```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
%matplotlib inline
```

```
covid_df = pd.read_csv("C:/Users/Asus/Downloads/Latest Covid-19 India (Active.csv")
```

```
covid_df.head()
```

| | State/UTs | Total Cases | Active | Discharged | Deaths | Active Ratio | Discharge Ratio | Death Ratio | Population |
|---|---|---|---|---|---|---|---|---|---|
| 0 | Maharashtra and Dadra | 13747 | 9 | 13547 | 191 | 0.07 | 98.55 | 1.39 | 12500000.0 |
| 1 | Andhra Pradesh | 2330670 | 7 | 2324168 | 14725 | 0.00 | 99.72 | 0.63 | 53000000.0 |
| 2 | Arunachal Pradesh | 50001 | 5 | 50000 | 206 | 0.00 | 99.99 | 0.41 | 1800.0 |
| 3 | Assam | 740158 | 8 | 738030 | 8015 | 0.00 | 99.65 | 1.08 | 2800000 |
| 4 | Bihar | 707918 | 2 | 706380 | 7121 | 0.00 | 99.78 | 1.01 | 10000000 |
| 5 | Chandigarh | 71756 | 6 | 70746 | 804 | 0.01 | 98.59 | 1.12 | 1000000.0 |
| 6 | Chhattisgarh | 1177828 | 7 | 1164019 | 14102 | 0.00 | 98.83 | 1.20 | 29000007 |
| 7 | Dadra and Nagar Haveli and Daman and Diu | 1391 | 0 | 1387 | 4 | 0.00 | 99.71 | 0.29 | 2220000.0 |
| 8 | Delhi | 2007152 | 19 | 1967750 | 20322 | 0.00 | 98.04 | 1.02 | 172000 |
| 9 | Goa | 250120 | 35 | 245104 | 4521 | 0.01 | 98.00 | 1.80 | 3771000 |

```
covid_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36 entries, 0 to 35
Data columns (total 9 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   State/UTs       36 non-null     object
 1   Total Cases     36 non-null     int64
 2   Active          36 non-null     int64
 3   Discharged      36 non-null     int64
 4   Deaths          36 non-null     int64
 5   Active Ratio    36 non-null     float64
 6   Discharge Ratio 36 non-null     float64
 7   Death Ratio     36 non-null     float64
 8   Population      36 non-null     float64
dtypes: float64(5), int64(4), object(1)
memory usage: 2.7 KB
```

```
covid_df.describe()
```

| | Total Cases | Active | Discharged | Deaths | Active Ratio | Discharge Ratio | Death Ratio | Population |
|---|---|---|---|---|---|---|---|---|
| **count** | 2.420000e+03 | 36.000000 | 2.420000e+03 | 36.000000 | 36.000000 | 36.000000 | 36.000000 | 2.420000e+03 |
| **mean** | 1.151824e+06 | 30.777778 | 1.151046e+06 | 14722.388889 | 0.010222 | 99.989333 | 1.210000 | 57.181818e+07 |
| **std** | 1.434111e+06 | 214.407762 | 1.433124e+06 | 27170.080767 | 0.007061 | 0.49744 | 0.496474 | 5.293114e+07 |
| **min** | 1.377500e+04 | 0.000000 | 1.361000e+04 | 4.000000 | 0.000000 | 97.440000 | 0.290000 | 0.000164e+00 |
| **25%** | 1.500000e+05 | 5.000000 | 1.496000e+05 | 1125.750000 | 0.000000 | 99.675000 | 0.962500 | 1.000175e+08 |
| **50%** | 7.000000e+05 | 7.000000 | 6.970000e+05 | 7121.000000 | 0.010000 | 99.980000 | 1.080000 | 2.000000e+08 |
| **75%** | 1.225000e+06 | 17.000000 | 1.211000e+06 | 14925.750000 | 0.014000 | 99.997500 | 1.292500 | 4.500000e+07 |
| **max** | 6.150000e+06 | 1283.000000 | 7.900000e+06 | 164616.000000 | 0.040000 | 99.970000 | 2.400000 | 2.310000e+08 |

```
covid_df = pd.read_csv("C:/Users/Asus/Downloads/COVID-19 India Statewise Vaccine Data.csv")
```

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

| | State/UTs | Total Vaccination Doses | Dose 1 | Dose 2 | Population |
|---|---|---|---|---|---|
| 0 | Andaman and Nicobar | 624504 | 311993 | 312511 | 380581 |
| 1 | Andhra Pradesh | 8414787 | 4918243 | 3496544 | 8175276 |
| 2 | Arunachal Pradesh | 1390148 | 850710 | 539438 | 1710387 |
| 3 | Assam | 11490104 | 6938462 | 4551642 | 30967791 |
| 4 | Bihar | 12480330 | 6938866 | 5541464 | 12456956 |
| 5 | Chandigarh | 1108 | 1108 | 0 | 1000000 |
| 6 | Chhattisgarh | 8667948 | 4965121 | 3702827 | 31387722 |
| 7 | Dadra and Nagar Haveli and Daman and Diu | 1130914 | 728015 | 402899 | 1739887 |
| 8 | Delhi | 8771555 | 4905040 | 3866515 | 20281000 |
| 9 | Goa | 1070500 | 1000456 | 1070044 | 1555000 |

```
covid_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36 entries, 0 to 35
Data columns (total 5 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   State/UTs               36 non-null     object
 1   Total Vaccination Doses 36 non-null     int64
 2   Dose 1                  36 non-null     int64
 3   Dose 2                  36 non-null     int64
 4   Population              36 non-null     int64
dtypes: int64(4), object(1)
memory usage: 1.5 KB
```

```
covid_df.isnull().sum()
```

```
State/UTs                 0
Total Vaccination Doses   0
Dose 1                    0
Dose 2                    0
Population                0
dtype: int64
```

```
covid_df.describe(include=['object'])
```

| | State/UTs | Total Vaccination Doses | Dose 1 | Dose 2 | Population |
|---|---|---|---|---|---|
| **count** | 36 | 36 | 36 | 36 | 36 |
| **unique** | 36 | 36 | 36 | 36 | 36 |
| **top** | Andaman and Nicobar | 624504 | 311993 | 312511 | 380581 |
| **freq** | 1 | 1 | 1 | 1 | 1 |

```
covid_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36 entries, 0 to 35
Data columns (total 5 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   State/UTs               36 non-null     object
 1   Total Vaccination Doses 36 non-null     int64
 2   Dose 1                  36 non-null     int64
 3   Dose 2                  36 non-null     int64
 4   Population              36 non-null     int64
dtypes: int64(4), object(1)
memory usage: 1.5 KB
```

```
covid_df.isnull().sum()
```

```
State/UTs                 0
Total Vaccination Doses   0
Dose 1                    0
Dose 2                    0
Population                0
dtype: int64
```

```
covid_df.describe(include=['object'])
```

| | State/UTs | Total Vaccination Doses | Dose 1 | Dose 2 | Population |
|---|---|---|---|---|---|
| **count** | 36 | 36 | 36 | 36 | 36 |
| **unique** | 36 | 36 | 36 | 36 | 36 |
| **top** | Andaman and Nicobar | 624504 | 311993 | 312511 | 380581 |
| **freq** | 1 | 1 | 1 | 1 | 1 |

```
covid_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36 entries, 0 to 35
Data columns (total 5 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   State/UTs               36 non-null     object
 1   Total Vaccination Doses 36 non-null     int64
 2   Dose 1                  36 non-null     int64
 3   Dose 2                  36 non-null     int64
 4   Population              36 non-null     int64
dtypes: int64(4), object(1)
memory usage: 1.5 KB
```

```
covid_df.isnull().sum()
```

```
State/UTs                 0
Total Vaccination Doses   0
Dose 1                    0
Dose 2                    0
Population                0
dtype: int64
```

```
covid_df.describe(include=['object'])
```

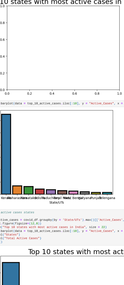
| | State/UTs | Total Vaccination Doses | Dose 1 | Dose 2 | Population |
|---|---|---|---|---|---|
| **count** | 36 | 36 | 36 | 36 | 36 |
| **unique** | 36 | 36 | 36 | 36 | 36 |
| **top** | Andaman and Nicobar | 624504 | 311993 | 312511 | 380581 |
| **freq** | 1 | 1 | 1 | 1 | 1 |

```
# Top 10 active cases states
top_10_active_cases = covid_df.groupby('State/UTs').max()['Active Cases', 'Discharged'].sort_values(by = ['Active Cases'], ascending = False).reset_index()
top_10_active_cases
```

```
Fig = px.figure(figsize=(12,6))
figure size 1200x600 with 0 Axes
```

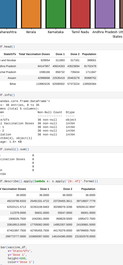
```
plt.title('Top 10 states with most active cases in India', size = 20)
Text(0.5, 1.0, 'Top 10 states with most active cases in India')
```

```
ax = sns.barplot(data = top_10_active_cases.iloc[:10], y = 'Active Cases', x = 'State/UTs', linewidth = 3, edgecolor = 'black')
```




```
# Top 10 active cases states
top_10_active_cases = covid_df.groupby('State/UTs').max()['Active Cases', 'Discharged'].sort_values(by = ['Active Cases'], ascending = False).reset_index()
Fig = px.figure(figsize=(14,6))
plt.title('Top 10 states with most active cases in India', size = 20)
ax = sns.barplot(data = top_10_active_cases.iloc[:10], y = 'Active Cases', x = 'State/UTs', linewidth = 3, edgecolor = 'black')
plt.xlabel('States')
plt.ylabel('Total Active Cases')
plt.show()
```

```
# Top states with highest deaths
top_10_deaths = covid_df.groupby('State/UTs').max()['Deaths', 'Total Cases'].sort_values(by = ['Deaths'], ascending = False).reset_index()
Fig = px.figure(figsize=(14,6))
plt.title('Top 10 states with most Deaths', size = 20)
ax = sns.barplot(data = top_10_deaths.iloc[:10], y = 'Deaths', x='State/UTs', linewidth = 3, edgecolor = 'black')
plt.xlabel('States')
plt.ylabel('Total Death Cases')
plt.show()
```




```
# Top 10 discharged Cases states
top_10_discharged = covid_df.groupby('State/UTs').max()['Discharged', 'Active Cases'].sort_values(by = ['Discharged'], ascending = False).reset_index()
Fig = px.figure(figsize=(14,6))
plt.title('Top 10 states with most Discharged cases in India', size = 20)
ax = sns.barplot(data = top_10_discharged.iloc[:10], y = 'Discharged', x = 'State/UTs', linewidth = 3, edgecolor = 'black')
plt.xlabel('States')
plt.ylabel('Total Discharged Cases')
plt.show()
```

```
covid_df.head()
```

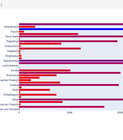
| | State/UTs | Total Vaccination Doses | Dose 1 | Dose 2 | Population |
|---|---|---|---|---|---|
| 0 | Andaman and Nicobar | 624504 | 311993 | 312511 | 380581 |
| 1 | Andhra Pradesh | 8414787 | 4918243 | 3496544 | 8175276 |
| 2 | Arunachal Pradesh | 1390148 | 850710 | 539438 | 1710387 |
| 3 | Assam | 11490104 | 6938462 | 4551642 | 30967791 |
| 4 | Bihar | 12480330 | 6938866 | 5541464 | 12456956 |

```
# Top 10 states with most Total Vaccination Doses
top_10_total = covid_df.groupby('State/UTs').max()['Total Vaccination Doses', 'Population'].sort_values(by = ['Total Vaccination Doses'], ascending = False).reset_index()
top_10_total
```

| | State/UTs | Total Vaccination Doses | Population |
|---|---|---|---|
| 0 | Uttar Pradesh | 48239700.0000 | 230400000.0000 |
| 1 | Maharashtra | 48239700.0000 | 230400000.0000 |
| 2 | Kerala | 42555611.0000 | 35699443.0000 |
| 3 | Karnataka | 34000000.0000 | 61095297.0000 |
| 4 | Odisha | 31387722.0000 | 31387722.0000 |
| 5 | Chhattisgarh | 31387722.0000 | 31387722.0000 |
| 6 | Punjab | 30967791.0000 | 30967791.0000 |
| 7 | Telangana | 30967791.0000 | 30967791.0000 |
| 8 | Andhra Pradesh | 8175276.0000 | 53000000.0000 |
| 9 | Andaman and Nicobar | 380581.0000 | 380581.0000 |

```
Fig = px.bar(data_frame=top_10_total,
             x='State/UTs',
             y='Dose 1',
             height=500,
             color='Dose 1')
Fig.show()
```

```
Fig = px.bar(data_frame=top_10_total,
             x='State/UTs',
             y='Dose 2',
             height=500,
             color='Dose 2')
Fig.show()
```




```
Fig = px.bar(data_frame=top_10_total,
             x='State/UTs',
             y='Total Vaccination Doses',
             height=500,
             color='Total Vaccination Doses')
Fig.show()
```




```
df_top = covid_df.sort_values(['Total Vaccination Doses', 'Dose 1', 'Dose 2'],
                              ascending = False).head(1)
```

```
Fig = px.barh(data_frame=df_top,
              x='Total Vaccination Doses', y='State/UTs', color='Total Vaccination Doses', orientation='h',
              color_continuous_scale='Blues',)[:, :, Dose1,Dose2]/[:,:]
```

```
Fig.show()
```