DMG - Assignment 0

Question 1 -

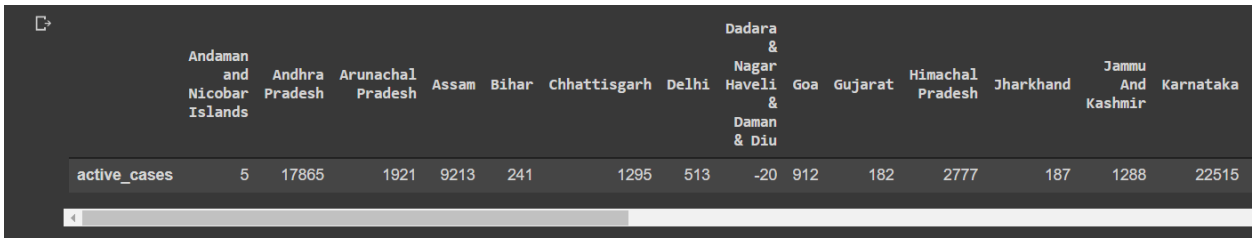
Assumptions -

1. Below are the states and ut's that we have considered.

```
states = {'an':'Andaman and Nicobar Islands',
          'ap':'Andhra Pradesh',
          'ar': 'Arunachal Pradesh',
          'as':'Assam',
          'br':'Bihar',
          'ct':'Chhattisgarh',
          'dl':'Delhi',
          'dn':'Dadara & Nagar Haveli & Daman & Diu',
          'ga':'Goa',
          'gj':'Gujarat',
          'hp':'Himachal Pradesh',
          'jh':'Jharkhand',
          'jk':'Jammu And Kashmir',
          'ka':'Karnataka',
          'kl':'Kerala',
          'la':'Ladakh',
          'ld':'Lakshadweep',
          'mh':'Maharashtra',
          'mn':'Manipur',
          'mp':'Madhya Pradesh',
          'nl':'Nagaland',
          'or':'Odisha',
          'pb':'Punjab',
          'py':'Puducherry',
          'rj':'Rajasthan',
          'sk':'Sikkim',
          'tg':'Telangana',
          'tn':'Tamil Nadu',
          'tr':'Tripura',
          'un':'States Unassigned',
          'up':'Uttar Pradesh',
          'ut':'Uttarakhand',
          'wb':'West Bengal'}
```

2. The number of active cases before the start date is assumed to be zero.
3. We have also assumed that the number of confirmed, Recovered, and Deceased cases before the start date are zero.

   This is also the reason the active cases of some states/UT's are negative on some dates.



	Andaman and Nicobar Islands	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chhattisgarh	Delhi	Dadara & Nagar Haveli & Daman & Diu	Goa	Gujarat	Himachal Pradesh	Jharkhand	Jammu And Kashmir	Karnataka
active_cases	5	17865	1921	9213	241	1295	513	-20	912	182	2777	187	1288	22515

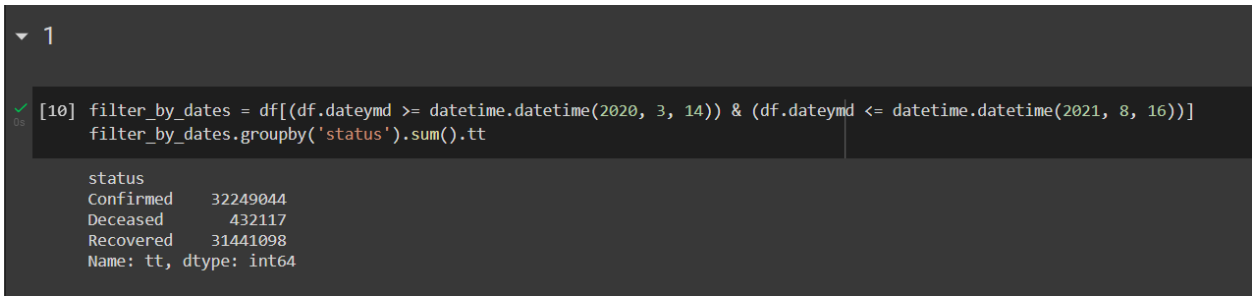
(See the active number of cases of `Dadara & Nagar Haveli & Daman & Diu` in the above pic)

4. The recovery rate for a state is taken as -

```
state_data_transpose["Recovery Rate"] = 100*state_data_transpose["Recovered"]/state_data_transpose["Confirmed"]
```

Outputs -

Part 1.

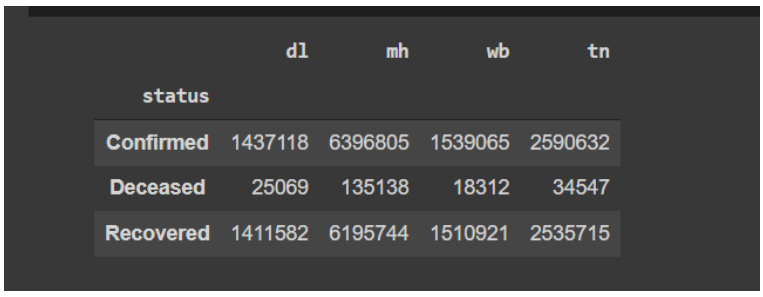


```
1

[10] filter_by_dates = df[(df.dateymd >= datetime.datetime(2020, 3, 14)) & (df.dateymd <= datetime.datetime(2021, 8, 16))]
     filter_by_dates.groupby('status').sum().tt

status
Confirmed    32249044
Deceased       432117
Recovered    31441098
Name: tt, dtype: int64
```

Part 2.



status	dl	mh	wb	tn
Confirmed	1437118	6396805	1539065	2590632
Deceased	25069	135138	18312	34547
Recovered	1411582	6195744	1510921	2535715

Part 3.

```
[14] state_data_transpose.nlargest(10,'Recovery Rate').index

     Index(['Dadara & Nagar Haveli & Daman & Diu', 'Rajasthan', 'Ladakh', 'Gujarat',
            'Lakshadweep', 'Madhya Pradesh', 'hr', 'Uttar Pradesh', 'Bihar',
            'Chhattisgarh'],
           dtype='object')

[15] state_data_transpose.nsmallest(10,'Recovery Rate').index

     Index(['Sikkim', 'Nagaland', 'Manipur', 'Kerala', 'Arunachal Pradesh',
            'Uttarakhand', 'Maharashtra', 'Himachal Pradesh', 'Tripura', 'Punjab'],
           dtype='object')
```

Part 4.

```
[17] #3 largest in terms of confirmed
     state_data_transpose.nlargest(3,'Confirmed')
```

status	Confirmed	Deceased	Recovered	Recovery Rate
Maharashtra	6396805	135138	6195744	96.856853
Kerala	3702417	18744	3510904	94.827352
Karnataka	2930529	37014	2871449	97.983982

```
[18] #3 largest in terms of recovered
     state_data_transpose.nlargest(3,'Recovered')
```

status	Confirmed	Deceased	Recovered	Recovery Rate
Maharashtra	6396805	135138	6195744	96.856853
Kerala	3702417	18744	3510904	94.827352
Karnataka	2930529	37014	2871449	97.983982

```
[19] #3 largest in terms of deceased
     state_data_transpose.nlargest(3,'Deceased')
```

status	Confirmed	Deceased	Recovered	Recovery Rate
Maharashtra	6396805	135138	6195744	96.856853
Karnataka	2930529	37014	2871449	97.983982
Tamil Nadu	2590632	34547	2535715	97.880170

Part 5.

```
#3 lowest in terms of confirmed
state_data_transpose.nsmallest(3,'Confirmed')
```

status	Confirmed	Deceased	Recovered	Recovery Rate
Andaman and Nicobar Islands	7545	129	7410	98.210736
Lakshadweep	10294	51	10164	98.737128
Dadara & Nagar Haveli & Daman & Diu	10520	4	10534	100.133080

```
[21] #3 lowest in terms of recovered
state_data_transpose.nsmallest(3,'Recovered')
```

status	Confirmed	Deceased	Recovered	Recovery Rate
Andaman and Nicobar Islands	7545	129	7410	98.210736
Lakshadweep	10294	51	10164	98.737128
Dadara & Nagar Haveli & Daman & Diu	10520	4	10534	100.133080

```
[22] #3 lowest in terms of deceased
state_data_transpose.nsmallest(3,'Deceased')
```

status	Confirmed	Deceased	Recovered	Recovery Rate
Dadara & Nagar Haveli & Daman & Diu	10520	4	10534	100.133080
Lakshadweep	10294	51	10164	98.737128
Andaman and Nicobar Islands	7545	129	7410	98.210736

Part 6. (For the remaining part please have a look at the .ipynb file)

	Andaman and Nicobar Islands	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chhattisgarh	Delhi	Dadara & Nagar Haveli & Daman & Diu	Goa	Gujarat	Himachal Pradesh	Jharkhand	Jammu And Kashmir
Confirmed	{2020-08-14 00:00:00: 149}	{2021-05-16 00:00:00: 24171}	{2021-07-12 00:00:00: 566}	{2021-05-20 00:00:00: 6573}	{2021-04-30 00:00:00: 15853}	{2021-04-23 00:00:00: 17397}	{2021-04-20 00:00:00: 28395}	{2021-04-22 00:00:00: 359}	{2021-05-07 00:00:00: 4195}	{2021-04-30 00:00:00: 14605}	{2021-05-08 00:00:00: 5424}	{2021-04-28 00:00:00: 8075}	{2021-05-07 00:00:00: 5443}
Recovered	{2020-08-09 00:00:00: 148}	{2021-05-19 00:00:00: 24819}	{2021-07-23 00:00:00: 517}	{2021-05-26 00:00:00: 6266}	{2021-05-10 00:00:00: 15800}	{2021-04-20 00:00:00: 18746}	{2021-05-01 00:00:00: 27421}	{2021-05-28 00:00:00: 339}	{2021-05-17 00:00:00: 4008}	{2021-05-14 00:00:00: 15365}	{2021-05-22 00:00:00: 5021}	{2021-05-12 00:00:00: 8331}	{2021-05-23 00:00:00: 4956}
Deceased	{2021-05-28 00:00:00: 4}	{2021-05-22 00:00:00: 118}	{2021-06-14 00:00:00: 6}	{2021-05-25 00:00:00: 92}	{2021-06-09 00:00:00: 3971}	{2021-04-28 00:00:00: 279}	{2021-05-03 00:00:00: 448}	{2021-04-15 00:00:00: 2}	{2021-06-07 00:00:00: 80}	{2021-04-29 00:00:00: 180}	{2021-05-18 00:00:00: 78}	{2021-05-01 00:00:00: 169}	{2021-05-17 00:00:00: 73}

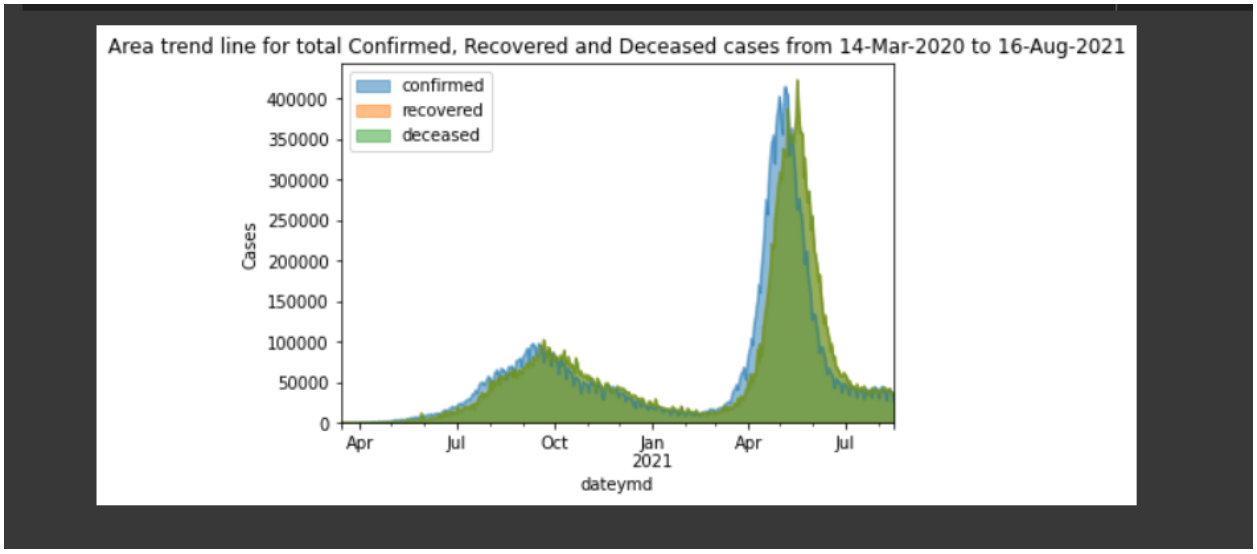
Part 7. (For the remaining part please have a look at the .ipynb file)

	Andaman and Nicobar Islands	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chhattisgarh	Delhi	Dadara & Nagar Haveli & Daman & Diu	Goa	Gujarat	Himachal Pradesh	Jharkhand	Jammu And Kashmir	Karnataka	Kerala
active_cases	5	17865	1921	9213	241	1295	513	-20	912	182	2777	187	1288	22515	179159

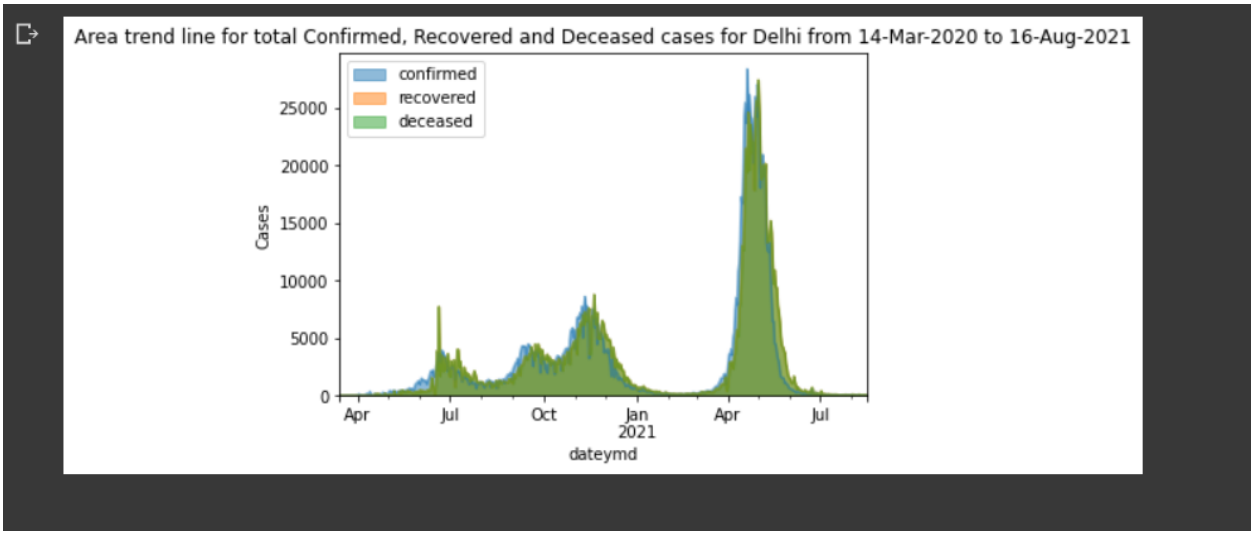
Question 2.

Results

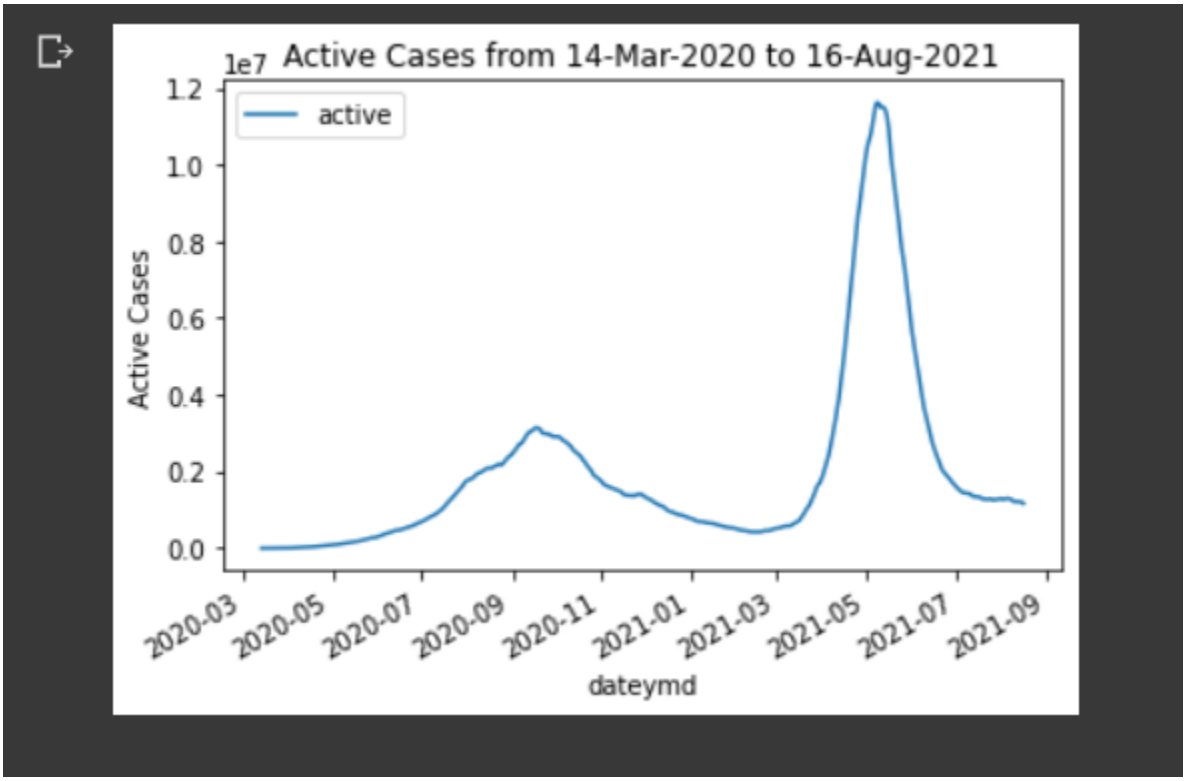
Part 1.



Part 2.


Area trend line for total Confirmed, Recovered and Deceased cases for Delhi from 14-Mar-2020 to 16-Aug-2021
confirmed
recovered
deceased
Cases
25000
20000
15000
10000
5000
0
Apr
Jul
Oct
Jan
2021
Apr
Jul
dateymd

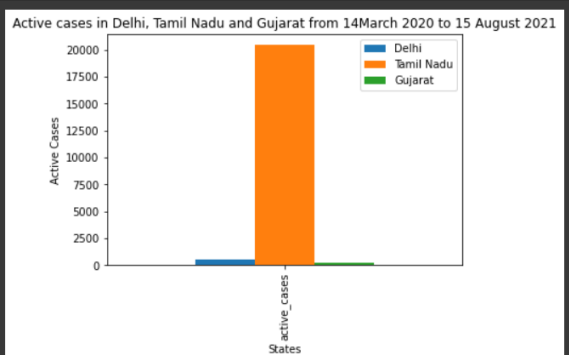
Part 3.


Active Cases from 14-Mar-2020 to 16-Aug-2021
1e7
active
Active Cases
1.2
1.0
0.8
0.6
0.4
0.2
0.0
2020-03
2020-05
2020-07
2020-09
2020-11
2021-01
2021-03
2021-05
2021-07
2021-09
dateymd

Part 4.

Date Range: 14 March 2020-15 Aug 2021

```
active_cases_3states = new_df[["Delhi","Tamil Nadu","Gujarat"]]
active_cases_3states
active_cases_3states.plot.bar()
plt.xlabel("States")
plt.ylabel("Active Cases")
plt.title("Active cases in Delhi, Tamil Nadu and Gujarat from 14March 2020 to 15 August 2021")
plt.show()
```



We have assumed the start date as 14th March and the end date as 15th August 2021. The Bar graph shows the active cases on the end date starting from the start date.