

Q1. Data Manipulation

1. Count the total number of "Confirmed", "Recovered" and "Deceased" from 14-Mar-2020 to 30-Sept-2020 and report the numbers.

Answer 1: Total Confirmed: 6309638, Recovered: 5269993, Deceased: 98718

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

Answer 2 - Delhi (Confirmed): 188193, Recovered: 163785, Deceased: 4538

3. Report total count of "Confirmed", "Recovered" and "Deceased" count from states Delhi + Karnataka (ka) (Sum of both states count) from 14-Mar-2020 to 05-Sept-2020.

Answer 3 - Combined Delhi and Karnataka (Confirmed): 577425, Recovered: 447084, Deceased: 10843

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

Answer 4 - Highest Affected State (Confirmed): mh, Count: 883862

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

Answer 5 - Lowest Affected State (Confirmed): dd, Count: 0

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.

Answer 6 - Highest spike in Delhi (Confirmed): 2020-06-23 00:00:00 with 3947 cases
Highest spike in Delhi (Recovered): 2020-06-20 00:00:00 with 7725 cases
Highest spike in Delhi (Deceased): 2020-06-16 00:00:00 with 437 cases

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.

Answer 7 - Active cases on 05-Sept-2020:

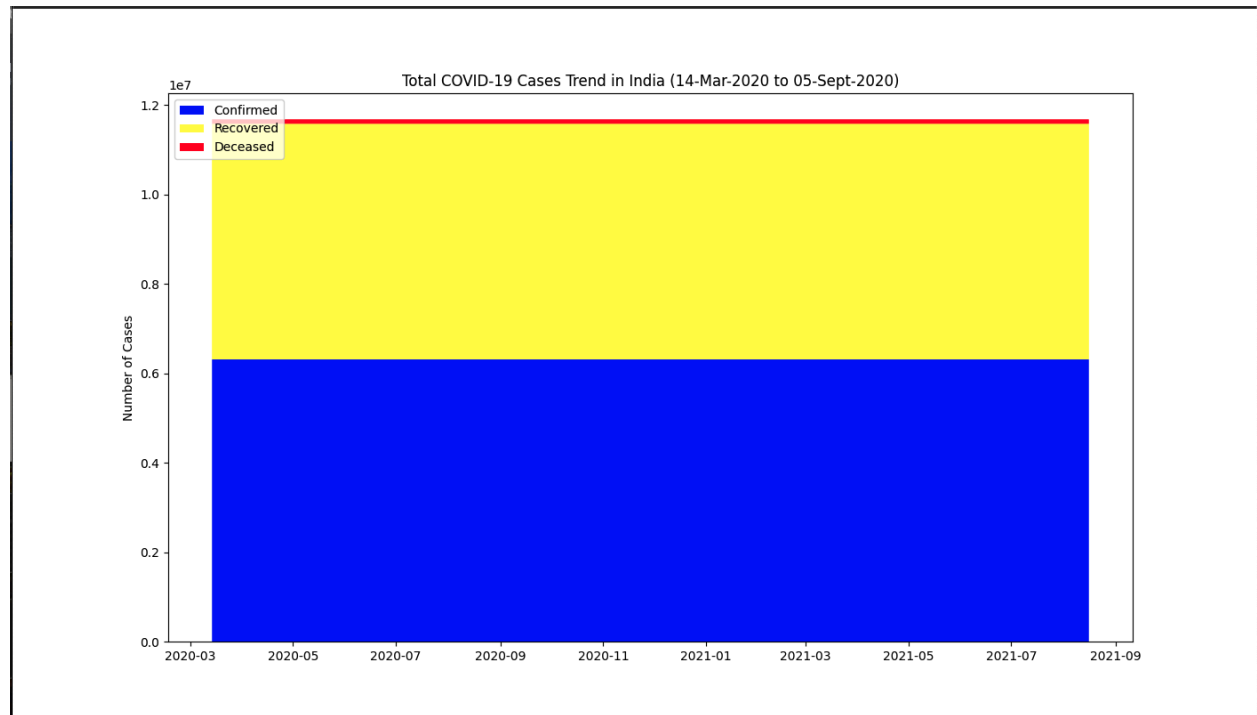
an: -7

ap: -1187

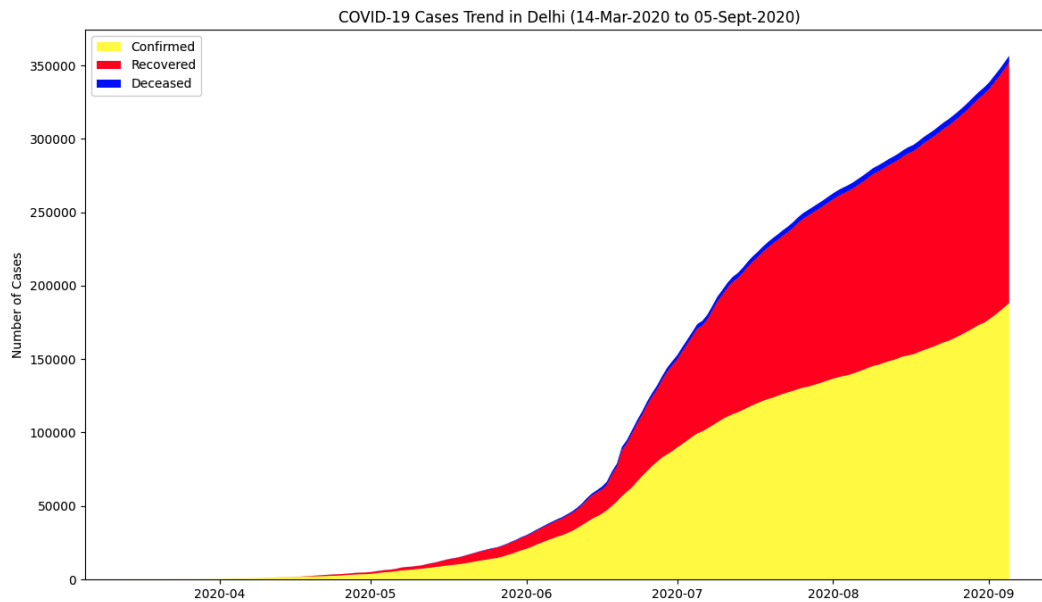
ar: 38
as: 348
br: -247
ch: 48
ct: 1631
dd: 0
dl: 1028
dn: -7
ga: 49
gj: 147
hp: 163
hr: 858
jh: -544
jk: 747
ka: 516
kl: 533
la: 19
ld: 0
mh: 9687
ml: 122
mn: 107
mp: 214
mz: -16
nl: -47
or: 93
pb: 139
py: -57
rj: 206
sk: 32
tg: -79
tn: -53
tr: 413
un: 0
up: 1368
ut: 397
wb: -264

Q2. Plotting

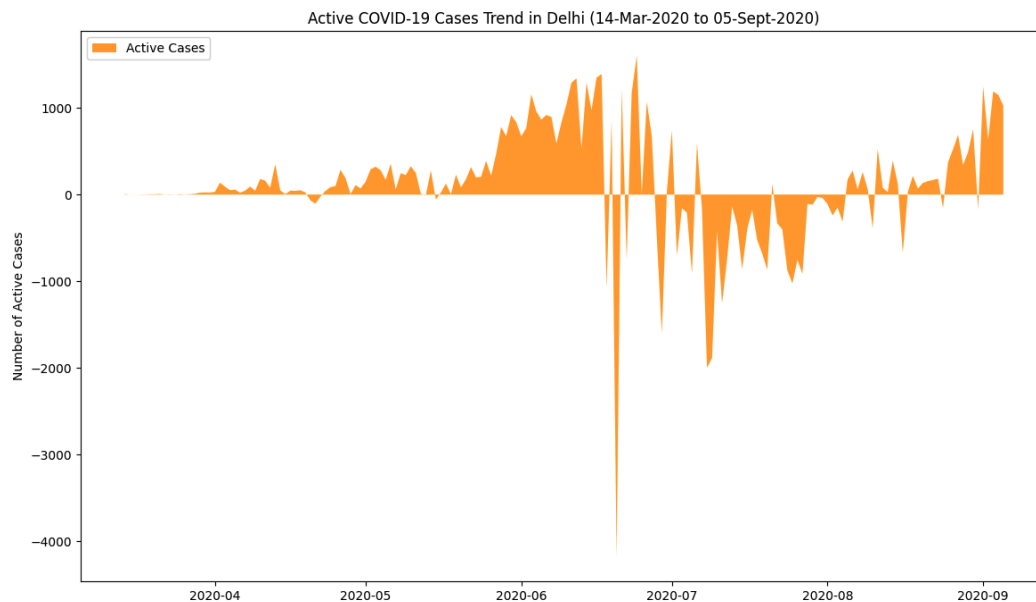
1. Plot the area trend line for total “Confirmed”, “Recovered” and “Deceased” cases from 14-Mar-2020 to 05-Sept-2010.



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

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: Linear Regression for Delhi - Confirmed cases: Intercept = 0.5298536209552367, Slope = 12.2142692053709

Linear Regression for Delhi - Recovered cases: Intercept = -146.13713405238866, Slope = 12.305528285274052

Linear Regression for Delhi - Deceased cases: Intercept = 9.138674884437595, Slope = 0.1902333259960379

