

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
from plotly.subplots import make_subplots
```

```
covid_df = pd.read_csv('/content/covid_19_india.csv')
```

```
covid_df.head(10)
```

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



```
covid_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18110 entries, 0 to 18109
Data columns (total 9 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Sno                                    18110 non-null  int64
1   Date                                  18110 non-null  object
2   Time                                  18110 non-null  object
3   State/UnionTerritory                 18110 non-null  object
4   ConfirmedIndianNational              18110 non-null  object
5   ConfirmedForeignNational             18110 non-null  object
6   Cured                                18110 non-null  int64
7   Deaths                               18110 non-null  int64
8   Confirmed                            18110 non-null  int64
dtypes: int64(4), object(5)
memory usage: 1.2+ MB
```

```
covid_df.describe()
```

	Sno	Cured	Deaths	Confirmed
count	18110.000000	1.811000e+04	18110.000000	1.811000e+04
mean	9055.500000	2.786375e+05	4052.402264	3.010314e+05
std	5228.051023	6.148909e+05	10919.076411	6.561489e+05
min	1.000000	0.000000e+00	0.000000	0.000000e+00

```
vaccine_df = pd.read_csv('/content/covid_vaccine_statewise.csv')
```

```
vaccine_df.head(7)
```

	Updated On	State	Total Doses Administered	Sessions	Sites	First Dose Administered	Second Dose Administered
0	16/01/2021	India	48276.0	3455.0	2957.0	48276.0	0.0
1	17/01/2021	India	58604.0	8532.0	4954.0	58604.0	0.0
2	18/01/2021	India	99449.0	13611.0	6583.0	99449.0	0.0
3	19/01/2021	India	195525.0	17855.0	7951.0	195525.0	0.0
4	20/01/2021	India	251280.0	25472.0	10504.0	251280.0	0.0
5	21/01/2021	India	365965.0	32226.0	12600.0	365965.0	0.0
6	22/01/2021	India	549381.0	36988.0	14115.0	549381.0	0.0

7 rows × 24 columns



```
covid_df.drop(["Sno","Time","ConfirmedIndianNational","ConfirmedForeignNational"], inplace = True, axis = 1)
covid_df.head()
```

	Date	State/UnionTerritory	Cured	Deaths	Confirmed
0	2020-01-30	Kerala	0	0	1
1	2020-01-31	Kerala	0	0	1
2	2020-02-01	Kerala	0	0	2
3	2020-02-02	Kerala	0	0	3
4	2020-02-03	Kerala	0	0	3

```
covid_df['Date'] = pd.to_datetime(covid_df['Date'],format = '%Y-%m-%d')
covid_df.head()
```

	Date	State/UnionTerritory	Cured	Deaths	Confirmed
0	2020-01-30	Kerala	0	0	1
1	2020-01-31	Kerala	0	0	1
2	2020-02-01	Kerala	0	0	2
3	2020-02-02	Kerala	0	0	3
4	2020-02-03	Kerala	0	0	3

```
covid_df['Active_cases'] = covid_df['Confirmed']-(covid_df['Cured']+covid_df['Deaths'])
covid_df.tail()
```

	Date	State/UnionTerritory	Cured	Deaths	Confirmed	Active_cases
18105	2021-08-11	Telangana	638410	3831	650353	8112
18106	2021-08-11	Tripura	77811	773	80660	2076

```
statewise = pd.pivot_table(covid_df, values = ["Confirmed", "Deaths", "Cured"], index = "State/UnionTerritory", aggfunc = m
```

18108	2021-08-11	Uttar Pradesh	1685492	22775	1708812	545
--------------	------------	---------------	---------	-------	---------	-----

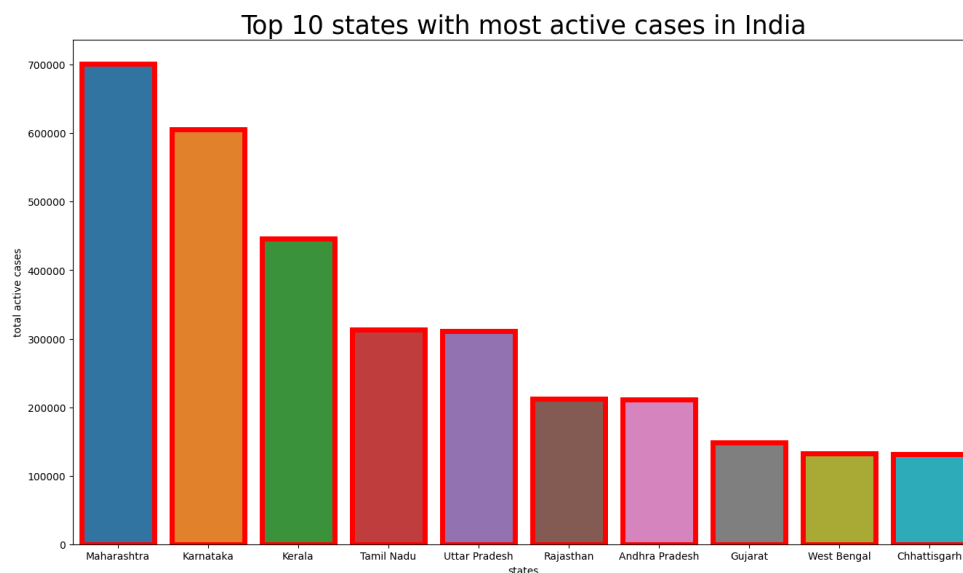
```
statewise["Recovery Rate"] = statewise["Cured"]*100/statewise['Confirmed']
statewise["Mortality Rate"] = statewise["Deaths"]*100/statewise['Confirmed']
statewise = statewise.sort_values(by = "Confirmed", ascending = False)
statewise.style.background_gradient(cmap = "cubehelix")
```

	Confirmed	Cured	Deaths	Recovery Rate	Mortality Rate
State/UnionTerritory					
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

```

top_10_active_cases = covid_df.groupby(by = 'State/UnionTerritory').max()[['Active_cases', 'Date']].sort_values(by = ['Active_cases', 'Date'], ascending = False)
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/UnionTerritory", linewidth = 5, edgecolor = 'b')
plt.xlabel("states")
plt.ylabel("total active cases")
plt.show()

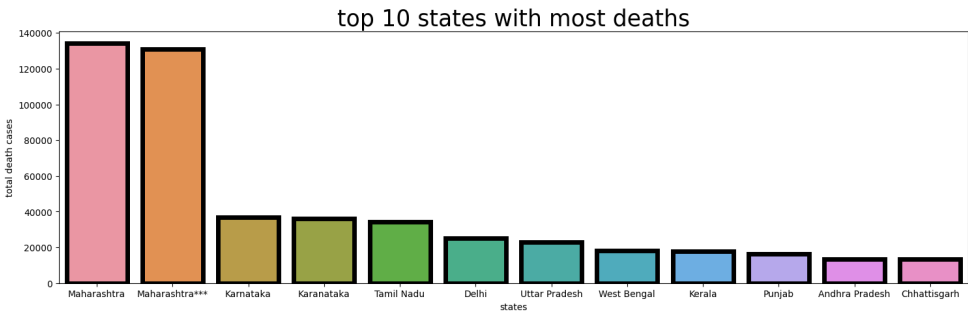
```



```

top_10_deaths = covid_df.groupby(by = 'State/UnionTerritory').max()[['Deaths', 'Date']].sort_values(by = ['Deaths', 'Date'], ascending = False)
fig = plt.figure(figsize=(18,5))
plt.title("top 10 states with most deaths", size = 25)
ax = sns.barplot(data = top_10_deaths.iloc[:12], y = "Deaths", x = "State/UnionTerritory", linewidth = 5, edgecolor = 'b')
plt.xlabel("states")
plt.ylabel("total death cases")
plt.show()

```



[Colab paid products](#) - [Cancel contracts here](#)

✓ 0s completed at 12:30 PM

