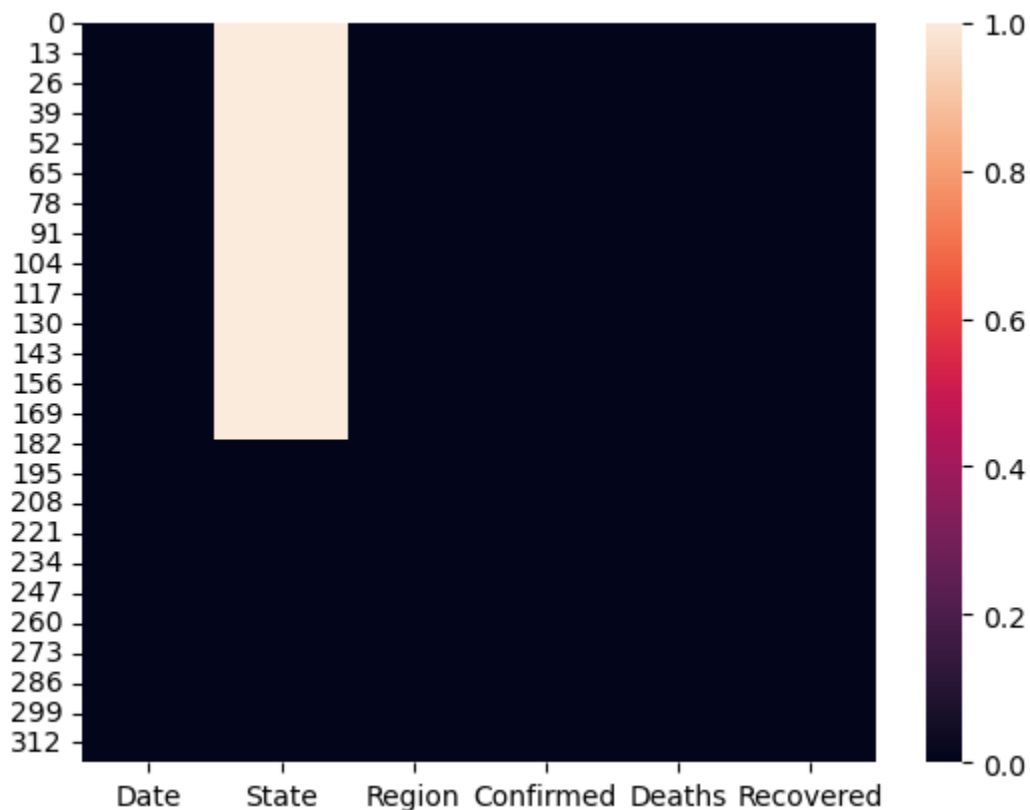
```
import seaborn as sns
```

In [15]:

```
import matplotlib.pyplot as plt
```

In [17]:

```
sns.heatmap(data.isnull())  
plt.show()
```



1) Show the number of Confirmed, Deaths and Recovered cases in each Region.

In [20]:

```
data.head(1)
```

Out[20]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252

In [34]:

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

Out[34]:

	Confirmed	Recovered
Region		
Afghanistan	1939	252
Albania	766	455
Algeria	3848	1702

	Confirmed	Recovered
Region		
Andorra	743	423
Angola	27	7
...
West Bank and Gaza	344	71
Western Sahara	6	5
Yemen	6	1
Zambia	97	54
Zimbabwe	32	5

187 rows × 2 columns

2) Remove all the records where the Confirmed Cases is Less Than 10.

In [37]:

```
data.head(1)
```

Out[37]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252

In [47]:

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

Out[47]:

	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

3. In which Region, maximum number of Confirmed cases were recorded ?

In [50]:

```
data.groupby('Region').Confirmed.sum().sort_values(ascending = False).head(20)
```

Out[50]:

Region	
US	1039909
Spain	236899
Italy	203591
France	166536
UK	166432
Germany	161539
Turkey	117589
Russia	99399
Iran	93657
Mainland China	82861
Brazil	79685
Canada	52860
Belgium	47859
Netherlands	38993
Peru	33931
India	33062
Switzerland	29407
Ecuador	24675
Portugal	24505
Saudi Arabia	21402

Name: Confirmed, dtype: int64

4) In which Region, minimum number of Deaths cases were recorded ?

In [59]:

```
data.groupby('Region').Deaths.sum().sort_values(ascending=True).head(20)
```

Out[59]:

Region	
Cambodia	0
Seychelles	0
Saint Lucia	0
Central African Republic	0
Saint Kitts and Nevis	0
South Sudan	0
Rwanda	0
Grenada	0
Macau	0
Madagascar	0
Nepal	0
Namibia	0
Saint Vincent and the Grenadines	0

```

Mozambique          0
Holy See            0
Timor-Leste         0
Mongolia            0
Uganda              0
Laos                0
Eritrea             0
Name: Deaths, dtype: int64

```

5) How many Confirmed, Deaths & Recovered cases were reported from India till 29 April 2020 ?

In [63]:

```
data.head(1)
```

Out[63]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252

In [67]:

```
data[data['Region'] == 'India']
```

Out[67]:

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

6-A) Sort the entire data wrt No. of Confirmed cases in ascending order.

In [72]:

```
data.sort_values(by = ['Confirmed'], ascending = True).head(50)
```

Out[72]:

	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
217	4/29/2020	Greenland	Denmark	11	0	11
256	4/29/2020	Montserrat	UK	11	1	2
144	4/29/2020	NaN	Seychelles	11	0	6
27	4/29/2020	NaN	Burundi	11	1	4
306	4/29/2020	Turks and Caicos Islands	UK	12	1	5
118	4/29/2020	NaN	Nicaragua	13	3	7

	Date	State	Region	Confirmed	Deaths	Recovered
215	4/29/2020	Grand Princess	Canada	13	0	0
206	4/29/2020	Falkland Islands (Malvinas)	UK	13	0	11
270	4/29/2020	Northern Mariana Islands	US	14	2	0
136	4/29/2020	NaN	Saint Kitts and Nevis	15	0	4
138	4/29/2020	NaN	Saint Vincent and the Grenadines	16	0	8
45	4/29/2020	NaN	Dominica	16	0	13
201	4/29/2020	Curacao	Netherlands	16	1	13
114	4/29/2020	NaN	Namibia	16	0	8
137	4/29/2020	NaN	Saint Lucia	17	0	15
260	4/29/2020	New Caledonia	France	18	0	17
16	4/29/2020	NaN	Belize	18	2	9
55	4/29/2020	NaN	Fiji	18	0	12
281	4/29/2020	Qinghai	Mainland China	18	0	18
90	4/29/2020	NaN	Laos	19	0	7
64	4/29/2020	NaN	Grenada	20	0	13
21	4/29/2020	NaN	Botswana	23	1	5
5	4/29/2020	NaN	Antigua and Barbuda	24	3	11
163	4/29/2020	NaN	Timor-Leste	24	0	6
279	4/29/2020	Prince Edward Island	Canada	27	0	0
4	4/29/2020	NaN	Angola	27	2	7
271	4/29/2020	Northern Territory	Australia	28	0	25
180	4/29/2020	NaN	Zimbabwe	32	4	5
152	4/29/2020	NaN	South Sudan	34	0	0
100	4/29/2020	NaN	Malawi	36	3	7
300	4/29/2020	St Martin	France	38	3	24
110	4/29/2020	NaN	Mongolia	38	0	10
51	4/29/2020	NaN	Eritrea	39	0	19
159	4/29/2020	NaN	Syria	43	3	21
244	4/29/2020	Macau	Macau	45	0	34
204	4/29/2020	Diamond Princess cruise ship	US	49	0	0
31	4/29/2020	NaN	Central African Republic	50	0	10
32	4/29/2020	NaN	Chad	52	2	19
310	4/29/2020	Virgin Islands	US	57	4	0

	Date	State	Region	Confirmed	Deaths	Recovered
115	4/29/2020	NaN	Nepal	57	0	16
210	4/29/2020	French Polynesia	France	58	0	50
94	4/29/2020	NaN	Libya	61	2	18
17	4/29/2020	NaN	Benin	64	1	33
196	4/29/2020	Cayman Islands	UK	73	1	10
296	4/29/2020	Sint Maarten	Netherlands	75	13	33
267	4/29/2020	Ningxia	Mainland China	75	0	75

6-B) Sort the entire data wrt No. of Recovered cases in descending order.

In [77]:

```
data.sort_values(by=['Recovered'], ascending = False).head(50)
```

Out[77]:

	Date	State	Region	Confirmed	Deaths	Recovered
153	4/29/2020	NaN	Spain	236899	24275	132929
61	4/29/2020	NaN	Germany	161539	6467	120400
76	4/29/2020	NaN	Iran	93657	5957	73791
80	4/29/2020	NaN	Italy	203591	27682	71252
229	4/29/2020	Hubei	Mainland China	68128	4512	63616
57	4/29/2020	NaN	France	165093	24087	48228
167	4/29/2020	NaN	Turkey	117589	3081	44040
22	4/29/2020	NaN	Brazil	79685	5513	34132
158	4/29/2020	NaN	Switzerland	29407	1716	22600
78	4/29/2020	NaN	Ireland	20253	1190	13386
8	4/29/2020	NaN	Austria	15402	580	12779
107	4/29/2020	NaN	Mexico	17799	1732	11423
15	4/29/2020	NaN	Belgium	47859	7501	11283
134	4/29/2020	NaN	Russia	99399	972	10286
128	4/29/2020	NaN	Peru	33931	943	10037
151	4/29/2020	NaN	South Korea	10765	247	9059
74	4/29/2020	NaN	India	33062	1079	8437
79	4/29/2020	NaN	Israel	15834	215	8233
33	4/29/2020	NaN	Chile	14885	216	8057
42	4/29/2020	NaN	Denmark	9008	443	6366

	Date	State	Region	Confirmed	Deaths	Recovered
101	4/29/2020	NaN	Malaysia	5945	100	4087
133	4/29/2020	NaN	Romania	11978	693	3569
124	4/29/2020	NaN	Pakistan	15525	343	3425
97	4/29/2020	NaN	Luxembourg	3769	89	3134
41	4/29/2020	NaN	Czech Republic	7579	227	3108
130	4/29/2020	NaN	Poland	12640	624	3025
141	4/29/2020	NaN	Saudi Arabia	21402	157	2953
56	4/29/2020	NaN	Finland	4906	206	2800
162	4/29/2020	NaN	Thailand	2947	54	2665
83	4/29/2020	NaN	Japan	13895	413	2368
171	4/29/2020	NaN	United Arab Emirates	11929	98	2329
264	4/29/2020	New South Wales	Australia	3016	40	2284
150	4/29/2020	NaN	South Africa	5350	103	2073
14	4/29/2020	NaN	Belarus	13181	84	2072
2	4/29/2020	NaN	Algeria	3848	444	1702
73	4/29/2020	NaN	Iceland	1797	10	1656
47	4/29/2020	NaN	Ecuador	24675	883	1557
220	4/29/2020	Guangdong	Mainland China	1588	8	1557
131	4/29/2020	NaN	Portugal	24505	973	1470
11	4/29/2020	NaN	Bahrain	2921	8	1455
34	4/29/2020	NaN	Colombia	6207	278	1411
75	4/29/2020	NaN	Indonesia	9771	784	1391
88	4/29/2020	NaN	Kuwait	3740	24	1389
77	4/29/2020	NaN	Iraq	2003	92	1346
48	4/29/2020	NaN	Egypt	5268	380	1335
309	4/29/2020	Victoria	Australia	1361	18	1291
38	4/29/2020	NaN	Croatia	2062	67	1288
9	4/29/2020	NaN	Azerbaijan	1766	23	1267
320	4/29/2020	Zhejiang	Mainland China	1268	1	1263
227	4/29/2020	Henan	Mainland China	1276	22	1254

In []: