In [42]:
```python
import pandas as pd
```

In [43]:
```python
data = pd.read_csv("D:\covid_19_data.csv")
```

In [44]:
```python
data
```

Out[44]:

	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 [21]:
```python
data.count()
```

Out[21]:
```
Date         321
State        140
Region       321
Confirmed    321
Deaths       321
Recovered    321
dtype: int64
```

In [22]: `data.isnull()`

Out[22]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	False	True	False	False	False	False
1	False	True	False	False	False	False
2	False	True	False	False	False	False
3	False	True	False	False	False	False
4	False	True	False	False	False	False
...	...	...	...	...	...	...
316	False	False	False	False	False	False
317	False	False	False	False	False	False
318	False	False	False	False	False	False
319	False	False	False	False	False	False
320	False	False	False	False	False	False

321 rows × 6 columns

In [23]: `data.isnull().sum()`

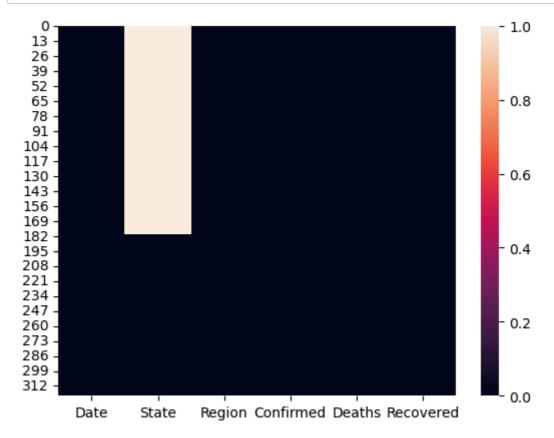Out[23]:
```
Date           0
State        181
Region         0
Confirmed      0
Deaths         0
Recovered      0
dtype: int64
```

In [24]: `import seaborn as sns`

In [25]: `import matplotlib.pyplot as plt`

In [26]:
```python
sns.heatmap(data.isnull())
plt.show()
```


In [27]:
```python
data.head(2)
```

Out[27]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252
1	4/29/2020	NaN	Albania	766	30	455

In [28]: `data.groupby('Region').sum()`

Out[28]:

Region	Date	State	Confirmed	Deaths	Recovered
Afghanistan	4/29/2020	0	1939	60	252
Albania	4/29/2020	0	766	30	455
Algeria	4/29/2020	0	3848	444	1702
Andorra	4/29/2020	0	743	42	423
Angola	4/29/2020	0	27	2	7
...	...	...	...	...	...
West Bank and Gaza	4/29/2020	0	344	2	71
Western Sahara	4/29/2020	0	6	0	5
Yemen	4/29/2020	0	6	0	1
Zambia	4/29/2020	0	97	3	54
Zimbabwe	4/29/2020	0	32	4	5

187 rows × 5 columns

In [29]: `data.groupby('Region').sum().head(20)`

Out[29]:

Region	Date	State	Confirmed	Deaths
Afghanistan	4/29/2020	0	1939	60
Albania	4/29/2020	0	766	30
Algeria	4/29/2020	0	3848	444
Andorra	4/29/2020	0	743	42
Angola	4/29/2020	0	27	2
Antigua and Barbuda	4/29/2020	0	24	3
Argentina	4/29/2020	0	4285	214
Armenia	4/29/2020	0	1932	30
Australia	4/29/20204/29/20204/29/20204/29/20204/29/20204...	Australian Capital TerritoryNew South WalesNor...	6752	91
Austria	4/29/2020	0	15402	580
Azerbaijan	4/29/2020	0	1766	23
Bahamas	4/29/2020	0	80	11
Bahrain	4/29/2020	0	2921	8
Bangladesh	4/29/2020	0	7103	163
Barbados	4/29/2020	0	80	7
Belarus	4/29/2020	0	13181	84
Belgium	4/29/2020	0	47859	7501
Belize	4/29/2020	0	18	2
Benin	4/29/2020	0	64	1
Bhutan	4/29/2020	0	7	0

In [30]: `data.groupby('Region')['Confirmed'].sum()`

Out[30]:
```
Region
Afghanistan          1939
Albania               766
Algeria              3848
Andorra               743
Angola                 27
                    ...
West Bank and Gaza    344
Western Sahara          6
Yemen                   6
Zambia                 97
Zimbabwe               32
Name: Confirmed, Length: 187, dtype: int64
```

In [31]: ```python
data.groupby('Region')['Confirmed'].sum().sort_values(ascending = Fals
```

Out[31]:
```
Region
US   1039909
Spain   236899
Italy   203591
France   166543
UK   166441
Germany   161539
Turkey   117589
Russia   99399
Iran   93657
Mainland China   82862
Name: Confirmed, dtype: int64
```

In [34]: ```python
data.groupby('Region')[['Confirmed', 'Recovered']].sum()
```

Out[34]:

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

In [35]: ```python
data.head(2)
```

Out[35]:

|  | Date | State | Region | Confirmed | Deaths | Recovered |
|---|---|---|---|---|---|---|
| **0** | 4/29/2020 | NaN | Afghanistan | 1939 | 60 | 252 |
| **1** | 4/29/2020 | NaN | Albania | 766 | 30 | 455 |

In [41]:
```python
data = data[~(data.Confirmed < 10)]
data
```

Out[41]:

	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 [45]:
```python
data.head(2)
```

Out[45]:

	Date	State	Region	Confirmed	Deaths	Recovered
0	4/29/2020	NaN	Afghanistan	1939	60	252
1	4/29/2020	NaN	Albania	766	30	455

In [47]:
```python
data.groupby('Region')['Confirmed'].sum().sort_values(ascending = Fals
```

Out[47]:
```
Region
US               1039909
Spain             236899
Italy             203591
France            166543
UK                166441
Germany           161539
Turkey            117589
Russia             99399
Iran               93657
Mainland China     82862
Brazil             79685
Canada             52865
Belgium            47859
Netherlands        38998
Peru               33931
India              33062
Switzerland        29407
Ecuador            24675
Portugal           24505
Saudi Arabia       21402
Name: Confirmed, dtype: int64
```

In [51]:
```python
data.groupby('Region').Deaths.sum().sort_values(ascending = True).head
```

Out[51]:
```
Region
Laos                                  0
Mongolia                              0
Mozambique                            0
Cambodia                              0
Fiji                                  0
Namibia                               0
Nepal                                 0
Madagascar                            0
Macau                                 0
Papua New Guinea                      0
Rwanda                                0
Saint Kitts and Nevis                 0
Bhutan                                0
Dominica                              0
Central African Republic              0
Saint Lucia                           0
Holy See                              0
Sao Tome and Principe                 0
Yemen                                 0
Western Sahara                        0
Eritrea                               0
Vietnam                               0
Saint Vincent and the Grenadines      0
Timor-Leste                           0
Uganda                                0
Grenada                               0
South Sudan                           0
Seychelles                            0
Liechtenstein                         1
Maldives                              1
Gambia                                1
Eswatini                              1
Guinea-Bissau                         1
Equatorial Guinea                     1
Mauritania                            1
Cabo Verde                            1
Benin                                 1
Burundi                               1
Suriname                              1
Brunei                                1
Botswana                              1
West Bank and Gaza                    2
Angola                                2
Belize                                2
Djibouti                              2
Chad                                  2
Libya                                 2
MS Zaandam                            2
Nicaragua                             3
Syria                                 3
Name: Deaths, dtype: int64
```

In [54]: `data.Region == 'India'`

Out[54]:
```
0        False
1        False
2        False
3        False
4        False
         ...
316      False
317      False
318      False
319      False
320      False
Name: Region, Length: 321, dtype: bool
```

In [55]: `data[data.Region == 'India']`

Out[55]:

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

In [56]:
```python
data[data.Region == 'US']
```

Out[56]:

	Date	State	Region	Confirmed	Deaths	Recovered
181	4/29/2020	Alabama	US	6912	256	0
182	4/29/2020	Alaska	US	355	9	0
186	4/29/2020	Arizona	US	7209	308	0
187	4/29/2020	Arkansas	US	3193	57	0
195	4/29/2020	California	US	48747	1946	0
199	4/29/2020	Colorado	US	14758	766	0
200	4/29/2020	Connecticut	US	26767	2169	0
202	4/29/2020	Delaware	US	4655	144	0
204	4/29/2020	Diamond Princess cruise ship	US	49	0	0
205	4/29/2020	District of Columbia	US	4106	205	0
208	4/29/2020	Florida	US	33193	1218	0
213	4/29/2020	Georgia	US	25775	1101	0
216	4/29/2020	Grand Princess	US	103	3	0
219	4/29/2020	Guam	US	141	5	0
224	4/29/2020	Hawaii	US	613	16	0
231	4/29/2020	Idaho	US	1952	60	0
232	4/29/2020	Illinois	US	50358	2215	0
233	4/29/2020	Indiana	US	17182	964	0
235	4/29/2020	Iowa	US	6843	148	0
240	4/29/2020	Kansas	US	3839	134	0
241	4/29/2020	Kentucky	US	4537	234	0
243	4/29/2020	Louisiana	US	27660	1845	0
245	4/29/2020	Maine	US	1056	52	0
248	4/29/2020	Maryland	US	20849	1078	0
249	4/29/2020	Massachusetts	US	60265	3405	0
251	4/29/2020	Michigan	US	40399	3670	0
252	4/29/2020	Minnesota	US	4644	319	0
253	4/29/2020	Mississippi	US	6569	250	0
254	4/29/2020	Missouri	US	7660	338	0
255	4/29/2020	Montana	US	451	16	0
257	4/29/2020	Nebraska	US	3851	56	0
258	4/29/2020	Nevada	US	4934	230	0
261	4/29/2020	New Hampshire	US	2058	60	0
262	4/29/2020	New Jersey	US	116365	6771	0
263	4/29/2020	New Mexico	US	3213	112	0
265	4/29/2020	New York	US	299691	23477	0
268	4/29/2020	North Carolina	US	10180	382	0

	Date	State	Region	Confirmed	Deaths	Recovered
269	4/29/2020	North Dakota	US	1033	19	0
270	4/29/2020	Northern Mariana Islands	US	14	2	0
274	4/29/2020	Ohio	US	17303	937	0
275	4/29/2020	Oklahoma	US	3473	214	0
277	4/29/2020	Oregon	US	2446	101	0
278	4/29/2020	Pennsylvania	US	46327	2373	0
280	4/29/2020	Puerto Rico	US	1433	86	0
285	4/29/2020	Recovered	US	0	0	120720
287	4/29/2020	Rhode Island	US	8247	251	0
298	4/29/2020	South Carolina	US	5882	231	0
299	4/29/2020	South Dakota	US	2373	13	0
302	4/29/2020	Tennessee	US	10366	195	0
303	4/29/2020	Texas	US	27257	754	0
307	4/29/2020	Utah	US	4497	45	0
308	4/29/2020	Vermont	US	862	47	0
310	4/29/2020	Virgin Islands	US	57	4	0
311	4/29/2020	Virginia	US	14962	522	0
312	4/29/2020	Washington	US	14070	801	0
313	4/29/2020	West Virginia	US	1110	38	0
315	4/29/2020	Wisconsin	US	6520	308	0
316	4/29/2020	Wyoming	US	545	7	0

In [59]:
```python
data.sort_values(by = ['Confirmed'], ascending = True).head(50)
```

Out[59]:

	Date	State	Region	Confirmed	Deaths	Recovered
285	4/29/2020	Recovered	US	0	0	120720
284	4/29/2020	Recovered	Canada	0	0	20327
203	4/29/2020	Diamond Princess cruise ship	Canada	0	1	0
305	4/29/2020	Tibet	Mainland China	1	0	1
289	4/29/2020	Saint Pierre and Miquelon	France	1	0	0
184	4/29/2020	Anguilla	UK	3	0	3
192	4/29/2020	Bonaire, Sint Eustatius and Saba	Netherlands	5	0	0
272	4/29/2020	Northwest Territories	Canada	5	0	0
288	4/29/2020	Saint Barthelemy	France	6	0	6
178	4/29/2020	NaN	Yemen	6	0	1
194	4/29/2020	British Virgin Islands	UK	6	1	3
177	4/29/2020	NaN	Western Sahara	6	0	5
18	4/29/2020	NaN	Bhutan	7	0	5
126	4/29/2020	NaN	Papua New Guinea	8	0	0
140	4/29/2020	NaN	Sao Tome and Principe	8	0	4
105	4/29/2020	NaN	Mauritania	8	1	6
98	4/29/2020	NaN	MS Zaandam	9	2	0
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
144	4/29/2020	NaN	Seychelles	11	0	6
27	4/29/2020	NaN	Burundi	11	1	4
256	4/29/2020	Montserrat	UK	11	1	2
318	4/29/2020	Yukon	Canada	11	0	0
217	4/29/2020	Greenland	Denmark	11	0	11
306	4/29/2020	Turks and Caicos Islands	UK	12	1	5
206	4/29/2020	Falkland Islands (Malvinas)	UK	13	0	11
118	4/29/2020	NaN	Nicaragua	13	3	7
215	4/29/2020	Grand Princess	Canada	13	0	0
270	4/29/2020	Northern Mariana Islands	US	14	2	0
136	4/29/2020	NaN	Saint Kitts and Nevis	15	0	4
45	4/29/2020	NaN	Dominica	16	0	13
138	4/29/2020	NaN	Saint Vincent and the Grenadines	16	0	8
114	4/29/2020	NaN	Namibia	16	0	8

	Date	State	Region	Confirmed	Deaths	Recovered
201	4/29/2020	Curacao	Netherlands	16	1	13
137	4/29/2020	NaN	Saint Lucia	17	0	15
260	4/29/2020	New Caledonia	France	18	0	17
281	4/29/2020	Qinghai	Mainland China	18	0	18
55	4/29/2020	NaN	Fiji	18	0	12
16	4/29/2020	NaN	Belize	18	2	9
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

In [60]:
```python
data.sort_values(by = ['Recovered'], ascending = False).head(50)
```

Out[60]:

	Date	State	Region	Confirmed	Deaths	Recovered
153	4/29/2020	NaN	Spain	236899	24275	132929
285	4/29/2020	Recovered	US	0	0	120720
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
284	4/29/2020	Recovered	Canada	0	0	20327
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
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

	Date	State	Region	Confirmed	Deaths	Recovered
73	4/29/2020	NaN	Iceland	1797	10	1656
220	4/29/2020	Guangdong	Mainland China	1588	8	1557
47	4/29/2020	NaN	Ecuador	24675	883	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