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

```
In [2]: covid=pd.read_csv(r'E:\Personal Python Projects\Covid data  set\Covid_19_data.csv')
```

```
In [3]: covid
```

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 [4]: covid.count()
```

```
Out[4]: Date         321
        State        140
        Region       321
        Confirmed    321
        Deaths       321
        Recovered    321
        dtype: int64
```

```
In [5]: covid.isnull().sum()
```

```
Out[5]: Date           0
        State        181
        Region         0
        Confirmed      0
        Deaths         0
        Recovered      0
        dtype: int64
```

using seborne lib to visualise null values via heatmap

In [6]: `import seaborn as sns`

In [7]: `sns.heatmap(covid.isnull())`

Out[7]: `<Axes: >`

show number of confirmed deaths and recovered case region wise

In [18]: `covid.groupby('Region')[['Confirmed','Recovered']].sum()`

Out[18]:

Region	Confirmed	Recovered
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

remove all data weher confirmed cases is less than 10

```
In [36]: covid=covid[~covid['Confirmed']<10]
```

```
In [37]: covid
```

Out[37]:

	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 which region were max no. of covid cases were recorded

```
In [46]: covid.groupby('Region')['Confirmed'].sum().sort_values(ascending=False).head(1)
```

```
Out[46]: Region
         US    1039909
         Name: Confirmed, dtype: int64
```

in which region min no. of death cases were recorded?

```
In [54]: death_min=covid.groupby('Region')['Deaths'].sum().reset_index()
         death_min[death_min['Deaths']==0]
```

Out[54]:

	Region	Deaths
19	Bhutan	0
30	Cambodia	0
33	Central African Republic	0
47	Dominica	0
53	Eritrea	0
57	Fiji	0
66	Grenada	0
72	Holy See	0
93	Laos	0
102	Macau	0
103	Madagascar	0
115	Mongolia	0
118	Mozambique	0
119	Namibia	0
120	Nepal	0
131	Papua New Guinea	0
140	Rwanda	0
141	Saint Kitts and Nevis	0
142	Saint Lucia	0
143	Saint Vincent and the Grenadines	0
145	Sao Tome and Principe	0
149	Seychelles	0
157	South Sudan	0
168	Timor-Leste	0
175	Uganda	0
181	Vietnam	0
183	Western Sahara	0
184	Yemen	0

confirmed, death & recovered cased from India till 29 apr 2020

In [62]:
```python
covid[(covid.Region=='India') & (covid['Date']<="2020-04-29")]
```

Out[62]:

Date	State	Region	Confirmed	Deaths	Recovered

sort entire data wrt confirmed cases in asc order

In [64]:
```python
covid.sort_values(by=['Confirmed'],ascending=True)
```

Out[64]:

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

321 rows × 6 columns

sort entire data wrt no of recovered cases in des order

In [65]:
```python
covid.sort_values(by=['Recovered'],ascending=False)
```

Out[65]:

	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
...	...	...	...	...	...	...
248	4/29/2020	Maryland	US	20849	1078	0
246	4/29/2020	Manitoba	Canada	275	6	0
243	4/29/2020	Louisiana	US	27660	1845	0
241	4/29/2020	Kentucky	US	4537	234	0
215	4/29/2020	Grand Princess	Canada	13	0	0

321 rows × 6 columns

In [ ]: