Q1. Data Manipulation

1. Count the total number of “Confirmed”, “Recovered” and “Deceased” from 14-Mar-2020 to 05-Sept-2020 and report the numbers.

Ans) confirmed_count: 4110211 recovered_count: 3177666 deceased_count: 70094

2. Count the total number of “Confirmed”, “Recovered” and “Deceased” from 14-Mar-2020 to 05-Sept-2020 for state Delhi (dl)

Ans) confirmed_count: 188193 recovered_count: 163785 deceased_count: 4538

3. Report total count of “Confirmed”, “Recovered” and “Deceased” count from states Delhi (dl)+ Maharasthra(mh) (Sum of both states count) from 14-Mar-2020 to 05-Sept-2020.

Ans) confirmed_count: 1072055 recovered_count: 1072055 deceased_count: 1072055

4. Report the highest affected state in terms of “Confirmed”, “Recovered” and “Deceased” with the count till 05-Sept-2020 from 14-Mar-2020.

Ans)
Confirmed :
Highest affected State is: 883862
Highest affected State count is: ['mh']

Recovered :
Highest affected State is: 636574
Highest affected State count is: ['mh']

Deceased :
Highest affected State is: 26275
Highest affected State count is: ['mh']

5. Report the lowest affected state in terms of “Confirmed”, “Recovered” and “Deceased” with the count till 05-Sept-2020 from 14-Mar-2020.

Ans) Confirmed :
Highest affected State is: 0
Highest affected State count is: ['dd', 'ld', 'un']

Recovered :
Highest affected State is: 0
Highest affected State count is: ['dd', 'ld', 'un']

Deceased :
Highest affected State is: 0
Highest affected State count is: ['dd', 'ld', 'mz', 'un']

6. Find the day and count with the highest spike in a day in the number of cases for the state Delhi for “Confirmed”, “Recovered” and “Deceased” between dates 14-Mar-2020 and 05-Sept-2020.

Ans)

Confirmed :
Day: 2020-06-23
Count: 3947

Recovered :
Day: 2020-06-20
Count: 7725

Deceased :
Day: 2020-06-16
Count: 437

7. Report active cases (Assume active = Confirmed - (Recovered + Deceased)) state wise for all states separately on date 05-Sept-2020 (This date only) starting from 14-March-2020.

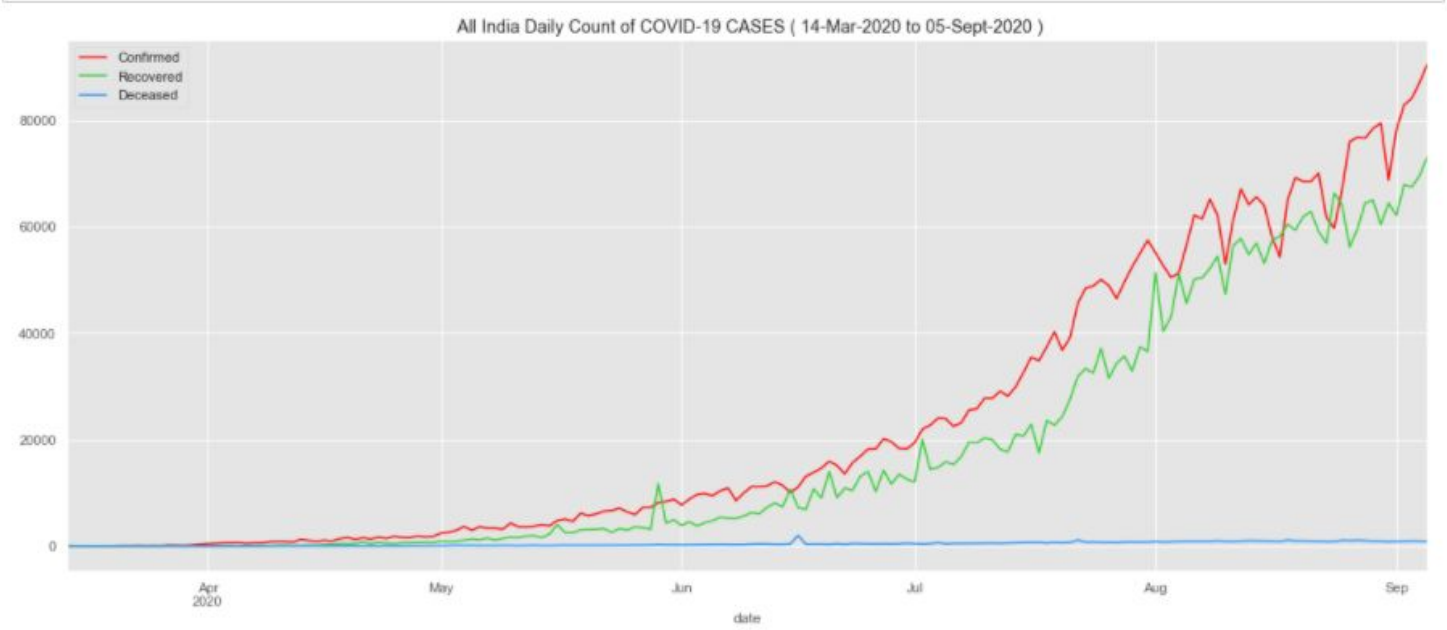
Ans)

an 343
ap 100880
ar 1525
as 28404
br 16735
ch 2143
ct 22320
dd 0
dl 19870
dn 301
ga 4945
gj 16266
hp 2023
hr 14912
jh 14980
jk 9547
ka 100224
kl 21867
la 834
ld 0
mh 221013
ml 1374
mn 1872
mp 15687
mz 349
nl 701
or 25856
pb 15870
py 5163

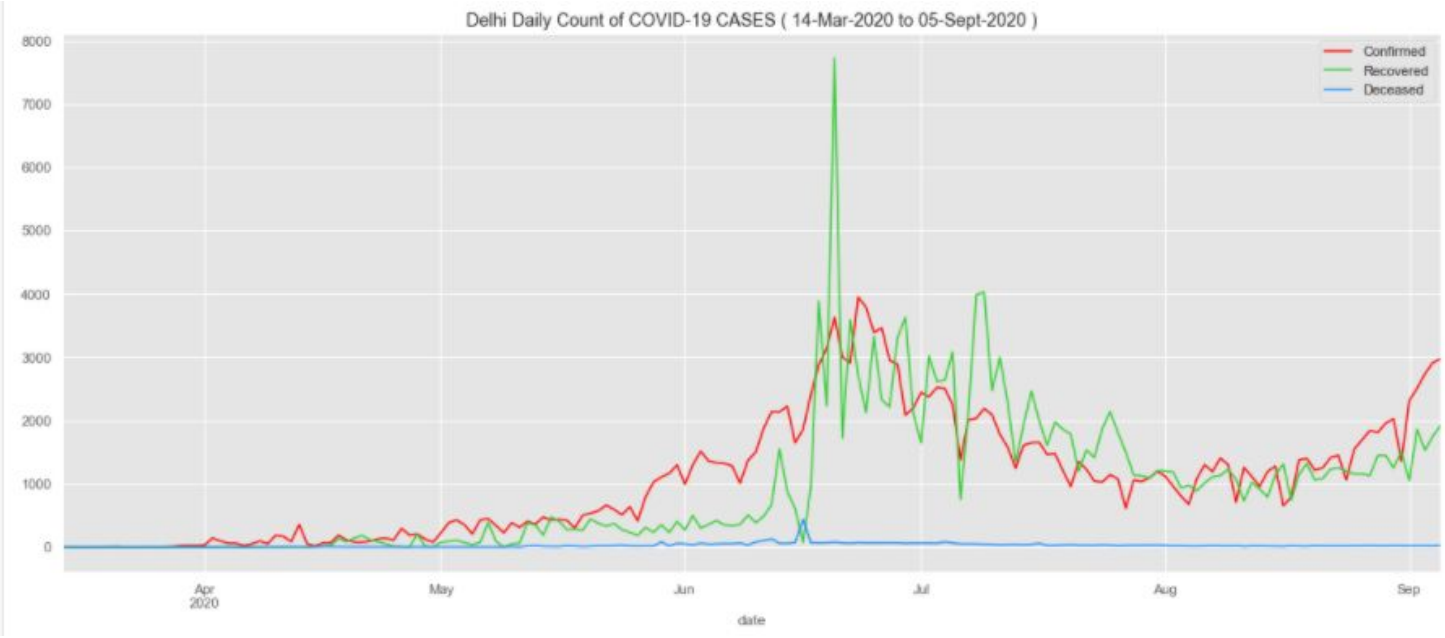
rj 14996
sk 561
tg 32405
tn 51580
tr 5905
tt 862451
un 0
up 59963
ut 7649
wb 23390

Q2. Plotting

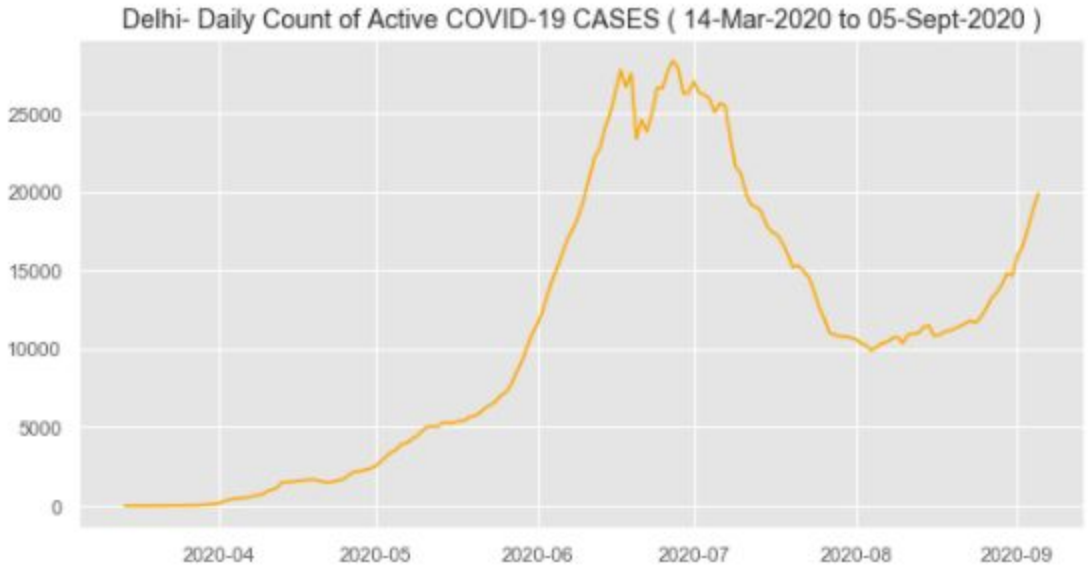
1. Plot the area trend line for total “Confirmed”, “Recovered” and “Deceased” cases from 14-Mar-2020 to 05-Sept-2010.



2. Plot the area trend line for total “Confirmed”, “Recovered” and “Deceased” cases for the state Delhi (dl) from 14-Mar-2020 to 05-Sept-2020.



3. Plot the area trend line for active cases. Assume active = Confirmed -(Recovered + Deceased) from 14-Mar-2020 to 05-Sept-2020.



Q3. Linear Regression

1. Implement a linear regression on the state Delhi data over dates, separately for “Confirmed”, “Recovered” or “Deceased” and report intercept and slope coefficients for all 3 cases from 14-Mar-2020 to 05-Sept-2020.

Ans)

Confirmed :
Intercept: 737572.4816363358 , Slope: 0.015658726862550116

Recovered :
Intercept: 737574.6884091854 , Slope: 0.016436943293872016

Deceased :
Intercept: 737586.638959881 , Slope: 0.22519668808977067

Taha Junaid
taha.junaid3000@gmail.com