```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import datetime as dt
```

```
covid= pd.read_csv("C:\\Users\\neha bani\\Documents\\rohit details\\covid_19_india.csv")
covid
```

| | Sno | Date | Time | State/UnionTerritory | ConfirmedIndianNational | ConfirmedForei |
|---|---|---|---|---|---|---|
| **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 | |
| **...** | ... | ... | ... | ... | ... | |
| **18105** | 18106 | 2021-08-11 | 8:00 AM | Telangana | - | |
| **18106** | 18107 | 2021-08-11 | 8:00 AM | Tripura | - | |

```
covid.head()
```

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

```
covid=covid[['Date','State/UnionTerritory','Cured','Deaths','Confirmed']]
covid.columns=['date','state','cured','deaths','confirmed']
```

```
covid.head()
```

| | date | state | 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.tail()
```

| date | state | cured | deaths | confirmed |
|---|---|---|---|---|

```
covid.describe()
```

| | cured | deaths | confirmed |
|---|---|---|---|
| **count** | 1.811000e+04 | 18110.000000 | 1.811000e+04 |
| **mean** | 2.786375e+05 | 4052.402264 | 3.010314e+05 |
| **std** | 6.148909e+05 | 10919.076411 | 6.561489e+05 |
| **min** | 0.000000e+00 | 0.000000 | 0.000000e+00 |
| **25%** | 3.360250e+03 | 32.000000 | 4.376750e+03 |
| **50%** | 3.336400e+04 | 588.000000 | 3.977350e+04 |
| **75%** | 2.788698e+05 | 3643.750000 | 3.001498e+05 |
| **max** | 6.159676e+06 | 134201.000000 | 6.363442e+06 |

```
today=covid[covid.date=="2021-08-11"]
```

```
today.shape
```

```
(36, 5)
```

```
today.head()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **18074** | 2021-08-11 | Andaman and Nicobar Islands | 7412 | 129 | 7548 |
| **18075** | 2021-08-11 | Andhra Pradesh | 1952736 | 13564 | 1985182 |
| **18076** | 2021-08-11 | Arunachal Pradesh | 47821 | 248 | 50605 |
| **18077** | 2021-08-11 | Assam | 559684 | 5420 | 576149 |
| **18078** | 2021-08-11 | Bihar | 715352 | 9646 | 725279 |

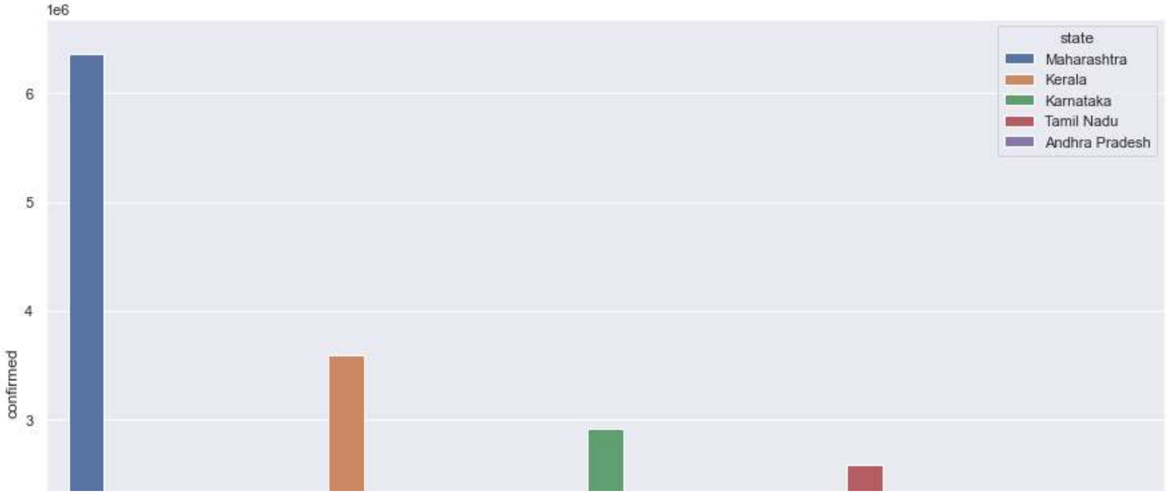
```
max_confirmed_cases=today.sort_values(by="confirmed",ascending=False)
```

```
max_confirmed_cases.head()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **18094** | 2021-08-11 | Maharashtra | 6159676 | 134201 | 6363442 |
| **18090** | 2021-08-11 | Kerala | 3396184 | 18004 | 3586693 |
| **18089** | 2021-08-11 | Karnataka | 2861499 | 36848 | 2921049 |
| **18104** | 2021-08-11 | Tamil Nadu | 2524400 | 34367 | 2579130 |
| **18075** | 2021-08-11 | Andhra Pradesh | 1952736 | 13564 | 1985182 |

```
top_states_confirmed=max_confirmed_cases[0:5]
```

```
sns.set(rc={'figure.figsize':(15,10)})
sns.barplot(x="state",y="confirmed",data=top_states_confirmed,hue='state')
plt.show()
```



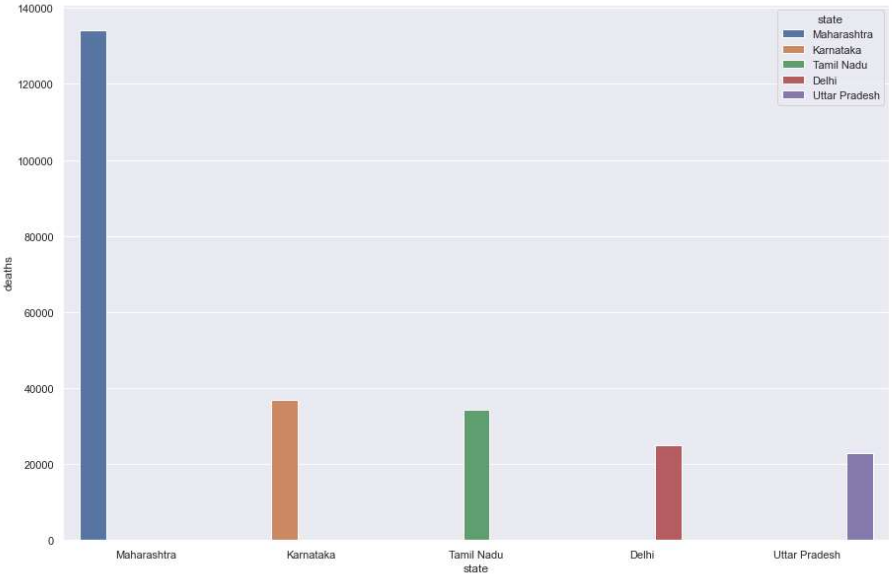
```
max_death_cases=today.sort_values(by='deaths',ascending=False)
```

```
max_death_cases.head()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **18094** | 2021-08-11 | Maharashtra | 6159676 | 134201 | 6363442 |
| **18089** | 2021-08-11 | Karnataka | 2861499 | 36848 | 2921049 |
| **18104** | 2021-08-11 | Tamil Nadu | 2524400 | 34367 | 2579130 |
| **18082** | 2021-08-11 | Delhi | 1411280 | 25068 | 1436852 |
| **18108** | 2021-08-11 | Uttar Pradesh | 1685492 | 22775 | 1708812 |

```
top_states_death=max_death_cases[0:5]
```

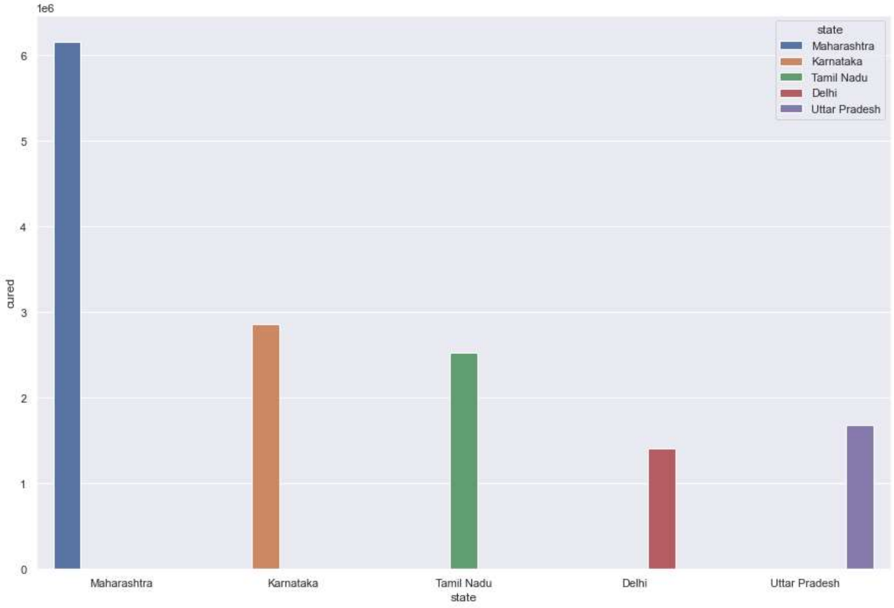
```
sns.set(rc={'figure.figsize':(15,10)})
sns.barplot(x="state",y="deaths",data=top_states_death,hue='state')
plt.show()
```



```
max_cured_cases=today.sort_values(by='deaths',ascending=False)
```

```
top_states_cured=max_cured_cases[0:5]
```

```
sns.set(rc={'figure.figsize':(15,10)})
sns.barplot(x="state",y="cured",data=top_states_death,hue='state')
plt.show()
```



```
#Maharashtra
```

```
maha=covid[covid.state=="Maharashtra"]
```

```
maha.head()
```

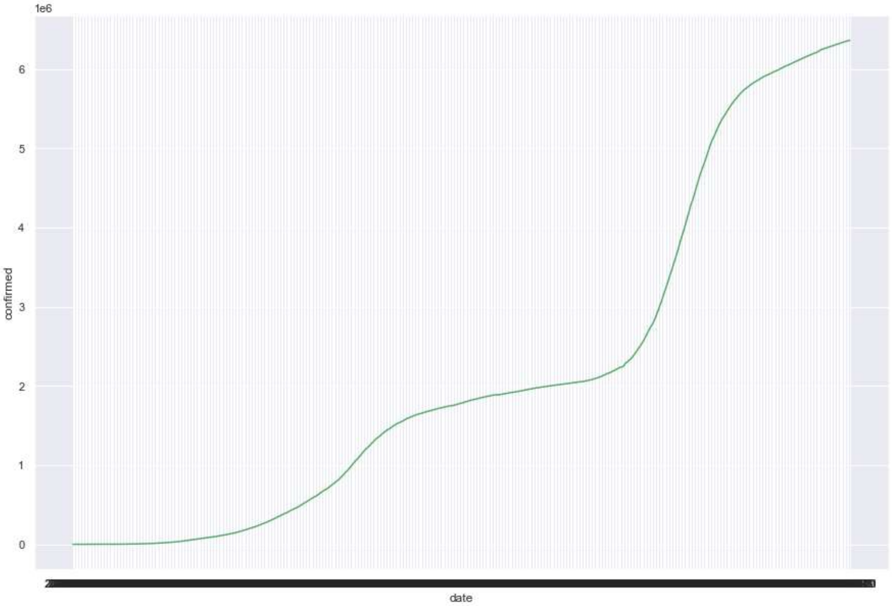
| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **76** | 2020-03-09 | Maharashtra | 0 | 0 | 2 |
| **91** | 2020-03-10 | Maharashtra | 0 | 0 | 5 |
| **97** | 2020-03-11 | Maharashtra | 0 | 0 | 2 |
| **120** | 2020-03-12 | Maharashtra | 0 | 0 | 11 |
| **133** | 2020-03-13 | Maharashtra | 0 | 0 | 14 |

```
maha.tail()
```

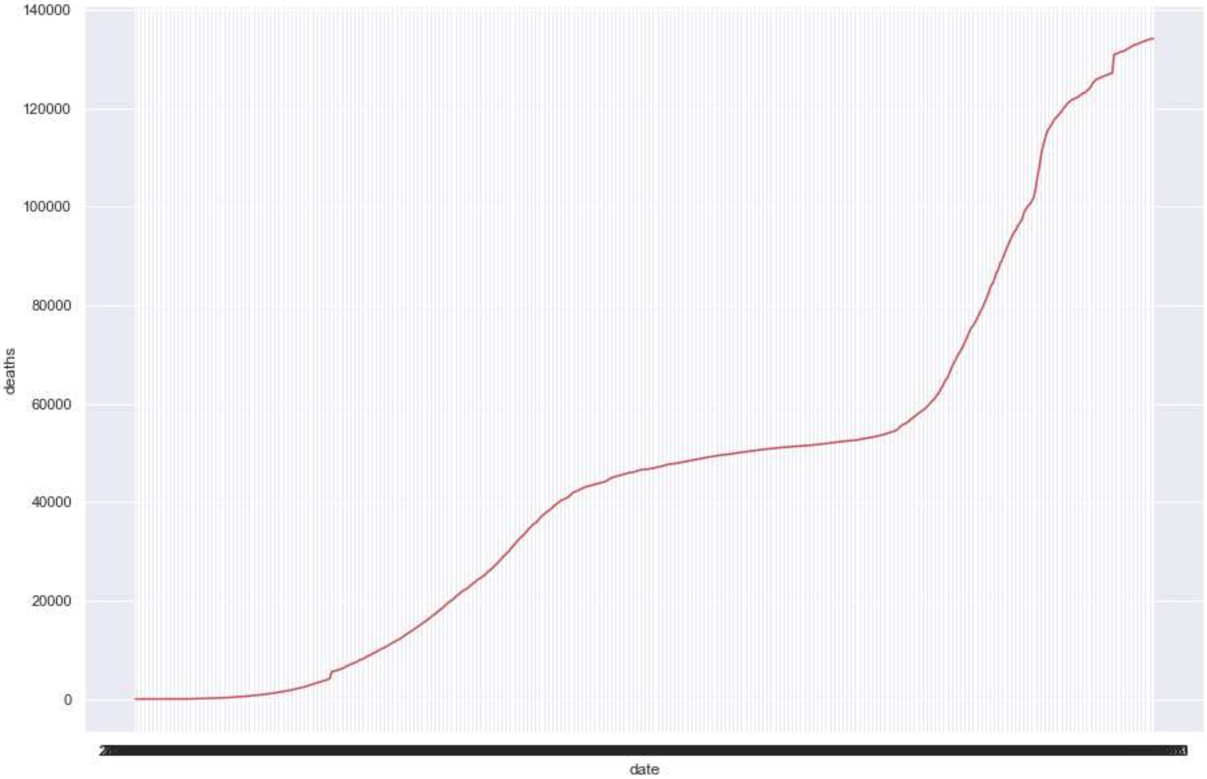
| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **17950** | 2021-08-07 | Maharashtra | 6130137 | 133717 | 6341759 |
| **17986** | 2021-08-08 | Maharashtra | 6139493 | 133845 | 6347820 |
| **18022** | 2021-08-09 | Maharashtra | 6144388 | 133996 | 6353328 |
| **18058** | 2021-08-10 | Maharashtra | 6151956 | 134064 | 6357833 |
| **18094** | 2021-08-11 | Maharashtra | 6159676 | 134201 | 6363442 |

```
sns.set(rc={'figure.figsize':(15,10)})
sns.lineplot(x="date",y="confirmed",data=maha,color='g')
```

```
plt.show()
```



```
sns.set(rc={'figure.figsize':(15,10)})
sns.lineplot(x="date",y="deaths",data=maha,color='r')
plt.show()
```



```
#kerla
```

```
kerala=covid[covid.state=="Kerala"]
```

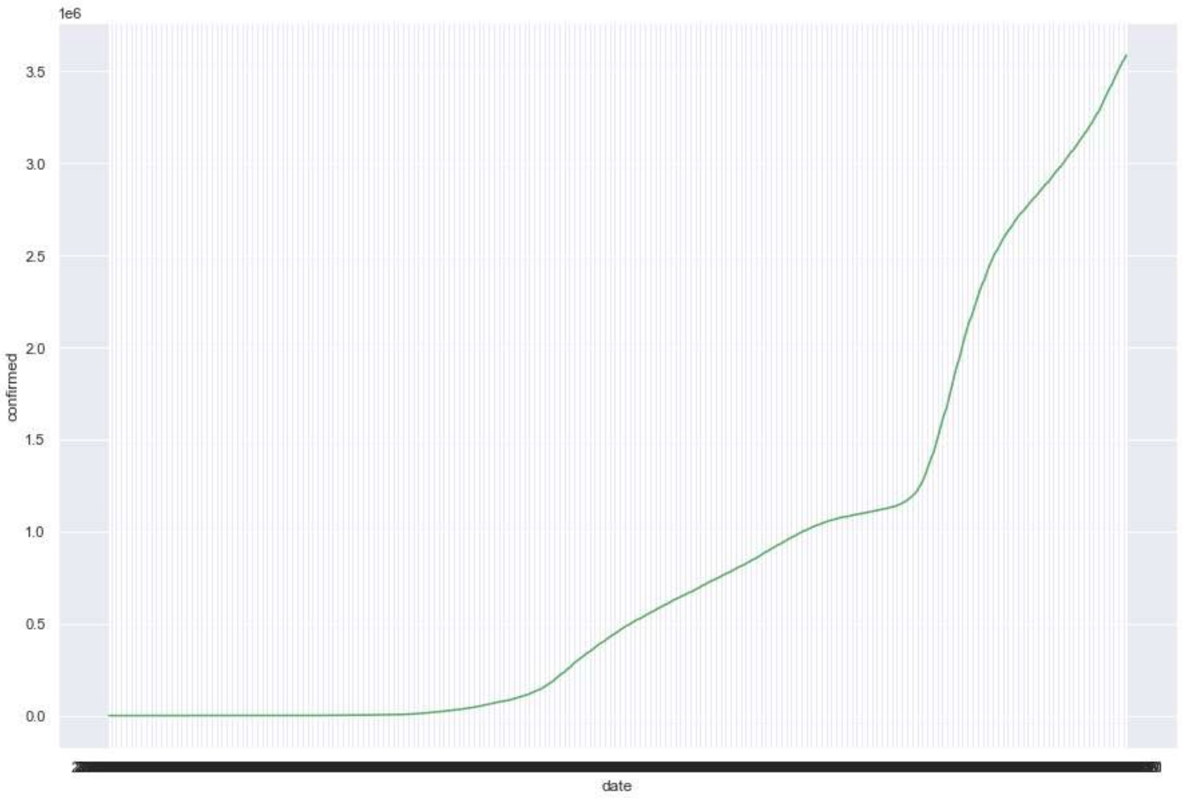
```
kerala.head()
```

| | date | state | 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 |

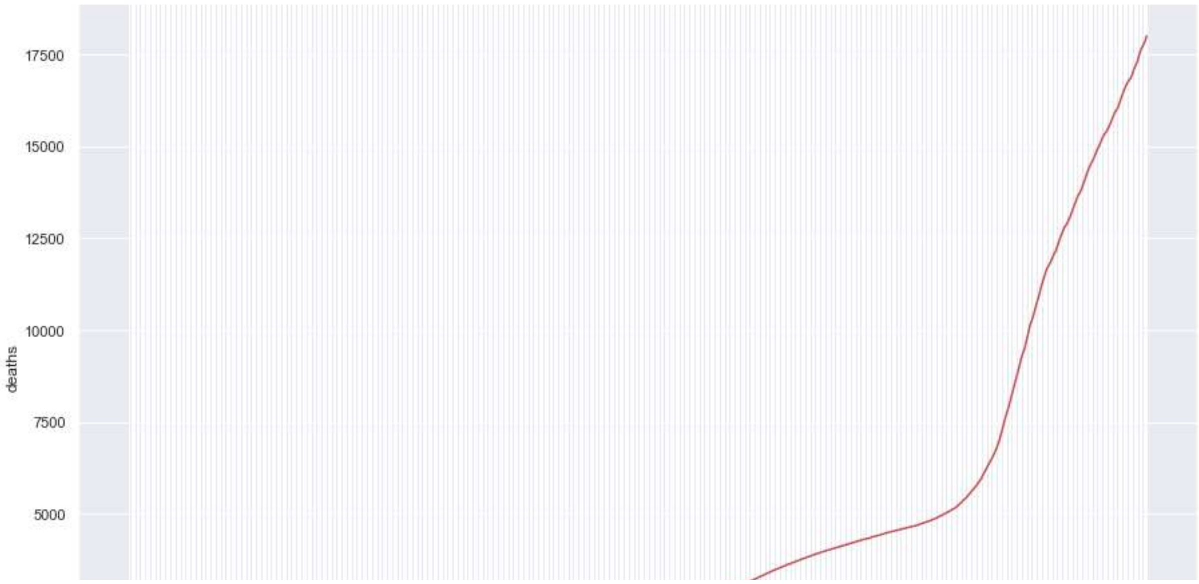
```
kerala.tail()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **17946** | 2021-08-07 | Kerala | 3317314 | 17515 | 3513551 |
| **17982** | 2021-08-08 | Kerala | 3337579 | 17654 | 3533918 |
| **18018** | 2021-08-09 | Kerala | 3357687 | 17747 | 3552525 |
| **18054** | 2021-08-10 | Kerala | 3377691 | 17852 | 3565574 |
| **18090** | 2021-08-11 | Kerala | 3396184 | 18004 | 3586693 |

```
sns.set(rc={'figure.figsize':(15,10)})
sns.lineplot(x="date",y="confirmed",data=kerala,color='g')
plt.show()
```



```
sns.set(rc={'figure.figsize':(15,10)})
sns.lineplot(x="date",y="deaths",data=kerala,color='r')
plt.show()
```



```
#jk
```

```
jk=covid[covid.state=="Jammu and Kashmir"]
```

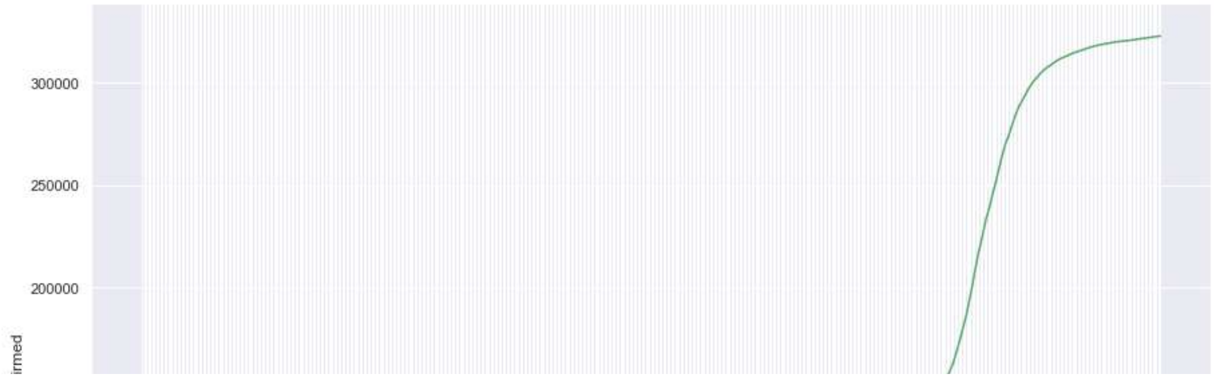
```
jk.head()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **81** | 2020-03-09 | Jammu and Kashmir | 0 | 0 | 1 |
| **96** | 2020-03-10 | Jammu and Kashmir | 0 | 0 | 1 |
| **106** | 2020-03-11 | Jammu and Kashmir | 0 | 0 | 1 |
| **117** | 2020-03-12 | Jammu and Kashmir | 0 | 0 | 1 |
| **130** | 2020-03-13 | Jammu and Kashmir | 0 | 0 | 1 |

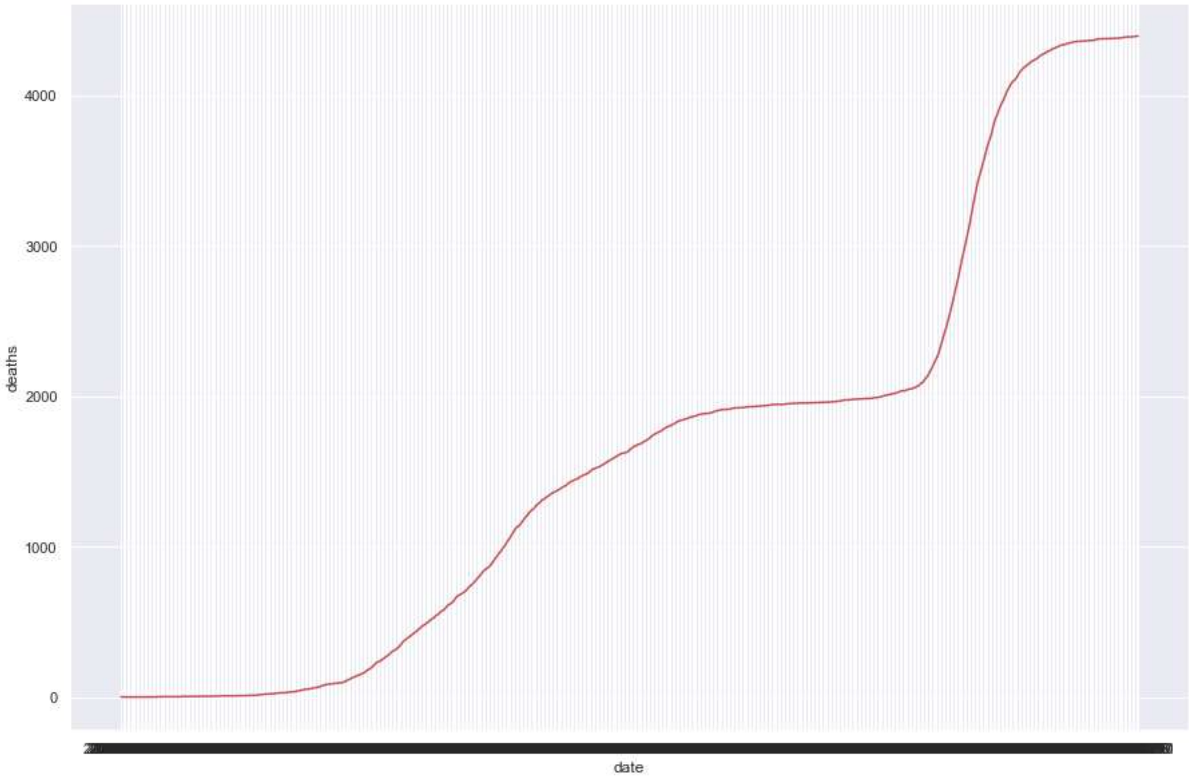
```
jk.tail()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **17943** | 2021-08-07 | Jammu and Kashmir | 316496 | 4386 | 322286 |
| **17979** | 2021-08-08 | Jammu and Kashmir | 316632 | 4386 | 322428 |
| **18015** | 2021-08-09 | Jammu and Kashmir | 316761 | 4389 | 322550 |
| **18051** | 2021-08-10 | Jammu and Kashmir | 316957 | 4390 | 322658 |
| **18087** | 2021-08-11 | Jammu and Kashmir | 317081 | 4392 | 322771 |

```
sns.set(rc={'figure.figsize':(15,10)})
sns.lineplot(x="date",y="confirmed",data=jk,color='g')
plt.show()
```



```
sns.set(rc={'figure.figsize':(15,10)})
sns.lineplot(x="date",y="deaths",data=jk,color='r')
plt.show()
```



```
#tests
```

```
tests=pd.read_csv("C:\\Users\\neha bani\\Documents\\rohit details\\StatewiseTestingDetails.csv")
```

```
tests.head()
```

| | Date | State | TotalSamples | Negative | Positive |
|---|---|---|---|---|---|
| **0** | 2020-04-17 | Andaman and Nicobar Islands | 1403.0 | 1210 | 12.0 |
| **1** | 2020-04-24 | Andaman and Nicobar Islands | 2679.0 | NaN | 27.0 |
| **2** | 2020-04-27 | Andaman and Nicobar Islands | 2848.0 | NaN | 33.0 |
| **3** | 2020-05-01 | Andaman and Nicobar Islands | 3754.0 | NaN | 33.0 |
| **4** | 2020-05-16 | Andaman and Nicobar Islands | 6677.0 | NaN | 33.0 |

```
tests.tail()
```

| | Date | State | TotalSamples | Negative | Positive |
|---|---|---|---|---|---|
| **16331** | 2021-08-06 | West Bengal | 15999961.0 | NaN | NaN |
| **16332** | 2021-08-07 | West Bengal | 16045662.0 | NaN | NaN |
| **16333** | 2021-08-08 | West Bengal | 16092192.0 | NaN | NaN |
| **16335** | 2021-08-10 | West Bengal | 16162814.0 | NaN | NaN |

```
from sklearn.model_selection import train_test_split
```

```
maha['date']=maha['date'].astype('datetime64[ns]')
```

```
<ipython-input-97-23b841640f05>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-c
  maha['date']=maha['date'].astype('datetime64[ns]')
```

```
maha['date']=maha['date'].map(dt.datetime.toordinal)
```

```
<ipython-input-98-595f711d6bae>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-c
  maha['date']=maha['date'].map(dt.datetime.toordinal)
```

```
maha.head()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **76** | 719163 | Maharashtra | 0 | 0 | 2 |
| **91** | 719163 | Maharashtra | 0 | 0 | 5 |
| **97** | 719163 | Maharashtra | 0 | 0 | 2 |
| **120** | 719163 | Maharashtra | 0 | 0 | 11 |
| **133** | 719163 | Maharashtra | 0 | 0 | 14 |

```
x=maha['date']
y=maha['confirmed']
```

```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3)
```

```
from sklearn.linear_model import LinearRegression
```

```
lr =LinearRegression()
```

```
y_train
```

```
2649       74860
13738    3407245
12982    2479682
12298    2169330
17374    6237755
          ...
6965     1430861
1592       12974
6895     1400922
14710    5053336
5250      584754
Name: confirmed, Length: 364, dtype: int64
```

```
lr.fit(np.array(x_train).reshape(-1,1),np.array(y_train).reshape(-1,1))
```

```
LinearRegression()
```

```
maha.tail()
```

| | date | state | cured | deaths | confirmed |
|---|---|---|---|---|---|
| **17950** | 719163 | Maharashtra | 6130137 | 133717 | 6341759 |
| **17986** | 719163 | Maharashtra | 6139493 | 133845 | 6347820 |
| **18022** | 719163 | Maharashtra | 6144388 | 133996 | 6353328 |
| **18058** | 719163 | Maharashtra | 6151956 | 134064 | 6357833 |
| **18094** | 719163 | Maharashtra | 6159676 | 134201 | 6363442 |

```
lr.predict(np.array([[719170]]))
```

```
array([[2180410.50824176]])
```