In [39]:
```
1  import pandas as pd
```

In [40]:
```
1  data = pd.read_csv(r"covid_19_data.csv")
```

In [72]:
```
1  data
```

Out[72]:

	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 [42]:
```
1  #
```

In [43]:
```
1  data.count()
```

Out[43]:
```
Date         321
State        140
Region       321
Confirmed    321
Deaths       321
Recovered    321
dtype: int64
```

In [44]:
```
1  data.isnull().sum()
```
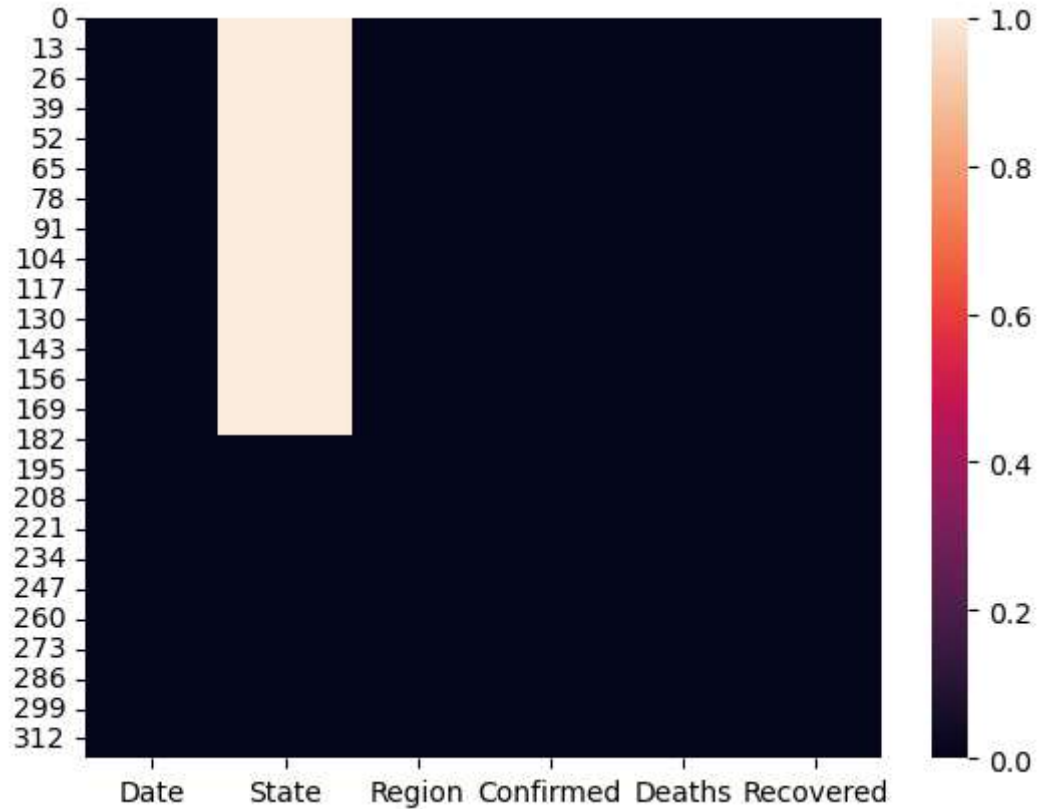
Out[44]:
```
Date          0
State        181
Region        0
Confirmed     0
Deaths        0
Recovered     0
dtype: int64
```

```
In [45]:    1  import seaborn as sns
```

```
In [46]:    1  import matplotlib.pyplot as plt
```

```
In [47]:    1  sns.heatmap(data.isnull())
            2  plt.show()
```



```
In [ ]:     1
```

The number of confirmed,deaths and recovered cases in each region.

```
In [48]:    1  data.head(2)
```

Out[48]:

	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 [54]:
```
1  data.groupby('Region').sum()
```

Out[54]:

	Date	State	Confirmed	Deaths	Recovered
Region					
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

Removal of record where confirmed cases isless than 10

In [56]:
```
1  data[data.Confirmed < 10]
```

Out[56]:

	Date	State	Region	Confirmed	Deaths	Recovered
18	4/29/2020	NaN	Bhutan	7	0	5
98	4/29/2020	NaN	MS Zaandam	9	2	0
105	4/29/2020	NaN	Mauritania	8	1	6
126	4/29/2020	NaN	Papua New Guinea	8	0	0
140	4/29/2020	NaN	Sao Tome and Principe	8	0	4
177	4/29/2020	NaN	Western Sahara	6	0	5
178	4/29/2020	NaN	Yemen	6	0	1
184	4/29/2020	Anguilla	UK	3	0	3
192	4/29/2020	Bonaire, Sint Eustatius and Saba	Netherlands	5	0	0
194	4/29/2020	British Virgin Islands	UK	6	1	3
203	4/29/2020	Diamond Princess cruise ship	Canada	0	1	0
272	4/29/2020	Northwest Territories	Canada	5	0	0
284	4/29/2020	Recovered	Canada	0	0	20327
285	4/29/2020	Recovered	US	0	0	120720
288	4/29/2020	Saint Barthelemy	France	6	0	6
289	4/29/2020	Saint Pierre and Miquelon	France	1	0	0
305	4/29/2020	Tibet	Mainland China	1	0	1

In [59]:
```python
1  data=data[~(data.Confirmed < 10)]
2  data
```

Out[59]:

	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 [60]:
```python
1  data.head(10)
```

Out[60]:

	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
5	4/29/2020	NaN	Antigua and Barbuda	24	3	11
6	4/29/2020	NaN	Argentina	4285	214	1192
7	4/29/2020	NaN	Armenia	1932	30	900
8	4/29/2020	NaN	Austria	15402	580	12779
9	4/29/2020	NaN	Azerbaijan	1766	23	1267

In which region, maximum number of Confirmed cases where recorded

In [62]:
```
1 data.groupby('Region').Confirmed.sum().sort_values(ascending=False).head(2
```

Out[62]:
```
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
```

In [ ]:
```
1
```

In which Region, minimum number of Deaths cases were recorded

In [65]:
```python
1  data.groupby('Region').Deaths.sum().sort_values().head(20)
```

Out[65]:
```
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
```

In [ ]:
```python
1
```

Confirmed,Deaths & REcovered cases reported from india till 29 April 2020

In [69]:
```python
1  data[data.Region=='India']
```

Out[69]:

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

In [ ]:
```python
1
```

Sort the number of data wrt no. of Confirmed cases in ascending order.

In [74]:
```
1  data.sort_values(by = ['Confirmed'],ascending=True)
```

Out[74]:

	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
...	...	...	...	...	...	...
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

Sort the entir of data wrt no. of Recovered cases in decending order.

In [75]:
```
1  data.sort_values(by=['Recovered'],ascending=False)
```

Out[75]:

	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
...	...	...	...	...	...	...
258	4/29/2020	Nevada	US	4934	230	0
257	4/29/2020	Nebraska	US	3851	56	0
255	4/29/2020	Montana	US	451	16	0
254	4/29/2020	Missouri	US	7660	338	0
274	4/29/2020	Ohio	US	17303	937	0

304 rows × 6 columns

In [ ]:

```
1
```