

In [127]:

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

In [128]:

```
#importing main data set
df =pd.read_csv('covid_19_india.csv',parse_dates=['Date'],dayfirst=True)
```

In [129]:

```
df.head()
```

Out[129]:

	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

Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignNational	Cured	Deaths	Confirmed
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

In [130]:

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

In [131]:

```
df.head()
```

Out[131]:

	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

In [132]:

```
df.tail()
```

Out[132]:

	date	state	cured	deaths	confirmed
18105	2021-08-11	Telangana	638410	3831	650353
18106	2021-08-11	Tripura	77811	773	80660
18107	2021-08-11	Uttarakhand	334650	7368	342462
18108	2021-08-11	Uttar Pradesh	1685492	22775	1708812
18109	2021-08-11	West Bengal	1506532	18252	1534999

In [133]:

```
#curred date  
today =df[df.date == '2020-07-17']
```

In [134]:

```
today.head()
```

Out[134]:

	date	state	cured	deaths	confirmed
4179	2020-07-17	Andaman and Nicobar Islands	133	0	180

	date	state	cured	deaths	confirmed
4180	2020-07-17	Andhra Pradesh	19393	492	38044
4181	2020-07-17	Arunachal Pradesh	153	3	543
4182	2020-07-17	Assam	12888	48	19754
4183	2020-07-17	Bihar	14018	197	21764

In [135]:

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

In [136]:

```
max_confirmed_cases
```

Out[136]:

	date	state	cured	deaths	confirmed
4198	2020-07-17	Maharashtra	158140	11194	284281

	date	state	cured	deaths	confirmed
4208	2020-07-17	Tamil Nadu	107416	2236	156369
4187	2020-07-17	Delhi	97693	3545	118645
4194	2020-07-17	Karnataka	19729	1032	51422
4189	2020-07-17	Gujarat	32103	2089	45481
4212	2020-07-17	Uttar Pradesh	26675	1046	43441
4209	2020-07-17	Telengana	27295	396	41018
4180	2020-07-17	Andhra Pradesh	19393	492	38044
4213	2020-07-17	West Bengal	21415	1023	36117

	date	state	cured	deaths	confirmed
4206	2020-07-17	Rajasthan	19970	538	27174
4190	2020-07-17	Haryana	18185	322	24002
4183	2020-07-17	Bihar	14018	197	21764
4197	2020-07-17	Madhya Pradesh	14127	689	20378
4182	2020-07-17	Assam	12888	48	19754
4203	2020-07-17	Odisha	10877	79	15392
4192	2020-07-17	Jammu and Kashmir	6446	222	12156
4195	2020-07-17	Kerala	4862	37	10275

	date	state	cured	deaths	confirmed
4205	2020-07-17	Punjab	6277	230	9094
4185	2020-07-17	Chhattisgarh	3451	21	4732
4193	2020-07-17	Jharkhand	2513	42	4624
4211	2020-07-17	Uttarakhand	2995	50	3982
4188	2020-07-17	Goa	1817	19	3108
4210	2020-07-17	Tripura	1604	3	2283
4199	2020-07-17	Manipur	1129	0	1764
4204	2020-07-17	Puducherry	947	22	1743

	date	state	cured	deaths	confirmed
4191	2020-07-17	Himachal Pradesh	984	11	1377
4196	2020-07-17	Ladakh	970	1	1147
4202	2020-07-17	Nagaland	391	0	916
4184	2020-07-17	Chandigarh	476	11	651
4186	2020-07-17	Dadra and Nagar Haveli and Daman and Diu	371	2	552
4181	2020-07-17	Arunachal Pradesh	153	3	543
4214	2020-07-17	Cases being reassigned to states	0	0	531
4200	2020-07-17	Meghalaya	66	2	377

	date	state	cured	deaths	confirmed
4201	2020-07-17	Mizoram	160	0	272
4207	2020-07-17	Sikkim	88	0	243
4179	2020-07-17	Andaman and Nicobar Islands	133	0	180

In [137]:

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

Out[137]:

	date	state	cured	deaths	confirmed
4198	2020-07-17	Maharashtra	158140	11194	284281
4208	2020-07-17	Tamil Nadu	107416	2236	156369
4187	2020-07-17	Delhi	97693	3545	118645

	date	state	cured	deaths	confirmed
4194	2020-07-17	Karnataka	19729	1032	51422
4189	2020-07-17	Gujarat	32103	2089	45481
4212	2020-07-17	Uttar Pradesh	26675	1046	43441
4209	2020-07-17	Telengana	27295	396	41018
4180	2020-07-17	Andhra Pradesh	19393	492	38044
4213	2020-07-17	West Bengal	21415	1023	36117
4206	2020-07-17	Rajasthan	19970	538	27174
4190	2020-07-17	Haryana	18185	322	24002

	date	state	cured	deaths	confirmed
4183	2020-07-17	Bihar	14018	197	21764
4197	2020-07-17	Madhya Pradesh	14127	689	20378
4182	2020-07-17	Assam	12888	48	19754
4203	2020-07-17	Odisha	10877	79	15392
4192	2020-07-17	Jammu and Kashmir	6446	222	12156
4195	2020-07-17	Kerala	4862	37	10275
4205	2020-07-17	Punjab	6277	230	9094
4185	2020-07-17	Chhattisgarh	3451	21	4732

	date	state	cured	deaths	confirmed
4193	2020-07-17	Jharkhand	2513	42	4624
4211	2020-07-17	Uttarakhand	2995	50	3982
4188	2020-07-17	Goa	1817	19	3108
4210	2020-07-17	Tripura	1604	3	2283
4199	2020-07-17	Manipur	1129	0	1764
4204	2020-07-17	Puducherry	947	22	1743
4191	2020-07-17	Himachal Pradesh	984	11	1377
4196	2020-07-17	Ladakh	970	1	1147

	date	state	cured	deaths	confirmed
4202	2020-07-17	Nagaland	391	0	916
4184	2020-07-17	Chandigarh	476	11	651
4186	2020-07-17	Dadra and Nagar Haveli and Daman and Diu	371	2	552
4181	2020-07-17	Arunachal Pradesh	153	3	543
4214	2020-07-17	Cases being reassigned to states	0	0	531
4200	2020-07-17	Meghalaya	66	2	377
4201	2020-07-17	Mizoram	160	0	272
4207	2020-07-17	Sikkim	88	0	243

	date	state	cured	deaths	confirmed
4179	2020-07-17	Andaman and Nicobar Islands	133	0	180

In [138]:

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

Out[138]:

	date	state	cured	deaths	confirmed
4198	2020-07-17	Maharashtra	158140	11194	284281
4208	2020-07-17	Tamil Nadu	107416	2236	156369
4187	2020-07-17	Delhi	97693	3545	118645
4194	2020-07-17	Karnataka	19729	1032	51422
4189	2020-07-17	Gujarat	32103	2089	45481

	date	state	cured	deaths	confirmed
4212	2020-07-17	Uttar Pradesh	26675	1046	43441
4209	2020-07-17	Telengana	27295	396	41018
4180	2020-07-17	Andhra Pradesh	19393	492	38044
4213	2020-07-17	West Bengal	21415	1023	36117
4206	2020-07-17	Rajasthan	19970	538	27174
4190	2020-07-17	Haryana	18185	322	24002
4183	2020-07-17	Bihar	14018	197	21764
4197	2020-07-17	Madhya Pradesh	14127	689	20378

	date	state	cured	deaths	confirmed
4182	2020-07-17	Assam	12888	48	19754
4203	2020-07-17	Odisha	10877	79	15392
4192	2020-07-17	Jammu and Kashmir	6446	222	12156
4195	2020-07-17	Kerala	4862	37	10275
4205	2020-07-17	Punjab	6277	230	9094
4185	2020-07-17	Chhattisgarh	3451	21	4732
4193	2020-07-17	Jharkhand	2513	42	4624
4211	2020-07-17	Uttarakhand	2995	50	3982

	date	state	cured	deaths	confirmed
4188	2020-07-17	Goa	1817	19	3108
4210	2020-07-17	Tripura	1604	3	2283
4199	2020-07-17	Manipur	1129	0	1764
4204	2020-07-17	Puducherry	947	22	1743
4191	2020-07-17	Himachal Pradesh	984	11	1377
4196	2020-07-17	Ladakh	970	1	1147
4202	2020-07-17	Nagaland	391	0	916
4184	2020-07-17	Chandigarh	476	11	651

	date	state	cured	deaths	confirmed
4186	2020-07-17	Dadra and Nagar Haveli and Daman and Diu	371	2	552
4181	2020-07-17	Arunachal Pradesh	153	3	543
4214	2020-07-17	Cases being reassigned to states	0	0	531
4200	2020-07-17	Meghalaya	66	2	377
4201	2020-07-17	Mizoram	160	0	272
4207	2020-07-17	Sikkim	88	0	243
4179	2020-07-17	Andaman and Nicobar Islands	133	0	180

In [139]:

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

In [140]:

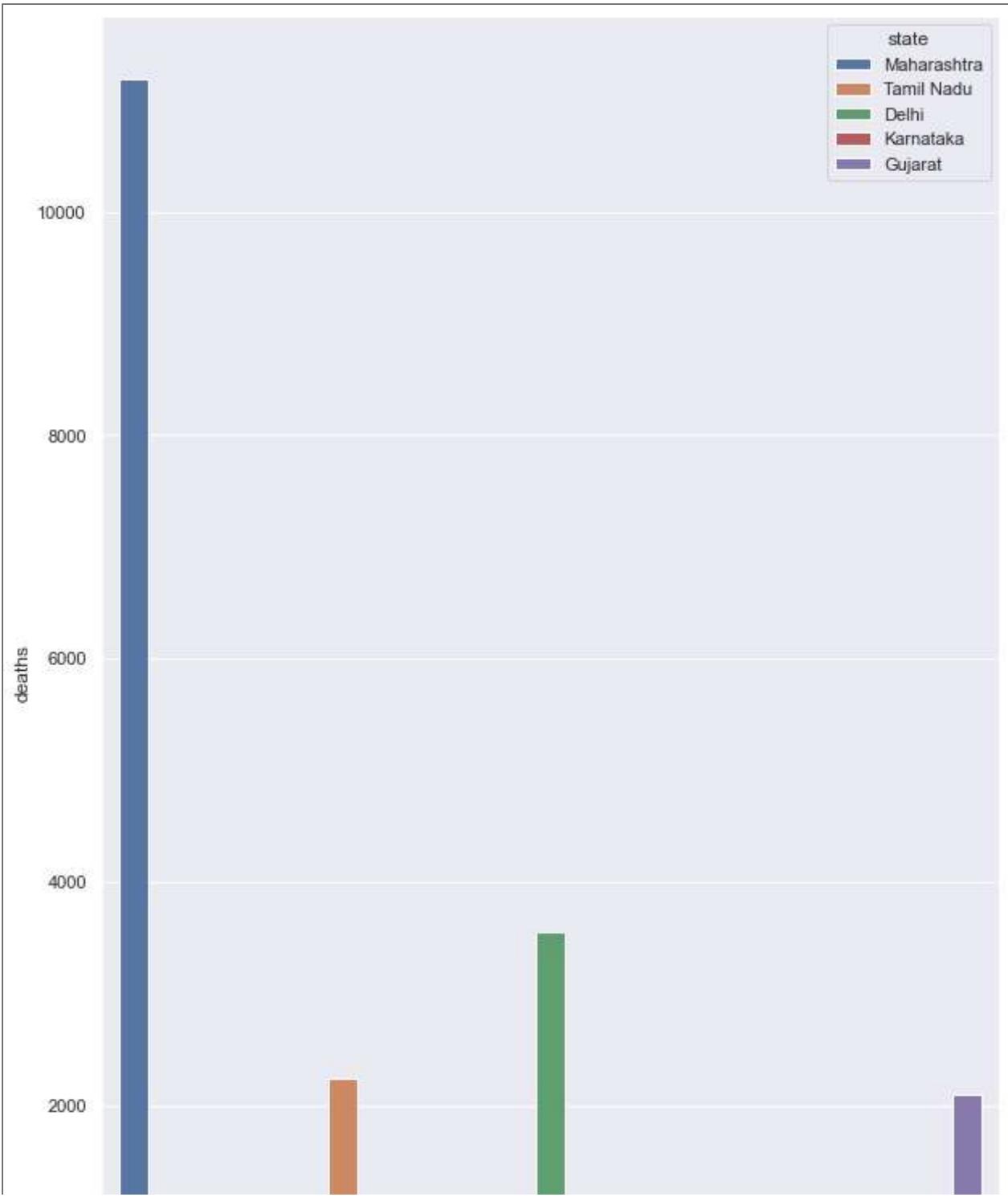
```
top_states_confirmed
```

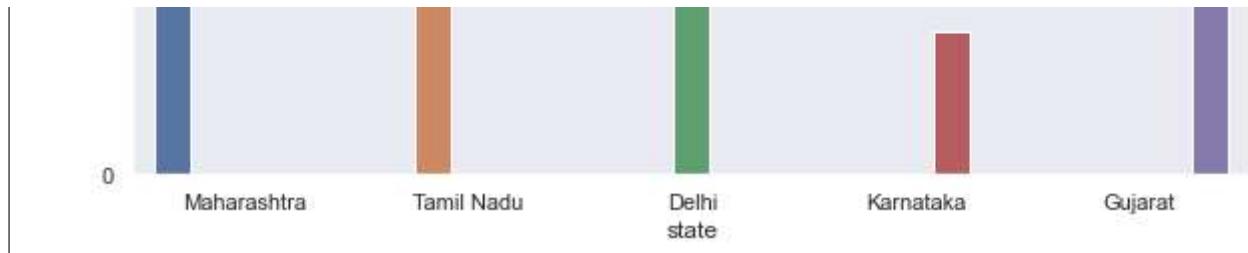
Out[140]:

	date	state	cured	deaths	confirmed
4198	2020-07-17	Maharashtra	158140	11194	284281
4208	2020-07-17	Tamil Nadu	107416	2236	156369
4187	2020-07-17	Delhi	97693	3545	118645
4194	2020-07-17	Karnataka	19729	1032	51422
4189	2020-07-17	Gujarat	32103	2089	45481

In [141]:

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





In [142]:

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

Out[142]:

	date	state	cured	deaths	confirmed
4198	2020-07-17	Maharashtra	158140	11194	284281
4187	2020-07-17	Delhi	97693	3545	118645
4208	2020-07-17	Tamil Nadu	107416	2236	156369
4189	2020-07-17	Gujarat	32103	2089	45481
4212	2020-07-17	Uttar Pradesh	26675	1046	43441

	date	state	cured	deaths	confirmed
4194	2020-07-17	Karnataka	19729	1032	51422
4213	2020-07-17	West Bengal	21415	1023	36117
4197	2020-07-17	Madhya Pradesh	14127	689	20378
4206	2020-07-17	Rajasthan	19970	538	27174
4180	2020-07-17	Andhra Pradesh	19393	492	38044
4209	2020-07-17	Telengana	27295	396	41018
4190	2020-07-17	Haryana	18185	322	24002
4205	2020-07-17	Punjab	6277	230	9094

	date	state	cured	deaths	confirmed
4192	2020-07-17	Jammu and Kashmir	6446	222	12156
4183	2020-07-17	Bihar	14018	197	21764
4203	2020-07-17	Odisha	10877	79	15392
4211	2020-07-17	Uttarakhand	2995	50	3982
4182	2020-07-17	Assam	12888	48	19754
4193	2020-07-17	Jharkhand	2513	42	4624
4195	2020-07-17	Kerala	4862	37	10275
4204	2020-07-17	Puducherry	947	22	1743

	date	state	cured	deaths	confirmed
4185	2020-07-17	Chhattisgarh	3451	21	4732
4188	2020-07-17	Goa	1817	19	3108
4191	2020-07-17	Himachal Pradesh	984	11	1377
4184	2020-07-17	Chandigarh	476	11	651
4181	2020-07-17	Arunachal Pradesh	153	3	543
4210	2020-07-17	Tripura	1604	3	2283
4200	2020-07-17	Meghalaya	66	2	377
4186	2020-07-17	Dadra and Nagar Haveli and Daman and Diu	371	2	552

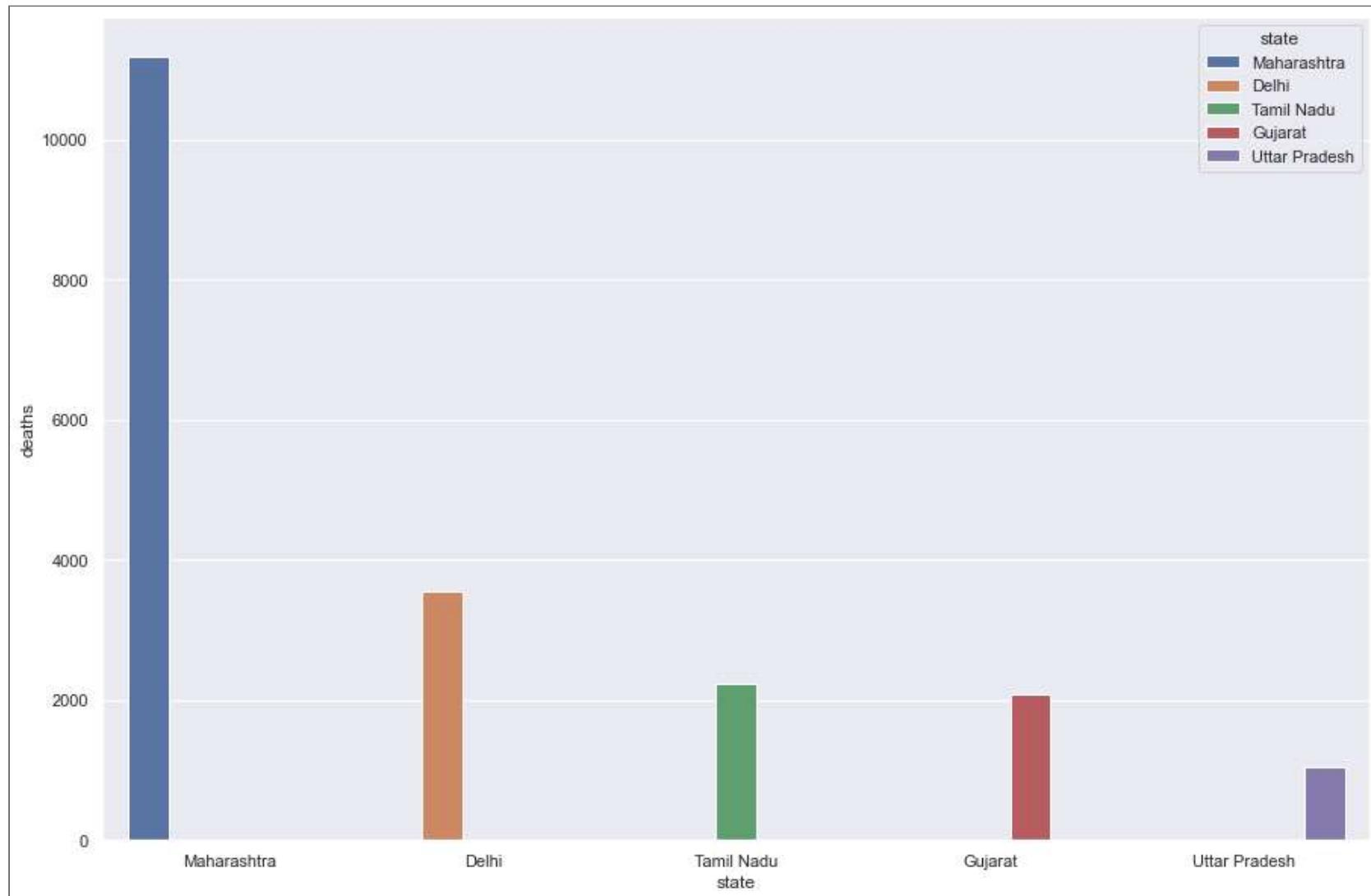
	date	state	cured	deaths	confirmed
4196	2020-07-17	Ladakh	970	1	1147
4179	2020-07-17	Andaman and Nicobar Islands	133	0	180
4207	2020-07-17	Sikkim	88	0	243
4202	2020-07-17	Nagaland	391	0	916
4201	2020-07-17	Mizoram	160	0	272
4199	2020-07-17	Manipur	1129	0	1764
4214	2020-07-17	Cases being reassigned to states	0	0	531

In [143]:

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

In [144]:

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



In [145]:

```
max_cured_cases=today.sort_values(by="cured",ascending=False)
max_cured_cases
```

Out[145]:

	date	state	cured	deaths	confirmed
4198	2020-07-17	Maharashtra	158140	11194	284281
4208	2020-07-17	Tamil Nadu	107416	2236	156369
4187	2020-07-17	Delhi	97693	3545	118645
4189	2020-07-17	Gujarat	32103	2089	45481
4209	2020-07-17	Telengana	27295	396	41018
4212	2020-07-17	Uttar Pradesh	26675	1046	43441
4213	2020-07-17	West Bengal	21415	1023	36117

	date	state	cured	deaths	confirmed
4206	2020-07-17	Rajasthan	19970	538	27174
4194	2020-07-17	Karnataka	19729	1032	51422
4180	2020-07-17	Andhra Pradesh	19393	492	38044
4190	2020-07-17	Haryana	18185	322	24002
4197	2020-07-17	Madhya Pradesh	14127	689	20378
4183	2020-07-17	Bihar	14018	197	21764
4182	2020-07-17	Assam	12888	48	19754
4203	2020-07-17	Odisha	10877	79	15392

	date	state	cured	deaths	confirmed
4192	2020-07-17	Jammu and Kashmir	6446	222	12156
4205	2020-07-17	Punjab	6277	230	9094
4195	2020-07-17	Kerala	4862	37	10275
4185	2020-07-17	Chhattisgarh	3451	21	4732
4211	2020-07-17	Uttarakhand	2995	50	3982
4193	2020-07-17	Jharkhand	2513	42	4624
4188	2020-07-17	Goa	1817	19	3108
4210	2020-07-17	Tripura	1604	3	2283

	date	state	cured	deaths	confirmed
4199	2020-07-17	Manipur	1129	0	1764
4191	2020-07-17	Himachal Pradesh	984	11	1377
4196	2020-07-17	Ladakh	970	1	1147
4204	2020-07-17	Puducherry	947	22	1743
4184	2020-07-17	Chandigarh	476	11	651
4202	2020-07-17	Nagaland	391	0	916
4186	2020-07-17	Dadra and Nagar Haveli and Daman and Diu	371	2	552
4201	2020-07-17	Mizoram	160	0	272

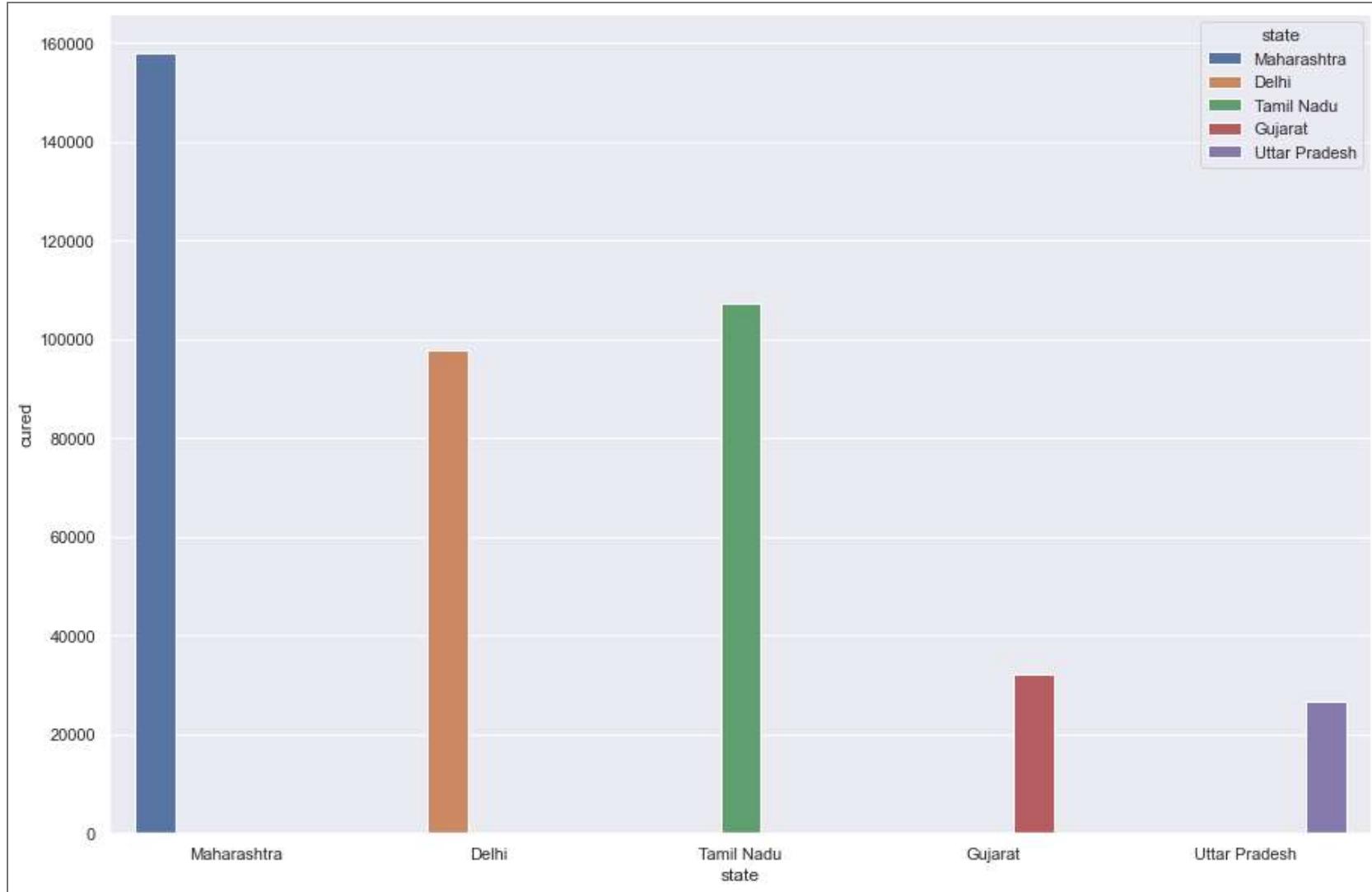
	date	state	cured	deaths	confirmed
4181	2020-07-17	Arunachal Pradesh	153	3	543
4179	2020-07-17	Andaman and Nicobar Islands	133	0	180
4207	2020-07-17	Sikkim	88	0	243
4200	2020-07-17	Meghalaya	66	2	377
4214	2020-07-17	Cases being reassigned to states	0	0	531

In [146]:

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

In [147]:

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



In [148]:

```
mah =df[df.state == 'Maharashtra']
```

In [149]:

```
mah
```

Out[149]:

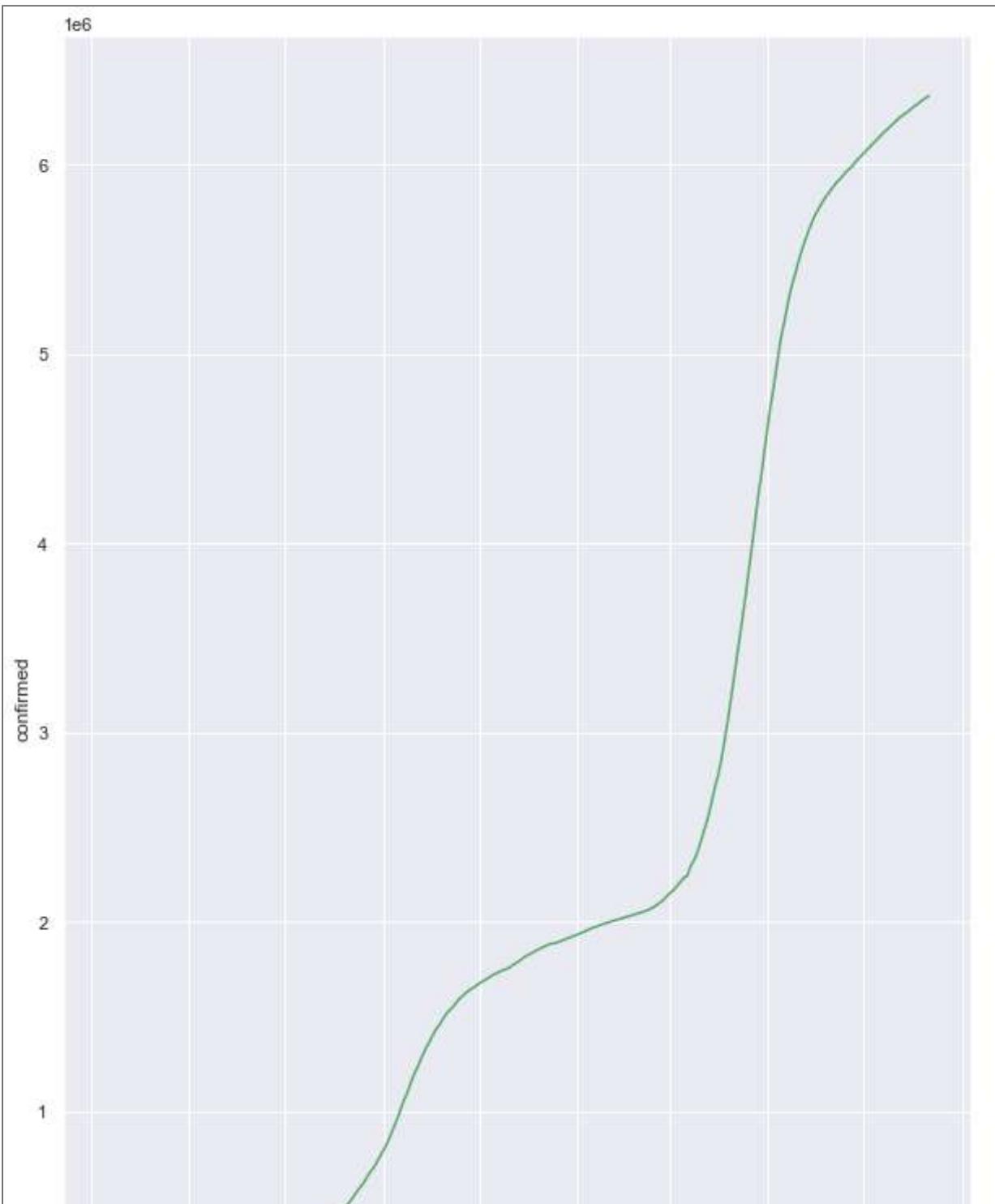
	date	state	cured	deaths	confirmed
76	2020-03-09	Maharashtra	0	0	2
91	2020-03-10	Maharashtra	0	0	5
97	2020-03-11	Maharashtra	0	0	2
120	2020-03-12	Maharashtra	0	0	11
133	2020-03-13	Maharashtra	0	0	14
...					
17950	2021-08-07	Maharashtra	6130137	133717	6341759
17986	2021-08-08	Maharashtra	6139493	133845	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

520 rows × 5 columns

In [150]:

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

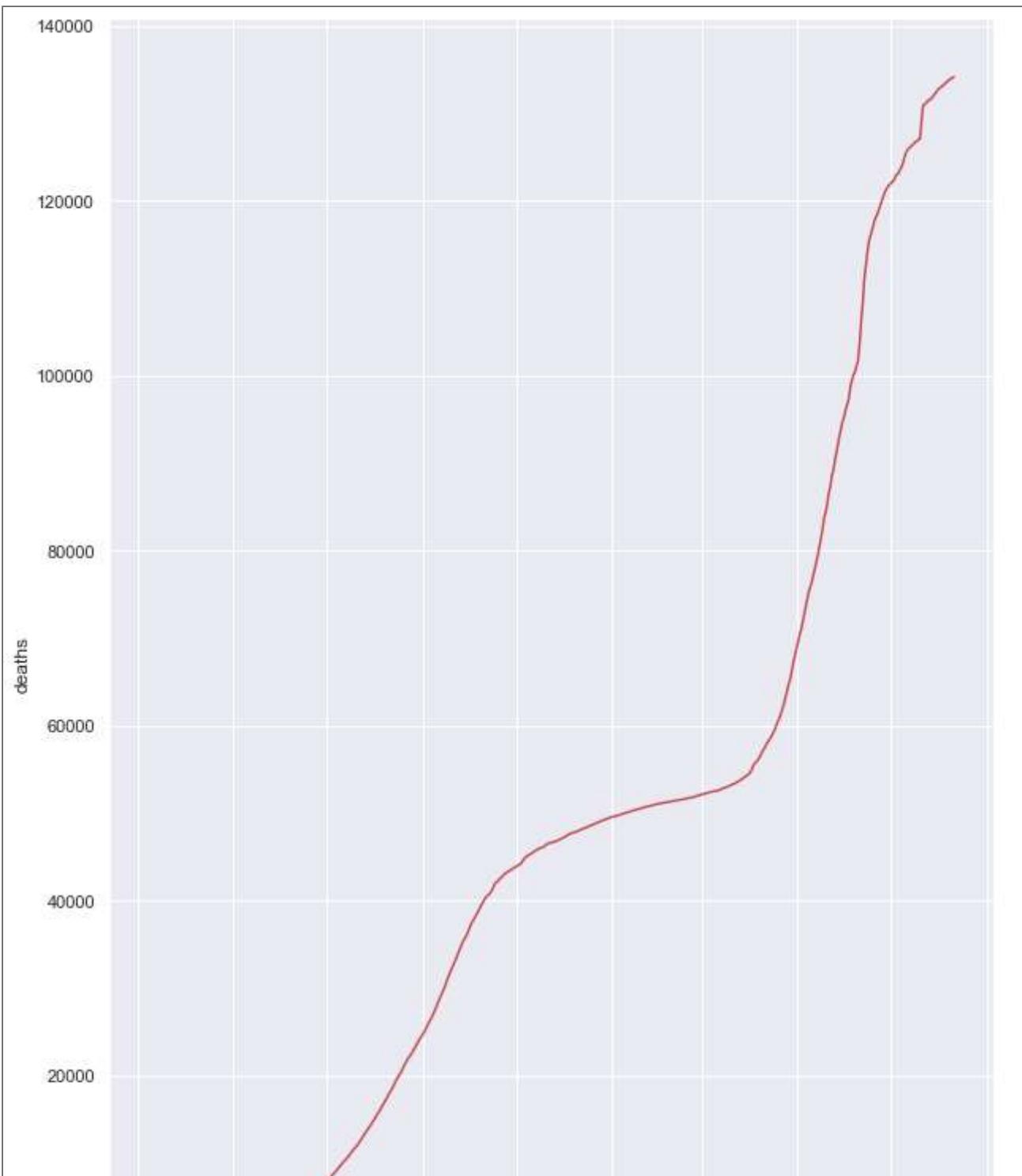
```
| plt.show()
```





In [151]:

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





In [152]:

```
Kerala=df[df.state =='Kerala']  
Kerala
```

Out[152]:

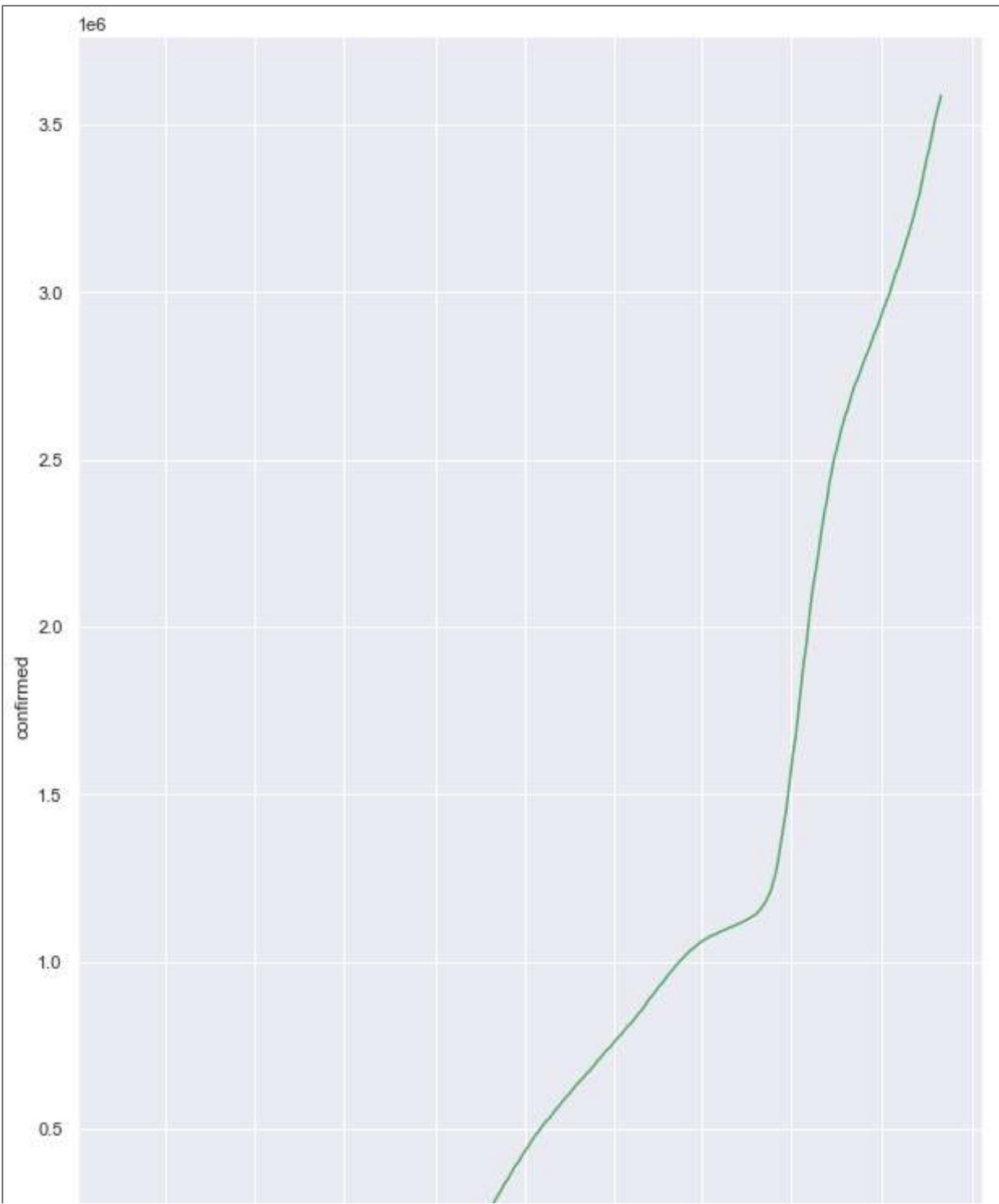
	date	state	cured	deaths	confirmed
0	2020-01-30	Kerala	0	0	1
1	2020-01-31	Kerala	0	0	1
2	2020-02-01	Kerala	0	0	2
3	2020-02-02	Kerala	0	0	3
4	2020-02-03	Kerala	0	0	3
...
17946	2021-08-07	Kerala	3317314	17515	3513551
17982	2021-08-08	Kerala	3337579	17654	3533918
18018	2021-08-09	Kerala	3357687	17747	3552525

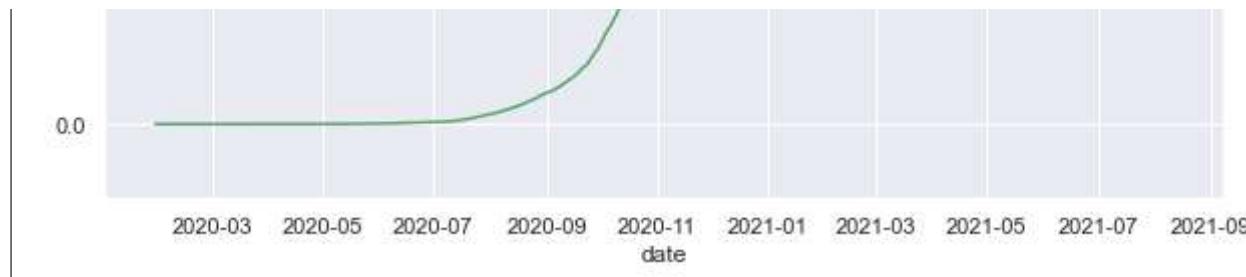
	date	state	cured	deaths	confirmed
18054	2021-08-10	Kerala	3377691	17852	3565574
18090	2021-08-11	Kerala	3396184	18004	3586693

560 rows × 5 columns

In [153]:

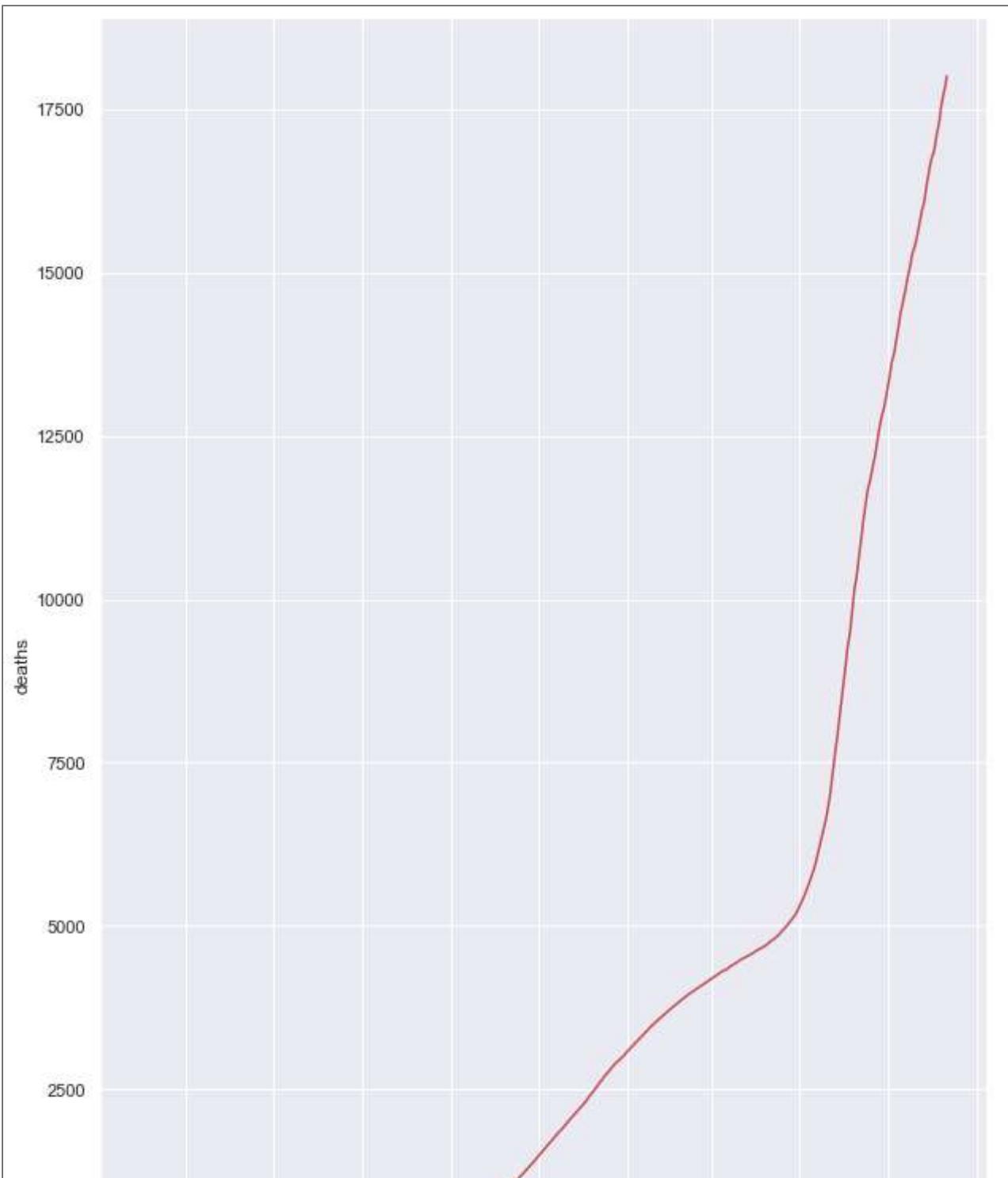
```
sns.set(rc={'figure.figsize':(10,15)})  
sns.lineplot(x="date",y="confirmed",data=Kerala,color="g")  
plt.show()
```

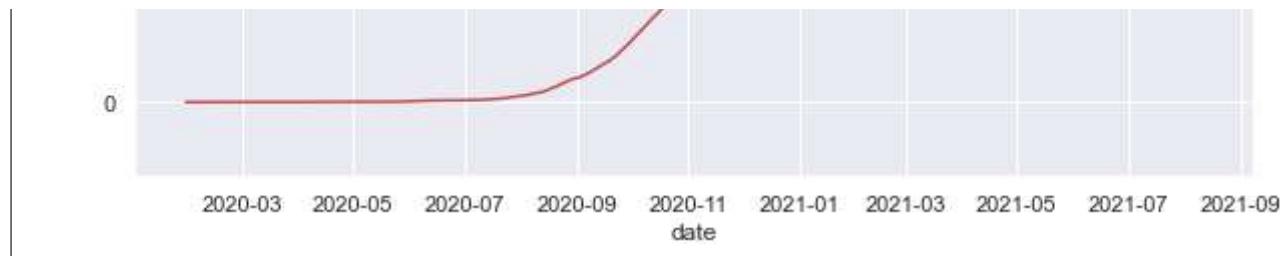




In [154]:

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





In [155]:

```
jk=df[df.state == 'Jammu and Kashmir']
jk
```

Out[155]:

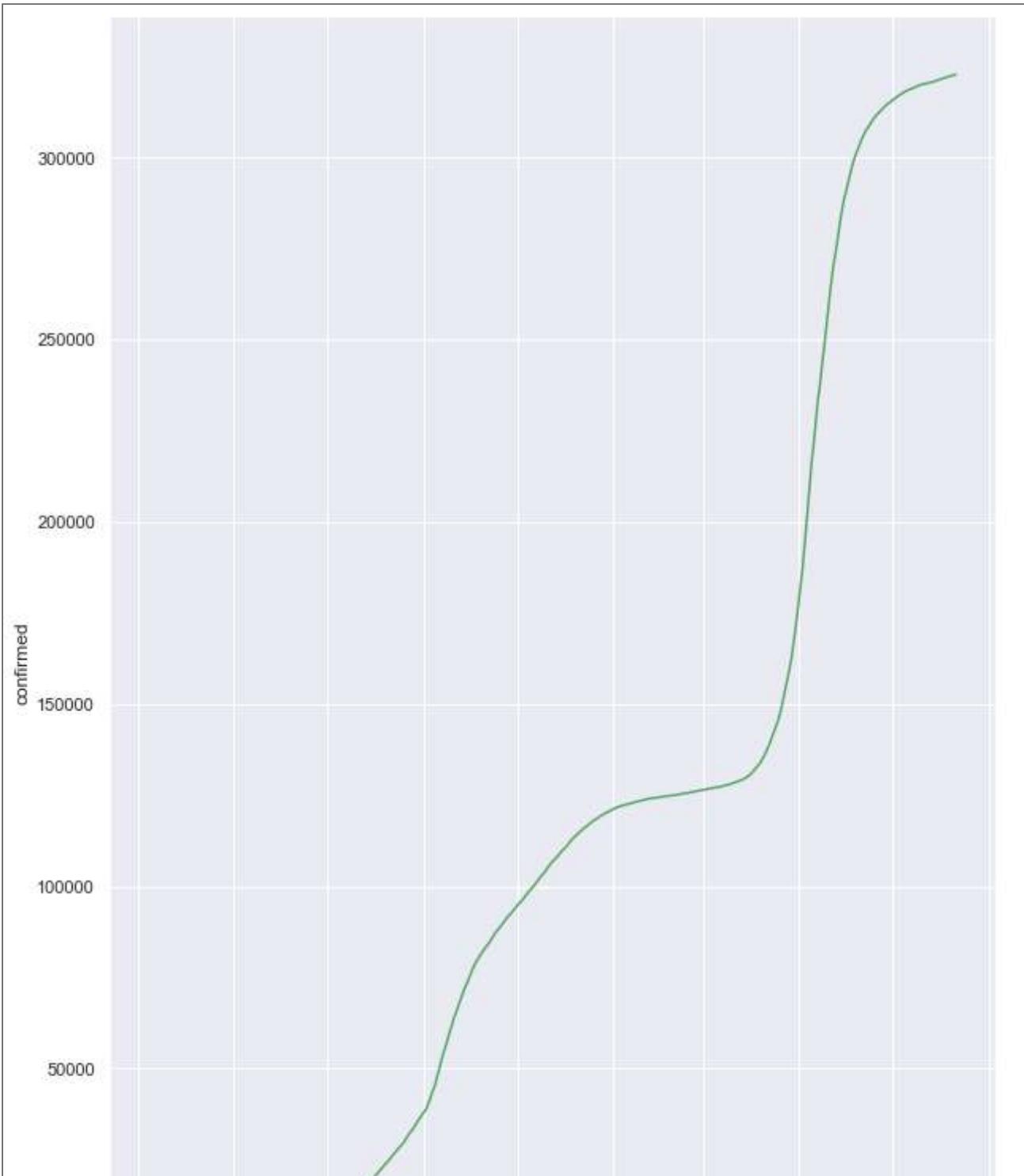
	date	state	cured	deaths	confirmed
81	2020-03-09	Jammu and Kashmir	0	0	1
96	2020-03-10	Jammu and Kashmir	0	0	1
106	2020-03-11	Jammu and Kashmir	0	0	1
117	2020-03-12	Jammu and Kashmir	0	0	1
130	2020-03-13	Jammu and Kashmir	0	0	1
...					
17943	2021-08-07	Jammu and Kashmir	316496	4386	322286
17979	2021-08-08	Jammu and Kashmir	316632	4386	322428
18015	2021-08-09	Jammu and Kashmir	316761	4389	322550

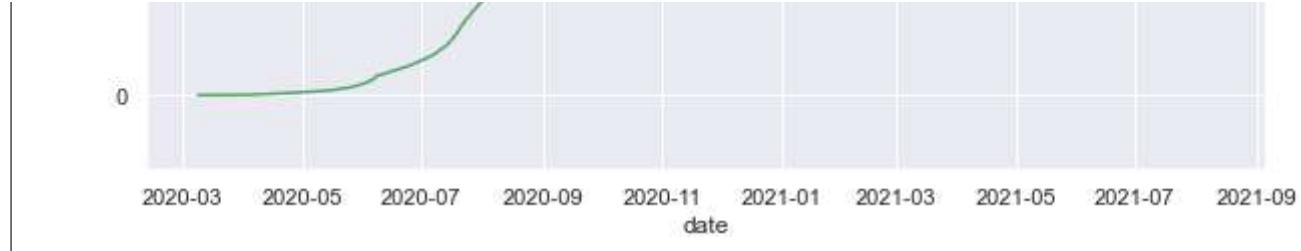
	date	state	cured	deaths	confirmed
18051	2021-08-10	Jammu and Kashmir	316957	4390	322658
18087	2021-08-11	Jammu and Kashmir	317081	4392	322771

521 rows × 5 columns

In [156]:

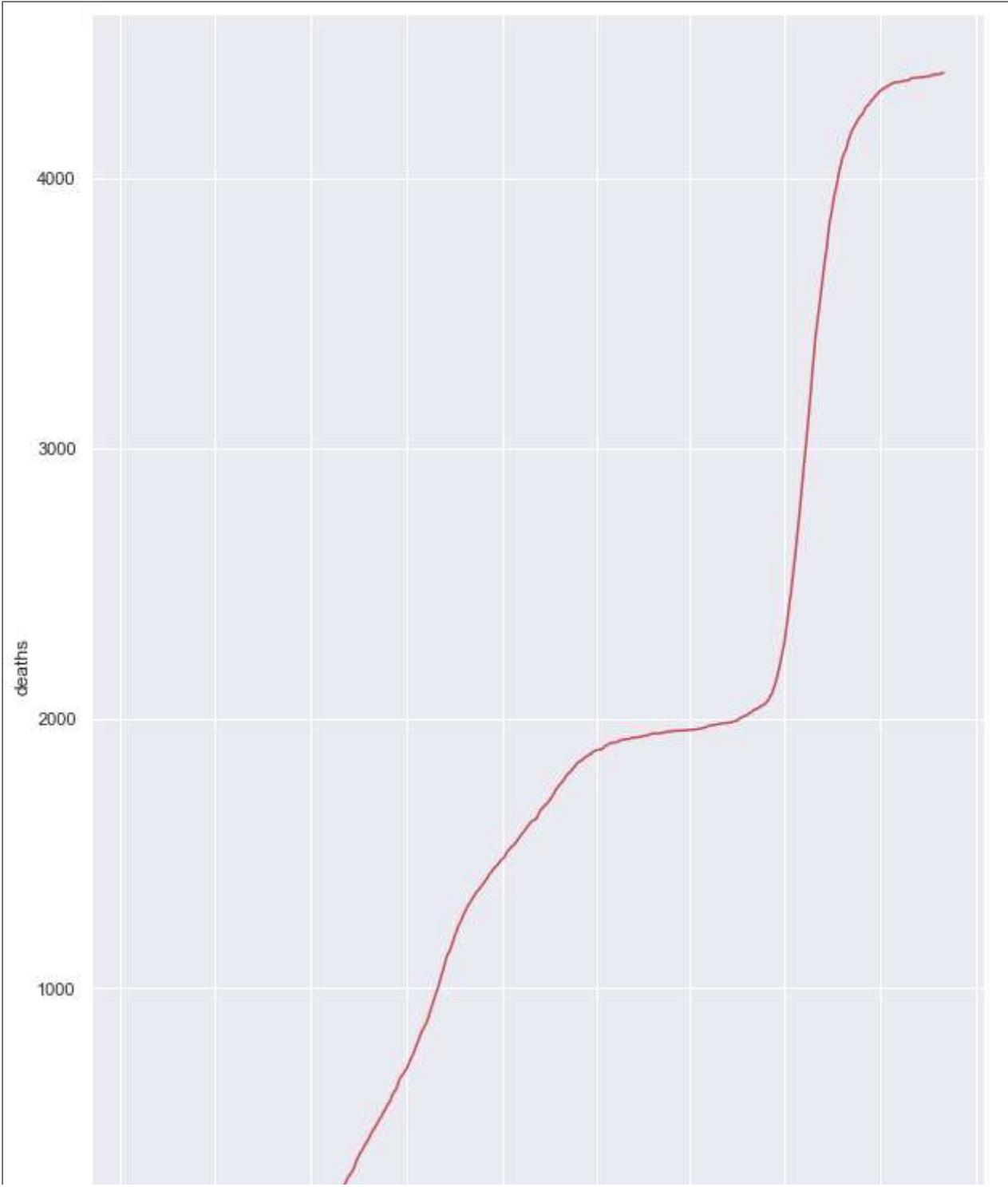
```
sns.set(rc={'figure.figsize':(10,15)})  
sns.lineplot(x="date",y="confirmed",data=jk,color="g")  
plt.show()
```





In [157]:

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





In [158]:

```
#linear regression
from sklearn.model_selection import train_test_split
```

In [159]:

```
mah
```

Out[159]:

	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
...
17950	2021-08-07	Maharashtra	6130137	133717	6341759

	date	state	cured	deaths	confirmed
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

520 rows × 5 columns

In [160]:

```
# mah['date']=mah['date'].map(dt.datetime.toordinal)
# mah.head()
```

In [161]:

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

In [162]:

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

In [164]:

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

In [165]:

```
lr=LinearRegression()
```

In [169]:

```
y_train
```

Out[169]:

7350	1554389
407	130
11794	2071306
97	2
14998	5378452
...	
9240	1855341
15034	5405068
14422	4602472
14242	4295027
10966	2009106

Name: confirmed, Length: 364, dtype: int64

In [171]:

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

Out[171]:

LinearRegression()

In [172]:

```
mah.tail()
```

Out[172]:

	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

In [176]:

```
lr.predict(np.array([[737625]]))
```

Out[176]:

array([[-2.33254044e+08]])

In []:

