CS 649 Big Data: Tools and Methods
Spring Semester, 2021
Doc 23 Exam Comments
Apr 27, 2021

	US Total per "United States" Location	US Total without "US" Location	Total for States, DC, Ind. Countries, Groups	Total for U.S Territories	Total for Groups	Total for Ind. Countries	Total for States, DC, Ind. Countries, U.S.Territories	Discreptancy between US total calculated and US total provided	Discreptancy between US total calculated without Groups and US total provided
Total Vaccinations	202282923.000	217536581.000	215798956.000	1737625.000	16071987.000	69346.000	201464594.000	15253658.000	-818329.000
Total Distributed	258502815.000	274217715.000	271236585.000	2981130.000	15714900.000	139100.000	258502815.000	15714900.000	0.000
People Vaccinated	127743096.000	132884741.000	131774949.000	1109792.000	6077763.000	40340.000	126806978.000	5141645.000	-936118.000
Fully Vaccinated per 100	24.280	1494.930	1375.830	119.100	22.700	72.880	1472.230	1470.650	1447.950
Total Vaccinations per 100	60.930	3673.190	3382.050	291.140	54.600	161.140	3618.590	3612.260	3557.660
Fully Vaccinated	80609818.000	83614702.000	82927331.000	687371.000	3540006.000	29470.000	80074696.000	3004884.000	-535122.000
People Vaccinated per 100	38.480	2261.740	2089.320	172.420	32.440	90.310	2229.300	2223.260	2190.820
Distributed per 100	77.860	4821.030	4402.650	418.380	79.150	286.180	4741.880	4743.170	4664.020
Daily Vaccinations Raw	4629928.000	7940750.000	7844456.000	96294.000	577318.000	3195.000	7363432.000	3310822.000	2733504.000

```
display(vaccination_data_totals_output_table.style.format("{:,.1f}"))
```

```
pd.set_option('display.max_rows', None)
pd.options.display.float_format = '{0:,.1f}'.format
```

	US Total per "United States" Location	US Total without "US" Location	Total for States, DC, Ind. Countries, Groups	Total for U.S Territories	Total for Groups	Total for Ind. Countries	Total for States, DC, Ind. Countries, U.S.Territories	Discreptancy between US total calculated and US total provided	Discreptancy between US total calculated without Groups and US total provided
Total Vaccinations	205,871,913.0	221,210,329.0	219,440,889.0	1,769,440.0	16,171,205.0	70,204.0	205,039,124.0	15,338,416.0	-832,789.0
Total Distributed	264,499,715.0	280,226,275.0	277,210,755.0	3,015,520.0	15,726,560.0	139,100.0	264,499,715.0	15,726,560.0	0.0
People Vaccinated	129,494,179.0	134,666,572.0	133,539,432.0	1,127,140.0	6,121,575.0	40,784.0	128,544,997.0	5,172,393.0	-949,182.0
Fully Vaccinated per 100	24.8	1,525.1	1,403.7	121.3	23.1	73.5	1,502.0	1,500.2	1,477.2
Total Vaccinations per 100	62.0	3,729.0	3,433.5	295.5	55.2	162.3	3,673.8	3,667.0	3,611.8
Fully Vaccinated	82,471,151.0	85,520,344.0	84,816,840.0	703,504.0	3,594,282.0	29,908.0	81,926,062.0	3,049,193.0	-545,089.0
People Vaccinated per 100	39.0	2,287.8	2,113.2	174.6	32.7	90.9	2,255.1	2,248.8	2,216.1
Distributed per 100	79.7	4,911.5	4,492.1	419.5	79.6	286.2	4,831.9	4,831.9	4,752.3
Daily Vaccinations Raw	4,629,928.0	7,940,750.0	7,844,456.0	96,294.0	577,318.0	3,195.0	7,363,432.0	3,310,822.0	2,733,504.0

What does this mean?

US Country Vaccination Count
total_vaccinations   118313818.0
dtype: float64

Sum Of All Other Groups Vaccination Count
total_vaccinations   131268371.0
dtype: float64

```
print(grouped_percentage_fully_vaccinated.max().sort_values(by=
                                ['people_fully_vaccinated_per_hundred'], ascending=False))
```

location	people_fully_vaccinated_per_hundred
Alaska	19.16
Republic of Palau	18.86
New Mexico	18.63
Guam	18.32
Northern Mariana Islands	17.52
...	...
Federated States of Micronesia	8.17
Bureau of Prisons	NaN
Dept of Defense	NaN
Long Term Care	NaN
Veterans Health	NaN

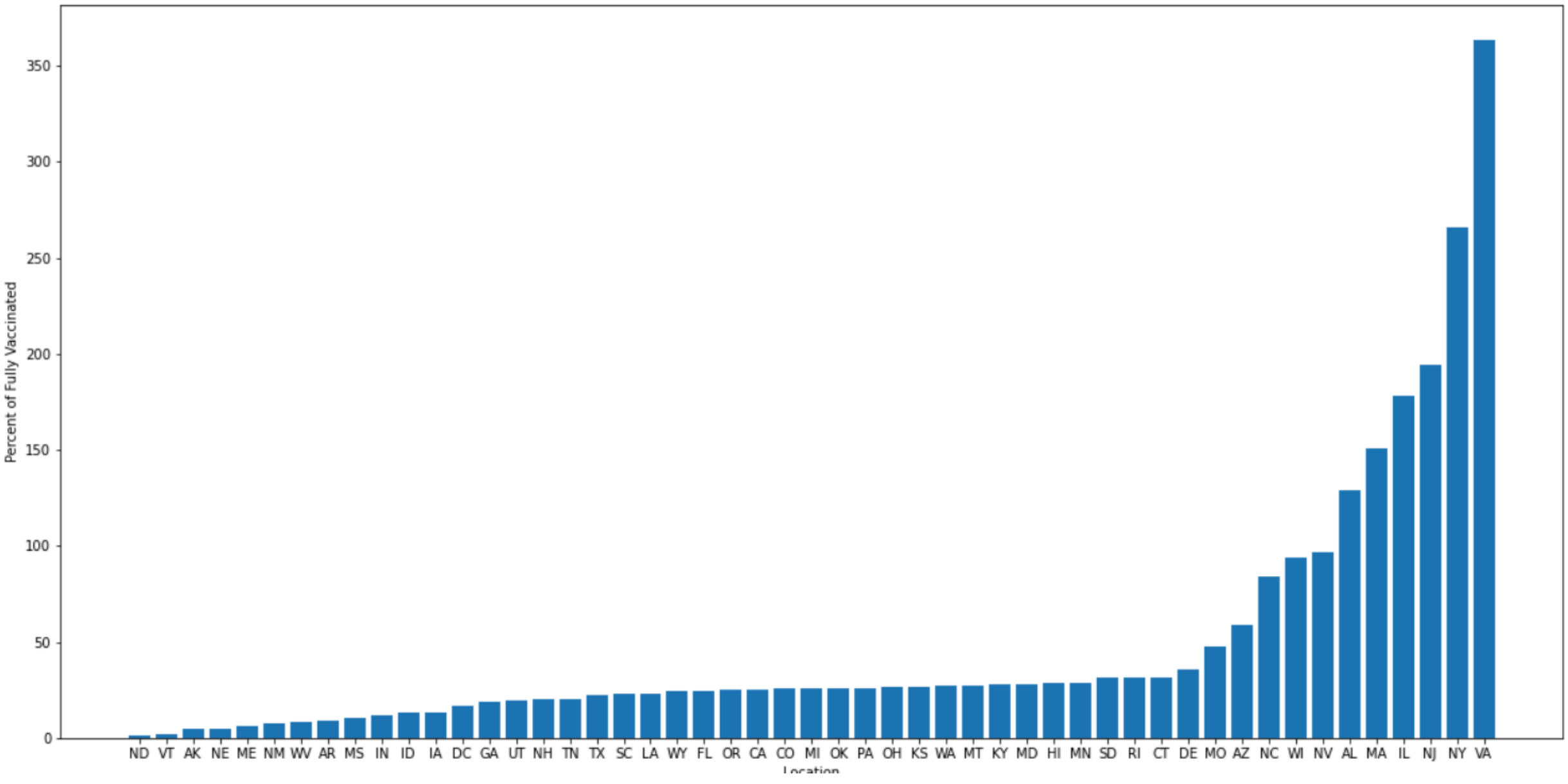
```
display(grouped_percentage_fully_vaccinated.max().sort_values(by=
                ['people_fully_vaccinated_per_hundred'], ascending=False).style.format("{:.1f}")
```

	people_fully_vaccinated_per_hundred
location	
Alaska	19.2
Republic of Palau	18.9
New Mexico	18.6
Guam	18.3
Northern Mariana Islands	17.5
South Dakota	17.2
American Samoa	16.2
North Dakota	15.9
Hawaii	15.8
Connecticut	15.7
West Virginia	15.4
Maine	15.2
Montana	15.0
Nebraska	14.8
Iowa	14.7
Wyoming	14.7

```
def states_abbreviations(state):
    if state == 'Alabama':
        return 'AL'
    elif state == 'Alaska':
        return 'AK'
    elif state == 'Arizona':
        return 'AZ'
    elif state == 'Arkansas':
        return 'AR'
    elif state == 'California':
        return 'CA'
    elif state == 'Colorado':
        return 'CO'
    elif state == 'Connecticut':
        return 'CT'
    elif state == 'Delaware':
        return 'DE'
    elif state == 'District of Columbia':
        return 'DC'
    elif state == 'Florida':
        return 'FL'
```

```
STATE_TO_ABV = {
    'Alabama': 'AL',
    'Alaska': 'AK',
    'Arizona': 'AZ',
    'Arkansas': 'AR',
    'California': 'CA',
    'Colorado': 'CO',
    'Connecticut': 'CT',
    'Delaware': 'DE',
    'District of Columbia': 'DC',
    'Florida': 'FL',
    'Georgia': 'GA',
    'Hawaii': 'HI',
    'Idaho': 'ID',
    'Illinois': 'IL',
```

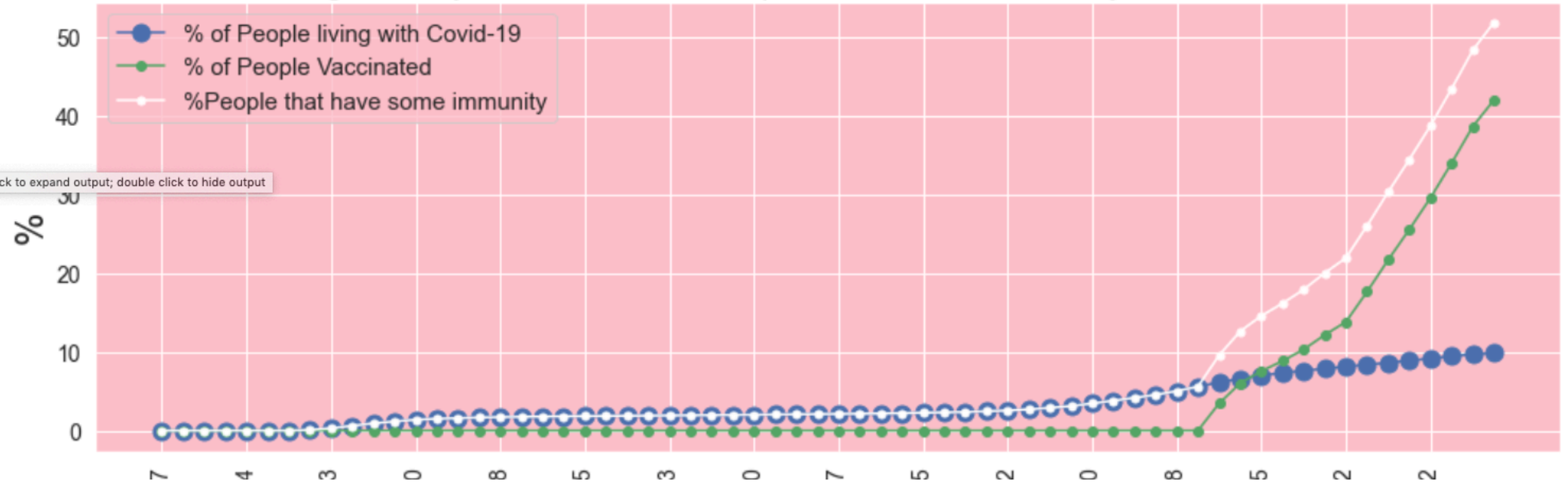
350%


Percent of Fully Vaccinated
0
50
100
150
200
250
300
350
ND VT AK NE ME NM WV AR MS IN ID IA DC GA UT NH TN TX SC LA WY FL OR CA CO MI OK PA OH KS WA MT KY MD HI MN SD RI CT DE MO AZ NC WI NV AL MA IL NJ NY VA
Location

8

```
us_states_recent_data['population_proportion'] = us_states_recent_data['population'].apply(lambda pop: pop/united_states_population)
#Add people fully vaccinated from special groups into states proportionally
us_states_recent_data['ammended_people_fully_vaccinated'] = us_states_recent_data['population_proportion']*special_recent_data['people_fully_vaccinated'].sum()+us_states_recent_data['people_fully_vaccinated']
#Get percent difference comparing previous people_fully_vaccinated with the ammended_people_fully_vaccinated
us_states_recent_data['fully_vaccinated_percent_diff'] = (us_states_recent_data['ammended_people_fully_vaccinated'] - us_states_recent_data['people_fully_vaccinated'])*100 /
((us_states_recent_data['ammended_people_fully_vaccinated'] + us_states_recent_data['people_fully_vaccinated'])/2)
print("Average percent difference of fully vaccinated people: " + str(us_states_recent_data['fully_vaccinated_percent_diff'].mean()))
print()
print(us_states_recent_data[['location', 'fully_vaccinated_percent_diff']])
```

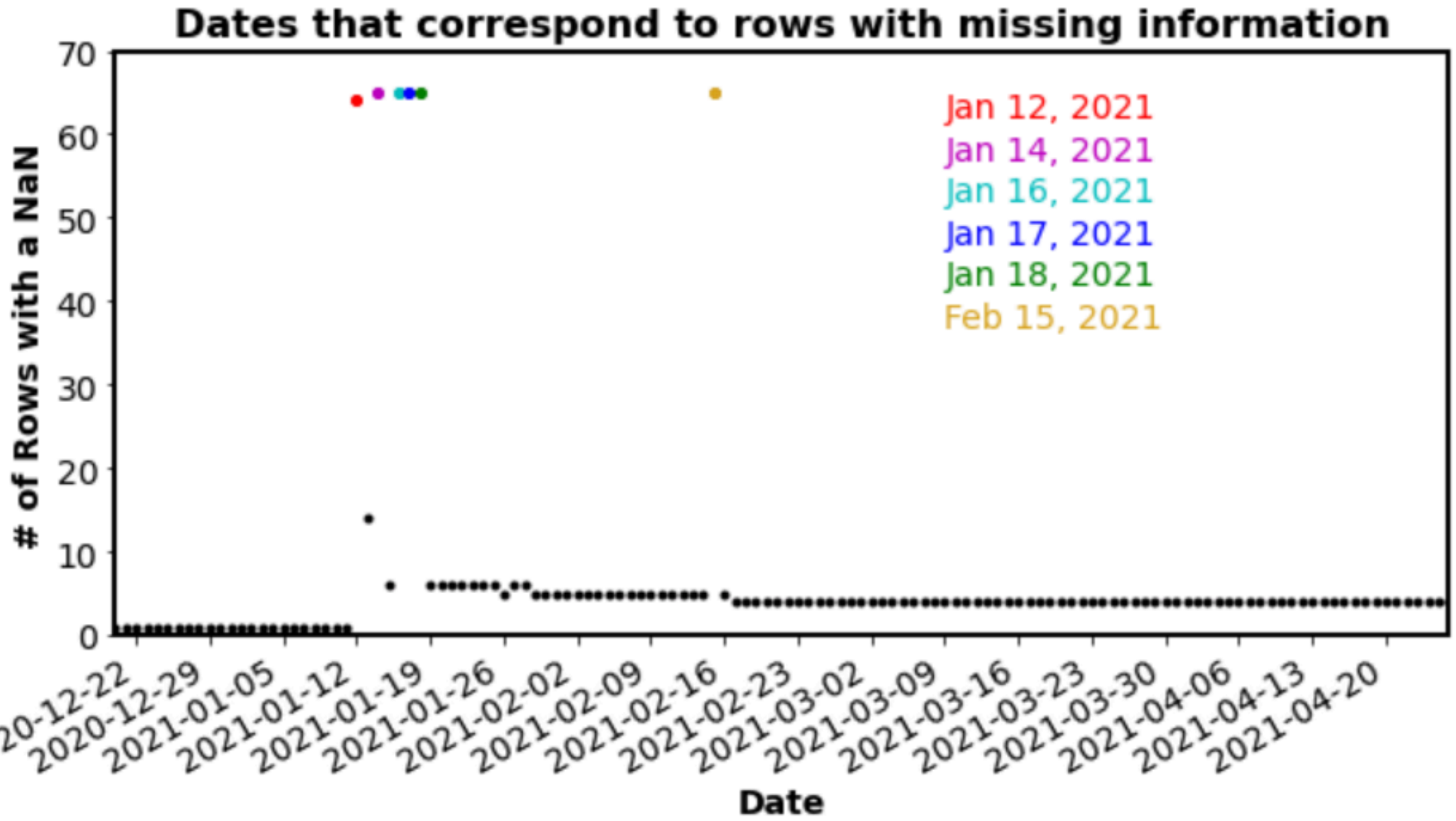
Fig.4 Weekly Data of Percent of People that have some immunity to Covid-19 in NY



10

	people_fully_vaccinated	%vaccinated	fully_vaccinated_total	%vaccinated_total
Maine	486,986.00	36.23	503,072.07	37.42
Connecticut	1,278,901.00	35.87	1,321,566.47	37.07

people_fully_vaccinated vs fully_vaccinated_total


Dates that correspond to rows with missing information
# of Rows with a NaN
0
10
20
30
40
50
60
70
Jan 12, 2021
Jan 14, 2021
Jan 16, 2021
Jan 17, 2021
Jan 18, 2021
Feb 15, 2021
2020-12-22
2020-12-29
2021-01-05
2021-01-12
2021-01-19
2021-01-26
2021-02-02
2021-02-09
2021-02-16
2021-02-23
2021-03-02
2021-03-09
2021-03-16
2021-03-23
2021-03-30
2021-04-06
2021-04-13
2021-04-20
Date

Out[16]:

	State	population	State Name	date	location	total_vaccinations	total_distributed	people_vaccinated	people_fully_vaccinated_per_hundred	total_va
0	AK	731545	Alaska	2021-04-26	Alaska	540,565.00	759,125.00	300,706.00	34.04	
1	AL	4903185	Alabama	2021-04-26	Alabama	2,478,227.00	3,960,530.00	1,549,782.00	21.75	
2	AR	3017804	Arkansas	2021-04-26	Arkansas	1,734,405.00	2,586,420.00	1,053,861.00	24.22	
3	AZ	7278717	Arizona	2021-04-26	Arizona	4,912,269.00	6,285,415.00	2,900,707.00	27.61	
4	CA	39512223	California	2021-04-26	California	29,151,065.00	36,101,260.00	18,667,924.00	28.78	
5	CO	5758736	Colorado	2021-04-26	Colorado	4,241,020.00	4,968,480.00	2,604,683.00	30.31	
6	CT	3565287	Connecticut	2021-04-26	Connecticut	3,094,177.00	3,716,115.00	1,897,401.00	36.43	

4. Produce a chart or table showing the percent of people vaccinated. Sort the data by the percent of people vaccina the dataset so when it is run we get the most recent data.

[18]:
```
df_state_vaccine_and_pop_data_sorted = df_state_vaccine_and_pop_data
df_state_vaccine_and_pop_data_sorted
```

[18]:

	State	State Name	% Population Fully Vaccinated	% Change
0	AL	Alabama	21.75	5.50
1	UT	Utah	22.01	5.44
2	MS	Mississippi	22.33	5.36
3	GA	Georgia	22.49	5.32

percentage of people fully vaccinated