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

In [2]:
```
data = pd.read_csv("covid_19_india.csv")
data.head()
```

Out[2]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignNational	Cur
0	1	2020-01-30	6:00 PM	Kerala	1	0	
1	2	2020-01-31	6:00 PM	Kerala	1	0	
2	3	2020-02-01	6:00 PM	Kerala	2	0	
3	4	2020-02-02	6:00 PM	Kerala	3	0	
4	5	2020-02-03	6:00 PM	Kerala	3	0	

In [3]:
```
data.head(10)
```

Out[3]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignNational	Cur
0	1	2020-01-30	6:00 PM	Kerala	1	0	
1	2	2020-01-31	6:00 PM	Kerala	1	0	
2	3	2020-02-01	6:00 PM	Kerala	2	0	
3	4	2020-02-02	6:00 PM	Kerala	3	0	
4	5	2020-02-03	6:00 PM	Kerala	3	0	
5	6	2020-02-04	6:00 PM	Kerala	3	0	
6	7	2020-02-05	6:00 PM	Kerala	3	0	
7	8	2020-02-06	6:00 PM	Kerala	3	0	
8	9	2020-02-07	6:00 PM	Kerala	3	0	
9	10	2020-02-08	6:00 PM	Kerala	3	0	

In [4]:
```
data.tail(10)
```

Out[4]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignNationa
18100	18101	2021-08-11	8:00 AM	Puducherry	-	
18101	18102	2021-08-11	8:00 AM	Punjab	-	
18102	18103	2021-08-11	8:00 AM	Rajasthan	-	
18103	18104	2021-08-11	8:00 AM	Sikkim	-	
18104	18105	2021-08-11	8:00 AM	Tamil Nadu	-	
18105	18106	2021-08-11	8:00 AM	Telangana	-	
18106	18107	2021-08-11	8:00 AM	Tripura	-	
18107	18108	2021-08-11	8:00 AM	Uttarakhand	-	
18108	18109	2021-08-11	8:00 AM	Uttar Pradesh	-	
18109	18110	2021-08-11	8:00 AM	West Bengal	-	

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

Out[5]:
```
Sno                        18110
Date                       18110
Time                       18110
State/UnionTerritory       18110
ConfirmedIndianNational    18110
ConfirmedForeignNational   18110
Cured                      18110
Deaths                     18110
Confirmed                  18110
dtype: int64
```

In [6]:
```
from sklearn.preprocessing import LabelEncoder
enc = LabelEncoder()
```

In [7]:
```
%matplotlib inline
import matplotlib.pyplot as plt
```

In [8]:
```
len(data)
```

Out[8]:
```
18110
```

In [9]:
```
len(data.columns)
```

Out[9]:
```
9
```

In [10]:
```
import seaborn as sns
```

In [11]:
```
statewise = pd.pivot_table(data, values=['Confirmed','Deaths','Cured'],index= 'State
```

In [12]:
```
statewise['Recovery Rate'] = statewise ['Cured']*100/ statewise['Confirmed']
statewise['Mortality Rate'] = statewise ['Deaths']*100/ statewise['Confirmed']
statewise = statewise.sort_values( by = 'Confirmed', ascending = False)
statewise.style.background_gradient(cmap="cubehelix")
```

Out[12]:

State/UnionTerritory	Confirmed	Cured	Deaths	Recovery Rate	Mortality Rate
Maharashtra	6363442	6159676	134201	96.797865	2.108937
Maharashtra***	6229596	6000911	130753	96.329056	2.098900
Kerala	3586693	3396184	18004	94.688450	0.501967
Karnataka	2921049	2861499	36848	97.961349	1.261465
Karanataka	2885238	2821491	36197	97.790581	1.254559
Tamil Nadu	2579130	2524400	34367	97.877967	1.332504
Andhra Pradesh	1985182	1952736	13564	98.365591	0.683262
Uttar Pradesh	1708812	1685492	22775	98.635309	1.332797
West Bengal	1534999	1506532	18252	98.145471	1.189056
Delhi	1436852	1411280	25068	98.220276	1.744647
Chhattisgarh	1003356	988189	13544	98.488373	1.349870
Odisha	988997	972710	6565	98.353180	0.663804
Rajasthan	953851	944700	8954	99.040626	0.938721
Gujarat	825085	814802	10077	98.753704	1.221329
Madhya Pradesh	791980	781330	10514	98.655269	1.327559
Madhya Pradesh***	791656	780735	10506	98.620487	1.327092
Haryana	770114	759790	9652	98.659419	1.253321
Bihar	725279	715352	9646	98.631285	1.329971
Bihar****	715730	701234	9452	97.974655	1.320610
Telangana	650353	638410	3831	98.163613	0.589065
Punjab	599573	582791	16322	97.201008	2.722271
Assam	576149	559684	5420	97.142232	0.940729
Telengana	443360	362160	2312	81.685312	0.521472
Jharkhand	347440	342102	5130	98.463620	1.476514
Uttarakhand	342462	334650	7368	97.718871	2.151480
Jammu and Kashmir	322771	317081	4392	98.237140	1.360717
Himachal Pradesh	208616	202761	3537	97.193408	1.695460

State/UnionTerritory	Confirmed	Cured	Deaths	Recovery Rate	Mortality Rate
Himanchal Pradesh	204516	200040	3507	97.811418	1.714780
Goa	172085	167978	3164	97.613389	1.838626
Puducherry	121766	119115	1800	97.822873	1.478245
Manipur	105424	96776	1664	91.796934	1.578388
Tripura	80660	77811	773	96.467890	0.958344
Meghalaya	69769	64157	1185	91.956313	1.698462
Chandigarh	61992	61150	811	98.641760	1.308233
Arunachal Pradesh	50605	47821	248	94.498567	0.490070
Mizoram	46320	33722	171	72.802245	0.369171
Nagaland	28811	26852	585	93.200514	2.030474
Sikkim	28018	25095	356	89.567421	1.270612
Ladakh	20411	20130	207	98.623291	1.014159
Dadra and Nagar Haveli and Daman and Diu	10654	10646	4	99.924911	0.037545
Dadra and Nagar Haveli	10377	10261	4	98.882143	0.038547
Lakshadweep	10263	10165	51	99.045114	0.496931
Cases being reassigned to states	9265	0	0	0.000000	0.000000
Andaman and Nicobar Islands	7548	7412	129	98.198198	1.709062
Unassigned	77	0	0	0.000000	0.000000
Daman & Diu	2	0	0	0.000000	0.000000

In [13]:
```
data['Active_Cases']=data['Confirmed']-(data['Cured']-data['Deaths'])
```

In [14]:
```
#Top 10 active cases state
Top_10_active_cases = data.groupby(by = 'State/UnionTerritory').max()[['Active_Cases
```

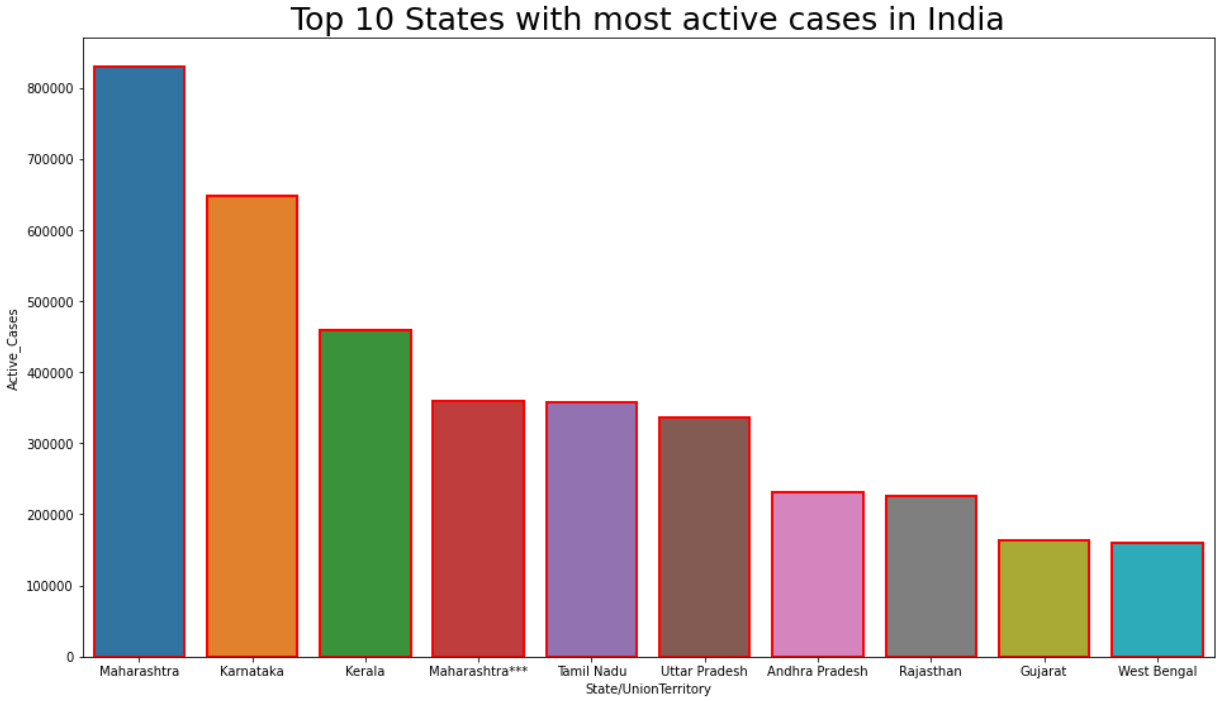
In [15]:
```
Top_10_active_cases.head()
```

Out[15]:

	State/UnionTerritory	Active_Cases	Date
0	Maharashtra	829727	2021-08-11
1	Karnataka	648383	2021-08-11
2	Kerala	458370	2021-08-11
3	Maharashtra***	359438	2021-07-21
4	Tamil Nadu	357936	2021-08-11

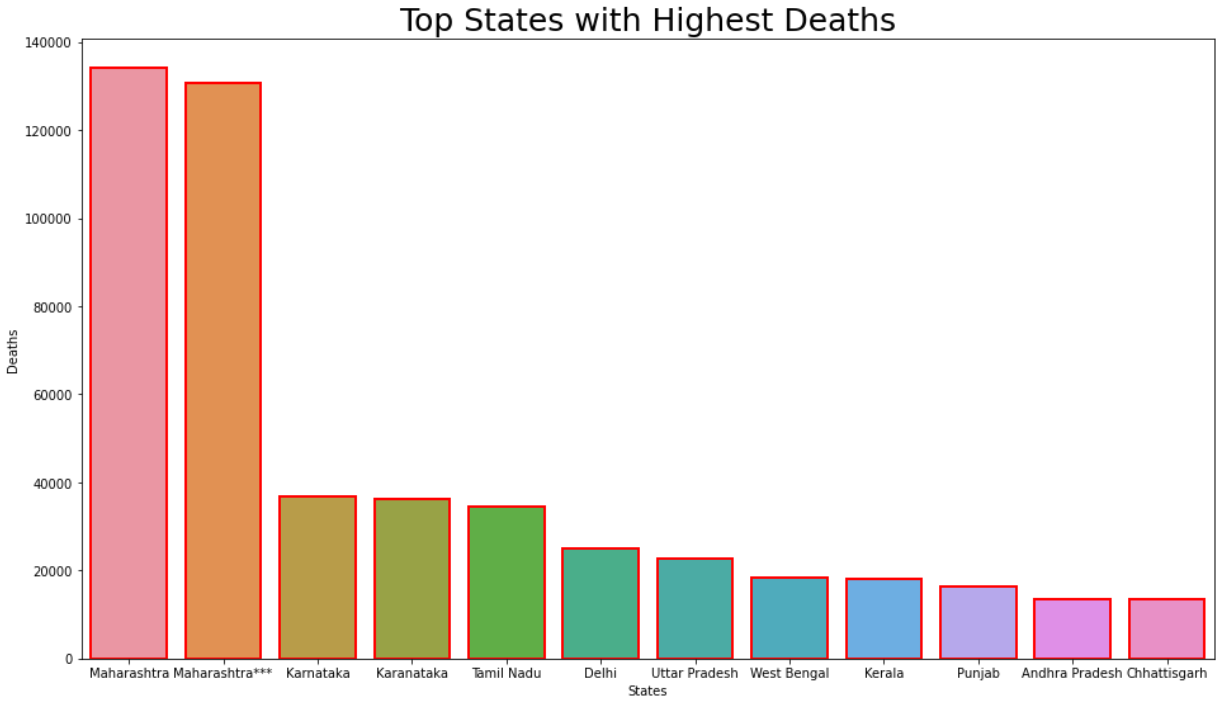
In [16]:
```
fig = plt.figure(figsize=(16,9))
```

```
plt.title("Top 10 States with most active cases in India",size =25)
ax= sns.barplot(data = Top_10_active_cases.iloc[:10],y = 'Active_Cases',x='State/Uni
```



In [17]:

```
Top_10_highest_deaths= data.groupby(by = 'State/UnionTerritory').max()[['Deaths','Da
fig = plt.figure(figsize=(16,9))
plt.title("Top States with Highest Deaths", size=25)
ax= sns.barplot(data = Top_10_highest_deaths.iloc[:12], x='State/UnionTerritory',y='
plt.xlabel('States')
plt.ylabel('Deaths')
plt.show()
```

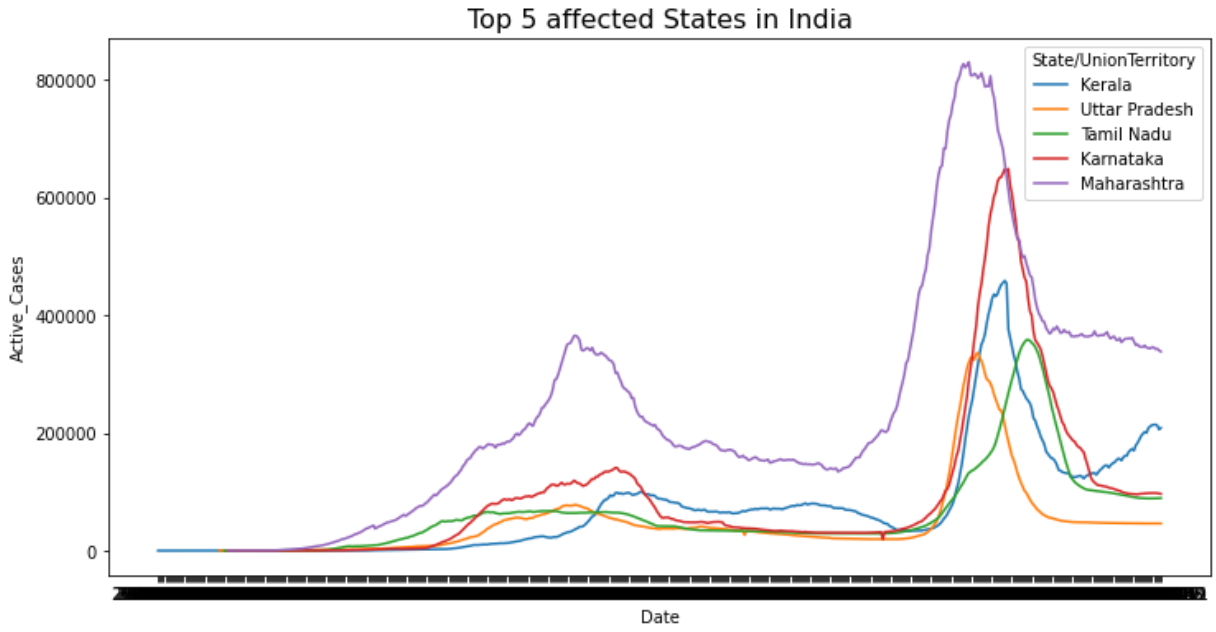


In [18]:

```
fig= plt.figure(figsize=(12,6))
ax = sns.lineplot(data= data[data['State/UnionTerritory'].isin(['Maharashtra','Karna
ax.set_title("Top 5 affected States in India", size=16)
```

Out[18]:

```
Text(0.5, 1.0, 'Top 5 affected States in India')
```



https://www.kaggle.com/code/sriramcs/covid-data-analysis-project/notebook

In [ ]: