Suraj Kadam - Covid19 India Project

In [129]:

```
import numpy as np
import pandas as pd  #importing numpy ,matplotlib ,seaborn,plotly & pandas
import matplotlib.pyplot as plt
import seaborn as sns
import plotly as pl
import plotly.graph_objs as go
import plotly.offline as py
```

In [130]:

```
covid_19_india=pd.read_csv('covid_19_india.csv')  #reading csv
```

In [131]:

```
covid_19_india
```

Out[131]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

In [132]:

```
covid_19_india_copy=covid_19_india.copy()   #creating copy of data
```

In [133]:

```
covid_19_india_copy
```

Out[133]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

16850 rows × 12 columns

In [134]:

```
covid_19_india_copy.head()  #first 5 rows
```

Out[134]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignN
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	

In [135]:

```
covid_19_india_copy.tail()  #last 5 rows
```

Out[135]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

In [136]:

```
covid_19_india_copy.info()   #all column name & data types
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 16850 entries, 0 to 16849
Data columns (total 12 columns):
 #   Column                    Non-Null Count  Dtype  
---  ------                    --------------  -----  
 0   Sno                       16850 non-null  int64  
 1   Date                      16850 non-null  object 
 2   Time                      16850 non-null  object 
 3   State/UnionTerritory      16850 non-null  object 
 4   ConfirmedIndianNational   16850 non-null  object 
 5   ConfirmedForeignNational  16850 non-null  object 
 6   Cured                     16850 non-null  int64  
 7   Deaths                    16850 non-null  int64  
 8   Confirmed                 16850 non-null  int64  
 9   Unnamed: 9                0 non-null      float64
 10  Unnamed: 10               12 non-null     object 
 11  Unnamed: 11               12 non-null     object 
dtypes: float64(1), int64(4), object(7)
memory usage: 1.5+ MB
```

In [137]:

```
covid_19_india_copy.describe()  #description of data
```

Out[137]:

	Sno	Cured	Deaths	Confirmed	Unnamed: 9
count	16850.000000	1.685000e+04	16850.000000	1.685000e+04	0.0
mean	8425.500000	2.360353e+05	3485.222552	2.583667e+05	NaN
std	4864.320353	5.225438e+05	9330.541749	5.672808e+05	NaN
min	1.000000	0.000000e+00	0.000000	0.000000e+00	NaN
25%	4213.250000	2.658500e+03	22.000000	3.644750e+03	NaN
50%	8425.500000	2.889500e+04	453.000000	3.336150e+04	NaN
75%	12637.750000	2.537510e+05	3071.250000	2.666530e+05	NaN
max	16850.000000	5.872268e+06	123531.000000	6.113335e+06	NaN

In [138]:

```
#desc_df=pd.DataFrame(covid_19_india_copy.describe())
#desc_df
```

In [ ]:

```
```

In [139]:

```
covid_19_india_copy=covid_19_india_copy.drop(columns=['Unnamed: 9','Unnamed: 10','U
#deleting unnecessary columns
```

In [140]:

```
covid_19_india_copy
```

Out[140]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

16850 rows × 9 columns

In [141]:

```
total_cured=covid_19_india['Cured'].sum()  #total cured
```

In [142]:

```
total_cured
```

Out[142]:

```
3977194136
```

In [143]:

```
total_deaths=covid_19_india['Deaths'].sum()  #total deaths
```

In [144]:

```
total_deaths
```

Out[144]:

```
58726000
```

In [145]:

```
total_confirmed=covid_19_india['Confirmed'].sum()  #total confirm cases
```

In [146]:

```
total_confirmed
```

Out[146]:

```
4353478074
```

In [147]:

```
total_covid_data=np.array([total_cured,total_deaths,total_confirmed])
total_covid_data
```

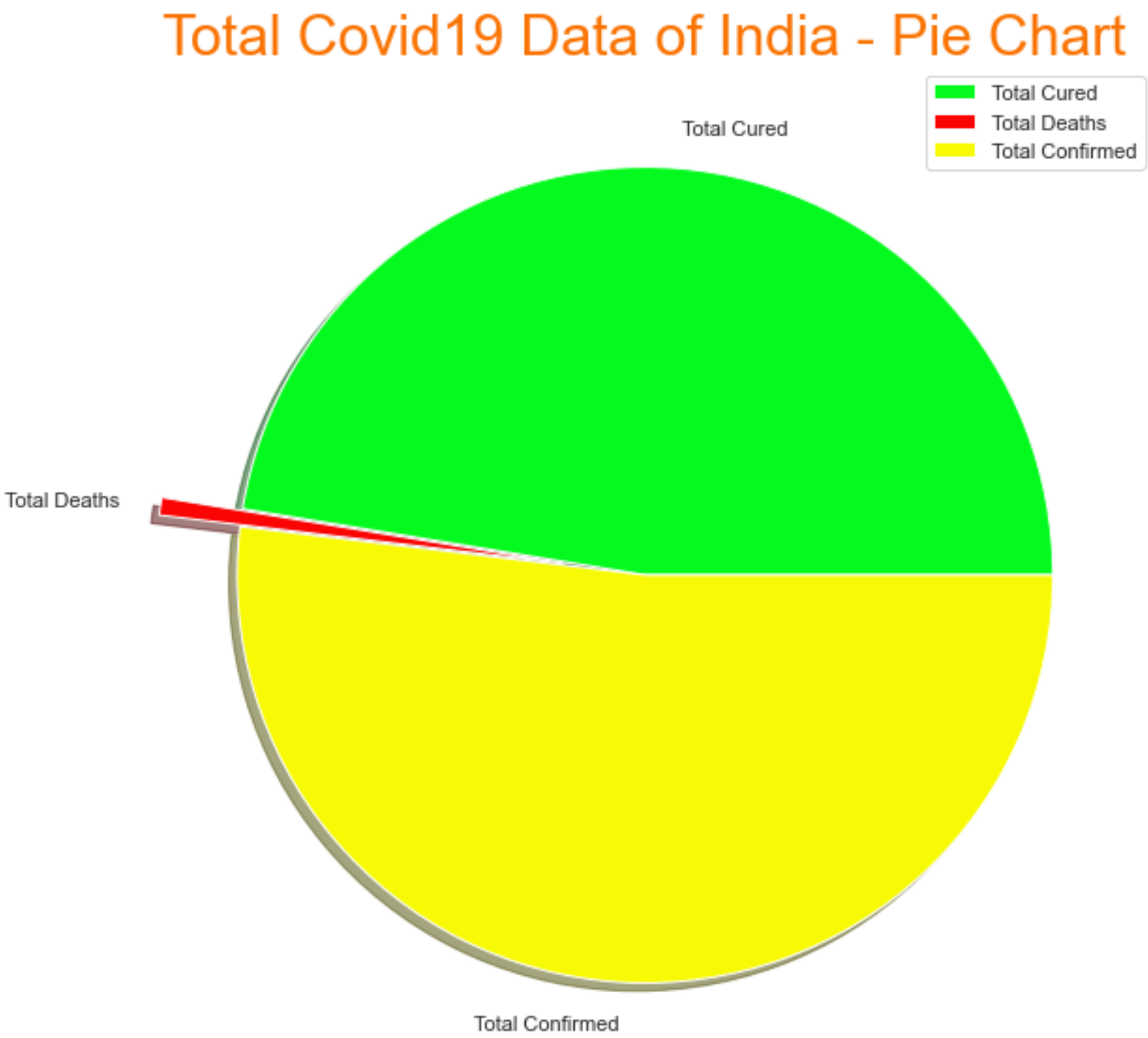
Out[147]:

```
array([3977194136,   58726000, 4353478074], dtype=int64)
```

Plotting data in bar plot graph

In [148]:

```
plt.figure(figsize=(10,10))
title_color='#FA7404'
color=['#04FA1E','#FA0404','#F7FA04']
label=['Total Cured','Total Deaths','Total Confirmed']
plt.pie(total_covid_data,labels=label,shadow=True,colors=color,explode=(0.0,0.2,0.0
plt.title('Total Covid19 Data of India - Pie Chart',size=30,color=title_color)
plt.legend(label)
plt.show()
```


Total Covid19 Data of India - Pie Chart
Total Cured
Total Deaths
Total Confirmed
Total Cured
Total Deaths
Total Confirmed

In [149]:

```
all_states=covid_19_india['State/UnionTerritory'].value_counts()
#checking cases count against states
```

In [150]:

```
all_states
```

Out[150]:

Kerala	525
Delhi	493
Rajasthan	492
Uttar Pradesh	491
Haryana	491
Ladakh	488
Tamil Nadu	488
Punjab	486
Maharashtra	486
Karnataka	486
Jammu and Kashmir	486
Andhra Pradesh	483
Uttarakhand	480
Odisha	479
Puducherry	477
West Bengal	477
Chhattisgarh	476
Chandigarh	476
Gujarat	475
Himachal Pradesh	474
Madhya Pradesh	474
Bihar	471
Manipur	471

```
Mizoram                                          470
Andaman and Nicobar Islands                      469
Goa                                              469
Assam                                            463
Jharkhand                                        463
Arunachal Pradesh                                461
Tripura                                          457
Meghalaya                                        450
Dadra and Nagar Haveli and Daman and Diu         426
Telengana                                        426
Nagaland                                         417
Sikkim                                           410
Lakshadweep                                      209
Telangana                                         67
Cases being reassigned to states                  60
Unassigned                                         3
Bihar****                                          2
Dadra and Nagar Haveli                             2
Daman & Diu                                        1
Name: State/UnionTerritory, dtype: int64
```

In [151]:

```
top5_statedata=all_states[0:5]
top5_statedata
```

Out[151]:

```
Kerala           525
Delhi            493
Rajasthan        492
Uttar Pradesh    491
Haryana          491
Name: State/UnionTerritory, dtype: int64
```

In [152]:

```
top5_statenames=top5_statedata.index
top5_statenames
```

Out[152]:

```
Index(['Kerala', 'Delhi', 'Rajasthan', 'Uttar Pradesh', 'Haryana'],
dtype='object')
```

In [153]:

```
top5_statecases=top5_statedata.values
top5_statecases
```

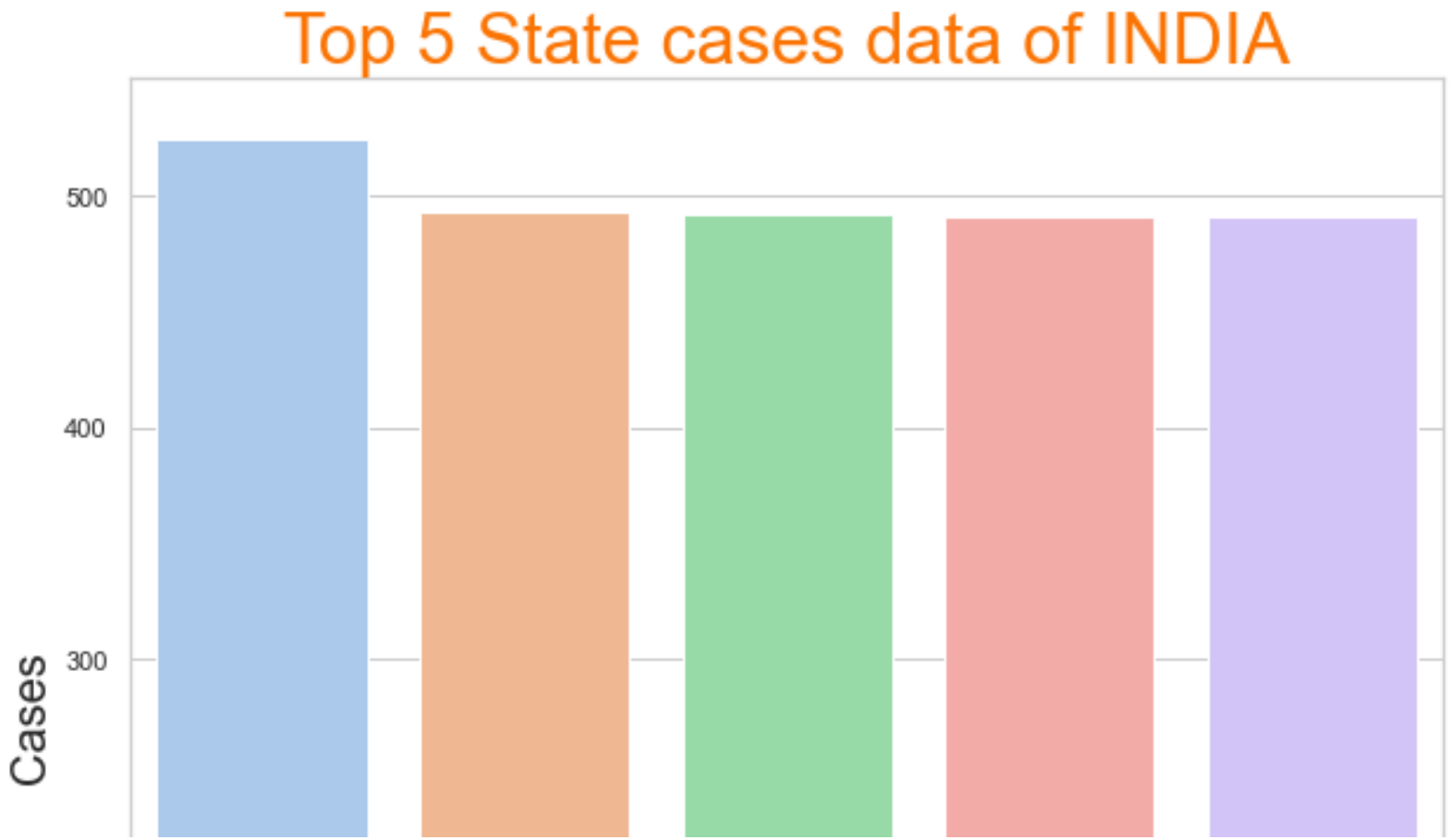
Out[153]:

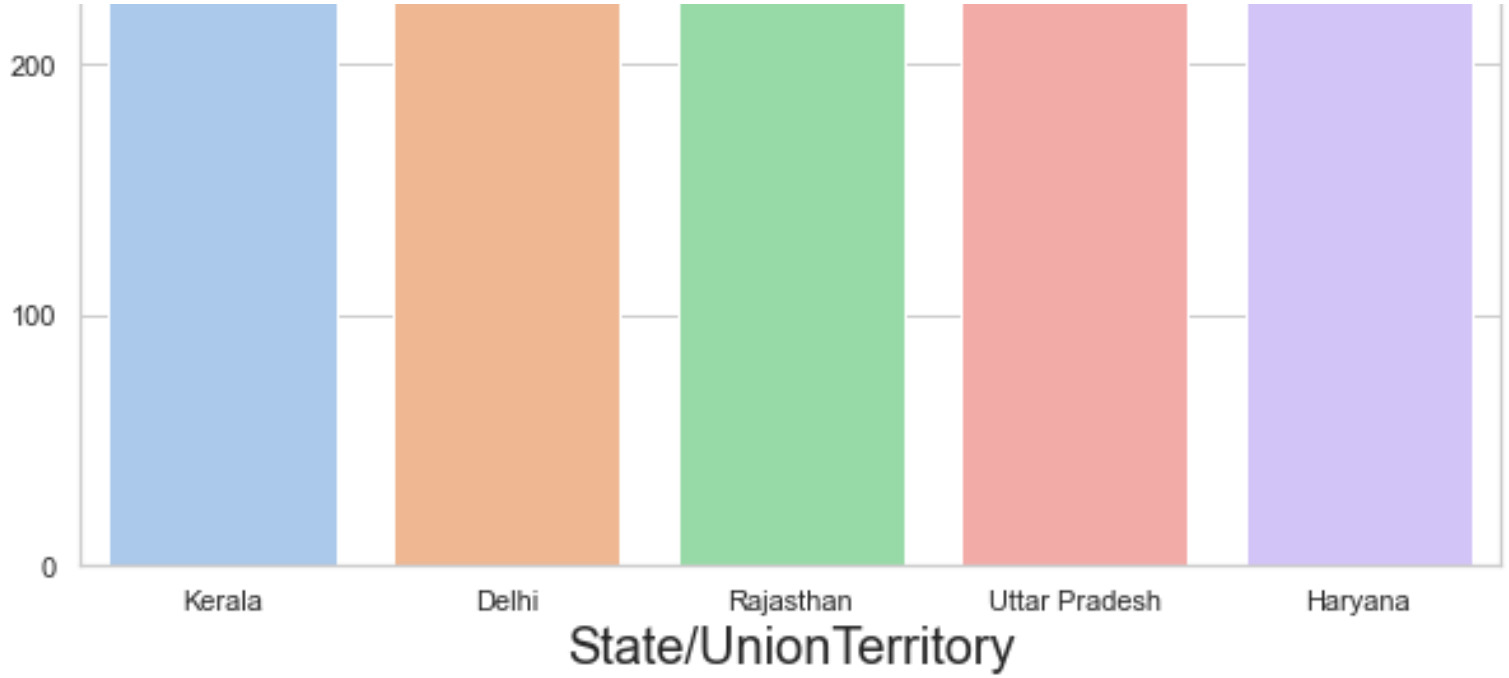
```
array([525, 493, 492, 491, 491], dtype=int64)
```

Top 5 State wise total cases data

In [258]:

```
plt.figure(figsize=(10,10))
sns.set(style='whitegrid')
#colors=['#33FF46','#33FFE9','#E8FC17','#FC3617','#1733FC']
sns.barplot(x=top5_statenames,y=top5_statecases,palette='pastel') #,color=colors)
plt.xlabel('State/UnionTerritory',size=20)
plt.ylabel('Cases',size=20)
plt.title("Top 5 State cases data of INDIA",size=30,color=title_color)
plt.show()
```





In [155]:

```
death_rate=(total_deaths/total_confirmed)*100   #calculating death rate
```

In [156]:

```
print('Death Rate for India Covid 19 is {:.2f} %'.format(death_rate))
```

```
Death Rate for India Covid 19 is 1.35 %
```

In [157]:

```
positiveity_rate=(total_confirmed/total_cured)*100  #calculating positivity rate
```

In [158]:

```
positiveity_rate
```

Out[158]:

```
109.46104024930605
```

In [159]:

```
covid_19_india_copy.head()
```

Out[159]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignN
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	

In [160]:

```
Statewise_data=covid_19_india_copy.groupby('State/UnionTerritory')[['Cured','Deaths
```

In [161]:

```
Statewise_data
```

Out[161]:

State/UnionTerritory	Cured	Deaths	Confirmed
Andaman and Nicobar Islands	1589935	22624	1675248
Andhra Pradesh	303427899	2475816	324146783
Arunachal Pradesh	5150519	19303	5598324
Assam	74011348	459575	80418492
Bihar	100131380	756282	106881540
Bihar****	1402468	18881	1430909
Cases being reassigned to states	0	0	345565
Chandigarh	7980284	119356	8691806
Chhattisgarh	117163544	1591126	128751782
Dadra and Nagar Haveli	20352	8	20722
Dadra and Nagar Haveli and Daman and Diu	1470986	874	1566846
Daman & Diu	0	0	2

State/UnionTerritory	Cured	Deaths	Confirmed
Delhi	224062704	4066907	236972842
Goa	20224042	338359	22280065
Gujarat	103995131	1866811	114557615
Haryana	100010131	1166573	107408371
Himachal Pradesh	20682770	371931	23052151
Jammu and Kashmir	42295048	686680	46899925
Jharkhand	46083978	569298	49971564
Karnataka	345648926	4819018	387597335
Kerala	311127643	1327754	344319045
Ladakh	3059045	38578	3344131
Lakshadweep	471712	2178	561459
Madhya Pradesh	100169697	1427780	108712983
Maharashtra	813788907	19314532	908892470
Manipur	8420223	122089	9440912
Meghalaya	4606548	66293	5221064
Mizoram	1534630	5073	1822190

State/UnionTerritory	Cured	Deaths	Confirmed
Nagaland	3628619	39420	4089547
Odisha	117984789	600149	126408397
Puducherry	14376916	249683	15858688
Punjab	71108712	2216735	78999515
Rajasthan	117312772	1159823	128998101
Sikkim	1983899	41530	2315519
Tamil Nadu	317067499	4731627	342829697
Telangana	35544978	217455	38162058
Telengana	64666267	400427	69990668
Tripura	10479169	124444	11397656
Unassigned	0	0	161
Uttar Pradesh	232529439	3347656	252843682
Uttarakhand	36684388	728512	41179396
West Bengal	195296839	3214840	209822848

In [162]:

```
statewise_df=pd.DataFrame(Statewise_data)
statewise_df
```

Out[162]:

State/UnionTerritory	Cured	Deaths	Confirmed
Andaman and Nicobar Islands	1589935	22624	1675248
Andhra Pradesh	303427899	2475816	324146783
Arunachal Pradesh	5150519	19303	5598324
Assam	74011348	459575	80418492
Bihar	100131380	756282	106881540
Bihar****	1402468	18881	1430909
Cases being reassigned to states	0	0	345565
Chandigarh	7980284	119356	8691806
Chhattisgarh	117163544	1591126	128751782
Dadra and Nagar Haveli	20352	8	20722
Dadra and Nagar Haveli and Daman and Diu	1470986	874	1566846

State/UnionTerritory	Cured	Deaths	Confirmed
Daman & Diu	0	0	2
Delhi	224062704	4066907	236972842
Goa	20224042	338359	22280065
Gujarat	103995131	1866811	114557615
Haryana	100010131	1166573	107408371
Himachal Pradesh	20682770	371931	23052151
Jammu and Kashmir	42295048	686680	46899925
Jharkhand	46083978	569298	49971564
Karnataka	345648926	4819018	387597335
Kerala	311127643	1327754	344319045
Ladakh	3059045	38578	3344131
Lakshadweep	471712	2178	561459
Madhya Pradesh	100169697	1427780	108712983
Maharashtra	813788907	19314532	908892470
Manipur	8420223	122089	9440912
Meghalaya	4606548	66293	5221064

State/UnionTerritory	Cured	Deaths	Confirmed
Mizoram	1534630	5073	1822190
Nagaland	3628619	39420	4089547
Odisha	117984789	600149	126408397
Puducherry	14376916	249683	15858688
Punjab	71108712	2216735	78999515
Rajasthan	117312772	1159823	128998101
Sikkim	1983899	41530	2315519
Tamil Nadu	317067499	4731627	342829697
Telangana	35544978	217455	38162058
Telengana	64666267	400427	69990668
Tripura	10479169	124444	11397656
Unassigned	0	0	161
Uttar Pradesh	232529439	3347656	252843682
Uttarakhand	36684388	728512	41179396
West Bengal	195296839	3214840	209822848

In [163]:

```
statewise_death=statewise_df.Deaths
statewise_death[:10]
```

Out[163]:

```
State/UnionTerritory
Andaman and Nicobar Islands          22624
Andhra Pradesh                     2475816
Arunachal Pradesh                    19303
Assam                               459575
Bihar                               756282
Bihar****                            18881
Cases being reassigned to states         0
Chandigarh                          119356
Chhattisgarh                       1591126
Dadra and Nagar Haveli                   8
Name: Deaths, dtype: int64
```

Maharashtra State Covid19 Data

In [164]:

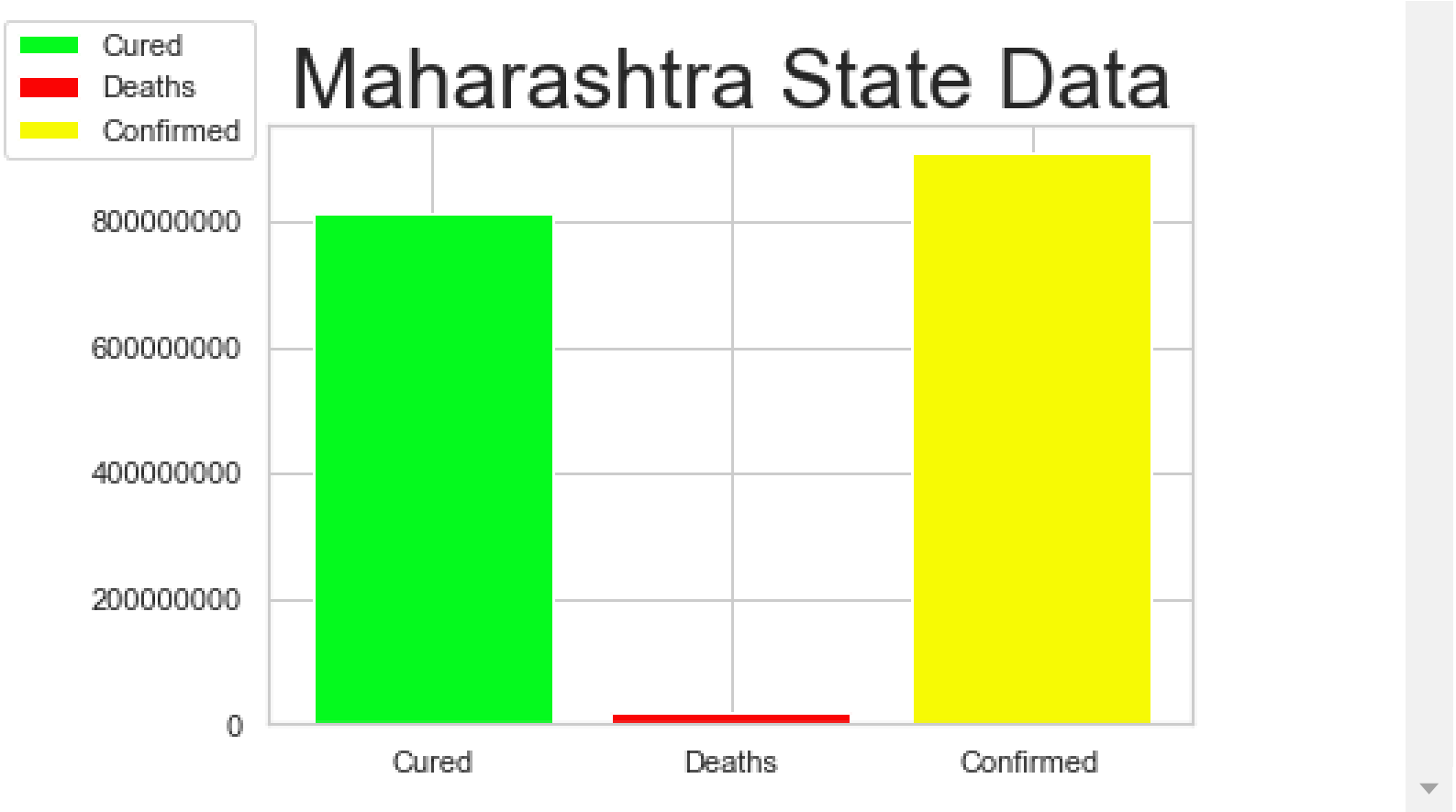
```
mh_data=statewise_df[statewise_df.index=='Maharashtra']
mh_data
```

Out[164]:

State/UnionTerritory	Cured	Deaths	Confirmed
Maharashtra	813788907	19314532	908892470

In [165]:

```
#plt.figure(figsize=(15,10))
sns.set(style='whitegrid')
plt.bar(x=0,height=mh_data.Cured,color='#04FA1E')
plt.bar(x=1,height=mh_data.Deaths,color='#FA0404')
plt.bar(x=2,height=mh_data.Confirmed,color='#F7FA04')
plt.ticklabel_format(useOffset=False,style='Plain')
plt.xticks([0,1,2],['Cured','Deaths','Confirmed'])
plt.title('Maharashtra State Data',size=30)
plt.legend(['Cured','Deaths','Confirmed'],loc='upper right',bbox_to_anchor=(0,1,0,0
plt.show()
```



In [166]:

```
covid_19_india_copy['Date']=pd.to_datetime(covid_19_india_copy['Date'])
#converting date column to datetime format
```

In [167]:

```
covid_19_india_copy['Date']
```

Out[167]:

```
0       2020-01-30
1       2020-01-31
2       2020-02-01
3       2020-02-02
4       2020-02-03
           ...
16845   2021-07-07
16846   2021-07-07
16847   2021-07-07
16848   2021-07-07
16849   2021-07-07
Name: Date, Length: 16850, dtype: datetime64[ns]
```

In [168]:

```
covid_19_india_copy['Year']=pd.DatetimeIndex(covid_19_india_copy['Date']).year
#extracting year and adding into new column
```

In [169]:

```
covid_19_india_copy
```

Out[169]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

16850 rows × 10 columns

In [170]:

```
covid_19_india_copy['Month']=pd.DatetimeIndex(covid_19_india_copy['Date']).month
#extracting month and adding into new column
```

In [171]:

```
covid_19_india_copy['Day']=pd.DatetimeIndex(covid_19_india_copy['Date']).day
#extracting day and adding into new column
```

In [172]:

```
covid_19_india_copy
```

Out[172]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

16850 rows × 12 columns

In [173]:

```
covid_19_india_copy['Year'].unique()
```

Out[173]:

```
array([2020, 2021], dtype=int64)
```

In [174]:

```
yearly_data=covid_19_india_copy.groupby('Year')[['Cured','Deaths','Confirmed']].sum
```

In [175]:

```
yearly_data
```

Out[175]:

Year	Cured	Deaths	Confirmed
2020	941314195	17022508	1074022781
2021	3035879941	41703492	3279455293

In [176]:

```
yearly_df=pd.DataFrame(yearly_data)
yearly_df
```

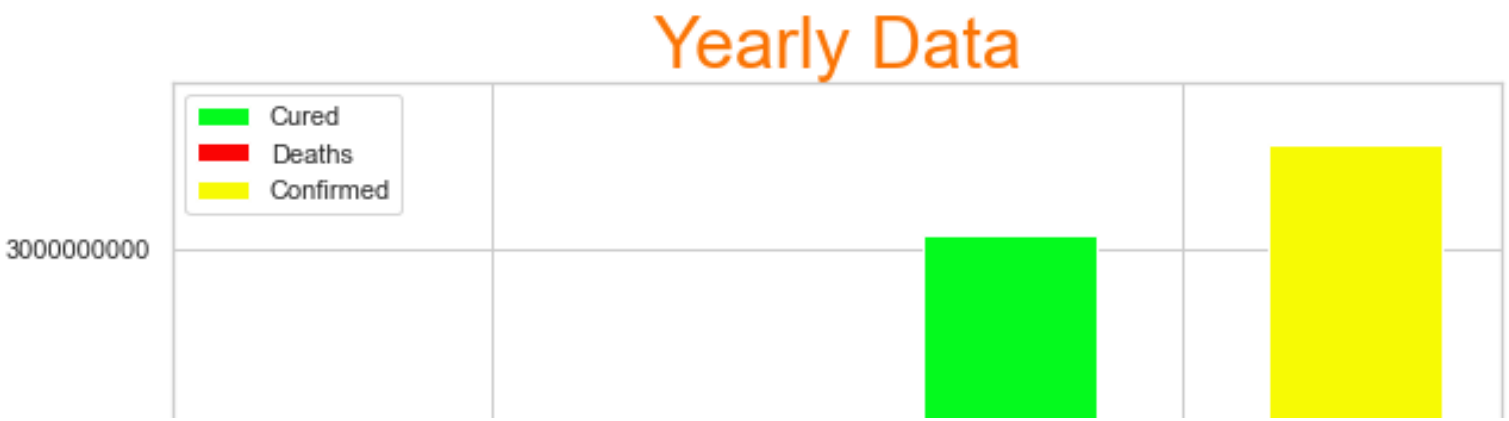
Out[176]:

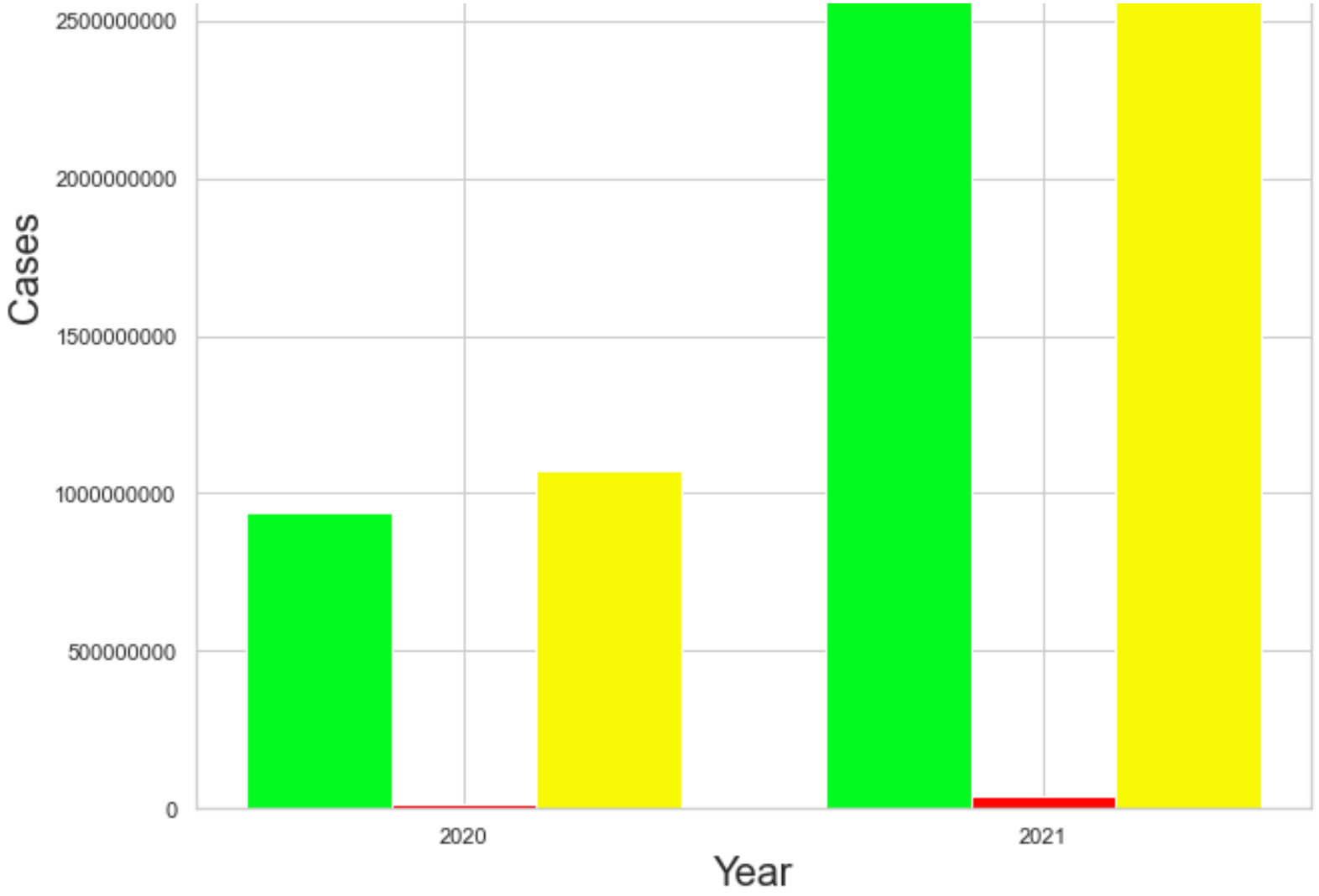
Year	Cured	Deaths	Confirmed
2020	941314195	17022508	1074022781
2021	3035879941	41703492	3279455293

Yearly data

In [177]:

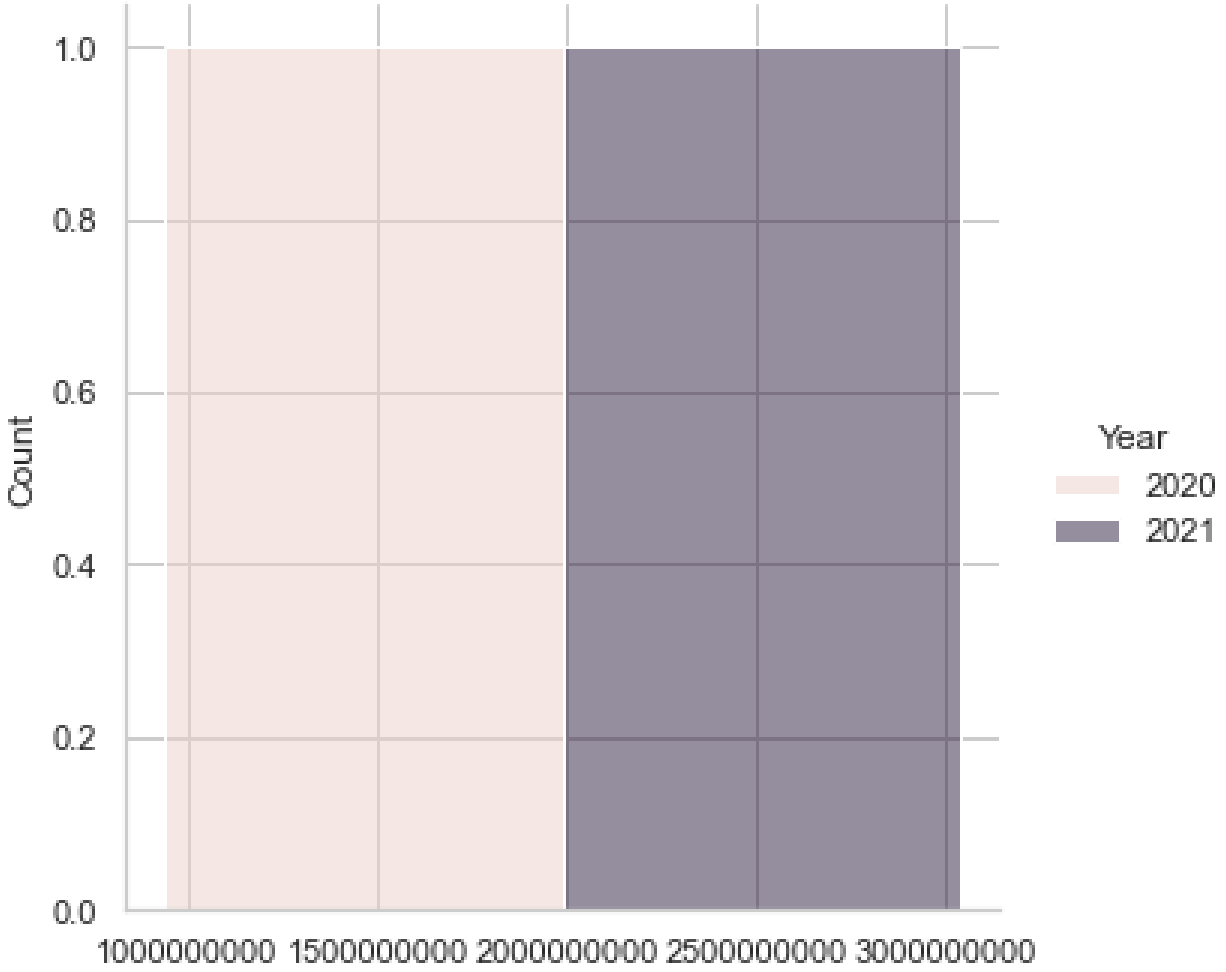
```
plt.figure(figsize=(10,10))
#labels=['Cured','Deaths','Confirmed']
labels=yearly_df.columns
bar_width=0.25
b1=np.arange(len(yearly_df.Cured.values))
b2=[x+bar_width for x in b1]
b3=[x+bar_width for x in b2]
plt.bar(b1,yearly_df.Cured.values,width=bar_width,color='#04FA1E')
plt.bar(b2,yearly_df.Deaths.values,width=bar_width,color='#FA0404')
plt.bar(b3,yearly_df.Confirmed.values,width=bar_width,color='#F7FA04')
plt.xlabel( 'Year',size=20)
plt.ylabel( 'Cases',size=20)
plt.ticklabel_format(useOffset=False, style='plain')
plt.xticks([ item + bar_width for item in range ( len ( yearly_df.Cured ))],[ '2020
plt.legend( labels )
plt.title('Yearly Data',size=30,color=title_color)
plt.show()
```




2500000000
2000000000
1500000000
1000000000
500000000
0
Cases
2020
2021
Year

In [178]:

```
sns.displot(x=yearly_df.Cured.values,hue=yearly_df.index)
plt.ticklabel_format(useOffset=False, style='plain')
plt.show()
```



In [ ]:

```
```

In [179]:

```
monthly_data=covid_19_india_copy.groupby('Month')[['Cured','Deaths','Confirmed']].s
```

In [180]:

```
monthly_data
```

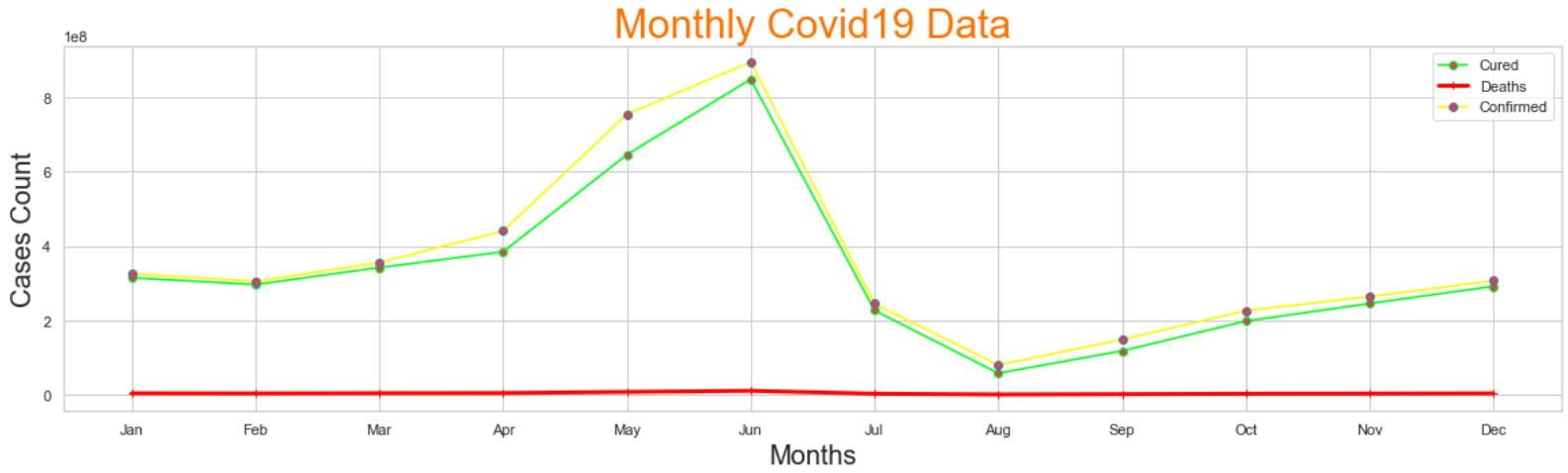
Out[180]:

Month	Cured	Deaths	Confirmed
1	315332019	4709167	326469749
2	297133802	4359434	305631889
3	342611205	4935455	356315303
4	385065633	5353568	441083113
5	646240106	8480751	754865720
6	848822379	11475067	895231838
7	227533465	3606557	245513007
8	58580895	1553468	80749620
9	118592934	2443374	149113758
10	198824412	3457615	226770312
11	246213201	3894165	264556412
12	292244085	4457379	307177353

Monthly data

In [181]:

```
plt.figure(figsize=(20,5))
labels=yearly_df.columns
month_label=['Jan','Feb','Mar','Apr','May','Jun','Jul','Aug','Sep','Oct','Nov','Dec
plt.plot(month_label,monthly_data.Cured,color='#04FA1E',marker='o',mfc='r')
plt.plot(month_label,monthly_data.Deaths,color='#FA0404',marker='+',linewidth=3)
plt.plot(month_label,monthly_data.Confirmed,color='#F7FA04',marker='o',mfc='r',mec=
plt.legend(labels)
plt.xlabel('Months',size=20)
plt.ylabel('Cases Count',size=20)
plt.title("Monthly Covid19 Data",size=30,color=title_color)
plt.show()
```

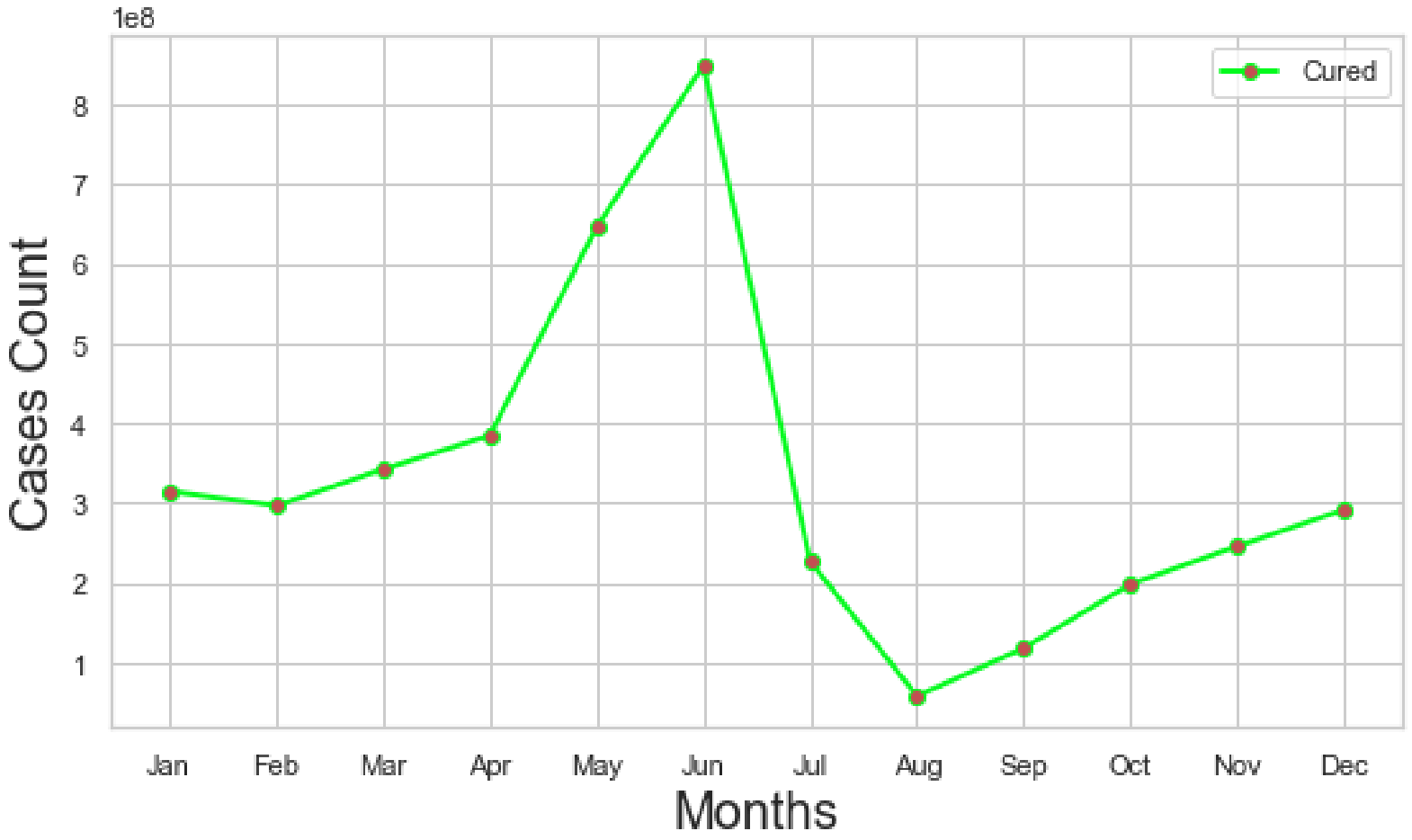


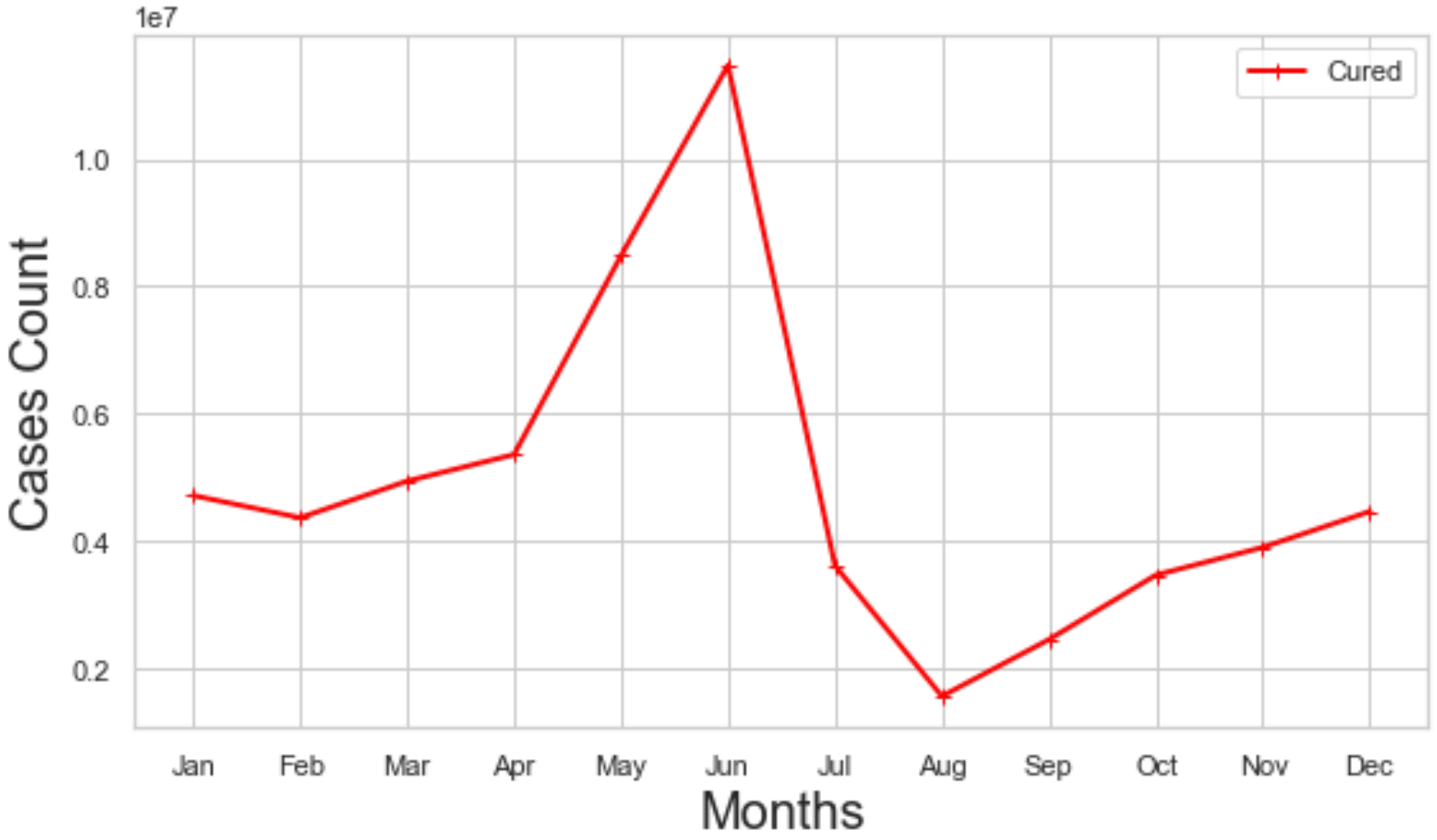
Monthly data - separate plots

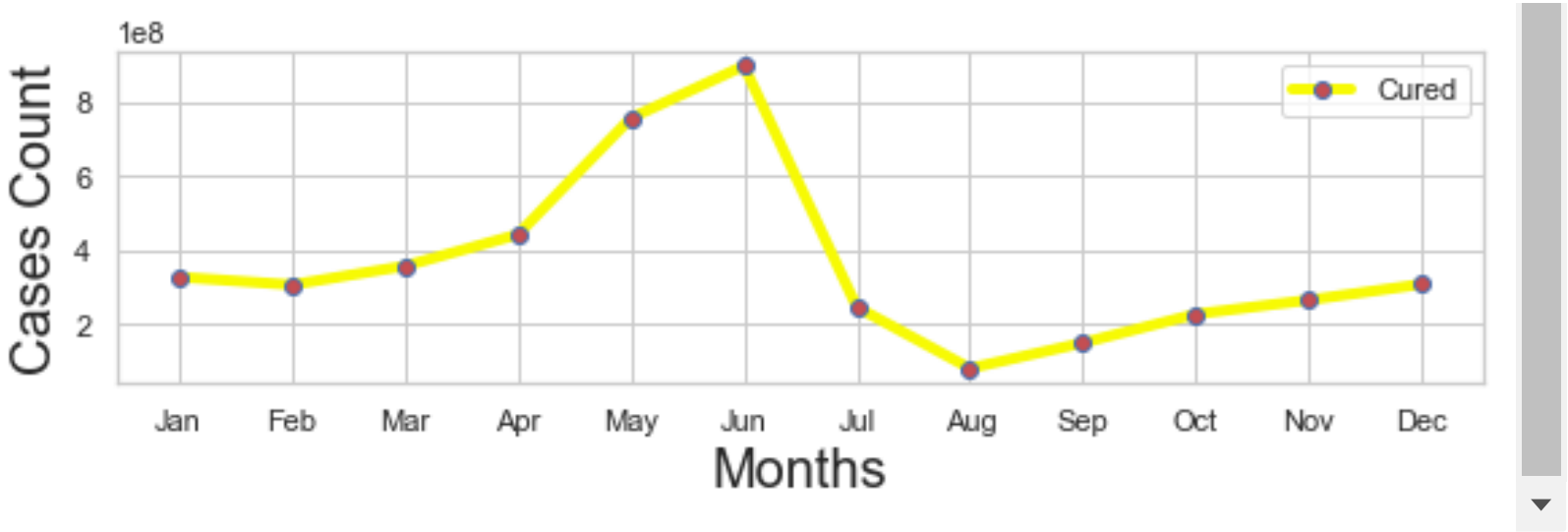
In [182]:

```
plt.figure(figsize=(20,5))
month_label=['Jan','Feb','Mar','Apr','May','Jun','Jul','Aug','Sep','Oct','Nov','Dec
#Plot 1
plt.title('Plot 1 - Monthly Cured',size=30,color=title_color)
plt.subplot(1,2,1)
plt.plot(month_label,monthly_data.Cured,color='#04FA1E',marker='o',mfc='r',linewidt
plt.legend(labels)
#plt.ticklabel_format(useOffset=False, style='plain')
plt.xlabel('Months',size=20)
plt.ylabel('Cases Count',size=20)
plt.show()


#Plot 2
plt.figure(figsize=(20,5))
plt.title('Plot 2 - Monthly Deaths')
plt.subplot(1,2,2)
plt.plot(month_label,monthly_data.Deaths,color='#FA0404',marker='+',linewidth=2)
plt.legend(labels)
plt.xlabel('Months',size=20)
plt.ylabel('Cases Count',size=20)
plt.show()

#Plot 3
plt.figure(figsize=(20,5))
plt.title('Plot 3 - Monthly Confirmed')
```

```
plt.subplot(2,2,1)
plt.plot(month_label,monthly_data.Confirmed,marker='o',mfc='r',mec='b',color='#F7FA
plt.legend(labels)
plt.xlabel('Months',size=20)
plt.ylabel('Cases Count',size=20)
plt.show()
```



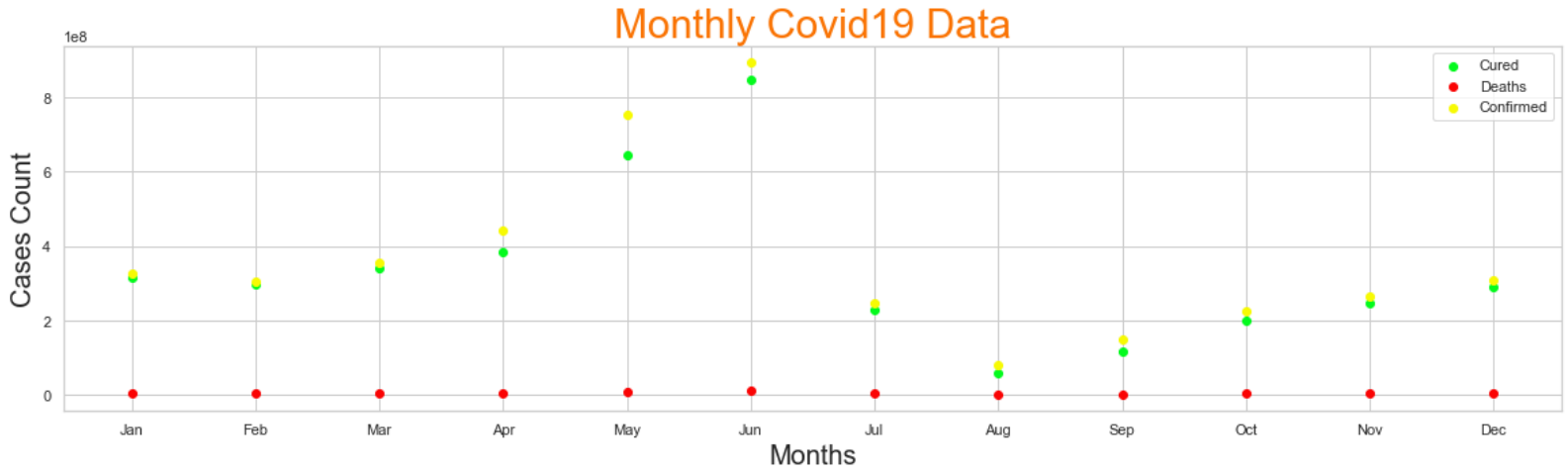

1e7
Cured
1.0
0.8
0.6
0.4
0.2
Cases Count
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Months



Monthly data - Scatter plot

In [183]:

```
plt.figure(figsize=(20,5))
labels=yearly_df.columns
month_label=['Jan','Feb','Mar','Apr','May','Jun','Jul','Aug','Sep','Oct','Nov','Dec
plt.scatter(month_label,monthly_data.Cured,color='#04FA1E')
plt.scatter(month_label,monthly_data.Deaths,color='#FA0404')
plt.scatter(month_label,monthly_data.Confirmed,color='#F7FA04')
plt.legend(labels)
plt.xlabel('Months',size=20)
plt.ylabel('Cases Count',size=20)
plt.title("Monthly Covid19 Data",size=30,color=title_color)
plt.show()
```



In [184]:

```
Statewise_yearly_data=covid_19_india_copy.groupby(['State/UnionTerritory','Year'])[
```

In [185]:

```
Statewise_yearly_data
```

Out[185]:

State/UnionTerritory	Year	Cured	Deaths	Confirmed
Andaman and Nicobar Islands	**2020**	534731	7772	590838
	2021	1055204	14852	1084410
Andhra Pradesh	**2020**	95094768	871178	104134066
	2021	208333131	1604638	220012717
Arunachal Pradesh	**2020**	1442769	4488	1679508
...	**...**	...	...	...
Uttar Pradesh	**2021**	177050153	2411300	189954972
Uttarakhand	**2020**	6864570	121701	7960257
	2021	29819818	606811	33219139
West Bengal	**2020**	44508487	941478	50095209
	2021	150788352	2273362	159727639

78 rows × 3 columns

In [ ]:

```
```

In [186]:

```
Statewise_monthly_data=covid_19_india_copy.groupby(['State/UnionTerritory','Month']
```

In [187]:

```
Statewise_monthly_data
```

Out[187]:

State/UnionTerritory	Month	Cured	Deaths	Confirmed
Andaman and Nicobar Islands	**1**	151473	1922	154187
	2	138309	1736	140209
	3	153888	1922	156090
	4	155615	1905	160840
	5	194448	2728	204112
...	**...**	...	...	...
West Bengal	**8**	2668216	73557	3537564
	9	5379943	121453	6232914
	10	8463405	181984	9643839
	11	11728981	228800	12862215
	12	15314155	282550	16211044

439 rows × 3 columns

In [188]:

```
kerala_data=covid_19_india_copy[covid_19_india_copy['State/UnionTerritory']=='Kera
```

In [189]:

```
kerala_data
```

Out[189]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16686	16687	2021-07-03	8:00 AM	Kerala	-	
16722	16723	2021-07-04	8:00 AM	Kerala	-	
16758	16759	2021-07-05	8:00 AM	Kerala	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16794	16795	2021-07-06	8:00 AM	Kerala	-	
16830	16831	2021-07-07	8:00 AM	Kerala	-	

525 rows × 12 columns

In [190]:

```
kerala_yearly_data=kerala_data.groupby('Year')[['Cured','Deaths','Confirmed']].sum(
```

In [191]:

```
kerala_yearly_data
```

Out[191]:

Year	Cured	Deaths	Confirmed
2020	42951434	193376	51854118
2021	268176209	1134378	292464927

In [192]:

```
kerala_yearly_df=pd.DataFrame(kerala_yearly_data)
kerala_yearly_df
```

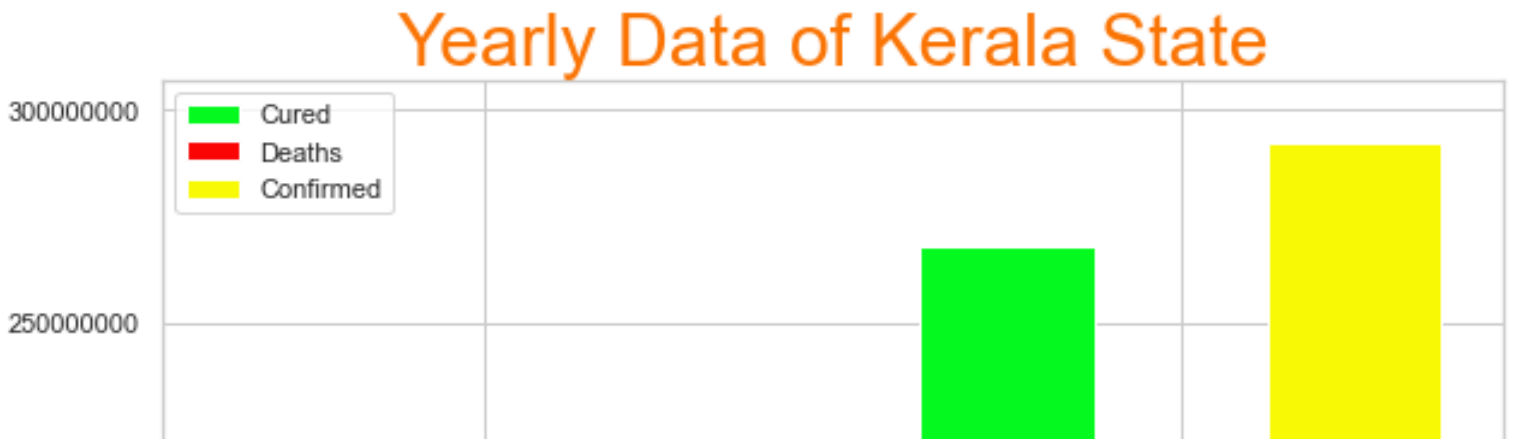
Out[192]:

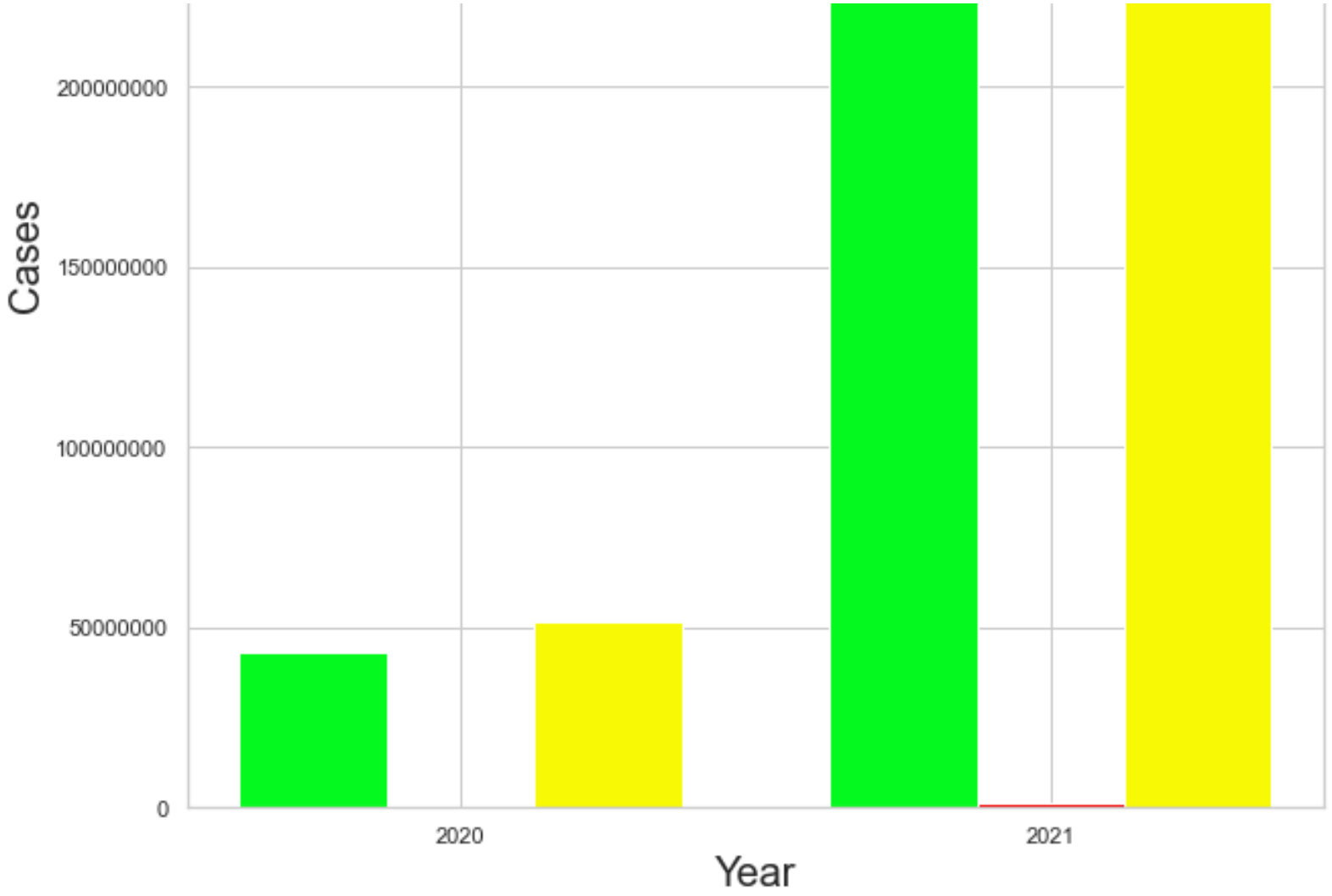
Year	Cured	Deaths	Confirmed
2020	42951434	193376	51854118
2021	268176209	1134378	292464927

Yearly Data of Kerala

In [193]:

```
plt.figure(figsize=(10,10))
#labels=['Cured','Deaths','Confirmed']
labels=kerala_yearly_df.columns
bar_width=0.25
b1=np.arange(len(kerala_yearly_df.Cured.values))
b2=[x+bar_width for x in b1]
b3=[x+bar_width for x in b2]
plt.bar(b1,kerala_yearly_df.Cured.values,width=bar_width,color='#04FA1E')
plt.bar(b2,kerala_yearly_df.Deaths.values,width=bar_width,color='#FA0404')
plt.bar(b3,kerala_yearly_df.Confirmed.values,width=bar_width,color='#F7FA04')
plt.xlabel( 'Year',size=20)
plt.ylabel( 'Cases',size=20)
plt.ticklabel_format(useOffset=False, style='plain')
plt.xticks([ item + bar_width for item in range ( len ( yearly_df.Cured ))],[ '2020
plt.legend( labels )
plt.title('Yearly Data of Kerala State',size=30,color=title_color)
plt.show()
```





In [194]:

```
kerala_monthly_data=kerala_data.groupby('Month')[['Cured','Deaths','Confirmed']].su
```

In [195]:

```
kerala_monthly_data
```

Out[195]:

Month	Cured	Deaths	Confirmed
1	23797725	105654	26008756
2	26076321	111157	27930718
3	32704468	136589	33852823
4	34171565	146410	37315759
5	53709432	203461	64585178
6	77870015	336646	82160691
7	20076034	96459	21068834
8	897214	5049	1399347
9	2528807	14165	3570633
10	6918689	33918	9763929
11	13288127	55658	15591120
12	19089246	82588	21071257

In [196]:

```
kerala_monthly_df=pd.DataFrame(kerala_monthly_data)
kerala_monthly_df
```

Out[196]:

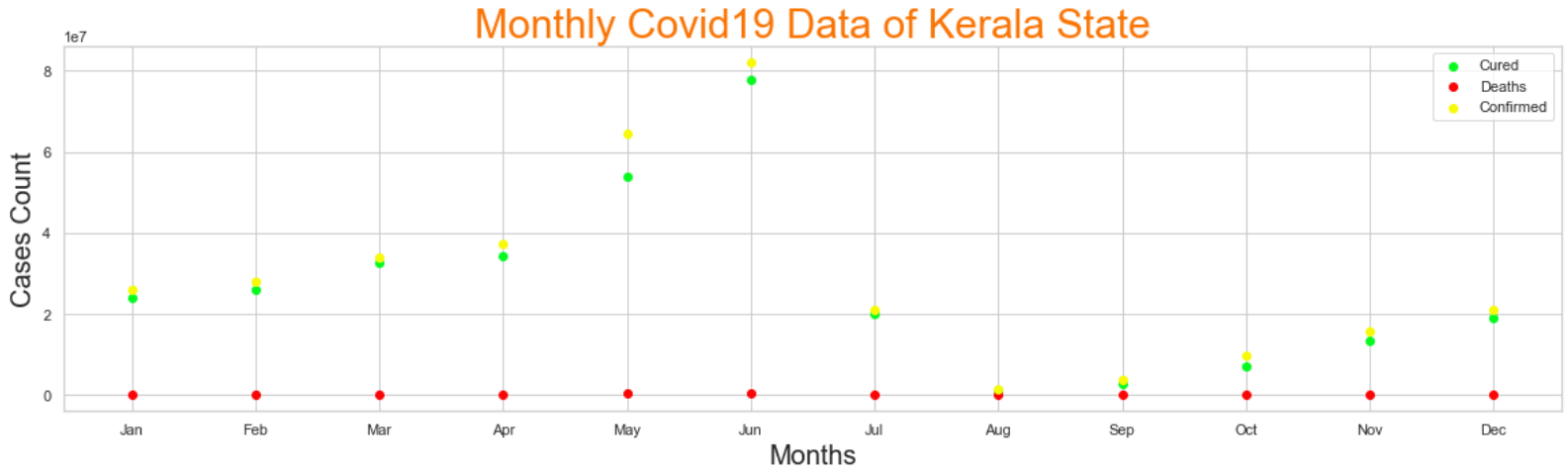
Month	Cured	Deaths	Confirmed
1	23797725	105654	26008756
2	26076321	111157	27930718
3	32704468	136589	33852823
4	34171565	146410	37315759
5	53709432	203461	64585178
6	77870015	336646	82160691
7	20076034	96459	21068834
8	897214	5049	1399347
9	2528807	14165	3570633
10	6918689	33918	9763929
11	13288127	55658	15591120

Month	Cured	Deaths	Confirmed
12	19089246	82588	21071257

Monthly data of Kerala state

In [197]:

```
plt.figure(figsize=(20,5))
labels=kerala_monthly_df.columns
month_label=['Jan','Feb','Mar','Apr','May','Jun','Jul','Aug','Sep','Oct','Nov','Dec
plt.scatter(month_label,kerala_monthly_df.Cured,color='#04FA1E')
plt.scatter(month_label,kerala_monthly_df.Deaths,color='#FA0404')
plt.scatter(month_label,kerala_monthly_df.Confirmed,color='#F7FA04')
plt.legend(labels)
plt.xlabel('Months',size=20)
plt.ylabel('Cases Count',size=20)
plt.title("Monthly Covid19 Data of Kerala State",size=30,color=title_color)
plt.show()
```



In [198]:

```
kerala_timely_data=kerala_data.groupby(['Date','Time'])[['Cured','Deaths','Confirme
```

In [199]:

```
kerala_timely_data
```

Out[199]:

Date	Time	Cured	Deaths	Confirmed
2020-01-30	**6:00 PM**	0	0	1
2020-01-31	**6:00 PM**	0	0	1
2020-02-01	**6:00 PM**	0	0	2
2020-02-02	**6:00 PM**	0	0	3
2020-02-03	**6:00 PM**	0	0	3
...	**...**	...	...	...
2021-07-03	**8:00 AM**	2831394	13505	2949128
2021-07-04	**8:00 AM**	2843909	13640	2961584
2021-07-05	**8:00 AM**	2855460	13716	2973684
2021-07-06	**8:00 AM**	2866806	13818	2981721
2021-07-07	**8:00 AM**	2877557	13960	2996094

525 rows × 3 columns

In [200]:

```
kerala_daily_data=kerala_data.groupby('Day')[['Cured','Deaths','Confirmed']].sum()
```

In [201]:

```
kerala_daily_data
```

Out[201]:

Day	Cured	Deaths	Confirmed
1	11103484	47867	12233670
2	11190809	48393	12338157
3	11283761	48947	12441060
4	11380976	49443	12537531
5	11484955	49868	12644285
6	11582147	50407	12749867
7	11679457	51006	12864350
8	8889279	37474	9957128
9	8975585	37828	10050080
10	9065799	38227	10142987
11	9149750	38642	10228882
12	9238144	39064	10324246

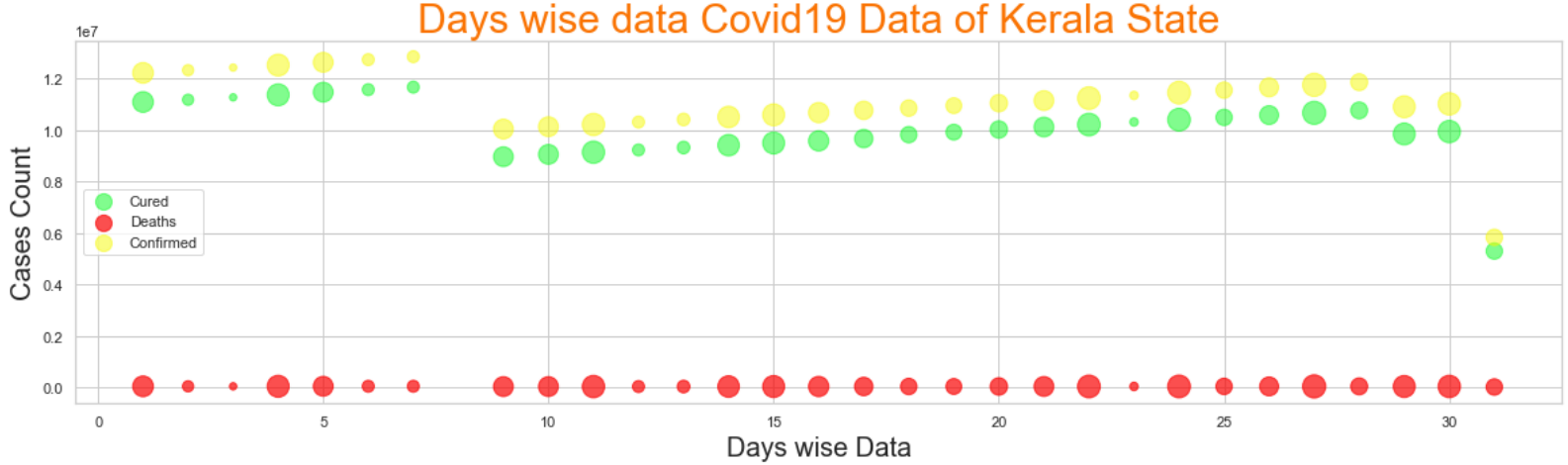
Day	Cured	Deaths	Confirmed
13	9332370	39515	10424151
14	9423198	39987	10521231
15	9508900	40408	10604432
16	9590161	40835	10690180
17	9680019	41248	10778412
18	9832677	41603	10864848
19	9931125	41955	10961786
20	10031545	42364	11058258
21	10130209	42786	11159493
22	10225993	43202	11254673
23	10325943	43706	11359329
24	10415791	44219	11466403
25	10505368	44731	11563107
26	10594052	45205	11670826
27	10682175	45654	11769269
28	10776568	46087	11873834

Day	Cured	Deaths	Confirmed
29	9859438	42379	10917613
30	9952906	42869	11030394
31	5305059	21835	5838563

Day wise Data of Kerala State

In [202]:

```
plt.figure(figsize=(20,5))
labels=kerala_daily_data.columns
sizes = 10*np.random.randint(31,size=(31))
plt.scatter(kerala_daily_data.index,kerala_daily_data.Cured,color='#04FA1E',alpha=0
plt.scatter(kerala_daily_data.index,kerala_daily_data.Deaths,color='#FA0404',alpha=
plt.scatter(kerala_daily_data.index,kerala_daily_data.Confirmed,color='#F7FA04',alp
plt.legend(labels)
plt.xlabel('Days wise Data',size=20)
plt.ylabel('Cases Count',size=20)
plt.title("Days wise data Covid19 Data of Kerala State",size=30,color=title_color)
plt.show()
```



In [203]:

```
kerala_data
```

Out[203]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
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	
...	...	...	...	...	...	
16686	16687	2021-07-03	8:00 AM	Kerala	-	
16722	16723	2021-07-04	8:00 AM	Kerala	-	
16758	16759	2021-07-05	8:00 AM	Kerala	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
16794	16795	2021-07-06	8:00 AM	Kerala	-	
16830	16831	2021-07-07	8:00 AM	Kerala	-	

525 rows × 12 columns

In [204]:

```
kerala_total_deaths=kerala_data['Deaths'].sum()
```

In [269]:

```
kerala_total_deaths
```

Out[269]:

```
1327754
```

In [206]:

```
kerala_total_confirmed=kerala_data['Confirmed'].sum()
```

In [207]:

```
kerala_total_confirmed
```

Out[207]:

344319045

In [266]:

```
kerala_total_df=pd.DataFrame([kerala_total_deaths,kerala_total_confirmed],index=['D
kerala_total_df
```

Out[266]:

	0
Deaths	1327754
Confirmed Cases	344319045

In [273]:

```
kerala_total_df.iloc[1]
```

Out[273]:

```
0    344319045
Name: Confirmed Cases, dtype: int64
```

In [281]:

```
kerala_data_sample=kerala_data.sample(20)
kerala_data_sample
```

Out[281]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
6017	6018	2020-09-07	8:00 AM	Kerala	-	
1654	1655	2020-05-06	8:00 AM	Kerala	-	
16470	16471	2021-06-27	8:00 AM	Kerala	-	
7102	7103	2020-10-08	8:00 AM	Kerala	-	
10962	10963	2021-01-25	8:00 AM	Kerala	-	
14958	14959	2021-05-16	8:00 AM	Kerala	-	
3979	3980	2020-07-11	8:00 AM	Kerala	-	
152	153	2020-03-15	6:00 PM	Kerala	22	

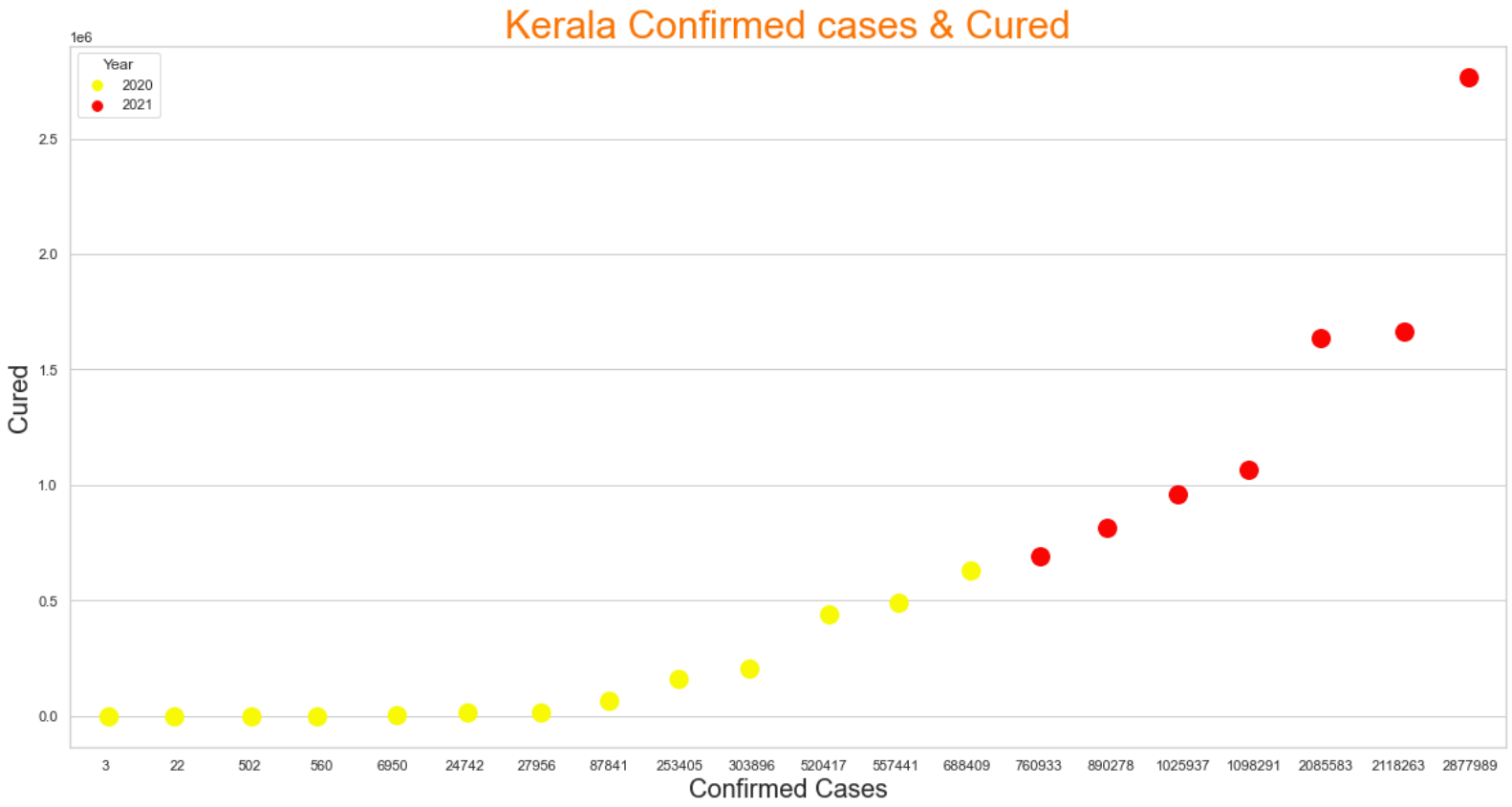
	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
14922	14923	2021-05-15	8:00 AM	Kerala	-	
9594	9595	2020-12-18	8:00 AM	Kerala	-	
1951	1952	2020-05-15	8:00 AM	Kerala	-	
7312	7313	2020-10-14	8:00 AM	Kerala	-	
11	12	2020-02-10	6:00 PM	Kerala	3	
8677	8678	2020-11-22	8:00 AM	Kerala	-	
4862	4863	2020-08-05	8:00 AM	Kerala	-	
8432	8433	2020-11-15	8:00 AM	Kerala	-	
4757	4758	2020-08-02	8:00 AM	Kerala	-	
10098	10099	2021-01-01	8:00 AM	Kerala	-	
12870	12871	2021-03-19	8:00 AM	Kerala	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
11898	11899	2021-02-20	8:00 AM	Kerala	-	

Kerala Sample data plotting

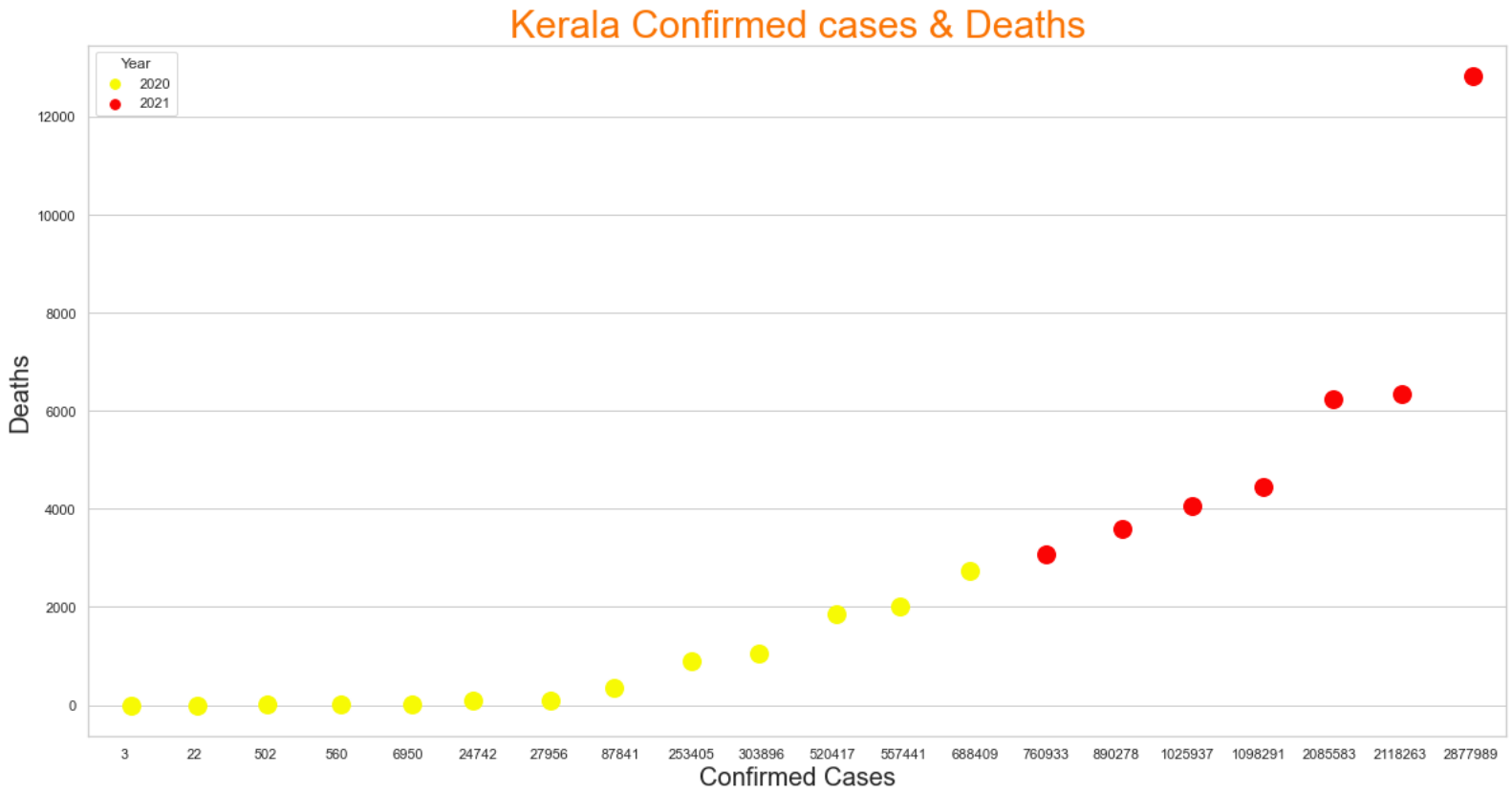
In [318]:

```
plt.figure(figsize=(20,10))
sns.stripplot(x=kerala_data_sample.Confirmed,y=kerala_data_sample.Cured
              ,data=kerala_data_sample,hue='Year',palette=['#F7FA04','#FA0404'],siz
plt.xlabel('Confirmed Cases',size=20)
plt.ylabel('Cured',size=20)
plt.title('Kerala Confirmed cases & Cured',size=30,color=title_color)
plt.show()
```



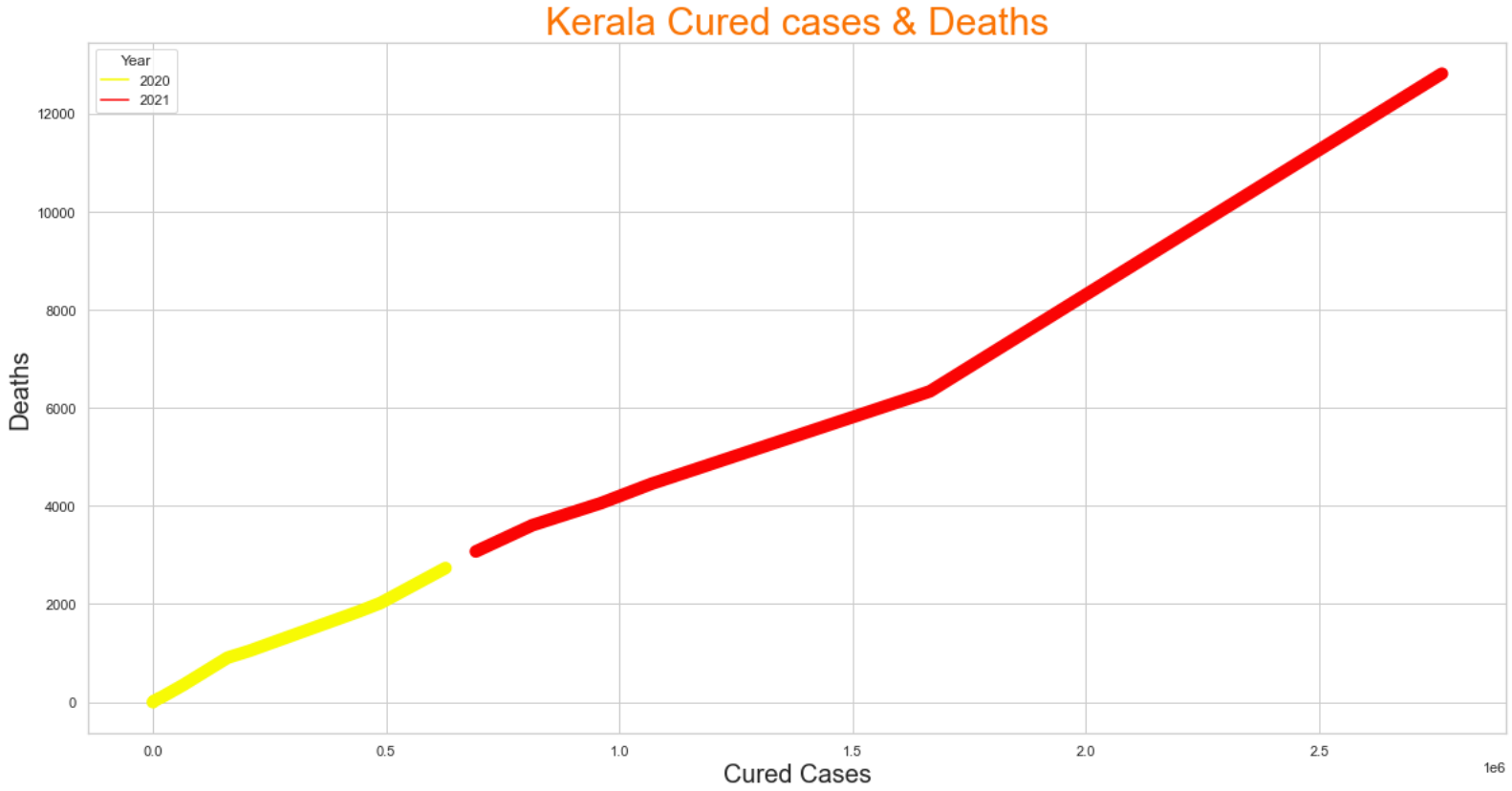
In [317]:

```
plt.figure(figsize=(20,10))
sns.stripplot(x=kerala_data_sample.Confirmed,y=kerala_data_sample.Deaths
              ,data=kerala_data_sample,hue='Year',palette=['#F7FA04','#FA0404'],siz
plt.xlabel('Confirmed Cases',size=20)
plt.ylabel('Deaths',size=20)
plt.title('Kerala Confirmed cases & Deaths',size=30,color=title_color)
plt.show()
```



In [320]:

```
plt.figure(figsize=(20,10))
sns.lineplot(x=kerala_data_sample.Cured,y=kerala_data_sample.Deaths
             ,data=kerala_data_sample,hue='Year',palette=['#F7FA04','#FA0404'],lir
plt.xlabel('Cured Cases',size=20)
plt.ylabel('Deaths',size=20)
plt.title('Kerala Cured cases & Deaths',size=30,color=title_color)
plt.show()
```

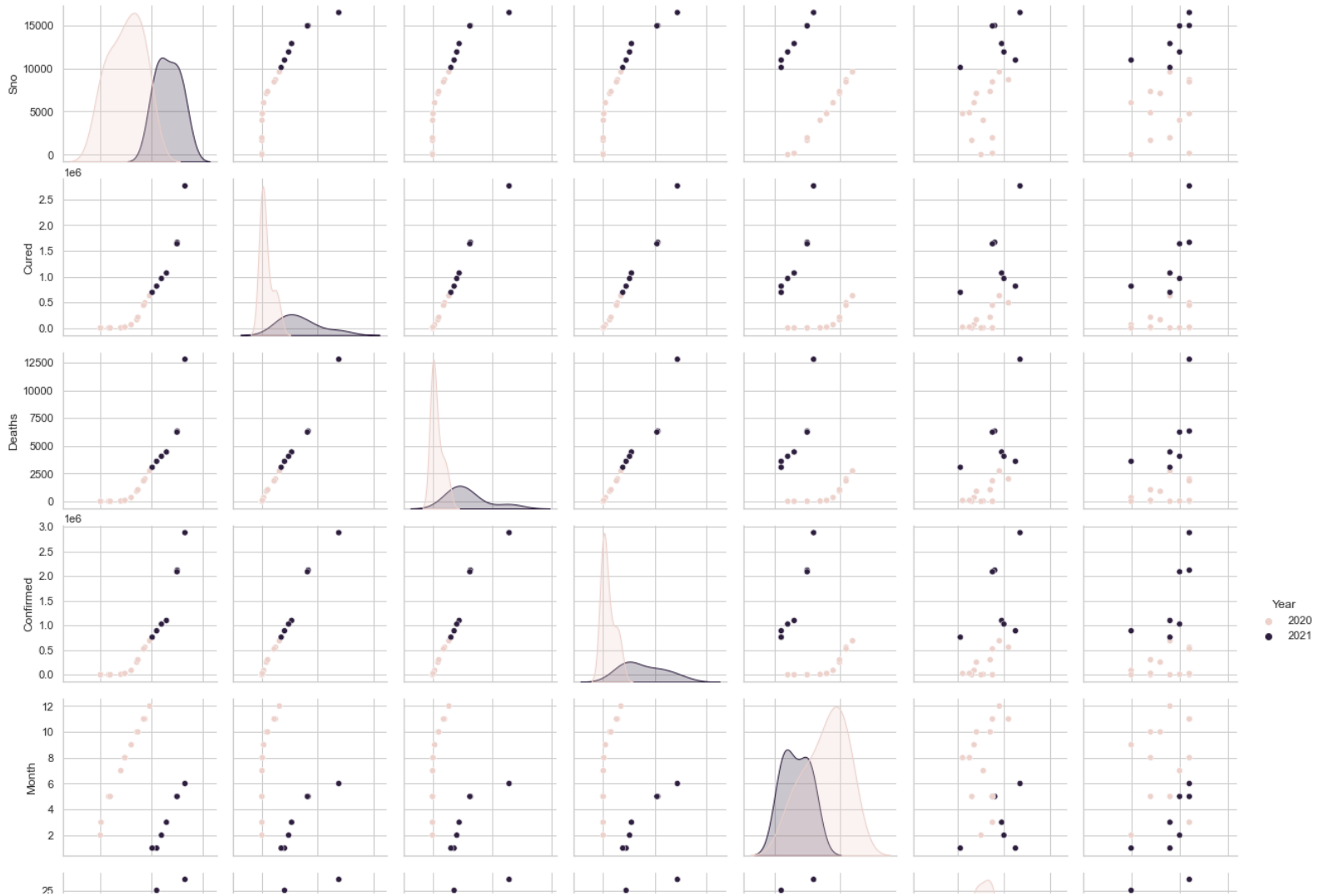


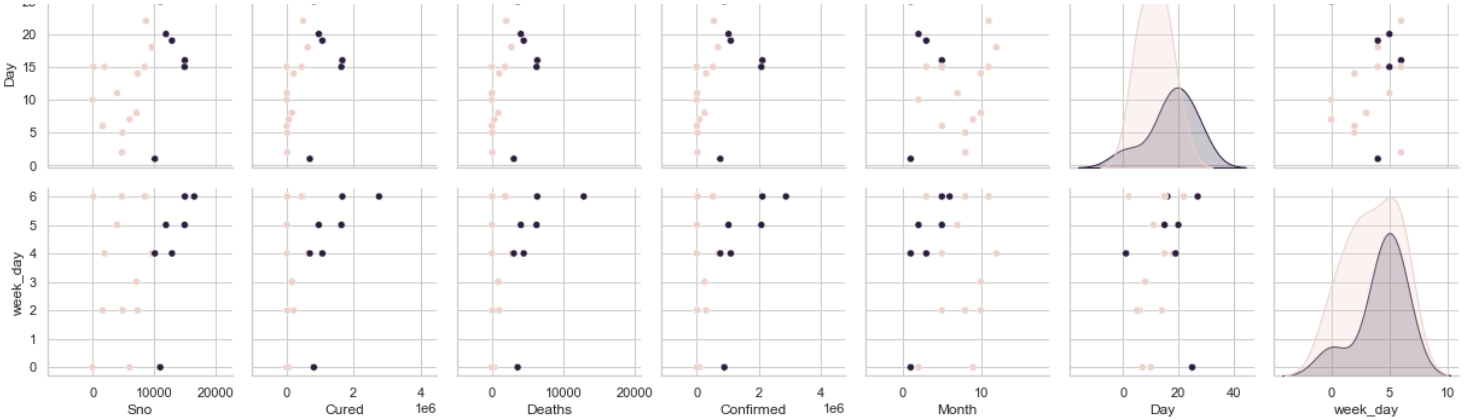
Pairplot against Year

In [296]:

```
plt.figure(figsize=(20,10))
sns.pairplot(hue='Year',data=kerala_data_sample,kind='scatter')
plt.show()
```

```
<Figure size 1440x720 with 0 Axes>
```





In [208]:

```
kerala_death_rate=(kerala_total_deaths/kerala_total_confirmed)*100
```

In [209]:

```
kerala_death_rate
```

Out[209]:

```
0.3856173567163559
```

In [210]:

```
print('Death Rate for Kerala Covid 19 is {:.2f} %'.format(kerala_death_rate))
```

```
Death Rate for Kerala Covid 19 is 0.39 %
```

In [211]:

```
kerala_positivity_rate=(kerala_total_confirmed/(kerala_data['Cured'].sum()))*100
```

In [212]:

```
print('Positivity Rate for Kerala Covid 19 is {:.2f} %'.format(kerala_positivity_ra
```

```
Positivity Rate for Kerala Covid 19 is 110.67 %
```

In [213]:

```
kerala_data.sample(10)
```

Out[213]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
11610	11611	2021-02-12	8:00 AM	Kerala	-	
10602	10603	2021-01-15	8:00 AM	Kerala	-	
14418	14419	2021-05-01	8:00 AM	Kerala	-	
10386	10387	2021-01-09	8:00 AM	Kerala	-	
23	24	2020-02-22	6:00 PM	Kerala	3	
11142	11143	2021-01-30	8:00 AM	Kerala	-	
6857	6858	2020-10-01	8:00 AM	Kerala	-	
5457	5458	2020-08-22	8:00 AM	Kerala	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	Confirme
6927	6928	2020-10-03	8:00 AM	Kerala	-	
485	486	2020-03-30	9:30 PM	Kerala	-	

In [214]:

```
covid_19_india_copy['positivity_rate']=covid_19_india_copy['Confirmed']/covid_19_ir
```

In [215]:

```
covid_19_india_copy
```

Out[215]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

16850 rows × 13 columns

In [216]:

```
covid_19_india_copy['death_rate']=covid_19_india_copy['Deaths']/covid_19_india_copy
```

In [217]:

```
covid_19_india_copy
```

Out[217]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

16850 rows × 14 columns

In [218]:

```
covid_19_india_copy['week_day'] = pd.DatetimeIndex(covid_19_india_copy['Date']).wee
```

```
c:\users\nikhil\appdata\local\programs\python\python37\lib\site-pack
ages\ipykernel_launcher.py:2: 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/pand
as-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-co
py (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexin
g.html#returning-a-view-versus-a-copy)
```

In [306]:

```
kerala_data['week_day'] = pd.DatetimeIndex(kerala_data['Date']).weekday
```

```
c:\users\nikhil\appdata\local\programs\python\python37\lib\site-pack
ages\ipykernel_launcher.py: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/pand
as-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-co
py (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexin
g.html#returning-a-view-versus-a-copy)
```

In [219]:

```
kerala_data
```

Out[219]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16686	16687	2021-07-03	8:00 AM	Kerala	-	
16722	16723	2021-07-04	8:00 AM	Kerala	-	
16758	16759	2021-07-05	8:00 AM	Kerala	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16794	16795	2021-07-06	8:00 AM	Kerala	-	
16830	16831	2021-07-07	8:00 AM	Kerala	-	

525 rows × 13 columns

In [220]:

```
covid_19_india_copy
```

Out[220]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
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	
...	...	...	...	...	...	
16845	16846	2021-07-07	8:00 AM	Telangana	-	
16846	16847	2021-07-07	8:00 AM	Tripura	-	
16847	16848	2021-07-07	8:00 AM	Uttarakhand	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16848	16849	2021-07-07	8:00 AM	Uttar Pradesh	-	
16849	16850	2021-07-07	8:00 AM	West Bengal	-	

16850 rows × 15 columns

In [221]:

```
kerala_weekly_data=kerala_data.groupby('week_day')[['Cured','Deaths','Confirmed']].
```

In [222]:

```
#covid_19_india_copy=covid_19_india_copy.drop(columns=['week_day','WeekDay'])
```

In [223]:

```
kerala_weekly_data
```

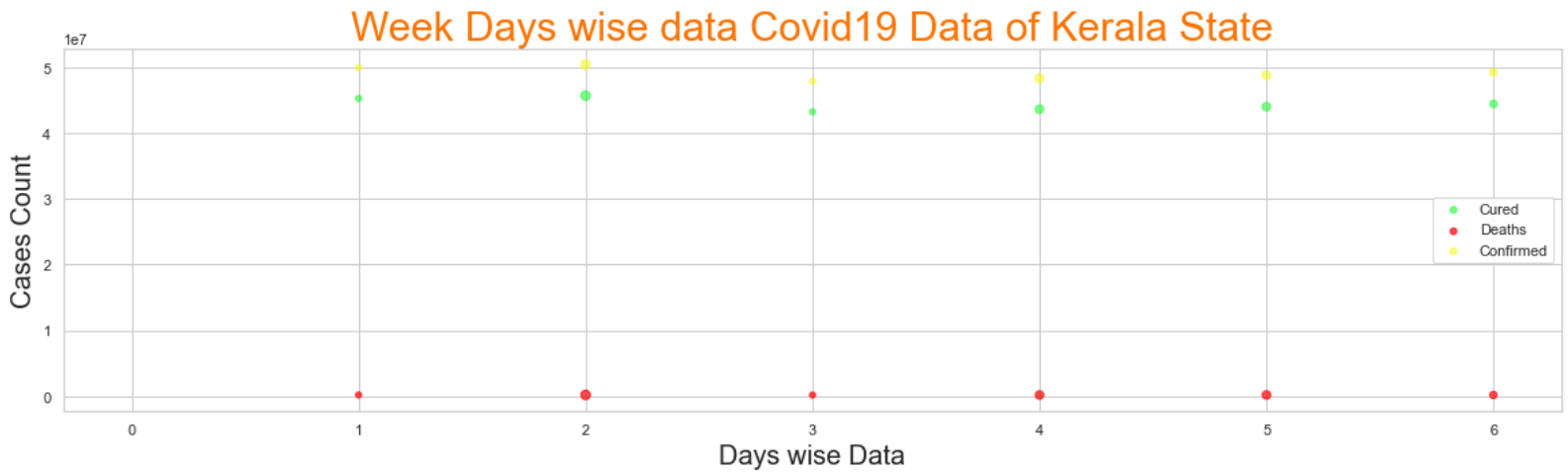
Out[223]:

week_day	Cured	Deaths	Confirmed
0	44824358	191720	49669831
1	45278940	193626	49962627
2	45691390	195681	50414261
3	43231396	183720	47892661
4	43640340	185665	48343334
5	44030953	187621	48788322
6	44430266	189721	49248009

Weekly Days wise data - Kerala state

In [224]:

```
plt.figure(figsize=(20,5))
labels=kerala_weekly_data.columns
sizes = 10*np.random.randint(7,size=(7))
plt.scatter(kerala_weekly_data.index,kerala_weekly_data.Cured,color='#04FA1E',alpha
plt.scatter(kerala_weekly_data.index,kerala_weekly_data.Deaths,color='#FA0404',alph
plt.scatter(kerala_weekly_data.index,kerala_weekly_data.Confirmed,color='#F7FA04',a
plt.legend(labels)
plt.xlabel('Days wise Data',size=20)
plt.ylabel('Cases Count',size=20)
plt.title("Week Days wise data Covid19 Data of Kerala State",size=30,color=title_co
plt.show()
```



In [225]:

```
#total confirmed cases detected on day basis
```

In [226]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==0]['Confirmed'].sum()
```

Out[226]:

```
626562686
```

In [227]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==1]['Confirmed'].sum()
```

Out[227]:

```
630436932
```

In [228]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==2]['Confirmed'].sum()
```

Out[228]:

```
634762271
```

In [229]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==3]['Confirmed'].sum()
```

Out[229]:

```
608618001
```

In [230]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==4]['Confirmed'].sum()
```

Out[230]:

```
613151011
```

In [231]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==5]['Confirmed'].sum()
```

Out[231]:

```
617711410
```

In [232]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==6]['Confirmed'].sum()
```

Out[232]:

```
622235763
```

In [233]:

```
#total deaths detected on day basis
```

In [234]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==0]['Deaths'].sum()
```

Out[234]:

```
8447800
```

In [235]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==1]['Deaths'].sum()
```

Out[235]:

```
8499965
```

In [236]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==2]['Deaths'].sum()
```

Out[236]:

```
8559975
```

In [237]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==3]['Deaths'].sum()
```

Out[237]:

```
8217023
```

In [238]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==4]['Deaths'].sum()
```

Out[238]:

```
8274842
```

In [239]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==5]['Deaths'].sum()
```

Out[239]:

```
8334502
```

In [240]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==6]['Deaths'].sum()
```

Out[240]:

```
8391893
```

In [241]:

```
#total cured cases detected on day basis
```

In [242]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==0]['Cured'].sum()
```

Out[242]:

```
572339316
```

In [243]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==1]['Cured'].sum()
```

Out[243]:

576725928

In [244]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==2]['Cured'].sum()
```

Out[244]:

581005853

In [245]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==3]['Cured'].sum()
```

Out[245]:

555391445

In [246]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==4]['Cured'].sum()
```

Out[246]:

559596627

In [247]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==5]['Cured'].sum()
```

Out[247]:

563883624

In [248]:

```
covid_19_india_copy[covid_19_india_copy['week_day']==6]['Cured'].sum()
```

Out[248]:

568251343

In [249]:

```
#Fetching data for the month of December
```

In [250]:

```
dec_data=covid_19_india_copy[covid_19_india_copy['Month']==12]
```

In [251]:

```
dec_data
```

Out[251]:

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
8976	8977	2020-12-01	8:00 AM	Andaman and Nicobar Islands	-	
8977	8978	2020-12-01	8:00 AM	Andhra Pradesh	-	
8978	8979	2020-12-01	8:00 AM	Arunachal Pradesh	-	
8979	8980	2020-12-01	8:00 AM	Assam	-	
8980	8981	2020-12-01	8:00 AM	Bihar	-	
...	...	...	...	...	...	
10077	10078	2020-12-31	8:00 AM	Telengana	-	
10078	10079	2020-12-31	8:00 AM	Tripura	-	
10079	10080	2020-12-31	8:00 AM	Uttarakhand	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
10080	10081	2020-12-31	8:00 AM	Uttar Pradesh	-	
10081	10082	2020-12-31	8:00 AM	West Bengal	-	

1106 rows × 15 columns

In [252]:

```
dec_data.groupby('State/UnionTerritory')[['Cured','Confirmed','Deaths']].sum() #std
```

Out[252]:

State/UnionTerritory	Cured	Confirmed	Deaths
Andaman and Nicobar Islands	145485	149755	1901
Andhra Pradesh	26782937	27154143	218804
Arunachal Pradesh	498669	512665	1711
Assam	6518436	6657671	31286
Bihar	7333313	7540194	41243
Chandigarh	549127	579205	9345
Chhattisgarh	7411448	8044501	96702
Dadra and Nagar Haveli and Daman and Diu	103598	104066	62
Delhi	17977943	18789163	310431
Goa	1479452	1536156	22082
Gujarat	6572851	7097154	129363
Haryana	7441429	7809674	84386

State/UnionTerritory	Cured	Confirmed	Deaths
Himachal Pradesh	1309890	1530320	25337
Jammu and Kashmir	3414908	3603650	55880
Jharkhand	3388287	3472129	30993
Karnataka	27068337	27994899	370482
Kerala	19089246	21071257	82588
Ladakh	261501	282403	3813
Lakshadweep	0	0	0
Madhya Pradesh	6490706	6974358	106252
Maharashtra	54632673	58306196	1499062
Manipur	758156	837188	10038
Meghalaya	377310	397526	3964
Mizoram	119872	125723	213
Nagaland	343262	361522	2197
Odisha	9902108	10054358	56156
Puducherry	1133568	1164162	19265
Punjab	4618235	4969915	157910

State/UnionTerritory	Cured	Confirmed	Deaths
Rajasthan	8440147	9042995	78781
Sikkim	152612	168048	3708
Tamil Nadu	24141778	24819908	369344
Telengana	8372522	8646516	46521
Tripura	1003331	1024869	11710
Uttar Pradesh	16732432	17564562	250784
Uttarakhand	2364361	2579458	42515
West Bengal	15314155	16211044	282550

In [253]:

```
dec_data['Cured'].mean()
```

Out[253]:

```
264235.1582278481
```

In [254]:

```
dec_data['Deaths'].mean()
```

Out[254]:

```
4030.1799276672696
```

In [255]:

```
dec_data['Confirmed'].mean()
```

Out[255]:

```
277737.2088607595
```

In [298]:

```
covid19_sample=covid_19_india_copy.sample(10)
covid19_sample
```

Out[298]:

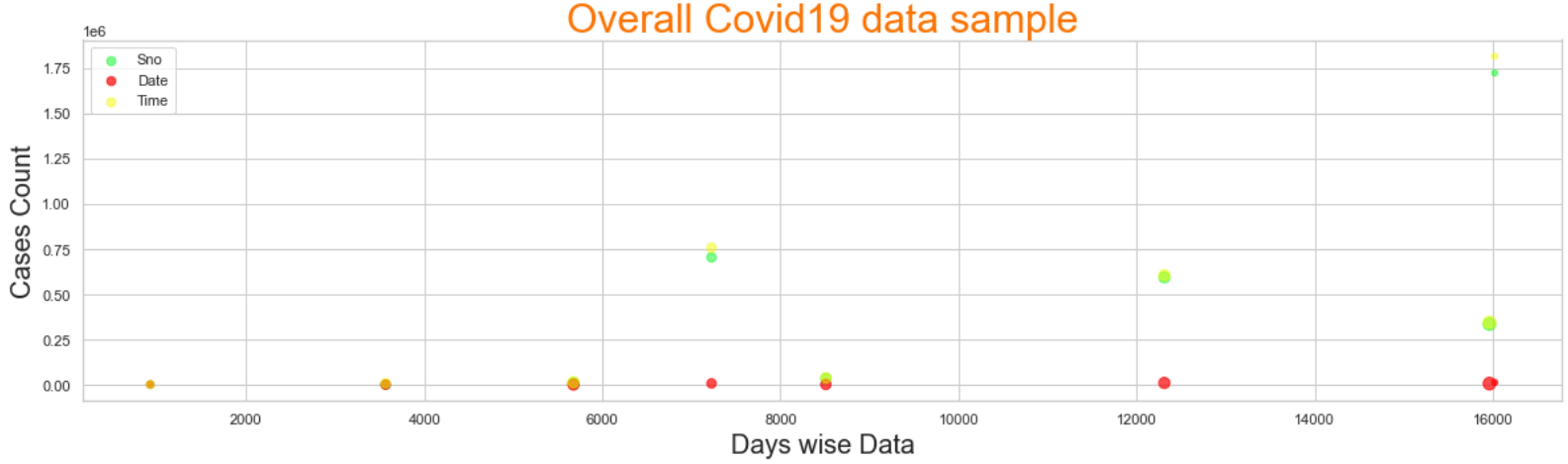
	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
16023	16024	2021-06-15	8:00 AM	Andhra Pradesh	-	
5676	5677	2020-08-28	8:00 AM	Puducherry	-	
12312	12313	2021-03-03	8:00 AM	Uttar Pradesh	-	
9218	9219	2020-12-07	8:00 AM	Uttarakhand	-	
923	924	2020-04-13	5:00 PM	Uttarakhand	-	
15964	15965	2021-06-13	8:00 AM	Jharkhand	-	
15499	15500	2021-05-31	8:00 AM	Ladakh	-	
8511	8512	2020-11-17	8:00 AM	Puducherry	-	

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedF
3566	3567	2020-06-29	8:00 AM	Cases being reassigned to states	-	
7227	7228	2020-10-12	8:00 AM	Andhra Pradesh	-	

Overall data sample

In [305]:

```
plt.figure(figsize=(20,5))
labels=covid19_sample.columns
sizes = 10*np.random.randint(10,size=(10))
plt.scatter(covid19_sample.index,covid19_sample.Cured,color='#04FA1E',alpha=0.5,s=s
plt.scatter(covid19_sample.index,covid19_sample.Deaths,color='#FA0404',alpha=0.7,s=
plt.scatter(covid19_sample.index,covid19_sample.Confirmed,color='#F7FA04',alpha=0.5
plt.legend(labels)
plt.xlabel('Days wise Data',size=20)
plt.ylabel('Cases Count',size=20)
plt.title("Overall Covid19 data sample",size=30,color=title_color)
plt.show()
```



In [ ]: